Fundamentals of Semigroup Theory

John M. Howie

Regius Professor of Mathematics
University of St Andrews

CLARENDON PRESS · OXFORD

LONDON MATHEMATICAL SOCIETY
MONOGRAPHS NEW SERIES

Series Editors

H. G. Dales P. M. Neumann

LONDON MATHEMATICAL SOCIETY MONOGRAPHS NEW SERIES

Previous volumes of the LMS Monographs were published by Academic Press, to whom all enquiries should be addressed. Volumes in the New Series will be published by Oxford University Press throughout the world.

NEW SERIES

1. *Diophantine inequalities* R. C. Baker
2. *The Schur multiplier* Gregory Karpilovsky
3. *Existentially closed groups* Graham Higman and Elizabeth Scott
4. *The asymptotic solution of linear diffential systems* M. S. P. Eastham
5. *The restricted Burnside problem* Michael Vaughan-Lee
6. *Pluripotential theory* Maciej Klimek
7. *Free Lie algebras* Christophe Reutenauer
8. *The restricted Burnside problem (2nd edition)* Michael Vaughan-Lee
9. *The geometry of topological stability* Andrew du Plessis and Terry Wall
10. *Spectral decompositions and analytic sheaves* J. Eschmeier and M. Putinar
11. *An atlas of Brauer characters* Christoph Jansen, Klaus Lux, Richard Parker, and Robert Wilson
12. *Fundamentals of semigroup theory* John M. Howie

This book has been printed digitally and produced in a standard specification in order to ensure its continuing availability

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

Published in the United States by Oxford University Press Inc., New York

Reprinted 2003

ISBN 0-19-851194-9

In memory of

Alfred Hoblitzelle Clifford (11 July, 1908 – 27 December, 1992)

Preface

Since the publication in 1976 of my earlier monograph, *An Introduction to Semigroup Theory*, much has happened in the theory of semigroups, too much, indeed, for a general introduction to do more than scrape the surface of the subject, for its scope has widened remarkably to embrace many aspects of theoretical computer science. In response to the growth of the subject in both pure and applied aspects, many more specialist books have appeared (Eilenberg 1974, 1976; Lallement 1979; Nambooripad 1979; Lothaire 1983; Berstel and Perrin 1985; Petrich 1984; Pastijn and Petrich 1985; Okniński 1990; Jürgensen *et al.* 1991; Pin 1986; Shyr 1991; Higgins 1992; Almeida 1995), and the list of Conference Proceedings volumes on my shelves (T. E. Hall *et al.* 1980; Jürgensen *et al.* 1981; Byleen *et al.* 1984; Hofmann *et al.* 1983; Pollák *et al.* 1985; Almeida *et al.* 1990; T. E. Hall *et al.* 1991; Rhodes 1991; Howie *et al.* 1992; Shum and Yuen 1993; Bonzini *et al.* 1993) is certainly not exhaustive.

There is, however, still a place for a general introduction, offering both an overview of the subject for specialist and non-specialist alike, and an entrée for the graduate student. This is what I have set out to provide. I have used my earlier volume as a basis, and have been gratified to find that most of the material there still earns its place in a general introduction to the subject. Most of that, however, has been substantially rewritten to provide the perspective that seems most relevant to contemporary research, and significant amounts of new (post 1976) material have been incorporated, especially in the exercises.

The emphasis throughout is unashamedly on what might be called 'pure' semigroup theory; the inclusion of significant amounts of applications, for example, to automata, languages and machines, would have involved a huge, and probably unacceptable, increase in the length (and the price) of the book. The only mild genuflection in the direction of applied semigroup theory occurs in Chapter 7, where the brief section on variable length codes gives some hint of a fascinating and continually developing field.

The first six chapters give what I hope is a reasonably coherent account of regular semigroups of various kinds. Chapters 1 and 2 develop the fundamental language and concepts of the subject, and on those foundations is then built a fairly natural edifice, consisting of completely (0-)simple semigroups (Chapter 3), completely regular semigroups (Chapter 4), and inverse semigroups (Chapter 5). The huge success of inverse semigroup theory has naturally given rise to a great deal of work (of varying degrees of interest, it must be said) on generalizations. In Chapter 6, somewhat lamely entitled 'Other classes of regular semigroups', I give a very brief account of some of the more interesting types of non-inverse regular semigroups.

Chapter 7, already mentioned above, heralds a change of theme. It give a brief account of free semigroups and monoids and of variable length codes, an account which is certainly superficial, but which may whet readers' appetites and lead them to the excellent specialized texts now available.

The use of module theory and homological algebra in the study of rings is well known. One of the most encouraging developments of the last two decades has been the use of what one might call 'non-additive' homological algebra in the study of semigroups. A notable success in this area has been in the study of semigroup amalgams, and these are the subject of the final chapter.

The layout of the book is entirely traditional, and the system of referencing is, I hope, self-explanatory. I attempt to guide the reader by making a distinction between theorems and propositions, the former term being reserved for results of greater depth or importance. The distinction is, however, merely a guide, and should not be taken too seriously. A few more specialized sections in the earlier chapters have been starred to indicate that they may safely be omitted in a first reading.

It is a pleasure to record thanks to the University of St Andrews for a period of research leave during which a significant part of this book was written, and to the university of Tasmania for taking me in during part of that leave. Thanks are due to T. E. Hall, P. G. Trotter and James Renshaw, who read and commented on parts of the book, and especially to Nikola Ruškuc, whose careful reading and frank comments were of enormous service.

The first tentative steps towards the writing of this book had just been taken when I learned of the death of A. H. Clifford. At an early stage of my mathematical career, in 1964–65, I spent a year with Alfred Clifford at Tulane University, and was much influenced by his penetrating mind and infectious enthusiasm. This book is dedicated, with deep respect, to his memory.

University of St Andrews J. M. H.
February 1995

Contents

1	**Introductory ideas**		1
	1.1	Basic definitions	1
	1.2	Monogenic semigroups	8
	1.3	Ordered sets, semilattices and lattices	13
	1.4	Binary relations; equivalences	16
	1.5	Congruences	22
	1.6	Free semigroups and monoids; presentations	29
	1.7	Ideals and Rees congruences	33
	1.8	Lattices of equivalences and congruences*	34
	1.9	Exercises	37
	1.10	Notes	44
2	**Green's equivalences; regular semigroups**		45
	2.1	Green's equivalences	45
	2.2	The structure of $\mathcal{D}$-classes	48
	2.3	Regular $\mathcal{D}$-classes	50
	2.4	Regular semigroups	54
	2.5	The sandwich set	58
	2.6	Exercises	60
	2.7	Notes	64
3	**0-simple semigroups**		66
	3.1	Simple and 0-simple semigroups; principal factors	66
	3.2	The Rees Theorem	69
	3.3	Completely simple semigroups	77
	3.4	Isomorphism and normalization	80
	3.5	Congruences on completely 0-simple semigroups*	83
	3.6	The lattice of congruences on a completely 0-simple semigroup*	91
	3.7	Finite congruence-free semigroups*	93
	3.8	Exercises	95
	3.9	Notes	101
4	**Completely regular semigroups**		102
	4.1	The Clifford decomposition	103
	4.2	Clifford semigroups	107
	4.3	Varieties	108
	4.4	Bands	113
	4.5	Free bands	119
	4.6	Varieties of bands*	124

	4.7	Exercises	138
	4.8	Notes	142
5	**Inverse semigroups**		144
	5.1	Preliminaries	145
	5.2	The natural order relation	152
	5.3	Congruences on inverse semigroups	154
	5.4	The Munn semigroup	162
	5.5	Anti-uniform semilattices	166
	5.6	Bisimple inverse semigroups	169
	5.7	Simple inverse semigroups	176
	5.8	Representations of inverse semigroups	185
	5.9	E-unitary inverse semigroups	192
	5.10	Free inverse monoids	200
	5.11	Exercises	211
	5.12	Notes	219
6	**Other classes of regular semigroups**		222
	6.1	Locally inverse semigroups	223
	6.2	Orthodox semigroups	226
	6.3	Semibands	230
	6.4	Exercises	235
	6.5	Notes	237
7	**Free semigroups**		238
	7.1	Properties of free semigroups	238
	7.2	Codes	243
	7.3	Exercises	248
	7.4	Notes	250
8	**Semigroup amalgams**		251
	8.1	Systems	252
	8.2	Free products	258
	8.3	Dominions and zigzags	266
	8.4	Direct limits, free extensions and free products	274
	8.5	The extension property	288
	8.6	Inverse semigroups and amalgamation	303
	8.7	Exercises	310
	8.8	Notes	315
	References		318
	List of symbols		341
	Index		345

1
Introductory ideas

In this chapter certain basic definitions and results are presented. Reference will be made to these throughout the book, though it should be noted that Section 1.8 is referred to only in Section 3.5.

1.1 BASIC DEFINITIONS

A *groupoid* (S, μ) is defined as a non-empty set S on which a *binary operation* μ—by which we mean a map $\mu : S \times S \to S$— is defined. We say that (S, μ) is a *semigroup* if the operation μ is *associative*, that is to say, if, for all x, y and z in S,

$$\big((x, y)\mu, z\big)\mu = \big(x, (y, z)\mu\big)\mu. \tag{1.1.1}$$

(Here, and throughout the book, we write mapping symbols on the *right*.) This notation is rather cumbersome, and we shall follow the usual algebraic practice of writing the binary operation as *multiplication*. Thus $(x, y)\mu$ becomes $x.y$ or (more usually) xy, and formula (1.1.1) takes the simple form

$$(xy)z = x(yz),$$

the familiar associative law of elementary algebra. Expressions such as xyz and $x_1x_2 \dots x_n$, where $x, y, z, x_1, x_2, \dots, x_n$ are elements of S, then have unambiguous meanings, and we can use the notation $x^n \quad (n \in \mathbf{N})$ to mean the product of n elements each equal to x. The cardinal number $|S|$—see Halmos (1960) for this and other items of basic set theory—will be called the *order* of S.

Where the semigroup is written multiplicatively and where the nature of the multiplication is clear from the context, we shall write simply S rather than $(S, .)$.

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. (The term *abelian* is also used, by analogy with the group theoretic term.) If a semigroup S contains an element 1 with the property that, for all x in S,

$$x1 = 1x = x,$$

we say that 1 is an *identity element* (or just an *identity*) of S, and that S is a *semigroup with identity* or (more usually) a *monoid*. A semigroup S has at most one identity element, since if $1'$ also has the property that $x1' = 1'x = x$ for all x in S, then

$$\begin{aligned} 1' &= 11' && \text{(since 1 is an identity)} \\ &= 1 && \text{(since } 1' \text{ is an identity).} \end{aligned}$$

If S has no identity element then it is very easy to adjoin an extra element 1 to S to form a monoid. We define

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

and it is a routine matter to check that $S \cup \{1\}$ becomes a monoid. We now define

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

We refer to S^1 as the *monoid obtained from S by adjoining an identity if necessary.*

If a semigroup S with at least two elements contains an element 0 such that, for all x in S,

$$0x = x0 = 0,$$

we say that 0 is a *zero element* (or just a *zero*) of S, and that S is a *semigroup with zero.* As with identity elements, it is a trivial matter to verify that there can be at most one such element in a semigroup. The proviso that S should have at least two elements means merely that we shall not want to refer to the single element of the *trivial* semigroup $\{e\}$ (in which $e^2 = e$) as a zero. (It *is* an identity!)

Again, if S has no zero it is easy to adjoin an extra element 0. We define

$$0s = s0 = 00 = 0 \text{ for all } s \text{ in } S,$$

and it is a routine matter to check that associativity survives in the extended set $S \cup \{0\}$. By analogy with the case of S^1, we define

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

and refer to S^0 as the *semigroup obtained from S by adjoining a zero if necessary.*

Despite the great ease with which we can adjoin an identity and a zero to a semigroup, we cannot altogether reduce the study of semigroups to the study of monoids with zero, for in adjoining the extra elements we may sacrifice some crucial property of the semigroup. To take a very trivial example, if we adjoin a zero element to a semigroup which is a group, we obtain a semigroup which is *not* a group.

Among semigroups with zero we find the very trivial *null* semigroups, in which the product of any two elements is zero. Only slightly less trivially, on any non-empty set S we can define a multiplication

$$ab = a \quad (a, b \in S),$$

and obtain what is called a *left zero semigroup. Right zero semigroups* are defined in an analogous way.

Another easy example arises if on the closed interval $I = [0, 1]$ we define

$$xy = \min(x, y) \quad (x, y \in I).$$

Associativity is clear, and it is easy to see that 0 is a zero element and 1 is an identity. Many other examples will emerge as we proceed.

If A and B are subsets of a semigroup S, then we write AB to mean $\{ab : a \in A,\ b \in B\}$. It is easy to verify that, for all subsets A, B, C of S,

$$(AB)C = A(BC);$$

hence once again notations such as ABC and $A_1A_2 \ldots A_n$ are meaningful. The usual hazard, namely that A^2 means $\{a_1a_2 : a_1, a_2 \in A\}$ rather than $\{a^2 : a \in A\}$, should be noted. When dealing with singleton sets we shall use the notational simplifications that are customary in algebra, writing (for example) Ab rather than $A\{b\}$.

If a is an element of a semigroup S without identity then Sa need not contain a. The following notations will be standard:

$$\left.\begin{aligned} S^1a &= Sa \cup \{a\}, \\ aS^1 &= aS \cup \{a\}, \\ S^1aS^1 &= SaS \cup Sa \cup aS \cup \{a\}. \end{aligned}\right\} \tag{1.1.2}$$

Notice that S^1a, aS^1 and S^1aS^1 are all subsets of S—they do not contain the element 1.

If a semigroup S has the property that

$$(\forall a \in S)\ aS = S \text{ and } Sa = S \tag{1.1.3}$$

we call it a *group*. This is not the commonest definition of a group, but it is an easy exercise to show that it is equivalent to the more usual definition of a group as a semigroup for which

$$\left.\begin{aligned} &(\exists e \in S)\ (\forall a \in S) && ea = a, \\ &(\forall a \in S)\ (\exists a^{-1} \in S) && a^{-1}a = e. \end{aligned}\right\} \tag{1.1.4}$$

The definition (1.1.3) is given first because it is the one that seems to occur most often in semigroup theory. It is clearly equivalent to

$$(\forall a, b \in S)\ (\exists x, y \in S)\ ax = b \text{ and } ya = b. \tag{1.1.5}$$

If G is a group, then $G^0 = G \cup \{0\}$ is a semigroup. We shall call a semigroup formed in this way a *0-group*, or *group-with-zero.*

Proposition 1.1.1 *A semigroup with zero is a* 0*-group if and only if*

$$(\forall a \in S \setminus \{0\})\ aS = S \text{ and } Sa = S.$$

Proof Suppose first that $S = G^0$, a 0-group, and let $a \in G = S \setminus \{0\}$. Certainly $aG = Ga = G$. Since $aS = aG \cup \{0\}$ and $Sa = Ga \cup \{0\}$, it follows that $aS = Sa = S$.

Conversely, suppose that S has the given property, and let $G = S \setminus \{0\}$. Since S by implication has more than two elements we have $G \neq \emptyset$. To show that G is a group we must first show that it is closed with respect to multiplication. So suppose, by way of contradiction, that there exist a, b in G such that $ab = 0$. Then

$$S^2 = (Sa)(bS) = S(ab)S = S0S = \{0\},$$

and so $S = aS \subseteq S^2 = \{0\}$. This is a contradiction, and so G has the desired closure property. The assumed property implies that for all a, b in G there exist x, y in S such that $ax = b$ and $ya = b$. The elements x and y cannot be zero and so are in G. Thus G satisfies (1.1.5) and so is a group. □

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

$$(\forall x, y \in T) \quad xy \in T \tag{1.1.6}$$

—a condition that can be expressed more compactly as $T^2 \subseteq T$. The associativity condition that holds throughout S certainly holds throughout T, and so T is itself a semigroup. Among special subsemigroups worth mentioning are S itself, $\{0\}$ and $\{1\}$ where appropriate, and also more generally $\{e\}$, where e is any element of S that is *idempotent*, that is to say, for which $e^2 = e$.

A subsemigroup of S which is a group with respect to the multiplication inherited from S will be called a *subgroup* of S; the one-element subsemigroups $\{0\}$, $\{1\}$ and $\{e\}$ mentioned in the last paragraph are all trivial examples. It is not hard to see that a non-empty subset T of S is a subgroup of S if and only if

$$(\forall a \in T)\ aT = T \text{ and } Ta = T. \tag{1.1.7}$$

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. Evidently every ideal (whether right, left or two-sided) is a subsemigroup, but the converse is not the case. Among the ideals are S itself and (if S has a zero element) $\{0\}$. An ideal I such that $\{0\} \subset I \subset S$ (strictly) is called *proper*.

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

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

If $(S, ., 1_S)$ and $(T, ., 1_T)$ are *monoids*, with identity elements 1_S, 1_T respectively, then ϕ will be called a morphism only if we have the additional property

$$1_S\phi = 1_T.$$

There is a possibility of confusion here, and if there seems any doubt we shall distinguish between a *semigroup morphism* and a *monoid morphism*.

In either event we refer to S as the *domain* of ϕ and T as the *codomain*. The *image* (or *range*) of ϕ is defined as $\{s\phi : s \in S\}$. If ϕ is one-one we shall call it a *monomorphism*; this definition is equivalent to the 'categorical' definition of a monomorphism as a right cancellative morphism, that is, $\phi : S \to T$ is a *monomorphism* if, for all semigroups U and for all morphisms $\alpha, \beta : U \to S$,

$$\alpha\phi = \beta\phi \;\Rightarrow\; \alpha = \beta.$$

(See Mitchell (1965).)

A morphism $\phi : S \to T$ is called an *isomorphism* if it is invertible, that is to say, if there exists a morphism $\phi^{-1} : T \to S$ such that $\phi\phi^{-1}$ is the identity map of S and $\phi^{-1}\phi$ is the identity map of T. It is not hard to show that a morphism $\phi : S \to T$ is an isomorphism if and only if it is bijective. If there exists an isomorphism $\phi : S \to T$ we say that S and T are *isomorphic*, and write $S \simeq T$.

A morphism ϕ from S into S is called an *endomorphism* of S, and if it is one-one and onto it is called an *automorphism*.

If S and T are semigroups, then the cartesian product $S \times T$ becomes a semigroup if we define

$$(s, t)(s', t') = (ss', tt').$$

We refer to this semigroup as the *direct product* of S and T.

We shall have occasion in Chapter 4 to make use of a more general notion of direct product. Let $\{S_i : i \in I\}$ be a family of semigroups indexed by the set I. Then P, the direct product of the family, is defined as the set of all maps $p : I \to \bigcup_{i \in I} S_i$ such that $ip \in S_i$ for each i in I. If we define the multiplication of two elements p, q in P by the 'componentwise' rule

$$i(pq) = (ip)(iq) \quad (i \in I),$$

then P becomes a semigroup. For each i in I there is a *projection morphism* π_i from P onto S_i given by the rule that

$$p\pi_i = ip \quad (p \in P).$$

Moreover, if T is a semigroup and if there are morphisms $\tau_i : T \to S_i$ $(i \in I)$, then there is a unique morphism $\gamma : T \to P$ with the property that $\gamma\pi_i = \tau_i$ for every i in I. The map γ is defined by the rule that, for every t in T,

$$(t\gamma)(i) = t\tau_i \quad (i \in I).$$

This is expressed in category theory by saying that P is the *product* of the semigroups S_i. (See Mitchell (1965).)

If $I = \{1, 2\}$, then the direct product P essentially coincides with the direct product $S_1 \times S_2$ as previously defined, for the map $p \mapsto (1p, 2p)$ from P onto $S_1 \times S_2$ is an isomorphism. More generally, if $I = \{1, 2, \ldots, n\}$, then in the same way we may identify the map p with the n-tuple $(1p, 2p, \ldots, np)$, and think of P as consisting of all n-tuples $(x_1, x_2, \ldots, x_n)$, in which $x_i \in S_i$ $(i = 1, 2, \ldots, n)$, with multiplication given by the componentwise formula

$$(x_1, x_2, \ldots, x_n)(y_1, y_2, \ldots, y_n) = (x_1y_1, x_2y_2, \ldots, x_ny_n).$$

Just as groups arise most naturally as groups of permutations of a set, so semigroups arise from more general mappings of a set into itself. The analogue of the *symmetric group* $(\mathcal{G}_X, \circ)$ of all permutations of a set X is the *full transformation semigroup* $(\mathcal{T}_X, \circ)$ consisting of all maps from X into X. The operation in both cases is composition of maps, sometimes written $\circ$, but often just written multiplicatively: if α and β are maps from X into X, then

$$x(\alpha \circ \beta) \ \big(= x(\alpha\beta)\big) = (x\alpha)\beta \quad (x \in X).$$

It is clear that $\mathcal{G}_X$, consisting of all *bijections* from X onto X, is a sub*group* of $\mathcal{T}_X$. Simple combinatorial arguments show that, if $|X| = n$, then

$$|\mathcal{G}_X| = n!, \quad |\mathcal{T}_X| = n^n.$$

If a semigroup S is, for some X, a subsemigroup of $\mathcal{T}_X$, we say that S is a *semigroup of maps*, or a *transformation semigroup*. A morphism ϕ from a semigroup S into some $\mathcal{T}_X$ is called a *representation of* S (*by maps*), and ϕ is called a *faithful* representation if it is one-one. In such a case the image $S\phi$ of ϕ is a transformation semigroup isomorphic to S.

The following theorem, closely analogous to Cayley's Theorem for groups—see M. Hall (1959) for this and other items of elementary group theory—shows that every semigroup is isomorphic to a transformation semigroup:

Theorem 1.1.2 *If S is a semigroup and $X = S^1$ then there is a faithful representation $\phi : S \to \mathcal{T}_X$.*

Proof For each a in S, define a map $\rho_a : S^1 \to S^1$ by

$$x\rho_a = xa \quad (x \in S^1).$$

Thus $\rho_a \in \mathcal{T}_X$, and so there is a map $\alpha : S \to \mathcal{T}_X$ given by

$$a\alpha = \rho_a \quad (a \in S).$$

The map α is one-one, since, for all a, b in S,

$$\begin{aligned} a\alpha = b\alpha \;&\Rightarrow\; \rho_a = \rho_b \;\Rightarrow\; xa = xb \text{ for all } x \text{ in } S^1 \\ &\Rightarrow\; 1a = 1b \;\Rightarrow\; a = b. \end{aligned}$$

(Notice that the argument might break down at this point if S rather than S^1 were used.) Also, α is a morphism, since, for all x in S^1,

$$x(\rho_a\rho_b) = (x\rho_a)\rho_b = (xa)b = x(ab) = x\rho_{ab},$$

and so $(a\alpha)(b\alpha) = (ab)\alpha$. □

The representation α introduced in this proof is called the *extended right regular representation* of S. The word 'extended' here signifies that S^1 is used for the set X rather than (as in group theory) the more obvious S. This is necessary to ensure the faithfulness of the representation. (See Exercise 6.)

To rescue this section from complete triviality, and also to underline the point that there are interesting semigroups that are totally different from groups, we close with a somewhat more substantial theorem. We say that a semigroup S is a *rectangular band* if $aba = a$ for all a, b in S. Then we have

Theorem 1.1.3 *Let S be a semigroup. Then the following conditions are equivalent:*

(1) *S is a rectangular band;*
(2) *every element of S is idempotent, and $abc = ac$ for all a, b, c in S;*
(3) *there exist a left zero semigroup L and a right zero semigroup R such that $S \simeq L \times R$;*
(4) *S is isomorphic to a semigroup of the form $A \times B$, where A and B are non-empty sets, and where multiplication is given by*

$$(a_1, b_1)(a_2, b_2) = (a_1, b_2).$$

Proof (1) $\Rightarrow$ (2). Let $a \in S$. Then by (1) we have $a^3 = a$, and so $a^4 = a^2$. Again by (1) we have $a = a(a^2)a = a^4$. Hence $a^2 = a$ as required.

Now let $a, b, c \in S$. From (1) we have $a = aba$, $c = cbc$ and $b = b(ac)b$. Hence

$$ac = (aba)(cbc) = a(bacb)c = abc,$$

as required.

(2) $\Rightarrow$ (3). Choose and fix an element c of S. Let $L = Sc$ and $R = cS$. Then, using (2), we see that, for all $x = zc$ and $y = tc$ in L,

$$xy = zctc = zc^2 = zc = x,$$

and so L is a left zero semigroup. Similarly, R is a right zero semigroup. Define $\phi : S \to L \times R$ by

$$x\phi = (xc, cx) \quad (x \in S).$$

Then ϕ is one-one, for if $(xc, cx) = (yc, cy)$ then

$$\begin{aligned} x = x^2 &= xcx \text{ by (2)} \\ &= ycx = ycy = y^2 = y. \end{aligned}$$

Also, ϕ is onto, since for all (ac, cb) in $L \times R$, we may use condition (2) to see that

$$(ac, cb) = (abc, cab) = (ab)\phi.$$

Finally, ϕ is a morphism, since, for all x, y in S,

$$(xy)\phi = (xyc, cxy) = (xc, cy) = (xcyc, cxcy) = (xc, cx)(yc, cy) = (x\phi)(y\phi).$$

(3) $\Rightarrow$ (4). Suppose that $S = L \times R$, where L is a left zero semigroup and Y is a right zero semigroup. Then the product of two elements (a, b) and (c, d) in S is given by

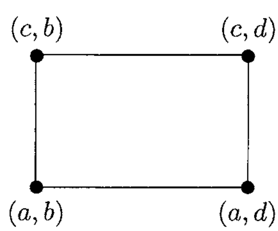
$$(a, b)(c, d) = (ac, bd) = (a, d).$$

Thus we need only take $A = L$ and $B = R$.

(4) $\Rightarrow$ (1). Let $S = A \times B$, with the given multiplication. Then for all $a = (x, y)$ and $b = (z, t)$ in S,

$$aba = (x, y)(z, t)(x, y) = (x, t)(x, y) = (x, y) = a. \qquad \square$$

The term 'rectangular' comes from the property (4). If we think of (a, b) and (c, d) as points in the cartesian plane, then the products $(a, b)(c, d)$ and $(c, d)(a, b)$ are placed at the vertices of the rectangle



The term 'band' is in general use for a semigroup consisting of idempotents.

1.2 MONOGENIC SEMIGROUPS

Let S be a semigroup, and let $\{U_i : i \in I\}$ (with $I \neq \emptyset$) be an indexed family of subsemigroups of S. It is easy to see that the intersection U of all the subsemigroups U_i, if non-empty, is again a subsemigroup of S. For every non-empty subset A of S there is at least one subsemigroup of S containing A, namely S itself. Hence the intersection of all the subsemigroups of S containing A is a subsemigroup of S containing A. We denote it by $\langle A \rangle$, and note that it is a subsemigroup defined by the two properties: (1) $A \subseteq \langle A \rangle$; and (2) if U is a subsemigroup of S containing A, then $\langle A \rangle \subseteq U$.

The subsemigroup $\langle A \rangle$ consists of all elements of S that can be expressed as finite products of elements in A. If $\langle A \rangle = S$ we say that A is a *set of generators*, or a *generating set*, of S.

Particular interest attaches to the case where A is finite. If $A = \{a_1, a_2, \ldots, a_n\}$ then we shall write $\langle A \rangle$ as $\langle a_1, a_2, \ldots, a_n \rangle$. Especially interesting is the case where $A = \{a\}$, a singleton set, when

$$\langle a \rangle = \{a, a^2, a^3, \ldots\}.$$

At this point it is worth pausing to note that if S is a monoid then we can equally well talk of the sub*monoid* of S generated by S. This will always contain 1, and in the case of a singleton generator we find that

$$\langle a \rangle = \{1, a, a^2, a^3, \ldots\}.$$

In what follows, however, it will be sufficient to consider the semigroup case.

We refer to $\langle a \rangle$ as the *monogenic* subsemigroup of S generated by the element a. The *order* of the element a is defined, as in group theory, as the order of the subsemigroup $\langle a \rangle$. If S is a semigroup in which there exists an element a such that $S = \langle a \rangle$, then S is said to be a *monogenic* semigroup.

Clifford and Preston (1961) followed the group-theoretic terminology, and referred to semigroups with one generator as 'cyclic'. From what follows, the reader may judge whether monogenic semigroups are 'round' enough to merit the description 'cyclic.'

Let a be an element of a semigroup S, and consider the monogenic subsemigroup

$$\langle a \rangle = \{a, a^2, a^3, \ldots\}$$

generated by a. If there are no repetitions in the list $a, a^2, a^3, \ldots$, that is, if

$$a^m = a^n \;\Rightarrow\; m = n,$$

then evidently $(\langle a \rangle, .)$ is isomorphic to the semigroup $(\mathbf{N}, +)$ of natural numbers with respect to addition. In such a case we say that a is an *infinite* monogenic semigroup, and that a has *infinite order* in S.

Suppose now that there are repetitions among the powers of a. Then the set

$$\{x \in \mathbf{N} : (\exists y \in N)\; a^x = a^y,\; x \neq y\}$$

is non-empty and so has a least element. Let us denote this least element by m and call it the *index* of the element a. Then the set

$$\{x \in \mathbf{N} : a^{m+x} = a^m\}$$

is non-empty, and so it too has a least element r, which we call the *period* of a. We shall also refer to m and r as the index and period, respectively, of the monogenic semigroup $\langle a \rangle$.

Let a be an element with index m and period r. Thus

$$a^m = a^{m+r}. \tag{1.2.1}$$

It follows that

$$a^m = a^{m+r} = a^m a^r = a^{m+r} a^r = a^{m+2r},$$

and, more generally, that

$$(\forall q \in \mathbf{N})\ a^m = a^{m+qr}.$$

By the minimality of m and r in (1.2.1) we may deduce that the powers

$$a, a^2, \dots, a^m, a^{m+1}, \dots, a^{m+r-1}$$

are all distinct. For every $s \geq m$ we can, by the division algorithm, write $s = m + qr + u$, where $q \geq 0$ and $0 \leq u \leq r-1$. It then follows that

$$a^s = a^{m+qr} a^u = a^m a^u = a^{m+u};$$

thus

$$\langle a \rangle = \{a, a^2, \dots, a^{m+r-1}\}, \text{ and } |\langle a \rangle| = m + r - 1.$$

We say that a has *finite order* in this case; the order is given by the rule

$$\text{order of } a = (\text{index of } a) + (\text{period of } a) - 1.$$

The subset $K_a = \{a^m, a^m + 1, \dots, a^{m+r-1}\}$ of $\langle a \rangle$ is a subsemigroup, indeed an ideal, of $\langle a \rangle$. We call it the *kernel* of $\langle a \rangle$, and we shall see in due course that this use of the word does not conflict with the more general use of 'kernel' in Chapter 3. In fact K_a is a *subgroup* of $\langle a \rangle$, for if a^{m+u} and a^{m+v} are elements of K_a, then we can find an element a^{m+x} in K_a for which

$$a^{m+u} a^{m+x} = a^{m+v}$$

simply by choosing x so that

$$x \equiv v - u - m \pmod{r} \ \text{ and } \ 0 \leq x \leq r - 1.$$

Indeed K_a is a cyclic group. To see this, notice that the integers

$$m,\ m+1,\ \dots,\ m+r-1$$

form a complete set of incongruent residues modulo r. (For this and other elementary number-theoretic ideas see, for example, Hardy and Wright (1979).) It follows that there exists g such that

$$0 \le g \le r-1 \quad \text{and} \quad m+g \equiv 1 \pmod r. \tag{1.2.2}$$

Hence $k(m+g) \equiv k \pmod r$ for every k in $\mathbf{N}$, and so the powers $(a^{m+g})^k$ of a^{m+g}, for $k = 1, 2, \ldots, r$, exhaust K_a. Thus K_a is a cyclic group of order r, generated by the element a^{m+g}.

If we choose z so that

$$0 \le z \le r-1 \quad \text{and} \quad m+z \equiv 0 \pmod r, \tag{1.2.3}$$

then a^{m+z} is idempotent, and so it is the identity of the group K_a.

Example 1.2.1 Let $X = \{1, 2, \ldots, 7\}$, and consider the element

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & 7 & 5 \end{pmatrix}$$

of $\mathcal{T}_X$. (The notation for α is an obvious generalization of the standard notation for permutations: the import is that $1\alpha = 2$, $2\alpha = 3$, ..., $6\alpha = 7$, $7\alpha = 5$.) It is easy to calculate that

$$\alpha^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 5 & 6 \end{pmatrix}, \quad \alpha^3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 5 & 6 & 7 \end{pmatrix},$$

$$\alpha^4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 5 & 6 & 7 & 5 \end{pmatrix}, \quad \alpha^5 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 6 & 7 & 5 & 6 & 7 & 5 & 6 \end{pmatrix},$$

$$\alpha^6 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 5 & 6 & 7 & 5 & 6 & 7 \end{pmatrix}, \quad \alpha^7 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 5 & 6 & 7 & 5 \end{pmatrix},$$

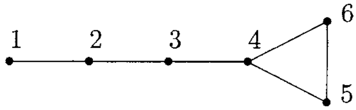
and so α has index 4 and period 3. The kernel K_α is equal to $\{\alpha^4, \alpha^5, \alpha^6\}$, and has Cayley table

	α^4	α^5	α^6
α^4	α^5	α^6	α^4
α^5	α^6	α^4	α^5
α^6	α^4	α^5	α^6

Thus α^6 is the identity of K_a, in accord with formula (1.2.3), since $6 \equiv 0 \pmod 3$. Also, in accord with formula (1.2.2), since $4 \equiv 1 \pmod 3$, a suitable generator of the cyclic group K_a is 4:

$$(\alpha^4)^2 = \alpha^5, \quad (\alpha^4)^3 = \alpha^6.$$

We can visualize $\langle \alpha \rangle$ as



It is useful to summarize the results in a theorem:

Theorem 1.2.2 *Let a be an element of a semigroup S. Then either:*

(1) *all powers of a are distinct, and the monogenic subsemigroup $\langle a \rangle$ of S is isomorphic to the semigroup $(\mathbf{N}, +)$ of natural numbers under addition; or*

(2) *there exist positive integers m (the* index *of a) and r (the* period *of a) with the following properties:*

   (a) $a^m = a^{m+r}$;
   (b) *for all u, v in $\mathbf{N}^0$, $a^{m+u} = a^{m+v}$ if and only if $u \equiv v \pmod r$;*
   (c) $\langle a \rangle = \{a, a^2, \ldots, a^{m+r-1}\}$;
   (d) $K_a = \{a^m, a^{m+1}, \ldots, a^{m+r-1}\}$ *is a cyclic subgroup of $\langle a \rangle$.* □

Nothing that we have said so far makes it clear that for every pair (m, r) of positive integers there does in fact exist a semigroup S containing an element a of index m and period r. This, however, is the case: it is a routine matter to verify that the element

$$a = \begin{pmatrix} 1\,2\,3 \ldots & m & m+1 \ldots & m+r-1 & m+r \\ 2\,3\,4 \ldots & m+1 & m+2 \ldots & m+r & m+1 \end{pmatrix}$$

of the semigroup $\mathcal{T}_{\{1,2,\ldots,m+r\}}$ has index m and period r.

It is easy to see that if a and b are elements of finite order in the same or in different semigroups, then $\langle a \rangle \simeq \langle b \rangle$ if and only if a and b have the same index and period. The conclusion is that for each (m, r) in $\mathbf{N} \times \mathbf{N}$ there is, up to isomorphism, exactly one monogenic semigroup with index m and period r. We shall feel free to talk of *the monogenic semigroup $M(m, r)$* with index m and period r. Notice that $M(1, r)$ is the cyclic group of order r.

A semigroup is called *periodic* if all its elements are of finite order. A finite semigroup is necessarily periodic.

Proposition 1.2.3 *In a periodic semigroup every element has a power which is idempotent. Hence in every periodic semigroup—in particular, in every finite semigroup—there is at least one idempotent.*

Proof If $a \in S$, a periodic semigroup, then $\langle a \rangle$ is finite, and so some power a^n of a serves as the identity of the group K_a. The result follows. □

If the hypothesis of periodicity is dropped then we can no longer guarantee the existence of an idempotent. The semigroup $(\mathbf{N}, +)$ is an obvious example.

1.3 ORDERED SETS, SEMILATTICES AND LATTICES

A binary relation ω on a set X (that is, a subset ω of $X \times X$) is called a (partial) *order* if

(O1) $(x, x) \in \omega$ for all x in X—that is, ω is *reflexive*;

(O2) $(\forall x, y \in X)\ (x, y) \in \omega$ and $(y, x) \in \omega \;\Rightarrow\; x = y$—that is, ω is *antisymmetric*;

(O3) $(\forall x, y, z \in X)\ (x, y) \in \omega$ and $(y, z) \in \omega \;\Rightarrow\; (x, z) \in \omega$—that is, ω is *transitive*.

Traditionally one writes $x \leq y$ rather than $(x, y) \in \omega$. We shall follow this convention, and also write $x \geq y$, $x < y$ and $x > y$ to mean (respectively) $(y, x) \in \omega$, $(x, y) \in \omega$ and $x \neq y$, and $(y, x) \in \omega$ and $x \neq y$.

A partial order having the extra property

(O4) $(\forall x, y \in X)\ x \leq y$ or $y \leq x$

will be called a *total* order. We shall refer to $(X, \leq)$, or just to X, as a *(partially) ordered set*, or (where appropriate) a *totally ordered set.*

Let Y be a non-empty subset of a partially ordered set $(X, \leq)$. An element a of Y is called *minimal* if there is no element of Y that is strictly less than a, that is to say, if

$$(\forall y \in Y)\ y \leq a \;\Rightarrow\; y = a.$$

An element b of Y is called *minimum* if

$$(\forall y \in Y)\ b \leq y.$$

Evidently a minimum element is minimal, but in a partially ordered set it is perfectly possible to have minimal elements that are not minimum. The following elementary facts are easily verified:

Proposition 1.3.1 *Let Y be a non-empty subset of a partially ordered set X. Then*

(1) *Y has at most one minimum element;*

(2) *if Y is totally ordered, then the terms 'minimal' and 'minimum' are equivalent.* □

We shall say that $(X, \leq)$ *satisfies the minimal condition* if every non-empty subset of X has a minimal element. A totally ordered set X satisfying the minimal condition is said to be *well-ordered.*

We leave it to the reader to provide the analogous definitions of *maximal, maximum* and the *maximal condition.*

If Y is a non-empty subset of $(X, \leq)$, we say that an element c of X is a *lower bound* of Y if $c \leq y$ for every y in Y. If the set of lower bounds of Y

is non-empty and has a maximum element d, we say that d is the *greatest lower bound*, or *meet*, of Y. The element d is unique if it exists; we write

$$d = \bigwedge\{y : y \in Y\}.$$

If $Y = \{a, b\}$ then we write $d = a \wedge b$.

If $(X, \leq)$ is such that $a \wedge b$ exists for all a, b in X, then we say that $(X, \leq)$ is a *lower semilattice*. If we have the stronger property that $\bigwedge\{y : y \in Y\}$ exists for every non-empty subset Y of X, then we say that $(X, \leq)$ is a *complete lower semilattice*. In a lower semilattice $(X, \leq)$ we have that, for all a, b in X,

$$a \leq b \text{ if and only if } a \wedge b = a. \tag{1.3.1}$$

Analogous definitions are easily given for the *least upper bound*, or *join*

$$\bigvee\{y : y \in Y\},$$

for $a \vee b$, for an *upper semilattice* and for a *complete upper semilattice*. If $(X, \leq)$ is both a (complete) upper semilattice and a (complete) lower semilattice we call it a (*complete*) *lattice*. In these circumstances we may wish to emphasize the lattice structure by writing $X = (X, \leq, \wedge, \vee)$. By a *sublattice* of X we shall then mean a non-empty subset Y of X such that

$$a, b \in Y \;\Rightarrow\; a \wedge b, a \vee b \in Y.$$

Let $(E, \leq)$ be a lower semilattice. Then it is not hard to verify, for a, b and c in E, that both $(a \wedge b) \wedge c$ and $a \wedge (b \wedge c)$ are greatest lower bounds of $\{a, b, c\}$, and we deduce that

$$(a \wedge b) \wedge c = a \wedge (b \wedge c).$$

Thus $(E, \wedge)$ is a semigroup. Since it is obvious that $a \wedge a = a$ for every a in E, and that $a \wedge b = b \wedge a$ for all a, b in E, we now invoke formula (1.3.1) and observe that we have proved half of the following proposition:

Proposition 1.3.2 *Let $(E, \leq)$ be a lower semilattice. Then $(E, \wedge)$ is a commutative semigroup consisting entirely of idempotents, and*

$$(\forall a, b \in E)\; a \leq b \textit{ if and only if } a \wedge b = a. \tag{1.3.2}$$

Conversely, suppose that $(E, .)$ is a commutative semigroup of idempotents. Then the relation $\leq$ on E defined by

$$a \leq b \textit{ if and only if } ab = a$$

is a partial order on E, with respect to which $(E, \leq)$ is a lower semilattice. In $(E, \leq)$, the meet of a and b is their product ab.

Proof Let $(E, .)$ be a commutative semigroup of idempotents, and let $\leq$ be defined by (1.3.2). Since $a^2 = a$ we have immediately that $a \leq a$ for every a in E. Suppose now that $a \leq b$ and $b \leq a$; then $ab = a$ and $ba = b$ and so

$$a = ab = ba = b.$$

Next, suppose that $a \leq b$ and $b \leq c$, so that $ab = a$, $bc = b$. Then

$$ac = (ab)c = a(bc) = ab = a,$$

and so $a \leq c$. We have shown that $\leq$ is a partial order.

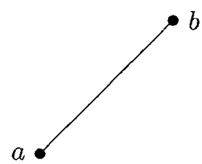
Since $a(ab) = a^2b = ab$ and $b(ab) = ab^2 = ab$, we have that $ab \leq a$, $ab \leq b$. If $c \leq a$ and $c \leq b$ then

$$c(ab) = (ca)b = cb = c,$$

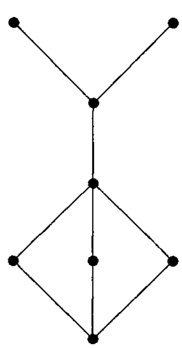
and so $c \leq ab$. It follows that ab is the unique greatest lower bound—the *meet*, as we have called it—of a and b. □

The effect of this proposition is that the notions of 'lower semilattice' and 'commutative semigroup of idempotents' are equivalent and interchangeable. We shall use the term 'semilattice' with either meaning, making free and frequent transfers between the semigroup and the ordered set points of view.

When describing an ordered set $(X, \leq)$, particularly when X is finite, we shall sometimes use so-called *Hasse diagrams*. In such a diagram, elements of the set are represented by small black circles, and two elements a and b in X for which $a < b$ and for which there is no x in X such that $a < x < b$ are depicted thus:



That is, b appears above a, and there is a line connecting the two. We thus build up diagrams such as



which we can label if necessary. Notice that the given diagram represents a semilattice.

1.4 BINARY RELATIONS; EQUIVALENCES

In our discussion of ordered sets we have already encountered the idea of a (*binary*) *relation* on a set X, by which we simply mean a subset ρ of the cartesian product $X \times X$. At this stage it is convenient to develop the theory of relations in a somewhat more general and abstract way. Intuitively we think of elements x and y for which $(x,y) \in \rho$ as being *related*, and we frequently prefer to write $x \ \rho \ y$ instead of $(x,y) \in \rho$. The empty subset $\emptyset$ of $X \times X$ is included among the binary relations on X; other special relations worthy of mention are the *universal* relation $X \times X$, in which everything is related to everything else, and the *equality* relation

$$1_X = \{(x,x) : x \in X\}, \tag{1.4.1}$$

also known as the *diagonal* relation, in which two elements are related if and only if they are equal.

Let us denote the set of all binary relations on X by $\mathcal{B}_X$. A binary operation $\circ$ is defined on $\mathcal{B}_X$ by the rule that, for all ρ, σ in $\mathcal{B}_X$,

$$\rho \circ \sigma = \{(x,y) \in X \times X : (\exists z \in X)\ (x,z) \in \rho \ \text{ and } \ (z,y) \in \sigma\}. \tag{1.4.2}$$

It is easy to see that, for all ρ, σ, τ in $\mathcal{B}_X$,

$$\rho \subseteq \sigma \ \Rightarrow \ \rho \circ \tau \subseteq \sigma \circ \tau \text{ and } \tau \circ \rho \subseteq \tau \circ \sigma. \tag{1.4.3}$$

It is easy to see also that the operation $\circ$ is associative: for all ρ, σ, τ in $\mathcal{B}_X$,

$$(\rho \circ \sigma) \circ \tau = \rho \circ (\sigma \circ \tau),$$

for

$$\begin{aligned}
(x,y) \quad \in \quad & (\rho \circ \sigma) \circ \tau \\
\iff \quad & (\exists z \in X)\ (x,z) \in \rho \circ \sigma \text{ and } (z,y) \in \tau, \\
\iff \quad & (\exists z \in X)(\exists u \in X)\ (x,u) \in \rho,\ (u,z) \in \sigma \text{ and } (z,y) \in \tau, \\
\iff \quad & (\exists u \in X)\ (x,u) \in \rho \text{ and } (u,y) \in \sigma \circ \tau, \\
\iff \quad & (x,y) \in \rho \circ (\sigma \circ \tau).
\end{aligned}$$

We have proved

Proposition 1.4.1 *Let $\mathcal{B}_X$ be the set of all binary relations on a set X, and define the operation $\circ$ on $\mathcal{B}_X$ by* (1.4.2). *Then $(\mathcal{B}_X, \circ)$ is a semigroup.*

□

Whilst we shall not normally revert to simple multiplicative notation when discussing the semigroup $(\mathcal{B}_X, \circ)$, we shall allow ourselves to write ρ^2, ρ^3, etc., instead of $\rho \circ \rho$, $\rho \circ \rho \circ \rho$, etc.

For each $\rho \in \mathcal{B}_X$ we define the *domain* $\operatorname{dom}\rho$ by

$$\operatorname{dom}\rho = \{x \in X : (\exists y \in X)\ (x,y) \in \rho\}, \tag{1.4.4}$$

and the *image* $\operatorname{im}\rho$ by

$$\operatorname{im}\rho = \{y \in X : (\exists x \in X)\ (x,y) \in \rho\}. \tag{1.4.5}$$

It is immediate that, for all ρ, σ in $\mathcal{B}_X$,

$$\rho \subseteq \sigma \ \Rightarrow\ \operatorname{dom}\rho \subseteq \operatorname{dom}\sigma \text{ and } \operatorname{im}\rho \subseteq \operatorname{im}\sigma. \tag{1.4.6}$$

For each x in X and ρ in $\mathcal{B}_X$ we define a subset $x\rho$ of X by

$$x\rho = \{y \in X : (x,y) \in \rho\}; \tag{1.4.7}$$

thus $x\rho \neq \emptyset$ if and only if $x \in \operatorname{dom}\rho$. If A is a subset of X we define

$$A\rho = \bigcup\{a\rho : a \in A\}. \tag{1.4.8}$$

For each ρ in $\mathcal{B}_X$, we define ρ^{-1}, the *converse* of ρ, by

$$\rho^{-1} = \{(x,y) \in X \times X : (y,x) \in \rho\}. \tag{1.4.9}$$

Certainly $\rho^{-1} \in \mathcal{B}_X$, and it is easy to see that, for all ρ, σ, $\rho_1, \rho_2, \ldots, \rho_n$ in $\mathcal{B}_X$,

$$(\rho^{-1})^{-1} = \rho, \tag{1.4.10}$$

$$(\rho_1 \circ \rho_2 \circ \cdots \circ \rho_n)^{-1} = \rho_n^{-1} \circ \cdots \circ \rho_2^{-1} \circ \rho_1^{-1}, \tag{1.4.11}$$

$$\rho \subseteq \sigma \ \Rightarrow\ \rho^{-1} \subseteq \sigma^{-1}. \tag{1.4.12}$$

Notice also that

$$\operatorname{dom}(\rho^{-1}) = \operatorname{im}\rho, \quad \operatorname{im}(\rho^{-1}) = \operatorname{dom}\rho, \tag{1.4.13}$$

and that

$$x\rho^{-1} \neq \emptyset \text{ if and only if } x \in \operatorname{im}\rho.$$

An element ϕ of $\mathcal{B}_X$ is called a *partial map* of X if $|x\phi| = 1$ for all x in $\operatorname{dom}\phi$, that is, if, for all x, y_1, y_2 in X,

$$[(x,y_1) \in \phi \text{ and } (x,y_2) \in \phi] \ \Rightarrow\ y_1 = y_2. \tag{1.4.14}$$

It will not conflict at all with (1.4.7) if we decide in such a case to let $x\phi$ denote the unique element y such that $(x,y) \in \phi$ (rather than the set consisting of that element). Notice that the condition (1.4.14) is fulfilled (vacuously) by the empty relation $\emptyset$, which is therefore included among the partial maps.

If ϕ, ψ are partial maps of X such that $\phi \subseteq \psi$, we sometimes say that ϕ is a *restriction* of ψ, or that ψ is an *extension* of ϕ. If, in these circumstances, $\operatorname{dom}\phi = A \subset \operatorname{dom}\psi$, then we denote ϕ by $\psi|_A$ (*ψ restricted to A*).

Proposition 1.4.2 *The subset $\mathcal{P}_X$ of $\mathcal{B}_X$ consisting of all partial maps of X is a subsemigroup of $\mathcal{B}_X$.*

Proof Let $\phi, \psi \in \mathcal{P}_X$, and suppose that $(x, y_1), (x, y_2) \in \phi \circ \psi$. Then there exist z_1, z_2 in X such that

$$(x, z_1) \in \phi,\ (z_1, y_1) \in \psi,\ (x, z_2) \in \phi,\ (z_2, y_2) \in \psi.$$

The condition (1.4.14) on ϕ implies that $z_1 = z_2$, and then the same condition on ψ implies that $y_1 = y_2$. Thus $\phi \circ \psi \in \mathcal{P}_X$. □

It is important to note that the converse ϕ^{-1} of a partial map ϕ need not be a partial map. For example, if $X = \{1, 2\}$, then $\phi = \{(1,1), (2,1)\}$ is a partial map, but ϕ^{-1} is not.

In view of Proposition 1.4.2 we can talk of $(\mathcal{P}_X, \circ)$ as the *semigroup of all partial maps of* X. The composition law $\circ$ in this semigroup is in fact a fairly natural composition law for partial maps:

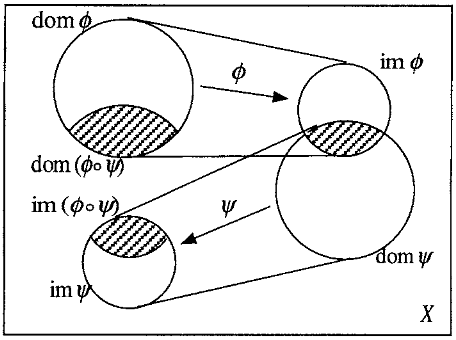
Proposition 1.4.3 *If $\phi, \psi \in \mathcal{P}_X$, then*

$$\operatorname{dom}(\phi \circ \psi) = [\operatorname{im} \phi \cap \operatorname{dom} \psi]\phi^{-1},$$
$$\operatorname{im}(\phi \circ \psi) = [\operatorname{im} \phi \cap \operatorname{dom} \psi]\psi,$$

and

$$(\forall x \in \operatorname{dom}(\phi \circ \psi)) \quad x(\phi \circ \psi) = (x\phi)\psi.$$

Proof Before proving this, we illustrate it in a diagram as follows:



Suppose first that $x \in \operatorname{dom}(\phi \circ \psi)$. Then there exist y and z in X such that $(x, z) \in \phi$, $(z, y) \in \psi$. Clearly $z \in \operatorname{im} \phi \cap \operatorname{dom} \psi$, and $(z, x) \in \phi^{-1}$. Hence

$$x \in z\phi^{-1} \subseteq [\operatorname{im} \phi \cap \operatorname{dom} \psi]\phi^{-1}.$$

Conversely, suppose that $x \in [\operatorname{im} \phi \cap \operatorname{dom} \psi]\phi^{-1}$. Then there exists z in $\operatorname{im} \phi \cap \operatorname{dom} \psi$ such that $x \in z\phi^{-1}$, that is, such that $(x, z) \in \phi$. Since $z \in \operatorname{dom} \psi$, there exists y in X such that $(z, y) \in \psi$. Hence $(x, y) \in \phi \circ \psi$ and so $x \in \operatorname{dom}(\phi \circ \psi)$. Thus

$$\operatorname{dom}(\phi \circ \psi) = [\operatorname{im} \phi \cap \operatorname{dom} \psi]\phi^{-1},$$

as required. The proof that $\text{im}(\phi \circ \psi) = [\text{im}\,\phi \cap \text{dom}\,\psi]\psi$ is similar. (Notice that as yet we have used no special properties of partial maps. The characterizations established for $\text{dom}(\phi \circ \psi)$ and $\text{im}(\phi \circ \psi)$ apply equally well to arbitrary relations.)

To complete the proof, notice that $(x, y) \in \phi \circ \psi$ if and only if there exists z in X such that $(x, z) \in \phi$ and $(z, y) \in \psi$. Since ϕ, ψ and $\phi \circ \psi$ are partial maps, we have $z = x\phi$, $y = z\psi$, and $y = x(\phi \circ \psi)$. Hence

$$x(\phi \circ \psi) = y = z\psi = (x\phi)\psi,$$

exactly as required. □

A partial map ϕ is called a *map*, or a *function*, if $\text{dom}\,\phi = X$. Thus a relation ϕ on X is a map if and only if $|x\phi| = 1$ for every x in X. If ϕ and ψ are maps, it is easy to see that $\phi \circ \psi$ is again a map, and that the operation (1.4.2) coincides with ordinary composition of maps. To put it formally, we have

Proposition 1.4.4 *The set $\mathcal{T}_X$ of all maps from X into itself is a subsemigroup of $(\mathcal{B}_X, \circ)$.* □

Once again it is important to note that the converse ϕ^{-1} of a map need not be a map. Indeed we have the following easily proved result:

Proposition 1.4.5 *Let X be a non-empty set.*

(1) *If $\phi \in \mathcal{P}_X$ then $\phi^{-1} \in \mathcal{P}_X$ if and only if ϕ is one-one.*
(2) *If $\phi \in \mathcal{T}_X$, then $\phi^{-1} \in \mathcal{T}_X$ if and only if ϕ is bijective (that is to say, ϕ is both one-one and onto).* □

In Section 1.3 we encountered relations, namely *order* relations, that are reflexive, anti-symmetric and transitive. We can now express these properties very compactly as follows. A relation ρ on a set X is

reflexive if and only if $1_X \subseteq \rho$,
anti-symmetric if and only if $\rho \cap \rho^{-1} = 1_X$, and
transitive if and only if $\rho \circ \rho \subseteq \rho$.

We now define an *equivalence* ρ on a set X to be a relation that is reflexive, transitive and *symmetric*, by which we mean that

$$(\forall x, y \in X)\ (x, y) \in \rho \ \Rightarrow\ (y, x) \in \rho.$$

In compact form this property is expressed as $\rho \subseteq \rho^{-1}$. Notice that, by (1.4.12) and (1.4.10) it then follows that $\rho^{-1} \subseteq \rho$; thus the symmetry condition can equally well be expressed as $\rho^{-1} = \rho$. On the same theme, if ρ is an equivalence, then by (1.4.3) we can deduce that

$$\rho = 1_X \circ \rho \subseteq \rho \circ \rho;$$

thus the transitivity condition can be replaced by $\rho \circ \rho = \rho$.

If ρ is an equivalence on X then, by (1.4.6),

$$\operatorname{dom}\rho \supseteq \operatorname{dom} 1_X = X, \quad \operatorname{im}\rho \supseteq \operatorname{im} 1_X = X;$$

hence $\operatorname{dom}\rho = \operatorname{im}\rho = X$.

A family $\pi = \{A_i : i \in I\}$ of subsets of a set X is said to form a *partition* of X if

(P1) each A_i is non-empty;
(P2) for all i, j in I, either $A_i = A_j$ or $A_i \cap A_j = \emptyset$;
(P3) $\bigcup\{A_i : i \in I\} = X$.

On the face of it, the notions of 'partition' and 'equivalence' are quite different, but in fact they are closely related. The proof of the following proposition is routine and is omitted.

Proposition 1.4.6 *Let ρ be an equivalence on a set X. Then the family*

$$\Phi(\rho) = \{x\rho : x \in X\}$$

of subsets of X is a partition of X.

Conversely, if $\pi = \{A_i : i \in I\}$ is a partition of X, then the relation

$$\Psi(\pi) = \{(x,y) \in X \times X : (\exists i \in I)\; x, y \in A_i\}$$

is an equivalence on X.

For every equivalence ρ on X, $\Psi\big(\Phi(\rho)\big) = \rho$, and for every partition π of X, $\Phi\big(\Psi(\pi)\big) = \pi$. □

If ρ is an equivalence on X, we shall sometimes write $x\;\rho\;y$ or $x \equiv y \pmod{\rho}$ as alternatives to $(x,y) \in \rho$. The sets $x\rho$ that form the partition associated with the equivalence are called *ρ-classes*, or *equivalence classes.* The set of ρ-classes, whose elements are the subsets $x\rho$, is called the *quotient set of X by ρ*, and is denoted by X/ρ. In the next section we shall have occasion to examine the natural map $\rho^\natural$ (read 'ρ natural') from X onto X/ρ defined by

$$x\rho^\natural = x\rho \quad (x \in X). \tag{1.4.15}$$

An important connection between maps and equivalences is given by

Proposition 1.4.7 *If $\phi : X \to Y$ is a map, then $\phi \circ \phi^{-1}$ is an equivalence.*

Proof The easiest way to see this is to note that

$$\begin{aligned}\phi \circ \phi^{-1} &= \{(x,y) \in X \times X : (\exists z \in X)\; (x,z) \in \phi,\; (y,z) \in \phi\} \\ &= \{(x,y) \in X \times X : x\phi = y\phi\}.\end{aligned}$$

It is then clear that $\phi \circ \phi^{-1}$ is reflexive, symmetric and transitive. □

We call the equivalence $\phi \circ \phi^{-1}$ the *kernel* of ϕ, and write $\phi \circ \phi^{-1} = \ker\phi$. Notice that $\ker\rho^\natural = \rho$.

If $\{\rho_i : i \in I\}$ is a non-empty family of equivalences on a set X, then, as may be verified in a routine manner, $\bigcap\{\rho_i : i \in I\}$ is again an equivalence.

If $\mathbf{R}$ is any relation at all on X—even the empty relation will do—then the family of equivalences containing $\mathbf{R}$ is non-empty, since $X \times X$ is one such; hence the intersection of all the equivalences containing $\mathbf{R}$ is an equivalence, the unique minimum equivalence on X containing $\mathbf{R}$. We call it the equivalence *generated by* $\mathbf{R}$, and denote it by $\mathbf{R}^e$.

It is frequently necessary to be able to describe $\mathbf{R}^e$ for a given $\mathbf{R}$, and the foregoing general description is not particularly useful. It is necessary therefore to develop an alternative description. First, let $\mathbf{S}$ be a relation on X such that $1_X \subseteq \mathbf{S}$—a reflexive relation, in fact. Then we have

$$\mathbf{S} \subseteq \mathbf{S} \circ \mathbf{S} \subseteq \mathbf{S} \circ \mathbf{S} \circ \mathbf{S} \subseteq \cdots,$$

which we can write in simpler notation as

$$\mathbf{S} \subseteq \mathbf{S}^2 \subseteq \mathbf{S}^3 \subseteq \cdots$$

The relation

$$\mathbf{S}^\infty = \bigcup \{\mathbf{S}^n : n \geq 1\} \tag{1.4.16}$$

is called the *transitive closure* of the relation $\mathbf{S}$, a term that is justified by the following lemma:

Lemma 1.4.8 *For every reflexive relation* $\mathbf{S}$ *on a set* X*, the relation* $\mathbf{S}^\infty$ *defined by* (1.4.16) *is the smallest transitive relation on* X *containing* $\mathbf{S}$.

Proof First, $\mathbf{S}^\infty$ is transitive. Suppose that $(x, y), (y, z) \in \mathbf{S}^\infty$. Then there exist positive integers m and n such that $(x, y) \in \mathbf{S}^m$ and $(y, z) \in \mathbf{S}^n$. It follows that

$$(x, z) \in \mathbf{S}^m \circ \mathbf{S}^n = \mathbf{S}^{m+n} \subseteq \mathbf{S}^\infty.$$

It is clear that $\mathbf{S}^\infty$ contains $\mathbf{S}^1 = \mathbf{S}$.

Finally, if $\mathbf{T}$ is a transitive relation containing $\mathbf{S}$, then

$$\mathbf{S}^2 = \mathbf{S} \circ \mathbf{S} \subseteq \mathbf{T} \circ \mathbf{T} \subseteq \mathbf{T},$$

and more generally $\mathbf{S}^n \subseteq \mathbf{T}$ for all $n \geq 1$. Hence $\mathbf{S}^\infty \subseteq \mathbf{T}$. □

We now have

Proposition 1.4.9 *For every relation* $\mathbf{R}$ *on a set* X*,*

$$\mathbf{R}^e = [\mathbf{R} \cup \mathbf{R}^{-1} \cup 1_X]^\infty.$$

Proof From Lemma 1.4.8 we see that the relation $\mathbf{E} = [\mathbf{R} \cup \mathbf{R}^{-1} \cup 1_X]^\infty$ is transitive and contains $\mathbf{R}$. Since

$$1_X \subseteq \mathbf{R} \cup \mathbf{R}^{-1} \cup 1_X \subseteq \mathbf{E},$$

$\mathbf{E}$ is also reflexive. Certainly the relation $\mathbf{S} = \mathbf{R} \cup \mathbf{R}^{-1} \cup 1_X$ is symmetric, and it follows that, for every n in $\mathbf{N}$,

$$\mathbf{S}^n = (\mathbf{S}^{-1})^n = (\mathbf{S}^n)^{-1},$$

the second equality being a specialization of (1.4.11). Hence $\mathbf{S}^n$ is symmetric. It now follows that $\mathbf{E} = \mathbf{S}^\infty$ is symmetric, since

$$\begin{aligned}(x,y) \in \mathbf{E} &\Rightarrow (\exists n \in \mathbf{N})\ (x,y) \in \mathbf{S}^n \\ &\Rightarrow (\exists n \in \mathbf{N})\ (y,x) \in \mathbf{S}^n \\ &\Rightarrow (y,x) \in \mathbf{E}.\end{aligned}$$

We have shown that $\mathbf{E}$ is an equivalence relation containing $\mathbf{R}$.

Suppose now that σ is an equivalence relation containing $\mathbf{R}$. Then $1_X \subseteq \sigma$, and $\mathbf{R}^{-1} \subseteq \sigma^{-1} = \sigma$. Hence

$$\mathbf{S} = \mathbf{R} \cup \mathbf{R}^{-1} \cup 1_X \subseteq \sigma.$$

Moreover,

$$\mathbf{S} \circ \mathbf{S} \subseteq \sigma \circ \sigma = \sigma,$$

and more generally $\mathbf{S}^n \subseteq \sigma$ for all $n \geq 1$. It follows that $\mathbf{E} = \mathbf{S}^\infty \subseteq \sigma$. We have shown that $\mathbf{E} = [\mathbf{R} \cup \mathbf{R}^{-1} \cup 1_X]^\infty$ is the smallest equivalence on X containing $\mathbf{R}$. Thus

$$\mathbf{R}^e = [\mathbf{R} \cup \mathbf{R}^{-1} \cup 1_X]^\infty$$

as required. □

In more down-to-earth terms, Proposition 1.4.9 can be rewritten thus:

Proposition 1.4.10 *If* $\mathbf{R}$ *is a relation on a set* X *and* $\mathbf{R}^e$ *is the smallest equivalence on* X *containing* $\mathbf{R}$*, then* $(x,y) \in \mathbf{R}^e$ *if and only if either* $x = y$ *or, for some* n *in* $\mathbf{N}$*, there is a sequence of transitions*

$$x = z_1 \longrightarrow z_2 \longrightarrow \cdots \longrightarrow z_n = y$$

in which, for each i *in* $\{1, 2, \ldots, n-1\}$*, either* $(z_i, z_{i+1}) \in \mathbf{R}$ *or* $(z_{i+1}, z_i) \in \mathbf{R}$. □

1.5 CONGRUENCES

Let S be a semigroup. A relation $\mathbf{R}$ on the set S is called *left compatible* (with the operation on S) if

$$(\forall s,t,a \in S) \quad (s,t) \in \mathbf{R} \Rightarrow (as, at) \in \mathbf{R},$$

and *right compatible* if

$$(\forall s,t,a \in S) \quad (s,t) \in \mathbf{R} \Rightarrow (sa, ta) \in \mathbf{R}.$$

It is called *compatible* if

$$(\forall s,t,s',t' \in S) \quad [(s,t) \in \mathbf{R} \text{ and } (s',t') \in \mathbf{R}] \Rightarrow (ss', tt') \in \mathbf{R}.$$

A left [right] compatible equivalence is called a *left* [*right*] *congruence*. A compatible equivalence relation is called a *congruence*.

Proposition 1.5.1 *A relation* ρ *on a semigroup* S *is a congruence if and only if it is both a left and a right congruence.*

Proof Suppose first that ρ is a congruence. If $(s,t) \in \rho$ and $a \in S$ then $(a,a) \in \rho$ by reflexivity and so $(as, at) \in \rho$ and $(sa, ta) \in \rho$ by compatibility. Thus ρ is both left and right compatible.

Conversely, suppose that ρ is both a left and a right congruence, and let $(s,t), (s',t') \in \rho$. Then $(ss', ts') \in \rho$ by right compatibility and $(ts', tt') \in \rho$ by left compatibility. Hence $(ss', tt') \in \rho$ by transitivity. Thus ρ is a congruence. □

Exercise 11 below explores in more detail the relationship between left and right compatibility on the one hand and compatibility on the other.

If ρ is a congruence on a semigroup S then we can define a binary operation on the quotient set S/ρ in a natural way as follows:

$$(a\rho)(b\rho) = (ab)\rho. \tag{1.5.1}$$

This is well-defined precisely because ρ is compatible: for all a, a', b, b' in S,

$$\begin{aligned} a\rho = a'\rho \text{ and } b\rho = b'\rho \quad &\Rightarrow \quad (a,a') \in \rho \text{ and } (b,b') \in \rho \\ &\Rightarrow \quad (ab, a'b') \in \rho \\ &\Rightarrow \quad (ab)\rho = (a'b')\rho. \end{aligned}$$

The operation is easily seen to be associative, and so S/ρ is a semigroup.

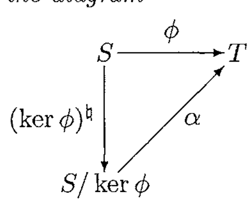
We have proved part of the following theorem:

Theorem 1.5.2 *Let S be a semigroup, and let ρ be a congruence on S. Then S/ρ is a semigroup with respect to the operation defined by* (1.5.1), *and the map $\rho^\natural$ from S onto S/ρ given by* (1.4.15) *is a morphism.*

Now let T be a semigroup and let $\phi : S \to T$ be a morphism. Then the relation

$$\ker\phi = \phi \circ \phi^{-1} = \{(a,b) \in S \times S : a\phi = b\phi\}$$

is a congruence on S, and there is a monomorphism $\alpha : S/\ker\phi \to T$ such that $\operatorname{im}\alpha = \operatorname{im}\phi$ *and the diagram*



commutes.

Proof It is easy to verify that $\rho^\natural$ is a morphism. For the second part, suppose that $\phi : S \to T$ is a morphism. That $\ker\phi$ is an equivalence follows from Proposition 1.4.7. To show that it is a congruence, suppose that $(a, a'), (b, b') \in \ker \phi$. Then $a\phi = a'\phi$ and $b\phi = b'\phi$, from which we deduce that

$$(ab)\phi = (a\phi)(b\phi) = (a'\phi)(b'\phi) = (a'b')\phi.$$

Hence $(ab, a'b') \in \ker \phi$, as required. For brevity, let us denote $\ker \phi$ by κ, and define $\alpha : S/\kappa \to T$ by

$$(a\kappa)\alpha = a\phi \quad (a \in S).$$

Then α is both well-defined and one-one, since

$$a\kappa = b\kappa \iff (a, b) \in \kappa \iff a\phi = b\phi.$$

It is also a morphism, since, for all a, b in S,

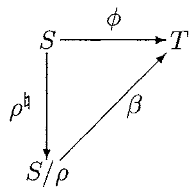
$$[(a\kappa)(b\kappa)]\alpha = [(ab)\kappa]\alpha = (ab)\phi$$
$$= (a\phi)(b\phi) = [(a\kappa)\alpha][(b\kappa)\alpha].$$

Clearly $\operatorname{im}\alpha = \operatorname{im}\phi$, and from the definition of α it is clear that, for all a in S,

$$a\kappa^\natural\alpha = a\phi. \qquad \square$$

The next theorem is concerned with a more general situation, and will frequently be useful:

Theorem 1.5.3 *Let ρ be a congruence on a semigroup S, and let $\phi : S \to T$ be a morphism such that $\rho \subseteq \ker \phi$. Then there is a unique morphism $\beta : S/\rho \to T$ such that $\operatorname{im}\beta = \operatorname{im}\phi$ and such that the diagram*



commutes.

Proof We define $\beta : S/\rho \to T$ by

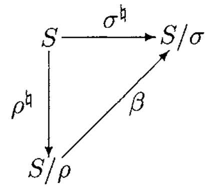
$$(a\rho)\beta = a\phi \quad (a \in S). \tag{1.5.2}$$

Then β is well-defined, since, for all a, b in S,

$$a\rho = b\rho \Rightarrow (a, b) \in \rho \Rightarrow (a, b) \in \ker \phi \Rightarrow a\phi = b\phi.$$

It is now a routine matter to show that β is a morphism, that $\operatorname{im}\beta = \operatorname{im}\phi$, and that $\rho^\natural \circ \beta = \phi$. The uniqueness of β is also clear, since any morphism satisfying $\rho^\natural \circ \beta = \phi$ *must* be defined by the rule (1.5.2). $\square$

One application of this theorem is to the situation where ρ and σ are congruences on S and where $\rho \subseteq \sigma$. The theorem implies that there is a morphism β from S/ρ onto S/σ such that the diagram



commutes. The morphism β is given by

$$(a\rho)\beta = a\sigma \quad (a \in S),$$

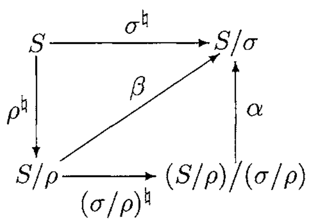
and the congruence $\ker\beta$ on S/ρ is given by

$$\ker\beta = \{(a\rho, b\rho) \in S/\rho \times S/\rho : (a,b) \in \sigma\}.$$

It is usual to write $\ker\beta$ as σ/ρ. From Theorem 1.5.2 it now follows that there is an isomorphism $\alpha : (S/\rho)/(\sigma/\rho) \to S/\sigma$ defined by

$$[(a\rho)(\sigma/\rho)]\alpha = a\sigma \quad (a \in S),$$

and such that the diagram



commutes. We summarize in a theorem:

Theorem 1.5.4 *Let ρ, σ be congruences on a semigroup S such that $\rho \subseteq \sigma$. Then*

$$\sigma/\rho = \{(x\rho, y\rho) \in S/\rho \times S/\rho : (x,y) \in \sigma\}$$

is a congruence on S/ρ, and $(S/\rho)/(\sigma/\rho) \simeq S/\sigma$. □

Since the intersection of a non-empty family of congruences on a semigroup S is a congruence on S, we can argue exactly as in Section 1.4 and deduce that for every relation $\mathbf{R}$ on S there is a unique smallest congruence $\mathbf{R}^{\#}$ (read '$\mathbf{R}$ sharp') on S containing $\mathbf{R}$, namely the intersection of the family of *all* congruences on S containing $\mathbf{R}$. We now seek a result, analogous to Proposition 1.4.9, that will give us a usable description of $\mathbf{R}^{\#}$.

First, for an arbitrary relation $\mathbf{R}$ on S we define

$$\mathbf{R}^c = \{(xay, xby) : x, y \in S^1,\ (a,b) \in \mathbf{R}\}.$$

Then

Lemma 1.5.5 $\mathbf{R}^c$ *is the smallest left and right compatible relation containing* $\mathbf{R}$.

Proof First, it is clear that $\mathbf{R}^c$ contains $\mathbf{R}$. To show that $\mathbf{R}^c$ is left compatible, suppose that $(u,v) \in \mathbf{R}^c$ and $w \in S$. Then $u = xay$, $v = xby$ for some x, y in S^1 and some (a,b) in $\mathbf{R}$. Hence $wu = (wx)ay$ and $wv = (wx)by$, and so $(wu, wv) \in \mathbf{R}^c$ as required. Right compatibility follows in a similar way.

Suppose now that $\mathbf{S}$ is a left and right compatible relation containing $\mathbf{R}$. Then for all x, y in S^1 and all (a,b) in $\mathbf{R}$ it follows that $(xay, xby) \in \mathbf{S}$. Hence $\mathbf{R}^c \subseteq \mathbf{S}$, as required. □

Some further easily proved properties of $\mathbf{R}^c$ are encapsulated in the following lemma:

Lemma 1.5.6 *Let* $\mathbf{R}$, $\mathbf{S}$ *be relations on a semigroup* S. *Then:*

(1) $\mathbf{R} \subseteq \mathbf{S} \;\Rightarrow\; \mathbf{R}^c \subseteq \mathbf{S}^c$;
(2) $(\mathbf{R}^{-1})^c = (\mathbf{R}^c)^{-1}$;
(3) $(\mathbf{R} \cup \mathbf{S})^c = \mathbf{R}^c \cup \mathbf{S}^c$. □

Next, we have

Lemma 1.5.7 *Let* $\mathbf{R}$ *be a left and right compatible relation on a semigroup* S. *Then* $\mathbf{R}^n$ $(= \mathbf{R} \circ \mathbf{R} \circ \cdots \circ \mathbf{R})$ *is left and right compatible for every* $n \geq 1$.

Proof Let $(s,t) \in \mathbf{R}^n$. Then there exist $z_1, z_2, \ldots, z_{n-1}$ in S such that

$$(s, z_1),\ (z_1, z_2),\ \ldots,\ (z_{n-1}, t)\ \in\ \mathbf{R}.$$

Since $\mathbf{R}$ is left and right compatible by assumption, it follows that, for every a in S,

$$(as, az_1),\ (az_1, az_2),\ \ldots,\ (az_{n-1}, at)\ \in\ \mathbf{R},$$
$$(sa, z_1a),\ (z_1a, z_2a),\ \ldots,\ (z_{n-1}a, ta)\ \in\ \mathbf{R}.$$

Hence (as, at), $(sa, ta) \in \mathbf{R}^n$, as required. □

We can now easily obtain the following characterization of $\mathbf{R}^\#$, the congruence on S generated by $\mathbf{R}$:

Proposition 1.5.8 *For every relation* $\mathbf{R}$ *on a semigroup* S, $\mathbf{R}^\# = (\mathbf{R}^c)^e$.

Proof From Proposition 1.4.9 we know that $(\mathbf{R}^c)^e$ is an equivalence relation containing $\mathbf{R}^c$, and so certainly containing $\mathbf{R}$. To show that $(\mathbf{R}^c)^e$ is a congruence, we must show that it is both left and right compatible. So suppose that $(s,t) \in (\mathbf{R}^c)^e$ and $a \in S$. Then, by Proposition 1.4.9, $(s,t) \in \mathbf{S}^n$ for some n in $\mathbf{N}$, where $\mathbf{S} = \mathbf{R}^c \cup (\mathbf{R}^c)^{-1} \cup 1_S^c$. Now, by Lemma 1.5.6, and by the easy observation that $1_S^c = 1_S$,

$$\mathbf{S} = \mathbf{R}^c \cup (\mathbf{R}^{-1})^c \cup 1_S^c = (\mathbf{R} \cup \mathbf{R}^{-1} \cup 1_S)^c.$$

Thus $\mathbf{S}$ is right and left compatible by Lemma 1.5.5, and hence, by Lemma 1.5.7, so also is $\mathbf{S}^n$. It follows that

$$(as, at) \in \mathbf{S}^n \subseteq (\mathbf{R}^c)^e, \quad (sa, ta) \in \mathbf{S}^n \subseteq (\mathbf{R}^c)^e,$$

and so $(\mathbf{R}^c)^e$ is a congruence on S containing $\mathbf{R}$.

To show that $(\mathbf{R}^c)^e$ is the smallest congruence on S containing $\mathbf{R}$, consider a congruence κ on S containing $\mathbf{R}$. Then $\kappa^c = \kappa$ by Lemma 1.5.5, and so

$$\mathbf{R}^c \subseteq \kappa^c = \kappa.$$

We now have that κ is an equivalence on S containing $\mathbf{R}^c$, and so from Proposition 1.4.9 it follows that $(\mathbf{R}^c)^e \subseteq \kappa$. □

By analogy with Propositions 1.4.9 and 1.4.10 we may write the last result in more elementary terms. First, if c, d in S are such that

$$c = xay, \quad d = xby,$$

for some x, y in S^1, where either (a, b) or (b, a) belongs to $\mathbf{R}$, we say that c is connected to d by an *elementary* $\mathbf{R}$*-transition*. Then we have

Proposition 1.5.9 *Let* $\mathbf{R}$ *be a relation on a semigroup* S*, and let* $a, b \in S$*. Then* $(a, b) \in \mathbf{R}^\#$ *if and only if either* $a = b$ *or, for some* n *in* $\mathbf{N}$*, there is a sequence*

$$a = z_1 \to z_2 \to \cdots \to z_n = b$$

of elementary $\mathbf{R}$*-transitions connecting* a *to* b. □

Having used the musical notations ♮ (natural) and # (sharp), we now complete the scene by using ♭ (flat). We define, for an arbitrary equivalence $\mathbf{E}$ on a semigroup S,

$$\mathbf{E}^\flat = \{(a, b) \in S \times S : (\forall x, y \in S^1)\ (xay, xby) \in \mathbf{E}\}. \tag{1.5.3}$$

Then we have

Proposition 1.5.10 *Let* $\mathbf{E}$ *be an equivalence on a semigroup* S*. Then* $\mathbf{E}^\flat$ *is the largest congruence on* S *contained in* $\mathbf{E}$.

Proof First, it is clear that $\mathbf{E}^\flat$ is an equivalence. Also, if $(a, b) \in \mathbf{E}^\flat$ and $c \in S$, then $(xcay, xcby) \in \mathbf{E}$ for all choices of x and y in S^1. Hence $(ca, cb) \in \mathbf{E}^\flat$, and similarly $(ac, bc) \in \mathbf{E}^\flat$. Thus $\mathbf{E}^\flat$ is a congruence, and clearly $\mathbf{E}^\flat \subseteq \mathbf{E}$, since

$$(a, b) \in \mathbf{E}^\flat \ \Rightarrow\ (1a1, 1b1) \in \mathbf{E} \ \Rightarrow\ (a, b) \in \mathbf{E}.$$

Finally, suppose that η is a congruence on S contained in $\mathbf{E}$. Then

$$\begin{aligned}(a, b) \in \eta \ &\Rightarrow\ (\forall x, y \in S^1)\ (xay, xby) \in \eta \\ &\Rightarrow\ (\forall x, y \in S^1)\ (xay, xby) \in \mathbf{E}\end{aligned}$$

which implies in turn that $(a, b) \in \mathbf{E}^\flat$. □

Notice carefully that, while $\mathbf{R}^\#$ is defined for an arbitrary relation $\mathbf{R}$, the congruence $\mathbf{E}^\flat$ is defined only for an *equivalence* $\mathbf{E}$.

One special case of $\mathbf{E}^\flat$ is worthy of particular mention. For every subset A of a semigroup S there is an associated equivalence π_A whose classes are A and $S \setminus A$. The congruence $\pi_A^\flat$, given by

$$\pi_A^\flat = \{(u,v) \in S \times S : (\forall x, y \in S^1)\ [xuy \in A \text{ if and only if } xvy \in A]\},$$

is called the *syntactic congruence* of A. This is of importance in the theory of automata and languages. See, for example, Howie (1991).

Let S be a semigroup. Then both the set $\mathcal{E}(S)$ of equivalences and the set $\mathcal{C}(S)$ of congruences on S are partially ordered by inclusion. In fact both are lattices. If $\rho, \sigma \in \mathcal{E}(S)$ then $\rho \cap \sigma \in \mathcal{E}(S)$ and is their greatest lower bound, while $(\rho \cup \sigma)^e$ is their least upper bound. If $\rho, \sigma \in \mathcal{C}(S)$ then $\rho \cap \sigma \in \mathcal{C}(S)$ and is their greatest lower bound, while $(\rho \cup \sigma)^\#$ is their least upper bound. Notice now that if $\rho, \sigma \in \mathcal{C}(S)$ then Lemma 1.5.6 gives

$$(\rho \cup \sigma)^c = \rho^c \cup \sigma^c = \rho \cup \sigma.$$

Hence, by Proposition 1.5.8, $(\rho \cup \sigma)^\# = (\rho \cup \sigma)^e$. We may therefore write $\rho \vee \sigma$ unambiguously for the join of ρ and σ either in $\mathcal{E}(S)$ or in $\mathcal{C}(S)$.

In the foregoing analysis there is no difficulty in replacing the set $\{\rho, \sigma\}$ by an arbitary family of equivalences or congruences, and so both the lattices $\mathcal{E}(S)$ and $\mathcal{C}(S)$ are complete. Both lattices have maximum element $S \times S$ and minimum element 1_S.

A specialization of Propositions 1.4.10 and 1.5.9 to the case where the relation $\mathbf{R}$ is the union of two equivalences or congruences is worth stating separately:

Proposition 1.5.11 *Let ρ, σ be equivalences on a set S [congruences on a semigroup S]. Then $(a,b) \in \rho \vee \sigma$ if and only if, for some n in $\mathbf{N}$, there exist elements $x_1, x_2, \ldots, x_{2n-1}$ in S such that*

$$(a, x_1) \in \rho,\ (x_1, x_2) \in \sigma,\ (x_2, x_3) \in \rho, \ldots,\ (x_{2n-1}, b) \in \sigma.$$

Proof The result says effectively that

$$\rho \vee \sigma = (\rho \circ \sigma)^\infty.$$

To see that this is so, notice first that the general approach in Propositions 1.4.9 and 1.5.8 gives

$$\rho \vee \sigma = \mathbf{R}^\infty,$$

where

$$\begin{aligned}\mathbf{R} &= (\rho \cup \sigma) \cup (\rho \cup \sigma)^{-1} \cup 1_S \\ &= \rho \cup \sigma \cup \rho^{-1} \cup \sigma^{-1} \cup 1_S \\ &= \rho \cup \sigma\end{aligned}$$

(since $\rho = \rho^{-1}$, $\sigma = \sigma^{-1}$, $1_S \subseteq \rho$ and $1_S \subseteq \sigma$ by assumption). Now $\rho \subseteq \rho \cup \sigma$ and $\sigma \subseteq \rho \cup \sigma$, and so $\rho \circ \sigma \subseteq (\rho \cup \sigma)^2$. Hence $(\rho \circ \sigma)^n \subseteq (\rho \cup \sigma)^{2n}$ for every $n \geq 1$, and so

$$(\rho \circ \sigma)^\infty \subseteq (\rho \cup \sigma)^\infty = \rho \vee \sigma.$$

To show the converse inclusion, notice that $\rho \subseteq \rho \circ \sigma$ and $\sigma \subseteq \rho \circ \sigma$, since ρ and σ are equivalences. Hence $\rho \cup \sigma \subseteq \rho \circ \sigma$, and so

$$\rho \vee \sigma = (\rho \cup \sigma)^\infty \subseteq (\rho \circ \sigma)^\infty. \qquad \square$$

A useful further specialization of this result is provided by

Corollary 1.5.12 *Let ρ, σ be equivalences on a set S [congruences on a semigroup S] such that $\rho \circ \sigma = \sigma \circ \rho$. Then $\rho \vee \sigma = \rho \circ \sigma$.*

Proof If $\rho \circ \sigma = \sigma \circ \rho$, then

$$(\rho \circ \sigma)^2 = \rho \circ (\sigma \circ \rho) \circ \sigma = (\rho \circ \rho) \circ (\sigma \circ \sigma) = \rho \circ \sigma,$$

and more generally $(\rho \circ \sigma)^n = \rho \circ \sigma$ for every n in $\mathbf{N}$. Hence $(\rho \circ \sigma)^\infty = \rho \circ \sigma$, and the result follows from Proposition 1.5.11. $\square$

1.6 FREE SEMIGROUPS AND MONOIDS; PRESENTATIONS

Let A be a non-empty set. Let A^+ be the set of all finite, non-empty words $a_1 a_2 \ldots a_m$ in the 'alphabet' A. A binary operation is defined on A^+ by juxtaposition:

$$(a_1 a_2 \ldots a_m)(b_1 b_2 \ldots b_n) = a_1 a_2 \ldots a_m b_1 b_2 \ldots b_n.$$

With respect to this operation, A^+ is a semigroup, called the *free semigroup* on A. The set A is a generating set for A^+. By contrast with the situation in group theory—see, for example, Kurosh (1956)—the set A is a unique minimum generating set for A^+, since $A = A^+ \setminus (A^+)^2$. In making this statement we are of course tacitly identifying each element a of A with the one-letter word a in A^+. This is certainly a reasonable identification, and we shall almost always want to make it. At other times we shall refer to the map $\alpha : A \to A^+$ that associates each a in A with the corresponding one-letter word in A^+ as the *standard embedding* of A in A^+.

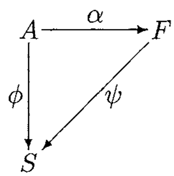
If we adjoin an identity 1 to A^+ we obtain the *free monoid* on A. We denote this by A^*, and think of 1 as the *empty word* (containing no letters at all).

If $A = \{a\}$ has just one letter, we write a^+ rather than $\{a\}^+$. Note that $a^+ = \{a, a^2, a^3, \ldots\}$, and is simply the infinite monogenic semigroup already encountered in Section 1.2. If $|A| > 1$ then A^+ is not commutative.

An abstract definition of 'a free semigroup on A' can be given as follows: F is a *free semigroup on A* if

(F1) there is a map $\alpha : A \to F$;

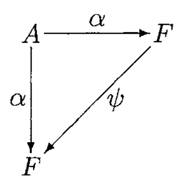
(F2) for every semigroup S and every map $\phi : A \to S$ there exists a unique morphism $\psi : F \to S$ such that the diagram



(1.6.1)

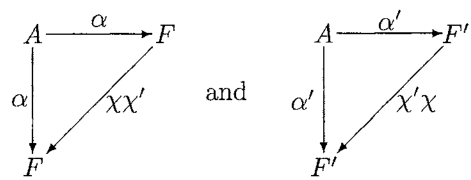
commutes.

These properties in fact define F (if it exists) up to isomorphism. Notice first that if we substitute F and α for S and ϕ in the above definition then we see that the unique ψ making the diagram



(1.6.2)

commute is the identity map 1_F. Suppose now that F' and α' also have the properties (F1) and (F2). Then we may substitute F' and α' for S and ϕ in (1.6.1) and obtain a unique morphism $\chi : F \to F'$ such that $\alpha\chi = \alpha'$. Then, interchanging the roles of F and F', we similarly obtain a unique morphism $\chi' : F' \to F$ such that $\alpha'\chi' = \alpha$. This gives us two commutative diagrams



and the uniqueness in (1.6.2) (which applies also to F') then implies that $\chi\chi' = 1_F$ and $\chi'\chi = 1_{F'}$. Thus $F' \simeq F$.

It is easy to see that the free semigroup A^+ as defined above does have the properties (F1) and (F2). For α we take the standard embedding, and for a given $\phi : A \to S$ we define the morphism $\psi : A^+ \to S$ by

$$(a_1a_2\ldots a_m)\psi = (a_1\phi)(a_2\phi)\ldots(a_m\phi) \quad (a_1a_2\ldots a_m \in A^+).$$

It is a routine matter to verify that ψ is a morphism and that $\alpha\psi = \phi$.

If S is a semigroup and A is a generating set for S then the property (F2) gives us a morphism ψ from A^+ onto S. Hence $S \simeq A^+/\ker\psi$. Now, since we can always find a generating set for S—if all else fails, then S itself will do—we deduce that every semigroup can be expressed up to isomorphism as a quotient of a free semigroup by a congruence. The expression is of course not unique.

If $A = \{a_1, a_2, \dots, a_n\}$ is finite, and if we can find a finite set

$$\mathbf{R} = \{(w_1, z_1), (w_2, z_2), \dots, (w_r, z_r)\}$$

of elements (w_i, z_i) in $A^+ \times A^+$ such that $\mathbf{R}^\# = \ker\psi$, then we say that S is *finitely presented*, and that it has a *finite presentation*

$$\langle a_1, a_2, \dots, a_n \mid w_1 = z_1, w_2 = z_2, \dots, w_r = z_r\rangle.$$

The semigroup has *generators* $a_1, a_2, \dots, a_n$ and *defining relations* $w_1 = z_1, w_2 = z_2, \dots, w_r = z_r$.

We have already in effect come across one important example: the monogenic semigroup $M(m, r)$ can be expressed as

$$\langle a \mid a^{m+r} = a^m\rangle.$$

What we have done for semigroups in this section can be done just as easily for monoids, with A^* replacing A^+. We conclude the section by giving an example of a monoid presentation. Consider the monoid

$$B = \langle a, b \mid ab = 1\rangle = \{a, b\}^*/\rho,$$

where $\rho = \{(ab, 1)\}^\#$. It is clear that for any word w in $\{a, b\}^*$ in which b follows a there is a word w' such that $w' \;\rho\; w$ and $|w'| < |w|$. The elements of B are ρ-classes $(b^m a^n)\rho \quad (m, n \geq 0)$, but it is not immediately clear that these ρ-classes are all distinct. To show that they are in fact all different we adopt a rather oblique approach.

Consider the submonoid $M = \langle A, B\rangle$ of the full transformation monoid $\mathcal{T}_{\mathbf{N}}$, where

$$A = \begin{pmatrix} 1\,2\,3\,4 \dots \\ 2\,3\,4\,5 \dots \end{pmatrix}, \quad B = \begin{pmatrix} 1\,2\,3\,4 \dots \\ 1\,1\,2\,3 \dots \end{pmatrix}.$$

Then by the property (F2) (interpreted for monoids) there is a unique morphism $\phi : \{a, b\}^* \to M$ such that $a\psi = A$, $b\psi = B$. Now

$$(ab)\psi = AB = 1 = 1\psi,$$

and so $(ab, 1) \in \ker\psi$. Since $\ker\psi$ is a congruence, we deduce that $\rho = \{(ab, 1)\}^\# \subseteq \ker\psi$, and so, by Theorem 1.5.3, there is a morphism $\bar{\psi} : B \to M$ given by

$$[(b^m a^n)\rho]\bar{\psi} = B^m A^n \quad (m, n \geq 0).$$

Now an easy calculation shows that

$$B^m A^n = \begin{pmatrix} 1 & 2 & \dots & m+1 & m+2 & m+3 & \dots \\ n+1 & n+1 & \dots & n+1 & n+2 & n+3 & \dots \end{pmatrix},$$

and from this it follows that the elements $B^m A^n$ are all distinct. Hence, since $\bar{\psi}$ is a well-defined map, the elements $(b^m a^n)\rho$ are also all distinct.

Notice now that, if we work modulo ρ, then for all $m, n, p, q \geq 0$,

$$(b^m a^n)(b^p a^q) \equiv \begin{cases} b^m a^{q-p+n} & \text{if } n \geq p \\ b^{m-n+p} a^q & \text{if } n \leq p. \end{cases}$$

We can put the two cases together as follows:

$$(b^m a^n)(b^p a^q) \equiv b^{m-n+t} a^{q-p+t} \quad \big(t = \max(n,p)\big).$$

The monoid B is thus isomorphic to the monoid $\mathbf{N}^0 \times \mathbf{N}^0$, with multiplication given by

$$(m,n)(p,q) = \big(m - n + \max(n,p), q - p + \max(n,p)\big). \tag{1.6.3}$$

This important semigroup, called the *bicyclic semigroup*, will feature frequently as we proceed. Notice that we have in effect described it in three different ways—by means of a presentation, as a semigroup of maps from $\mathbf{N}$ into itself, and as a set of ordered pairs of non-negative integers. For the most part it will be convenient to use the last of these three descriptions.

Though we have no special word corresponding to 'monoid' for a semigroup with zero, we can repeat the 'free object' analysis for semigroups with zero—or indeed for monoids with zero. The appropriate free object is $A^+ \cup \{0\}$, with the obvious multiplication (or $A^* \cup \{0\}$ in the case of monoids). The following example introduces another important semigroup.

Let

$$B_2 = \langle a, b \mid a^2 = b^2 = 0,\ aba = a,\ bab = b \rangle = \big(\{a,b\}^+ \cup \{0\}\big)/\rho,$$

where $\rho = \{(a^2,0),\ (b^2,0),\ (aba,a),\ (bab,b)\}^\#$. Then every word in $\{a,b\}^+ \cup \{0\}$ is reducible modulo ρ to exactly one of the five words 0, a, b, ab, ba. If we denote ab, ba by e, f, respectively, we obtain the Cayley table

	0	a	b	e	f
0	0	0	0	0	0
a	0	0	e	0	a
b	0	f	0	b	0
e	0	a	0	e	0
f	0	0	b	0	f

The semigroup B_2 can alternatively be realised by means of 2×2 matrices as $\{O, A, B, E, F\}$, where

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

where the operation is matrix multiplication, or by means of partial maps of the set $\{1, 2\}$, where

$$0 = \emptyset,\ a = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ b = \begin{pmatrix} 2 \\ 1 \end{pmatrix},\ e = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ f = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

1.7 IDEALS AND REES CONGRUENCES

The reader is probably familiar with the way in which morphisms in ring theory are associated with ideals. The corresponding link in semigroup theory is, as we shall see, less satisfactory, but there is one type of morphism, here called a *Rees morphism*, that does correspond very closely to an ideal.

First, if I is a proper ideal of a semigroup S, then

$$\rho_I = (I \times I) \cup 1_S$$

is a congruence on S. To see this, notice that $x\ \rho_I\ y$ if and only if either $x = y$ or both x and y belong to I. It is then easy to verify that ρ_I is reflexive, symmetric, transitive and compatible. The quotient semigroup is

$$S/\rho_I = \{I\} \cup \big\{\{x\} : x \in S \setminus I\big\},$$

which it is convenient to regard as consisting of I together with the elements of $S \setminus I$. In S/ρ_I the product of two elements in $S \setminus I$ is the same as their product in S if this lies in $S \setminus I$; otherwise the product is I. Since the element I of S/ρ_I is the zero element of the semigroup, another useful way of thinking of S/ρ_I is as $(S \setminus I) \cup \{0\}$, where all products not falling in $S \setminus I$ are zero.

We shall call a congruence of this type a *Rees congruence*, and if a morphism $\phi : S \to T$ is such that $\ker \phi$ is a Rees congruence we shall say that ϕ is a *Rees morphism*. We shall normally write S/I rather than S/ρ_I, and when we talk of the *kernel* of a Rees morphism we shall mean the ideal I rather than the congruence ρ_I.

It is important to note that not every semigroup morphism is of this type. Groups are semigroups in good standing, but a morphism $\phi : G \to H$ between two non-trivial groups cannot possibly be a Rees morphism, since G has no proper ideals and H has no zero element.

The next result requires only routine verification and its proof is omitted:

Proposition 1.7.1 *Let I be a proper ideal of a semigroup S. Let $\mathcal{A}$ be the set of ideals of S containing I and let $\mathcal{B}$ be the set of ideals of S/I. Then the map $\theta : J \mapsto J/I \quad (J \in \mathcal{A})$ is an inclusion-preserving bijection from $\mathcal{A}$ onto $\mathcal{B}$.* □

1.8 LATTICES OF EQUIVALENCES AND CONGRUENCES*

We have seen that, while the study of lattices of equivalences is in general hampered by the somewhat complicated way (Proposition 1.5.11) in which the join of two equivalences is formed, a useful simplification (Corollary 1.5.12) takes place if the equivalences commute. For this reason it is of interest to record the following result:

Proposition 1.8.1 *If G is a group, then $\rho \circ \sigma = \sigma \circ \rho$ for any two congruences ρ, σ on G.*

Proof Let $(a, b) \in \rho \circ \sigma$. Then there exists c in G such that $(a, c) \in \rho$, $(c, b) \in \sigma$. It follows that

$$a = cc^{-1}a \equiv bc^{-1}a \ (\mathrm{mod}\ \sigma),$$

and

$$bc^{-1}a \equiv bc^{-1}c = b \ (\mathrm{mod}\ \rho).$$

Thus $(a, b) \in \sigma \circ \rho$. We have shown that $\rho \circ \sigma \subseteq \sigma \circ \rho$, and by interchanging the roles of ρ and σ we may equally well show the opposite inclusion. □

This result can of course be proved in a more traditionally group-theoretic way. The connection between the general approach *via* congruences and the traditional approach *via* normal subgroups is outlined in the following proposition, whose proof is left to the reader:

Proposition 1.8.2 *Let G be a group with identity element e.*

(1) *If N is a normal subgroup of G, then*

$$\rho_N = \{(a, b) \in G \times G : ab^{-1} \in N\}$$

is a congruence on G. For each g in G the ρ_N-class $g\rho_N$ coincides with the coset Ng.

(2) *If ρ is a congruence on G then $N = e\rho$ is a normal subgroup of G, and $\rho = \rho_N$.*

(3) *If M, N are normal subgroups of G, then*

$$\rho_M \cap \rho_N = \rho_{M \cap N}, \quad \rho_M \vee \rho_N = \rho_M \circ \rho_N = \rho_{MN}. \qquad \square$$

The commuting of congruences in a group can therefore be derived as a consequence of the well-known group-theoretic result—see, for example, M. Hall (1959)—that normal subgroups commute.

A lattice $(L, \leq, \wedge, \vee)$ is called *modular* if, for all a, b, c in L,

$$a \leq c \ \Rightarrow \ (a \vee b) \wedge c = a \vee (b \wedge c). \tag{1.8.1}$$

Notice that in any lattice whatever, if $a \leq c$ then

$$a \leq a \vee b \ \text{ and } \ a \leq c,$$

and so $a \leq (a \vee b) \wedge c$; also

$$b \wedge c \leq b \leq a \vee b \ \text{ and } \ b \wedge c \leq c,$$

and so $b \wedge c \leq (a \vee b) \wedge c$. Hence

$$a \vee (b \wedge c) \leq (a \vee b) \wedge c,$$

and so we establish that a lattice is modular if we merely show the opposite inequality, that for all a, b, c in L,

$$a \leq c \Rightarrow (a \vee b) \wedge c \leq a \vee (b \wedge c). \qquad (1.8.2)$$

Proposition 1.8.3 *Let $\mathcal{K}$ be a sublattice of the lattice $(\mathcal{C}(S), \subseteq, \cap, \vee)$ of congruences of a semigroup S, and suppose that $\rho \circ \sigma = \sigma \circ \rho$ for all ρ, σ in $\mathcal{K}$. Then $\mathcal{K}$ is a modular lattice.*

Proof By saying that $\mathcal{K}$ is a sublattice of $(\mathcal{C}(S), \subseteq, \cap, \vee)$ we mean that $\mathcal{K}$ is closed under the operations $\cap$ and $\vee$. Let $\alpha, \beta, \gamma \in \mathcal{K}$ be such that $\alpha \subseteq \gamma$, and let

$$(x, y) \in (\alpha \vee \beta) \cap \gamma = (\alpha \circ \beta) \cap \gamma.$$

Then $(x, y) \in \gamma$, and there exists z in S such that $(x, z) \in \alpha$ and $(z, y) \in \beta$. Since $\alpha \subseteq \gamma$ and γ is an equivalence, we deduce that $(z, y) \in \gamma$. We thus have $(x, z) \in \alpha$ and $(z, y) \in \beta \cap \gamma$, from which we deduce that

$$(x, y) \in \alpha \circ (\beta \cap \gamma) = \alpha \vee (\beta \cap \gamma).$$

We have shown that

$$(\alpha \vee \beta) \cap \gamma \subseteq \alpha \vee (\beta \cap \gamma),$$

and so $\mathcal{K}$ is modular as required. □

Corollary 1.8.4 *The lattice of congruences on a group is modular.* □

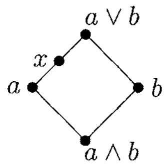
This result breaks down even for relatively small semigroups. In Exercise 21 below an example is given of a semigroup with four elements having a non-modular lattice of congruences. Indeed the lattice of congruences of that semigroup does not even have the weaker property of being *semimodular*. To explain the concept of semimodularity, first consider two elements a and b in a lattice L. We say that a *covers* b, and write $a \succ b$, if $a > b$ and if there is no x in L such that $a > x > b$. A lattice L is said to be (*upper*) *semimodular* if, for all a, b in L,

$$a \succ a \wedge b \text{ and } b \succ a \wedge b \Rightarrow a \vee b \succ a \text{ and } a \vee b \succ b.$$

We now justify our description of semimodularity as a weaker property than modularity:

Proposition 1.8.5 *Every modular lattice is semimodular.*

Proof Let L be a modular lattice, and suppose that a, b in L are such that $a \succ a \wedge b$ and $b \succ a \wedge b$. Suppose that x (in L) is such that $a \leq x < a \vee b$.



To show that $a \vee b \succ a$ we must show that $x = a$. Since $x \geq a$ it is clear that $x \wedge b \geq a \wedge b$. Also $x \wedge b \leq (a \vee b) \wedge b = b$, and so

$$a \wedge b \leq x \wedge b \leq b.$$

Now $x \wedge b \neq b$, for $x \wedge b = b$ would imply that $x \geq b$, and this, together with the assumption that $x \geq a$, would give the false conclusion that $x \geq a \vee b$. Hence

$$a \wedge b \leq x \wedge b < b,$$

and from the assumption that $b \succ a \wedge b$ we deduce that $x \wedge b = a \wedge b$. By modularity it now follows, since $a \leq x$, that

$$x = (a \vee b) \wedge x = a \vee (b \wedge x) = a \vee (a \wedge b) = a.$$

We have shown that $a \vee b \succ a$, and it follows in a similar way that $a \vee b \succ b$. □

For further lattice-theoretic results see, for example Dubreil-Jacotin *et al.* (1953). We shall here pursue only those aspects of the subject that have a direct bearing on the aspects of semigroup theory we shall be studying. The concept of semimodularity is important for us because of the following result:

Proposition 1.8.6 *The lattice* $(\mathcal{E}(X), \subseteq, \cap, \vee)$ *of equivalences on a set* X *is semimodular.*

Proof Let ρ, σ be equivalences on X such that $\rho \succ \rho \cap \sigma$ and $\sigma \succ \rho \cap \sigma$, and suppose that the $\rho \cap \sigma$-classes are

$$A_1, A_2, A_3, \ldots$$

Then there is exactly one ρ-class that is the union of two $\rho \cap \sigma$-classes, and so there is no loss of generality in supposing that the ρ-classes are

$$A_1 \cup A_2, A_3, A_4, \ldots$$

The same argument also applies to σ, but $\sigma \neq \rho$, and so there are two distinct cases: either the σ-classes are

$$A_1, A_2, A_3 \cup A_4, A_5, \ldots,$$

or they are

$$A_1, A_2 \cup A_3, A_4, \ldots$$

In the former case the $\rho \vee \sigma$-classes are

$$A_1 \cup A_2, A_3 \cup A_4, A_5, \ldots,$$

while in the latter case they are

$$A_1 \cup A_2 \cup A_3, A_4, A_5, \ldots$$

In either case $\rho \vee \sigma \succ \rho$ and $\rho \vee \sigma \succ \sigma$. □

If $|X| \geq 5$ the lattice $\mathcal{E}(X)$ is not modular. See Exercise 22.

If $L_1, L_2, \ldots, L_n$ are lattices, the cartesian product $L = L_1 \times L_2 \times \ldots \times L_n$ becomes a partially ordered set if we define

$$(x_1, x_2, \ldots, x_n) \leq (y_1, y_2, \ldots, y_n)$$

if and only if $x_1 \leq y_1$, $x_2 \leq y_2$, $\ldots$, $x_n \leq y_n$. It is even a lattice, with

$$(x_1, x_2, \ldots, x_n) \wedge (y_1, y_2, \ldots, y_n) = (x_1 \wedge y_1, x_2 \wedge y_2, \ldots, x_n \wedge y_n),$$
$$(x_1, x_2, \ldots, x_n) \vee (y_1, y_2, \ldots, y_n) = (x_1 \vee y_1, x_2 \vee y_2, \ldots, x_n \vee y_n).$$

We say that L is the *direct product* of $L_1, L_2, \ldots, L_n$. It is easy to see that

$$(a_1, a_2, \ldots, a_n) \succ (b_1, b_2, \ldots, b_n)$$

in L if and only if $a_i \succ b_i$ in L_i for some i in $\{1, 2, \ldots, n\}$ and $a_j = b_j$ for all $j \neq i$. It is now not hard to deduce

Proposition 1.8.7 *The direct product $L_1 \times L_2 \times \ldots \times L_n$ of semimodular lattices $L_1, L_2, \ldots, L_n$ is semimodular.*

Proof Let x, y be elements of $L = L_1 \times L_2 \times \ldots \times L_n$ such that $x \succ x \wedge y$, $y \succ x \wedge y$, and suppose that

$$x \wedge y = (z_1, z_2, \ldots, z_n).$$

Then $x = (x_1, x_2, \ldots, x_n)$, where $x_i \succ z_i$ in L_i for some i in $\{1, 2, \ldots, n\}$, and $x_k = z_k$ otherwise. Equally, $y = (y_1, y_2, \ldots, y_n)$, where $y_j \succ z_j$ for some j in $\{1, 2, \ldots, n\}$, and $x_k = z_k$ otherwise. If $i \neq j$ then $x \vee y = (t_1, t_2, \ldots, t_n)$, where

$$t_i = x_i,\ t_j = y_j \ \text{ and } \ t_k = z_k \text{ otherwise;}$$

in this case it is immediate that $x \vee y \succ x$, $x \vee y \succ y$. If $i = j$ then $x \vee y = (u_1, u_2, \ldots, u_n)$, where

$$u_i = x_i \vee y_i \ \text{ and } \ u_k = z_k \text{ otherwise.}$$

In this case we have $u_i \succ x_i$, $u_i \succ y_i$ in L_i, since L_i is semimodular, and it again follows that $x \vee y \succ x$, $x \vee y \succ y$ in L. □

1.9 EXERCISES

1. An element e of a semigroup S is called a *left identity* if $ex = x$ for every x in S, and a *right identity* if $xe = x$ for every x in S. An

element z of S is called a *left zero* if $zx = z$ for all x in S, and a *right zero* if $xz = z$ for all x in S.

(a) Show that, if S has a left identity e and a right identity f, then $e = f$ and e is the unique two-sided identity for S.

(b) Show that, if S has a left zero z and a right zero u, then $z = u$ and z is the unique two-sided zero for S.

(c) Give an example of a semigroup having two (at least) left identities and two (at least) right zeros.

2. Show that the definitions of a group given by equations (1.1.3) and (1.1.4) are equivalent.

3. Let X be a countably infinite set and let S be the set of one-one maps $\alpha : X \to X$ with the property that $X \setminus X\alpha$ is infinite.
   (a) Show that S is a subsemigroup of $\mathcal{T}_X$.
   (b) Show that for each α in S there is a one-one correspondence between $X \setminus X\alpha$ and $X\alpha \setminus X\alpha^2$.
   (c) Deduce that S has no idempotent elements.

   S is called the *Baer–Levi semigroup*.

4. Show that a semigroup S is a rectangular band if and only if

$$(\forall a, b \in S)\ ab = ba \ \Rightarrow\ a = b.$$

5. Let $\phi : S \to T$ be a morphism, where S and T are semigroups.
   (a) Show that the image under ϕ of an idempotent in S is an idempotent in T.
   (b) Show that the image under ϕ of a subsemigroup of S is a subsemigroup of T.
   (c) Show that if ϕ is onto then the image under ϕ of a right ideal [left ideal, ideal] of S is a right ideal [left ideal, ideal] of T.
   (d) Show that if ϕ is onto then the image under ϕ of the identity [zero] of S is the identity [zero] of T.
   (e) Show that in (c) and (d) the hypothesis that ϕ is onto cannot be removed.

6. A *permutation* σ of $\{1, 2, \ldots, n\}$ is defined as a bijection

$$\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}.$$

Denote by S_n the symmetric group $\mathcal{G}_{\{1,2,\ldots,n\}}$ consisting of all such permutations. In elementary books on group theory it is shown that every permutation is expressible as a composition of 'disjoint cycles'. (A *cycle*, written $(a_1\, a_2 \ldots a_k)$, where $a_1, a_2, \ldots, a_k$ are distinct elements of $\{1, 2, \ldots, n\}$, is a map ϕ defined by

$$a_i\phi = a_{i+1} \quad (i = 1, 2, \ldots, k-1), \quad a_k\phi = a_1,$$

$$x\phi = x \quad (x \notin \{a_1, a_2, \ldots, a_k\}),$$

and two cycles $(a_1 a_2 \ldots a_k)$ and $(b_1 b_2 \ldots b_l)$ are said to be *disjoint* if the sets $\{a_1, a_2, \ldots, a_k\}$ and $\{b_1, b_2, \ldots, b_l\}$ are disjoint.) A cycle $(a_1\, a_2)$ of length 2 is called a *transposition*. To avoid trivialities, suppose that $n \geq 3$.

(a) Show that if $k \geq 3$ then

$$(a_1\, a_2 \ldots a_k) = (a_1\, a_2) \circ (a_1\, a_3) \circ \cdots \circ (a_1\, a_k),$$

and deduce that the group S_n is generated by the set of all transpositions.

(b) Consider the cycles

$$\tau = (1\,2), \quad \zeta = (1\,2 \ldots n).$$

Show that

$$\zeta^{-1} = \zeta^{n-1}.$$

Show that

$$\zeta^{-1} \circ \tau \circ \zeta = (23),$$

and more generally that

$$\zeta^{-i+1} \circ \tau \circ \zeta^{i-1} = (i\; i+1) \quad (i = 1, 2, \ldots, n-1).$$

Next, show that, for $j = 2, 3, \ldots, n-1$,

$$(j\; j+1) \circ (j-1\; j) \circ \cdots \circ (23) \circ (12) \circ (23) \circ \cdots \circ (j\; j+1) = (1\; j+1),$$

and that, for $i = 1, 2, \ldots, n-1$ and $j = 1, 2, \ldots, n-i$,

$$\zeta^{-i+1} \circ (1\; j+1) \circ \zeta^{i-1} = (i\; i+j).$$

(c) Deduce that $S_n = \langle \tau, \zeta \rangle$.

7. Let T_n be the full transformation semigroup $\mathcal{T}_{\{1,2,\ldots,n\}}$, where $n \geq 3$. Let π denote the element of T_n given by

$$1\pi = 2, \quad x\pi = x \quad (x = 2, 3, \ldots, n),$$

and let ζ, τ be the permutations defined in the previous exercise. For $i \neq j$ in $\{1, 2, \ldots, n\}$, let $\|i\, j\|$ denote the map ϕ for which

$$i\phi = j, \quad x\phi = x \quad (x \neq i).$$

Thus, in particular, $\pi = \|1\,2\|$.

(a) Prove the identities

$$\begin{aligned}
(1\, i) \circ \|1\,2\| \circ (1\, i) &= \|i\,2\| \quad (i \geq 3),\\
(2j) \circ \|1\,2\| \circ (2\, j) &= \|1\, j\| \quad (j \geq 3),\\
(1i) \circ (2\, j) \circ \|i\, j\| \circ (2\, j) \circ (1\, i) &= \|i\, j\| \quad (i, j \geq 3,\ i \neq j),\\
(i\, j) \circ \|i\, j\| \circ (i\, j) &= \|j\, i\| \quad (i, j \geq 1,\ i \neq j).
\end{aligned}$$

(b) Let $\phi \in T_n$ and let $|\operatorname{im}\phi| = r \leq n-1$. Let i, j (with $i \neq j$) be such that $i\phi = j\phi$, and let

$$z \in \{1, 2, \ldots, n\} \setminus \operatorname{im}\phi.$$

Show that

$$\phi = \|i\,j\| \circ \hat{\phi},$$

where

$$i\hat{\phi} = z, \quad k\hat{\phi} = k\phi \quad (k \neq i).$$

(c) Deduce that $T_n = \langle \zeta, \tau, \pi \rangle$.

8. Show that if S is a semigroup then the (unextended) right regular representation $\alpha : S \to \mathcal{T}_S$ may or may not be faithful. More specifically, show that it is faithful if S is a right zero semigroup but is not faithful if S is a left zero semigroup.

9. A semigroup may be isomorphic to a semigroup of maps in more than one way. For example, for the semigroup $S = \{e, a, x, y\}$ with Cayley table

	e	a	x	y
e	e	a	x	y
a	a	e	x	y
x	x	y	x	y
y	y	x	x	y

show that the map ϕ given by

$$e\phi = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \; a\phi = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \; x\phi = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \; y\phi = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix},$$

embeds S in $\mathcal{T}_{\{1,2\}}$. Show also that ψ, defined by

$$e\psi = \begin{pmatrix} e & a & x & y \\ e & a & x & y \end{pmatrix}, \quad a\psi = \begin{pmatrix} e & a & x & y \\ a & e & y & x \end{pmatrix},$$
$$x\psi = \begin{pmatrix} e & a & x & y \\ x & x & x & x \end{pmatrix}, \quad y\psi = \begin{pmatrix} e & a & x & y \\ y & y & y & y \end{pmatrix},$$

embeds S in $\mathcal{T}_{\{e,a,x,y\}}$. (Notice that ψ is the right regular representation of S.)

10. Let $S = \langle a \rangle = M(m, r)$, where $m > 1$. Show that $S \setminus S^2 = \{a\}$, and deduce that the generator a of a finite monogenic semigroup S is uniquely determined by S unless S is a group.

11. Show that, if $|X| = n$, then

$$|\mathcal{B}_X| = 2^{n^2}.$$

12. Let ϕ be an element of $\mathcal{P}_X$, the semigroup of all partial maps of the set X. Let $Z = X \cup \{0\}$, where $0 \notin X$, and define $\phi^* : Z \to Z$ by

$$x\phi^* = \begin{cases} x\phi \text{ if } x \in \operatorname{dom}\phi \\ 0 \quad \text{otherwise.} \end{cases}$$

Show that $\phi \mapsto \phi^*$ is an isomorphism from $\mathcal{P}_X$ onto the subsemigroup of $\mathcal{T}_Z$ consisting of all maps $\alpha : Z \to Z$ for which $0\alpha = 0$.

Deduce that if X is a finite set containing n elements, then

$$|\mathcal{P}_X| = (n+1)^n.$$

13. Consider the partial transformation semigroup $P_n = \mathcal{P}_{\{1,2,\ldots,n\}}$, let ζ, τ and π be as in Exercises 6 and 7, and let ξ be the partial map

$$\begin{pmatrix} 2\;3\;\ldots\;n \\ 2\;3\;\ldots\;n \end{pmatrix}.$$

For each subset Y of $X = \{1,2,\ldots,n\}$, define ξ_Y to be $1_{X\setminus Y}$; thus $\xi = \xi_{\{1\}}$.

(a) Show that $(1\,i)\xi(1\,i) = \xi_{\{i\}}$, $\quad (i = 2,3,\ldots,n)$, and that

$$\xi_Y = \prod_{i\in Y} \xi_{\{i\}}.$$

(b) For each α in P_n, let $\hat{\alpha}$, the *completion* of α, be given by

$$x\hat{\alpha} = \begin{cases} x\alpha & \text{if } x \in \operatorname{dom}\alpha \\ x & \text{otherwise.} \end{cases}$$

Show that, if $X \setminus \operatorname{dom}\alpha = Y$, then $\alpha = \xi_Y\hat{\alpha}$.

(c) Deduce that $P_n = \langle \zeta, \tau, \pi, \xi \rangle$.

14. Let S be a semigroup, and let R, C^l, C^r, C and T denote, respectively, the sets of reflexive, left compatible, right compatible, compatible and transitive relations on S.

(a) Show that

$$C \cap R \subseteq (C^l \cap C^r) \cap R, \quad (C^l \cap C^r) \cap T \subseteq C \cap T.$$

(b) Deduce that

$$C \cap (R \cap T) = (C^l \cap C^r) \cap (R \cap T),$$

that is to say, show that a reflexive and transitive relation is compatible if and only if it is both left and right compatible.

(c) Show that each of the inclusions in (a) can be strict. [Hint. Let $S = \{1,2\} \times \{1,2\}$ be a rectangular band. Write $x = (1,1)$, $y = (1,2)$, $z = (2,1)$, $t = (2,2)$, and consider the relations

$$\mathbf{R} = 1_S \cup \{(x,y),\ (x,z),\ (y,t),\ (z,t)\}, \quad \mathbf{S} = \{(x,y)\}$$

on S.]

15. Let X be a finite set with n elements, and let $S(n,r)$ $(1 \leq r \leq n)$ be the number of equivalences ρ on X such that $|X/\rho| = r$. Show that

$$S(n,1) = S(n,n) = 1,$$

$$S(n,r) = S(n-1,r-1) + r\,S(n-1,r) \quad (2 \leq r \leq n-1),$$

and use this information to calculate $S(n,r)$ for $1 \leq r \leq n \leq 6$. (The numbers $S(n,r)$ are the *Stirling numbers of the second kind.* See, for example, van Lint and Wilson (1992).)

16. If A and B are sets then a *relation ρ from A to B* may be defined as a subset of $A \times B$. For each a in A we then define $a\rho$ in the obvious way:

$$a\rho = \{b \in B : (a,b) \in \rho\}.$$

If S and T are semigroups, then a subset μ of $S \times T$ is called a *relational morphism from S to T* if

(RM1) $(\forall a \in S)\ a\mu \neq \emptyset$;

(RM2) $(\forall a,b \in S)\ (a\mu)(b\mu) \subseteq (ab)\mu$.

It is called *injective* if in addition

(RM3) $(\forall a,b \in S)\ a\mu \cap b\mu \neq \emptyset \ \Rightarrow\ a\mu = b\mu$.

Show that every relational morphism is a subsemigroup of the direct product $S \times T$.

We say that S *divides* T if there exists a subsemigroup U of T and a morphism ψ from U onto S. (Thus S is a quotient of a subsemigroup of T.) Show that S divides T if and only if there exists an injective relational morphism from S to T.

17. For a commutative semigroup S, define the relation θ_n^S $(n \geq 1)$ by

$$a\,\theta_n^S\,b \text{ if and only if } (\forall x \in S^n)\ xa = xb.$$

(a) Show that θ_n^S is a congruence on S, and that

$$\theta_1^S \subseteq \theta_2^S \subseteq \cdots$$

(b) Show that $\theta_n^S = 1_S$ for all n if S is a monoid.

(c) For $n = 1,2,\ldots$, denote S/θ_n^S by S_n. Show that, for all $n \geq 2$,

$$S_n \simeq S_{n-1}/\theta_1^{S_{n-1}}.$$

(d) Let $S = \langle a \rangle = M(m,r)$ be a finite monogenic semigroup, where $m > 1$. Show that

$$S/\theta_1^S \simeq M(m-1,r),$$

and deduce that

$$S/\theta_n^S \simeq M(m-n,r)$$

for all $n < m$. Show also that S/θ_n^S is isomorphic to the cyclic group of order r for all $n \geq m$.

18. Let I, J be ideals of a semigroup S such that $I \subseteq J$. Show that

$$S/J \simeq (S/I)/(J/I).$$

19. Let I, J be ideals of a semigroup S. Show that $I \cap J$, $I \cup J$ are ideals of S. (Notice that $IJ \subseteq I \cap J$ and so $I \cap J \neq \emptyset$.) Show also that

$$(I \cup J)/J \simeq I/(I \cap J).$$

20. Let $\rho_{m,r}$ $(m, r \geq 1)$ be the congruence $\{(a^m, a^{m+r})\}^\#$ on the free monogenic semigroup a^+. (Thus a^+/ρ is the monogenic semigroup $M(m,r)$.)
    (a) Show that $(a^p, a^q) \in \rho$ if and only if $p, q \geq m$ and $p \equiv q \pmod r$.
    (b) Show that, for all $m, n, r, s \geq 1$, $\rho_{m,r} \subseteq \rho_{n,s}$ if and only if $m \geq n$ and s divides r.
    (c) Deduce that, for all $m, n, r, s \geq 1$,

$$\rho_{m,r} \cap \rho_{n,s} = \rho_{\max(m,n),\mathrm{lcm}(r,s)},$$
$$\rho_{m,r} \vee \rho_{n,s} = \rho_{\min(m,n),\mathrm{hcf}(r,s)}.$$

(Here lcm stands for the least common multiple, and hcf for the highest common factor.)

21. Let $S = \{e, f, a, b\}$ be the semigroup with multiplication table

	e	a	f	b
e	e	a	f	b
a	a	e	b	f
f	f	b	f	b
b	b	f	b	f

Verify that the congruences on S are as follows:

1_S, with classes $\{e\}$, $\{a\}$, $\{f\}$, $\{b\}$;
σ, with classes $\{e, f\}$, $\{a, b\}$;
μ, with classes $\{e, a\}$, $\{f, b\}$;
ν, with classes $\{e\}$, $\{a\}$, $\{f, b\}$;
ω, with the single class $\{e, f, a, b\}$.

Draw a Hasse diagram for the lattice $(\mathcal{C}(S), \subseteq, \cap, \vee)$, and deduce that $\mathcal{C}(S)$ is not semimodular.

22. In Proposition 1.8.6 it is shown that the lattice of equivalences on a set is semimodular. Let $X = \{x_1, x_2, x_3, \ldots\}$, with $|X| \geq 5$. Let α, β, γ be equivalences on X with classes as follows

$$\begin{aligned}
\alpha &: \{x_1, x_2\},\ \{x_3, x_4\},\ \{x_5\},\ \{x_6\}, \ldots;\\
\beta &: \{x_1, x_3\},\ \{x_2, x_5\},\ \{x_4\},\ \{x_6\}, \ldots;\\
\gamma &: \{x_1, x_2\},\ \{x_3, x_4, x_5\},\ \{x_6\}, \ldots
\end{aligned}$$

Show that $\alpha \subseteq \gamma$, but that $(\alpha \vee \beta) \cap \gamma \neq \alpha \vee (\beta \cap \gamma)$, and deduce that $\mathcal{E}(X)$ is not modular.

1.10 NOTES

Much of this chapter is 'folklore', and there is little to be gained in attempting to track down original sources. The notion of the syntactic congruence (Section 1.5) goes back in effect to Dubreil (1941), and the idea of a Rees quotient (Section 1.7) to Rees (1940). The bicyclic semigroup (Section 1.6) can be traced to Lyapin (1953) and Andersen (1952) The semimodularity of the lattice of equivalences on a set appears in Dubreil-Jacotin *et al.*(1953).

The semigroup described in Exercise 3, usually called the *Baer–Levi semigroup*, appears in Baer and Levi (1932), where it resolved a question in group axiomatics. A fuller and more general account of Baer–Levi semigroups appears in Clifford and Preston (1967). The isomorphism described in Exercise 12 was first observed by Vagner (1956). The idea of a relational morphism appears in Eilenberg (1976) within the chapters written by Tilson, and is much used in applications of semigroups to language theory—see Pin (1986). It might have been more logical to call it a morphic relation, but the term 'relational morphism' is now standard.

The congruences θ_n^S featuring in Exercise 17 have been studied by Kopamu (1995).

2

Green's equivalences; regular semigroups

The notion of *ideal* mentioned in the last chapter leads naturally to the consideration of certain equivalence relations on a semigroup. These equivalences, first studied by J. A. Green (1951), have played a fundamental role in the development of semigroup theory. They are concerned with mutual divisibility of various kinds, and all of them reduce to the universal equivalence in a group.

Regular semigroups—the definition is copied from von Neumann's (1936) definition of a regular ring—are particularly amenable to analysis using Green's equivalences. The chapter goes on to consider such semigroups, and ends with a brief account of 'sandwich sets', a concept due to Nambooripad (1974, 1975).

2.1 GREEN'S EQUIVALENCES

If a is an element of a semigroup S, the smallest left ideal of S containing a is $Sa \cup \{a\}$, which, as in (1.1.2), it is convenient to denote by S^1a. We shall call it the *principal left ideal generated by* a. An equivalence $\mathcal{L}$ on S is defined by the rule that $a \mathrel{\mathcal{L}} b$ if and only if a and b generate the same principal left ideal, that is, if and only if $S^1a = S^1b$.

Similarly, we define the equivalence $\mathcal{R}$ by the rule that $a \mathrel{\mathcal{R}} b$ if and only if $aS^1 = bS^1$.

An alternative characterization, making the 'mutual divisibility' aspect of these equivalences more apparent, is given in the following proposition:

Proposition 2.1.1 *Let a, b be elements of a semigroup S. Then $a \mathrel{\mathcal{L}} b$ if and only if there exist x, y in S^1 such that $xa = b$, $yb = a$. Also, $a \mathrel{\mathcal{R}} b$ if and only if there exist u, v in S^1 such that $au = b$, $bv = a$.* □

Another immediate property of $\mathcal{L}$ and $\mathcal{R}$ is as follows (see Section 1.5):

Proposition 2.1.2 *$\mathcal{L}$ is a right congruence and $\mathcal{R}$ is a left congruence.* □

We have seen in Section 1.4 that the intersection of two equivalences is again an equivalence. Since the intersection of $\mathcal{L}$ and $\mathcal{R}$ is of great importance in the development of the theory, we reserve for it the letter $\mathcal{H}$. The join $\mathcal{L} \vee \mathcal{R}$ is also of great importance, and we denote it by $\mathcal{D}$. As we saw in Section 1.5, the join of two equivalences can be rather hard to describe, but we are saved from these difficulties by the following fortunate occurrence:

Proposition 2.1.3 *The relations $\mathcal{L}$ and $\mathcal{R}$ commute.*

Proof Let S be a semigroup, let $a, b \in S$, and suppose that $(a, b) \in \mathcal{L} \circ \mathcal{R}$. Then there exists c in S such that $a \,\mathcal{L}\, c$ and $c \,\mathcal{R}\, b$. That is, there exist x, y, u, v in S^1 such that

$$xa = c, \quad cu = b,$$
$$yc = a, \quad bv = c.$$

If we now write d for the element ycu of S, we see that

$$au = ycu = d, \quad dv = ycuv = ybv = yc = a;$$

hence $a \,\mathcal{R}\, d$. Also,

$$yb = ycu = d, \quad xd = xycu = xau = cu = b,$$

and so $d \,\mathcal{L}\, b$. We deduce that $(a, b) \in \mathcal{R} \circ \mathcal{L}$. We have shown that $\mathcal{L} \circ \mathcal{R} \subseteq \mathcal{R} \circ \mathcal{L}$; the reverse inclusion follows in a similar way. □

By Corollary 1.5.12 it now follows that

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

and this means that $\mathcal{D}$ is a great deal easier to handle than one might have expected.

Our final equivalence is the two-sided analogue of $\mathcal{L}$ and $\mathcal{R}$. The principal two-sided ideal of S generated by a is S^1aS^1, and we define the equivalence $\mathcal{J}$ by the rule that $a \,\mathcal{J}\, b$ if and only if $S^1aS^1 = S^1bS^1$, that is to say, if and only if there exist x, y, u, v in S^1 such that

$$xay = b, \quad ubv = a.$$

It is immediate that $\mathcal{L} \subseteq \mathcal{J}$ and $\mathcal{R} \subseteq \mathcal{J}$. Hence, since $\mathcal{D}$ is the smallest equivalence containing $\mathcal{L}$ and $\mathcal{R}$, we must have

$$\mathcal{D} \subseteq \mathcal{J}. \tag{2.1.1}$$

An example showing that this inclusion may be strict was given by Green (1951); a more striking example, from Andersen (1952), appears as Exercise 1 below. In certain classes of semigroups we do have equality: in a group G we have

$$\mathcal{H} = \mathcal{L} = \mathcal{R} = \mathcal{D} = \mathcal{J} = G \times G,$$

and in a commutative semigroup we have

$$\mathcal{H} = \mathcal{L} = \mathcal{R} = \mathcal{D} = \mathcal{J}.$$

Less trivially, we have a result which implies in particular that $\mathcal{D} = \mathcal{J}$ in every finite semigroup:

Proposition 2.1.4 *If S is a periodic semigroup then $\mathcal{D} = \mathcal{J}$.*

Proof Suppose that a, b in S are such that $a \mathcal{J} b$. Then there exist x, y, u, v in S^1 such that

$$xay = b, \quad ubv = a. \tag{2.1.2}$$

To prove the desired result we need to find an element c in S such that $a \mathcal{L} c$, $c \mathcal{R} b$. It follows easily from the equations (2.1.2) that

$$a = (ux)a(yv) = (ux)^2 a(yv)^2 = (ux)^3 a(yv)^3 = \cdots,$$
$$b = (xu)b(vy) = (xu)^2 b(vy)^2 = (xu)^3 b(vy)^3 = \cdots$$

Since S is periodic we can, by Proposition 1.2.3, find an m for which $(ux)^m$ is idempotent. Then, if we let $c = xa$, we find that

$$a = (ux)^m a(yv)^m = (ux)^m (ux)^m a(yv)^m = (ux)^m a = (ux)^{m-1} uc,$$

and so $a \mathcal{L} c$. Also, $cy = xay = b$, and if we choose n so that $(vy)^n$ is idempotent, we have

$$\begin{aligned} c = xa = x(ux)^{n+1} a(yv)^{n+1} &= (xu)^{n+1} xay(vy)^n v \\ = (xu)^{n+1} b(vy)^{2n} v &= (xu)^{n+1} b(vy)^{n+1}(vy)^{n-1} v \\ = b(vy)^{n-1} v. & \end{aligned}$$

Hence $c \mathcal{R} b$ as required. □

Some preliminaries are necessary before we can describe another important class of semigroups in which $\mathcal{D} = \mathcal{J}$. First, since the Green equivalences occur so often, it is worth making an exception to the general notational convention announced in Section 1.4 about equivalence classes: the $\mathcal{L}$-class [$\mathcal{R}$-class, $\mathcal{H}$-class, $\mathcal{D}$-class, $\mathcal{J}$-class] containing the element a will be denoted by L_a [R_a, H_a, D_a, J_a]. Since $\mathcal{L}$, $\mathcal{R}$ and $\mathcal{J}$ are defined in terms of ideals, the inclusion order among these ideals induces a partial order among the equivalence classes:

$$\left.\begin{array}{lll} L_a \leq L_b & \text{if} & S^1 a \subseteq S^1 b; \\ R_a \leq R_b & \text{if} & aS^1 \subseteq bS^1; \\ J_a \leq J_b & \text{if} & S^1 a S^1 \subseteq S^1 b S^1. \end{array}\right\} \tag{2.1.3}$$

We may thus regard $S/\mathcal{L}$, $S/\mathcal{R}$ and $S/\mathcal{J}$ as partially ordered sets. Notice that, for all a in S and for all x, y in S^1,

$$L_{xa} \leq L_a, \quad R_{ax} \leq R_a, \quad J_{xay} \leq J_a. \tag{2.1.4}$$

Notice also that

$$L_a \leq L_b \Rightarrow J_a \leq J_b, \quad R_a \leq R_b \Rightarrow J_a \leq J_b. \tag{2.1.5}$$

We shall say that a semigroup S *satisfies the condition* $\min_L$, $\min_R$ or $\min_J$ according as the partially ordered set $S/\mathcal{L}$, $S/\mathcal{R}$ or $S/\mathcal{J}$ satisfies the minimal condition. (See Section 1.3.) These conditions are of course equivalent, respectively, to the minimal conditions on principal left ideals, principal right ideals and principal two-sided ideals, and are weaker than the corresponding conditions on (not necessarily principal) ideals. See the notes at the end of this chapter for further information on minimal conditions.

Proposition 2.1.5 *If S is a semigroup satisfying $\min_L$ and $\min_R$, then $\mathcal{D} = \mathcal{J}$.*

Proof Suppose that S satisfies $\min_L$ and $\min_R$. Then so also does S^1, for S^1 has exactly the same principal left and right ideals as S, except (in the case where $S^1 \neq S$) for S^1 itself. So we may assume that S has an identity element.

Suppose now that $a \mathcal{J} b$, so that there exist p, q, r, s in S such that

$$paq = b, \quad rbs = a.$$

It follows that the set

$$X = \{x \in S : (\exists y \in S)\ xay = b\}$$

is non-empty, and hence so also is the subset

$$\Lambda = \{L_x : x \in X\}$$

of $S/\mathcal{L}$. The condition $\min_L$ allows us to select a minimal element L_u in Λ, and we can then choose an element v in S such that $uav = b$.

Now $uruavsv = b$, and so $L_{uru} \in \Lambda$. From (2.1.4) we have $L_{uru} \leq L_u$, and it then follows by the minimality of L_u in Λ that

$$L_u = L_{uru} \leq L_{ru} \leq L_u.$$

Hence $ru \mathcal{L} u$. By Proposition 2.1.2 it follows that $ruav \mathcal{L} uav$, that is, $rb \mathcal{L} b$.

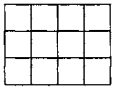
A similar argument using the condition $\min_R$ establishes that $bs \mathcal{R} b$, and then by Proposition 2.1.2 we deduce that $rbs \mathcal{R} rb$. We thus have $a \mathcal{R} rb$, $rb \mathcal{L} b$, and so $a \mathcal{D} b$ as required. □

2.2 THE STRUCTURE OF $\mathcal{D}$-CLASSES

Each $\mathcal{D}$-class in a semigroup S is a union of $\mathcal{L}$-classes and also a union of $\mathcal{R}$-classes. The intersection of an $\mathcal{L}$-class and an $\mathcal{R}$-class is either empty or is an $\mathcal{H}$-class. However, by the very definition of $\mathcal{D}$,

$$a \mathcal{D} b \iff R_a \cap L_b \neq \emptyset \iff L_a \cap R_b \neq \emptyset.$$

Hence it is convenient to visualize a $\mathcal{D}$-class as what Clifford and Preston (1961) have called an 'eggbox', in which each row represents an $\mathcal{R}$-class, each column represents an $\mathcal{L}$-class, and each cell represents an $\mathcal{H}$-class. (It is of course possible for the 'eggbox' to contain a single row or a single column of cells, or even to contain only one cell. Also, it may well be an infinite eggbox.)



If D is an arbitrary $\mathcal{D}$-class in a semigroup S, and if $a, b \in D$ are such that $a \; \mathcal{R} \; b$ (so that a and b are in the same row of the eggbox), then by definition of $\mathcal{R}$ there exist s, s' in S^1 such that

$$as = b, \quad bs' = a.$$

The right translation $\rho_s : S \to S$ thus maps a to b. In fact it maps L_a into L_b, for if $x \in L_a$ then $xs \; \mathcal{L} \; as$ by Proposition 2.1.2, and so $xs \in L_{as} = L_b$. Now it is just as easy to show that $\rho_{s'}$ maps L_b into L_a, and if we investigate the composition $\rho_s \rho_{s'} : L_a \to L_a$ we find that for any $x = ua$ in L_a,

$$x\rho_s\rho_{s'} = uass' = ubs' = ua = x.$$

Thus $\rho_s\rho_{s'}$ is the identity map of L_a, and we can show in a closely similar way that $\rho_{s'}\rho_s$ is the identity map on L_b. We deduce that $\rho_s|_{L_a}$ and $\rho_{s'}|_{L_b}$ are mutually inverse bijections from L_a onto L_b and L_b onto L_a, respectively.

We can say even more about these maps: if $x \in L_a$ then the element $y = x\rho_s$ of L_b has the property that

$$y = xs, \quad x = ys'.$$

Thus $y \; \mathcal{R} \; x$, and so the map ρ_s is *$\mathcal{R}$-class preserving.* It maps each $\mathcal{H}$-class in L_a in a one-one manner onto the corresponding ($\mathcal{R}$-equivalent) $\mathcal{H}$-class in L_b. Similar remarks apply to $\rho_{s'}$.

To summarize, we have proved the following result, usually known as Green's Lemma:

Lemma 2.2.1 *Let a, b be $\mathcal{R}$-equivalent elements in a semigroup S, and let s, s' in S^1 be such that*

$$as = b, \quad bs' = a.$$

Then the right translations $\rho_s|_{L_a}$, $\rho_{s'}|_{L_b}$ are mutually inverse $\mathcal{R}$-class preserving bijections from L_a onto L_b and L_b onto L_a, respectively. □

It is clear that this result has a left-right dual (also known as Green's Lemma) featuring the *left* translations $\lambda_t : x \mapsto tx$ and $\lambda_{t'} : y \mapsto t'y$:

Lemma 2.2.2 *Let a, b be $\mathcal{L}$-equivalent elements in a semigroup S, and let t, t' in S^1 be such that*

$$ta = b, \quad t'b = a.$$

Then the left translations $\lambda_t|_{R_a}$, $\lambda_{t'}|_{R_b}$ are mutually inverse $\mathcal{L}$-class preserving bijections from R_a onto R_b and R_b onto R_a, respectively. □

The combined effect of these two lemmas is as follows:

Lemma 2.2.3 *If a, b are $\mathcal{D}$-equivalent elements in a semigroup S, then $|H_a| = |H_b|$.*

Proof Let c be such that $a \,\mathcal{R}\, c$ and $c \,\mathcal{L}\, b$, with

$$as = c, \quad cs' = a, \quad tc = b, \quad t'b = c.$$

Then by the preceding lemmas $\rho_s|_{H_a}$ is a bijection onto H_c and $\lambda_t|_{H_c}$ is a bijection onto H_b. Thus $\rho_s\lambda_t : x \mapsto txs$ is a bijection from H_a onto H_b (with inverse $\lambda_{t'}\rho_{s'} : y \mapsto t'ys'$), and it follows that $|H_a| = |H_b|$. □

A more striking consequence of Lemmas 2.2.1 and 2.2.2 is concerned with the multiplicative properties of an $\mathcal{H}$-class. A preliminary lemma, which is just a specialization of Lemmas 2.2.1 and 2.2.2, is useful. We have seen that if $as \,\mathcal{R}\, a$, then $x \mapsto xs$ is a bijection from H_a onto H_{as}. So, in particular, if $as \,\mathcal{H}\, a$ then $x \mapsto xs$ is a bijection of H_a onto itself. This, together with the dual argument, gives

Lemma 2.2.4 *Let x, y be elements of a semigroup S. If $xy \in H_x$ then $\rho_y|_{H_x}$ is a bijection of H_x onto itself. If $xy \in H_y$ then $\lambda_x|_{H_y}$ is a bijection of H_y onto itself.* □

Then we have the following result, usually called Green's Theorem:

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

Proof Suppose that $H^2 \cap H \neq \emptyset$, so that there exist a, b in H such that $ab \in H$. By the lemma, ρ_b and λ_a are bijections of H onto itself. Hence $hb \in H$ and $ah \in H$ for every h in H. Again by the lemma it follows that λ_h and ρ_h are bijections of H onto itself. Hence $Hh = hH = H$ for every h in H. Certainly $H^2 = H$, and it follows from (1.1.7) that H is a subgroup of S. □

We now immediately deduce

Corollary 2.2.6 *If e is an idempotent in a semigroup S, then H_e is a subgroup of S. No $\mathcal{H}$-class in S can contain more than one idempotent.* □

2.3 REGULAR $\mathcal{D}$-CLASSES

An element a of a semigroup S is called *regular* if there exists x in S such that $axa = a$. The semigroup S is called *regular* if all its elements are regular. Groups are of course regular semigroups, but the class of regular semigroups is vastly more extensive than the class of groups. For example, every rectangular band B (see Theorem 1.1.3) is trivially regular, since $axa = a$ for all a, x in B.

If a is a regular element, with $axa = a$, and if $b \in L_a$, then there exist u, v in S^1 such that $ua = b$, $vb = a$, and so

$$b = ua = uaxa = bxa = b(xv)b,$$

and so b also is regular. The same conclusion applies to any element in R_a, and so we have

Proposition 2.3.1 *If a is a regular element of a semigroup S, then every element of D_a is regular.* □

If, then, D is a $\mathcal{D}$-class then either every element of D is regular or no element of D is regular; we call the $\mathcal{D}$-class *regular* if all its elements are regular, and *irregular* otherwise. This dichotomy does not apply to $\mathcal{J}$-classes in general, and of course a single semigroup may contain both regular and irregular elements. (See Exercise 25 in Chapter 5.)

Since idempotents e are regular ($eee = e$) it follows that every $\mathcal{D}$-class containing an idempotent is regular. Conversely, we can show that every regular $\mathcal{D}$-class must contain at least one idempotent. Indeed we can show more:

Proposition 2.3.2 *In a regular $\mathcal{D}$-class, each $\mathcal{L}$-class and each $\mathcal{R}$-class contains an idempotent.*

Proof If a is a member of a regular $\mathcal{D}$-class, and if x is such that $axa = a$, then xa is idempotent and $xa \mathrel{\mathcal{L}} a$. Similarly, ax is idempotent, and $ax \mathrel{\mathcal{R}} a$. □

Proposition 2.3.3 *Every idempotent e in a semigroup S is a left identity for R_e and a right identity for L_e.*

Proof If $a \in R_e$ then $a = ex$ for some x in S^1, and so

$$ea = e(ex) = e^2x = ex = a.$$

Similarly $be = b$ for all b in L_e. □

An idea of great importance in semigroup theory is that of an *inverse* of an element. If a is an element of a semigroup S, we say that a' is an *inverse* of a if

$$aa'a = a, \quad a'aa' = a'. \tag{2.3.1}$$

Notice that an element with an inverse is necessarily regular. Less obviously, every regular element has an inverse: if there exists x such that $axa = a$, then define $a' = xax$ and observe that

$$aa'a = axaxa = axa = a, \quad a'aa' = xaxaxax = xaxax = xax = a'.$$

An element a may well have more than one inverse. Indeed, in a rectangular band (see Theorem 1.1.3) every element is an inverse of every other element. So the idea of inverse under discussion here is substantially more general than a group inverse. (Of course, a group inverse *is* an inverse in the sense of (2.3.1).) We shall denote the set of inverses of an element a by $V(a)$.

The eggbox picture of a $\mathcal{D}$-class is exceedingly useful in the location of inverses. First, notice that if a is an element of a regular $\mathcal{D}$-class D then every inverse a' of a must lie in D, for $a \mathcal{R} aa'$ and $aa' \mathcal{L} a'$.

	L_a		$L_{a'}$	
R_a	a		aa'	
$R_{a'}$	$a'a$		a'	

We have proved part of

Theorem 2.3.4 *Let a be an element of a regular $\mathcal{D}$-class D in a semigroup S.*

(1) *If $a' \in V(a)$, then $a' \in D$ and the two $\mathcal{H}$-classes $R_a \cap L_{a'}$, $L_a \cap R_{a'}$ contain, respectively, the idempotents aa' and $a'a$.*
(2) *If b in D is such that $R_a \cap L_b$ and $L_a \cap R_b$ contain idempotents e, f, respectively, then H_b contains an inverse a^* of a such that $aa^* = e$, $a^*a = f$.*
(3) *No $\mathcal{H}$-class contains more than one inverse of a.*

Proof It remains to establish the second and third parts.

	L_a		L_b	
R_a	a		e	
R_b	f		a^*, b	

To prove (2), notice that from $a \mathcal{R} e$ it follows by Proposition 2.3.3 that $ea = a$. Similarly, from $a \mathcal{L} f$ it follows that $af = a$. Again from $a \mathcal{R} e$ it follows that there exists x in S^1 such that $ax = e$. Let $a^* = fxe$. Then

$$aa^*a = (af)x(ea) = axa = ea = a,$$

$$a^*aa^* = fx(eaf)xe = fx(ax)e = fxe^2 = fxe = a^*,$$

and so $a^* \in V(a)$. Also

$$aa^* = (af)xe = (ax)e = e^2 = e.$$

Further, since $a \mathcal{L} f$, there exists y in S^1 such that $ya = f$. Hence

$$a^*a = fxea = fxa = yaxa = yea = ya = f.$$

It now follows easily that

$$a^* \in L_e \cap L_f = L_b \cap R_b = H_b.$$

To prove part (3), suppose that both a' and a^* are inverses of a inside the single $\mathcal{H}$-class H_b. It follows that both aa' and aa^* are idempotents in the $\mathcal{H}$-class $R_a \cap L_b$. Hence $aa' = aa^*$ by Corollary 2.2.6. Similarly $a'a = a^*a$, and it now follows that

$$a^* = a^*aa^* = a^*aa' = a'aa' = a'. \qquad \square$$

Notice that this theorem allows us to locate the inverses of a regular element provided we know where the idempotents are. For example, in a finite semigroup we can say immediately that the number of inverses of an element s is the number of idempotents in R_a multiplied by the number of idempotents in L_a. This idea of deducing properties of the semigroup from facts about its idempotents is a recurring theme in the study of regular semigroups, and we shall have occasion to mention it several times. The following easy consequence of Theorem 2.3.4 will be of considerable use in later chapters:

Proposition 2.3.5 *Let e, f be idempotents in a semigroup S. Then $(e,f) \in \mathcal{D}$ if and only if there exist an element a in S and an inverse a' of a such that $aa' = e$, $a'a = f$.*

Proof Suppose first that $(e,f) \in \mathcal{D}$. Then e and f are members of the same regular $\mathcal{D}$-class.

	L_f		L_e	
R_e	a		e	
R_f	f		a'	

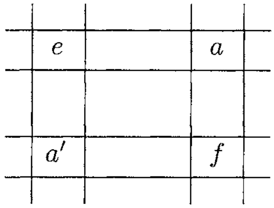
Let $a \in R_e \cap L_f$. Then by Theorem 2.3.4(2) there is an inverse a' of a in $R_f \cap L_e$ such that $aa' = e$, $a'a = f$.

Conversely, if there exist mutually inverse elements a, a' such that $aa' = e$ and $a'a = f$, then from Theorem 2.3.4(1) it follows that $e \mathcal{R} a$, $a \mathcal{L} f$. Hence $e \mathcal{D} f$ as required. $\square$

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

Proof The method of proof is essentially that used in the proof of Lemma 2.2.3, with a careful (yet obvious) choice of translation maps. Suppose that

$H = H_e$ and $K = H_f$, where e, f are idempotents and are, respectively, the identities of the groups H_e and H_f.



Let $a \in R_e \cap L_f$ (which is non-empty since $e \mathcal{D} f$ by assumption); then by Theorem 2.3.4 there is a unique inverse a' of a in $L_e \cap R_f$, and we have

$$aa' = e, \quad a'a = f, \quad ea = af = a, \quad a'e = fa' = a'.$$

By Green's Lemmas (2.2.1 and 2.2.2) it follows that $\rho_a|_{H_e}$ is a bijection from H_e onto H_a and $\lambda_{a'}|_{H_a}$ is a bijection from H_a onto H_f. Thus the map $\phi = \rho_a \lambda_{a'}$ given by

$$x\phi = a'xa \quad (x \in H_e)$$

is a bijection from H_e onto H_f, with inverse given by

$$y\phi^{-1} = aya' \quad (y \in H_f).$$

The bijection ϕ is even an isomorphism, for if $x_1, x_2 \in H_e$ then

$$(x_1\phi)(x_2\phi) = a'x_1aa'x_2a = a'x_1ex_2a = a'(x_1x_2)a = (x_1x_2)\phi,$$

since e is the identity element of the group H_e. □

We close this section with a result that will be required in Chapter 3.

Proposition 2.3.7 *Let a, b be elements in a $\mathcal{D}$-class D. Then $ab \in R_a \cap L_b$ if and only if $L_a \cap R_b$ contains an idempotent.*

Proof Suppose first that $ab \in R_a \cap L_b$. Then there exists c such that $abc = a$, and by Green's Lemma (2.2.1) the right translation $\rho_c : x \mapsto xc$ maps H_b onto $L_a \cap R_b$. Thus in particular $bc \in L_a \cap R_b$. The translation $\rho_b : y \to yb$ maps $L_a \cap R_b$ onto H_b, and is the inverse of ρ_c, and it then follows that bc is idempotent, since

$$(bc)^2 = b\rho_c\rho_b\rho_c = b\rho_c = bc.$$

Conversely, suppose that $L_a \cap R_b$ contains an idempotent e. Then $eb = b$ by Proposition 2.3.3, and hence the translation $x \mapsto xb$ maps H_a onto $R_a \cap L_b$. In particular, $ab \in R_a \cap L_b$. □

2.4 REGULAR SEMIGROUPS

In a regular semigroup S we have a particularly useful way of looking at the equivalences $\mathcal{L}$ and $\mathcal{R}$. First, notice that if S is regular then $a = axa \in aS$, and similarly $a \in Sa$, $a \in SaS$. Hence in describing the Green equivalences for a regular semigroup we can drop all reference to S^1, and assert simply that

$$\begin{array}{lll} a\,\mathcal{L}\,b & \text{if and only if} & Sa = Sb, \\ a\,\mathcal{R}\,b & \text{if and only if} & aS = bS, \\ a\,\mathcal{J}\,b & \text{if and only if} & SaS = SbS. \end{array}$$

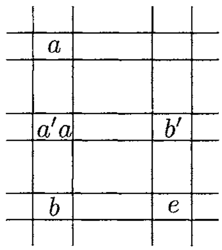
Next, we have

Proposition 2.4.1 *Let a, b be elements of a regular semigroup S. Then*

(1) *$(a,b) \in \mathcal{L}$ if and only if there exist a' in $V(a)$ and b' in $V(b)$ such that $a'a = b'b$;*

(2) *$(a,b) \in \mathcal{R}$ if and only if there exist a' in $V(a)$ and b' in $V(b)$ such that $aa' = bb'$;*

(3) *$(a,b) \in \mathcal{H}$ if and only if there exist a' in $V(a)$ and b' in $V(b)$ such that $a'a = b'b$ and $aa' = bb'$.*



Proof (1) Suppose first that $(a,b) \in \mathcal{L}$. If $a' \in V(a)$ then $a'a$ is an idempotent in $L_a = L_b$. The $\mathcal{R}$-class R_b contains at least one idempotent e by Proposition 2.3.2, and then, by Theorem 2.3.4(2), the $\mathcal{H}$-class $R_{a'a} \cap R_e$ contains an inverse b' of b with the property that $b'b = a'a$ (and $bb' = e$). Notice that we have shown the stronger implication that

$$(a,b) \in \mathcal{L} \ \Rightarrow\ \big(\forall a' \in V(a)\big)\big(\exists b' \in V(b)\big)\ a'a = b'b. \qquad (2.4.1)$$

Conversely, if $a'a = b'b$ for some a' in $V(a)$ and some b' in $V(b)$ then $a\,\mathcal{L}\,a'a$ and $b'b\,\mathcal{L}\,b$ by Theorem 2.3.4(1), and so $a\,\mathcal{L}\,b$ by transitivity.

Part (2) is similar, and once again we can prove

$$(a,b) \in \mathcal{L} \ \Rightarrow\ \big(\forall a' \in V(a)\big)\big(\exists b' \in V(b)\big)\ aa' = bb'. \qquad (2.4.2)$$

To prove part (3), suppose that $a\,\mathcal{H}\,b$ and that $a' \in V(a)$. Then $aa' \in R_a = R_b$ and $a'a \in L_a = L_b$. Hence, by Theorem 2.3.4(2), the $\mathcal{H}$-class $L_{aa'} \cap R_{a'a}$ contains an inverse b' of b such that $bb' = aa'$ and $b'b = a'a$. Once again we have proved the implication

$$(a,b) \in \mathcal{H} \ \Rightarrow\ \big(\forall a' \in V(a)\big)\big(\exists b' \in V(b)\big)\ a'a = b'b \text{ and } aa' = bb'. \quad (2.4.3)$$

The converse half is clear. □

If U is a subsemigroup of a (not necessarily regular) semigroup S, and if $a, b \in U$, there can be some ambiguity about the meaning of (for example) $a\ \mathcal{L}\ b$, since $\mathcal{L}$ may stand for the appropriate Green equivalence either in S or in U. When confusion of this sort is likely to arise we shall put superscripts on the symbols to distinguish between the two equivalences. Thus $(a, b) \in \mathcal{L}^U$ means that there exist u, v in U^1 such that $ua = b$, $vb = a$, while $(a, b) \in \mathcal{L}^S$ means that there exist s, t in S^1 such that $sa = b$, $tb = a$. We shall also use the notations

$$L_a^U = \{u \in U : (a, u) \in \mathcal{L}^U\}, \quad L_a^S = \{s \in S : (a, s) \in \mathcal{L}^S\}.$$

It is clear that

$$\mathcal{L}^U \subseteq \mathcal{L}^S \cap (U \times U),$$

and also, with the obvious notations, that

$$\begin{aligned} \mathcal{R}^U &\subseteq \mathcal{R}^S \cap (U \times U), \quad & \mathcal{H}^U &\subseteq \mathcal{H}^S \cap (U \times U), \\ \mathcal{D}^U &\subseteq \mathcal{D}^S \cap (U \times U), \quad & \mathcal{J}^U &\subseteq \mathcal{J}^S \cap (U \times U). \end{aligned}$$

These inclusions may well be proper: for example, if S is the infinite cyclic group generated by a and if U is the infinite monogenic semigroup of S consisting of all positive powers of a, then

$$\mathcal{L}^U = \mathcal{R}^U = \mathcal{H}^U = \mathcal{D}^U = \mathcal{J}^U = 1_U,$$

while

$$\begin{aligned} \mathcal{L}^S \cap (U \times U) &= \mathcal{R}^S \cap (U \times U) = \mathcal{H}^S \cap (U \times U) \\ &= \mathcal{D}^S \cap (U \times U) = \mathcal{J}^S \cap (U \times U) = U \times U. \end{aligned}$$

However, we have the following useful result:

Proposition 2.4.2 *If U is a regular subsemigroup of a semigroup S, then*

$$\mathcal{L}^U = \mathcal{L}^S \cap (U \times U), \quad \mathcal{R}^U = \mathcal{R}^S \cap (U \times U), \quad \mathcal{H}^U = \mathcal{H}^S \cap (U \times U).$$

Proof Suppose that $(a, b) \in \mathcal{L}^S \cap (U \times U)$, and let a' and b' be inverses in U of a and b, respectively. Then

$$(a'a, a) \in \mathcal{L}^U \subseteq \mathcal{L}^S, \quad (b'b, b) \in \mathcal{L}^U \subseteq \mathcal{L}^S,$$

and so $(a'a, b'b) \in \mathcal{L}^S$. By Proposition 2.3.3 each of $a'a$ and $b'b$ is a right identity for $L_{a'a}^S = L_{b'b}^S$; hence in particular

$$a'ab'b = a'a, \quad b'ba'a = b'b.$$

These equations involve only elements of U, and so may be interpreted as implying that $(a'a, b'b) \in \mathcal{L}^U$. But we now have

$$a\ \mathcal{L}^U\ a'a, \quad a'a\ \mathcal{L}^U\ b'b, \quad b'b\ \mathcal{L}^U\ b,$$

and so $a\ \mathcal{L}^U\ b$ as required.

The proof for $\mathcal{R}$ is exactly dual, and the result for $\mathcal{H}$ is a consequence of the results for $\mathcal{L}$ and $\mathcal{R}$. □

The corresponding assertions for $\mathcal{D}$ and $\mathcal{J}$ are not true. Consider, for example, the semigroup $S = B_2$ described at the end of Section 1.6. It is clear that B_2 is regular: the idempotents 0, e, f are certainly regular elements, and a and b are inverses of each other. The regular subsemigroup $U = \{0, e, f\}$ is a semilattice, and so

$$\mathcal{L}^U = \mathcal{R}^U = \mathcal{H}^U = \mathcal{D}^U = \mathcal{J}^U = 1_U.$$

On the other hand, it is easy to check that the $\mathcal{R}^S$-classes are $\{0\}$, $\{e, a\}$, $\{f, b\}$, that the $\mathcal{L}^S$-classes are $\{0\}$, $\{e, b\}$, $\{f, a\}$, and hence that the $\mathcal{D}^S$-classes are $\{0\}$ and $\{e, f, a, b\}$. Now $\mathcal{J}^S = \mathcal{D}^S$ since S is finite (Proposition 2.1.4) and so

$$(e, f) \in \mathcal{D}^S \cap (U \times U) = \mathcal{J}^S \cap (U \times U).$$

If S is a regular semigroup and ρ is a congruence on S, then S/ρ is regular. Indeed, if a' is an inverse in S of an element a then $a'\rho$ is an inverse of $a\rho$ in S/ρ, since

$$(a\rho)(a'\rho)(a\rho) = (aa'a)\rho = a\rho,\ (a'\rho)(a\rho)(a'\rho) = (a'aa')\rho = a'\rho.$$

The following result, usually known as Lallement's Lemma, is crucial in the study of regular semigroups:

Lemma 2.4.3 *Let ρ be a congruence on a regular semigroup S, and let $a\rho$ be an idempotent in S/ρ. Then there exists an idempotent e in S such that $e\rho = a\rho$. Moreover, e can be chosen so that $R_e \leq R_a$, $L_e \leq L_a$.*

Proof Let $a\rho$ be an idempotent in S/ρ. Then $(a, a^2) \in \rho$. Let x be an inverse of a^2:

$$a^2xa^2 = a^2, \quad xa^2x = x.$$

Let $e = axa$. Then

$$e^2 = a(xa^2x)a = axa = e,$$

and so e is idempotent. Also, modulo ρ,

$$e = axa \equiv a^2xa^2 = a^2 \equiv a,$$

and so $e\rho = a\rho$. It is clear from equation (2.1.4) that $R_e \leq R_a$, $L_e \leq L_a$. □

The close correspondence between congruences and morphisms discussed in Section 1.5 enables us to obtain the following alternative version of Lallement's Lemma:

Lemma 2.4.4 *Let $\phi : S \to T$ be a morphism from a regular semigroup S into a semigroup T. Then* $\operatorname{im}\phi$ *is regular. If f is an idempotent in* $\operatorname{im}\phi$ *then there exists an idempotent e in S such that $e\phi = f$.* □

We remark that this result may be untrue if we drop the hypothesis of regularity. There is an obvious morphism from the free monogenic semigroup a^+ onto the finite monogenic semigroup $M(m, r)$, but the idempotent in $M(m,r)$ cannot have an idempotent counterimage in a^+, for a^+ has no idempotents.

If S is a semigroup and E is the set of idempotents of S, then we shall say that an equivalence ρ on S is *idempotent-separating* if

$$\rho \cap (E \times E) = 1_E,$$

that is, if no ρ-class contains more than one idempotent. By Corollary 2.2.6 we have that $\mathcal{H}$ is an idempotent-separating equivalence on any semigroup S, and so every congruence contained in $\mathcal{H}$ is idempotent-separating. In fact we have:

Proposition 2.4.5 *If S is a regular semigroup, then a congruence ρ on S is idempotent-separating if and only if $\rho \subseteq \mathcal{H}$. Hence $\mathcal{H}^\flat$ is the maximum idempotent-separating congruence on S.*

Proof We have already seen that one half of this result holds in any semigroup whatever. So suppose that S is regular, and that $a \; \rho \; b$, where ρ is an idempotent-separating congruence. We need to show that $a \; \mathcal{H} \; b$. Certainly $(ba', aa') \in \rho$, and so, by Lallement's Lemma (2.4.3) there exists an idempotent e such that $(e, ba') \in \rho$ and $R_e \leq R_{ba'}$. But then $(e, aa') \in \rho$, and so $e = aa'$, since ρ is idempotent-separating. Hence

$$R_a = R_{aa'} \leq R_{ba'} \leq R_b.$$

A similar argument establishes that $R_b \leq R_a$, and we deduce that $a \; \mathcal{R} \; b$. But then an exactly dual argument shows that $a \; \mathcal{L} \; b$, and so $a \; \mathcal{H} \; b$ as required. □

Remark The hypothesis that S is regular cannot be removed in the above proposition. If $S = \{x, 0\}$ is the two element null semigroup, in which

$$x^2 = x0 = 0x = 00 = 0,$$

then 0 is the only idempotent and $\mathcal{H} = 1_S$. The universal congruence $S \times S$ is then idempotent-separating, but is not contained in $\mathcal{H}$.

2.5 THE SANDWICH SET

The concept of inverse that has been a main theme of this chapter is of course a generalization of the inverse a^{-1} of an element a in a group G. Within a group we have the very useful property that $(ab)^{-1} = b^{-1}a^{-1}$, and it is reasonable to ask whether some version of this holds in a regular semigroup. In general it is not the case that $V(ab) = V(b)V(a)$, and indeed we shall see in Chapter 8 that $V(b)V(a) \subseteq V(ab)$ for all a and b in S if and only if S is *orthodox*, that is, if and only if the set E of idempotents of S is closed under multiplication. For a generalization that works for every regular semigroup we need a new concept.

Proposition 2.5.1 *Let S be a regular semigroup with set E of idempotents, and let $e, f \in E$. Then the set $S(e, f)$, defined by*

$$S(e,f) = \{g \in V(ef) \cap E : ge = fg = g\} \tag{2.5.1}$$

is non-empty.

Proof Let $x \in V(ef)$, and let $g = fxe$. Then

$$(ef)g(ef) = ef^2xe^2f = efxef = ef,$$
$$g(ef)g = fxe^2f^2xe = f(xefx)e = fxe = g,$$

and so $g \in V(ef)$. Also,

$$g^2 = f(xefx)e = fxe = g,$$

and so $g \in E$. Finally, it is clear that $ge = fg = g$, and so $g \in S(e, f)$.

□

The set $S(e, f)$ is called the *sandwich set* of e and f. It has an obvious alternative characterization

$$S(e,f) = \{g \in E : ge = fg = g,\ egf = ef\}. \tag{2.5.2}$$

The next result gives a connection with Green's equivalences that will prove useful in Chapter 6:

Proposition 2.5.2 *Let e, f and g be idempotents in a regular semigroup.*
(1) *If $e \mathcal{L} f$ then $S(e, g) = S(f, g)$;*
(2) *if $e \mathcal{R} f$ then $S(g, e) = S(g, f)$.*

Proof It is obviously sufficient to prove (1), and indeed it will be enough to show that $S(e, g) \subseteq S(f, g)$. Suppose therefore that $e \mathcal{L} f$ and that $h \in S(e, g)$, so that

$$ef = e,\ fe = f,\ he = gh = h,\ ehg = eg.$$

Then

$$h = he = hef = hf, \quad fhg = fehg = feg = fg,$$

exactly as required. □

Proposition 2.5.3 *Let e, f be idempotents in a regular semigroup S. Then $S(e, f)$ is a subsemigroup of S and is a rectangular band.*

Proof Let $g, h \in S(e, f)$. Then

$$ghg = (ge)h(fg) = g(ehf)g = g(ef)g = (ge)(fg) = g^2 = g. \tag{2.5.3}$$

It follows that $(gh)^2 = gh$ and so gh is idempotent. Also

$$(gh)e = g(he) = gh,\ f(gh) = (fg)h = gh,$$
$$e(gh)f = egfhf = efhf = ehf = ef,$$

and so $gh \in S(e, f)$. From (2.5.3) we deduce that $S(e, f)$ is a rectangular band. □

Finally, we have

Theorem 2.5.4 *Let $a, b \in S$, where S is a regular semigroup. Let $a' \in V(a)$, $b' \in V(b)$ and $g \in S(a'a, bb')$. Then $b'ga' \in V(ab)$.*

Proof Write $a'a = e$, $bb' = f$, and let $g \in S(e, f)$. Then

$$\begin{aligned}(ab)(b'ga')(ab) &= afgeb = agb = aa'agbb'b = a(egf)b \\ &= a(ef)b = aa'abb'b = ab,\end{aligned}$$

$$(b'ga')(ab)(b'ga') = b'gefga' = b'g^2a' = b'ga',$$

and so $b'ga' \in V(ab)$, as required. □

2.6 EXERCISES

1. Let C be a *cancellative* semigroup (that is, a semigroup in which, for all a, b, c, $ca = cb \Rightarrow a = b$ and $ac = bc \Rightarrow a = b$), and suppose that C has no identity. Show that there cannot be any pair of elements e, a in C for which $ea = a$ or for which $ae = e$, and deduce that $\mathcal{L} = \mathcal{R} = \mathcal{D} = 1_C$.

   Show that, with respect to matrix multiplication, the set

   $$S = \left\{ \begin{pmatrix} a & 0 \\ b & 1 \end{pmatrix} : a, b \in \mathbf{R},\ a, b > 0 \right\}$$

   is a cancellative semigroup without identity.

   Show that $\mathcal{J} = S \times S$, and deduce that $\mathcal{D}$ is properly contained in $\mathcal{J}$.

   Observe that S in not periodic, since (for example)

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

   Observe also that S does not satisfy $\min_L$. More precisely, denote $\left(\begin{smallmatrix} 1 & 0 \\ n & 1 \end{smallmatrix}\right)$ by s_n, and observe that

   $$S^1 s_1 \supset S^1 s_2 \supset S^1 s_3 \supset \cdots$$

2. In the *bicyclic semigroup* $B = \mathbf{N}^0 \times \mathbf{N}^0$, with multiplication given by equation (1.6.3), show that
   (a) $(m, n)\ \mathcal{R}\ (p, q)$ if and only if $m = p$;
   (b) $(m, n)\ \mathcal{L}\ (p, q)$ if and only if $n = q$;
   (c) $\mathcal{D} = \mathcal{J} = B \times B$.

3. Let e, f be idempotents in a semigroup S. Show that

   $$\begin{aligned} e\,\mathcal{L}\,f &\quad \text{if and only if} \quad ef = e,\ fe = f, \\ e\,\mathcal{R}\,f &\quad \text{if and only if} \quad ef = f,\ fe = e. \end{aligned}$$

4. Let S be a semigroup, let λ be a right congruence on S contained in $\mathcal{L}$ and let ρ be a left congruence on S contained in $\mathcal{R}$. Show that $\lambda \circ \rho = \rho \circ \lambda$. Deduce, using Proposition 1.8.3, that the sublattice $[1_S, \mathcal{H}^b]$, consisting of all congruences on S contained in $\mathcal{H}$, is modular.

5. A semigroup S is called *right simple* if $\mathcal{R} = S \times S$, and *left simple* if $\mathcal{L} = S \times S$. It is called *right cancellative* if $(\forall a, b, c \in S)\ ac = bc \Rightarrow a = b$ and *left cancellative* if $(\forall a, b, c \in S)\ ca = cb \Rightarrow a = b$.
   (a) Show that a left zero semigroup is left simple and right cancellative, but is neither right simple nor left cancellative.
   (b) Show that the Baer–Levi semigroup (Exercise 1(3)) is right simple and right cancellative, but is neither left simple nor left cancellative.
   (c) Show that a semigroup is right simple and left simple if and only if it is a group.
   (d) Show that a finite semigroup is right cancellative and left cancellative if and only if it is a group.
   (e) Show that the word 'finite' cannot be removed from (d). In other words, give an example of an infinite semigroup which is both right and left cancellative, but is not a group.

6. A semigroup which is right simple and left cancellative is called a *right group*. Show that if G is a group and E is a right zero semigroup, then the direct product $G \times E$ is a right group.

   Conversely, suppose that S is a right group.
   (a) Show that the set E of idempotents in S is non-empty.
   (b) Show that E is a right zero subsemigroup of S.
   (c) Show that $eb = b$ for every b in S.
   (d) Show that Se is a subgroup of S for every e in E.
   (e) Let f be a fixed element of E, and denote the group Sf by G. Show that the map $\phi : G \times E \to S$ defined by
   $$(a, e)\phi = ae$$
   is an isomorphism.
   (f) Deduce that a semigroup is a right group if and only if it is isomorphic to the direct product of a group and a right zero semigroup.

7. Define an equivalence $\mathcal{L}^*$ on a semigroup S by the rule that $a\ \mathcal{L}^*\ b$ if and only if
   $$(\forall x, y \in S^1)\ ax = ay \iff bx = by.$$
   (a) Show that $\mathcal{L} \subseteq \mathcal{L}^*$, and that $\mathcal{L} = \mathcal{L}^*$ if S is regular.
   (b) Show that $\mathcal{L}^*$ is a right congruence on S.

(c) Show that, for every idempotent e in S, $a\ \mathcal{L}^*\ e$ if and only if $ae = a$ and $(\forall x, y \in S^1)\ ax = ay \Rightarrow ex = ey$. In particular, note that e acts as a right identity within its $\mathcal{L}^*$-class.

8. The containment $\mathcal{L} \subseteq \mathcal{L}^*$ noted in the previous exercise may well be proper. For example, in the cancellative semigroup S defined in Exercise 1 above, show that $\mathcal{L} = 1_S$, $\mathcal{L}^* = S \times S$.

9. The equivalences $\mathcal{R}^*$ and $\mathcal{H}^*$ are defined by analogy with $\mathcal{L}^*$. (See Exercise 7 above.) Show that every $\mathcal{H}^*$-class containing an idempotent is a subsemigroup of S and is a cancellative semigroup with identity element e.

10. Let S be an arbitrary semigroup, and let ρ be a congruence on S such that $\rho \subseteq \mathcal{L}$. Show that $(a, b) \in \mathcal{L}$ in S if and only if $(a\rho, b\rho) \in \mathcal{L}$ in S/ρ.

11. Let S be a regular semigroup. Show that the following statements are equivalent:
    (a) S has exactly one idempotent;
    (b) S is cancellative;
    (c) S is a group.

12. Define a *cancellative congruence* on a semigroup S to be a congruence ρ with the property that S/ρ is a cancellative semigroup. Show that ρ is a cancellative congruence if and only if, for all a, b, c in S

$$(ca, cb) \in \rho \Rightarrow (a, b) \in \rho, \quad (ac, bc) \in \rho \Rightarrow (a, b) \in \rho,$$

and deduce that the intersection of a family $\{\rho_i : i \in I\}$ of cancellative congruences is cancellative. Deduce that there is a minimum cancellative congruence on S, and hence deduce that every *regular* semigroup possesses a minimum group congruence.

13. Show now that not every semigroup has a minimum group congruence. Specifically, consider $S = a^+$, the free monogenic semigroup. Show that the relation

$$\rho_n = \{(a^p, a^q) : p \equiv q \pmod{n}\}$$

is a group congruence on S, and show conversely that if ρ is a group congruence then there exists n $\big(= \min\{|r - s| : (a^r, a^s) \in \rho \setminus \{1_S\}\}\big)$ such that $\rho = \rho_n$. Deduce that there is no minimum group congruence on S.

14. Show that, if R is a right ideal and L is a left ideal of a semigroup S, then $R \cap L \supseteq RL$. Show that equality holds if S is regular.

15. Show that the full transformation semigroup $\mathcal{T}_X$ and the partial transformation semigroup $\mathcal{P}_X$ are both regular. Show, however, that the semigroup $\mathcal{B}_X$ of all binary relations on X is not regular. More precisely, show that the relation $\rho = \{(1,1),(1,2),(2,1)\}$ in $\mathcal{B}_{\{1,2\}}$ is not a regular element of $\mathcal{B}_{\{1,2\}}$.

16. Show that, in $\mathcal{T}_X$,
    (a) $(\alpha, \beta) \in \mathcal{L}$ if and only if $\operatorname{im}\alpha = \operatorname{im}\beta$;
    (b) $(\alpha, \beta) \in \mathcal{R}$ if and only if $\ker\alpha = \ker\beta$;
    (c) $(\alpha, \beta) \in \mathcal{D}$ if and only if $|\operatorname{im}\alpha| = |\operatorname{im}\beta|$;
    (d) $\mathcal{D} = \mathcal{J}$.

17. For a partial map ϕ of X, define $\ker\phi$ to be the relation $\{(x,y) \in X \times X : x, y \in \operatorname{dom}\phi,\ x\phi = y\phi\}$. This relation is symmetric and transitive, but is not reflexive unless $\phi \in \mathcal{T}_X$. Show that, with this interpretation of kernel, parts (a) to (d) of the previous exercise hold also in $\mathcal{P}_X$.

18. Let ρ be an equivalence on a set X. A subset A of X is said to be a *cross-section* (or a *transversal*) of ρ if each ρ-class contains exactly one element of A.

    In the full transformation semigroup $\mathcal{T}_X$, denote the $\mathcal{H}$-class consisting of elements α such that $\operatorname{im}\alpha = A$ and $\ker\alpha = \rho$ by $H_{A,\rho}$. Show that $H_{A,\rho}$ is a group if and only if A is a cross-section of ρ.

19. Let V be a vector space over a field and let $\mathcal{L}(V)$ be the semigroup of all linear maps $\alpha : V \to V$. Let $\ker\alpha = \{v \in V : v\alpha = 0\}$ (the *kernel*, or *nullspace*, of α). Show that $\mathcal{L}(V)$ is regular.

    Let $\alpha, \beta \in \mathcal{L}(V)$. Show that
    (a) $(\alpha, \beta) \in \mathcal{L}$ if and only if $\operatorname{im}\alpha = \operatorname{im}\beta$;
    (b) $(\alpha, \beta) \in \mathcal{R}$ if and only if $\ker\alpha = \ker\beta$;
    (c) $(\alpha, \beta) \in \mathcal{D}$ if and only if $\operatorname{rank}\alpha = \operatorname{rank}\beta$;
    (d) $\mathcal{D} = \mathcal{J}$.
    (e) Denote the $\mathcal{H}$-class consisting of elements α for which $\operatorname{im}\alpha = A$ and $\ker\alpha = B$ by $H_{A,B}$. Show that $H_{A,B}$ is a group if and only if $A \cap B = \{0\}$.

20. A subset A of a semigroup S is called *right unitary* if

    $$(\forall a \in A)(\forall s \in S)\ sa \in A \ \Rightarrow\ s \in A,$$

    left unitary if

    $$(\forall a \in A)(\forall s \in S)\ as \in A \ \Rightarrow\ s \in A,$$

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

Let E be the set of idempotents of a regular semigroup S, and suppose that E is a right unitary subset. Show that E is a unitary subsemigroup. [Hint: show first that if $es \in E$ with e in E, then $sess' \in E$ for every s' in $V(s)$.]

21. Define a regular semigroup to be *E-unitary* if the set E of idempotents is a unitary subsemigroup of S. Show that if S is E-unitary then, for all a, b in S,

$$ab \in E \Rightarrow ba \in E.$$

[Hint: show first that $babb' \in E$.]

22. Let S be a regular semigroup, let ϕ be a morphism from S onto T, and let c, d be mutually inverse elements of T. Show that, if x, y in S are such that $x\phi = c$ and $y\phi = d$, and if $v \in V(xyxy)$, then

$$a = xyvxyx \quad \text{and} \quad b = yvxy$$

are such that $a\phi = c$, $b\phi = d$, and a and b are mutually inverse in S.

23. Let S be a regular semigroup with set E of idempotents. Show that, for all $n \geq 1$,

$$V(E^n) = E^{n+1}.$$

[Hint: if $x = e_1e_2 \dots e_n \in E^n$ and $y \in V(x)$, let

$$f_j = e_j \dots e_n y e_1 \dots e_{j-1};$$

show that f_j is idempotent and that $y = yxf_n \dots f_2xy \in E^{n+1}$. Conversely, if $x = e_1 \dots e_{n+1} \in E^{n+1}$, let

$$g_j = e_{j+1} \dots e_{n+1} y e_1 \dots e_j,$$

where $y \in V(x)$, and show that $x \in V(g_n \dots g_1)$.]

2.7 NOTES

The results in Sections 2.1 and 2.2 are from Green (1951). Section 2.3 is from Miller and Clifford (1956). Proposition 2.4.2 is due to T. E. Hall (1972), while Lemmas 2.4.3 and 2.4.4, together with Proposition 2.4.5 are from Lallement (1966). The concept of the sandwich set is due to Nambooripad (1974, 1975).

Minimal conditions, mentioned briefly in Section 1, are the subject of Section 6.6 in Clifford and Preston (1967). A very thorough study, developing the ideas in that section and in Munn (1957), appears in Hall and Munn (1979).

Exercise 1 is from Andersen (1952). Exercise 6 appears in the book by Clifford and Preston (1961, Section 1.11), who trace it back to Suschkewitsch (1928). The starred Green equivalences featuring in Exercises 7, 8 and 9 have been extensively studied by Fountain and others: see, for example, Fountain (1979, 1982). Exercises 10 and 22 are from Hall (1972).

Exercises 20 and 21 are from Howie and Lallement (1966). Exercise 23 reports work done independently by Fitz-Gerald (1972) and Eberhart *et al.* (1973).

3

0-simple semigroups

The chapter begins with some elementary results on simple and 0-simple semigroups. A decomposition theorem for semigroups in general then indicates why an understanding of simple and 0-simple semigroups is important. The main result of the chapter is due to Rees (1940), and applies to 0-simple semigroups containing primitive idempotents. The Rees Theorem, strongly motivated by the Wedderburn–Artin Theorem for rings (see, for example, Cohn (1989)) has played a dominant role in the development of the subject.

Perhaps unfortunately, the word 'simple' as used in semigroup theory does not have the same import as in group theory or ring theory, where it implies the total absence of proper homomorphic images. By contrast with ring theory, not every congruence on a semigroup is associated with an ideal, and so it is normally the case that a simple (or 0-simple) semigroup has non-trivial congruences. Using the Rees structure theorem we then describe a classification of the congruences on a completely 0-simple semigroup. A result on the nature of the lattice of congruences readily follows, and so also does a classification of the finite *congruence-free* (i.e. truly simple) semigroups.

3.1 SIMPLE AND 0-SIMPLE SEMIGROUPS; PRINCIPAL FACTORS

A semigroup without zero is called *simple* if it has no proper ideals. A semigroup S with zero is called *0-simple* if

(i) $\{0\}$ and S are its only ideals;
(ii) $S^2 \neq \{0\}$.

The latter condition serves only to exclude the two-element null semigroup, since any larger null semigroup fails to qualify on the grounds of having proper ideals.

It is easy to see that S is simple if and only if $\mathcal{J} = S \times S$. The corresponding criterion for 0-simplicity is that $S^2 \neq \{0\}$ and that $\{0\}$ and $S \setminus \{0\}$ are the only $\mathcal{J}$-classes.

A simple semigroup can be made into a 0-simple semigroup by merely adjoining a zero element. Not all 0-simple semigroups arise in this way, however; the zero element of a 0-simple semigroup cannot always be removed to leave behind a simple semigroup. The removal is in fact possible only if 0 is a 'prime' element in S, in the sense that

$$ab = 0 \Rightarrow a = 0 \text{ or } b = 0, \tag{3.1.1}$$

and we shall see that this is not always the case. The implication (3.1.1), familiar in ring theory as the criterion for a ring to be an integral domain, is often expressed by saying that S *has no proper zero-divisors.* Since it will always be possible, by specializing to the case where the semigroup has no proper zero-divisors, to deduce a theorem about simple semigroups from one about 0-simple semigroups, we shall focus attention primarily on the 0-simple case.

The first significant result gives an important alternative characterization of 0-simple semigroups:

Proposition 3.1.1 *A semigroup S is 0-simple if and only if $SaS = S$ for every $a \neq 0$ in S, that is, if and only if for every a, b in $S \setminus \{0\}$ there exist x, y in S such that $xay = b$.*

Proof Suppose first that S is 0-simple. Then S^2, being an ideal of S, and being by definition distinct from $\{0\}$, must coincide with S, and it follows that $S^3 = S^2.S = S.S = S$ also. Let a be a non-zero element of S. Then SaS is an ideal of S and so either $SaS = S$ or $SaS = \{0\}$. If $SaS = \{0\}$ then the set $I = \{x \in S : SxS = \{0\}\}$ contains the non-zero element a. Since I is easily seen to be an ideal of S it follows that $I = S$, and so $SxS = \{0\}$ for *every* x in S. But this implies that $S^3 = \{0\}$, in contradiction to the already noted fact that $S^3 = S$. Hence $SaS = S$ as required.

Conversely, suppose that $SaS = S$ for all $a \neq 0$ in S. Then certainly $S^2 \neq \{0\}$. If A is an ideal of S containing a non-zero element a then

$$S = SaS \subseteq SAS \subseteq A,$$

and so $A = S$. Thus S is 0-simple. □

As a corollary we have the corresponding statement for simple semigroups.

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

By a 0-*minimal* ideal in a semigroup S we mean an ideal minimal within the set of all non-zero ideals. The next result shows that 0-simple semigroups can occur inside more general semigroups.

Proposition 3.1.3 *If M is a 0-minimal ideal of S then either $M^2 = \{0\}$ or M is a 0-simple semigroup.*

Proof Since M^2 is an ideal of S contained in M, we must have either $M^2 = M$ or $M^2 = \{0\}$. Suppose that $M^2 = M$. Then $M^3 = M$. If a is a non-zero element of M then S^1aS^1, being a non-zero ideal of S contained in M, must coincide with M. Hence

$$MaM \subseteq S^1aS^1 = M = M^3 = M(S^1aS^1)M = (MS^1)a(S^1M) \subseteq MaM,$$

and so in fact $MaM = M$. Thus M is 0-simple by Proposition 3.1.1. □

The specialization of this result to the case of semigroups without zero is of special interest, since such a semigroup can have at most one *minimal* ideal. To see this, suppose that M and N are both minimal ideals of S. Then MN, being an ideal contained in both M and N, must be *equal* to both M and N, which must therefore be equal to each other. Thus either there are no minimal ideals at all (which can happen—see Exercise 3(1)—) or there is a unique *minimum* ideal, which we call the *kernel* $K = K(S)$ of S. A simplified version of the proof of Proposition 3.1.3 now gives

Proposition 3.1.4 *Let S be a semigroup without zero. If S has a kernel K, then K is a simple semigroup.* □

It is worth remarking that a semigroup with zero does have a kernel, namely the unique minimum ideal $\{0\}$, but in this case the notion is not particularly useful. More significantly, every finite semigroup has a kernel, since the alternative to having a kernel is to have infinite descending chains of ideals, and this alternative is not open to a finite semigroup.

The next result describes another context in which 0-simple semigroups can occur.

Proposition 3.1.5 *If I, J are ideals of a semigroup S such that $I \subset J$ and there is no ideal B of S such that $I \subset B \subset J$, then J/I is either 0-simple or null.*

Proof By virtue of Proposition 1.7.1, J/I is a 0-minimal ideal of S/I, and so the result is a direct consequence of Proposition 3.1.3. □

This result, while not of any great depth, is the basis of an important decomposition method for an arbitrary semigroup S. For each a in S, write the principal ideal S^1aS^1 generated by the element a as $J(a)$, and recall the natural partial order (2.1.3) among the $\mathcal{J}$-classes of S, whereby $J_x < J_y$ if and only if $J(x) \subset J(y)$. Suppose first that J_a is minimal among the $\mathcal{J}$-classes. Then $J(a)$ is a minimal ideal of S and so it is the unique minimum ideal $K(S)$. If $b \in J(a)$, then S^1bS^1 is an ideal of S contained in $J(a)$ and so $S^1bS^1 = J(a) = S^1aS^1$. Thus $b \mathcal{J} a$. We deduce that

$$J(a) = J_a = K(S) \tag{3.1.2}$$

(where of course $K(S) = \{0\}$ in the case where S has a zero element).

If J_a is not minimal in $S/\mathcal{J}$ then the set

$$I(a) = \{b \in J(a) : J_b < J_a\}$$

is non-empty. In fact $I(a)$ is an ideal of S, for if $b \in I(a)$ and $u, v \in S$ then

$$J_{ub} \leq J_b < J_a, \quad J_{bv} \leq J_b < J_a,$$

and so $ub, bv \in I(a)$. Notice that

$$J(a) = J_a \cup I(a), \quad I(a) = \bigcup \{J_b : b <_{\mathcal{J}} a\}, \tag{3.1.3}$$

and that both unions are disjoint.

Suppose now that B is an ideal of S such that

$$I(a) \subseteq B \subset J(a).$$

If $b \in B$ then $J(b) \subseteq B$, since $J(b)$ is the smallest ideal containing b. Certainly $J(b) \subset J(a)$ and so $J_b < J_a$. Thus $b \in I(a)$. We have shown that $B = I(a)$, and so $J(a)$ and $I(a)$ satisfy the criteria for Proposition 3.1.5. We deduce that the Rees quotient $J(a)/I(a)$ is either 0-simple or null. The semigroups $K(S)$, $J(a)/I(a)$ $(a \in S)$ are called the *principal factors* of S.

Obviously, then, if we can find out something about 0-simple semigroups, we shall have gone some way towards an understanding of the 'local' structure of a semigroup. By virtue of (3.1.3) we can think of the principal factor $J(a)/I(a)$ as consisting of the $\mathcal{J}$-class $J_a = J(a) \setminus I(a)$ with a zero adjoined. If the factor is null then the product of two elements in J_a always falls into a lower $\mathcal{J}$-class. If the factor is 0-simple then the product may lie in J_a or may fall into a lower $\mathcal{J}$-class.

We summarize these observations in a theorem as follows

Theorem 3.1.6 *If a is an element of a semigroup S then either*

(1) *J_a is the kernel of S; or*
(2) *the set $I(a) = \{b \in S : J_b < J_a\}$ is non-empty and is an ideal of $J(a)$ $(= S^1aS^1)$ such that $J(a)/I(a)$ is either 0-simple or null.* □

3.2 THE REES THEOREM

Among idempotents in an arbitrary semigroup there is a natural (partial) order relation defined by the rule that $e \leq f$ if and only if $ef = fe = e$. It is easy to verify that the given relation has the properties (O1), (O2) and (O3) (see Section 1.3) that define an order relation. Certainly it is clear that $e \leq e$, and that $e \leq f$ and $f \leq e$ together imply that $e = f$. To show transitivity, notice that if $e \leq f$ and $f \leq g$, so that $ef = fe = e$ and $fg = gf = f$, then

$$eg = efg = ef = e \quad \text{and} \quad ge = gfe = fe = e,$$

and so $e \leq g$.

If S is a semigroup with zero, then the defining properties of a zero element immediately imply that 0 is the unique minimum idempotent. The idempotents that are minimal within the set of non-zero idempotents are called *primitive*. Thus a primitive idempotent e has the property that

$$ef = fe = f \neq 0 \Rightarrow e = f. \tag{3.2.1}$$

A semigroup will be called *completely 0-simple* if it is 0-simple and has a primitive idempotent. We immediately record that

Proposition 3.2.1 *Every finite 0-simple semigroup is completely 0-simple.*

Proof Let S be a finite 0-simple semigroup. By Proposition 1.2.3 every element of S has a power which is idempotent. Suppose first that 0 is the only idempotent. Then every element s in S is *nilpotent*, in the sense that there exists a positive integer n with the property that

$$s^n = s^{n+1} = s^{n+2} = \cdots = 0.$$

Let a be a non-zero element of S. Then by Proposition 3.1.2 there exist x, y in S such that

$$a = xay = x^2ay^2 = x^3ay^3 = \cdots$$

Since x and y are nilpotent, we deduce that $a = 0$, and so we have a contradiction. Thus the set of *non-zero* idempotents of S is non-empty. It now follows that primitive idempotents exist, for otherwise there would exist infinite descending chains $e_1 > e_2 > e_3 > \cdots$ of non-zero idempotents, and this is not possible in a finite semigroup. □

We shall see very shortly that infinite completely 0-simple semigroups also exist. Indeed we have a fairly straightforward recipe, due to Rees (1940), for constructing completely 0-simple semigroups, which we now describe. Let G be a group with identity element e, and let I, Λ be non-empty sets. Let $P = (p_{\lambda i})$ be a $\Lambda \times I$ matrix with entries in the 0-group G^0 ($= G \cup \{0\}$), and suppose that P is *regular*, in the sense that no row or column of P consists entirely of zeros. Formally,

$$\begin{aligned} &(\forall i \in I)(\exists \lambda \in \Lambda)\ p_{\lambda i} \neq 0, \\ &(\forall \lambda \in \Lambda)(\exists i \in I)\ p_{\lambda i} \neq 0. \end{aligned} \tag{3.2.2}$$

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

$$\begin{aligned} (i, a, \lambda)(j, b, \mu) &= \begin{cases} (i, ap_{\lambda j}b, \mu) & \text{if } p_{\lambda j} \neq 0, \\ 0 & \text{if } p_{\lambda j} = 0, \end{cases} \\ (i, a, \lambda)0 &= 0(i, a, \lambda) = 00 = 0. \end{aligned} \tag{3.2.3}$$

Then we have

Lemma 3.2.2 *S is a completely 0-simple semigroup.*

Proof It is fairly easy to give a direct verification of the associativity of the composition (3.2.3), but it is probably more illuminating (and recalls Rees's own approach) to observe that $S \setminus \{0\}$ is in one-one correspondence with the set of $I \times \Lambda$ matrices $(a)_{i\lambda}$ $(a \in G)$, where $(a)_{i\lambda}$ denotes the matrix with entry a in the (i, λ) position and zeros elsewhere. Since $(0)_{i\lambda}$ is independent of i and λ we may simply write it as 0. Thus the correspondence extends to a correspondence between S and the set

$$T = \{(a)_{i\lambda} : a \in G^0,\ i \in I,\ \lambda \in \Lambda\}.$$

It is now a routine matter to verify that

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

where the juxtaposition on the left denotes matrix multiplication in the usual sense. (The low density of non-zero entries in the matrices means that no additions become necessary and so no trouble arises from the fact that the matrix entries lie in a 0-group rather than the more usual ring.) Thus the composition (3.2.3) in S corresponds in T to the evidently associative composition $\circ$ given by

$$(a)_{i\lambda} \circ (b)_{j\mu} = (a)_{i\lambda} P(b)_{j\mu},$$

and so is itself associative.

To verify that S is 0-simple, note that for any two non-zero elements (i, a, λ) and (j, b, μ) of S we may, by the regularity of P, choose ν in Λ and k in I such that $p_{\nu i} \neq 0$, $p_{\lambda k} \neq 0$, and then easily show that

$$(j, a^{-1}p_{\nu i}^{-1}, \nu)(i, a, \lambda)(k, p_{\lambda k}^{-1}b, \mu) = (j, b, \mu).$$

Hence by Proposition 3.1.1, S is 0-simple.

To complete the proof that S is completely 0-simple we must first identify its idempotents. The non-zero element (i, a, λ) is idempotent if and only if

$$(i, a, \lambda) = (i, a, \lambda)(i, a, \lambda) = (i, ap_{\lambda i}a, \lambda),$$

that is, if and only if $p_{\lambda i} \neq 0$ and $a = p_{\lambda i}^{-1}$. If we now take two non-zero idempotents $e = (i, p_{\lambda i}^{-1}, \lambda)$ and $f = (j, p_{\mu j}^{-1}, \mu)$, then $e \leq f$ if and only if $ef = fe = e$, that is, if and only if

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

that is, if and only if $j = i$ and $\lambda = \mu$, that is if and only if $e = f$. The conclusion is that *every* idempotent is primitive. Certainly there exists a primitive idempotent, and so S is completely 0-simple. □

The semigroup constructed in accordance with this recipe will, as in Clifford and Preston (1961), be denoted by $\mathcal{M}^0[G; I, \Lambda; P]$, and will be called the $I \times \Lambda$ *Rees matrix semigroup over the* 0-*group* G^0 *with the regular sandwich matrix* P.

The importance of Rees's recipe lies in its universality: *every* completely 0-simple semigroup is isomorphic to some $\mathcal{M}^0[G; I, \Lambda; P]$. To spell it out formally, we have the following theorem, of which the easy half has been proved:

Theorem 3.2.3 (The Rees Theorem) *Let G^0 be a 0-group, let I, Λ be non-empty sets and let $P = (p_{\lambda i})$ be a $\Lambda \times I$ matrix with entries in G^0. Suppose that P is regular in the sense of* (3.2.2). *Let $S = (I \times G \times \Lambda) \cup \{0\}$, and define a multiplication on S by* (3.2.3). *Then S is a completely 0-simple semigroup.*

Conversely, every completely 0-simple semigroup is isomorphic to one constructed in this way.

Proof Let S be a completely 0-simple semigroup containing a primitive idempotent e. As a first step towards proving that S is isomorphic to a Rees matrix semigroup we establish

Lemma 3.2.4 $R_e = eS \setminus \{0\}$.

Proof Certainly any element in R_e is a non-zero right multiple of e and so lies in $eS \setminus \{0\}$. Thus $R_e \subseteq eS \setminus \{0\}$.

Conversely, let $a = es$ be a non-zero element in eS. Then

$$ea = e^2 s = es = a. \tag{3.2.4}$$

Since S is 0-simple, there exist (by Proposition 3.1.1) elements z, t in S such that $e = zat$. We can in fact replace the elements z and t by 'improved' versions, writing $e = xay$, where $x = eze$, $y = te$. The verification that this is so is entirely routine, and we also have

$$ex = xe = x, \quad ye = y. \tag{3.2.5}$$

Now let $f = ayx$. Then

$$f^2 = ayxayx = ayex = ayx = f,$$

and so f is idempotent. Moreover $f \neq 0$, since $f = 0$ would imply that

$$e = e^2 = xayxay = xfay = 0,$$

and we know that this is not so. Now notice that from (3.2.4) and (3.2.5) it follows that $ef = fe = f$. That is, $f \leq e$, and since e is primitive we deduce that $f = e$. Thus $e = ayx$, and this together with $a \in eS$ gives $a \in R_e$. Thus $eS \setminus \{0\} \subseteq R_e$ as required. □

In fact the property described in Lemma 3.2.4 holds for every non-zero element of S:

Lemma 3.2.5 *For every $a \neq 0$ in S, $R_a = aS \setminus \{0\}$.*

Proof Again, any element of R_a is a non-zero right multiple of a, and so $R_a \subseteq aS \setminus \{0\}$. Conversely, suppose that b is a non-zero element of aS. Now by 0-simplicity there exist z and t in S such that $a = zet$, where e is a primitive idempotent. Hence $b = zeu$ for some u in S. Now $eu \, \mathcal{R} \, et$ by the previous lemma, and so, by the left congruence property of $\mathcal{R}$ (Proposition 2.1.2), $zeu \mathcal{R} zet$, that is, $b \mathcal{R} a$. We have shown that $aS \setminus \{0\} \subseteq R_a$, as required. □

A dual version of this last lemma can be shown in a similar manner:

Lemma 3.2.6 *For every $a \neq 0$ in S, $L_a = Sa \setminus \{0\}$.* □

The next result brings us a step nearer to showing the required connection between completely 0-simple semigroups and Rees matrix semigroups.

Lemma 3.2.7 *S is regular, and has exactly two $\mathcal{D}$-classes, namely $\{0\}$ and $D = S \setminus \{0\}$. If $a, b \in D$, then either $ab = 0$ or $ab \in R_a \cap L_b$. The latter occurs if and only if $L_a \cap R_b$ contains an idempotent.*

Proof Let $a, b \in S \setminus \{0\}$. Then $aSb \neq \{0\}$, since $aSb = \{0\}$ would imply that

$$S^2 = (SaS)(SbS) = S(aSb)S = \{0\},$$

and we know that this is not the case. Let u in S be such that $aub = c \neq 0$. Then, by Lemmas 3.2.5 and 3.2.6,

$$c \in \big(aS \setminus \{0\}\big) \cap \big(Sb \setminus \{0\}\big) = R_a \cap L_b,$$

and so $a \mathcal{D} b$. Since the $\mathcal{D}$-class $D = S \setminus \{0\}$ contains a (primitive) idempotent, it must consist entirely of regular elements (see Section 2.3), and since 0 also is a regular element we conclude that S is regular.

Finally, if $ab \neq 0$ then

$$ab \in \big(aS \setminus \{0\}\big) \cap \big(Sb \setminus \{0\}\big) = R_a \cap L_b.$$

By Proposition 2.3.7, this happens precisely when $L_a \cap R_b$ contains an idempotent. □

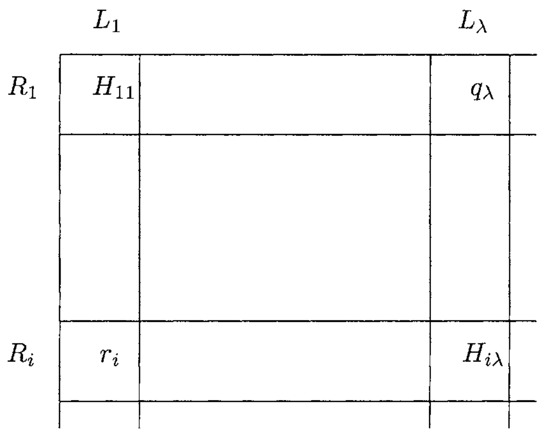
Let H be an $\mathcal{H}$-class of S, contained in the $\mathcal{D}$-class $D = S \setminus \{0\}$, and let $a, b \in H$. Then either $ab \in R_a \cap L_b = H$ or $ab = 0$. In the former case H is a group by Green's Theorem (Theorem 2.2.5). In the latter case we have $H^2 = \{0\}$, for if c and d are arbitrary members of H there exist x, y in S such that $c = xa$, $d = by$, and it then follows that

$$cd = (xa)(by) = x(ab)y = 0.$$

We may therefore refer to the $\mathcal{H}$-classes within D as either *group* $\mathcal{H}$-classes or *zero* $\mathcal{H}$-classes.

We are now ready to begin the process of shaping S into the form of a Rees matrix semigroup. Denote the set of non-zero $\mathcal{R}$-classes of S by I and the set of non-zero $\mathcal{L}$-classes by Λ. As a matter of notation we shall treat I and Λ as index sets and write the $\mathcal{R}$-classes as R_i ($i \in I$) and the $\mathcal{L}$-classes as L_λ ($\lambda \in \Lambda$). The $\mathcal{H}$-class $R_i \cap L_\lambda$ is denoted by $H_{i\lambda}$.

Since D is a regular $\mathcal{D}$-class, it follows by Proposition 2.3.2 that each R_i contains at least one group $\mathcal{H}$-class $H_{i\lambda}$. Equally each L_λ contains at least one group $\mathcal{H}$-class. Without loss of generality we may suppose that there is an element $1 \in I \cap \Lambda$ such that H_{11} is a group $\mathcal{H}$-class. Denote the identity of H_{11} by e. Notice that the choice of one particular group $\mathcal{H}$-class rather than another does not affect the abstract properties of the group, for we have seen (Proposition 2.3.6) that all the group $\mathcal{H}$-classes in a $\mathcal{D}$-class are isomorphic. The group H_{11} will turn out to be the group figuring in the Rees matrix semigroup we are looking for.



Now, again in a quite arbitrary way, we select for each i in I and λ in Λ an element r_i in H_{i1} and an element q_λ in $H_{1\lambda}$.

Since $r_i \;\mathcal{L}\; e$ we have by Proposition 2.3.3 that $r_ie = r_i$. Hence by Green's Lemma (2.2.2) it follows that $x \mapsto r_ix$ maps H_{11} onto H_{i1} in a one-one fashion. Similarly, we have $eq_\lambda = q_\lambda$, and again by Green's Lemma (2.2.1) it follows that $y \mapsto yq_\lambda$ maps H_{i1} onto $H_{i\lambda}$ in a one-one fashion. Thus (once we have chosen the elements r_i and q_λ) we have a unique expression r_iaq_λ $(a \in H_{11})$ for each element of $H_{i\lambda}$. Since

$$S \setminus \{0\} = \bigcup \{H_{i\lambda} : i \in I,\ \lambda \in \Lambda\}$$

and since the union is disjoint, we thus have a bijection

$$\phi : (I \times H_{11} \times \Lambda) \cup \{0\} \to S$$

given by

$$(i, a, \lambda)\phi = r_iaq_\lambda, \quad 0\phi = 0.$$

The final stage in the argument is to introduce a multiplication into $(I \times H_{11} \times \Lambda) \cup \{0\}$ so as to make it into a regular Rees matrix semigroup, and so that the bijection ϕ becomes an isomorphism. This amounts to defining a suitable sandwich matrix, and since

$$(r_iaq_\lambda)(r_jbq_\mu) = r_i(aq_\lambda r_jb)q_\mu,$$

it seems reasonable to define $p_{\lambda i}$ as $q_\lambda r_i \quad (i \in I,\ \lambda \in \Lambda)$. By Lemma 3.2.7,

$$p_{\lambda i} \in R_{q_\lambda} \cap L_{r_i} = H_{11}$$

if and only if the $\mathcal{H}$-class

$$L_{q_\lambda} \cap R_{r_i} = H_{i\lambda}$$

contains an idempotent, that is, if and only if $H_{i\lambda}$ is a group $\mathcal{H}$-class; otherwise $p_{\lambda i} = 0$. Hence $P = (p_{\lambda i})$ is a matrix with entries in H_{11}^0. It is, moreover, regular in the sense of (3.2.2), since the already observed property that each $\mathcal{R}$-class and each $\mathcal{L}$-class in D contains at least one group $\mathcal{H}$-class translates into exactly the property of regularity for P.

It is now entirely routine to check that ϕ is an isomorphism from the Rees matrix semigroup $\mathcal{M}^0[H_{11}; I, \Lambda; P]$ onto S. □

Another possible approach to the idea of a completely 0-simple semigroup is via the idea of 0-minimal one-sided ideals. A left [right] ideal is called 0-minimal if it is minimal within the set of all non-zero left [right] ideals. Then we have

Lemma 3.2.8 *Let S be a 0-simple semigroup, and suppose that S contains a 0-minimal left ideal L such that $L^2 \neq \{0\}$. Then $L = Sa$ for every a in $L \setminus \{0\}$.*

Proof The set Sa is easily seen to be a left ideal of S contained in L, and so by the 0-minimality of L either $Sa = L$ or $Sa = \{0\}$. Suppose first that $Sa = \{0\}$. Then $\{a, 0\}$ is a non-zero left ideal of S contained in L and so must coincide with L. But then $L^2 = \{0, a\}^2 = \{0\}$, contrary to assumption, and so we conclude that $Sa = L$. □

Next, we have

Lemma 3.2.9 *If S is 0-simple and contains a 0-minimal left ideal, then S is the union of its 0-minimal left ideals.*

Proof Let S be a 0-simple semigroup, and let L be a 0-minimal left ideal. Then LS is an ideal of S and so either $LS = \{0\}$ or $LS = S$. Suppose first that $LS = \{0\}$. Then certainly $LS \subseteq L$, and so L is in fact a two-sided ideal of S. Since $L \neq \{0\}$ it follows that $L = S$, and hence

$$S^2 = LS = \{0\},$$

a contradiction. We conclude that $LS = S$, and so there exists s in S such that $Ls \neq \{0\}$. Such a 'translate' Ls of L is evidently a left ideal, and it is even 0-minimal, for if B is a non-zero left ideal of S contained in Ls, then $A = \{a \in L : as \in B\}$, being a non-zero left ideal of S contained in L, coincides with L, and so $B = As = Ls$.

Now let

$$M = \bigcup \{L_s : s \in S\}.$$

Then certainly M is a non-zero left ideal. It is actually a two-sided ideal, for if $m \in Ls \subseteq M$ then $mt \in L(st) \subseteq M$ for all t in S. Hence, since S is 0-simple, $M = S$, and we have shown that S is a union of 0-minimal left ideals Ls. □

Lemma 3.2.10 *Let S be a 0-simple semigroup containing at least one 0-minimal left ideal and at least one 0-minimal right ideal. For every 0-minimal left ideal L of S there exists a 0-minimal right ideal R such that:*

(1) *$LR = S$;*

(2) *RL is a 0-group;*

(3) *the identity element e of RL is a primitive idempotent.*

Thus S is completely 0-simple.

Proof (1) Let L be a 0-minimal left ideal. Now for every 0-minimal right ideal R the set LR is a two-sided ideal and so either $LR = \{0\}$ or $LR = S$. From the proof of the previous lemma we know that there exists s in S such that $Ls \neq \{0\}$. From the dual of Lemma 3.2.9 we know that s lies in some 0-minimal right ideal R, and it follows that LR, which contains Ls, must coincide with S.

(2) We have now chosen R such that $LR = S$. Notice that $RL \subseteq R \cap L$. We show that it is a 0-group by showing that $RLa = aRL = RL$ for all $a \neq 0$ in RL. Such an a clearly belongs to $R \setminus \{0\}$, and so $R = aS$ by the dual to Lemma 3.2.8. Since $S = LR = LaS$ we certainly have $La \neq \{0\}$; hence, as observed in the proof of Lemma 3.2.9, La is a 0-minimal left ideal. However, $La \subseteq L$ since $a \in RL \subseteq L$. Hence $La = L$, and it follows that $RLa = RL$. The proof that $aRL = RL$ proceeds in just the same way.

(3) Let e be the identity of the 0-group RL, and suppose that f is a non-zero idempotent in S such that $f \leq e$, that is, such that $ef = fe = f$. Now since $e \in R \cap L$ it follows by Lemma 3.2.8 that $R = eS$ and $L = Se$. Hence $eSe = eS^2e = (eS)(Se) = RL$. Hence $f = efe \in RL$, and so f coincides with e, the unique idempotent in the group $RL \setminus \{0\}$. □

In effect we have now proved part of the following theorem. By a *group-bound* semigroup we mean a semigroup S with the property that every element a in S has a power a^n $(n \geq 1)$ lying in a subgroup of S. Notice that every finite semigroup is group-bound, by Proposition 1.2.3.

Theorem 3.2.11 *Let S be a 0-simple semigroup. Then the following conditions are equivalent:*

(1) *S is completely 0-simple;*

(2) *S is group-bound;*

(3) *S satisfies $\min_L$ and $\min_R$;*

(4) *S contains at least one 0-minimal left ideal and at least one 0-minimal right ideal.*

Proof We have in fact established that (4) $\Rightarrow$ (1). That (1) $\Rightarrow$ (2) is an easy consequence of the remarks following the proof of Lemma 3.2.7, since every non-zero element a either lies in a group $\mathcal{H}$-class or has the property that a^2 belongs to the trivial group $\{0\}$.

To show that (2) $\Rightarrow$ (3), suppose that S is group-bound, and let a, b in $S \setminus \{0\}$ be such that $L_a \leq L_b$. Then there exist x, y and u in S such that $a = ub$ and $b = xay$. Thus

$$b = (xu)by = \cdots = (xu)^n by^n,$$

and we may assume that n has been chosen so that $(xu)^n$ belongs to a subgroup of S with identity e. For notational simplicity write $(xu)^n = g$ and denote its inverse within the subgroup by g^{-1}. Then

$$eb = egby^n = gby^n = b,$$

and so

$$b = g^{-1}gb = g^{-1}(xu)^n b = g^{-1}(xu)^{n-1}x(ub) = g^{-1}(xu)^{n-1}xa.$$

We deduce that $L_a = L_b$. We have shown that in the semigroup S we have the implication $L_a \leq L_b \Rightarrow L_a = L_b$ among the non-zero $\mathcal{L}$-classes, and this certainly implies the condition $\min_L$. The property $\min_R$ is established in a similar way.

The implication (3) $\Rightarrow$ (4) is clear. □

3.3 COMPLETELY SIMPLE SEMIGROUPS

Let S be a semigroup without zero. We shall say that S is *completely simple* if S is simple and if it contains a primitive idempotent (by which we now mean an idempotent which is minimal within the set of *all* idempotents of S). The semigroup S^0 is then a completely 0-simple semigroup with a prime zero element, as defined by equation (3.1.1), and so for the most part we can deduce results about completely simple semigroups from corresponding results about completely 0-simple semigroups. Certainly we have the following simplified version of the Rees Theorem, due in effect to Suschkewitsch (1928):

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

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

Then S is a completely simple semigroup.

Conversely, every completely simple semigroup is isomorphic to a semigroup constructed in this way. □

We denote the semigroup $I \times G \times \Lambda$ with the given multiplication by

$$\mathcal{M}[G; I, \Lambda; P].$$

We also have a direct analogue of Theorem 3.2.11, but before stating it we require another definition. The next chapter will be devoted to what are called 'completely regular' semigroups. One possible definition of this concept (equivalent to the one we shall adopt in Chapter 4 as our main definition) is that a semigroup is *completely regular* if every element a of S lies in a subgroup of S. Notice that a completely regular semigroup is certainly group-bound.

Then we have the following analogue of Theorem 3.2.11:

Theorem 3.3.2 *Let S be a simple semigroup (without zero). Then the following conditions are equivalent:*

(1) *S is completely simple;*
(2) *S is completely regular;*
(3) *S satisfies $\min_L$ and $\min_R$;*
(4) *S contains at least one minimal left ideal and at least one minimal right ideal.* □

Further characterizations are available:

Theorem 3.3.3 *Let S be a semigroup without zero. Then the following conditions are equivalent:*

(1) *S is completely simple;*
(2) *S is regular, and has the 'weak cancellation' property: for all a, b, c in S,*

$$[ca = cb \ \text{ and } \ ac = bc] \ \Rightarrow \ a = b;$$

(3) *S is regular, and for all a in S*

$$aba = a \ \Rightarrow \ bab = b;$$

(4) *S is regular and every idempotent is primitive.*

Proof (1) $\Rightarrow$ (2). Certainly a completely simple semigroup S is regular, since every element of S lies in a subgroup. By the Rees–Suschkewitsch Theorem (Theorem 3.3.1) we may assume that $S = \mathcal{M}[G; I, \Lambda; P]$. Let $a = (i, x, \lambda)$, $b = (j, y, \mu)$, $c = (k, z, \nu)$, and suppose that $ca = cb$ and $ac = bc$. Then

$$(k, zp_{\nu i}x, \lambda) = (k, zp_{\nu j}y, \mu), \quad (i, xp_{\lambda k}z, \nu) = (j, yp_{\mu k}z, \nu),$$

from which it follows that $i = j$, $\lambda = \mu$ and $x = y$. Thus $a = b$, as required.

(2) $\Rightarrow$ (3). Suppose that S is regular, and that $aba = a$. Then

$$a(bab) = ab \ \text{ and } \ (bab)a = ba;$$

hence $bab = b$ by the weak cancellative property.

(3) $\Rightarrow$ (4). Since S is regular, it contains an idempotent e. Suppose that f is an idempotent such that $f \leq e$, that is, such that $ef = fe = f$. It follows that $fef = f$ and hence $efe = e$ by (3). But the condition that

$f \leq e$ gives $efe = f$. Hence $f = e$, and so e, which was an arbitrary idempotent of S, is primitive.

(4) $\Rightarrow$ (1). In effect we have to show that S is simple. Since S is regular, every $\mathcal{D}$-class (and so certainly every $\mathcal{J}$-class) contains an idempotent. Consider a typical $\mathcal{J}$-class J_e where e is a (necessarily primitive) idempotent. We show that J_e is a minimal $\mathcal{J}$-class. For suppose that $J_f \leq J_e$, where f is another idempotent. Then $f = xey$ for some x, y in S^1, and the element $g = eyfxe$ has the properties

$$g^2 = eyfxe^2yfxe = eyf^3xe = eyfxe = g, \quad eg = ge = g.$$

Thus $g \leq e$ and so $g = e$. We now have

$$xey = f, \quad eyfxe = e,$$

and so $J_f = J_e$.

By virtue of the remark preceding the statement of Proposition 3.1.4, we now conclude that *every* $\mathcal{J}$-class in S coincides with the kernel $K(S)$. Thus $S = K(S)$, and is simple by Proposition 3.1.4. □

It is reasonable to ask whether a regular semigroup with zero in which every non-zero idempotent is primitive is necessarily completely 0-simple. In fact it is not, but it is a semigroup whose structure is relatively easy to describe. The argument above applies to non-zero $\mathcal{J}$-classes J_e and J_f, for we can easily see that if the idempotent $g = eyfxe$ were equal to 0 then we could deduce the obviously false result that

$$f = f^3 = xeyfxey = xgy = 0.$$

We deduce that if S has the property that all of its non-zero idempotents are primitive then every non-zero $\mathcal{J}$-class is minimal within the set of all non-zero $\mathcal{J}$-classes. Suppose now that $J_e \neq J_f$, which is equivalent to saying that $J_e \cap J_f = \emptyset$. Then for all x in J_e and y in J_f,

$$J_{xy} \leq J_e, \quad J_{xy} \leq J_f.$$

If $xy \neq 0$ then the minimality of J_e and J_f gives

$$J_e = J_{xy} = J_f,$$

which is impossible. Hence $xy = 0$. We have shown that $J_e \neq J_f$ implies that $J_e J_f = \{0\}$.

Now each $J_e \cup \{0\}$ is a principal factor of S and is clearly not a null semigroup. Hence it is 0-simple by Theorem 3.1.6, and since it contains the primitive idempotent e it must be completely 0-simple. In effect we have shown that

$$S = \bigcup_{i \in I} S_i \,,$$

where each S_i is a completely 0-simple semigroup, and where

$$S_i \cap S_j = S_i S_j = \{0\}$$

if $i \neq j$. We say that S is a *0-direct union* of completely 0-simple semigroups.

In effect we have now proved half of the following theorem. The converse half is an easy direct verification.

Theorem 3.3.4 *Let S be a semigroup with zero. The following statements are equivalent:*

(1) *S is regular and every non-zero idempotent of S is primitive;*

(2) *S is a 0-direct union of completely 0-simple semigroups.* □

3.4 ISOMORPHISM AND NORMALIZATION

There is often room for argument as to what constitutes a satisfactory structure theory in algebra. In the context of abelian group theory Kaplansky (1954) proposed certain test questions that a truly satisfactory structure theory ought to enable one to answer. Kaplansky's questions do not seem very appropriate to semigroup theory, and two questions have traditionally been asked in the semigroup context. The first is 'Does there exist an associated isomorphism theorem?', and the second is 'Does the theory enable one to give an explicit description of the congruences?' The first question is inescapable, and applies in every algebraic context, for an alleged structure theorem that does not have an associated isomorphism theorem is of little use. If (to be fanciful) we show that some mathematical object W called a 'wall' is built out of n 'bricks' $B_1, B_2, \ldots, B_n$, then our theorem is undoubtedly more useful if one can say that any other decomposition of the same wall W into bricks $B'_1, B'_2, \ldots, B'_m$ must be such that $m = n$ and $B'_i \simeq B_i$ for all i. The second question, regarding the description of congruences, is much more characteristic of semigroup theory. Its answer often involves work of a tediously detailed nature, and in a book of modest size it will usually be sufficient to give a reference to the literature. In the case of the Rees Theorem for completely 0-simple semigroups, however, it seems reasonable to make an exception, partly because of the great importance of the theorem in semigroup theory, but also because the classification of congruences on completely 0-simple semigroups has attracted the attention of several authors. Differing accounts have been given by Gluskin (1956), Tamura (1960), Preston (1961, 1965), Lallement (1967, 1974) and Kapp and Schneider (1969). The account given in Section 3.5 most closely resembles Lallement's (1974) version.

First, however, we prove an isomorphism theorem. It is fairly clear that in our proof that every completely 0-simple semigroup S is isomorphic to some Rees matrix semigroup $\mathcal{M}^0[G; I, \Lambda; P]$, the group G we encountered was 'intrinsic' and so were the sets I and Λ (since they index the $\mathcal{R}$- and $\mathcal{L}$-classes of S, respectively). The elements r_i and q_λ are, by contrast, open to some choice, and so we should expect the entries $p_{\lambda i} = q_\lambda r_i$ of the sandwich matrix to be subject to some variation. In fact we have

Theorem 3.4.1 *Two regular Rees matrix semigroups*

$$S = \mathcal{M}^0[G; I, \Lambda; P] \quad \textit{and} \quad T = \mathcal{M}^0[K; J, \mathrm{M}; Q]$$

are isomorphic if and only if there exist an isomorphism $\theta : G \to K$, *bijections* $\psi : I \to J$, $\chi : \Lambda \to \mathrm{M}$ *and elements* u_i $(i \in I)$, v_λ $(\lambda \in \Lambda)$ *such that*

$$p_{\lambda i}\theta = v_\lambda q_{\lambda\chi, i\psi} u_i \tag{3.4.1}$$

for all i *in* I *and* λ *in* Λ.

Proof Certainly if we are given θ, ψ, χ, $\{u_i : i \in I\}$ and $\{v_\lambda : \lambda \in \Lambda\}$ with the given properties, then it is easy to check that the map $\phi : S \to T$ given by

$$(i, a, \lambda)\phi = (i\psi, u_i(a\theta)v_\lambda, \lambda\chi) \quad \big((i, a, \lambda) \in S\big) \tag{3.4.2}$$

is an isomorphism.

Conversely, if $\phi : S \to T$ is an isomorphism, it maps the non-zero $\mathcal{R}$-classes of S in a one-one fashion onto the non-zero $\mathcal{R}$-classes of T, and so there is a bijection $\psi : I \to J$ such that $(i, a, \lambda)\phi \in R_{i\psi}$. Similarly there is a bijection $\chi : \Lambda \to \mathrm{M}$ such that $(i, a, \lambda)\phi \in L_{\lambda\chi}$. Moreover, ϕ maps group $\mathcal{H}$-classes onto group $\mathcal{H}$-classes, and so $p_{\lambda i} \neq 0$ if and only if $q_{\lambda\chi, i\psi} \neq 0$.

Now choose a group $\mathcal{H}$-class of S, and as in the proof of the Rees Theorem denote it by H_{11} (with $1 \in I \cap \Lambda$) without essential loss of generality. Its image under the isomorphism ϕ is the group $\mathcal{H}$-class $H_{1\chi, 1\psi}$. We now have three isomorphisms, an isomorphism $\alpha : G \to H_{11}$ sending the element x of G to $(1, p_{11}^{-1}x, 1)$, an isomorphism $\beta : K \to H_{1\psi, 1\chi}$ sending the element y of K to $(1\psi, q_{1\chi, 1\psi}^{-1}y, 1\chi)$, and the isomorphism

$$\phi|_{H_{11}} : H_{11} \to H_{1\psi, 1\chi}.$$

Let

$$\theta = \alpha\phi|_{H_{11}}\beta^{-1} : G \to H.$$

Then θ is an isomorphism, and has the property that for all x in G

$$(1, p_{11}^{-1}x, 1)\phi = x\alpha\phi|_{H_{11}} = x\theta\beta = (1\psi, q_{1\chi, 1\psi}^{-1}(x\theta), 1\chi).$$

Now notice that for an arbitrary element (i, a, λ) of S we can write

$$(i, a, \lambda) = (i, e, 1)(1, p_{11}^{-1}a, 1)(1, p_{11}^{-1}, \lambda).$$

Define the elements u_i, v_λ in K by

$$(i, e, 1)\phi = (i\psi, u_i, 1\chi), \quad (1, p_{11}^{-1}, \lambda)\phi = (1\psi, q_{1\chi, 1\psi}^{-1}v_\lambda, \lambda\chi).$$

Then by the morphism property of ϕ it follows that

$$(i,a,\lambda)\phi = (i\psi, u_i, 1\chi)(1\psi, q_{1\chi,1\psi}^{-1}(a\theta), 1\chi)(1\psi, q_{1\chi,1\psi}^{-1}v_\lambda, \lambda\chi)$$
$$= (i\psi, u_i(a\theta)v_\lambda, \lambda\chi).$$

Now if $p_{\lambda i} \neq 0$ then

$$\begin{aligned}(i\psi, u_i(p_{\lambda i}\theta)v_\lambda, \lambda\chi) &= (i, p_{\lambda i}, \lambda)\phi \\ &= [(i,e,\lambda)(i,e,\lambda)]\phi = [(i,e,\lambda)\phi][(i,e,\lambda)\phi] \\ &= (i\psi, u_iv_\lambda, \lambda\chi)(i\psi, u_iv_\lambda, \lambda\chi) \\ &= (i\psi, u_iv_\lambda q_{\lambda\chi,i\psi}u_iv_\lambda, \lambda\chi),\end{aligned}$$

and from this it follows, since K is a group, that

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

as required. Notice finally that if $p_{\lambda i} = 0$ then $q_{\lambda\chi,i\psi} = 0$ also. Thus the equality (3.4.1) holds in all cases. □

An alternative way of expressing the result is sometimes useful. Here we consider two Rees matrix semigroups $\mathcal{M}^0[G; I, \Lambda; P]$ and $\mathcal{M}^0[G; I, \Lambda; Q]$. If there exist a $\Lambda \times \Lambda$ diagonal matrix V with entries v_λ in G and an $I \times I$ diagonal matrix U with entries u_i in G such that $P = VQU$, then the map $(i,a,\lambda) \mapsto (i, u_iav_\lambda, \lambda)$ is an isomorphism from $\mathcal{M}^0[G; I, \Lambda; P]$ onto $\mathcal{M}^0[G; I, \Lambda; Q]$.

The isomorphism theorem can be seen as specifying the degree to which there is variability in the sandwich matrix P attaching to a completely 0-simple semigroup. In the case where S is completely simple (without zero) and where every $\mathcal{H}$-class is a group there is an obvious and natural choice for the elements r_i in H_{i1} $(i \in I)$ and q_λ in $H_{1\lambda}$ $(\lambda \in \Lambda)$: we simply take r_i as the identity of the group H_{i1} and q_λ as the identity of the group $H_{1\lambda}$. In particular $r_1 = q_1 = e$, the identity of the group H_{11}. Then for all i in I

$$p_{i1} = q_1r_i = er_i = e,$$

since the idempotent r_i acts as a right identity within the $\mathcal{L}$-class L_1, and similarly $p_{\lambda 1} = e$ for all λ in Λ. We obtain a sandwich matrix $P = (p_{\lambda i})$ which is *normal*, in the sense that every entry in the first row and the first column of P is equal to e. We have proved

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

A normalization theorem is available for the completely 0-simple case, but is necessarily more complicated. Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely 0-simple semigroup, and as usual let us denote the identity element of G by e. For each i in I let $\Lambda_i = \{\lambda \in \Lambda : p_{\lambda i} \neq 0\}$, and for each λ in Λ

let $I_\lambda = \{i \in I : p_{\lambda i} \neq 0\}$. Then define equivalence relations $\mathcal{E}_I$, $\mathcal{E}_\Lambda$ on I and Λ, respectively, by

$$\mathcal{E}_I = \{(i,j) \in I \times I : \Lambda_i = \Lambda_j\}, \quad \mathcal{E}_\Lambda = \{(\lambda,\mu) \in \Lambda \times \Lambda : I_\lambda = I_\mu\}.$$

Denote the $\mathcal{E}_I$-class containing i by i^* and the $\mathcal{E}_\Lambda$-class containing λ by λ^*. We can think of the sets I and Λ as ordered in such a way that the matrix $P = (p_{\lambda i})$ is partitioned into blocks $i^* \times \lambda^*$, and the definition of the equivalences $\mathcal{E}_I$ and $\mathcal{E}_\Lambda$ ensures that either $p_{\lambda i} = 0$ for all i in i^* and λ in λ^*, or $p_{\lambda i} \neq 0$ for all i in i^* and λ in λ^*. In the latter case we say that $i^* \times \lambda^*$ is a *non-zero block.* The regularity of the matrix P (see equation (3.2.2)) ensures that for every i^* there exists λ^* such that $i^* \times \lambda^*$ is a non-zero block, and for every λ^* there exists i^* such that $i^* \times \lambda^*$ is a non-zero block.

Now for every i^* choose an $\mathcal{E}_\Lambda$-class $\nu^*(i^*)$ such that $i^* \times \nu^*(i^*)$ is a non-zero block, and choose an element $\nu(i^*)$ within the $\mathcal{E}_\Lambda$-class $\nu^*(i^*)$. Thus $p_{\nu(i^*),i} \neq 0$ for all i in i^*. Similarly, for every λ^* choose an $\mathcal{E}_I$-class $n^*(\lambda^*)$ such that $n^*(\lambda^*) \times \lambda^*$ is a non-zero block, and choose an element $n(\lambda^*)$ within the $\mathcal{E}_I$-class $n^*(\lambda^*)$. Thus $p_{\lambda,n(\lambda^*)} \neq 0$ for all λ in λ^*.

The matrix P is called *normal* if (i) for every $\mathcal{E}_I$-class i^* there exists λ such that $p_{\lambda i} = e$ for all i in i^*, and (ii) for every $\mathcal{E}_\Lambda$-class λ^* there exists i such that $p_{\lambda i} = e$ for all λ in λ^*. Then we have the following theorem:

Theorem 3.4.3 *Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely 0-simple semigroup. Then $S \simeq \mathcal{M}^0[G; I, \Lambda; R]$, where $R = (r_{\lambda i})$ is a normal matrix.*

Proof We use the remark following Theorem 3.4.1. Premultiplying P by a diagonal matrix V^{-1} with entries v_λ^{-1} in G in effect multiplies row λ of P by v_λ^{-1} for each λ in Λ. Define v_λ to be $p_{\lambda,n(\lambda^*)}$ for each λ in λ^*. Then the $\Lambda \times I$ matrix $Q = (q_{\lambda i}) = V^{-1}P$ has the property that $q_{\lambda,n(\lambda^*)} = e$ for every λ in λ^* and every λ^* in $\Lambda/\mathcal{E}_\Lambda$. Now define u_i to be $q_{\nu(i^*),i}$ for each i in i^*. Then $R = V^{-1}PU^{-1} = (r_{\lambda i})$ is a normal matrix, and $(i, a, \lambda) \mapsto (i, u_i a v_\lambda, \lambda)$ is an isomorphism from $\mathcal{M}^0[G; I, \Lambda; P]$ onto $\mathcal{M}^0[G; I, \Lambda; R]$. □

3.5 CONGRUENCES ON COMPLETELY 0-SIMPLE SEMIGROUPS*

We now give a description of the congruences on a completely 0-simple semigroup, or rather (what by the Rees Theorem amounts to the same thing) on a Rees matrix semigroup $\mathcal{M}^0[G; I, \Lambda; P]$, in the case where the sandwich matrix $P = (p_{\lambda i})$ is regular in the sense of (3.2.2).

First, if ρ is a congruence on a completely 0-simple semigroup S then 0ρ, the ρ-class containing 0, is easily seen to be an ideal of S. It follows by the 0-simplicity of S that either $0\rho = \{0\}$ or $0\rho = S$. In the latter case $\rho = S \times S$, the universal congruence, and so there will be little loss if we restrict attention for the moment to what we shall call *proper* congruences ρ, for which $0\rho = \{0\}$.

Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely 0-simple semigroup. As in the previous section we define an equivalence relation $\mathcal{E}_I$ on I by the rule that

$$(i,j) \in \mathcal{E}_I \text{ if } \{\lambda \in \Lambda : p_{\lambda i} = 0\} = \{\lambda \in \Lambda : p_{\lambda j} = 0\}, \tag{3.5.1}$$

and an equivalence relation $\mathcal{E}_\Lambda$ on Λ by the rule that

$$(\lambda, \mu) \in \mathcal{E}_\Lambda \text{ if } \{i \in I : p_{\lambda i} = 0\} = \{i \in I : p_{\mu i} = 0\}. \tag{3.5.2}$$

If ρ is a proper congruence on S, we define a relation ρ_I on I by the rule that $(i,j) \in \rho_I$ if $(i,j) \in \mathcal{E}_I$ and if

$$(i, p_{\lambda i}^{-1}, \lambda) \; \rho \; (j, p_{\lambda j}^{-1}, \lambda) \tag{3.5.3}$$

for every λ such that $p_{\lambda i}$ (and hence also $p_{\lambda j}$) is non-zero.

It is clear that ρ_I is reflexive and symmetric. To show that it is transitive, note first that if $(i,j) \in \rho_I \subseteq \mathcal{E}_I$ and $(j,k) \in \rho_I \subseteq \mathcal{E}_I$ then $(i,k) \in \mathcal{E}_I$ by the transitivity of $\mathcal{E}_I$. If λ is such that $p_{\lambda i} \neq 0$ then $p_{\lambda j}$ and $p_{\lambda k}$ are non-zero also, and

$$(i, p_{\lambda i}^{-1}, \lambda) \; \rho \; (j, p_{\lambda j}^{-1}, \lambda), \quad (j, p_{\lambda j}^{-1}, \lambda) \; \rho \; (k, p_{\lambda k}^{-1}, \lambda).$$

It now follows by the transitivity of ρ that ρ_I is an equivalence relation on I.

By analogy, we have an equivalence relation ρ_Λ on Λ defined by the rule that $(\lambda, \mu) \in \rho_\Lambda$ if $(\lambda, \mu) \in \mathcal{E}_\Lambda$ and

$$(i, p_{\lambda i}^{-1}, \lambda) \; \rho \; (i, p_{\mu i}^{-1}, \mu) \tag{3.5.4}$$

for every i in I such that $p_{\lambda i}$ (and hence also $p_{\mu i}$) is non-zero.

Once again we select an arbitrary group $\mathcal{H}$-class of S and call it (without loss of generality) H_{11}. Let

$$N_\rho = \{a \in G : (1, a, 1) \; \rho \; (1, e, 1)\}, \tag{3.5.5}$$

where e is the identity of G. Then certainly $e \in N_\rho$, and so $N_\rho \neq \emptyset$. In fact,

Lemma 3.5.1 *N_ρ is a normal subgroup of G.*

Proof Suppose that $a, b \in N_\rho$, so that

$$(1, a, 1) \; \rho \; (1, e, 1) \text{ and } (1, b, 1) \; \rho \; (1, e, 1).$$

Then

$$(1, a, 1)(1, p_{11}^{-2}, 1)(1, b, 1) \; \rho \; (1, e, 1)(1, p_{11}^{-2}, 1)(1, e, 1),$$

that is,

$$(1, ab, 1) \; \rho \; (1, e, 1).$$

Thus $ab \in N_\rho$. Again, if $a \in N_\rho$ then

$$(1, a, 1)(1, p_{11}^{-1}a^{-1}, 1) \; \rho \; (1, e, 1)(1, p_{11}^{-1}a^{-1}, 1),$$

that is, $(1, e, 1)\ \rho\ (1, a^{-1}, 1)$. Thus $a^{-1} \in N_\rho$. We have shown that N_ρ is a subgroup of G. To show that it is normal, consider a in N_ρ and g in G. Then

$$(1, g^{-1}p_{11}^{-1}, 1)(1, a, 1)(1, p_{11}^{-1}g, 1)\ \rho\ (1, g^{-1}p_{11}^{-1}, 1)(1, e, 1)(1, p_{11}^{-1}g, 1),$$

that is, $(1, g^{-1}ag, 1)\ \rho\ (1, e, 1)$. Thus $g^{-1}ag \in N_\rho$. □

Before moving on, let us record the easily verified fact that, for all a, b in G,

$$(1, a, 1)\ \rho\ (1, b, 1) \text{ if and only if } ab^{-1} \in N_\rho. \tag{3.5.6}$$

So far, then, we have seen that a proper congruence ρ determines equivalence relations ρ_I $(\subseteq \mathcal{E}_I)$ and ρ_Λ $(\subseteq \mathcal{E}_\Lambda)$ on the sets I, Λ, respectively, and a normal subgroup N_ρ of G. These three objects are not independent, and to describe the nature of their interdependence we now require a notion due to Lallement (1967). If $p_{\lambda i}$, $p_{\mu i}$, $p_{\lambda j}$ and $p_{\mu j}$ are all non-zero, then $p_{\lambda i}p_{\mu i}^{-1}p_{\mu j}p_{\lambda j}^{-1}$ is an element of G. We call it an *extract* of the matrix P, and write

$$q_{\lambda\mu ij} = p_{\lambda i}p_{\mu i}^{-1}p_{\mu j}p_{\lambda j}^{-1}. \tag{3.5.7}$$

It is easy to verify the following identities, which will be required later:

$$q_{\lambda\mu ij}q_{\lambda\mu jk} = q_{\lambda\mu ik}, \tag{3.5.8}$$

and

$$p_{\lambda i}^{-1}q_{\lambda\mu ij}^{-1}p_{\lambda i} \cdot p_{\mu i}^{-1}q_{\mu\nu ij}^{-1}p_{\mu i} = p_{\lambda i}^{-1}q_{\lambda\nu ij}^{-1}p_{\lambda i}. \tag{3.5.9}$$

Then we can prove

Lemma 3.5.2 *Let ρ be a proper congruence on a completely 0-simple semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$, and let i, j, λ, μ be such that $p_{\lambda i}$, $p_{\mu i}$, $p_{\lambda j}$ and $p_{\mu j}$ are all non-zero. If either $(i, j) \in \rho_I$ or $(\lambda, \mu) \in \rho_\Lambda$ then $q_{\lambda\mu ij} \in N_\rho$.*

Proof Suppose first that $(i, j) \in \rho_I$. Then

$$(i, p_{\mu i}^{-1}, \mu)\ \rho\ (j, p_{\mu j}^{-1}, \mu),$$

and so

$$(1, e, \lambda)(i, p_{\mu i}^{-1}, \mu)(j, p_{\lambda j}^{-1}, 1)\ \rho\ (1, e, \lambda)(j, p_{\mu j}^{-1}, \mu)(j, p_{\lambda j}^{-1}, 1).$$

That is,

$$(1, p_{\lambda i}p_{\mu i}^{-1}p_{\mu j}p_{\lambda j}^{-1}, 1)\ \rho\ (1, e, 1),$$

and so $q_{\lambda\mu ij} \in N_\rho$. Similarly, if $(\lambda, \mu) \in \rho_\Lambda$ then

$$(1, e, \lambda)(i, p_{\lambda i}^{-1}, \lambda)(j, p_{\lambda j}^{-1}, 1)\ \rho\ (1, e, \lambda)(i, p_{\mu i}^{-1}, \mu)(j, p_{\lambda j}^{-1}, 1).$$

That is,

$$(1, e, 1)\ \rho\ (1, p_{\lambda i}p_{\mu i}^{-1}p_{\mu j}p_{\lambda j}^{-1}, 1),$$

and so again $q_{\lambda\mu ij} \in N_\rho$. □

This result motivates the following definition. A triple $(N, \mathcal{S}, \mathcal{T})$ consisting of a normal subgroup N of G, an equivalence relation $\mathcal{S}$ on I and an equivalence relation $\mathcal{T}$ on Λ will be called *linked* if

(L1) $\mathcal{S} \subseteq \mathcal{E}_I$, $\mathcal{T} \subseteq \mathcal{E}_\Lambda$;
(L2) if i, j in I and λ, μ in Λ are such that $p_{\lambda i}$, $p_{\lambda j}$, $p_{\mu i}$ and $p_{\mu j}$ are all non-zero, then $q_{\lambda\mu ij} \in N$ whenever either $(i,j) \in \mathcal{S}$ or $(\lambda,\mu) \in \mathcal{T}$.

What we have shown thus far is that there exists a map Γ from the set of proper congruences of S into the set of linked triples, associating ρ with the linked triple $(N_\rho, \rho_I, \rho_\Lambda)$. We shall eventually see that Γ is a bijection.

As a first step towards a full examination of the map Γ we have

Lemma 3.5.3 *If the elements (i,a,λ) and (j,b,μ) of S are such that $(i,a,\lambda)\ \rho\ (j,b,\mu)$ then $(i,j) \in \rho_I$.*

Proof We need to show that $(i,j) \in \mathcal{E}_I$ and that

$$(i, p_{\xi i}^{-1}, \xi)\ \rho\ (j, p_{\xi j}^{-1}, \xi)$$

for every ξ in Λ such that $p_{\xi i}$ (and hence also $p_{\xi j}$) is non-zero. To show the first of these properties, suppose that ξ in Λ is such that $p_{\xi i} = 0$. Then

$$0 = (1,e,\xi)(i,a,\lambda)\ \rho\ (1,e,\xi)(j,b,\mu),$$

and so, since ρ is a proper congruence, $(1,e,\xi)(j,b,\mu) = 0$. We have shown that $p_{\xi i} = 0 \ \Rightarrow\ p_{\xi j} = 0$, and since the opposite implication is provable in exactly the same way we deduce that $(i,j) \in \mathcal{E}_I$. A similar argument shows that $(\lambda,\mu) \in \mathcal{E}_\Lambda$.

To prove the second part we begin with the given information that $(i,a,\lambda)\ \rho\ (j,b,\mu)$. We premultiply both sides by $(1,e,\xi)$, where ξ is any element of Λ such that $p_{\xi i}$ and $p_{\xi j}$ are non-zero, and we postmultiply both sides by $(k, p_{11}, 1)$, where k is any element of I such that $p_{\lambda k}$ and $p_{\mu k}$ are both non-zero. A routine calculation then gives

$$(1, p_{\xi i} a p_{\lambda k} p_{11}, 1)\ \rho\ (1, p_{\xi j} b p_{\mu k} p_{11}, 1).$$

Writing this for brevity as $(1,c,1)\ \rho\ (1,d,1)$, we deduce from (3.5.6) that $cd^{-1} \in N_\rho$. Hence $c^{-1}(cd^{-1})^{-1}c = c^{-1}d \in N_\rho$ by the normality of N_ρ, and we deduce that $(1,c^{-1},1)\ \rho\ (1,d^{-1},1)$. In the original notation this gives

$$(1, p_{11}^{-1} p_{\lambda k}^{-1} a^{-1} p_{\xi i}^{-1}, 1)\ \rho\ (1, p_{11}^{-1} p_{\mu k}^{-1} a^{-1} p_{\xi j}^{-1}, 1).$$

Now premultiply the left-hand side of this relation by $(i,a,\lambda)(k,e,1)$ and the right-hand side by the ρ-equivalent element $(j,b,\mu)(k,e,1)$, and postmultiply both sides by $(1, p_{11}^{-1}, \xi)$. A routine calculation gives

$$(i, p_{\xi i}^{-1}, \xi)\ \rho\ (j, p_{\xi j}^{-1}, \xi),$$

exactly as required. □

A similar argument establishes

Lemma 3.5.4 *If the elements (i,a,λ) and (j,b,μ) of S are such that $(i,a,\lambda)\ \rho\ (j,b,\mu)$, then $(\lambda,\mu)\in\rho_\Lambda$.* □

The next result shows that the map Γ from the set of proper congruences on S into the set of linked triples is order-preserving:

Lemma 3.5.5 *If ρ and σ are proper congruences on the completely 0-simple semigroup $S=\mathcal{M}^0[G;I,\Lambda;P]$, then $\rho\subseteq\sigma$ if and only if*

$$\rho_I\subseteq\sigma_I,\ \rho_\Lambda\subseteq\sigma_\Lambda \text{ and } N_\rho\subseteq N_\sigma.$$

Proof It is clear from the definitions of ρ_I, ρ_Λ and N_ρ that if $\rho\subseteq\sigma$ then $\rho_I\subseteq\sigma_I$, $\rho_\Lambda\subseteq\sigma_\Lambda$ and $N_\rho\subseteq N_\sigma$. Conversely, suppose that ρ and σ are congruences on S such that $\rho_I\subseteq\sigma_I$, $\rho_\Lambda\subseteq\sigma_\Lambda$ and $N_\rho\subseteq N_\sigma$, and suppose that $(i,a,\lambda)\ \rho\ (j,b,\mu)$. By the last two lemmas we immediately deduce that

$$(i,j)\in\rho_I\subseteq\sigma_I \text{ and } (\lambda,\mu)\in\rho_\Lambda\subseteq\sigma_\Lambda. \tag{3.5.10}$$

Also, if we choose x in I and ξ in Λ such that $p_{\xi i}$ and $p_{\lambda x}$ (and hence also $p_{\xi j}$ and $p_{\mu x}$) are non-zero, and if we premultiply by $(1,e,\xi)$ and postmultiply by $(x,e,1)$, we obtain

$$(1,p_{\xi i}ap_{\lambda x},1)\ \rho\ (1,p_{\xi j}bp_{\mu x},1).$$

Hence $p_{\xi i}ap_{\lambda x}(p_{\xi j}bp_{\mu x})^{-1}\in N_\rho\subseteq N_\sigma$, and so

$$(1,p_{\xi i}ap_{\lambda x},1)\ \sigma\ (1,p_{\xi j}bp_{\mu x},1).$$

Now, since $(i,j)\in\sigma_I$ we have that $(i,p_{\xi i}^{-1},\xi)\ \sigma\ (j,p_{\xi j}^{-1},\xi)$; hence, postmultiplying by $(i,p_{\xi i}^{-1}p_{11}^{-1},1)$, we obtain that

$$(i,p_{\xi i}^{-1}p_{11}^{-1},1)\ \sigma\ (j,p_{\xi j}^{-1}p_{11}^{-1},1).$$

By a similar argument we have

$$(1,p_{11}^{-1}p_{\lambda x}^{-1},\lambda)\ \sigma\ (1,p_{11}^{-1}p_{\mu x}^{-1},\mu),$$

and since σ is a congruence it now follows that the elements

$$(i,p_{\xi i}^{-1}p_{11}^{-1},1)(1,p_{\xi i}ap_{\lambda x},1)(1,p_{11}^{-1}p_{\lambda x}^{-1},\lambda)$$

and

$$(j,p_{\xi j}^{-1}p_{11}^{-1},1)(1,p_{\xi j}bp_{\mu x},1)(1,p_{11}^{-1}p_{\mu x}^{-1},\mu)$$

are σ-equivalent. That is, $(i,a,\lambda)\ \sigma\ (j,b,\mu)$. □

As a corollary to this result we have that if $\rho_I=\sigma_I$, $\rho_\Lambda=\sigma_\Lambda$ and $N_\rho=N_\sigma$ then $\rho=\sigma$; thus the map Γ is one-one. That Γ is onto follows from the next result:

Lemma 3.5.6 *Let $(N,\mathcal{S},\mathcal{T})$ be a linked triple and let the relation ρ on $S\setminus\{0\}$ be defined by the rule that $(i,a,\lambda)\ \rho\ (j,b,\mu)$ if and only if:*

(1) $(i,j)\in\mathcal{S}$ and $(\lambda,\mu)\in\mathcal{T}$;

(2) $p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} \in N$ *for some* x *in* I *and* ξ *in* Λ *such that* $p_{\xi i}$, $p_{\xi j}$, $p_{\lambda x}$ *and* $p_{\mu x}$ *are all non-zero.*

Then $\rho \cup \{(0,0)\}$ *is a proper congruence on* S *such that* $\rho_I = \mathcal{S}$, $\rho_\Lambda = \mathcal{T}$ *and* $N_\rho = N$.

Proof It is clear that ρ is reflexive. It is also symmetric, since $\mathcal{S}$ and $\mathcal{T}$ are symmetric and since $p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} \in N$ implies that

$$p_{\xi j} b p_{\mu x} p_{\lambda x}^{-1} a^{-1} p_{\xi i}^{-1} = (p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1})^{-1} \in N.$$

The verification of transitivity presents us with some difficulties, and it pays first to develop a slightly different characterization of ρ in which the words 'for some' are replaced by 'for every':

Lemma 3.5.7 *If* ρ *is as defined in the statement of Lemma* 3.5.6, *then* $(i, a, \lambda)\ \rho\ (j, b, \mu)$ *if and only if*

(1) $(i, j) \in \mathcal{S}$ *and* $(\lambda, \mu) \in \mathcal{T}$;

(2) $p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} \in N$ *for EVERY* x *in* I *and* ξ *in* Λ *such that* $p_{\xi i}$, $p_{\xi j}$, $p_{\lambda x}$ *and* $p_{\mu x}$ *are all non-zero.*

Proof It is of course obvious that the new relation defined in this way is contained in ρ. To show the reverse conclusion we must show that if $p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} \in N$ and if y in I and η in Λ are such that $p_{\eta i}$, $p_{\eta j}$, $p_{\lambda y}$ and $p_{\mu y}$ are all non-zero, then it is also the case that

$$p_{\eta i} a p_{\lambda y} p_{\mu y}^{-1} b^{-1} p_{\eta j}^{-1} \in N.$$

Now, a normal subgroup of G has the 'symmetric' property that for all c, d in G

$$cd \in N \Rightarrow dc \in N$$

(for clearly $dc = c^{-1}(cd)c$). Hence from the hypothesis that

$$p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} \in N$$

we can deduce that

$$p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} p_{\xi i} a \in N.$$

Now $(\lambda, \mu) \in \mathcal{T}$ and the triple $(N, \mathcal{S}, \mathcal{T})$ is linked. Hence

$$q_{\lambda \mu y x} = p_{\lambda y} p_{\mu y}^{-1} p_{\mu x} p_{\lambda x}^{-1} \in N.$$

By forming the product of these two elements of N we deduce that

$$p_{\lambda y} p_{\mu y}^{-1} b^{-1} p_{\xi j}^{-1} p_{\xi i} a \in N.$$

Now, again by the symmetry property of N, we deduce that

$$a p_{\lambda y} p_{\mu y}^{-1} b^{-1} p_{\xi j}^{-1} p_{\xi i} \in N.$$

Also, since $(j,i) \in \mathcal{S}$ and since $(N,\mathcal{S},\mathcal{T})$ is a linked triple, we have that the element $p_{\xi i}^{-1} q_{\xi\eta j i} p_{\xi i} \in N$, that is,

$$p_{\xi i}^{-1} p_{\xi j} p_{\eta j}^{-1} p_{\eta i} \in N.$$

Multiplying these two elements of N together then gives

$$a p_{\lambda y} p_{\mu y}^{-1} b^{-1} p_{\eta j}^{-1} p_{\eta i} \in N,$$

from which it follows by the symmetry property that

$$p_{\eta i} a p_{\lambda y} p_{\mu y}^{-1} b^{-1} p_{\eta j}^{-1} \in N,$$

exactly as required. □

Returning now to the proof of Lemma 3.5.6, let us suppose that

$$(i,a,\lambda)\ \rho\ (j,b,\mu) \ \text{ and } \ (j,b,\mu)\ \rho\ (k,c,\nu).$$

Then certainly $(i,k) \in \mathcal{S}$ and $(\lambda,\nu) \in \mathcal{T}$. Also, if x in I and ξ in Λ are such that $p_{\xi i}$, $p_{\xi k}$, $p_{\lambda x}$ and $p_{\nu x}$ are all non-zero, then because $\mathcal{S} \subseteq \mathcal{E}_I$ and $\mathcal{T} \subseteq \mathcal{E}_\Lambda$ it follows that $p_{\xi j}$ and $p_{\mu x}$ are non-zero also. By the lemma just proved we have

$$p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} \in N \ \text{ and } \ p_{\xi j} b p_{\mu x} p_{\nu x}^{-1} c^{-1} p_{\xi k}^{-1} \in N.$$

Multiplying these gives

$$p_{\xi i} a p_{\lambda x} p_{\nu x}^{-1} c^{-1} p_{\xi k}^{-1} \in N,$$

which is exactly what we require in order to show that $(i,a,\lambda)\ \rho\ (k,c,\nu)$.

We have shown that ρ is an equivalence relation. To show that it is a congruence, let $(i,a,\lambda), (j,b,\mu), (k,c,\nu) \in S$ and suppose that $(i,a,\lambda)\ \rho\ (j,b,\mu)$. Then $(i,j) \in \mathcal{S} \subseteq \mathcal{E}_I$ and so either $p_{\nu i} = p_{\nu j} = 0$ (in which case both products $(k,c,\nu)(i,a,\lambda)$ and $(k,c,\nu)(j,b,\mu)$ are zero and so certainly ρ-equivalent) or $p_{\nu i}$ and $p_{\nu j}$ are both non-zero, in which case we have to prove that

$$(k, c p_{\nu i} a, \lambda)\ \rho\ (k, c p_{\nu j} b, \mu).$$

This amounts to proving that, for some x in I and ξ in Λ such that $p_{\xi k}$, $p_{\lambda x}$ and $p_{\mu x}$ are all non-zero, N contains

$$p_{\xi k} c p_{\nu i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\nu j}^{-1} c^{-1} p_{\xi k}^{-1}. \tag{3.5.11}$$

Now $(i,a,\lambda)\ \rho\ (j,b,\mu)$, and so by Lemma 3.5.7 we have that

$$p_{\xi i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\xi j}^{-1} \in N$$

for *every* x in I and ξ in Λ such that $p_{\xi i}$, $p_{\xi j}$, $p_{\lambda x}$ and $p_{\mu x}$ are all non-zero. Since we are presently assuming that $p_{\nu i}$ and $p_{\nu j}$ are non-zero, we may substitute ν for ξ in at this point and deduce that

$$p_{\nu i} a p_{\lambda x} p_{\mu x}^{-1} b^{-1} p_{\nu j}^{-1} \in N$$

whenever $p_{\lambda x}$ and $p_{\mu x}$ are non-zero. The desired formula (3.5.11) follows from this by conjugation. Thus ρ is a left congruence.

To show that ρ is a right congruence, once again let

$$(i, a, \lambda), (j, b, \mu), (k, c, \nu) \in S,$$

and suppose that $(i, a, \lambda)\ \rho\ (j, b, \mu)$. If $p_{\lambda k} = p_{\mu k} = 0$, then

$$(i, a, \lambda)(k, c, \nu) = (j, b, \mu)(k, c, \nu) = 0,$$

and so the products are certainly ρ-equivalent. Since $(\lambda, \mu) \in \mathcal{T} \subseteq \mathcal{E}_\Lambda$ the only other possibility is that $p_{\lambda k}$ and $p_{\mu k}$ are both non-zero, and in this case we begin by noting that

$$p_{\xi i} a p_{\lambda k} p_{\mu k}^{-1} b^{-1} p_{\xi j}^{-1} \in N$$

for every ξ in Λ such that $p_{\xi i}$ and $p_{\xi j}$ are both non-zero. (Here we have chosen $x = k$, as we may.) Hence, if x in I is such that $p_{\nu x} \neq 0$, then

$$p_{\xi i} a p_{\lambda k} c p_{\nu x} p_{\nu x}^{-1} c^{-1} p_{\mu k}^{-1} b^{-1} p_{\xi j}^{-1} \in N,$$

from which it follows that $(i, a p_{\lambda k} c, \nu)\ \rho\ (j, b p_{\mu k} c, \nu)$ as required. Thus $\rho \cup \{(0,0)\}$ is a proper congruence on S.

It remains to show that $\rho_I = \mathcal{S}$, $\rho_\Lambda = \mathcal{T}$ and $N_\rho = N$. First, if $(i, j) \in \rho_I$ then $(i, p_{\lambda i}^{-1}, \lambda)\ \rho\ (j, p_{\lambda j}^{-1}, \lambda)$ for every λ such that $p_{\lambda i} \neq 0$. Hence by the definition of ρ we have $(i, j) \in \mathcal{S}$. Conversely, if $(i, j) \in \mathcal{S}$ and $p_{\lambda i} \neq 0$ then, by the linked property of the triple, N contains the element

$$q_{\xi \lambda i j} = p_{\xi i} p_{\lambda i}^{-1} p_{\lambda x} p_{\lambda x}^{-1} (p_{\lambda j}^{-1})^{-1} p_{\xi j}^{-1}$$

for every x in I and ξ in Λ such that $p_{\xi i}$, $p_{\xi j}$ and $p_{\lambda x}$ are all non-zero. It follows that $(i, p_{\lambda i}^{-1}, \lambda)\ \rho\ (j, p_{\lambda j}^{-1}, \lambda)$, and so $(i, j) \in \rho_I$. Thus $\rho_I = \mathcal{S}$, and a similar argument shows that $\rho_\Lambda = \mathcal{T}$.

Finally, recall that $a \in N_\rho$ if and only if $(1, a, 1)\ \rho\ (1, e, 1)$, that is, if and only if for every x in I and ξ in Λ such that $p_{\xi 1}$ and p_{1x} are both non-zero,

$$p_{\xi 1} a p_{1x} p_{1x}^{-1} e p_{\xi 1}^{-1} \in N,$$

that is, if and only if $a \in N$. Thus $N_\rho = N$. □

We summarize the results of this section in a theorem as follows:

Theorem 3.5.8 *Let* $S = \mathcal{M}^0[G; I, \Lambda; P]$ *be a completely* 0*-simple semigroup. Then the mapping* $\Gamma : \rho \mapsto (N_\rho, \rho_I, \rho_\Lambda)$ *is an order-preserving bijection from the set of proper congruences on* S *onto the set of linked triples.* □

If we study completely simple semigroups (without zero) many of the complications in the foregoing account disappear. In particular, we do not need to restrict to proper congruences, and the equivalences $\mathcal{E}_I$ and $\mathcal{E}_\Lambda$ are unnecessary. A *linked* triple is then simply a triple $(N, \mathcal{S}, \mathcal{T})$ consisting of a normal subgroup N of G and equivalence relations $\mathcal{S}$, $\mathcal{T}$ on I, Λ, respectively, with the property that if $(i,j) \in \mathcal{S}$ or $(\lambda,\mu) \in \mathcal{T}$ then $q_{\lambda\mu ij} \in N$. We then have

Theorem 3.5.9 *Let $S = \mathcal{M}[G; I, \Lambda; P]$ be a completely simple semigroup. Then the mapping $\Gamma : \rho \mapsto (N_\rho, \rho_I, \rho_\Lambda)$ is an order-preserving bijection from the set of proper congruences on S onto the set of linked triples.* □

3.6 THE LATTICE OF CONGRUENCES ON A COMPLETELY 0-SIMPLE SEMIGROUP*

In this section we use the correspondence proved above between the proper congruences on a completely 0-simple semigroup and linked triples to derive information on the nature of the lattice of congruences. If $(N, \mathcal{S}, \mathcal{T})$ is a linked triple of $S = \mathcal{M}^0[G; I, \Lambda; P]$ we shall write the congruence $(N, \mathcal{S}, \mathcal{T})\Gamma^{-1}$ corresponding to the triple as $[N, \mathcal{S}, \mathcal{T}]$ (with square brackets). Thus $\rho = [N_\rho, \rho_I, \rho_\Lambda]$.

Lemma 3.6.1 *If ρ and σ are proper congruences on a completely 0-simple semigroup $S = \mathcal{M}^0[G; I, \Lambda; P]$, then*

$$\rho \cap \sigma = [N_\rho \cap N_\sigma, \rho_I \cap \sigma_I, \rho_\Lambda \cap \sigma_\Lambda],$$
$$\rho \vee \sigma = [N_\rho N_\sigma, \rho_I \vee \sigma_I, \rho_\Lambda \vee \sigma_\Lambda].$$

Proof To prove the first of these statements, notice first that $(N_\rho \cap N_\sigma, \rho_I \cap \sigma_I, \rho_\Lambda \cap \sigma_\Lambda)$ is a linked triple. For it is clear that $\rho_I \cap \sigma_I \subseteq \mathcal{E}_I$ and $\rho_\Lambda \cap \sigma_\Lambda \subseteq \mathcal{E}_\Lambda$, and if i, j, λ and μ are such that $p_{\lambda i}$, $p_{\lambda j}$, $p_{\mu i}$, $p_{\mu j}$ are all non-zero, then

$$\begin{aligned}(i,j) \in \rho_I \cap \sigma_I \;&\Rightarrow\; (i,j) \in \rho_I \text{ and } (i,j) \in \sigma_I\\ &\Rightarrow\; q_{\lambda\mu ij} \in N_\rho \text{ and } q_{\lambda\mu ij} \in N_\sigma\\ &\Rightarrow\; q_{\lambda\mu ij} \in N_\rho \cap N_\sigma,\end{aligned}$$

and similarly

$$(\lambda,\mu) \in \rho_\Lambda \cap \sigma_\Lambda \;\Rightarrow\; q_{\lambda\mu ij} \in N_\rho \cap N_\sigma.$$

Thus there is a congruence $[N_\rho \cap N_\sigma, \rho_I \cap \sigma_I, \rho_\Lambda \cap \sigma_\Lambda]$ which, by Lemma 3.5.6, is contained in ρ and in σ and is the largest congruence with these properties.

The same approach works for $\rho \vee \sigma$, except that there is more difficulty in establishing that $(N_\rho N_\sigma, \rho_I \vee \sigma_I, \rho_\Lambda \vee \sigma_\Lambda)$ is a linked triple. Since $\rho_I, \sigma_I \subseteq \mathcal{E}_I$ we do have that $\rho_I \vee \sigma_I \subseteq \mathcal{E}_I$, and similarly $\rho_\Lambda \vee \sigma_\Lambda \subseteq \mathcal{E}_\Lambda$. Also, by elementary group-theoretic results—see, for example, M. Hall (1959)—$N_\rho N_\sigma$ is a normal subgroup of G and is the smallest normal subgroup containing both N_ρ and N_σ. If $p_{\lambda i}$, $p_{\lambda j}$, $p_{\mu i}$, $p_{\mu j}$ are all non-zero, and if $(i,j) \in \rho_I \vee \sigma_I$, then by Proposition 1.5.11 there exist $i_1, i_2, \dots, i_{2n-1}$ in I such that

$$(i, i_1) \in \rho_I,\ (i_1, i_2) \in \sigma_I, \dots, (i_{2n-2}, i_{2n-1}) \in \rho_I,\ (i_{2n-1}, j) \in \sigma_I\,.$$

Since ρ_I and σ_I are both contained in $\mathcal{E}_I$ we have that $p_{\lambda i_r}$ and $p_{\mu i_r}$ are non-zero for every r, and by the linked property of the triples $(N_\rho, \rho_I, \rho_\Lambda)$ and $(N_\sigma, \sigma_I, \sigma_\Lambda)$ we deduce that

$$q_{\lambda\mu i i_1} \in N_\rho,\ q_{\lambda\mu i_1 i_2} \in N_\sigma, \ldots, q_{\lambda\mu i_{2n-2} i_{2n-1}} \in N_\rho,\ q_{\lambda\mu i_{2n-1} j} \in N_\sigma.$$

Hence the product

$$q_{\lambda\mu i i_1} q_{\lambda\mu i_1 i_2} \cdots q_{\lambda\mu i_{2n-2} i_{2n-1}} q_{\lambda\mu i_{2n-1} j}$$

belongs to $N_\rho N_\sigma$, that is, $q_{\lambda\mu ij} \in N_\rho N_\sigma$, by virtue of equation (3.5.8).

If $(\lambda, \mu) \in \rho_\Lambda \vee \sigma_\Lambda$ then there exist $\lambda_1, \lambda_2, \ldots, \lambda_{2n-1}$ such that

$$(\lambda, \lambda_1) \in \rho_\Lambda,\ (\lambda_1, \lambda_2) \in \sigma_\Lambda, \ldots, (\lambda_{2n-2}, \lambda_{2n-1}) \in \rho_\Lambda,\ (\lambda_{2n-1}, \mu) \in \sigma_\Lambda\,.$$

Again, if i, j in I are such that $p_{\lambda i}$, $p_{\lambda j}$, $p_{\mu i}$, $p_{\mu j}$ are all non-zero, then $p_{\lambda_r i}$ and $p_{\lambda_r j}$ are non-zero for every r, and

$$q_{\lambda\lambda_1 ij} \in N_\rho,\ q_{\lambda_1\lambda_2 ij} \in N_\sigma, \ldots, q_{\lambda_{2n-2}\lambda_{2n-1} ij} \in N_\rho, q_{\lambda_{2n-1}\mu ij} \in N_\sigma.$$

Hence, since N_ρ and N_σ are normal subgroups,

$$p_{\lambda i}^{-1} q_{\lambda\lambda_1 ij} p_{\lambda i} \in N_\rho,\ p_{\lambda_1 i}^{-1} q_{\lambda_1\lambda_2 ij} p_{\lambda_1 i} \in N_\sigma, \ldots, p_{\lambda_{2n-1} i}^{-1} q_{\lambda_{2n-1}\mu ij} p_{\lambda_{2n-1} i} \in N_\sigma,$$

and so the product of these elements is in $N_\rho N_\sigma$. That is, by virtue of equation (3.5.9), $p_{\lambda i}^{-1} q_{\lambda\mu ij}^{-1} p_{\lambda i} \in N_\rho N_\sigma$, from which it follows immediately that $q_{\lambda\mu ij} \in N_\rho N_\sigma$, as required.

We have shown that $(N_\rho N_\sigma, \rho_I \vee \sigma_I, \rho_\Lambda \vee \sigma_\Lambda)$ is a linked triple, and it now follows from Lemma 3.5.6 that $[N_\rho N_\sigma, \rho_I \vee \sigma_I, \rho_\Lambda \vee \sigma_\Lambda] = \rho \vee \sigma$. □

We can now establish the following result:

Theorem 3.6.2 *The lattice of congruences on a completely* 0*-simple semigroup is semimodular.*

Proof Let $S = \mathcal{M}^0[G; I, \Lambda; P]$. Let $(\mathcal{N}, \cap, .)$ be the lattice of all normal subgroups of G, let $(\mathcal{P}, \cap, \vee)$ be the lattice consisting of all equivalences on I contained in $\mathcal{E}_I$, and let $(\mathcal{Q}, \cap, \vee)$ be the lattice of all equivalences on Λ contained in $\mathcal{E}_\Lambda$. Let $\mathcal{X} = \mathcal{N} \times \mathcal{P} \times \mathcal{Q}$ be the direct product of these three lattices. By Lemma 3.6.1 the subset $\mathcal{Y}$ of $\mathcal{X}$ consisting of all *linked* triples $(N, \mathcal{S}, \mathcal{T})$ in $\mathcal{X}$ is a sublattice of $\mathcal{X}$. The effect of Theorem 3.5.9 is to give us an isomorphism Γ between the lattice $(\mathcal{K}, \cap, \vee)$ of proper congruences on S and the lattice $\mathcal{Y}$. Now $\mathcal{X}$ is semimodular by Corollary 1.8.4 and Propositions 1.8.5, 1.8.6 and 1.8.7. It will follow that $\mathcal{Y}$ is semimodular if we establish the following result:

Lemma 3.6.3 *If $(N_1, \mathcal{S}_1, \mathcal{T}_1)$ and $(N_2, \mathcal{S}_2, \mathcal{T}_2)$ belong to $\mathcal{Y}$, then $(N_1, \mathcal{S}_1, \mathcal{T}_1)$ covers $(N_2, \mathcal{S}_2, \mathcal{T}_2)$ in $\mathcal{Y}$ if and only if $(N_1, \mathcal{S}_1, \mathcal{T}_1)$ covers $(N_2, \mathcal{S}_2, \mathcal{T}_2)$ in $\mathcal{X}$.*

Proof One way round this is obvious. The difficulty arises because it is conceivable that $(N_1, \mathcal{S}_1, \mathcal{T}_1)$ could cover $(N_2, \mathcal{S}_2, \mathcal{T}_2)$ in $\mathcal{Y}$, but that there might be an element $(N, \mathcal{S}, \mathcal{T})$ in $\mathcal{X} \setminus \mathcal{Y}$ strictly in between. To see that this cannot in fact happen, notice that if $N_1 \supset N_2$ then $(N_1, \mathcal{S}_2, \mathcal{T}_2)$ is a linked triple such that

$$(N_1, \mathcal{S}_1, \mathcal{T}_1) \geq (N_1, \mathcal{S}_2, \mathcal{T}_2) > (N_2, \mathcal{S}_2, \mathcal{T}_2);$$

hence, by the covering property, $\mathcal{S}_1 = \mathcal{S}_2$, $\mathcal{T}_1 = \mathcal{T}_2$.

Equally, if $\mathcal{S}_1 \supset \mathcal{S}_2$ then $(N_1, \mathcal{S}_2, \mathcal{T}_1)$ is a linked triple such that

$$(N_1, \mathcal{S}_1, \mathcal{T}_1) > (N_1, \mathcal{S}_2, \mathcal{T}_1) \geq (N_2, \mathcal{S}_2, \mathcal{T}_2),$$

and hence $N_1 = N_2$, $\mathcal{T}_1 = \mathcal{T}_2$.

Applying a similar argument to the case where $\mathcal{T}_1 \supset \mathcal{T}_2$ we draw the general conclusion that $(N_1, \mathcal{S}_1, \mathcal{T}_1)$ covers $(N_2, \mathcal{S}_2, \mathcal{T}_2)$ in $\mathcal{Y}$ if and only if either

(a) $N_1 \succ N_2$ in $\mathcal{N}$, $\mathcal{S}_1 = \mathcal{S}_2$ and $\mathcal{T}_1 = \mathcal{T}_2$,

or

(b) $N_1 = N_2$, $\mathcal{S}_1 \succ \mathcal{S}_2$ in $\mathcal{P}$ and $\mathcal{T}_1 = \mathcal{T}_2$,

or

(c) $N_1 = N_2$, $\mathcal{S}_1 = \mathcal{S}_2$ and $\mathcal{T}_1 \succ \mathcal{T}_2$ in $\mathcal{Q}$,

that is, if and only if $(N_1, \mathcal{S}_1, \mathcal{T}_1)$ covers $(N_2, \mathcal{S}_2, \mathcal{T}_2)$ in $\mathcal{X}$. □

We may now conclude that the lattice $\mathcal{K}$ of proper congruences on S is semimodular, since it is isomorphic to $\mathcal{Y}$. The lattice $\mathcal{C}(S)$ of *all* congruences on S is obtained from $\mathcal{K}$ by adjoining a single extra maximum element, the universal congruence $S \times S$. Since the maximum element can never figure either as a or as b in the hypothesis 'if $a \succ a \wedge b$ and $b \succ a \wedge b$' appearing in the definition of semimodularity, the adjunction of this element does not destroy the semimodularity property of the lattice. The proof of Theorem 3.6.2 is thus complete. □

3.7 FINITE CONGRUENCE-FREE SEMIGROUPS*

If a semigroup S has a proper ideal I then the Rees quotient S/I is a proper homomorphic image of S. Thus if S is to be *congruence-free*, that is to say, if S is to have no congruences other than 1_S and $S \times S$, then it must in the first place be simple or 0-simple. In the finite case this implies that S is completely simple or completely 0-simple. (This need not be so in the infinite case: see the Notes at the end of the chapter.)

If S has a zero we may thus identify it with $\mathcal{M}^0[G; I, \Lambda; P]$. Now the triple $(G, 1_I, 1_\Lambda)$ is certainly linked, and gives rise to a non-identical congruence unless $G = \{e\}$. Thus $G = \{e\}$ if S is congruence-free. Hence every $p_{\lambda i}$ is equal either to e or to 0 and every extract $q_{\lambda\mu ij}$ is equal to e. Thus $(\{e\}, \mathcal{E}_I, \mathcal{E}_\Lambda)$ is a linked triple, and determines a non-identical congruence unless $\mathcal{E}_I = 1_I$ and $\mathcal{E}_\Lambda = 1_\Lambda$.

Now in the case we are considering, where $G = \{e\}$, to say that $(i, j) \in \mathcal{E}_I$ is to say that $p_{\lambda i} = e$ if and only if $p_{\lambda j} = e$; so in effect $(i, j) \in \mathcal{E}$ if and only if columns i and j of P are identical. Thus to say that $\mathcal{E}_I = 1_I$ is to say that no two columns of P are identical. Similarly, to say that $\mathcal{E}_\Lambda = 1_\Lambda$ is to say that no two rows of P are identical.

We can summarize what we have found in the following theorem:

Theorem 3.7.1 *Let $I = \{1, 2, \ldots, m\}$ and $\Lambda = \{1, 2, \ldots, n\}$ be finite sets and let P be a regular $n \times m$ matrix of 1s and 0s such that no two rows are identical and no two columns are identical. Let $S = (I \times \Lambda) \cup \{0\}$ and suppose that a binary operation is defined on S by*

$$(i, \lambda)(j, \mu) = \begin{cases} (i, \mu) & \text{if } p_{\lambda j} = 1 \\ 0 & \text{if } p_{\lambda j} = 0, \end{cases}$$
$$(i, \lambda)0 = 0(i, \lambda) = 00 = 0.$$

Then S is a congruence-free semigroup of order $mn + 1$.

Conversely, every finite congruence-free semigroup with zero is isomorphic to one of this kind. □

This constitutes a fairly reasonable classification of finite congruence-free semigroups with zero. It does not appear to be possible to write down a usable formula for the number of congruence-free semigroups of a given order, but for small orders it is not too hard to compute. (See Exercise 3.14.)

In the case where S has no zero, a simpler, more striking statement can be made:

Theorem 3.7.2 *If S is a finite congruence-free semigroup without zero and if $|S| > 2$, then S is a simple group.*

Proof We may take $S = \mathcal{M}[G; I, \Lambda; P]$. If N is a proper normal subgroup of G then $(N, 1_I, 1_\Lambda)$ is a linked triple (since $q_{\lambda\lambda ij} = q_{\lambda\mu ii} = e$) and so $[N, 1_I, 1_\Lambda]$ is a congruence on S distinct from both 1_S $\big(= [\{e\}, 1_I, 1_\Lambda]\big)$ and $S \times S$ $\big(= [G, I \times I, \Lambda \times \Lambda]\big)$. Hence if S is congruence-free then either G is simple or $G = \{e\}$. If $G = \{e\}$ then $|I \times \Lambda| = |S| > 2$ and so either $|I| = |\Lambda| = 2$ or at least one of I, Λ (say I) has more than two elements. In the first case we find linked triples $\big(\{e\}, 1_I, \Lambda \times \Lambda\big)$ and $\big(\{e\}, I \times I, 1_\Lambda\big)$ giving rise to non-trivial congruences, while in the second case there exists an equivalence $\mathcal{S}$ on I such that $1_I \subset \mathcal{S} \subset I \times I$, and this gives rise to a non-trivial congruence $[\{e\}, \mathcal{S}, 1_\Lambda]$ on S. Hence if S is congruence-free then $G \neq \{e\}$.

Thus G is a simple group. If either of I, Λ contains more than one element then $[G, 1_I, 1_\Lambda]$ is a congruence on S distinct from both $S \times S$ and 1_S. Hence if S is congruence-free we must have $|I| = |\Lambda| = 1$ and so $S \simeq G$, a simple group. □

The classification of finite congruence-free semigroups without zero is thus reduced to the group-theoretic problem of the classification of finite simple groups. The solving of this group-theoretic problem is one of the major achievements in algebra in recent decades. See, for example, Gorenstein (1982).

Remark If in Theorem 3.7.2 we drop the proviso that $|S| > 2$ then we cannot altogether rule out the possibility that $G = \{e\}$. In fact we obtain a congruence-free semigroup of order 2 when $G = \{e\}$ if either $|I| = 2$ and $|\Lambda| = 1$ or $|I| = 1$ and $|\Lambda| = 2$. In the former case we obtain the left zero semigroup of order 2, and in the latter case the corresponding right zero semigroup.

3.8 EXERCISES

1. In the infinite monogenic semigroup $S = \langle a \rangle$, show that every proper ideal is of the form Sa^m, and that
$$S \supset Sa \supset Sa^2 \supset \cdots$$
Deduce that S has no minimal ideals.

2. If $S = \langle a \rangle = M(m, r)$, a monogenic semigroup with index m and period r, show that
$$K(S) = \{a^m, a^{m+1}, \ldots, a^{m+r-1}\}.$$

3. Show that if I and J are ideals of a semigroup S then $I \cap J$ and $I \cup J$ are ideals. Show also that $(I \cup J)/J \simeq I/(I \cap J)$.

   Let A, B be ideals of a semigroup S such that $A \subset B$ and such that there does not exist any ideal C of S for which $A \subset C \subset B$. Let $b \in B \setminus A$.
   (a) Show that $A \cup J(b) = B$.
   (b) Show that $I(b) = A \cap J(b)$.
   (c) Show that $B/A \simeq J(b)/I(b)$.

4. Define
$$S_1 \supset S_2 \supset \cdots \supset S_m$$
to be a *principal series* of a semigroup S if
   (a) each S_i is a (two-sided) ideal of S;
   (b) there is no ideal of S strictly between S_i and S_{i+1} ($i = 1, 2, \ldots, m-1$);
   (c) $S_1 = S$, $S_m = K(S)$.

   Show that the factors S_1/S_2, S_2/S_3, $\ldots S_{m-1}/S_m$, S_m are isomorphic, in some order, to the principal factors of the semigroup. Deduce that any two principal series have isomorphic factors.

5. A semigroup S is called *semisimple* if none of its principal factors is null. Show that S is semisimple if and only if $A^2 = A$ for every ideal A of S.

6. Use the Rees Theorem to show that if S is completely simple and $\mathcal{H} = 1_S$ then S is isomorphic to a rectangular band $I \times \Lambda$.

   Now let S be completely 0-simple and suppose that $\mathcal{H} = 1_S$. Show that S is isomorphic to a semigroup $(I \times \Lambda) \cup \{0\}$ whose multiplication is given in terms of a regular $\Lambda \times I$ matrix $P = (p_{\lambda i})$ with entries in $\{1, 0\}$ as follows:

$$(i,\lambda)(j,\mu) = \begin{cases} (i,\mu) & \text{if } p_{\lambda j} = 1 \\ 0 & \text{if } p_{\lambda j} = 0 \end{cases}$$
$$(i,\lambda)0 = 0(i,\lambda) = 00 = 0.$$

   Such a semigroup is sometimes called a *rectangular* 0-*band.*

7. Let $S = \mathcal{M}[G; I, \Lambda; P]$ be a completely simple semigroup. Show that the idempotents of S form a subsemigroup of S if and only if

$$(\forall i, j \in I)(\forall \lambda, \mu \in \Lambda)\ q_{\lambda\mu ij} = e.$$

8. Let $S = \mathcal{M}[G; I, \Lambda; P]$ be a completely simple semigroup in which the idempotents form a subsemigroup. Choose a fixed i_0 in I and a fixed λ_0 in Λ and for simplicity of notation write $i_0 = \lambda_0 = 1$. Show that if E is the rectangular band $I \times \Lambda$ then the mapping $\phi : G \times E \to S$ given by

$$(g, (i,\lambda))\phi = (i, p_{1i}^{-1} g p_{11} p_{\lambda 1}^{-1}, \lambda)$$

   is an isomorphism.

9. Let T_n be the full transformation semigroup $\mathcal{T}_{[n]}$, where $[n]$ is the finite set $\{1, 2, \ldots, n\}$.
   (a) Show that T_n has n $\mathcal{J}$-classes $J_1, J_2, \ldots, J_n$, where

$$J_r = \{\alpha \in T_n : |\operatorname{im}\alpha| = r\} \quad (r = 1, 2, \ldots, n),$$

   and that

$$J_1 < J_2 < \cdots < J_n.$$

   (b) Let $K_r = \{\alpha \in T_n : |\operatorname{im}\alpha| \le r\}$ $(r = 1, 2, \ldots, n)$, so that $J_r = K_r \backslash K_{r-1}$. Show that the principal factor K_r/K_{r-1} is completely 0-simple $(r = 2, 3, \ldots, n)$. Show also that K_n/K_{n-1} is a 0-group, and that the kernel J_1 is a right zero semigroup of order n.
   (c) Show that the $\mathcal{J}$-class J_r contains $\binom{n}{r}$ $\mathcal{L}$-classes and $S(n,r)$ $\mathcal{R}$-classes (where $S(n,r)$ is the Stirling number of the second kind — see Exercise 1(15)) . Show also that each $\mathcal{H}$-class in J_r contains $r!$ elements.

(d) Let H be an $\mathcal{H}$-class within the completely 0-simple principal factor K_r/K_{r+1}, and suppose that H is the intersection of the $\mathcal{L}$-class consisting of all elements with image $B = \{b_1, b_2, \dots, b_r\}$ and the $\mathcal{R}$-class consisting of all elements whose kernel is the equivalence on $[n]$ with classes $A_1, A_2, \dots, A_r$. Show that every element α of H is then described completely by a permutation σ of $[r]$, where each element of A_i maps to $b_{i\sigma}$ $(i = 1, 2, \dots, r)$. Show that H is a group $\mathcal{H}$-class if and only if the elements $b_{i\sigma}$ $(i = 1, 2, \dots, r)$ all belong to different classes A_j. Deduce that if H is a group $\mathcal{H}$-class then H is isomorphic to the symmetric group S_r on r symbols.

10. (See Exercise 9 above, and also Exercise 2(15).) If $X = \{1, 2, 3, 4\}$ then the semigroup $\mathcal{T}_X$ has four $\mathcal{J}$-classes $J^{(1)}$, $J^{(2)}$, $J^{(3)}$ and $J^{(4)}$, where $J^{(i)} = \{\alpha \in \mathcal{T}_X : |\operatorname{im}\alpha| = i\}$. Show that

$$|J^{(1)}| = 4, \quad |J^{(2)}| = 84, \quad |J^{(3)}| = 144, \quad |J^{(4)}| = 24.$$

The $\mathcal{J}$-class $J^{(2)}$ can be enumerated in eggbox fashion as follows:

1222 2111	1333 3111	1444 4111	2333 3222	2444 4222	3444 4333
2122 1211	3133 1311	4144 1411	3233 2322	4244 2422	4344 3433
2212 1121	3313 11331	4414 1141	3323 2232	4424 2242	4434 3343
2221 1112	3331 1113	4441 1114	3332 2223	4442 2224	4443 3334
1122 2211	1133 3311	1144 4411	2233 3322	2244 4422	3344 4433
1212 2121	1313 3131	1414 4141	2323 3232	2424 4242	3434 4343
1221 2112	1331 3113	1441 4114	2332 3223	2442 4224	3443 4334

(We are here using the convention that $a_1a_2a_3a_4$ denotes the map α for which $i\alpha = a_i$ $\quad (i = 1, 2, 3, 4)$.)

The principal factor $J^{(2)}/J^{(1)}$ is a 0-simple semigroup and so, being finite, is completely 0-simple. To express it as a Rees matrix semigroup, note that the $\mathcal{H}$-class in the top left-hand corner is a group $\mathcal{H}$-class. Treat it as H_{11} and denote its elements by

$$e = \begin{pmatrix} 1 & 234 \\ 1 & 2 \end{pmatrix}, \qquad a = \begin{pmatrix} 1 & 234 \\ 2 & 1 \end{pmatrix}.$$

Call the $\mathcal{R}$-classes $R_1, \dots, R_7$ (reading from top to bottom) and the $\mathcal{L}$-classes $L_1, \dots, L_6$ (reading from left to right). For $\lambda = 1, \dots, 6$, choose q_λ as the *first-named* element in the $\mathcal{H}$-class $H_{1\lambda}$, and for $i = 1, \dots, 7$

let r_i be the first-named element of H_{i1}. Then the matrix $P = (p_{\lambda i}) = (q_\lambda r_i)$ is

$$\begin{bmatrix} e & a & 0 & 0 & 0 & e & e \\ e & 0 & a & 0 & e & 0 & e \\ e & 0 & 0 & a & e & e & 0 \\ 0 & e & a & 0 & e & a & 0 \\ 0 & e & 0 & a & e & 0 & a \\ 0 & 0 & e & a & 0 & e & a \end{bmatrix}$$

and $J^{(2)}/J^{(1)} \simeq \mathcal{M}^0[H_{11}; \{1, \ldots, 7\}, \{1, \ldots, 6\}; P]$.

The principal factors corresponding to $J^{(1)}$ and $J^{(4)}$ are easily described, since the former is a right zero semigroup and the latter is a 0-group. The principal factor $J^{(3)}/(J^{(1)} \cup J^{(2)})$ is a completely 0-simple semigroup with 6 $\mathcal{R}$-classes, 4 $\mathcal{L}$-classes, and in which each $\mathcal{H}$-class is of order 6. Find G, I, Λ and P such that $J^{(3)}/(J^{(1)} \cup J^{(2)}) \simeq \mathcal{M}^0[G; I, \Lambda; P]$.

11. Use the Rees Theorem to show that every completely simple right simple semigroup is a right group. (See Exercise 2(5).)

12. Let $a_1, a_2, \ldots, a_n \in S$, where S is a completely 0-simple semigroup. Show that if a_1a_2, a_2a_3, $\ldots$, $a_{n-1}a_n$ are all non-zero, then $a_1a_2 \ldots a_n \neq 0$. Show also that $a_1a_2 \ldots a_n \in R_{a_1} \cap L_{a_n}$.

13. Let ρ be a proper congruence on a completely 0-simple semigroup S (that is to say, a congruence for which $0\rho = \{0\}$).
    (a) Use Proposition 3.1.1 to show that S/ρ is 0-simple.
    (b) Show that if e is a primitive idempotent in S then $e\rho$ is a primitive idempotent in S/ρ. [Hint: if $h \in S$ is such that $h\rho$ is an idempotent in S/ρ and $h\rho \leq e\rho$, show that $(ehe)\rho = h\rho$. Then use Lallement's Lemma (Lemma 2.4.3) to find an idempotent g in S such that $g\rho = (ehe)\rho$, $L_g \leq L_{ehe}$, $R_g \leq R_{ehe}$.]
    (c) Deduce that S/ρ is completely 0-simple.

14. (a) If $K(n)$ denotes the number of non-isomorphic congruence-free semigroups with 0 having order n, show that $K(5) = 2$, $K(7) = 2$.
    (b) Show that the two congruence-free semigroups with 0 having order 7 are anti-isomorphic.
    (c) Show that $K(p + 1) = 0$ for every prime p.

15. An alternative, direct approach to the proof of Theorem 3.7.2 is given below. Let S be a finite, congruence-free semigroup without zero, and suppose that $|S| > 2$.
    (a) Let $m = \min\{|xS| : x \in S\}$. Show that X, defined as $\{x \in S : |xS| = m\}$, is a two-sided ideal of S, and deduce that $X = S$.

(b) Show that the relation $\mathcal{R}$ is a congruence on S, and deduce that $\mathcal{R} = S \times S$.

(c) From this, and from the dual result concerning $\mathcal{L}$, deduce that S is a group.

16. Let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely 0-simple semigroup. Let $\Gamma(S)$ be the graph whose set of vertices is $\{(i, \lambda) \in I \times \Lambda : H_{i\lambda} \text{ is a group}\}$ and where there is an edge between (i, λ) and (j, μ) if and only if $i = j$ or $\lambda = \mu$. If $H_{i\lambda}$ is a group, denote its identity by $e_{i\lambda}$. Let $\langle E \rangle$ be the subsemigroup of S generated by the idempotents.

    (a) Show that if $e_{i\lambda}e_{j\mu} \neq 0$ then there is a path
    $$(i, \lambda) \to (j, \lambda) \to (j, \mu)$$
    in $\Gamma(S)$.

    (b) Show that the graph $\Gamma(S)$ is connected if and only if
    $$(\forall i \in I)(\forall \lambda \in \Lambda)\langle E \rangle \cap H_{i\lambda} \neq \emptyset.$$

The Rees matrix construction $\mathcal{M}[S; I, \Lambda; P]$ works even if S is not a group. If S is a semigroup with zero, then we have to be a little more careful in defining $\mathcal{M}^0[S; I, \Lambda; P]$. We form the cartesian product $T = I \times S \times \Lambda$ and define an associative multiplication
$$(i, a, \lambda)(j, b, \mu) = (i, ap_{\lambda j}b, \mu).$$
The subset $Z = \{(i, 0, \lambda) : i \in I, \lambda \in \Lambda\}$ is an ideal of T, and we define $\mathcal{M}^0[S; I, \Lambda; P]$ to be the Rees quotient $(I \times S \times \Lambda)/Z$. Equivalently, we can regard $\mathcal{M}^0[S; I, \Lambda; P]$ as consisting of the elements of $I \times \big(S \setminus \{0\}\big) \times \Lambda$ together with 0, where
$$(i, a, \lambda)(j, b, \mu) = \begin{cases} (i, ap_{\lambda j}b, \mu) & \text{if } ap_{\lambda j}b \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$
The remaining exercises explore this idea.

17. Consider a Rees matrix semigroup $\mathcal{M}[S; I, \Lambda; P]$. Show that the element (i, a, λ) is regular if and only if
$$V(a) \cap p_{\lambda j}Sp_{\mu i} \neq \emptyset$$
for some j in I and μ in Λ.

18. Let $T = \mathcal{M}^0[S; I, \Lambda; P]$, where S is a monoid with zero. Say that the matrix $P = (p_{\lambda i})$ is 'semi-regular' if every row of P contains a right unit of S and every column of P contains a left unit of S. Show that T is a regular semigroup if and only if S is regular and P is semi-regular. Show also that T is 0-bisimple if and only if S is 0-bisimple and P is semi-regular.

19. Let B denote the bicyclic monoid
$$\langle p, q \mid pq = 1\rangle = \{q^m p^n : m, n \geq 0\},$$
and let $S = \mathcal{M}[B; \{1,2\}, \{1,2\}; P]$, where
$$P = \begin{pmatrix} 1 & q \\ 1 & 1 \end{pmatrix}.$$
Show that S is a bisimple monoid and that it is generated by the four idempotents $(1,1,1)$, $(2,1,1)$, $(1,1,2)$ and $(1,p,2)$.

20. Let S be a finite monoid, let $I = \{1,\ldots,r\}$, $\Lambda = \{1,\ldots s\}$ and let $P = (p_{ji})$ be an $s \times r$ matrix over S such that:
   (a) the first row and first column of P consist entirely of 1s;
   (b) the entries p_{ji} $(i, j > 1)$ include a set of generators for S.

   Let $T = \mathcal{M}[S; I, \Lambda; P]$. Show that:
   (a) the elements $(i,1,1)$, $(1,1,j)$ $(1 \leq i \leq r,\ 1 \leq j \leq s)$ are idempotents in T;
   (b) the idempotents $(i,1,1)$, $(1,1,j)$ generate T;
   (c) the map $s \mapsto (1,s,1)$ embeds S in T;
   (d) Deduce that every finite semigroup can be embedded in a finite semigroup generated by idempotents.

21. Let
$$S = \{1, a_1, a_2, \ldots, a_k\}$$
be a finite regular monoid, and let
$$T = \mathcal{M}[S; \{1,\ldots,k+1\}, \{1,\ldots,k+1\}; P],$$
where
$$P = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & a_1 & a_2 & \ldots & a_k \\ 1 & a_2 & a_3 & \ldots & a_1 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & a_k & a_1 & \ldots & a_{k-1} \end{pmatrix}$$
Denote the set of idempotents of T by E.
   (a) Show that the map $s \mapsto (1,s,1)$ embeds S in T.
   (b) Show that the elements $(i,1,1)$, $(1,1,j)$ are idempotent $(1 \leq i \leq k+1,\ 1 \leq j \leq k+1)$.
   (c) Show that every element of T is expressible as a product of two idempotents.
   (d) Show that T is regular.
   (e) Deduce that every finite semigroup can be embedded in a finite regular semigroup T such that $T = E^2$.

3.9 NOTES

The fundamental ideas in Section 3.1 go back to Suschkewitsch (1928) and Rees (1940), though the treatment of principal factors is much influenced by J. A. Green (1951).

The main result in Section 3.2 is due to Rees (1940), but is given in a version essentially due to Clifford and Preston (1961). The approach via 0-minimal ideals dates back to Clifford (1948). For an exhaustive study of minimal conditions see Hall and Munn (1979).

As a result of the Rees Theorem certain aspects of the study of completely 0-simple semigroups acquire an essentially combinatorial flavour. See, for example, Graham (1968), Houghton (1977, 1979), Houghton and Sullivan (1984), and Howie (1978).

Theorem 3.3.1 is in essence due to Suschkewitsch (1928), and is a special case of the result of Rees (1940). References for the remainder of Section 3.3 are Croisot (1953), Lallement and Petrich (1966), Steinfeld (1966), and Venkatesan (1966).

In Section 3.4, the isomorphism theorem and the normalization theorem for completely simple semigroups are due to Rees (1940). The normalization theorem for completely 0-simple semigroups is due to Tamura (1956).

In Sections 3.5 and 3.6 the material is drawn from a variety of sources. See Gluskin (1956, 1957), Tamura (1960), Preston (1961), Lallement (1967, 1974), and Kapp and Schneider (1969).

The main theorem in Section 3.6 is due to Tamura (1956). The problem of describing infinite congruence-free semigroups is much harder, and many diverse examples of such semigroups have been described. For example, see Trotter (1974), Munn (1972, 1974a,1975), T. E. Hall (1979), Howie (1981a,b), and Marques (1983).

Exercises 3 and 4 are from J. A. Green (1951), and Exercise 5 is from Munn (1955).

Meakin (1985a) has written a very useful survey article on the Rees matrix construction. See also Márki (1975), Tran Quy Tien (1975), and Meakin (1985b). Exercise 15 is due to P. M. Neumann (private communication). Exercise 16 is from Ruškuc (1994), and Exercise 17 is from McAlister (1985). Exercise 18, from (Byleen 1981), describes the 'fundamental four-spiral semigroup' Sp_4, characterized in a different way by Byleen *et al.* (1978). For Exercise 20 see Byleen (1981), Pastijn (1977), Howie (1966), and Giraldes and Howie (1991). Exercise 21 is due to T. E. Hall (private communication), and is quoted by Giraldes and Howie (1984).

4

Completely regular semigroups

It is well known that a group (G, μ) can alternatively be regarded as having three operations, namely the binary operation $\mu : (a, b) \mapsto ab$, the unary operation $a \mapsto a^{-1}$, and the 0-ary operation (the constant) 1. If we wish to emphasize this aspect, we write $G = (G, \mu, ^{-1}, 1)$. From this point of view, a morphism $\phi : G \to H$ between two groups is defined by the properties

$$(ab)\phi = (a\phi)(b\phi), \quad (a^{-1})\phi = (a\phi)^{-1}, \quad 1\phi = 1;$$

but it is an elementary exercise in group theory to show that the first of these properties implies the other two, and it is for this reason that one does not have to be completely precise about one's point of view.

A semigroup (S, μ) will be called a *U-semigroup* if a unary operation $a \mapsto a'$ is defined on S, with the property that

$$(a')' = a$$

for every a in S. We write $S = (S, \mu, ')$.

It is of course clear that every semigroup may be regarded as a U-semigroup: the most obvious approach is to define $a' = a$ for every a in S. To be interesting, the unary operation must interact in some way with the binary operation. Two versions of interaction have been studied. The first, in which a' is usually denoted by a^*, gives us a $*$-*semigroup*, or a *semigroup with involution*; here the properties of the unary operation are given by

$$(a^*)^* = a, \quad (ab)^* = b^*a^*.$$

The second, in which we shall write a' as a^{-1}, gives us what we shall call an *I-semigroup*; here the properties are

$$(a^{-1})^{-1} = a, \quad aa^{-1}a = a.$$

Notice that, since these equations are to hold for *every* element of S, it follows that

$$a^{-1}aa^{-1} = a^{-1}(a^{-1})^{-1}a^{-1} = a^{-1},$$

and so a^{-1} is an inverse of a.

Notice that a group $(G, ., ^{-1})$ is both a $*$-semigroup and an I-semigroup. The class of U-semigroups for which the unary operation satisfies the conditions both for a $*$-semigroup and for an I-semigroup is in fact the class of *inverse semigroups*, which will be the subject of Chapter 5. The subject of this chapter is another important class of I-semigroups, called *completely regular* semigroups.

4.1 THE CLIFFORD DECOMPOSITION

A semigroup S will be called *completely regular* if there exists a unary operation $a \mapsto a^{-1}$ on S with the properties

$$(a^{-1})^{-1} = a, \quad aa^{-1}a = a, \quad aa^{-1} = a^{-1}a. \tag{4.1.1}$$

More briefly, a completely regular semigroup is an I-semigroup S in which, for every a in S, $aa^{-1} = a^{-1}a$.

The following result gives us two alternative definitions:

Proposition 4.1.1 *Let S be a semigroup. Then the following statements are equivalent:*

(1) *S is completely regular;*

(2) *every element of S lies in a subgroup of S;*

(3) *every $\mathcal{H}$-class in S is a group.*

Proof (1) $\Rightarrow$ (2). Let $a \in S$, and let $aa^{-1} = a^{-1}a = e$. Then, by Theorem 2.3.4, $a \in R_e \cap L_e = H_e$, and H_e is a subgroup of S by Corollary 2.2.6.

(2) $\Rightarrow$ (3). Let $a \in S$; then $a \in G$ for some subgroup G of S. Denote the identity element of G by e, and the inverse of a within G by a^*. Then from

$$ea = ae = a \text{ and } aa^* = a^*a = e$$

it follows that $a \,\mathcal{H}\, e$, and hence $H_a = H_e$, a group.

(3) $\Rightarrow$ (1). For each a in S, define a^{-1} to be the unique inverse of a within the group H_a. (Notice that the element a may have several inverses in S, but only one of them lies in H_a.) Then it is clear that

$$(a^{-1})^{-1} = a, \quad aa^{-1}a = a, \quad aa^{-1} = a^{-1}a,$$

and so S is completely regular. □

We have already encountered one important class of completely regular semigroups, namely the class of completely simple semigroups. (Completely 0-simple semigroups are *not* in general completely regular, since not every $\mathcal{H}$-class in such a semigroup is a group.) The next result gives in effect two alternative definitions of a completely simple semigroup:

Proposition 4.1.2 *Let S be a semigroup. Then the following statements are equivalent:*

(1) *S is completely simple;*

(2) *S is completely regular, and, for all x, y in S,*

$$xx^{-1} = (xyx)(xyx)^{-1}.$$

(3) *S is completely regular and simple.*

Proof (1) $\Rightarrow$ (2). Let S be completely simple, and for each a in S, let a^{-1} be the unique inverse of a lying inside H_a. Let $x, y \in S$. Then by Lemma 3.2.7, applied to the case where 0 is indecomposable, we deduce that $xyx \mathcal{H} x$, and it then follows that $xx^{-1} = (xyx)(xyx)^{-1}$, as required.

(2) $\Rightarrow$ (3). Let $a, b \in S$. Then

$$a = aa^{-1}a = a.b.a(aba)^{-1}a,$$

and so $J_a \leq J_b$. By interchanging the roles of a and b we may equally well show that $J_b \leq J_a$. It follows that $\mathcal{J} = S \times S$, and so S is simple.

(3) $\Rightarrow$ (1). Suppose that S is completely regular and simple. We shall show that every idempotent of S is primitive, from which it will follow, by Theorem 3.3.3, that S is completely simple. Accordingly, let e, f be idempotents in S, and suppose that $f \leq e$, so that $ef = fe = f$. Then, since S is simple, there exist z, t in S such that $e = zft$. (See Corollary 3.1.2.) We now produce 'improved' versions of z and t by defining $x = ezf$ and $y = fte$; we still have

$$xfy = (ezf)f(fte) = e(zft)e = e^3 = e,$$

but now have the extra advantage that $ex = xf = x$ and $fy = ye = y$.

Now S is completely regular and so, by Proposition 4.1.1, the element x belongs to H_g for some idempotent g. Thus $gx = xg = x$, and there exists x^{-1} in H_g such that $xx^{-1} = x^{-1}x = g$. As a consequence, $gf = x^{-1}xf = x^{-1}x = g$. But we also have

$$gf = gef = gxfyf = xfyf = ef = f,$$

and so $g = f$. Hence

$$f = fe = ge = gxfy = xfy = e.$$

We have shown that $f \leq e$ implies $f = e$ for every pair of idempotents in S. Thus every idempotent in the non-empty set of idempotents of S is primitive, and so S is completely simple as required. □

Let S be a completely regular semigroup. Then $a \mathcal{H} a^2$ for every a in S, and so certainly

$$a \mathcal{J} a^2. \tag{4.1.2}$$

Hence, using equation (2.1.3), we see that for every a, b in S,

$$J_{ab} = J_{(ab)^2} = J_{a(ba)b} \leq J_{ba}.$$

Since by the same token we have $J_{ba} \leq J_{ab}$, we conclude that

$$J_{ab} = J_{ba}. \tag{4.1.3}$$

Next, if $a \mathcal{J} b$ then $b = xay$, $a = ubv$ for some x, y, u, v in S^1. If $c \in S$ then

$$\begin{aligned} J_{ca} &= J_{cubv} \leq J_{cub} \text{ by (2.1.3)} \\ &= J_{ubc} \text{ by (4.1.3)} \\ &\leq J_{bc} = J_{cb}. \end{aligned}$$

Similarly $J_{cb} \leq J_{ca}$ and so $ca \mathcal{J} cb$. By virtue of (4.1.3) we also have $ac \mathcal{J} bc$, and so $\mathcal{J}$ is a congruence. It follows from (4.1.2) and (4.1.3) that $S/\mathcal{J}$ is a semilattice.

Now consider a typical $\mathcal{J}$-class $J = J_a$ of S. It is in fact a subsemigroup of S, since the congruence property gives

$$(J_a)^2 \subseteq J_{a^2} = J_a.$$

If a, b are elements of J then there exist x, y, u, v in S such that $xay = b$, $ubv = a$. Now there exist idempotents e and f (in J) such that $a \in H_e$, $b \in H_f$. Hence

$$(fx)a(yf) = fbf = b, \quad (eu)b(vf) = eae = a.$$

It is clear that

$$J_{fx} \geq J_{(fx)a(yf)} = J_b = J,$$

and we also have

$$J_{fx} \leq J_f = J.$$

Hence $fx \in J$, and similarly yf, eu, ve are all in J. By Theorem 3.1.6, the subsemigroup J is simple, and since it is also completely regular it must be completely simple.

At this point it pays to change both the notation and our point of view. We denote the semilattice $S/\mathcal{J}$ by Y, and for each α in Y we denote $\alpha(\mathcal{J}^\natural)^{-1}$ by S_α. Each S_α is a $\mathcal{J}$-class of S and is a completely simple subsemigroup. Thus S is the disjoint union of the completely simple semigroups S_α $(\alpha \in Y)$, and the congruence property of $\mathcal{J}$ gives us that

$$S_\alpha S_\beta \subseteq S_{\alpha\beta}. \tag{4.1.4}$$

We say that S is a *semilattice of completely simple semigroups.* We have proved

Theorem 4.1.3 *Every completely regular semigroup is a semilattice of completely simple semigroups.* □

At a first glance, perhaps, this does not look like progress at all, since we already know that a completely regular semigroup is a disjoint union of groups, and we have now expressed it as a disjoint union of completely simple semigroups, more complicated objects than groups. The progress lies in the 'gross' multiplication formula (4.1.4). Previously we knew that

$$S = \bigcup_{e \in E} H_e,$$

where E is the set of idempotents in S and each H_e is a group, but we had no idea at all about the location of the product of an element x in H_e and an element y in H_f, or even whether the product of H_e and H_f was contained in a single $\mathcal{H}$-class. Now we know that the product of x in S_α and y in S_β lies in $S_{\alpha\beta}$. On the other hand, what we have is not entirely satisfactory, for if we know (in terms of the Rees Theorem, say) the location of x and y within the completely simple semigroups S_α and S_β, we do not know the location of xy within $S_{\alpha\beta}$. We know the 'gross' structure of S but not its 'fine' structure.

One possible 'fine structure' is as follows. Suppose that we have a semilattice Y and a set of completely simple semigroups S_α indexed by Y, and suppose that, for all $\alpha \geq \beta$ in Y there exists a morphism $\phi_{\alpha,\beta} : S_\alpha \to S_\beta$ such that:

(S1) $(\forall \alpha \in Y)\ \phi_{\alpha,\alpha} = 1_{S_\alpha}$;
(S2) for all α, β, γ in Y such that $\alpha \geq \beta \geq \gamma$,

$$\phi_{\alpha,\beta}\phi_{\beta,\gamma} = \phi_{\alpha,\gamma}.$$

Now define a multiplication in $S = \bigcup_{\alpha \in Y} S_\alpha$, in terms of the multiplications in the components S_α and the morphisms $\phi_{\alpha,\beta}$, by the rule that, for each x in S_α and y in S_β,

$$xy = (x\phi_{\alpha,\alpha\beta})(y\phi_{\beta,\alpha\beta}). \tag{4.1.5}$$

If $x \in S_\alpha$, $y \in S_\beta$ and $z \in S_\gamma$, then by the morphism properties and by the transitivity condition (S2),

$$\begin{aligned}(xy)z &= (x\phi_{\alpha,\alpha\beta})(y\phi_{\beta,\alpha\beta}).z \\ &= [(x\phi_{\alpha,\alpha\beta})(y\phi_{\beta,\alpha\beta})\phi_{\alpha\beta,\alpha\beta\gamma}](z\phi_{\gamma,\alpha\beta\gamma}) \\ &= (x\phi_{\alpha,\alpha\beta\gamma})(y\phi_{\beta,\alpha\beta\gamma})(z\phi_{\gamma,\alpha\beta\gamma}),\end{aligned}$$

and similarly

$$x(yz) = (x\phi_{\alpha,\alpha\beta\gamma})(y\phi_{\beta,\alpha\beta\gamma})(z\phi_{\gamma,\alpha\beta\gamma}).$$

Thus the multiplication (4.1.5) is associative. It follows that S is a particular type of completely regular semigroup, called a *strong semilattice of completely simple semigroups.* We write

$$S = \mathcal{S}[Y; S_\alpha; \phi_{\alpha,\beta}].$$

We shall see that not every completely regular semigroup is of this kind.

4.2 CLIFFORD SEMIGROUPS

A *Clifford semigroup* is defined as a completely regular semigroup $(S, \mu, ^{-1})$ in which, for all x, y in S

$$(xx^{-1})(yy^{-1}) = (yy^{-1})(xx^{-1}). \tag{4.2.1}$$

In an arbitrary semigroup S, let us say that an element c is *central* if $cs = sc$ for every s in S. The set of central elements forms a subsemigroup of S, called the *centre* of S.

The following theorem gives in effect several alternative definitions of a Clifford semigroup:

Theorem 4.2.1 *Let S be a semigroup with set E of idempotents. Then the following statements are equivalent:*

(1) *S is a Clifford semigroup;*
(2) *S is a semilattice of groups;*
(3) *S is a strong semilattice of groups;*
(4) *S is regular, and the idempotents of S are central;*
(5) *S is regular, and $\mathcal{D}^S \cap (E \times E) = 1_E$.*

Proof (1) $\Rightarrow$ (2). Let S be a Clifford semigroup. Then S is completely regular, and so is a semilattice Y of completely simple semigroups S_α. Now every idempotent e in S is expressible as xx^{-1} for some x—the obvious choice for x is e itself—and so the condition (4.2.1) in effect says that idempotents commute. This happens within each of the components S_α, and so each S_α, being a completely simple semigroup in which idempotents commute, is a group. Thus S is a semilattice of groups.

(2) $\Rightarrow$ (3). For each α in Y let e_α be the identity element of S_α $(\alpha \in Y)$. Suppose now that $\alpha \geq \beta$. Then for each a_α in S_α the product $e_\beta a_\alpha$ belongs to $S_{\alpha\beta} = S_\beta$, and so it makes sense to define a map $\phi_{\alpha,\beta} : S_\alpha \to S_\beta$ by the rule that $a_\alpha\phi_{\alpha,\beta} = e_\beta a_\alpha$. It is clear that $\phi_{\alpha,\alpha}$ is the identity map on S_α. Also $\phi_{\alpha,\beta}$ is a morphism. To see this, notice that for every a_α, b_α in S_α,

$$(a_\alpha\phi_{\alpha,\beta})(b_\alpha\phi_{\alpha,\beta}) = (e_\beta a_\alpha)(e_\beta b_\alpha) = \big((e_\beta a_\alpha)e_\beta\big)b_\alpha.$$

Now $e_\beta a_\alpha \in S_\beta$, and e_β is the identity of S_β. So

$$(a_\alpha\phi_{\alpha,\beta})(b_\alpha\phi_{\alpha,\beta}) = e_\beta a_\alpha b_\alpha = (a_\alpha b_\alpha)\phi_{\alpha,\beta},$$

as required.

Next, suppose that $\alpha \geq \beta \geq \gamma$, and notice, by a standard property of group morphisms, that, for all a_α in S_α,

$$\begin{aligned}(a_\alpha\phi_{\alpha,\beta})\phi_{\beta,\gamma} &= e_\gamma(e_\beta a_\alpha)\\ &= (e_\gamma e_\beta)a_\alpha = (e_\beta\phi_{\beta,\gamma})a_\alpha\\ &= e_\gamma a_\alpha = a_\alpha\phi_{\alpha,\gamma};\end{aligned}$$

thus $\phi_{\alpha,\beta}\phi_{\beta,\gamma} = \phi_{\alpha,\gamma}$, as required.

Finally, notice that, for arbitrary α and β in Y and for elements a_α in S_α and b_β in S_β, the product $a_\alpha b_\beta$ lies in S_γ, where $\gamma = \alpha\beta$. Hence

$$\begin{aligned} a_\alpha b_\beta &= e_\gamma(a_\alpha b_\beta) = (e_\gamma a_\alpha)b_\beta \\ &= \big((e_\gamma a_\alpha)e_\gamma\big)b_\beta \text{ (since } e_\gamma a_\alpha \in S_\gamma) \\ &= (e_\gamma a_\alpha)(e_\gamma b_\beta) = (a_\alpha\phi_{\alpha,\gamma})(b_\beta\phi_{\beta,\gamma}), \end{aligned}$$

and so S is indeed isomorphic to the strong semilattice of groups $\mathcal{S}[Y; S_\alpha; \phi_{\alpha,\beta}]$.

(3) $\Rightarrow$ (4). Certainly every strong semilattice of groups $\mathcal{S}[Y; G_\alpha; \phi_{\alpha,\beta}]$ is a regular semigroup. Its idempotents are the identity elements e_α of the groups G_α, and it is easy to calculate that, for all β in Y and all g_β in G_β,

$$\begin{aligned} e_\alpha g_\beta &= (e_\alpha\phi_{\alpha,\alpha\beta})(g_\beta\phi_{\beta,\alpha\beta}) = e_{\alpha\beta}(g_\beta\phi_{\beta,\alpha\beta}) = g_\beta\phi_{\beta,\alpha\beta}, \\ g_\beta e_\alpha &= (g_\beta\phi_{\beta,\alpha\beta})(e_\alpha\phi_{\alpha,\alpha\beta}) = (g_\beta\phi_{\beta,\alpha\beta})e_{\alpha\beta} = g_\beta\phi_{\beta,\alpha\beta}; \end{aligned}$$

thus idempotents are central.

(4) $\Rightarrow$ (5). Suppose that $e \mathcal{D}^S f$, where e and f are idempotents. Then, by Theorem 2.3.4, there exists an element a and an inverse a' of a such that $aa' = e$, $a'a = f$. Hence, using the centrality of the idempotents e and f, we have

$$\begin{aligned} e = e^2 &= a(a'a)a' = afa' = faa' = a'aaa' \\ &= a'ae = a'ea = a'aa'a = f^2 = f, \end{aligned}$$

and we deduce that $\mathcal{D}^S \cap (E \times E) = 1_E$.

(5) $\Rightarrow$ (1). Each $\mathcal{D}$-class contains a single idempotent, and so is a group. Thus $\mathcal{D} = \mathcal{H}$, and so each element a has exactly one inverse a^{-1}, with the properties

$$(a^{-1})^{-1} = a, \quad aa^{-1}a = a, \quad aa^{-1} = a^{-1}a.$$

Thus S is completely regular, and so is a semilattice Y of completely simple semigroups S_α. Now for all x, y in S_α we have $xy \in R_x \cap L_y$, and so $x \mathcal{D} y$. Thus each S_α is contained in a single $\mathcal{D}$-class, and so has a single idempotent. Hence each S_α is a group.

From (2) $\Rightarrow$ (3) we now deduce that S is a strong semilattice of groups $\mathcal{S}[Y; S_\alpha; \phi_{\alpha,\beta}]$, and it then follows easily that for an arbitrary x in S_α and y in S_β,

$$xx^{-1}yy^{-1} = e_\alpha e_\beta = e_{\alpha\beta} = e_\beta e_\alpha = yy^{-1}xx^{-1}.$$

Thus S is a Clifford semigroup. □

4.3 VARIETIES

Much of this section could be treated in a more general context, as part of the theory of Ω-algebras—see, for example, Cohn (1965) or Grätzer (1968). Here it will be sufficient to consider the case of $(2,1)$-algebras, systems on which a binary operation $(x, y) \mapsto xy$ and a unary operation $x \mapsto x'$ are defined. A *subalgebra* U of a $(2,1)$-algebra S is defined as a non-empty subset U of S such that, for all x, y in S,

$$x, y \in U \;\Rightarrow\; xy \in U, \quad x \in U \;\Rightarrow\; x' \in U.$$

If S and T are two such algebras, then a *morphism* from S into T is defined in the obvious way as a map $\phi : S \to T$ such that, for all x, y in S,

$$(xy)\phi = (x\phi)(y\phi), \quad (x')\phi = (x\phi)'.$$

If ϕ is onto, we say that T is a *morphic image* of S.

Let S_i $(i \in I)$ be a family of $(2,1)$-algebras, and let P be the cartesian product of the sets S_i. The elements of P are maps $\gamma : I \to \bigcup\{S_i : i \in I\}$ with the property that $i\gamma \in S_i$ for each i in I. If we now define, for γ, δ in P and i in I,

$$i(\gamma\delta) = (i\gamma)(i\delta), \quad i\gamma' = (i\gamma)',$$

we give P the structure of a $(2,1)$-algebra, the *direct product* of the algebras S_i. We write $P = \prod\{S_i : i \in I\}$.

Let X be a non-empty set, and consider the set $F_{2,1}(X)$ of all formal expressions in the alphabet $X \cup \{(,),'\}$ defined by the rules:

(1) $X \subseteq F_{2,1}(X)$;
(2) if $u \in F_{2,1}(X)$ then $(u)' \in T$;
(3) if $u, v \in F_{2,1}(X)$ then $(u)(v) \in F_{2,1}(X)$.

Thus, if $X = \{x, y, z\}$, then a typical element of $F_{2,1}(X)$ might be

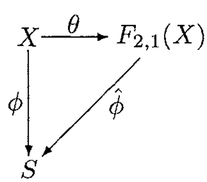
$$((((x)(y))'((z)(y)))')((((y)(((z)(x))'))')').$$

The set $F_{2,1}(X)$ can be made into a $(2,1)$-algebra in an obvious way by defining

$$uv = (u)(v), \quad u' = (u)'.$$

It is the *free* $(2,1)$-*algebra* (sometimes called the *absolutely free* $(2,1)$-*algebra*) on the set X, which is to say that it has the properties:

(F1) there is a map $\theta : X \to F_{2,1}(X)$—the obvious inclusion map;
(F2) for every $(2,1)$-algebra S and every map $\phi : X \to S$ there is a unique extension of ϕ to a morphism $\hat{\phi} : F_{2,1}(X) \to S$, that is to say, the diagram



commutes.

Let $\mathcal{V}$ be a non-empty class of $(2,1)$-algebras, and suppose that $\mathcal{V}$ has the following properties:

(V1) if $S \in \mathcal{V}$ and T is a subalgebra of S, then $T \in \mathcal{V}$;

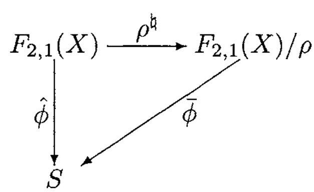
(V2) if $S \in \mathcal{V}$ and T is a morphic image of S, then $T \in \mathcal{V}$;
(V3) if $S_i \in \mathcal{V}$ $(i \in I)$, then the direct product $\prod\{S_i : i \in I\}$ also belongs to $\mathcal{V}$.

Then we say that $\mathcal{V}$ is a *variety* of $(2,1)$-algebras. To put it succinctly, a variety is a class of algebras closed under the taking of subalgebras, morphic images and direct products.

Let $\mathcal{V}$ be a variety of $(2,1)$-algebras, and let X be a non-empty set. We now establish the existence of a 'relatively free', or $\mathcal{V}$-*free*, $(2,1)$-algebra $F_{\mathcal{V}}(X)$ in the variety $\mathcal{V}$. Let $S \in \mathcal{V}$, and let S^X denote the set of all maps from X into S. For each ϕ in S^X there is a unique extension $\hat{\phi} : F_{2,1}(X) \to S$, a $(2,1)$-morphism whose image, being a subalgebra of S, is in $\mathcal{V}$. Let us denote the congruence $\ker\hat{\phi}$ on $F_{2,1}(X)$ by ρ_ϕ. Thus we see that each S in $\mathcal{V}$ and each ϕ in S^X determines a congruence ρ_ϕ on $F_{2,1}(X)$. Let us denote the set of all congruences on $F_{2,1}(X)$ obtained in this way by $\{\rho_i : i \in I\}$, and define

$$\rho = \bigcap\{\rho_i : i \in I\}. \tag{4.3.1}$$

Then $\rho \subseteq \rho_i$ for all i, and so for each S in $\mathcal{V}$ there exists a unique morphism $\bar{\phi} : F_{2,1}(X)/\rho \to S$ such that the diagram



commutes.

We now show that $F_{2,1}(X)/\rho \in \mathcal{V}$. By construction we know that $F_{2,1}(X)/\rho_i \in \mathcal{V}$ for every i in I. Hence $P \in \mathcal{V}$, where P is the direct product

$$P = \prod\{F_{2,1}(X)/\rho_i : i \in I\},$$

defined as in Section 1.1. This consists of all maps

$$\delta : I \to \bigcup\{F_{2,1}(X)/\rho_i : i \in I\},$$

with the property that

$$i\delta \in F_{2,1}(X)/\rho_i \quad (i \in I).$$

Define a morphism $\chi : F_{2,1}(X) \to P$ by the rule that, for each u in $F_{2,1}(X)$, $u\chi$ is the map from I into $\bigcup\{F_{2,1}(X)/\rho_i : i \in I\}$ given by

$$i(u\chi) = u\rho_i \quad (i \in I).$$

Then, for all u, v in $F_{2,1}(X)$,

$$u\chi = v\chi \iff i(u\chi) = i(v\chi) \text{ for all } i \text{ in } I,$$

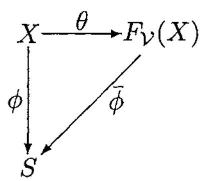
$$\iff u\rho_i = v\rho_i \text{ for all } i \text{ in } I,$$
$$\iff (u, v) \in \rho.$$

Thus $F_{2,1}(X)/\rho \simeq \operatorname{im}\chi$, a subalgebra of P, and so is in the variety $\mathcal{V}$.

We denote the $(2,1)$-algebra $F_{2,1}(X)/\rho$ by $F_\mathcal{V}(X)$, and call it the *$\mathcal{V}$-free algebra generated by X*. It is defined up to isomorphism by the properties:

(F1) there is a map $\theta : X \to F_\mathcal{V}(X)$, given by $x\theta = x\rho$;

(F2) for every S in $\mathcal{V}$ and every map $\phi : X \to S$ there is a unique extension of ϕ to a morphism $\bar{\phi} : F_\mathcal{V}(X) \to S$, that is to say, the diagram



commutes.

Let $u, v \in F_{2,1}(X)$, where X is a non-empty set, and let S be a $(2,1)$-algebra. As we have seen, every map $\phi : X \to S$ extends to a morphism $\hat{\phi} : F_{2,1}(X) \to S$. We say that S satisfies the *identical relation* (or *law*) $u = v$ if $u\hat{\phi} = v\hat{\phi}$ for every choice of $\phi : X \to S$. Informally, a $(2,1)$-algebra S satisfies a given law if we obtain equality in S for every possible substitution of the variables in u and v by elements of S.

Let X be a countably infinite set, and let $\mathcal{E}$ be a class of $(2,1)$-algebras. Suppose that there is a subset $\mathbf{R}$ of $F_{2,1}(X) \times F_{2,1}(X)$ such that $S \in \mathcal{E}$ if and only if S satisfies the identical relations $u = v$ for every (u, v) in $\mathbf{R}$. Then we say that $\mathcal{E}$ is an *equational class*, defined by the identical relations $u = v$ for (u, v) in $\mathbf{R}$. The following theorem, a special case of a general result due to Birkhoff (1935), is the key to the study of varieties:

Theorem 4.3.1 *Let $\mathcal{V}$ be a class of $(2,1)$-algebras. Then $\mathcal{V}$ is an equational class if and only if it is a variety.*

Proof It is a routine matter to verify that every equational class is a variety. To show the converse, let $\mathcal{V}$ be a variety, let X be an infinite set, and let $F_\mathcal{V}(X)$ be the $\mathcal{V}$-free algebra generated by X. Thus, as in equation (4.3.1), $F_\mathcal{V}(X) = F_{2,1}(X)/\rho$, where

$$\rho = \bigcap\{\rho_i : i \in I\}.$$

Let $S \in \mathcal{V}$ and let $(u, v) \in \rho$. Then certainly $(u, v) \in \rho_\phi$ for every ϕ in S^X and so for every map $\phi : X \to S$ we have $u\hat{\phi} = v\hat{\phi}$, where $\hat{\phi} : F_{2,1}(X) \to S$ is the unique extension of ϕ. Thus $u = v$ is an identical relation in S. We have shown that every S in $\mathcal{V}$ satisfies the law $u = v$ for every (u, v) in ρ.

Conversely, suppose that $S = (S, ., ')$ satisfies the law $u = v$ for every (u, v) in ρ. Choose an infinite set Y such that $|Y| \geq |S|$ and $|Y| \geq |X|$, and let $F_{\mathcal{V}}(Y) = F_{2,1}(Y)/\pi$ be the $\mathcal{V}$-free algebra generated by Y. We shall show that every morphism ψ from $F_{2,1}(Y)$ onto S factors through $F_{\mathcal{V}}(Y)$, and from this it will follow that $S \in \mathcal{V}$.

Let $(u, v) \in \pi$. We require to show that $u\psi = v\psi$. Now the set $C(u) \cup C(v)$ of all letters in Y featuring in u or in v is a finite set Y_0 of Y, and we may suppose that there exists a finite subset X_0 of X and an injection $\xi_0 : X \to Y$ such that $X_0\xi_0 = Y_0$. Let η_0 be a one-sided inverse of ξ_0, that is, $y\eta_0$ is the unique x such that $x\xi_0 = y$ if $y \in \operatorname{im}\xi_0$, and $y\eta_0$ is an arbitrary element of X otherwise. We can extend ξ_0 to a monomorphism $\xi : F_{2,1}(X) \to F_{2,1}(Y)$, and there are uniquely defined elements u_0, v_0 in $F_{2,1}(X)$ such that $u_0\xi = u$, $v_0\xi = v$. By the same token, we can extend η_0 to a morphism $\eta : F_{2,1}(Y) \to F_{2,1}(X)$, and our mode of construction ensures that $u\eta = u_0$, $v\eta = v_0$.

Now consider the map $\eta\rho^\natural : F_{2,1}(Y) \to F_{\mathcal{V}}(X)$. Then, since $F_{\mathcal{V}}(X) \in \mathcal{V}$, we must have that $\pi \subseteq \ker(\eta\rho^\natural)$. In particular, $u\eta\rho^\natural = v\eta\rho^\natural$, and so $(u_0, v_0) \in \rho$.

Let us return now to our arbitrary morphism ψ from $F_{2,1}(Y)$ onto S, and consider the morphism $\xi\psi : F_{2,1}(X) \to S$. Our assumption is that S satisfies all of the identities from ρ, and so in particular S satisfies $u_0 = v_0$. It follows that $u_0\xi\psi = v_0\xi\psi$, and hence that $u\psi = v\psi$, exactly as required. □

We shall here be concerned solely with varieties of I-semigroups, that is to say, with varieties of $(2,1)$-algebras contained in the variety of I-semigroups. We shall accordingly take the identities

$$(xy)z = x(yz), \quad (x')' = x \quad \text{and} \quad xx'x = x$$

as read, and we can list varieties of I-semigroups so far encountered, along with their associated laws. All the varieties here listed are *finitely based*, in the sense that the associated laws are all consequences of a finite collection of laws. We cannot assume that every variety has this property—see, for example, Austin (1966).

$\mathcal{CR}$: completely regular semigroups : $xx' = x'x$;
$\mathcal{CS}$: completely simple semigroups : $xx' = x'x, xyx(xyx)' = xx'$;
$\mathcal{CL}$: Clifford semigroups : $xx' = x'x,\ xx'yy' = yy'xx'$;
$\mathcal{G}$: Groups : $xx' = yy'$.

(See equations (4.1.1) and (4.2.1), and Proposition 4.1.2(2).)

We shall also want to refer to certain varieties of *semigroups*. It is easy to adapt—indeed to simplify—our analysis of $(2,1)$-algebras so as to deal with systems having just one binary operation, and so once again we can specify varieties by means of laws. The law $(xy)z = x(yz)$ is of course taken as read, since we are concerned solely with semigroups. Familiar examples of semigroup varieties are

$\mathcal{C}$	: commutative semigroups	$: xy = yx;$
$\mathcal{Z}$	: null semigroups	$: xy = zt;$
$\mathcal{B}$	: bands (idempotent semigroups)	$: x^2 = x;$
$\mathcal{RZ}$	: right zero semigroups	$: xy = y;$
$\mathcal{LZ}$	: left zero semigroups	$: xy = x;$
$\mathcal{RB}$	: rectangular bands	$: xyx = x;$
$\mathcal{SL}$	: semilattices	$: x^2 = x,\ xy = yx;$
$\mathcal{T}$	: trivial semigroups	$: x = y.$

(See Theorem 1.1.3.)

We shall often want to denote the variety defined by laws $w_1 = z_1$, $w_2 = z_2, \ldots$ by $[w_1 = z_1, w_2 = z_2, \ldots]$. Thus $\mathcal{C} = [xy = yx]$, $\mathcal{B} = [x^2 = x]$, etc.

Some further examples of varieties of semigroups and I-semigroups are explored in Exercises 9 to 11, but for the rest of this chapter we shall be concerned with the variety $\mathcal{B}$ of bands. This, as we have seen, is a variety of semigroups, but it may also be regarded as a variety of completely regular semigroups, for if we define $x' = x$ then the laws

$$(x')' = x, \quad xx'x = x, \quad xx' = x'x$$

follow automatically.

4.4 BANDS

Let B be a band. Since B is completely regular, it decomposes by Theorem 4.1.3 into a semilattice Y of completely simple semigroups S_α $(\alpha \in Y)$. Each of these completely simple semigroups, being a subsemigroup of B, is a band, and it is a band satifying the law $(xyx)(xyx)' = xx'$, by Proposition 4.1.2. Since $x' = x$ for every x in a band, this identity reduces to $xyx = x$, and so we conclude that each S_α is a rectangular band. We have shown

Theorem 4.4.1 *Every band is a semilattice of rectangular bands.* □

This is certainly a useful result, but it cannot be interpreted as a complete solution to the problem of describing the structure of bands, for in general we do not obtain a *strong* semilattice. We shall eventually identify (in Proposition 4.6.14) the precise class of bands for which we do obtain a strong semilattice, but in this section our aim will be to establish a general structure theorem for bands.

From Theorem 1.1.3 we know that every rectangular band is isomorphic to a cartesian product $I \times \Lambda$, with multiplication given by

$$(i, \lambda)(j, \mu) = (i, \mu).$$

In terms of this description, we now investigate the nature of morphisms between rectangular bands.

Proposition 4.4.2 *If ϕ is a morphism from a rectangular band $I_1 \times \Lambda_1$ into a rectangular band $I_2 \times \Lambda_2$, then there exist maps $\phi^l : I_1 \to I_2$ and $\phi^r : \Lambda_1 \to \Lambda_2$ such that, for all (x_1, ξ_1) in $I_1 \times \Lambda_1$,*

$$(x_1, \xi_1)\phi = (x_1\phi^l, \xi_1\phi^r). \tag{4.4.1}$$

Conversely, if $\phi^l : I_1 \to I_2$ and $\phi^r : \Lambda_1 \to \Lambda_2$ are arbitrary maps, then the formula (4.4.1) *defines a morphism from $I_1 \times \Lambda_1$ into $I_2 \times \Lambda_2$.*

Proof Let $\phi : I_1 \times \Lambda_1 \to I_2 \times \Lambda_2$ be a morphism. Choose a fixed λ_1 in Λ_1, and for every x_1 in I_1 define $x_1\phi^l$ by

$$(x_1, \lambda_1)\phi = (x_1\phi^l, \lambda_2).$$

Similarly, choose a fixed i_1 in I_1, and for every ξ_1 in Λ_1 define $\xi_1\phi^r$ by

$$(i_1, \xi_1)\phi = (i_2, \xi_1\phi^r).$$

Then for all (x_1, ξ_1) in $I_1 \times \Lambda_1$,

$$\begin{aligned}(x_1, \xi_1)\phi &= [(x_1, \lambda_1)(i_1, \xi_1)]\phi = [(x_1, \lambda_1)\phi][(i_1, \xi_1)\phi] \\ &= (x_1\phi^l, \lambda_2)(i_2, \xi_1\phi^r) = (x_1\phi^l, \xi_1\phi^l).\end{aligned}$$

Conversely, if ϕ is defined by (4.4.1) then, for all (x_1, ξ_1), (y_1, η_1) in $I_1 \times \Lambda_1$,

$$\begin{aligned}[(x_1, \xi_1)(y_1, \eta_1)]\phi &= (x_1, \eta_1)\phi = (x_1\phi^l, \eta_1\phi^l) = (x_1\phi^l, \xi_1\phi^r)(y_1\phi^l, \eta_1\phi^r) \\ &= [(x_1, \xi_1)\phi][(y_1, \eta_1)\phi].\end{aligned}$$

Thus ϕ is a morphism. □

If, as in Theorem 1.1.3, we choose to regard a rectangular band $I \times \Lambda$ as the direct product of the left zero semigroup I and the right zero semigroup Λ, then we may interpret the maps $\phi^l : I_1 \to I_2$ and $\phi^r : \Lambda_1 \to \Lambda_2$ as semigroup *morphisms.* There is no extra information involved in this version, since *every* map between left (or right) zero semigroups is a morphism. It is, however, useful to state the alternative version as a corollary:

Corollary 4.4.3 *Let L_1, L_2 be left zero semigroups and let R_1, R_2 be right zero semigroups. If ϕ is a morphism from the rectangular band $L_1 \times R_1$ into the rectangular band $L_2 \times R_2$, then there exist morphisms $\phi^l : L_1 \to L_2$, $\phi^r : R_1 \to R_2$ such that*

$$(l_1, r_1)\phi = (l_1\phi^l, r_1\phi^r) \tag{4.4.2}$$

for all (l_1, r_1) in $L_1 \times R_1$.

Conversely, for every pair of morphisms $\phi^l : L_1 \to L_2$, $\phi^r : R_1 \to R_2$, the formula (4.4.2) *defines a morphism from $L_1 \times R_1$ into $L_2 \times R_2$.* □

The key idea in the description of bands is that of a *translation.* We have already come across the maps λ_a and ρ_a associated with each element a of a semigroup S:

$$\lambda_a s = as, \quad s\rho_a = sa \quad (s \in S).$$

Here, for convenience, we are writing λ_a as a left map and ρ_a as a right map. We refer to λ_a and ρ_a as the *inner left and right translations* of S associated with the element a. Because of associativity, we have

$$\lambda_a(st) = (\lambda_a s)t, \quad (st)\rho_a = s(t\rho_a), \quad s(\lambda_a t) = (s\rho_a)t$$

for all a, s, t in S. These observations motivate the following definitions. The left map $\lambda : S \to S$ is called a *left translation* of S if

$$\lambda(st) = (\lambda s)t$$

for all s, t in S; the right map $\rho : S \to T$ is called a *right translation* of S if

$$(st)\rho = s(t\rho)$$

for all s, t in S; the left translation λ and the right translation ρ are said to be *linked* if

$$s(\lambda t) = (s\rho)t$$

for all s, t in S. The set of all linked pairs (λ, ρ) of left and right translations is called the *translational hull* of S and is denoted by $\Omega(S)$. It is a semigroup under the obvious multiplication

$$(\lambda, \rho)(\lambda', \rho') = (\lambda\lambda', \rho\rho'),$$

where $\lambda\lambda'$ denoted the composition of the *left* maps λ and λ' (that is, first λ', then λ), while $\rho\rho'$ denotes the composition of the right maps ρ and ρ' (that is, first ρ, then ρ'). The proof of this assertion is routine.

Within $\Omega(S)$ there is a linked pair (λ_a, ρ_a) for each a in S, and it is easy to verify that, for all a, b in S,

$$(\lambda_a, \rho_a)(\lambda_b, \rho_b) = (\lambda_{ab}, \rho_{ab}).$$

Thus we have a morphism $a \mapsto (\lambda_a, \rho_a)$ from S into $\Omega(S)$. In general this is not a monomorphism—see Exercise 7—but it is a monomorphism in the cases that concern us:

Lemma 4.4.4 *If a and b are elements in a regular semigroup S, then*

$$[\lambda_a = \lambda_b \text{ and } \rho_a = \rho_b] \Rightarrow a = b.$$

Proof Suppose that $\lambda_a = \lambda_b$ and $\rho_a = \rho_b$, and let $a' \in V(a)$, $b' \in V(b)$. Then

$$a = aa'a = (\lambda_a a')a = (\lambda_b a')a = ba'a,$$

and so (by (2.1.4)) $R_a \leq R_b$. Similar arguments show that $L_a \leq L_b$, $R_b \leq R_a$, $L_b \leq L_a$, and so $a \mathcal{H} b$. By Proposition 2.4.1 we may now suppose that a' and b' have been chosen so that $aa' = bb'$ and $a'a = b'b$, and it then easily follows that $a = ba'a = bb'b = b$. □

Consider now a completely regular semigroup S that has been expressed as a semilattice Y of completely simple semigroups S_α ($\alpha \in Y$), and consider two elements α, β in Y such that $\alpha \geq \beta$. Then $S_\alpha S_\beta$ and $S_\beta S_\alpha$ are both contained in $S_{\alpha\beta} = S_\beta$, and so every element a in S_α induces maps λ_a, ρ_a from S_β into S_β:

$$\lambda_a x = ax, \quad x\rho_a = xa.$$

The associativity of S guarantees that λ_a and ρ_a are linked left and right translations, and hence $a \mapsto (\lambda_a, \rho_a)$ defines a map (indeed a morphism) from S_α into $\Omega(S_\beta)$, the translational hull of S_β.

In the case considered in Section 4.2, where S is a Clifford semigroup and each S_α is a group, the translational hull of S_β can be shown to be isomorphic to S_β itself—see Exercise 6—and the morphism $a \mapsto (\lambda_a, \rho_a)$ essentially reduces to the morphism $\phi_{\alpha,\beta} : S_\alpha \to S_\beta$ already described. In general, however, the translations λ_a and ρ_a are not inner translations of S_β.

The problem of describing the translational hull of a completely simple semigroup has been tackled by Petrich (1968). The description is complicated, and here we shall confine ourselves to the case of a rectangular band $E = I \times \Lambda$. Let λ be a left translation of E, let $(i, \mu) \in E$, and suppose that $\lambda(i, \mu) = (i^*, \mu^*)$, an element of E. Then

$$\lambda(i, \mu) = \lambda[(i, \mu)(i, \mu)] = [\lambda(i, \mu)](i, \mu) = (i^*, \mu^*)(i, \mu) = (i^*, \mu),$$

and so $\mu^* = \mu$. Moreover, for every ξ in Λ,

$$\lambda(i, \xi) = \lambda[(i, \mu)(i, \xi)] = [\lambda(i, \mu)](i, \xi) = (i^*, \mu)(i, \xi) = (i^*, \xi).$$

Thus λ determines a map $\phi : I \to I$ (which we shall find it convenient to write as a left map) such that, for all (i, ξ) in E,

$$\lambda(i, \xi) = (\phi i, \xi). \tag{4.4.3}$$

It is easy to verify the converse, that for every map $\phi : I \to I$ the formula (4.4.3) defines a left translation of E.

A closely similar argument establishes that every right translation ρ of E determines and is determined by a right map $\psi : \Lambda \to \Lambda$ such that, for all (x, μ) in E,

$$(x, \mu)\rho = (x, \mu\psi). \tag{4.4.4}$$

Now notice that if λ, defined by (4.4.3), is a left translation, and ρ, defined by (4.4.4), is a right translation, then, for all (i, μ), (j, ν) in E,

$$(i, \mu)[\lambda(j, \nu)] = (i, \mu)(\phi j, \nu) = (i, \nu) = (i, \mu\psi)(j, \nu) = [(i, \mu)\rho](j, \nu);$$

thus every pair (λ, ρ) is linked. It is now easy to verify that the map $(\lambda, \rho) \mapsto (\phi, \psi)$ gives an isomorphism from the translational hull $\Omega(I \times \Lambda)$

onto the direct product $\mathcal{T}_I^* \times \mathcal{T}_\Lambda$ of the semigroup $\mathcal{T}_I^*$ of all left maps from I into I with the semigroup $\mathcal{T}_\Lambda$ of all right maps from Λ into Λ.

Consider now a general band B, and suppose that we have expressed it as a semilattice Y of rectangular bands E_α $(\alpha \in Y)$. Write E_α as $I_\alpha \times \Lambda_\alpha$. We have seen that for all $\alpha \geq \beta$ in Y each a in E_α induces a linked pair (λ_a, ρ_a) of left and right translations of E_β, and that the map $a \mapsto (\lambda_a, \rho_a)$ is a morphism from S_α into $\Omega(S_\beta)$. As a result of our investigations into the translational hull of a rectangular band, we can now be a bit more explicit and assert that a induces a left map $\phi_\beta^a : I_\beta \to I_\beta$ and a right map $\psi_\beta^a : \Lambda_\beta \to \Lambda_\beta$, in accordance with the formulae

$$a(x_\beta, \xi_\beta) = (\phi_\beta^a x_\beta, \xi_\beta), \quad (x_\beta, \xi_\beta)a = (x_\beta, \xi_\beta \psi_\beta^a). \tag{4.4.5}$$

To put the same thing more globally, we have, whenever $\alpha \geq \beta$, a morphism

$$\Phi_{\alpha,\beta} : E_\alpha \to \mathcal{T}_{I_\beta}^* \times \mathcal{T}_{\Lambda_\beta}$$

given by

$$a\Phi_{\alpha,\beta} = (\phi_\beta^a, \psi_\beta^a) \quad (a \in E_\alpha).$$

Consider now what happens when $\beta = \alpha$, and when $a = (i, \mu)$. From the formula (4.4.5) we have

$$a(x_\alpha, \xi_\alpha) = (\phi_\alpha^{(i,\mu)} x_\alpha, \xi_\alpha), \quad (x_\alpha, \xi_\alpha)a = (x_\alpha, \xi_\alpha \psi_\alpha^{(i,\mu)}),$$

while from the rule for multiplication within E_α we have

$$(i, \mu)(x_\alpha, \xi_\alpha) = (i, \xi_\alpha), \quad (x_\alpha, \xi_\alpha)(i, \mu) = (x_\alpha, \mu).$$

Hence the map $\phi^{(i,\mu)} : I_\alpha \to I_\alpha$ has the property that $\phi_\alpha^{(i,\mu)} x_\alpha = i$ for all i in I_α—and similarly $\xi_\alpha \psi_\alpha^{(i,\mu)} = \mu$ for all μ in Λ. Thus, if we adopt a notation whereby the constant value of a constant map χ is written as $\langle \chi \rangle$, we may write

$$\langle \phi_\alpha^{(i,\mu)} \rangle = i, \quad \langle \psi_\alpha^{(i,\mu)} \rangle = \mu \tag{4.4.6}$$

whenever $(i, \mu) \in E_\alpha$.

We are now ready to consider a more general product in B. Let $a \in E_\alpha$, $b \in E_\beta$, let $\alpha\beta = \gamma$, and let $z = (x_\gamma, \xi_\gamma)$ be an arbitrary element of E_γ. Then $ab \in E_\gamma$; let us write $ab = (i_\gamma, \mu_\gamma)$. Then

$$abz = (ab)z = (i_\gamma, \mu_\gamma)(x, \xi) = (i_\gamma, \xi_\gamma),$$

and

$$abz = a[b(x_\gamma, \xi_\gamma)] = a(\phi_\gamma^b x_\gamma, \xi_\gamma) = (\phi_\gamma^a \phi_\gamma^b x_\gamma, \xi_\gamma).$$

We deduce that the left map $\phi_\gamma^a \phi_\gamma^b$ of I_γ has the property that

$$\phi_\gamma^a \phi_\gamma^b x_\gamma = i_\gamma$$

for every x_γ in I_γ. Thus $\phi_\gamma^a \phi_\gamma^b$ is a constant map, with constant value i_γ. Dually, by considering zab in two different ways, we see that the right map $\psi^a_\gamma \psi^b_\gamma$ of Λ_γ has the constant value μ_γ. We thus obtain the product $ab = (i_\gamma, \mu_\gamma)$ of a and b in terms of the maps ϕ^a_γ, ϕ^b_γ, ψ^a_γ, ψ^b_γ as follows:

$$ab = \big(\langle \phi^a_\gamma \phi^b_\gamma \rangle, \langle \psi^a_\gamma \psi^b_\gamma \rangle\big). \tag{4.4.7}$$

If, then, we think of the morphisms $\Phi_{\alpha,\beta}$ as 'known', then formula (4.4.7) shows how the product ab of two arbitrary elements of B is determined by these morphisms.

Consider now what happens when we multiply the product ab on the right by an element $d = (x_\delta, \xi_\delta)$ of E_δ, where $\delta \leq \alpha\beta$. On the one hand

$$abd = (ab)d = (\phi^{ab}_\delta x_\delta, \xi_\delta),$$

while on the other hand

$$abd = a(bd) = a(\phi^b_\delta x_\delta, \xi_\delta) = (\phi^a_\delta \phi^b_\delta x_\delta, \xi_\delta).$$

We deduce that

$$\phi^{ab}_\delta = \phi^a_\delta \phi^b_\delta, \tag{4.4.8}$$

and a similar argument based on the two computations of dab gives us the corresponding formula

$$\psi^{ab}_\delta = \psi^a_\delta \psi^b_\delta. \tag{4.4.9}$$

It is convenient now to state the theorem towards which we are working:

Theorem 4.4.5 *Let Y be a semilattice and let $\{E_\alpha : \alpha \in Y\}$ be a family of pairwise disjoint rectangular bands indexed by Y. For each α, let $E_\alpha = I_\alpha \times \Lambda_\alpha$, and for each pair α, β of elements of Y such that $\alpha \geq \beta$ let $\Phi_{\alpha,\beta} : E_\alpha \to \mathcal{T}^*_{I_\beta} \times \mathcal{T}_{\Lambda_\beta}$ be a morphism, where*

$$a\Phi_{\alpha,\beta} = (\phi^a_\beta, \psi^b_\beta) \quad (a \in E_\alpha).$$

Suppose also that

(1) *if $(i, \mu) \in E_\alpha$, then $\phi^{(i,\mu)}_\alpha$ and $\psi^{(i,\mu)}_\alpha$ are constant maps, and*
$$\langle \phi^{(i,\mu)}_\alpha \rangle = i, \quad \langle \psi^{(i,\mu)}_\alpha \rangle = \mu;$$
(2) *if $a \in S_\alpha$, $b \in S_\beta$ and $\alpha\beta = \gamma$, then $\phi^a_\gamma \phi^b_\gamma$ and $\psi^a_\gamma \psi^b_\gamma$ are constant maps;*
(3) *if $\langle \phi^a_\gamma \phi^b_\gamma \rangle$ is denoted by j and $\langle \psi^a_\gamma \psi^b_\gamma \rangle$ by ν, then, for all $\delta \leq \gamma$,*
$$\phi^{(j,\nu)}_\delta = \phi^a_\delta \phi^b_\delta, \quad \psi^{(j,\nu)}_\delta = \psi^a_\delta \psi^b_\delta.$$

Let $B = \bigcup\{E_\alpha : \alpha \in Y\}$ and define the product of a in E_α and b in E_β by

$$a * b = \big(\langle \phi^a_\gamma \phi^b_\gamma \rangle, \langle \psi^a_\gamma \psi^b_\gamma \rangle\big),$$

where $\gamma = \alpha\beta$. Then $(B, *)$ is a band, whose $\mathcal{J}$-classes are the rectangular bands E_α.

Conversely, every band is determined in this way by a semilattice Y, a family of rectangular bands $E_\alpha = I_\alpha \times \Lambda_\alpha$ indexed by Y, and a family of *morphisms* $\Phi_{\alpha,\beta} : E_\alpha \to \mathcal{T}^*_{I_\beta} \times \mathcal{T}_{\Lambda_\beta}$ $(\alpha, \beta \in Y,\ \alpha \geq \beta)$ *satisfying* (1), (2) *and* (3).

Proof We have already established the more difficult converse half of this result, by virtue of formulae (4.4.6) to (4.4.9). To prove the direct half we begin by showing that the given multiplication is associative. Let $a \in E_\alpha$, $b \in E_\beta$ and $c \in E_\gamma$, with $\alpha\beta = \delta$, $\beta\gamma = \epsilon$, and $\alpha\beta\gamma = \zeta$. Then

$$a * b = \big(\langle \phi^a_\delta \phi^b_\delta \rangle, \langle \psi^a_\delta \psi^b_\delta \rangle\big) = (j, \nu), \text{ say,}$$

and

$$b * c = \big(\langle \phi^b_\epsilon \phi^c_\epsilon \rangle, \langle \psi^b_\epsilon \psi^c_\epsilon \rangle\big) = (k, \pi), \text{ say.}$$

Hence

$$\begin{aligned}(a * b) * c &= \big(\langle \phi^{(j,\nu)}_\zeta \phi^c_\zeta \rangle, \langle \psi^{(j,\nu)}_\zeta \psi^c_\zeta \rangle\big) = \big(\langle \phi^a_\zeta \phi^b_\zeta \phi^c_\zeta \rangle, \langle \psi^a_\zeta \psi^b_\zeta \psi^c_\zeta \rangle\big) \\ &= \big(\langle \phi^a_\zeta \phi^{(k,\pi)}_\zeta \rangle, \langle \psi^a_\zeta \psi^{(k,\pi)}_\zeta \rangle\big) = a * (b * c).\end{aligned}$$

Next, note that if $a = (i, \mu)$ and $b = (j, \nu)$ both belong to E_α, then the multiplication formula gives

$$a * b = \big(\langle \phi^{(i,\mu)}_\alpha \phi^{(j,\nu)}_\alpha \rangle, \langle \psi^{(i,\mu)}_\alpha \psi^{(j,\nu)}_\alpha \rangle\big) = (i, \nu),$$

by property (1) and by the properties of constant left and right maps. This coincides exactly with the product of a and b in the rectangular band E_α. In particular it follows that $a * a = a$, and so B is a band.

The multiplication formula implies that

$$E_\alpha E_\beta \subseteq E_{\alpha\beta},$$

and so two elements can be $\mathcal{J}$-equivalent only if they fall in the same E_α. Since any two elements in E_α are easily seen to be $\mathcal{J}$-equivalent, we thus conclude that the $\mathcal{J}$-classes of B are the rectangular bands E_α. □

4.5 FREE BANDS

From the general arguments given in Section 4.3 we know that (relatively) free objects exist in every variety. The arguments do not, however, give any clue as to how to obtain a usable description of the free objects, and in practice this can be very difficult. For the variety of bands, however, we do have a good (though far from trivial) description, due to Green and Rees (1952).

In one sense it is easy to give a description of the *free band* B_A on a set A of generators. If β is the congruence relation on the free semigroup A^+ generated by the subset

$$\mathbf{B} = \{(w^2, w) : w \in A^+\}$$

of $A^+ \times A^+$, then it is a routine matter to verify that $B_A = A^+/\beta$ has precisely the properties that we require of the free band on A. On the other hand, this is for most purposes not a very useful description, for it is not at all clear how to determine whether or not two words u and v in A^+ are β-equivalent.

For each w in A^+ we define the *content* $C(w)$ to be the (necessarily finite) set of elements of A appearing in w. By Proposition 1.5.9, two distinct β-equivalent elements of A^+ are connected by a sequence of elementary **B**-transitions, that is to say, by steps of one or other of the types

$$pxq \to px^2q, \quad px^2q \to pxq \quad (p, q \in A^*,\ x \in A^+);$$

hence β-equivalent elements have the same content, and it is possible to talk in an unambiguous way of the *content* $C(w\beta)$ of an element $w\beta$ of B_A. For each non-empty finite subset P of A we denote by U_P the set of elements of B_A with content P. Each U_P is fairly obviously a sub-band of B_A, and $U_P \cap U_{P'} = \emptyset$ if $P \neq P'$.

Suppose now that w, z in A^+ are such that $w\beta\ \mathcal{J}\ z\beta$ in B_A, so that there exist x, y, u, v in A^* such that

$$(x\beta)(w\beta)(y\beta) = z\beta, \quad (u\beta)(z\beta)(v\beta) = w\beta.$$

The first of these equations implies that $C(w\beta) \subseteq C(z\beta)$, and the second implies that $C(z\beta) \subseteq C(w\beta)$. Thus

$$w\beta\ \mathcal{J}\ z\beta \ \Rightarrow\ C(w\beta) = C(z\beta).$$

In fact the converse implication is true also. For suppose that

$$C(w\beta) = C(z\beta) = \{a_1, a_2, \ldots, a_n\}.$$

Then, since B_A, being a band, is completely regular, formulae (4.1.2) and (4.1.3) apply, and imply that each of $w\beta$ and $z\beta$ is $\mathcal{J}$-equivalent to $(a_1a_2 \ldots a_n)\beta$. Thus

$$w\beta\ \mathcal{J}\ z\beta \text{ if and only if } C(w\beta) = C(z\beta), \tag{4.5.1}$$

and we conclude that each U_P is a $\mathcal{J}$-class of B_A. The gross structure of B_A is thus fairly clear, the semilattice involved being effectively the set of finite subsets of A under the operation $\cup$.

By Theorem 4.4.1, each U_P is a rectangular band, and so we have the not altogether obvious result that if $x\beta$, $y\beta$ and $z\beta$ are elements of B_A with the same content then

$$(x\beta)(y\beta)(z\beta) = (x\beta)(z\beta). \tag{4.5.2}$$

This equality generalizes to the case where

$$C(y\beta) \subseteq C(x\beta) = C(z\beta),$$

for in this case $C[(x\beta)(y\beta)(z\beta)] = C(x\beta) = C(z\beta)$, and so

$$(x\beta)(y\beta)(z\beta) = (x\beta)[(x\beta)(y\beta)(z\beta)](z\beta) = (x\beta)(z\beta).$$

Formula (4.5.1) gives us a straightforward algorithm for determining whether two elements of B_A are $\mathcal{J}$-equivalent. We now describe the more complicated algorithm by which we can determine whether or not they are equal. Let $u \in A^+$, where $|C(u)| = n \geq 1$. We define $\overline{u}(0)$ to be the letter in u that is *last to make its first appearance*, and $u(0)$ to be the subword of u that *precedes the first appearance of* $\overline{u}(0)$. (Thus, for example, if $n = 4$ and $u = abacabcdbcbd$, then $\overline{u}(0) = d$ and $u(0) = abacabc$.) Dually, we define $\overline{u}(1)$ to be the letter in u that is *first to make its last appearance*, and $u(1)$ to be the subword that *follows the last appearance of* $\overline{u}(1)$. (In our example, $\overline{u}(1) = a$ and $u(1) = bcdbcbd$.) If $|C(u)| = 1$ then the meaning of $u(0)$ and $u(1)$ is unclear: it is convenient to define $u(0) = u(1) = 1$ (the empty word) in this case.

In the terminology of Green and Rees (1952), $\overline{u}(0)$ is the *initial mark*, $u(0)$ is the *initial*, $\overline{u}(1)$ is the *terminal mark*, and $u(1)$ is the *terminal*. Since we shall require to iterate the process, it is more convenient here to adopt the general notation of Gerhard (1970).

Lemma 4.5.1 *Let $u, v \in A^+$. Then $(u, v) \in \beta$ if and only if*

(1) $C(u) = C(v)$;
(2) $\overline{u}(0) = \overline{v}(0), \quad \overline{u}(1) = \overline{v}(1)$;
(3) $\big(u(0), v(0)\big) \in \beta, \quad \big(u(1), v(1)\big) \in \beta$.

Proof Suppose first that (1), (2) and (3) are given. Then

$$C\big(u(0)\overline{u}(0)\big) = C(u) = C(v) = C\big(v(0)\overline{v}(0)\big),$$
$$C\big(\overline{u}(1)u(1)\big) = C(u) = C(v) = C\big(\overline{v}(1)v(1)\big).$$

Using (2), (3) and equation (4.5.2), we deduce that (modulo β)

$$\begin{aligned} u &\equiv u(0)\overline{u}(0)u \equiv u(0)\overline{u}(0)u\overline{u}(1)u(1) \equiv u(0)\overline{u}(0)\overline{u}(1)u(1) \\ &\equiv v(0)\overline{v}(0)\overline{v}(1)v(1) \equiv v(0)\overline{v}(0)v\overline{v}(1)v(1) \equiv v. \end{aligned}$$

Conversely, suppose that $(u, v) \in \beta$, so that u and v are connected by a finite sequence of elementary **B**-transitions. For simplicity, let us suppose first that u and v are connected by a single elementary **B**-transition, and suppose, without essential loss of generality, that

$$u = puq, \quad v = pu^2q \quad (u \in A^+,\ p, q \in A^*).$$

Then it is clear that $\overline{u}(0) = \overline{v}(0)$. If the first occurrence of $\overline{u}(0)$ is in pu then we even have *equality* between $u(0)$ and $v(0)$; if $\overline{u}(0)$ occurs first in q then all we can say is that $u(0) \to v(0)$ by an elementary **B**-transition. This, however, is enough, for we certainly conclude in both cases that

$$\overline{u}(0) = \overline{v}(0) \text{ and } \big(u(0), v(0)\big) \in \beta.$$

Similarly,

$$\overline{u}(1) = \overline{v}(1) \text{ and } \big(u(1), v(1)\big) \in \beta.$$

Clearly these results then extend to the case where u and v are connected by a finite sequence of elementary **B**-transitions, and so the lemma is proved. □

If the β-equivalent elements u and v considered in the lemma are such that $|C(u)| = |C(v)| = k$, then the elements $u(0)$, $u(1)$, $v(0)$, $v(1)$ featuring in part (3) of the conditions all have contents of cardinality $k-1$. The lemma does therefore give useful information, particularly if we repeat the process. Before we do this, however, it is helpful to have some further notation. Let P_j be the set of words of length j in the alphabet $\{0,1\}$, and let $\Sigma_k = \bigcup\{P_j : 1 \le j \le k\}$. For each u in A^+ and each α in Σ_k we now make recursive definitions of $\overline{u}(\alpha)$ and $u(\alpha)$ as follows:

$$\overline{u}(\alpha 0) = \overline{u(\alpha)}(0), \quad u(\alpha 0) = u(\alpha)(0),$$
$$\overline{u}(\alpha 1) = \overline{u(\alpha)}(1), \quad u(\alpha 1) = u(\alpha)(1).$$

(Thus, in the example $u = abacabcdbcbd$ already considered,

$$\begin{array}{llll} \overline{u}(00) = c, & u(00) = aba, & \overline{u}(01) = a, & u(01) = bc, \\ \overline{u}(10) = d, & u(10) = bc, & \overline{u}(11) = c & u(11) = bd, \end{array}$$

and it is a routine matter to go on to calculate $\overline{u}(\alpha)$ and $u(\alpha)$ for words of length 3 and 4.)

We now apply Lemma 4.5.1 repeatedly, recalling the convention that $w(0) = w(1) = 1$ for every w for which $|C(w)| = 1$. We obtain

Theorem 4.5.2 *Let A be a non-empty set, and let β be the congruence $\mathbf{B}^{\#}$ on A^+, where*

$$\mathbf{B} = \{(w^2, w) : w \in A^+\}.$$

If u and v are elements of A^+ such that $C(u) = C(v) = \{a_1, a_2, \ldots, a_k\}$, then $(u,v) \in \beta$ if and only if $\overline{u}(\alpha) = \overline{u}(\beta)$ for every α in Σ_k. □

One consequence of this analysis is that each U_P is finite. First, it is certainly the case that $|U_P|$ is determined only by $|P|$; so if $|P| = k$ we shall write $|U_P| = c_k$. An element $u\beta$ in U_P uniquely determines a quadruple $\big(\overline{u}(0), \overline{u}(1), [u(0)]\beta, [u(1)]\beta\big)$, and every quadruple $(a_1, a_1, w_0\beta, w_1\beta)$ for which $a_0, a_1 \in P$ and $w_0\beta \in U_{P\setminus\{a_0\}}$, $w_1\beta \in U_{P\setminus\{a_1\}}$ determines an element $(w_0 a_0 a_1 w_1)\beta$ of U_P. The number of elements in U_P is thus the number of quadruples of the kind described, and this leads to the recursion formula

$$c_k = k^2 c_{k-1}^2.$$

Hence, observing that $c_1 = 1$, we obtain by iteration that

$$c_k = k^2(k-1)^4(k-2)^8 \ldots = \prod_{i=1}^{k-1} (k-i+1)^{2^i}. \tag{4.5.3}$$

If A is finite—say $|A| = n$—we now readily obtain a formula for the number of elements of B_A:

Theorem 4.5.3 *The free band on a finite set A is finite. Specifically, if $|A| = n$, then*

$$|B_A| = \sum_{k=1}^{n} \binom{n}{k} c_k,$$

where c_k is given by (4.5.3). □

The number of elements in the free band on n generators increases with great rapidity as n increases. If we denote the number by b_n, then by direct calculation we obtain

$$\begin{array}{ll} c_1 = 1 & b_1 = 1 \\ c_2 = 4 & b_2 = 6 \\ c_3 = 144 & b_3 = 159 \\ c_4 = 331,776 & b_4 = 332,380 \\ c_5 = 2,751,882,854,400 & b_5 = 2,751,884,514,765, \end{array}$$

while c_6 is approximately 2.7×10^{26}.

Before we leave the topic of free bands, it is of interest to record that our analysis enables us to characterize the Green equivalences $\mathcal{R}$ and $\mathcal{L}$:

Proposition 4.5.4 *Let $x\beta, y\beta \in B_A$. Then*

(1) $(x\beta, y\beta) \in \mathcal{R}$ *if and only if*

$$C(x) = C(y),\ \overline{x}(0) = \overline{y}(0) \ \text{ and } \ \big(x(0), y(0)\big) \in \beta;$$

(2) $(x\beta, y\beta) \in \mathcal{L}$ *if and only if*

$$C(x) = C(y),\ \overline{x}(1) = \overline{y}(1) \ \text{ and } \ \big(x(1), y(1)\big) \in \beta.$$

Proof It will be sufficient to prove the first assertion. Accordingly, suppose first that $(x\beta, y\beta) \in \mathcal{R}$. Then certainly $(x\beta, y\beta) \in \mathcal{J}$ and so, by (4.5.1), $C(x) = C(y)$. Moreover, there exist u, v in A^+ such that (modulo β)

$$xu \equiv y, \quad yv \equiv x.$$

By Lemma 4.5.1(1) we deduce that $C(x) = C(y) = C(xu)$, and so it follows that $C(u) \subseteq C(x)$. Hence any letter that is going to appear in xu has already appeared in x, and so

$$\overline{y}(0) = \overline{xu}(0) = \overline{x}(0), \quad y(0) \equiv (xu)(0) = x(0),$$

as required.

Conversely, suppose that x, y in A^+ are such that $\overline{x}(0) = \overline{y}(0) = a$ and

$$x(0) \equiv y(0) \equiv w \pmod{\beta}.$$

Then (modulo β)

$$x \equiv x(0)\overline{x}(0)x \equiv wax,$$

where $C(wa) = C(x)$. Hence, by (4.5.2),

$$xwa \equiv waxwa \equiv (wa)^2 \equiv wa,$$

and so $x\beta \; \mathcal{R} \; (wa)\beta$ in B_A. A similar argument shows that $y\beta \; \mathcal{R} \; (wa)\beta$, and so $x\beta \; \mathcal{R} \; y\beta$ as required. □

4.6 VARIETIES OF BANDS*

It is necessary to begin this section with some observations on varieties further to those already made in Section 4.3. Since we are now concerned solely with varieties of bands, and since these can be regarded either as varieties of $(2,1)$-algebras or varieties of semigroups, we can pursue the ideas in the simpler context of semigroups. It is not at all hard to modify the ideas so as to cope with algebras of more complicated type.

We have seen that every variety $\mathcal{V}$ is determined by a set of identical relations $\mathbf{R}$, and we have decided to write $\mathcal{V} = [\mathbf{R}]$. If we have a family of varieties $\mathcal{V}_i = [\mathbf{R}_i] \quad (i \in I)$, then it is clear from the definition of a variety in Section 4.3 that $\mathcal{V} = \bigcap\{\mathcal{V}_i : i \in I\}$ is again a variety, and that the set of identical relations defining $\mathcal{V}$ is $\bigcup\{\mathbf{R}_i : i \in I\}$. So, for example, the intersection $\mathcal{B} \cap \mathcal{C}$ of the variety $\mathcal{B}$ of bands and the variety $\mathcal{C}$ of commutative semigroups is the variety $[x^2 = x,\; xy = yx]$ of semilattices.

By contrast, the union of a family of varieties need not be a variety. On the other hand, for any given family $\{\mathcal{V}_i : i \in I\}$ there exists a *join* $\mathcal{U} = \bigvee\{\mathcal{V}_i : i \in I\}$, namely the intersection of the collection of all varieties containing $\bigcup\{\mathcal{V}_i : i \in I\}$. (The collection is non-empty, since the variety of *all* semigroups necessarily contains every $\mathcal{V}_i$.) The determination of a set of identities characterizing $\mathcal{U}$ is not a trivial process, and to get any idea at all of how to proceed we need a new idea.

A congruence ρ on a semigroup S is called *fully invariant* if, for every endomorphism α of S, and for all x, y in S,

$$(x,y) \in \rho \;\Rightarrow\; (x\alpha, y\alpha) \in \rho.$$

It is a routine matter to verify that the intersection of a non-empty family of fully invariant congruences is a fully invariant congruence.

Now let $\mathbf{R}$ be a set of identical relations. That is to say, let $\mathbf{R}$ be a subset of $A^+ \times A^+$, where A is a countable set. Define $\mathbf{R}^v$ to be the smallest fully invariant congruence containing $\mathbf{R}$, namely the intersection of the collection of all fully invariant congruences containing $\mathbf{R}$. (Since not every congruence is fully invariant, $\mathbf{R}^v$ is potentially larger than $\mathbf{R}^\#$.)

For an arbitrary semigroup S, let $\mathbf{I}(S)$ be the subset of $A^+ \times A^+$ consisting of all pairs (w_1, w_2) for which the identical relation $w_1 = w_2$ is satisfied in S. Then we have

Proposition 4.6.1 *The relation $\mathbf{I}(S)$ is a fully invariant congruence on A^+. If $\mathcal{V} = [\mathbf{R}]$ is a variety, then $S \in \mathcal{V}$ if and only if $\mathbf{I}(S) \supseteq \mathbf{R}^v$.*

Proof It is obvious that $\mathbf{I}(S)$ is an equivalence. Suppose now that

$$(w_1, w_2), (z_1, z_2) \in \mathbf{I}(S).$$

Then $w_1\phi = w_2\phi$ and $z_1\phi = z_2\phi$ for every morphism $\phi : A^+ \to S$, and so

$$(w_1z_1)\phi = (w_1\phi)(z_1\phi) = (w_2\phi)(z_2\phi) = (w_2z_2)\phi.$$

Thus $(w_1z_1, w_2z_2) \in \mathbf{I}(S)$, and so $\mathbf{I}(S)$ is a congruence.

To show that $\mathbf{I}(S)$ is fully invariant, consider an endomorphism $\alpha : A^+ \to A^+$, and let $\phi : A^+ \to S$ be an arbitrary morphism. Then $\alpha\phi : A^+ \to S$ is a morphism, and so $w_1\alpha\phi = w_2\alpha\phi$ for every (w_1, w_2) in $\mathbf{I}(S)$. It follows that $(w_1\alpha, w_2\alpha) \in \mathbf{I}(S)$, and hence that $\mathbf{I}(S)$ is fully invariant.

Consider now a variety $\mathcal{V} = [\mathbf{R}]$, and suppose that $S \in \mathcal{V}$. Then certainly $\mathbf{R} \subseteq \mathbf{I}(S)$. Since $\mathbf{I}(S)$ is then a fully invariant congruence on A^+ containing $\mathbf{R}$, it must contain the *smallest* fully invariant congruence containing $\mathbf{R}$, that is, $\mathbf{R}^v \subseteq \mathbf{I}(S)$.

Conversely, if $\mathbf{R}^v \subseteq \mathbf{I}(S)$ then certainly $\mathbf{R} \subseteq \mathbf{I}(S)$; hence the identical relations in $\mathbf{R}$ are all satisfied in S, and so $S \in \mathcal{V}$. □

For a given S in a variety $\mathcal{V} = [\mathbf{R}]$ it may well be the case that $\mathbf{I}(S)$ *properly* contains $\mathbf{R}^v$, for we may have chosen an S in $[\mathbf{R}]$ that 'accidentally' satisfies more identical relations than are implied by the original set $\mathbf{R}$. For example, if $\mathbf{R} = \{(xy, yx)\}$, we may happen to choose an S in $[\mathbf{R}]$ that is a semilattice, and so contained in a variety strictly smaller than $[\mathbf{R}]$. However, we can reasonably regard $\mathbf{R}^v$ as providing us with the totality of identical relations associated with the variety $[\mathbf{R}]$, since, as the next result makes clear, there is always at least one semigroup S in $[\mathbf{R}]$ for which $\mathbf{I}(S) = \mathbf{R}^v$:

Proposition 4.6.2 *Let A be a countable set, let $\mathbf{R} \subseteq A^+ \times A^+$, and let $\rho = \mathbf{R}^v$. Then A^+/ρ is the (relatively) free semigroup in the variety $[\mathbf{R}]$, and $\mathbf{I}(A^+/\rho) = \mathbf{R}^v$.*

Proof We show first that $A^+/\rho \in [\mathbf{R}]$, which amounts to showing that every (w_1, w_2) in $\mathbf{R}$ is satisfied identically in A^+/ρ. So let $(w_1, w_2) \in \mathbf{R}$, and let $\theta : A^+ \to A^+/\rho$ be a morphism. For each x in A, choose an element in $x\theta(\rho^\natural)^{-1}$ and call it $x\gamma$. Since A^+ is free, the map $x \to x\gamma$ from A into A^+ can be extended uniquely to a morphism $\overline{\gamma} : A^+ \to A^+$, where

$$(x_1x_2 \dots x_n)\overline{\gamma} = (x_1\gamma)(x_2\gamma)\dots(x_n\gamma) \quad (x_1x_2\dots x_n \in A^+).$$

Now, for every $w = x_1x_2\dots x_n$ in A^+,

$$\begin{aligned} w\theta &= (x_1\theta)(x_2\theta)\dots(x_n\theta) = (x_1\gamma\rho^\natural)(x_2\gamma\rho^\natural)\dots(x_n\gamma\rho^\natural) \\ &= [(x_1\gamma)(x_2\gamma)\dots(x_n\gamma)]\rho^\natural = w\overline{\gamma}\rho^\natural. \end{aligned}$$

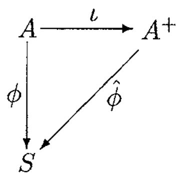
Since ρ is fully invariant, $(w_1\overline{\gamma}, w_2\overline{\gamma}) \in \rho$. Hence

$$w_1\theta = w_1\overline{\gamma}\rho^\natural = w_2\overline{\gamma}\rho^\natural = w_2\theta,$$

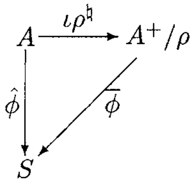
as required. We have shown that $A^+/\rho \in [\mathbf{R}]$, and so $\mathbf{I}(A^+/\rho) \supseteq \mathbf{R}^v$.

To show equality, suppose that $(w_1, w_2) \in \mathbf{I}(A^+/\rho)$. Then $w_1\theta = w_2\theta$ for every morphism θ from A^+ into A^+/ρ. This holds in particular for the morphism $\rho^\natural$, and from this it follows that $(w_1, w_2) \in \rho = \mathbf{R}^v$.

To show that A^+/ρ is the $[\mathbf{R}]$-free semigroup on A, suppose that $S \in [\mathbf{R}]$, and consider an arbitrary map $\phi : A \to S$. There is a unique extension of this map to a morphism $\hat{\phi} : A^+ \to S$, that is, if we denote the inclusion map from A into A^+ by ι, we have a commutative diagram



Now $\mathbf{I}(S) \supseteq \rho$, and so $w_1\hat{\phi} = w_2\hat{\phi}$ for every (w_1, w_2) in ρ. It follows that the map $\hat{\phi}$ factors through A^+/ρ, in the sense that there is a unique morphism $\overline{\phi} : A^+/\rho \to S$ such that $\hat{\phi} = \rho^\natural\overline{\phi}$. It follows that the diagram



commutes, and this is precisely the property we require. □

In a similar way we can obtain the $[\mathbf{R}]$-free semigroup on a more general infinite set A of generators, and also on a finite set A, provided A has enough symbols to spell out the words in $\mathbf{R}$. The free band B_A studied in the previous section is an example of the phenomenon described in the proposition, with $\mathbf{R} = \{(x^2, x)\}$ and $\mathbf{R}^v = \beta$.

Proposition 4.6.3 *Let A be a countable set, and let $\mathbf{R}_1$, $\mathbf{R}_2$ be subsets of $A^+ \times A^+$ determining varieties $[\mathbf{R}_1]$, $[\mathbf{R}_2]$, respectively. Then*

$$[\mathbf{R}_1] \subseteq [\mathbf{R}_2] \quad \textit{if and only if} \quad \mathbf{R}_1^v \supseteq \mathbf{R}_2^v.$$

Proof We use Propositions 4.6.1 and 4.6.2. If $[\mathbf{R}_1] \subseteq [\mathbf{R}_2]$, then $A^+/\mathbf{R}_1^v \in [\mathbf{R}_2]$, and so

$$\mathbf{R}_1^v = \mathbf{I}(A^+/\mathbf{R}_1^v) \supseteq \mathbf{R}_2^v.$$

Conversely, if $\mathbf{R}_1^v \supseteq \mathbf{R}_2^v$, then

$$S \in [\mathbf{R}_1] \;\Rightarrow\; \mathbf{I}(S) \supseteq \mathbf{R}_1^v \;\Rightarrow\; \mathbf{I}(S) \supseteq \mathbf{R}_2^v \;\Rightarrow\; S \in [\mathbf{R}_2].$$

Thus $[\mathbf{R}_1] \subseteq [\mathbf{R}_2]$. □

It is important to realise that two different sets $\mathbf{R}_1$ and $\mathbf{R}_2$ of identical relations may determine the same variety. We have already encountered such a situation in Theorem 1.1.3, which we can interpret as saying that the variety of rectangular bands can be described either as $[xyx = x]$ or as $[x^2 = x,\ xyz = xz]$. What we can deduce from Proposition 4.6.3 is that

Corollary 4.6.4 $[\mathbf{R}_1] = [\mathbf{R}_2]$ *if and only if* $\mathbf{R}_1^v = \mathbf{R}_2^v$. □

Suppose now that we have a family $\{\mathcal{V}_i : i \in I\}$ of varieties of semigroups, where $\mathcal{V}_i = [\mathbf{R}_i]$ $(i \in I)$. Then $\mathbf{R} = \bigcap\{\mathbf{R}_i^v : i \in I\}$ is a fully invariant congruence on A^+, and the correspondence established in Proposition 4.6.3 implies that $[\mathbf{R}] = \bigvee\{\mathcal{V}_i : i \in I\}$. In a sense we have solved the problem of finding identical relations defining $\mathcal{U} = \bigvee\{\mathcal{V}_i : i \in I\}$, but in the case where the sets $\mathbf{R}_i$ are all finite we might well be interested in finding a *finite* set of relations defining $\mathcal{U}$. There is no general method of doing this, and indeed it may not be possible. (See Taylor (1979).)

With respect to the operations $\cap$ and $\vee$, the set $\mathbf{V}(\mathcal{S})$ of varieties of semigroups becomes a complete lattice with greatest element the variety $[x = x]$ of *all* semigroups and least element the variety $[x = y]$ of *trivial* semigroups. The set $\mathbf{V}(\mathcal{B})$ of varieties of bands, consisting of all varieties of semigroups contained in the variety $\mathcal{B}$, is a sublattice of $\mathbf{V}(\mathcal{S})$. Any logical difficulty involved here in talking of a set of classes is only apparent, for we can if we wish identify each variety with a fully invariant congruence on a fixed A^+, and talk instead of the lattice $\mathbf{C}$ of such congruences on A^+ and the sublattice of $\mathbf{C}$ consisting of all fully invariant congruences containing β.

By an *atom* in a lattice $(L, \wedge, \vee)$ with least element 0 we mean a minimal element a in the set $L \setminus \{0\}$. We now proceed to identify the atoms of the lattice $\mathbf{V}(\mathcal{B})$, and to this end we examine three very specific fully invariant congruences on A^+, namely

$$\sigma = \{(xy, yx),\ (x^2, x)\}^v, \quad \lambda = \{(xy, x)\}^v, \quad \rho = \{(xy, y)\}^v. \qquad (4.6.1)$$

These correspond to the varieties $\mathcal{SL}$ (semilattices), $\mathcal{LZ}$ (left zero semigroups) and $\mathcal{RZ}$ (right zero semigroups), respectively. In each case we have a useful alternative description. As before $C(w)$ denotes the *content* of w, the set of letters in A appearing in w; also, $h(w)$ and $t(w)$, the *head* and the *tail* of w, respectively denote the first and the last letter in w.

Proposition 4.6.5 *Let σ, λ and ρ be the fully invariant congruences on A^+ given by* (4.6.1). *Then, for all w, z in A^+,*

(1) $(w, z) \in \sigma$ *if and only if* $C(w) = C(z)$;
(2) $(w, z) \in \lambda$ *if and only if* $h(w) = h(z)$;
(3) $(w, z) \in \rho$ *if and only if* $t(w) = t(z)$.

Proof (1) Denote the relation $\{(w, z) : C(w) = C(z)\}$ on A^+ by κ, and let

$$\mathbf{S} = \{(uv, vu), (u^2, u) : u, v \in A^+\}.$$

Certainly κ is a congruence on A^+. It is even fully invariant, since for every endomorphism α on A^+, and every (w, z) in κ,

$$C(w\alpha) = \bigcup\{C(x\alpha) : x \in C(w)\} = \bigcup\{C(x\alpha) : x \in C(z)\} = C(z\alpha).$$

Since $\mathbf{S} \subseteq \kappa$, it is clear that $\sigma \subseteq \kappa$. Conversely, it is clear that if $C(w) = C(z) = \{x_1, x_2, \ldots, x_n\}$, then each of w and z can be connected to $x_1x_2\ldots x_n$ by a sequence of elementary $\mathbf{S}$-transitions, and so $(w, z) \in \sigma$.

(2) The approach is very similar. Denote the relation $\{(w, z) : h(w) = h(z)\}$ by η, and let

$$\mathbf{E} = \bigcup\{(uv, u) : u, v \in A^+\}.$$

Then η is a congruence, and is fully invariant since, for every endomorphism α of A^+ and every (w, z) in η,

$$h(w\alpha) = h\big(h(w)\alpha\big) = h\big(h(z)\alpha\big) = h(z\alpha).$$

Since $\mathbf{E} \subseteq \eta$, it follows that $\lambda \subseteq \eta$. Conversely, if $h(w) = h(z) = x$, then, modulo λ,

$$w \equiv x \equiv z.$$

Thus $\eta = \lambda$, as required.

(3) The proof is the exact left-right dual of the previous case. □

Theorem 4.6.6 *The set of atoms in the lattice $\mathbf{V}(\mathcal{B})$ of varieties of bands is $\{\mathcal{SL}, \mathcal{LZ}, \mathcal{RZ}\}$.*

Proof To show that $\mathcal{SL}$, $\mathcal{LZ}$ and $\mathcal{RZ}$ are atoms we must show that the corresponding congruences σ, λ and ρ are maximal among the non-universal fully invariant congruences on A^+. So let us suppose that τ is a fully invariant congruence on A^+ properly containing σ, so that there exists (u, v) in τ such that $C(u) \neq C(v)$. Without loss of generality, suppose that there exists x in A such that $x \in C(u)$, $x \notin C(v)$. Since $\sigma \subseteq \tau$, we may assume that x appears just once in u, and at the beginning. Let α be an endomorphism of A^+ mapping x to itself and all other elements of A to some y $(\neq x)$ in A. Then $u\alpha = xy^k$, $v\alpha = y^l$, and so (modulo τ)

$$xy \equiv xy^k \equiv y^l \equiv y.$$

By the fully invariant property and from the assumption that $\sigma \subseteq \tau$ it now follows that, for all w_1, w_2 in A^+,

$$w_1 \equiv w_2w_1 \equiv w_1w_2 \equiv w_2 \pmod{\tau}.$$

Thus $\tau = A^+ \times A^+$, and so σ is maximal, as required.

Next, suppose that μ is a fully invariant congruence on A^+ properly containing λ. Then there exist elements $u = xu'$, $v = yv'$ $(x, y \in A;\ u', v' \in A^*)$ such that $x \neq y$ and $(u, v) \in \mu$. From $\mu \supseteq \lambda$ we then deduce that

$$x \equiv xu' \equiv yv' \equiv y \pmod{\mu},$$

and from the fully invariant property it then follows that $(w_1, w_2) \in \mu$ for all w_1, w_2 in A^+. Thus λ is maximal. The proof that ρ is maximal is effectively identical.

We have now established that $\mathcal{SL}$, $\mathcal{LZ}$ and $\mathcal{RZ}$ are atoms within the lattice $\mathbf{V}(\mathcal{B})$. It remains to show that these are the only atoms in the lattice. So suppose, by way of contradiction, that $\mathcal{V}$ is an atom in $\mathbf{V}(\mathcal{B})$ and that $\mathcal{V}$ is distinct from each of $\mathcal{SL}$, LZ and $\mathcal{RZ}$. Then

$$\mathcal{V} \cap \mathcal{SL} = \mathcal{V} \cap \mathcal{LZ} = \mathcal{V} \cap \mathcal{RZ} = \mathcal{T}, \tag{4.6.2}$$

where $\mathcal{T}$ is the class of trivial bands. Let B be a non-trivial band in $\mathcal{V}$. If B is not a rectangular band, then by Theorem 4.4.1 it has a non-trivial semilattice morphic image Y. Then $Y \in \mathcal{V} \cap \mathcal{SL}$, and we have a contradiction. Hence B is a rectangular band, and $\mathcal{V}$ consists entirely of rectangular bands. Now $B \notin \mathcal{LZ}$ and $B \notin \mathcal{RZ}$. However, by Theorem 1.1.3, B is then expressible non-trivially as a direct product of a left zero semigroup L and a right zero semigroup R. Since L and R are then morphic images of B, both L and R belong to $\mathcal{V}$, and we have a contradiction to (4.6.2). □

It is natural now to seek to identify the varieties that constitute the sublattice of $\mathbf{V}(\mathcal{B})$ generated by the atoms $\mathcal{SL}$, $\mathcal{LZ}$ and $\mathcal{RZ}$. Up to a point it is possible to do this with informal, largely verbal arguments of the following sort. The fully invariant congruence σ on A^+ corresponding to $\mathcal{SL}$ consists of all pairs (w_1, w_2) for which w_1 and w_2 have the same content. Similarly, the fully invariant congruence λ corresponding to $\mathcal{LZ}$ consists of all pairs (w_1, w_2) such that w_1 and w_2 have the same first letter. Hence $\sigma \cap \lambda$ consists of all pairs (w_1, w_2) having the same content *and* the same first letter. 'Clearly' $\sigma \cap \lambda = \{(x^2, x), (xyz, xzy)\}^v$, and so $\mathcal{SL} \vee \mathcal{LZ} = [x^2 = x,\ xyz = xzy]$, the variety of *left normal* bands.

There is nothing fundamentally wrong with this argument, though the use of 'clearly' involves the common, but ultimately disreputable, mathematical practice of handwaving. More formal arguments are available, and both the methods we shall use and the results that we shall obtain are sufficiently interesting to justify the more lengthy procedure.

First, we shall require an alternative version of Birkhoff's Theorem (Theorem 4.3.1), involving the notion of a *subdirect product.* Again, this is a concept belonging to general algebra, but (just as we proved Birkhoff's Theorem for $(2,1)$-algebras) we shall consider the alternative version only for semigroups.

There are in fact several ways of approaching the idea of a subdirect product. First, let P be the direct product of a family $\{S_i : i \in I\}$ of semigroups, consisting of maps $\phi : I \to \bigcup\{S_i : i \in I\}$ such that $i\phi \in S_i$ for each i in I, with the 'componentwise' multiplication

$$i(\phi\psi) = (i\phi)(i\psi) \quad (i \in I,\ \phi, \psi \in P).$$

For each i in I there is a *projection* morphism π_i from P onto S_i, given by

$$\phi\pi_i = i\phi \quad (\phi \in P).$$

A subsemigroup Q of P is called a *subdirect product* of the family $\{S_i : i \in I\}$ if $Q\pi_i = S_i$ for every i in I.

Examples include P itself and (if each S_i has an identity element 1) the so-called *weak* direct product consisting of those ϕ in P for which all but finitely many of the elements $i\phi$ are equal to 1. If all of the semigroups S_i are equal to a single semigroup S, and if, for all s in S, we define $\langle s \rangle$ in P by the rule that

$$i\langle s \rangle = s \quad (i \in I), \tag{4.6.3}$$

then the *diagonal*

$$D = \{\langle s \rangle : s \in S\}$$

is easily seen to be a subdirect product.

An alternative approach to subdirect products is given by

Theorem 4.6.7 *Let $\{S_i : i \in I\}$ be a family of semigroups, and let Q be a semigroup with the property that, for every i in I, there exists a morphism λ_i from Q onto S_i. If, for all x, y in Q,*

$$[(\forall i \in I)\ x\lambda_i = y\lambda_i] \ \Rightarrow\ x = y, \tag{4.6.4}$$

then S is isomorphic to a subdirect product of the family $\{S_i : i \in I\}$.

Conversely, if Q is a subdirect product of $\{S_i : i \in I\}$, then for each i in I there is a morphism λ_i from Q onto S_i, and the implication (4.6.4) *holds.*

Proof Let P be the direct product of the family $\{S_i : i \in I\}$, and consider the map $\lambda : Q \to P$ given by

$$i(q\lambda) = q\lambda_i \quad (i \in I,\ q \in Q).$$

Then λ is a morphism, since, for all q_1, q_2 in Q and all i in I,

$$i[(q_1q_2)\lambda] = (q_1q_2)\lambda_i = (q_1\lambda_i)(q_2\lambda_i) = [i(q_1\lambda)][i(q_2\lambda)] = i[(q_1\lambda)(q_2\lambda)].$$

Also, λ is one-one, since, for all x, y in Q,

$$x\lambda = y\lambda \ \Rightarrow\ (\forall i \in I)\ x\lambda_i = y\lambda_i \ \Rightarrow\ x = y.$$

Thus Q is isomorphic to the subsemigroup $Q\lambda$ of P. We show finally that $Q\lambda$ is a subdirect product. Since each λ_i maps *onto* S_i, we can be sure that for every s in S_i there exists q in Q such that $q\lambda_i = s$. Hence

$$(q\lambda)\pi_i = i(q\lambda) = q\lambda_i = s,$$

and so $(Q\lambda)\pi_i = S_i$, as required.

Conversely, suppose that Q is a subdirect product of the family $\{S_i : i \in I\}$. Then all the projection maps $\pi_i|_Q : Q \to S_i$ are onto. Moreover, for all ξ, η in Q,

$$[(\forall i \in I)\ \xi\pi_i = \eta\pi_i] \ \Rightarrow\ [(\forall i \in I)\ i\xi = i\eta] \ \Rightarrow\ \xi = \eta,$$

and so (4.6.4) is satisfied. □

From the correspondence between morphisms and congruences given by Theorem 1.5.2 we have the following alternative version:

Theorem 4.6.8 *If Q is a semigroup having a family of congruences $\{\rho_i : i \in I\}$ such that $\bigcap\{\rho_i : i \in I\} = 1_Q$, then Q is isomorphic to a subdirect product of the family $\{Q/\rho_i : i \in I\}$.* □

All this is leading to the following result:

Theorem 4.6.9 *A non-empty class $\mathcal{V}$ of semigroups is a variety if and only if*

(1) *every morphic image of a semigroup in $\mathcal{V}$ is in $\mathcal{V}$;*

(2) *every subdirect product of a family of semigroups in $\mathcal{V}$ is in $\mathcal{V}$.*

Proof It is clear that if $\mathcal{V}$ is a variety then both (1) and (2) hold.

Conversely, suppose that $\mathcal{V}$ is a non-empty class of semigroups satisfying (1) and (2). Then it is clear that $\mathcal{V}$ is closed under the taking of morphic images and direct products. To show that it is closed under the taking of subsemigroups, let $S \in \mathcal{V}$ and let U be a subsemigroup of S. Let I be a countable set, and let P be the direct product S^I of $|I|$ copies of S. Let Q be the set of 'almost constant' maps ϕ in S^I with the property that, for some u in U, $i\phi = u$ for all but finitely many i in I. That is,

$$\phi \in Q \ \text{ if and only if }\ (\exists u \in U)\ |\{i \in I : i\phi \neq u\}| < \infty.$$

Among the elements of Q are the constant maps $\langle u \rangle$ $(u \in U)$ given by

$$i\langle u \rangle = u \quad (i \in I),$$

and also the maps $\phi_{i,s,u}$ $(i \in I,\ s \in S,\ u \in U)$ given by

$$i\phi_{i,s,u} = s, \quad j\phi_{i,s,u} = u \quad (j \neq i).$$

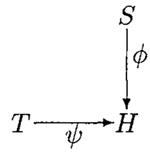
Since the projection map π_i sends $\phi_{i,s,u}$ to s, we have $Q\pi_i = S$ for each i, and so Q, being a subdirect product of copies of S, belongs to $\mathcal{V}$.

Now define a map $\Gamma : Q \to U$ by the rule that, for each ϕ in Q, $\phi\Gamma$ is the unique element u in U such that $\{i \in I : i\phi \neq u\}$ is finite. Then it is easy to verify that Γ is a morphism from Q onto U, and so $U \in \mathcal{V}$. □

The subdirect product is a very general construction, and we shall have reason to consider a special case. Let S and T be semigroups having a common morphic image H, and let ϕ, ψ be morphisms from S onto H and from T onto H respectively. The *spined product of* S *and* T *with respect to* H, ϕ *and* ψ is defined as

$$Y = \{(s,t) \in S \times T : s\phi = t\psi\}.$$

This is the standard terminology in semigroup theory, but the idea occurs elsewhere in mathematics under different titles. The semigroup Y is the *pullback*, or *fibre product*—see, for example, Schubert (1972)—of the diagram



It is a routine matter to verify that Y, as defined, is a subdirect product of S and T. It can usefully be thought of as a generalization of the diagonal, and reduces to the diagonal if $H = S = T$ and $\phi = \psi = 1$. At the other extreme, if H is the trivial semigroup, then Y becomes the direct product of S and T. For our purposes the important point to note is that if S and T belong to some variety $\mathcal{V}$ then so does their spined product.

We now show that the strong semilattice construction exemplified in Section 4.2 is relevant to the study of varieties. But first we must cope with a detail concerning the adjunction of a zero to a semigroup. Let us denote by ${}^0S$ the semigroup obtained from S by adjoining a zero element 0 *whether or not it already has a zero.* (This carries the assumption that if S does have a zero we are denoting it by a symbol other than 0.) In ${}^0S$ we always have the implication

$$xy = 0 \;\Rightarrow\; x = 0 \text{ or } y = 0.$$

Lemma 4.6.10 *Let* $\mathcal{V}$ *be a variety of semigroups containing the variety* $\mathcal{SL}$ *of semilattices. If* $S \in \mathcal{V}$, *then* ${}^0S \in \mathcal{V}$.

Proof The multiplicative semigroup $\{0,1\}$, being a semilattice, is in $\mathcal{V}$, and so $T = S \times \{0,1\} \in \mathcal{V}$. The subset $I = \{(x,0) : x \in S\}$ is an ideal of T, and so the Rees quotient T/I, being a morphic image of T, belongs to $\mathcal{V}$. But T/I is essentially $\{(s,1) : s \in S\} \cup \{0\}$, and so is isomorphic to ${}^0S$. $\square$

Proposition 4.6.11 *Let* $\mathcal{V}$ *be a variety of semigroups containing the variety of semilattices, and let* $S = \mathcal{S}(Y; S_\alpha; \phi_{\alpha,\beta})$ *be a strong semilattice of semigroups* S_α. *If each* S_α *is in the variety* $\mathcal{V}$, *then* $S \in \mathcal{V}$.

Proof Define $\psi_\alpha : S \to {}^0S_\alpha$ by the rule that

$$x\psi_\alpha = \begin{cases} x\phi_{\beta,\alpha} & \text{if } x \in S_\beta \text{ and } \beta \geq \alpha \\ 0 & \text{otherwise.} \end{cases}$$

To see that this is a morphism, suppose that x, y in S are such that $x \in S_\beta$, $y \in S_\gamma$. Then either $\beta\gamma \geq \alpha$ or $\beta\gamma \not\geq \alpha$. If $\beta\gamma \geq \alpha$ then $\beta \geq \beta\gamma \geq \alpha$ and $\gamma \geq \beta\gamma \geq \alpha$, and so

$$(xy)\psi_\alpha = [(x\phi_{\beta,\beta\gamma})(y\phi_{\gamma,\beta\gamma})]\phi_{\beta\gamma,\alpha} = (x\phi_{\beta,\alpha})(y\phi_{\gamma,\alpha}) = (x\psi_\alpha)(y\psi_\alpha).$$

If $\beta\gamma \not\geq \alpha$ then $(xy)\psi_\alpha = 0$. Also, either $\beta \not\geq \alpha$ or $\gamma \not\geq \alpha$, for $\beta \geq \alpha$ and $\gamma \geq \alpha$ would imply that $\beta\gamma \geq \alpha$, contrary to assumption. Hence one or other of $x\psi_\alpha$ and $y\psi_\alpha$ is equal to 0, and so $(xy)\psi_\alpha = (x\psi_\alpha)(y\psi_\alpha)$ in this case also.

Next, notice that $S\psi_\alpha \supseteq S_\alpha\psi_\alpha = S_\alpha$. In fact, unless $\beta \geq \alpha$ for every β in Y, we have $S\psi_\alpha = {}^0S_\alpha$. If there exists a unique minimum element ω of Y, then $S\psi_\omega = S_\omega$.

Suppose now that x in S_β and y in S_γ are such that $x\psi_\alpha = y\psi_\alpha$ for every α in Y. In particular $x\psi_\beta = x\phi_{\beta,\beta} = x \neq 0$, and so $y\psi_\beta \neq 0$. Hence $\gamma \geq \beta$. We can equally well show that $\beta \geq \gamma$, and so $\beta = \gamma$. Hence

$$x = x\psi_\beta = y\psi_\beta = y.$$

It now follows from Theorem 4.6.7 that S is a subdirect product of the semigroups ${}^0S_\alpha \quad (\alpha \in Y \setminus \{\omega\})$ and S_ω. Since $S_\omega \in \mathcal{V}$ and (by Lemma 4.6.10) ${}^0S_\alpha \in \mathcal{V}$ for all $\alpha \neq \omega$, we deduce that $S \in \mathcal{V}$. □

To the list of varieties in Section 4.3 we now add the following new members:

$\mathcal{LN}$: left normal bands : $xyz = xzy$,
$\mathcal{RN}$: right normal bands : $xyz = yxz$,
$\mathcal{N}$: normal bands : $xyzx = xzyx$,

where in each case the given identical relation characterizes the variety *within the variety of bands*. The relation $x^2 = x$ is taken as read.

We already know that $\mathcal{LZ} \subseteq \mathcal{RB}$, $\mathcal{RZ} \subseteq \mathcal{RB}$. Hence

$$\mathcal{LZ} \vee \mathcal{RZ} \subseteq \mathcal{RB}.$$

On the other hand, every rectangular band, by Theorem 1.1.3, is a direct product of a left zero semigroup L and a right zero semigroup R. Since

$$L \in \mathcal{LZ} \subseteq \mathcal{LZ} \vee \mathcal{RZ} \quad\text{and}\quad R \in \mathcal{RZ} \subseteq \mathcal{LZ} \vee \mathcal{RZ},$$

it follows that $L \times R \in \mathcal{LZ} \vee \mathcal{RZ}$. We deduce that

$$\mathcal{LZ} \vee \mathcal{RZ} = \mathcal{RB}.$$

In considering normal bands, it is useful to establish first the following alternative characterization:

Proposition 4.6.12 $\mathcal{N} = [x^2 = x,\ xyzt = xzyt]$.

Proof It is clear that the identical relation $xyzt = xzyt$ implies the identical relation $xyzx = xzyx$. What we require to prove is that $x^2 = x$ and

$xyzx = xzyx$ imply $xyzt = xzyt$. From the given identities we can deduce that

$$\begin{aligned} xyzt &= xyztxyzt = [xy(zt)x]yzt = xztyxyzt \\ &= x[z(tyx)yz]t = xzytyxzt \\ &= xzy[ty(xz)t] = xzytxzyt \\ &= xzyt. \end{aligned}$$

□

Now let B be an arbitrary band, and let us express it in the usual way as a semilattice Y of rectangular bands E_α ($\alpha \in Y$). Within any band there is a natural order relation—which we have already come across in Chapter 3 in the context of primitive idempotents—given by

$$y \leq x \text{ if and only if } xy = yx = y.$$

It is clear from this definition that if $x \leq y$, where $x \in E_\alpha$ and $y \in E_\beta$, then $\beta \leq \alpha$ in Y. It is also easy to see that if x is an arbitrary element of E_α and $\beta \leq \alpha$, then, for every z in E_β, the element $y = xzx$ is in E_β, and is such that $y \leq x$. To put it formally, we have shown that

$$(\forall \alpha \in Y)(\forall x \in E_\alpha)(\forall \beta \leq \alpha)(\exists y \in E_\beta)\ y \leq x.$$

We now show that for B to be a normal band it is necessary and sufficient that the element y be *unique*:

Proposition 4.6.13 *With the above definitions, a band B is normal if and only if*

$$(\forall \alpha \in Y)(\forall x \in E_\alpha)(\forall \beta \leq \alpha)(\exists! y \in E_\beta)\ y \leq x.$$

Proof Suppose first that B is a normal band. On the face of it, the element $y = xzx$ described above depends on the choice of z in E_β. However, if t is another element of E_β, then $ztz = z$ and $tzt = t$, since E_β is a rectangular band, and

$$\begin{aligned} xzx &= xztzx = xtz^2x \quad \text{(by normality)} \\ &= xtzx = xt^2zx = xtztx = xtx. \end{aligned}$$

Thus the value of xzx is independent of the choice of z.

Suppose now that y and y' are elements of E_β such that $y \leq x$, $z \leq x$. Then from the argument of the last paragraph, $xyx = xy'x$. But we also have $xy = yx = y$, $xy' = y'x = y'$, and so

$$y = xyx = xy'x = y'.$$

Conversely, suppose that we have the given uniqueness condition, and let $a \in E_\alpha$, $b \in E_\beta$, $c \in E_\gamma$. Then $abca, acba \in E_\delta$, where $\delta = \alpha\beta\gamma$. It is clear that $abca \leq a$ and $acba \leq a$, and so by the uniqueness assumption $abca = acba$. □

Proposition 4.6.14 *A band B is normal if and only if it is a strong semilattice of rectangular bands.*

Proof Suppose first that $B = \mathcal{S}(Y; E_\alpha; \phi_{\alpha,\beta})$ is a strong semilattice of rectangular bands, and let a, b, c be arbitrary elements of B, with $a \in E_\alpha$, $b \in E_\beta$ and $c \in E_\gamma$. Then, writing δ for the product $\alpha\beta\gamma$ in Y, we have

$$\begin{aligned} abca &= (a\phi_{\alpha,\delta})(b\phi_{\beta,\delta})(c\phi_{\gamma,\delta})(a\phi_{\alpha,\delta}) \\ &= a\phi_{\alpha,\delta}, \quad \text{since } E_\delta \text{ is a rectangular band,} \end{aligned}$$

and similarly $acba = a\phi_{\alpha,\delta}$. Thus B is a normal band.

Conversely, suppose that B is normal. For each pair α, β in Y such that $\alpha \geq \beta$, and for each x in E_α, define $x\phi_{\alpha,\beta}$ to be the unique y in E_β such that $y \leq x$. Then it is immediate that $\phi_{\alpha,\alpha}$ is the identity map for every α, and that $\phi_{\alpha,\beta}\phi_{\beta,\gamma} = \phi_{\alpha,\gamma}$ whenever $\alpha \geq \beta \geq \gamma$. To show that each $\phi_{\alpha,\beta}$ is a morphism, consider x, y in E_α and let z be an arbitrary element of E_β. Then

$$\begin{aligned} (xy)\phi_{\alpha,\beta} &= xyzxy = xzxyy \quad \text{(by normality)} \\ &= x(z.zxy)y = x(zxy.z)y \quad \text{(again by normality)} \\ &= (xzx)(yzy) = (x\phi_{\alpha,\beta})(y\phi_{\alpha,\beta}). \end{aligned}$$

Finally, consider arbitrary elements α and β in Y, and let $x \in E_\alpha$, $y \in E_\beta$. Then $xy \in E_\gamma$, where $\gamma = \alpha\beta$, and, for an arbitrarily chosen z in E_γ,

$$\begin{aligned} xy &= xyzxy = xzxy^2 \quad \text{(by normality)} \\ &= x(z.zxy)y = x(zxy.z)y \quad \text{(again by normality)} \\ &= (xzx)(yzy) = (x\phi_{\alpha,\gamma})(y\phi_{\beta,\gamma}). \end{aligned}$$

That, is, B is a strong semilattice of rectangular bands. □

An alternative approach to the proof of Proposition 4.6.14, based on the structure theorem (Theorem 4.4.5) is outlined in Exercise 19.

Before proceeding with our discussion of varieties, it is convenient to mention a corollary to Proposition 4.6.14. It is evident that every left normal band B is normal, and is therefore a strong semilattice of rectangular bands E_α. But each of the rectangular bands E_α, being a sub-band of B, is also left normal, and from the identical relations $xyx = x$ and $xyz = xzy$ in E_α we easily deduce that

$$xy = xxy = xyx = x;$$

thus E_α is a left zero semigroup.

We have shown half of

Corollary 4.6.15 *A band is left normal if and only if it is a strong semilattice of left zero semigroups.*

Proof Let $B = \mathcal{S}(Y; E_\alpha; \phi_{\alpha,\beta})$ be a strong semilattice of left zero semigroups, and let a (in E_α), b (in E_β), c (in E_γ) be arbitrary elements of B. Then

$$abc = (a\phi_{\alpha,\delta})(b\phi_{\beta,\delta})(c\phi_{\gamma,\delta}) = a\phi_{\alpha,\delta},$$

where $\delta = \alpha\beta\gamma$, since E_δ is a left zero semigroup. Since we may equally well show that $acb = a\phi_{\alpha,\delta}$, we deduce that B is a left normal band. □

Dually, we have

Corollary 4.6.16 *A band is right normal if and only if it is a strong semilattice of right zero semigroups.* □

We record that in a strong semilattice $\mathcal{S}(Y; L_\alpha; \phi_{\alpha,\beta})$ of left zero semigroups, the multiplication is given by

$$l_\alpha l_\beta = l_\alpha \phi_{\alpha,\alpha\beta} \quad (l_\alpha \in L_\alpha,\ l_\beta \in L_\beta), \tag{4.6.5}$$

while in a strong semilattice $\mathcal{S}(Y; R_\alpha; \phi_{\alpha,\beta})$ of right zero semigroups, the multiplication is given by

$$r_\alpha r_\beta = r_\beta \phi_{\beta,\alpha\beta} \quad (r_\alpha \in R_\alpha,\ r_\beta \in R_\beta). \tag{4.6.6}$$

Notice now that from $\mathcal{SL} \subseteq \mathcal{N}$ and $\mathcal{RB} \subseteq \mathcal{N}$ we deduce that $\mathcal{SL} \vee \mathcal{RB} \subseteq \mathcal{N}$. Conversely, if $B \in \mathcal{N}$, then, by Proposition 4.6.14, B is a strong semilattice of semigroups in the variety $\mathcal{SL} \vee \mathcal{RB}$. Hence, by Proposition 4.6.11, $B \in \mathcal{SL} \vee \mathcal{RB}$. We have shown that

$$\mathcal{SL} \vee \mathcal{RB} = \mathcal{N}. \tag{4.6.7}$$

Similar arguments, based on Corollaries 4.6.15 and 4.6.16, show that

$$\mathcal{SL} \vee \mathcal{LZ} = \mathcal{LN}, \quad \mathcal{SL} \vee \mathcal{RZ} = \mathcal{RN}. \tag{4.6.8}$$

Since $\mathcal{LN} \subseteq \mathcal{N}$ and $\mathcal{RN} \subseteq \mathcal{N}$, it is clear that

$$\mathcal{LN} \vee \mathcal{RN} \subseteq \mathcal{N}.$$

To show that this is in fact an equality we require the following result:

Proposition 4.6.17 *Every normal band B is isomorphic to a spined product of a left normal and a right normal band.*

Proof Let $B = \mathcal{S}(Y; E_\alpha; \phi_{\alpha,\beta})$ be a normal band. Each of the rectangular bands E_α is a direct product $L_\alpha \times R_\alpha$ of a left zero semigroup L_α and a right zero semigroup R_α. Moreover, as shown in Corollary 4.4.3, the morphism $\phi_{\alpha,\beta}$ determines morphisms $\phi^l_{\alpha,\beta} : L_\alpha \to L_\beta$, $\phi^r_{\alpha,\beta} : R_\alpha \to R_\beta$ such that

$$(l_\alpha, r_\alpha)\phi_{\alpha,\beta} = (l_\alpha \phi^l_{\alpha,\beta}, r_\alpha \phi^r_{\alpha,\beta})$$

for every (l_α, r_α) in E_α.

Now, $L = \bigcup\{L_\alpha : \alpha \in Y\}$ becomes a strong semilattice of left zero semigroups if we define

$$l_\alpha \circ l_\beta = l_\alpha \phi^l_{\alpha,\alpha\beta},$$

and similarly $R = \bigcup\{R_\alpha : \alpha \in Y\}$ becomes a strong semilattice of right zero semigroups if we define

$$r_\alpha * r_\beta = r_\beta \phi^r_{\beta,\alpha\beta}.$$

(Compare these formulae with (4.6.5) and (4.6.6).) Thus L is a left normal band and R is a right normal band.

Certainly the semigroups L and R have a common morphic image, namely the semilattice Y. If ϕ and ψ are the obvious morphisms, given by

$$l_\alpha\phi = \alpha, \quad r_\alpha\psi = \alpha,$$

then the spined product of L and R, consisting of those pairs (l, r) for which $l\phi = r\psi$, coincides with

$$\bigcup\{L_\alpha \times R_\alpha : \alpha \in Y\}.$$

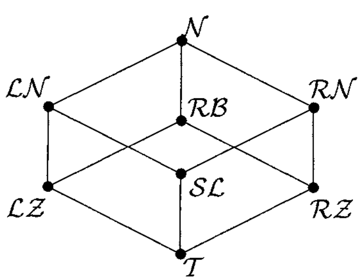
Moreover, the multiplication in the spined product is given by

$$\begin{aligned}(l_\alpha, r_\alpha)(l_\beta, r_\beta) &= (l_\alpha \circ l_\beta, r_\alpha * r_\beta) = (l_\alpha\phi^l_{\alpha,\alpha\beta}, r_\beta\phi^r_{\beta,\alpha\beta}) \\ &= (l_\alpha\phi^l_{\alpha,\alpha\beta}, r_\alpha\phi^r_{\alpha,\alpha\beta})(l_\beta\phi^l_{\beta,\alpha\beta}, r_\beta\phi^r_{\beta,\alpha\beta}) \\ &= [(l_\alpha, r_\alpha)\phi_{\alpha,\alpha\beta}][(l_\beta, r_\beta)\phi_{\beta,\alpha\beta}],\end{aligned}$$

and so coincides with the multiplication in the normal band B. □

It is now immediate that

$$\mathcal{LN} \vee \mathcal{RN} = \mathcal{N}.$$



We can summarize what we have learned about the sublattice of $\mathbf{V}(\mathcal{B})$ generated by the atoms by means of the diagram above. Not all of the statements implied by the diagram have been proved, but it is easy to fill in the gaps. For example, to show that $\mathcal{LN} \vee \mathcal{RB} = \mathcal{N}$, note that from $\mathcal{RB} \subseteq \mathcal{N}$ and $\mathcal{SL} \subseteq \mathcal{LN} \subseteq \mathcal{N}$ we can deduce that

$$\mathcal{N} = \mathcal{SL} \vee \mathcal{RB} \subseteq \mathcal{LN} \vee \mathcal{RB} \subseteq \mathcal{N}.$$

The various statements about intersections implied by the diagram are all obvious. For example, it is clear that $\mathcal{SL} \subseteq \mathcal{LN} \cap \mathcal{RN}$, but, conversely, if $B \in \mathcal{LN} \cap \mathcal{RN}$ then for all a, b in B

$$\begin{aligned} ab = aab &= aba \quad \text{(by left normality)} \\ &= baa \quad \text{(by right normality)} \\ &= ba, \end{aligned}$$

and so $B \in \mathcal{SL}$.

4.7 EXERCISES

1. Show that a semigroup S is completely regular if and only if for every a in S there exists x in S such that $axa = a$ and $ax = xa$.

2. Let S, T be completely regular semigroups.
   (a) Let $\phi : S \to T$ be a map satisfying
   $$(\forall a, b \in S)\ (ab)\phi = (a\phi)(b\phi).$$
   Show that ϕ also has the property
   $$(\forall a \in S)\ (a\phi)^{-1} = (a^{-1})\phi.$$
   (b) Show by an example that this deduction is invalid if S and T are merely I-semigroups.

3. Let S be a commutative semigroup, and let $a, b \in S$. Write $a \mid b$ (a *divides* b) if there exists x in S such that $ax = b$. Define η as the set of pairs (a, b) in $S \times S$ for which
   $$(\exists m \in \mathbf{N})\ a \mid b^m \ \text{ and } \ (\exists n \in \mathbf{N})\ b \mid a^n.$$
   (a) Show that η is a congruence on S.
   (b) Show that S/η is a semilattice.
   (c) Show that η is the least semilattice congruence on S, in the sense that if ρ is a congruence on S such that S/ρ is a semilattice, then $\eta \subseteq \rho$.

   A commutative semigroup A is called *archimedean* if
   $$(\forall a, b \in A)(\exists m, n \in \mathbf{N})\ a \mid b^m \ \text{ and } \ b \mid a^n.$$
   Show that each η-class of a commutative semigroup S is an archimedean subsemigroup of S. Deduce that every commutative semigroup can be expressed as a semilattice of archimedean semigroups.

4. Let S be a completely regular semigroup, expressed as a semilattice Y of completely simple semigroups S_α $(\alpha \in Y)$.

(a) Show that, if L is a left ideal of S_α, then

$$L \cup \left[\bigcup\{S_\beta : \beta < \alpha\}\right]$$

is a left ideal of S.

(b) Suppose now that $S = \mathcal{S}(Y; G_\alpha; \phi_{\alpha,\beta})$ is a strong semilattice of groups and that L is a left ideal of S. Show that

$$L \cap G_\alpha \neq \emptyset \Rightarrow \bigcup\{G_\beta : \beta \leq \alpha\} \subseteq L.$$

(c) Show that S is a Clifford semigroup if and only if it is regular and every one-sided ideal is a two-sided ideal.

5. Recall that a regular semigroup is called *E-unitary* if its idempotents form a unitary subsemigroup. Show that a strong semilattice $S = \mathcal{S}(Y; G_\alpha; \phi_{\alpha,\beta})$ of groups is E-unitary if and only if $\phi_{\alpha,\beta}$ is one-one for all $\alpha \geq \beta$ in Y.

6. Let G be a group with identity element e. Show that every left translation λ of G coincides with the inner left translation $\lambda_{\lambda e}$. Show similarly that all right translations of G are inner, and deduce that G is isomorphic to its translational hull $\Omega(G)$.

7. Let S be the two-element null semigroup. Show that the homomorphism $a \mapsto (\lambda_a, \rho_a)$ from S into $\Omega(S)$ is not one-one.

8. A semigroup S is called a *band of groups* if there exists a congruence β such that S/β is a band and each β-class is a group. Show that a semigroup S is a band of groups if and only if it is completely regular and $\mathcal{H}$ is a congruence.

9. A semigroup is called *orthodox* if it is regular and the idempotents form a subsemigroup. Let $\mathcal{OCR}$ denote the class of orthodox completely regular semigroups and let $\mathcal{OBG}$ be the class of orthodox bands of groups. Show that both these classes are varieties of completely regular semigroups determined, respectively, by the laws

$$xy = xyy^{-1}x^{-1}xy, \quad \text{and} \quad (xy)(xy)^{-1} = xx^{-1}yy^{-1}.$$

10. A semigroup S is called a *rectangular group* if it is isomorphic to the direct product of a group and a rectangular band.

(a) Show that a semigroup is completely simple and orthodox if and only if it is a rectangular group.

(b) Show that S is a rectangular group if and only if it is completely regular and satisfies the law

$$x^{-1}yy^{-1}x = x^{-1}x.$$

(c) Denote the class of rectangular groups by $\mathcal{RG}$ and the class of groups by $\mathcal{G}$. Show that, within the lattice of varieties of completely regular semigroups,

$$\mathcal{RG} = \mathcal{G} \vee \mathcal{RB}.$$

11. Denote the class of bands of groups by $\mathcal{BG}$.
  (a) Show that every semigroup in $\mathcal{BG}$ satisfies the law
  $$(x^2yz^2)(x^2yz^2)^{-1} = (xyz)(xyz)^{-1}.$$
  (b) Suppose conversely that S is a completely regular semigroup satisfying the above law. Let $(a, b) \in \mathcal{H}$, and let $c \in S$. By writing $ac = a^2(a^{-1}c^{-1})c^2$, show that $(ac, bc) \in \mathcal{H}$. Deduce from this result and its dual that S is a band of groups.

12. Show that a band B is

	rectangular	if and only if	$\mathcal{D} = B \times B$,
	left zero	if and only if	$\mathcal{L} = B \times B$,
	right zero	if and only if	$\mathcal{R} = B \times B$,
and	trivial	if and only if	$\mathcal{H} = B \times B$.

13. Let B be a band. Show that, for all a, b in B,
$$ab \,\mathcal{L}\, bab, \quad ab \,\mathcal{R}\, aba.$$

14. Give an example of a semigroup that satisfies the law $xyz = xz$ but is not a rectangular band.

15. The free band B_A is expressed in the usual way as a semilattice of rectangular bands. The underlying semilattice is the semilattice of finite non-empty subsets of A, and the 'product' of two subsets P and Q is their *union*. Thus, if $\leq$ is defined in the usual way in a semilattice,
$$Q \leq P \text{ if and only if } Q \supseteq P.$$
The rectangular bands are the $\mathcal{J}$-classes of B_A. There is one such $\mathcal{J}$-class U_P corresponding to each finite non-empty subset P of A, and U_P is the set of all elements of B_A with content P.

If $|P| > 1$, let
$$I_P = \bigcup \{U_{P \setminus \{a\}} : a \in P\},$$
and let $\alpha : U_P \to I_P \times I_P$ be defined by
$$x\alpha = \big(x(0), x(1)\big) \quad (x \in U_P).$$
Show that α is an isomorphism from U_P onto the rectangular band $I_P \times I_P$. Define $I_P = \{1\}$ if $|P| = 1$.

Now let $Q \supseteq P$, so that $Q \leq P$ in the semilattice. Show that, for each x in U_P, the maps $\phi_Q^x : I_Q \to I_Q$ and $\psi_Q^x : I_Q \to I_Q$—see the equations (4.4.5)—are given by

$$\phi_Q^x y(0) = (xy)(0), \quad y(1)\psi_Q^x = (yx)(1) \quad (y \in U_Q).$$

16. Let $A = \{a, b, c\}$. Then the $\mathcal{J}$-classes of the free band B_A are

$$U_{\{a\}},\ U_{\{b\}},\ U_{\{c\}},\ U_{\{a,b\}},\ U_{\{a,c\}},\ U_{\{b,c\}},\ U_A.$$

For simplicity of notation, denote each β-class by any chosen representative of it in A^+. Then

$$U_{\{a\}} = \{a\}, \quad U_{\{a,b\}} = \{ab, aba, ba, bab\},$$

while U_A may be identified with the 12×12 rectangular band $I_A \times I_A$, where

$$I_A = \{ab, aba, ba, bab, ac, aca, ca, cac, bc, bcb, cb, cbc\}.$$

(a) Determine the isomorphism α from $U_{\{a,b\}}$ onto the retangular band $\{a, b\} \times \{a, b\}$. (See the previous exercise for the definition of α.)

(b) Show that, in the notation of (4.4.5),

$$\phi^a_{\{a,b\}} = \begin{pmatrix} a & b \\ a & a \end{pmatrix}, \quad \phi^b_{\{a,b\}} = \begin{pmatrix} a & b \\ b & b \end{pmatrix},$$

and similarly compute $\psi^a_{\{a,b\}}$ and $\psi^b_{\{a,b\}}$.

(c) Show that

$$\phi_A^a = \begin{pmatrix} ab & aba & ba & bab & ac & aca & ca & cac & bc & bcb & cb & cbc \\ ab & aba & aba & ab & ac & aca & aca & ac & ab & ab & ac & ac \end{pmatrix},$$

and similarly compute ϕ_A^b, ψ_A^a, ψ_A^b, ϕ_A^{ab}, ψ_A^{ab}.

(d) Verify that ϕ_A^{abc} and ψ_A^{abc} are constant maps, and that

$$\langle \phi_A^{abc} \rangle = ab, \quad \langle \psi_A^{abc} \rangle = bc.$$

(e) Verify that $\phi_A^{ab} = \phi_A^a \phi_A^b$ (composed as left maps), and that $\psi_A^{ab} = \psi_A^a \psi_A^b$ (composed as right maps).

17. Let S be a non-trivial semigroup with the property that no proper subset of S generates S. Show that each $\mathcal{J}$-class of S is either a left zero semigroup or a right zero semigroup, and that the semilattice $S/\mathcal{J}$ is a chain.

18. Show that a band B is normal if and only if the natural order relation, defined by

$$x \leq y \ \text{ if and only if } \ xy = yx = x,$$

is compatible.

19. Let B be a band, expressed as a semilattice Y of rectangular bands $E_\alpha = I_\alpha \times \Lambda_\alpha$.
    (a) Show that B is normal if and only if the maps ϕ^a_β and ψ^a_β are constant maps for every α in Y, every a in E_α and every $\beta \leq \alpha$.
    (b) Suppose that B is normal. For all $\alpha \geq \beta$, define $\phi_{\alpha,\beta} : E_\alpha \to E_\beta$ $(= I_\beta \times \Lambda_\beta)$ by
$$a\phi_{\alpha,\beta} = \big(\langle \phi^a_\beta \rangle, \langle \psi^a_\beta \rangle\big) \quad (a \in E_\alpha).$$
    Show that $\phi_{\alpha,\beta}$ is a morphism.
    (c) Show that B is a strong semilattice $\mathcal{S}(Y; E_\alpha; \phi_{\alpha,\beta})$ of rectangular bands. That is to say, show that $\phi_{\alpha,\alpha} = 1_{E_\alpha}$ for every α, that $\phi_{\alpha,\beta}\phi_{\beta,\gamma} = \phi_{\alpha,\gamma}$ whenever $\alpha \geq \beta \geq \gamma$, and that, for all α, β in Y and all a in E_α, b in E_β,
$$ab = (a\phi_{\alpha,\gamma})(b\phi_{\beta,\gamma}),$$
    where $\gamma = \alpha\beta$.

20. Let $N_A = A^+/\nu$ be the (relatively) free normal band on the set A of generators, where $|A| \geq 3$. Show that, for $u, v \in A^+$, $(u, v) \in \nu$ if and only if u and v have the same initial letter, the same final letter, and the same content. Deduce that if $|A| = n$ then
$$|N_A| = \sum_{k=1}^{n} \binom{n}{k} k^2 = 2^{n-2}n(n+1).$$

21. Describe the free semilattice, the free left zero semigroup and the free left normal semigroup on a set A.

4.8 NOTES

The study of completely regular semigroups goes back to Clifford (1941), who proved the basic gross structure theorem (Theorem 4.1.3) and the fine structure theorem (Theorem 4.2.1) for what we have called Clifford semigroups. The term 'completely regular' was introduced by Petrich (1973), and is now standard. The idea of regarding a completely regular semigroup explicitly as a $(2, 1)$-algebra dates also from the 70s. It appears in Petrich (1975), and has been much developed, in the context of varieties of U-semigroups, by Petrich (1977, 1982), Hall and Jones (1980), Petrich and Reilly (1981a, 1982, 1983a,b,c, 1984a,b, 1988), Rasin (1979, 1981), Gerhard and Petrich (1983, 1985), Jones (1983) and Polák (1985, 1987, 1988).

The study of bands was begun in the 50s. D. McLean (1954) proved Theorem 4.4.1, and the general structure theorem (Theorem 4.4.5) is due to Petrich (1971), who also classified all varieties of bands determined by identities in at most three letters, and related the identities to properties of the maps ϕ^{α}_{β}, ψ^{α}_{β} described in Section 4.4. (See, for example, Exercise 19(a).) Petrich's structure theorem is in effect a specialization of an even more complicated result due to Lallement (1967) on the fine structure of completely regular semigroups. The results on free bands in Section 4.5 were obtained by Green and Rees (1952).

Varieties of bands have been studied by Yamada and Kimura (1958), who proved Proposition 4.6.14. Other references are Kimura (1957a,b, 1958a,b,c) and Yamada (1958). Complete descriptions of the lattice $\mathbf{V}(\mathcal{B})$ of all varieties of bands have been given independently by Biryukov (1970), Fennemore (1971a,b) and Gerhard (1970), and the smaller lattice of varieties of band monoids has been described by Wismath (1986). These descriptions have been used very effectively by Sezinando (1992a,b).

Useful survey articles on varieties of semigroups have been written by Evans (1971), and by Shevrin and Volkov (1985).

Exercise 3 is from Tamura and Kimura (1954). In connection with Exercise 4, we remark that many different characterizations of Clifford semigroups have been given by Lajos (1969, 1970a,b,c, 1971a,b, 1972a,b). Exercises 8, 9, 10 and 11 are extracted from a very thorough study of varieties of completely regular semigroups by Petrich (1975). Exercise 17 is from Giraldes and Howie (1985).

5
Inverse semigroups

In planning the chapter on inverse semigroups the main problem has been one of selection. As long ago as 1961, Clifford and Preston offered the opinion that inverse semigroups were the most promising class of semigroups for future study, and the intervening years have amply justified their forecast. It is inappropriate in a general introductory book to attempt a comprehensive updating of Petrich's (1984) very complete account. Rather I shall cover the basic ideas and attempt to explore one or two avenues in more detail.

Inverse semigroups were studied first by Vagner (1952, 1953) and independently by Preston (1954a,b,c). Vagner called them 'generalized groups', and for many years this nomenclature was standard in the Russian literature. In both cases the origin of the idea was the study of semigroups of partial one-one mappings of a set, and one of the earliest results was a representation theorem (analogous to Cayley's Theorem in group theory) to the effect that every inverse semigroup has a faithful representation as an inverse semigroup of partial one-one mappings. This important result has become known as the Vagner–Preston Theorem, and is proved as Theorem 5.1.7 below.

The theory of inverse semigroups has many features in common with the theory of groups, but there are some important differences. Among the new features is the natural partial order that exists in each inverse semigroup. This is discussed in Section 5.2.

Section 5.3 is devoted to congruences, and in particular to a description of congruences in terms of *trace* and *kernel*. Here group theory provides a strong motivation, but there are several new ideas. Two particular congruences, namely the minimum group congruence and the maximum idempotent-separating congruence, are discussed in detail. The existence of *fundamental* inverse semigroups (those having no non-trivial idempotent-separating congruences) is established.

Section 5.4 is devoted to a study of the *Munn semigroup* of a semilattice, and in Sections 5.5, 5.6 and 5.7 this idea is used to investigate inverse semigroups whose semilattices of idempotents are anti-uniform, and to describe fundamental simple and bisimple inverse semigroups. If the restriction 'fundamental' is dropped then the problem becomes vastly more complicated, and both in Section 5.6 and in Section 5.7 we confine ourselves to the case of ω-semigroups, where the semilattice is $\{e_0, e_1, e_2, \ldots\}$, with $e_0 > e_1 > e_2 > \cdots$

In Section 5.8 the study of representations of inverse semigroups by partial one-one mapping is resumed.

The theory of inverse semigroups entered a new phase with the work of McAlister (1974a,b) and others on E-unitary inverse semigroups. In Section 5.9 we give an account of some of this work, and in Section 5.10 we apply it to give a description of free inverse semigroups.

5.1 PRELIMINARIES

Recall that a completely regular semigroup is specified within the class of I-semigroups by the property

$$xx^{-1} = x^{-1}x,$$

and that a Clifford semigroup is similarly specified by the properties

$$xx^{-1} = x^{-1}x, \quad xx^{-1}yy^{-1} = yy^{-1}xx^{-1}.$$

If we take the second of these properties but not the first, we get what is called an *inverse semigroup.* To be precise, an *inverse semigroup* is an I-semigroup S such that, for all x, y in S,

$$xx^{-1}yy^{-1} = yy^{-1}xx^{-1}. \tag{5.1.1}$$

Then we have

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.*

Proof (1) $\Rightarrow$ (2). This will follow if we show that every idempotent in S can be expressed in the form xx^{-1}. Let e be an idempotent in S. Then the I-semigroup property ensures that there is an element e^{-1} in S such that $ee^{-1}e = e$ and $(e^{-1})^{-1} = e$. Hence

$$\begin{aligned} e^{-1} &= e^{-1}(e^{-1})^{-1}e^{-1} = e^{-1}ee^{-1} = e^{-1}e^2e^{-1} \\ &= (e^{-1}e)(ee^{-1}) = (ee^{-1})(e^{-1}e). \end{aligned}$$

It then follows that

$$e = ee^{-1}e = e(ee^{-1})(e^{-1}e)e = (e^2e^{-1})(e^{-1}e^2) = (ee^{-1})(e^{-1}e) = e^{-1},$$

and hence that $e = ee = ee^{-1}$.

(2) $\Rightarrow$ (3). Suppose that S is regular, and that its idempotents commute. From Proposition 2.3.2 we know that every $\mathcal{L}$-class contains at least one idempotent. If e, f are idempotents in a single $\mathcal{L}$-class, then, by Proposition 2.3.3, $ef = e$ and $fe = f$. Hence, since idempotents commute, $e = f$. A similar argument applies to $\mathcal{R}$-classes.

(3) $\Rightarrow$ (4). Let x', x'' be inverses of x in S. Then $xx', xx'' \in R_x$, and so $xx' = xx''$. Similarly, $x'x, x''x \in L_x$ and so $x'x = x''x$. Hence

$$x' = x'xx' = x'xx'' = x''xx'' = x'',$$

and we conclude that inverses are unique.

(4) $\Rightarrow$ (1). Denote the unique inverse of x by x^{-1}. Then certainly we have $xx^{-1}x$ for every x. We also have that x is the unique inverse of x^{-1}, and so $(x^{-1})^{-1} = x$ for every x. It remains to establish the identity (5.1.1). Denote xx^{-1} by e, yy^{-1} by f, and let z be the unique inverse of ef. It follows that

$$(ef)(fze)(ef) = ef^2ze^2f = efzef = ef,$$
$$(fze)(ef)(fze) = f(zefz)e = fze,$$

so that fze is also an inverse of ef. By uniqueness it follows that $z = fze$. Hence z is idempotent, since

$$(fze)^2 = f(zefz)e = fze.$$

But now it follows, since both ef and z are inverses of z, that $z = ef$, and so ef is idempotent. Thus ef is its own unique inverse. By a similar argument, fe is idempotent also. Since we now have

$$(ef)(fe)(ef) = (ef)^2 = ef \quad \text{and} \quad (fe)(ef)(fe) = (fe)^2 = fe,$$

both fe and ef are inverses of ef. Hence, by uniqueness, $ef = fe$. □

The set of idempotents of an inverse semigroup S is a commutative subsemigroup of S. It is indeed a semilattice, as defined in Section 1.3. We shall talk of the *semilattice of idempotents* of an inverse semigroup S, and shall consistently denote it by E_S, or simply by E. We have already observed that $e^{-1} = e$ for every e in E. Some further elementary properties of inverse semigroups are listed in the following theorem:

Proposition 5.1.2 *Let S be an inverse semigroup with semilattice E of idempotents. Then:*

(1) $(ab)^{-1} = b^{-1}a^{-1}$ *for every a, b in S;*

(2) *both aea^{-1} and $a^{-1}ea$ are idempotent for every a in A, e in E;*

(3) *$a \mathcal{L} b$ if and only if $a^{-1}a = b^{-1}b$; $a \mathcal{R} b$ if and only if $aa^{-1} = bb^{-1}$;*

(4) *if $e, f \in E$, then $e \mathcal{D} f$ if and only if there exists a in S such that $aa^{-1} = e$, $a^{-1}a = f$.*

Proof (1) Notice that, since $a^{-1}a$ and bb^{-1} are idempotents,

$$(ab)(b^{-1}a^{-1})(ab) = a(bb^{-1})(a^{-1}a)b = a(a^{-1}a)(bb^{-1})b = ab,$$
$$(b^{-1}a^{-1})(ab)(b^{-1}a^{-1}) = b^{-1}(a^{-1}a)(bb^{-1})a^{-1} = b^{-1}(bb^{-1})(a^{-1}a)a^{-1}$$
$$= b^{-1}a^{-1};$$

thus $b^{-1}a^{-1}$ is the unique inverse of ab:

$$b^{-1}a^{-1} = (ab)^{-1}.$$

(2) Again by the commuting of idempotents,

$$(aea^{-1})^2 = ae(a^{-1}a)ea^{-1} = aa^{-1}ae^2a^{-1} = aea^{-1},$$

and similarly $(a^{-1}ea)^2 = a^{-1}ea$.

Part (3) is simply a specialization of Proposition 2.4.1 to the case where inverses are unique. As for Part (4), if $e \mathcal{D} f$ then there exists a in S such that $e \mathcal{R} a$ and $a \mathcal{L} f$, that is, such that $aa^{-1} = e$, $a^{-1}a = f$. □

Part (1) generalizes to more than two factors in the obvious way:

Corollary 5.1.3 *Let $a_1, a_2, \ldots, a_n$ be elements of an inverse semigroup. Then*

$$(a_1a_2 \ldots a_n)^{-1} = a_n^{-1}a_{n-1}^{-1} \ldots a_1^{-1}.$$ □

In particular we have $(a^n)^{-1} = (a^{-1})^n$ for every element a in an inverse semigroup, and so the notation a^{-n} can be used unambiguously. On the other hand, it is important to note that the index law $a^pa^q = a^{p+q}$ cannot be assumed for all p, q in $\mathbf{Z}$.

Usually the best way to establish that a semigroup is an inverse semigroup is to show that it is regular and that idempotents commute. For example, we have

Theorem 5.1.4 *Let S be an inverse semigroup, let T be a semigroup and let ϕ be a semigroup morphism from S onto T. Then T is an inverse semigroup. Moreover, ϕ is an inverse semigroup morphism.*

Proof By a *semigroup morphism* we mean a map that respects the binary operation of multiplication. An *inverse semigroup morphism* has the additional property that $s^{-1}\phi = (s\phi)^{-1}$. It is part of the import of our theorem that there is in fact no distinction between the two kinds of morphism.

Every t in T is expressible (not necessarily uniquely) as $s\phi$, where $s \in S$. If s^{-1} is the inverse of s in S, then

$$(s\phi)\big((s^{-1})\phi\big)(s\phi) = (ss^{-1}s)\phi = s\phi,$$
$$\big((s^{-1})\phi\big)(s\phi)\big((s^{-1})\phi\big) = (s^{-1}ss^{-1})\phi = (s^{-1})\phi,$$

and so $(s^{-1})\phi$ is an inverse of $s\phi$ in T. Thus T is regular. Also, if g, h are idempotents in T then by Lemma 2.4.4 there exist idempotents e, f in S such that $e\phi = g$, $f\phi = h$. Hence

$$gh = (e\phi)(f\phi) = (ef)\phi = (fe)\phi = (f\phi)(e\phi) = hg,$$

and so T is an inverse semigroup as required. Notice that we have also shown that

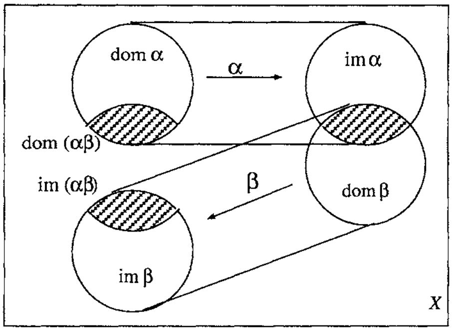
$$(s\phi)^{-1} = (s^{-1})\phi$$

for each s in S, and so ϕ is indeed an inverse semigroup morphism. □

We have seen that there is an intimate relation between morphisms and congruences. The result just proved implies that, for every semigroup congruence ρ on an inverse semigroup S, the quotient S/ρ is again an inverse semigroup. It implies moreover that ρ is an I-semigroup congruence, in the sense that, for all s, t in S,

$$(s,t) \in \rho \Rightarrow (s^{-1}, t^{-1}) \in \rho. \tag{5.1.2}$$

So far we have encountered only groups, semilattices and Clifford semigroups as examples of inverse semigroups. We find a more representative example in the so-called *symmetric inverse semigroup*. Given a non-empty set X, we define $\mathcal{I}_X$ to consist of all partial one-one maps of X. This becomes a semigroup under the standard operation $\circ$ of composition of relations : if $\alpha, \beta \in \mathcal{I}_X$, then $(x,y) \in \alpha \circ \beta$ if and only if there exists z in X such that $(x,z) \in \alpha$ and $(z,y) \in \beta$.



Thus $z = x\alpha$ and $y = z\beta$, and so $y = (x\alpha)\beta$. It follows that, for all $x_1, x_2 \in \text{dom}(\alpha \circ \beta)$,

$$\begin{aligned} x_1(\alpha \circ \beta) = x_2(\alpha \circ \beta) &\Rightarrow (x_1\alpha)\beta = (x_2\alpha)\beta \\ &\Rightarrow x_1\alpha = x_2\alpha \Rightarrow x_1 = x_2. \end{aligned}$$

Notice also that $x \in \text{dom}(\alpha\beta)$ if and only if there exist z and y such that $(x,z) \in \alpha$ and $(z,y) \in \beta$. Thus $z \in \text{im}\,\alpha \cap \text{dom}\,\beta$, and so

$$\text{dom}(\alpha\beta) = (\text{im}\,\alpha \cap \text{dom}\,\beta)\alpha^{-1}, \quad \text{im}(\alpha\beta) = (\text{im}\,\alpha \cap \text{dom}\,\beta)\beta.$$

Then we have the following theorem:

Theorem 5.1.5 *$\mathcal{I}_X$ is an inverse semigroup.*

Proof We have seen that $\mathcal{I}_X$ is closed under the associative operation $\circ$. From now on we shall write simply $\alpha\beta$ rather than $\alpha \circ \beta$. Each α in $\mathcal{I}_X$ is a bijection from $\operatorname{dom}\alpha$ onto $\operatorname{im}\alpha$, and so there is an inverse map α^{-1}, also an element of $\mathcal{I}_X$, such that

$$\operatorname{dom}(\alpha^{-1}) = \operatorname{im}\alpha, \quad \operatorname{im}(\alpha^{-1}) = \operatorname{dom}\alpha,$$
$$\alpha\alpha^{-1} = 1_{\operatorname{dom}\alpha}, \quad \alpha^{-1}\alpha = 1_{\operatorname{im}\alpha}.$$

Certainly $\alpha\alpha^{-1}\alpha = \alpha$ and $\alpha^{-1}\alpha\alpha^{-1} = \alpha^{-1}$, and so $\mathcal{I}_X$ is regular. We complete the proof by showing that idempotents commute, which of course requires us to be able to identify idempotents. If α is idempotent then

$$\operatorname{dom}(\alpha^2) = (\operatorname{dom}\alpha \cap \operatorname{im}\alpha)\alpha^{-1} = \operatorname{dom}\alpha = (\operatorname{im}\alpha)\alpha^{-1}.$$

Since α^{-1} is one-one it follows that $\operatorname{dom}\alpha \cap \operatorname{im}\alpha = \operatorname{im}\alpha$, i.e., that $\operatorname{im}\alpha \subseteq \operatorname{dom}\alpha$. Similarly,

$$\operatorname{im}(\alpha^2) = (\operatorname{dom}\alpha \cap \operatorname{im}\alpha)\alpha = \operatorname{im}\alpha = (\operatorname{dom}\alpha)\alpha,$$

and from this we deduce that $\operatorname{dom}\alpha \subseteq \operatorname{im}\alpha$. Thus $\operatorname{dom}\alpha = \operatorname{im}\alpha = A$ (say), and $x\alpha^2 = x\alpha$ for each x in A. Since α is one-one, it follows that $x\alpha = x$ for all x in A; thus $\alpha = 1_A$, the identity map of the subset A.

It now easily follows that

$$1_A 1_B = 1_{A\cap B} \quad (A, B \subseteq X), \tag{5.1.3}$$

and it is then automatic that $1_A 1_B = 1_B 1_A$. Thus $\mathcal{I}_X$ is an inverse semigroup, and the inverse α^{-1} exhibited above is in fact the unique inverse of α in the inverse semigroup. □

We call $\mathcal{I}_X$ the *symmetric inverse semigroup* on the set X. We shall see that in a real sense it is the appropriate analogue in inverse semigroup theory of the symmetric group in group theory and the full transformation semigroup in semigroup theory. If $X = \{1, 2\}$ it consists of the maps

$$I = \begin{pmatrix} 1\,2 \\ 1\,2 \end{pmatrix}, \quad A = \begin{pmatrix} 1\,2 \\ 2\,1 \end{pmatrix},$$
$$E = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad F = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad Y = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

together with the empty map, which we denote by 0. The multiplication is easily computed.

If $|X| = 3$ then $\mathcal{I}_X$ contains 34 elements. For a general formula for $|\mathcal{I}_X|$, see Exercise 3.

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 (Theorem 1.1.2), so every inverse semigroup can be embedded up to isomorphism in a symmetric inverse semigroup. The inverse semigroup result is proved in a way that is certainly similar in spirit to the earlier proofs, but is a little more complicated in detail. Before stating the theorem, it will be convenient to record a fairly simple lemma. In fact, for future use, we shall prove something a little more general than our immediate requirements warrant.

Lemma 5.1.6 *Let V be a semigroup containing an inverse semigroup S as a subsemigroup. Then:*

(1) *for every pair e, f of idempotents in S, $Ve = Vf \Rightarrow e = f$, and $eV = fV \Rightarrow e = f$;*

(2) *for every pair e, f of idempotents in S, $Ve \cap Vf = Vef$, and $eV \cap fV = efV$;*

(3) *for every a in S, $Vaa^{-1} = Va^{-1}$, $Va^{-1}a = Va$, $aa^{-1}V = aV$, $a^{-1}aV = a^{-1}V$.*

Proof (1) If $Ve = Vf$, then $e = ee = xf$ for some x in V, and so $ef = xf^2 = xf = e$. Similarly $fe = f$, and so $e = f$, since idempotents commute. The dual result is proved in a similar way.

(2) Certainly $Vef \subseteq Vf$, $Vef = Vfe \subseteq Ve$. Thus $Vef \subseteq Ve \cap Vf$. Also, if $z = xe = yf \in Ve \cap Vf$, then $zef = xe^2f = xef = zf = yf^2 = yf = z$, and so $z = zef \in Vef$. The dual result follows in a similar way

(3) We have

$$Vaa^{-1} \subseteq Va^{-1} = Va^{-1}aa^{-1} \subseteq Vaa^{-1},$$
$$Va^{-1}a \subseteq Va = Vaa^{-1}a \subseteq Va^{-1}a,$$

and the required results follow. Once again, the proof of the dual result is similar. □

Now we show the analogue of Cayley's Theorem, a result due to Vagner (1952) and (independently) to Preston (1954c):

Theorem 5.1.7 *Let S be an inverse semigroup. Then there exists a symmetric inverse semigroup $\mathcal{I}_X$ and a monomorphism ϕ from S into $\mathcal{I}_X$.*

Proof Let $X = S$, and for each a in S define the partial map ρ_a to have domain $Sa^{-1} = Saa^{-1}$ and to be such that

$$x\rho_a = xa \quad (x \in Sa^{-1}).$$

The image of the map ρ_a is $Sa^{-1}a = Sa$, and the map is in fact one-one, since if $x = sa^{-1}$ and $y = ta^{-1}$ are in $Sa^{-1} = \operatorname{dom} \rho_a$, then

$$\begin{aligned} x\rho_a = y\rho_a &\Rightarrow sa^{-1}a = ta^{-1}a \\ &\Rightarrow x = sa^{-1} = sa^{-1}aa^{-1} = ta^{-1}aa^{-1} = ta^{-1} = y. \end{aligned}$$

Thus $\rho_a \in \mathcal{I}_S$.

Define the map $\phi : S \to \mathcal{I}_S$ by the rule that $a\phi = \rho_a$ for all a in S. Certainly ϕ is one-one. For suppose that $a\phi = b\phi$. Then in the first place it follows that the domains of ρ_a and ρ_b are equal, and so $Saa^{-1} = Sbb^{-1}$. By Lemma 5.1.6 we deduce that $aa^{-1} = bb^{-1}$, and hence

$$a = aa^{-1}a = (aa^{-1})\rho_a = (aa^{-1})\rho_b = (bb^{-1})\rho_b = b.$$

To show that ϕ is a morphism we must show that $\rho_a\rho_b = \rho_{ab}$ for all a, b in S. It is useful to establish first that the inverse $(\rho_a)^{-1}$ of ρ_a in $\mathcal{I}_S$ is $\rho_{a^{-1}}$; for $\rho_{a^{-1}}$ has domain $Sa = \operatorname{im}\rho_a$ and image $Sa^{-1} = \operatorname{dom}\rho_a$, and

$$x\rho_a\rho_{a^{-1}}\rho_a = xaa^{-1}a = xa = x\rho_a \quad (x \in Sa^{-1}),$$
$$x\rho_{a^{-1}}\rho_a\rho_{a^{-1}} = xa^{-1}aa^{-1} = xa^{-1} = x\rho_{a^{-1}} \quad (x \in Sa).$$

Now

$$\begin{aligned}\operatorname{dom}(\rho_a\rho_b) &= (Sa^{-1}a \cap Sbb^{-1})(\rho_a)^{-1} = Sa^{-1}abb^{-1}\rho_{a^{-1}} \\ &= Sa^{-1}abb^{-1}a^{-1} = Sab(ab)^{-1} = \operatorname{dom}\rho_{ab},\end{aligned}$$

and

$$\begin{aligned}\operatorname{im}(\rho_a\rho_b) &= (Sa^{-1}a \cap Sbb^{-1})\rho_b = Sa^{-1}abb^{-1}\rho_b \\ &= Sa^{-1}abb^{-1}b = Sab = \operatorname{im}\rho_{ab}.\end{aligned}$$

Also, for all x in $\operatorname{dom}\rho_{ab}$,

$$x(\rho_a\rho_b) = (xa)b = x(ab) = x\rho_{ab}.$$

This completes the proof. We note finally that when S is a group the above proof reduces to the proof of Cayley's Theorem in group theory, since $Saa^{-1} = Sa^{-1}a = S$ in this case. □

It is convenient at this stage to mention another important class of inverse semigroups. In Chapter 3 we investigated completely 0-simple semigroups, and showed that every such semigroup is isomorphic to a regular Rees matrix semigroup $\mathcal{M}^0[G; I, \Lambda; P]$. Completely 0-simple semigroups are certainly regular, and it is reasonable to ask under what circumstances they are inverse semigroups. If we ask the same question of completely simple semigroups (without zero) then the answer is relatively uninteresting: the only completely simple inverse semigroups are groups. Where we have a zero, however, we obtain a new class of inverse semigroups.

Accordingly, let $S = \mathcal{M}^0[G; I, \Lambda; P]$ be a completely 0-simple inverse semigroup. Since each $\mathcal{R}$-class and each $\mathcal{L}$-class contains just one idempotent, there is exactly one non-zero entry in each row and each column of the sandwich matrix P. There is thus a bijection from I onto Λ defined by the rule that $i \mapsto \lambda$ if and only if $p_{\lambda i} \neq 0$. Hence $|\Lambda| = |I|$, and we may suppose that Λ and I are ordered so that the non-zero entries occur on the main diagonal. Since I and Λ are merely index sets we may in effect suppose that $\Lambda = I$. Thus $S = \mathcal{M}^0[G; I, I; P]$, where P is a diagonal matrix.

Now let $\Delta = [\delta_{ij}]$ be the $I \times I$ matrix given by

$$\delta_{ij} = \begin{cases} e & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

By Theorem 3.4.1 we see that $S \simeq \mathcal{M}^0[G; I, I; \Delta]$, since there are elements $u_i = p_{ii}$, $v_j = e$ $(i, j \in I)$ such that

$$p_{ij} = v_j \delta_{ij} u_i$$

for all i, j in I.

We have thus proved the difficult half of

Theorem 5.1.8 *A semigroup S is both completely simple and an inverse semigroup if and only if $S \simeq \mathcal{M}^0[G; I, I; \Delta]$ for some group G and some index set I.*

Proof It remains to verify that idempotents commute within the semigroup $S = \mathcal{M}^0[G; I, I; \Delta]$. The non-zero idempotents of S are the elements (i, e, i), where e is the identity of G, and it is clear that if (i, e, i) and (j, e, j) are distinct idempotents then

$$(i, e, i)(j, e, j) = (j, e, j)(i, e, i) = 0. \qquad \square$$

The semigroup $\mathcal{M}^0[G; I, I; \Delta]$ is called a *Brandt semigroup*. If I is a finite set, with $|I| = n$, we usually denote the Brandt semigroup by $B(G, n)$. In particular, we write $B(\{e\}, n)$, a semigroup of order $2n + 1$, as B_n. We encountered the semigroup B_2 in Section 1.6.

5.2 THE NATURAL ORDER RELATION

On an inverse semigroup S it is possible to define a partial order relation in a very natural way: given a, b in S, we define $a \leq b$ if there exists an idempotent e in S such that $a = eb$. The first thing to check, of course, is that this relation is indeed a partial order. Certainly $a \leq a$ for every a in S, for we have $a = (aa^{-1})a$. Next, if $a \leq b$ and $b \leq a$ then $a = eb$ and $b = fa$, where e and f are idempotents, and it follows that

$$a = eb = efa = fea = fe^2b = feb = fa = b.$$

Finally, if $a \leq b$ and $b \leq c$ then there are idempotents e, f such that $a = eb$ and $b = fc$. It follows that $a = (ef)c$, where ef is idempotent, and hence $a \leq c$ as required.

The order relation $\leq$ is in fact *compatible* with the multiplication of S, in the sense that

$$a \leq b \text{ and } c \in S \;\Rightarrow\; ac \leq bc \text{ and } ca \leq cb. \tag{5.2.1}$$

The first implication is obvious, since $a = eb$ implies that $ac = e(bc)$. For the second implication, notice that if $a = eb$ then

$$ca = ceb = c\big((c^{-1}c)e\big)b = c(ec^{-1}c)b = (cec^{-1})cb,$$

where cec^{-1} is idempotent.

The order is also compatible with inversion, in the sense that

$$a \leq b \Rightarrow a^{-1} \leq b^{-1}; \tag{5.2.2}$$

for $a = eb$ implies that

$$a^{-1} = b^{-1}e = b^{-1}bb^{-1}e = b^{-1}ebb^{-1} = (b^{-1}eb)b^{-1},$$

and $b^{-1}eb$ is idempotent.

If S is a group then it is clear that $\leq$ reduces to equality. In the case where S is a semilattice the order coincides with the order relation we already have on a semilattice. This is defined by $x \leq y$ if and only if $x = xy$, and on a commutative semigroup E of idempotents it is easy to see that this is equivalent to the statement that $x = ey$ for some e in E. In the case of the inverse semigroup $\mathcal{I}_X$ the interpretation of $\leq$ is a very natural one: $\alpha \leq \beta$ if and only if $\alpha \subseteq \beta$. Here we are thinking of α and β as subsets of $X \times X$; if we want to think of them in a more intuitive way we have that $\alpha \leq \beta$ if and only if α is a *restriction* of β, i.e., if and only if $\operatorname{dom}\alpha \subseteq \operatorname{dom}\beta$ and $x\alpha = x\beta$ for all x in $\operatorname{dom}\alpha$. The verification of this is straightforward.

It is perhaps already clear that the one-sidedness in the definition of $\leq$ is only apparent. The next result gathers together a number of different characterizations of $\leq$. The list is by no means exhaustive.

Proposition 5.2.1 *Let S be an inverse semigroup with semilattice E of idempotents, and let $a, b \in S$. The following statements are equivalent:*

(1) $a \leq b$;	(2) $(\exists e \in E)\ a = be$;
(3) $aa^{-1} = ba^{-1}$;	(4) $aa^{-1} = ab^{-1}$;
(5) $a^{-1}a = b^{-1}a$;	(6) $a^{-1}a = a^{-1}b$;
(7) $a = ab^{-1}a$;	(8) $a = aa^{-1}b$.

Proof Only sample proofs are necessary. We shall show that (1) $\Leftrightarrow$ (3) $\Leftrightarrow$ (7).

(1) $\Rightarrow$ (3). Suppose that $a = eb$, where f is idempotent. Then

$$aa^{-1} = ebb^{-1}e = bb^{-1}e = b(eb)^{-1} = ba^{-1}.$$

(3) $\Rightarrow$ (7). Suppose that $aa^{-1} = ba^{-1}$. Then $a^{-1} = a^{-1}aa^{-1} = a^{-1}ba^{-1}$, and taking inverses of both sides we obtain $a = ab^{-1}a$.

(7) $\Rightarrow$ (1). Let $a = ab^{-1}a$. Then $(ab^{-1})^2 = ab^{-1}$, and so $a = eb$, where $e = ab^{-1}$ is idempotent. □

Let S be an inverse semigroup, and let H be a subset of S. The *upper saturation* or *closure* $H\omega$ of H in S is defined by

$$H\omega = \{s \in S : (\exists h \in H)h \leq s\}.$$

The term 'closure' is justified by the following easily verified facts (where H, K are subsets of S):

$$H \subseteq H\omega, \tag{5.2.3}$$

$$H \subseteq K \Rightarrow H\omega \subseteq K\omega, \tag{5.2.4}$$

$$(H\omega)\omega = H\omega. \tag{5.2.5}$$

The subset H will be called *closed* if $H\omega = H$.

It is clear that not every subsemigroup of an inverse semigroup is an inverse semigroup. A subsemigroup H of an inverse semigroup S is an inverse semigroup if and only if

$$(\forall x \in S)\, x \in H \Rightarrow x^{-1} \in H.$$

In such a case we say that H is an *inverse subsemigroup* of S.

Proposition 5.2.2 *If H is an inverse subsemigroup of an inverse semigroup S, then $H\omega$ is a closed inverse subsemigroup of S.*

Proof That $H\omega$ is closed follows immediately from (5.2.5). To see that $H\omega$ is a subsemigroup, consider two elements x, y in $H\omega$. Then by definition there exist h, k in H such that $x \geq h$, $y \geq k$. By (5.2.1) it follows that $xy \geq hk \in H$, and so $xy \in H\omega$. To see that $H\omega$ is an inverse subsemigroup, suppose that $x \in H\omega$. Then there exists h in H such that $x \geq h$, and it follows by (5.2.2) that $x^{-1} \geq h^{-1} \in H$. Thus $x^{-1} \in H\omega$, and so $H\omega$ is an inverse subsemigroup, as required. □

5.3 CONGRUENCES ON INVERSE SEMIGROUPS

Congruences are rarely mentioned explicitly in group theory, but they are present in the background. Given a congruence ρ on a group G with identity e, it is easy to verify that $e\rho$ is a normal subgroup of G. If we denote $e\rho$ by N then we easily see that

$$(x, y) \in \rho \text{ if and only if } xy^{-1} \in N.$$

For each x in G the ρ-class $x\rho$ is simply the coset Nx (or equivalently xN, since N is normal). Since the congruence ρ is completely determined by N, it is always sufficient to deal with N, and we write G/N rather than G/ρ for the quotient group.

A similar situation arises in a ring R. If ρ is a congruence on R then 0ρ is a two-sided ideal of R. Denoting 0ρ by I, we see that

$$(x, y) \in \rho \text{ if and only if } x - y \in I.$$

The ρ-classes in this case are the residue classes $x + I$. Here again the congruence is wholly determined by I, and the standard notation for the quotient ring is R/I.

In semigroup theory it is not possible to avoid the explicit study of congruences. It is, however, possible in the context of inverse semigroups to mimic to some extent the group-theoretical treatment of congruences.

Let ρ be a congruence on an inverse semigroup S with semilattice E of idempotents. The restriction of ρ to E is a congruence on E, which we call the *trace* of ρ and write as $\tau = \operatorname{tr}\rho$. Each τ-class $e\tau$ is equal to $e\rho \cap E$. The congruence τ is *normal*, in the sense that

$$e \,\tau\, f \;\Rightarrow\; (\forall a \in S)\, a^{-1}ea \,\tau\, a^{-1}fa,$$

for we certainly have $a^{-1}ea \,\rho\, a^{-1}fa$, and both $a^{-1}ea$ and $a^{-1}fa$ are in E. For future reference we note at this stage the following result:

Proposition 5.3.1 *Let ρ be a congruence on an inverse semigroup S. Then S/ρ is a group if and only if* $\operatorname{tr}\rho = E \times E$.

Proof Suppose first that S/ρ is a group. Then every idempotent of S, since it must map to an idempotent of S/ρ, maps to the identity of S/ρ. It follows that $\operatorname{tr}\rho = E \times E$. Conversely, suppose that $\operatorname{tr}\rho = E \times E$. By Lallement's Lemma (Lemma 2.4.3) every idempotent in S/ρ is the image under $\rho^\natural$ of an idempotent of S. It follows that S/ρ is an inverse semigroup containing only one idempotent, and hence (as is easy to verify) S/ρ is a group. □

Next, by analogy with the group-theoretic case, we define $N = \operatorname{Ker}\rho$, the *kernel* of ρ, to be the union of all the idempotent ρ-classes:

$$N = \operatorname{Ker}\rho = \bigcup_{e \in E} e\rho.$$

(We use a capital letter in order to distinguish this from the earlier and more general notion of the kernel of a morphism.) The set N is a subsemigroup of S: if $x \in e\rho$ and $y \in f\rho$ then the congruence property gives us that $xy \in (ef)\rho \subseteq N$. It is even an inverse subsemigroup: if $x \in e\rho$ then $x \,\rho\, e$; hence $x^{-1} \,\rho\, e^{-1} = e$ and so $x^{-1} \in e\rho$. Certainly N is a *full* inverse subsemigroup, in the sense that it contains all the idempotents of S. Next, N is *self-conjugate*, in the sense that

$$a \in N \;\Rightarrow\; (\forall x \in S)\, x^{-1}ax \in N,$$

for $a \in e\rho$ implies that $x^{-1}ax \in (x^{-1}ex)\rho \subseteq N$.

Let us decide to call a subsemigroup *normal* if it is a full and self-conjugate inverse subsemigroup. Thus we have established that $N = \operatorname{Ker}\rho$ is a normal subsemigroup. We might hope that the properties of a congruence ρ could be completely described in terms of its kernel $\operatorname{Ker}\rho$ and its trace $\operatorname{tr}\rho$. This turns out to be the case, but before we can state a theorem we observe two connections between the kernel and the trace. First, for all a in S and e in E,

$$ae \in \operatorname{Ker}\rho \text{ and } (e, a^{-1}a) \in \operatorname{tr}\rho \;\Rightarrow\; a \in \operatorname{Ker}\rho. \tag{5.3.1}$$

For suppose that $ae \, \rho \, f$, where $f \in E$. Then

$$a = aa^{-1}a \, \rho \, ae \, \rho \, f,$$

and so $a \in \operatorname{Ker} \rho$. Also,

$$a \in \operatorname{Ker} \rho \;\Rightarrow\; (aa^{-1}, a^{-1}a) \in \operatorname{tr} \rho. \tag{5.3.2}$$

For suppose that $a \in e\rho$, where $e \in E$. Then, as already observed, $a^{-1} \in e\rho$, and it then follows that $aa^{-1}, a^{-1}a \in e\rho$. Thus $(aa^{-1}, a^{-1}a) \in \operatorname{tr} \rho$.

We now produce an abstract version of the kernel and trace. As usual, S is an inverse semigroup with semilattice E of idempotents. Let N be a normal subsemigroup of S and let τ be a normal congruence on E. We shall call the pair (N, τ) a *congruence pair* of S if (for all a in S and e in E)

(C1) $ae \in N$ and $(e, a^{-1}a) \in \tau \;\Rightarrow\; a \in N$;

(C2) $a \in N \;\Rightarrow\; (aa^{-1}, a^{-1}a) \in \tau$.

Before we state our main theorem, it pays to prove the following technical lemma:

Lemma 5.3.2 *Let S be an inverse semigroup, let (N, τ) be a congruence pair, and let $a, b \in S$, $e \in E$.*

(1) *If $aeb \in N$ and $e \, \tau \, a^{-1}a$, then $ab \in N$.*

(2) *If $(a^{-1}a, b^{-1}b) \in \tau$ and $ab^{-1} \in N$, then $(a^{-1}ea, b^{-1}eb) \in \tau$ for every e in E.*

Proof (1) Suppose that $aeb \in N$ and that $e \, \tau \, a^{-1}a$. Then

$$aeb = aebb^{-1}b = ab.b^{-1}eb = ab.f,$$

where $f = b^{-1}eb \in E$, and so $(ab)f \in N$. Now from $e \, \tau \, a^{-1}a$ we deduce that $f \, \tau \, b^{-1}a^{-1}ab = (ab)^{-1}ab$, since τ is normal. Hence by (C1) it follows that $ab \in N$.

(2) Suppose that $(a^{-1}a, b^{-1}b) \in \tau$, $ab^{-1} \in N$ and $e \in E$. Then (modulo τ)

$$\begin{aligned}
a^{-1}ea &= (a^{-1}ea)(a^{-1}a)(a^{-1}ea) \\
&\equiv (a^{-1}ea)(b^{-1}b)(a^{-1}ea) \quad (\text{since } a^{-1}a \, \tau \, b^{-1}b \text{ and } \tau \text{ is normal}) \\
&= a^{-1}e(ab^{-1})(ab^{-1})^{-1}ea \\
&\equiv a^{-1}eba^{-1}ab^{-1}ea \quad (\text{by (C2), since } ab^{-1} \in N \text{ and } \tau \text{ is normal}) \\
&= a^{-1}(ab^{-1}e)^{-1}(ab^{-1}e)a \\
&\equiv a^{-1}ab^{-1}eba^{-1}a \quad (\text{by (C2), since } ab^{-1}e \in N \text{ and } \tau \text{ is normal}) \\
&\equiv b^{-1}bb^{-1}ebb^{-1}b \quad (\text{since } \tau \text{ is a congruence on } E) \\
&= b^{-1}eb.
\end{aligned}$$

Thus the lemma is proved. □

We now have our main result:

Theorem 5.3.3 *Let S be an inverse semigroup with semilattice E of idempotents. If ρ is a congruence on S then* $(\operatorname{Ker}\rho, \operatorname{tr}\rho)$ *is a congruence pair. Conversely, if (N, τ) is a congruence pair, then the relation*

$$\rho_{(N,\tau)} = \{(a,b) \in S \times S : (a^{-1}a, b^{-1}b) \in \tau,\ ab^{-1} \in N\}$$

is a congruence on S. Moreover, $\operatorname{Ker}\rho_{(N,\tau)} = N$, $\operatorname{tr}\rho_{(N,\tau)} = \tau$, *and* $\rho_{(\operatorname{Ker}\rho, \operatorname{tr}\rho)} = \rho$.

Proof By virtue of (5.3.1) and (5.3.2) we have in fact already established the first statement in the theorem. So suppose now that (N, τ) is a congruence pair, and let $\rho = \rho_{(N,\tau)}$ be as defined. Then ρ is a reflexive relation since N is full, and is a symmetric relation since τ is symmetric and N is an inverse subsemigroup. To show that ρ is transitive, let $(a,b), (b,c) \in \rho$. Then $(a^{-1}a, b^{-1}b), (b^{-1}b, c^{-1}c) \in \tau$, and so $(a^{-1}a, c^{-1}c) \in \tau$. Also $ab^{-1}, bc^{-1} \in N$, from which we deduce that N contains $a(b^{-1}b)c^{-1} = aec^{-1}$, where $e = b^{-1}b$. Since $e \mathrel{\tau} a^{-1}a$, we deduce from Lemma 5.3.2(1) that $ac^{-1} \in N$. Hence $(a,c) \in \rho$, and we have now established that ρ is an equivalence relation.

To show that ρ is a congruence, suppose that $a \mathrel{\rho} b$ and let $c \in S$. Then $a^{-1}a \mathrel{\tau} b^{-1}b$ and so, modulo τ,

$$(ac)^{-1}(ac) = c^{-1}(a^{-1}a)c \equiv c^{-1}(b^{-1}b)c = (bc)^{-1}(bc).$$

Also

$$(ac)(bc)^{-1} = a(cc^{-1})b^{-1} = a(cc^{-1})(b^{-1}b)b^{-1} = ab^{-1}(bcc^{-1}b^{-1}) \in N.$$

Thus $ac \mathrel{\rho} bc$. Also, again modulo τ,

$$(ca)^{-1}(ca) = a^{-1}(c^{-1}c)a \equiv b^{-1}(c^{-1}c)b = (cb)^{-1}(cb),$$

by Lemma 5.3.2(2), and $(ca)(cb)^{-1} = c(ab^{-1})c^{-1} \in N$, since N is normal. Hence $ca \mathrel{\rho} cb$.

We have shown that $\rho = \rho_{(N,\tau)}$ is a congruence on S. It is now clear that if $a \in e\rho$ with e in E, then $a^{-1}a \mathrel{\tau} e$ and $ae \in N$. By (C1) it follows that $a \in N$. Thus $\operatorname{Ker}\rho \subseteq N$. Conversely, if $a \in N$ then $ae^{-1} \in N$ with $e = a^{-1}a$, and certainly $a^{-1}a \mathrel{\tau} e^{-1}e$ (since the two idempotents are in fact equal); hence $a \in e\rho \subseteq \operatorname{Ker}\rho$. Hence $\operatorname{Ker}\rho_{(N,\tau)} = N$.

Next, to show that $\operatorname{tr}\rho_{(N,\tau)} = \tau$, let $e, f \in E$ and suppose that $(e,f) \in \rho = \rho_{(N,\tau)}$. Then

$$e = e^{-1}e \mathrel{\tau} f^{-1}f = f.$$

Hence $\operatorname{tr}\rho \subseteq \tau$. Conversely, if $e \mathrel{\tau} f$, then

$$e^{-1}e = e \mathrel{\tau} f = f^{-1}f$$

and $ef^{-1} = ef \in E \subseteq N$; hence $(e,f) \in \rho \cap (E \times E) = \operatorname{tr}\rho$. We have shown that $\operatorname{tr}\rho_{(N,\tau)} = \tau$.

To establish the final statement in the theorem, suppose first that $(a,b) \in \rho$. Then $(a^{-1},b^{-1}) \in \rho$, and so $(a^{-1}a, b^{-1}b) \in \rho$. Since $a^{-1}a$, $b^{-1}b$ are idempotents, we in fact have $(a^{-1}a, b^{-1}b) \in \operatorname{tr}\rho$. Also $(ab^{-1}, bb^{-1}) \in \rho$, and so $ab^{-1} \in (bb^{-1})\rho \subseteq \operatorname{Ker}\rho$. We have shown that $\rho \subseteq \rho_{(\operatorname{Ker}\rho, \operatorname{tr}\rho)}$.

Conversely, suppose that $(a,b) \in \rho_{(\operatorname{Ker}\rho,\operatorname{tr}\rho)}$, so that $(a^{-1}a, b^{-1}b) \in \operatorname{tr}\rho$ and $ab^{-1} \in \operatorname{Ker}\rho$. Then $(ab^{-1})\rho$ is an idempotent in S/ρ, and so

$$(ab^{-1})\rho = \Big(\big((ab^{-1})^{-1}\big)\rho\Big)\Big((ab^{-1})\rho\Big) = (ba^{-1}ab^{-1})\rho.$$

Then, modulo ρ,

$$a = aa^{-1}a \equiv ab^{-1}b \equiv ba^{-1}ab^{-1}b \equiv bb^{-1}bb^{-1}b = b.$$

That is, $(a,b) \in \rho$, and this, together with the inclusion proved in the previous paragraph, gives us $\rho = \rho_{(\operatorname{Ker}\rho,\operatorname{tr}\rho)}$, as required. □

In looking at special cases it is helpful first to establish the following result:

Proposition 5.3.4 *Let S be an inverse semigroup with semilattice E of idempotents, and let τ be a normal congruence on E. Then:*

(1) *the relation $\tau_{\min}$ defined as*

$$\{(a,b) \in S \times S : aa^{-1} \,\tau\, bb^{-1} \text{ and } (\exists e \in E)\, \big(e \,\tau\, aa^{-1} \text{ and } ea = eb\big)\}$$

is the smallest congruence on S with trace τ;

(2) *the relation*

$$\tau_{\max} = \{(a,b) \in S \times S : (\forall e \in E)\, a^{-1}ea \,\tau\, b^{-1}eb\}$$

is the largest congruence on S with trace τ.

Proof (1) We show first that $\tau_{\min}$ is an equivalence. It is clear that $\tau_{\min}$ is reflexive and symmetric. To show that it is transitive, let $(a,b),(b,c) \in \tau_{\min}$. Then $aa^{-1} \,\tau\, bb^{-1} \,\tau\, cc^{-1}$, and there exist e, f in $(aa^{-1})\tau$ such that $ea = eb$, $fb = fc$. Then $ef \in (aa^{-1})\tau$ and $efa = efc$, and so $(a,c) \in \tau_{\min}$ as required.

To show that $\tau_{\min}$ is a congruence, suppose that $(a,b) \in \tau_{\min}$ and that $c \in S$. Thus $aa^{-1} \,\tau\, bb^{-1}$ and there exists e in $(aa^{-1})\tau$ such that $ea = eb$. It follows that $(ca)(ca)^{-1} = c(aa^{-1})c^{-1} \,\tau\, c(bb^{-1})c^{-1} = (cb)(cb)^{-1}$, since τ is normal, and we have cec^{-1} in $\big((ca)(ca)^{-1}\big)\tau$ such that $(cec^{-1})ca = cea = ceb = (cec^{-1})cb$. Thus $(ca,cb) \in \tau_{\min}$. To show that $(ac,bc) \in \tau_{\min}$, notice that (modulo τ)

$$\begin{aligned}(ac)(ac)^{-1} &= a(cc^{-1})a^{-1} = aa^{-1}a(cc^{-1})a^{-1}\\ &\equiv eacc^{-1}a^{-1} = eacc^{-1}a^{-1}e = ebcc^{-1}b^{-1}e = ebcc^{-1}b^{-1}\\ &\equiv bb^{-1}bcc^{-1}b^{-1} = (bc)(bc)^{-1}.\end{aligned}$$

To complete the proof that $\tau_{\min}$ is a congruence, denote the idempotent $e(ac)(ac)^{-1}$ by f. Then $f \; \tau \; aa^{-1}(ac)(ac)^{-1} = (ac)(ac)^{-1}$, and

$$\begin{aligned} f(ac) &= e(ac)(ac)^{-1}(ac) = (ac)(ac)^{-1}(ea)c = (ac)(ac)^{-1}(eb)c \\ &= e(ac)(ac)^{-1}(bc) = f(bc). \end{aligned}$$

We now verify that $\operatorname{tr}\tau_{\min} = \tau$. Suppose first that $(e,f) \in \tau$. Then $ee^{-1} \; \tau \; ff^{-1}$, and there is an element ef in E with the property that $ef \; \tau \; ee^{-1}$ and $(ef)e = (ef)f$. Hence $(e,f) \in \operatorname{tr}\tau_{\min}$. Conversely, suppose that $(e,f) \in \tau_{\min} \cap (E \times E)$. Then $e = ee^{-1} \; \tau \; ff^{-1} = f$.

Finally, we show that $\tau_{\min}$ is the least congruence with trace equal to τ. Let ρ be a congruence on S with trace τ, and suppose that $(a,b) \in \tau_{\min}$. Then $aa^{-1} \; \rho \; bb^{-1}$, and there is an idempotent e such that $e \; \rho \; aa^{-1}$ and $ea = eb$. It follows that (modulo ρ)

$$a = aa^{-1}a \equiv ea = eb \equiv bb^{-1}b = b.$$

(2) Turning now to $\tau_{\max}$, we easily see that this relation is an equivalence. To show that it is a congruence, let $(a,b) \in \tau_{\max}$ and let $c \in S$. Then, using the normality of τ we have, for all e in E,

$$(ac)^{-1}e(ac) = c^{-1}(a^{-1}ea)c \equiv c^{-1}(b^{-1}eb)c = (bc)^{-1}e(bc) \pmod{\tau}$$

and so $(ac, bc) \in \tau_{\max}$. Also, since $c^{-1}ec \in E$ for all e in E, we have, for all e in E,

$$(ca)^{-1}e(ca) = a^{-1}(c^{-1}ec)a \equiv b^{-1}(c^{-1}ec)b = (cb)^{-1}e(cb);$$

thus $(ca, cb) \in \tau_{\max}$.

To show that $\operatorname{tr}\tau_{\max} = \tau$, suppose first that $(e,f) \in \tau$. Then, for every i in E,

$$e^{-1}ie = ie \equiv if = f^{-1}if \pmod{\tau},$$

and so $(e,f) \in \operatorname{tr}\tau_{\max}$. Conversely, suppose that $(e,f) \in \tau_{\max} \cap (E \times E)$. Then $ie \; \tau \; if$ for every i in E. In particular $e = ee \; \tau \; ef$ and $fe \; \tau \; ff = f$, and so $(e,f) \in \tau$ as required.

Finally, we show that $\tau_{\max}$ is the greatest congruence with trace τ. Let ρ be a congruence on S whose trace is τ and suppose that $(a,b) \in \rho$. Then $(a^{-1}, b^{-1}) \in \rho$, and so it follows that, for all e in E,

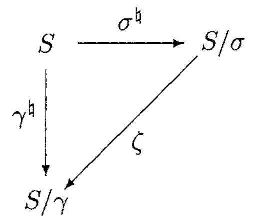
$$(a^{-1}ea, b^{-1}eb) \in \rho \cap (E \times E) = \tau.$$

Thus $(a,b) \in \tau_{\max}$, and our proof is complete. □

Two congruences are of particular interest. First, let $\tau = E \times E$, the universal congruence on E. Then, recalling Proposition 5.3.1, we see that

$$\tau_{\min} = \{(a,b) \in S \times S : (\exists e \in E) \; ea = eb\} \tag{5.3.3}$$

is the smallest congruence ρ on S such that S/ρ is a group. We shall consistently denote this congruence by σ and call it the *minimum group* *congruence* on S. Thus, to put the same thing another way, S/σ is the *maximum group morphic image* of S, in the sense that for every congruence γ on S such that S/γ is a group there is a morphism $\zeta : S/\sigma \to S/\gamma$ such that the diagram



commutes.

It is easy to determine the kernel of σ:

Theorem 5.3.5 *Let S be an inverse semigroup with semilattice E of idempotents and let σ be the minimum group congruence on S. Then* $\operatorname{Ker}\sigma = E\omega$*, and so*

$$\sigma = \{(a,b) \in S \times S : ab^{-1} \in E\omega\}.$$

Proof Since $e\sigma = f\sigma$ for all e, f in E, we deduce that $\operatorname{Ker}\sigma = i\sigma$ for some arbitrarily chosen idempotent i in E. Suppose now that $a \in \operatorname{Ker}\sigma$. Then $(a,i) \in \sigma$ and so, by equation (5.3.3), $ea = ei$ for some e in E. Thus $a \geq ea = ei \in E$ and so $a \in E\omega$.

Conversely, suppose that $a \in E\omega$, so that $a \geq e$ for some e in E. Then $ea = e = ee$, and so $a \; \sigma \; e$. Thus $a \in \operatorname{Ker}\sigma$. The remaining statement follows from Theorem 5.3.3. □

The other special case of interest is the congruence $1_{\max}$. A congruence whose trace is the identical congruence 1 is called *idempotent-separating*. We shall consistently denote $1_{\max}$ by μ, and call it the *maximum idempotent-separating congruence* on S. It is given by

$$\mu = \{(a,b) \in S \times S : (\forall e \in E)\; a^{-1}ea = b^{-1}eb\}. \tag{5.3.4}$$

To determine $\operatorname{Ker}\mu$ we need a definition. Let $E\zeta$, the *centralizer of E in S*, be given by

$$E\zeta = \{a \in S : (\forall e \in E)\; ae = ea\}. \tag{5.3.5}$$

Then we have the following result:

Theorem 5.3.6 *Let S be an inverse semigroup with semilattice E of idempotents, and let μ be the maximum idempotent-separating congruence on S. Then* $\operatorname{Ker}\mu = E\zeta$*, and so*

$$\mu = \{(a,b) \in S \times S : a^{-1}a = b^{-1}b \text{ and } ab^{-1} \in E\zeta\}.$$

Proof Suppose first that $a \in \operatorname{Ker}\mu$, so that $a \mathrel{\mu} i$ for some i in E. Hence $a^{-1} \equiv i^{-1} = i$ and so $a^{-1}a \mathrel{\mu} i^2 = i$. Thus $a \equiv a^{-1}a$, and so, for all e in E,

$$a^{-1}ea = a^{-1}aea^{-1}a = a^{-1}ae.$$

Hence

$$ea = eaa^{-1}a = aa^{-1}ea = aa^{-1}ae = ae,$$

and so $a \in E\zeta$.

Conversely, suppose that $a \in E\zeta$. Then, for all e in E,

$$a^{-1}ea = a^{-1}ae = a^{-1}aea^{-1}a = (a^{-1}a)^{-1}ea^{-1}a.$$

Thus $a \mathrel{\mu} a^{-1}a$ and so $a \in \operatorname{Ker}\mu$. The final statement follows from Theorem 5.3.3. □

A further characterization of μ is of interest:

Proposition 5.3.7 *Let S be an inverse semigroup with semilattice E of idempotents, and let μ be the maximum idempotent-separating congruence on S. Then $\mu = \mathcal{H}^\flat$, the largest congruence on S contained in $\mathcal{H}$.*

Proof Let $(a, b) \in \mu$. Then from Theorem 5.3.6 we see that $a^{-1}a = b^{-1}b$. Since $(a^{-1}, b^{-1}) \in \mu$ we also have $aa^{-1} = bb^{-1}$, and so $(a, b) \in \mathcal{H}$.

We have shown that $\mu \subseteq \mathcal{H}$. To see that μ is the largest congruence contained in $\mathcal{H}$, consider a congruence $\rho \subseteq \mathcal{H}$, and suppose that $(a, b) \in \rho$. Then $(a^{-1}, b^{-1}) \in \rho$, and so

$$(a^{-1}ea, b^{-1}eb) \in \rho \subseteq \mathcal{H}$$

for all e in E. By Corollary 2.2.6 we deduce that $a^{-1}ea = b^{-1}eb$ for all e in E, and hence that $(a, b) \in \mu$. □

An inverse semigroup in which the maximum idempotent-separating congruence is the identical congruence 1_S is called *fundamental.*

Theorem 5.3.8 *Let S be an inverse semigroup with semilattice E of idempotents, and let μ be the maximum idempotent-separating congruence on S. Then S/μ is fundamental, and has semilattice of idempotents isomorphic to E.*

Proof Every idempotent in S/μ has the form $e\mu$, where $e \in E$. Suppose, to use an obvious notation, that $(a\mu, b\mu) \in \mu_{S/\mu}$. Then, for every e in E,

$$(a\mu)^{-1}(e\mu)(a\mu) = (b\mu)^{-1}(e\mu)(b\mu),$$

and so $(a^{-1}ea, b^{-1}eb) \in \mu$. Since μ is idempotent-separating, it follows that $a^{-1}ea = b^{-1}eb$, i.e., that $a\mu = b\mu$. Thus S/μ is fundamental. Since the morphism $\mu^\natural$ separates idempotents, the rest of the proposition is immediate. □

5.4 THE MUNN SEMIGROUP

The main aim of this section is to carry out a construction, due to Munn (1966b), which produces from an arbitrary semilattice E a fundamental inverse semigroup with an important universal property.

Let E be a semilattice. For each e in E the set $Ee = \{i \in E : i \leq e\}$ is a principal ideal of E. The *uniformity relation* $\mathcal{U}$ on E is given by

$$\mathcal{U} = \{(e, f) \in E \times E : Ee \simeq Ef\}. \tag{5.4.1}$$

For each (e, f) in U we define $T_{e,f}$ to be the set of all isomorphisms from Ee onto Ef. Let

$$T_E = \bigcup \{T_{e,f} : (e, f) \in \mathcal{U}\}. \tag{5.4.2}$$

We call T_E the *Munn semigroup* of the semilattice E.

This terminology presupposes that we can define an associative multiplication on T_E. We now show that this is indeed possible. Clearly $T_E \subseteq \mathcal{I}_E$, since all the elements are partial one-one maps of E. We shall show that T_E is in fact an inverse subsemigroup of $\mathcal{I}_E$. Let $\alpha : Ee \to Ef$ and $\beta : Eg \to Eh$ be elements of T_E. The product of α and β in $\mathcal{I}_E$ maps $(Ef \cap Eg)\alpha^{-1} = (Efg)\alpha^{-1}$ onto $(Ef \cap Eg)\beta = (Efg)\beta$. If we write $(ef)\alpha^{-1}$ as i and $(ef)\beta$ as j, we see that

$$\begin{aligned} x \in (Efg)\alpha^{-1} &\Leftrightarrow x\alpha \in Efg \Leftrightarrow x\alpha \leq fg \\ &\Leftrightarrow x \leq (fg)\alpha^{-1} \Leftrightarrow x \in Ei, \end{aligned}$$

and similarly $(Efg)\beta = Ej$. Thus α maps the principal ideal Ei onto the principal ideal Ej by the rule that $x(\alpha\beta) = (x\alpha)\beta$ for all x in Ei. Since $\alpha\beta$ is clearly an isomorphism, we deduce that $\alpha\beta \in T_E$. It is clear that, for every $\alpha : Ee \to Ef$ in T_E, the inverse $\alpha^{-1} : Ef \to Ee$ is also in T_E.

We have established part of the following result:

Theorem 5.4.1 *For every semilattice E, the Munn semigroup T_E is an inverse semigroup whose semilattice of idempotents is isomorphic to E.*

Proof It remains to consider the semilattice of idempotents of T_E. A typical idempotent of T_E is the identical map 1_{Ee} of Ee onto itself. We have a one-one map $e \mapsto 1_{Ee}$ from E onto the set of idempotents of T_E, and since (by (5.1.3))

$$1_{Ee} 1_{Ef} = 1_{Ee \cap Ef} = 1_{Eef},$$

the map $e \mapsto 1_{Ee}$ is an isomorphism.

It is useful to note at this stage that, for each $\alpha : Ee \to Ef$ in T_E,

$$\alpha\alpha^{-1} = 1_{Ee}, \quad \alpha^{-1}\alpha = 1_{Ef}.$$

Hence, in T_E, $\alpha \mathrel{\mathcal{R}} \beta$ if and only if $\operatorname{dom}\alpha = \operatorname{dom}\beta$, and $\alpha \mathrel{\mathcal{L}} \beta$ if and only if $\operatorname{im}\alpha = \operatorname{im}\beta$. □

We shall very often want to identify 1_{Ee} with e and to think of T_E as an inverse semigroup having E as semilattice of idempotents.

Before investigating the general consequences of this construction, we pause to consider two examples.

Example 5.4.2 Let $E = \{0, 1, 2, \ldots\}$, with the natural order, given by $0 < 1 < 2 < \cdots$ Then, for each n,

$$En = \{0, 1, \ldots, n\},$$

and so $Em \simeq En$ if and only if $m = n$. In this case $\mathcal{U} = 1_E$ and we say that E is *anti-uniform*. The only isomorphism in $T_{n,n}$ is 1_{En}, and so

$$T_E = \{1_{E0}, 1_{E1}, 1_{E2}, \ldots\} \simeq E.$$

Example 5.4.3 Let $E = C_\omega = \{e_0, e_1, e_2, \ldots\}$, with

$$e_0 > e_1 > e_2 > \cdots$$

Then

$$Ee_n = \{e_n, e_{n+1}, e_{n+2}, \ldots\},$$

and $Ee_m \simeq Ee_n$ for every m, n in $\mathbf{N}^0$. Here $\mathcal{U} = E \times E$ and we say that E is *uniform*. The only isomorphism from Ee_m onto Ee_n is $\alpha_{m,n}$, given by

$$e_k\alpha_{m,n} = e_{k-m+n} \quad (k \geq m),$$

the inverse being $\alpha_{n,m} : Ee_n \to Ee_m$, defined by

$$e_l\alpha_{n,m} = e_{l-n+m} \quad (l \geq n).$$

If $\alpha_{m,n}$ and $\alpha_{p,q}$ are elements of T_E, then their product maps $(Ee_n \cap Ee_p)\alpha_{m,n}^{-1}$ onto $(Ee_n \cap Ee_p)\alpha_{p,q}$. If we write $t = \max(n,p)$, we can say that $\alpha_{m,n}\alpha_{p,q}$ maps Ee_{t-n+m} onto Ee_{t+p-q}. That is,

$$\alpha_{m,n}\alpha_{p,q} = \alpha_{m-n+t,q-p+t}. \tag{5.4.3}$$

We can thus identify the Munn semigroup of the semilattice C_ω with the *bicyclic semigroup* already encountered in (1.6.3), that is to say, with the semigroup $\mathbf{N}^0 \times \mathbf{N}^0$, in which multiplication is given by

$$(m,n)(p,q) = \big(m - n + \max(n,p), q - p + \max(n,p)\big). \tag{5.4.4}$$

The importance of the Munn semigroup lies in the following result:

Theorem 5.4.4 *For every inverse semigroup S with semilattice of idempotents E there is a morphism $\phi : S \to T_E$ whose kernel is μ, the maximum idempotent-separating congruence on S.*

Proof Let $a \in S$. We define a map $\alpha_a : Eaa^{-1} \to Ea^{-1}a$ by the rule that

$$x\alpha_a = a^{-1}ea \quad (e \in Eaa^{-1}).$$

Notice that the image of this map is certainly contained in $Ea^{-1}a$, since $a^{-1}ea = (a^{-1}ea)a^{-1}a$. To show that α_a is a bijection we note that the map $\alpha_{a^{-1}} : Ea^{-1}a \to Eaa^{-1}$ is a two-sided inverse of α_a:

$$(eaa^{-1})\alpha_a\alpha_{a^{-1}} = aa^{-1}eaa^{-1}aa^{-1} = eaa^{-1},$$
$$(ea^{-1}a)\alpha_{a^{-1}}\alpha_a = a^{-1}aea^{-1}aa^{-1}a = ea^{-1}a.$$

Suppose now that $x = eaa^{-1}$ and $y = faa^{-1}$ belong to Eaa^{-1}. Then

$$(x\alpha_a)(y\alpha_a) = a^{-1}eaa^{-1}aa^{-1}faa^{-1}a = a^{-1}(eaa^{-1}faa^{-1})a = (xy)\alpha_a;$$

hence α_a is an isomorphism.

We have shown that $\alpha_a \in T_E$ for every a in S. Now define a map ϕ from S into T_E by the rule that

$$a\phi = \alpha_a \quad (a \in S).$$

To see that ϕ is a morphism, suppose that $a, b \in S$. Then by the multiplication rule in T_E, the product $\alpha_a\alpha_b$ is an isomorphism from Ei onto Ej, where

$$i = (a^{-1}abb^{-1})\alpha_a^{-1} = aa^{-1}abb^{-1}a^{-1} = (ab)(ab)^{-1},$$
$$j = (a^{-1}abb^{-1})\alpha_b = b^{-1}a^{-1}abb^{-1}b = (ab)^{-1}(ab).$$

Moreover, for every x in Ei,

$$(x\alpha_a)\alpha_b = b^{-1}a^{-1}xab = (ab)^{-1}x(ab),$$

and so $\alpha_a\alpha_b = \alpha_{ab}$. This shows that ϕ is a morphism.

To determine the kernel of ϕ, suppose first that $(a, b) \in \mu$. Then $(a, b) \in \mathcal{H}$ by Theorem 5.3.6 and so

$$\operatorname{dom}\alpha_a = Eaa^{-1} = Ebb^{-1} = \operatorname{dom}\alpha_b,$$
$$\operatorname{im}\alpha_a = Ea^{-1}a = Eb^{-1}b = \operatorname{im}\alpha_b.$$

Also, for all $x = eaa^{-1}$ in $\operatorname{dom}\alpha_a = Eaa^{-1}$,

$$x\alpha_a = a^{-1}(eaa^{-1})a = a^{-1}ea$$
$$= b^{-1}eb = b^{-1}(ebb^{-1})b = x\alpha_b,$$

and so $a\phi = b\phi$. Thus $\mu \subseteq \ker\phi$. Conversely, suppose that $a\phi = b\phi$. Then

$$Eaa^{-1} = \operatorname{dom}\alpha_a = \operatorname{dom}\alpha_b = Ebb^{-1},$$
$$Ea^{-1}a = \operatorname{im}\alpha_a = \operatorname{im}\alpha_b = Eb^{-1}b,$$

and so $aa^{-1} = bb^{-1}$, $a^{-1}a = b^{-1}b$. Also, for all e in E,

$$a^{-1}ea = a^{-1}(eaa^{-1})a = (eaa^{-1})\alpha_a$$
$$= (ebb^{-1})\alpha_b = b^{-1}eb,$$

and so $(a, b) \in \mu$. □

In the case where the inverse semigroup S is fundamental, the morphism ϕ obtained in Theorem 5.4.4 is one-one. In general it is not onto, but its image is in a certain sense a 'large' inverse subsemigroup of T_E. Recalling from Section 5.3 that an inverse subsemigroup of an inverse semigroup S is said to be *full* if it contains all the idempotents of S, we now have

Theorem 5.4.5 *An inverse semigroup with semilattice of idempotents E is fundamental if and only if it is isomorphic to a full inverse subsemigroup of T_E.*

Proof Suppose first that S is a fundamental inverse semigroup with semilattice of idempotents E. Then, by Theorem 5.4.4, the image $S\phi$ of the map $\phi : S \to T_E$ is an inverse subsemigroup of T_E and is isomorphic to S. Moreover, for each idempotent e of E,

$$e\phi = \alpha_e : Eee^{-1} \to Ee^{-1}e,$$

where

$$x\alpha_e = e^{-1}xe \quad (x \in Eee^{-1}).$$

That is, α_e maps Ee onto Ee, and $x\alpha_e = xe = x$ for all x in Ee. That is, $\alpha_e = 1_{Ee}$, and so ϕ maps the semilattice E of idempotents of S onto the isomorphic semilattice $\{1_{Ee} : e \in E\}$ of idempotents of T_E. Thus $S\phi$ is full.

Conversely, suppose that S is a full inverse subsemigroup of T_E. We must show that S is fundamental. Accordingly, suppose that α, β are elements of S and (in an obvious notation) that $(\alpha, \beta) \in \mu_S$. Then $\alpha \mathcal{H} \beta$ in S and so

$$1_{\text{dom}\,\alpha} = \alpha\alpha^{-1} = \beta\beta^{-1} = 1_{\text{dom}\,\beta},$$
$$1_{\text{im}\,\alpha} = \alpha^{-1}\alpha = \beta^{-1}\beta = 1_{\text{im}\,\beta}.$$

Thus $\text{dom}\,\alpha = \text{dom}\,\beta = Ee$ (say), and $\text{im}\,\alpha = \text{im}\,\beta = Ef$. Also, for every x in E (since S is full),

$$\alpha^{-1}1_{Ex}\alpha = \beta^{-1}1_{Ex}\beta.$$

In particular these maps have the same domain, and so

$$E\big((xe)\alpha\big) = E\big((xe)\beta\big)$$

for every x in E. Hence $(xe)\alpha = (xe)\beta$ for all xe in Ee, and so $\alpha = \beta$ as required. □

In particular, we note that

Corollary 5.4.6 *The inverse semigroup T_E is fundamental for every semilattice E.* □

5.5 ANTI-UNIFORM SEMILATTICES

Recall that a semilattice is *uniform* if

$$(\forall e, f \in E)\ Ee \simeq Ef,$$

and *anti-uniform* if, for all e, f in E,

$$Ee \simeq Ef \ \Rightarrow\ e = f.$$

Now, by Proposition 5.1.2(4), two idempotents e, f in an inverse semigroup S are $\mathcal{D}$-equivalent if and only if there exists a in S such that $aa^{-1} = e$, $a^{-1}a = f$. In such a case it is easy to verify that the map $x \mapsto a^{-1}xa$ $(x \in E)$ is an isomorphism from Ee onto Ef, and so $(e, f) \in \mathcal{U}$. We have shown that

$$\mathcal{D} \cap (E \times E) \subseteq \mathcal{U}. \tag{5.5.1}$$

Among the inverse semigroups with semilattice of idempotents (isomorphic to) E is the Munn semigroup T_E. In this semigroup $(e, f) \in \mathcal{U}$ implies the existence of an element α such that $\alpha\alpha^{-1} = 1_{Ee}$, $\alpha^{-1}\alpha = 1_{Ef}$. Hence (if we agree to identify 1_{Ee} with e for each e in E) $e\ \mathcal{D}\ f$ in T_E. We see that, for every semilattice E, the Munn semigroup T_E has the property that

$$\mathcal{D} \cap (E \times E) = \mathcal{U}. \tag{5.5.2}$$

Suppose now that S is an inverse semigroup whose semilattice of idempotents E is anti-uniform. Then, by (5.5.1),

$$\mathcal{D} \cap (E \times E) = 1_E.$$

Now, by Theorem 4.2.1, we know that an inverse semigroup—indeed a regular semigroup—has this property if and only if it is a Clifford semigroup, and so we have proved half of the following theorem:

Theorem 5.5.1 *Let E be a semilattice. Then E has the property that EVERY inverse semigroup with E as semilattice of idempotents is a Clifford semigroup if and only if E is anti-uniform.*

Proof If E is *not* anti-uniform, then, by (5.5.2), T_E is an inverse semigroup with semilattice of idempotents E, and T_E is *not* a Clifford semigroup. □

In view of Theorem 5.5.1 it is of interest to identify anti-uniform semilattices. We have seen (Example 5.4.2) that a chain order-isomorphic to the non-negative integers is anti-uniform. So is a finite chain. Indeed we have

Proposition 5.5.2 *Every well-ordered chain is anti-uniform.*

Proof Let E be a well-ordered chain and let $\phi : Ee \to Ef$ be an isomorphism, where $e, f \in E$. Without loss of generality we may suppose that

$e \leq f$, from which it follows that $Ee \subseteq Ef$. It is clear that $e\phi = f$, since maximum elements must correspond. Suppose now that the set

$$Y = \{y \in Ee : y\phi \neq y\}$$

is non-empty. By the well-ordered property there exists a least element x in Y, that is, an element x with the property that $x\phi \neq x$ and $z\phi = z$ for all $z < x$.

Now either $x\phi < x$ or $x\phi > x$. If $x\phi < x$ then $x\phi \in Ee$ and $(x\phi)\phi = x\phi$, which contradicts the hypothesis that ϕ is one-one. If $x\phi > x$ then, for each y in Ee,

$$y \geq x \Rightarrow y\phi \geq x\phi > x,$$

while

$$y < x \Rightarrow y\phi = y < x.$$

It follows that there is no element y of Ee such that $y\phi = x$, and this contradicts the assumption that ϕ maps onto Ef.

The assumption that $Y \neq \emptyset$ has led to a contradiction. Hence $Y = \emptyset$, and so $x\phi = x$ for every x in Ee. In particular, $f = e\phi = e$, and so E is anti-uniform, as required. □

Schein (1964) effectively conjectured that the *only* anti-uniform semilattices are the well-ordered chains. This proved over-optimistic, as we shall see. However, we do have

Proposition 5.5.3 *Let E be a semilattice with the minimal condition. Then E is anti-uniform if and only if E is a well-ordered chain.*

Proof We have already proved half of this. To prove the other half, suppose that E satisfies the minimal condition and is anti-uniform. Thus every non-empty subset of E contains a minimal element. If E is totally ordered then it is a well-ordered chain. We shall show that if E is *not* totally ordered then E is not anti-uniform. So suppose that E is not totally ordered. Following the usual terminology for ordered sets, we say that a and b are *comparable* if $a \leq b$ or $b \leq a$; otherwise we say that a and b are *incomparable*.

We define a subset K of E by the rule that $x \in K$ if there exist elements of E that are incomparable with x. Since E is not totally ordered, $K \neq \emptyset$, and so by the minimal condition there exists at least one minimal member e of K. Then $K_e \neq \emptyset$, where K_e is the set of elements of E that are incomparable with e. Choose a minimal member f of K_e. Then e and f are incomparable, and so $ef < e$, $ef < f$.

In fact e *covers* ef, in the sense that there is no g in E such that $ef < g < e$. To see this, suppose that $ef \leq g < e$. Now g, being less than e, is comparable with every element of E, and in particular is comparable with f. If $f \leq g$ then

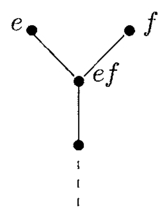
$$f \leq g < e,$$

and so e and f are comparable. This is a contradiction, and so $g < f$. But then from $g < e$ and $g < f$ we deduce that $g \leq ef$. Hence $g = ef$.

We have shown that there is no g such that $ef < g < e$. Thus e covers ef, and a closely similar argument shows that f covers ef. Hence

$$Ee = Eef \cup \{e\}, \quad Ef = Eef \cup \{f\},$$

where Eef is an ordered chain.



Now clearly $(e, f) \in \mathcal{U}$, an obvious isomorphism between Ee and Ef being that which sends e to f and every other element to itself. Thus E is not anti-uniform. □

It is not possible to remove all mention of the minimal condition in this proposition. Consider, for example, the following semilattice, described by Howie and Schein (1969). The set $\mathbf{Q}$ of rational numbers is known to be countable: let ϵ be a one-one map from $\mathbf{Q}$ onto the set $\mathbf{N}^0 = \{0, 1, 2, \ldots\}$ of non-negative integers, and let

$$E = \bigcup_{q \in \mathbf{Q}} \{(q, 0), (q, 1), \ldots, (q, q\epsilon)\},$$

with the *lexicographic* order: $(q, m) \leq (r, n)$ if and only if either $q < r$ or $q = r$ and $m \leq n$.

The set of elements of E having no immediate predecessor is $\{(q, 0) : q \in \mathbf{Q}\}$, and the set of elements having no immediate successor is $\{(q, q\epsilon) : q \in \mathbf{Q}\}$. Suppose that E is *not* anti-uniform. Then there exist distinct elements (q, m), (r, n) of E for which there is an isomorphism $\phi : E(q, m) \to E(r, n)$. Then

$$(q, m)\phi = (r, n), \ (q, n-1)\phi = (r, n-1), \ldots,$$

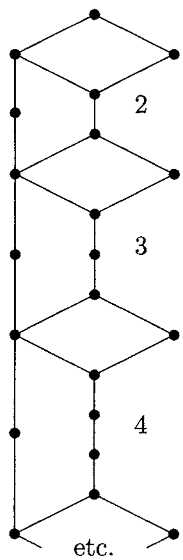
and if $m \neq n$ then we reach a situation where the isomorphism ϕ links two elements, one of which has an immediate successor and the other has not. Hence $m = n$, and ϕ maps (q, k) to (r, k) for $0 \leq k \leq m$.

Without loss of generality, suppose that $q > r$, and choose s in $\mathbf{Q}$ such that $q > s > r$. Then $(s, 0) \in E(q, m)$ and, having no immediate predecessor, must map to some $(t, 0)$ in $E(r, m)$, where $t < r$. Certainly $t < s$ and so $t\epsilon \neq s\epsilon$. But then

$$(s, 1)\phi = (t, 1), \quad (s, 2)\phi = (t, 2), \ldots,$$

and we reach a contradictory situation in which ϕ links two elements, one of which has an immediate successor and the other does not. We conclude that E is anti-uniform. It is certainly not well-ordered, since, for every q in $\mathbf{Q}$, the set $\{x \in E : x > (q, q\epsilon)\}$ has no least element.

Anti-uniform semilattices need not be chains, as is shown in the following example. A verbal description is tedious, and perhaps it is best to let the diagram speak for itself.



5.6 BISIMPLE INVERSE SEMIGROUPS

If S is a bisimple inverse semigroup with semilattice of idempotents E, then in particular all the idempotents are $\mathcal{D}$-equivalent, and it follows by (5.5.1) that $\mathcal{U} = E \times E$. That is, E is *uniform.*

If, conversely, we start with a uniform semilattice E, then we cannot expect that every inverse semigroup having E as semilattice of idempotents will be bisimple: E itself is one such inverse semigroup, and is assuredly not bisimple, since in E we have $\mathcal{D} = \mathcal{H} = 1_E$. On the other hand, we know that there is always at least one bisimple inverse semigroup having E as its semilattice, for it follows from (5.5.2) that

Proposition 5.6.1 *If E is a uniform semilattice, then the Munn semigroup T_E is bisimple.* □

This certainly gives a useful recipe for constructing bisimple inverse semigroups. It is not of course a universal recipe, for we have seen that T_E is fundamental, and we can easily produce a non-fundamental bisimple inverse semigroup simply by taking the direct product of some bisimple T_E with a non-trivial group. Recall, however, that we can embed any fundamental inverse semigroup as a full subsemigroup S' of the appropriate

T_E (Theorem 5.4.5). In the bisimple case we can be somewhat more precise, for if the subsemigroup S' of T_E is itself to be bisimple, there must exist for each e, f in E at least one element α *contained in* S' for which $\operatorname{dom}\alpha = Ee$ and $\operatorname{im}\alpha = Ef$. That is, for all e, f in E,

$$S' \cap T_{e,f} \neq \emptyset. \tag{5.6.1}$$

If we decide to apply the term *transitive* to subsemigroups of T_E having this property, it is easy to see that transitivity is not only a necessary but also a sufficient condition for bisimplicity. Noting finally that a transitive inverse subsemigroup of T_E is necessarily full, we deduce

Theorem 5.6.2 *An inverse semigroup S with semilattice of idempotents E is fundamental and bisimple if and only if E is uniform and S is isomorphic to a transitive inverse subsemigroup of T_E.* □

Let us apply this result to the uniform semlattice C_ω (Example 5.4.3). Here we have $|T_{e,f}| = 1$ for every pair (e, f), and so the condition (5.6.1) can be satisfied only if, for all pairs (e, f),

$$S' \cap T_{e,f} = T_{e,f},$$

that is, only if $S' = T_E$. We thus have

Corollary 5.6.3 *Up to isomorphism, the only fundamental bisimple inverse semigroup having C_ω as semilattice of idempotents is the bicyclic semigroup.* □

A further observation that can be made fairly easily at this stage will prove useful later.

Proposition 5.6.4 *Let S be an inverse semigroup with semilattice of idempotents E, and suppose that E has the property that $|T_{e,f}| = 1$ for all (e, f) in $\mathcal{U}$. Then the semigroup S has the property that $\mu = \mathcal{H}$, and (consequently) $\mathcal{H}$ is a congruence.*

Proof Since $\mu = \mathcal{H}^\flat$ by Proposition 5.3.7, we certainly have $\mu \subseteq \mathcal{H}$. To show the opposite inclusion, suppose that $a \mathcal{H} b$, so that $aa^{-1} = bb^{-1}$, $a^{-1}a = b^{-1}b$. Then, in the notation of the proof of Theorem 5.4.4,

$$\operatorname{dom}\alpha_a = Eaa^{-1} = Ebb^{-1} = \operatorname{dom}\alpha_b,$$

$$\operatorname{im}\alpha_a = Ea^{-1}a = Eb^{-1}b = \operatorname{im}\alpha_b.$$

By the given property of E, it now follows that $\alpha_a = \alpha_b$, and so $a \mu b$ by Theorem 5.4.4. □

Corollary 5.6.5 *If S is an inverse semigroup with semilattice of idempotents C_ω, then $\mathcal{H} = \mu$.* □

We should not expect the simple situation indicated by Corollaries 5.6.3 and 5.6.5 to obtain for more general uniform semilattices, and indeed it does not, as is illustrated by Exercise 23 below.

It is, however, reasonable to think of a general bisimple inverse semigroup as in some sense an 'inflation' of a fundamental bisimple inverse semigroup. The Notes at the end of the chapter refer to various attempts to make this vague notion precise. Here we shall home in on what are usually called *ω-semigroups*: semigroups whose semilattices of idempotents are isomorphic to C_ω. Here we have the advantage of Corollaries 5.6.3 and 5.6.5, and we also have an ingenious 'inflation' method, first considered in special cases by Bruck (1958) and Reilly (1966), and in general form by Munn (1970c). This we now describe.

Let T be a monoid with identity 1 and let θ be a morphism from T into H_1, the 'group of units' of T. Then we can make $\mathbf{N}^0 \times T \times \mathbf{N}^0$ into a semigroup by defining

$$(m,a,n)(p,b,q) = (m-n+t, (a\theta^{t-n})(b\theta^{t-p}), q-p+t), \tag{5.6.2}$$

where $t = \max(n,p)$ and where θ^0 is interpreted as the identity map of T. We must of course check that the given composition is associative. Observe that, if we write

$$\left.\begin{aligned} u &= \max\big(q-p+\max(n,p),\ r\big), \\ w &= \max\big(n,\ p-q+\max(q,r)\big), \end{aligned}\right\} \tag{5.6.3}$$

then

$$\begin{aligned} &\big((m,a,n)(p,b,q)\big)(r,c,s) \\ &= (m-n-q+p+u, (a\theta^{u-n-q+p})(b\theta^{u-q})(c\theta^{u-r}), s-r+u), \end{aligned}$$

while

$$\begin{aligned} &(m,a,n)\big((p,b,q)(r,c,s)\big) \\ &= (m-n+w, (a\theta^{w-n})(b\theta^{w-p})(c\theta^{w-r-p+q}), s-r-p+q+w). \end{aligned}$$

Now the outer coordinates in the multiplication (5.6.2) combine exactly as in the bicyclic semigroup, which we know to be associative (for example, because it is isomorphic to T_{C_ω}). Hence, equating first coordinates (or equivalently third coordinates) in the two products, we obtain

$$w = u + p - q,$$

a result that could of course be obtained directly (but tediously) from the definitions (5.6.3). From this it now easily follows that the middle coordinates in the two products are equal, and so the composition (5.6.2) is indeed associative. We shall denote the semigroup obtained in this way from T by $BR(T,\theta)$, and call it the *Bruck–Reilly extension of T determined by θ.*

Proposition 5.6.6 *Let T be a monoid with identity element* 1, *and let S be the Bruck–Reilly extension of T determined by θ. Then:*

(1) *S is a simple semigroup with identity* $(0,1,0)$;

(2) *two elements* (m,a,n) *and* (p,b,q) *are* $\mathcal{D}$*-equivalent in* S *if and only if* a *and* b *are* $\mathcal{D}$*-equivalent in* T*;*
(3) *the element* (m,a,n) *of* S *is idempotent if and only if* $m=n$ *and* a *is idempotent in* T*;*
(4) S *is an inverse semigroup if and only if* T *is an inverse semigroup.*

Proof (1) We show that if (m,a,n) and (p,b,q) are arbitrary elements of S, then there exist (r,x,s) and (t,y,u) such that

$$(r,x,s)(m,a,n)(t,y,u)=(p,b,q);$$

by Corollary 3.1.2 this is just what we require. If we take

$$(r,x,s)=(p,(a\theta)^{-1},m+1) \text{ and } (t,y,u)=(n+1,b,q),$$

where $(a\theta)^{-1}$ is the inverse of $a\theta$ in the group H_1, then we obtain the desired equality. The identity of S is easily seen to be $(0,1,0)$.

(2) We begin by considering the relation $\mathcal{R}$. Since there is some scope for confusion here, we shall use superscripts to distinguish between the Green equivalences on S and those on T. Suppose first that $(m,a,n)\ \mathcal{R}^S\ (p,b,q)$. Then

$$(m,a,n)(r,x,s)=(p,b,q) \tag{5.6.4}$$

for some (r,x,s) in S, and so

$$p=m-n+\max(n,r)\geq m.$$

But equally we may show that $m\geq p$. Hence $m=p$, and from

$$m=m-n+\max(n,r)$$

it follows that $n\geq r$. Hence, equating the middle coordinates in (5.6.4), we obtain

$$a(x\theta^{n-r})=b.$$

Thus $R_a\leq R_b$ in T. Since we can just as easily show that $R_b\leq R_a$, we deduce that $a\ \mathcal{R}^T\ b$.

Suppose conversely that $a\ \mathcal{R}^T\ b$. Then $ax=b$ and $bx'=a$ for some x, x' in T $(=T^1)$, and so, in S,

$$(m,a,n)(n,x,q)=(m,b,q), \quad (m,b,q)(q,x',n)=(m,a,n).$$

We deduce that $(m,a,n)\ \mathcal{R}^S\ (m,b,q)$.

We have shown in fact that

$$(m,a,n)\ \mathcal{R}^S\ (m,b,q) \text{ if and only if } m=p \text{ and } a\ \mathcal{R}^T\ b. \tag{5.6.5}$$

A dual argument establishes that

$$(m,a,n)\ \mathcal{L}^S\ (m,b,q) \text{ if and only if } n=q \text{ and } a\ \mathcal{L}^T\ b. \tag{5.6.6}$$

Now suppose that $(m,a,n)\ \mathcal{D}^S\ (p,b,q)$, so that there exists (r,c,s) in S such that $(m,a,n)\ \mathcal{R}^S\ (r,c,s)$ and $(r,c,s)\ \mathcal{L}^S\ (p,b,q)$. By (5.6.5) and

(5.6.6) it follows that $a\ \mathcal{R}^T\ c$ and $c\ \mathcal{L}^T\ b$ (and $r = m$, $s = q$). Hence $a\ \mathcal{D}^T\ b$.

Conversely, if $a\ \mathcal{D}^T\ b$ then for some c in T we have $a\ \mathcal{R}^T\ c$ and $c\ \mathcal{L}^T\ b$. Hence, for all m, n, p, q in $\mathbf{N}^0$,

$$(m, a, n)\ \mathcal{R}^S\ (m, c, q), \quad (m, c, q)\ \mathcal{L}^S\ (p, b, q),$$

and so $(m, a, n)\ \mathcal{D}^S\ (p, b, q)$, as required.

(3) Certainly every (m, e, m) (where e is an idempotent of T) is idempotent in S. For an arbitrary (m, a, n) in S, we have

$$(m, a, n)^2 = (m - n + t, (a\theta^{t-n})(a\theta^{t-m}), n - m + t),$$

where $t = \max(m, n)$. Hence (m, a, n) can be idempotent only if $m = n$. Then, since $(m, a, m)^2 = (m, a^2, m)$, it follows that the only idempotents in S are the elements (m, e, m) for which e is idempotent in T.

(4) Suppose first that T is an inverse semigroup. Then it is easy to verify that each (m, a, n) in S has an inverse (n, a^{-1}, m). Thus S is regular. If (m, e, m) and (n, f, n) are two idempotents in S (with (say) $m \geq n$), then

$$\left.\begin{aligned}(m, e, m)(n, f, n) &= (m, e(f\theta^{m-n}), m),\\ (n, f, n)(m, e, m) &= (m, (f\theta^{m-n})e, m).\end{aligned}\right\} \tag{5.6.7}$$

Now $f\theta^{m-n}$ is an idempotent in T—indeed if $m > n$ we must have $f\theta^{m-1} = 1$. So certainly $e(f\theta^{m-n}) = (f\theta^{m-n})e$, and we deduce that idempotents commute in S.

Conversely, suppose that S is an inverse semigroup. Let $(m, a, n)^{-1} = (p, b, q)$. Then

$$(m, a, n)(p, b, q) = (m - n + t, (a\theta^{t-n})(b\theta^{t-p}), q - p + t)$$

(with $t = \max(n, p)$) is an idempotent $\mathcal{R}^S$-equivalent to (m, a, n) and $\mathcal{L}^S$-equivalent to (p, b, q). It follows that

$$m = m - n + t = q - p + t = q,$$

and so $n = p$ $(= t)$ and $m = q$. The inverse property now gives

$$(m, a, n) = (m, a, n)(n, b, m)(m, a, n) = (m, aba, n),$$
$$(n, b, m) = (n, b, m)(m, a, n)(n, b, m),$$

and so $aba = a$, $bab = b$ in T. Hence T is regular. Finally, if e, f are idempotents in T, then the commuting of the idempotents $(0, e, 0)$ and $(0, f, 0)$ in S immediately implies that $ef = fe$. Hence T is an inverse semigroup. □

As a consequence of this theorem, if T is a group with identity e (so that θ is simply an endomorphism of T) then $BR(T, \theta)$ becomes a bisimple inverse semigroup with idempotents (m, e, m) $(m \in \mathbf{N}^0)$. From (5.6.7) it is easy to see, moreover, that

$$(0, e, 0) > (1, e, 1) > (2, e, 2) > \cdots;$$

thus $BR(T, \theta)$ is a bisimple inverse ω-semigroup.

In fact the converse also holds. We summarize the situation in a theorem as follows:

Theorem 5.6.7 *Let G be a group and let θ be an endomorphism of G. Let $S = BR(G, \theta)$ be the Bruck–Reilly extension of G determined by θ. Then S is a bisimple inverse ω-semigroup. Conversely, every bisimple inverse ω-semigroup is isomorphic to some $BR(G, \theta)$.*

Proof We have already shown the direct half. Let S be a bisimple inverse semigroup whose semilattice of idempotents is

$$E = C_\omega = \{e_0, e_1, e_2, \ldots\}.$$

Then, by Example 5.4.3, T_E is isomorphic to the bicyclic semigroup first encountered in Section 1.6. That is,

$$T_E = \{\alpha_{m,n} : m, n \in \mathbf{N}^0\},$$

where $\alpha_{m,n}$ is the unique isomorphism from Ee_m onto Ee_n given by

$$e_k\alpha_{m,n} = e_{k-m+n} \quad (k \geq m)$$

and where

$$\alpha_{m,n}\alpha_{p,q} = \alpha_{m-n+t,q-p+t},$$

with $t = \max(n, p)$. By Theorem 5.4.4, there is a morphism $\phi : S \to T_E$ whose kernel is μ. In fact, by Corollary 5.6.5, $\ker\phi = \mathcal{H}$.

A typical $\mathcal{H}$-class of S is

$$H_{m,n} = \{a \in S : aa^{-1} = e_m,\ a^{-1}a = e_n\},$$

and each element a of $H_{m,n}$ maps by ϕ to an element α_a of T_E whose domain is $Eaa^{-1} = Ee_m$ and whose image is $Ea^{-1}a = Ee_n$. There is only one such element, and so

$$H_{m,n}\phi = \alpha_{m,n}.$$

It follows that, within S,

$$H_{m,n}H_{p,q} \subseteq H_{m-n+t,q-p+t}. \tag{5.6.8}$$

Let us denote the group $H_{0,0}$, with identity e_0, by G. Choose and fix an element a in $H_{0,1}$. Then $a^{-1} \in H_{1,0}$, by the standard eggbox argument. (See Section 2.3.) Also $a^2 \in H_{0,2}$ by (5.6.8), and more generally by induction we find that

$$a^n \in H_{0,n}, \quad a^{-n} \in H_{n,0} \quad (n \in \mathbf{N}^0),$$

(where a^0 is defined as e_0). Notice also that, for all n in $\mathbf{N}^0$,

$$a^n a^{-n} = e_0, \quad a^{-n} a^n = e_n.$$

	L_0	L_1	$\cdots$	L_n
R_0	e_0	a		a^n
R_1	a^{-1}	e_1		
$\vdots$				
R_m	a^{-m}			$H_{m,n}$

Lemma 5.6.8 *The map*

$$g \mapsto a^{-m} g a^n \quad (g \in G)$$

is a bijection from $G = H_{0,0}$ *onto* $H_{m,n}$*, with inverse given by*

$$x \mapsto a^m x a^{-n} \quad (x \in H_{m,n}).$$

Proof This is a simple application of Green's Lemmas (Lemmas 2.2.1 and 2.2.2). We have $e_0 a^n = a^n$ and $a^n a^{-1} = e_0$, and so the map $g \mapsto ga^n$ is a bijection from $G = H_{0,0}$ onto $H_{0,n}$, with inverse $y \mapsto ya^{-n}$. Also, $a^{-m}e_0 = a^{-m}$ and $a^m a^{-m} = e_0$, and so the map $y \mapsto a^{-m}y$ is a bijection from $H_{0,n}$ onto $H_{m,n}$, with inverse $x \mapsto a^m x$. Combining these two bijections, we obtain the required result. □

Returning now to the proof of Theorem 5.6.7, we see that, once a is chosen, we have a bijection $\psi : S \to \mathbf{N}^0 \times G \times \mathbf{N}^0$ given by

$$(a^{-m} g a^n)\psi = (m, g, n). \tag{5.6.9}$$

For each g in G, using (5.6.8), we have

$$ag \in H_{0,1} H_{0,0} = H_{0,1},$$

and so, by Lemma 5.6.8, there is a unique expression $a^0 g' a^1 = g'a$ for ag. Accordingly, we define a map $\theta : G \to G$ by the rule

$$ag = (g\theta)a \quad (g \in G). \tag{5.6.10}$$

For all g_1, g_2 in G,

$$\begin{aligned}[(g_1g_2)\theta]a = a(g_1g_2) &= (ag_1)g_2 = [(g_1\theta)a]g_2 \\ &= (g_1\theta)(ag_2) = (g_1\theta)[(g_2\theta)a] = [(g_1\theta)(g_2\theta)]a;\end{aligned}$$

hence, postmultiplying by a^{-1}, and noting that $aa^{-1} = e_0$, we find that

$$(g_1g_2)\theta = (g_1\theta)(g_2\theta).$$

Thus θ is an endomorphism of G.

Notice next that

$$a^2g = a(ag) = a(g\theta)a = (g\theta^2)a^2;$$

indeed, by induction, we have

$$a^ng = (g\theta^n)a^n \quad (n \in \mathbf{N}^0),$$

where θ^0 is interpreted as the identity map of G. Also, from (5.6.10) we can deduce that

$$g^{-1}a^{-1} = a^{-1}(g\theta)^{-1}$$

for every g in G. Hence, changing notation, we deduce that

$$ga^{-1} = a^{-1}(g\theta) \quad (g \in G),$$

and, more generally, we have

$$ga^{-n} = a^{-n}(g\theta^n) \quad (g \in G,\ n \in \mathbf{N}^0).$$

Lemma 5.6.8 yields a unique expression of the type $a^{-m}ga^n$ for every element of S. We are now in a position to describe the manner in which two such expressions multiply. If $n \geq p$, then

$$(a^{-m}ga^n)(a^{-p}ha^q) = a^{-m}g(a^{n-p}h)a^q = a^{-m}g(h\theta^{n-p})a^{q-p+n},$$

while if $n \leq p$ then

$$(a^{-m}ga^n)(a^{-p}ha^q) = a^{-m}(ga^{-(p-n)})ha^q = a^{-(m-n+p)}(g\theta^{p-n})ha^q.$$

We can summarize these two formulae as follows:

$$(a^{-m}ga^n)(a^{-p}ha^q) = a^{-(m-n+t)}(g\theta^{t-n})(h\theta^{t-p})a^{q-p+t},$$

where $t = \max(n, p)$. Thus the bijection $\psi : S \to \mathbf{N}^0 \times G \times \mathbf{N}^0$ given by (5.6.9) is an isomorphism from S onto the Bruck–Reilly extension $BR(G, \theta)$. □

5.7 SIMPLE INVERSE SEMIGROUPS

We now briefly describe a theory for simple inverse semigroups that parallels in many respects the theory we have developed for the bisimple case. In the bisimple case we were able to show that a semilattice E can be the semilattice of idempotents of a bisimple inverse semigroup if and only if it is uniform. The first stage in generalizing the theory must be to give an analogous result for simple inverse semigroups.

We begin with

Lemma 5.7.1 *Let S be an inverse semigroup with semilattice of idempotents E. Then S is simple if and only if*

$$(\forall e, f \in E)(\exists g \in E)\ [g \leq f \text{ and } e\,\mathcal{D}\,g]. \qquad (5.7.1)$$

Proof Suppose first that S is a simple inverse semigroup with semilattice of idempotents E, and consider two elements e, f of E. Since $e \mathcal{J} f$, there exist x, y in S such that $e = xfy$. If g is defined as $fyex$, then

$$g^2 = fye(xfy)ex = fye^3x = fyex = g,$$

and so $g \in E$. It is clear that $fg = g$, and so $g \leq f$. Denoting $x^{-1}e$ by z, we see that

$$xz = xx^{-1}e = xx^{-1}xfy = xfy = e,$$

and so $e \mathcal{L} z$. Also

$$\begin{aligned} zx &= x^{-1}ex = x^{-1}e^2x = x^{-1}xfyex = x^{-1}xg \\ &= gx^{-1}x = fyexx^{-1}x = fyex = g, \end{aligned}$$

and

$$\begin{aligned} gx^{-1} &= gx^{-1}xx^{-1} = x^{-1}xgx^{-1} = x^{-1}xfyexx^{-1} \\ &= x^{-1}e^2xx^{-1} = x^{-1}xx^{-1}e = x^{-1}e = z, \end{aligned}$$

and so $z \mathcal{R} g$. Hence $e \mathcal{D} g$.

Conversely, suppose that we have the property (5.7.1), let S be an inverse semigroup with semilattice of idempotents E, and let e, f be arbitrary elements of E. Then there exists g in E such that $g \leq f$ and $e \mathcal{D} g$, from which we deduce that

$$J_e = J_g \leq J_f.$$

Equally, by interchanging the rôles of e and f, we have an idempotent h such that

$$J_f = J_h \leq J_e.$$

It follows that $e \mathcal{J} f$ for every pair e, f of idempotents in S, and hence (since every element of S is $\mathcal{J}$-equivalent—indeed $\mathcal{R}$- or $\mathcal{L}$-equivalent—to an idempotent) S is simple. □

This motivates the following definition: a semilattice E is said to be *subuniform* if

$$(\forall e, f \in E)(\exists g \in E)\ \big(g \leq f \text{ and } Ee \simeq Eg\big). \tag{5.7.2}$$

It is clear that every uniform semilattice is subuniform: simply take $g = f$ in the definition. Exercise 26 demonstrates that there exist subuniform semilattices that are not uniform.

Since $e \mathcal{D} g$ (for idempotents e, g in an inverse semigroup) implies that $Ee \simeq Eg$, we have in fact proved half of the following result:

Proposition 5.7.2 *A semilattice E is the semilattice of idempotents of a simple inverse semigroup if and only if it is subuniform.*

Proof To establish the other half, consider T_E, where E is subuniform. If $e, f \in E$ then $Ee \simeq Eg$ for some $g \leq f$, and so $e \mathcal{D} g$ in T_E. It follows that T_E is simple. □

If S (with semilattice of idempotents E) is a simple inverse semigroup which is also fundamental, then, by Theorem 5.4.5, S is in effect a full inverse subsemigroup of T_E. As in the bisimple case (Theorem 5.6.2), we can be a little more precise than this. In effect we defined an inverse subsemigroup S of T_E to be *transitive* if, for all e, f in E, there exists α in S such that

$$\operatorname{dom}\alpha = Ee, \quad \operatorname{im}\alpha = Ef.$$

We now define S to be *subtransitive* if, for all e, f in E, there exists α in S such that

$$\operatorname{dom}\alpha = Ee, \quad \operatorname{im}\alpha \subseteq Ef.$$

Then we have

Theorem 5.7.3 *If E is a subuniform semilattice, then every subtransitive inverse subsemigroup of T_E is a fundamental inverse semigroup with semilattice of idempotents isomorphic to E. Conversely, if S is a fundamental simple inverse semigroup with (necessarily subuniform) semilattice of idempotents E, then S is isomorphic to a subtransitive inverse subsemigroup of T_E.*

Proof Let S be a subtransitive inverse subsemigroup of T_E, where E is subuniform. If $e \in E$ then by subtransitivity there exists α in S with $\operatorname{dom}\alpha = Ee$ (and $\operatorname{im}\alpha \subseteq Ee$). Hence S contains the element $\alpha\alpha^{-1} = 1_{Ee}$. We have shown that S is full, and so, by Theorem 5.4.5, S is a fundamental inverse semigroup with semilattice of idempotents isomorphic to E. To show that S is simple, note that, for all e, f in E, there exists α in S such that

$$\operatorname{dom}\alpha = Ee, \quad \operatorname{im}\alpha = Eg \subseteq Ef.$$

Then $g \leq f$ and, since $\alpha\alpha^{-1} = 1_{Ee}$ and $\alpha^{-1}\alpha = 1_{Eg}$, it follows that e and g (strictly 1_{Ee} and 1_{Eg}) are $\mathcal{D}$-equivalent in S. By Lemma 5.7.1 we deduce that S is simple.

Conversely, suppose that S is a fundamental simple inverse semigroup. Then in the usual way—see Theorem 5.4.5—we have that $S \simeq S'$, where S' is a full inverse subsemigroup of T_E. Now S' is simple, and so, by Lemma 5.7.1, for all e, f in E there exists g in E such that $g \leq f$ and $1_{Ee}\ \mathcal{D}\ 1_{Eg}$ in S'. Hence there exists α in S' such that $\alpha\alpha^{-1} = 1_{Ee}$ and $\alpha^{-1}\alpha = 1_{Eg}$. Hence

$$\operatorname{dom}\alpha = Ee, \quad \operatorname{im}\alpha = Eg \subseteq Ef,$$

and so S' is subtransitive as required. □

In Corollary 5.6.3 we saw that, up to isomorphism, the only fundamental bisimple inverse ω-semigroup is the bicyclic semigroup. We can be just as precise in describing fundamental simple inverse ω-semigroups, but the answer is necessarily more complicated. Let us begin by defining, for $d = 1, 2, 3, \ldots$,

$$B_d = \{(m, n) \in \mathbf{N}^0 \times \mathbf{N}^0 : m \equiv n \pmod{d}\}.$$

We can regard B_d as a subset of the bicyclic semigroup $B = \mathbf{N}^0 \times \mathbf{N}^0$, and, regarded in this way, it is in fact a subsemigroup of B, for it is clear that from $m \equiv n \pmod{d}$ and $p \equiv q \pmod{d}$ we can deduce that

$$m - n + \max(n, p) \equiv q - p + \max(n, p) \pmod{d}.$$

B_d is even an inverse subsemigroup of B, since it is clear that if (m, n) belongs to B_d then the inverse (n, m) of (m, n) in B also belongs to B_d. It is also full, since, for every d, the idempotents (m, m) of B all belong to B_d.

We shall see shortly that B_d is simple, but first we describe the Green relations $\mathcal{R}$ and $\mathcal{L}$:

Lemma 5.7.4 *Let $(m, n), (p, q) \in B_d$. Then*

$$(m, n) \mathrel{\mathcal{R}} (p, q) \text{ if and only if } m = p,$$

$$(m, n) \mathrel{\mathcal{L}} (p, q) \text{ if and only if } n = q.$$

Proof By Proposition 5.1.2, if (m, n) and (p, q) are in B_d, then $(m, n) \mathrel{\mathcal{R}} (p, q)$ if and only if

$$(m, n)(m, n)^{-1} = (p, q)(p, q)^{-1},$$

that is, if and only if $(m, m) = (p, p)$, that is, if and only if $m = p$. The result for $\mathcal{L}$ is proved in exactly the same way. □

Two idempotents (m, m) and (n, n) in B_d are $\mathcal{D}$-equivalent in B_d if and only if there exists an element of B_d which is $\mathcal{R}$-equivalent to (m, m) and $\mathcal{L}$-equivalent to (n, n). This element can only be (m, n), and belongs to B_d if and only if $m \equiv n \pmod{d}$. We deduce that in B_d there are precisely d $\mathcal{D}$-classes, namely

$$D_{(0,0)}, D_{(1,1)}, \ldots, D_{(d-1,d-1)}.$$

We can now prove the analogue of Corollary 5.6.3:

Proposition 5.7.5 *Up to isomorphism, the only fundamental simple inverse ω-semigroups are the semigroups B_d $(d = 1, 2, 3, \ldots)$.*

Proof By virtue of Theorem 5.7.3, what we have to show is that the semigroups

$$B'_d = \{\alpha_{m,n} : m \equiv n \pmod{d}\}$$

are the only subtransitive inverse subsemigroups of the Munn semigroup

$$T_{C_\omega} = \{\alpha_{m,n} : m, n \in \mathbf{N}^0\}.$$

First, to show that B'_d is subtransitive, consider two idempotents e_m, e_n in C_ω. Certainly there exists $p \geq n$ such that $m \equiv p \pmod{d}$. Hence $\alpha_{m,p} \in B'_d$, and

$$\operatorname{dom} \alpha_{m,p} = Ee_m, \quad \operatorname{im} \alpha_{m,p} = Ee_p \subseteq Ee_n.$$

Conversely, suppose that S is a subtransitive inverse subsemigroup of T_{C_ω}. Then there exists α in S such that

$$\operatorname{dom} \alpha = Ee_0, \quad \operatorname{im} \alpha \subseteq Ee_1,$$

that is, such that $\operatorname{dom} \alpha = Ee_0$ and $\operatorname{im} \alpha = Ee_d$ for some $d \geq 1$. We deduce that S contains $\alpha_{0,d}$ for some $d \geq 1$. Suppose that d is the *least* positive integer for which this holds: thus $\alpha_{0,r} \in S$ with $0 \leq r < d$ implies that $r = 0$.

The standard multiplication rule (5.4.3) in T_{C_ω} gives $\alpha_{0,d}^2 = \alpha_{0,2d}$; indeed

$$\alpha_{0,d}^k = \alpha_{0,kd} \quad (k \geq 1).$$

Now, as remarked in the proof of Theorem 5.7.3, a subtransitive inverse subsemigroup of T_{C_ω} is necessarily full; hence for every $m, k \geq 0$,

$$\alpha_{m,m}\alpha_{0,kd} = \alpha_{m,m+kd} \in S.$$

It follows that $\alpha_{m+kd,m}$ is also in S, and so in fact $B'_d \subseteq S$. To show that this is in fact an equality, suppose, by way of contradiction, that there exists $\alpha_{m,n}$ in S for which $m \not\equiv n \pmod{d}$. Writing n as $m + kd + r$, with $0 < r < d$, by the division algorithm, we deduce that S contains

$$\alpha_{m,m+kd+r}\alpha_{m+kd,m} = \alpha_{m,m+r}.$$

If $m = 0$ we have already obtained our contradiction. If $m > 0$ we next observe that S contains

$$\alpha_{m-1,m-1+d}\alpha_{m,m+r}\alpha_{m-1+d,m-1} = \alpha_{m-1,m-1+r}.$$

We can continue this argument until we reach the false statement that $\alpha_{0,r} \in S$, and we are forced to conclude that $S = B'_d$. □

In the last section we observed that the Bruck–Reilly extension $BR(T, \theta)$ of a *group* T is a bisimple ω-semigroup. More generally, we now take T to be a semilattice of groups—a Clifford semigroup—of a special form. More precisely, let T be the union of disjoint groups $G_0, G_1, \ldots, G_{d-1}$ and, for $i = 0, 1, \ldots, d-2$, let $\gamma_i : G_i \to G_{i+1}$ be a morphism. Then, for all i, j in $\{0, 1, \ldots, d-1\}$ such that $i < j$ there is a morphism $\alpha_{i,j} : G_i \to G_j$ defined by

$$\alpha_{i,j} = \gamma_i\gamma_{i+1} \cdots \gamma_{j-1}.$$

If, for each i in $\{0, 1, \ldots, d-1\}$, we define $\alpha_{i,i}$ to be the identity map of G_i, then we certainly have

$$\alpha_{i,j}\alpha_{j,k} = \alpha_{i,k}$$

whenever $i \leq j \leq k$, and so we have a Clifford semigroup $T = \mathcal{S}(Y; G_i; \alpha_{i,j})$, in which the semilattice Y is the chain $\{0, 1, \ldots, d-1\}$. We denote the idempotents of T by $e_0, e_1, \ldots, e_{d-1}$ (the identity elements of $G_0, G_1, \ldots, G_{d-1}$, respectively), and note that, in T,

$$e_0 > e_1 > \cdots > e_{d-1}.$$

The element e_0 is the identity of G. We shall refer to T as a *finite chain of groups* (of *length* d).

Let $T = \mathcal{S}(Y; G_i; \alpha_{i,j})$ be such a semigroup, and let $S = BR(T, \theta)$, where θ is a morphism from T into G_0, the group of units of T. By Proposition 5.6.6, S is a simple inverse semigroup in which the $\mathcal{D}$-classes are the subsets $\mathbf{N}^0 \times G_i \times \mathbf{N}^0$ ($i = 0, 1, \ldots, d-1$). For two idempotents (m, e_i, m) and (m, e_j, m) in S, it is clear that $(m, e_i, m) \leq (m, e_j, m)$ if and only if $e_i \leq e_j$ in T, that is, if and only if $i \geq j$. More generally, if we consider two idempotents (m, e_i, m) and (n, e_j, n) for which $m > n$, then $e_j\theta^{m-n} = e_0$, the identity of T, and so

$$(m, e_i, m)(n, e_j, n) = (m, e_i(e_j\theta^{m-n}), m) = (m, e_i, m),$$

and so $(m, e_i, m) > (n, e_j, n)$ irrespective of the values of i and j. In summary, the idempotents of S form a chain

$$\begin{aligned} &(0, e_0, 0) > (0, e_1, 0) > \cdots > (0, e_{d-1}, 0)\\ > &(1, e_0, 1) > (1, e_1, 1) > \cdots > (1, e_{d-1}, 1)\\ > &(2, e_0, 2) > (2, e_1, 2) > \cdots > (2, e_{d-1}, 2)\\ > &\cdots \end{aligned}$$

Thus $S = BR(T, \theta)$ is a simple inverse ω-semigroup.

More remarkably, we can show that every simple inverse ω-semigroup is of this form:

Theorem 5.7.6 *Let T be a finite chain of groups of length d (≥ 1). If θ is a morphism from T into the group of units of T then the Bruck–Reilly extension $BR(T, \theta)$ of T determined by θ is a simple inverse ω-semigroup with d $\mathcal{D}$-classes. Conversely, every simple inverse ω-semigroup is isomorphic to one of this type.*

Proof It remains to prove the converse half. Let S be a simple inverse ω-semigroup whose semilattice of idempotents is

$$E = C_\omega = \{f_0, f_1, f_2, \ldots\}.$$

(The change in notation from e to f is made in order to avoid confusion with the notation used above in describing the chain of groups $\mathcal{S}(Y; G_i; \alpha_{i,j})$.)

By Theorem 5.4.4 there is a morphism $\phi : S \to T_E$ mapping each a in S to the element $\alpha_a : Eaa^{-1} \to Ea^{-1}a$ of T_E, where

$$x\alpha_a = a^{-1}xa \quad (x \in Eaa^{-1}).$$

The kernel of ϕ is μ, which in this case coincides with $\mathcal{H}$, by Corollary 5.6.5. The subsemigroup $S\phi$ of T_E is a subtransitive inverse subsemigroup of T_E, and so, as shown in the proof of Proposition 5.7.5, $S\phi$ coincides with B_d for some $(d \geq 1)$. We write

$$S\phi = \{\alpha_{m,n} : m, n \in \mathbf{N}^0,\ m \equiv n \ (\mathrm{mod}\, d)\},$$

where $\alpha_{m,n} : Ef_m \to Ef_n$ is given as usual by

$$f_k\alpha_{m,n} = f_{k-m+n} \quad (k \geq m),$$

and where

$$\alpha_{m,n}\alpha_{p,q} = \alpha_{m-n+t,q-p+t} \quad \big(t = \max(n,p)\big).$$

As in the bisimple case we define

$$H_{m,n} = \{a \in S : aa^{-1} = f_m,\ a^{-1}a = f_n\};$$

then either $H_{m,n} = \emptyset$ or $H_{m,n}$ is an $\mathcal{H}$-class of S. The essential difference between this case and the bisimple case now appears, for the former situation can arise here. Indeed we can say precisely when it does arise, for if $H_{m,n} \neq \emptyset$ then $H_{m,n}\phi = \alpha_{m,n} \in S\phi$; hence $H_{m,n} \neq \emptyset$ if and only if $m \equiv n$ (mod d). We easily deduce that the semigroup S has precisely d $\mathcal{D}$-classes

$$D^0, D^1, \ldots, D^{d-1},$$

where

$$D^i = \bigcup\{H_{m,n} : m \equiv n \equiv i \ (\mathrm{mod}\, d)\} = \bigcup\{H_{pd+i,qd+i} : p, q \in \mathbf{N}^0\}.$$

Lemma 5.7.7 *For $i = 0, 1, \ldots, d-1$, the $\mathcal{D}$-class D^i is a bisimple inverse ω-semigroup with identity element f_i.*

Proof Since $H_{m,n}\phi = \alpha_{m,n}$ (whenever $m \equiv n$ (mod d)), we can deduce from the multiplication formula for the elements $\alpha_{m,n}$ that

$$H_{m,n}H_{p,q} \subseteq H_{m-n+t,q-p+t} \quad \big(t = \max(n,p)\big). \tag{5.7.3}$$

If $H_{m,n}, H_{p,q} \subseteq D^i$, that is, if $m \equiv n \equiv i$ (mod d) and $p \equiv q \equiv i$ (mod d), then

$$m - n + t \equiv q - p + t \equiv t \equiv i \ (\mathrm{mod}\, d)$$

and so $H_{m,n}H_{p,q} \subseteq D^i$. Thus D^i is a subsemigroup of S, and is even an inverse subsemigroup, since it is clear that $H_{m,n} \subseteq D^i$ if and only if $H_{n,m} \subseteq D^i$.

Of the $\mathcal{H}$-classes

$$H_{pd+i,qd+i} = \{a \in S : aa^{-1} = f_{pd+i},\ a^{-1}a = f_{qd+i}\}$$

that make up D^i, those that contain idempotents are precisely the ones for which $p = q$. Indeed the $\mathcal{H}$-class $H_{pd+i,pd+i}$ contains the idempotent f_{pd+i}. The idempotents of D^i form an infinite chain

$$f_i > f_{d+i} > f_{2d+i} > \cdots,$$

and so D^i is an inverse ω-semigroup.

Finally, all the idempotents D^i are $\mathcal{D}$-equivalent in D^i, since for all p, q in $\mathbf{N}^0$ there exists an element a in D^i (in fact in $H_{pd+i,qd+i}$) such that $aa^{-1} = f_{pd+i}$, $a^{-1}a = f_{qd+i}$. Thus D^i is a bisimple inverse ω-semigroup.

□

Returning now to the proof of Theorem 5.7.6, we observe first that the group of units of D^i is $H_{i,i}$, and for simplicity of notation we shall denote this by G_i. The identity of G_i is the idempotent f_i. Let

$$T = \bigcup_{i=o}^{d-1} G_i.$$

If $i, j \in \{0, 1, \ldots, d-1\}$, with $i \leq j$, then by (5.7.3) and by the fact that $H_{m,n}^{-1} = H_{n,m}$ we have

$$G_iG_j \subseteq G_j, \quad G_jG_i \subseteq G_j,$$
$$G_i^{-1} = G_i.$$

Hence T is an inverse subsemigroup of S. Since it is evidently also completely regular, T is in fact a Clifford semigroup in which the idempotents form the finite chain

$$f_0 > f_1 > f_2 > \cdots > f_{d-1}.$$

Thus T is a finite chain of groups.

In view of the result stated in Lemma 5.7.7, it would seem reasonable to apply Theorem 5.6.7 to each D^i, deducing that $D^i \simeq BR(G_i, \theta_i)$, where θ_i is an endomorphism of G_i. This in fact is almost what we do. Our method of proof in Theorem 5.6.7 suggests that we ought for each i to choose an element a_i in each $H_{i,d+i}$ and then express each element of D^i as $a_i^{-m}g_ia_i^n$ ($\in H_{md+i,nd+i}$), with $m, n \in \mathbf{N}^0$ and $g_i \in G_i$. It pays, however, not to choose the elements a_i independently, but to make them all depend on the chosen element a_0 in $H_{0,d}$.

Simplifying the notation, let us choose and fix an element a in $H_{0,d}$. Then $a^{-1} \in H_{d,0}$, and

$$aa^{-1} = f_0, \quad a^{-1}a = f_d.$$

For $i = 0, 1, \ldots, d-1$,

$$f_ia \in H_{i,i}H_{0,d} = H_{i,d+i}.$$

Since f_i is the identity of D^i it follows that $(f_ia)f_i = f_ia$, from which we easily deduce that $(f_ia)^2 = f_ia^2$. Indeed, by induction, we have

$$(f_ia)^n = f_ia^n \quad (n \in \mathbf{N}),$$

and by taking inverses we easily deduce that

$$(f_ia)^{-n} = a^{-n}f_i \quad (n \in \mathbf{N}).$$

Using f_ia for a_i in the way indicated above, we express each element of D^i uniquely as

$$(f_ia)^{-m}g_i(f_ia)^n = a^{-m}f_ig_if_ia^n = a^{-m}g_ia^{-n},$$

(since f_i is the identity of G_i), where $m, n \in \mathbf{N}^0$ and $g_i \in G_i$. The element $a^{-m}g_ia^n$ belongs to the $\mathcal{H}$-class $H_{md+i,nd+i}$. (There is no difficulty about zero values of m and n if we interpret a^0 formally as 1.)

We thus have a bijection Ψ from S onto $\mathbf{N}^0 \times T \times \mathbf{N}^0$, defined by

$$(a^{-m}g_ia^n)\Psi = (m, g_i, n).$$

It remains to find a morphism $\theta : T \to G_0$ such that Ψ becomes an isomorphism.

Again following the lead given by the bisimple case, we note that for all g_i in G_i $(i = 0, 1, \ldots, d-1)$,

$$ag_i \in H_{0,d}H_{i,i} \subseteq H_{0,d};$$

hence ag_i is uniquely expressible as $g_0'a$, with $g_0' \in G_0$. We can regard the resulting formula $ag_i = g_0'a$ as defining a map $\phi_i : G_i \to G_0$:

$$ag_i = (g_i\phi_i)a. \tag{5.7.4}$$

Since the groups $G_0, G_1, \ldots, G_{d-1}$ are disjoint, we can use the maps $\phi_0, \phi_1, \ldots, \phi_{d-1}$ thus obtained to piece together a map $\theta : T \to G_0$:

$$t\theta = t\phi_i \quad (t \in G_i,\ i = 0, 1, \ldots, d-1).$$

This θ is the morphism we are looking for. To show the morphic property, consider x, y in T, with $x \in G_i$, $y \in G_j$ and $i \leq j$. Then $xy \in G_j$ and so, by (5.7.4),

$$a(xy) = [(xy)\phi_j]a = [(xy)\theta]a.$$

On the other hand,

$$a(xy) = (ax)y = (x\phi_i)ay = (x\phi_i)(y\phi_j)a = [(x\theta)(y\theta)]a.$$

Thus $[(xy)\theta]a = [(x\theta)(y\theta)]a$, and so, since $aa^{-1} = f_0$, we deduce that $(xy)\theta = (x\theta)(y\theta)$, as required.

Exactly as in the proof of Theorem 5.6.7 we have that, for every x in T and every k in $\mathbf{N}^0$,

$$a^kx = (x\theta^k)a^k, \quad xa^{-k} = a^{-k}(x\theta^k),$$

where a^0 is interpreted as f_0 and $x\theta^0$ as x.

Now let $x, y \in T$ and let $m, n, p, q \in \mathbf{N}^0$. If $n \geq p$, then

$$(a^{-m}xa^n)(a^{-p}ya^q) = a^{-m}x(y\theta^{n-p})a^{q-p+n},$$

while if $n \leq p$,

$$(a^{-m}xa^n)(a^{-p}ya^q) = a^{m-n+p}(x\theta^{p-n})ya^q.$$

If we write $t = \max(n, p)$ then we can combine these two statements into a single statement:

$$(a^{-m}xa^n)(a^{-p}ya^q) = a^{-(m-n+t)}(x\theta^{t-n})(y\theta^{t-p})a^{q-p+t},$$

and it is now clear that Ψ is an isomorphism from S onto $BR(T, \theta)$. □

5.8 REPRESENTATIONS OF INVERSE SEMIGROUPS

In this section we describe an analysis of representations of inverse semigroups by partial one-one mappings. The idea is essentially the same as in group theory, where one shows that every transitive representation of a group by permutations is equivalent to a representation by permutations of the cosets of a certain subgroup of the group. (See, for example, M. Hall (1959).)

Let us begin by reminding ourselves that a *representation* of an inverse semigroup S is a morphism ϕ of S into some symmetric inverse semigroup $\mathcal{I}_X$. If ϕ is one-one we call the representation *faithful*. The particular representation $\phi : S \to \mathcal{I}_S$ described in Theorem 5.1.7 we shall call the *Vagner–Preston* representation of S.

Theorem 5.1.4 assures us that for every representation $\phi : S \to \mathcal{I}_X$ the image $S\phi$ is an inverse subsemigroup of $\mathcal{I}_X$. Accordingly, let us turn our attention to an arbitrarily chosen inverse subsemigroup H of $\mathcal{I}_X$, where X is a non-empty set. Let τ_H be the relation

$$\{(a, b) \in X \times X : (\exists \kappa \in H)\ a \in \operatorname{dom} \kappa \text{ and } a\kappa = b\}.$$

We call τ_H the *transitivity relation* of H. Then we have

Lemma 5.8.1 *If H is an inverse subsemigroup of a symmetric inverse semigroup $\mathcal{I}_X$, then τ_H is a symmetric and transitive relation on X.*

Proof If $(x, y) \in \tau_H$ then $x\kappa = y$ for some κ in H. Hence $y \in \operatorname{im} \kappa = \operatorname{dom} \kappa^{-1}$ and $y\kappa^{-1} = x$. Since $\kappa^{-1} \in H$, we deduce that $(y, x) \in \kappa$. Suppose now that $(x, y), (y, z) \in \tau_H$. Then $x\kappa = y$, $y\lambda = z$ for some κ, λ in H. It follows that $x \in \operatorname{dom} \kappa\lambda$ and $x(\kappa\lambda) = z$. Thus $(x, z) \in \tau_H$ as required. □

We cannot in general assert that τ_H is an equivalence relation, since there may exist elements x in X that are not in the domain of any κ in H. What this amounts to is that there exist elements x on X for which $(x, x) \notin \tau_H$, for if we had $(x, y) \in \tau_H$ for some y then by symmetry and transitivity we could deduce that $(x,x) \in \tau_H$. We define the *domain* of τ_H to be the set

$$X\tau_H = \{x \in X : (x,x) \in \tau_H\},$$

and then we can conclude that τ_H is an equivalence relation on its domain $X\tau_H$. We say that H is an *effective* inverse subsemigroup of $\mathcal{I}_X$ if $X\tau_H = X$; in this case τ is an equivalence relation on X.

The τ_H-classes in $X\tau_H$ are called the *transitivity classes* of H, and H is called *transitive* if τ_H is the universal relation on $X\tau_H$. Thus H is effective and transitive if and only if for all a, b in X there exists κ in H such that $a\kappa = b$. We say that $\phi : S \to \mathcal{I}_X$ is an *effective [transitive]* representation if $S\phi$ is an effective [transitive] inverse subsemigroup of $\mathcal{I}_X$.

Let $\{X_i : i \in I\}$ be a family of pairwise disjoint sets and let

$$X = \bigcup_{i\in I} X_i.$$

Let S be an inverse semigroup, and suppose that for each i in I we have a representation $\phi_i : S \to \mathcal{I}_{X_i}$. For each s in S we may regard the one-one partial map $s\phi_i$ as a subset of $X_i \times X_i$. Then

$$\bigcup_{i\in I} s\phi_i$$

is a partial one-one map of X, whose domain is $\bigcup_{i\in I} \operatorname{dom} \phi_i$; we denote this map by ϕ, and call it the *sum* of the representations ϕ_i. We write

$$\phi = \bigoplus_{i\in I} \phi_i. \tag{5.8.1}$$

If $I = \{1, 2, \ldots, n\}$ we write

$$\phi = \phi_1 \oplus \phi_2 \oplus \cdots \oplus \phi_n.$$

Because the definition (5.8.1) is in terms of set-theoretic union, the infinite commutative and associative laws hold for $\oplus$.

If $\phi : S \to \mathcal{I}_X$ and $\psi : S \to \mathcal{I}_Y$ are representations of an inverse semigroup S, we say that ϕ and ψ are *equivalent* if there exists a bijection $\theta : X \to Y$ with the property that, for each s in S,

$$s\psi = \{(x\theta, x'\theta) \in Y \times Y : (x,x') \in s\phi\}.$$

In other words, $\operatorname{dom}(s\psi) = \big(\operatorname{dom}(s\phi)\big)\theta$, and, for all x in $\operatorname{dom}(s\phi)$,

$$\big(x(s\phi)\big)\theta = (x\theta)(s\psi).$$

This amounts in practice to saying that the two representations differ 'in name only'. It is important to be able to replace a representation by an equivalent one when we want to form the sum $\phi_1 \oplus \phi_2$ of representations $\phi_1 : S \to \mathcal{I}_{X_1}$ and $\phi_2 : S \to \mathcal{I}_{X_2}$ in the case where X_1 and X_2 are not disjoint. We cannot do this under the rules of addition, but what we can do is to form a representation $\phi_1 \oplus \psi_2$, where $\psi_2 : S \to \mathcal{I}_{Y_2}$ is a representation equivalent to ϕ_2 and where $X_1 \cap Y_2 = \emptyset$—and for most purposes this is just as good.

The importance of effective and transitive representations lies in the following result:

Theorem 5.8.2 *Every effective representation of an inverse semigroup S is the sum of a uniquely determined family of effective transitive representations of S.*

Proof Let $\phi : S \to \mathcal{I}_X$ be an effective representation of S, and denote $S\phi$ by H. The transitivity relation τ_H is an equivalence relation on X; let us denote the τ_H-classes by X_i $(i \in I)$. Then

$$\tau_H = \bigcup_{i \in I}(X_i \times X_i) \text{ and } \bigcup_{i \in I} X_i = X.$$

We now define for each i in I the representation ϕ_i by the rule that

$$s\phi_i = s\phi \cap (X_i \times X_i) \quad (s \in S).$$

That is, the partial map $s\phi_i$ is the restriction to $\operatorname{dom}(s\phi) \cap X_i$ of the partial map $s\phi$. Since X_i is a transitivity class of H it is then automatic that $\operatorname{im}(s\phi_i) \subseteq X_i$; indeed

$$\operatorname{im}(s\phi_i) = \operatorname{im}(s\phi) \cap X_i.$$

To show that ϕ_i is a representation, observe that for any choice of s and t in S,

$$(st)\phi_i = (st)\phi \cap (X_i \times X_i) = (s\phi)(t\phi) \cap (X_i \times X_i).$$

Now $(x, y) \in (s\phi)(t\phi)$ if and only if there exists z in X such that $(x, z) \in s\phi$ and $(z, y) \in t\phi$. In fact, if x and y are in X_i then it follows, since X_i is a transitivity class of $H = S\phi$, that z also belongs to X_i. Hence

$$(s\phi)(t\phi) \cap (X_i \times X_i) = \big(s\phi \cap (X_i \times X_i)\big)\big(t\phi \cap (X_i \times X_i)\big),$$

and so $(st)\phi_i = (s\phi_i)(t\phi_i)$, as required.

Next, each ϕ_i is transitive and effective, since for each (x, y) in $X_i \times X_i$ there exists, by definition of τ_H, an element $s\phi$ such that $(x, y) \in s\phi$. Thus $(x, y) \in s\phi \cap (X_i \times X_i) = s\phi_i$.

For each s in S,

$$\begin{aligned}\bigcup\{s\phi_i : i \in I\} &= \bigcup\{s\phi \cap (X_i \times X_i) : i \in I\} \\ &= (s\phi) \cap \bigcup\{X_i \times X_i : i \in I\} \\ &= s\phi, \quad \text{since } s\phi \subseteq \bigcup\{X_i \times X_i : i \in I\}.\end{aligned}$$

Hence ϕ is the sum of the representations ϕ_i.

Finally, to show that the family $\{\phi_i : i \in I\}$ is unique, suppose that ϕ is the sum of a family $\{\psi_j : j \in J\}$ of effective transitive representations

$\psi_j : S \to Y_j$, where X is the union of the pairwise disjoint sets Y_j. The transitivity classes of the sum $\phi = \bigoplus_{j \in J} \psi_j$ are the sets Y_j, and, for each s in S,

$$(s\phi) \cap (Y_j \times Y_j) = s\psi_j \, .$$

It thus follows that the sets Y_j are just the sets X_i in some order, and that each ψ_j is equal to the appropriate ϕ_i. □

The next stage in the investigation is to discover more about effective transitive representations. We begin by describing a particular kind of representation of an inverse semigroup associated with an inverse subsemigroup H that is *closed*, in the sense introduced in Section 5.2.

The first step in this process is to generalize the group-theoretic notion of 'right coset'. Let S be an inverse semigroup and let H be a (not necessarily closed) inverse subsemigroup of S. If $s \in S$ then the subset Hs may fail to contain s, but it certainly contains s if $ss^{-1} \in H$. We define a *right coset* of H to be a set Hs ($s \in S$) for which $ss^{-1} \in H$. Even if H is closed, the right coset Hs may fail to be closed. The closure $(Hs)\omega$ of a right coset Hs will be called a *right ω-coset* of H. Notice carefully that a subset Hs is not deemed to be a right coset unless $ss^{-1} \in H$.

Among the right ω-cosets of H is $H\omega$ itself, for, as is not hard to verify, $H\omega = (Hh)\omega$ for every h in H. Let $\mathcal{X}$ denote the set of all right ω-cosets of H.

Proposition 5.8.3 *Let H be an inverse subsemigroup of an inverse semigroup S, and let $(Ha)\omega$, $(Hb)\omega$ be right ω-cosets of H. Then the following statements are equivalent:*

(1) $(Ha)\omega = (Hb)\omega$; (2) $ab^{-1} \in H\omega$;
(3) $a \in (Hb)\omega$; (4) $b \in (Ha)\omega$.

Proof (1) ⇒ (2). Suppose that $(Ha)\omega = (Hb)\omega$. Then

$$a = aa^{-1}a \in Ha \subseteq (Ha)\omega = (Hb)\omega,$$

and so $a \geq hb$ for some h in H. It follows that

$$ab^{-1} \geq hbb^{-1} \in H,$$

and so $ab^{-1} \in H\omega$.

(2) ⇒ (3). If $ab^{-1} \in H\omega$ then $ab^{-1} \geq h$ for some h in H, and so

$$a \geq ab^{-1}b \geq hb.$$

Thus $a \in (Hb)\omega$.

(3) ⇒ (1). Suppose that $a \in (Hb)\omega$. Then $a \geq hb$ for some h in H. For all s in $(Ha)\omega$ we have $s \geq ka$ for some k in H, and it follows that $s \geq khb$. Since $kh \in H$, we have shown that $s \in (Hb)\omega$, and hence that $(Ha)\omega \subseteq (Hb)\omega$. To establish the reverse conclusion, suppose that

$t \in (Hb)\omega$, so that $t \geq kb$ for some k in H. From our assumption that $a \geq hb$ we deduce by Proposition 5.2.1(8) that

$$hb = hbb^{-1}h^{-1}a,$$

and hence that

$$t \geq kb \geq kh^{-1}hb = kh^{-1}hbb^{-1}h^{-1}a \in Ha.$$

Thus $t \in (Ha)\omega$, as required.

The incorporation of (4) into the list of equivalent statements follows immediately when we observe that condition (1) is symmetrical in a and b. □

Now define, for each s in S, an element $s\phi_H$ of $\mathcal{I}_\mathcal{C}$ by

$$s\phi_H = \{((Hx)\omega, (Hxs)\omega) : (Hx)\omega, (Hxs)\omega \in \mathcal{C}\}. \tag{5.8.2}$$

Thus the domain of $s\phi_H$ is

$$\{(Hx)\omega \in \mathcal{C} : (Hxs)\omega \in \mathcal{C}\},$$

and, for each $(Hs)\omega$ in this domain,

$$((Hx)\omega)(s\phi_H) = (Hxs)\omega.$$

To verify that $s\phi_H$ does indeed belong to $\mathcal{I}_\mathcal{C}$, notice first that it is a well-defined partial mapping of $\mathcal{C}$. For suppose that $(Hx)\omega = (Hy)\omega$, and that $(Hxs)\omega \in \mathcal{C}$. Then $xx^{-1}, yy^{-1}, xss^{-1}x^{-1} \in H$. By Proposition 5.8.3, $xy^{-1} \in H\omega$, and so

$$(xs)(ys)^{-1} = xss^{-1}y^{-1} = xx^{-1}xss^{-1}y^{-1} = xss^{-1}x^{-1}.xy^{-1} \in H\omega.$$

Thus, again by Proposition 5.8.3, $(Hxs)\omega = (Hys)\omega$.

Also, $s\phi_H$ is one-one, for if $(Hxs)\omega = (Hys)\omega$, then

$$xy^{-1} \geq xss^{-1}y^{-1} \in H\omega.$$

Thus $xy^{-1} \in (H\omega)\omega = H\omega$, and so $(Hx)\omega = (Hy)\omega$, as required.

We have shown that $s\phi_H \in \mathcal{I}_\mathcal{C}$. Indeed we have

Proposition 5.8.4 *If H is a closed inverse subsemigroup of an inverse semigroup S, then the mapping $\phi_H : S \to \mathcal{I}_\mathcal{C}$ defined by* (5.8.2) *is an effective transitive representation of S.*

Proof To show that ϕ_H is a representation, consider an element

$$((Hx)\omega, (Hxst)\omega)$$

of $(st)\phi_H$. Then xx^{-1} and $xstt^{-1}s^{-1}x^{-1}$ belong to H. Now $xss^{-1}x^{-1} \geq xstt^{-1}s^{-1}x^{-1}$, and so $xss^{-1}x^{-1} \in H\omega = H$. Hence

$$((Hx)\omega, (Hxs)\omega) \in s\phi_H, \quad ((Hxs)\omega, (Hxst)\omega) \in t\phi_H,$$

which gives that $\big((Hx)\omega, (Hxst)\omega\big) \in (s\phi_H)(t\phi_H)$. We have shown that $(st)\phi_H \subseteq (s\phi_H)(t\phi_H)$. To show the converse, suppose that

$$\big((Hx)\omega, (Hy)\omega\big) \in (s\phi_H)(t\phi_H),$$

so that there exists $(Hz)\omega$ in $\mathcal{C}$ such that

$$\big((Hx)\omega, (Hz)\omega\big) \in s\phi_H, \quad \big((Hz)\omega, (Hy)\omega\big) \in t\phi_H.$$

Then $(Hz)\omega = (Hxs)\omega$ and $(Hy)\omega = (Hxst)\omega$, and we deduce that $\big((Hx)\omega, (Hy)\omega\big) \in (st)\phi_H$, as required. Thus ϕ_H is a representation.

To show that ϕ_H is effective and transitive, we establish that for any $(Hx)\omega$, $(Hy)\omega$ in $\mathcal{C}$ there exists s in S, namely $s = x^{-1}y$, such that $\big((Hx)\omega, (Hy)\omega\big) \in s\phi_H$. First, notice that

$$(xs)(xs)^{-1} = xx^{-1}yy^{-1}xx^{-1} \in H,$$

since xx^{-1} and yy^{-1} are in H by assumption. Thus $(Hxs)\omega \in \mathcal{C}$. Then observe that

$$(xs)y^{-1} = xx^{-1}yy^{-1} \in H = H\omega,$$

and so, by Proposition 5.8.3, $(Hy)\omega = (Hxs)\omega$. □

We now show that *every* effective transitive representation $\psi : S \to \mathcal{I}_{\mathcal{C}}$ of an inverse semigroup S is equivalent to one of type ϕ_H:

Proposition 5.8.5 *Let X be a set, and let $\psi : S \to \mathcal{I}_X$ be an effective transitive representation of the inverse semigroup S. Let z be an arbitrary fixed element of X and let*

$$H = \{s \in S : (z,z) \in s\psi\}.$$

Then H is a closed inverse subsemigroup of S, and ψ is equivalent to the representation ϕ_H defined by (5.8.2).

Proof Certainly H is a subsemigroup, since $(z,z) \in s\psi$ and $(z,z) \in t\psi$ implies that $(z,z) \in (s\psi)(t\psi) = (st)\psi$. Since ψ is a morphism we have $s^{-1}\psi = (s\psi)^{-1}$; hence

$$s \in H \;\Rightarrow\; (z,z) \in s\psi \;\Rightarrow\; (z,z) \in s^{-1}\psi \;\Rightarrow\; s^{-1} \in H.$$

Thus H is an inverse subsemigroup of S.

To show that H is closed, suppose that $k \in H\omega$, so that $k \geq h$ for some h in H. Since the order relation is defined in terms of multiplication, it follows that $k\psi \geq h\psi$ in $\mathcal{I}_X$, and so, by the remarks preceding Proposition 5.2.1, $k\psi \supseteq h\psi$ (as subsets of $X \times X$). Now $(z,z) \in h\psi$, since $h \in H$, and it is then immediate that $(z,z) \in k\psi$. Thus $k \in H$, and so H is closed.

To show that ψ is equivalent to ϕ_H we must begin by defining a bijection θ from X into the set $\mathcal{C}$ of right ω-cosets of H. If $x \in X$, then, since ψ is effective and transitive, there exists a_x in S such that $(z,x) \in a_x\psi$. The element a_x of S then necessarily has the property that $a_x a_x^{-1} \in H$, for we have

$$(z, x) \in a_x\psi \text{ and } (x, z) \in a_x^{-1}\psi,$$

and so $(z, z) \in a_x a_x^{-1}\psi$. Thus $(Ha_x)\omega$ is a right ω-coset of H, characterized as follows:

$$(Ha_x)\omega = \{s \in S : (z, x) \in s\psi\}. \tag{5.8.3}$$

To see this, suppose first that $s \in (Ha_x)\omega$. Then $s \geq ha_x$ for some h in H, and so $s\psi \supseteq (h\psi)(a_x\psi)$ (as subsets of $X \times X$). Now $(z, z) \in h\psi$ and $(z, x) \in a_x\psi$, and so it follows that $(z, x) \in s\psi$, as required.

Conversely, if $(z, x) \in s\psi$, then from $(x, z) \in a_x^{-1}\psi$ it follows that $sa_x^{-1} \in H$. Hence $s \in (Ha_x)\omega$, by Proposition 5.8.3.

The element a_x is not uniquely determined by x. However, if b_x also has the property that $(z, x) \in b_x\psi$, then (5.8.3) assures us that $b_x \in (Ha_x)\omega$, and hence, by Proposition 5.8.3,

$$(Ha_x)\omega = (Hb_x)\omega.$$

It is therefore correct to say that the right ω-coset $(Ha_x)\omega$ is uniquely determined by x.

We now define a map $\theta : X \to \mathcal{C}$ by the rule that

$$x\theta = (Ha_x)\omega \quad (x \in X),$$

where a_x is any element of S such that $(z, x) \in a_x\psi$. The conclusion of the previous paragraph implies that θ is well defined. Moreover, it is immediate that θ is one-one, for if $x\theta = x'\theta = (Ha)\omega$ then

$$(z, x) \in a\psi, \quad (z, x') \in a\psi,$$

from which it immediately follows that $x = x'$, since $a\psi$ is a partial one-one map of X. In fact θ is also onto, for if $(Ha)\omega$ is a right ω-coset of H, then by definition $aa^{-1} \in H$, and so

$$(z, z) \in (aa^{-1})\psi = (a\psi)(a^{-1}\psi).$$

Hence there exists x in X such that $(z, x) \in a\psi$, $(x, z) \in a^{-1}\psi$, and it follows that $x\theta = (Ha)\omega$.

Suppose finally that x, y in X and s in S are such that $(x, y) \in s\psi$; we must show that $(x\theta, y\theta) \in s\phi_H$. Write $x\theta = (Ha)\omega$, $y\theta = (Hb)\omega$, so that $(z, x) \in a\psi$, $(z, y) \in b\psi$. Then from

$$(z, x) \in a\psi,\ (x, y) \in s\psi,\ (y, x) \in s^{-1}\psi \text{ and } (x, z) \in a^{-1}\psi$$

we deduce that $(z, z) \in (ass^{-1}a^{-1})\psi$, and hence that $ass^{-1}a^{-1} \in H$. Again, from $(z, x) \in a\psi$ and $(x, y) \in s\psi$ we deduce that $(z, y) \in (as)\psi$. From (5.8.3) and Proposition 5.8.3 it now follows that $(Hb)\omega = \big(H(as)\big)\omega$, and hence

$$(x\theta, y\theta) = \big((Ha)\omega, (Has)\omega\big) \in s\phi_H.$$

We have shown that $s\psi \subseteq s\phi_H$. Suppose now that

$$(x\theta, y\theta) = \big((Ha)\omega, (Has)\omega\big) \in s\phi_H.$$

By the definition of θ, we have $(z,x) \in a\psi$, $(z,y) \in (as)\psi$. It follows that

$$(x,y) \in (a^{-1}as)\psi \subseteq s\psi. \qquad \square$$

We summarize the results in a theorem as follows:

Theorem 5.8.6 *Let S be an inverse semigroup. Then every effective representation of S is uniquely expressible as a sum of effective, transitive representations ψ_i $(i \in I)$, each of which is equivalent to ϕ_{H_i} for some closed inverse subsemigroup H_i of S.* $\square$

5.9 E-UNITARY INVERSE SEMIGROUPS

Let S be an inverse semigroup with semilattice E of idempotents. We say that S is *E-unitary* if E is a unitary subsemigroup of E, that is, if, for all e in E and s in S,

$$es \in E \ \Rightarrow\ s \in E\,,$$
$$se \in E \ \Rightarrow\ s \in E\,.$$

In fact to show that S is E-unitary we need only show one of these implications: if, for example, we assume the first implication, and if we suppose that $e, se \in E$, then ses^{-1} and $(ses^{-1})s$ $(= se)$ both belong to E, and we deduce that $s \in E$.

Further characterizations of E-unitary inverse semigroups are available:

Proposition 5.9.1 *Let S be an inverse semigroup with semilattice E of idempotents. Let σ be the minimum group congruence on S. The following statements are equivalent:*

(1) *S is E-unitary;*
(2) *$E\omega$ $(= \operatorname{Ker}\sigma) = E$;*
(3) *$\sigma \cap \mathcal{L} = 1_S$.*

Proof (1) $\Rightarrow$ (2). Suppose that S is E-unitary, and let $a \in E\omega$. Then $e \leq a$ for some e in E and so $e = fa$ for some f in E. From $f, fa \in E$ we deduce that $a \in E$.

(2) $\Rightarrow$ (3). Suppose that $E\omega = E$, and let a, b in S be such that $a\ \sigma\ b$ and $a\ \mathcal{L}\ b$. Thus $a^{-1}a = b^{-1}b$ by Proposition 5.1.2 and $ab^{-1} \in E\omega = E$ by Theorem 5.3.5. Hence $ab^{-1} = (ab^{-1})^{-1}ab^{-1} = ba^{-1}ab^{-1}$. It now follows that

$$a = aa^{-1}a = ab^{-1}b = ba^{-1}ab^{-1}b = bb^{-1}bb^{-1}b = b,$$

and so $\sigma \cap \mathcal{L} = 1_S$.

(3) $\Rightarrow$ (1). Suppose that $\sigma \cap \mathcal{L} = 1_S$, and suppose that $e, ea \in E$. Certainly $(a, a^{-1}a) \in \mathcal{L}$. Also $a \geq ea \in E$ and so $a \in E\omega$. From Theorem

5.3.5 it follows that $a \; \sigma \; i$ for every i in E. In particular $(a, a^{-1}a) \in \sigma$. Thus $(a, a^{-1}a) \in \sigma \cap \mathcal{L}$, and so $a = a^{-1}a \in E$. □

Remark The one-sided nature of condition (3) is only apparent. It is not hard to show that the condition $\sigma \cap \mathcal{R} = 1_S$ is equivalent.

We now describe a recipe, seemingly rather arbitrary, for constructing E-unitary inverse semigroups. First, let $\mathcal{X}$ be a set furnished with a partial order relation $\leq$, and let $\mathcal{Y}$ be a subset of $\mathcal{X}$ such that:

(P1) $\mathcal{Y}$ is a *lower semilattice* with respect to $\leq$, in the sense that for every A and B in $\mathcal{Y}$ there is a greatest lower bound $A \wedge B$, also in $\mathcal{Y}$;

(P2) $\mathcal{Y}$ is an *order ideal*, in the sense that, for all A, X in $\mathcal{X}$,

$$A \in \mathcal{Y} \text{ and } X \leq A \;\Rightarrow\; X \in \mathcal{Y}.$$

We shall say that a bijection $\alpha : \mathcal{X} \to \mathcal{X}$ is an *order automorphism* if, for all A, B in $\mathcal{X}$, $A \leq B \;\Leftrightarrow\; \alpha A \leq \alpha B$. (It is convenient at this stage to write the mapping symbol on the left.) Let $\operatorname{Aut} \mathcal{X}$ be the group of all *order automorphisms* of $\mathcal{X}$. Now let G be a group, and suppose that G acts on $\mathcal{X}$ by order automorphisms. By this we mean that there is a morphism θ from G into the group $\operatorname{Aut} \mathcal{X}$ of order automorphisms of $\mathcal{X}$. The fact that θ is a morphism means that, for all g, h in G and for all A in $\mathcal{X}$,

$$(g\theta)\big((h\theta)A\big) = \big((g\theta)(h\theta)\big)A = \big((gh)\theta\big)A.$$

For the most part we can simplify the notation by suppressing any explicit mention of θ, and think of G itself as acting on $\mathcal{X}$ (on the left). The fact that it does so by means of order automorphisms is expressed by saying that

$$gA = gB \;\Leftrightarrow\; A = B, \quad (\forall B \in \mathcal{X})(\exists A \in \mathcal{X}) gA = B,$$

$$A \leq B \;\Leftrightarrow\; gA \leq gB,$$

and the morphism property becomes

$$g(hA) = (gh)A.$$

(Here $g, h \in G$, $A, B \in \mathcal{X}$.) It is precisely in this last property that the difference between a left and a right action occurs: the action of gh results from the action of h followed by the action of g. Were we to use a right action then the two actions would be in the opposite order.

For each g in G the order automorphism property of the map $A \mapsto gA$ implies that if A and B are in $\mathcal{X}$ and if $A \wedge B$ exists then $gA \wedge gB$ exists, and

$$gA \wedge gB = g(A \wedge B).$$

Suppose finally that the triple $(G, \mathcal{X}, \mathcal{Y})$ has the properties:

(P3) $G\mathcal{Y} = \mathcal{X}$; that is, for every X in $\mathcal{X}$ there exist g in G and A in $\mathcal{Y}$ such that $gA = X$;

(P4) for all g in G, $g\mathcal{Y} \cap \mathcal{Y} \neq \emptyset$.

Let us say that a triple $(G, \mathcal{X}, \mathcal{Y})$ having the properties (P1)–(P4) is a *McAlister triple.* Given such a triple, let

$$S = \mathcal{M}(G, \mathcal{X}, \mathcal{Y}) = \{(A, g) \in \mathcal{Y} \times G : g^{-1}A \in \mathcal{Y}\}, \tag{5.9.1}$$

and define a multiplication on S by the rule that

$$(A, g)(B, h) = (A \wedge gB, gh).$$

We begin by verifying that S is closed under this operation. First, $g^{-1}A \wedge B$ exists, since both $g^{-1}A$ and B are in $\mathcal{Y}$. Hence $g(g^{-1}A \wedge B) = A \wedge gB$ exists, and is in $\mathcal{Y}$, since $A \wedge gB \leq A \in \mathcal{Y}$. Also

$$(gh)^{-1}(A \wedge gB) = h^{-1}g^{-1}A \wedge h^{-1}B \leq h^{-1}B \in \mathcal{Y},$$

and so $(gh)^{-1}(A \wedge gB) \in \mathcal{Y}$.

Next, the operation is associative, for if $(A, g), (B, h), (C, k) \in S$, then

$$(A \wedge gB) \wedge (gh)C = A \wedge g(B \wedge hC).$$

Thus S is a semigroup. It is even a regular semigroup, since for each (A, g) in S the element $(g^{-1}A, g^{-1})$ is also in S, and

$$(A, g)(g^{-1}A, g^{-1})(A, g) = (A, 1)(A, g) = (A, g),$$
$$(g^{-1}A, g^{-1})(A, g)(g^{-1}A, g^{-1}) = (g^{-1}A, g^{-1})(A, 1) = (g^{-1}A, g^{-1}).$$

Moreover, it is evident that (A, g) is idempotent if and only if $g = 1$, and for any two idempotents $(A, 1)$ and $(B, 1)$, it is easy to see that

$$(A, 1)(B, 1) = (B, 1)(A, 1) = (A \wedge B, 1).$$

Thus S is an inverse semigroup.

Finally, notice that in the inverse semigroup S the natural order relation is given by

$$(A, g) \leq (B, h) \text{ if and only if } A \leq B \text{ and } g = h.$$

Consequently, $(A, g) \in E\omega$ if and only if there exists an idempotent $(B, 1)$ such that $(B, 1) \leq (A, g)$. This can happen only if $g = 1$, that is, only if $(A, g) \in E$. Thus $E\omega = E$, and so S is an E-unitary inverse semigroup, by Proposition 5.9.1.

We have proved the easy half of the following result.

Theorem 5.9.2 *Let $(G, \mathcal{X}, \mathcal{Y})$ be a McAlister triple. Then $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is an E-unitary inverse semigroup. Conversely, every E-unitary inverse semigroup is isomorphic to one of this kind.*

Proof Let S be an E-unitary inverse semigroup. We must produce a McAlister triple $(G, \mathcal{X}, \mathcal{Y})$ such that $S \simeq \mathcal{M}(G, \mathcal{X}, \mathcal{Y})$.

Let $G = S/\sigma$, and for each s in S define s^0 in $E \times G$ by

$$s^0 = (s^{-1}s, s\sigma).$$

Notice that, by Proposition 5.9.1(3) above,

$$s^0 = t^0 \Rightarrow s = t.$$

If T is a subset of S, then T^0 will mean $\{t^0 : t \in T\}$.

Let $\mathbf{R}$ be the set of all principal right ideals of S. Since every principal right ideal of S can be written uniquely as eS, with e in E, the set $\mathbf{R}$ is in one-one correspondence with E, and so also is $\mathcal{Y}$, which we define by

$$\mathcal{Y} = \{A^0 : A \in \mathbf{R}\}.$$

The correspondence is even a semilattice isomorphism between E and $(\mathcal{Y}, \cap)$, since

$$efS = eS \cap fS.$$

There is an obvious action of G on $E \times G$ given by

$$g(e, h) = (e, gh).$$

We define $\mathcal{X}$ to be $G\mathcal{Y}$, by which we mean $\{gA^0 : g \in G, A^0 \in \mathcal{Y}\}$. Both $\mathcal{Y}$ and $\mathcal{X}$ are sets of subsets of $E \times G$, partially ordered by inclusion, and $\mathcal{Y} \subseteq \mathcal{X}$.

The action of G on $\mathcal{Y}$ extends naturally to $\mathcal{X}$: for all g in G and all hA^0 in $\mathcal{X}$ (with h in G and A^0 in $\mathcal{Y}$), we define

$$g(hA^0) = (gh)A^0.$$

Of course this definition conceals a difficulty, for in general the expression of an element of $\mathcal{X}$ as hA^0 is not unique. However, if $hA^0 = kB^0$ then within $\mathcal{Y}$ we have

$$A^0 = (h^{-1}k)B^0 = \big((gh)^{-1})(gk)\big)B^0,$$

and so $(gh)A^0 = (gk)B^0$ as required.

Next we have a lemma:

Lemma 5.9.3 *With the above definitions, $\mathcal{Y}$ is an order ideal of $\mathcal{X}$.*

Proof Let $g(eS)^0$ be a typical element of $\mathcal{X}$, and suppose that $g(eS)^0 \subseteq (fS)^0$ for some f in E. Among the elements of $(eS)^0$ is the element $(e, 1)$, where 1 is the identity of the group G; so in particular we have

$$(e, g) = g(e, 1) = u^0 = (u^{-1}u, u\sigma)$$

for some u in fS. Hence $g = u\sigma$ and $e = u^{-1}u$, and so

$$\begin{aligned} g(eS)^0 &= (u\sigma)(u^{-1}uS)^0 \\ &= (u\sigma)\{(u^{-1}us)^0 : s \in S\} \\ &= (u\sigma)\{(s^{-1}u^{-1}us, (u^{-1}us)\sigma) : s \in S\} \end{aligned}$$

$$= \{((us)^{-1}us, (us)\sigma) \colon s \in S\} = (uS)^0.$$

Thus $g(eS)^0 \in \mathcal{Y}$, as required. □

In effect we have now verified that the triple $(G, \mathcal{X}, \mathcal{Y})$ satisfies conditions (P1) to (P3). To show that it is a McAlister triple we need to show that it also satisfies (P4). Suppose therefore that $g \in G$. Let s in S be such that $s\sigma = g^{-1}$. Then

$$\begin{aligned} g(sS)^0 &= (s\sigma)^{-1}\{((st)^{-1}(st), (st)\sigma) : t \in S\} \\ &= \{((s^{-1}st)^{-1}(s^{-1}st), (s^{-1}st)\sigma) : t \in S\} \\ &= (s^{-1}sS)^0 \in \mathcal{Y}. \end{aligned}$$

Thus $g\mathcal{Y} \cap \mathcal{Y} \neq \emptyset$, and so $(G, \mathcal{X}, \mathcal{Y})$ is a McAlister triple as required.

It remains to show that $S \simeq \mathcal{M}(G, \mathcal{X}, \mathcal{Y})$. A pair $((sS)^0, g)$ in $\mathcal{Y} \times G$ belongs to $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ if and only if $g^{-1}(sS^0) \in \mathcal{Y}$. From the argument of the last paragraph this holds if $g = s\sigma$. We now proceed to show that these are in effect the only circumstances in which it holds: we show that if $g^{-1}(sS)^0 \in \mathcal{Y}$ then there exists v in S such that $((sS)^0, g) = ((vS)^0, v\sigma)$. So suppose that $g^{-1}(sS)^0 = (tS)^0 \in \mathcal{Y}$. Then in particular $g^{-1}(ss^{-1})^0 = u^0$ for some u in tS. That is,

$$(ss^{-1}, g^{-1}) = u^0 = (u^{-1}u, u\sigma),$$

and so $ss^{-1} = u^{-1}u$, $g = (u^{-1})\sigma$. We have shown that

$$((sS)^0, g) = ((u^{-1}S)^0, (u^{-1})\sigma);$$

hence u^{-1} is the element v we have been looking for.

We now define a map $\phi : S \to \mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ by

$$s\phi = ((sS)^0, s\sigma) \quad (s \in S);$$

the conclusion of the last paragraph in effect is that ϕ is onto. It is clear that ϕ is also one-one, since (for all s, t in S)

$$\begin{aligned} s\phi = t\phi &\Rightarrow (sS)^0 = (tS)^0,\ s\sigma = t\sigma \\ &\Rightarrow sS = tS,\ s\sigma = t\sigma \\ &\Rightarrow (s,t) \in \sigma \cap \mathcal{R} \\ &\Rightarrow s = t, \end{aligned}$$

by the remark following the proof of Proposition 5.9.1.

It remains to show that ϕ is a morphism. In $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$,

$$((sS)^0, s\sigma)((tS)^0, t\sigma) = ((sS)^0 \cap (s\sigma).(tS)^0, (st)\sigma),$$

and so we must show that

$$(sS)^0 \cap (s\sigma)(tS)^0 = (stS)^0.$$

Suppose first that $(e,g) \in (sS)^0 \cap (s\sigma)(tS)^0$. Then

$$(e,g) = (su)^0 = (s\sigma)(tv)^0$$

for some u, v in S. Hence $(s\sigma)^{-1}(su)^0 = (tv)^0$ and so

$$\begin{aligned}(tv)^0 &= \big((su)^{-1}(su),(s^{-1}su)\sigma\big) = \big((s^{-1}su)^{-1}(s^{-1}su),(s^{-1}su)\sigma\big)\\ &= (s^{-1}su)^0\,.\end{aligned}$$

It follows that $tv = s^{-1}su$. Hence $su = stv$ and so $(e,g) = (stv)^0 \in (stS)^0$ as required.

Conversely, suppose that $(e,g) \in (stS)^0$. Then certainly $(e,g) \in (sS)^0$, and $(e,g) = (stu)^0$ for some u in S. Let $v = (st)^{-1}(st)u$. Then $v\sigma = u\sigma$ and

$$(tv)^{-1}(tv) = (stu)^{-1}(stu).$$

Hence

$$\begin{aligned}(e,g) &= \big((stu)^{-1}(stu),(stu)\sigma\big) = \big((tv)^{-1}(tv),(stv)\sigma\big)\\ &= (s\sigma).\big((tv)^{-1}(tv),(tv)\sigma\big) \in (s\sigma)(tS)^0\,.\end{aligned}$$

Thus $(e,g) \in (sS)^0 \cap (s\sigma)(tS)^0$, exactly as required. □

Before proceeding to consider the isomorphism theorem associated with the McAlister structure theorem, we pause to make a formal record of some easily verified facts about the semigroup $\mathcal{M}(G,\mathcal{X},\mathcal{Y})$:

Proposition 5.9.4 *Let (A,g), (B,h) be elements of $S = \mathcal{M}(G,\mathcal{X},\mathcal{Y})$. Then*

(1) $(A,g)^{-1} = (g^{-1}A, g^{-1})$;
(2) $(A,g)\ \mathcal{R}\ (B,h)$ *if and only if* $A = B$;
(3) $(A,g)\ \mathcal{L}\ (B,h)$ *if and only if* $g^{-1}A = h^{-1}B$;
(4) $(A,g)\ \mathcal{D}\ (B,h)$ *if and only if there exists z in G such that* $zA = B$;
(5) $(A,g)\ \mathcal{J}\ (B,h)$ *if and only if there exist z, t in G such that* $zA \leq B$, $tB \leq A$;
(6) $(A,g) \leq (B,h)$ *if and only if* $A \leq B$ *and* $g = h$;
(7) $(A,g)\ \sigma(B,h)$ *if and only if* $g = h$.

Proof Most of these statements are obvious. To prove (5), suppose that $(A,g)\ \mathcal{J}\ (B,h)$. Then $(A,g) = (C,x)(B,h)(D,y)$, so that

$$A = C \wedge xB \wedge xhD, \quad g = xhy.$$

Thus $A \leq xB$, and if we put $z = x^{-1}$ we obtain $zA \leq B$ as required. Similarly we find t such that $tB \leq A$.

Conversely, suppose that we have elements z, t in G such that $zA \leq B$, $tB \leq A$. Then $(A,z^{-1}), (zA,z) \in S$, and

$$(A,1) = (A,z^{-1})(B,1)(zA,z).$$

Similarly

$$(B,1) = (B,t^{-1})(A,1)(tB,t),$$

and so $(A,1)\ \mathcal{J}\ (B,1)$. Since $(A,g)\ \mathcal{R}\ (A,1)$ and $(B,h)\ \mathcal{R}(B,1)$, it follows that $(A,g)\ \mathcal{J}\ (B,h)$. □

We now have the following isomorphism theorem:

Theorem 5.9.5 *Let $(G,\mathcal{X},\mathcal{Y})$, $(G',\mathcal{X}',\mathcal{Y}')$ be McAlister triples. Let $\theta : G \to G'$ be a group isomorphism and let $\psi : \mathcal{X} \to \mathcal{X}'$ be an order-isomorphism such that $\psi|_{\mathcal{Y}}$ is an isomorphism from the semilattice $\mathcal{Y}$ onto the semilattice $\mathcal{Y}'$. Suppose also that, for all g in G and X in $\mathcal{X}$,*

$$(gX)\psi = (g\theta)(X\psi). \tag{5.9.2}$$

Then the map $\phi : \mathcal{M}(G,\mathcal{X},\mathcal{Y}) \to \mathcal{M}(G',\mathcal{X}',\mathcal{Y}')$ defined by

$$(A,g)\phi = (A\psi, g\theta) \quad \big((A,g) \in \mathcal{M}(G,\mathcal{X},\mathcal{Y})\big)$$

is an isomorphism. Conversely, every isomorphism from $\mathcal{M}(G,\mathcal{X},\mathcal{Y})$ onto $\mathcal{M}(G',\mathcal{X}',\mathcal{Y}')$ is of this type.

Proof The direct half of the proof is a matter of routine verification and is omitted.

For brevity let us write $\mathcal{M}(G,\mathcal{X},\mathcal{Y}) = M$, $\mathcal{M}(G',\mathcal{X}',\mathcal{Y}') = M'$, and let us suppose that there is an isomorphism $\phi : M \to M'$. Then ϕ maps the idempotents of M isomorphically onto the idempotents of M' and so induces a semilattice isomorphism $\psi : \mathcal{Y} \to \mathcal{Y}'$, given by

$$(A,1)\phi = (A\psi, 1) \quad (A \in \mathcal{Y}).$$

Since G and G' are, respectively, the maximum group morphic images of M and M', there is an isomorphism $\theta : G \to G'$ such that the diagram

$$\begin{array}{ccc} M & \xrightarrow{\ \phi\ } & M' \\ \sigma^\natural \big\downarrow & & \big\downarrow \sigma^\natural \\ G & \xrightarrow[\ \theta\ ]{} & G' \end{array}$$

commutes. If we write $(A,g)\phi$ as (A',g'), we see that

$$g' = (A',g')\sigma^\natural = (A,g)\phi\sigma^\natural = (A,g)\sigma^\natural\theta = g\theta;$$

hence $(A,g)\phi = (A',g\theta)$ for some A' in $\mathcal{Y}'$. Now $(A,g)\ \mathcal{R}\ (A,1)$ in M, and so $(A,g)\phi\ \mathcal{R}\ (A,1)\phi$ in M'. That is, $(A',g\theta)\ \mathcal{R}\ (A\psi,1)$, and so

$$(A,g)\phi = (A\psi, g\theta)$$

as required. If $A, gA \in \mathcal{Y}$ then $(gA,g) = (A,g^{-1})^{-1}$. Hence

$$(gA,g)\phi = \big((A,g^{-1})\phi\big)^{-1},$$

and it follows that

$$(gA)\psi = (g\theta)(A\psi) \tag{5.9.3}$$

whenever A and gA are in $\mathcal{Y}$.

The obvious way to extend ψ from $\mathcal{Y}$ to $\mathcal{X}$ is to use the property (P3) to express each X in $\mathcal{X}$ in the form gA, with g in G and A in $\mathcal{Y}$, and to define

$$(gA)\psi = (g\theta)(A\psi).$$

To see that this is well-defined, suppose that $gA = hB$. Then $A = g^{-1}hB$ in $\mathcal{Y}$, and so

$$(A,1)\phi = (g^{-1}hB,1)\phi.$$

Hence, by (5.9.3),

$$A\psi = \big((g^{-1}h)\theta\big)(B\psi) = (g\theta)^{-1}(h\theta)(B\psi),$$

and so $(g\theta)(A\psi) = (h\theta)(B\psi)$ as required. It is now easy to verify that the map $\psi : \mathcal{X} \to \mathcal{X}'$ is a bijection, that it preserves order, and that, for all g in G and X in $\mathcal{X}$,

$$(gX)\psi = (g\theta)(X\psi). \qquad \square$$

The remainder of this section is devoted to a proof of a universal property of E-unitary inverse semigroups. Let S be an arbitrary inverse semigroup, and let G be a group. We say that an E-unitary inverse semigroup P is an *E-unitary cover of S over G* if

(1) $P/\sigma \simeq G$;

(2) there is an idempotent-separating morphism from P onto S.

We shall show shortly that there always exists an inverse semigroup P with these properties. It is, however, useful to introduce another concept at this stage. An inverse semigroup S with group of units G is said to be *factorizable* if

$$(\forall a \in S)(\exists g \in G)\ a \leq g.$$

(Notice that $S = EG$ for such a semigroup.) Then we have

Proposition 5.9.6 *Every inverse semigroup is embeddable in a factorizable inverse semigroup.*

Proof By the Vagner–Preston Theorem we may suppose that S is an inverse subsemigroup of $\mathcal{I}_X$ for some X. Suppose first that X is finite (which will be the case if S is finite). Then for each α in $\mathcal{I}_X$ we can deduce from $|\operatorname{dom}\alpha| = |\operatorname{im}\alpha|$ that $|X \setminus \operatorname{dom}\alpha| = |X \setminus \operatorname{im}\alpha|$, and so α can be extended (usually not uniquely) to a permutation γ of $\mathcal{I}_X$. Thus $\alpha \leq \gamma$, and so $\mathcal{I}_X$ is factorizable.

If X is infinite, with cardinality $\mathbf{k} \geq \aleph_0$, then we define $Y = X \cup X'$, where $X \cap X' = \emptyset$ and $|X'| = \mathbf{k}$. It is a routine matter to show that the subset T of $\mathcal{I}_Y$ defined by

$$T = \{\alpha \in \mathcal{I}_Y : (\exists \gamma \in \mathcal{G}_Y)\ \alpha \leq \gamma\}$$

is a factorizable inverse semigroup. Also T contains $\mathcal{I}_X$ (and hence also contains S), since for every α in $\mathcal{I}_X$ we have (within $\mathcal{I}_Y$)

$$|Y \setminus \operatorname{dom} \alpha| = |Y \setminus \operatorname{im} \alpha| = \mathbf{k},$$

and so α is extendable to a permutation of Y. □

We now have

Theorem 5.9.7 *Let S be an inverse semigroup. Let F be a factorizable inverse semigroup with group of units G and let $\theta : S \to F$ be a monomorphism. Then the inverse subsemigroup P of $S \times G$ defined by*

$$P = \{(s, g) \in S \times G : s\theta \leq g\}$$

is an E-unitary cover of S over G.

Proof First, it is a routine matter to verify that P is an inverse subsemigroup of the inverse semigroup $S \times G$. The idempotents of $S \times G$ are of the form $(e, 1)$, where 1 is the identity of G and e is an idempotent in S. Indeed every element of this form is actually an idempotent of P, since $e\theta$, being an idempotent in F, always has the property $e\theta \leq 1$. We have shown that $E_P = E_S \times \{1\}$.

To show that P is E-unitary, suppose that $(s, g)(e, 1) = (f, 1)$ in the semigroup P. Then $g = 1$ and so $s\theta \leq 1$. It follows that $s\theta$ is idempotent in F, and hence (since θ is one-one) that s is idempotent in S. We have shown that $(s, g) \in E_S \times \{1\} = E_P$, and it follows that P is E-unitary.

The minimum group congruence σ on P is now given by the rule that $(s, g)\ \sigma\ (t, h)$ if and only if $(st^{-1}, gh^{-1}) \in E_P$. (See Theorem 5.3.5.) This happens if and only if $g = h$, for in such a case we have $(st^{-1})\theta \leq 1$ and so $st^{-1} \in E_S$. It follows that $P/\sigma \simeq G$.

The projection morphism $\pi : P \to S$ given by

$$(s, g)\pi = s \quad \big((s, g) \in P\big)$$

maps onto S, by the factorizable property of F. It separates idempotents, since, for all e, f in E_S,

$$(e, 1)\pi = (f, 1)\pi \ \Rightarrow\ e = f.$$ □

5.10 FREE INVERSE MONOIDS

By a *free inverse monoid* on a non-empty set X we mean an inverse monoid FI_X together with a map $\theta : X \to FI_X$ with the property that, for every inverse monoid S and every map $\alpha : X \to S$, there is a unique morphism

$\overline{\alpha} : FI_X \to S$ such that $\theta\overline{\alpha} = \alpha$. A standard argument similar to that in Section 1.6 shows that such an object is unique if it exists. In one sense it is easy to see that the free inverse monoid exists for every X, and to describe it. Let

$$X' = \{x^{-1} : x \in X\}$$

be a set in one-one correspondence with X and disjoint from it. Let $Y = X \cup X'$ and let Y^* be the free monoid on Y. Define formal inverses for elements of Y^* by the rules

$$\begin{aligned} 1^{-1} &= 1, \\ (x^{-1})^{-1} &= x \quad (x \in X), \\ (y_1 y_2 \ldots y_n)^{-1} &= y_n^{-1} \ldots y_2^{-1} y_1^{-1} \quad (y_1, y_2, \ldots, y_n \in Y). \end{aligned}$$

Notice that $(w^{-1})^{-1} = w$ for all w in Y^*.

Let τ be the congruence on Y^* generated by the set

$$\mathbf{T} = \{(ww^{-1}w, w) : w \in Y^*\} \cup \{(ww^{-1}zz^{-1}, zz^{-1}ww^{-1}) : w, z \in Y^*\}.$$

Then we have

Theorem 5.10.1 *With the above definitions, Y^*/τ is the free inverse monoid on X.*

Proof To show that Y^*/τ is an inverse monoid we use the original definition in (5.1.1). Certainly we have two operations on Y^*/τ, given by

$$(w\tau)(z\tau) = (wz)\tau, \quad (w\tau)^{-1} = w^{-1}\tau.$$

The latter operation is in fact well-defined: from the nature of the relation $\mathbf{T}$ generating τ it is not hard to see that if a sequence of $\mathbf{T}$-transitions connects w and z then a modified series of $\mathbf{T}$-transitions connects w^{-1} and z^{-1}. Then, again from the nature of $\mathbf{T}$, we easily see that, for all $w\tau$, $z\tau$ in Y^*/τ,

$$(w\tau)(w^{-1}\tau)(w\tau) = w\tau, \quad \left((w\tau)^{-1}\right)^{-1} = w\tau,$$
$$(w\tau)(w\tau)^{-1}(z\tau)(z\tau)^{-1} = (z\tau)(z\tau)^{-1}(w\tau)(w\tau)^{-1}.$$

Hence Y^*/τ is an inverse monoid, in which the inverse of a typical element $(y_1 y_2 \ldots y_n)\tau$ is $(y_n^{-1} \ldots y_2^{-1} y_1^{-1})\tau$.

We have an obvious map $\theta : x \mapsto x\tau$ from X into Y^*/τ. Suppose now that S is an inverse monoid and that there is a map $\alpha : X \to S$. Then α extends to $Y = X \cup X'$ in an obvious way if we define $(x^{-1})\alpha = (x\alpha)^{-1}$ for every x in X. Since Y^* is the free monoid on Y we can define a monoid morphism $\hat{\alpha} : Y^* \to S$ by

$$(y_1 y_2 \ldots y_n)\hat{\alpha} = (y_1\alpha)(y_2\alpha) \ldots (y_n\alpha).$$

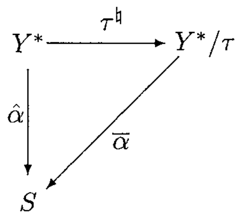
Then the inverse monoid property of S implies that

$$(ww^{-1}w)\hat{\alpha} = w\hat{\alpha}$$

for all w in Y^*, and

$$(ww^{-1}zz^{-1})\hat{\alpha} = (zz^{-1}ww^{-1})\hat{\alpha}$$

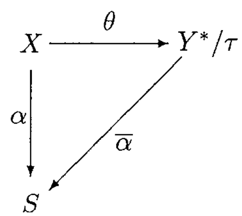
for all w, z in Y^*. It follows that the morphism $\hat{\alpha}$ factors through Y^*/τ, in the sense that there exists a morphism $\overline{\alpha} : Y^*/\tau \to S$ such that the diagram



commutes. The morphism $\overline{\alpha}$ is given by

$$\big((y_1y_2\dots y_n)\tau\big)\overline{\alpha} = (y_1\alpha)(y_2\alpha)\dots(y_n\alpha).$$

It is now clear that the diagram



(5.10.1)

commutes, since $(x\tau)\overline{\alpha} = x\alpha$ for every x in X.

Finally, we show that the morphism $\overline{\alpha}$ is unique with respect to the property of making the diagram (5.10.1) commute. If β is a morphism from Y^*/τ to S such that $\theta\beta = \alpha$, then $(x\tau)\beta = x\alpha$ and so, by the morphism property of β,

$$(x^{-1}\tau)\beta = (x\tau)^{-1}\beta = \big((x\tau)\beta\big)^{-1} = (x\alpha)^{-1}.$$

Hence

$$\big((y_1y_2\dots y_n)\tau\big)\beta = (y_1\alpha)(y_2\alpha)\dots(y_n\alpha)$$

for every $(y_1y_2\dots y_n)\tau$ in Y^*/τ, and so $\beta = \overline{\alpha}$. □

Remark By using Y^+ rather than Y^* throughout the above treatment we obtain the free inverse *semigroup* on X.

The description we have obtained for a free inverse monoid is not in practice very useful. We now give an alternative approach, based primarily on the work of Scheiblich (1973a,b). The idea is to express FI_X as $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ for a suitable McAlister triple $(G, \mathcal{X}, \mathcal{Y})$.

Let X be a non-empty set, and let $G = FG_X$ be the *free group* on the set X. The elements of G are 'group-reduced' words in the alphabet $Y = X \cup X'$, where a word is called *group-reduced* if, for each x in X, it contains no occurrences of xx^{-1} or $x^{-1}x$. (Among the group-reduced words is the empty word 1.) The product in G of two such words is obtained by taking their product in the free monoid Y^* and then excising (stage by stage if necessary) any occurrences of the forbidden subwords. Thus, for example, the product in FG_X of $x_1x_2x_3^{-1}$ and $x_3x_2^{-1}x_3^{-1}x_2$ is $x_1x_3^{-1}x_2$. We shall sometimes want to think of words of this type as elements of G and sometimes as 'group-reduced words' within Y^*, and it will sometimes be useful to make a notational distinction between the two. We shall denote the set of group-reduced words in Y^* by R.

We shall also attempt to avoid confusion by denoting the product in FG_X of two group-reduced words u and v by $u \cdot v$, while denoting their product in Y^* by uv. The unique group-reduced word obtained from a word w in Y^* by the group-reduction process will be denoted by $\overline{w}$. Notice that $\overline{w} = w$ if and only if w is group-reduced, and that for all w, z in Y^*,

$$\overline{wz} = \overline{w} \cdot \overline{z}.$$

To specify the set $\mathcal{Y}$ we need a new concept. Let $w = y_1y_2\dots y_n$ be a group-reduced word in Y^*. Then we define

$$w^{\downarrow} = \{1, y_1, y_1y_2, \dots, y_1y_2\cdots y_n\} \tag{5.10.2}$$

to be the set of all left factors of w (including 1 and w itself). A non-empty set A of group-reduced words is called *saturated* if, for every w,

$$w \in A \;\Rightarrow\; w^{\downarrow} \subseteq A.$$

(Although we require A to be non-empty, the case $A = \{1\}$ is a legitimate one.) It is obvious that, for every pair w, z of elements of R,

$$(w \cdot z)^{\downarrow} \subseteq w^{\downarrow} \cup w \cdot (z^{\downarrow}), \tag{5.10.3}$$

and only slightly less obvious that

$$w^{-1} \cdot (w^{\downarrow}) = (w^{-1})^{\downarrow} \tag{5.10.4}$$

for every w in R. Let

$$\mathcal{Y} = \{A \subseteq R : A \text{ is finite and saturated}\}. \tag{5.10.5}$$

Notice that $w^{\downarrow} \in \mathcal{Y}$ for every w in R.

Somewhat perversely—but there is a good reason for it—we define an order $\leq$ on $\mathcal{Y}$ by the rule that

$$A \leq B \text{ if and only if } A \supseteq B.$$

Then $\mathcal{Y}$ is a lower semilattice. (In effect we are observing that if $A, B \in \mathcal{Y}$ then $A \cup B \in \mathcal{Y}$.) The action of G on $\mathcal{Y}$, by contrast, is exactly as expected: for g in G and A in $\mathcal{Y}$ we define

$$g \cdot A = \{g \cdot w : w \in A\}.$$

We define

$$\mathcal{X} = \{g \cdot A : g \in G, A \in \mathcal{Y}\}.$$

Then $\mathcal{X}$ is a partially ordered set, where, as in $\mathcal{Y}$, we define $A \leq B$ to mean $A \supseteq B$. It is a routine matter to verify that G acts on $\mathcal{X}$ by order automorphisms.

We now show that $(G, \mathcal{X}, \mathcal{Y})$ is a McAlister triple. We have already observed that (P1) holds, and our construction of $\mathcal{X}$ ensures that (P3) also is satisfied. To show that $\mathcal{Y}$ is an order ideal of $\mathcal{X}$, let $Z = g \cdot A \in \mathcal{X}$, where $g \in G$ and $A \in \mathcal{Y}$, and suppose that $Z \leq B$, where $B \in \mathcal{Y}$. Thus $B \subseteq Z = g \cdot A$. Since B is saturated, we have in particular that $1 \in B \subseteq g \cdot A$. Hence $g^{-1} \in A$ and so, since A also is saturated, $(g^{-1})^{\downarrow} \subseteq A$. We deduce by (5.10.4) that

$$g^{\downarrow} = g \cdot (g^{-1})^{\downarrow} \subseteq g \cdot A = Z. \tag{5.10.6}$$

Now let $v \in Z$. Then $v = g \cdot u$, where $u \in A$. Since $u^{\downarrow} \subseteq A$ we may conclude from (5.10.3) and (5.10.6) that

$$v^{\downarrow} \subseteq g^{\downarrow} \cup g \cdot (u^{\downarrow}) \subseteq g \cdot A \subseteq Z.$$

Thus Z is saturated and so $Z \in \mathcal{Y}$. We have verified (P2).

To verify (P4), consider g in G. Then, by (5.10.4),

$$g \cdot (g^{-1})^{\downarrow} = g^{\downarrow} \in g \cdot \mathcal{Y} \cap \mathcal{Y}.$$

The E-unitary semigroup $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is given, in accordance with (5.9.1), by

$$\mathcal{M}(G, \mathcal{X}, \mathcal{Y}) = \{(A, g) \in \mathcal{Y} \times G : g^{-1} \cdot A \in \mathcal{Y}\}, \tag{5.10.7}$$

and the multiplication is given by

$$(A, g)(B, h) = (A \cup g \cdot B, gh). \tag{5.10.8}$$

The condition $g^{-1} \cdot A \in \mathcal{Y}$ in (5.10.7) can be expressed more simply:

$$g^{-1} \cdot A \in \mathcal{Y} \text{ if and only if } g \in A. \tag{5.10.9}$$

To see this, suppose first that $g^{-1} \cdot A \in \mathcal{Y}$. Then every $w = g^{-1} \cdot u$ in $g^{-1} \cdot A$ is such that $w^{\downarrow} \subseteq g^{-1} \cdot A$. In particular, since $1 \in w^{\downarrow}$, $1 = g^{-1} \cdot v$ for some v in A. Hence $g = v \in A$. Conversely, suppose that $g \in A$. Then $g^{\downarrow} \subseteq A$ and so, by (5.10.3) and (5.10.4), for every $w = g^{-1} \cdot u$ in $g^{-1}A$,

$$w^{\downarrow} \subseteq (g^{-1})^{\downarrow} \cup g^{-1} \cdot (u^{\downarrow}) = g^{-1} \cdot (g^{\downarrow} \cup u^{\downarrow}) \subseteq g^{-1} \cdot A.$$

Hence $g^{-1} \cdot A$, being saturated, is an element of $\mathcal{Y}$.

We may thus redefine the semigroup $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$, which for convenience we shall call M_X, by

$$M_X = \{(A, g) \in \mathcal{Y} \times G : g \in A\}. \tag{5.10.10}$$

At this point it is useful to consider the structure of a typical element A of $\mathcal{Y}$. Given A in $\mathcal{Y}$ we define w in A to be *maximal* if it is not a proper left factor of any element of A. Since A is finite there is no problem over the existence of maximal elements. Indeed, if A has maximal elements $w_1, w_2, \ldots, w_m$, then

$$A = w_1^{\downarrow} \cup w_2^{\downarrow} \cup \ldots \cup w_m^{\downarrow}.$$

The condition in (5.10.10) that $g \in A$ means that g is a left factor of some maximal w in A.

We now show that M_X is the free inverse semigroup on X. Certainly there is a map $\theta : X \to M_X$ given by

$$x\theta = (x^{\downarrow}, x) \quad (x \in X).$$

(Notice the 'doublethink' involved in the notation here: the first x appearing on the right hand side is to be thought of as a group-reduced word in Y^*, while the second x $(= \overline{x})$ is a member of the free group on X.) Suppose now that S is an inverse semigroup and that there is a map $\alpha : X \to S$. We extend α to Y by defining $(x^{-1})\alpha = (x\alpha)^{-1}$, and then to Y^* by defining

$$(y_1 y_2 \ldots y_n)\alpha = (y_1\alpha)(y_2\alpha)\ldots(y_n\alpha).$$

Let $A = w_1^{\downarrow} \cup w_2^{\downarrow} \cup \ldots \cup w_m^{\downarrow}$ be an element of $\mathcal{Y}$. Define the idempotent element e_A of S by

$$e_A = \big((w_1 w_1^{-1})(w_2 w_2^{-1}) \ldots (w_m w_m^{-1})\big)\alpha. \tag{5.10.11}$$

Notice that e_A is independent of the order in which $w_1, w_2, \ldots, w_m$ appear, and depends only on the set A. We can alternatively write e_A as e_M, where $M = \{w_1, w_2, \ldots, w_m\}$ is the set of maximal elements in A, for e_A depends only on the maximal elements. It is clear that, for all A, B in $\mathcal{Y}$,

$$e_A e_B = e_{A \cup B}. \tag{5.10.12}$$

We can in fact define e_Z for any finite subset Z of Y^*, simply by choosing the maximal elements $w_1, w_2, \ldots, w_m$ in Z and using formula (5.10.11). Notice now that for all A in $\mathcal{Y}$ and g in G,

$$(g\alpha)e_A = e_{gA}(g\alpha). \tag{5.10.13}$$

To see this, consider first a typical product $(g\alpha)(w\alpha)(w\alpha)^{-1}$, where g and w are group-reduced words in Y^*, and suppose that $g = hu$, $w = u^{-1}v$, and $g \cdot w = hv$. Then

$$\begin{aligned}(g\alpha)(w\alpha)(w\alpha)^{-1} &= (h\alpha)(u\alpha)(u\alpha)^{-1}(v\alpha)(v\alpha)^{-1}(u\alpha)\\ &= (h\alpha)(v\alpha)(v\alpha)^{-1}(u\alpha)\\ &= [(h\alpha)(v\alpha)(v\alpha)^{-1}(h\alpha)^{-1}](h\alpha)(u\alpha)\\ &= \big([(g \cdot w)(g \cdot w)^{-1}]\alpha\big)(g\alpha).\end{aligned}$$

Suppose now that A has a set $M = \{w_1, \dots, w_m\}$ of maximal elements. Then within the inverse semigroup S we have

$$\begin{aligned}(g\alpha)e_A &= (g\alpha)\big((w_1\alpha)(w_1^{-1}\alpha)\big)\dots\big((w_m\alpha)(w_m^{-1}\alpha)\big)\\ &= \big([(g\cdot w_1)(g\cdot w_1)^{-1}]\alpha\big)(g\alpha)e_{\{w_2,\dots,w_m\}}\\ &= \big([(g\cdot w_1)(g\cdot w_1)^{-1}]\alpha\big)\big([(g\cdot w_2)(g\cdot w_2)^{-1}]\alpha\big)(g\alpha)e_{\{w_3,\dots,w_m\}}\\ &= \dots = e_{g\cdot A}(g\alpha),\end{aligned}$$

since the maximal elements of $g \cdot A$ are $g \cdot w_1, \dots, g \cdot w_m$.

Now let $(A, g) \in M_X$, and define $\overline{\alpha} : M_X \to S$ by

$$(A,g)\overline{\alpha} = e_A(g\alpha) \quad \big((A,g) \in M_X\big). \tag{5.10.14}$$

To see that this is a morphism, notice first that, by (5.10.12) and (5.10.14),

$$\begin{aligned}[(A,g)\overline{\alpha}][(B,h)\overline{\alpha}] &= e_A[(g\alpha)e_B](h\alpha)\\ &= e_Ae_{gB}(g\alpha)(h\alpha) = e_{A\cup g\cdot B}[(gh)\alpha].\end{aligned}$$

Suppose now that $g = ac$, $h = c^{-1}b$ and $g \cdot h = ab$. Since $g \in A$, there is a factor $[(acd)(acd)^{-1}]\alpha$ in $e_{A\cup g\cdot B}$. Then

$$\begin{aligned}\big([(acd)(acd)^{-1}]\alpha\big)(g\alpha)(h\alpha) &= [(acdd^{-1}c^{-1}a^{-1})(acc^{-1}b)]\alpha\\ &= [(acdd^{-1}c^{-1}a^{-1})(ab)]\alpha\\ &= \big([(acd)(acd)^{-1}]\alpha\big)[(g\cdot h)\alpha].\end{aligned}$$

It now follows easily that

$$\begin{aligned}[(A,g)\overline{\alpha}][(B,h)\overline{\alpha}] &= e_{A\cup g\cdot B}[(g\cdot h)\alpha]\\ &= (A\cup g\cdot B, g\cdot h)\overline{\alpha} = [(A,g)(B,h)]\overline{\alpha}.\end{aligned}$$

Certainly we have

$$x\theta\overline{\alpha} = (x^{\downarrow}, x)\overline{\alpha} = [(xx^{-1})\alpha](x\alpha) = x\alpha \tag{5.10.15}$$

for all x in X. It will follow that M_X is the free inverse monoid on X if we show that $\overline{\alpha}$ is the only morphism from M_X into S satisfying (5.10.15). This in turn will follow if we show that the elements $x\theta = (x^{\downarrow}, x)$ generate M_X, for it will then follow that if some morphism $\beta : M_X \to S$ coincides with $\overline{\alpha}$ on the elements $x\theta$ then it coincides with $\overline{\alpha}$ over the whole of M_X.

So let us denote by T_X the inverse submonoid of M_X generated by the elements $x\theta = (x^{\downarrow}, x)$ $(x \in X)$. Using Proposition 5.9.4(1) and equation (5.10.4), we see that

$$(x^{\downarrow}, x)^{-1} = ((x^{-1})^{\downarrow}, x^{-1}) \in T_X$$

for each x, and so the products

$$(x^{\downarrow}, x)\big((x^{-1})^{\downarrow}, x^{-1}\big) = (x^{\downarrow}, 1), \tag{5.10.16}$$

$$\big((x^{-1})^{\downarrow}, x^{-1}\big)(x^{\downarrow}, x) = \big((x^{-1})^{\downarrow}, 1\big) \tag{5.10.17}$$

both belong to T_X.

We next show that $(w^{\downarrow}, 1) \in T_X$ for every group-reduced word w in Y^*, This we do by induction on the length $|w|$ of w, it being clear from (5.10.16) and (5.10.17) that the result holds if $|w| = 1$. If $w = y_1 y_2 \dots y_n$ is of length n, then $w = y_1 z$, where $z = y_2 \dots y_n$, and the inductive step depends on the observation that

$$(y_1^{\downarrow}, y_1)(z^{\downarrow}, 1)\big((y_1^{-1})^{\downarrow}, y^{-1}\big) = \big((y_1 z)^{\downarrow}, 1\big).$$

Next, we show that $(w^{\downarrow}, u) \in T_X$ for every group-reduced word w in Y^* and every left factor u of w. Let $w = y_1 y_2 \dots y_n$ and let $u = y_1 y_2 \dots y_j$, where $0 \le j \le n$. We prove the result by induction on j, it being clear from the last paragraph that it holds if $j = 0$. The inductive step is clear once we observe that

$$(w^{\downarrow}, y_1 y_2 \dots y_j) = (w^{\downarrow}, y_1 y_2 \dots y_{j-1})(y_j^{\downarrow}, y_j).$$

Finally, if we take an arbitrary $A = w_1^{\downarrow} \cup w_2^{\downarrow} \cup \dots \cup w_m^{\downarrow}$ and (without loss of generality) an arbitrary left factor u of w_m, then

$$(A, u) = (w_1^{\downarrow}, 1)(w_2^{\downarrow}, 1) \dots (w_{m-1}^{\downarrow}, 1)(w_m^{\downarrow}, u),$$

and so $(A, u) \in T_X$.

We have now proved our main result:

Theorem 5.10.2 *Let X be a non-empty set. Let $X' = \{x^{-1} : x \in X\}$ be a set disjoint from X and in one-one correspondence with X. Let G be the free group on X. Let $\mathcal{Y}$ be the set of finite saturated sets of group-reduced words in $(X \cup X')^*$, and let $\mathcal{X} = G\mathcal{Y}$. Then $(G, \mathcal{X}, \mathcal{Y})$ is a McAlister triple, and $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is the free inverse monoid on X.* □

The uniqueness of the free inverse monoid on X implies that there is an isomorphism between M_X and Y^*/τ. If we substitute Y^*/τ for S in (5.10.13) we see that the isomorphism $\phi : \mathcal{M}(G, \mathcal{X}, \mathcal{Y}) \to Y^*/\tau$ is given by

$$(w_1^{\downarrow} \cup \dots \cup w_m^{\downarrow}, u)\phi = [(w_1 w_1^{-1}) \dots (w_m w_m^{-1})u]\tau.$$

In effect this gives us a canonical form for words in Y^* modulo τ. To be more specific, the existence of the isomorphism ϕ shows that every element w of Y^* is equivalent to a word

$$(w_1 w_1^{-1}) \dots (w_m w_m^{-1})u, \tag{5.10.18}$$

where

(1) each w_i is a group-reduced word in Y^+;
(2) no w_i is a proper left factor of any w_j;
(3) u is a left factor of some w_i, and $u = \overline{w}$.

The word (5.10.18) is in effect unique: only the order of the factors $w_i w_i^{-1}$ is open to variation.

The isomorphism ϕ gives us little or no clue as to how to find the canonical word (5.10.18) associated with a given element of Y^*, and our final task in this section will be to repair this defect. For a given word $w = y_1y_2\dots y_n$ in Y^* consider the set $\{1, y_1, y_1y_2, \dots, y_1y_2\dots y_n\}$ of left factors. Now compute the group-reduced word associated with each of these left factors, and obtain the set $A(w)$ of *group-reduced left factors* of w. The set $A(w)$ is saturated. We define $M(w)$ as the set of *maximal* group-reduced left factors of w.

Example 5.10.3 Let $X = \{a, b, c\}$ and let

$$w = ab^{-1}bcaa^{-1}b^{-1}a^{-1}abc^{-1}c.$$

Then

$$A(w) = \{a, ab^{-1}, ac, aca, acb^{-1}, acb^{-1}a^{-1}\},$$

and $M(w) = \{ab^{-1}, aca, acb^{-1}a^{-1}\}$.

It is easy to see that, for all w_1, w_2 in Y^*,

$$A(w_1w_2) = A(w_1) \cup \overline{w_1} \cdot A(w_2). \tag{5.10.19}$$

There is no comparably simple formula for $M(w_1w_2)$.

For any set $D = \{d_1, d_2, \dots, d_k\}$ of group-reduced elements of Y^*, let

$$e_D = (d_1d_1^{-1})(d_2d_2^{-1})\dots(d_kd_k^{-1}).$$

The element e_D is not well-defined in Y^*, since it depends on the order of the elements $d_1, d_2, \dots, d_k$, but it is well-defined modulo τ, and this is what will matter. For an arbitrary word w in Y^*, we have

Lemma 5.10.4 . *With the above definitions,*

$$e_{A(w)} \;\tau\; e_{M(w)}.$$

Proof This is clear once we observe that for two words u and $v = uz$,

$$(uu^{-1})(vv^{-1}) = (uu^{-1})(uzz^{-1}u^{-1}) \;\tau\; uzz^{-1}u^{-1} = vv^{-1}.$$

That is, if u is a left factor of v, then uu^{-1} is superfluous. □

Theorem 5.10.5 *Let $w \in Y^*$ and let $M(w) = \{w_1, w_2, \dots, w_m\}$. Let $\overline{w}$ be the group-reduced word associated with w. Then*

$$w \;\tau\; e_{M(w)}\overline{w}.$$

Proof We shall show that $w \;\tau\; e_{A(w)}\overline{w}$. By Lemma 5.10.4 this is equivalent. If $w = 1$, the empty word, then $A(w) = \{1\}$, $\overline{w} = 1$, and the result is clear. If $w = x \in X$, then $A(w) = \{1, x\}$, $\overline{x} = x$, and it is clear that $x \;\tau\; (11^{-1})(xx^{-1})x$. The case where $w = x^{-1}$ is equally clear. So suppose now that $|w| > 1$,and let u be the longest group-reduced left factor of w. If $v = w$ then the process is complete: w is already a canonical word. Otherwise $w = uv$, with $|u|, |v| \geq 1$. From the choice of u it now follows that we have factorizations $u = u_1 z$, $v = z^{-1}v_1$, where $|z| \geq 1$, and we may suppose that we have chosen z to be as long as possible. Then (modulo τ)

$$\begin{aligned} w &\equiv (u_1 z z^{-1} u_1^{-1}) u_1 z z^{-1} v_1 \\ &\equiv (u_1 z z^{-1} u_1^{-1}) u_1 v_1 \text{ by commuting idempotents} \\ &\equiv (uu^{-1}) u_1 v_1 = (uu^{-1}) w_1 \text{ (say).} \end{aligned}$$

Now from $w = u_1 z z^{-1} v_1$ it is clear that $\overline{w_1} = \overline{w}$. By Lemma 5.10.4, $uu^{-1} \;\tau\; e_{A(u)}$. Now, by (5.10.19),

$$A(w_1) = A(u_1) \cup \overline{u_1} \cdot A(v_1),$$

while

$$A(w) = A(u_1 z z^{-1}) \cup \overline{u_1} \cdot A(v_1) = A(u) \cup \overline{u_1} \cdot A(v_1).$$

Since $A(u_1) \subseteq A(u)$ we deduce that

$$A(w) = A(u) \cup A(w_1).$$

Now $|w_1| = |w| - 2|z| < |w|$, and so we may suppose inductively that $w_1 \;\tau\; e_{A(w_1)}\overline{w_1}$. Hence (modulo τ)

$$w \equiv (uu^{-1}) w_1 \equiv e_{A(u)} e_{A(w_1)} \overline{w_1} \equiv e_{A(u) \cup A(w_1)} \overline{w} = e_{A(w)} \overline{w}. \qquad \square$$

Example 5.10.6 Look again at the word

$$w = ab^{-1}bcaa^{-1}b^{-1}a^{-1}abc^{-1}c$$

considered in Example 5.10.3. Here $u = ab^{-1}$, $z = b^{-1}$, $u_1 = a$, and so the first stage in the reduction is

$$w \to (ab^{-1}ba^{-1}) w_1,$$

where $w_1 = acaa^{-1}b^{-1}a^{-1}abc^{-1}c$. Repeating this process on w_1 we obtain

$$w_1 \to (acaa^{-1}c^{-1}a^{-1}) w_2,$$

where $w_2 = acb^{-1}a^{-1}abc^{-1}c$. Repeating the process on w_2, we obtain

$$w_2 \to (acb^{-1}a^{-1}abc^{-1}a^{-1})ac,$$

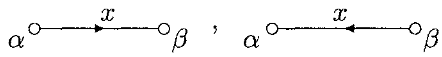
and the conclusion is that

$$w \;\tau\; (ab^{-1}ba^{-1})(acaa^{-1}c^{-1}a^{-1})(acb^{-1}a^{-1}abc^{-1}a^{-1})ac.$$

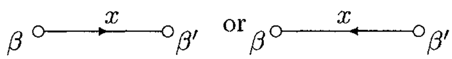
Of course it is not necessary to go through this procedure: by virtue of Theorem 5.10.5 we can simply observe that the three maximal reduced left factors of w are ab^{-1}, aca and $acb^{-1}a^{-1}$, and that $\overline{w} = ac$.

There is a convenient graphical way of identifying the sets $A(w)$ and $M(w)$ associated with a word w in Y^*. We draw a labelled 'word tree' with a labelled *initial* vertex α and a labelled *final* vertex β (possibly the same as α). We shall do this inductively, first associating with the empty word

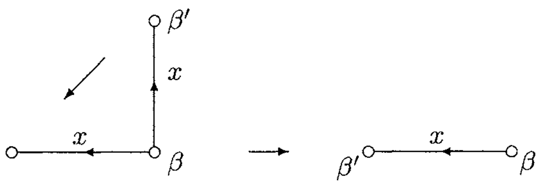
1 the trivial word tree with one vertex and no edges. (Here $\beta = \alpha$.) Less trivially, the words x and x^{-1} are associated with the word trees



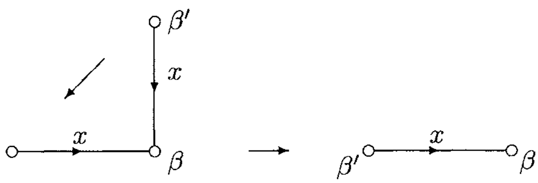
Suppose now that $w = y_1 y_2 \dots y_m$, and that we have constructed a word tree for $y_1 y_2 \dots y_{m-1}$ with vertices α and β, respectively. To construct the word tree for w we begin by 'tacking on' a new edge



according as $y_m = x \in X$ or $y_m = x^{-1} \in X'$. If $y_{m-1} \neq y_m^{-1}$ then our task is complete. If $y_{m-1} = y_m^{-1}$ then we 'fold over' the new edge so as to coincide with the edge corresponding to y_{m-1}:



or

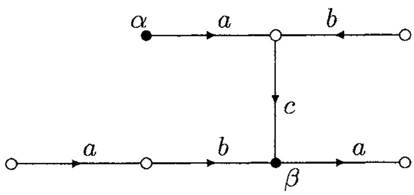


In all cases the new final vertex is β'.

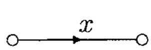
Example 5.10.7 Let us look yet again at the word

$$w = ab^{-1}bcaa^{-1}b^{-1}a^{-1}abc^{-1}c$$

we studied in Examples 5.10.3 and 5.10.6. The associated word tree is



By an *α-walk* in a word tree we mean a path (which need not respect the arrows) beginning at α and visiting no vertex more than once. Each step



in an α-walk is labelled with x if it goes with the arrow and with x^{-1} if it goes against the arrow. The *label* of an α-walk is the product of the labels of the edges traversed. It is not hard to see that the reduced left factors of w are the labels of the α-walks in the word tree, that the maximal reduced left factors of w are the labels of those α-walks that end in an extremal vertex, and that the free group element associated with w is the label of the α-walk terminating in β.

5.11 EXERCISES

1. Show that the $\mathcal{D}$-classes in an inverse semigroup are 'square'. More precisely, show that there is a bijection from the set of $\mathcal{L}$-classes in a $\mathcal{D}$-class D onto the set of $\mathcal{R}$-classes in D, defined by the rule that L_a maps to $R_{a^{-1}}$.

2. In the symmetric inverse semigroup, show that
   (a) $\alpha \mathcal{L} \beta$ if and only if $\operatorname{im}\alpha = \operatorname{im}\beta$;
   (b) $\alpha \mathcal{R} \beta$ if and only if $\operatorname{dom}\alpha = \operatorname{dom}\beta$;
   (c) $\alpha \mathcal{D} \beta$ if and only if $|\operatorname{dom}\alpha| = |\operatorname{dom}\beta|$;
   (d) $\mathcal{D} = \mathcal{J}$.

3. If $|X| = n$, show that
$$|\mathcal{I}_X| = \sum_{r=0}^{n} \binom{n}{r}^2 r!.$$

4. Show that every ideal of an inverse semigroup S is an inverse subsemigroup of S, and deduce that the principal factors of an inverse semigroup are all Brandt semigroups, except for the kernel, which is a group.

   Let I_n be the symmetric inverse semigroup $\mathcal{I}_{\{1,2,\dots,n\}}$, and, for $1 \le r \le n$ let
$$K_r = \{\alpha \in I_n : |\operatorname{dom}\alpha| \le r\}, \quad L_r = \{\alpha \in I_n : |\operatorname{dom}\alpha| < r\}.$$
   Show that the principal factor K_r/L_r is isomorphic to the Brandt semigroup $B\big(S_r, \binom{n}{r}\big)$.

5. If A is a subset of an inverse semigroup, the notation $\langle A\rangle$ is potentially ambiguous, since it might mean the subsemigroup generated by A or the *inverse* subsemigroup generated by A. Show, by considering $A = \{(1, e, 2)\}$ in the Brandt semigroup B_2, that this is a real ambiguity.

6. It is known (Exercise 1(6)) that (for all $n \ge 3$) the symmetric group S_n, consisting of all permutations of the set $X = \{1, 2, \dots, n\}$, is generated by the cycles
$$\tau = (1\,2), \quad \zeta = (1\,2 \dots n).$$

For $r = 0, 1, \ldots, n$, denote by J_r the set

$$\{\alpha \in I_n : |\operatorname{dom}\alpha| = r\},$$

and let β be an arbitrarily chosen, but fixed member of J_{n-1}.

(a) Let $\xi \in J_{n-1}$ and let η be a permutation mapping $\operatorname{dom}\xi$ onto $\operatorname{dom}\beta$. Show that there exists a permutation ϕ such that $\xi = \eta\beta\phi$, and deduce that $J_{n-1} \subseteq \langle \tau, \zeta, \beta \rangle$.

(b) Show that, for $r = 0, 1, \ldots, n-2$,

$$J_r \subseteq J_{r+1}J_{n-1}.$$

(c) Deduce that $I_n = \langle \tau, \zeta, \beta \rangle$.

7. In the finite symmetric group S_n the decomposition of permutations into products (compositions) of disjoint cycles is well known. Let α be a member of the finite symmetric inverse semigroup I_n, and define a relation $\mathbf{E}$ on $X = \{1, 2, \ldots, n\}$ by the rule that

$$x \mathbf{E} y \text{ if and only if } (\exists m, n \in \mathbf{Z})\ x\alpha^m = y\alpha^n.$$

(Interpret $x\alpha^0$ as x, and notice that $x\alpha^0$ is defined even if $x \notin \operatorname{dom}\alpha \cup \operatorname{im}\alpha$.)

(a) Show that $\mathbf{E}$ is an equivalence on X, and that the equivalence class $x\mathbf{E}$ is the set $x\alpha^{\mathbf{Z}}$, defined by

$$x\alpha^{\mathbf{Z}} = \{x\alpha^n : n \in \mathbf{Z}\}.$$

(b) Show that either $x\alpha^{\mathbf{Z}} \subseteq \operatorname{dom}\alpha$, or there exists a unique x_1 in $x\alpha^{\mathbf{Z}} \setminus \operatorname{im}\alpha$ and a unique x_2 in $x\alpha^{\mathbf{Z}} \setminus \operatorname{dom}\alpha$. Show that $x_1 = x_2$ if and only if $x \notin \operatorname{dom}\alpha \cup \operatorname{im}\alpha$.

(c) Deduce that every α in I_n can be expressed as a product (composition) of disjoint partial one-one maps of two types:

i. *cycles* $\zeta = (x_1\, x_2 \ldots x_m)$ $(m \geq 1)$, where $x_i\zeta = x_{i+1}$ $(i = 1, 2, \ldots, m-1)$ and $x_m\zeta = x_1$;

ii. *chains* $\chi = (y_1\, y_2 \ldots y_p]$ $(p \geq 2)$, where $y_1 \notin \operatorname{im}\chi$, $y_p \notin \operatorname{dom}\chi$, and $y_i\chi = y_{i+1}$ $(i = 1, 2, \ldots, p-1)$.

8. Let G be a group, and let S be the set of all right cosets Ha in G. Here H runs over all subgroups of G (including G itself and $\{e\}$) and a over all elements of G. Define a binary operation $*$ on S by the rule that

$$Ha * Kb = (H \vee aKa^{-1})ab.$$

Here $H \vee aKa^{-1}$ denotes the smallest subgroup of G containing both H and aKa^{-1}.

(a) Show that $Ha * Kb$ is the smallest coset containing the product $HaKb$.

(b) Show that the operation $*$ is associative.
(c) Show that Ha has an inverse $(a^{-1}Ha)a^{-1}$ in the semigroup $\mathcal{S}$.
(d) Show that the idempotents of $\mathcal{S}$ are the subgroups of G, and deduce that $\mathcal{S}$ is an inverse semigroup.
(e) Show that the subgroup H is a central idempotent in $\mathcal{S}$ if and only if it is a normal subgroup of G.

9. For every subset K of an inverse semigroup S, and for every s in S, show that $(Ks)\omega = \big((K\omega)s\big)\omega$.

10. Let S be an inverse semigroup with semilattice of idempotents E, and let σ be the minimum group congruence on S. Show that the following statements are equivalent:

(a) $x \,\sigma\, y$;	(b) $(\exists e \in E)\ xe = ye$;
(c) $(\exists a \in S)\ ax = ay$;	(d) $(\exists a \in S)\ xa = ya$;
(e) $(\exists e, f \in E)\ ex = fy$;	(f) $(\exists e, f \in E)\ xe = yf$;
(g) $x^{-1}y \in E\omega$;	(h) $(Ex)\omega = (Ey)\omega$;
(i) $y \in (Ex)\omega$;	(j) $x \in (Ey)\omega$.

11. Let U be a subsemigroup of an inverse semigroup S.
(a) Show that if U is full and unitary then U is an inverse subsemigroup of S.
(b) Show that if U is a left unitary inverse subsemigroup of S then U is unitary.
(c) Show that an inverse subsemigroup of S is unitary if and only if it is closed.

12. Let U be a closed inverse subsemigroup of T, where T is a closed inverse subsemigroup of the inverse semigroup S. Show that U is a closed inverse subsemigroup of S.

Show that the assumption that T is closed is necessary. Specifically, let S be the Clifford semigroup which is the disjoint union of two groups U and V with identities e, f, respectively, where $e < f$, and with connecting morphism $\phi : V \to U$. Show that, if H is a subgroup of U, then H is a closed inverse subsemigroup of U, but that the closure of H in S is $H \cup H\phi^{-1}$.

13. Let S be an inverse semigroup and let σ be the minimum group congruence on S. Show that, for every congruence ξ on S,

$$\xi \vee \sigma = \sigma \circ \xi \circ \sigma.$$

[Hint. The essential point is to show that the relation $\sigma \circ \xi \circ \sigma$ is transitive. If

$$ea = ep, \quad p \,\xi\, q, \quad fq = fb,$$

and

$$gb = gr, \quad r \ \xi \ s, \quad hs = hc,$$

where e, f, g, h are idempotents, show that

$$ia = it, \quad t \ \xi \ u, \quad ju = jc,$$

where $i = efg$, $t = fgp$ and $j = hfg$.]

14. Let S be an inverse semigroup with semilattice of idempotents E, and let μ be the maximum idempotent-separating congruence on S.
    (a) Show that

    $$\mu = \min\{\rho \in \mathcal{C}(S) : S/\rho \text{ is fundamental}\}.$$

    (b) Show that $S/\mu \simeq E$ if and only if S is a Clifford semigroup.

15. Let S be an inverse semigroup with semilattice of idempotents E. A congruence ρ on S is called *idempotent-pure* if $\operatorname{Ker}\rho = E$. Show that ρ is idempotent-pure if and only if $\rho \cap \mathcal{L} = 1_S$.

In Proposition 5.3.4 we determined the smallest and the largest congruences on an inverse semigroup S having a given trace τ. We can also describe the smallest and the largest congruences on S having a given kernel K. The next three exercises explore this idea.

16. A full inverse subsemigroup K of an inverse semigroup S is said to have the *kernel property* if

    $$(\forall a, b \in K) \ ab \in K \ \Rightarrow \ aKb \subseteq K.$$

    Show that $\operatorname{Ker}\rho$ has the kernel property for every congruence ρ on S.

17. Let K be a full inverse subsemigroup of S having the kernel property, and let σ_K be the syntactic congruence of K.
    (a) Show that $\operatorname{Ker}\sigma_K \subseteq K$.
    (b) Show that if $a \in K$ then, for all x, y in S,

    $$xay \in K \text{ if and only if } xaa^{-1}y \in K,$$

    and deduce that $a \in \operatorname{Ker}\sigma_K$.
    (c) Let γ be a congruence on K such that $\operatorname{Ker}\gamma = K$. Show that $\gamma \subseteq \sigma_K$.
    (d) Deduce that

    $$\sigma_K = \max\{\rho \in \mathcal{C}(S) : \operatorname{Ker}\rho = K\}.$$

18. Let K be a full inverse subsemigroup of S having the kernel property, and let $\tau_K = (\sigma_K \cap \mathcal{L})^\#$.

(a) Let γ be a congruence on S such that $\operatorname{Ker}\sigma = K$. Show that $\sigma \cap \mathcal{L} \subseteq \gamma$, and deduce that $\tau_K \subseteq \gamma$.
(b) Show that $K \subseteq \operatorname{Ker}\tau_K$.
(c) Deduce that

$$\tau_K = \min\{\rho \in \mathcal{C}(S) : \operatorname{Ker}\rho = K\}.$$

19. Let S be an inverse semigroup.
   (a) Show that, if ρ is a congruence on S, then ρ is a *semilattice congruence* (that is to say, S/ρ is a semilattice) if and only if $(a^2, a) \in \rho$ for every a in S. Deduce that the intersection of a non-empty family $\{\rho_i : i \in I\}$ of semilattice congruences is a semilattice congruence, and hence that there exists a minimum semilattice congruence η on S.
   (b) Show that $\mathcal{J} \subseteq \eta$ and that $\eta \subseteq \mathcal{H}^{\#}$. Deduce that

$$\mathcal{H}^{\#} = \mathcal{L}^{\#} = \mathcal{R}^{\#} = \mathcal{D}^{\#} = \mathcal{J}^{\#} = \eta.$$

20. Show that the Clifford semigroup $\mathcal{SL}[Y; G_\alpha; \phi_{\alpha,\beta}]$ is E-unitary if and only if the morphisms $\phi_{\alpha,\beta}$ are all one-one.
    Let T be an inverse semigroup with identity, and let θ be a morphism from T into the group of units H_1 of T. Show that the Bruck–Reilly extension $BR(T, \theta)$ is E-unitary if and only if T is E-unitary and $\ker\theta = \sigma$ (the minimum group congruence on T). Deduce that the bisimple inverse ω-semigroup is E-unitary if and only if θ is one-one.

21. Let S be an inverse semigroup, and let σ, η be (respectively) the minimum group congruence and the minimum semilattice congruence on S. Show that $\sigma \cap \eta = 1_S$ if and only if S is an E-unitary Clifford semigroup. Deduce that, on an arbitrary inverse semigroup S, $\sigma \cap \eta$ is the minimum congruence ρ such that S/ρ is an E-unitary Clifford semigroup.

22. Let α and β be $\mathcal{H}$-equivalent elements of the symmetric inverse semigroup $\mathcal{I}_X$. Suppose, in other words, that

$$\operatorname{dom}\alpha = \operatorname{dom}\beta = A, \quad \operatorname{im}\alpha = \operatorname{im}\beta = B.$$

Show that, for every idempotent 1_C in $\mathcal{I}_X$,

$$\alpha^{-1}1_C\alpha = 1_{(A\cap C)\alpha}, \quad \beta^{-1}1_C\beta = 1_{(A\cap C)\beta}.$$

Deduce that $\mathcal{I}_X$ is fundamental. Hence deduce that every inverse semigroup is embeddable in a fundamental inverse semigroup.

23. Let X be a set and let $\mathcal{E} = \mathcal{P}(X)$ be the semilattice (under intersection) of all subsets of X. Show that the Munn semigroup $T_{\mathcal{E}}$ is isomorphic to $\mathcal{I}_X$. (Notice that from this we can deduce the result of the previous exercise, that $\mathcal{I}_X$ is fundamental.)

24. Let E be the semilattice $C_\omega \times C_\omega$. Then E may be identified with $\{e_{m,n} : m, n \in \mathbf{N}^0\}$, where

$$e_{m,n}e_{p,q} = e_{\max(m,p),\max(n,q)}.$$

(a) Show that the group of automorphisms of E is $\{1_E, \gamma\}$, where

$$e_{m,n}\gamma = e_{n,m} \quad (m, n \in \mathbf{N}^0).$$

(b) Show that E is uniform and that, if $e = e_{m,n}$ and $f = e_{p,q}$, then $T_{e,f} = \{\alpha, \beta\}$, where

$$e_{r,s}\alpha = e_{r-m+p,s-n+q} \quad (r \geq m,\ s \geq n),$$
$$e_{r,s}\beta = e_{s-n+p,r-m+q} \quad (r \geq m,\ s \geq n).$$

(c) Deduce that $\mu \neq \mathcal{H}$ in T_E.

25. By considering $BR(S^1, \theta)$, where θ maps every element of S^1 to 1, show that every semigroup S can be embedded up to isomorphism in a simple monoid.

26. From Section 2.3 we know that a $\mathcal{D}$-class in a semigroup cannot contain both regular and irregular elements. Show that this does not apply to $\mathcal{J}$-classes. Specifically, consider the monoid $T = \{1, x, 0\}$, where $x^2 = 0$, let $\theta : T \to H_1$ be given by $1\theta = x\theta = 0\theta = 1$, and let $S = BR(T, \theta)$.

(a) Show that the $\mathcal{D}$-classes of S are

$$D^1 = \mathbf{N}^0 \times \{1\} \times \mathbf{N}^0,\ D^x = \mathbf{N}^0 \times \{x\} \times \mathbf{N}^0,\ D^0 = \mathbf{N}^0 \times \{0\} \times \mathbf{N}^0.$$

(b) Show that D^1 and D^0 are regular, but that D^x is irregular.

(c) Deduce that the single $\mathcal{J}$-class of the semigroup S contains both regular and irregular elements.

27. Let $S_1 = BR(G_1, \theta_1)$ and $S_2 = BR(G_2, \theta_2)$ be bisimple ω-semigroups, and let $\alpha : G_1 \to G_2$ be an isomorphism such that $\theta_1\alpha = \alpha\theta_2\lambda_z$, where, for some z in G_2, λ_z is the inner automorphism $g_2 \mapsto zg_2z^{-1}$. Show that the map $\phi : S_1 \to S_2$, defined by

$$(m, g_1, n)\phi = (1, z^{-1}, 0)^m(0, g\alpha, 0)(0, z, 1)^n \quad \big((m, g_1, n) \in S_1\big),$$

is an isomorphism. Show conversely that every isomorphism from S_1 onto S_2 is of this type.

28. Let A be the semilattice $\{a, b, z\}$, in which z is the minimum element and $ab = z$. Let $E = C_\omega \times A$, with the lexicographic ordering

$$(e_m, x) \leq (e_n, y) \text{ if and only if either } m > n, \text{ or } m = n \text{ and } x \leq y.$$

Show that E (known as the 'lazy tongs' semilattice) is subuniform but not uniform.

29. More generally, if A is an arbitrary semilattice and U is a uniform semilattice, let E be the *ordinal product* $U \circ A$ (in the sense of Birkhoff (1948)). That is, let $E = U \times A$, with the lexicographic ordering

$$(u, a) \leq (v, b) \text{ if and only if either } u < v, \text{ or } u = v \text{ and } a \leq b.$$

Show that E is subuniform. (As the previous exercise shows, it need not be uniform.)

30. Let T be the Clifford semigroup consisting of the chain of groups G_0, $G_1, \ldots, G_{d-1}$, with identity elements $e_0 > e_1 > \cdots > e_{d-1}$ and with structure morphisms $\gamma_i : G_i \to G_{i+1}$ $(i = 0, 1, \ldots, d-2)$. Let θ be an endomorphism of T with $S\theta \subseteq G_0$ (as in a Bruck–Reilly extension), and let

$$\theta|_{G_{d-1}} = \delta : G_{d-1} \to G_0.$$

Show that

$$\theta|_{G_i} = \gamma_i \ldots \gamma_{d-2}\delta \quad (i = 0, 1, \ldots, d-2).$$

Conversely, let $\delta : G_{d-1} \to G_0$ be a morphism. Show that $\theta : S \to G_0$, defined by

$$\theta|_{G_i} = \gamma_i \cdots y_{d-1}\delta \quad (i = 0, 1, \ldots, d-2), \quad \theta_{G_{d-1}} = \delta,$$

is an endomorphism of S such that $S\theta \subseteq G_0$. (This enables one to recover Munn's original (1968) version of Theorem 5.7.6.)

31. Show that:
    (a) $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is simple if and only if for each A, B in $\mathcal{Y}$ there exists g in G such that $gA \geq B$;
    (b) $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is bisimple if and only if for each A, B in $\mathcal{Y}$ there exists g in G such that $gA = B$. (This property is usually described by saying that G acts *transitively* on $\mathcal{Y}$.)

32. Let μ be the maximum idempotent-separating congruence on $\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$, and for each A in $\mathcal{Y}$ let

$$C_A = \{g \in G : gB = B \text{ for all } B \leq A\}.$$

Show that $(A, g)\ \mu\ (B, h)$ if and only if $A = B$ and $gh^{-1} \in C_A$.

33. Let $S = \mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ be an E-unitary inverse semigroup. Since such semigroups have been called 'proper', let us say that S is *prim* if it has the additional property that $\mathcal{L} \circ \sigma = S \times S$. Show that S is prim if and only if $\mathcal{Y} = \mathcal{X}$.

34. Let S be the inverse subsemigroup of $\mathcal{I}_{\{1,2,3,4,5\}}$ generated by

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 5 \end{pmatrix}.$$

(a) Show that $|S| = 7$, $|E| = 3$.
(b) Show that S is E-unitary.
(c) Show that the σ-classes of S are E, $\{\alpha, \alpha^{-2}\}$, $\{\alpha^{-1}, \alpha^2\}$, and that the $\mathcal{L}$-classes are $\{\alpha, \alpha^{-1}\alpha\}$, $\{\alpha^{-1}, \alpha\alpha^{-1}\}$, $\{\alpha^2, \alpha^{-2}, \alpha^2\alpha^{-2}\}$. Deduce that S is not prim.
(d) Show that, in S, $\mathcal{L} \circ \sigma \circ \mathcal{L} = \sigma \circ \mathcal{L} \circ \sigma = S \times S$.

35. Show that $S = \mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is a Clifford semigroup if and only if the action of G on $\mathcal{Y}$ is trivial, in the sense that $gA = A$ for every g in G and A in $\mathcal{Y}$. (Note that $\mathcal{X} = \mathcal{Y}$ in this case.)

36. An inverse semigroup S is said to be an *F-inverse semigroup* if and only if every σ-class contains a greatest element. Let S be an F-inverse semigroup. Show that S contains an identity element and that it is E-unitary.

37. Show that $S = \mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is an F-inverse semigroup if and only if $\mathcal{Y}$ is a principal ideal of $\mathcal{X}$ with greatest element E, and $gE \wedge E$ exists for every g in G.

38. Show that $S = \mathcal{M}(G, \mathcal{X}, \mathcal{Y})$ is an F-inverse semigroup if and only if $\mathcal{Y}$ is a principal ideal of $\mathcal{X}$ and $\mathcal{X}$ is a semilattice.

39. Let $\mathcal{S}$ be the set of all (finite) saturated sets of group-reduced words in $(X \cup X')^*$. Show that the semilattice E of idempotents of FI_X is isomorphic to $(\mathcal{S}, \cup)$.

40. Let $\mathcal{P}(X)$ be the set of all subsets of X, and let FG_X be the free group on X, consisting of all group-reduced words in the alphabet $Y = X \cup X'$. For each w in Y^* and x in X, say that $x \in C(w)$ (the *content* of w) if $w \in Y^*xY^* \cup Y^*x^{-1}Y^*$, that is, if the word w contains x or x^{-1}. Then $C(w) \subseteq X$. For a subset A of Y^*, define

$$C(A) = \bigcup_{w \in A} C(w).$$

Show that, for all w, w' in Y^*,

$$C(ww') = C(w) \cup C(w').$$

(a) Let

$$C_X = \{(Z, g) \in \mathcal{P}(X) \times FG_X : c(g) \subseteq Z\},$$

and define a multiplication on C_X by

$$(Z, g)(T, h) = (Z \cup T, gh).$$

Show that C_X is a Clifford semigroup.

(b) Show that C_X is the free Clifford semigroup on the set X. Show in fact that if $\theta : X \to C_X$ is given by

$$x\theta = \big(\{x\}, x\big) \quad (x \in X),$$

then for every Clifford semigroup S and every map $\alpha : X \to S$ there is a unique morphism $\bar{\alpha} : C_X \to S$ such that $\theta\bar{\alpha} = \alpha$. [Hint. In a Clifford semigroup, two products of elements and their inverses are $\mathcal{H}$-equivalent if they have the same content.]

(c) Show that $C_X = M_X/\kappa$, where, for (A, g), (B, h) in M_X,

$$(A, g) \ \kappa \ (B, h) \text{ if and only if } g = h \text{ and } c(A) = c(B).$$

41. Let A be a saturated set of group-reduced words in $(X \cup X')^*$ and let $g \in A$. Show that $g^{-1}A = A$ if and only if $g = 1$. Show also that:
 (a) $\mathcal{H} = 1$;
 (b) $\mathcal{D} = \mathcal{J}$.

42. In the case where $X = \{x\}$ has cardinality 1, every saturated set has the form $(x^r)^{\downarrow} \cup (x^{-s})^{\downarrow}$. Show that $FI_{\{x\}}$ is isomorphic to

$$T = \{(r, s, k) \in \mathbf{Z}^3 : r \geq 0, \ s \geq 0, \ -s \leq k \leq r\},$$

with multiplication given by

$$(r, s, k)(r', s', k') = \big(\max(r, r' + k), \max(s, s' - k), k + k'\big).$$

Determine the inverse of (r, s, k) in T. Show that:
 (a) $(r, s, k) \ \mathcal{R} \ (r', s', k')$ if and only if $r = r'$ and $s = s'$;
 (b) $(r, s, k) \ \mathcal{L} \ (r', s', k')$ if and only if $r - k = r' - k'$ and $s + k = s' + k'$;
 (c) $(r, s, k) \ \mathcal{D} \ (r', s', k')$ if and only if $r + s = r' + s'$;
 (d) $(r, s, k) \ \sigma \ (r', s', k')$ if and only if $k = k'$.

5.12 NOTES

The origin of inverse semigroup theory is in papers by Vagner (1952) and Preston (1954a,b,c), and Sections 5.1 and 5.2 are drawn from their fundamental work. Both Vagner and Preston considered the natural order relation $\leq$, but the first substantial use of the closure operation ω is in Schein (1962).

The notion of what is now called a *Brandt groupoid* goes back to Brandt (1927); a Brandt semigroup arises by adjoining a zero. Clifford and Preston (1961) give a full account of the relationship between the two ideas.

The idea of the kernel and the trace goes back to Scheiblich (1974) and was developed by D. G. Green (1975) and Petrich (1978). For an application of the same idea to completely simple semigroups, see Petrich and Reilly (1981b).

The congruence σ was studied first by Munn (1961), who gave the characterization (5.3.3). The alternative description of σ in Theorem 5.3.5

is from Howie (1964c), a paper that also saw the first appearance of the congruence μ and of its characterizations in (5.3.4) and Theorem 5.3.6. The observation that $\mu = \mathcal{H}^\flat$ is due to Lallement (1966).

The results of Section 5.4 are from (Munn 1970a). Section 5.5 is due to Howie and Schein (1969), except for the final example, which was contributed by Munn (private communication).

The Bruck–Reilly extension (Section 5.6) was devised in its general form by Munn (1970b), having been invented in special cases by Bruck (1958) and Reilly (1966). The Bruck case is described in Exercise 25; the Reilly case is given in Theorem 5.6.7. See also Warne (1966). The results of Section 5.7 are due to Munn (1968, 1970b) and Kochin (1968).

Reilly's (1966) structure theorem for bisimple ω-semigroups has been the model for many subsequent generalizations. See, for example, Warne (1968) and the PhD theses of Hickey (1970) and P. McLean (1973). As long as the uniform semilattice E has the crucial property that $|T_{e,f}| = 1$ for all e, f in E then the difficulties of classifying the bisimple semigroups appear to be manageable.

The isomorphism theorem associated with Reilly's result is given as Exercise 27. Munn and Reilly (1966) and Munn (1966a) studied congruences on a bisimple ω-semigroup; and Baird (1972) extended this study to simple ω-semigroups.

The results of Section 5.8 are from Schein (1962).

The structure theorem for E-unitary inverse semigroups is due to McAlister (1974a,b), but the proof we have given is essentially that of Schein (1975b). Alternative proofs have been given by Munn (1976) and Wilkinson (1983). Another approach to the description of E-unitary semigroups can be found in Petrich and Reilly (1979). The notion of a factorizable inverse semigroup comes from Chen and Hsieh (1974), and the proof of Theorem 5.9.7 comes from McAlister and Reilly (1977). A more thorough study of E-unitary covers has been made by Petrich and Reilly (1983a).

The McAlister theorem (Theorem 5.9.2) has been generalized to $\mathcal{R}$-unipotent semigroups (regular semigroups in which $efe = ef$ for all idempotents e and f) by Takizawa (1978).

The serious study of free inverse semigroups seems to have begun in the early seventies, with papers by Reilly (1972, 1973) and Scheiblich (1973a,b). We have presented Scheiblich's approach as an application of McAlister's theorem. Historically, however, it happened the other way round: Scheiblich's treatment of free inverse semigroups provided important motivation for McAlister's more general work. The canonical form (5.10.18) was derived by Preston (1973) and Schein (1975a). For a survey of free inverse semigroups, see Reilly (1979). The graphical approach briefly alluded to at the end of Section 5.10 was pioneered by Munn (1974b), and put to use by O'Carroll (1974). It has been greatly developed by later authors, such as Margolis *et al.* (1990) and Stephen (1990). Jones (1981, 1982, 1984) and

Jones *et al.* (1991) have used graphical techniques to study free products within the category of inverse semigroups. Further information, and many extra references, can be found in an excellent survey article by Meakin (1993).

F-inverse semigroups were studied by McFadden and O'Carroll (1971) and McAlister and McFadden (1974).

Free monogenic inverse semigroups (see Exercise 42) were studied first by Gluskin (1957). Later authors gave alternative versions. For an exhaustive study, see Petrich (1984).

Free objects for other varieties of U-semigroups have also been studied. Free objects in certain varieties of inverse semigroups were studied by Trotter (1986a) and Reilly and Trotter (1986). Free completely regular semigroups on two generators were described by Clifford (1979), and the general case was treated by Gerhard (1983a,b) and Trotter (1988). Free completely simple semigroups have been studied by Rasin (1979). For a description of free bands, free semilattices and free normal bands, see Chapter 4 (Section 4.5, Exercises 4(20) and 4(21)). For a survey of free objects in varieties of inverse semigroups, see Reilly (1987).

The idea of chains and cycles in Exercise 7 was used by Gomes and Howie (1987a,b). See also Lipscomb (1986, 1992a,b). Of the other exercises not already mentioned, Exercise 13 is from Howie (1964c). Exercise 15 is from D. G. Green (1973), and references for Exercises 16, 17 and 18 are D. G. Green (1975) and Petrich and Reilly (1982). Exercise 19 is in Howie and Lallement (1966), and Exercises 28 and 29 are from (Munn 1970b). A description of free Clifford semigroups (see Exercise 40) was given by Liber (1954). Exercise 41, concerning the Green relations in FI_X, comes from Reilly (1972).

6

Other classes of regular semigroups

The success of inverse semigroup theory has naturally prompted the study of more general regular semigroups, very often quite explicitly defined as generalizations. In the Exercises and in the Notes at the end of the chapter we make brief reference to one or two pieces of work of this kind, but in the text we shall confine ourselves to three fairly natural types of regular semigroup.

We may regard an inverse semigroup as a regular semigroup in which the idempotents form a semilattice. What have been called *R-unipotent* semigroups—see Exercise 6 for an explanation of the terminology—may be defined as regular semigroups in which the idempotents form a right normal band. More generally, several authors have studied regular semigroups in which the idempotents form a normal band; these are usually, though in the event somewhat confusingly, called *generalized inverse semigroups.*

A generalized inverse semigroup S has the property of being *locally inverse*, which is to say that eSe is an inverse semigroup for every idempotent e, for if $f = efe$ and $g = ege$ are idempotents in eSe, then

$$fg = efege = egefe = gf.$$

However, not every locally inverse semigroup is a generalized inverse semigroup: every completely simple semigroup is certainly locally inverse, but its idempotents do not necessarily form a subsemigroup. The class of generalized inverse semigroups is in fact the intersection of the class of locally inverse semigroups and the class of *orthodox* semigroups, where a semigroup is called *orthodox* if it is regular and its idempotents form a subsemigroup.

In Sections 6.1 and 6.2 we study locally inverse semigroups and orthodox semigroups, respectively. Both these types of semigroup do appear quite naturally in the development of the theory. Orthodox semigroups make an appearance in the study of congruences by Howie and Lallement (1966), and have featured strongly in subsequent work on varieties of completely regular semigroups. Locally inverse semigroups play a part in the work of Nambooripad (1980) and McAlister (1983, 1985).

The final section is devoted to a class of semigroups that are not in any way a generalization of inverse semigroups. If S is a regular semigroup with set E of idempotents, then S is orthodox if and only if $\langle E \rangle = E$. The opposite case, when S is regular and when $\langle E \rangle = S$, is in some ways even more natural; we call S a *semiband* in this case.

6.1 LOCALLY INVERSE SEMIGROUPS

In every regular semigroup S the subset eSe is clearly a subsemigroup for every idempotent e. It is even a regular subsemigroup, since for every $x = ese$ in eSe and every inverse x' of x,

$$x = xx'x = (xe)x'(ex) = x(ex'e)x.$$

A regular semigroup S with set E of idempotents will be called *locally inverse* if eSe is an inverse semigroup for every e in E.

Before discussing these semigroups, we pause to examine an idea that does not at first seem to be related. The order relation $\leq$ that is such an important tool in inverse semigroups can in fact be generalized to arbitrary regular semigroups. If a, b are elements of a regular semigroup S with set E of idempotents, then we define $a \leq b$ if

$$R_a \leq R_b \text{ and } (\exists e \in E \cap R_a)\ a = eb. \tag{6.1.1}$$

Then we have

Proposition 6.1.1 *Let S be a regular semigroup with set E of idempotents. Then the relation $\leq$ defined by* (6.1.1) *is a partial order relation. Within E the order coincides with the natural order among idempotents:*

$$e \leq f \text{ if and only if } ef = fe = e.$$

Proof It is clear that $a \leq a$ for every a in S—simply choose $e = aa'$. Suppose now that $a \leq b$ and $b \leq a$. Then certainly $a \mathcal{R} b$. Also, there exist idempotents e, f in $R_a = R_b$ such that $a = eb$ and $b = fa$. Since $e \mathcal{R} f$ we have $fe = e$, and it then easily follows that

$$a = eb = feb = fa = b.$$

To show that $\leq$ is transitive, suppose that $a \leq b$ and $b \leq c$. Certainly $R_a \leq R_b \leq R_c$, and there exist e in $E \cap R_a$ and f in $E \cap R_b$ such that $a = eb$ and $b = fc$. Now $R_e = R_a \leq R_b = R_f$, and so $fe = e$. Hence

$$(ef)^2 = e(fe)f = e^2 f = ef.$$

We now have $a = (ef)c$, and from

$$R_a = R_{efc} \leq R_{ef} \leq R_e = R_a$$

we have that $ef \in E \cap R_a$.

To prove the final assertion, observe that, for all e, f in E, $e \leq f$ if and only if $R_e \leq R_f$ and there exists i in $E \cap R_e$ such that $e = if$, that is, if and only if $fe = e$ and $ef = e$. □

The one-sided nature of the definition (6.1.1) is only apparent. The following result makes this clear, since conditions (3), (4) and (5) are left/right symmetric.

Theorem 6.1.2 *Let a, b be elements of a regular semigroup S with set E of idempotents. Then the following statements are equivalent:*

(1) $a \leq b$;
(2) $a \in bS$ *and* $\big(\exists a' \in V(a)\big)\ a = aa'b$;
(3) $(\exists e, f \in E)\ a = eb = bf$;
(4) $H_a \leq H_b$ *and* $\big(\forall b' \in V(b)\big)\ a = ab'a$;
(5) $H_a \leq H_b$ *and* $\big(\exists b' \in V(b)\big)\ a = ab'a$.

Proof (1) $\Rightarrow$ (2) is clear, since $e \in E \cap R_a$ if and only if there exists a' in $V(a)$ such that $aa' = e$.

(2) $\Rightarrow$ (3). We are supposing that $a = bu$ for some u in S, and that $a = (aa')b$. Clearly we take e as aa'. Now notice that

$$(ua'b)^2 = ua'bua'b = ua'aa'b = ua'b.$$

So define f as $ua'b$, and observe that

$$bf = bua'b = aa'b = a.$$

(3) $\Rightarrow$ (4). Suppose that $a = eb = bf$, with $e, f \in E$. Then $R_a \leq R_b$ and $L_a \leq L_b$, and so $H_a \leq H_b$. Also, for every b' in $V(b)$,

$$ab'a = ebb'bf = ebf = a.$$

(4) $\Rightarrow$ (5) is clear.

(5) $\Rightarrow$ (1). Suppose that $H_a \leq H_b$ and that there exists an inverse b' of b for which $a = ab'a$. Certainly $R_a \leq R_b$. For every inverse a' of a we see that

$$a(a'ab')a = ab'a = a \ \text{ and } \ (a'ab')a(a'ab') = a'(ab'a)a'ab' = a'ab';$$

hence $a'ab' \in V(a)$. Let $e = aa'ab'$; then $e \in E \cap R_a$. From $L_a \leq L_b$ we deduce that $a = ub$ for some u in S. Then

$$eb = aa'ab'b = ab'b = ubb'b = ub = a.$$ □

Remarks It is a consequence of this theorem that the order $\leq$ can be defined also by the left/right duals of the one-sided definitions. Thus, for example,

$$a \leq b \text{ if and only if } L_a \leq L_b \text{ and } (\exists e \in E \cap L_a)\ a = be. \qquad (6.1.2)$$

It is clear that the order $\leq$ reduces to the usual order defined in Section 5.2 in the case where S is an inverse semigroup.

In an inverse semigroup S the order relation $\leq$ is compatible with the multiplication:

$$a \leq b \text{ and } c \in S \ \Rightarrow\ ca \leq cb \text{ and } ac \leq bc.$$

This is not the case for regular semigroups in general. Indeed we have the following result:

Theorem 6.1.3 *Let S be a regular semigroup with set E of idempotents. Then the following statements are equivalent:*

(1) *S is locally inverse;*
(2) *$\leq$ is compatible;*
(3) *$|S(e,f)| = 1$ for all e, f in E.*

Proof (1) $\Rightarrow$ (2). Let $a \leq b$ and let $c \in S$. Thus $R_a \leq R_b$, and there exists e in $E \cap R_a$ such that $a = eb$. Let a' in $V(a)$ be such that $aa' = e$, choose c' in $V(c)$, and let g be an element of the sandwich set $S(a'a, cc')$. (Thus $ga'a = cc'g = g$ and $a'agcc' = a'acc'$.) Also $c'ga' \in V(ac)$ by Proposition 2.5.3, and so the element $f = acc'ga' \in E \cap R_{ac}$. Also

$$f(bc) = acc'ga'bc = aga'bc = aga'aa'bc = aga'ebc = aga'ac = agc = ac.$$

We must now show that $R_{ac} \leq R_{bc}$. From $R_a \leq R_b$ we deduce that $a = bu$ for some u in S. Hence for all b' in $V(b)$ we have

$$(b'a)^2 = b'ab'a = b'ebb'bu = b'ebu = b'ea = b'a;$$

thus $b'a \in E$. Moreover,

$$b'b.b'a = b'a, \quad b'a.b'b = b'ebb'b = b'eb = b'a,$$

and so $b'a \leq b'b$.

From

$$a = bu = bb'bu = bb'a \tag{6.1.3}$$

we deduce that $a \mathcal{L} b'a$, and it follows that there exists an inverse a'' of a such that $a''a = b'a$. To summarize, we now have

$$a''a = b'a \leq b'b. \tag{6.1.4}$$

Also, from (6.1.3) we deduce that

$$a = ba''a. \tag{6.1.5}$$

As before, let $c' \in V(c)$, and let $h \in S(a''a, cc')$. Then from (6.1.4) we have

$$(a''ah)^2 = a''a(ha''a)h = a''ah^2 = a''ah,$$
$$(b'bh)^2 = b'bha''ab'bh = b'bha''ah = b'bh^2 = b'bh,$$

and so $a''ah, b'bh \in E$. In fact

$$a''ah = a''aha''a = b'ba''aha''ab'b \in b'bSb'b,$$

$$b'bh = b'bha''a = b'bha''ab'b \in b'bSb'b,$$

and so both $a''ah$ and $b'bh$ are idempotents within the inverse semigroup $b'bSb'b$. We deduce that

$$a''ah = a''aha''ah = a''aha''ab'bh = (a''ah)(b'bh) = (b'bh)(a''ah) = b'bh.$$

Finally, denoting the idempotent $c'ha''ac$ by f, we conclude, using (6.1.5), that

$$(bc)f = bcc'ha''ac = bhc = bb'bhc = ba''ahc = ahc = ac,$$

and so $R_{ac} \le R_{bc}$ as required.

(2) $\Rightarrow$ (3). Let $g, h \in S(e, f)$, where $e, f \in E$. Then in particular $fg = g$ and so $(gf)^2 = g(fg)f = g^2f = gf$. Moreover,

$$f(gf) = gf, \quad (gf)f = gf,$$

and so $gf \le f$. Similarly $eg \in E$ and $eg \le e$. By compatibility we deduce that

$$gh = g(fh) = (gf)h \le fh = h, \quad hg = (he)g = h(eg) \le he = h.$$

That is,

$$(gh)h = h(gh) = gh, \quad (hg)h = h(hg) = hg,$$

and so $gh = hg$. However, by Proposition 2.5.3, $S(e, f)$ is a rectangular band. Hence

$$g = ghg = g^2h = gh = hg = h(hg) = hgh = h.$$

We conclude that $|S(e, f)| = 1$.

(3) $\Rightarrow$ (1). Let $e \in E$, let $a \in eSe$, and let $a' \in V(a) \cap eSe$. Then $a'a \in S(a'a, e)$, for $a'aa'a = a'a$, $ea'a = a'a$ and $a'a(a'a)e = a'ae$. Hence in fact, by our assumption, $a'a$ is the *only* element in $S(a'a, e)$. By the same token, if a'' is another inverse of a in eSe then $S(a''a, e) = \{a''a\}$. But, by Proposition 2.5.2, $S(a'a, e) = S(a''a, e)$, and so it follows that $a''a = a'a$. Similarly, by considering $S(e, aa')$ and $S(e, aa'')$, we deduce that $aa'' = aa'$, and it now follows that

$$a'' = a''aa'' = a'aa'' = a'aa' = a'.$$

Hence eSe is an inverse semigroup. □

6.2 ORTHODOX SEMIGROUPS

Recall from the introduction that a semigroup is called *orthodox* if it is regular and if its idempotents form a subsemigroup. Several alternative definitions are available:

Theorem 6.2.1 *Let S be a regular semigroup with set E of idempotents. Then the following statements are equivalent:*

(1) *S is orthodox;*

(2) $(\forall e, f \in E)\ fe \in S(e, f)$;
(3) $(\forall a, b \in S)\ V(b)V(a) \subseteq V(ab)$;
(4) $(\forall e \in E)\ V(e) \subseteq E$.

Proof (1) $\Rightarrow$ (2). Suppose that S is orthodox, let $e, f \in E$, and let $g = fe$. Then

$$ge = fg = g, \quad egf = (ef)^2 = ef,$$

and so $g = fe \in S(e, f)$ by (2.5.2).

(2) $\Rightarrow$ (3). Let $a, b \in S$ and let $a' \in V(a)$, $b' \in V(b)$. Then, by Proposition 2.5.3, $b'ga' \in V(ab)$ for all g in $S(a'a, bb')$. From (2) it thus follows that

$$b'a' = b'(bb'a'a)a' \in V(ab),$$

exactly as required.

(3) $\Rightarrow$ (4). Let $e \in E$ and let x be an inverse of e:

$$xex = x, \quad exe = e.$$

Now xe and ex are both idempotent, and so each is an inverse of itself. By (3) we deduce that $(ex)(xe)$ is an inverse of $(xe)(ex)$, that is to say, that ex^2e is an inverse of $xe^2x = xex = x$. Hence

$$x = x(ex^2e)x = (xex)(xex) = (xex)^2 = x^2,$$

and so x is idempotent as required.

(4) $\Rightarrow$ (1). Let $e, f \in E$. By Proposition 2.5.1 there exists an idempotent g in $V(ef)$ (an element of the sandwich set $S(e, f)$). But then ef, being an inverse of the idempotent g, must itself be idempotent. Hence S is orthodox. □

The property (3), a generalization of the property

$$(ab)^{-1} = b^{-1}a^{-1}$$

possessed by inverse semigroups, enables us to produce in modified form some of the key properties of inverse semigroups. Certainly the echo of inverse semigroup theory is clear in the next result:

Proposition 6.2.2 *Let S be an orthodox semigroup with set E of idempotents. For all a in S, e in E and a' in $V(a)$, the elements aea' and $a'ea$ are idempotent.*

Proof With the given notation,

$$(aea')^2 = aea'aea' = aea'aea'aa' = a(ea'a)^2a' = aea'aa' = aea'.$$

Thus aea' is idempotent. The proof for $a'ea$ is similar. □

The set E of idempotents in an orthodox semigroup S forms a band under multiplication, and this is, by Theorem 4.4.1, expressible as a semilattice Y of rectangular bands E_α $(\alpha \in Y)$. Certainly $E_\alpha \cap E_\beta = \emptyset$ if $\alpha \neq \beta$, and we also have

$$E_\alpha E_\beta \subseteq E_{\alpha\beta} \quad (\alpha, \beta \in Y). \tag{6.2.1}$$

Each E_α is a $\mathcal{J}^E$-class, and it will be consistent with our previous notation to write J_e^E for the rectangular band E_α containing e. The formula (6.2.1) 'translates' to

$$J_e^E J_f^E \subseteq J_{ef}^E = J_{fe}^E \quad (e, f \in E). \tag{6.2.2}$$

The equivalence $\mathcal{J}^E$ is the minimum semilattice congruence on E.

From Theorem 6.2.1 we know that $V(e) \subset E$ for every e in E. In fact, if $f \in V(e)$ then

$$efe = e, \quad fef = f,$$

and it is clear that $f \in J_e^E$. Conversely, if f belongs to the rectangular band J_e^E then certainly $f \in V(e)$, since any two elements of a rectangular band are mutually inverse. Hence

$$V(e) = J_e^E \quad (e \in E). \tag{6.2.3}$$

Thus $V(e)$ is determined solely by the nature of the band E.

We cannot expect the set $V(a)$ of inverses of an arbitrary element of S to be solely determined by properties of E, but in fact the properties of E are highly influential, in the sense that if we know a single inverse a' of a then $V(a)$ is wholly determined by a' and by E:

Proposition 6.2.3 *Let $a \in S$, an orthodox semigroup with band E of idempotents. If a' is an inverse of a, then*

$$V(a) = J_{a'a}^E a' J_{aa'}^E.$$

Proof Let $e \in J_{a'a}^E$ and $f \in J_{aa'}^E$. Then

$$a'aea'a = a'a, \quad aa'faa' = aa',$$

and so

$$\begin{aligned} a(ea'f)a &= aa'aea'aa'aa'faa'a = a(a'aea'a)a'(aa'faa')a \\ &= aa'aa'aa'a = a, \end{aligned}$$

and

$$\begin{aligned} (ea'f)a(ea'f) &= ea'aa'faa'aa'aea'aa'f = ea'(aa'faa')a(a'aea'a)a'f \\ &= ea'(aa')a(a'a)a'f = ea'f. \end{aligned}$$

Thus $J_{a'a}^E a' J_{aa'}^E \subseteq V(a)$.

Conversely, suppose that $a^* \in V(a)$. Then

$$a^* = a^*aa^* = a^*aa'aa^*. \tag{6.2.4}$$

Now, from

$$(a^*a)(a'a)(a^*a) = a^*(aa'a)a^*a = a^*aa^*a = a^*a$$

and

$$(a'a)(a^*a)(a'a) = a'(aa^*a)a'a = a'aa'a = a'a$$

we deduce that $a^*a \in J^E_{a'a}$. A similar argument shows that $aa^* \in J^E_{aa'}$, and it is now immediate from (6.2.4) that $V(a) \subseteq J^E_{a'a}a'J^E_{aa'}$. □

We can now give yet another characterization of orthodox semigroups:

Theorem 6.2.4 *A regular semigroup S is orthodox if and only if*

$$(\forall a, b \in S)\ [V(a) \cap V(b) \neq \emptyset \ \Rightarrow\ V(a) = V(b)].$$

Proof Suppose first that S is orthodox, and that a, b in S are such that $x \in V(a) \cap V(b)$. Then a and b both belong to $V(x)$ and so, by Theorem 2.3.4, $xa\ \mathcal{R}^S\ xb$ and $ax\ \mathcal{L}^S\ bx$. Now $xa, xb, ax, bx \in E$, and so, by Proposition 2.4.2, $xa\ \mathcal{R}^E\ xb$ and $ax\ \mathcal{L}^E\ bx$. Certainly $xa\ \mathcal{J}^E\ xb$ and $ax\ \mathcal{J}^E\ bx$, and so

$$V(a) = J^E_{xa}xJ^E_{ax} = J^E_{xb}xJ^E_{bx} = V(b).$$

Conversely, suppose that S is regular and that we have the given implication. Let $e, f \in E$ and let $g \in S(e, f)$. Then from $ge = g$ we may deduce that eg is idempotent. Also

$$g(eg)g = g,\ (eg)g(eg) = eg,$$

and so we have that $g \in V(g) \cap V(eg)$. From our assumption we deduce that $V(g) = V(eg)$. Hence in particular $ef \in V(eg)$, and so

$$ef = (ef)(eg)(ef) = (ef)(efg)(ef) = (ef)(efgef) = (ef)^2.$$

Thus S is orthodox. □

The equivalence relation

$$\gamma = \{(x, y) \in S \times S : V(x) = V(y)\} \tag{6.2.5}$$

on an orthodox semigroup S turns out to be a congruence. Indeed, we have

Theorem 6.2.5 *Let S be an orthodox semigroup with set E of idempotents. Then the equivalence γ defined by (6.2.5) is the smallest inverse semigroup congruence on S. Moreover, for each a in S and each a' in $V(a)$,*

$$a\gamma = J^E_{aa'}aJ^E_{a'a}.$$

Proof To show that γ is a congruence, consider (a, b) in γ and let $c \in S$. Then, for every x in $V(a)$ $(= V(b))$ and for every c' in $V(c)$, we have $xc' \in V(ca) \cap V(cb)$. Hence $V(ca) = V(cb)$ by Theorem 6.2.4. A similar argument shows that $V(ac) = V(bc)$, and so γ is a congruence.

The quotient S/γ is certainly regular. By Lallement's Lemma (Lemma 2.4.3) each idempotent of S/γ is of the form $e\gamma$, where e is an idempotent of S. Now, for any two idempotents e, f in E,

$$\begin{aligned} V(ef) &= J^E_{ef} = J^E_{fe} \text{ (by (6.2.2))} \\ &= V(fe), \end{aligned}$$

and from this we deduce that $(e\gamma)(f\gamma) = (f\gamma)(e\gamma)$ in S/γ.

Finally, to show that γ is the *least* inverse semigroup congruence, let ρ be a congruence on S such that S/ρ is an inverse semigroup, let $(a,b) \in \gamma$, and let $x \in V(a)$ $\big(= V(b)\big)$. Then both $a\rho$ and $b\rho$ are inverses of $x\rho$ in the inverse semigroup S/ρ, and so $a\rho = b\rho$. We have shown that $\gamma \subseteq \rho$.

To prove the final statement of the theorem, suppose that $b \in a\gamma$. Then $V(a) = V(b)$, and so $a' \in V(b)$ for every a' in $V(a)$. It now follows from Proposition 6.2.3 that

$$b \in V(a') = J^E_{aa'} a J^E_{a'a}.$$

Conversely, if $b \in J^E_{aa'} a J^E_{a'a} = V(a')$, then $V(a) \cap V(b) \neq \emptyset$, and so $V(a) = V(b)$ by Theorem 6.2.4. Thus $b \in a\gamma$, as required. □

6.3 SEMIBANDS

A regular semigroup generated by its idempotents is called a *semiband.* Semibands differ from locally inverse and orthodox semigroups in the sense that they are not generalizations of inverse semigroups. A regular semigroup is orthodox and a semiband if and only if it is a band, and it is both an inverse semigroup and a semiband if and only if it is a semilattice. It is, however, possible for a semigroup to be non-trivially both a semiband and a locally inverse semigroup—see Exercise 11.

No substantial theory of semibands is available as yet. The justification for this section lies in the frequency with which such semigroups occur 'in nature', and in the universal property (Theorem 6.3.4) they possess.

Consider the set Sing_n of all singular maps from the set

$$[n] = \{1, 2, \ldots, n\}$$

into itself. (By a *singular* map we mean one that is not a bijection.) This is a finite semigroup, of order $n^n - n!$. Then we have

Theorem 6.3.1 *For all $n \geq 2$, the semigroup Sing_n is a semiband.*

Proof To show that Sing_n is regular, let $\alpha \in \mathrm{Sing}_n$, and define $\xi : [n] \to [n]$ as follows: if $j \in \operatorname{im}\alpha$, let $j\xi$ be an arbitrarily chosen element of $j\alpha^{-1}$; if $j \notin \operatorname{im}\alpha$, let $j\xi$ be an arbitrarily chosen element of $[n]$. Then it is clear that $i\alpha\xi\alpha = i\alpha$ for all i in $[n]$. Of course ξ may be a permutation, but certainly $\eta = \xi\alpha\xi$ is singular, and

$$\alpha\eta\alpha = \alpha\xi\alpha\xi\alpha = \alpha\xi\alpha = \alpha.$$

The semigroup Sing_n has $n-1$ $\mathcal{J}$-classes $J_1, \ldots, J_{n-1}$, where

$$J_r = \{\alpha \in \mathrm{Sing}_n : |\operatorname{im}\alpha| = r\} \quad (r = 1, \ldots, n-1).$$

Let E_{n-1} denote the set of idempotents in J_{n-1}. A typical element ϵ of E_{n-1} has image $[n] \setminus \{i\}$ of cardinality $n-1$. The map ϵ acts identically on $[n] \setminus \{i\}$, and sends i to some element $j \neq i$. We denote this map by $\binom{i}{j}$; it maps i to j and all other elements identically. Notice that we can easily deduce that $|E_{n-1}| = n(n-1)$.

Lemma 6.3.2 *Let $\alpha \in J_r$, where $1 \leq r \leq n-1$. Then there exist ϵ in E_{n-1} and β in J_{r+1} such that $\alpha = \epsilon\beta$.*

Proof Write $\operatorname{im}\alpha = \{b_1, b_2, \ldots, b_r\}$, and let $b_i\alpha^{-1} = A_i \quad (i = 1, 2, \ldots, r)$. It is convenient to write

$$\alpha = \begin{pmatrix} A_1 & A_2 & \ldots & A_r \\ b_1 & b_2 & \ldots & b_r \end{pmatrix},$$

in an obvious extension of a familiar notation. The sets A_i form a partition of $[n]$. Since not all of the sets A_i are singletons, we may assume without loss of generality that $A_1 = \{a_1, a_1', \ldots\}$ has at least two elements. Then let

$$\epsilon = \begin{pmatrix} a_1 \\ a_1' \end{pmatrix}, \quad \beta = \begin{pmatrix} A_1 \setminus \{a_1\} & A_2 & \ldots & A_r & \{a_1\} \\ b_1 & b_2 & \ldots & b_r & b_{r+1} \end{pmatrix},$$

where $b_{r+1} \notin \operatorname{im}\alpha$, and verify that $\alpha = \epsilon\beta$. □

As a consequence of this lemma we easily deduce that Sing_n is generated by J_{n-1}. The task remaining is to show that every element of J_{n-1} is expressible as a product of elements in E_{n-1}.

The most illuminating way to prove this is to make use of the directed graph $\Gamma(\alpha)$ associated with any $\alpha : [n] \to [n]$. The vertices of $\Gamma(\alpha)$ are labelled $1, 2, \ldots, n$, and there is a directed edge $i \longrightarrow j$ if and only if $i\alpha = j$. The graph may fail to be connected: its connected components are the ω-classes, where ω is the equivalence relation given by

$$\omega = \{(i, j) \in [n] \times [n] : (\exists r, s \geq 0)\ i\alpha^r = j\alpha^s\}.$$

Each ω-class Ω has a *kernel* $K(\Omega)$, defined by

$$K(\Omega) = \{i \in \Omega : (\exists r > 0)\ i\alpha^r = i\}.$$

To see that $K(\Omega)$ is non-empty, consider an element i in Ω. The elements

$$i, i\alpha, i\alpha^2, \ldots$$

cannot all be distinct, and so there exist $m \geq 0$ and $r \geq 1$ such that $i\alpha^{m+r} = i\alpha^m$. Thus $i\alpha^m \in K(\Omega)$.

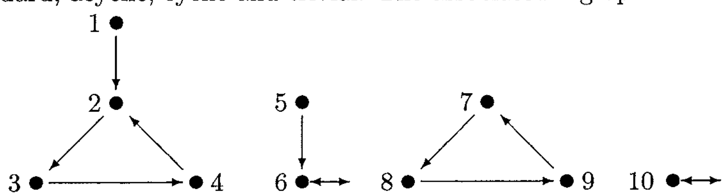
It is useful to distinguish four types of component Ω:

$$\begin{array}{rcl} \textit{standard}\text{ components} & : & 2 \le |K(\Omega)| < |\Omega|; \\ \textit{acyclic}\text{ components} & : & 1 = |K(\Omega)| < |\Omega|; \\ \textit{cyclic}\text{ components} & : & 2 \le |K(\Omega)| = |\Omega|; \\ \textit{trivial}\text{ components} & : & 1 = |K(\Omega)| = |\Omega|. \end{array}$$

Every component falls into exactly one of these categories, and all four cases can arise: for example, the map

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 3 & 4 & 2 & 6 & 6 & 8 & 9 & 7 & 10 \end{pmatrix}$$

has components $\{1,2,3,4\}$, $\{5,6\}$, $\{7,8,9\}$, $\{10\}$ which are, respectively, standard, acyclic, cyclic and trivial. The associated digraph is



Let α be an arbitrary element of Sing_n. Each standard or acyclic component Ω contains at least one element not in $\operatorname{im}\alpha$. To see this, suppose by way of contradiction that $i\alpha^{-1} \neq \emptyset$ for all $i \in \Omega$, and consider an element j_0 in $\Omega \setminus K(\Omega)$. Let

$$j_1 \in j_0\alpha^{-1},\ j_2 \in j_1\alpha^{-1}, \ldots, j_{n+1} \in j_n\alpha^{-1}, \ldots$$

The elements $j_1, j_2, \ldots$ cannot all be distinct, and so there exist r, $r+m$ with $m \ge 1$ and $j_{r+m} = j_r$. It follows that

$$j_0 = j_r\alpha^r = j_r\alpha^{m+r} = j_0\alpha^m.$$

Thus $j \in K(\Omega)$, contrary to assumption.

Now consider an element α in J_{n-1}. Since $|\operatorname{im}\alpha| = n-1$, all the components of α are cyclic or trivial except one, which is either acyclic or standard. Let us suppose first that the 'rogue' component is acyclic, so that the components are

$$A;\ Z_1, Z_2, \ldots, Z_c;\ \{u_1\}, \{u_2\}, \ldots, \{u_t\},$$

where A is acyclic, $Z_1, Z_2, \ldots, Z_c$ are cyclic, and $\{u_1\}, \{u_2\}, \ldots, \{u_t\}$ are trivial. Suppose also that $|A| = a$ and $|Z_i| = z_i$ for $i = 1, 2, \ldots, c$. Thus

$$a + \sum_{i=1}^{c} z_i + t = n. \tag{6.3.1}$$

Let x_1 be the unique element in $A \setminus \operatorname{im}\alpha$. Then

$$A = \{x_1, x_2, \ldots, x_a\},$$

where $x_{j+1} = x_j\alpha$ for $j = 1, 2, \ldots, a-1$ and $x_a\alpha = x_a$. Also, for $i = 1, 2, \ldots, c$, let

$$Z_i = \{v_{i1}, v_{i2}, \ldots, v_{iz_i}\},$$

where $v_{i,j+1} = v_{ij}\alpha$ for $j = 1, 2, \ldots, z_i - 1$ and $v_{iz_i}\alpha = v_{i1}$. Then

$$\alpha = \beta\gamma_1\gamma_2\ldots\gamma_c, \tag{6.3.2}$$

where

$$\beta = \begin{pmatrix} x_{a-1} \\ x_a \end{pmatrix}\begin{pmatrix} x_{a-2} \\ x_{a-1} \end{pmatrix}\cdots\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

and, for $i = 1, 2, \ldots, c$,

$$\gamma_i = \begin{pmatrix} v_{iz_i} \\ x_1 \end{pmatrix}\begin{pmatrix} v_{i,z_i-1} \\ v_{iz_i} \end{pmatrix}\begin{pmatrix} v_{i,z_i-2} \\ v_{i,z_i-1} \end{pmatrix}\cdots\begin{pmatrix} v_{i1} \\ v_{i2} \end{pmatrix}\begin{pmatrix} x_1 \\ v_{i1} \end{pmatrix}. \tag{6.3.3}$$

Thus α is a product of idempotents in E_{n-1}.

Suppose now that the components of α are

$$S;\ Z_1, Z_2, \ldots, Z_c;\ \{u_1\}, \{u_2\}, \ldots, \{u_t\},$$

where S is now a standard component, and the other components are as before. Again, let x_1 be the unique element in $S \setminus \operatorname{im}\alpha$, and suppose that $x_m = x_1\alpha^{m-1} \in K(S)$, with m chosen as small as possible. Then

$$S = \{x_1, x_2, \ldots, x_{m+p}\},$$

with $x_i\alpha = x_{i+1}$ for $i = 1, 2, \ldots, m+p-1$, and $x_{m+p}\alpha = x_m$. If we now define

$$\beta' = \begin{pmatrix} x_{m+p} \\ x_{m-1} \end{pmatrix}\begin{pmatrix} x_{m+p-1} \\ x_{m+p} \end{pmatrix}\begin{pmatrix} x_{m+p-2} \\ x_{m+p-1} \end{pmatrix}\cdots\begin{pmatrix} x_{m-1} \\ x_m \end{pmatrix}\begin{pmatrix} x_{m-2} \\ x_{m-1} \end{pmatrix}\cdots\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

we can verify that

$$\alpha = \beta'\gamma_1\gamma_2\ldots\gamma_c, \tag{6.3.4}$$

where $\gamma_1, \gamma_2, \ldots, \gamma_c$ are as in (6.3.3). Thus again α is expressed as a product of idempotents from E_{n-1}. □

Remark The number of idempotents in the product (6.3.2) is

$$\begin{aligned}(a-1) + \sum_{i=1}^{c}(z_i+1) &= a - 1 + \sum_{i=1}^{c} z_i + c \\ &= n + c - t - 1;\end{aligned}$$

the number of idempotents in the product (6.3.4) is

$$m + p + \sum_{i=1}^{c}(z_i+1) = n + c - t.$$

In both cases, the length of the product is

$$n + \operatorname{cycl}\alpha - \operatorname{fix}\alpha,$$

where $\operatorname{cycl}\alpha$ is the number of cyclic components and

$$\operatorname{fix}\alpha = |\{i \in [n] : i\alpha = i\}|.$$

This has been shown by Iwahori (1977) (see also Howie (1980)) to be best possible, in the sense that α cannot be expressed as a shorter product of elements of E_{n-1}.

Corollary 6.3.3 *Every finite semigroup is embeddable in a finite semiband.*

Proof Let S be a finite semigroup and let $X = S^1 \cup \{y, z\}$, where $y, z \notin S^1$. Define a map $\alpha : S \to \mathcal{T}_X$ by $s\alpha = \rho_s$, where

$$x\rho_s = xs \ \text{ if } \ x \in S^1,$$

$$y\rho_s = z\rho_s = z.$$

It is a routine matter to verify that α is a monomorphism. Moreover, it is clear that $s\alpha$ is a singular element of $\mathcal{T}_X$ for every s in S, and so α embeds S in the finite semiband $\operatorname{Sing}_{|X|}$. □

The removal of 'finite' from Corollary 6.3.3 presents some difficulty, for the fact that $\operatorname{Sing}_{|X|}$ is a semiband depends heavily on the finiteness of X. It is possible to identify the idempotent-generated part $\langle E \rangle$ of $\mathcal{T}_X$, and then to use a modified regular representation along the lines of the proof of Corollary 6.3.3—see Howie (1966). An alternative approach is given below.

Theorem 6.3.4 *Every semigroup is embeddable in a semiband.*

Proof Let S be a semigroup, and let T be a regular semigroup containing S. It is always possible to find such a semigroup T: for example, take $T = \mathcal{T}_{S^1}$. Let I be a set containing a named element 1, and such that $|I \setminus \{1\}|^2 \geq |T|$, and define B to be the Rees matrix semigroup

$$\mathcal{M}[T^1; I, I; P],$$

where the matrix $P = (p_{ij})$ over T^1 has the properties that

$$p_{i1} = p_{1i} = 1 \quad (i \in I)$$

and

$$T \subseteq \{p_{ij} : i, j \neq 1\}.$$

The elements $(1, 1, i)$ and $(i, 1, 1)$ of B are evidently idempotent for all values of i. Also, since each t in T is equal to some p_{kl}, we have

$$(i, t, j) = (i, 1, 1)(1, 1, k)(l, 1, 1)(1, 1, j),$$

a product of idempotents. Thus B is generated by its idempotents.

Next, B is regular, for if $(i, t, j) \in T$ and if t' is an inverse of t in the regular semigroup T, then

$$(i, t, j)(1, t', 1)(i, t, j) = (i, tt't, j) = (i, t, j).$$

Finally, it is clear that the map $t \mapsto (1, t, 1)$ embeds T in B, and so S, as required, is embedded in a semiband B. $\square$

Remark The foregoing argument is of course available in the finite case, and so we have an alternative proof of Corollary 6.3.3.

6.4 EXERCISES

1. The *Lallement order* λ on a regular semigroup S is defined by the rule that $a \,\lambda\, b$ if and only if, for all x, y in S,

$$x \,\mathcal{R}\, xa \;\Rightarrow\; xa = xb, \qquad y \,\mathcal{L}\, ay \;\Rightarrow\; ay = by.$$

Show that λ is a partial order relation, and that it is compatible with multiplication.

2. Given an order relation ω on a semigroup S, define $\omega^\flat$ by the rule that $a \,\omega^\flat\, b$ if and only if $xay \,\omega\, xby$ for all x, y in S^1. Show that $\omega^\flat$ is the largest compatible order relation contained in ω.

3. On a regular semigroup S, show that $\lambda = \nu^\flat$, where ν is the *Nambooripad order* given by (6.1.1) and λ is the Lallement order defined in Exercise 1. Deduce that $\lambda = \nu$ if and only if S is locally inverse.

4. Let S be a regular semigroup with set E of idempotents, and let the relation λ' be defined by the rule that $a \,\lambda'\, b$ if and only if

$$(\forall a' \in V(a))(\forall e \in E \cap aa'S)(\forall f \in E \cap Sa'a)\; ea = eb,\; af = bf.$$

   (a) Show that $\lambda \subseteq \lambda'$.
   (b) Suppose that $(a, b) \in \lambda'$. Show that $x \,\mathcal{R}\, xa$ implies that $ga = gb$, where g is in the sandwich set $S(x'x, aa')$, and that $y \,\mathcal{L}\, ay$ implies that $ah = bh$, where $h \in S(a'a, yy')$.
   (c) Deduce that $\lambda = \lambda'$.
   (d) Show that, if S is orthodox, then λ is equal to

$$\{(a, b) \in S \times S : (\forall a' \in V(a))(\forall e \in E)\; aea' = bea',\; a'ea = a'eb\}.$$

5. On an arbitrary semigroup S the *Mitsch order* μ is defined by the rule that $a \,\mu\, b$ if and only if

$$(\exists s, t \in S^1)\; sa = sb = a = at = bt.$$

Show that μ is an order relation. Show also that μ coincides with ν, the Nambooripad order, when S is regular.

6. A semigroup S with set E of idempotents is called *R-unipotent* if it is regular and if

$$(\forall e, f \in E)\; efe = ef.$$

Show that the following statements about a regular semigroup S are equivalent:

(a) S is R-unipotent;
(b) each $\mathcal{R}$-class of S contains exactly one idempotent;
(c) $(\forall e \in E)(\forall a \in S)(\forall a' \in V(a))\ aea'a = ae$;
(d) $(\forall a \in S)(\forall a', a'' \in V(a))\ aa' = aa''$.

7. Let γ be the least inverse congruence on an orthodox semigroup S, as described by (6.2.5). Show that $\gamma \cap \mathcal{H} = 1_S$.

8. Let S be an orthodox semigroup, with band E of idempotents. Show that, for all a in S and all a' in $V(a)$,
$$aV(a) = R^E_{aa'}, \quad V(a)a = L^E_{a'a}.$$

9. Let S be an orthodox semigroup, and let
$$T = \{(a, a') \in S \times S : a' \in V(a)\}.$$
Show that, relative to the multiplication
$$(a, a')(b, b') = (ab, b'a'),$$
T is an orthodox semigroup.

10. Let S be an orthodox semigroup, and suppose that E, the band of idempotents of S, is a rectangular band. Show that $S \simeq G \times E$, where G is a group. [Hint: show first that S is simple, and use Exercise 3(8).]

11. Consider a completely simple semigroup $S = \mathcal{M}[G; I, \Lambda; P]$, where G is a group with identity e, where $1 \in I \cap \Lambda$ and where the $\Lambda \times I$ matrix $P = (p_{\lambda i})$ has the property that
$$p_{\lambda 1} = p_{1i} = 1 \quad (\lambda \in \Lambda,\ i \in I).$$
(See Theorem 3.4.2, where such a P is called *normal.*) Show that S is generated by idempotents if and only if the entries $p_{\lambda i}$ $(\lambda \neq 1,\ i \neq 1)$ generate G.

    Deduce that in such a case S is a locally inverse semiband.

12. Let S be a semiband. If
$$a = e_1 e_2 \ldots e_k,$$
a product of idempotents in S, show that, for each a' in $V(a)$,
$$f_i = e_i e_{i+1} \ldots e_k a' e_1 e_2 \ldots e_i$$
is idempotent, and that
$$a = f_1 f_2 \ldots f_k.$$
Deduce that every element a of S is expressible as a product of idempotents from within its own $\mathcal{J}$-class J_a.

13. Show that if $a = eb$ in a regular semigroup, where e is idempotent, then there exists an idempotent in R_a such that $a = fb$. Deduce from this and its dual that if $a = e_1 e_2 \dots e_n$, a product of idempotents, then we may assume that $e_1 \mathcal{R} a$ and $e_n \mathcal{L} a$.

14. Let E be the set of idempotents in Sing_n. If $j \in \mathrm{im}\,\alpha$, where $\alpha \in \mathrm{Sing}_n$, we say that the $(\ker\alpha)$-class $j\alpha^{-1}$ is *stationary* if $j \in j\alpha^{-1}$. Show that $\alpha \in E^2$ if and only if every non-stationary $(\ker\alpha)$-class contains an element of $[n] \setminus \mathrm{im}\,\alpha$.

    Deduce that $\alpha \in E^2$ implies that

$$\mathrm{rank}\,\alpha \leq (n + \mathrm{fix}\,\alpha)/2,$$

where $\mathrm{rank}\,\alpha = |\,\mathrm{im}\,\alpha|$, $\mathrm{fix}\,\alpha = |\{i \in [n] : i\alpha = i\}|$.

6.5 NOTES

The main reference for Section 6.1 is Nambooripad (1980). For Section 6.2 the principal references are T. E. Hall (1969, 1970, 1971). Hall (1970) goes on to generalize the Munn semigroup (see Section 5.4) to the orthodox case, and to prove a quite detailed structure result for orthodox semigroups. See also T. E. Hall (1989) for an associated isomorphism theorem. $\mathcal{R}$-unipotent semigroups (see Exercise 6), which are necessarily orthodox, have been studied by various authors, including Edwards (1977), Feigenbaum (1979), La Torre (1981) and Gomes (1985, 1986a,b).

References for Section 6.3 are Howie (1966, 1980) and Iwahori (1977). The construction in the proof of Theorem 6.3.4 appears in Giraldes and Howie (1984), where it is attributed to T. E. Hall.

The order λ featuring in Exercises 1 to 4 is from Lallement (1966), while ν is from Nambooripad (1980), and μ (see Exercise 5) is from Mitsch (1986). See also Higgins (1994). Exercises 3 and 4 are from Gomes (1983), and Exercise 6 is from Venkatesan (1974) and Edwards (1977). Exercises 7 and 8 are from T. E. Hall (1970), while Exercise 9 is from Newton (1993).

Exercise 11 is from Howie (1978), Exercises 12 and 13 are from T. E. Hall (1973), and Exercise 14 is from Howie *et al.* (1988).

7

Free semigroups

In Chapter 1 we came across the notion of the free semigroup A^+ and the free monoid A^* on a set A. While the emphasis of this book is unquestionably on the study of various classes of regular semigroups, it would be wrong in a general introduction not to include a short chapter on free semigroups, for in the wide and growing applications of semigroup theory to automata, languages and machines it is the free semigroups that play the major role.

This will be a fairly brief account of a large and vital area, the intention being to convey the flavour and some of the methods. For more information, the reader is referred to several more specialized texts, mentioned in the Notes at the end of the chapter The flavour and the methods are certainly very different from those of the previous chapters, for the Green equivalences that have been such an important tool in studying regular semigroups all reduce to the identity relation 1 in a free semigroup or monoid.

Section 7.2 is devoted to a brief account of codes. The notion of a (variable length) *code* is a natural one both for language theory and for algebra, and codes of various kinds have been extensively studied in recent years.

7.1 PROPERTIES OF FREE SEMIGROUPS

The importance in applications of free semigroups and monoids lies in the fact that the elements are *words* in an alphabet. They are *strings of symbols*, and such strings are encountered in a huge variety of contexts. Any spoken statement, any written statement, any input to a computer is a string of symbols. Every Shakespeare play is an element of A^+, where A is the set consisting of the 52 upper- and lower-case letters of the alphabet together with various space and punctuation symbols. This is not to claim that the study of A^+ includes the literary study of Shakespeare's plays, but it does indicate that in studying A^+ we are studying something both general and ubiquitous.

From the fact that two words in the alphabet A represent the same element of A^* if and only if they are identical, we easily deduce that

Proposition 7.1.1 *The free monoid A^* is cancellative.* □

A semigroup (or monoid) S is called *equidivisible* if, for all s, t, u, v in S, $st = uv$ implies either

(i) there exists x in S^1 such that $s = ux$ and $v = xt$; or

(ii) there exists y in S^1 such that $u = sy$ and $t = yv$.

Notice that every *group* is equidivisible: simply define $x = u^{-1}s$ in (i), or $y = s^{-1}u$ in (ii). More generally, every completely simple semigroup is equidivisible—see Exercise 1. From our point of view the important result is

Proposition 7.1.2 *The free monoid A^* is equidivisible.*

Proof Let $st = uv = a_1a_2\dots a_m$, where $a_1, a_2, \dots, a_m \in A$. Then, for some k, l in $\{0, 1, \dots, m\}$,

$$s = a_1a_2\dots a_k,\ t = a_{k+1}a_{k+2}\dots a_m,\ u = a_1a_2\dots a_l,\ t = a_{l+1}a_{l+2}\dots a_m.$$

If $k \geq l$, then (i) holds, with $x = a_{l+1}\dots a_k$; if $k \leq l$, then (ii) holds, with $y = a_{k+1}\dots a_l$. □

A semigroup (or monoid) S is said to be a *semigroup (or monoid) with length* if there is a map $s \mapsto |s|$ from S into $\mathbf{N}^0$ such that, for all s, t in S,

$$|st| = |s| + |t|. \tag{7.1.1}$$

Our notations are natural and useful, but notice the initially disconcerting conclusion that in a monoid S it follows from (7.1.1) that $|1| = 0$. The map $s \mapsto |s|$ is a morphism from S into $(\mathbf{N}^0, +)$. We say that S is a monoid with *proper* length if, for all s in S,

$$|s| = 0 \ \Rightarrow\ s = 1.$$

It is clear that every free monoid A^* is a monoid with proper length: simply define $|w| = m$ for each $w = a_1a_2\dots a_m$ in A^* (and $|1| = 0$). Since it is clear that every subsemigroup of a semigroup with (proper) length is a semigroup with (proper) length, it follows that every subsemigroup of A^* is a semigroup with proper length.

Another example of a monoid with length is provided by the multiplicative monoid $F[x]$ of all non-zero polynomials

$$f = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

with coefficients in a field F. Here we define $|f|$ to be the *degree* of f. The length is not proper unless $F = \mathbf{Z}_2 = \{0, 1\}$, but even in this case the monoid $F[x]$ is not free, for it is commutative, and a free monoid A^* is commutative only if $|A| = 1$.

Proposition 7.1.3 *Let U be a subsemigroup or submonoid of a free monoid A^*, and let $V = U \setminus \{1\}$. Then $V \setminus V^2$ is the unique minimum generating set of U.*

Proof It is clear that every generating set of U must contain $V \setminus V^2$. To show that $\langle V \setminus V^2 \rangle = U$, suppose that $l = \min\{|v| : v \in V\}$. Let $v \in V$. If $|v| = l$, then $v \in V \setminus V^2 \subseteq \langle V \setminus V^2 \rangle$. We complete the proof by induction on $|v|$. If $v \in V \setminus V^2$ then there is nothing to prove. Otherwise there exist v_1, v_2 in V such that $v = v_1 v_2$. Now $|v_1|, |v_2| > 0$, since U has proper length, and $|v_1| + |v_2| = |v|$, by (7.1.1). Hence $|v_1|, |v_2| < |v|$, and so we may assume that $v_1, v_2 \in \langle V \setminus V^2 \rangle$. It now follows immediately that $v \in \langle V \setminus V^2 \rangle$. □

Corollary 7.1.4 *The set A is the unique minimum set of generators of A^+ (or A^*), in the sense that every set of generators must contain A.* □

This is in strong contrast to the situation in group theory. For example, the free group generated by $\{x, y\}$ is equally well generated freely by the set $\{x, xy\}$.

The *base* of a submonoid or subsemigroup U of the free monoid A^* is defined to be $V \setminus V^2$, where $V = U \setminus \{1\}$. In particular, the base of A^* (or of A^+) is A.

If $|A| \geq 2$, the free monoid is in fact as far from being commutative as it is possible to be:

Proposition 7.1.5 *Let $|A| \geq 2$, and let $u, v \in A^+$. Then $uv = vu$ if and only if u and v are powers of the same element w.*

Proof Actually this proposition is trivially true even if $|A| = 1$, but is of interest only for larger A.

One way round the result is obvious. Suppose therefore that $uv = vu$, where $|uv|$ ($= |vu|$) $= n$, and suppose inductively that the proposition is true whenever $|uv| < n$. There is no problem about anchoring this induction: if $|uv| = 2$ then it is clear that $uv = vu$ implies that $u = v$. We may suppose, without loss of generality, that $|u| \geq |v|$, and so by equidivisibility it follows from $uv = vu$ that there exists z in A^* such that $u = vz$, $u = zv$. If $z = 1$ then $u = v$. Otherwise, we have $vz = zv$, where $v, z \in A^+$, and $|vz| = |u| < |uv| = n$. By the induction hypothesis, there exists w such that $v = w^p$, $z = w^q$, and it now immediately follows that $u = w^{p+q}$ and $v = w^p$, exactly as required. □

Proposition 7.1.5 says in essence that two elements of A^* commute only if they must. Another example of the same phenomenon is as follows:

Proposition 7.1.6 *Let u, v be elements of a free semigroup A^+. If $u^m = v^n$ for some $m, n \geq 1$, then u and v are both expressible as powers of some w in A^+.*

Proof If $|A| = 1$ or $m = 1$ or $n = 1$ the result is immediate. So suppose that $|A|, m, n \geq 2$. Then

$$|u^m| = \frac{1}{2}\big(|u^m| + |v^n|\big) \geq \frac{1}{2}\big(2|u| + 2|v|\big) = |u| + |v|. \tag{7.1.2}$$

Now

$$uv^n = u^{m+1}, \quad vu^m = v^{n+1} = u^m v,$$

and so uv^n and vu^m have a common initial segment of length at least $|u^m|$. From (7.1.2) we deduce that uv^n and vu^m have a common initial segment of length at least $|u| + |v|$, and from this it follows that $uv = vu$. The required result is now a consequence of Proposition 7.1.5. □

So far we have been listing properties of free semigroups and monoids. It is reasonable to ask whether any of these properties, either singly or in combination, actually *characterize* free semigroups. Our first result is fairly superficial, but provides a useful step on the way to less transparent characterizations.

Lemma 7.1.7 *A semigroup S is free if and only if every element of S has a unique expression as a product of elements of $S \setminus S^2$.*

Proof It is clear that if $S = A^+$, a free semigroup, then $S \setminus S^2 = A$, and every element has a unique expression as a product of elements of A. Conversely, suppose that S has the given property, and denote $S \setminus S^2$ by A. We show that S has the defining property (1.6.1) of A^+. Let T be a semigroup, and let $\alpha : A \to T$ be an arbitrary map. For each s in S consider the unique expression $s = a_1 a_2 \dots a_m$ of s as a product of elements in A, and define $\overline{\alpha} : S \to T$ by the rule that

$$s\overline{\alpha} = (a_1\alpha)(a_2\alpha)\dots(a_m\alpha).$$

If $t = a'_1 a'_2 \dots a'_n \in S$, then it is clear that the unique expression for st must be $a_1 a_2 \dots a_m a'_1 a'_2 \dots a'_n$, and from this remark it follows that $\overline{\alpha}$ is a morphism. It is clear that $\overline{\alpha}$ is the unique extension of α to S, and so $S \simeq A^+$, the free semigroup on A. □

Proposition 7.1.8 *Let S be an equidivisible monoid with proper length. Then S is a free monoid.*

Proof Let $T = S \setminus \{1\}$ and let $A = T \setminus T^2$. Notice that T is closed under multiplication, for

$$xy = 1 \Rightarrow |x| + |y| = 0 \Rightarrow |x| = |y| = 0 \Rightarrow x = y = 1.$$

Let $s \in T^2$. Then $s = uv$, with $u, v \in T$ and $0 < |u| < |s|$, $0 < |v| < |s|$. If $u \in T^2$ or $v \in T^2$, the process can be repeated, and since lengths cannot diminish indefinitely we conclude that

$$s = a_1 a_2 \dots a_m,$$

where $a_1, a_2, \ldots, a_m \in A$. To see that this expression is unique, suppose that

$$a_1 a_2 \ldots a_m = a'_1 a'_2 \ldots a'_n,$$

where $a_1, a_2, \ldots, a_m, a'_1, a'_2, \ldots, a'_n \in A$, and where $m, n \geq 1$. We proceed by induction on $m + n$. If $m + n = 2$ then $a_1 = a'_1$, and the result is clear. So suppose that $m + n \geq 3$. By equidivisibility we may suppose without loss of generality that there exists z in S such that

$$a_1 z = a'_1 \text{ and } a_2 \ldots a_m = z a'_2 \ldots a'_n.$$

Since $a'_1 \in T \setminus T^2$, we must have $z = 1$; hence $a_1 = a'_1$ and $a_2 \ldots a_m = a'_2 \ldots a'_n$. By the induction hypothesis we then deduce that $m = n$ and $a_2 = a'_2, \ldots, a_n = a'_n$. The result now follows by Lemma 7.1.7. □

Finally, let us define, for an element s of a monoid S, a *non-trivial left factor* of s to be an element $x \neq 1$ in S for which there exists y in S such that $xy = s$. Then we have

Proposition 7.1.9 *A free monoid $S = A^*$ has the properties:*

(1) *S has a trivial group of units;*
(2) *S is cancellative and equidivisible;*
(3) *every s in S has only finitely many non-trivial left factors.*

Conversely, every monoid S satisfying (1), (2) *and* (3) *is a free monoid.*

Proof The direct half is clear. To prove the converse half, we begin by showing that $T = S \setminus \{1\}$ is a subsemigroup of S, which amounts to showing that

$$(\forall x, y \in S)\; xy = 1 \;\Rightarrow\; x = y = 1.$$

Accordingly, suppose that $xy = 1$. Then $(yx)^2 = y(xy)x = yx$, and so $yx = e$, an idempotent. From $ee = e1$ we then deduce by cancellativity that $e = 1$. Thus $xy = yx = 1$, and so $x, y \in H_1$, the group of units of S. By (1) it follows that $x = y = 1$.

Next, we show that $T \setminus T^2 \neq \emptyset$. For suppose, by way of contradiction, that $T = T^2$. Then $T^n = T$ for every n, and so every t in T has arbitrarily long factorizations

$$t = u_1 u_2 \ldots u_n \quad (u_1, u_2, \ldots, u_n \in T). \tag{7.1.3}$$

Now, by (3), t has only $l(t)$ non-trivial left factors, where $l(t)$ is finite. If in (7.1.3) we choose $n > l(t)$ then we obtain left factors

$$u_1, u_1 u_2, \ldots, u_1 u_2 \ldots u_n,$$

which are all non-trivial, since T is a subsemigroup. Moreover, they are all distinct, for if $u_1 u_2 \ldots u_k = u_1 u_2 \ldots u_{k+r}$, with $r \geq 1$, then by cancellation we deduce that $u_{k+1} \ldots u_{k+r} = 1$, which is not possible. From the contradiction we have now obtained we conclude that $A = T \setminus T^2 \neq \emptyset$.

A very similar argument then shows that $\bigcap\{T^n : n \geq 1\} = \emptyset$, for otherwise there would exist t in T with arbitrarily long factorizations. Hence

$$T \supset T^2 \supset T^3 \supset \cdots,$$

and every element t of T is in $T^m \setminus T^{m+1}$ for some uniquely determined m. Thus t has an expression $t = a_1a_2 \dots a_m$ as a product of elements of A. By the argument used in the previous proof we may conclude from equidivisibility that the expression is unique, and the result now follows from Lemma 7.1.7. □

7.2 CODES

One of the cornerstones of infinite group theory is the Nielsen–Schreier Theorem, which states that every non-trivial subgroup of a free group is free. (See, for example, M. Hall (1959).) It is not at all hard to see that this is not true in general for semigroups or monoids. Even in the free monoid a^* generated by a single element a, the submonoid

$$S = a^* \setminus \{a\} = \{1, a^2, a^3, a^4, \dots\}$$

is not free. It is generated by its base $\{a^2, a^3\}$, but is certainly not freely generated by those elements, since it is commutative. Relative to the generators $x = a^2$ and $y = a^3$, it has presentation

$$\langle x, y \mid xy = yx,\ x^3 = y^2\rangle.$$

A somewhat more substantial example is provided by the subsemigroup $S = \langle ab, ba, aba, bab\rangle$ of the free semigroup $\{a, b\}^+$. Here again S is generated by its base $\{ab, ba, aba, bab\}$, but among the generators

$$x_1 = ab,\ x_2 = ba,\ x_3 = aba,\ x_4 = bab$$

there are the obvious relations

$$x_1x_3 = x_3x_2, \quad x_2x_4 = x_4x_1.$$

We have the following necessary and sufficient condition for a subsemigroup to be free:

Proposition 7.2.1 *Let A^* be a free monoid and let U be a subsemigroup or submonoid of A^*. Then U is free if and only if*

$$(\forall w \in A^+)\ [wU \cap U \neq \emptyset \ \textit{and}\ Uw \cap U \neq \emptyset] \Rightarrow w \in U. \qquad (7.2.1)$$

Proof Suppose first that U is free, and let w in A^+ be such that $wu_1 = u_2$, $u_3w = u_4$, where $u_1, u_2, u_3, u_4 \in U$. Then

$$u_3u_2 = u_3wu_1 = u_4u_1,$$

and so, by equidivisibility in U, either there exists v in U such that $u_3v = u_4$ (and $u_2 = vu_1$) or there exists z in U such that $u_4z = u_3$ and $zu_2 = u_1$. In the former case we deduce by cancellativity in A^+ that $w = v \in U$. In the latter case it follows that $wzu_2 = u_2$, and this cannot occur, since $|w| \geq 1$.

Conversely, suppose that U has the property (7.2.1). Certainly every element u of U is expressible as a product $u = c_1c_2\dots c_m$ of elements $c_1, c_2, \dots, c_m$ in its base $C = V \setminus V^2$, where $V = U \setminus \{1\}$. It remains to show that the expression is unique. So suppose that

$$c_1c_2\dots c_m = c_1'c_2'\dots c_n',$$

with $c_1, c_2, \dots, c_m, c_1', c_2', \dots c_n' \in C$, and that we have chosen these elements so that $m + n$ is as small as possible. Then, by equidivisibility in A^*, we may suppose, without loss of generality, that there exists z in A^* such that

$$c_1' = c_1z, \quad zc_2'\dots c_n' = c_2\dots c_m.$$

If $z \neq 1$ then we conclude from (7.2.1) that $z \in V$, and hence that $c_1' \in V^2$, contrary to assumption. Hence $z = 1$, and so $c_1 = c_1'$. It then follows by cancellation that $c_2\dots c_m = c_2'\dots c_n'$, and hence, by the minimality of $m+n$, that $m = n$ and $c_i = c_i'$ for $i = 2, \dots, m$. Thus U is freely generated by C. □

It follows in particular that if U is a left or right unitary subsemigroup or submonoid of A^* then U is free.

By a *code* (or *variable length code*) in the alphabet A we mean a subset C of A^+ with the property that C is a set of free generators for $\langle C\rangle$. The idea is simple enough. We think of the elements $c_1, c_2, \dots$ of C as being encoded by their expressions as words in the alphabet A. The C is a code if and only if no word in A^* can be decoded in two different ways. In our second example above, we would not know whether to decode *ababab* as x_1^3 or as x_3x_4.

In our definition of a code we have not insisted that C be finite, but we shall mostly be interested in the finite case.

The celebrated Morse code, devised last century for use in the electric telegraph, based on dots and dashes

A	$\cdot -$	D	$- \cdot \cdot$	
B	$- \cdot \cdot \cdot$	E	$\cdot$	etc.,
C	$- \cdot - \cdot$	F	$\cdot \cdot - \cdot$	

is not, as it stands, a code: the sequence $\cdot\ \cdot - \cdot$ could, for example, be read as F or as EAE. But if an extra symbol $|$ is included as a 'space' symbol, and if we terminate each of the Morse letter codings with $|$, then we do obtain a code, and a message such as

$$\cdot\cdot\cdot\cdot|\cdot|\cdot-\cdot\cdot|\cdot--\cdot||--|\cdot|$$

can be uniquely decoded as 'HELP ME'.

We say that a non-empty subset C of A^+ has the *prefix property* if $CA^+ \cap C = \emptyset$. Then we have

Proposition 7.2.2 *Let U be a submonoid of a free monoid A^* and let $C = V \setminus V^2$ (where $V = U \setminus \{1\}$) be its base. Then C has the prefix property if and only if U is a left unitary submonoid of A^*.*

Proof Suppose first that C has the prefix property, and suppose that $u, uw \in U$, where $w \neq 1$. That is,

$$c_1c_2 \dots c_m w = c_1'c_2' \dots c_n',$$

where $c_1, c_2, \dots, c_m, c_1', c_2', \dots, c_n' \in C$. By equidivisibility and the prefix property we deduce that $c_1 = c_1'$, and then by cancellativity and by repeating the argument we deduce that either

$$c_{n+1} \dots c_m w = 1 \quad \text{or} \quad w = c_{m+1}' \dots c_n'.$$

The former case cannot arise, since $w \neq 1$, and so $w = c_{m+1}' \dots c_n' \in U$.

Conversely, suppose that U is left unitary, and suppose that $cw = c'$ for some c, c' in C and some w in A^+. Then $w \in U$ by the left unitary property, and so $w = c_1c_2 \dots c_m$ for some $m \geq 1$ and some $c_1, c_2, \dots, c_m$ in C. It follows that $c' = cc_1c_2 \dots c_m \in V^{m+1} \subseteq V^2$, which is a contradiction. Thus C has the prefix property. □

It is an immediate consequence of the result, and of the remark following the proof of Proposition 7.2.1, that a subset C with the prefix property is a code. We call it a *prefix code*. The modified Morse code described above is a prefix code, since any element of CA^+ must contain the symbol | in a non-terminal position, and so cannot lie in C. The advantage of a prefix code is that there are no 'false starts' in the decoding process. By contrast, with a code such as $\{c_1, c_2\}$, where $c_1 = ab$ and $c_2 = aba$, which is clearly *not* a prefix code, we do (for example) have a unique decoding of

$$abababababa$$

as $c_1^3c_2$, but we might, in reading from the left, have tried $c_2(?)$, or $c_1^2c_2(?)$, before hitting on the correct answer.

It is evident that *suffix codes* C, with the property that $C \cap A^+C = \emptyset$, have the same advantage as prefix codes provided we are willing to read from the right. A code, such as $\{ab, ba\}$, which is both a prefix and a suffix code, is called a *biprefix code.*

While it is easy to determine whether or not a finite set C has the prefix property, it is not altogether clear how to discover whether or not a more general finite set C is a code. Various algorithms are available. We consider only one, usually called the *Sardinas–Patterson algorithm.*

In describing this algorithm, it is convenient first to introduce a piece of notation: if P, Q are subsets of a semigroup S, then

$$P^{-1}Q = \{s \in S : Ps \cap Q \neq \emptyset\}.$$

As usual, if P or Q or both are singletons, we simplify the notation, writing $p^{-1}Q$, $P^{-1}q$, $p^{-1}q$.

Let C be a subset of a free monoid A^*. Define subsets $D_0, D_1, \ldots$ inductively by

$$D_0 = C, \quad D_1 = C^{-1}C \setminus \{1\}, \quad D_i = C^{-1}D_{i-1} \cup D_{i-1}^{-1}C \quad (i \geq 2).$$

Thus, for example, if $C = \{ab, ba, aba, bab\}$, then

$$\begin{aligned} D_1 &= \{w \in A^+ : Cw \cap C \neq \emptyset\} = \{a, b\}; \\ D_2 &= \{w \in A^* : Cw \cap D_1 \neq \emptyset\} \cup \{w \in A^* : D_1 w \cap C \neq \emptyset\} \\ &= \emptyset \cup \{a, b, ab, ba\} = \{a, b, ab, ba\}; \\ D_3 &= \{w \in A^* : Cw \cap D_2 \neq \emptyset\} \cup \{w \in A^* : D_2 w \cap C \neq \emptyset\} \\ &= \{1\} \cup \{1, a, b, ab, ba\} = \{1, a, b, ab, ba\}; \\ D_4 &= \{w \in A^* : Cw \cap D_3 \neq \emptyset\} \cup \{w \in A^* : D_3 w \cap C \neq \emptyset\} \\ &= \{1\} \cup \{1, a, b, ab, ba, aba, bab\} = \{1, a, b, ab, ba, aba, bab\}; \\ D_5 &= \{w \in A^* : Cw \cap D_4 \neq \emptyset\} \cup \{w \in A^* : D_4 w \cap C \neq \emptyset\} \\ &= \{1, a, b\} \cup \{1, a, b, ab, ba, aba, bab\} = \{1, a, b, ab, ba, aba, bab\} = D_4, \end{aligned}$$

and so $D_n = D_4$ for all $n \geq 4$.

The fact that $1 \in D_3$ is crucial here. The set D_1 gives the 'remnants' when we attempt a double factorization within C itself:

$$aba = (ab)a, \quad bab = (ba)b.$$

Then D_2 describes how we must multiply these remnants (a and b) to obtain elements of C:

$$\begin{aligned} (aba)b &= (ab)(ab), \quad & (aba)(ba) &= (ab)(aba), \\ (bab)a &= (ba)(ba), \quad & (bab)(ab) &= (ba)(bab). \end{aligned}$$

The second and fourth of these equations give genuine double factorizations within $\langle C \rangle$, and this is reflected in the fact that $1 \in D_3$.

The fact that, in our example, the sets $D_0, D_1, D_2, \ldots$ are not all distinct is not an accident. If C is a finite set, and if $\max\{|c| : c \in C\} = m$, then all words in the sets $D_0, D_1, D_2, \ldots$ are of length at most m, and so there are at most $2^{|A|^m}$ distinct sets D_n. Thus the theorem we are about to state does give a genuine algorithm in the case where C is a finite subset. If C is a prefix code, the algorithm determines the fact very quickly, since in this case $D_1 = \emptyset$.

Theorem 7.2.3 *Let C be a non-empty subset of a free monoid A^*. Then C is a code if and only if $1 \notin D_i$ for all $i \geq 0$.*

Proof The key to the proof is the following lemma:

Lemma 7.2.4 *Let C be a non-empty subset of A^* and let $n \geq 1$ be an integer. Then the following statements are equivalent:*

(1) $1 \in D_n$;

(2) *for all k in $\{1, 2, \ldots, n\}$ there exist $i, j \geq 0$ such that $i + j + k = n$, and an element u of D_k such that $uC^i \cap C^j \neq \emptyset$;*

(3) *there exist k in $\{1, 2, \ldots, n\}$ and $i, j \geq 0$ such that $i + j + k = n$, and an element u of D_k such that $uC^i \cap C^j \neq \emptyset$.*

Proof (1) $\Rightarrow$ (2). Suppose that $1 \in D_n$. We prove (2) by induction on $n - k$, the result being clear if $n - k = 0$: simply take $i = j = 0$ and $u = 1$. Suppose therefore that $n - k \geq 1$, and suppose inductively that there exist u in D_{k+1} and integers $i, j \geq 0$ such that $i+j+k+1 = n$ and $uC^i \cap C^j \neq \emptyset$. Specifically, suppose that $v \in C^i$, $w \in C^j$ and $uv = w$. Now

$$u \in D_{k+1} = C^{-1}D_k \cup D_k^{-1}C,$$

and so either there exists c in C such that $cu \in D_k$ or there exists d in D_k such that $du \in C$. In the former case,

$$cuv = cw \in cuC^i \cap C^{j+1};$$

in the latter case,

$$duv = dw \in dC^j \cap C^{i+1};$$

in either case we have found an element z (equal to cu or to d) in D_k and integers $i', j' \geq 0$ such that $i' + j' + k = n$ and $zC^{i'} \cap C^{j'} \neq \emptyset$.

Since (2) $\Rightarrow$ (3) is immediate, we now turn our attention to (3) $\Rightarrow$ (1). Suppose, therefore, that for some k in $\{1, 2, \ldots, n\}$ there exist u in D_k and integers $i, j \geq 0$ such that $i + j + k = n$ and $uC^i \cap C^j \neq \emptyset$. That is to say,

$$uc_1c_2 \ldots c_i = c_1'c_2' \ldots c_j'$$

for some $c_1, c_2, \ldots, c_i, c_1', c_2', \ldots, c_j'$ in C. By equidivisibility in A^*, either: (i) $u = c_1'v$ for some v in A^*; or (ii) $c_1' = uv$ for some v in A^+. In case (i), $v \in C^{-1}D_k \subseteq D_{k+1}$, and

$$vc_1c_2 \ldots c_i = c_2' \ldots c_j';$$

hence $vC^i \cap C^{j-1} \neq \emptyset$. In case (ii), $v \in D_k^{-1}C \subseteq D_{k+1}$, and

$$c_1c_2 \ldots c_i = vc_2' \ldots c_j';$$

hence $vC^{j-1} \cap C_i \neq \emptyset$. Notice that $i + (j - 1) + (k + 1) = n$.

We can repeat this argument as often as necessary, and eventually obtain an element u in D_n and integers $i, j \geq 0$ such that $i + j + n = n$ and $uC^i \cap C^j \neq \emptyset$. Clearly $i = j = 0$ and $u = 1$. Hence $1 \in D_n$, as required.

□

Returning now to the proof of Theorem 7.2.3, suppose first that $1 \in D_n$. Then, by the lemma, there exist u in D_1 and integers $i, j \geq 0$ such that $i + j + 1 = n$ and $uC^i \cap C^j \neq \emptyset$. That is,

$$uc_1c_2 \ldots c_i = c_1'c_2' \ldots c_j'$$

for some $c_1, c_2, \ldots, c_i, c'_1, c'_2, \ldots, c'_j$ in C. Now $u \in D_1 = C^{-1}C \setminus \{1\}$, and so there exist c, c' in C such that $c \neq c'$ and $cu = c'$. It follows that

$$c'c_1c_2 \ldots c_i = cc'_1c'_2 \ldots c_j,$$

and so C is not a code.

Conversely, suppose that C is not a code, so that we have distinct factorizations within $\langle C \rangle$:

$$c_1c_2 \ldots c_i = c'_1c'_2 \ldots c'_j.$$

We may assume that $c_1 \neq c'_1$ and indeed that $c'_1 = c_1u$ for some $u \neq 1$ in A^*. But then $u \in C^{-1}C \subseteq D_1$, and

$$uc_2 \ldots c_i = c'_2 \ldots c'_j$$

gives $uC^{i-1} \cap C^{j-1} \neq \emptyset$. Hence, by the lemma, $1 \in D_{i+j-1}$. □

7.3 EXERCISES

1. Show that every completely simple semigroup is equidivisible.
2. Let a^+ be the free monogenic semigroup. Let $m, n \geq 0$ be coprime integers, and let $U = \langle a^m, a^n \rangle$.
   (a) Show that $a^{mn-m-n} \notin U$.
   (b) Let $k > mn - m - n$, let s, t in $\mathbf{Z}$ be such that $sm + tn = 1$, and suppose, without loss of generality, that $s \geq 0$. Observing that
   $$(ks - ln)m + (kt + lm)n = k$$
   for every l, and choosing l so that $0 \leq ks - ln \leq n - 1$, show that $kt + lm \geq 0$.
   (c) Deduce that U contains the ideal $\{a^k : k \geq (m-1)(n-1)\}$.
3. Show, by induction on $|w|$, that if u, v, w in the free semigroup A^+ are such that $uw = wv$ then there exist z, t in A^* and an integer $k \geq 0$ such that
   $$u = zt, \quad v = tz, \quad w = (zt)^k z.$$
4. Let S be an equidivisible semigroup. Say that two elements u and v of S are *conjugate* (and write $u \sim v$) if there exist x, y in S^1 such that $u = xy$ and $v = yx$.
   (a) Show that $\sim$ is an equivalence relation.
   (b) Show that, if S is a group, conjugacy has its usual meaning.
   (c) Show that, in the completely simple semigroup $\mathcal{M}[G; I, \Lambda; P]$, $(i, a, \lambda) \sim (j, b, \mu)$ if and only if the elements $ap_{\lambda i}$ and $bp_{\mu j}$ are conjugate in G.
   (d) Show that, in the free monoid A^*, $u \sim v$ if and only if there exists w in A^* such that $uw = wv$.

5. Let $u, v \in A^+$. Show that $\{u, v\}$ is a code if and only if $uv \neq vu$. [Hint: to prove $\Leftarrow$, let u, v (with $|u| \leq |v|$) be such that $uv \neq vu$ and $\{u, v\}$ is *not* a code, and suppose that u and v are chosen so that $|u| + |v|$ is as small as possible, and let w be a word of minimal length having two different factorizations in $\langle u, v \rangle$.]

6. An element u of A^+ is called *primitive* if, for all v in A^+,

$$u = v^n \Rightarrow n = 1 \text{ and } v = u.$$

   (a) Show that every element of A^+ is uniquely expressible as a power of a primitive element.
   (b) Show that every conjugate of a primitive element is primitive.
   (c) Let $w = uv$ (where $u, v \in A^+$) be a primitive word. Show that $\{u, v\}$ is a code.

7. Show that the free monoid $\{a, b\}^*$ contains prefix codes of arbitrary finite or countably infinite cardinality.

8. Let C, D be prefix codes over A. Show that CD is a prefix code.

9. Determine whether or not the following subsets of $\{a, b\}^*$ are codes:
   (a) $\{a^2, ab, a^2b, ab^2, b^2\}$;
   (b) $\{ab^2, ab, a^2b, ab^2a, ba^2b\}$;
   (c) $\{ab, ab^2, ab^3, \ldots\}$;
   (d) $\{ab, a^2b^2, a^3b^3, \ldots\}$.

10. A prefix code C is called *maximal* if there is no prefix code properly containing C.
   (a) Show that the 'uniform' code $A^n = \{w \in A^* : |w| = n\}$ is a maximal prefix code.
   (b) Show that, if C is a maximal prefix code, then, for every w in A^*, either w has a left factor in C or some c in C has w as a left factor.

11. Show that a prefix code C over A is maximal if and only if

$$(\forall w \in A^*)\ wA^* \cap C^* \neq \emptyset.$$

12. Consider the sequence (u_n) of words in $\{a, b\}^*$ defined by

$$u_1 = b, \quad u_2 = a, \quad u_{n+1} = u_n u_{n-1} \quad (n \geq 2).$$

   (Thus the sequence $|u_n|$ is the *Fibonacci sequence* (f_n):

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13, \ldots)$$

   For $n \geq 3$, let v_n be the left factor of u_n of length $f_n - 2$.
   (a) Show that, for all $n \geq 4$,

$$v_{n+1} = u_n v_{n-1}.$$

(b) Show, by induction on n, that, for all $n \geq 5$,

$$v_{n+1} = u_{n-1}^2 v_{n-2}.$$

7.4 NOTES

A much fuller account of free semigroups and codes can be found in Lallement (1979). See also Lothaire (1983), Berstel and Perrin (1985), and Shyr (1991). Proposition 7.1.6 is from Lyndon and Schützenberger (1962), which is also a reference for the notions of conjugacy and primitive words in Exercises 3, 4 and 6. Proposition 7.1.8 is from Levi (1944), and Proposition 7.1.9 is from Dubreil-Jacotin (1947).

Proposition 7.2.1 seems to have appeared first in Schützenberger (1956); see also Shevrin (1960), Cohn (1962), and Blum (1965). The Sardinas–Patterson algorithm (Theorem 7.2.3) appears in Sardinas and Patterson (1953); see also Bandyopadhyay (1963), Levenshtein (1966), Riley (1967), and de Luca (1976). An alternative algorithm, with the advantage of giving a presentation of $\langle C \rangle$ as $\langle C \mid R \rangle$ when C is not a code, has been given by Spehner (1975), and is described by Lallement (1979).

Exercise 1 is from MacKnight and Storey (1969), Exercise 5 is from Shyr and Thierrin (1976), and Exercise 12 is from Lothaire (1983).

8
Semigroup amalgams

The use of module theory in the study of rings is long established, but only in relatively recent years has serious use been made in semigroup theory of the non-additive analogue of a module. Perhaps surprisingly for such a simple structure, what we shall call S-systems (where S is a monoid) have proved a useful tool in the study of monoids and semigroups. The particular use that we shall highlight in this chapter is the application to amalgamation theory and related topics. The technique is available only for monoids, but, as we shall see, it is usually possible to extend the results to semigroups in a fairly straightforward way.

A (*semigroup*) *amalgam* will be defined more carefully below, but may conveniently be thought of as an indexed family $\{S_i : i \in I\}$ of semigroups intersecting pairwise in a common subsemigroup U; thus $S_i \cap S_j = U$ if $i \neq j$. Effectively $A = \bigcup_{i \in I} S_i$ is a partial semigroup: that is to say, if x and y are members of the set then the product xy may or may not be defined; and if x, y, z are in A then $(xy)z = x(yz)$ whenever both products are meaningful. The central question concerning a semigroup amalgam is whether or not the partial semigroup A can be embedded in a semigroup, that is, whether there exists a semigroup T containing A in which the product of two elements is *always* defined and in which previously defined multiplications within A take place as before. It was shown by Schreier (1927) that a *group* amalgam is always embeddable in a group, but it has long been known that a simple answer of this sort will not suffice for monoids or semigroups. The situation is in fact much more similar to that obtaining in ring theory, where important work by Cohn (1959) established necessary conditions for an amalgam to be embeddable. Only recently has the homological apparatus used by Cohn been developed in the non-additive setting to the point where real comparisons could be made between the ring and semigroup cases.

8.1 SYSTEMS

Let S be a monoid with identity element 1 and let X be a non-empty set. We say that X is a *left S-system* if there is an action $(s,x) \mapsto sx$ from $S \times X$ into X with the properties

$$(st)x = s(tx) \quad (s,t \in S,\ x \in X),$$
$$1x = x \quad (x \in X).$$

Various alternative names have been used, such as *S-act*, *S-set* and *S-operand.* We shall stick to the term 'left S-system', feeling free to drop one or both of 'left' and 'S-' if the context permits.

Dually, a non-empty set X is a *right S-system* if there is an action $(x,s) \mapsto xs$ from $X \times S$ into X such that

$$x(st) = (xs)t \quad (s,t \in S,\ x \in X),$$
$$x1 = x \quad (x \in X).$$

Also, if S and T are (not necessarily different) monoids we say that X is an (S,T)*-bisystem* if it is a left S-system, a right T-system, and if

$$(sx)t = s(xt)$$

for all s in S, t in T and x in X.

These definitions are of course closely modelled on the definitions of left modules, right modules and bimodules. We shall often find it useful to express the statement that X is a left S-system by writing $X \in S\text{-}\textsc{Ens}$. The meanings to be attached to the statements $X \in \textsc{Ens}\text{-}S$ and $X \in S\text{-}\textsc{Ens}\text{-}T$ are clear by analogy. (The use of $\textsc{Ens}$, an abbreviation of the French *ensemble*, is fairly standard in this context.)

Remark If S is a commutative monoid then there is no distinction between a left and a right S-system. For if $X \in S\text{-}\textsc{Ens}$ we may define a right action $*$ of S on X by

$$x * s = sx \quad (x \in X,\ s \in S).$$

Then certainly $x * 1 = x$ for all x. Also, for all s, t in S,

$$x * (st) = x * (ts) = (ts)x = t(sx) = (x * s) * t.$$

Indeed we can regard X as an (S,S)-bisystem, since for all x in X and s, t in S,

$$(sx) * t = t(sx) = (ts)x = (st)x = s(tx) = s(x * t).$$

It is clear that any set X whatever can be regarded as a $(\{1\},\{1\})$-bisystem, where $\{1\}$ is the trivial monoid. It will therefore occasionally be convenient to state and prove results for (S,T)-bisystems, deducing results regarding one-sided systems by taking either S or T as the trivial monoid. At other times it will be sufficient to consider the case of a left S-system, since the analogous results for systems of other kinds will be obvious.

Systems are algebras of a very rudimentary kind, but much of the standard apparatus of abstract algebra applies to them. We shall give the merest sketch. A *subsystem* of a left S-system X is a subset Y of X with the property that (in an obvious notation) $SY \subseteq Y$. By a *morphism* (or S-*morphism* or S-*map*) from a left S-system X into a left S-system Y we mean a map $\phi : X \to Y$ with the property that

$$(sx)\phi = s(x\phi) \quad (s \in S,\ x \in X).$$

A *congruence* on a left S-system X is an equivalence on X with the property that, for all x, y in X and all s in S,

$$x \,\rho\, y \quad \Rightarrow \quad sx \,\rho\, sy.$$

The quotient X/ρ then inherits a left S-system structure by means of the definition

$$s(x\rho) = (sx)\rho,$$

and there is a morphism $\rho^\natural : S \to S/\rho$ defined by the rule that $x\rho^\natural = x\rho$ for every x in X.

Since the intersection of any non-empty collection of congruences is a congruence, the notion of the congruence $\mathbf{R}^\#$ generated by a given relation $\mathbf{R}$ is available here, just as in the theory of semigroup congruences (Section 1.5). If, by analogy with the definition in Section 1.5, we define $\mathbf{R}^c$ to be

$$\{(sa, sb) : s \in S,\ (a, b) \in \mathbf{R}\},$$

then it is not hard to prove that, in the notation of that section,

$$\mathbf{R}^\# = (\mathbf{R}^c)^e.$$

In more elementary terms, we have that $(x, y) \in \mathbf{R}^\#$ if and only if either $x = y$ or for some $n \geq 1$ there is a sequence

$$x = s_1a_1 \to s_1b_1 = s_2a_2 \to s_2b_2 = s_3a_3 \to \cdots \to s_nb_n = y,$$

in which each (a_i, b_i) belongs either to $\mathbf{R}$ or to $\mathbf{R}^{-1}$.

Example 8.1.1 Every monoid S is an (S, S)-bisystem, where the actions of S on S are defined by means of the multiplication.

Example 8.1.2 If U is a submonoid of S then S is a (U, U)-bisystem in the obvious way.

Example 8.1.3 If I is a left ideal of S then $I \in S$-ENS.

Example 8.1.4 If λ is a left congruence on S then there is a well-defined action of S on S/λ given by $s(x\lambda) = (sx)\lambda$. With this definition, $S/\lambda \in$ S-ENS.

Example 8.1.5 Let U be a submonoid of a monoid S. Then the Rees quotient S/U (consisting of U, considered as a single element, and of the elements of $S \setminus U$) inherits the structure of a monoid if U is an ideal. If U is not an ideal then this is not necessarily the case. What does always happen is that S/U has the structure of a left U-system: we define, for u in U and s in $S \setminus U$,

$$us = \begin{cases} us, \text{ the product in } S, \text{ if } us \notin U, \\ U \quad \text{otherwise}, \end{cases}$$
$$uU = U.$$

By making analogous definitions for a right action one can show that $S/U \in U\text{-}\textsc{ens}\text{-}U$.

More generally, for every U-system X, and for every U-subsystem Y of X, we can define X/Y as the quotient of X by the equivalence $(Y \times Y) \cup 1_X$ and define an inherited U-system structure in a natural way.

Example 8.1.6 Let $S = \mathcal{T}_X$, the full transformation semigroup on a non-empty set X. Then we may regard each α in $\mathcal{T}_X$ as acting on elements of X on the right. Denoting the identity element id_X of $\mathcal{T}_X$ by 1, we see that $x(\alpha\beta) = (x\alpha)\beta$ for every x in X and for every α and β in S, and that $x1 = x$ for every x in X. Thus $X \in \textsc{ens}\text{-}S$.

More generally, if T is any submonoid of $\mathcal{T}_X$ then in the same way X may be regarded as a right T-system. There is, however, a distinction between the statements 'X is a right T-system' and 'T is a monoid of mappings from X into X', for in the latter case we have the implication

$$[(\forall x \in X)\, x\alpha = x\beta] \quad \Rightarrow \quad \alpha = \beta,$$

and this may fail to be the case for a T-system in general. See Exercise 1 for further comment on the relation between the two ideas.

It is clear that the cartesian product $X \times Y$ of a left S-system X and a right T-system Y becomes an (S, T)-bisystem if we make the obvious definitions

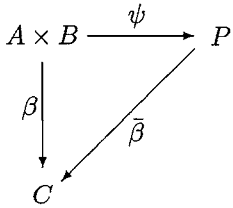
$$s(x, y) = (sx, y), \quad (x, y)t = (x, yt).$$

Now we introduce a construction that is crucially important in homological algebra (see Rotman (1979)) but which here appears in a non-additive version. Let $A \in T\text{-}\textsc{ens}\text{-}S$, $B \in S\text{-}\textsc{ens}\text{-}U$ and $C \in T\text{-}\textsc{ens}\text{-}U$. By the remark in the last paragraph we may give $A \times B$ the structure of a (T, U)-bisystem. A (T, U)-map $\beta : A \times B \to C$ will be called a *bimap* if, for all a in A, s in S and b in B,

$$(as, b)\beta = (a, sb)\beta. \tag{8.1.1}$$

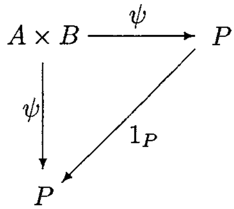
A pair (P, ψ) consisting of a (T, U)-bisystem P and a bimap $\psi : A \times B \to P$ will be called a *tensor product of A and B over S* if for every (T, U)-

bisystem C and every bimap $\beta : A \times B \to C$ there exists a unique (T, U)-map $\bar{\beta} : P \to C$ such that the diagram



(8.1.2)

commutes. This holds in particular in the case where $C = P$ and $\beta = \psi$, in which case the appropriate unique $\bar{\beta}$ is 1_P:

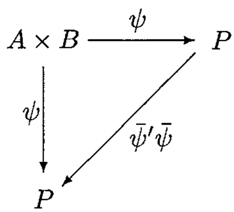


(8.1.3)

This trivial remark will be useful in proving the following lemma:

Lemma 8.1.7 *If a tensor product of A and B over S exists, then it is unique up to isomorphism.*

Proof Suppose that (P, ψ), (P', ψ') are tensor products of A and B. Then by putting $C = P'$ in the diagram (8.1.2) we find a unique $\bar{\psi}' : P \to P'$ such that $\psi\bar{\psi}' = \psi'$. Then by substituting P' for P in (8.1.2) and putting $C = P$ we find a unique $\bar{\psi} : P' \to P$ such that $\psi'\bar{\psi} = \psi$. Thus $\psi\bar{\psi}'\bar{\psi} = \psi$, and so the diagram



commutes. Hence, by the uniqueness property in the diagram (8.1.3), $\bar{\psi}'\bar{\psi} = \mathrm{id}_P$. By a similar argument, $\bar{\psi}\bar{\psi}' = \mathrm{id}_{P'}$, and so $P \simeq P'$ as required.

□

We must now assure ourselves that a tensor product exists. Let us define $A \otimes_S B$ to be $(A \times B)/\tau$, where τ is the equivalence on $A \times B$ generated by the relation

$$\mathbf{T} = \{\big((as, b), (a, sb)\big) : a \in A,\ b \in B,\ s \in S\}.$$

This is a non-additive modification of a classical construction in module theory. (See Rotman (1979).) We denote a typical element $(a, b)\tau$ of $A \otimes_S B$

by $a \otimes b$, and note that by the definition of τ we immediately have that $as \otimes b = a \otimes sb$ for all a in A, s in S and b in B. Then we have

Proposition 8.1.8 *Two elements $a \otimes b$ and $c \otimes d$ in $A \otimes_S B$ are equal if and only if either $(a, b) = (c, d)$ or there exist $a_1, \dots, a_{n-1}$ in A, $b_1, \dots, b_{n-1}$ in B, $s_1, \dots, s_n$, $t_1, \dots, t_{n-1}$ in S such that*

$$\begin{aligned} a &= a_1 s_1, & s_1 b &= t_1 b_1, \\ a_1 t_1 &= a_2 s_2, & s_2 b_1 &= t_2 b_2, \\ a_i t_i &= a_{i+1} s_{i+1}, & s_{i+1} b_i &= t_{i+1} b_{i+1} \ (i = 2, \dots, n-2), \\ a_{n-1} t_{n-1} &= c s_n, & s_n b_{n-1} &= d. \end{aligned} \tag{8.1.4}$$

Proof Suppose first that we have the given sequence of equations. Then

$$\begin{aligned} a \otimes b &= a_1 s_1 \otimes b = a_1 \otimes s_1 b = a_1 \otimes t_1 b_1 \\ &= a_1 t_1 \otimes b_1 = a_2 s_2 \otimes b_1 = a_2 \otimes s_2 b_1 = a_2 \otimes t_2 b_2 \\ &= \cdots \\ &= a_{n-1} t_{n-1} \otimes b_{n-1} = c s_n \otimes b_{n-1} = c \otimes s_n b_{n-1} = c \otimes d. \end{aligned}$$

Conversely, suppose that $a \otimes b = c \otimes d$. Then by Proposition 1.4.10 there is a sequence

$$(a, b) = (p_1, q_1) \to (p_2, q_2) \to \cdots \to (p_n, q_n) = (c, d)$$

in which for each $i \in \{1, \dots, n-1\}$ either $\big((p_i, q_i), (p_{i+1}, q_{i+1})\big) \in \mathbf{T}$ or $\big((p_{i+1}, q_{i+1}), (p_i, q_i)\big) \in \mathbf{T}$. At each transition, an element of S either moves right, as in $(as, b) \to (a, sb)$, or moves left, as in $(a, sb) \to (as, b)$. Two successive 'move right' transitions, such as

$$(as, b) \to (a, sb) = (a's', sb) \to (a', s'sb),$$

can be combined into a single transition

$$(as, b) = (a's's, b) \to (a', s'sb).$$

The same applies to two successive 'move left' transitions, and so we may assume that the 'move right' and 'move left' transitions alternate. We may also suppose that the sequence begins and ends with a 'move right' transition, since otherwise we can tack on 'move right' transitions in which the moving element of S is the identity element. □

Recall now that we have assumed that $A \in T\text{-}\textsc{Ens}\text{-}S$ and $B \in S\text{-}\textsc{Ens}\text{-}U$. From the equations (8.1.4) it is clear that from $a \otimes b = c \otimes d$ we can deduce that $ta \otimes b = tc \otimes d$ for all t in T, and that $a \otimes bu = c \otimes du$ for all u in U. Hence we have

Corollary 8.1.9 *The equivalence τ is a (T, U)-congruence on $A \times B$.* □

Consequently $A \otimes_S B$ inherits the (T, U)-bisytem structure of $A \times B$, and

$$t(a \otimes b) = (ta) \otimes b, \quad (a \otimes b)u = a \otimes (bu). \tag{8.1.5}$$

Next, we have

Proposition 8.1.10 *Let $A \in T$-ENS-S, $B \in S$-ENS-U. Then $(A \otimes_S B, \tau^\natural)$ is a tensor product of A and B over S.*

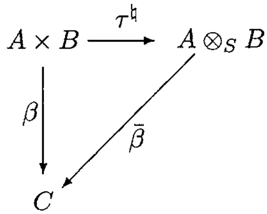
Proof It is clear that $\tau^\natural$ is a bimap. Now let $C \in T$-ENS-U, and let $\beta : A \times B \to C$ be a bimap. Define $\bar{\beta} : A \otimes_S B \to C$ by

$$(a \otimes b)\bar{\beta} = (a,b)\beta \quad (a \in A,\ b \in B).$$

To prove that this is a well-defined map, suppose that $a \otimes b = c \otimes d$. Then we have a system of equations as in (8.1.4), and it follows that

$$\begin{aligned}(a,b)\beta &= (a_1 s_1, b)\beta = (a_1, s_1 b)\beta = (a_1, t_1 b_1)\beta \\ &= (a_1 t_1, b_1)\beta = (a_2 s_2, b_1)\beta = (a_2, s_2 b_1)\beta = (a_2, t_2 b_2)\beta = (a_2 t_2, b_2)\beta \\ &= \cdots \\ &= (a_{n-1} t_{n-1}, b_{n-1})\beta = (c s_n, b_{n-1})\beta = (c, s_n b_{n-1})\beta = (c,d)\beta;\end{aligned}$$

hence $(a \otimes b)\bar{\beta} = (c \otimes d)\bar{\beta}$ as required. It is now routine to establish that $\bar{\beta}$ is a (T,U)-map, that the diagram



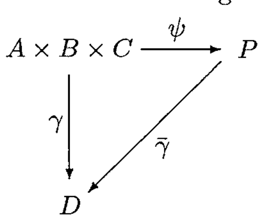
commutes and that $\bar{\beta}$ is unique with respect to these properties. □

We shall often simply refer to $A \otimes_S B$ as 'the tensor product of A and B over S', dropping explicit reference to the associated bimap $\tau^\natural$.

We can generalize the notion of tensor product to the case of three (or indeed more) bisystems. Let $A \in S$-ENS-T, $B \in T$-ENS-U, $C \in U$-ENS-V and $D \in S$-ENS-V. An (S,V)-map $\gamma : A \times B \times C \to D$ is called a *trimap* if, for all a in A, b in B, c in C, t in T and u in U,

$$(at, b, c)\gamma = (a, tb, c)\gamma, \quad (a, bu, c)\gamma = (a, b, uc)\gamma.$$

Then a pair (P, ψ), where $P \in S$-ENS-V and $\psi : A \times B \times C$ is a trimap, is said to be a tensor product over T and U of A, B and C if for every (S,V)-bisystem D and every trimap $\gamma : A \times B \times C \to D$ there is a unique (S,V)-map $\bar{\gamma} : P \to D$ such that the diagram



commutes. An argument closely similar to that used in the proof of Lemma 8.1.7 establishes that the tensor product is unique up to isomorphism. One candidate for such a tensor product is $\big((A \otimes_T B) \otimes_U C, \psi\big)$, where

$$(a, b, c)\psi = (a \otimes b) \otimes c.$$

Using (8.1.5) one may easily verify that ψ is a trimap. Given a trimap $\gamma : A \times B \times C \to D$, we can attempt to define $\bar{\gamma} : (A \otimes_T B) \otimes_U C \to D$ by

$$\big((a \otimes b) \otimes c\big)\bar{\gamma} = (a, b, c)\gamma. \tag{8.1.6}$$

If $(a \otimes b) \otimes c = (a' \otimes b') \otimes c'$ then there is a sequence of transitions from (a, b, c) to (a', b', c') based on moves of the types

$$(at, b, c) \rightleftharpoons (a, tb, c), \quad (a, bu, c) \rightleftharpoons (a, b, uc).$$

The trimap property of γ means that each such move can be duplicated on the right-hand side of (8.1.6); hence $(a, b, c)\gamma = (a', b', c')\gamma$, and the map given by (8.1.6) is therefore well-defined. It is now routine to show that $\bar{\gamma}$ is an (S, V)-map, that $\psi\bar{\gamma} = \gamma$, and that $\bar{\gamma}$ is unique relative to these properties.

A very similar argument shows that $A \otimes_T (B \otimes_U C)$, together with the obvious trimap $(a, b, c) \mapsto a \otimes (b \otimes c)$, is also a tensor product of A, B and C. Hence we have

Proposition 8.1.11 *Let* $A \in S$-ENS-T, $B \in T$-ENS-U, $C \in U$-ENS-V. *Then* $(A \otimes_T B) \otimes_U C \simeq A \otimes_T (B \otimes_U C)$. □

Since the isomorphism sends $(a \otimes b) \otimes c$ to $a \otimes (b \otimes c)$, we may from now on safely use notations such as $A \otimes_T B \otimes_U C$ and $a \otimes b \otimes c$.

8.2 FREE PRODUCTS

Given an indexed family $\{S_i : i \in I\}$ of pairwise disjoint semigroups, we show how to form a semigroup $F = \Pi^*\{S_i : i \in I\}$, the *free product* of the family $\{S_i : i \in I\}$. It is helpful first to introduce the following notational device: if $a \in \bigcup\{S_i : i \in I\}$ then there is a unique k in I such that $a \in S_k$; we refer to k as the *index* of a and write $k = \sigma(a)$.

Now let F consist of all finite 'strings'

$$(a_1, a_2, \ldots, a_m),$$

where m (≥ 1) is an integer, where $a_r \in \bigcup\{S_i : i \in I\}$ for $r = 1, 2, \ldots, m$ and where $\sigma(a_r) \neq \sigma(a_{r+1})$ for $r = 1, \ldots, m-1$. Let $\mathbf{a} = (a_1, a_2, \ldots, a_m)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ be elements of F. Then we define the product of $\mathbf{a}$ and $\mathbf{b}$ in F by

$$\mathbf{ab} = \begin{cases} (a_1, a_2, \dots, a_m, b_1, b_2, \dots, b_n) \text{ if } \sigma(a_m) \neq \sigma(b_1) \\ (a_1, a_2, \dots, a_m b_1, b_2, \dots, b_n) \text{ if } \sigma(a_m) = \sigma(b_1). \end{cases} \quad (8.2.1)$$

It is a routine matter to check that if

$$\mathbf{a} = (a_1, a_2, \dots, a_m), \quad \mathbf{b} = (b_1, b_2, \dots, b_n), \quad \mathbf{c} = (c_1, c_2, \dots, c_p)$$

are in F, then $(\mathbf{ab})\mathbf{c} = \mathbf{a}(\mathbf{bc})$. Four distinct cases arise, depending on whether $\sigma(a_m)$ equals or does not equal $\sigma(b_1)$ and whether $\sigma(b_n)$ equals or does not equal $\sigma(c_1)$, but no difficulties are encountered. Indeed it is worth noting that the verification of associativity here is considerably easier than in the corresponding group-theoretic construction—see, for example, Kurosh (1956). It must be emphasized that if the semigroups S_i are all groups then the semigroup $F = \Pi^*\{S_i : i \in I\}$ we are describing here is *not* the free product of the groups S_i as normally understood in group theory. Indeed, it is not even a group. This is because in the *group* free product all the identity elements of the individual groups are identified, whereas in the *semigroup* free product they remain distinct. For the same reason, F is not a monoid even if all the semigroups S_i are monoids. We shall see shortly how to derive the *monoid* free product from F.

We have established that F is a semigroup. Among the elements of F are strings (s_i) of length one, where $s_i \in S_i$. In fact F is generated by strings of length one, since

$$(a_1, a_2, \dots, a_m) = (a_1)(a_2)\dots(a_m)$$

for every $(a_1, a_2, \dots, a_m)$ in F. It is customary to leave out the brackets in writing down a string of length one, and hence to regard F as consisting of finite non-empty *words* in the alphabet $\bigcup\{S_i : i \in I\}$. Multiplication of words $a_1 a_2 \dots a_m$ and $b_1 b_2 \dots b_n$ is then simply a matter of juxtaposition if $\sigma(a_m) \neq \sigma(b_1)$:

$$(a_1 a_2 \dots a_m)(b_1 b_2 \dots b_n) = a_1 a_2 \dots a_m b_1 b_2 \dots b_n,$$

while if $\sigma(a_m) = \sigma(b_1) = i$ (say), then

$$(a_1 a_2 \dots a_m)(b_1 b_2 \dots b_n) = a_1 a_2 \dots a_{m-1} c b_2 \dots b_n,$$

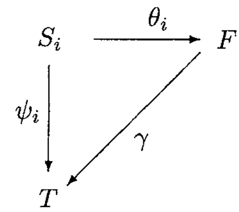
where c is the product in S_i of a_m and b_1.

If the context ensures that no confusion will arise, we shall abbreviate $\Pi^*\{S_i : i \in I\}$ to $\Pi^* S_i$. Also, if I is finite—say $I = \{1, 2, \dots, n\}$—we shall often prefer to write $S_1 * S_2 * \dots * S_n$ for the free product $\Pi^*\{S_i : i \in I\}$. The crucial property of free products is contained in the following result:

Proposition 8.2.1 *Let $F = \Pi^*\{S_i : i \in I\}$ be the free product of a family $\{S_i : i \in I\}$ of disjoint semigroups. Then for each i in I there exists a monomorphism $\theta_i : S_i \to F$ given by*

$$s_i \theta_i = (s_i) \quad (s_i \in S_i),$$

associating the element s_i of S_i with the one-letter word (s_i). If T is a semigroup for which there is a morphism $\psi_i : S_i \to T$ for each i then there is a unique morphism $\gamma : F \to T$ with the property that the diagram



(8.2.2)

commutes for every i in I.

Proof We shall frequently wish to drop the distinction between s_i and $s_i\theta_i = (s_i)$, and to regard S_i as a subsemigroup of F.

If T and the morphisms ψ_i $(i \in I)$ are given, we define $\gamma : F \to T$ by

$$(a_1a_2\dots a_m)\gamma = (a_1\psi_{\sigma(a_1)})(a_2\psi_{\sigma(a_2)})\dots(a_m\psi_{\sigma(a_m)}).$$

Since the expression on the right is a product of elements of T, we conclude that γ maps F into T. To show that γ is a morphism, let $a = a_1a_2\dots a_m$ and $b = b_1b_2\dots b_n$ be elements of F. Then if $\sigma(a_m) \neq \sigma(b_1)$,

$$\begin{aligned}(ab)\gamma &= (a_1a_2\dots a_mb_1b_2\dots b_n)\gamma\\ &= (a_1\psi_{\sigma(a_1)})\dots(a_m\psi_{\sigma(a_m)})(b_1\psi_{\sigma(b_1)})\dots(b_n\psi_{\sigma(b_n)})\\ &= [(a_1\psi_{\sigma(a_1)})\dots(a_m\psi_{\sigma(a_m)})][(b_1\psi_{\sigma(b_1)})\dots(b_n\psi_{\sigma(b_n)})]\\ &= (a\gamma)(b\gamma),\end{aligned}$$

while if $\sigma(a_m) = \sigma(b_1) = i$ (say) and $a_mb_1 = c$ in S_i, then

$$\begin{aligned}&(ab)\gamma = (a_1a_2\dots a_{m-1}cb_2\dots b_n)\gamma\\ &= (a_1\psi_{\sigma(a_1)})\dots(a_{m-1}\psi_{\sigma(a_{m-1})})(c\psi_i)(b_2\psi_{\sigma(b_2)})\dots(b_n\psi_{\sigma(b_n)})\\ &= (a_1\psi_{\sigma(a_1)})\dots(a_{m-1}\psi_{\sigma(a_{m-1})})(a_m\psi_i)(b_1\psi_i)(b_2\psi_{\sigma(b_2)})\dots(b_n\psi_{\sigma(b_n)})\\ &\qquad \text{(since } \psi_i \text{ is a morphism)}\\ &= [(a_1\psi_{\sigma(a_1)})\dots(a_m\psi_{\sigma(a_m)})][(b_1\psi_{\sigma(b_1)})\dots(b_n\psi_{\sigma(b_m)})]\\ &= (a\gamma)(b\gamma).\end{aligned}$$

To see that the diagram (8.2.2) commutes for every i, observe that the definitions of θ_i and γ imply that

$$s_i(\theta_i\gamma) = (s_i\theta_i)\gamma = (s_i)\gamma = s_i\psi_i$$

for every s_i in S_i. Hence $\theta_i\gamma = \psi_i$, as required.

The uniqueness of γ follows from the fact that F is generated by words of length one. If γ is to make the diagram (8.2.2) commute then we *must* have $(s_i)\gamma = s_i\psi_i$ for every s_i in S_i and every i in I. Then, if γ is to be a morphism we *must* have

$$(a_1a_2\dots a_m)\gamma = \big((a_1)\gamma\big)\big((a_2)\gamma\big)\dots\big((a_m)\gamma\big)$$

$$= (a_1\psi_{\sigma(a_1)})(a_2\psi_{\sigma(a_2)})\dots(a_m\psi_{\sigma(a_m)})$$

for every $a_1a_2\dots a_m$ in F. That is, γ must be exactly as we defined it.

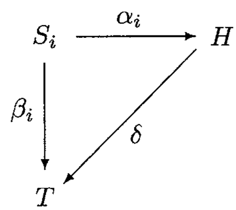
□

The property described in Proposition 8.2.1 does in fact *characterize* the free product. More precisely, we have

Proposition 8.2.2 *Let $\{S_i : i \in I\}$ be a family of semigroups, and let H be a semigroup such that*

(1) *there exists a monomorphism $\alpha_i : S_i \to H$ for each i in I;*

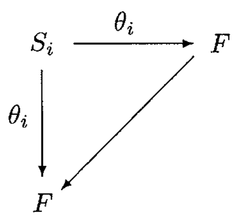
(2) *if T is a semigroup and if there exists a morphism $\beta_i : S_i \to T$ for every i in I, then there exists a unique morphism $\delta : H \to T$ such that the diagram*



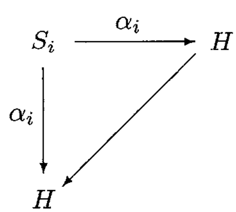
commutes for every i in I.

Then H is isomorphic to $F = \Pi^*\{S_i : i \in I\}$.

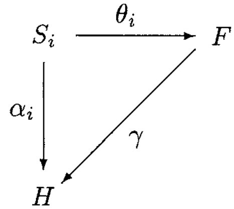
Proof From the property of F established in Proposition 8.2.1, applied to the case when $T = F$ and $\psi_i = \theta_i$ $(i \in I)$, there is a unique morphism $F \to F$ making the diagram



commute for every i in I. It is evident that the identity map 1_F has this property, and so by uniqueness it is the *only* morphism from F into F having the property. By the same token, the identity map 1_H is the only morphism from H into H with the property that the diagram

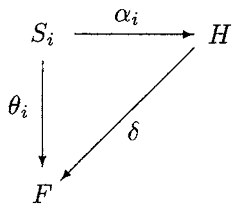


commutes for every i. Now we apply Proposition 8.2.1 with $T = H$ and $\psi_i = \alpha_i \quad (i \in I)$ and obtain a morphism $\gamma : F \to H$ such that the diagram



(8.2.3)

commutes. Then by the assumed property of H, with $T = F$ and $\beta_i = \theta_i \quad (i \in I)$ we obtain a morphism $\delta : H \to F$ such that the diagram

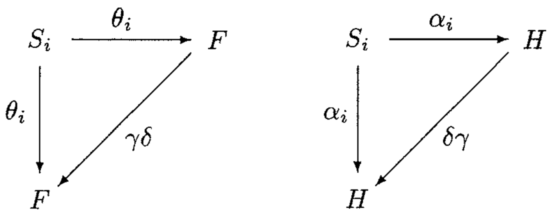


(8.2.4)

commutes. It follows that

$$\theta_i\gamma\delta = \alpha_i\delta = \theta_i, \quad \alpha_i\delta\gamma = \theta_i\gamma = \alpha_i.$$

That is, if we 'tack together' the diagams (8.2.3) and (8.2.4) in both of the possible ways we obtain commutative diagrams



for every i in I. From our earlier remarks about uniqueness we now immediately deduce that $\gamma\delta = 1_F$, $\delta\gamma = 1_H$. Thus γ and δ are mutually inverse isomorphisms, and $H \simeq F$ as required. □

The conclusions of Propositions 8.2.1 and 8.2.2 can be summarized by saying that $F = \Pi^* S_i$ is the unique *coproduct*, in the sense of category theory, of the objects S_i. (See, for example, Mitchell (1965).)

We turn now to the more complex situation that faces us when we consider a semigroup amalgam. The intuitive definition given in the introduction is not now precise enough, and so we begin by giving a more careful definition as follows. A *(semigroup) amalgam*

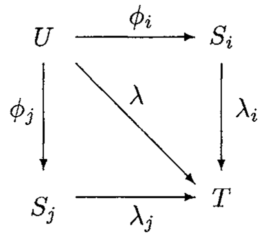
$$\mathcal{A} = [U; \{S_i : i \in I\}; \{\phi_i : i \in I\}]$$

consists of a semigroup U (called the *core* of the amalgam), a family $\{S_i : i \in I\}$ of semigroups disjoint from each other and from U, and a family of monomorphisms $\phi_i : U \to S_i \quad (i \in I)$. We shall simplify the notation to $\mathcal{A} = [U; S_i; \phi_i]$ when the context allows.

The amalgam $\mathcal{A} = [U; S_i; \phi_i]$ is said to be *embedded* in a semigroup T if there exist a monomorphism $\lambda : U \to T$ and, for each i in I, a monomorphism $\lambda_i : S_i \to T$ such that

(a) $\phi_i \lambda_i = \lambda$ for each i in I;

(b) $S_i\lambda_i \cap S_j\lambda_j = U\lambda$ for all i, j in I such that $i \neq j$.



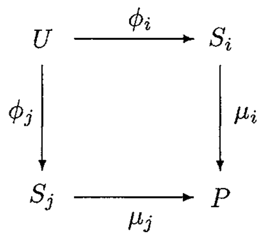
The earlier intuitive definition differs from this only in that all the monomorphisms are regarded as inclusion maps.

The *free product* $\Pi^*_U S_i$ of the amalgam $\mathcal{A}$ is defined as a quotient semigroup of the ordinary free product $\Pi^* S_i$ in which for each i and j in I the image $u\phi_i$ of an element u of U in S_i is identified with its image $u\phi_j$ in S_j. More precisely, if as before we denote by θ_i the natural monomorphism from S_i into $\Pi^* S_i$, then we define $P = \Pi^*_U S_i$ to be $(\Pi^* S_i)/\rho$, where ρ is the congruence on $\Pi^* S_i$ generated by the subset

$$\mathbf{R} = \{(u\phi_i\theta_i, u\phi_j\theta_j) : u \in U\} \tag{8.2.5}$$

of $(\Pi^* S_i) \times (\Pi^* S_i)$.

It is clear that for each i in I there is a morphism $\mu_i = \theta_i \rho^\natural$ from S_i into P. It is also clear from the definition of ρ that we have a commutative diagram



for every i and j. So there exists a morphism $\mu : U \to P$ such that $\phi_i\mu_i = \mu$ for every i in I. Since $U\mu \subseteq S_i\mu_i$ for every i in I, we necessarily have that $U\mu \subseteq S_i\mu_i \cap S_j\mu_j$ for all i and j. Hence the amalgam $\mathcal{A}$ is embedded in its free product P if and only if

$$\text{each } \mu_i \text{ is one-one;} \tag{8.2.6}$$

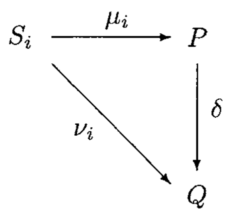
$$S_i\mu_i \cap S_j\mu_j \subseteq U\mu \text{ for all } i,\ j \text{ in } I \text{ such that } i \neq j. \tag{8.2.7}$$

If conditions (8.2.6) and (8.2.7) hold, we say that the amalgam $\mathcal{A}$ is *naturally embedded in its free product.* We shall see shortly that if an amalgam is embeddable at all then it is naturally embedded in its free product. First, however, we have

Proposition 8.2.3 *If $\mathcal{A} = [U; S_i; \phi_i]$ is a semigroup amalgam, then the free product $P = \Pi^*_U S_i$ is the 'pushout' of the diagram $\{U \to S_i\}_{i \in I}$. That is:*

(1) *there exists for each i in I a morphism $\mu_i : S_i \to P$ such that the diagram $\{U \to S_i \to P\}_{i \in I}$ commutes, that is, such that $\phi_i \mu_i = \phi_j \mu_j$ for all i and j in I;*

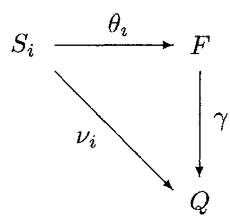
(2) *if Q is a semigroup for which morphisms $\nu_i : S_i \to Q$ exist such that $\phi_i \nu_i = \phi_j \nu_j$ for all i and j in I, then there exists a unique morphism $\delta : P \to Q$ such that the diagram*



(8.2.8)

commutes for each i in I.

Proof Property (1) has already been observed. To see that (2) holds, notice that by Proposition 8.2.1 there is a unique morphism $\gamma : F \to Q$ such that



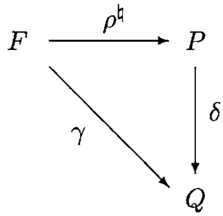
is a commutative diagram for each i in I. If $i, j \in I$ then, for all u in U,

$$u\phi_i\theta_i\gamma = u\phi_i\nu_i = u\phi_j\nu_j = u\phi_j\theta_j\gamma;$$

hence $(u\phi_i\theta_i, u\phi_j\theta_j) \in \gamma \circ \gamma^{-1}$. Recalling formula (8.2.5), we deduce that $\mathbf{R} \subseteq \gamma \circ \gamma^{-1}$; hence, since $\gamma \circ \gamma^{-1}$ is a congruence,

$$\rho = \mathbf{R}^{\#} \subseteq \gamma \circ \gamma^{-1}.$$

By Theorem 1.5.3 it follows that the morphism $\gamma : F \to Q$ factors through $P = F/\rho$, that is, there is a unique morphism $\delta : P \to Q$ such that the diagram



commutes. It now follows from the definition of μ_i and from the commutativity of these two diagrams that

$$\mu_i\delta = \theta_i\rho^\natural\delta = \theta_i\gamma = \nu_i$$

for each i in I; hence the diagram (8.2.8) commutes as required. □

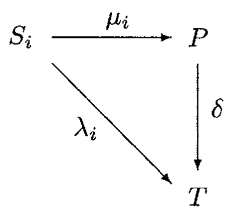
From this we can readily deduce

Theorem 8.2.4 *The semigroup amalgam*

$$\mathcal{A} = [U; \{S_i : i \in I\}; \{\phi_i : i \in I\}]$$

is embeddable in a semigroup if and only if it is naturally embedded in its free product.

Proof One way round this is obvious. Suppose therefore that $\mathcal{A}$ is embedded in a semigroup T, so that there exist monomorphisms $\lambda : U \to T$, $\lambda_i : S_i \to T \quad (i \in I)$ such that $\phi_i\lambda_i = \lambda$ for every i in I and such that $S_i\lambda_i \cap S_j\lambda_j = U\lambda$ whenever $i \neq j$. By Proposition 8.2.3 there exists a unique morphism $\delta : P \to T$ such that



is a commutative diagram for each i in I. It follows that each μ_i is a monomorphism, since for all x, y in S_i

$$x\mu_i = y\mu_i \Rightarrow x\mu_i\delta = y\mu_i\delta \Rightarrow x\lambda_i = y\lambda_i \Rightarrow x = y.$$

Suppose now that $i \neq j$ and that $x = s_i\mu_i = s_j\mu_j \in S_i\mu_i \cap S_j\mu_j$. Then $x\delta = s_i\mu_i\delta = s_i\lambda_i \in S_i\lambda_i$, and similarly $x\delta \in S_j\lambda_j$. Thus $x\delta \in S_i\lambda_i \cap S_j\lambda_j = U\lambda$, and so there exists u in U such that $x\delta = u\lambda$. That is,

$$s_i\lambda_i = s_i\mu_i\delta = x\delta = u\lambda = u\phi_i\lambda_i.$$

Since λ_i is by assumption a monomorphism, it follows that $s_i = u\phi_i$, and so $x = s_i\mu_i = u\phi_i\mu_i \in U\mu$. We have shown that both (8.2.6) and (8.2.7) hold. Thus $\mathcal{A}$ is naturally embedded in P as required. □

One case of a free product $\Pi_U^* S_i$ deserves special mention. If the semigroups S_i are all monoids with identity elements 1_i and if $U = \{1\}$ then $\Pi_U^* S_i$ is the *monoid free product* of the monoids S_i. This consists in effect of the element 1 together with all words $a_1 a_2 \dots a_m$ in the free product $\Pi^* S_i$ for which each a_r belongs to some $S_i \setminus \{1_i\}$. Notice that the monoid free product of a family of groups is the same as their group free product.

8.3 DOMINIONS AND ZIGZAGS

The ideas of tensor product and free product (with amalgamation) introduced in the last two sections have no very obvious connection, but eventually we shall see that the two ideas are closely linked. In this section we get the first hint of a connection, arising from an idea that seems at first entirely different.

If U is a submonoid of a monoid S and $d \in S$, then, following Isbell (1966), we shall say that U *dominates* d if for all monoids T and all (monoid) morphisms $\beta, \gamma : S \to T$,

$$[(\forall u \in U)\ u\beta = u\gamma] \ \Rightarrow\ d\beta = d\gamma.$$

More informally, U dominates d if any two morphisms of S that coincide on elements of U coincide also on d. The set of elements dominated by U is called the *dominion* of U in S and is written $\mathrm{Dom}_S(U)$. It is clear that $U \subseteq \mathrm{Dom}_S(U)$. Also, $\mathrm{Dom}_S(U)$ is a submonoid of S. For suppose that $d, d' \in \mathrm{Dom}_S(U)$ and that $\beta, \gamma : S \to T$ are such that $u\beta = u\gamma$ for all u in U. Then $d\beta = d\gamma$ and $d'\beta = d'\gamma$, and so

$$(dd')\beta = (d\beta)(d'\beta) = (d\gamma)(d'\gamma) = (dd')\gamma;$$

hence $dd' \in \mathrm{Dom}_S(U)$.

Since the mapping $U \mapsto \mathrm{Dom}_S(U)$ is a closure operation in the sense that is usual in algebra — see Exercise 2—we say that U is a *closed* submonoid of S if $\mathrm{Dom}_S(U) = U$, while if $\mathrm{Dom}_S(U) = S$ we say that U is a *dense* submonoid of S. Both types of submonoid can arise, and it is possible also to have

$$U \subset \mathrm{Dom}_S(U) \subset S.$$

See Exercise 6.

The idea of a dominion is closely linked to that of an epimorphism. According to the general theory of categories, a monoid morphism $\alpha : S \to T$ is an *epimorphism* if, for all monoids U and all morphisms $\beta, \gamma : T \to U$,

$$\alpha\beta = \alpha\gamma \ \Rightarrow\ \beta = \gamma.$$

It is easy to see that if α is onto then α is an epimorphism, and in the category of groups one has the converse result that every epimorphism is onto. In the category of semigroups or of monoids, however, this is not the case: indeed it is easy to see that $\alpha : S \to T$ is an epimorphism if and only if $\mathrm{Dom}_T(\mathrm{im}\,\alpha) = T$.

It is of course potentially very hard to discover whether or not a given element belongs to the dominion of a submonoid, since the definition of dominion involves phrases such as 'for all monoids' and 'for all morphisms'. For this reason it is important to characterize the notion of dominion in some more accessible way. With this end in view, consider a monoid S and a submonoid U of S, and let S' be an isomorphic copy of S, disjoint from S. To be more specific, let $\alpha : S \to S'$ be an isomorphism, and for each s in S let us write $s\alpha$ as s'. We now examine the amalgam $\mathcal{A} = [U; \{S, S'\}; \{\iota, \alpha|_U\}]$, where ι is the inclusion map from U into S. We show that this amalgam is *weakly embeddable*, which means precisely that it satisfies (8.2.6) but not necessarily (8.2.7). More formally, we have the following result, in which $\theta : S \to S * S'$ and $\theta' : S' \to S * S'$ are the standard inclusion maps, and the congruence ρ on $S * S'$ is generated by

$$\{(u\theta, u\alpha\theta') : u \in U\}.$$

Proposition 8.3.1 *Let S be a monoid and let U be a submonoid of S. Let S' be a monoid disjoint from S and let $\alpha : S \to S'$ be an isomorphism. Let $P = S *_U S' = (S * S')/\rho$ be the free product of the amalgam*

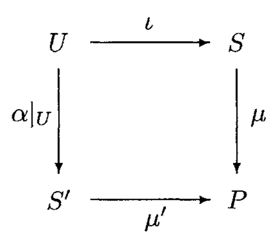
$$\mathcal{A} = [U; \{S, S'\}; \{\iota, \alpha|_U\}],$$

where ι is the inclusion map of U into S. Then both the natural maps $\mu : S \to P$, $\mu' : S' \to P$ defined by

$$s\mu = (s\theta)\rho, \quad s'\mu' = (s'\theta')\rho \quad (s \in S,\ s' \in S')$$

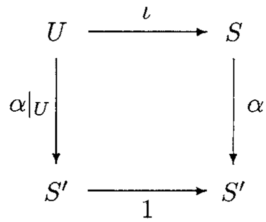
are monomorphisms.

Proof Certainly we have a commutative diagram

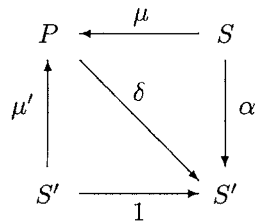


(8.3.1)

We also have a commutative diagram



and so, by the pushout property of the first diagram (more precisely, by Proposition 8.2.3) there is a unique morphism $\delta : P \to S'$ such that the diagram



(8.3.2)

commutes. From this it readily follows that μ and μ' are monomorphisms, for if $s_1, s_2 \in S$ then

$$s_1\mu = s_2\mu \;\Rightarrow\; s_1\mu\delta = s_2\mu\delta \;\Rightarrow\; s_1\alpha = s_2\alpha \;\Rightarrow\; s_1 = s_2,$$

and if $s'_1, s'_2 \in S'$ then

$$s'_1\mu' = s'_2\mu' \;\Rightarrow\; s'_1\mu'\delta = s'_2\mu'\delta \;\Rightarrow\; s'_1 = s'_2. \qquad \square$$

It is natural now to ask whether the amalgam $\mathcal{A} = [U; \{S, S'\}; \{\iota, \alpha|_U\}]$ has the second property (8.2.7) required for embeddability. In general it does not. In fact Isbell (1966) proved the following result.

Theorem 8.3.2 *Let S be a monoid and let U be a submonoid of S. Let S' be a monoid disjoint from S and let $\alpha : S \to S'$ be an isomorphism. Let $P = S *_U S'$ be the free product of the amalgam*

$$\mathcal{A} = [U; \{S, S'\}; \{\iota, \alpha|_U\}],$$

where ι is the inclusion map of U into S, and let μ, μ' be the natural maps from S, S', respectively, into P. Then

$$(S\mu \cap S'\mu')\mu^{-1} = \mathrm{Dom}_S(U).$$

Proof The commutativity of the diagram (8.3.1) ensures that the monomorphisms $\mu : S \to P$ and $\alpha\mu' : S \to P$ have the property that $u\mu = u\alpha\mu'$ for every u in U. Hence for every d in $\mathrm{Dom}_S(U)$ we have

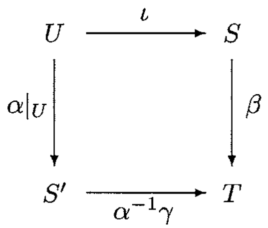
$$d\mu = d\alpha\mu' \in S\mu \cap S'\mu'.$$

Thus $\mathrm{Dom}_S(U) \subseteq (S\mu \cap S'\mu')\mu^{-1}$.

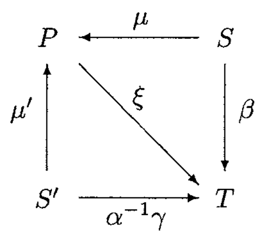
To show the opposite inclusion, consider an element d in $(S\mu \cap S'\mu')\mu^{-1}$. Then there exists s' in S' such that $d\mu = s'\mu'$, and from the commutativity of the diagram (8.3.2) it follows that

$$s' = s'\mu'\delta = d\mu\delta = d\alpha.$$

Now let T be a monoid and let $\beta : S \to T$ and $\gamma : S \to T$ be morphisms such that $u\beta = u\gamma$ for all u in U. Then the diagram



commutes and so, by Proposition 8.2.3 there exists a unique $\xi : P \to T$ such that



commutes. Hence

$$d\beta = d\mu\xi = s'\mu'\xi = d\alpha\mu'\xi = d\alpha\alpha^{-1}\gamma = d\gamma,$$

and so $d \in \mathrm{Dom}_S(U)$. □

While Theorem 8.3.2 provides an interesting characterization of the dominion, and an interesting connection with amalgamation theory, it may still be very hard to determine whether or not a given element d of S belongs to $(S\mu \cap S'\mu')\mu^{-1}$. We proceed now to find a more usable characterization of the dominion.

The following result is due in essence to Isbell (1966), but we present it in a form due to Stenström (1971).

Theorem 8.3.3 *Let U be a submonoid of a monoid S and let $d \in S$. Then $d \in \mathrm{Dom}_S(U)$ if and only if $d \otimes 1 = 1 \otimes d$ in the tensor product $A = S \otimes_U S$.*

Proof Suppose first that $d \otimes 1 = 1 \otimes d$, and suppose that we have a monoid T and morphisms $\beta, \gamma : S \to T$ such that $u\beta = u\gamma$ for all u in U. We may regard T as a (U, U)-bisystem if we define

$$u.t = (u\beta)t\big(= (u\gamma)t\big), \quad t.u = t(u\beta)\big(= t(u\gamma)\big).$$

Define $\psi : S \times S \to T$ by the rule that

$$(s, s')\psi = (s\beta)(s'\gamma) \quad \big((s, s') \in S \times S\big).$$

Then ψ is a (U,U)-map, and is even a bimap (in the sense of Section 8.1), since, for all s, s' in S and all u in U,

$$\begin{aligned}(su,s')\psi &= \big((su)\beta\big)(s'\gamma) = (s\beta)(u\beta)(s'\gamma)\\ &= (s\beta)(u\gamma)(s'\gamma) = (s\beta)\big((us')\gamma\big) = (s,us')\psi.\end{aligned}$$

It follows that the map ψ factors through the tensor product $S\otimes_U S$. That is to say, there is a map $\overline{\psi} : S\otimes_U S \to T$ such that

$$(s\otimes s')\overline{\psi} = (s,s')\psi = (s\beta)(s\gamma)$$

for all $s\otimes s'$ in $S\otimes_U S'$. Now, our assumption is that $d\otimes 1 = 1\otimes d$, and it now follows that

$$d\beta = (d\beta)(1\gamma) = (d\otimes 1)\overline{\psi} = (1\otimes d)\overline{\psi} = (1\beta)(d\gamma) = d\gamma.$$

Thus $d \in \mathrm{Dom}_S(U)$ as required.

To prove the converse, notice first that we can regard $A = S\otimes_U S$ as an (S,S)-bisystem in an obvious way:

$$s(x\otimes y) = sx\otimes y, \quad (x\otimes y)s = x\otimes ys \quad (x,y,s\in S).$$

Let $(\mathbf{Z}(A),+)$ be the free abelian group on the set A, that is to say, the set of all finite linear combinations $\sum z_i a_i$ of elements of A with integral coefficients, and with the obvious addition. (More formally, we could regard $\mathbf{Z}(A)$ as the set of all maps $\zeta : A \to \mathbf{Z}$ such that $a\zeta = 0$ for all but finitely many a in A, and define $\zeta+\tau$, $-\zeta$ by

$$a(\zeta+\tau) = a\zeta + a\tau, \quad a(\zeta) = -a\zeta \quad (a\in A).)$$

The abelian group $\mathbf{Z}(A)$ inherits a natural (S,S)-bisystem structure from A if we define, for s in S and $\sum z_i a_i$ in $\mathbf{Z}(A)$,

$$s\big(\sum z_i a_i\big) = \sum z_i(sa_i), \quad \big(\sum z_i a_i\big)s = \sum z_i(a_i s).$$

Indeed, $\mathbf{Z}(A)$ comes close to being a bimodule, since

$$s(x+y) = sx+sy \text{ and } (x+y)s = xs+ys \qquad (8.3.3)$$

for all s in S and all x, y in $\mathbf{Z}(A)$. (It is not actually a bimodule, since S is not a ring.)

We now define a binary operation on $S\times\mathbf{Z}(A)$ by the rule that

$$(p,x)(q,y) = (pq, py+xq).$$

(Notice that this is in essence the rule for the addition of fractions x/p and y/q.) Certainly the operation is associative, since from the bisystem laws in $\mathbf{Z}(A)$ and from (8.3.3) we have that

$$\begin{aligned}[(p,x)(q,y)](r,z) &= (pq,py+xq)(r,z) = \big((pq)r,(pq)z+(py+xq)r\big)\\ &= \big(p(qr),p(qz+yr)+x(qr)\big) = (p,x)(qr,qz+yr)\\ &= (p,x)[(q,y)(r,z)].\end{aligned}$$

In fact $S \times \mathbf{Z}(A)$ is a monoid with identity $(1, 0)$.

Suppose now that $d \in \mathrm{Dom}_S(U)$. We consider two morphisms β, γ from S into $S \times \mathbf{Z}(A)$, and show that they coincide on elements of U. First, it is easy to verify that β, defined by

$$s\beta = (s, 0) \quad (s \in S),$$

is a morphism from S into $S \times \mathbf{Z}(A)$. Less obviously, so is γ, where

$$s\gamma = (s, s \otimes 1 - 1 \otimes s) \quad (s \in S).$$

To see this, notice that we may write $s \otimes 1 - 1 \otimes s$ as $sa - as$, where $a = 1 \otimes 1$. Then observe that

$$\begin{aligned}(s\gamma)(t\gamma) &= (s, sa - as)(t, ta - at) = \big(st, s(ta - at) + (sa - as)t\big) \\ &= \big(st, (st)a - a(st)\big) = (st)\gamma.\end{aligned}$$

If $u \in U$ then $u \otimes 1 = 1 \otimes u$ in $A = S \otimes_U S$, and so $u\beta = u\gamma$. Since $d \in \mathrm{Dom}_S(U)$ it follows that $d\beta = d\gamma$, and so $d \otimes 1 = 1 \otimes d$, as required. □

Stenström's version of the theorem has the advantage of showing the strong analogy between the monoid result and the corresponding result for rings. When it comes to applications of the result, however, it is the original version due to Isbell that is the more useful. In effect we recover Isbell's (1966) version by applying Proposition 8.1.8. If U is a submonoid of a monoid S, then a system of equalities

$$\begin{aligned} d &= x_1 u_1 & u_1 &= v_1 y_1 \\ x_{i-1} v_{i-1} &= x_i u_i & u_i y_{i-1} &= v_i y_i \ (i = 2, \ldots, m-1) \\ x_{m-1} v_{m-1} &= u_m & u_m y_{m-1} &= d \end{aligned} \tag{8.3.4}$$

in which

$$u_1, \ldots, u_m, v_1, \ldots, v_{m-1} \in U$$

and

$$x_1, x_2, \ldots, x_{m-1}, y_1, y_2, \ldots, y_{m-1} \in S,$$

is called a *zigzag of length m in S over U with value d*. By the *spine* of the zigzag we mean the ordered $(2m-1)$-tuple $(u_1, v_1, u_2, v_2, \ldots, u_m)$. Then we have *Isbell's Zigzag Theorem for Monoids*:

Theorem 8.3.4 *If U is a submonoid of a monoid S then $d \in \mathrm{Dom}_S(U)$ if and only if either $d \in U$ or there exists a zigzag in S over U with value d.* □

In fact Isbell's theorem applies to semigroups as well as to monoids, and we shall want to make use of it in that form. To obtain the semigroup theorem we begin by adjoining an identity element 1 to S *whether or not it already has one.* If we call the resulting monoid S^* and write $U \cup \{1\}$ as U^*, then we easily verify that an element d in $S \setminus U$ is in $\mathrm{Dom}_S(U)$ if and only if it is in $\mathrm{Dom}_{S^*}(U^*)$, and hence if and only if $d \otimes 1 = 1 \otimes d$

in the tensor product $S^* \otimes_{U^*} S^*$. Then to obtain the Isbell theorem we simply observe that zigzags in which the extraneous element 1 appears can be shortened. For example, if $u_i = 1$, then the equalities

$$\begin{array}{rlrl} x_{i-2}v_{i-2} &= x_{i-1}u_{i-1} & u_{i-1}y_{i-2} &= v_{i-1}y_{i-1} \\ x_{i-1}v_{i-1} &= x_i u_i & u_i y_{i-1} &= v_i y_i \\ x_i v_i &= x_{i+1}u_{i+1} & u_{i+1}y_i &= v_{i+1}y_{i+1} \end{array}$$

can be collapsed to

$$\begin{array}{rlrl} x_{i-2}v_{i-2} &= x_{i-1}u_{i-1} & u_{i-1}y_{i-2} &= v_{i-1}v_i y_i \\ x_{i-1}v_{i-1}v_i &= x_{i+1}u_{i+1} & u_{i+1}y_i &= v_{i+1}y_{i+1}. \end{array}$$

Slightly less obviously, if some x_i is equal to 1, suppose that x_k is the *first* such x_i. If $k = 1$ then $d = x_1u_1 = u_1 \in U$. So suppose that $k \neq 1$, so that $x_1, \ldots, x_{k-1}$ are all in S. Then

$$\begin{aligned} d &= x_1u_1 = x_1v_1y_1 = x_2u_2y_1 = x_2v_2y_2 = x_3u_3y_2 \\ &= \cdots = x_k u_k y_{k-1} = u_k y_{k-1}, \end{aligned} \tag{8.3.5}$$

and so

$$\begin{array}{rlrl} d &= x_1u_1 & u_1 &= v_1y_1 \\ x_1v_1 &= x_2u_2 & u_2y_1 &= v_2y_2 \\ & \cdots & & \cdots \\ x_{k-1}v_{k-1} &= u_k & u_k y_{k-1} &= d \end{array}$$

is a zigzag of length k with value d in which every x_i is in S.

So we may suppose that in the zigzag (8.3.4) all the xs and us are in S. If some y_i is equal to 1, then an argument very similar to the one just employed, in which we consider y_k as the *last* occurrence of 1 among the ys, leads to a zigzag in which all xs, ys and us are in S.

We thus have *Isbell's Zigzag Theorem for Semigroups*:

Theorem 8.3.5 *If U is a subsemigroup of a semigroup S then $d \in \mathrm{Dom}_S(U)$ if and only if either $d \in U$ or there exists a zigzag in S over U with value d.* □

We conclude this section with one application of the Zigzag Theorem. This and other applications can be found in Howie and Isbell (1967). A semigroup U is called *absolutely closed* if $\mathrm{Dom}_S(U) = U$ for *every* semigroup S containing U.

Theorem 8.3.6 *Inverse semigroups are absolutely closed.*

Proof Let U be an inverse semigroup and let S be a semigroup containing U as a subsemigroup. Let d be an element of $\mathrm{Dom}_S(U) \setminus U$. Then there exists a zigzag (8.3.4) in S over U with value d. The method of proof is to show that (8.3.4) can be replaced by a zigzag

$$\begin{aligned} d &= x_1u_1 & u_1 &= v_1t_1 \\ x_{i-1}v_{i-1} &= x_iu_i & u_it_{i-1} &= v_it_i \ (i=2,\dots,m-1) \\ x_{m-1}v_{m-1} &= u_m & u_mt_{m-1} &= d \end{aligned} \tag{8.3.6}$$

which is *right-inner*, that is to say, is such that $t_1, t_2, \dots, t_{m-1} \in U$. From this it immediately follows that $d = u_mt_{m-1} \in U$.

In defining the elements $t_1, t_2, \dots, t_{m-1}$ it is convenient to begin by defining elements $w_1, w_2, \dots, w_{m-1}$ of U as follows:

$$w_1 = v_1^{-1}, \quad w_i = v_i^{-1}u_iw_{i-1} \quad (i = 2, \dots, m-1).$$

Now define $t_i = w_iu_1 \quad (i = 1, 2, \dots, m-1)$. The elements t_i evidently belong to U. Also

$$u_1 = v_1y_1 = v_1v_1^{-1}v_1y_1 = v_1v_1^{-1}u_1 = v_1t_1$$

and so the top right-hand equality of (8.3.6) is satisfied. To show that the other equalities hold, it is convenient first to prove by induction on i that

$$t_i = w_iw_i^{-1}y_i \quad (i = 1, 2, \dots, m-1). \tag{8.3.7}$$

This is obvious for $i = 1$. If we assume that it holds for $i = r-1$ then

$$\begin{aligned} t_r &= w_ru_1 = v_r^{-1}u_rw_{r-1}u_1 \\ &= v_r^{-1}u_rt_{r-1} = v_r^{-1}u_rw_{r-1}w_{r-1}^{-1}y_{r-1} \quad \text{(by hypothesis)} \\ &= v_r^{-1}u_r(u_r^{-1}u_r)(w_{r-1}w_{r-1}^{-1})y_{r-1} \\ &= v_r^{-1}u_rw_{r-1}w_{r-1}^{-1}u_r^{-1}u_ry_{r-1} \quad \text{(by commuting idempotents)} \\ &= v_r^{-1}u_rw_{r-1}w_{r-1}^{-1}u_r^{-1}v_ry_r \quad \text{(by (8.3.4))} \\ &= w_rw_r^{-1}y_r, \end{aligned}$$

exactly as required. We now have that, for $i = 1, 2, \dots, m-1$,

$$\begin{aligned} u_it_{i-1} &= u_iw_{i-1}w_{i-1}^{-1}y_{i-1} = u_i(u_i^{-1}u_i)(w_{i-1}w_{i-1}^{-1})y_{i-1} \\ &= u_iw_{i-1}w_{i-1}^{-1}u_i^{-1}u_iy_{i-1} = u_iw_{i-1}w_{i-1}^{-1}u_i^{-1}v_iy_i \\ &= (u_iw_{i-1}w_{i-1}^{-1}u_i^{-1})(v_iv_i^{-1})v_iy_i \\ &= v_i(v_i^{-1}u_iw_{i-1}w_{i-1}^{-1}u_i^{-1}v_i)y_i \\ &= v_i(w_iw_i^{-1})y_i = v_it_i. \end{aligned}$$

The final equality in (8.3.6) now follows easily, since

$$d = x_1u_1 = x_1v_1t_1 = x_2u_2t_1 = x_2v_2t_2 = \cdots = x_{m-1}v_{m-1}t_{m-1} = u_mt_{m-1}.$$

□

We shall see in Section 8.5 that inverse semigroups have an even stronger property.

8.4 DIRECT LIMITS, FREE EXTENSIONS AND FREE PRODUCTS

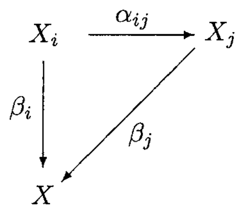
Let $(I, \leq)$ be a partially ordered set, let X_i $(i \in I)$ be a collection of (S,T)-bisystems indexed by I, and for all i, j in I such that $i \leq j$ let $\alpha_{ij} : X_i \to X_j$ be an (S,T)-morphism. Suppose also that

(1) $\alpha_{ii} = 1_{X_i}$ for all i in I;

(2) $\alpha_{ij}\alpha_{jk} = \alpha_{ik}$ whenever $i \leq j \leq k$.

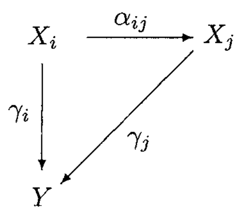
Then we say that (X_i, α_{ij}) is a *direct system* (of (S,T)-bisystems).

An (S,T)-bisystem X is called a *direct limit* of (X_i, α_{ij}) if there exist (S,T)-morphisms $\beta_i : X_i \to X$ such that:
(a) for all $i \leq j$ in I the diagram



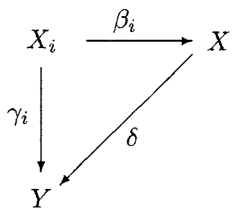
(8.4.1)

commutes;
(b) if an (S,T)-bisystem Y has the property that there exist (S,T)-morphisms $\gamma_i : X_i \to Y$ $(i \in I)$ for which the diagrams



(8.4.2)

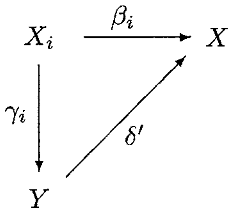
(where $i \leq j$) all commute, then there is a unique (S,T)-morphism $\delta : X \to Y$ such that the diagram



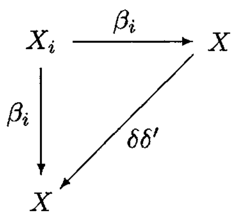
(8.4.3)

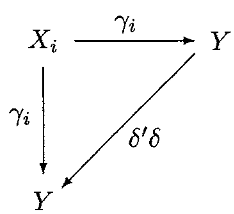
commutes for every i in I.

If a direct limit exists then it is unique up to isomorphism. To see this we first remark that if we take $Y = X$ and $\gamma_i = \beta_i$ in the above then the unique δ such that the diagram (8.4.3) commutes is the identity map 1_X. Suppose now that both X and Y are direct limits of (X_i, α_{ij}). Then we have the diagram (8.4.3) as before, and we also have a unique $\delta' : Y \to X$ such that



commutes. That is, $\beta_i\delta = \gamma_i$, $\gamma_i\delta' = \beta_i$. Hence $\beta_i\delta\delta' = \beta_i$, $\gamma_i\delta'\delta = \gamma_i$, and so we have commutative diagrams





But then by the above remark $\delta\delta' = 1_X$, $\delta'\delta = 1_Y$, and so $X \simeq Y$ as required.

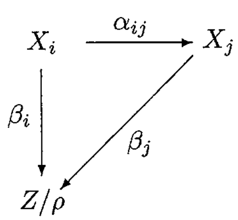
The next thing to establish is that direct limits exist for (S,T)-bisystems. Consider a direct system (X_i, α_{ij}). Then there is no loss of generality in supposing that the sets X_i $(i \in I)$ are pairwise disjoint. If we define Z to be the union of all the sets X_i then Z is an (S,T)-bisystem in an obvious way. Let ρ be the (S,T)-congruence on Z generated by the relation

$$R = \{(x_i, x_i\alpha_{ij}) : i, j \in I,\ x_i \in X_i \text{ and } j \geq i\}.$$

Then Z/ρ is the direct limit of (X_i, α_{ij}). To see this, notice first that we have a morphism $\beta_i : X_i \to Z/\rho$ for each i, defined by

$$x_i\beta_i = x_i\rho \quad (x_i \in X_i).$$

Then $x_i\alpha_{ij}\beta_j = (x_i\alpha_{ij})\rho = x_i\rho = x_i\beta_i$ for every x_i, and so the diagram



commutes. Also, if Y, with associated morphisms γ_i, is such that we have a commutative diagram (8.4.2) for each i, then certainly we have a morphism ζ from $Z = \bigcup_{i\in I} X_i$ into Y given by

$$x_i\zeta = x_i\gamma_i \quad (x_i \in X_i,\ i \in I).$$

Now $R \subseteq \ker \zeta$, for

$$(x_i\alpha_{ij})\zeta = x_i\alpha_{ij}\gamma_j = x_i\gamma_i = x_i\zeta;$$

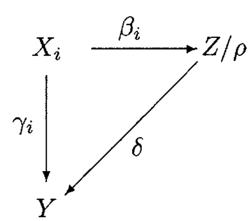
hence there exists a morphism $\delta : Z/\rho \to Y$, defined by

$$(x_i\rho)\delta = x_i\gamma_i \quad (x_i \in X_i,\ i \in I).$$

Moreover,

$$x_i\beta_i\delta = (x_i\rho)\delta = x_i\gamma_i$$

for all x_i in X_i, and so the diagram



commutes. Finally, the morphism δ is unique, since if δ' is another morphism with the same properties, then for all x_i in X_i and all i in I

$$(x_i\rho)\delta' = x_i\beta_i\delta' = x_i\gamma_i = (x_i\rho)\delta,$$

that is, $\delta' = \delta$.

We deduce that Z/ρ is the direct limit (unique up to isomorphism) of the direct system (X_i, α_{ij}).

Two special cases of direct limits are worth mentioning. The first is where $I = \mathbf{N} = \{1, 2, 3, \dots\}$, with $1 < 2 < 3 < \cdots$ Here we need only specify morphisms $\alpha_n : X_n \to X_{n+1}$, since for all $m \le n$ we can *define* $\alpha_{mn} = \alpha_m\alpha_{m+1}\dots\alpha_{n-1}$. We denote this direct system by (X_n, α_n).

We shall have occasion to make use of the following result:

Proposition 8.4.1 *With the notation above, let Z/ρ be the direct limit of the direct system (X_n, α_n). Then the maps β_n $(n \ge 1)$ are all one-one if and only if the maps α_n $(n \ge 1)$ are all one-one.*

Proof Suppose that all the maps $\alpha_n : X_n \to X_{n+1}$ are one-one, and suppose that $x_m\beta_m = x'_m\beta_m$, where $x_m, x'_m \in X_m$. Then $x_m\rho = x'_m\rho$, and this is possible if and only if

$$x_m\alpha_m\alpha_{m+1}\dots\alpha_n = x_m\alpha_m\alpha_{m+1}\dots\alpha_n$$

for some $n \ge m$. Since all the α_k are one-one, it follows that $x_m = x'_m$, and so β_m is one-one. Conversely, suppose that some α_m is *not* one-one, so that $x_m\alpha_m = x'_m\alpha_m$ for some $x_m \ne x'_m$ in X_m. Then $x_m\beta_m = x'_i\beta_m$, and so β_m is not one-one. □

The other special case is where $I = \{0, 1, 2\}$, where $0 < 1$, $0 < 2$ and 1 and 2 are incomparable. Here the direct limit is an (S, T)-system Z for which we have a commutative diagram

$$\begin{array}{ccc} X_0 & \xrightarrow{\alpha_{01}} & X_1 \\ {\scriptstyle\alpha_{02}}\downarrow & & \downarrow{\scriptstyle\beta_1} \\ X_2 & \xrightarrow[\beta_2]{} & Z \end{array}$$

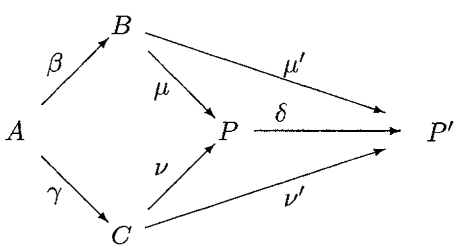
This is what is called a *pushout diagram*, a notion we have already encountered in Proposition 8.2.3. Simplifying the notation, we define a commutative diagram

$$\begin{array}{ccc} A & \xrightarrow{\beta} & B \\ {\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\mu} \\ C & \xrightarrow[\nu]{} & P \end{array} \qquad (8.4.4)$$

to be a *pushout diagram* if, whenever

$$\begin{array}{ccc} A & \xrightarrow{\beta} & B \\ {\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\mu'} \\ C & \xrightarrow[\nu']{} & P' \end{array} \qquad (8.4.5)$$

is commutative, there exists a unique $\delta : P \to P'$ such that the diagram



commutes.

Notice that the pushout P of the diagram

$$\begin{array}{ccc} A & \xrightarrow{\ \beta\ } & B \\ {\scriptstyle\gamma}\big\downarrow & & \\ C & & \end{array}$$

is equal to the quotient $(A \dot\cup B \dot\cup C)/\rho$ of the disjoint union $A \dot\cup B \dot\cup C$ of A, B and C, where ρ is the (S,T)-congruence generated by all pairs $(a, a\beta)$, $(a, a\gamma)$, where $a \in A$. The maps $\mu : B \to P$, $\nu : C \to P$ are given by

$$b\mu = b\rho, \qquad c\nu = c\rho.$$

We shall require to make repeated use of the following lemma:

Lemma 8.4.2 *Consider the pushout diagram* (8.4.4). *Let $b \in B$, $c \in C$ and suppose that $b\mu = c\nu$. Then $b \in \operatorname{im}\beta$.*

Proof We have $(b,c) \in \rho$, and so b is connected to c by a sequence of transitions based on the pairs $(a, a\beta)$, $(a, a\gamma)$. Such a sequence cannot even begin unless $b \in \operatorname{im}\beta$. □

The other salient fact that we shall require later applies to direct limits in general:

Proposition 8.4.3 *Let D, with associated morphisms β_i, be the direct limit of (X_i, α_{ij}) in S-ENS-T. Let $A \in U$-ENS-S and let $B \in T$-ENS-V. Then $A \otimes_S D \otimes_T B$ is the direct limit of $(A \otimes_S X_i \otimes_T B, 1 \otimes \alpha_{ij} \otimes 1)$ in U-ENS-V.*

Proof Certainly the diagram

$$\begin{array}{ccc} A \otimes X_i \otimes B & \xrightarrow{\ 1 \otimes \alpha_{ij} \otimes 1\ } & A \otimes X_j \otimes B \\ {\scriptstyle 1 \otimes \beta_i \otimes 1}\big\downarrow & \swarrow {\scriptstyle 1 \otimes \beta_j \otimes 1} & \\ A \otimes D \otimes B & & \end{array}$$

commutes for all $i \leq j$ in I. Suppose now that we have a (U,V)-bisystem Q and (U,V)-morphisms $\gamma_i : A \otimes X_i \otimes B \to Q$ such that $(1 \otimes \alpha_{ij} \otimes 1)\gamma_j = \gamma_i$ for all $i \leq j$. Recall now that $D = Z/\rho$, where Z is the disjoint union $\bigcup\{X_i i \in I\}$ and ρ is the (S,T)-congruence generated by all the pairs $(x_i, x_i\alpha_{ij})$ for x_i in X_i and $i \leq j$ in I. Recall also that $x_i\beta_i = x_i\rho$ for all x_i in X_i. Now define $\delta_0 : A \times Z \times B \to Q$ by

$$(a, x_i, b)\delta_0 = (a \otimes x_i \otimes b)\gamma_i \quad (x_i \in X_i).$$

Since for all $j \geq i$ we have

$$\begin{aligned}(a, x_i\alpha_{ij}, b)\delta_0 &= (a \otimes x_i\alpha_{ij} \otimes b)\gamma_j \\ &= (a \otimes x_i \otimes b)(1 \otimes \alpha_{ij} \otimes 1)\gamma_j \\ &= (a \otimes x_i \otimes b)\gamma_i \\ &= (a, x_i, b)\delta_0,\end{aligned}$$

δ_0 induces a map $\delta_1 : A \times D \times B \to Q$ given by

$$(a, x_i\rho, b)\delta_1 = (a \otimes x_i \otimes b)\gamma_i.$$

Next, for all s in S and all t in T we have

$$(as, x_i\rho, b)\delta_1 = (as \otimes x_i \otimes b)\gamma_i = (a \otimes sx_i \otimes b)\gamma_i = \big(a, s(x_i\rho), b\big)\delta_1,$$

and similarly

$$(a, x_i\rho, tb)\delta_1 = \big(a, (x_i\rho)t, b\big)\delta_1;$$

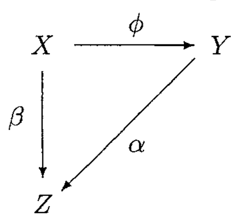
hence δ_1 induces a map $\delta : A \otimes D \otimes B \to Q$ given by

$$(a \otimes x_i\rho \otimes b)\delta = (a \otimes x_i \otimes b)\gamma_i \quad (x_i \in X_i). \tag{8.4.6}$$

It is now clear that δ is a (U, V)-morphism and that $(1 \otimes \beta_i \otimes 1)\delta = \gamma_i$ for all i. The uniqueness of δ is clear from (8.4.6). □

Suppose now that U is a submonoid of a monoid S. Let $Y \in$ ENS-U and let X be a sub-U-system of Y. Suppose in fact that $X \in$ ENS-S. Informally, we have that Y is bigger whereas X is better (in that it is acted upon by S rather than merely by U). We ask the question as to whether we can find some Z in ENS-S which is bigger *and* better.

It pays in fact to ask a slightly different question, in which $Y \in$ ENS-U, $X \in$ ENS-S and there exists a U-morphism $\phi : X \to Y$. (In the previous paragraph ϕ is simply the inclusion map.) We wish to find Z in ENS-S with the property that there exists a U-morphism $\alpha : Y \to Z$ and an S-morphism $\beta : X \to Z$ such that the diagram



(8.4.7)

commutes.

The required Z is not hard to find, the technique being a modification of one used for rings and modules by Cohn (1959). Let $T = Y \otimes_U S$, and let σ be the equivalence on T generated by the relation

$$R = \{(x\phi \otimes s, x'\phi \otimes s') : x, x' \in X,\ s, s' \in S,\ xs = x's'\}. \tag{8.4.8}$$

Then, since $xs = x's'$ implies $(xs)s'' = (x's')s''$ for all s'' in S, it follows that σ is a right S-congruence on T, and so $Z = T/\sigma$ is a right S-system. Now define $\alpha : Y \to Z$ by

$$y\alpha = (y \otimes 1)\sigma \quad (y \in Y). \tag{8.4.9}$$

Then for all y in Y and u in U

$$(yu)\alpha = (yu \otimes 1)\sigma = (y \otimes u)\sigma = (y \otimes 1)\sigma.u = (y\alpha)u,$$

and so α is a U-morphism. In essentially the same way we define $\beta : X \to Z$ by

$$x\beta = (x\phi \otimes 1)\sigma \quad (x \in X), \tag{8.4.10}$$

but now we can show that β is a right S-morphism, since for all x in X and s in S

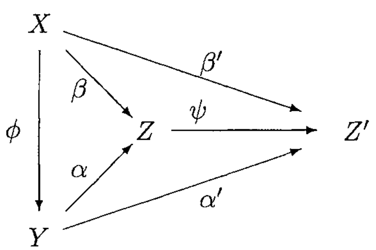
$$\begin{aligned}(xs)\beta &= \big((xs)\phi \otimes 1\big)\sigma \\ &= (x\phi \otimes s)\sigma \ \big(\text{by (8.4.8), since } (xs)1 = xs\big) \\ &= (x\phi \otimes 1)\sigma.s = (x\beta)s.\end{aligned}$$

Finally, since it is clear that for all x in X

$$x\phi\alpha = (x\phi \otimes 1)\sigma = x\beta,$$

we see that the diagram (8.4.7) commutes.

The S-system $Z = (Y \otimes_U S)/\sigma$ is in fact best possible, in the sense that if $Z' \in$ ENS-S has the property that there exist a U-morphism $\alpha' : Y \to Z'$ and an S-morphism $\beta' : X \to Z'$ such that $\phi\alpha' = \beta'$, then there exists a unique S-morphism $\psi : Z \to Z'$ such that the diagram



(8.4.11)

commutes. To see this, define a map $\psi'' : Y \times S \to Z'$ by

$$(y, s)\psi'' = (y\alpha')s.$$

It is easy to check that ψ'' is a right S-morphism. It is, moreover, a bimap, since for all u in U, y in Y and s in S

$$(yu, s)\psi'' = \big((yu)\alpha'\big)s = (y\alpha')(us) = (y, us)\psi'';$$

hence there is an S-morphism $\psi' : Y \otimes_U S \to Z'$ given by

$$(y \otimes s)\psi' = (y\alpha')s.$$

Again, let us suppose that x, x' in X and s, s' in S are such that $xs = x's'$, so that $(x\phi \otimes s, x'\phi \otimes s') \in R$. Then

$$\begin{aligned}(x\phi \otimes s)\psi' &= \big((x\phi)\alpha'\big)s = (x\beta')s \\ &= (xs)\beta' = (x's')\beta' = \cdots = (x'\phi \otimes s')\psi';\end{aligned}$$

thus $\sigma \subseteq \ker \psi'$, and so ψ' induces an S-morphism $\psi : Z \to Z'$ given by

$$\big((y \otimes s)\sigma\big)\psi = (y\alpha')s \quad (y \in Y,\ s \in S).$$

We now easily see that for all x in X

$$x\beta\psi = \big((x\phi \otimes 1)\sigma\big)\psi = (x\phi)\alpha' = x\beta',$$

and for all y in Y

$$y\alpha\psi = \big((y \otimes 1)\sigma\big)\psi = y\alpha';$$

thus the diagram (8.4.11) commutes.

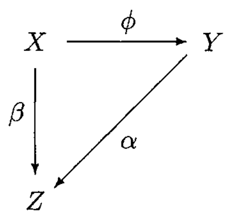
To show the uniqueness of ψ, suppose that we could substitute an alternative S-morphism $\chi : Z \to Z'$ in the diagram (8.4.11). Then for all $(y \otimes 1)\sigma$ in Z

$$\big((y \otimes 1)\sigma\big)\chi = y\alpha\chi = y\alpha' = y\alpha\psi = \big((y \otimes 1)\sigma\big)\psi,$$

and so $\chi = \psi$.

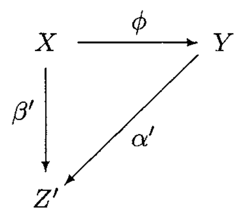
The right S-system $Z = (Y \otimes_U S)/\sigma$ is called the *free S-extension of X and Y* and is denoted by $F(S, X, Y)$. To summarize what we have learned, $Z = F(S, X, Y)$ is a right S-system and is characterized up to isomorphism by the properties:

(E1) there exists a U-morphism $\alpha : Y \to Z$ and an S-morphism $\beta : X \to Z$ such that the diagram

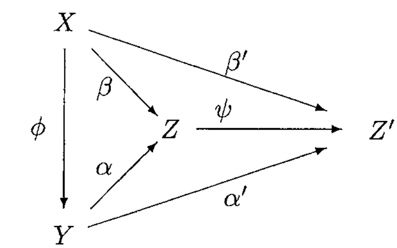


commutes;

(E2) If $Z' \in$ ENS-S is such that there exist a U-morphism $\alpha' : Y \to Z'$ and an S-morphism $\beta' : X \to Z'$ such that



commutes, then there is a unique S-morphism $\psi : Z \to Z'$ such that the diagram



commutes.

The fact that these properties do indeed characterize $F(S, X, Y)$ up to isomorphism is proved by a standard argument, closely analogous to the argument used to establish uniqueness of direct limits.

There is indeed a close connection between free extensions and direct limits, as evidenced by the following result:

Proposition 8.4.4 *Let U be a submonoid of a monoid S. Let $X \in$ ENS-S, $Y \in$ ENS-U, and let $\phi : X \to Y$ be a U-morphism. Let $Z = (Y \otimes_U S)/\sigma$ be the free S-extension of X and Y. Let $\gamma : X \otimes S \to X$ be given by*

$$(x \otimes s)\gamma = xs \quad (x \in X,\ s \in S).$$

Then the diagram

$$\begin{array}{ccc} X \otimes S & \xrightarrow{\phi \otimes 1} & Y \otimes S \\ {\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\sigma^\natural} \\ X & \xrightarrow[\beta]{} & Z \end{array}$$

is a pushout in ENS-S.

Proof It is clear from the definition of σ that the diagram commutes. Suppose now that Z' is such that there exist β' and τ such that the diagram

$$\begin{array}{ccc} X \otimes S & \xrightarrow{\phi \otimes 1} & Y \otimes S \\ {\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\tau} \\ X & \xrightarrow[\beta']{} & Z' \end{array}$$

commutes. Let $(x\phi \otimes s, x'\phi \otimes s')$ be a typical element of the set R defined by (8.4.8); thus $xs = x's'$ in X. Then

$$(x\phi \otimes s)\tau = (x \otimes s)(\phi \otimes 1)\tau = (x \otimes s)\gamma\beta'$$

$$= (xs)\beta' = (x's')\beta' = \cdots = (x'\phi \otimes s')\tau.$$

It follows that τ induces a unique S-morphism $\delta : Z \to Z'$ given by

$$[(y \otimes s)\sigma]\delta = (y \otimes s)\tau.$$

The verification that for all x in X

$$\begin{aligned} x\beta\delta &= [(x\phi \otimes 1)\sigma]\delta = (x\phi \otimes 1)\tau \\ &= (x \otimes 1)(\phi \otimes 1)\tau = (x \otimes 1)\gamma\beta' = x\beta' \end{aligned}$$

completes the proof. □

We have seen in Section 8.2 how the free product $P = S_1 *_U S_2$ of two monoids S_1, S_2 amalgamating a common submonoid U arises as a quotient by a certain congruence ρ of the free product $S_1 * S_2$. (Since the identity elements are amalgamated in P it matters not at all whether we start with the semigroup free product or the monoid free product.) Each element of P is an equivalence class $(x_1x_2\dots x_n)\rho$, where $x_1x_2\dots x_n \in S_1 * S_2$. Following Renshaw (1986a) we shall write this as

$$(x_1, x_2, \dots, x_n).$$

Notice that if we make the obvious definitions (in which y belongs to S_1 or S_2)

$$(x_1, x_2, \dots, x_n)y = \begin{cases} (x_1, x_2, \dots, x_n, y) & \text{if } \sigma(x_n) \neq \sigma(y) \\ (x_1, x_2, \dots, x_{n-1}, x_n y) & \text{if } \sigma(x_n) = \sigma(y) \end{cases} \qquad (8.4.12)$$

(with similar definitions for left multiplication), then P is an (S_i, S_j)-bisystem for $i, j \in \{1, 2\}$. Our final aim in this section is to present an alternative characterization of $S_1 *_U S_2$ as a direct limit of U-systems.

Let $[U; S_1, S_2; \phi_1, \phi_2]$ be a monoid amalgam. We now construct a direct system (W_n, α_n). First, let $W_1 = S_1$, an (S_1, S_1)-bisystem, let $W_2 = S_1 \otimes_U S_2$, an (S_1, S_2)-bisystem, and let α_1 be the (S_1, U)-morphism given by

$$s_1\alpha_1 = s_1 \otimes 1 \quad (s_1 \in S_1).$$

Next, let

$$W_3 = F(S_1, W_1, W_2) = (S_1 \otimes_U S_2 \otimes_U S_1)/\sigma_1,$$

where σ_1, in accordance with equations (8.4.8) and (8.4.9), is the congruence on $S_1 \otimes S_2 \otimes S_1$ generated by

$$R_1 = \{(s_1 \otimes 1 \otimes t_1,\ s_1' \otimes 1 \otimes t_1') : s_1t_1 = s_1't_1'\}.$$

Let $\alpha_2 : W_2 \to W_3$ be the (S_1, U)-morphism given by

$$(s_1 \otimes s_2)\alpha_2 = (s_1 \otimes s_2 \otimes 1)\sigma_1,$$

again in accordance with the standard map from Y into $F(S, X, Y)$.

In general, suppose that we have constructed a sequence

$$W_1 \xrightarrow{\alpha_1} W_2 \xrightarrow{\alpha_2} \cdots \xrightarrow{\alpha_{n-1}} W_n,$$

and suppose that W_k belongs to S_1-ENS-S_2 if k is even, and to S_1-ENS-S_1 if k is odd. Suppose also that α_k is an (S_1, U)-map for $k = 1, 2, \dots, n-1$. Let $i \equiv n+1 \pmod 2$ and let W_{n+1} be the (S_1, S_i)-bisystem

$$F(S_i, W_{n-1}, W_n) = (W_n \otimes S_i)/\sigma_{n-1},$$

where σ_{n-1} is the congruence on $W_n \otimes S_i$ generated by

$$R_{n-1} = \{(w_{n-1}\alpha_{n-1} \otimes s_i,\ w'_{n-1}\alpha_{n-1} \otimes s'_i : w_{n-1}s_i = w'_{n-1}s'_i\}.$$

Then $\alpha_n : W_n \to W_{n+1}$ is given, as in (8.4.9), by

$$w_n\alpha_n = (w_n \otimes 1)\sigma_{n-1} \quad (w_n \in W_n),$$

and is an (S_1, U)-map.

A more explicit notation for elements of W_n will in due course be helpful. We have

$$W_n = \big(\cdots\big((S_1 \otimes S_2 \otimes S_1)/\sigma_1 \otimes S_2\big)/\sigma_2 \otimes \cdots \otimes S_i\big)/\sigma_{n-2},$$

and we shall denote a typical element

$$\big(\cdots\big((a_1 \otimes a_2 \otimes a_3)\sigma_1 \otimes a_4\big)\sigma_2 \otimes \cdots \otimes a_n\big)\sigma_{n-2}$$

by $[a_1, a_2, \dots, a_n]$. This, as we shall see, is not altogether unconnected with the element $(a_1, a_2, \dots, a_n)$ of $S_1 *_U S_2$, but it is important to note that in the square brackets formula we must have $a_1, a_3, \dots \in S_1$ and $a_2, a_4, \dots \in S_2$. If $n = 1$ then $[a_1]$ is simply the element a_1 of S_1, and if $n = 2$ then $[a_1, a_2]$ is the element $a_1 \otimes a_2$ of $W_2 = S_1 \otimes S_2$. Notice that, for all $n \geq 3$,

$$[a_1, a_2, \dots, a_n] = \big([a_1, a_2, \dots, a_{n-1}] \otimes a_n\big)\sigma_{n-2}. \tag{8.4.13}$$

Our aim is to prove the following result:

Theorem 8.4.5 *Let $[U; S_1, S_2; \phi_1, \phi_2]$ be a monoid amalgam. Then $P = S_1 *_U S_2$ is the direct limit in U-ENS-U of the direct system (W_n, α_n) defined above.*

Proof Recall from Section 8.2 that there exist monoid morphisms $\mu_1 : S_1 \to P$, $\mu_2 : S_2 \to P$, given in the Renshaw notation by

$$s_1\mu_1 = s_1\rho = (s_1), \quad s_2\mu_2 = s_2\rho = (s_2).$$

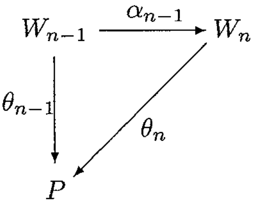
We define maps $\theta_n : W_n \to P$ inductively as follows. First $\theta_1 = \mu_1$, an (S_1, S_1)-map, and θ_2, an (S_1, S_2)-map, is given by

$$(s_1 \otimes s_2)\theta_2 = (s_1, s_2).$$

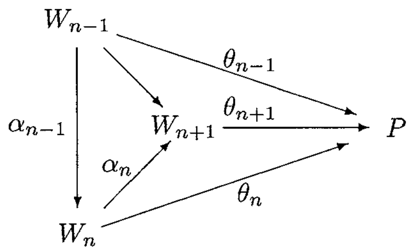
Suppose now that $n \geq 2$ and that we have defined $\theta_k : W_k \to P$ for $k = 1, \dots, n$ such that

(a) θ_k is an (S_1, S_i)-map, where $i \equiv k \pmod 2$;
(b) $\alpha_{k-1}\theta_k = \theta_{k-1} \quad (2 \leq k \leq n)$.

We then have a commutative diagram



in which θ_{n-1} is an (S_1, S_i)-map and θ_n is an (S_1, U)-map (in fact an (S_1, S_j)-map, where $j \neq i$. Here $i \equiv n-1 \equiv n+1 \pmod 2$.) Since $W_{n+1} = F(S_i, W_{n-1}, W_n)$, it follows by (E2) that there is a unique (S_1, S_i)-map $\theta_{n+1} : W_{n+1} \to P$ such that the diagram



commutes.

We have in fact proved half of what we need, namely that there are (U, U)-maps $\theta_n : W_n \to P$ $(n \geq 1)$ such that $\alpha_{n-1}\theta_n = \theta_{n-1}$ for all $n \geq 2$. It is, however, worth taking time at this point to describe the maps θ_n in a more explicit way. Let us suppose inductively that $\theta_{n-1} : W_{n-1} \to P$ is given by

$$[a_1, a_2, \ldots, a_{n-1}]\theta_{n-1} = (a_1, a_2, \ldots, a_{n-1}).$$

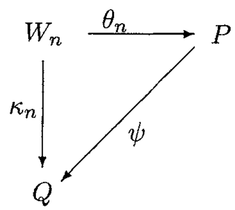
(We have already anchored our induction at $n = 2$ and $n = 3$.) Then, from (8.4.13) and (8.4.12),

$$\begin{aligned}
[a_1, a_2, \ldots, a_n]\theta_n &= \Big(\big([a_1, a_2, \ldots, a_{n-1}] \otimes a_n\big)\sigma_{n-2}\Big)\theta_n \\
&= \bigg(\Big(\big([a_1, a_2, \ldots, a_{n-1}] \otimes 1\big)\sigma_{n-2}\Big)\theta_n\bigg)a_n \\
&= \big([a_1, a_2, \ldots, a_{n-1}]\alpha_{n-1}\theta_n\big)a_n \\
&= \big([a_1, a_2, \ldots, a_{n-1}]\theta_{n-1}\big)a_n \\
&= (a_1, a_2, \ldots, a_{n-1})a_n \\
&= (a_1, a_2, \ldots, a_n).
\end{aligned}$$

We have proved that for all $n \geq 1$

$$[a_1, a_2, \ldots, a_n]\theta_n = (a_1, a_2, \ldots, a_n). \tag{8.4.14}$$

To complete the proof that P is the direct limit of (W_n, α_n) we must show that it is 'best possible' relative to the given properties. More precisely, suppose that Q in U-ENS-U is such that there exist (U, U)-maps $\kappa_n : W_n \to Q$ such that $\alpha_{n-1}\kappa_n = \kappa_{n-1}$ for all $n \geq 2$. Then we must find a unique $\psi : P \to Q$ such that the diagram



(8.4.15)

commutes for every n. Let us define $\psi : P \to Q$ by

$$(a_1, a_2, \ldots, a_n)\psi = \begin{cases} [a_1, a_2, \ldots, a_n]\kappa_n & \text{if } a_1 \in S_1 \\ [1, a_1, a_2, \ldots, a_n]\kappa_{n+1} & \text{if } a_1 \in S_2. \end{cases}$$

Supposing for the moment that ψ is well-defined, we see that it is a (U, U)-map, and certainly the diagrams (8.4.15) commute for every n. So the crucial task is to show that ψ is well-defined.

Now $P = (S_1 * S_2)/\rho$, where ρ is the congruence generated by

$$\{(u\phi_1, u\phi_2) : u \in U\}.$$

In fact, as pointed out in Howie (1962), two elements $(a_1, a_2, \ldots, a_m)$ and $(b_1, b_2, \ldots, b_n)$ of P are equal if and only if they are identical or are connected by a finite sequence of transformations of three main types:

1. E-steps (E for edge). There are six cases:
   (a) $(z_1, \ldots, z_i u, z_{i+1}, \ldots, z_n) \to (z_1, \ldots, z_i, u z_{i+1}, \ldots, z_n)$;
   (b) $(z_1, \ldots, z_i, u z_{i+1}, \ldots, z_n) \to (z_1, \ldots, z_i u, z_{i+1}, \ldots, z_n)$;
   (c) $(z_1, \ldots, z_n u) \to (z_1, \ldots, z_n, u)$;
   (d) $(z_1, \ldots, z_n, u) \to (z_1, \ldots, z_n u)$;
   (e) $(u z_1, \ldots, z_n) \to (u, z_1, \ldots, z_n)$;
   (f) $(u, z_1, \ldots, z_n) \to (u z_1, \ldots, z_n)$.
2. M-steps (M for middle):
$$(z_1, \ldots, z_i' u z_i'', \ldots, z_n) \to (z_1, \ldots, z_i', u, z_i'', \ldots, z_n).$$
3. S-steps (S for syllable):
$$(z_1, \ldots, z_i', u, z_i'', \ldots, z_n) \to (z_1, \ldots, z_i' u z_i'', \ldots, z_n).$$

There is a certain amount of abuse of notation here: for example, in case 1(a), if we assume that $z_i \in S_1$, then $z_{i+1} \in S_2$ and we are transforming $(\ldots, z_i(u\phi_1), z_{i+1}, \ldots)$ into $(\ldots, z_i, (u\phi_2)z_{i+1}, \ldots)$. What we therefore need to show is that if two elements of $P = S_1 *_U S_2$ are linked by a step of any of the above types then their images under ψ are equal. In cases 1(a) and 1(b) this is clear, since the defining property of tensor products certainly gives

$$[z_1, \ldots, z_i u, z_{i+1}, \ldots, z_n] = [z_1, \ldots, z_i, u z_{i+1}, \ldots, z_n],$$

in the case where $z_1 \in S_1$, and just as easily gives

$$[1, z_1, \ldots, z_i u, z_{i+1}, \ldots, z_n] = [1, z_1, \ldots, z_i, u z_{i+1}, \ldots, z_n],$$

in the case where $z_1 \in S_2$. The remaining cases are less trivial, and the following lemma is of assistance.

Lemma 8.4.6 *For all $n \geq 3$ and for all i in $\{2, \ldots, n-1\}$*

$$[a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n] = [a_1, \ldots, a_{i-1}a_{i+1}, \ldots, a_n, 1, 1].$$

Proof Let us write $[a_1, \ldots, a_{i-1}]$ as w_i. Then

$$\begin{aligned}
&[a_1, \ldots, a_{i-1}, 1, a_{i+1}] = \big((w_i \otimes 1)\sigma_{i-2} \otimes a_{i+1}\big)\sigma_{i-1} \\
&= (w_i \alpha_{i-1} \otimes a_{i+1})\sigma_{i-1} \\
&= \big((w_i a_{i+1})\alpha_{i-1} \otimes 1\big)\sigma_{i-1} \\
&\qquad \text{by definition of } \sigma_{i-1}, \text{ since } w_i a_{i+1} = (w_i a_{i+1})1, \\
&= (w_i a_{i+1})\alpha_{i-1}\alpha_i = [a_1, \ldots, a_{i-1}a_{i+1}, 1, 1].
\end{aligned}$$

What we have shown is the case $k = 1$ of the more general statement that

$$[a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_{i+k}] = [a_1, \ldots, a_{i-1}a_{i+1}, \ldots, a_{i+k}, 1, 1],$$

and we proceed now to prove the more general statement by induction. If we assume the result for k, then

$$\begin{aligned}
&[a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_{i+k+1}] \\
&= \big([a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_{i+k}] \otimes a_{i+k+1}\big)\sigma_{i+k-1} \\
&= \big([a_1, \ldots, a_{i-1}a_{i+1}, \ldots, a_{i+k}, 1, 1] \otimes a_{i+k+1}\big)\sigma_{i+k-1} \\
&= \big([a_1, \ldots, a_{i-1}a_{i+1}, \ldots, a_{i+k}, 1]\alpha_{i+k-1} \otimes a_{i+k+1}\big)\sigma_{i+k-1} \\
&= \Big(\big([a_1, \ldots, a_{i-1}a_{i+1}, \ldots, a_{i+k}, 1]a_{i+k+1}\big)\alpha_{i+k-1} \otimes 1\Big)\sigma_{i+k-1} \\
&\qquad (\text{by definition of } \sigma_{i+k-1}) \\
&= [a_1, \ldots, a_{i-1}a_{i+1}, \ldots, a_{i+k}, a_{i+k+1}]\alpha_{i+k-1}\alpha_{i+k} \\
&= [a_1, \ldots, a_{i-1}a_{i+1}, \ldots, a_{i+k}, a_{i+k+1}, 1, 1].
\end{aligned}$$

This completes the proof of the lemma. □

Returning now to the main verification, we consider case 1(c), where we have two elements

$$w = (z_1, \ldots, z_{n-1}, z_n u) \text{ and } w' = (z_1, \ldots, z_{n-1}, z_n, u)$$

of P. If $z_1 \in S_1$ then

$$w\psi = [z_1, \ldots, z_{n-1}, z_n u]\kappa_n = [z_1, \ldots, z_{n-1}, z_n u]\alpha_n \kappa_{n+1}$$
$$= [z_1, \ldots, z_{n-1}, z_n u, 1]\kappa_{n+1} = [z_1, \ldots, z_n, u]\kappa_{n+1} = w'\psi.$$

The case where $z_1 \in S_2$, and where $w\psi = [1, z_1, \ldots, z_n u]$, is verified in exactly the same way.

Case 1(d) is the inverse of 1(c). So suppose now that we have case 1(e), where we have two elements

$$w = (uz_1, z_2, \ldots, z_n) \text{ and } w' = (u, z_1, z_2, \ldots, z_n)$$

of P. If $uz_1 \in S_1$ then in the second element $u \in S_2$, and we have

$$\begin{aligned} w\psi &= [uz_1, \ldots, uz_n]\kappa_n \\ &= [uz_1, z_2, \ldots, z_n]\alpha_n \alpha_{n+1}\kappa_{n+2} = [uz_1, z_2, \ldots, z_n, 1, 1]\kappa_{n+2} \\ &= [u, 1, z_1, z_2, \ldots, z_n] \text{ (by Lemma 8.4.6)} \\ &= [1, u, z_1, z_2, \ldots, z_n] = w'\psi. \end{aligned}$$

If $uz_1 \in S_2$ then

$$w\psi = [1, u, z_1, z_2, \ldots, z_n]\kappa_{n+1} = [u, z_1, z_2, \ldots, z_n]\kappa_{n+1} = w'\psi.$$

Case 1(f) is the inverse of case 1(e).

Suppose now that we have case 2, where we are considering two elements

$$w = (z_1, \ldots, z_i, \ldots, z_n) \text{ and } w' = (z_1, \ldots, z_i', u, z_i'', \ldots, z_n),$$

with $z_i = z_i' u z_i''$. If $z_1 \in S_1$ then

$$\begin{aligned} w\psi &= [z_1, \ldots, z_i, \ldots, z_n]\kappa_n \\ &= [z_1, \ldots, z_i, \ldots, z_n]\alpha_n \alpha_{n+1}\kappa_{n+2} = [z_1, \ldots, z_i' u z_i'', \ldots, z_n, 1, 1]\kappa_{n+2} \\ &= [z_1, \ldots, z_i' u, 1, z_i'', \ldots, z_n]\kappa_{n+2} \text{ (by Lemma 8.4.6)} \\ &= [z_1, \ldots, z_i', u, z_i'', \ldots, z_n]\kappa_{n+2} = w'\psi. \end{aligned}$$

The case where $z_1 \in S_2$, and where $w\psi = [1, z_1, \ldots, z_i, \ldots, z_n]\kappa_{n+1}$, is very similar.

Case 3 is the inverse of case 2.

This completes the proof that the map ψ is well-defined. Hence Theorem 8.4.5 is established. □

8.5 THE EXTENSION PROPERTY

Let U be a submonoid of a monoid S. We say that U has the *extension property* in S if for every X in ENS-U and every Y in U-ENS the map $X \otimes_U Y \to X \otimes_U S \otimes_U Y$ defined by $x \otimes y \mapsto x \otimes 1 \otimes y$ is one-one. It follows in particular (by putting $Y = U$) that the map $x \mapsto x \otimes 1$ from X into $X \otimes_U S$ is one-one (the *left extension property*) and (by putting $X = U$)

that the map $y \mapsto 1 \otimes y$ from Y into $S \otimes_U Y$ is one-one (the *right extension property*).

A closely related and somewhat more general notion is that of *purity.* A monomorphism $\phi : A \to B$ in S-ENS-T is called *pure* if for every X in ENS-S and Y in T-ENS the induced map $1 \otimes \phi \otimes 1 : X \otimes_S A \otimes_T Y \to X \otimes_S B \otimes_T Y$ is one-one. It is not hard to see that U has the extension property in S if and only if the inclusion monomorphism from U into S is pure.

A further related notion is that of *flatness.* If $A \in S$-ENS then we say that A is *flat* if for every monomorphism $\phi : X \to Y$ of right S-systems the induced map $\phi \otimes 1 : X \otimes_S A \to Y \otimes_S A$ is one-one. Dually, a right S-system A is *flat* if for every monomorphism $\psi : X \to Y$ of left S-systems the induced map $1 \otimes \psi : A \otimes_S X \to A \otimes_S Y$ is one-one.

A monoid S is called *absolutely extendable* if it has the extension property in every monoid T containing it as a submonoid. A monoid S is called *absolutely flat* if all left and right S-systems are flat. It is clear that the extension property and flatness are related in the sense that both are concerned with the preservation of monomorphisms. A more precise connection is given by the following result.

Proposition 8.5.1 *Every absolutely flat monoid is absolutely extendable.*

Proof Suppose that S is absolutely flat, and let T be a monoid containing S as a submonoid. We must show that if $X \in$ ENS-S and $Y \in S$-ENS then the map $X \otimes_S Y \to X \otimes_S T \otimes_S Y$ given by $x \otimes y \mapsto x \otimes 1 \otimes y$ is one-one. Now the map

$$X \simeq X \otimes_S S \to X \otimes_S T$$

induced by the inclusion monomorphism from S into T is one-one, since X is flat, and it then follows by the flatness of Y that the map

$$X \otimes_S Y \simeq X \otimes_S S \otimes_S Y \to X \otimes_S T \otimes_S Y,$$

sending $x \otimes y$ to $x \otimes 1 \otimes y$, is one-one. Hence S has the extension property in T. □

Recall now our previous notation (introduced following the statement of Theorem 8.3.4) that for a semigroup S the monoid S^* is defined to be $S \cup \{1\}$, with the obvious multiplication, where 1 is adjoined whether or not S already has an identity. We say that a subsemigroup U of a semigroup S has the *semigroup extension property* in S if the monoid U^* has the (monoid) extension property in S^*. Now every monoid is also a semigroup, and so we have a possible source of confusion. That this is not a serious confusion follows from

Proposition 8.5.2 *Let U be a submonoid of a monoid S. Then U has the semigroup extension property in S if and only if U has the monoid extension property in S.*

Proof Suppose first that U has the semigroup extension property in S. Let $X \in$ ENS-U, $Y \in U$-ENS. Then we can make X and Y into U^*-systems by defining $x.1 = x$ $(x \in X)$ and $1.y = y$ $(y \in Y)$. Moreover $X \otimes_{U^*} Y = X \otimes_U Y$, for the identification of $(x.1, y)$ with $(x, 1.y)$ within $X \otimes_{U^*} Y$ adds nothing to the identifications already present in $X \otimes_U Y$. Denote the identity of U (and of S) by e, and suppose that $x \otimes e \otimes y = x' \otimes e \otimes y'$ in $X \otimes_U S \otimes_U Y$. Then certainly $x \otimes e \otimes y = x' \otimes e \otimes y'$ in $X \otimes_{U^*} S^* \otimes_{U^*} Y$. Indeed in $X \otimes_{U^*} S^* \otimes_{U^*} Y$ we have

$$\begin{aligned} x \otimes 1 \otimes y &= xe \otimes 1 \otimes y = x \otimes e1 \otimes y \\ &= x \otimes e \otimes y = x' \otimes e \otimes y' = \cdots = x' \otimes 1 \otimes y'. \end{aligned}$$

Hence $x \otimes y = x' \otimes y'$ in $X \otimes_{U^*} Y = X \otimes_U Y$, and so U has the monoid extension property in S.

Conversely, suppose that U has the monoid extension property in S. Denote the identity element of S by e; then $e \in U$ since U is a submonoid. Let $X \in$ ENS-U^*, $Y \in U^*$-ENS. We must show that the map $x \otimes y \mapsto x \otimes 1 \otimes y$ from $X \otimes_{U^*} Y$ into $X \otimes_{U^*} S^* \otimes_{U^*} Y$ is one-one.

We cannot assume that $X \in$ ENS-U, since we may have $xe \neq x$ for some x in X. Equally, we cannot assume that $Y \in U$-ENS. However, we do have $X' = Xe = XU \in$ ENS-U and $Y' = eY = UY \in U$-ENS. Suppose now that

$$x_1 \otimes 1 \otimes y_1 = x_2 \otimes 1 \otimes y_2$$

in $X \otimes_{U^*} S^* \otimes_{U^*} Y$. Then

$$x_1 e \otimes e \otimes e y_1 = x_2 e \otimes e \otimes e y_2$$

in $X \otimes_U S \otimes_U Y$, for every transition

$$zu \otimes s \otimes t \ \leftrightarrow \ z \otimes us \otimes t \quad \text{or} \quad z \otimes su \otimes t \leftrightarrow z \otimes s \otimes ut$$

in $X \otimes_{U^*} S^* \otimes_{U^*} Y$ can be mirrored by a transition

$$\begin{aligned} &(ze)u \otimes ese \otimes et \ \leftrightarrow \ ze \otimes u(ese) \otimes et \\ \text{or} \quad &ze \otimes (ese)u \otimes et \ \leftrightarrow \ ze \otimes ese \otimes u(et) \end{aligned}$$

in $X \otimes_U S \otimes_U Y$. By the monoid extension property we deduce that $x_1 e \otimes ey_1 = x_2 e \otimes ey_2$ in $Xe \otimes_U eY$.

Now if $x_1 \notin XU$ and $y_1 \notin UY$ then

$$x_1 \otimes 1 \otimes y_1 = x_2 \otimes 1 \otimes y_2 \ \Rightarrow \ x_1 = x_2,\ y_1 = y_2,$$

for no non-trivial transition from $x_1 \otimes 1 \otimes y_1$ is possible. Similar considerations apply to $x_2 \otimes 1 \otimes y_2$, and so there are in essence three non-trivial cases:

(1) $x_1, x_2 \in XU$;

(2) $y_1, y_2 \in UY$;

(3) $x_1 \in XU$, $y_2 \in UY$.

In case (1) we have

$$x_1 \otimes y_1 = x_1e^2 \otimes y_1 = x_1e \otimes ey_1 = x_2e \otimes ey_2 = x_2e^2 \otimes y_2 = x_2 \otimes y_2$$

in $X \otimes_{U^*} Y$. Case (2) is similar, while in case (3) we have

$$x_1 \otimes y_1 = x_1e^2 \otimes y_1 = x_1e \otimes ey_1 = x_2e \otimes ey_2 = x_2 \otimes e^2y_2 = x_2 \otimes y_2$$

in $X \otimes_{U^*} Y$. □

As a consequence of this result we do not need to be over-precise about which of the extension properties we mean.

We have already come across the idea of a *unitary* subsemigroup, particularly in Chapter 7, but for convenience we repeat the definition here: a subsemigroup U of a semigroup S is called *unitary* if

$$(\forall u \in U)(\forall s \in S)\ us \in U \ \Rightarrow\ s \in U;$$

$$(\forall u \in U)(\forall s \in S)\ su \in U \ \Rightarrow\ s \in U.$$

More generally, we say that U is a *quasi-unitary* subsemigroup of S if there is a (U^*, U^*)-morphism $\phi : S^* \to S^*$ such that

$$\phi^2 = \phi, \quad 1\phi = 1;$$

$$(\forall u \in U)(\forall s \in S)\ us \in U \ \Rightarrow\ s\phi \in U;$$

$$(\forall u \in U)(\forall s \in S)\ su \in U \ \Rightarrow\ s\phi \in U.$$

Every unitary subsemigroup is quasi-unitary, with ϕ as the identity map of S^*. Notice that by the (U^*, U^*)-morphism property of ϕ it follows that for all u in U

$$u\phi = (u1)\phi = u(1\phi) = u1 = u. \tag{8.5.1}$$

Notice also that for all u in U and s in S

$$(us)\phi = u(s\phi), \quad (su)\phi = (s\phi)u. \tag{8.5.2}$$

If S is a monoid then a submonoid U of S is called *quasi-unitary* if there exists a (U, U)-morphism $\phi : S \to S$ such that

$$\phi^2 = \phi, \quad 1\phi = 1;$$

$$(\forall u \in U)(\forall s \in S)\ us \in U \ \Rightarrow\ s\phi \in U;$$

$$(\forall u \in U)(\forall s \in S)\ su \in U \ \Rightarrow\ s\phi \in U.$$

It is not hard to verify that this 'monoid quasi-unitary property' is equivalent to the semigroup property defined previously.

The connection between this notion and the extension property given earlier is as follows:

Proposition 8.5.3 *If U is a quasi-unitary subsemigroup of a semigroup S, then U^* has the extension property in S^*.*

Proof Let U be a quasi-unitary subsemigroup of a semigroup S. Let $X \in$ ENS-U^*, $Y \in U^*$-ENS, and suppose that

$$x \otimes 1 \otimes y = x' \otimes 1 \otimes y'$$

in $X \otimes_{U^*} S^* \otimes_{U^*} Y$. We must show that $x \otimes y = x' \otimes y'$ in $X \otimes_{U^*} Y$. We first establish a preliminary lemma:

Lemma 8.5.4 *If $x \otimes 1 = x' \otimes su$ in $X \otimes_{U^*} S^*$, then $s\phi \in U^*$ and $x = x'(s\phi)u$.*

Proof Suppose that $x \otimes 1 = x' \otimes su$. Then by Proposition 8.1.8 we have equations

$$\begin{aligned} x &= x_1u_1 & u_1 &= v_1y_1 \\ x_1v_1 &= x_2u_2 & u_2y_1 &= v_2y_2 \\ x_iv_i &= x_{i+1}u_{i+1} & u_{i+1}y_i &= v_{i+1}y_{i+1} \\ &(i = 2, \ldots, n-1) & & \\ x_{n-1}v_{n-1} &= x'u_n & u_ny_{n-1} &= su, \end{aligned}$$

in which $u_1, \ldots, u_n, v_1, \ldots, v_{n-1} \in U^*$, $x_1, \ldots, x_{n-1} \in X$, $y_1, \ldots, y_{n-1} \in Y$.

By virtue of (8.5.1) and (8.5.2) the equalities on the right imply equalities

$$\begin{aligned} u_1 &= u_1\phi = (v_1y_1)\phi = v_1(y_1\phi), \\ u_2(y_1\phi) &= v_2(y_2\phi), \\ u_{i+1}(y_i\phi) &= v_{i+1}(y_{i+1}\phi) \ (i = 1, \ldots, n-2), \\ u_n(y_{n-1}\phi) &= (s\phi)u, \end{aligned}$$

and then from the definition of a quasi-unitary subsemigroup we deduce successively that

$$y_1\phi \in U^*, \ y_2\phi \in U^*, \ldots, y_{n-1}\phi \in U^*, \quad s\phi \in U^*.$$

Also

$$\begin{aligned} x &= x_1u_1 = x_1v_1(y_1\phi) = x_2u_2(y_1\phi) \\ &= x_2v_2(y_2\phi) = \cdots = x'u_n(y_{n-1}\phi) = x'(s\phi)u, \end{aligned}$$

as required. □

We return now to the proof of Proposition 8.5.3, interpreting $x \otimes 1 \otimes y$ as $(x \otimes 1) \otimes y$ (as we may, by the remark following Proposition 8.1.11). We may suppose that we have equations

$$\begin{array}{rlrl} x \otimes 1 &= x_1 \otimes s_1 u_1 & u_1 y &= v_1 y_1 \\ x_1 \otimes s_1 v_1 &= x_2 \otimes s_2 u_2 & u_2 y_1 &= v_2 y_2 \\ x_i \otimes s_i v_i &= x_{i+1} \otimes s_{i+1} u_{i+1} & u_{i+1} y_i &= v_{i+1} y_{i+1} \\ & (i = 2, \dots, n-2) & & \\ x_{n-1} \otimes s_{n-1} v_{n-1} &= x' \otimes u_n & u_n y_{n-1} &= y' \end{array}$$

in which $u_1, \dots, u_n, v_1, \dots, v_{n-1} \in U^*$, $x_1, \dots, x_{n-1} \in X$, $y_1, \dots, y_{n-1} \in Y$, and $s_1, \dots, s_{n-1} \in S^*$. By Lemma 8.5.4 we have $s_1\phi \in U^*$ and $x = x_1(s_1\phi)u_1$. Also, by applying $1 \otimes \phi$ to the equation $x_1 \otimes s_1 v_1 = x_2 \otimes s_2 u_2$ we deduce that

$$x_1(s_1\phi)v_1 \otimes 1 = x_1 \otimes (s_1\phi)v_1 = x_2 \otimes (s_2\phi)u_2;$$

hence again by Lemma 8.5.4 we deduce that $s_2\phi \in U^*$ and $x_1(s_1\phi)v_1 = x_2(s_2\phi)u_2$.

Continuing in this way we show that $s_i\phi \in U^*$ $(i = 1, \dots, n-1)$ and that

$$x_i(s_i\phi)v_i = x_{i+1}(s_{i+1}\phi)u_{i+1}.$$

Hence

$$\begin{array}{rlrl} x &= \big(x_1(s_1\phi)\big)u_1 & u_1 y &= v_1 y_1 \\ \big(x_1(s_1\phi)\big)v_1 &= \big(x_2(s_2\phi)\big)u_2 & u_2 y_1 &= v_2 y_2 \\ \big(x_i(s_i\phi)\big)v_i &= \big(x_{i+1}(s_{i+1}\phi)\big)u_{i+1} & u_{i+1} y_i &= v_{i+1} y_{i+1} \\ & (i = 2, \dots, n-2) & & \\ \big(x_{n-1}(s_{n-1}\phi)\big)v_{n-1} &= x' u_n & u_n y_{n-1} &= y', \end{array}$$

which gives us $x \otimes y = x' \otimes y'$ as required. □

The connection between the extension property and subsemigroups of 'unitary type' is further exhibited by the next result. A subsemigroup U of a semigroup S is called *relatively unitary* if, for all u in U and s in S,

$$us \in U \ \Rightarrow\ (\exists s' \in U^*)\ us = us',$$

$$su \in U \ \Rightarrow\ (\exists s' \in U^*)\ su = s'u.$$

Every quasi-unitary subsemigroup is relatively unitary (with $s' = s\phi$), but the converse is not true. (See Exercise 18.)

Proposition 8.5.5 *Let S be a semigroup and let U be a subsemigroup of S having the extension property. Then U is relatively unitary.*

Proof Suppose that $us = v \in U$ for some u in U and s in S. Let $x_1U^* \dot\cup x_2U^*$, a disjoint union, be the free right U^*-system on two generators—see Exercise 2 below—and let X be the quotient of $x_1U^* \dot\cup x_2U^*$ by the U^*-congruence generated by the singleton set

$\{(x_1u, x_2u)\}$. In effect we identify x_1w and x_2w if and only if $w \in uU^*$. Notice now that in $X \otimes_{U^*} S^*$

$$\begin{aligned} x_1v \otimes 1 &= x_1 \otimes v = x_1 \otimes us = x_1u \otimes s \\ &= x_2u \otimes s = x_2 \otimes us = x_2 \otimes v = x_2v \otimes 1. \end{aligned}$$

However, since U has the extension property, the map

$$X \simeq X \otimes U^* \to X \otimes S^*$$

induced by the inclusion $U^* \subseteq S^*$, and given by $x \to x \otimes 1 \quad (x \in X)$, is one-one. We deduce that $x_1v = x_2v$ in X, and hence $v \in uU^*$. We have verified half of the relatively unitary property. The other half follows in a similar manner. □

We shall see shortly that the extension property has a further connection with previously encountered ideas. First we establish the following lemma.

Lemma 8.5.6 *Let U be a submonoid of a monoid S, and suppose that U has the extension property in S. Let $\lambda : X \to Y$ be a morphism of right U-sets, and let $y \otimes 1 = x\lambda \otimes s$ in $Y \otimes S$. Then $y \in \operatorname{im}\lambda$.*

Proof Consider the pushout diagram

$$\begin{array}{ccc} X & \xrightarrow{\lambda} & Y \\ {\scriptstyle\lambda}\downarrow & & \downarrow{\scriptstyle\alpha} \\ Y & \xrightarrow[\beta]{} & P \end{array}$$

By Proposition 8.4.3 the diagram

$$\begin{array}{ccc} X \otimes S & \xrightarrow{\lambda \otimes 1} & Y \otimes S \\ {\scriptstyle\lambda \otimes 1}\downarrow & & \downarrow{\scriptstyle\alpha \otimes 1} \\ Y \otimes S & \xrightarrow[\beta \otimes 1]{} & P \otimes S \end{array}$$

is also a pushout. Suppose that $y \otimes 1 = x\lambda \otimes s$ in $Y \otimes S$. Then in $P \otimes S$

$$\begin{aligned} y\alpha \otimes 1 &= (y \otimes 1)(\alpha \otimes 1) = (x\lambda \otimes s)(\alpha \otimes 1) = x\lambda\alpha \otimes s \\ &= x\lambda\beta \otimes s = (x\lambda \otimes s)(\beta \otimes 1) = (y \otimes 1)(\beta \otimes 1) = y\beta \otimes 1. \end{aligned}$$

By the extension property the map $y \mapsto y \otimes 1$ from $Y \simeq Y \otimes U$ to $Y \otimes S$ is one-one. Hence $y\alpha = y\beta$, and it now follows by Lemma 8.4.2 that $y \in \operatorname{im}\lambda$. □

We are now in a position to prove the following result:

Proposition 8.5.7 *Let U be a submonoid of a monoid S. If U has the extension property in S then U is closed.*

Proof In the lemma above, put $X = U$, $Y = S$ and take λ as the inclusion map from U into S. If $d \in \mathrm{Dom}_S(U)$ then by Theorem 8.3.3 we have $d \otimes 1 = 1 \otimes d = 1\lambda \otimes d$ in $S \otimes S$. The lemma allows us to conclude that $d \in \operatorname{im} \lambda$. That is, $d \in U$, and so U is closed. □

It is convenient to establish two further technical lemmas before stating the main theorem of this section:

Lemma 8.5.8 *Let U be a submonoid of a monoid S and suppose that U has the extension property in S. Let $X, Y \in$ ENS-U and let $\lambda : X \to Y$ be a U-monomorphism. Let $Z \in U$-ENS and suppose that $\lambda \otimes 1 : X \otimes Z \to Y \otimes Z$ is also a monomorphism. If $y \otimes 1 \otimes z = x\lambda \otimes s \otimes z'$ in $Y \otimes S \otimes Z$, then there exist x_1 in X and z_1 in Z such that $y \otimes 1 \otimes z = x_1\lambda \otimes 1 \otimes z_1$.*

Proof Suppose that $y \otimes 1 \otimes z = x\lambda \otimes s \otimes z'$ in $Y \otimes S \otimes Z$, and consider the pushout diagram

$$\begin{array}{ccc} X & \xrightarrow{\lambda} & Y \\ {\scriptstyle\lambda}\downarrow & & \downarrow{\scriptstyle\alpha} \\ Y & \xrightarrow[\beta]{} & P \end{array}$$

By Proposition 8.4.3 the diagram

$$\begin{array}{ccc} X \otimes Z & \xrightarrow{\lambda \otimes 1} & Y \otimes Z \\ {\scriptstyle\lambda \otimes 1}\downarrow & & \downarrow{\scriptstyle\alpha \otimes 1} \\ Y \otimes Z & \xrightarrow[\beta \otimes 1]{} & P \otimes Z \end{array} \tag{8.5.3}$$

is also a pushout. We now have that in $P \otimes S \otimes Z$

$$\begin{aligned} y\alpha \otimes 1 \otimes z &= (y \otimes 1 \otimes z)(\alpha \otimes 1 \otimes 1) = (x\lambda \otimes s \otimes z')(\alpha \otimes 1 \otimes 1) \\ &= x\lambda\alpha \otimes s \otimes z' = x\lambda\beta \otimes s \otimes z' \\ &= \cdots = y\beta \otimes 1 \otimes z. \end{aligned}$$

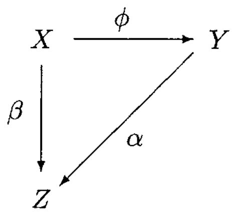
By the extension property we deduce that $y\alpha \otimes z = y\beta \otimes z$ in $P \otimes Z$. Hence by the pushout property of the diagram (8.5.3) and by Lemma 8.4.2 there exist $x_1 \otimes z_1$ in $X \otimes Z$ such that

$$y \otimes z = (x_1 \otimes z_1)(\lambda \otimes 1) = x_1\lambda \otimes z_1.$$

Thus $y \otimes 1 \otimes z = x_1\lambda \otimes 1 \otimes z_1$ as required. □

Recall now, from the beginning of the section, that a monomorphism $\pi : X \to Y$ of (U,U)-bisystems is called *pure* if for every A in ENS-U and every B in U-ENS the induced map $1 \otimes \pi \otimes 1 : A \otimes X \otimes B \to A \otimes Y \otimes B$ is one-one.

Lemma 8.5.9 *Let U be a submonoid of a monoid S and suppose that U has the extension property in S. Let $X \in U$-ENS-S, $Y \in U$-ENS-U and let $\phi : X \to Y$ be a pure (U,U)-monomorphism. Let $Z = F(S,X,Y)$ be the free S-extension of X and Y. Then in the diagram*



—see diagram (8.4.7)—both α and β are monomorphisms and α is pure.

Proof We have $Z = (Y \otimes_U S)/\sigma$, where σ is the (U,S)-congruence generated by

$$R = \{(x\phi \otimes s, x'\phi \otimes s') : xs = x's'\}.$$

It is easy to see that the relation R is symmetric, and that it is compatible with the left action of U and the right action of S. With the assumptions of the lemma we shall see that it is also transitive. Suppose in fact that

$$(x_1\phi \otimes s_1, x_2\phi \otimes s_2) \in R,$$

that

$$x_2\phi \otimes s_2 = x_3\phi \otimes s_3,$$

and that

$$(x_3\phi \otimes s_3, x_4\phi \otimes s_4) \in R,$$

so that

$$x_1s_1 = x_2s_2, \quad x_3s_3 = x_4s_4. \tag{8.5.4}$$

Then from $x_2\phi \otimes s_2 = x_3\phi \otimes s_3$ we deduce that $x_2 \otimes s_2 = x_3 \otimes s_3$, since ϕ is pure. Let $\psi_0 : X \times S \to X$ be given by

$$(x,s)\psi_0 = xs.$$

Then ψ_0, being a bimap (with respect to U) induces a (U,S)-morphism $\psi : X \otimes_U S \to X$, defined by

$$(x \otimes s)\psi = xs.$$

Applying ψ to the equality $x_2 \otimes s_2 = x_3 \otimes s_3$ now gives $x_2s_2 = x_3s_3$, which together with (8.5.4) gives $x_1s_1 = x_4s_4$. Thus $(x_1\phi \otimes s_1, x_4\phi \otimes s_4) \in R$, as required.

We deduce that σ coincides with $R \cup 1_{Y \otimes S}$.

Suppose now that y_1, y_2 in Y are such that $y_1\alpha = y_2\alpha$. Then $(y_1 \otimes 1)\sigma = (y_2 \otimes 1)\sigma$, and so either

(i) $y_1 \otimes 1 = y_2 \otimes 1$, or

(ii) $y_1 \otimes 1 = x_1\phi \otimes s_1$ and $y_2 \otimes 1 = x_2\phi \otimes s_2$, with $x_1 s_1 = x_2 s_2$.

In case (i) it follows by the extension property that $y_1 = y_2$, for the map $Y \simeq Y \otimes_U U \to Y \otimes_U S$ is one-one. In case (ii) we deduce from Lemma 8.5.6 that there exist x_1', x_2' in X such that $y_1 = x_1'\phi$, $y_2 = x_2'\phi$. Thus

$$x_1'\phi \otimes 1 = x_1\phi \otimes s_1, \quad x_2'\phi \otimes 1 = x_2\phi \otimes s_2,$$

and since $\phi \otimes 1$ is one-one it follows that

$$x_1' \otimes 1 = x_1 \otimes s_1, \quad x_2' \otimes 1 = x_2 \otimes s_2.$$

Then applying the map $\psi : x \otimes s \mapsto xs$ gives

$$x_1' = x_1 s_1 = x_2 s_2 = x_2',$$

and so $y_1 = y_2$ as required.

We have shown that α is a monomorphism. Since $\beta = \phi\alpha$ it follows that β too is a monomorphism,

It remains to show that α is pure. Let $A \in$ ENS-U and $B \in U$-ENS. We require to show that the map

$$1 \otimes \alpha \otimes 1 : A \otimes Y \otimes B \to A \otimes Z \otimes B$$

is one-one. From Propositions 8.4.3 and 8.4.4 we deduce that

$$\begin{array}{ccc} A \otimes (X \otimes S) \otimes B & \longrightarrow & A \otimes (Y \otimes S) \otimes B \\ \theta \big\downarrow & & \big\downarrow \\ A \otimes X \otimes B & \longrightarrow & A \otimes Z \otimes B \end{array} \tag{8.5.5}$$

is a pushout diagram, where Z, as before, is $(Y \otimes S)/\sigma$. All tensor products are over U, and the map θ from $A \otimes X \otimes S \otimes B$ to $A \otimes X \otimes B$ is given by

$$(a \otimes x \otimes s \otimes b)\theta = a \otimes xs \otimes b.$$

We now show that $A \otimes Z \otimes B$ is the quotient of $A \otimes Y \otimes S \otimes B$ by the equivalence $\rho = T \cup \{1\}$, where T is the set of all pairs

$$(a_1 \otimes x_1\phi \otimes s_1 \otimes b_1, a_2 \otimes x_2\phi \otimes s_2 \otimes b_2) \tag{8.5.6}$$

for which $a_1 \otimes x_1 s_1 \otimes b_1 = a_2 \otimes x_2 s_2 \otimes b_2$ in $A \otimes X \otimes B$. We do this by verifying that

$$\begin{array}{ccc} A \otimes X \otimes S \otimes B & \longrightarrow & A \otimes Y \otimes S \otimes B \\ \theta \big\downarrow & & \big\downarrow \\ A \otimes X \otimes B & \xrightarrow[\chi]{} & (A \otimes Y \otimes S \otimes B)/\rho \end{array} \qquad (8.5.7)$$

(where $(a \otimes x \otimes b)\chi = (a \otimes x\phi \otimes 1 \otimes b)\rho)$ is a pushout diagram. The commuting of the diagram presents no problems: if $a \otimes x \otimes s \otimes b \in A \otimes X \otimes S \otimes B$, then one way round the square gives $(a \otimes x\phi \otimes s \otimes b)\rho$ while the other way gives $(a \otimes (xs)\phi \otimes 1 \otimes b)\rho$, and these are evidently equal. So suppose now that we have a set Q and maps λ, μ such that the diagram

$$\begin{array}{ccc} A \otimes X \otimes S \otimes B & \longrightarrow & A \otimes Y \otimes S \otimes B \\ \theta \big\downarrow & & \big\downarrow \lambda \\ A \otimes X \otimes B & \xrightarrow[\mu]{} & Q \end{array}$$

commutes. Then T, as defined by (8.5.6), is contained in $\ker\lambda$, for if $a_1 \otimes x_1 s_1 \otimes b_1 = a_2 \otimes x_2 s_2 \otimes b_2$ in $A \otimes X \otimes B$, then

$$\begin{aligned} (a_1 \otimes x_1\phi \otimes s_1 \otimes b_1)\lambda &= (a_1 \otimes x_1 \otimes s_1 \otimes b_1)(1 \otimes \phi \otimes 1 \otimes 1)\lambda \\ &= (a_1 \otimes x_1 \otimes s_1 \otimes b_1)\theta\mu \\ &= (a_1 \otimes x_1 s_1 \otimes b_1)\mu \\ &= (a_2 \otimes x_2 s_2 \otimes b_2)\mu \\ &= \cdots = (a_2 \otimes x_2\phi \otimes s_2 \otimes b_2)\lambda. \end{aligned}$$

Hence there is a unique map $\delta : (A \otimes Y \otimes S \otimes B)/\rho \to Q$, defined by

$$[(a \otimes y \otimes s \otimes b)\rho]\delta = (a \otimes y \otimes s \otimes b)\lambda.$$

Moreover, the triangle

$$\begin{array}{ccc} A \otimes X \otimes B & \xrightarrow{\mu} & Q \\ \chi \big\downarrow & \nearrow_{\delta} & \\ (A \otimes Y \otimes S \otimes B)/\rho & & \end{array}$$

commutes, since

$$(a \otimes x \otimes b)\chi\delta = [(a \otimes x\phi \otimes 1 \otimes b)\rho]\delta$$

$$\begin{aligned}
&= (a \otimes x\phi \otimes 1 \otimes b)\lambda \\
&= (a \otimes x \otimes 1 \otimes b)(1 \otimes \phi \otimes 1 \otimes 1)\lambda \\
&= (a \otimes x \otimes 1 \otimes b)\theta\mu \\
&= (a \otimes x \otimes b)\mu.
\end{aligned}$$

Now from the fact that both (8.5.5) and (8.5.7) are pushout diagrams we conclude that the map

$$a \otimes (y \otimes s)\sigma \otimes b \mapsto (a \otimes y \otimes s \otimes b)\rho$$

is a bijection. Hence we can focus attention on the map

$$a \otimes y \otimes b \mapsto (a \otimes y \otimes 1 \otimes b)\rho;$$

to show that α is pure we require to show that this map is one-one.

Suppose therefore that

$$(a \otimes y \otimes 1 \otimes b, a' \otimes y' \otimes 1 \otimes b') \in \rho \setminus \{1\} = T.$$

Then

$$\begin{aligned}
a \otimes y \otimes 1 \otimes b &= a_1 \otimes x_1\phi \otimes s_1 \otimes b_1, \\
a' \otimes y' \otimes 1 \otimes b' &= a_2 \otimes x_2\phi \otimes s_2 \otimes b_2,
\end{aligned}$$

where $a_1 \otimes x_1 s_1 \otimes b_1 = a_2 \otimes x_2 s_2 \otimes b_2$ in $A \otimes X \otimes B$. Now if in Lemma 8.5.8 we replace X by $A \otimes X$, Y by $A \otimes Y$, λ by $1 \otimes \phi$ and Z by B we can deduce that there exist x_3 and x_4 such that

$$\begin{aligned}
a \otimes y \otimes 1 \otimes b &= a_3 \otimes x_3\phi \otimes 1 \otimes b_3, \\
a' \otimes y' \otimes 1 \otimes b' &= a_4 \otimes x_4\phi \otimes 1 \otimes b_4
\end{aligned} \tag{8.5.8}$$

in $A \otimes Y \otimes S \otimes B$. We now have

$$\begin{aligned}
a_1 \otimes x_1\phi \otimes s_1 \otimes b_1 &= a_3 \otimes x_3\phi \otimes 1 \otimes b_3, \\
a_2 \otimes x_2\phi \otimes s_2 \otimes b_2 &= a_4 \otimes x_4\phi \otimes 1 \otimes b_4,
\end{aligned}$$

from which it follows, since ϕ is by assumption pure, that

$$\begin{aligned}
a_1 \otimes x_1 \otimes s_1 \otimes b_1 &= a_3 \otimes x_3 \otimes 1 \otimes b_3, \\
a_2 \otimes x_2 \otimes s_2 \otimes b_2 &= a_4 \otimes x_4 \otimes 1 \otimes b_4.
\end{aligned}$$

Hence, by applying θ to these equations, we deduce that

$$\begin{aligned}
a_3 \otimes x_3 \otimes b_3 &= a_1 \otimes x_1 s_1 \otimes b_1 \\
&= a_2 \otimes x_2 s_2 \otimes b_2 \\
&= a_4 \otimes x_4 \otimes b_4.
\end{aligned} \tag{8.5.9}$$

Now, by the extension property of U in S, we can deduce from (8.5.8) that

$$\begin{aligned}
a \otimes y \otimes b &= a_3 \otimes x_3\phi \otimes b_3 \\
a_4 \otimes x_4\phi \otimes b_4 &= a' \otimes y \otimes b'.
\end{aligned}$$

Hence from (8.5.9) it follows finally that

$$a \otimes y \otimes b = a' \otimes y' \otimes b'.$$

Thus α is pure. □

We now state the main theorem of this section:

Theorem 8.5.10 *Let $\mathcal{A} = [U, S_1, S_2]$ be a monoid amalgam. If U has the extension property in S_1 and S_2 then $\mathcal{A}$ is embeddable. Moreover, U has the extension property in $S_1 *_U S_2$.*

Proof We construct the direct system (W_n, α_n) as in Section 8.4. Recall that $W_1 = S_1$, $W_2 = S_1 \otimes S_2$, and

$$W_n = F(S; W_{n-2}, W_{n-1}) \quad (n \geq 3).$$

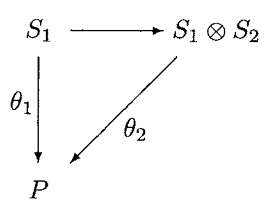
The map $\alpha_1 : W_1 \to W_2$ is one-one by the extension property of U in S_2. Moreover α_1 is pure, since if $X \in$ ENS-U and $Y \in U$-ENS the map

$$X \otimes S_1 \otimes Y \to X \otimes S_1 \otimes S_2 \otimes Y$$

is one-one, again by the extension property of U in S_2. If we now suppose inductively that $\alpha_k : W_k \to W_{k+1}$ is one-one and pure for all $k < n$, then by Lemma 8.5.9 it follows that α_n is both one-one and pure. By Proposition 8.4.1 and Theorem 8.4.5 it follows that if $P = S_1 *_U S_2$ then the maps $\theta_1 : S_1 \to P$ and $\theta_2 : S_1 \otimes S_2 \to P$ are both one-one. Since the map $S_2 \to S_1 \otimes S_2$ is one-one by the extension property of U in S_1 we deduce that the composition θ_2^* of this map with θ_2 embeds S_2 in P. In the notation developed in Section 8.4 we have

$$s_1\theta_1 = [s_1], \quad s_2\theta_2^* = [1, s_2].$$

Suppose now that $s_1\theta_1 = s_2\theta_2^*$. Then from the commuting of the diagram



we deduce that

$$(s_1 \otimes 1)\theta_2 = s_1\theta_1 = s_2\theta_2^* = (1 \otimes s_2)\theta_2.$$

Since θ_2 is one-one we deduce that $s_1 \otimes 1 = 1 \otimes s_2$ in $S_1 \otimes S_2$. Then, exactly as in the proof of Proposition 8.5.7, we deduce that $s_1 \in U$. Thus

$$1 \otimes s_1 = s_1 \otimes 1 = 1 \otimes s_2$$

in $S_1 \otimes S_2$ and so, again by the extension property of U in S_2, it follows that $s_2 = s_1 \in U$. Thus $S_1\theta_1 \cap S_2\theta_2^* = U$ as required.

To prove the final statement we must show that, for every X in ENS-U and every Y in U-ENS, the map $x \otimes y \mapsto x \otimes 1 \otimes y$ from $X \otimes Y$ into $X \otimes (S_1 *_U S_2) \otimes Y$ is one-one. Now $S_1 *_U S_2$ is the direct limit of (W_n, α_n), and so, by Proposition 8.4.3, $X \otimes (S_1 *_U S_2) \otimes Y$ is the direct limit of the direct system $(X \otimes W_n \otimes Y, 1 \otimes \alpha_n, \otimes 1)$. Since each α_n is pure and one-one, it follows that each of the maps $1 \otimes \alpha_n \otimes 1$ is one-one. Hence, by Proposition 8.4.1, the map

$$1 \otimes \theta_1 \otimes 1 : X \otimes W_1 \otimes Y \to X \otimes (S_1 *_U S_2) \otimes Y$$

is one-one. Recall now that $W_1 = S_1$, and that, by the extension property of U in S_1, the map $X \otimes Y \to X \otimes S_1 \otimes Y$ is one-one. Hence the composite map

$$X \otimes Y \to X \otimes (S_1 *_U S_2) \otimes Y$$

is one-one, and the proof is complete. □

Remark Free products with amalgamation are known to be associative (Howie 1968). The final sentence of Theorem 8.5.10 enables us to extend the result to amalgams of more than two monoids. In fact one can show that for an arbitrary index set I the amalgam $[U; \{S_i : i \in I\}]$ is embeddable if U has the extension property in each S_i.

Theorem 8.5.10 applies to monoids, but we can fairly easily get a corresponding result for semigroups. Let $\mathcal{A} = [U; S_1, S_2]$ be a semigroup amalgam in which U has the extension property in S_1 and S_2, and consider the monoid amalgam $\mathcal{A}^* = [U^*; S_1^*, S_2^*]$. By Proposition 8.5.2, U^* has the extension property in S_1^* and S_2^* and so we can embed $\mathcal{A}^*$ in its free product P. Now P contains S_1 and S_2, and from

$$\begin{aligned} U \dot{\cup} \{1\} &= U^* = S_1^* \cap S_2^* \\ &= \big(S_1 \dot{\cup} \{1\}\big) \cap \big(S_2 \dot{\cup} \{1\}\big) = (S_1 \cap S_2) \dot{\cup} \{1\} \end{aligned}$$

we deduce that $S_1 \cap S_2 = U$. Hence we have the corresponding semigroup result:

Theorem 8.5.11 *Let $\mathcal{A} = [U; S_1, S_2]$ be a semigroup amalgam. If U has the extension property in S_1 and S_2 then $\mathcal{A}$ is embeddable.* □

It is not reasonable to ask for precise converses to Theorems 8.5.10 and 8.5.11, since an amalgam $[U; S_1, S_2]$ may be embeddable for a variety of 'accidental' reasons. We can, however, obtain a partial converse if we introduce a new notion: a monoid or semigroup U is said to be an *amalgamation base* if every amalgam $[U; S_1, S_2]$ having U as core is embeddable. Then we have the result:

Theorem 8.5.12 *A monoid [semigroup] U is an amalgamation base if and only if it is absolutely extendable.*

Proof We shall prove this for monoids. By means of the usual device we can extend the proof to semigroups.

One way round this is a direct consequence of Theorem 8.5.10: if U is absolutely extendable then U has the extension property in every S_1 and S_2 containing U, and hence the amalgam $[U; S_1, S_2]$ is embeddable.

Conversely, suppose that U is an amalgamation base, and let S be a monoid containing U as a submonoid. Let $X \in \textsc{ens}\text{-}U$, $Y \in U\text{-}\textsc{ens}$; we must show that the map $X \otimes Y \to X \otimes S \otimes Y$ is one-one. We construct a monoid T containing U, and use the fact that $[U; S, T]$ is embeddable. To construct T we begin by defining $W = X \dot{\cup} Y$. Then W is a (U, U)-bisystem in an obvious way if we define the left action of U on X and the right action of U on Y trivially: if $u \in U$, $x \in X$ and $y \in Y$, then

$$ux = x, \qquad yu = y.$$

Otherwise W inherits the given right action of U on X and the given left action of U on Y.

Let

$$W^{(0)} = U, \quad W^{(1)} = W, \quad W^{(n)} = W^{(n-1)} \otimes_U W \ (n \geq 2).$$

Let T be the disjoint union of all the sets $W^{(n)}$ $(n \geq 0)$, and define a multiplication on T by

$$(w_1 \otimes \cdots \otimes w_r)(z_1 \otimes \cdots \otimes z_s) = w_1 \otimes \cdots \otimes w_r \otimes z_1 \otimes \cdots \otimes z_s,$$

(where $r, s \geq 1$) and

$$u(w_1 \otimes \cdots \otimes w_r) = (uw_1) \otimes \cdots \otimes w_r,$$
$$(w_1 \otimes \cdots \otimes w_r)u = w_1 \otimes \cdots \otimes (w_r u).$$

Then T is a monoid containing U as a submonoid.

The map $X \otimes_U Y \to W \otimes_U W = W_2 \subseteq T$ is one-one, for if $x \otimes y = x' \otimes y'$ in $W \otimes_U W$ (with $x, x' \in X$, $y, y' \in Y$) then the sequence of transitions connecting $x \otimes y$ to $x' \otimes y'$ in $W \otimes_U W$ can never involve elements of y on the left of $\otimes$ or elements of X on the right of $\otimes$, and so may equally well take place in $X \otimes_U Y$.

Now consider the pushout diagram

$$\begin{array}{ccc} U & \longrightarrow & S \\ \big\downarrow & & \big\downarrow \alpha \\ T & \underset{\beta}{\longrightarrow} & P \end{array}$$

and the map $\phi : T \otimes S \otimes T \to P$ given by

$$(t \otimes s \otimes t')\phi = (t\beta)(s\alpha)(t'\beta).$$

Suppose that $x \otimes 1 \otimes y = x' \otimes 1 \otimes y'$ in $X \otimes S \otimes Y$. Then certainly $x \otimes 1 \otimes y = x' \otimes 1 \otimes y'$ in $T \otimes S \otimes T$, and so $(x\beta)(y\beta) = (x'\beta)(y'\beta)$ in P. Now β is one-one since by assumption U is an amalgamation base. Hence $xy = x'y'$ in T. By definition this means that $x \otimes y = x' \otimes y'$ in $W \otimes W$, and we have already observed that this carries the implication that $x \otimes y = x' \otimes y'$ in $X \otimes Y$. Hence U has the extension property in S, and since S was an arbitrary monoid containing U as a submonoid, it follows that U is absolutely extendable. □

8.6 INVERSE SEMIGROUPS AND AMALGAMATION

It was shown by Howie (1962) that every group is an amalgamation base. Much more generally, we have

Theorem 8.6.1 *Every inverse monoid [semigroup] is an amalgamation base.*

Proof We shall confine ourselves to the monoid case. The semigroup case can be tackled by the usual device of adjoining an identity. Let U be an inverse monoid. We show that U is absolutely flat, from which it follows by Proposition 8.5.1 and Theorem 8.5.12 that U is an amalgamation base.

The definition of flatness was given at the beginning of Section 8.5. The first step in our proof is to establish a more easily verified criterion for flatness:

Lemma 8.6.2 *Let U be a monoid and let $X \in U$-ENS. Then X is flat if and only if, for all B in ENS-U and for all a, a' in B and x, x' in X, the equality $a \otimes x = a' \otimes x'$ in $B \otimes_U X$ implies the same equality in $(aU \cup a'U) \otimes_U X$.*

Proof It is clear from the definition of flatness that if X is flat then the inclusion $aU \cup a'U \subseteq B$ induces a monomorphism $(aU \cup a'U) \otimes_U X \to A \otimes_U X$.

Conversely, suppose that we have the given property, and consider a monomorphism $\lambda : A \to B$ of right U-sets. We must show that $\lambda \otimes 1 : A \otimes_U X \to B \otimes_U X$ is one-one. So suppose that $a, a' \in A$, $x, x' \in X$, and that (in $B \otimes X$)

$$(a\lambda) \otimes x = (a'\lambda) \otimes x'.$$

By our assumption, this equality holds also in $\big((a\lambda)U \cup (a'\lambda)U\big) \otimes X$, and so we have equations

$$\begin{aligned} a\lambda &= b_1u_1 & u_1x &= v_1x_1 \\ b_1v_1 &= b_2u_2 & u_2x_1 &= v_2x_2 \\ b_2v_2 &= b_3u_3 & u_3x_2 &= v_3x_3 \\ &\cdots & &\cdots \\ b_{n-1}v_{n-1} &= b_nu_n & u_nx_{n-1} &= v_nx' \\ b_nv_n &= a'\lambda \end{aligned} \qquad (8.6.1)$$

in which $u_1, \ldots, u_n, v_1, \ldots, v_n \in U$, $x_1, \ldots, x_{n-1} \in X$, and in which each of $b_1, \ldots, b_n$ is either in $(a\lambda)U$ or in $(a'\lambda)U$. Now any element $b = (a\lambda)u$ in $(a\lambda)U$ can be expressed as $(au)\lambda$, where $au \in A$. A similar observation applies to elements of $(a'\lambda)U$, and so there exist elements a_i in A such that $a_i\lambda = b_i$ for $i = 1, 2, \ldots, n$. Since λ is a monomorphism of right U-systems, we thus have a sequence of equalities

$$a = a_1u_1,\ a_1v_1 = a_2u_2,\ a_2v_2 = a_3u_3, \ldots, a_nv_n = a',$$

and these, together with the right hand column of equalities in (8.6.1), imply that $a \otimes x = a' \otimes x'$ in $A \otimes X$. □

Returning to the proof of Theorem 8.6.1, consider an inverse monoid U and a left U-system X. We aim to show that X is flat. Accordingly, let $B \in$ ENS-U and suppose that $a \otimes x = a' \otimes x'$ in $B \otimes_U X$. Then without loss of generality we may suppose that there is a series of equalities

$$\begin{array}{rclrcl} a &=& b_1u_1 & u_1x &=& v_1x_1 \\ b_1v_1 &=& b_2u_2 & u_2x_1 &=& v_2x_2 \\ b_2v_2 &=& b_3u_3 & u_3x_2 &=& v_3x_3 \\ & \cdots & & & \cdots & \\ b_{2n-1}v_{2n-1} &=& b_{2n}u_{2n} & u_{2n}x_{2n-1} &=& v_{2n}x' \\ b_{2n}v_{2n} &=& a' & & & \end{array}$$

where $b_1, \ldots, b_{2n} \in B$, $x_1, \ldots, x_{2n-1} \in X$, and $u_1, \ldots, u_{2n}, v_1, \ldots, v_{2n} \in U$. There is no real loss of generality in supposing that the number of us is even, for a scheme of this kind can always be extended trivially with equations in which u and v are equal to 1. We aim to show that there is a sequence of equalities demonstrating that $a \otimes x = a' \otimes x'$ within $(aU \cup a'U) \otimes X$. The sequence will be three times as long as the given one, and to describe it we require some preliminaries. Define

$$\begin{array}{ll} z_0 = 1, & z_i = z_{i-1}u_i^{-1}v_i\ (i = 1, \ldots, 2n), \\ t_0 = 1, & t_i = t_{i-1}v_{2n-i+1}^{-1}u_{2n-i+1}\ (i = 1, \ldots, 2n). \end{array}$$

Thus

$$\begin{aligned} z_i &= u_1^{-1}v_1u_2^{-1}v_2 \ldots u_i^{-1}v_i \quad (i \geq 1), \\ t_i &= v_{2n}^{-1}u_{2n}v_{2n-1}^{-1}u_{2n-1} \ldots v_{2n-i+1}^{-1}u_{2n-i+1} \quad (i \geq 1). \end{aligned}$$

We now gather together some useful equalities into the following lemma.

Lemma 8.6.3 *With the above definitions:*

(1) $z_{2n-i}t_i^{-1} = z_{2n} = t_{2n}^{-1}$;
(2) $az_i = b_iv_iz_i^{-1}z_i \quad (1 \leq i \leq 2n)$;
(3) $a't_i = b_{2n-i+1}u_{2n-i+1}t_i^{-1}t_i \quad (1 \leq i \leq 2n)$;
(4) $a = au_1^{-1}u_1$;
(5) $(az_{i-1}u_i^{-1})v_i = (az_iu_{i+1}^{-1})u_{i+1} \quad (1 \leq i \leq 2n-1)$;
(6) $(az_{2n-1}u_{2n}^{-1})v_{2n} = (az_{2n}v_{2n}^{-1})v_{2n}$;

(7) *for* $1 \le i \le n-1$,

$$(az_{2n}t_{i-1}v_{2n-i+1}^{-1})u_{2n-i+1} = (az_{2n}t_iv_{2n-i}^{-1})v_{2n-i};$$

(8) $(az_{2n}t_{n-1}v_{n+1}^{-1})u_{n+1} = (a't_{2n}z_{n-1}u_n^{-1})v_n;$

(9) $(a't_{2n}z_iu_{i+1}^{-1})u_{i+1} = (a't_{2n}z_{i-1}u_i^{-1})v_i \quad (1 \le i \le n-1);$

(10) $(a't_{2n}u_1^{-1})u_1 = (a't_{2n-1}v_1^{-1})u_1;$

(11) $(a't_{2n-i-1}v_{i+1}^{-1})u_{i+1} = (a't_{2n-i}v_i^{-1})v_i \quad (1 \le i \le 2n-1);$

(12) $a'v_{2n}^{-1}v_{2n} = a'$.

Proof (1) This is clear from the definitions.

(2) We prove this by induction. First, notice that

$$az_1 = b_1u_1u_1^{-1}v_1 = b_1v_1v_1^{-1}u_1u_1^{-1}v_1 = b_1v_1z_1^{-1}z_1.$$

Then, inductively, we have

$$\begin{aligned} az_i &= az_{i-1}u_i^{-1}v_i = b_{i-1}v_{i-1}z_{i-1}^{-1}z_{i-1}u_i^{-1}v_i \\ &= b_i(u_iz_{i-1}^{-1}z_{i-1}u_i^{-1})(v_iv_i^{-1})v_i = b_iv_iv_i^{-1}u_iz_{i-1}^{-1}z_{i-1}u_i^{-1}v_i \\ &= b_iv_iz_i^{-1}z_i. \end{aligned}$$

(3) The proof is similar to that of (2).

(4) This is immediate.

(5) Using standard techniques involving commuting idempotents, we have

$$\begin{aligned} (az_{i-1}u_i^{-1})v_i &= az_i = b_iv_iz_i^{-1}z_i \text{ (by (2))} \\ &= b_{i+1}u_{i+1}z_i^{-1}z_i = b_{i+1}u_{i+1}z_i^{-1}z_iu_{i+1}^{-1}u_{i+1} \\ &= b_iv_iz_i^{-1}z_iu_{i+1}^{-1}u_{i+1} = (az_iu_{i+1}^{-1})u_{i+1}. \end{aligned}$$

(6) We have

$$(az_{2n-1}u_{2n}^{-1})v_{2n} = az_{2n-1}u_{2n}^{-1}v_{2n}v_{2n}^{-1}v_{2n} = az_{2n}v_{2n}^{-1}v_{2n}.$$

(7) For $1 \le i \le n-1$ we have

$$\begin{aligned} az_{2n}t_{i-1}v_{2n-i+1}^{-1}u_{2n-i+1} &= az_{2n}t_i = az_{2n-i}t_i^{-1}t_i \text{ (by (1))} \\ &= b_{2n-i}v_{2n-i}z_{2n-i}^{-1}z_{2n-i}t_i^{-1}t_i \text{ (by (2))} \\ &= b_{2n-i}v_{2n-i}z_{2n-i}^{-1}z_{2n-i}t_i^{-1}t_iv_{2n-i}^{-1}v_{2n-i} \\ &= (az_{2n}t_iv_{2n-i}^{-1})v_{2n-i}. \end{aligned}$$

(8) Using (1) and (2) above, we have

$$\begin{aligned} (az_{2n}t_{n-1}v_{n+1}^{-1})u_{n+1} &= az_{2n}t_n = b_{2n}v_{2n}z_{2n}^{-1}z_{2n}t_n \\ &= a'z_{2n}^{-1}z_{2n}t_n = a't_nz_n^{-1}z_nt_n^{-1}t_n = a't_nz_n^{-1}z_n \\ &= a't_{2n}z_n = (a't_{2n}z_{n-1}u_n^{-1})v_n. \end{aligned}$$

(9) Using (1) and (3) above, we have (for $1 \leq i \leq n-1$)

$$\begin{aligned}(a't_{2n}z_iu_{i+1}^{-1})u_{i+1} &= a't_{2n-i}z_i^{-1}z_iu_{i+1}^{-1}u_{i+1}\\ &= b_{i+1}u_{i+1}t_{2n-i}^{-1}t_{2n-i}z_i^{-1}z_iu_{i+1}^{-1}u_{i+1}\\ &= b_{i+1}u_{i+1}t_{2n-i}^{-1}t_{2n-i}z_i^{-1}z_i\\ &= a't_{2n}z_i = (a't_{2n}z_{i-1}u_i^{-1})v_i.\end{aligned}$$

(10) Since $t_{2n}u_1^{-1}u_1 = t_{2n}$, we have

$$(a't_{2n}u_1^{-1})u_1 = a't_{2n} = (a't_{2n-1}v_1^{-1})u_1.$$

(11) Using (3) above, we have

$$\begin{aligned}(a't_{2n-i}v_i^{-1})v_i &= b_{i+1}u_{i+1}t_{2n-i}^{-1}t_{2n-i}v_i^{-1}v_i = b_iv_it_{2n-i}^{-1}t_{2n-i}v_i^{-1}v_i\\ &= b_iv_it_{2n-i}^{-1}t_{2n-i} = a't_{2n-i} = (a't_{2n-i-1}v_{i+1}^{-1})u_{i+1}.\end{aligned}$$

(12) This is immediate. □

With this information we can now fairly quickly show that $a \otimes x = a' \otimes x'$ in $(aU \cup a'U) \otimes X$. For we have a system of equations

$$\begin{array}{rclcrcl}
a &=& (au_1^{-1})u_1 & & u_1x &=& v_1x_1\\
(au_1^{-1})v_1 &=& (az_1u_2^{-1})u_2 & & u_2x_1 &=& v_2x_2\\
(az_1u_2^{-1})v_2 &=& (az_2u_3^{-1})u_3 & & u_3x_2 &=& v_3x_3\\
& \cdots & & & & \cdots &\\
(az_{2n-2}u_{2n-1}^{-1})v_{2n-1} &=& (az_{2n-1}u_{2n}^{-1})u_{2n} & & u_{2n}x_{2n-1} &=& v_{2n}x'\\
(az_{2n-1}u_{2n}^{-1})v_{2n} &=& (az_{2n}v_{2n}^{-1})v_{2n} & & v_{2n}x' &=& u_{2n}x_{2n-1}\\
(az_{2n}v_{2n}^{-1})u_{2n} &=& (az_{2n}t_1v_{2n-1}^{-1})v_{2n-1} & & v_{2n-1}x_{2n-2} &=& u_{2n-1}x_{2n-2}\\
(az_{2n}t_1v_{2n-1}^{-1})u_{2n-1} &=& (az_{2n}t_2v_{2n-2}^{-1})v_{2n-2} & & v_{2n-2}x_{2n-2} &=& u_{2n-2}x_{2n-3}\\
& \cdots & & & & \cdots &\\
(az_{2n}t_{n-1}v_{n+1}^{-1})u_{n+1} &=& (a't_{2n}z_{n-1}u_n^{-1})v_n & & v_nx_n &=& u_nx_{n-1}\\
(a't_{2n}z_{n-1}u_n^{-1})u_n &=& (a't_{2n}z_{n-2}u_{n-1}^{-1})v_{n-1} & & v_{n-1}x_{n-1} &=& u_{n-1}x_{n-2}\\
& \cdots & & & & \cdots &\\
(a't_{2n}z_1u_2^{-1})u_2 &=& (a't_{2n}z_0u_1^{-1})v_1 & & v_1x_1 &=& u_1x\\
(a't_{2n}z_0u_1^{-1})u_1 &=& (a't_{2n-1}v_1^{-1})u_1 & & u_1x &=& v_1x_1\\
(a't_{2n-1}v_1^{-1})v_1 &=& (a't_{2n-2}v_2^{-1})u_2 & & u_2x_1 &=& v_2x_2\\
(a't_{2n-2}v_2^{-1})v_2 &=& (a't_{2n-3}v_3^{-1})u_3 & & u_3x_2 &=& v_3x_3\\
& \cdots & & & & \cdots &\\
(a't_1v_{2n-1}^{-1})v_{2n-1} &=& (a't_0v_{2n}^{-1})u_{2n} & & u_{2n}x_{2n-1} &=& v_{2n}x'\\
(a't_0v_{2n}^{-1})v_{2n} &=& a'. & & & &
\end{array}$$

Since all the left factors of the left hand column of equations are visibly in $aU \cup a'U$ it follows that $a \otimes x = a' \otimes x'$ in $(aU \cup a'U) \otimes_U X$. Thus X is flat, and since X was an arbitrary left U-system we deduce that all left U-systems are flat. A similar argument shows that all right U-systems are flat, and hence U, an arbitrary inverse monoid, is absolutely flat. □

By an *inverse semigroup amalgam* we mean a semigroup amalgam

$$\mathcal{A} = [U; \{S_i : i \in I\}; \{\phi_i : i \in I\}]$$

in which U and all of the semigroups S_i are inverse semigroups. From Theorem 8.6.1 we know that every inverse semigroup amalgam $\mathcal{A}$ is embeddable in a semigroup V. It is, however, more natural to seek to embed $\mathcal{A}$ in an *inverse semigroup*. Let us say that a class $\mathbf{C}$ of semigroups *has the amalgamation property* if every amalgam $[U; \{S_i : i \in I\}; \{\phi_i : i \in I\}]$ in which U and S_i $(i \in I)$ all belong to $\mathbf{C}$ is embeddable in a semigroup from the class $\mathbf{C}$. Then we have

Theorem 8.6.4 *The class of inverse semigroups has the amalgamation property.*

Proof It will be sufficient to consider an inverse semigroup amalgam $\mathcal{A} = [U; S, T]$, and by Theorem 8.6.1 we may suppose that S and T are subsemigroups of a semigroup V intersecting in U.

For each s in S, let λ_s be the map with domain $s^{-1}V$ defined by

$$x\lambda_s = sx \quad (x \in s^{-1}V).$$

The image of λ_s is clearly contained in sV. In fact $\operatorname{im}\lambda_s = sV$, since, for all sv in sV,

$$sv = ss^{-1}sv = (s^{-1}sv)\lambda_s.$$

The map $\lambda_s : s^{-1}V \to sV$ is one-one, since, for all $x_1 = s^{-1}y_1$ and $x_2 = s^{-1}y_2$ in $s^{-1}V$,

$$\begin{aligned} x_1\lambda_s &= x_2\lambda_s \Rightarrow ss^{-1}y_1 = ss^{-1}y_2 \\ &\Rightarrow x_1 = s^{-1}y_1 = s^{-1}ss^{-1}y_1 = s^{-1}ss^{-1}y_2 = x_2. \end{aligned}$$

Thus $\lambda_s \in \mathcal{I}_V$.

Next, we show that the map $s \mapsto \lambda_s$ is one-one. Let $s, t \in S$ and suppose that $\lambda_s = \lambda_t$. Then, making use of Lemma 5.1.6, we see that

$$s^{-1}sV = s^{-1}V = \operatorname{dom}\lambda_s = \operatorname{dom}\lambda_t = t^{-1}tV,$$

and so $s^{-1}s = t^{-1}t$. Hence

$$s = ss^{-1}s = (s^{-1}s)\lambda_s = (s^{-1}s)\lambda_t = (t^{-1}t)\lambda_t = tt^{-1}t = t.$$

It is easy to see that the inverse of the bijection $\lambda_s : s^{-1}V \to sV$ is $\lambda_{s^{-1}}$, since for all $x = s^{-1}y$ in $\operatorname{dom}\lambda_s$,

$$x\lambda_s\lambda_{s^{-1}} = s^{-1}ss^{-1}y = s^{-1}y = x,$$

and for all $u = sv$ in $\operatorname{dom}\lambda_{s^{-1}}$

$$u\lambda_{s^{-1}}\lambda_s = ss^{-1}sv = sv = u.$$

The map $s \mapsto \lambda_s$ is a morphism, not into $\mathcal{I}_V$, but into the dual semigroup $\mathcal{I}_V^*$. In other words, for all s, t in S

$$\lambda_s \lambda_t = \lambda_{ts}.$$

To see this, notice (again using Lemma 5.1.6) that

$$\begin{aligned}\operatorname{dom}(\lambda_s\lambda_t) &= (sV \cap t^{-1}V)\lambda_{s^{-1}} \\ &= (ss^{-1}V \cap t^{-1}tV)\lambda_{s^{-1}} = (t^{-1}tss^{-1}V)\lambda_{s^{-1}} \\ &= s^{-1}t^{-1}tss^{-1}V = s^{-1}t^{-1}tsV = (ts)^{-1}(ts)V = \operatorname{dom}\lambda_{ts}.\end{aligned}$$

Also, for all x in $\operatorname{dom}(\lambda_s\lambda_t)$,

$$x\lambda_s\lambda_t = (sx)\lambda_t = t(sx) = (ts)x = x\lambda_{ts}.$$

An exactly dual argument shows that for each s in S there is an element ρ_s of $\mathcal{I}_V$ with domain Vs^{-1} and image Vs, given by

$$x\rho_s = xs \quad (x \in Vs^{-1}),$$

and that $s \mapsto \rho_s$ is a monomorphism from S into $\mathcal{I}_V$. (No reversal of the multiplication is required in this case.)

We thus have a monomorphism ϕ from S into the inverse semigroup $\mathcal{I}_V^* \times \mathcal{I}_V$, given by

$$s\phi = (\lambda_s, \rho_s) \quad (s \in S).$$

By the same token we have a monomorphism $\psi : T \to \mathcal{I}_V^* \times \mathcal{I}_V$, given by

$$t\psi = (\lambda_t, \rho_t) \quad (t \in T).$$

It is clear that $u\phi = u\psi$ for all u in U, and so

$$S\phi \cap T\psi \supseteq U\phi.$$

To complete the proof of the theorem we need to show that this containment is in fact an equality.

Suppose, therefore, that $s\phi = t\psi$, where $s \in S$ and $t \in T$. Then $\lambda_s = \lambda_t$, and so

$$ss^{-1}V = \operatorname{im}\lambda_s = \operatorname{im}\lambda_t = tt^{-1}V;$$

hence $ss^{-1} \mathcal{R}\, tt^{-1}$ in V. Also, $\rho_s = \rho_t$, and so

$$Vss^{-1} = \operatorname{dom}\rho_s = \operatorname{dom}\rho_t = Vtt^{-1};$$

thus $ss^{-1} \mathcal{L}\, tt^{-1}$ in V. Since $\mathcal{H}$ separates idempotents in any semigroup (Corollary 2.2.6), it follows that $ss^{-1} = tt^{-1}$, and from this we easily deduce that

$$s = ss^{-1}s = (ss^{-1})\rho_s = (tt^{-1})\rho_t = t.$$

Since $S \cap T = U$, this implies that $s = t \in U$, exactly as required. □

It follows from our method of proof that if an inverse semigroup amalgam is embeddable in a finite semigroup then it is embeddable in a finite inverse semigroup. Given the result from group theory (see B. H. Neumann (1954)) that a finite group amalgam $[U; S, T]$ is embeddable in a finite group, it is reasonable to speculate that a finite inverse semigroup amalgam is embeddable in a finite inverse semigroup. The following example shows that this is not true in general.

Example 8.6.5 Let $S = \{0, e, f, a, b\}$, $T = \{0, e, f, g, x, y\}$ be the semigroups whose multiplication tables are

	0	e	f	a	b
0	0	0	0	0	0
e	0	e	0	a	0
f	0	0	f	0	b
a	0	0	a	0	e
b	0	b	0	f	0

	0	e	f	g	x	y
0	0	0	0	0	0	0
e	0	e	0	0	x	0
f	0	0	f	g	0	y
g	0	0	g	g	0	y
x	0	0	x	x	0	e
y	0	y	0	0	g	0

The first of these semigroups is the Brandt semigroup B_2 already encountered in Section 1.6. The second can be seen to be a semigroup if we represent it by matrices

$$0 = \begin{bmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&0 \end{bmatrix}, e = \begin{bmatrix} 1&0&0 \\ 0&0&0 \\ 0&0&0 \end{bmatrix}, f = \begin{bmatrix} 0&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix},$$

$$g = \begin{bmatrix} 0&0&0 \\ 0&1&0 \\ 0&0&0 \end{bmatrix}, x = \begin{bmatrix} 0&1&0 \\ 0&0&0 \\ 0&0&0 \end{bmatrix}, y = \begin{bmatrix} 0&0&0 \\ 1&0&0 \\ 0&0&0 \end{bmatrix}.$$

In fact both S and T are inverse semigroups, and $U = \{0, e, f\}$ is a common inverse subsemigroup.

Suppose now that the inverse semigroup amalgam $[U; S, T]$ is embedded in a semigroup Q. Thus Q contains subsemigroups S and T intersecting in U. Since $g < f$ in T we certainly have $g < f$ in Q. On the other hand, in Q we have

$$g = gfg, \quad f = ba = bea = bxya = bxgya,$$

and so $g \mathcal{J} f$ in Q. The principal factor $J_f \cup \{0\}$ of Q is certainly not null, since it contains the idempotent f. Hence, by Theorem 3.1.6, it is 0-simple. If it were finite then it would be *completely* 0-simple, and all its idempotents would be primitive. Since $g < f$ in $J_f \cup \{0\}$ this is not so; hence $J_f \cup \{0\}$, and so *a fortiori* Q, is infinite.

This example of a finite semigroup amalgam that is embeddable but not finitely embeddable incidentally answers negatively a question in Howie (1964a).

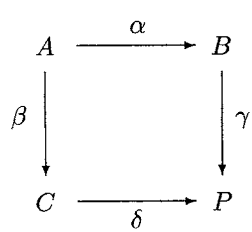
8.7 EXERCISES

1. Let X be a right S-system. Define a relation ρ on S by the rule that
$$a \,\rho\, b \text{ if and only if } (\forall x \in X)\ xa = xb.$$
Show that ρ is a congruence on S. Define a map $\phi : S \to \mathcal{T}_X$ by
$$x(a\phi) = xa \quad (x \in X,\ a \in S).$$
Show that ϕ is a morphism whose kernel is the congruence λ.

2. Let $X = \{x_1, x_2, \ldots, x_n\}$, and let U be a monoid. Let F be the cartesian product $X \times U$. Verify that F becomes a right U-system if we define, for (x, u) in F and v in U,
$$(x, u)v = (x, uv).$$
Show that F is the *free* right U-system on the set X. More precisely, show that there is a map $\theta : X \to F$ such that for every right U-system A and every map $\phi : X \to A$ there is a unique right U-map $\psi : F \to A$ such that $\theta\psi = \phi$.
Notice that $(x\theta)u = (x, u)$ for all x in X and u in U. If we identify $x\theta$ with x, we may regard F as the disjoint union
$$x_1U \,\dot\cup\, x_2U \,\dot\cup\, \ldots \,\dot\cup\, x_nU.$$

3. Let S be a commutative monoid, let X, Y be S-systems and let $\mathrm{Mor}(X, Y)$ be the set of all morphisms from X into Y.
   (a) Show that $\mathrm{Mor}(X, Y)$ becomes an S-system if, for every morphism $\alpha : X \to Y$ and every s in S the morphism αs is defined by
   $$x(\alpha s) = (x\alpha)s \quad (x \in X).$$
   (b) Show that, if X, Y and Z are S-systems then
   $$\mathrm{Mor}(X \otimes_S Y, Z) \simeq \mathrm{Mor}\big(X, \mathrm{Mor}(Y, Z)\big).$$
This is connected with the idea of an 'adjoint functor'. See, for example, Schubert (1972).

4. Let



be a pushout diagram in S-ENS. Show that $P = (B \dot{\cup} C)/\sigma$, where $\sigma = \{(a\alpha, a\beta : a \in A\}^{\#}$. Show that if α is one-one then δ is one-one, and that if α is onto then δ is onto. Show that in the latter case $P = C/\rho$, where $\rho = \{(a\beta, a'\beta) : a\alpha = a'\alpha\}^{\#}$.

5. Show that $U \mapsto \mathrm{Dom}_S(U)$ is a closure operation on the set of subsemigroups of a semigroup S. That is, show that

$$U \subseteq \mathrm{Dom}_S(U),$$
$$U \subseteq V \Rightarrow \mathrm{Dom}_S(U) \subseteq \mathrm{Dom}_S(V),$$
$$\mathrm{Dom}_S\big(\mathrm{Dom}_S(U)\big) = \mathrm{Dom}_S(U).$$

6. Let $U = a^+$, an infinite monogenic semigroup.
   (a) If S is the infinite cyclic group generated by a, show that
   $$U \subset \mathrm{Dom}_S(U) = S.$$
   (b) If S is the direct product $a^+ \times b^*$ of $U = a^+$ with the free monoid b^*, show that
   $$U = \mathrm{Dom}_S(U) \subset S.$$
   (c) If P and Q are infinite cyclic groups generated by a, b, respectively, and if S is the 0-direct union of P^0 and Q^0, show that
   $$U \subset \mathrm{Dom}_S(U) \subset S.$$

7. Show that left simple semigroups (semigroups S in which $\mathcal{L} = S \times S$) are absolutely closed. [Hint: show that any zigzag in any T over S can be replaced by a left-inner zigzag.]

8. Let S be a semigroup in which there exist elements a_1, a_2, a_3 for which
$$a_1S \cap a_2S = Sa_2 \cap Sa_3 = \emptyset.$$
(This is a property possessed in particular by the 2×2 rectangular band, with $a_1 = (1,1)$, $a_2 = (2,1)$ and $a_3 = (1,2)$.) Let $F = \{x_1, x_2, y_1, y_2\}^+$, and let $T = (S * F)/\mathbf{R}^{\#}$, where
$$\mathbf{R} = \{(a_1, x_1a_1),\ (a_1y_1, a_2y_2),\ (x_1a_2, x_2a_3),\ (a_3y_2, a_3)\}.$$
   (a) Show by induction that any element of $S * F$ obtained from an element s of S by elementary $\mathbf{R}$-transitions is of the form
   $$w_1z_1w_2z_2 \dots w_{n-1}z_{n-1}w_n,$$
   where:
   (i) $w_1, w_2, \dots, w_n \in S^1$;
   (ii) each of $z_1, z_2, \dots, z_{n-1}$ is either x_1 or y_2;
   (iii) if $z_i = x_1$ and $w_{i+1} \neq 1$ then $w_{i+1} \in a_1S$;

(iv) if $z_i = y_2$ and $w_i \neq 1$ then $w_i \in Sa_3$;
(v) $s = w_1 w_2 \dots w_n$.

(b) Deduce that $\mathbf{R}^\# \cap (S \times S) = 1_S$, and hence that T contains S up to isomorphism.

(c) Show that the element $d = (a_1 y_1)\mathbf{R}^\#$ of T is such that $d \in S$, $d \in \mathrm{Dom}_T(S)$, and deduce that S is not absolutely closed.

(This example shows that Theorem 8.3.6 cannot be generalized to regular semigroups. Indeed it fails even for rectangular bands.)

9. Within the full transformation semigroup $T = \mathcal{T}_{\{1,2,3,4,5,6\}}$, write $(x_1\, x_2\, x_3\, x_4\, x_5\, x_6)$ to denote the map

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{pmatrix}.$$

Let $u_1 = (2\,2\,3\,4\,5\,6)$, $u_2 = (1\,1\,3\,4\,4\,6)$, $u_3 = (1\,1\,1\,4\,4\,4)$, $u_4 = (2\,2\,4\,5\,5\,5)$, $x = (3\,3\,3\,6\,6\,6)$, $y = (3\,2\,3\,6\,5\,6)$, $z = (3\,1\,3\,6\,4\,6)$. Show that $S = \{u_1, u_2, u_3, u_4, x, y, z\}$ is a subsemigroup of T, and that $U = \{u_1, u_2, u_3, u_4\}$ is a right normal sub-band of S. Show that

$$\begin{aligned} x &= xu_1 & u_1 &= u_1 y \\ xu_1 &= xu_2 & u_2 y &= u_3 y \\ xu_3 &= u_3 & u_3 y &= x \end{aligned}$$

is a zigzag over U with value x, and deduce that $x \in \mathrm{Dom}_S(U)$.

10. Within the same full transformation semigroup T as in the previous exercise, show that the 14 elements

$$\begin{aligned} &u_0 = (4\,4\,6\,4\,6\,6), \quad u_1 = (2\,2\,2\,4\,5\,6), \quad u_2 = (2\,2\,3\,4\,3\,6), \\ &u_3 = (3\,2\,2\,4\,2\,6), \quad u_4 = (3\,3\,3\,3\,3\,3), \quad u_5 = (5\,2\,5\,4\,5\,6), \\ &u_6 = (6\,4\,6\,4\,6\,6), \quad u_7 = (5\,5\,5\,5\,5\,5), \quad u_8 = (6\,6\,6\,6\,6\,6), \\ &x = (1\,1\,1\,1\,1), \quad y = (4\,4\,4\,4\,6\,6), \quad d = (4\,4\,4\,4\,4\,4), \\ &z = (4\,4\,4\,4\,4\,6), \quad t = (2\,2\,2\,2\,2\,2), \end{aligned}$$

form a sub-band S of T, and that $U = \{u_0, u_2, \dots, u_8\}$ is a sub-band of S. Show that

$$\begin{aligned} d &= xu_0 & u_0 &= u_1 y \\ xu_1 &= xu_2 & u_2 y &= u_3 y \\ xu_3 &= u_4 & u_4 y &= d \end{aligned}$$

is a zigzag over U with value d, and deduce that U is not closed in S.

11. Let B be a band, let U be a proper sub-band of B, and suppose that $\mathrm{Dom}_B(U) = B$. Let $d \in B \setminus U$, and suppose that

$$\begin{aligned} d &= x_1u_1 & u_1 &= v_1y_1 \\ x_1v_1 &= x_2u_2 & u_2y_1 &= v_2y_2 \\ &\dots & &\dots \\ x_{m-1}v_{m-1} &= u_m & u_my_{m-1} &= d \end{aligned}$$

is a zigzag of minimum length over U with value d. Thus y_1 and x_{m-1} belong to $B \setminus U$, since otherwise the zigzag may be shortened.

(a) Using the fact that B is a band, show that $d = du_1 = u_md$.

(b) Show that

$$J_d \le J_{u_1} \le J_{y_1}, \quad J_d \le J_{u_m} \le J_{x_{m-1}}.$$

(c) Show that if $J_{y_1} = J_d$ and $J_{x_{m-1}} = J_d$ then $d = u_1du_m = u_1u_m \in U$.

(d) Deduce that B contains an infinite sequence of elements $d_1, d_2, d_3, \dots$ (where d_1 is either y_1 or y_{m-1}) such that

$$J_d < J_{d_1} < J_{d_2} < J_{d_3} < \cdots$$

(e) Deduce that epimorphisms are onto within the class of finite bands.

12. Let D be the infinite null semigroup $\{0, u_1, u_2, u_3, \dots\}$. Let $S = D \cup a^*$, and define

$$a0 = 0a = 0, \quad u_ia^j = a^ju_i = u_{i+j}.$$

Show that S is a commutative semigroup, and that $U = \{0, u_1, u_2\}$ is a subsemigroup of S. Write down a zigzag in S over U with value u_i, and deduce that $\mathrm{Dom}_S(U) = D$.

13. Let G be the infinite cyclic group generated by a, and with identity element e. Let S be the Brandt semigroup $B(G, 2)$, and let

$$U = \{0, i, j, x, y\},$$

where $i = (1, e, 1)$, $j = (2, e, 2)$, $x = (1, a, 2)$ and $y = (1, e, 2)$. Show that $x^{-1}, y^{-1} \in \mathrm{Dom}_S(U)$, and deduce that U is epimorphically embedded in S.

14. Let $U = \{u, v, 0\}$ be a three element null semigroup. Let $S = U \cup \{a\}$, with $au = ua = v$ and all other products equal to 0. Let $T = U \cup \{b\}$, with $bv = vb = u$ and all other products equal to 0. Show that S and T are semigroups. If P is a semigroup containing the amalgam $[U; S, T]$, show that $u = 0$ in P, and deduce that $[U; S, T]$ is not embeddable.

15. Let $[U; S, T]$ be an amalgam. Let

$$N = U \cup (S \setminus U) \cup (T \setminus U) \cup \{0\},$$

where $0 \notin S \cup T$, and define a binary operation $\circ$ on N by

$$x \circ y = \begin{cases} xy & \text{if } x,\ y \text{ are both in } S \text{ or both in } T \\ 0 & \text{otherwise.} \end{cases}$$

Show that $(N, \circ)$ is a semigroup if and only if N is *consistent* in S and in T, that is, if and only if, in both S and T,

$$xy \in U \;\Rightarrow\; x \in U \text{ and } y \in U.$$

Deduce that a (finite) amalgam $[U; S, T]$ in which U is consistent in both S and T is embeddable in a (finite) semigroup.

16. Show that a subsemigroup of a group is a subgroup if and only if it is a (left) unitary subsemigroup.

17. A subsemigroup U of a semigroup S is said to be *almost unitary* if there exist maps $\lambda, \rho : S \to S$ such that

    (a) $\lambda^2 = \lambda$, $\rho^2 = \rho$;
    (b) $\lambda(st) = (\lambda s)t$, $(st)\rho = s(t\rho)$ for all s, t in S; that is, λ is a *left translation* and ρ is a *right translation*;
    (c) $\lambda(s\rho) = (\lambda s)\rho$ for all s in S; that is, λ and ρ *commute*, and the notation $\lambda s\rho$ is unambiguous;
    (d) $s(\lambda t) = (s\rho)t$ for all s, t in S; that is, λ and ρ are *linked*;
    (e) $\lambda|_U = \rho|_U = 1_U$;
    (f) U is unitary in $\lambda S\rho$.

    Show that if U has an identity element e, then U is almost unitary if and only if U is unitary in eSe.

18. It is clear that

    unitary $\Rightarrow$ almost unitary $\Rightarrow$ quasi-unitary $\Rightarrow$ relatively unitary.

    All the implications are proper:

    (a) Show that every sub*group* of a semigroup is almost unitary. Let S be a semigroup with zero element 0. Show that $\{0\}$ is almost unitary but not unitary.
    (b) Let S be a Clifford semigroup with semilattice E of idempotents. Show that E is quasi-unitary in S. Suppose now that S has an identity and that S is not E-unitary. Deduce that E is not almost unitary.
    (c) Let U be a subsemigroup of a semigroup S, and suppose that U is an inverse semigroup. Show that U is relatively unitary. Now consider the subsemigroup $\mathcal{I}_X$ of $\mathcal{PT}_X$, where $|X| \geq 3$. Show that $\mathcal{I}_X$ is not quasi-unitary in $\mathcal{PT}_X$. More particularly,

consider $X = \{1,2,3\}$, and suppose that $\mathcal{I}_X$ is quasi-unitary in $\mathcal{PT}_X$, with associated map ϕ. Consider the elements

$$\alpha = \begin{pmatrix} 1\,2 \\ 1\,2 \end{pmatrix}, \quad \beta = \begin{pmatrix} 2\,3 \\ 2\,3 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 1\,2\,3 \\ 2\,3\,2 \end{pmatrix},$$

note that $\alpha\gamma$ and $\beta\gamma$ belong to $\mathcal{I}_X$, and attempt to find $\gamma\phi$.

19. Let U be a submonoid of a monoid S. Show that if the inclusion map $\iota : U \to S$ *splits* in the category U-ENS-U (by which is meant that there exists a (U,U)-map $\phi : S \to U$ such that $\iota\phi = 1_U$), then U is quasi-unitary in S.

20. Let $[U; \{S_i : i \in I\}]$ be an amalgam of monoids, let F be a finite subset of I and let P_F be the free product of the amalgam $[U; \{S_i : i \in F\}]$. It is clear that if F, G are finite subsets of I such that $F \subseteq G$, then there is a canonical map $\phi_{FG} : P_F \to P_G$ such that (P_F, ϕ_{FG}) is a direct system of monoids and monoid morphisms. Show that $\Pi^*\{S_i : i \in I\}$ is the direct limit in the category of monoids of this system.

21. It is known (Neumann 1954) that if

$$\mathcal{A} = [U; S, T]$$

is a group amalgam in which $U \subset S$ and $U \subset T$ (properly), then the free product of $\mathcal{A}$ is infinite. Show that this is not the case for semigroups. More precisely, let S be a finite semigroup with identity element 1, let T be a finite semigroup with zero element 0, and form the free product amalgamating the identity of S with the 0 of T. Show that the free product is a finite semigroup of order $|S|+|T|-1$. (Notice that the amalgamated subsemigroup is a group, and so the amalgam is naturally embedded in its free product.)

8.8 NOTES

We have given the merest sketch of the theory of S-systems. For further information see (for example) Berthiaume (1967), Feller and Gantos (1969), Knauer (1972, 1983), Kilp (1970, 1972, 1985), Kilp and Knauer (1982, 1986), Knauer and Petrich (1981), Fleischer (1982), Normak (1987), Bulman-Fleming and McDowell (1980a,b, 1983, 1984, 1985, 1987, 1988), Bulman-Fleming and Gould (1990), and Bulman-Fleming *et al.* (1990). The bibliography produced by Kilp and Knauer (1986) lists more than 200 items.

Isbell's (1966) proof of the Zigzag Theorem (Theorems 8.3.4 and 8.3.5) involves a degree of 'handwaving' that is (to use his own word) objectionable. A successful amplification of his approach was provided by Philip (1974), but we have preferred to give the proof due to Stenström (1971). Much more recently Higgins (1990, 1992) has provided a different approach.

In category theory, a morphism $\phi : A \to B$ is an *epimorphism* if for all objects C and all morphisms $\beta, \gamma : B \to C$,

$$\phi\beta = \phi\gamma \;\Rightarrow\; \beta = \gamma.$$

As was pointed out by Isbell (1966), a semigroup (or monoid) morphism $\phi : S \to T$ is an epimorphism if and only if $\mathrm{Dom}_T(\mathrm{im}\,\phi) = T$. (In particular this implies the existence of non-surjective epimorphisms in the categories of semigroups and monoids.) The Zigzag Theorem has been extensively used in the study of epimorphisms. See, for example, Scheiblich (1976), Higgins (1981, 1983a,b,c, 1984a,b, 1985a,b, 1986a,b), Hall (1982), Hall and Jones (1983), Khan (1982, 1985a,b), and Trotter (1986b). Some crucial differences between ring and semigroup epimorphisms emerge from a paper of Gardner (1975).

A number of interesting classes of semigroups have the property of being absolutely closed, and several of these are correctly identified in Howie and Isbell (1967). It is worth mentioning, however, that their proof that full transformation semigroups are absolutely closed is incorrect. A correct proof was given by Scheiblich and Moore (1973). See also Shoji (1980) and Hall (1982).

The direct limit is a standard construction in universal algebra. See Cohn (1965). The connection between direct limits and free products was established for rings by Cohn (1959) and extended to semigroups by Renshaw (1986a).

The systematic study of semigroup amalgams was initiated by Howie (1962, 1963a,b,c, 1964a,b, 1968). His methods were combinatorial. Preston (1976), Hall (1975, 1978, 1980) and Imaoka (1976) developed a representational approach. Renshaw (1986a,b) showed how this approach, expressed in 'homological' terms, could be seen as similar to Cohn's (1959) work on amalgamation theory for rings, and it is Renshaw's approach that we have adopted in Sections 8.4 and 8.5. See also Howie (1985). More recently Dekov (1991, 1993, 1994) has re-proved some of the early results using an approach based on work of Baer (1949, 1950a,b). A survey article on amalgamations, with an extensive bibliography, has been written by Kiss *et al.* (1983).

Theorem 8.6.1 was proved by Howie (1975) using a combinatorial argument, and later by Hall (1978) using representational techniques. Theorem 8.6.4 is due to Hall (1978). The proof that inverse semigroups are absolutely flat is due to Bulman-Fleming and McDowell (1983).

Exercises 5 to 8 are from Howie and Isbell (1967), and Exercise 9 is from Higgins (1983a). Exercises 10 and 11 are from Scheiblich (1976), Exercises 12 and 13 from Hall (1982). Exercise 14 is based on Kimura (1957a). Exercise 15 is from Howie (1964a).

The concept 'almost unitary' appears in Howie (1962), where it is proved that an amalgam $[U; S, T]$ is embeddable if U is almost unitary in S and T.

References

Almeida, Jorge (1995). *Finite semigroups and universal algebra.* World Scientific, Singapore.

Almeida, Jorge, Bordalo, Gabriela and Dwinger, Philip (eds.) (1990). *Lattice, semigroups and universal algebra.* Plenum, New York.

Andersen, O. (1952). *Eine Bericht über die Struktur abstrakter Halbgruppen.* Staatsexamensarbeit, Hamburg.

Austin, A. K. (1966). A closed set of laws which is not generated by a finite set of laws. *Quarterly Journal of Mathematics Oxford* (2), **17**, 11–13.

Baer, R. (1949). Free sums of groups and their generalizations. *American Journal of Mathematics*, **71**, 706–742.

Baer, R. (1950a). Free sums of groups and their generalizations II. *American Journal of Mathematics*, **72**, 625–646.

Baer, R. (1950b). Free sums of groups and their generalizations III. *American Journal of Mathematics*, **72**, 647–670.

Baer, R. and Levi, F. W. (1932). Vollständige irreduzibele Systeme von Gruppenaxiomen (Beiträge zur Algebra, No. 18). *Sitzungberichte der Heidelberger Akademie Wissenschaften*, Abh. 2, 1–12.

Baird, G. R. (1972). On a sublattice of the lattice of congruences on a simple regular ω-semigroup. *Journal of the Australian Mathematical Society*, **13**, 461–471.

Bandyopadhyay, G. (1963). A simple proof of the decipherability criterion of Sardinas and Patterson. *Information and Control*, **6**, 331–336.

Berstel, Jean and Perrin, Dominique (1985). *Theory of codes.* Academic Press, New York.

Berthiaume, P. (1967). The injective envelope of S-sets. *Canadian Mathematical Bulletin*, **10**, 261–273.

Birkhoff, G. (1935). On the structure of abstract algebras. *Proceedings of the Cambridge Philosophical Society*, **31**, 433–454.

Birkhoff, G. (1948). *Lattice theory.* American Mathematical Society Colloquium Publications, Volume 25, Second Edition. New York.

Biryukov, A. P. (1970). Varieties of idempotent semigroups. *Algebra i Logika*, **9**, 255–273 (Russian).

Blum, E. K. (1965). Free subsemigroups of a free semigroup. *Michigan Mathematical Journal*, **12**, 179–182.

Bonzini, C., Cherubini, A. and Tibiletti, C. (eds.) (1993). *Semigroups: algebraic theory and applications to formal languages and codes.* World Scientific, Singapore.

Brandt, H. (1927). Über eine Verallgemeinerung des Gruppenbegriffes. *Mathematische Annalen*, **96**, 360–366.

Bruck, R. H. (1958). *A survey of binary systems.* Ergebnisse der Mathematik, Heft 20. Springer, Berlin.

Bulman-Fleming, S. and Gould, V. A. R. (1990). On left absolutely flat monoids. *Semigroup Forum*, **41**, 55–59.

Bulman-Fleming, S. and McDowell, K. (1980a). Flatness in varieties of normal bands. *Semigroup Forum*, **19**, 139–149.

Bulman-Fleming, S. and McDowell, K. (1980b). Flatness and projectivity in commutative regular semigroups. *Semigroup Forum*, **19**, 247–259.

Bulman-Fleming, S. and McDowell, K. (1983). Absolutely flat semigroups. *Pacific Journal of Mathematics*, **107**, 319–333.

Bulman-Fleming, S. and McDowell, K. (1984). Flatness and amalgamation in semigroups. *Semigroup Forum*, **29**, 337–342.

Bulman-Fleming, S. and McDowell, K. (1985). Left absolutely flat generalized inverse semigroups. *Proceedings of the American Mathematical Society*, **94**, 553–561.

Bulman-Fleming, S. and McDowell, K. (1987). On left absolutely flat bands. *Proceedings of the American Mathematical Society*, **101**, 613–618.

Bulman-Fleming, S. and McDowell, K. (1988). On V. Fleischer's characterization of absolutely flat monoids. *Algebra Universalis*, **25**, 394–399.

Bulman-Fleming, S., McDowell, K. and Renshaw, J. (1990). Some observations on left absolutely flat monoids. *Semigroup Forum*, **41**, 165–171.

Byleen, K. (1981). Regular four-spiral semigroups, idempotent-generated semigroups and the Rees construction. *Semigroup Forum*, **22**, 97–100.

Byleen, K., Meakin, J. and Pastijn, F. (1978). The fundamental four-spiral semigroup. *Journal of Algebra*, **65**, 6–26.

Byleen, K., Jones, P. and Pastijn, F. (eds.) (1984). *Proceedings of the 1984 Marquette conference on semigroups.* Marquette University.

Chen, S. Y. and Hsieh, S. C. (1974). Factorizable inverse semigroups. *Semigroup Forum*, **8**, 283–297.

Cohn, P. M. (1959). On the free product of associative rings. *Mathematische Zeitschrift*, **71**, 380–398.

Cohn, P. M. (1962). On subsemigroups of free semigroups. *Proceedings of the American Mathematical Society*, **13**, 347–351.

Cohn, P. M. (1965). *Universal algebra.* Harper and Row, New York.

Cohn, P. M. (1989). *Algebra*, Volume 2. Wiley, Chichester.

Clifford, A. H. (1941). Semigroups admitting relative inverses. *Annals of Mathematics*, **42**, 1037–1049.

Clifford, A. H. (1948). Semigroups containing minimal ideals. *American Journal of Mathematics*, **70**, 521–526.

Clifford, A. H. (1979). The free completely regular semigroup on a set. *Journal of Algebra*, **59**, 434–451.

Clifford, A. H. and Preston, G. B. (1961). *The algebraic theory of semigroups, Vol. 1, Mathematical Surveys of the American Mathematical Society, no. 7.* Providence, R. I.

Clifford, A. H. and Preston, G. B. (1967). *The algebraic theory of semigroups, Vol. 2, Mathematical Surveys of the American Mathematical Society, no. 7.* Providence, R. I.

Croisot (1953). Demi-groupes inversifs et demi-groupes réunions de demi-groupes simples. *Annales de Science de l'Ecole Normale Supérieure* (3), **70**, 361–379.

Dekov, Deko V. (1991). The embedding of semigroup amalgams. *Journal of Algebra*, **141**, 158–161.

Dekov, Deko V. (1993). Free products with amalgamation of semigroups. *Semigroup Forum*, **46**, 54–61.

Dekov, Deko V. (1994). HNN extensions of semigroups. *Semigroup Forum*, **49**, 83–87.

Dubreil, P. (1941). Contribution à la théorie des demi-groupes. *Mémoires de l'Académie des Sciences de l'Institut de France* (2), **63**, No. 3, 1–52.

Dubreil-Jacotin, M.-L. (1947). Sur l'immersion d'un demi-groupe dans un groupe. *Comptes Rendus de l'Académie des Sciences de Paris*, **225**, 787–788.

Dubreil-Jacotin, M.-L., Lesieur, L. and Croisot, R. (1953). *Leçons sur la théorie des treillis.* Gauthier-Villars, Paris

Eberhart, C., Williams, W. W. and Kinch, L. (1973). Idempotent-generated regular semigroups. *Journal of the Australian Mathematical Society,* **15**, 27–34.

Edwards, Constance C. (1977). The greatest idempotent separating congruence on an L-unipotent semigroup. *Semigroup Forum*, **14**, 127–135.

Eilenberg, S. (1974). *Automata, languages and machines, Vol. A.* Academic Press, New York.

Eilenberg, S. (1976). *Automata, languages and machines, Vol. B.* Academic Press, New York.

Evans, T. (1971). The lattice of semigroup varieties. *Semigroup Forum*, **2**, 1–43.

Feigenbaum, Ruth (1979). Regular semigroup congruences. *Semigroup Forum*, **17**, 373–377.

Feller, E. H. and Gantos, R. L. (1969). Indecomposable and injective S-systems with zero. *Mathematische Nachrichten*, **4**, 37–48.

Fennemore, C. F. (1971a). All varieties of bands I. *Mathematische Nachrichten*, **48**, 237–252.

Fennemore, C. F. (1971b). All varieties of bands II. *Mathematische Nachrichten*, **48**, 253–262.

Fleischer, V. (1982). Completely flat monoids. *Uchenye Zapiski Tartu Un-ta*, **610**, 38–52 (Russian).

Fitz-Gerald, D. G. (1972). On inverses of products of idempotents in regular semigroups. *Journal of the Australian Mathematical Society*, **13**, 335–337.

Fountain, J. B. (1979). Adequate semigroups. *Proceedings of the Edinburgh Mathematical Society* (2), **22**, 113–125.

Fountain, J. B. (1982). Abundant semigroups. *Proceedings of the London Mathematical Society* (3), **44**, 103–129.

Gardner, B. J. (1975). Epimorphisms of regular rings. *Commentationes Mathematicae Universitatis Carolinae*, **16**, 151–160.

Gerhard, J. A. (1970). The lattice of equational classes of idempotent semigroups. *Journal of Algebra*, **15**, 195–224.

Gerhard, J. (1983a). Free completely regular semigroups. I. Representation. *Journal of Algebra*, **82**, 135–142.

Gerhard, J. (1983b). Free completely regular semigroups. II. Word problem. *Journal of Algebra*, **82**, 143–156.

Gerhard, J. and Petrich, M. (1983). Word problems for free objects in certain varieties of completely regular semigroups. *Pacific Journal of Mathematics*, **104**, 351–359.

Gerhard, J. and Petrich, M. (1985). All varieties of regular orthogroups. *Semigroup Forum*, **31**, 311–351.

Giraldes, E. and Howie, John M. (1984). Embedding finite semigroups in finite semibands of minimal depth. *Semigroup Forum*, **28**, 135–142.

Giraldes, E. and Howie, John M. (1985). Semigroups of high rank. *Proceedings of the Edinburgh Mathematical Society* (2), **28**, 13–34.

Giraldes, E. and Howie, John M. (1991). Embeddings into finite idempotent-generated semigroups: some arithmetical results. *Proceedings of the Edinburgh Mathematical Society* (2), **34**, 259–269.

Gluskin, L. M. (1956). Completely simple semigroups. *Uchenye Zapiski Khar'kovkogo Pedagogicheskogo Instituta*, **18**, 41–55 (Russian).

Gluskin, L. M. (1957). Elementary generalized groups. *Matematicheskiĭ Sbornik*, **41**, 23–36 (Russian).

Gomes, G. M. S. (1983). *Aspects of order and congruence relations on regular semigroups.* Ph.D. Thesis, University of St Andrews.

Gomes, G. M. S. (1985). R-unipotent congruences on regular semigroups. *Semigroup Forum*, **31**, 265–280.

Gomes, G. M. S. (1986a). The lattice of R-unipotent congruences on a regular semigroup. *Semigroup Forum*, **33**, 159–173.

Gomes, G. M. S. (1986b). Some results on congruences on R-unipotent semigroups. *Semigroup Forum*, **33**, 175–185.

Gomes, G. M. S. and Howie, John M. (1987a). Nilpotents in finite symmetric inverse semigroups. *Proceedings of the Edinburgh Mathematical Society* (2), **30**, 383–395.

Gomes, G. M. S. and Howie, John M. (1987b). On the ranks of certain finite semigroups of transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, **101**, 395–403.

Gorenstein, Daniel (1982). *Finite simple groups: an introduction to their classification.* Plenum, New York.

Graham, R. L. (1968). On finite 0-simple semigroups and graph theory. *Mathematical Systems Theory*, **2**, 325–339.

Grätzer, George (1968). *Universal algebra.* Van Nostrand, Princeton.

Green, D. G. (1973). Extension of a semilattice by an inverse semigroup. *Bulletin of the Australian Mathematical Society*, **9**, 21–31.

Green, D. G. (1975). The lattice of congruences on an inverse semigroup. *Pacific Journal of Mathematics*, **57**, 141–152.

Green, J. A. (1951). On the structure of semigroups. *Annals of Mathematics*, **54**, 163–172

Green, J. A. and Rees, D. (1952). On semigroups in which $x^r = x$. *Proceedings of the Cambridge Philosophical Society*, **48**, 35–40.

Hall, M. (1959). *The theory of groups.* Macmillan, New York.

Hall, T. E. (1969). On regular semigroups whose idempotents form a subsemigroup. *Bulletin of the Australian Mathematical Society*, **1**, 195–208.

Hall, T. E. (1970). On orthodox semigroups and uniform and anti-uniform bands. *Journal of Algebra*, **16**, 204–217.

Hall, T. E. (1971). Orthodox semigroups. *Pacific Journal of Mathematics*, **39**, 677–686.

Hall, T. E. (1972). Congruences and Green's relations on regular semigroups. *Glasgow Mathematical Journal*, **13**, 167–175.

Hall, T. E. (1973). On regular semigroups. *Journal of Algebra*, **24**, 1–24.

Hall, T. E. (1975). Free products with amalgamation of inverse semigroups, *Journal of Algebra*, **34**, 375–385.

Hall, T. E. (1978). Representation extension and amalgamation for semigroups. *Quarterly Journal of Mathematics Oxford* (2), **29**, 309–334.

Hall, T. E. (1979). Inverse and regular semigroups and amalgamation. *Proceedings of the Conference on Regular Semigroups*, DeKalb, Illinois.

Hall, T. E. (1980). Generalized inverse semigroups and amalgamation. In *Semigroups*, (ed. T. E. Hall *et al*), pp. 145–157. Academic Press, New York.

Hall, T. E. (1982). Epimorphisms and dominions. *Semigroup Forum* **24**, 271–283.

Hall, T. E. (1989). The isomorphism problem for orthodox semigroups. *Pacific Journal of Mathematics*, **136**, 123–128.

Hall, T. E. and Jones, P. R. (1980). On the lattice of varieties of bands of groups. *Pacific Journal of Mathematics*, **91**, 327–337.

Hall, T. E. and Jones, P. R. (1983). Epis are onto for finite regular semigroups. *Proceedings of the Edinburgh Mathematical Society* (2), **26**, 1–7.

Hall, T. E. and Munn, W. D. (1979). Semigroups satisfying minimal conditions II. *Glasgow Mathematical Journal*, **20**, 133–140.

Hall, T. E., Jones, P. R. and Meakin, J. C. (eds.) (1991). *Monash conference in semigroup theory.* World Scientific, Singapore.

Hall, T. E., Jones, P. R. and Preston, G. B. (eds.) (1980). *Semigroups.* Academic Press, New York.

Halmos, P. R. (1960). *Naive set theory.* Van Nostrand, New York.

Hardy, G. H. and Wright, E. M. (1979). *An introduction to the theory of numbers*, (Fifth Edition). Oxford University Press.

Hickey, J. B. (1970). *Bisimple inverse semigroups and uniform semilattices.* Ph.D. Thesis, University of Glasgow.

Higgins, Peter M. (1981). Epis are onto for generalised inverse semigroups. *Semigroup Forum*, **23**, 151–162.

Higgins, Peter M. (1983a). The determination of all varieties consisting of absolutely closed semigroups. *Proceedings of the American Mathematical Society*, **87**, 419–421.

Higgins, Peter M. (1983b). The varieties of commutative semigroups for which epis are onto. *Proceedings of the Royal Society of Edinburgh A*, **94**, 1–7.

Higgins, Peter M. (1983c). A semigroup with an epimorphically embedded subband. *Bulletin of the Australian Mathematical Society*, **27**, 231–242.

Higgins, Peter M. (1984a). Saturated and epimorphically closed varieties of semigroups. *Journal of the Australian Mathematical Society A*, **36**, 153–175.

Higgins, Peter M. (1984b). Epimorphisms, permutation identities and finite semigroups. *Semigroup Forum*, **29**, 87–97.

Higgins, Peter M. (1985a). Semigroup epimorphisms. *Proceedings of the Marquette conference on semigroups, Milwaukee 1984*, 51–64. Marquette University, Milwaukee, Wisconsin.

Higgins, Peter M. (1985b). Epimorphisms, dominions and semigroups. *Algebra Universalis*, **21**, 225–233.

Higgins, Peter M. (1986a). Completely semisimple semigroups and epimorphisms. *Proceedings of the American Mathematical Society*, **96**, 387–390.

Higgins, Peter M. (1986b). Dense subsets of some common classes of semigroups. *Semigroup Forum*, **34**, 5–19.

Higgins, Peter M. (1990). A short proof of Isbell's zigzag theorem. *Pacific Journal of Mathematics*, **144**, 47–50.

Higgins, Peter M. (1992). *Techniques of semigroup theory.* Oxford University Press.

Higgins, Peter M. (1994). The Mitsch order on a semigroup. *Semigroup Forum*, **49**, 261–266.

Hofmann, K. H., Jürgensen, H. and Weinert, H. J. (1983). *Recent developments in the algebraic, analytical, and topological theory of semigroups.* (Proceedings of the conference at Oberwolfach, 1981.) Springer, Berlin.

Houghton, C. H. (1977). Completely 0-simple semigroups and their associated graphs and groups. *Semigroup Forum*, **14**, 41–67.

Houghton, C. H. (1979). Anti-isomorphisms of completely 0-simple semigroups. *Proceedings of the Royal Society of Edinburgh A*, **83**, 55–62.

Houghton, C. H. and Sullivan, R. P. (1984). Completely 0-simple semigroups generated by nilpotent elements. *Glasgow Mathematical Journal*, **25**, 163–165.

Howie, John M. (1962). Embedding theorems with amalgamation for semigroups. *Proceedings of the London Mathematical Society*(3), **12**, 511–534.

Howie, John M. (1963a). An embedding theorem with amalgamation for cancellative semigroups. *Proceedings of the Glasgow Mathematical Association*, **6**, 19–26.

Howie, John M. (1963b). Subsemigroups of amalgamated free products of semigroups. *Proceedings of the London Mathematical Society* (3), **13**, 672–686.

Howie, John M. (1963c). Embedding theorems for semigroups. *Quarterly Journal of Mathematics* (2), **14**, 254–258.

Howie, John M. (1964a). The embedding of semigroup amalgams. *Quarterly Journal of Mathematics* (2), **15**, 55–68.

Howie, John M. (1964b). Subsemigroups of amalgamated free products of semigroups. II. *Proceedings of the London Mathematical Society* (3), **14**, 537–544.

Howie, John M. (1964c). The maximum idempotent-separating congruence on an inverse semigroup. *Proceedings of the Edinburgh Mathematical Society* (2), **14**, 71–79.

Howie, John M. (1966). The subsemigroup generated by the idempotents of a full transformation semigroup. *Journal of the London Mathematical Society*, **41**, 707–716.

Howie, John M. (1968). Commutative semigroup amalgams. *Journal of the Australian Mathematical Society*, **8**, 609–630.

Howie, John M. (1975). Semigroup amalgams whose cores are inverse semigroups. *Quarterly Journal of Mathematics Oxford* (2), **26**, 23–45.

Howie, John M. (1976). *An introduction to semigroup theory.* Academic Press, London.

Howie, John M. (1978). Idempotents in completely 0-simple semigroups. *Glasgow Mathematical Journal*, **19**, 109–113.

Howie, John M. (1980). Products of idempotents in finite full transformation semigroups. *Proceedings of the Royal Society of Edinburgh A*, **86**, 243–254.

Howie, John M. (1981a). A class of bisimple, idempotent-generated congruence-free semigroups. *Proceedings of the Royal Society of Edinburgh A*, **88**, 169–184.

Howie, John M. (1981b). Some subsemigroups of infinite full transformation semigroups. *Proceedings of the Royal Society of Edinburgh A*, **88**, 159–167.

Howie, John M. (1985). Epimorphisms and amalgamations: a survey of recent progress. *Semigroups (Szeged 1981), Colloquia Mathematica Societatis János Bolyai*, **39**, 63–82. North-Holland, Amsterdam.

Howie, John M. (1991). *Automata and languages.* Oxford University Press.

Howie, John M. and Isbell, John R. (1967). Epimorphisms and dominions. II. *Journal of Algebra*, **6**, 7–21.

Howie, John M. and Lallement, G. (1966). Certain fundamental congruences on a regular semigroup. *Proceedings of the Glasgow Mathematical Association*, **7**, 145–159.

Howie, John M. and Schein, Boris M. (1969). Anti-uniform semilattices. *Bulletin of the Australian Mathematical Society*, **1**, 263–268.

Howie, John M., Munn, W. D. and Weinert, H. J. (eds.) (1992). *Semigroups with Applications.* (Proceedings of the Conference at Oberwolfach, 1991). World Scientific, Singapore.

Howie, John M., Robertson, Edmund F. and Schein, Boris M. (1988). A combinatorial property of finite full transformation semigroups. *Proceedings of the Royal Society of Edinburgh A*, **109**, 319–328.

Imaoka, Teruo (1976). Free products with amalgamation of bands. *Memoirs of the Faculty of Literature and Science. Shimane University. Natural Sciences*, **10**, 7–17.

Isbell, John R. (1966). Epimorphisms and dominions. *Proceedings of the Conference on Categorical Algebra, La Jolla, California, 1965.* Springer, New York.

Iwahori, Nobuko (1977). A length formula in a semigroup of mappings. *Journal of the Faculty of Science of the University of Tokyo, Section 1A (Mathematics)*, **24**, 255–260.

Jones, P. R. (1981). Free products, free semigroups and varieties. *Proceedings of the Royal Society of Edinburgh A*, **88**, 293–313.

Jones, P. R. (1982). A graphical representation for the free product of E-unitary inverse semigroups. *Semigroup Forum*, **24**, 195–221.

Jones, P. R. (1983). On the lattice of varieties of completely regular semigroups. *Journal of the Australian Mathematical Society A*, **35**, 227–235.

Jones, P. R. (1984). Free products of inverse semigroups. *Transactions of the American Mathematical Society*, **282**, 293–317.

Jones, P. R., Margolis, S. W., Meakin, J. and Stephen, J. B. (1991). Free products of inverse semigroups. II. *Glasgow Mathematical Journal*, **33**, 373–387.

Jürgensen, H., Migliorini, F. and Szép, J. (1991). *Semigroups.* Akadémiai Kiadó, Budapest.

Jürgensen, H., Petrich, M. and Weinert, H. J. (eds.) (1981). *Semigroups: Proceedings, Oberwolfach, Germany, 1978.* Springer, Berlin.

Kaplansky, I. (1954). *Infinite abelian groups.* University of Michigan.

Kapp, K. M. and Schneider, H. (1969). *Completely* 0*-simple semigroups.* Benjamin, New York.

Khan, N. M. (1982). Epimorphisms, dominions and varieties of semigroups. *Semigroup Forum*, **25**, 331–337.

Khan, N. M. (1985a). On saturated permutative varieties and consequences of permutation identities. *Journal of the Australian Mathematical Society A*, **38**, 186–197.

Khan, N. M. (1985b). Epimorphically closed permutative varieties. *Transactions of the American Mathematical Society*, **287**, 507–528.

Kilp, M. (1970). On flat acts. *Uchenye Zapiski Tartu Un-ta*, **253**, 66–72.

Kilp, M. (1972). On the homological classification of monoids. *Sibirskiĭ Matematicheskiĭ Zhurnal*, **13**, 578–586.

Kilp, M. (1985). Strong flatness of flat cyclic left acts. *Uchenye Zapiski Tartu Un-ta*, **700**, 38–41.

Kilp, M. and Knauer, U. (1982). Acts over monoids: a bibliographic survey of publications 1975–1981. *Bibliotheks- und Informationssystem der Universität Oldenburg.*

Kilp, M. and Knauer, U. (1986). On free, projective, and strongly flat acts. *Archiv für Mathematik*, **47**, 17–23.

Kimura, N. (1957a). *On semigroups.* Doctoral dissertation, The Tulane University of Louisiana.

Kimura, N. (1957b). Note on idempotent semigroups. I. *Proceedings, Japan Academy*, **33**, 642–645.

Kimura, N. (1958a). Note on idempotent semigroups. III. *Proceedings, Japan Academy*, **34**, 113–114.

Kimura, N. (1958b). Note on idempotent semigroups. IV, Identities of three variables. *Proceedings, Japan Academy*, **34**, 121–123.

Kimura, N. (1958c). The structure of idempotent semigroups. (I). *Pacific Journal of Mathematics*, **8**, 257–275.

Kiss, E. W., Márki, L., Pröhle, P. and Tholen, W. (1983). Categorical algebraic properties; a compendium on amalgamation, congruence extension, residual smallness and injectivity. *Studia Scientiarum Mathematicarum Hungarica*, **18**, 79–141.

Knauer, U. (1972). Projectivity of acts and Morita equivalence of monoids. *Semigroup Forum*, **3**, 359–370.

Knauer, U (1983). Characterization of monoids by properties of their cyclic or finitely generated acta and their right ideals. In *Recent developments in the algebraic, analytic and topological theory of semigroups: Proceedings of the 1981 Oberwolfach Conference*, pp. 310–332. Springer-Verlag, Berlin.

Knauer, U. and Petrich, M. (1981). Characterization of monoids by torsion-free, flat, projective and free acts. *Archiv für Mathematik*, **36**, 289–294.

Kochin, B. P. (1968). The structure of inverse ideal-simple ω-semigroups. *Vestnik Leningradskogo Universiteta*, **23**, 41–50 (Russian).

Kopamu, Samuel J. L. (1995). On semigroup species (submitted).

Kurosh, A. G. (1956). *The theory of groups*, Volumes I and II (translated from the Russian by K. A. Hirsch). Chesea, New York.

Lajos, S. (1969). On semilattices of groups. *Proceedings of the Japan Academy*, **45**, 383–384.

Lajos, S. (1970a). Characterization of completely regular inverse semigroups. *Acta Universitatis Szegediensis. Acta Scientiarum Mathematicarum*, **31**, 229–231.

Lajos, S. (1970b). Notes on regular semigroups. *Proceedings of the Japan Academy*, **46**, 253–254.

Lajos, S. (1970c). Notes on regular semigroups. II. *Proceedings of the Japan Academy*, **46**, 929–931.

Lajos, S. (1971a). On semilattices of groups. II. *Proceedings of the Japan Academy*, **47**, 36–37.

Lajos, S. (1971b). Notes on regular semigroups. III. *Proceedings of the Japan Academy*, **47**, 185–186.

Lajos, S. (1972a). A remark on semilattices of groups. *Annales Universitatis Scientiarum Budapestinensis*, **15**, 121–122.

Lajos, S. (1972b). A note on semilattices of groups. *Acta Universitatis Szegediensis. Acta Scientiarum Mathematicarum*, **33**, 315–317.

Lallement, G. (1966). Congruences et équivalences de Green sur un demigroupe régulier. *Comptes Rendus de l'Académie des Sciences Paris, Sér. A*, **262**, 613–616.

Lallement, G. (1967). Demi-groupes réguliers. *Annali di Matematica Pura ed Applicada*, **77**, 47–129.

Lallement, G. (1974). A note on congruences on Rees matrix semigroups. *Semigroup Forum*, **8**, 89–92.

Lallement, G. (1979). *Semigroups and combinatorial applications.* Wiley, New York.

Lallement, G. and Petrich, M. (1966). Décompositions I-matricielles d'un demi-groupe. *Journal de Mathématiques Pures et Appliquées*, **45**, 67–117.

La Torre, D. R. (1981). θ-classes in L-unipotent semigroups. *Semigroup Forum*, **22**, 355–365.

Levenshtein, V. I. (1964). Some problems of coding and self-adjusting automata for decoding messages. *Problemy Kibernetiki*, **11**, 63–121.

Levi, F. W. (1944). On semigroups. *Bulletin of the Calcutta Mathematical Society*, **36**, 141–146.

Liber, S. A. (1954). On the theory of generalized groups. *Doklady Akademiĭ Nauk SSSR*, **97**, 25–28.

Lint, J. H. van and Wilson, R. M. (1992). *A course in combinatorics.* Cambridge University Press.

Lipscomb, S. L. (1986). Cyclic subsemigroups of symmetric inverse semigroups. *Semigroup Forum*, **34**, 243–248.

Lipscomb, S. L. (1992a). The alternating semigroups: generators and congruences. *Semigroup Forum*, **44**, 96–106.

Lipscomb, S. L. (1992b). Centralizers in symmetric inverse semigroups: structure and order. *Semigroup Forum*, **44**, 347–355.

Lothaire, M. (1983). *Combinatorics on words.* Encyclopedia of Mathematics and its Applications, Volume 17. Addison-Wesley, Reading, Massachusetts.

Luca, A. de (1976). A note on variable length codes. *Information and Control*, **32**, 263–271.

Lyapin, E. S. (1953). Canonical form of elements of an associative system given by defining relations. *Leningradskiĭ Gosudarstvennyĭ Pedagogicheskiĭ Institut Uchenye Zapiski*, **89**, 45–54 (Russian).

Lyndon, R. C. and Schützenberger, M. P. (1962). The equation $a^M = b^N c^P$ in a free group. *Michigan Mathematical Journal*, **9**, 289–298.

McAlister, D. B. (1974a). Groups, semilattices and inverse semigroups. *Transactions of the American Mathematical Society*, **192**, 227–244.

McAlister, D. B. (1974b). Groups, semilattices and inverse semigroups, II. *Transactions of the American Mathematical Society*, **196**, 351–370.

McAlister, D. B. (1983). Rees matrix covers for locally inverse semigroups. *Transactions of the American Mathematical Society*, **277**, 727–738.

McAlister, D. B. (1985). Some covering theorems for locally inverse semigroups. *Journal of the Australian Mathematical Society A*, **39**, 63–74.

McAlister, D. B. and McFadden, R. (1974). Zig-zag representations and inverse semigroups. *Journal of Algebra*, **32**, 178–206.

McAlister, D. B. and Reilly, N. R. (1977). E-unitary covers for inverse semigroups. *Pacific Journal of Mathematics*, **68**, 161–174.

McFadden, R. and O'Carroll, L. (1971). F-inverse semigroups. *Proceedings of the London Mathematical Society* (3), **22**, 652–666.

MacKnight, J. D., Jr., and Storey, J. (1969). Equidivisible semigroups. *Journal of Algebra*, **12**, 24–48.

McLean, D. (1954). Idempotent semigroups. *American Mathematical Monthly*, **61**, 110–113.

McLean, P. (1973). *Contributions to the theory of* 0*-simple inverse semigroups.* Ph.D. Thesis, University of Stirling.

Margolis, S. W., Meakin, J. and Stephen, J. B. (1990). Free objects in certain varieties of inverse semigroups. *Canadian Journal of Mathematics*, **42**, 1084–1097.

Márki, L. (1975). On locally regular Rees matrix semigroups. *Acta Scientiae Mathematicae (Szeged)*, **37**, 95–102.

Marques, M. Paula O. (1983). A congruence-free semigroup associated with an infinite cardinal number. *Proceedings of the Royal Society of Edinburgh A*, **93**, 245–257.

Meakin, J. (1985a). The Rees construction in regular semigroups. In *Semigroups (Szeged 1981), Colloquia Mathematica Societatis János Bolyai*, **39**, pp. 115–155. North-Holland, Amsterdam.

Meakin, J. (1985b). Fundamental regular semigroups and the Rees construction. *Quarterly Journal of Mathematics Oxford* (2), **36**, 91–103.

Meakin, J. (1993). An invitation to inverse semigroup theory. *Ordered structures and algebra of computer languages*, 91-115. World Scientific, Singapore.

Miller, D. D. and Clifford, A. H. (1956). Regular $\mathcal{D}$-classes in semigroups. *Transactions of the American Mathematical Society*, **82**, 270–280.

Mitchell, B. (1965). *Theory of categories.* Academic Press, New York.

Mitsch, H. (1986). A natural partial order for semigroups. *Proceedings of the American Mathematical Society*, **97**, 384–388.

Munn, W. D. (1955). On semigroup algebras. *Proceedings of the Cambridge Philosophical Society*, **51**, 1–15.

Munn, W. D. (1957). Semigroups satisfying minimal conditions. *Proceedings of the Glasgow Mathematical Association*, **3**, 145–152.

Munn, W. D. (1961). A class of irreducible representations of an arbitrary inverse semigroup. *Proceedings of the Glasgow Mathematical Association*, **5**, 41–48.

Munn, W. D. (1966a). The lattice of congruences on a bisimple ω-semigroup. *Proceedings of the Royal Society of Edinburgh A*, **67**, 175–184.

Munn, W. D. (1966b). Uniform semilattices and bisimple inverse semigroups. *Quarterly Journal of Mathematics Oxford* (2), **17**, 151–159.

Munn, W. D. (1968). Regular ω-semigroups. *Glasgow Mathematical Journal*, **9**, 46–66.

Munn, W. D. (1970a). Fundamental inverse semigroups. *Quarterly Journal of Mathematics Oxford* (2), **21**, 157–170.

Munn, W. D. (1970b). 0-Bisimple inverse semigroups. *Journal of Algebra*, **15**, 570–588.

Munn, W. D. (1970c). On simple inverse semigroups. *Semigroup Forum*, **1**, 63–74.

Munn, W. D. (1972). Embedding semigroups in congruence-free semigroups. *Semigroup Forum*, **4**, 46–60.

Munn, W. D. (1974a). Congruence-free inverse semigroups. *Quarterly Journal of Mathematics Oxford* (2), **25**, 463–484.

Munn, W. D. (1974b). Free inverse semigroups. *Proceedings of the London Mathematical Society* (3), **29**, 385–404.

Munn, W. D. (1975). A note on congruence-free inverse semigroups. *Quarterly Journal of Mathematics Oxford* (2), **26**, 385–387.

Munn, W. D. (1976). A note on E-unitary inverse semigroups. *Bulletin of the London Mathematical Society*, **8**, 71–76.

Munn, W. D. and Reilly, N. R. (1966). Congruences on a bisimple ω-semigroup. *Proceedings of the Glasgow Mathematical Association*, **7**, 184–192.

Nambooripad, K. S. S. (1974). Structure of regular semigroups I. *Semigroup Forum*, **9**, 354–363.

Nambooripad, K. S. S. (1975). Structure of regular semigroups II. The general case. *Semigroup Forum*, **9**, 364–371.

Nambooripad, K. S. S. (1979). *Structure of regular semigroups. I.* Memoirs of the American Mathematical Society, no. 224.

Nambooripad, K. S. S. (1980). The natural partial order on a regular semigroup. *Proceedings of the Edinburgh Mathematical Society* (2), **23**, 249–260.

Neumann, B. H. (1954). An essay on free products of groups with amalgamation. *Philosophical Transactions of the Royal Society*, **246**, 503–554.

Neumann, J. von (1936). On regular rings. *Proceedings of the National Academy of Sciences, U.S.A.*, **22**, 503–554.

Newton, Tony (1993). *Applications of the Hall–Yamada approach to orthodox semigroups.* Ph.D. Thesis, Monash University.

Normak, P. (1987). On equalizer-flat and pullback-flat acts. *Semigroup Forum*, **36**, 293–313.

O'Carroll, L. (1974). A note on free inverse semigroups. *Proceedings of the Edinburgh Mathematical Society* (2), **19**, 17–23.

Okniński, J. (1990). *Semigroup algebras.* Dekker, New York.

Pastijn, F. (1977). Embedding semigroups in semibands. *Semigroup Forum*, **14**, 247–263.

Pastijn, F. and Petrich, M. (1985). *Regular semigroups as extensions.* Pitman, Boston.

Pastijn, F. and Petrich, M. (1986). Congruences on regular semigroups. *Transactions of the American Mathematical Society*, **285**, 607–633.

Petrich, M. (1968). The translational hull of a completely 0-simple semigroup. *Glasgow Mathematical Journal*, **9**, 1–11.

Petrich, M. (1971). A construction and a classification of bands. *Mathematische Nachrichten*, **48**, 263–274.

Petrich, M. (1973). *Introduction to semigroups.* Merrill, Columbus, Ohio.

Petrich, M. (1975). Varieties of orthodox bands of groups. *Pacific Journal of Mathematics*, **58**, 209–217.

Petrich, M. (1977). Certain varieties and quasivarieties of completely regular semigroups. *Canadian Journal of Mathematics*, **29**, 1171–1197.

Petrich, M. (1978). Congruences on inverse semigroups. *Journal of Algebra*, **55**, 231–256.

Petrich, M. (1982). On the varieties of completely regular semigroups. *Semigroup Forum*, **25**, 153–169

Petrich, M. (1984). *Inverse semigroups.* Wiley, New York.

Petrich, M. and Reilly, N. R. (1979). A representation of E-unitary semigroups. *Quarterly Journal of Mathematics Oxford* (2), **30**, 339–350.

Petrich, M. and Reilly, N. R. (1981a). Varieties of groups and completely simple semigroups. *Bulletin of the Australian Mathematical Society*, **23**, 339–359.

Petrich, M. and Reilly, N. R. (1981b). The kernel-trace approach to congruences on completely simple semigroups. *Studia Scientiarum Mathematicarum Hungarica*, **16**, 103–114.

Petrich, M. and Reilly, N. R. (1982). A network of congruences on an inverse semigroup. *Transactions of the American Mathematical Society*, **270**, 309–325.

Petrich, M. and Reilly, N. R. (1983a). E-unitary covers and varieties of inverse semigroups. *Acta Scientiae Mathematicae (Szeged)*, **46**, 59–72.

Petrich, M. and Reilly, N. R. (1983b). Near varieties of idempotent-generated completely simple semigroups. *Algebra Universalis*, **16**, 83–104.

Petrich, M. and Reilly, N. R. (1983c). All varieties of central completely simple semigroups. *Transactions of the American Mathematical Society*, **280**, 623–636.

Petrich, M. and Reilly, N. R. (1984a). The join of the varieties of strict regular semigroups and rectangular bands. *Glasgow Mathematical Journal*, **25**, 59–74.

Petrich, M. and Reilly, N. R. (1984b). Certain homomorphisms of the lattice of varieties of completely simple semigroups. *Journal of the Australian Mathematical Society A*, **37**, 287–306.

Petrich, M. and Reilly, N. R. (1988). Semigroups generated by certain operators on varieties of completely regular semigroups. *Pacific Journal of Mathematics*, **132**, 151–175.

Philip, J. M. (1974). A proof of Isbell's zigzag theorem. *Journal of Algebra*, **32**, 328–331.

Pin, J.-E. (1986). *Varieties of formal languages*, North Oxford Academic Publishers, London. (A translation of *Variétés de langages formels.* Masson, Paris, 1984.)

Polák, Libor (1985). On varieties of completely regular semigroups I. *Semigroup Forum*, **32**, 97–123.

Polák, Libor (1987). On varieties of completely regular semigroups II. *Semigroup Forum*, **36**, 253–284

Polák, Libor (1988). On varieties of completely regular semigroups III. *Semigroup Forum*, **37**, 1–30.

Pollák, G., Schwarz, Št. and Steinfeld, O. (eds.) (1985). *Semigroups*. Colloquia Mathematica Societatis János Bolyai, 39. North-Holland, Amsterdam.

Preston, G. B. (1954a). Inverse semi-groups. *Journal of the London Mathematical Society*, **29**, 396–403.

Preston, G. B. (1954b). Inverse semi-groups with minimal right ideals. *Journal of the London Mathematical Society*, **29**, 404–411.

Preston, G. B. (1954c). Representations of inverse semi-groups. *Journal of the London Mathematical Society*, **29**, 411–419.

Preston, G. B. (1961). Congruences on completely 0-simple semigroups. *Proceedings of the London Mathematicak Society* (3), **11**, 557–576.

Preston, G. B. (1965). Chains of congruences on a completely 0-simple semigroup. *Journal of the Australian Mathematical Society*, **5**, 557–576.

Preston, G. B. (1973). Free inverse semigroups. *Journal of the Australian Mathematical Society*, **16**, 443–453.

Preston, G. B. (1976). Free products amalgamating unitary subsemigroups. *Bulletin of the Australian Mathematical Society*, **15**, 117–124.

Rasin, V. V. (1979). On the lattice of varieties of completely simple semigroups. *Semigroup Forum*, **17**, 114–122.

Rasin, V. V. (1981). On the varieties of Cliffordian semigroups. *Semigroup Forum*, **23**, 201–220.

Rees. D. (1940). On semi-groups. *Proceedings of the Cambridge Philosophical Society*, **36**, 387–400.

Reilly, N. R. (1966). Bisimple ω-semigroups. *Proceedings opf the Glasgow Mathematical Association*, **7**, 160–167.

Reilly, N. R. (1972). Free generators in free inverse semigroups. *Bulletin of the Australian Mathematical Society*, **8**, 407–424.

Reilly, N. R. (1973). Corrigenda: free generators in free inverse semigroups. *Bulletin of the Australian Mathematical Society*, **9**, 479–480.

Reilly, N. R. (1979). Free inverse semigroups. In *Semigroups (Szeged 1976), Colloquia Mathematica Societatis János Bolyai*, **20**, North-Holland, Amsterdam.

Reilly, N. R. (1987). Free objects in certain varieties of semigroups. *Proceedings of the Second Seminar on Non-Commutative Algebra*, 39–69, University of Siena.

Reilly, N. R. and Trotter, P. G. (1986). Properties of relatively free inverse semigroups. *Transactions of the American Mathematical Society*, **294**, 243–268.

Renshaw, J. (1986a). Extension and amalgamation in monoids and semigroups. *Proceedings of the London Mathematical Society* (3), **52**, 119–141.

Renshaw, J. (1986b). Flatness and amalgamation in monoids, *Journal of the London Mathematical Society* (2) **33**, 73–88.

Rhodes, John (ed.) (1991). *Monoids and semigroups with applications.* (Proceedings of the Berkeley Workshop on Monoids, 1989.) World Scientific, Singapore.

Riley, J. A. (1967). The Sardinas–Patterson and Levenshtein theorems. *Information and Control*, **10**, 120–136.

Rotman, J. J. (1979). *An introduction to homological algebra.* Academic Press, New York.

Ruškuc, N. (1994). On the rank of completely 0-simple semigroups. *Mathematical Proceedings of the Cambridge Philosophical Society*, **116**, 325–338.

Sardinas, A. A. and Patterson, C. W. (1953). A necessary and sufficient condition for the unique decomposition of coded messages. *Institute of Radio Engineers International Convention Records*, **8**, 104–108.

Scheiblich, H. E. (1973a). Free inverse semigroups. *Semigroup Forum*, **4**, 352–359.

Scheiblich, H. E. (1973b). Free inverse semigroups. *Proceedings of the American Mathematical Society*, **38**, 1–7.

Scheiblich, H. E. (1974). Kernels of inverse semigroup homomorphisms. *Journal of the Australian Mathematical Society*, **18**, 289–292.

Scheiblich, H. E. (1976). On epics and dominions in bands, *Semigroup Forum*, **13**, 103–114.

Scheiblich, H. E. and Moore, K. C. (1973). $\mathcal{T}_X$ is absolutely closed. *Semigroup Forum*, **6**, 216–226.

Schein, B. M. (1962). Representations of generalized groups. *Izvestiya Vysshikh Uchebnykh Zavedeniĭ Matematika*, No. 3, **28**, 164–176 (Russian).

Schein, B. M. (1964). Generalized groups with the well-ordered set of idempotents. *Matematicko-Fyzikálny Časopis. Slovenska Akadémia Vied*, **14**, 259–262.

Schein, B. M. (1975a). Free inverse semigroups are not finitely presentable. *Acta Mathematica Academiae Scientiarum Hungaricae*, **26**, 41–52.

Schein, B. M. (1975b). A new proof for the McAlister P-theorem. *Semigroup Forum*, **10**, 185–188.

Schreier, O. (1927). Die Untergruppen der freien Gruppen. *Abhandlungen aus dem mathematischen Seminar der hansischen Universität Hamburg*, **5**, 161–183.

Schubert, H. (1972). *Categories.* Springer, New York.

Schützenberger, M. P. (1956). Une théorie algébrique du codage. *Comptes Rendus de l'Académie des Sciences de Paris*, **242**, 862–864.

Sezinando, H. (1992a). The cardinalities of the Green classes of the free objects in varieties of bands. *Semigroup Forum*, **44**, 164–198.

Sezinando, H. (1992b). The varieties of languages corresponding to the varieties of finite band monoids. *Semigroup Forum*, **44**, 283–305.

Shevrin, L. N. (1960). Subsemigroups of free semigroups. *Doklady Akademiĭ Nauk SSSR*, **133**, 537–539 (Russian); translated in *Soviet Mathematics Doklady*, **1**, 892–894.

Shevrin, L. N. and Volkov, M. V. (1985). Identities of semigroups. *Izvestiya Vysshikh Uchebnikh Zavedeniĭ Matematika*, **39**, 3–47 (Russian).

Shoji, Kunitaka (1980). Right self-injective semigroups are absolutely closed. *Memoirs of the Faculty of Science, Shimane University*, **14**, 35-39.

Shum, K. P. and Yuen, P. C. (eds.) (1993). *Ordered structures and algebra of computer languages.* World Scientific, Singapore.

Shyr, H. J. (1991). *Free monoids and languages.* Hon Min Book Company, Taichung, Taiwan.

Shyr, H. J. and Thierrin, G. (1976). Codes, languages and MOL schemes. *Theoretical Computer Science*, **1**, 293–301.

Spehner, J.-C. (1975). Quelques constructions et algorithmes relatifs aux sous-monoïdes d'un monoïde libre. *Semigroup Forum*, **9**, 334–353.

Steinfeld, O. (1966). On semigroups which are unions of completely 0-simple semigroups. *Czechoslovak Mathematical Journal*, **16 (91)**, 63–69.

Stenström, B. (1971). Flatness and localization over monoids. *Mathematische Nachrichten*, **48**, 315–333.

Stephen, J. B. (1990). Presentations of inverse monoids. *Journal of Pure and Applied Algebra*, **63**, 81–112.

Suschkewitz, A. K. (1928). Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit. *Mathematische Annalen*, **99**, 30–50.

Takizawa, K. (1978). E-unitary $\mathcal{R}$-unipotent semigroups. *Bulletin of the Tokyo Gakugei University* (4), **30**, 21–33.

Tamura, T. (1956). Indecomposable completely simple semigroups except groups. *Osaka Mathematical Journal*, **8**, 35–42.

Tamura, T. (1960). Decompositions of a completely simple semigroup. *Osaka Mathematical Journal*, **12**, 269–275.

Tamura, T. and Kimura, N. (1954). On decompositions of a commutative semigroup. *Kodai Mathematical Seminar Reports*, **1954**, 109–112.

Taylor, W. (1979). Equational logic. *Houston Journal of Mathematics Survey*, (iii) + 83pp.

Tran Quy Tien (1975). *Investigations in the theory of Rees matrix semigroups and Rees matrix rings*, Doctoral Thesis, Hungarian Academy of Sciences, Budapest (in Hungarian).

Trotter, P. G. (1974). Congruence-free inverse semigroups. *Semigroup Forum*, **9**, 109–116.

Trotter, P. G. (1986a). Relatively free inverse semigroups. *Quarterly Journal of Mathematics*, **37**, 357–374.

Trotter, P. G. (1986b). A non-surjective epimorphism of bands. *Algebra Universalis*, **22**, 109–116.

Trotter, P. G. (1988). Free generators in relatively free completely regular semigroups. *Proceedings of the Royal Society of Edinburgh A*, **109**, 329–339.

Vagner, V. V. (1952). Generalized groups. *Doklady Akademiĭ Nauk SSSR*, **84**, 1119-1122 (Russian).

Vagner, V. V. (1953). Theory of generalized heaps and generalized groups. *Matematicheskiĭ Sbornik (N. S.)*, **32**, 545–632 (Russian).

Vagner, V. V. (1956). Representations of ordered semigroups. *Matematicheskiĭ Sbornik*, **38**, 203–240 (Russian); *American Mathematical Society Translations* (2), **36** (1964), 295–336.

Venkatesan, P. S. (1966). On decomposition of semigroups with zero. *Mathematische Zeitschrift*, **92**, 164–174.

Venkatesan, P. S. (1974). Right (Left) inverse semigroups. *Journal of Algebra*, **31**, 209–217.

Warne, R. J. (1966). A class of bisimple inverse semigroups. *Pacific Journal of Mathematics*, **18**, 563–577.

Warne, R. J. (1968). I-bisimple semigroups. *Transactions of the American Mathematical Society*, **130**, 367–386.

Wilkinson, R. (1983). A description of E-unitary inverse semigroups. *Proceedings of the Royal Society of Edinburgh A*, **95**, 239–242.

Wismath, S. (1986). The lattice of varieties and pseudovarieties of band monoids. *Semigroup Forum*, **33**, 187–198.

Yamada, M. (1958). Note on idempotent semigroups. V. *Proceedings, Japan Academy*, **34**, 668–671.

Yamada, M. and Kimura, N. (1958). Note on idempotent semigroups. *Proceedings, Japan Academy*, **34**, 110–112.

List of symbols

$\mathcal{A}$	a semigroup amalgam	263
A^+	the free semigroup on A	29
A^*	the free monoid on A	29
a^+	the free semigroup on $\{a\}$	29
a^*	the free monoid on $\{a\}$	29
$a \succ b$	a covers b	35
$a \wedge b$	the meet of a and b	13
$a \vee b$	the join of a and b	13
B_A	the free band on A	118
$B(G,n)$	a Brandt semigroup	152
B_n	the Brandt semigroup $B(\{e\},n)$	152
B_2	the five element Brandt semigroup	32
$BR(T,\theta)$	the Bruck–Reilly extension	171
$\mathcal{B}_X$	the semigroup of binary relations on X	17
$\mathcal{B}$	the variety of bands	112
$\mathcal{C}$	the variety of commutative semigroups	112
$\mathcal{CL}$	the variety of Clifford semigroups	111
$\mathcal{CR}$	the variety of completely regular semigroups	111
$\mathcal{CS}$	the variety of completely simple semigroups	111
C_ω	the semilattice $\{e_0 > e_1 > e_2 > \cdots\}$	163
$C(w)$	the contents of a word w	119
$\mathcal{D}$	Green's relation	46
$\mathcal{D}^U$	Green's relation in U	56
D_a	the $\mathcal{D}$-class containing a	47
D_a^U	the $\mathcal{D}^U$-class containing a	56
$\operatorname{dom}\phi$	the domain of ϕ	18
$\operatorname{Dom}_S(U)$	the dominion of U in S	266
E, or E_S	the set of idempotents of a semigroup S	146
$E^\flat$	the largest congruence contained in E	27
ENS-S	the class of right S-systems	252
$\operatorname{fix}\alpha$	the set $\{x : x\alpha = x\}$, where $\alpha \in \mathcal{T}_X$	237
FI_X	the free inverse monoid on X	200
$F(S,X,Y)$	the free S-extension of X and Y	281
$F_\mathcal{V}(X)$	the relatively free algebra on X	110
$\mathcal{G}$	the variety of groups	111
$\mathcal{G}_X$	the symmetric group on X	6

$(G, \mathcal{X}, \mathcal{Y})$	a McAlister triple	194
γ	the least inverse semigroup congruence	229
$\mathcal{H}$	Green's relation	46
$\mathcal{H}^U$	Green's relation in U	56
H_a	the $\mathcal{H}$-class containing a	47
H_a^U	the $\mathcal{H}^U$-class containing a	56
$H\omega$	the closure of H	154
$h(w)$	the head (first letter) of w	126
$I(a)$	the ideal $\{b \in J(a) : J_b < J_a\}$	67
$\operatorname{im} \phi$	the image of ϕ	18
$\mathcal{I}_X$	the symmetric inverse semigroup	148
$\mathcal{J}$	Green's relation	46
$\mathcal{J}^U$	Green's relation in U	56
$J(a)$	the principal ideal generated by a	67
$J(a)/I(a)$	the principal factor containing a	68
J_a	the $\mathcal{J}$-class containing a	47
J_a^U	the $\mathcal{J}^U$-class containing a	56
$\operatorname{Ker} \rho$	the kernel of an inverse semigroup congruence	155
$\ker \phi$	the kernel of a map ϕ	20
$K(S)$	the kernel of a semigroup S	67
$K(\Omega)$	the kernel of a component of the function graph	231
$\mathcal{L}$	Green's relation	45
$\mathcal{L}^U$	Green's relation in U	56
L_a	the $\mathcal{L}$-class containing a	47
L_a^U	the $\mathcal{L}^U$-class containing a	56
λ_s	left translation by s	50
$\mathcal{LN}$	the variety of left normal bands	132
$\mathcal{LZ}$	the variety of left zero semigroups	112
$M(m, r)$	a monogenic semigroup	12
μ	the maximum idempotent-separating congruence	160
$\mathcal{M}[G; I, \Lambda; P]$	the Rees matrix semigroup over G	77
$\mathcal{M}^0[G; I, \Lambda; P]$	the Rees matrix semigroup over G^0	70
$\mathcal{M}(G, \mathcal{X}, \mathcal{Y})$	an E-unitary inverse semigroup	194
$\mathcal{N}$	the variety of normal bands	132
$\mathbf{N}$	the natural numbers $\{1, 2, 3, \ldots\}$	
$\mathbf{N}^{\mathbf{0}}$	the set $\{0, 1, 2, 3, \ldots\}$	
$(N, \mathcal{S}, \mathcal{T})$	a linked triple	85
$[N, \mathcal{S}, \mathcal{T}]$	the congruence determined by a linked triple	90
$\Omega(S)$	the translational hull of S	114
1_X	the diagonal relation (identity map) on X	15
$p_{\lambda i}$	an entry of the sandwich matrix	69
$\mathcal{P}_X$	the semigroup of all partial maps of X	17
$\Pi^* S_i$	the free product of the semigroups S_i	256

$\Pi^*_U S_i$	the free product of the amalgam $[U; S_i]$	263
$\phi \oplus \psi$	the sum of representations ϕ and ψ	186
$\mathbf{Q}$	the rational numbers	
$q_{\lambda\mu ij}$	an extract from $P = (p_{\lambda i})$	84
$\mathcal{R}$	Green's relation	45
$\mathcal{R}^U$	Green's relation in U	56
R_a	the $\mathcal{R}$-class containing a	47
R_a^U	the $\mathcal{R}^U$-class containing a	56
rank α	$\lvert \operatorname{im} \alpha \rvert$, where $\alpha \in \mathcal{T}_X$	237
$\mathcal{RB}$	the variety of rectangular bands	112
$\mathbf{R}^e$	the equivalence generated by $\mathbf{R}$	21
$\mathbf{R}^\infty$	the transitive closure of $\mathbf{R}$	21
$\mathbf{R}^{-1}$	the converse of $\mathbf{R}$	16
$\mathbf{R}^\#$	the congruence generated by $\mathbf{R}$	25
$\rho^\natural$	the natural map from X onto X/ρ	20
$\rho \circ \sigma$	the composition of ρ and σ	16
ρ_s	right translation by s	49
$\mathcal{RN}$	the variety of right normal bands	132
$\mathcal{RZ}$	the variety of right zero semigroups	112
$\lvert S \rvert$	the order of a semigroup S	1
S-ENS	the class of left S-systems	252
S-ENS-T	the class of S,T-bisystems	252
σ	the minimum group congruence	160
$\mathcal{SL}$	the variety of semilattices	112
S^1	the semigroup S with adjoined identity	2
S^0	the semigroup S with adjoined zero	2
S/I	the Rees quotient of S by I	33
$S_1 *_U S_2$	the free product of S_1 and S_2 amalgamating U	263
$S(e,f)$	the sandwich set	58
$\mathrm{Sing}_{\lvert X \rvert}$	the semigroup of singular selfmaps of X	234
$\mathcal{S}[Y; S_\alpha; \phi_{\alpha,\beta}]$	a strong semilattice of semigroups	105
$S \simeq T$	S is isomorphic to T	5
$S \times T$	the cartesian product of S and T	5
$S \otimes_U T$	the tensor product of S and T over U	255
$\mathcal{T}$	the variety of trivial semigroups	112
T_E	the Munn semigroup	162
$T_{e,f}$	the set of isomorphisms from Ee onto Ef	162
$\mathcal{T}_X$	the full transformation semigroup on X	6
$\tau_{\max}$	the greatest congruence with trace τ	158
$\tau_{\min}$	the least congruence with trace τ	158
$\operatorname{tr} \rho$	the trace of an inverse semigroup congruence	155
$t(w)$	the tail (last letter) of w	126
$\mathcal{U}$	the uniformity relation	162

$[U; S_i; \phi_i]$	a semigroup amalgam	263
$u(0)$	the initial of u	120
$\overline{u}(0)$	the initial mark of u	120
$u(1)$	the terminal of u	120
$\overline{u}(1)$	the terminal mark of u	120
$V(a)$	the set of inverses of a	52
$\mathcal{Z}$	the variety of null semigroups	112
$\lvert w \rvert$	the length of a word	239
$w^{\downarrow}$	the set of left factors of w	203
$\overline{w}$	the group-reduced word associated with w	203
$[w_1 = z_1, \ldots]$	a variety defined by laws	112
(X_i, α_{ij})	a direct system	273
(X_n, α_n)	a direct system	276
$x \otimes y$	an element of a tensor product	255
$\mathbf{Z}$	the set of integers	

Index

α-walk (in a word tree), 210
abelian semigroup, 2
absolutely closed semigroup, 272
absolutely extendable
    monoid, 289
    semigroup, 301
absolutely flat monoid, 289
absolutely free algebra, 109
acyclic component, 232
adjunction
    of identity, 2
    of zero, 2
almost unitary subsemigroup, 314
alphabet, 238
amalgamation base, 301, 303
amalgamation property, 307
anti-symmetric relation, 13
anti-uniform semilattice, 163, 166
archimedean semigroup, 138
associative law, 1
atom (in a lattice), 127
automaton, 238
automorphism, 5

Baer–Levi semigroup, 38, 61
band, 113
    of groups, 139
base (of a submonoid of A^*), 240
bicyclic semigroup, 32, 60, 100, 163, 170, 171
bimap, 254
binary operation, 1
binary relation, 16
biprefix code, 245
bisimple inverse semigroup, 169, 174, 217
bisystem, 252
Brandt groupoid, 219
Brandt semigroup, 152, 309
Bruck–Reilly extension, 171, 174, 215, 216

cancellative congruence, 62
cancellative semigroup, 60, 239
cartesian product, 254
central element of a semigroup, 107
centralizer, 160
centre of a semigroup, 107
chain, 212
Clifford semigroup, 107, 215, 217
closed
    inverse subsemigroup, 188
    set, 154
    submonoid, 266
    subsemigroup, 295
closure, 153
code, 244
codomain (of a map), 5
commutative semigroup, 2, 113
comparable elements, 167
compatible relation, 22, 152
complete
    lattice, 14
    lower semilattice, 14
    upper semilattice, 14
completely 0-simple semigroup, 70
completely regular semigroup, 78, 103
completely simple semigroup, 77
component (of a digraph), 231
congruence, 22
    on an inverse semigroup, 154
    on an S-system, 253
    pair, 156

congruence-free semigroup, 93, 98, 101
conjugate elements, 248
consistent subsemigroup, 314
content of a word, 120
converse (of a relation), 17
coproduct, 262
core (of an amalgam), 263
covering relation, 35, 167
cross-section of an equivalence, 63
cycle, 38, 212
cyclic
  component, 232
  group, 10
  semigroup, 9

defining relations, 31
degree (of a polynomial), 239
dense submonoid, 266
diagonal relation, 16
direct
  limit, 274
  product, 5, 6, 109
  system, 274
0-direct union, 80
disjoint union, 278
division relation, 42
domain
  of a map, 5
  of a relation, 17
domination, 266
dominion, 266, 268

E-step, 286
E-unitary
  cover, 199
  inverse semigroup, 192, 215, 217
  regular semigroup, 64, 139
effective representation, 186, 192
element
  of finite order, 10
  of infinite order, 9
elementary **R**-transition, 27
embedding of an amalgam, 263, 265
empty word, 29
endomorphism, 5
epimorphism, 266
equality relation, 16
equational class of algebras, 111
equidivisible semigroup, 239, 241
equivalence, 19
  classes, 20
  generated by a relation, 21
equivalent representations, 186
extended right regular representation, 7
extension (of a map), 17
extension property
  for monoids, 288
  for semigroups, 289, 292
extract (of a sandwich matrix), 85

F-inverse semigroup, 218
factorizable inverse semigroup, 199
faithful representation, 6, 185
Fibonacci sequence, 249
fibre product, 132
finite chain of groups, 181
finitely based variety, 112
finitely presented semigroup, 31
finite presentation, 31
flat S-system, 289, 303
four-spiral semigroup, 101
free
  band, 119, 140
  Clifford semigroup, 218
  group, 202
  inverse monoid, 200
  inverse semigroup, 202
  left normal semigroup, 142
  left zero semigroup, 142
  monoid, 29, 238
  normal band, 142
  product of an amalgam, 263
  product (of groups), 259

product (of semigroups), 258
S-extension, 281, 296
S-system, 310
semigroup, 29, 238
semilattice, 142
full inverse subsemigroup, 155, 165
full transformation semigroup, 6, 39, 63, 96, 254
fully invariant congruence, 124
function, 19
fundamental inverse semigroup, 161

generalized inverse semigroup, 222
generating set, 9
generators, 9, 31
greatest lower bound, 14
Green's equivalences, 45, 55, 59, 123, 172, 179, 238
Green's Lemmas, 49, 54, 74, 175
Green's Theorem, 50, 73
group, 3
amalgam, 251
as amalgamation base, 303
free product, 266
group-bound semigroup, 76
group-reduced word, 202
groupoid, 1
group-with-zero, 4

Hasse diagram, 15
head (of a word), 127
homological algebra, 254
homomorphism, 5

I-semigroup, 102
ideal, 4
idempotent, 4
idempotent semigroup, 113
idempotent-separating congruence, 160
idempotent-separating equivalence, 58
identical relation, 111
identity element, 2

image
of a map, 5
of a relation, 17
incomparable elements, 167
index
of an element, 10
of a factor in a free product, 258
infinite monogenic semigroup, 9
initial, 121
initial mark, 121
inner left and right translations, 115
inverse of an element, 51
inverse semigroup, 103, 145, 272, 303, 307
amalgam, 307
morphism, 147
inverse subsemigroup, 154, 211
irregular $\mathcal{D}$-class, 51
Isbell's Zigzag Theorem, 271, 272
isomorphic, 5
isomorphism, 5

join, 14

kernel, 10
of a component, 231
of a congruence, 155
of a map, 20
of a semigroup, 68
of a Rees morphism, 33

Lallement order, 235
Lallement's Lemma, 57, 98, 230
languages, 238
lattice, 14
law, 111
lazy tongs semilattice, 216
least inverse semigroup congruence, 230, 236
least upper bound, 14

left
cancellative semigroup, 61
compatible relation, 22
congruence, 22, 253
factor (of an element of a monoid), 242
ideal, 4
identity element, 37
normal band, 133
simple semigroup, 61
translation, 115
unitary subsemigroup, 63
zero element, 38
zero semigroup, 3, 113
lexicographic order, 168, 216
linked left and right translations, 115
linked triple, 86, 90
locally inverse semigroup, 222, 223, 225
lower bound, 13
lower semilattice, 14, 193

M-step, 286
McAlister triple, 194
map, 19
maximal
condition, 13
element, 13
prefix code, 249
maximum
element, 13
group morphic image, 160
idempotent-separating congruence, 160, 217
meet, 14
0-minimal ideal, 67
minimal
condition, 13, 167
conditions on principal ideals, 48
element, 13
ideal, 68, 95
minimum
element, 13
group congruence, 62, 160
ideal, 68
Mitsch order, 235
modular lattice, 34
monogenic semigroup, 9, 95
monoid, 2
extension property, 289
free product, 266
morphism, 5
monomorphism, 5
morphic image, 109
morphism, 5
of Ω-algebras, 109
Morse code, 244
Munn semigroup, 162
Munn word tree, 210

Nambooripad order, 223, 235
natural embedding of an amalgam, 264, 265
natural order (on an inverse semigroup), 152
nilpotent element, 70
normal
band, 133
congruence, 155
inverse subsemigroup, 155
sandwich matrix, 82, 83, 236
null semigroup, 3, 113, 139

Ω-algebra, 108
ω-semigroup, 171
order
of a semigroup, 1
order of an element, 9
automorphism, 193
ideal, 193
relation, 13
ordered set, 13
ordinal product, 216
orthodox band of groups, 139
orthodox semigroup, 58, 139, 222, 226

partially ordered set, 13
partial map, 17
partial order relation, 13
partial transformation semigroup, 63
partition, 20
period (of an element), 10
periodic semigroup, 12
polynomial, 239
prefix code, 245
prefix property, 244
prim inverse semigroup, 217
primitive element, 249
primitive idempotent, 70
principal
  factor, 69, 95, 98
  ideal, 45
  series of a semigroup, 95
proper
  congruence, 83
  ideal, 5
  inverse semigroup, 217
pullback, 132
pure monomorphism, 289, 296
pushout, 264, 277

quasi-unitary
  submonoid, 291
  subsemigroup, 291, 292
quotient semigroup (by a congruence), 23
quotient set (by an equivalence), 20

$\mathcal{R}$-unipotent semigroup, 222, 235
range (of a map), 5
rectangular
  0-band, 96
  band, 7, 96, 113
  group, 139
Rees
  congruence, 33
  Isomorphism Theorem, 81
  matrix semigroup, 71, 97, 99, 151, 234
  morphism, 33
  quotient, 69, 253
  Theorem, 72
  Theorem (normalized version), 82, 83
Rees–Suschkewitsch Theorem, 77
reflexive relation, 13
regular
  $\mathcal{D}$-class, 51, 216
  element, 50
  sandwich matrix, 70
  semigroup, 50
relational morphism, 42
relatively free algebra, 110
relatively unitary subsemigroup, 293
representation, 6
representation (of an inverse semigroup), 185
restriction (of a map), 17, 153
right
  cancellative semigroup, 61
  compatible relation, 22
  congruence, 22
  coset, 188, 212
  group, 61, 98
  ideal, 4
  identity element, 37
  normal band, 133
  ω-coset, 188
  simple semigroup, 61, 98
  translation, 115
  unitary subsemigroup, 63
  zero element, 38
  zero semigroup, 3, 113
right-inner zigzag, 273

S-act, 252
S-morphism, 253
S-operand, 252
S-set, 252
S-step, 286

S-system, 252
sandwich set, 59
Sardinas–Patterson algorithm, 245
saturated set of words, 203
self-conjugate inverse subsemigroup, 155
semiband, 223, 230
semigroup, 1
  amalgam, 251, 262
  of binary relations, 63
  with involution, 102
  with length, 239, 241
  of linear maps, 63
  morphism, 5, 147
  $\mathcal{P}_X$ of all partial maps, 18
  $\mathcal{T}_X$ of all maps, 19
  with zero, 2
semilattice, 113
  of completely simple semigroups, 105
  of idempotents, 146
semimodular lattice, 35
semisimple semigroup, 95
simple
  inverse semigroup, 176, 178, 217
  monoid, 216
  semigroup, 66
0-simple semigroup, 66
spined product, 132
splitting monomorphism, 315
standard component, 232
Stirling number, 42, 96
strong semilattice of completely simple semigroups, 106
subalgebra of an Ω-algebra, 109
subdirect product, 130
subgroup, 4
sublattice (of a lattice), 14
subsemigroup, 4
subsystem, 253
subtransitive inverse subsemigroup, 178
subuniform semilattice, 177, 216
suffix code, 245
sum (of representations), 186
symmetric
  group, 6, 38
  inverse semigroup, 149, 211, 215
  relation, 19
syntactic congruence, 28
system, 252

tail (of a word), 127
tensor product, 254, 269
terminal, 121
terminal mark, 121
totally ordered set, 13
total order relation, 13
trace (of a congruence), 155
transformation semigroup, 6
transitive
  closure (of a relation), 21
  relation, 13
  representation, 186, 192
  subsemigroup, 170
transitivity relation, 185
translation, 115
translational hull, 115, 139
transposition, 39
transversal of an equivalence, 63
trimap, 257
trivial component, 232
trivial semigroup, 113

U-semigroup, 102
uniformity relation, 162
uniform semilattice, 163, 166, 169, 216
unitary subsemigroup, 63, 291
universal relation, 16
upper saturation, 154
upper semilattice, 14

Vagner–Preston representation, 185
variable length code, 244
variety of algebras, 110

weak direct product, 130
weakly embeddable amalgam, 267
well-ordered set, 13, 167
word
    in an alphabet, 238
    in a free product, 259
    tree, 210

zero divisors, 67
zigzag, 271
Zigzag Theorem, 271, 272

9 780198 511946