# The Classification of Dualisable Aperiodic Semigroups

A thesis submitted in total fulfilment of the requirements for the degree of Doctor of Philosophy

College of Science, Health, and Engineering
Department of Mathematics and Statistics
La Trobe University
Victoria, Australia

Submitted in May 2021 by

Timothy James Koussas
Bachelor of Science (Honours)

# Abstract

A fundamental problem in the theory of natural dualities is to determine when a finite algebra is dualisable; this is called the dualisability problem. While we are still a long way from a general solution to this problem, it has been solved in certain classes. In the realm of finite semigroups, we have characterised the dualisable members within the classes of groups, bands, nilpotent semigroups, and completely (zero-)simple semigroups. In each of these classes, the dualisable members turned out to be precisely those that generate a residually small variety, but there has been no direct explanation for why this is the case.

The main result of this thesis concerns aperiodic semigroups—semigroups whose subgroups are all trivial. The result completely classifies the dualisable finite aperiodic semigroups; it is therefore a significant extension of the results for bands and nilpotent semigroups, and provides a solid foundation for tackling the dualisability problem in the class of all finite semigroups.

The thesis is split into two parts. Part 1 provides the semigroup theory needed to prove the main result, and also gives a detailed account of the classification theorem for residually small semigroup varieties (originally proved by McKenzie, Golubov, and Sapir). Part 2 is dedicated to the new classification theorem; of course, much of the theory developed in Part 1 will be used in the proof. The new results reveal a connection with the finite q-basis problem, and also make the connection with residual smallness even more intriguing.

∘ ∘ ∘

This work was supported by an Australian Government Research Training Program Scholarship, the La Trobe University David Myers Research Scholarship, and the La Trobe University Professor Edgar Smith Scholarship.

Except where reference is made in the text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis accepted for the award of any other degree or diploma. No other person's work has been used without due acknowledgment in the main text of the thesis. This thesis has not been submitted for the award of any degree or diploma in any other tertiary institution.

Tim Koussas

25 May 2021

# Acknowledgements

Mum. Thank you for literally everything. I hope this makes it somewhat worth having me.

Dad. Thank you for teaching me the most important skill I have: learning how to learn.

James. Thank you for going on this mathematics journey with me—together but separately. I'll be honoured to graduate with you.

Stufful. Thank you for keeping my heart alive.

I have many friends to thank, but for good politics I will save the specifics for personal communication. Thank you all for the support, patience, and all of the side-adventures (i.e., life). I must, however, give a special mention to Zack, who had the hardest job of all: living with me while I was writing this. You have my eternal love and gratitude.

To my supervisors, Marcel and Brian. I got there in the end, but I must have had you worried more than a few times. Thank you for putting so much faith in me and letting me do things my way.

Finally, to any PhD students that might be reading this. You are good enough, and you will get through it—if it is what you really want. But I think, deep down, you knew what you were getting into, so don't let the difficult times stop you from achieving what you want.

# Contents

# Part 1

# Semigroup Theory

# Preface to Part 1

The main result of this thesis is a classification theorem for the dualisable finite aperiodic semigroups. Naturally, understanding this result requires a significant amount of background knowledge, which is part of the reason that the thesis is divided into halves. However, while this first half of the thesis is closely connected to the main result presented in Part 2, it is intended to stand alone as it own story—that story being the classification of residually small varieties of semigroups.

Residually small semigroup varieties were first characterised, modulo certain group-theoretic aspects, by McKenzie [**47**]. Independently, Golubov and Sapir [**32**] characterised the closely related residually finite semigroup varieties, which in fact coincide with the residually small varieties in the finitely generated case. In Chapter 3, we give an account of the theory developed in these two papers. Our presentation takes the best elements from each approach, and, with the benefit of hindsight, simplifies many of the more challenging aspects of the main proof.

The purpose of the first two chapters, then, is to provide a foundation for Chapter 3. In view of the mathematical backgrounds of the author, the supervisors, and the anticipated readership, the development of semigroup theory is approached from the perspective of universal algebra. Thus, Chapter 1 briefly covers the universal algebraic concepts that will be used throughout the thesis, with the assumption that the reader is already familiar with the content. The thesis begins properly in Chapter 2, where we present a thorough development of the basic semigroup theory needed for Chapter 3 and beyond.

CHAPTER 1

# Universal Algebra

To orient the reader, we will cover in this short introductory chapter the required background from universal algebra. Our goal is to fix notation and record some results, so we will not explain the theory in detail, nor will we provide any substantial proofs. To the reader unfamiliar with universal algebra, we recommend a standard text such as Bergman [**4**] or Grätzer [**34**], where the concepts in this chapter are covered in much more detail.

## 1.1. Orders and lattices

Order theory is the bread and butter of universal algebra, though it is of course a well-established theory in its own right, extending to domains far outside of universal algebra. The small part of order theory included in this section adopts its notation and general philosophy from Davey and Priestley [**19**], which continues to be the gold standard for introductory order theory.

Let $\mathbf{P} = \langle P; \leqslant \rangle$ be an ordered set (so $\leqslant$ is reflexive, antisymmetric, and transitive). One of the most basic ideas in order theory is the fact that the dual ordered set $\mathbf{P}^{\partial} := \langle P; \geqslant \rangle$ is also an ordered set, which allows order-theoretic notions to be defined in dual pairs. It is worth mentioning at the outset that order-theoretic duality is distinct from the notion of natural duality that we will focus on in the latter half of the thesis.

We say that $\mathbf{P}$ is a **chain** if for every $x, y \in P$, we have $x \leqslant y$ or $y \leqslant x$ (that is, each pair of elements in $\mathbf{P}$ is comparable). We call $\mathbf{P}$ an **antichain** if there are no strict order relations within $\mathbf{P}$; in other words, for all $x, y \in P$ with $x \leqslant y$, we have $x = y$.

We use the term **down-set** for a decreasing subset of $\mathbf{P}$; that is, $U \subseteq P$ is a down-set of $\mathbf{P}$ if for all $x, y \in P$ with $x \leqslant y$ and $y \in U$, we have $x \in U$. **Up-sets** are defined dually. The set of all down-sets of $\mathbf{P}$ forms an ordered set under inclusion, denoted by $\mathcal{O}(\mathbf{P})$. The **principal down-set** generated by $a \in P$ is defined as ${\downarrow}a := \{x \in P \mid x \leqslant a\}$.

For $x, y \in P$, we write $x \prec y$, and say $y$ **covers** $x$, if $x < y$ and there is no $z \in P$ such that $x < z < y$. Our first result describes the covering relation in $\mathcal{O}(\mathbf{P})$.

**Proposition 1.1.** *Let $\mathbf{P}$ be an ordered set, and let $U, V \in \mathcal{O}(\mathbf{P})$ with $U \subseteq V$. Then $U \prec V$ in $\mathcal{O}(\mathbf{P})$ if and only if $V \backslash U$ is a singleton set.*

Let $U \subseteq P$. The **join** of $U$ is defined as the least element of $\{x \in P \mid (\forall y \in U)\ \ x \geqslant y\}$, if such a least element exists (it is unique if it does exist). The join of $U$ is denoted by $\bigvee U$. We define the **meet** of $U$ dually, and denote it by $\bigwedge U$ when it exists.

We call an ordered set $\mathbf{L}$ a **lattice** if $x \vee y := \bigvee\{x, y\}$ and $x \wedge y := \bigwedge\{x, y\}$ exist for all $x, y \in L$. If $\mathbf{L}$ has the stronger property that $\bigvee U$ and $\bigwedge U$ exist for all $U \subseteq L$, then $\mathbf{L}$ is called a **complete lattice**. The ordered set $\mathcal{O}(\mathbf{P})$ is an example of a complete lattice, in which join and meet are given by union and intersection, respectively.

Since union and intersection can be viewed as special cases of join and meet, we will often use notation analogous to $\bigvee U$ and $\bigwedge U$ for union and intersection. If $\mathcal{U}$ is a set of subsets of some set $X$, we write $\bigcup \mathcal{U}$ for the union of all sets in $\mathcal{U}$ and $\bigcap \mathcal{U}$ for the intersection of all sets in $\mathcal{U}$. (Interpreting $\bigcap$ as meet in the power set of $X$, this gives $\bigcap \varnothing = X$.)

We assume the reader is familiar with the appropriate notions of isomorphisms for ordered sets and lattices. We will usually use the following characterisation (or definition) of order-isomorphisms: if $\mathbf{P}$ and $\mathbf{Q}$ are ordered sets and $\varphi \colon P \to Q$ is a map, then $\varphi$ is an an order-isomorphism if and only if it is surjective and satisfies $x \leqslant y \Leftrightarrow \varphi(x) \leqslant \varphi(y)$ for all $x, y \in P$. Often, the easiest way to verify that two lattices are isomorphic (as lattices) is to use the following elementary result (see [**19**, Proposition 2.19(ii)]).

**Proposition 1.2.** *Let $\mathbf{L}$ and $\mathbf{K}$ be lattices, and let $\varphi \colon \mathbf{L} \to \mathbf{K}$ be an order-isomorphism. Then $\varphi$ is a lattice isomorphism.*

Let $\mathbf{L}$ be a complete lattice, and let $a \in L$. We say that $a$ is **completely meet irreducible** in $\mathbf{L}$ if $a = \bigwedge A$ implies $a \in A$, for every $A \subseteq L$; **completely join irreducible** elements are defined dually. We will often use the fact that $a$ is completely meet irreducible in $\mathbf{L}$ if and only if $a < \bigwedge\{x \in L \mid a < x\}$.

We say that $a$ is **compact** in $\mathbf{L}$ if for every $A \subseteq L$ with $a \leqslant \bigvee A$, there exists a finite subset $B$ of $A$ such that $a \leqslant \bigvee B$. A complete lattice $\mathbf{L}$ is called **algebraic** if every element of $\mathbf{L}$ is a join of compact elements of $\mathbf{L}$, and $\mathbf{L}$ is called **dually algebraic** if the dual lattice $\mathbf{L}^\partial$ is algebraic. The following result concerning algebraic lattices is a simple application of Zorn's Lemma (see [**19**, Exercise 10.13]).

**Lemma 1.3.** *Let $\mathbf{L}$ be a complete lattice. If $\mathbf{L}$ is algebraic, then every element of $\mathbf{L}$ is a meet of completely meet irreducible elements of $\mathbf{L}$.*

## 1.2. Equivalence relations

Here, we will fix some notation regarding equivalence relations. Let $A$ be a set, and let $\theta$ be an equivalence relation on $A$. For each $a \in A$, we denote by $a/\theta$ the equivalence class of $a$ with respect to $\theta$, and we write $A/\theta$ for the partition $\{a/\theta \mid a \in A\}$ corresponding to $\theta$.

We denote by $\Delta_A$ the trivial equivalence relation $\{(a, a) \mid a \in A\}$. If $\theta$ is an equivalence relation on $A$ and $B \subseteq A$, we denote by $\theta{\restriction}_B$ the equivalence relation $\theta \cap B^2$ (i.e., the restriction of $\theta$ to $B$). We will also use the same notation for restrictions of maps; i.e., if $f$ is a function with domain $A$ and $B \subseteq A$, then $f{\restriction}_B$ is the restriction of $f$ to $B$.

Let $f$ be a function with domain $A$. Then the **kernel** of $f$, denote by $\ker(f)$, is the equivalence relation $\{(x, y) \in A^2 \mid f(x) = f(y)\}$.

The set of all equivalence relations on $A$ of course forms a complete lattice, with meet given by intersection. Joins are less straightforward to describe, but we will not need to do so in general. The next result describes joins in a special case. Note that for equivalence relations $\alpha$ and $\beta$ on $A$, we define $\alpha \circ \beta := \{(x, y) \in A^2 \mid (\exists z \in A)\ \ x \mathrel{\alpha} z \mathrel{\beta} y\}$.

**Proposition 1.4.** *Let $\alpha$ and $\beta$ be equivalence relations on a set $A$ such that $\alpha \circ \beta = \beta \circ \alpha$. Then the join $\alpha \vee \beta$ in the lattice of equivalence relations on $A$ equals $\alpha \circ \beta$.*

## 1.3. Preorders

Along with order relations, the weaker notion of a **preorder** (a reflexive and transitive binary relation) occurs quite commonly in semigroup theory. As we will see, all semigroups are equipped with three very naturally-defined preorders, fundamental to semigroup theory in general and this thesis in particular. We will consider here one aspect of preorders that is best understood in the abstract.

Let $S$ be a set and $\preccurlyeq$ a preorder on $S$. The **natural equivalence relation induced by** $\preccurlyeq$ is the equivalence relation $\theta$ on $S$ defined by

$$(\forall x, y \in S)\ \ x\ \theta\ y \iff (x \preccurlyeq y\ \ \&\ \ y \preccurlyeq x).$$

The following result is easily proved.

**Proposition 1.5.** *Let $\preccurlyeq$ be a preorder on a set $S$, and let $\theta$ be the natural equivalence relation induced by $\preccurlyeq$. Define the binary relation $\leqslant$ on $S/\theta$ by*

$$(\forall x, y \in S)\ \ x/\theta \leqslant y/\theta \iff x \preccurlyeq y.$$

*Then $\leqslant$ is a well-defined order relation on $S/\theta$.*

## 1.4. Varieties and quasivarieties

We assume the reader has internalised the concept of algebraic signature or type, as well as the concepts of homomorphism, isomorphism, embedding, endomorphism, automorphism, subalgebra, congruence, and direct product. Algebras will usually be denoted by bold upper case Roman letters, and their underlying sets by the corresponding unbolded symbols.

The isomorphism relation will be written as $\mathbf{A} \cong \mathbf{B}$, and the subalgebra relation will be written as $\mathbf{A} \leqslant \mathbf{B}$. If $\varphi\colon \mathbf{A} \to \mathbf{B}$ is a homomorphism, then we have $\mathbf{A}/\ker(\varphi) \cong \varphi(\mathbf{A})$; this fundamental isomorphism theorem will be used freely without mention.

The standard class operators will be denoted as follows. Let $\boldsymbol{\mathcal{K}}$ be a class of (similar) algebras; then

$$\begin{aligned}
\mathbb{H}(\boldsymbol{\mathcal{K}}) &:= \{\mathbf{A} \mid \mathbf{A} \text{ is a homomorphic image of a member of } \boldsymbol{\mathcal{K}}\};\\
\mathbb{I}(\boldsymbol{\mathcal{K}}) &:= \{\mathbf{A} \mid \mathbf{A} \text{ is isomorphic to a member of } \boldsymbol{\mathcal{K}}\};\\
\mathbb{S}(\boldsymbol{\mathcal{K}}) &:= \{\mathbf{A} \mid \mathbf{A} \text{ is isomorphic to a subalgebra of a member of } \boldsymbol{\mathcal{K}}\};\\
\mathbb{P}(\boldsymbol{\mathcal{K}}) &:= \{\mathbf{A} \mid \mathbf{A} \text{ is isomorphic to a direct product of members of } \boldsymbol{\mathcal{K}}\}.
\end{aligned}$$

We will omit braces when applying these class operators to finite sets given in list form; for example, $\mathbb{I}(\{\mathbf{A}, \mathbf{B}\})$ will be written as $\mathbb{I}(\mathbf{A}, \mathbf{B})$.

A class of (similar) algebras is called a **variety** if it is closed under $\mathbb{H}$, $\mathbb{S}$, and $\mathbb{P}$. The smallest variety containing the class $\boldsymbol{\mathcal{K}}$ is precisely $\mathbb{HSP}(\boldsymbol{\mathcal{K}})$. Accordingly, the composite operator $\mathbb{HSP}$ will be denoted by $\mathbb{V}$; thus, $\mathbb{V}(\boldsymbol{\mathcal{K}}) = \mathbb{HSP}(\boldsymbol{\mathcal{K}})$ is the variety generated by $\boldsymbol{\mathcal{K}}$.

The smallest class closed under $\mathbb{S}$ and $\mathbb{P}$ that contains $\boldsymbol{\mathcal{K}}$ is $\mathbb{SP}(\boldsymbol{\mathcal{K}})$. We will frequently use the following well-known characterisation of $\mathbb{SP}(\boldsymbol{\mathcal{K}})$ in terms of point separation. Recall that, for a set $A$ and a set $F$ of maps with domain $A$, the set $F$ is said to **separate the points of** $A$ if, for all $a, b \in A$ with $a \neq b$, there exists $f \in F$ with $f(a) \neq f(b)$. If $\boldsymbol{\mathcal{B}}$ is a set of algebras, then an algebra $\mathbf{A}$ lies in $\mathbb{SP}(\mathcal{B})$ if and only if the set $\{\mathbf{A} \to \mathbf{B} \mid \mathbf{B} \in \mathcal{B}\}$ of homomorphisms separates the points of $A$. This fact will be used freely without mention.

Let $\mathbf{A}$ be an algebra and $\{\mathbf{A}_i \mid i \in I\}$ a set of algebras similar to $\mathbf{A}$. Denote by $\pi_j$ the projection map $\prod_{i\in I} \mathbf{A}_i \to \mathbf{A}_j$, for each $j \in I$. If $\mathbf{A}$ is isomorphic to some $\mathbf{A}' \leqslant \prod_{i\in I} \mathbf{A}_i$ such that $\pi_j(A') = A_j$ for each $j \in I$, then $\mathbf{A}$ is called a **subdirect product of** $\{\mathbf{A}_i \mid i \in I\}$. It is easy to see that $\mathbf{A}$ is a subdirect product of $\{\mathbf{A}_i \mid i \in I\}$ if and only if there is a set $\{\varphi_i \colon \mathbf{A} \to \mathbf{A}_i \mid i \in I\}$ of surjective homomorphisms that separates the points of $A$.

An algebra $\mathbf{A}$ is **subdirectly irreducible** if the trivial congruence $\Delta_A$ is completely meet irreducible in the lattice of all congruences on $\mathbf{A}$; this ensures that $\mathbf{A}$ is not properly decomposable as a subdirect product [**4**, §3.3]. Consequently, a subdirectly irreducible algebra $\mathbf{A}$ has a least non-trivial congruence, which we will call the **monolith** of $\mathbf{A}$. An algebra $\mathbf{A}$ is subdirectly irreducible if and only if there is a pair $(a, b) \in A^2 \backslash \Delta_A$ such that $a\ \theta\ b$ for every congruence $\theta \neq \Delta_A$, in which case we say that $(a, b)$ **generates the monolith of $\mathbf{A}$**. Also, a subdirectly irreducible algebra $\mathbf{A}$ satisfies $\mathbf{A} \in \mathbb{SP}(\mathcal{K}) \Rightarrow \mathbf{A} \in \mathbb{S}(\mathcal{K})$ for every class $\mathcal{K}$.

Though we expect the reader is completely familiar with the following fundamental result of Garrett Birkhoff (proved in [**4**, Theorem 3.24]), it will be useful to refer to an explicit statement while we are deep into certain proofs later on.

**Theorem 1.6** (Birkhoff's Subdirect Decomposition Theorem)**.** *Let $\mathbf{A}$ be an algebra. For every $a, b \in A$ with $a \neq b$, there is a congruence $\theta$ on $\mathbf{A}$ such that $\mathbf{A}/\theta$ is subdirectly irreducible and $(a/\theta, b/\theta)$ generates the monolith of $\mathbf{A}/\theta$. Consequently, $\mathbf{A}$ is a subdirect product of its subdirectly irreducible homomorphic images.*

If $\mathcal{K}$ is a class of algebras, then $\mathbf{si}(\mathcal{K})$ will denote the class of all subdirectly irreducible members of $\mathcal{K}$. The following corollary of Theorem 1.6 expresses the crucial fact that varieties are determined by their subdirectly irreducible members.

**Corollary 1.7.** *If $\mathcal{V}$ is a variety, then $\mathcal{V} = \mathbb{SP}(\mathbf{si}(\mathcal{V})) = \mathbb{V}(\mathbf{si}(\mathcal{V}))$.*

A variety $\mathcal{V}$ is called residually small if there is a cardinal bound on the sizes of the subdirectly irreducibles in $\mathcal{V}$. We will study such varieties in Chapter 3, so subdirectly irreducibles will be particularly important. Since the definition of residual smallness also involves cardinals, we will take a moment to specify some conventions. First, an ordinal number will be thought of as being equal to the set of all smaller ordinals; so, for example, when thinking of 3 as an ordinal number, it will be identified with the set $\{0, 1, 2\}$. Importantly, $\omega$ will denote the smallest infinite ordinal $\{0, 1, \dots\}$, which is essentially the set $\mathbb{N}_0$ of non-negative integers. Now, a **cardinal** is an ordinal number that is not in bijection with any of its elements. By the Well-Ordering Principle, every set is in bijection with precisely one cardinal, called its cardinality. For a thorough treatment of ordinals and cardinals, see Dugundji [**25**, Chapter II].

Next, we move on to the important concept of a free algebra. We will not give the definition of a free algebra here since it will not be used explicitly, but we assume the reader is aware of the usual definition by universal mapping property; see [**4**, §4.3]. Let the class $\mathcal{K}$ be closed under $\mathbb{S}$ and $\mathbb{P}$, and assume that $\mathcal{K}$ is non-trivial (i.e., $\mathcal{K}$ contains an algebra with more than one element). Then, for every non-empty set $X$, there is a unique algebra in $\mathcal{K}$ (up to isomorphism) that is freely generated by $X$, which will be denoted by $\mathbf{F}_{\mathcal{K}}(X)$.

One of the main reasons that free algebras are so important is that they connect semantics with syntax. It is a consequence of the existence of free algebras in varieties that a class of algebras is a variety if and only if it can be defined by universally quantified equations (or identities) [**4**, Theorem 4.41]. This is another incredibly important result of Birkhoff's, typically used unconsciously by universal algebraists.

The following result ([**4**, Corollary 4.39]) allows us to show that an equation is satisfied by an entire variety simply by showing that two elements in a free algebra are equal. It will be used several times in our study of semigroup varieties in Chapter 3.

**Lemma 1.8.** *Let $\mathcal{V}$ be a variety, and let $s(x_1, \ldots, x_n) \approx t(x_1, \ldots, x_n)$ be an identity in the signature of $\mathcal{V}$, where $x_1, \ldots, x_n$ are variables belonging to some set $X$. If the elements $s^{\mathbf{F}}(x_1, \ldots, x_n)$ and $t^{\mathbf{F}}(x_1, \ldots, x_n)$ are equal in the free algebra $\mathbf{F} := \mathbf{F}_{\mathcal{V}}(X)$, then $\mathcal{V}$ satisfies the identity $s(x_1, \ldots, x_n) \approx t(x_1, \ldots, x_n)$.*

In Lemma 1.8 and several other situations we will encounter, it will be important to distinguish between terms and term functions (also called term operations; see [**4**, §4.3]). Given an $n$-ary term $s(x_1, \ldots, x_n)$ and an algebra $\mathbf{A}$ (in the same signature), the induced term function $s^{\mathbf{A}} \colon A^n \to A$ will be indicated by a superscript $\mathbf{A}$ (as in Lemma 1.8, where $s^{\mathbf{F}}$ is the function $F^n \to F$ induced by the formal expression $s(x_1, \ldots, x_n)$). The superscripts on term functions will often be omitted when there is no danger of ambiguity.

A **quasivariety** is a class defined by a set of **quasiequations** (also called **quasi-identities**), which are universally quantified implications of the form

$$s_1 \approx t_1 \ \ \& \ \ \cdots \ \ \& \ \ s_k \approx t_k \to s \approx t,$$

where $s_1, \ldots, s_k, s$ and $t_1, \ldots, t_k, t$ are (finitary) terms in some set of variables. For example, the quasiequation $xy \approx 1 \to xy \approx yx$ holds in all groups (in the signature $\{\cdot, {}^{-1}, 1\}$), and expresses the fact that inverse pairs commute. We regard equations as quasiequations for which $k = 0$ in the above template, so all varieties are quasivarieties.

If $\varepsilon$ is a quasiequation (or an equation) and $\mathbf{A}$ is an algebra, we write $\mathbf{A} \models \varepsilon$ if $\mathbf{A}$ satisfies $\varepsilon$. If $\mathcal{K}$ is a class of algebras, then $\mathcal{K} \models \varepsilon$ means that $\mathbf{A} \models \varepsilon$ for every $\mathbf{A} \in \mathcal{K}$. For example, if $\mathcal{V}$ is a variety of groups in the signature $\{\cdot, {}^{-1}, 1\}$, then $\mathcal{V} \models xy \approx 1 \to xy \approx yx$. The quasivariety generated by the class $\mathcal{K}$ is defined as the class of all algebras satisfying all quasiequations $\varepsilon$ for which $\mathcal{K} \models \varepsilon$.

Quasivarieties are closed under $\mathbb{S}$ and $\mathbb{P}$, but not necessarily under $\mathbb{H}$. If $\mathcal{K}$ is a finite set of finite algebras, then the quasivariety generated by $\mathcal{K}$ is $\mathbb{SP}(\mathcal{K})$, by [**4**, Theorem 5.6] and [**33**, Corollary 2.3.4]. This fact will be critically important in the second half of the thesis, and we will use it without mention.

Note that an algebra $\mathbf{A}$ satisfies a quasiequation $\varepsilon$ if and only if every finitely generated subalgebra of $\mathbf{A}$ satisfies $\varepsilon$; this is simply because quasiequations are quantified over finitely many variables. The following result is an immediate consequence of this observation.

**Theorem 1.9.** *Let $\mathbf{A}$ be an algebra. Then $\mathbf{A}$ is in the quasivariety* (*and therefore in the variety*) *generated by the finitely generated subalgebras of $\mathbf{A}$.*

An algebra $\mathbf{A}$ is called **locally finite** if every finitely generated subalgebra of $\mathbf{A}$ is finite, and a class $\mathcal{K}$ of algebras is called **locally finite** if every member of $\mathcal{K}$ is locally finite. If $\mathcal{K}$ is closed under $\mathbb{S}$ (e.g., if $\mathcal{K}$ is a quasivariety), then $\mathcal{K}$ is locally finite if and only if every finitely generated member of $\mathcal{K}$ is finite. The following result is proved in [**4**, Theorem 3.49].

**Theorem 1.10.** *If $\mathcal{K}$ is a finite set of finite algebras, then $\mathbb{V}(\mathcal{K})$ is locally finite.*

A variety will be called **finitely generated** if it is of the form $\mathbb{V}(\mathcal{K})$ for some finite set $\mathcal{K}$ of finite algebras; or, equivalently, if it is of the form $\mathbb{V}(\mathbf{A})$ for some finite algebra $\mathbf{A}$. Analogously, a quasivariety is called **finitely generated** if it is of the form $\mathbb{SP}(\mathcal{K})$ for some finite set $\mathcal{K}$ of finite algebras. Theorem 1.10 says, then, that finitely generated quasivarieties and varieties are locally finite.

The last general result of universal algebra that we will need is the following simple lemma; the proof may be extracted from the proof of [**4**, Theorem 3.49].

**Lemma 1.11.** *Let $\mathbf{A}$ and $\mathbf{B}$ be finite algebras with $\mathbf{A} \in \mathbb{HSP}(\mathbf{B})$. Then $\mathbf{A}$ is a homomorphic image of some finite algebra in $\mathbb{SP}(\mathbf{B})$.*

### 1.5. Semilattices

To transition into our study of semigroups, we will briefly discuss semilattices. An ordered set $\mathbf{S}$ is called a **semilattice** if $x \wedge y := \bigwedge\{x, y\}$ exists for all $x, y \in S$. We note that this really defines a meet semilattice; join semilattices are defined dually, but, in semigroup theory, we almost exclusively think of semilattices as meet semilattices.

Since meet operations are associative, semilattices can be regarded as semigroups. In this situation, we will denote the meet operation multiplicatively, as we will with most semigroups. Thus, from the point of view of semigroup theory, we can define a **semilattice** to be a semigroup that is commutative and idempotent; i.e., a semigroup $\mathbf{S} = \langle S; \cdot \rangle$ such that $\mathbf{S} \models xy \approx yx$ and $\mathbf{S} \models x^2 \approx x$.

As with lattices, one can translate between semilattices as ordered sets and semilattices as algebras. If $\mathbf{S}$ is a semilattice regarded as a semigroup, then the **semilattice order** on $\mathbf{S}$ is the order relation defined by $x \leqslant y \Leftrightarrow xy = x$, for all $x, y \in S$. If $\mathbf{S}$ was originally obtained from an ordered set, then this semilattice order agrees with the original order.

In our study of semigroup varieties, we will often define (quasi)varieties directly from (quasi)equations. We will use the standard square bracket notation, and we will take associativity as given. For example, $[xy \approx yx,\ x^2 \approx x]$ denotes the semigroup variety defined by $xy \approx yx$ and $x^2 \approx x$, which is of course the variety of semilattices.

We will close the chapter with a simple application of Corollary 1.7.

**Theorem 1.12.** *The two-element semilattice $\mathbf{I}$ is the only subdirectly irreducible semilattice. Consequently, $\mathbb{SP}(\mathbf{I})$ is the variety $[xy \approx yx,\ x^2 \approx x]$ of all semilattices.*

Proof. Let $\mathbf{S}$ be a subdirectly irreducible semilattice, and let $(a, b)$ generate its monolith. We can assume by symmetry that $a \nleqslant b$. Now, define $F := \{x \in S \mid a \leqslant x\}$. Then $\{F, S \backslash F\}$ is a partition of $\mathbf{S}$ with $a \in F$ and $b \in S \backslash F$, and the corresponding equivalence relation is a congruence on $\mathbf{S}$ and so must be trivial. Thus, $F$ and $S \backslash F$ are singleton sets, so $\mathbf{S} \cong \mathbf{I}$. $\square$

CHAPTER 2

# Semigroups for the Working Algebraist

The development of semigroup theory in this chapter is geared towards those that are already familiar with the universal algebra and order theory outlined in Chapter 1. Though such a reader likely has some prior knowledge of semigroups, no such knowledge is assumed.

This chapter was painstakingly arranged in service of the ultimate goal of the thesis, which is not only to prove the main result, but to give the reader a deep understanding of the result and its significance in the broader dualisability problem for semigroups. While the semigroup theorist is not likely to find any new results in this chapter, the author hopes that they might read it and find something interesting in the presentation.

It would perhaps have been permissible to condense this chapter and the next into a few logically minimalistic pages, giving references for all of the proofs. Indeed, the substantial results in this chapter can be found in standard texts such as Howie [**37**] or Clifford and Preston [**9**]. However, the relationships between these results are just as important as the results themselves, and in understanding these relationships, the devil is truly in the detail. Some of the tiniest arguments took weeks or even months for the author to uncover, all for the sake of presenting a seamless and efficient development of the required background.

A lot of my time would have been saved if this chapter had already existed. If nothing else, these notes will be of use to anyone who might tackle the general dualisability problem for semigroups. For such a person, I will gladly give up some more of my time.

## 2.1. Ideals and Green's $\mathcal{J}$ relation

The reader has most likely encountered the group-theoretic composition series, which allows groups to be broken down into simple groups. The composition series is defined in terms of normal subgroups, so there is no direct generalisation for semigroups. However, in his 1940 paper [**59**], David Rees explored the idea of using ideals in place of normal subgroups. This allows any semigroup to be broken down into more well-behaved parts, and while the decomposition is somewhat coarser than that available to groups, it turns out to be exactly the right concept for semigroups.

In this section, we will explain the fundamental ideas of Rees's decomposition. We take advantage of the reader's assumed familiarity with order theory, resulting in a somewhat unconventional introduction to semigroups, but perhaps, for the universal algebraist, a more natural one.

The first ingredient in Rees's decomposition was a suitable analogue of a quotient by a normal subgroup. This is now called a Rees quotient, where a semigroup is factored by an ideal. The construction should be transparent to those acquainted with abstract algebra, so we will just give the definitions.

An **ideal** of a semigroup $\mathbf{S}$ is a non-empty subset $I$ of $S$ such that for every $x \in I$ and every $a \in S$, we have $ax, xa \in I$. More compactly, $I$ is an ideal of $\mathbf{S}$ if $I \neq \varnothing$ and $SI, IS \subseteq I$.

If $I$ is an ideal of a semigroup $\mathbf{S}$, then $\theta_I := \Delta_S \cup I^2$ is a congruence on $\mathbf{S}$ whose only non-trivial equivalence class is $I$. In the factor semigroup $\mathbf{S}/\theta_I$, called a **Rees quotient**, the element $I$ acts as a multiplicative zero element. It is customary to write $\mathbf{S}/I$ rather than $\mathbf{S}/\theta_I$, with the understanding that we are factoring by the congruence $\Delta_S \cup I^2$.

In his semigroup-theoretic version of the composition series, Rees's original approach in [**59**] was to consider chains of ideals. However, concepts introduced in 1951 by James Alexander Green [**35**] allowed for a simpler approach that expanded the utility of Rees's theory. As is now tradition, we will develop Rees's ideas using Green's relations, which are now a hallmark of semigroup theory.

We will introduce Green's $\mathcal{J}$ relation in this section; the rest of Green's relations will be defined in the next section. For the following discussion, we will work within a fixed semigroup $\mathbf{S}$.

Any union of ideals of $\mathbf{S}$ is an ideal; any intersection of ideals is an ideal, provided only that the intersection is non-empty. Thus, for each $a \in S$, there exists a smallest ideal containing $a$, given by the intersection of all ideals of $\mathbf{S}$ that contain $a$. We will denote this ideal by $\langle a \rangle$, and call it the **principal ideal** generated by $a$. Any non-empty subset can generate an ideal, but we will exclusively be starting from singleton sets.

We define the preorder $\preccurlyeq_J$ on $S$ as follows. Given $a, b \in S$, we write $a \preccurlyeq_J b$ if $\langle a \rangle \subseteq \langle b \rangle$; equivalently, if $a \in \langle b \rangle$. We then define **Green's $\mathcal{J}$ relation** to be the equivalence relation on $S$ induced by $\preccurlyeq_J$ (see Section 1.3); that is, for all $a, b \in S$, we have

$$a \mathrel{\mathcal{J}} b \iff (a \preccurlyeq_J b \ \ \& \ \ b \preccurlyeq_J a) \iff \langle a \rangle = \langle b \rangle.$$

Thus, two elements of $\mathbf{S}$ are $\mathcal{J}$-equivalent if they generate the same principal ideal.

The equivalence class of $a \in S$ with respect to $\mathcal{J}$ is denoted by $J_a$, as is tradition. When we need to specify which semigroup we are working in, we add the semigroup as a superscript; i.e., we will write expressions such as $\preccurlyeq_J^{\mathbf{S}}$, $\mathcal{J}^{\mathbf{S}}$, and $J_a^{\mathbf{S}}$.

Observe that if $I$ is an ideal of $\mathbf{S}$ containing some $a \in S$, then $\langle a \rangle \subseteq I$ by the definition of $\langle a \rangle$, and so $I$ must also contain any element that generates the same ideal as $a$; that is, $J_a \subseteq I$. It follows that any ideal is a union of $\mathcal{J}$-classes. However, a non-empty union of $\mathcal{J}$-classes need not be an ideal. To understand the relationship between ideals and $\mathcal{J}$-classes, we would like to know exactly when a union of $\mathcal{J}$-classes forms an ideal.

Proposition 1.5 gives us a natural order on the set $S/\mathcal{J}$ of $\mathcal{J}$-classes of $\mathbf{S}$. Given $a, b \in S$, we will write $J_a \leqslant J_b$ if $a \preccurlyeq_J b$. Equipping the $\mathcal{J}$-classes with this order, which we call the **$\mathcal{J}$-order**, we can readily describe the collections of $\mathcal{J}$-classes whose unions form ideals.

**Proposition 2.1.** *Let $\mathbf{S}$ be a semigroup, and let $\mathcal{U}$ be a non-empty set of $\mathcal{J}$-classes of $\mathbf{S}$. Then $\bigcup \mathcal{U}$ is an ideal of $\mathbf{S}$ if and only if $\mathcal{U}$ is a down-set of $\langle S/\mathcal{J}; \leqslant \rangle$.*

PROOF. Assume that $\bigcup \mathcal{U}$ is an ideal of $\mathbf{S}$. To show that $\mathcal{U}$ is a down-set, let $a, b \in S$ with $J_b \leqslant J_a$ and $J_a \in \mathcal{U}$. Then, since $b \preccurlyeq_J a$ and $\bigcup \mathcal{U}$ is an ideal, we have $b \in \langle a \rangle \subseteq \bigcup \mathcal{U}$, so $b \in \bigcup \mathcal{U}$, and therefore $J_b$ must be one of the $\mathcal{J}$-classes in $\mathcal{U}$.

Conversely, assume that $\mathcal{U}$ is a down-set of $\langle S/\mathcal{J}; \leqslant\rangle$. To show that $I := \bigcup\mathcal{U}$ is an ideal, let $a \in I$, and let $x \in \langle a\rangle$, so $x \preccurlyeq_J a$. Then $J_x \leqslant J_a \in \mathcal{U}$, so $J_x \in \mathcal{U}$, and hence $x \in I$. This shows that $\langle a\rangle \subseteq I$ for all $a \in I$, so $I$ is a union of principal ideals, hence an ideal. □

The bijective correspondence in Proposition 2.1 is in fact an order-isomorphism: given two down-sets of $\mathcal{J}$-classes $\mathcal{U}$ and $\mathcal{V}$, we have $\bigcup\mathcal{U} \subseteq \bigcup\mathcal{V} \Rightarrow \mathcal{U} \subseteq \mathcal{V}$ because the $\mathcal{J}$-classes partition $S$, while the reverse implication is basic set theory. Hence, by Proposition 1.2, we have the following result.

**Proposition 2.2.** *Let* $\mathbf{S}$ *be a semigroup, let* $\mathcal{I}_0(\mathbf{S})$ *denote the set of ideals of* $\mathbf{S}$ *along with* $\varnothing$*, and let* $\mathbf{J}$ *denote the ordered set* $\langle S/\mathcal{J}; \leqslant\rangle$ *of* $\mathcal{J}$*-classes of* $\mathbf{S}$*. Then* $\mathcal{I}_0(\mathbf{S})$ *and* $\mathcal{O}(\mathbf{J})$ *are isomorphic lattices; an isomorphism* $\mathcal{O}(\mathbf{J}) \to \mathcal{I}_0(\mathbf{S})$ *is given by the mapping* $\mathcal{U} \mapsto \bigcup\mathcal{U}$.

**Example 2.3.** As a brief illustration of Proposition 2.2, we will find all ideals of the six-element semigroup, $\mathbf{S}$, given in Figure 2.1.

| $\cdot$ | $a$ | $b$ | $c$ | $u$ | $v$ | $d$ |
|---|---|---|---|---|---|---|
| $a$ | $a$ | $a$ | $c$ | $u$ | $u$ | $d$ |
| $b$ | $b$ | $b$ | $c$ | $v$ | $v$ | $d$ |
| $c$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |
| $u$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |
| $v$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |
| $d$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |

FIGURE 2.1. A six-element semigroup.

It can easily be verified that the principal ideals are

$$\langle c\rangle = \langle d\rangle = \{c, d\}, \quad \langle u\rangle = \langle v\rangle = \{u, v, c, d\}, \quad \langle a\rangle = \langle b\rangle = \{a, b, u, v, c, d\} = S.$$

It follows that the partition into $\mathcal{J}$-classes is $S/\mathcal{J} = \{\{c, d\}, \{u, v\}, \{a, b\}\}$, and the $\mathcal{J}$-order is given by $\{c, d\} < \{u, v\} < \{a, b\}$; thus, $\langle S/\mathcal{J}; \leqslant\rangle$ is a three-element chain. The non-empty down-sets of this ordered set are $\{\{c, d\}\}$, $\{\{c, d\}, \{u, v\}\}$, and $\{\{c, d\}, \{u, v\}, \{a, b\}\}$, so by taking the unions of these we find that $\{c, d\}$, $\{c, d, u, v\}$, and $S$ are the only ideals of $\mathbf{S}$. In this semigroup, all ideals happen to be principal, though this fact does not have anywhere near the significance it does in ring theory.

In Example 2.3, we note the presence of a smallest ideal. Given the size of the example, it is perhaps unclear whether this is a coincidence, but minimum ideals are more common than the universal algebraist may first expect.

An infinite intersection of ideals of a semigroup $\mathbf{S}$ may be empty and therefore not an ideal. However, a *finite* intersection of ideals is always non-empty, for given ideals $I, J$ of $\mathbf{S}$, we have $IJ \subseteq I \cap J$. So, while the set of ideals of $\mathbf{S}$ may not form a complete lattice, it does always form a lattice, which implies the following result.

**Theorem 2.4.** *If a semigroup* $\mathbf{S}$ *has a minimal ideal* $M$*, then* $M$ *is the least ideal of* $\mathbf{S}$*. In particular, any finite semigroup has a minimum ideal.*

Does the minimum ideal have any special properties? Certainly it is a subsemigroup, as is any ideal, but more can be said. To be precise, let **S** be a semigroup with a minimum ideal, **M**, regarded as a semigroup in its own right. Then **M** has no proper ideals (i.e., no ideals other than $M$). This is not difficult to prove, but it is perhaps somewhat surprising. While it is tautologically true that **M** contains no smaller ideals of **S**, it is not immediately obvious that **M** contains no smaller ideals within itself.

Rather than proving this now, it will pay to extend our attention to the other $\mathcal{J}$-classes. We are not so lucky that every $\mathcal{J}$-class is a subsemigroup (though this is perhaps one of the reasons why semigroup theory is interesting). We therefore cannot study any given $\mathcal{J}$-class in exactly the same way that we might study the minimum ideal. However, by forming an appropriate Rees quotient, we can force any $\mathcal{J}$-class into a position of minimality.

We know that a principal ideal $\langle a\rangle$ in a semigroup **S** is, by virtue of being an ideal, a union of $\mathcal{J}$-classes. But, in the case of principal ideals, there is a particularly neat description of which $\mathcal{J}$-classes they comprise. From the down-set point of view, $\langle a\rangle$ corresponds to the principal down-set $\downarrow J_a$; precisely, we have

$$\langle a\rangle = \bigcup\{J_x \mid x \in S \ \ \& \ \ J_x \leqslant J_a\} = \{x \in S \mid J_x \leqslant J_a\}.$$

The fact that principal ideals correspond to principal down-sets is a direct consequence of Proposition 2.2.

Now, let $\mathcal{U} := \{J_x \mid x \in S \ \ \& \ \ J_x < J_a\}$, the set of $\mathcal{J}$-classes *strictly* below $J_a$. Clearly, $\mathcal{U}$ is a down-set of $\mathcal{J}$-classes. If $\mathcal{U}$ is empty, then $J_a$ is the minimum ideal of **S**, but otherwise, $I := \bigcup\mathcal{U}$ is an ideal of **S**, and the set difference $\langle a\rangle\backslash I$ is the $\mathcal{J}$-class $J_a$.

So the situation is that our desired object of study, $J_a$, lives inside the semigroup $\langle a\rangle$, and the remaining elements form an ideal $I$. If we then form the Rees quotient $\langle a\rangle/I$, what we end up with is essentially the $\mathcal{J}$-class $J_a$ with an additional zero element that represents the products that fall into lower $\mathcal{J}$-classes. The factor semigroups so constructed will enable us to study the $\mathcal{J}$-classes of a semigroup in isolation, almost as if they were subsemigroups. We call them principal factors.

For a precise definition, let **S** be a semigroup, let $a \in S$, and let $I := \{x \in S \mid J_x < J_a\}$, the union of the $\mathcal{J}$-classes strictly below $J_a$. If $I$ is non-empty, then we define the **principal factor of S determined by** $J_a$ to be the Rees quotient $\langle a\rangle/I$. Of course, $I$ could be empty, in which case we define the principal factor of **S** determined by $J_a$ to be the minimum ideal $J_a = \langle a\rangle$.

**Example 2.5.** Let us calculate the three principal factors of the semigroup **S** from Example 2.3. One is the minimum ideal, $\{c, d\}$, which forms a **right-zero semigroup**; i.e., a semigroup satisfying $xy \approx y$.

The principal factor determined by $\{u, v\}$ is the quotient $\{u, v, c, d\}/\{c, d\}$. The elements are $\{u\}$, $\{v\}$, and $\{c, d\}$. As always, the ideal $\{c, d\}$ is a zero element in the Rees quotient, but all products in this quotient equal the zero element. In general, a semigroup with a zero element in which all products equal zero is called a **null semigroup**. Thus, the principal factor $\{u, v, c, d\}/\{c, d\}$ is the three-element null semigroup.

Finally, the principal factor determined by $\{a, b\}$ equals $\mathbf{S}/\{u, v, c, d\}$, which is isomorphic to the following three-element semigroup.

| $\cdot$ | $a$ | $b$ | $0$ |
|---|---|---|---|
| $a$ | $a$ | $a$ | $0$ |
| $b$ | $b$ | $b$ | $0$ |
| $0$ | $0$ | $0$ | $0$ |

This is a two-element left-zero semigroup $\{a, b\}$ with an adjoined zero. For this principal factor, it happens that the zero element is indecomposable; in other words, the $\mathcal{J}$-class $\{a, b\}$ is a subsemigroup. The minimum $\mathcal{J}$-class $\{c, d\}$ is also a subsemigroup, of course. In contrast, all products in the $\mathcal{J}$-class $\{u, v\}$ fall into strictly lower $\mathcal{J}$-classes (corresponding to the principal factor being null). This kind of dichotomy occurs in a surprisingly broad class of semigroups, including almost all semigroups we study in the second half of this thesis. We will elaborate on this in Chapter 3.

In forming principal factors, a broader view of the situation is that we have ideals $I \subset J$ such that the set difference $J \backslash I$ is a single $\mathcal{J}$-class. By combining Proposition 2.2 and Proposition 1.1, it is easy to see that this implies $I \prec J$ in the lattice of ideals. Conversely, if $I \prec J$ in the lattice of ideals, then $J \backslash I$ is a $\mathcal{J}$-class, by the same results.

This tells us in particular that the principal ideal $\langle a \rangle$ is minimal with respect to strictly containing the ideal $I = \{x \in S \mid J_x < J_a\}$. The next result, which is easily proved, shows that this minimality property carries through when we factor the entire semigroup by $I$.

**Proposition 2.6.** *Let $\mathbf{S}$ be a semigroup, let $I$ be an ideal of $\mathbf{S}$, and let $\mathbf{L}$ denote the lattice of ideals of $\mathbf{S}$ that contain $I$. Then $\mathbf{L}$ is isomorphic to the lattice $\mathcal{I}(\mathbf{S}/I)$ of ideals of $\mathbf{S}/I$; an isomorphism $\mathbf{L} \to \mathcal{I}(\mathbf{S}/I)$ is given by the mapping $J \mapsto J/I$.*

We now arrive at the crucial property of principal factors. From Proposition 2.6, we deduce that a principal factor $\langle a \rangle / I$ is an atom (i.e., covers the least element) in the lattice of ideals of $\mathbf{S}/I$, since $I \prec \langle a \rangle$ in the lattice of ideals of $\mathbf{S}$. This leads us to an important piece of terminology: given an arbitrary semigroup $\mathbf{S}$ with a zero element, an ideal of $\mathbf{S}$ is called **zero-minimal** if it is minimal in the ordered set of non-zero ideals of $\mathbf{S}$.

So, in other words, we have the following.

**Proposition 2.7.** *Let $\mathbf{S}$ be a semigroup, let $a \in S$, and let $I := \{x \in S \mid J_x < J_a\}$. If $I \neq \varnothing$, then the principal factor $\langle a \rangle / I$ is a zero-minimal ideal of $\mathbf{S}/I$.*

Knowing that a principal factor exists as a zero-minimal ideal in some naturally-chosen semigroup, we would now like to know what special properties zero-minimal ideals possess in general.

If $\mathbf{S}$ is a semigroup with a zero element 0, then $\mathbf{S}$ is called **zero-simple** if $\mathbf{S}$ is not a null semigroup (i.e., $SS \neq \{0\}$) and has precisely two ideals ($S$ and $\{0\}$). A **simple** semigroup is a semigroup $\mathbf{S}$ such that $S$ is the only ideal of $\mathbf{S}$. (A simple semigroup can be trivial.)

**Remark.** The semigroup-theoretic definition of simplicity is less restrictive than the 'usual' definition in, say, group theory or universal algebra. A simple semigroup need not be simple in the sense of having precisely two congruences; for example, every group is simple in the semigroup-theoretic sense, but certainly not in the group-theoretic sense. In semigroup theory, a semigroup with precisely two congruences is called 'congruence-free'.

Our goal is to prove is that a zero-minimal ideal of a semigroup is either a null semigroup (all products equal zero) or a zero-simple semigroup. Proposition 2.7 then tells us that principal factors have one of these properties, with the exception of the minimal ideal. Conveniently, we can study the minimum ideal at the same time by using the following construction (which we essentially encountered in Example 2.5).

Given a semigroup $\mathbf{S}$, we define $\mathbf{S}^0$ to be the semigroup obtained from $\mathbf{S}$ by adding a new zero element, which, context permitting, we will denote by 0. Importantly, if $I \subseteq S$, then $I$ is a minimal ideal of $\mathbf{S}$ if and only if $I \cup \{0\}$ is a zero-minimal ideal of $\mathbf{S}^0$. Thus, if we can prove zero minimal ideals are zero-simple, it will follow that minimum ideals must be simple.

The result that we wish to prove can be stated as follows.

**Theorem 2.8.** *Let $\mathbf{S}$ be a semigroup with zero, and let $M$ be a zero-minimal ideal of $\mathbf{S}$. Then the subsemigroup $\mathbf{M}$ formed by $M$ is either zero-simple or null.*

The remainder of this section will be devoted to proving Theorem 2.8. The proof will be assisted by the following characterisation of zero-simple semigroups.

**Lemma 2.9.** *Let $\mathbf{S}$ be a non-trivial semigroup with zero. Then $\mathbf{S}$ is zero-simple if and only if $SaS = S$ for all $a \in S \backslash \{0\}$.*

Proof. Assume that $\mathbf{S}$ is zero-simple. Since $SS$ is a non-zero ideal of $\mathbf{S}$, we have $SS = S$, and hence $SSS = S$. Now, consider the set $I := \{x \in S \mid SxS = \{0\}\}$, which is easily seen to be an ideal of $\mathbf{S}$. Then we must have $I = \{0\}$, because $I = S$ would contradict $SSS = S$. Thus, if $a \in S\backslash\{0\} = S\backslash I$, then $SaS \neq \{0\}$, so $SaS = S$.

Conversely, assume that $SaS = S$ for all $a \in S\backslash\{0\}$. Let $I$ be a non-zero ideal of $S$, and choose some $a \in I\backslash\{0\}$. Then $S = SaS \subseteq SIS \subseteq I$, so $I = S$. Hence $\mathbf{S}$ is zero-simple. □

Because a semigroup $\mathbf{S}$ is simple if and only if $\mathbf{S}^0$ is zero-simple, a result concerning zero-simple semigroups will usually imply a result concerning simple semigroups. For example, if $\mathbf{S}$ is a semigroup, then applying Lemma 2.9 to $\mathbf{S}^0$ tells us that $\mathbf{S}$ is simple if and only if $SaS = S$ for all $a \in S$. We will usually state such results only for zero-simple semigroups, but apply them freely to simple semigroups in this manner.

Our final ingredient for the proof of Theorem 2.8 is the concrete description of principal ideals, which we have delayed discussing as long as possible. The description is by no means deep nor complicated, but we wished to highlight the way in which the theory of $\mathscr{J}$-classes can be developed by more abstract means.

First, we define $\mathbf{S}^1$ to be the semigroup obtained from the semigroup $\mathbf{S}$ by adding a new identity element, 1. Naturally, we denote the underlying set of $\mathbf{S}^1$ by $S^1$. We will always use a superscript '1' to refer to this construction (as opposed to the first Cartesian or direct power). Now, given $a \in S$, the set $S^1aS^1$ is an ideal of $\mathbf{S}$ containing $a$, and every ideal containing $a$ must contain $S^1aS^1$; thus, $\langle a \rangle = S^1aS^1$. The use of the notation $S^1$ here is merely to abbreviate the description of principal ideals; without this notation, we would have to write $\langle a\rangle = \{a\} \cup aS \cup Sa \cup SaS$.

**Remark.** Our definition of $\mathbf{S}^1$ is non-standard; usually, $\mathbf{S}^1$ is defined to be $\mathbf{S}$ in the case that $\mathbf{S}$ already has an identity. This makes essentially no difference in practice, so we prefer the less complicated definition. Those familiar with some category theory may also find it somewhat pleasant that, with our definition, the construction $\mathbf{S} \mapsto \mathbf{S}^1$ defines a left adjoint to the forgetful functor from monoids to semigroups.

Proof of Theorem 2.8. Assume that $M$ is not null, so $MM \neq \{0\}$; we will show that $\mathbf{M}$ is zero-simple.

First, since $MM \subseteq M$ is a non-zero ideal of $\mathbf{S}$, we must have $MM = M$ by zero-minimality, and so $MMM = M$. To apply Lemma 2.9, let $a \in M\backslash\{0\}$. Then the principal ideal $\langle a\rangle = S^1aS^1$ is an ideal of $\mathbf{S}$ contained in $M$, so $M = S^1aS^1$ by zero-minimality. Now,

$$M = MMM = M(S^1aS^1)M = (MS^1)a(S^1M) \subseteq MaM \subseteq M,$$

so $M = MaM$. By Lemma 2.9, $\mathbf{M}$ is zero-simple. $\square$

**Remark.** If a zero-minimal ideal is a null semigroup, it does not necessarily have only two ideals. The only null semigroup with precisely two ideals is the two-element one, but any null semigroup can arise as a zero-minimal ideal of some semigroup. The attentive reader will be able to construct relevant examples by the end of Chapter 3.

**Corollary 2.10.** *Let $J$ be a $\mathcal{J}$-class of a semigroup $\mathbf{S}$. If $J$ forms a subsemigroup of $\mathbf{S}$, then it is simple.*

Proof. By adding a zero to $\mathbf{S}$ (which does not affect the abstract properties of $J$ as a subsemigroup), we can assume that $J$ is not the minimum ideal. Let $\mathbf{K}$ denote the principal factor of $\mathbf{S}$ determined by $J$. Clearly $\mathbf{K}$ is not null since $J$ is a subsemigroup, so $\mathbf{K}$ is zero-simple by Proposition 2.7 and Theorem 2.8. Moreover, the zero of $\mathbf{K}$ is indecomposable, and removing this zero yields an isomorphic copy of $J$. This shows that $J$ is simple. $\square$

The results of this section can now be summarised as follows.

**Theorem 2.11.** *Let $\mathbf{S}$ be a semigroup. Then the minimum ideal of $\mathbf{S}$ is simple (if it exists), and every other principal factor of $\mathbf{S}$ is either zero-simple or null.*

A null semigroup is determined up to isomorphism by the cardinality of its underlying set; these objects are indeed quite easily described. Simple and zero-simple semigroups, on the other hand, can be rather complicated. This is hardly surprising upon considering that all groups are simple as semigroups. Less obviously, every semigroup (!) can be embedded into a simple semigroup [**10**, Theorem 8.45].

For our purposes, we will not need to study these objects in complete generality. The zero-simple semigroups we will encounter will always satisfy a certain minimality condition which affords them with a particularly rigid and tangible structure. One of the main goals of this chapter is to expose this structure theory, which was the focus of the latter half of Rees's paper [**59**]. To understand this structure, we must look deeper within the $\mathcal{J}$-classes.

## 2.2. Green's relations $\mathcal{L}$, $\mathcal{R}$, $\mathcal{H}$, and $\mathcal{D}$

In Section 2.1, we took care to develop the theory of $\mathcal{J}$-classes in a certain abstract way. Let us focus in particular on Proposition 2.1 and its immediate consequence Proposition 2.2. In the proofs, both the implicit and the explicit, there was no need to work directly with the definition of an ideal. All we really needed was the fact that a union of ideals is an ideal. In more detail, the abstract set-up was as follows.

Suppose that, for some set $S$, we have a set $\mathcal{A}$ of subsets of $S$ that is closed under arbitrary union and intersection. Then, for each $a \in S$, we may define $\langle a \rangle$ to be the intersection of all members of $\mathcal{A}$ that contain $a$. Now, define a preorder $\preccurlyeq$ on $S$ by putting $a \preccurlyeq b$ if $\langle a \rangle \subseteq \langle b \rangle$, for every $a, b \in S$. From this, we get a natural equivalence relation $\theta$, defined by $a \;\theta\; b \Leftrightarrow a \preccurlyeq b \;\;\&\;\; b \preccurlyeq a$. In turn, this gives us a natural order on the $\theta$-classes, as in Proposition 1.5.

The point here is that this abstract framework was all that we needed to prove Proposition 2.2. So, in the general setting of the previous paragraph, every set in $\mathcal{A}$ is a union of a down-set of $\theta$-classes, with the natural correspondence being a lattice isomorphism.

So, with little effort, we can get analogous results for left and right ideals. A **left ideal** of a semigroup $\mathbf{S}$ is a non-empty subset $L$ of $S$ such that $SL \subseteq L$. A **right ideal** is defined dually, and so a subset of $S$ is an ideal precisely if it is both a left ideal and a right ideal. (An ideal is sometimes called a **two-sided ideal** for emphasis.)

Within a fixed semigroup $\mathbf{S}$, a union of left ideals is a left ideal, and a non-empty intersection of left ideals is a left ideal. The **principal left ideal** of an element $a \in S$ is defined as the smallest left ideal of $\mathbf{S}$ containing $a$, which is easily seen to equal $S^1a$.

Analogously to the relation $\preccurlyeq_J$, we define the preorder $\preccurlyeq_L$ by $a \preccurlyeq_L b \Leftrightarrow S^1a \subseteq S^1b$, for all $a, b \in S$. **Green's $\mathcal{L}$ relation** is the induced equivalence relation, so we have $a \;\mathcal{L}\; b$ in $\mathbf{S}$ if and only if $S^1a = S^1b$. Similarly to the notation for $\mathcal{J}$-classes, we write $L_a$ for the $\mathcal{L}$-class of $a \in S$. And, finally, the set $S/\mathcal{L}$ of $\mathcal{L}$-classes inherits a natural order relation, given by $L_a \leqslant L_b \Leftrightarrow a \preccurlyeq_L b$, for all $a, b \in S$.

The proof of Proposition 2.2 carries over to give us the following result.

**Proposition 2.12.** *Let $\mathbf{S}$ be a semigroup, let $\mathcal{L}_0(\mathbf{S})$ denote the set of left ideals of $\mathbf{S}$ along with $\varnothing$, and let $\mathbf{L}$ denote the ordered set $\langle S/\mathcal{L}; \leqslant\rangle$ of $\mathcal{L}$-classes of $\mathbf{S}$. Then $\mathcal{L}_0(\mathbf{S})$ and $\mathcal{O}(\mathbf{L})$ are isomorphic lattices; an isomorphism $\mathcal{O}(\mathbf{L}) \to \mathcal{L}_0(\mathbf{S})$ is given by the mapping $\mathcal{U} \mapsto \bigcup\mathcal{U}$.*

We of course have the dual definitions and results concerning right ideals. The symbols $\preccurlyeq_R$, $R_a$, and $\mathcal{R}$ are used analogously to $\preccurlyeq_L$, $L_a$, and $\mathcal{L}$.

**Example 2.13.** We will describe the relations $\mathcal{L}$ and $\mathcal{R}$ for the semigroup $\mathbf{S}$ from Examples 2.3 and 2.5. First, the principal left ideals are

$$S^1a = S^1b = \{a, b, c\}, \quad S^1u = S^1v = \{u, v, d\}, \quad S^1c = \{c\}, \quad S^1d = \{d\}.$$

It follows that the partition into $\mathcal{L}$-classes is $S/\mathcal{L} = \{\{c\}, \{d\}, \{u, v\}, \{a, b\}\}$.

The principal right ideals are

$$aS^1 = \{a, u, c, d\}, \quad bS^1 = \{b, v, c, d\}, \quad uS^1 = \{u, c, d\}, \quad vS^1 = \{v, c, d\},$$

and $cS^1 = dS^1 = \{c, d\}$. We therefore have $S/\mathcal{R} = \{\{c, d\}, \{u\}, \{v\}, \{a\}, \{b\}\}$.

From the definitions and the preceding example, it is clear that $\mathcal{L}, \mathcal{R} \subseteq \mathcal{J}$. In particular, we have $\mathcal{L} \vee \mathcal{R} \subseteq \mathcal{J}$ in the lattice of equivalence relations. In Example 2.13, it turned out that $\mathcal{L} \vee \mathcal{R}$ coincides with $\mathcal{J}$, though we will see that this is no coincidence.

The join $\mathcal{D} := \mathcal{L} \vee \mathcal{R}$ is called **Green's $\mathcal{D}$ relation**. We use the notation $D_a$ for the $\mathcal{D}$-class of an element $a$, though the $\mathcal{D}$ relation does not come with a natural preorder in the same way that $\mathcal{L}$, $\mathcal{R}$, and $\mathcal{J}$ do.

For our purposes, the primary significance of the $\mathcal{D}$ relation is that it is a means of studying the $\mathcal{J}$ relation, which is really the object of fundamental interest, but this approach is feasible only when $\mathcal{D} = \mathcal{J}$. This identity will be available to us when needed, but we will give an example to show that it need not hold in general.

**Example 2.14.** The following example is from Howie [**37**, Exercise 2.6.1]. Consider the multiplicative semigroup $\mathbf{S}$ of $2 \times 2$ real matrices with underlying set

$$S = \left\{ \begin{pmatrix} a & 0 \\ b & 1 \end{pmatrix} \,\middle|\, a, b \in \mathbb{R}^+ \right\}.$$

It can be shown that $\mathcal{L} = \mathcal{R} = \mathcal{D} = \Delta_S$ and $\mathcal{J} = S^2$, so here $\mathcal{D}$ and $\mathcal{J}$ are as different as they could possibly be. Such extreme situations are uncommon (particularly for us), but nonetheless, it is important to be aware that $\mathcal{D}$ and $\mathcal{J}$ need not coincide.

The remainder of the present section will be devoted to exposing some of the basic structure of $\mathcal{D}$-classes. We proceed with the next result, which follows immediately from the definitions of its subjects.

**Proposition 2.15.** *Let* $\mathbf{S}$ *be a semigroup. The relations* $\preccurlyeq_L$ *and* $\mathcal{L}$ *are right-compatible; that is, for all* $a, b, c \in S$, *we have*

$$a \preccurlyeq_L b \implies ac \preccurlyeq_L bc \quad \textit{and} \quad a \mathrel{\mathcal{L}} b \implies ac \mathrel{\mathcal{L}} bc.$$

*Dually,* $\preccurlyeq_R$ *and* $\mathcal{R}$ *are left-compatible.*

Among other things, Proposition 2.15 assists us in proving the following fundamental result concerning the $\mathcal{D}$ relation.

**Proposition 2.16.** *Let* $\mathbf{S}$ *be a semigroup. Then* $\mathcal{L} \circ \mathcal{R} = \mathcal{R} \circ \mathcal{L} = \mathcal{D}$.

PROOF. Let $a, b \in S$ with $a \mathrel{\mathcal{L} \circ \mathcal{R}} b$. Then there exists $c \in S$ with $a \mathrel{\mathcal{L}} c \mathrel{\mathcal{R}} b$. Since we have in particular that $a \in S^1 c$ and $b \in cS^1$, we may choose $x, y \in S^1$ with $a = xc$ and $b = cy$. Then $a \mathrel{\mathcal{L}} c$ implies $ay \mathrel{\mathcal{L}} cy = b$ as $\mathcal{L}$ is right-compatible, and by symmetry we have $xb \mathrel{\mathcal{R}} a$. Thus, $a \mathrel{\mathcal{R}} xb = xcy = ay \mathrel{\mathcal{L}} b$, so $a \mathrel{\mathcal{R} \circ \mathcal{L}} b$. This shows that $\mathcal{L} \circ \mathcal{R} \subseteq \mathcal{R} \circ \mathcal{L}$, and the reverse inclusion is by symmetry. By Proposition 1.4 we have $\mathcal{L} \circ \mathcal{R} = \mathcal{L} \vee \mathcal{R} = \mathcal{D}$. □

Owing to Proposition 2.16, it is possible to represent a given $\mathcal{D}$-class using a so-called **eggbox diagram** (Figure 2.2). In such a diagram, the elements of the $\mathcal{D}$-class are grouped into cells arranged in a rectangular grid so that each row of the grid makes up an $\mathcal{R}$-class and each column an $\mathcal{L}$-class. The cells themselves are classes of the equivalence relation $\mathcal{L} \cap \mathcal{R}$. This is the last of Green's relations, known as **Green's $\mathcal{H}$ relation**; that is, $\mathcal{H} := \mathcal{L} \cap \mathcal{R}$. As would be hoped for at this point, we denote by $H_a$ the $\mathcal{H}$-class of a semigroup element $a$. These $\mathcal{H}$-classes are indeed important objects, and we will have more to say about them shortly.

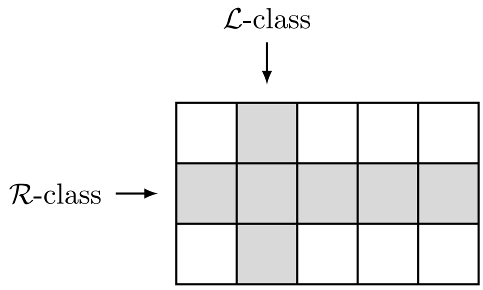



FIGURE 2.2. A generic eggbox diagram.

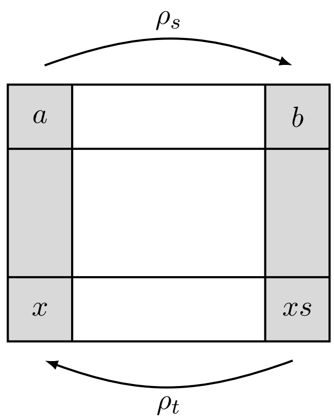
The general situation of Proposition 2.16 is what allows for the existence of eggbox diagrams. However, these diagrams are of particular value for $\mathcal{D}$-classes because they can provide useful information about the locations of products between $\mathcal{D}$-related elements.

The inner workings of $\mathcal{D}$-classes were of course discovered by Green in his seminal paper [**35**]. The following result of Green's distills the interaction between the semigroup operation and the eggbox diagrams. For the statement and proof, we will use the symbol $\rho_s$ to denote right translation by a given semigroup element $s$; that is $\rho_s(x) := xs$, for all $x$ in the given semigroup.

**Lemma 2.17** (Green's Lemma)**.** *Let* $\mathbf{S}$ *be a semigroup, let* $a, b \in S$ *with* $a \mathrel{\mathcal{R}} b$*, and choose any* $s \in S^1$ *with* $as = b$*. Then the right translation* $\rho_s$ *maps* $L_a$ *bijectively onto* $L_b$*. Moreover,* $\rho_s$ *maps each* $\mathcal{H}$*-class contained in* $L_a$ *bijectively onto the* $\mathcal{R}$*-related* $\mathcal{H}$*-class contained in* $L_b$*, so in particular,* $\rho_s$ *maps* $H_a$ *bijectively onto* $H_b$*.*

PROOF. Choose $t \in S^1$ such that $bt = a$. Let $x \in L_a$. By right-compatibility of $\mathcal{L}$ (Proposition 2.15) we have $xs \mathrel{\mathcal{L}} as = b$, so $\rho_s(x) \in L_b$. This shows that $\rho_s$ maps $L_a$ into $L_b$, and by symmetry, the right translation $\rho_t$ maps $L_b$ into $L_a$.

Since $ast = bt = a$, it follows that $st$ is a right identity for $a$ and therefore for every element of $S^1a$. Thus, if $x \in L_a \subseteq S^1a$, then $\rho_t(\rho_s(x)) = xst = x$, showing that $\rho_t \circ \rho_s$ acts as identity on $L_a$. By symmetry $\rho_s \circ \rho_t$ acts as identity on $L_b$, so $\rho_s \colon L_a \to L_b$ is a bijection.



Now, note that for any $x \in L_a$, we have $x \mathcal{R} \rho_s(x)$, as $\rho_s(x) = xs$ can be translated back to $x$ by $\rho_t$. Thus, if $x, y \in L_a$ with $x \mathcal{R} y$, then $\rho_s(x) \mathcal{R} x \mathcal{R} y \mathcal{R} \rho_s(y)$, so the restriction of $\rho_s$ to $L_a$ preserves $\mathcal{R}$. As it also preserves $\mathcal{L}$, the $\mathcal{H}$-classes in $L_a$ are mapped into the $\mathcal{R}$-related $\mathcal{H}$-classes in $L_b$. Symmetric statements hold for $\rho_t$, so the result follows. $\square$

Of course, Green's Lemma implicitly contains the corresponding dual result concerning left translations. Thus, we have the following immediate consequence.

**Corollary 2.18.** *Let* $\mathbf{S}$ *be a semigroup, and let* $a, b \in S$ *with* $a \mathcal{D} b$. *Then*

$$|H_a| = |H_b|, \qquad |L_a| = |L_b|, \qquad |R_a| = |R_b|.$$

A further development of the theory of $\mathcal{D}$-classes will benefit from a discussion of idempotents. Recall that a semigroup element $e$ is **idempotent** if $ee = e$. It is a decidedly important concept in semigroup theory, though its usefulness depends to an extent on a relative abundance of idempotents within a semigroup—enough so that they are able to serve as a skeleton of sorts.

We will be particularly concerned with the way that idempotents govern the structure of their $\mathcal{D}$-classes. First of all, looking within the smaller $\mathcal{L}$-, $\mathcal{R}$-, and $\mathcal{H}$-classes, we can immediately deduce the following.

**Proposition 2.19.** *Let* $\mathbf{S}$ *be a semigroup, and let* $e$ *be an idempotent element of* $\mathbf{S}$. *Then* $e$ *is a left identity for* $R_e$ *and a right identity for* $L_e$. *Consequently,* $e$ *is an identity for* $H_e$.

PROOF. Simply observe that $L_e \subseteq S^1 e$ and $R_e \subseteq eS^1$. $\square$

The relationship between idempotents and $\mathcal{D}$-classes is best explained via the notion of a regular element. An element $a$ of a semigroup $\mathbf{S}$ is called **regular** if there exists $x \in S$ such that $a = axa$. (A semigroup is called **regular** if all of its elements are regular.) The definition of a regular element comes from ring theory, but in the context of semigroup theory, one can naturally arrive at the concept by considering the $\mathcal{L}$- and $\mathcal{R}$-classes of idempotents.

Specifically, let $\mathbf{S}$ be a semigroup, and let $e \in S$ be idempotent. How does $e$ interact with the elements of its $\mathcal{L}$-class?

Take any $a \in L_e$. By Proposition 2.19, we have $ae = a$. Now, we also have $e \in Sa$ (more specifically than $S^1 a$, since $e$ is idempotent), so $e = xa$ for some $x \in S$. Combining this with $ae = a$ gives $a = axa$.

This shows that every element of $L_e$ is regular. Conversely, take a regular element $a \in S$, with $x \in S$ chosen so that $a = axa$. Multiplying on the left by $x$ gives $xa = xaxa$, so $e := xa$ is idempotent, and clearly $a \in Se$ and $e \in Sa$, so $e \mathcal{L} a$.

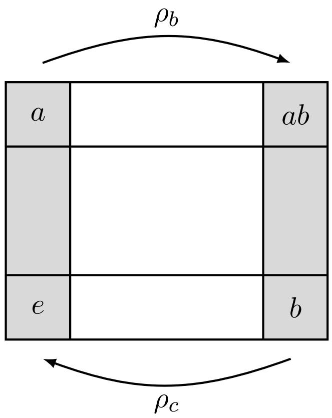
We have shown that a semigroup element is regular if and only if it has an idempotent in its $\mathcal{L}$-class; clearly, then, the $\mathcal{L}$-class of a regular element contains only regular elements. But the definition of a regular element is self-dual, so we can replace '$\mathcal{L}$' by '$\mathcal{R}$' in the previous sentence.

The regularity property therefore spreads throughout the rows and columns of a $\mathcal{D}$-class, saturating it. That is, if a $\mathcal{D}$-class contains an idempotent, every element is regular, leading to the obvious notion of a **regular $\mathcal{D}$-class**: a $\mathcal{D}$-class that contains only regular elements, or equivalently, at least one regular element, or equivalently, at least one idempotent. Within a regular $\mathcal{D}$-class, every $\mathcal{L}$- and $\mathcal{R}$-class contains an idempotent—thus appears an idempotent skeleton.

Using Green's Lemma, we can now derive a result which reveals how idempotents influence the location of products between elements within the same $\mathcal{D}$-class.

**Lemma 2.20.** *Let* $\mathbf{S}$ *be a semigroup, and let* $a, b \in S$ *with* $a \,\mathcal{D}\, b$*. Then* $ab \in R_a \cap L_b$ *if and only if there is an idempotent* $e \in L_a \cap R_b$*.*

PROOF. Assume that $ab \in R_a \cap L_b$, and choose $c \in S^1$ with $abc = a$. Then, by Green's Lemma 2.17, $\rho_c$ maps $H_b$ into $L_a \cap R_b$, with the inverse mapping given by $\rho_b$.




Thus, $bcbc = \rho_c(\rho_b(\rho_c(b))) = \rho_c(b) = bc$, so $e := bc \in L_a \cap R_b$ is idempotent, as required. Conversely, assume there is an idempotent $e \in L_a \cap R_b$. Since $eb = b$ (Proposition 2.19), the right translation $\rho_b$ maps $H_a$ into $R_a \cap L_b$ by Green's Lemma, so $ab \in R_a \cap L_b$. □

To close this section, we will use Lemma 2.20 in conjunction with Green's Lemma 2.17 to prove the following powerful theorem, in which groups make their inevitable entrance into semigroup theory.

**Theorem 2.21** (Green's Theorem)**.** *Let* $\mathbf{S}$ *be a semigroup, and let* $H$ *be an* $\mathcal{H}$*-class of* $\mathbf{S}$*. The following are equivalent:*

(i) $HH \cap H \neq \varnothing$*;*
(ii) $H$ *contains an idempotent element;*
(iii) $H$ *is a subgroup of* $\mathbf{S}$*.*

PROOF. (i) $\Rightarrow$ (ii): Assume that $HH \cap H \neq \varnothing$, so there are $a, b \in H$ with $ab \in H$. Note that $L_a \cap R_b$ and $R_a \cap L_b$ both equal $H$ because $a \,\mathcal{H}\, b$. Thus, by Lemma 2.20, $H$ contains an idempotent element.

(ii) $\Rightarrow$ (iii): Assume there is an idempotent $e \in H$, and let $x, y \in H$. Then $L_x \cap R_y = H$ contains an idempotent, so $xy \in R_x \cap L_y = H$ by Lemma 2.20, showing that $H$ is a subsemigroup.

We must prove that the subsemigroup $H$ is a group. By Proposition 2.19, $e$ is an identity for $H$, so it remains to prove that each element of $H$ has an inverse with respect to $e$.

Let $x \in H$. Then $ex = x$; i.e., right translation by $x$ sends $e$ to $x$. By Green's Lemma 2.17, right translation by $x$ is a permutation of $H$. Thus, $x'x = e$ for some $x' \in H$, which is then a left inverse of $x$. Dually, $x$ also has a right inverse $x^* \in H$, which coincides with $x'$ because $x' = x'e = x'xx^* = ex^* = x^*$. Thus, $x$ has a two-sided inverse, $x'$, in $H$.

Clearly (iii) implies (i), so the proof is complete. □

**Remark.** We have seen that in a regular $\mathcal{D}$-class, every row and column contains an idempotent, and we now know that each idempotent-containing cell must be a group. We also know from Corollary 2.18 that all of these groups have the same order. It can in fact be shown that all of the group $\mathcal{H}$-classes within a givin $\mathcal{D}$-class are isomorphic to each other. We will not prove this fact since we will not need it, but it would be remiss not to mention it.

### 2.3. Completely zero-simple semigroups and Rees's Theorem

Having covered the necessary rudiments of Green's relations, we will now continue the thread of Section 2.1. As we have already mentioned, we will not need to study zero-simple and simple semigroups in full generality, since we will always have a further property available to us. The objects we study in this section will be called completely (zero-)simple semigroups. Like the real number system, there are several definitions which are arguably as natural as each other, and the objects somehow precede their definition. The 'best' definition, then, comes down to aesthetics and context.

We will follow the approach of Clifford and Preston [**9**] and define these semigroups in terms of zero-minimal one-sided ideals. In a semigroup $\mathbf{S}$ with zero, a **zero-minimal left ideal** is a left ideal that is minimal in the set of non-zero left ideals of $\mathbf{S}$. A **zero-minimal right ideal** is defined dually.

A **completely zero-simple semigroup** is a zero-simple semigroup $\mathbf{S}$ with at least one zero-minimal left ideal and at least one zero-minimal right ideal. By removing all instances of the string 'zero-' in the previous sentence, we get the definition of a **completely simple semigroup**. Clearly, a simple semigroup $\mathbf{S}$ is completely simple if and only if $\mathbf{S}^0$ is completely zero-simple. Consequently, we can draw conclusions about completely simple semigroups from results concerning completely zero-simple semigroups.

Our main aim in this section is to understand the structure of these semigroups. The groundwork will be laid by the following three lemmas.

**Lemma 2.22.** *Let $\mathbf{S}$ be a semigroup with zero, and let $L$ be a left ideal of $\mathbf{S}$. Then $L$ is a zero-minimal left ideal if and only if $L \backslash \{0\}$ is an $\mathcal{L}$-class of $\mathbf{S}$.*

Proof. There is a short direct argument, but we instead appeal to Proposition 2.12 and Proposition 1.1. □

**Lemma 2.23.** *Let $\mathbf{S}$ be a zero-simple semigroup. Then*

(i) *$LS = S$ for every non-zero left ideal $L$;*
(ii) *if $\mathbf{S}$ has at least one zero-minimal left ideal, then $\mathbf{S}$ is the union of its zero-minimal left ideals.*

Proof. (i): Let $L$ be a non-zero left ideal of $\mathbf{S}$. Since $LS$ is evidently a two-sided ideal, the only alternative to $LS = S$ is $LS = \{0\}$. However, $LS = \{0\} \subseteq L$ would imply that $L$ is a two-sided ideal, which in turn would imply that $L = S$, so $SS = LS = \{0\}$, contradicting the zero-simplicity of $\mathbf{S}$. We must therefore have $LS = S$.

(ii): Let $L$ be a zero-minimal left ideal of $\mathbf{S}$. Certainly $0 \in L$, so 0 is contained in a zero-minimal left ideal. Let $a \in S\backslash\{0\}$. Since $LS = S$ by (i), we have $a \in Ls$ for some $s \in S$. Evidently, then, $Ls$ is a non-zero left ideal containing $a$. To complete the proof, it will suffice to show that $Ls$ is zero-minimal.

By Lemma 2.22, the non-zero elements of $L$ are $\mathcal{L}$-related. Since $\mathcal{L}$ is right-compatible (Proposition 2.15), it follows that $Ls\backslash\{0\}$ is contained in a single $\mathcal{L}$-class. But $Ls$ is a union of $\mathcal{L}$-classes by Proposition 2.12, so $Ls\backslash\{0\}$ must be a complete $\mathcal{L}$-class. It now follows from Lemma 2.22 that $Ls$ is zero-minimal, as required. □

**Lemma 2.24.** *Let* $\mathbf{S}$ *be a completely zero-simple semigroup. Then, for all* $a \in S\backslash\{0\}$*, we have* $L_a = Sa\backslash\{0\}$ *and* $R_a = aS\backslash\{0\}$*.*

Proof. Let $a \in S\backslash\{0\}$. By Lemma 2.23(ii), we have $a \in K$ for some zero-minimal left ideal $K$. Note that $Sa$ is a left ideal of $\mathbf{S}$ contained in $K$, because $Sa \subseteq SK \subseteq K$.

We claim that $Sa \neq \{0\}$. Suppose, by way of contradiction, that $Sa = \{0\}$. Then $\{0, a\}$ is a non-zero left ideal of $\mathbf{S}$ contained in $K$, so $K = \{0, a\}$. Now $Sa = \{0\}$ implies $KK = \{0\}$. However, we have $KS = S$ by Lemma 2.23(i), and this gives

$$SS = KSKS \subseteq KKS = \{0\}S = \{0\},$$

which is a contradiction. It follows that $Sa \neq \{0\}$, as required. Now, since $\{0\} \subset Sa \subseteq K$, we may conclude that $Sa = K$. By Lemma 2.22, we have $Sa = K = L_a \cup \{0\}$, so by duality we are done. □

We now come to what can be thought of as the essential structural theorem of completely (zero-)simple semigroups. Lemma 2.20 gave us our first hint at how idempotents impose structure on a $\mathcal{D}$-class. The next theorem shows that the property of being completely (zero-)simple imposes just enough extra structure for the idempotents to exert complete control over the locations of products modulo the $\mathcal{H}$ relation.

**Theorem 2.25.** *Let* $\mathbf{S}$ *be a completely zero-simple semigroup. Then*

(i) $S\backslash\{0\}$ *is a regular* $\mathcal{D}$*-class of* $\mathbf{S}$*, and so* $\mathbf{S}$ *is regular;*
(ii) *if* $a, b \in S\backslash\{0\}$*, then* $ab \neq 0$ *if and only if* $ab \in R_a \cap L_b$*.*

Proof. We will first prove item (ii). If $a, b \in S\backslash\{0\}$ with $ab \in R_a \cap L_b$, then clearly $ab \neq 0$. Conversely, let $a, b \in S\backslash\{0\}$ with $ab \neq 0$. Then, by Lemma 2.24, we have $ab \in aS\backslash\{0\} = R_a$ and $ab \in Sb\backslash\{0\} = L_b$, as required.

Now we will prove (i). To show that $S\backslash\{0\}$ is a $\mathcal{D}$-class, let $a, b \in S\backslash\{0\}$; we must show that $a \mathrel{\mathcal{D}} b$. To prove this, we will show that $aSb \neq \{0\}$. Suppose this is not the case, so $aSb = \{0\}$. Then, since $\mathbf{S}$ is zero-simple, we have $SS = S$, and so by Lemma 2.9, we have

$$S = SS = SaSSbS = SaSbS = S0S = \{0\},$$

which is a contradiction. Hence, $aSb \neq \{0\}$, so there is a non-zero element $c \in aSb$. By Lemma 2.24, we then have $c \in aS\backslash\{0\} = R_a$ and $c \in Sb\backslash\{0\} = L_b$, so $a \mathrel{\mathcal{D}} b$.

Since $SS \neq \{0\}$, there exist $a, b \in S \backslash \{0\}$ with $ab \neq 0$. This implies that $ab \in R_a \cap L_b$, so $L_a \cap R_b$ contains an idempotent by Lemma 2.20. The $\mathcal{D}$-class $S \backslash \{0\}$ is therefore regular, and evidently 0 is a regular element, so $\mathbf{S}$ is regular. $\square$

Theorem 2.25 specialises quite neatly to completely simple semigroups. In this case, it will be worth stating a separate result.

**Theorem 2.26.** *Let* $\mathbf{S}$ *be a completely simple semigroup. Then*

(i) $S$ *is a regular* $\mathcal{D}$*-class of* $\mathbf{S}$*, and so* $\mathbf{S}$ *is regular;*
(ii) *if* $a, b \in S$*, then* $ab \in R_a \cap L_b$*;*
(iii) *every* $\mathcal{H}$*-class of* $\mathbf{S}$ *is a subgroup.*

Proof. Statements (i) and (ii) follow from applying Theorem 2.25 to $\mathbf{S}^0$. Now, (ii) implies that every $\mathcal{H}$-class is a subsemigroup, so (iii) follows from Green's Theorem 2.21. $\square$

While Rees's Theorem 2.27 below gives a far more explicit description of the structure of these semigroups, Theorem 2.25 is really the underlying cause of this structure. It is common to appeal to Rees's Theorem, but it is often possible to use Theorem 2.25 instead, and this usually leads to a shorter and cleaner proof. Nonetheless, we will have occasion to use Rees's Theorem, so we will present it in this section.

Rees's Theorem is a representation theorem, so we must define the objects that will represent our completely (zero-)simple semigroups. We know that the non-zero elements of a completely zero-simple semigroup constitute a $\mathcal{D}$-class, and we know that a $\mathcal{D}$-class can be thought of as a rectangular grid of $\mathcal{H}$-classes. We can also determine the $\mathcal{H}$-classes of products provided only that we know which $\mathcal{H}$-classes are groups. One ingredient of the representation, then, is a matrix $P$ that, among other things, records which $\mathcal{H}$-classes in this $\mathcal{D}$-class are groups. A binary matrix would suffice for this purpose, but more information is needed if we are to capture the multiplication perfectly.

As noted in the remark following Theorem 2.21, it can be shown that all group $\mathcal{H}$-classes within a $\mathcal{D}$-class are isomorphic to each other. In particular, if $\mathbf{S}$ is completely zero-simple, then the group $\mathcal{H}$-classes in $S \backslash \{0\}$ are all isomorphic to a single group $\mathbf{G}$. By carefully choosing elements of $\mathbf{G}$ as non-zero entries of the matrix $P$, we can capture enough information to recover $\mathbf{S}$ from $\mathbf{G}$ and $P$. For this reason, $\mathbf{G}$ is often referred to as the 'structure group' of $\mathbf{S}$.

Thus, the representing objects will require two ingredients in their construction: a group $\mathbf{G}$, and a matrix $P$ with entries in $G \,\dot{\cup}\, \{0\}$ (where $\dot{\cup}$ denotes disjoint union). As is tradition, we will use sets $I$ and $\Lambda$ to serve as the dimensions of the non-zero $\mathcal{D}$-class. The set $I$ $(\Lambda)$ should be thought of as indexing the set of $\mathcal{R}$-classes ($\mathcal{L}$-classes), so that the non-zero $\mathcal{D}$-class has dimensions $I \times \Lambda$ when thought of as an eggbox diagram. For convenience, the matrix $P$ will have the 'transpose' dimensions; i.e., $\Lambda \times I$.

So much for the intuition. Let us now detail the construction of the **Rees matrix semigroups**. Let $I$ and $\Lambda$ be non-empty sets, let $\mathbf{G}$ be any group, and let $P$ be any $\Lambda \times I$ matrix with entries in $G \,\dot{\cup}\, \{0\}$. The $(\lambda, i)$-entry of $P$ will be denoted by $p_{\lambda i}$. We define the semigroup $\mathcal{M}^0[\mathbf{G}, P]$ to have underlying set $(I \times \Lambda \times G) \,\dot{\cup}\, \{0\}$; the multiplication is defined so that 0 is a multiplicative zero, and so that the product of the non-zero elements $(i, \lambda, a)$

and $(j, \mu, b)$ is given by

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

Associativity is easily verified: in either bracketing, the product $(i, \lambda, a)(j, \mu, b)(k, \nu, c)$ is non-zero if and only if $p_{\lambda j}$ and $p_{\mu k}$ are both non-zero, in which case both bracketings yield the element $(i, \nu, ap_{\lambda j}bp_{\mu k}c)$.

Now, $\mathcal{M}^0[\mathbf{G}, P]$ is a perfectly good semigroup, but it is not necessarily completely zero-simple, or even zero-simple. We could simply choose $P$ to have all entries equal to 0, which would make $\mathcal{M}^0[\mathbf{G}, P]$ a null semigroup. To remedy this, we obviously need to disbar the zero matrix, but others must also be excluded.

The non-zero entries in $P$ are intended to correspond to the group $\mathcal{H}$-classes. Recalling that a regular $\mathcal{D}$-class has a group $\mathcal{H}$-class in every row and column, we must require that every row and column of $P$ has a non-zero entry. This requirement turns out to be enough to yield a completely zero-simple semigroup.

Assume that $P$ has no zero row or column. Take any non-zero elements $(i, \lambda, a)$, $(j, \mu, b)$ of $\mathcal{M}^0[\mathbf{G}, P]$. To understand the $\mathcal{R}$ relation, we would like to know precisely when there exists a third element $(k, \nu, c)$ such that $(i, \lambda, a)(k, \nu, c) = (j, \mu, b)$. By definition of the product, certainly $i = j$ is necessary; we will show it is also sufficient. Assume that $i = j$, and choose any $k$ such that $p_{\lambda k} \neq 0$ (which is possible because $P$ has no zero row). Then, letting $\nu = \mu$ and $c = (ap_{\lambda k})^{-1}b$, we get

$$(i, \lambda, a)(k, \nu, c) = (i, \lambda, a)(k, \mu, (ap_{\lambda k})^{-1}b) = (i, \mu, ap_{\lambda k}(ap_{\lambda k})^{-1}b) = (i, \mu, b) = (j, \mu, b).$$

By symmetry, we can conclude that $(i, \lambda, a)\ \mathcal{R}\ (j, \mu, b)$ if and only if $i = j$. We can also use symmetry to conclude that $(i, \lambda, a)\ \mathcal{L}\ (j, \mu, b)$ if and only if $\lambda = \mu$. It is now evident that all non-zero elements are $\mathcal{D}$-related, which implies zero-simplicity. Now, by nature of the multiplication, each non-zero $\mathcal{L}$-class, when adjoined with the zero, forms a left ideal that is zero-minimal by Lemma 2.22. Dually, the non-zero $\mathcal{R}$-classes yield zero-minimal right ideals, so $\mathcal{M}^0[\mathbf{G}, P]$ is completely zero-simple, as claimed.

We now present Rees's Theorem, which says that all completely zero-simple semigroups arise from this construction.

**Theorem 2.27** (Rees's Theorem)**.** *Let* $\mathbf{S}$ *be a completely zero-simple semigroup. Then there exist a group* $\mathbf{G}$*, non-empty sets* $I$ *and* $\Lambda$*, and a* $\Lambda \times I$ *matrix* $P$ *over* $G \,\dot\cup\, \{0\}$*, with no zero column or row, such that* $\mathbf{S} \cong \mathcal{M}^0[\mathbf{G}, P]$*.*

We will be applying Rees's Theorem only to completely simple semigroups, so we will instead prove a restricted version of the theorem. First, we require a simpler variant of the Rees matrix semigroup.

In the situation where $P$ has no zero entries, the zero element of $\mathcal{M}^0[\mathbf{G}, P]$ is indecomposable, by definition of the multiplication. If we then remove the zero element, we obtain a completely simple semigroup. We can of course construct such semigroups more directly by starting with a group $\mathbf{G}$ and a $\Lambda \times I$ matrix $P$ over $G$. The Rees matrix semigroup $\mathcal{M}[\mathbf{G}, P]$ is then defined on the underlying set $I \times \Lambda \times G$ with multiplication given by $(i, \lambda, a)(j, \mu, b) := (i, \mu, ap_{\lambda j}b)$, as before.

The following specialisation of Rees's Theorem was essentially uncovered by Suschkewitsch [**64**], over a decade before Rees's result and two decades before Green's introduction of his now-indispensable relations. Needless to say, the modern proof does not at all resemble the original proof of Suschkewitsch (even if the latter were in English). Rees's more general Theorem 2.27 can be proved along the same lines as the following, but there are of course more details to consider in the general case. For said details, see [**37**, Theorem 3.2.3].

**Theorem 2.28.** *Let* $\mathbf{S}$ *be a completely simple semigroup. Then there exist a group* $\mathbf{G}$*, non-empty sets* $I$ *and* $\Lambda$*, and a* $\Lambda \times I$ *matrix* $P$ *over* $G$ *such that* $\mathbf{S} \cong \mathcal{M}[\mathbf{G}, P]$*. Moreover,* $P$ *can be chosen to have a row and a column containing only the identity element of* $\mathbf{G}$*.*

PROOF. Let $G$ be any $\mathcal{H}$-class of $\mathbf{S}$. Then $G$ forms a subgroup $\mathbf{G}$ of $\mathbf{S}$ by Theorem 2.26(iii). Denote the identity element of $\mathbf{G}$ by $e$.

Let $I$ and $\Lambda$ be the sets of $\mathcal{R}$- and $\mathcal{L}$-classes of $\mathbf{S}$, respectively. We will think of $I$ and $\Lambda$ as indexing sets. For each $(i, \lambda) \in I \times \Lambda$, denote by $H_{i\lambda}$ the $\mathcal{H}$-class $i \cap \lambda$, which is non-empty as $S$ is a $\mathcal{D}$-class (Theorem 2.26(i)).

Choose $(i_0, \lambda_0) \in I \times \Lambda$ such that $G = H_{i_0\lambda_0}$. For each $i \in I$, let $r_i$ be the idempotent element of $H_{i\lambda_0}$, and for each $\lambda \in \Lambda$, let $\ell_\lambda$ be the idempotent element of $H_{i_0\lambda}$ (see Figure 2.3). We then have $\ell_{\lambda_0} = r_{i_0} = e$, since $e$ is the idempotent element of $G = H_{i_0\lambda_0}$.

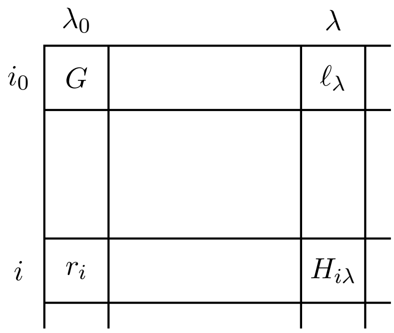



FIGURE 2.3.

Define the $\Lambda \times I$ matrix $P$ by setting $p_{\lambda i} := \ell_\lambda r_i$ for each $(\lambda, i) \in \Lambda \times I$. By Theorem 2.26(ii), the entries of $P$ lie in $G$. Moreover, for all $i \in I$ we have $p_{\lambda_0 i} = \ell_{\lambda_0} r_i = er_i = e$ by Proposition 2.19, and similarly, $p_{\lambda i_0} = e$ for all $\lambda \in \Lambda$. That is, the entries in the $\lambda_0$-row and $i_0$-column of $P$ all equal $e$.

Define the map $\varphi \colon I \times \Lambda \times G \to S$ by

$$\varphi(i, \lambda, a) := r_i a \ell_\lambda$$

for all $(i, \lambda, a) \in I \times \Lambda \times G$. We will show that $\varphi$ is an isomorphism $\mathcal{M}[\mathbf{G}, P] \to \mathbf{S}$.

Let $(i, \lambda) \in I \times \Lambda$. By Green's Lemma 2.17 and the fact that $e\ell_\lambda = \ell_\lambda$, right translation by $\ell_\lambda$ maps $H_{i_0\lambda_0} = G$ bijectively onto $H_{i_0\lambda}$, and since $r_i e = r_i$, left translation by $r_i$ maps $H_{i_0\lambda}$ bijectively onto $H_{i\lambda}$. Thus, $\varphi$ restricts to a bijection from $\{i\} \times \{\lambda\} \times G$ onto $H_{i\lambda}$.

Since this is true for all $(i,\lambda) \in I \times \Lambda$, it follows that $\varphi$ is a bijection. To see that $\varphi$ is a homomorphism, let $(i,\lambda,a),(j,\mu,b) \in I \times \Lambda \times G$. Then

$$\varphi((i,\lambda,a)(j,\mu,b)) = \varphi(i,\mu,ap_{\lambda j}b) = r_i a p_{\lambda j} b \ell_\mu = r_i a \ell_\lambda r_j b \ell_\mu = \varphi(i,\lambda,a)\varphi(j,\mu,b),$$

as required. □

To close this section, we will discuss a corollary of Theorem 2.28. First, let us consider the following result, which tells us when idempotents in a completely zero-simple semigroup are related by $\mathcal{L}$ or $\mathcal{R}$. We essentially already know the content, but this particular way of expressing it will be important.

**Proposition 2.29.** *Let* $\mathbf{S}$ *be a completely zero-simple semigroup, and let* $e, f \in S\backslash\{0\}$ *be idempotent. Then*

(i) *$e$ is a left zero for $f$ and and only if $e \mathcal{L} f$;*
(ii) *$e$ is a right zero for $f$ and and only if $e \mathcal{R} f$;*
(iii) *$e$ is a zero for $f$ if and only if $e = f$.*

Proof. To prove (i), assume that $e$ is a left zero for $f$; i.e., $ef = e$. Then $e \in Sf\backslash\{0\}$, so by Lemma 2.24 we have $e \mathcal{L} f$. Conversely, if $e \mathcal{L} f$, then $f$ is a right identity for $e$ by Lemma 2.19, so $e$ is a left zero for $f$. This proves (i), and (ii) follows by duality.

Now (i) and (ii) imply that $e$ is a zero for $f$ if and only if $e \mathcal{H} f$. By Green's Theorem 2.21, an $\mathcal{H}$-class contains at most one idempotent, so $e$ is a zero for $f$ if and only if $e = f$. □

From Proposition 2.29(i), we see that if $A$ is a set of idempotents of a completely zero-simple semigroup $\mathbf{S}$ such that $A$ is contained in a single $\mathcal{L}$-class of $\mathbf{S}$, then every element of $A$ is a left zero for every other element of $A$. In particular, $A$ forms a subsemigroup, which evidently satisfies $xy \approx x$.

In general, a semigroup satisfying $xy \approx x$ is called a **left-zero semigroup**. Dually, a **right-zero semigroup** is defined to be a semigroup satisfying $xy \approx y$. We encountered such semigroups in Example 2.5, but we can now see how these semigroups arise in much more general settings: the idempotents within an $\mathcal{L}$-class of a completely zero-simple semigroup (in fact, any semigroup) form a left-zero semigroup, and dually.

It should be no surprise, then, that we can see the shadows of these semigroups in the structure of Rees matrix semigroups. Consider the multiplication in a completely simple semigroup $\mathcal{M}[\mathbf{G}, P]$, where $\mathbf{G}$ is a group and $P$ is a $\Lambda \times I$ matrix. The multiplication is defined as

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

Observe that the first coordinate of the product depends only on the first coordinates of the two factors, and similarly for the second coordinate, but that the third coordinate may depend on multiple coordinates of the two factors. This should hopefully remind the reader of the semidirect product in group theory. In fact, we can make this connection stronger. Let us view $I$ as a left-zero semigroup, so that $ij = i$ for all $i, j \in I$. Similarly, view $\Lambda$ as a right-zero semigroup. Then the multiplication of $\mathcal{M}[\mathbf{G}, P]$ can be written as

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

Thus, the multiplication behaves as in the direct product $I \times \Lambda \times G$ in the first two coordinates, with a 'twist' in the third coordinate. It should now be apparent that if every entry of $P$ is the identity of $\mathbf{G}$, then $\mathcal{M}[\mathbf{G}, P]$ is precisely the direct product $I \times \Lambda \times G$.

When does this happen? Examining the proof of Theorem 2.28, we see that $P$ can be chosen so that every entry is a product of two idempotent elements of $\mathbf{S}$. Therefore, if the set of idempotent elements of $\mathbf{S}$ is closed under multiplication, then the entries of $P$ will be idempotent, and hence equal to the identity of $\mathbf{G}$. In such an event, the multiplication of $\mathcal{M}[\mathbf{G}, P]$ reduces to the direct product operation on $I \times \Lambda \times G$.

The leads us to define orthodox semigroups. A semigroup is **orthodox** if its set of idempotents is closed under multiplication. (Note that this definition is non-standard; orthodox semigroups are usually assumed to be regular, but we will not follow this convention.) From the above considerations, we have the following decomposition theorem.

**Theorem 2.30.** *Let* $\mathbf{S}$ *be a completely simple semigroup. Then* $\mathbf{S}$ *is orthodox if and only if* $\mathbf{S}$ *is isomorphic to the direct product of a left-zero semigroup, a right-zero semigroup, and a group.*

Proof. If $\mathbf{S}$ is orthodox, then our discussion above shows that $\mathbf{S}$ has the required decomposition. For the converse, note that left-zero semigroups, right-zero semigroups, and groups are always orthodox, and a direct product of orthodox semigroups is orthodox. □

Now is a good time to reflect. In Section 2.1, we saw how a semigroup can be broken up into $\mathcal{J}$-classes, and how these $\mathcal{J}$-classes can be converted into the so-called principal factors, which are either simple, zero-simple, or null. We have studied the special classes of completely simple and completely zero-simple semigroups, and developed some very powerful structure theorems. Although the 'completely' modifier is quite restrictive, it comes for free in the finite case; our next goal is to show that we can encompass more than the finite case. After this, we will reap the rewards of our general study when we investigate some well-known and well-behaved classes of semigroups.

## 2.4. Characterisations of complete zero-simplicity

Under what conditions is a (zero-)simple semigroup completely (zero-)simple? Based on the last two sections, it should be no surprise that idempotents will play a role in answering this question.

Let us take another look at Proposition 2.29(iii); in particular, the notion of one idempotent being a zero for another idempotent. This in fact gives us an order relation. To be precise, let $\mathbf{S}$ be a semigroup, and let $E$ be the set of all idempotents of $\mathbf{S}$. Given $e, f \in E$, we put $e \leqslant f$ if $ef = fe = e$ (that is, if $e$ is a zero for $f$, or equivalently, if $f$ is an identity for $e$). This is easily seen to be an order relation on $E$. The relation $\leqslant$ so defined does not seem to have a standard name, but we will refer to it as the **idempotent order**.

Framed in terms of the idempotent order, Proposition 2.29(iii) tells us that the non-zero idempotents of a completely zero-simple semigroup $\mathbf{S}$ form an antichain; that is, if $e$ and $f$ are non-zero idempotents of $\mathbf{S}$ with $e \leqslant f$, then $e = f$. It is precisely this lack of order that characterises completely zero-simple semigroups. Certainly, if a zero-simple semigroup $\mathbf{S}$

has non-zero idempotents $e, f \in S$ with $e < f$, then **S** cannot be completely zero-simple. The essence of this section is that the converse holds, assuming that **S** has at least one non-zero idempotent.

In a zero-simple semigroup, a non-trivial order relation among the non-zero idempotents will turn out to imply the existence of an infinite descending chain $e_0 > e_1 > e_2 > \cdots$ in the set of non-zero idempotents. Furthermore, this infinite chain of idempotents will live inside a subsemigroup isomorphic to what is known as the **bicyclic semigroup** (or **bicyclic monoid**). The bicyclic semigroup is most cleanly defined by means of the monoid presentation $\langle a, b \mid ab = 1\rangle$; so it is the 'most general' monoid that can be generated by two elements $a$ and $b$ such that $ab$ is the identity element.

Of course, if **M** is a monoid generated by some elements $a$ and $b$, then every element of $M$ is expressible as a string of '$a$'s and '$b$'s. Subject to the equality $ab = 1$, any occurrence of $ab$ can be deleted, so every element can be represented in the form $b^m a^n$, where $m, n \in \mathbb{N}_0$. (Here, of course, we interpret $b^0$ and $a^0$ as being equal to 1.) The next result shows, in a slightly more general setting, that each element can be *uniquely* expressed in the form $b^m a^n$, provided only that $ba \neq 1$.

**Lemma 2.31.** *Let* **S** *be a semigroup, and suppose there are elements $a, b, e \in S$ with the following properties:*

(i) *$e$ is idempotent;*
(ii) *$e$ is an identity element for $a$ and $b$; that is, $ea = ae = a$ and $eb = be = b$;*
(iii) *$ab = e$;*
(iv) *$ba \neq e$.*

*Then the mapping $(m, n) \mapsto b^m a^n$ of $\mathbb{N}_0 \times \mathbb{N}_0$ into $S$ is one-to-one (where $a^0 = b^0 = e$).*

Proof. First, we show that $b^h a^k = e$ implies $h = k = 0$, for all $h, k \geqslant 0$. Assume that $b^h a^k = e$ for some $h, k \geqslant 0$. If we had $k > 0$, then properties (ii) and (iii) would give

$$ba = eba = b^h a^k ba = b^h a^{k-1} ea = b^h a^k = e,$$

a contradiction to (iv). By symmetry, $h > 0$ is also impossible.

Now, suppose that $a^h = b^k$ for some $h, k \geqslant 0$. Multiplying on the left by $a^k$ then gives $a^{h+k} = e$, so by the previous paragraph we we must have $h = k = 0$.

Finally, let $m, n, i, j \geqslant 0$ with $b^m a^n = b^i a^j$. We may assume without loss of generality that $m \leqslant i$, and multiplying on the left by $a^m$ then gives $a^n = b^{i-m} a^j$. If $n \leqslant j$, then multiplying on the right by $b^n$ gives $e = b^{i-m} a^{j-n}$, and if $n \geqslant j$, multiplying $a^n = b^{i-m} a^j$ on the right by $b^j$ gives $a^{n-j} = b^{i-m}$. By the previous two paragraphs, both cases imply that $(m, n) = (i, j)$. □

It would be good to be sure, then, that there exists a semigroup with elements satisfying the conditions of Lemma 2.31. As an intuition, one may think of differentiation and anti-differentiation. This intuition is distilled into Figure 2.4, showing that such a configuration exists in the semigroup of maps $\mathbb{N}_0 \to \mathbb{N}_0$.

We now see that the bicyclic semigroup is uniquely defined up to isomorphism as any semigroup generated by $\{a, b, e\}$ satisfying (i)–(iv) of Lemma 2.31 (in fact, this can serve

$$\begin{tikzcd}[row sep=small]
\alpha & 0 \arrow[loop left] & 1 \arrow[l] & 2 \arrow[l] & 3 \arrow[l] & 4 \arrow[l] & \cdots \arrow[l] \\
\beta & 0 \arrow[r] & 1 \arrow[r] & 2 \arrow[r] & 3 \arrow[r] & 4 \arrow[r] & \cdots
\end{tikzcd}$$

FIGURE 2.4. Maps $\alpha$ and $\beta$ on $\mathbb{N}_0$ such that $\alpha \circ \beta = \mathrm{id}$ and $\beta \circ \alpha \neq \mathrm{id}$.

as a definition that does not require the use of presentations). Capturing this idea, the following result tells us when the bicyclic semigroup embeds into a given semigroup.

**Lemma 2.32.** *Let* $\mathbf{S}$ *be a semigroup. Then the bicyclic semigroup embeds into* $\mathbf{S}$ *if and only if there are elements* $a, b, e \in S$ *satisfying conditions* (i)–(iv) *of Lemma 2.31.*

We close this section with the next theorem, which shows how the idempotent order and the bicyclic semigroup can be used to characterise completely zero-simple semigroups.

**Theorem 2.33.** *Let* $\mathbf{S}$ *be a zero-simple semigroup with at least one non-zero idempotent. The following are equivalent:*

(i) $\mathbf{S}$ *is completely zero-simple;*
(ii) *the non-zero idempotents of* $\mathbf{S}$ *form an antichain;*
(iii) $\mathbf{S}$ *has a minimal non-zero idempotent;*
(iv) *the bicyclic semigroup does not embed into* $\mathbf{S}$.

PROOF. (i) $\Rightarrow$ (ii): This follows from Proposition 2.29(iii) and the definition of the idempotent order.

(iii) $\Rightarrow$ (i): Let $e$ be a minimal non-zero idempotent of $\mathbf{S}$; by this, we mean an idempotent that is minimal in the set of non-zero idempotents. We will show that $eS$ is a zero-minimal right ideal. By Lemma 2.22, it suffices to show that $eS = R_e \cup \{0\}$. Since $e$ is idempotent, we have $e \in eS$, so $R_e \subseteq eS \backslash \{0\}$. We must show that $eS \backslash \{0\} \subseteq R_e$.

Let $a \in eS \backslash \{0\}$. To show that $a \in R_e$, it suffices to show that $e \in aS$. Note that $ea = a$ from the choice of $a$; we will use this fact throughout. By Lemma 2.9, there are $x, y \in S$ with $xay = e$. Now, define $f := ayexe$. Then $f$ is idempotent:

$$ff = ayex\underline{ea}yexe = aye\underline{xay}exe = ayeeexe = ayexe = f.$$

Also, $f$ is non-zero because

$$xfay = xayex\underline{ea}y = \underline{xay}e\underline{xay} = eee = e \neq 0.$$

Thus, $f$ is a non-zero idempotent. Using the fact that $ea = a$, it is easily seen from the definition of $f$ that $ef = fe = f$, so $f \leqslant e$. By the choice of $e$, we have $e = f = ayexe \in aS$. This shows that $a \in R_e$, as required.

We have shown that $eS \backslash \{0\} = R_e$, so $eS$ is a zero-minimal right ideal by the dual of Lemma 2.22. By symmetry, $Se$ is a zero-minimal left ideal. This shows that $\mathbf{S}$ is completely zero-simple.

Now (ii) $\Rightarrow$ (iii) is trivial since $\mathbf{S}$ has a non-zero idempotent, so conditions (i)–(iii) are equivalent.

(iv) $\Rightarrow$ (ii): Let $e, f \in S\backslash\{0\}$ be idempotents with with $e \geqslant f$; we must show that $e = f$. By Lemma 2.9, there exist $x, y \in S$ with $e = xfy$. Define $a := exf$ and $b := fye$. Then $e$ is an identity for $a$ and $b$ since $e$ is an identity for $f$. We also have

$$ab = exffye = exfye = eee = e.$$

By (iv) and Lemma 2.32, we must have $ba = e$. But $f$ is an identity for $ba$ from the definition of $a$ and $b$, so $f$ is an identity for $e$; i.e., $f \geqslant e$. Hence, $f = e$, which proves (ii).

(ii) $\Rightarrow$ (iv): We will prove the contrapositive. Suppose that (iv) fails, so the bicyclic semigroup $\langle a, b \mid ab = 1\rangle$ embeds into $\mathbf{S}$. Define $e_n := b^n a^n$ for each $n \in \mathbb{N}_0$. Then $e_n$ is idempotent for every $n \in \mathbb{N}_0$, and furthermore, we have $e_n \geqslant e_m$ whenever $n \leqslant m$. Using the uniqueness of representation given by Lemma 2.31, we also have $e_n \neq e_m$ whenever $n \neq m$. It follows that $e_0 > e_1 > e_2 > \cdots$ is an infinite descending chain of non-zero idempotents in $\mathbf{S}$. Thus, (ii) fails. $\square$

## 2.5. Exponents and periodicity

If a finite semigroup is (zero-)simple, then it is automatically completely (zero-)simple. Using the results of Section 2.4, we will show that this implication extends to periodic semigroups (to be defined shortly). This class is sufficiently broad to encompass all of the semigroups we will encounter in later chapters.

We will begin with a brief discussion of **monogenic** (1-generated) semigroups. Naturally, such basic objects are ubiquitous in semigroup theory, so we will meet them very often, if only in the background.

Let $\mathbf{S}$ be a semigroup, and let $a \in S$. Then $\{a^k \mid k \in \mathbb{N}\}$ is the subsemigroup generated by $a$. Some authors denote this subsemigroup by $\langle a\rangle$, but we will reserve this notation for the ideal generated by $a$.

The **order** of $a$ is the cardinality of $\{a^k \mid k \in \mathbb{N}\}$. For us, the precise order of an element will not be particularly important; we will usually be interested only in whether the order is finite or infinite. It is easy to see that $a$ has finite order in $\mathbf{S}$ if and only if there is a repetition in the list $a^1, a^2, a^3, \ldots$ (i.e., $a^m = a^n$ for some $m, n \in \mathbb{N}$ with $m \neq n$).

An element of infinite order generates a subsemigroup isomorphic to the semigroup $\langle \mathbb{N}; +\rangle$ of positive integers. To describe finite monogenic subsemigroups, we associate two important numbers to an element $a$ of finite order, called the index and the period of $a$. These numbers determine the structure of the subsemigroup generated by $a$, just as the structure of a finite cyclic group is determined by the order of its generator.

If $a$ is a semigroup element of finite order, the **index** of $a$ is defined as the smallest $i \in \mathbb{N}$ such that $a^i = a^k$ for some $k > i$. If $a$ is of finite order with index $i$, we define the **period** of $a$ to be the smallest $p \in \mathbb{N}$ such that $a^i = a^{i+p}$.

Figure 2.5 depicts the subsemigroup generated by an element of index 3 and period 6. Displayed is the obvious fact that the powers of an element of finite order eventually loop. The index of the element $a$ is the smallest power of $a$ that lies in the loop, while the period of $a$ is the number of distinct elements in the loop.

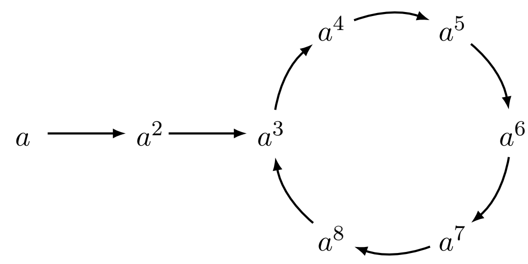



FIGURE 2.5. A subsemigroup generated by an element $a$ of index 3 and period 6. The arrows indicate the action of translation by $a$.

As with cyclic groups, we can compute products in $\{a^k \mid k \in \mathbb{N}\}$ using modular arithmetic. In the situation of Figure 2.5, we have $a^9 = a^3$, allowing us to reduce powers modulo 6. For example, we have $a^5 \cdot a^8 = a^{13} = a^7$, since $13 \equiv 7 \pmod 6$. But in contrast to cyclic groups, we cannot reduce 7 further to 1 since $a^1 \neq a^7$. Formally, if $a$ has index $i$ and period $p$, we have $a^k = a^\ell$ provided that $k, \ell \geqslant i$ and $k \equiv \ell \pmod p$.

It should be no surprise that modular arithmetic occurs in the study of monogenic semigroups, as it does with cyclic groups. Of course, a monogenic semigroup need not be a group, but it turns out that the 'loop' part of a finite monogenic semigroup is always a cyclic group.

**Theorem 2.34.** *Let* $\mathbf{S}$ *be a semigroup, let* $a \in S$ *have finite order, and let* $i$ *and* $p$ *be the index and period of* $a$*, respectively. Then the set* $\{a^\ell \mid \ell \geqslant i\}$ *forms a* $p$*-element cyclic subgroup of* $\mathbf{S}$*.*

PROOF. Clearly the subsemigroup $G := \{a^\ell \mid \ell \geqslant i\}$ has cardinality $p$, by the definition of index and period. Let $\oplus_p$ denote addition modulo $p$ on $\mathbb{Z}_p := \{0, \dots, p-1\}$. We will show that $\langle G; \cdot \rangle$ is isomorphic to $\langle \mathbb{Z}_p; \oplus_p \rangle$.

Define $\varphi\colon \mathbb{Z}_p \to \{a^\ell \mid \ell \geqslant i\}$ by $\varphi(k) := a^{ip+k}$ for each $k \in \mathbb{Z}_p$. Then, given $k, \ell \in \mathbb{Z}_p$, we have $\varphi(k)\varphi(\ell) = a^{2ip+k+\ell}$, and reducing modulo $p$ gives $a^{2ip+k+\ell} = a^{ip+k\oplus_p \ell}$, so we have $\varphi(k)\varphi(\ell) = \varphi(k \oplus_p \ell)$. This shows that $\varphi$ is a homomorphism. Since $\{ip + k \mid k \in \mathbb{Z}_p\}$ intersects all congruence classes modulo $p$, it follows that $\varphi$ is surjective, and is therefore an isomorphism, since $|G| = |\mathbb{Z}_p| = p$. $\square$

Theorem 2.34 will prove to be quite useful, but for now, the most important point is the following obvious consequence: if $a$ is a semigroup element of finite order, then there is some $n \in \mathbb{N}$ such that $a^n$ is idempotent. More generally, we have the following.

**Theorem 2.35.** *Let* $\mathbf{S}$ *be a semigroup, and let* $A$ *be a finite subset of* $S$ *such that every element of* $A$ *has finite order. Then there is some* $n \in \mathbb{N}$ *such that* $a^n$ *is idempotent for all* $a \in A$*.*

PROOF. By Theorem 2.34, we may choose, for each $a \in A$, some $n_a \in \mathbb{N}$ such that $a^{n_a}$ is idempotent. Let $n \in \mathbb{N}$ be any number divisible by $n_a$ for every $a \in A$. Then, for all $a \in A$, we have that $a^n$ is a power of $a^{n_a}$, and is therefore equal to $a^{n_a}$, so $a^n$ is idempotent. $\square$

In group theory, we say that $n$ is an exponent of a group $\mathbf{G}$ if $x^n$ is the identity element of $\mathbf{G}$ for all $x \in G$. Analogously, we say that $n$ is an **exponent** of a semigroup $\mathbf{S}$ if $x^n$

is idempotent for all $x \in S$; i.e., if $\mathbf{S} \models x^{2n} \approx x^n$. This agrees with the group-theoretic definition when $\mathbf{S}$ happens to be a group.

We say that **S has finite exponent** if there is some $n \in \mathbb{N}$ such that $n$ is an exponent of $\mathbf{S}$. Importantly, by Theorem 2.35, every finite semigroup has finite exponent. However, as with groups, having finite exponent is not guaranteed by all elements having finite order.

Extending the definition from group theory, we say that a semigroup is **periodic** if all of its elements have finite order; i.e., if every element generates a finite subsemigroup. Of course, every finite semigroup is periodic, as is every semigroup with finite exponent. The following is an obvious consequence of Theorem 2.35.

**Theorem 2.36.** *Let* $\mathbf{S}$ *be a periodic semigroup. Then* $\mathbf{S}$ *has at least one idempotent element.*

We will now prove two powerful theorems regarding periodic semigroups. The proofs will use the following trivial lemma. (We chose the name as a homage to the pumping lemmas in formal language theory.)

**Lemma 2.37** (Pumping Lemma)**.** *Let* $\mathbf{S}$ *be a semigroup, let* $a \in S$*, and assume that there are* $u, v \in S^1$ *with* $a = uav$*. Then* $a = u^k a v^k$ *for all* $k \in \mathbb{N}$*.*

PROOF. By induction. □

**Theorem 2.38.** *Let* $\mathbf{S}$ *be a periodic semigroup. Then* $\mathcal{D} = \mathcal{J}$*.*

PROOF. We always have $\mathcal{D} \subseteq \mathcal{J}$. Let $a, b \in S$ with $a \mathrel{\mathcal{J}} b$, so there exist $u, v, x, y \in S^1$ with $uav = b$ and $xby = a$. Then $a = xby = x(uav)y = (xu)a(vy)$. By Theorem 2.35, we may choose $n \in \mathbb{N}$ such that $(xu)^n$ is idempotent. By the Pumping Lemma 2.37, we have

$$a = (xu)^n a(vy)^n = (xu)^n (xu)^n a(vy)^n = (xu)^n a \in S^1 ua,$$

and clearly $ua \in S^1 a$, so $a \mathrel{\mathcal{L}} ua$. By symmetry, we have $a \mathrel{\mathcal{R}} av$, and so because $\mathcal{R}$ is left-compatible by Proposition 2.15, we have $ua \mathrel{\mathcal{R}} uav = b$. Thus, $a \mathrel{\mathcal{L}} ua \mathrel{\mathcal{R}} b$, showing that $a \mathrel{\mathcal{D}} b$. □

**Theorem 2.39.** *Let* $\mathbf{S}$ *be a periodic zero-simple semigroup. Then* $\mathbf{S}$ *is completely zero-simple.*

PROOF. We will use Theorem 2.33. Note that the generators of the bicyclic semigroup have infinite order, so the bicyclic group does not embed into $\mathbf{S}$. Thus, we need only show that $\mathbf{S}$ has a non-zero idempotent.

Choose any $a \in S \backslash \{0\}$. By Lemma 2.9, there exist $x, y \in S$ such that $a = xay$. By Theorem 2.35, there is some $n \geqslant 1$ such that $x^n$ is idempotent, and by the Pumping Lemma 2.37 we have $x^n a y^n = a \neq 0$; thus, $x^n$ is a non-zero idempotent of $\mathbf{S}$. □

As an immediate consequence of Theorem 2.39, we have the following.

**Theorem 2.40.** *Let* $\mathbf{S}$ *be a periodic semigroup. Then the minimum ideal of* $\mathbf{S}$ *is completely simple* (*if it exists*)*, and every other principal factor of* $\mathbf{S}$ *is either completely zero-simple or null.*

Proof. Noting that the class of periodic semigroups is closed under forming subsemigroups and quotients, the result follows from Theorem 2.11 and Theorem 2.39. □

Finally, to close this section, we will give a strengthened version of Theorem 2.39 for the finite exponent case. The following result shows in particular that the completely simple semigroups of exponent $n$ form a variety—not an altogether obvious fact.

**Theorem 2.41.** *Let* $\mathbf{S}$ *be semigroup of exponent* $n$ *for some* $n \in \mathbb{N}$. *Then the following are equivalent:*

(i) $\mathbf{S}$ *is simple;*
(ii) $\mathbf{S}$ *is completely simple;*
(iii) $\mathbf{S} \models (xy)^n x \approx x$.

Proof. Theorem 2.39 (applied to $\mathbf{S}^0$) gives (i) ⇒ (ii).

(ii) ⇒ (iii): Let $x, y \in S$. Then the idempotent $(xy)^n \in xS$ is in the $\mathcal{R}$-class of $x$ by Lemma 2.24. By Proposition 2.19, we have $(xy)^n x = x$.

(iii) ⇒ (i): Let $I$ be an ideal of $\mathbf{S}$, and fix some $y \in I$. If $x \in S$, then $x = (xy)^n x \in I$, so $S = I$. Thus, $\mathbf{S}$ is simple. □

## 2.6. Completely regular semigroups and bands

In this final section, we will study some important special classes of semigroups using the structure theory we have developed so far. All the hard work is done, so we will be able to deduce some very strong structural results with relative ease.

Let $\mathbf{S}$ be a semigroup, and let $a \in S$. We say that $a$ is a **group element** of $\mathbf{S}$ if $a$ lies in some subgroup of $\mathbf{S}$. Using Green's Theorem 2.21, one easily sees that $a$ is a group element of $\mathbf{S}$ if and only if $H_a$ is a subgroup of $\mathbf{S}$. The term 'group element' is not usually used in other texts, but it will be used uncountably many times in this thesis.

We open this section with a result characterising group elements within periodic semigroups. This result is quite simple to state, but it uses much of the machinery developed in this chapter.

**Theorem 2.42.** *Let* $\mathbf{S}$ *be a periodic semigroup, and let* $a \in S$. *Then* $a$ *is a group element of* $\mathbf{S}$ *if and only if* $a \mathrel{\mathcal{J}} a^2$.

Proof. If $a$ is a group element, then $a \mathrel{\mathcal{H}} a^2$, so certainly $a \mathrel{\mathcal{J}} a^2$. For the converse, assume that $a \mathrel{\mathcal{J}} a^2$.

First, if $\langle a \rangle$ happens to be the minimum ideal of $\mathbf{S}$, then $\langle a \rangle$ is completely simple by Theorem 2.40, and so $a$ is a group element by Theorem 2.26(iii). Thus, we may assume that $I := \{x \in S \mid J_x < J_a\}$ is non-empty. Consider now the principal factor $\langle a \rangle / I$. We may represent this semigroup on the set $J_a \cup \{0\}$, where $0 \notin J_a$ represents the ideal $I$. The product of $x, y \in J_a \cup \{0\}$ is given by the usual product $xy$ in $\mathbf{S}$ if $x, y, xy \in J_a$, and otherwise the product is 0. Since $a^2$ is a non-zero element of the principal factor by assumption, $\langle a \rangle / I$ is not null, so it is completely zero-simple by Theorem 2.40. The non-zero elements form a $\mathcal{J}$-class of the principal factor, so the relation $a \mathrel{\mathcal{J}} a^2$ holds in the principal factor.

Now, if the $\mathcal{H}$-class of $a$ in $\langle a\rangle/I$ is a subgroup, this will then be a subgroup of $\mathbf{S}$. So, without loss of generality, we can assume that $\mathbf{S}$ is completely zero-simple and that $a$ and $a^2$ lie in the non-zero $\mathcal{J}$-class of $\mathbf{S}$. In particular, $a^2 \neq 0$, so by Theorem 2.25(ii), we have $a^2 \in R_a \cap L_a = H_a$. By Green's Theorem 2.21, $H_a$ is a subgroup, and so $a$ is a group element of $\mathbf{S}$, as required. □

For the remainder of this section, we will be concerned with the situation where all elements in a given semigroup happen to be group elements. We say that a semigroup $\mathbf{S}$ is **completely regular** if every element of $\mathbf{S}$ is a group element, or equivalently, if every $\mathcal{H}$-class of $\mathbf{S}$ is a group.

The property of being completely regular is of course quite a strong one, though we have seen an important class of such semigroups already. By Theorem 2.26(iii), all completely simple semigroups are completely regular. To contrast, we mentioned in Section 2.1 that every semigroup can be embedded into a simple semigroup. These two facts demonstrate the vast difference between completely simple semigroups and the wilder class of simple semigroups.

Although complete regularity is a such a strong property, these semigroups arise naturally within more general semigroups. In the next chapter, we will encounter some broad classes of semigroups with the property that their group elements form a subsemigroup, which is obviously completely regular. Much of the original semigroup can be understood in terms of this subsemigroup of group elements, which leads one to study the internal structure of completely regular semigroups.

In a completely regular semigroup $\mathbf{S}$, we are afforded a notion of inverse, inherited from the $\mathcal{H}$-classes. For each $a \in S$, we define the **inverse**, $a^{-1}$, of $a$ to be the group-theoretic inverse of $a$ in its $\mathcal{H}$-class $H_a$. It is clear that

$$(a^{-1})^{-1} = a, \qquad a^{-1}a = aa^{-1}, \qquad aa^{-1}a = a.$$

In fact, if one treats the inverse operation on $\mathbf{S}$ as a fundamental operation, then completely regular semigroups are defined by these three equations (and associativity): if the above equations hold, then one easily sees that $a^{-1}a$ is an idempotent element in the $\mathcal{H}$-class of $a$, so $H_a$ is a subgroup by Green's Theorem 2.21.

We have seen that, in the class of periodic semigroups, all simple semigroups are completely simple. Using the notion of inverse elements, we can easily show that this implication also holds for completely regular semigroups (which need not be periodic).

It is an elementary fact that if $a$ and $b$ are elements of some group, then $ab$ is the identity if and only if $ba$ is the identity. In relation to Lemma 2.31, this tells us that the bicyclic semigroup cannot embed into a group (of course). Generalising this to completely regular semigroups gives us the following result.

**Theorem 2.43.** *Let* $\mathbf{S}$ *be a simple semigroup. Then* $\mathbf{S}$ *is completely simple if and only if* $\mathbf{S}$ *is completely regular.*

PROOF. By Theorem 2.26(iii), a completely simple semigroup is completely regular. Conversely, assume that $\mathbf{S}$ is completely regular, and let $a, b, e \in S$ satisfy (i)–(iii) of Lemma 2.31. We will show that (iv) fails, by showing that $a \mathrel{\mathcal{H}} b \mathrel{\mathcal{H}} e$.

We have $e = ab = a^{-1}aab = a^{-1}ae = a^{-1}a$, so $e$ is the identity element of $H_a$, and hence $a \in H_e$. By symmetry, we have $b \in H_e$. Since $H_e \ni a, b$ is a group and $ab = e$, we have $ba = e$ also. By Lemma 2.32 and Theorem 2.33, $\mathbf{S}$ is completely simple. □

In Section 1.5, we introduced semilattices, and explained how they can be regarded as commutative semigroups in which all elements are idempotent. We will now see that they arise naturally when studying more general classes.

We have seen how a general semigroup can be broken down into principal factors, which are either simple, zero-simple, or null (Theorem 2.11). In the class of completely regular semigroups, we can say much more: every completely regular semigroup is a "semilattice of completely simple semigroups", in the following sense.

**Theorem 2.44.** *Let* $\mathbf{S}$ *be a completely regular semigroup. Then* $\preccurlyeq_J$ *is compatible with* $\mathbf{S}$*, and so* $\mathcal{J}$ *is a congruence on* $\mathbf{S}$*. Moreover,* $\mathbf{S}/\mathcal{J}$ *is a semilattice, and each* $\mathcal{J}$*-class of* $\mathbf{S}$ *is a completely simple subsemigroup of* $\mathbf{S}$*.*

Proof. For all $a \in S$, we have $a \mathrel{\mathcal{H}} a^2$, so $a \mathrel{\mathcal{J}} a^2$. Hence, for all $a, b \in S$, we have

$$ab \mathrel{\mathcal{J}} (ab)^2 = a(ba)b \preccurlyeq_J ba,$$

and by symmetry $ba \preccurlyeq_J ab$, so $ab \mathrel{\mathcal{J}} ba$.

To show that $\preccurlyeq_J$ is compatible, it is enough by duality and transitivity to show that it is right-compatible. Let $a, b, c \in S$ with $a \preccurlyeq_J b$. Then there are $x, y \in S^1$ with $a = xby$, so, using the commutativity property just established, we have

$$ac = xbyc \preccurlyeq_J byc \mathrel{\mathcal{J}} cby \preccurlyeq_J cb \mathrel{\mathcal{J}} bc,$$

so $ac \preccurlyeq_J bc$, as required. Hence, $\mathcal{J}$ is a congruence. Since $a \mathrel{\mathcal{J}} a^2$ and $ab \mathrel{\mathcal{J}} ba$ for all $a, b \in S$, the factor semigroup $\mathbf{S}/\mathcal{J}$ is a semilattice. Moreover, if $a \mathrel{\mathcal{J}} b$ in $\mathbf{S}$, then $ab \mathrel{\mathcal{J}} bb \mathrel{\mathcal{J}} b$, so $ab$ is in the $\mathcal{J}$-class of $a$ and $b$, showing that each $\mathcal{J}$-class of $\mathbf{S}$ is a subsemigroup of $\mathbf{S}$

Now, by Corollary 2.10, each $\mathcal{J}$-class is simple, and since each $\mathcal{J}$-class is a union of $\mathcal{H}$-classes of $\mathbf{S}$, the $\mathcal{J}$-classes are also completely regular subsemigroups. By Theorem 2.43, each $\mathcal{J}$-class is completely simple. □

**Remark.** If $\mathbf{S}$ is completely regular, we have two natural orders on $S/\mathcal{J}$. One is the $\mathcal{J}$-order defined in Section 2.1, and the other is the semilattice order (Section 1.5). Fortunately, these two order relations coincide. This amounts to saying that $ab \mathrel{\mathcal{J}} a \Leftrightarrow a \preccurlyeq_J b$, which follows easily from the compatibility of $\preccurlyeq_J$.

In a completely regular semigroup $\mathbf{S}$, every $\mathcal{H}$-class contains an idempotent, so it is inevitable that the idempotents play an important role in their study. Of particular importants is the case where $\mathbf{S}$ is orthodox. In such an event, a completely regular semigroup can be studied by means of the idempotent subsemigroup. The next result gives a sufficient condition for a completely regular semigroup to be orthodox, which we will encounter in the next chapter. The exceedingly neat proof is due to Clifford [**8**, Proposition 1].

**Lemma 2.45.** *Let* $\mathbf{S}$ *be a completely regular semigroup such that each* $\mathcal{J}$*-class of* $\mathbf{S}$ *is an orthodox subsemigroup. Then* $\mathbf{S}$ *is orthodox.*

Proof. Let $e, f \in S$ be idempotent, and let $a := ef$. We must show that $a$ is idempotent.

Defining $b := fe$, we have $a \mathrel{\mathcal{J}} b$, since $\mathbf{S}/\mathcal{J}$ is commutative by Theorem 2.44. Let $\mathbf{J}$ be the subsemigroup on the $\mathcal{J}$-class $J_a = J_b$; so $\mathbf{J}$ is completely simple. Being orthodox, $\mathbf{J}$ is isomorphic to the direct product of a left-zero semigroup, a right-zero semigroup, and a group, by Theorem 2.30. These three direct factors satisfy $yxy \approx y^2 \to x^2 \approx x$, so we find that $\mathbf{J}$ also satisfies this quasiequation. Now, we have $bab = feeffe = fefe = b^2$, so $a^2 = a$ by the quasiequation, showing that $a = ef$ is idempotent, as required. □

We are now brought to the final class of semigroups that we will study in this chapter. We say that a semigroup $\mathbf{S}$ is a **band** if $\mathbf{S} \models x^2 \approx x$; i.e., if every element of $\mathbf{S}$ is idempotent. We have already encountered some important classes of bands: left- and right-zero semigroups are bands, as are semilattices. Bands are certainly worthy of study in their own right, but they are also important for studying semigroups whose idempotents form a subsemigroup.

Of course, bands are completely regular: every element lies in a trivial subgroup. Thus, Theorem 2.44 tells us that a band is a semilattice of completely simple bands. With the results we have presented in this chapter, these latter objects are easily described.

**Theorem 2.46.** *Let* $\mathbf{S}$ *be a semigroup. The following are equivalent:*

(i) $\mathbf{S}$ *is a simple band;*
(ii) $\mathbf{S}$ *is a completely simple band;*
(iii) $\mathbf{S} \models xyx \approx x$;
(iv) $\mathbf{S}$ *is isomorphic to the direct product of a left-zero and a right-zero semigroup.*

Proof. Note that if $\mathbf{S} \models xyx \approx x$ and $a \in S$, then $a = aa^3a = a^2aa^2 = a^2$, so $\mathbf{S}$ is a band. Thus, the equivalence of (i)–(iii) follows from Theorem 2.41 by taking $n = 1$.

The implication (iv) $\Rightarrow$ (iii) is true because $xyx \approx x$ holds in all left-zero and right-zero semigroups. Finally, assume (ii) to prove (iv). By Theorem 2.30, $\mathbf{S}$ is isomorphic to the direct product of a left-zero semigroup, a right-zero semigroup, and a group $\mathbf{G}$. But $\mathbf{G}$ embeds into the direct product and therefore into $\mathbf{S}$, so $\mathbf{G}$ must be trivial. □

A band $\mathbf{S}$ is called **rectangular** if it satisfies the equivalent conditions of Theorem 2.46. The reader with even a passing familiarity with semigroups has almost certainly encountered rectangular bands before, being as ubiquitous as they are. It is more usual in semigroup texts to define rectangular bands much earlier on (via condition (iii) or (iv)), but introducing them at this stage shows that, beyond simply having a nice decomposition theorem, rectangular bands are fundamental objects that arise inevitably in the study of much broader classes.

We conclude the chapter with the specialisation of Theorem 2.44 to bands: every band is a semilattice of rectangular bands.

**Theorem 2.47.** *Let* $\mathbf{S}$ *be a band. Then* $\mathcal{J}$ *is a congruence on* $\mathbf{S}$ *and* $\mathbf{S}/\mathcal{J}$ *is a semilattice. Moreover, each* $\mathcal{J}$*-class of* $\mathbf{S}$ *is a rectangular band.*

CHAPTER 3

# Residually Small Varieties of Semigroups

For reasons that are still not clear, there is an apparent connection between the residual smallness property (to be defined shortly) and the dualisability property (which we consider in Part 2). The connection is not straightforward; in some classes of algebras, it appears to be non-existent, but for others, the connection is as strong as it can be. Based on the available evidence at the commencement of the author's candidature, the situation for semigroups was the latter one. Motivated by the dualisability problem, this naturally led to an in-depth study of the known classification of residually small semigroup varieties. Presenting this classification is the aim of this chapter. As in Chapter 2, there are no new substantial results, but we offer an original simplified presentation.

The concept of a residual property appears to have its origins in group theory; Karl Gruenberg [**36**] attributes the terminology to Philip Hall. Given a property $P$, we say that an algebra $\mathbf{A}$ is **residually** $P$ if every distinct pair of elements of $\mathbf{A}$ can be separated by a homomorphism into some algebra with property $P$. For example, $\mathbf{A}$ is **residually finite** if every distinct pair of points in $\mathbf{A}$ can be separated by a homomorphism into a finite algebra.

Residual properties can be naturally extended to varieties in the obvious manner: a variety is residually $P$ if all of its members are residually $P$. One easily deduces from Birkhoff's Subdirect Decomposition Theorem 1.6 that a variety $\boldsymbol{\mathcal{V}}$ is residually finite precisely if every subdirectly irreducible member of $\boldsymbol{\mathcal{V}}$ is finite.

A landmark result in the study of residually finite groups came in 1969 when Ol'šanskiĭ characterised the residually finite varieties of groups [**50**]. Once a property of groups has been characterised, it is often natural for semigroup theorists to attempt to characterise the same property within the class of semigroups. In the case of residual finiteness, this was done by Golubov and Sapir [**32**] some ten years after Ol'šanskiĭ's result was published.

Meanwhile, variants and generalisations of the residual finiteness property were being considered by universal algebraists.

A variety $\boldsymbol{\mathcal{V}}$ is called **residually small** if there is a cardinal bound on the sizes of its subdirectly irreducible members; i.e., if there exists a cardinal $\kappa$ such that $|A| < \kappa$ for every $\mathbf{A} \in \mathbf{si}(\boldsymbol{\mathcal{V}})$. A variety that is not residually small is called **residually large**. Residual smallness is not really a residual property in the sense defined above, since it is sensibly formulated only in terms of varieties rather than individual algebras. Note that a variety is residually finite if we can choose $\kappa = \omega$ in the definition of residual smallness, so residual smallness generalises residual finiteness at the level of varieties.

The definition of residual smallness first appeared explicitly in Walter Taylor's 1971 paper [**65**]. The concept arose from a natural question: when can a variety $\boldsymbol{\mathcal{V}}$ be written as $\mathbb{SP}(\boldsymbol{\mathcal{K}})$ for some *set* $\boldsymbol{\mathcal{K}} \subseteq \boldsymbol{\mathcal{V}}$? This is possible precisely if $\boldsymbol{\mathcal{V}}$ is residually small (another corollary of Birkhoff's Theorem 1.6).

Following Taylor's foundational study of residually small varieties, the concept gained much attention from universal algebraists. Many articles were devoted in particular to answering a question posed by Robert Quackenbush in [**56**]. To state this problem, we define a variety $\boldsymbol{\mathcal{V}}$ to be **residually very finite** if there is a finite bound on the cardinalities of the members of $\mathbf{si}(\boldsymbol{\mathcal{V}})$. Quackenbush's question can then be stated as: if $\mathbf{A}$ is a finite algebra with $\mathbb{V}(\mathbf{A})$ residually finite, must $\mathbb{V}(\mathbf{A})$ be residually very finite?

Taylor proved in [**66**] that the answer to Quackenbush's question is 'yes' when $\mathbf{A}$ is of finite type and $\mathbb{V}(\mathbf{A})$ is both congruence permutable and congruence regular. As researchers progressed towards answering Quackenbush's question, it evolved into the stronger RS conjecture: if $\mathbf{A}$ is finite and $\mathbb{V}(\mathbf{A})$ is residually small, then $\mathbb{V}(\mathbf{A})$ is residually very finite.

Ralph Freese and Ralph McKenzie proved in [**27**] that the RS conjecture holds when $\mathbb{V}(\mathbf{A})$ is congruence modular (in particular, if $\mathbf{A}$ is a finite group), thus generalising Taylor's result. Shortly after, McKenzie proved that the RS conjecture holds in the case where $\mathbf{A}$ is a finite semigroup [**47**]. The proof was carried out by characterising the residually small varieties of semigroups, modulo certain aspects involving groups.

Although McKenzie does not address exactly the same problem in [**47**] that Golubov and Sapir do in [**32**], the overlap in content is significant. The collision of these two research threads is the topic of this chapter. (Though, to resolve one of the threads, McKenzie eventually showed that the general RS conjecture is false, and moreover gave a negative answer to Quackenbush's question [**49**].)

An interesting point is that it is still unknown as to whether 'residually small' is the same as 'residually finite' for semigroup varieties. As a follow-up to [**32**], and in light of McKenzie's parallel consideration of residually small semigroup varieties, Sapir and Shevrin showed in [**61**] that the question of whether these two properties coincide for semigroup varieties can be reduced to the same question for group varieties of finite exponent.

It is not presently known whether there exists a variety of groups, satisfying $x^n \approx 1$ for some $n \in \mathbb{N}$, that is residually small but not residually finite. As shown by Sapir and Shevrin [**61**], a proof that no such group variety exists would imply that every residually small semigroup variety is also residually finite. Sapir and Shevrin also show that such a group variety cannot be locally finite, and so the problem enters into the realm of Burnside groups, a notoriously difficult area of group theory. We will not attempt to explore this group-theoretic problem here, but it is certainly worth bringing to the reader's attention.

McKenzie's characterisation of residually small semigroup varieties [**47**], while mostly self-contained, does not use the techniques of semigroup theory and omits the occasional proof. It is certainly impressive that McKenzie derived these results without using the established theory of semigroups, but an unfortunate consequence is that the semigroup theorist must navigate 35 fairly dense pages using non-standard notation, spending considerable effort in connecting it to standard semigroup-theoretic notions as well as filling in the details of a number of proofs. Golubov and Sapir's approach in [**32**] is much shorter since it makes use of semigroup theory, but it relies on some rather lengthy papers.

Our account here will be almost entirely self-contained. We will require only the semigroup theory introduced in Chapter 2, some group-theoretic results (including that of

Ol'šanskiĭ [**50**]), and a short section of a paper of Rasin [**58**]. The proofs have been made accessible to semigroup theorists while also removing the dependence on certain longer papers, all without significantly increasing the combined length.

This chapter grew from the author's extensive notes on [**32**] and [**47**]. These notes were originally made purely for the author's reference, but eventually proved invaluable in deriving the novel results of this thesis. Considering that the broader dualisability problem for semigroups will need to be tackled using this theory, the inclusion of a refined version of these notes was decidedly worthwhile.

## 3.1. Nilpotent semigroups

In describing the residually small semigroup varieties, it is most natural to begin here. As we will see very shortly, any nilpotent semigroup of class at least 3 generates a residually large variety, and we will also encounter a plethora of situations forcing a variety to contain such nilpotent semigroups. From this alone, the assumption that a variety is residually small removes many properties that might be considered pathological, or at the very least inconvenient.

Let $u$ be a semigroup term. We interpret the equation $u \approx 0$ as an abbreviation for the equations $ux \approx u$ and $xu \approx u$, where $x$ is any variable not in $u$. If $\mathbf{S}$ is a semigroup, then we have $\mathbf{S} \models u \approx 0$ if and only if the term function of $\mathbf{S}$ induced by $u$ is constant and its value is a zero element of $\mathbf{S}$.

Let $\mathbf{S}$ be a semigroup. If $n \geqslant 1$ and $\mathbf{S} \models x_1 \cdots x_n \approx 0$, then $\mathbf{S}$ is called **$n$-nilpotent**. If $n \geqslant 2$ and $\mathbf{S}$ is $n$-nilpotent but not $(n-1)$-nilpotent, then $\mathbf{S}$ is called **proper $n$-nilpotent**, or **nilpotent of class $n$**.

We encountered 2-nilpotent semigroups in Chapter 2; these are precisely the null semigroups. The next result shows in particular that the variety of null semigroups is residually small; its only subdirectly irreducible member is the two-element null semigroup, which we denote by $\mathbf{N}$.

**Proposition 3.1.** *The semigroup* $\mathbf{N}$ *is the only subdirectly irreducible null semigroup. The variety* $[xy \approx 0] = [xy \approx zt]$ *of all null semigroups therefore equals* $\mathbb{SP}(\mathbf{N})$.

Proof. Let $\mathbf{S}$ be a subdirectly irreducible null semigroup, and let $a, b \in S$ with $a \neq b$. Then $\{0, a\}^2 \cup \Delta_S$ and $\{0, b\}^2 \cup \Delta_S$ are congruences on $\mathbf{S}$ which intersect to $\Delta_S$, so we have either $a = 0$ or $b = 0$. This shows that $\mathbf{S}$ can have only one non-zero element. $\square$

| $\cdot$ | $a$ | $b$ | $c$ | $0$ |
|---|---|---|---|---|
| $a$ | $0$ | $c$ | $0$ | $0$ |
| $b$ | $c$ | $0$ | $0$ | $0$ |
| $c$ | $0$ | $0$ | $0$ | $0$ |
| $0$ | $0$ | $0$ | $0$ | $0$ |

Figure 3.1. The semigroup $\mathbf{N}_4$.

Among the nilpotent semigroups of class at least 3, a special role is played by the 3-nilpotent semigroup, $\mathbf{N}_4$, defined in Figure 3.1. The next result shows that $\mathbf{N}_4$ occurs in every variety containing a nilpotent semigroup of class $n \geqslant 3$.

**Lemma 3.2.** *Let* $\mathbf{A}$ *be a commutative* 3*-nilpotent semigroup satisfying* $x^2 \approx 0$*. Then, for any nilpotent semigroup* $\mathbf{S}$ *of class at least* 3*, we have* $\mathbf{A} \in \mathbb{V}(\mathbf{S})$*.*

*Consequently, if a semigroup variety* $\boldsymbol{\mathcal{V}}$ *contains a nilpotent semigroup of class* $n$ *for some* $n \geqslant 3$*, then* $\boldsymbol{\mathcal{V}}$ *contains* $\mathbf{N}_4$*.*

Proof. Let $\mathbf{S}$ be a nilpotent semigroup of class $n \geqslant 3$. Then there exist $a_1, \ldots, a_{n-1} \in S$ with $a_1 \cdots a_{n-1} \neq 0$, but every product of length $n$ in $\mathbf{S}$ equals 0. Suppose that $a_1 a_2$ can be written as a product of length 3; i.e., $a_1 a_2 \in SSS$. Then $a_1 \cdots a_{n-1}$ can be written as a product of length $n$ and hence equals 0, which is a contradiction. Thus, $a_1 a_2 \notin SSS$, so the Rees quotient $\mathbf{S}/SSS$ is a nilpotent semigroup of class 3 contained in $\mathbb{V}(\mathbf{S})$. Since it suffices to show that $\mathbf{A} \in \mathbb{V}(\mathbf{S}/SSS)$, we may assume that $\mathbf{S}$ is nilpotent of class 3.

To show that $\mathbf{A} \in \mathbb{V}(\mathbf{S})$, we will show that $\mathbf{A}$ satisfies every equation that holds in $\mathbf{S}$. Let $u \approx v$ be an equation satisfied by $\mathbf{S}$, where $u$ and $v$ are distinct words. Clearly, $u$ and $v$ cannot both have length 1. If one of $u$ or $v$ has length 1 while the other does not, then by identifying all of the variables, $u \approx v$ implies $x^k \approx x$ for some $k > 1$, which fails in any non-trivial nilpotent semigroup. Thus, $u$ and $v$ must have length at least 2.

First, consider the case where one of $u$ or $v$ has length at least 3; by symmetry, we may assume that $u$ does. If $v$ also has length at least 3, then $u \approx v$ trivially holds in $\mathbf{A}$, so assume that $v$ has length 2. If $v$ is a product of two distinct variables, then $u \approx v$ implies $\mathbf{S}$ is null, which is false, so $v$ is the square of a variable. Then $u \approx v$ holds in $\mathbf{A}$ by assumption.

Finally, assume that $u$ and $v$ both have length exactly 2. If $u$ and $v$ contain exactly the same variables, then $\mathbf{A} \models u \approx v$ by commutativity. By symmetry, we can assume there is a variable $t$ occurring in $v$ but not $u$. Then, replacing $t$ with $t^2$ in $u \approx v$ gives $\mathbf{S} \models u \approx 0$. Because $\mathbf{S}$ is not null, $u$ must be the square of a variable $x$. By way of contradiction, suppose that $v = yz$ for some distinct variables $y$, $z$. Then $\mathbf{S} \models x^2 \approx yz$ implies $\mathbf{S}$ has an element $a$ with $a^2 \neq 0$. Now, if we let $x = a$ and set the remaining variables to 0, we get $x^2 = a^2 \neq 0$ and $yz = 0$, which is a contradiction, so we conclude that $v$ is the square of a variable. This gives $\mathbf{A} \models u \approx 0 \approx v$. □

In the next theorem, we will show that $\mathbb{V}(\mathbf{N}_4)$ is residually large, which, by Lemma 3.2, implies that any variety containing a nilpotent semigroup of class $n \geqslant 3$ is residually large. To this end, an equational basis for $\mathbb{V}(\mathbf{N}_4)$ will be useful, and Lemma 3.2 provides one with little further effort.

**Proposition 3.3.** *The variety* $\mathbb{V}(\mathbf{N}_4)$ *is axiomatised by* $x^2 \approx 0 \approx xyz$ *and* $xy \approx yx$*.*

Proof. Clearly $\mathbf{N}_4$ satisfies $x^2 \approx 0 \approx xyz$ and $xy \approx yx$. If $\mathbf{A}$ is a semigroup satisfying these identities, then $\mathbf{A} \in \mathbb{V}(\mathbf{N}_4)$ by Lemma 3.2. □

The following theorem appears to be the first non-residual finiteness result on semigroups outside of the class of groups. Based on remarks in the proof of [**32**, Lemma 1], it seems that it was first proved by Lesohin and Golubov in [**43**], though there does not appear to be an English translation of this paper. McKenzie [**47**, Lemma 15] was technically the first to prove residual largeness (rather than just non-residual finiteness), though our proof is somewhat simpler since the subdirectly irreducibles are constructed directly.

**Theorem 3.4.** *The semigroup $\mathbf{N}_4$ generates a residually large variety. Consequently, if a variety $\mathcal{V}$ contains a nilpotent semigroup of class at least 3, then $\mathcal{V}$ is residually large.*

Proof. To show that $\mathbb{V}(\mathbf{N}_4)$ is residually large, let $\kappa > 1$ be a cardinal, and let $A$ be a set of cardinality $\kappa$. Adjoin to $A$ two distinct elements $c, 0 \notin A$ to form $S := A \cup \{0, c\}$, and define $\mathbf{S}$ to be the semigroup on $S$ with multiplication defined by

$$xy := \begin{cases} c & \text{if } x, y \in A \text{ and } x \neq y, \\ 0 & \text{otherwise,} \end{cases}$$

for all $x, y \in S$. It easily verified that $\mathbf{S}$ is in the variety of $\mathbf{N}_4$ using Proposition 3.3. We will show that $\mathbf{S}$ is subdirectly irreducible with monolith generated by $(0, c)$.

Let $\theta \neq \Delta_S$ be a congruence on $\mathbf{S}$; we will show that $c \; \theta \; 0$. Choose distinct $x, y \in S$ with $x \; \theta \; y$. If $x, y \in \{0, c\}$, then we are done, so we can assume by symmetry that $x \in A$. Now, if $y \in \{0, c\}$, then taking any $z \in A \backslash \{x\}$, we get $c = xz \; \theta \; yz = 0$. On the other hand, if $y \in A$, then $c = xy \; \theta \; xx = 0$. Thus, in all cases, we have $c \; \theta \; 0$, so $\mathbf{S}$ is subdirectly irreducible. Since $|S| \geqslant \kappa$, it follows that $\mathbb{V}(\mathbf{N}_4)$ is residually large. The 'consequently' part now follows from Lemma 3.2. $\square$

With Theorem 3.4 proved, we can now derive various properties that must be satisfied by varieties not containing $\mathbf{N}_4$, which turns out to be quite a strong restriction. The most important of these properties is the following result, which will give us access to the powerful results of Section 2.5.

**Lemma 3.5.** *Let $\mathcal{V}$ be a semigroup variety with $\mathbf{N}_4 \notin \mathcal{V}$. Then $\mathcal{V}$ satisfies $x^2 \approx x^{n+2}$ for some $n \geqslant 2$. Consequently, if $\mathbf{S} \in \mathcal{V}$, then $\mathbf{S}$ has exponent $n$, and every element of $\mathbf{S}$ has index at most 2.*

Proof. Let $\mathbf{F} := \mathbf{F}_{\mathcal{V}}(\{x\})$. Then $J := \{x^k \mid k \geqslant 3\}$ is an ideal of $\mathbf{F}$ and $\mathbf{F}/J$ is 3-nilpotent. As $\mathbf{N}_4 \notin \mathcal{V}$, Lemma 3.2 implies that $\mathbf{F}/J$ is null, so $x^2 \in J$, and therefore $x^2 = x^{2+n}$ for some $n \geqslant 1$. If it happens that $n = 1$, then $x^2 = x^3$, which implies $x^2 = x^4$. Thus, we can choose $n \geqslant 2$. By Lemma 1.8, $\mathcal{V} \models x^2 \approx x^{2+n}$.

For the 'consequently' part, the statement about indices is trivial. The statement that each $\mathbf{S} \in \mathcal{V}$ has exponent $n$ amounts to saying that $\mathcal{V} \models x^{2n} \approx x^n$, which is trivial if $n = 2$. If $n > 2$, then multiplying the equation $x^{n+2} \approx x^2$ by $x^{n-2}$ gives $x^{2n} \approx x^n$. $\square$

**Remark.** The semigroup $\mathbf{N}$ satisfies $x^3 \approx x^2$ but not $x^2 \approx x$, which shows that the 'consequently' part of Lemma 3.5 does not hold when $n = 1$. This is why we forced $n \geqslant 2$ in Lemma 3.5; for simpler arguments, we would like $n$ to be an exponent of the varieties we consider. That being said, the $n = 1$ case will ultimately be the most important to us in later chapters, but we can easily obtain the relevant results after addressing the $n \geqslant 2$ case.

Next, we consider conditions for a monoid to generate a semigroup variety containing $\mathbf{N}_4$. These will serve as building blocks for a more general result for semigroups.

By adjoining an identity element to the two-element null semigroup, we of course obtain a monoid $\mathbf{N}^1$. This monoid arises naturally in the following manner.

**Proposition 3.6.** *Let* $\mathbf{M}$ *be a monoid that is not completely regular. Then the variety of* $\mathbf{M}$ *contains* $\mathbf{N}^1$.

PROOF. Take $a \in M$ such that $a$ is not a group element of $\mathbf{M}$. By Theorem 2.34, the index of $a$ is at least 2. Let $\mathbf{A}$ be the submonoid of $\mathbf{M}$ generated by $a$, which has underlying set $\{a^k \mid k \geqslant 0\}$. Then $1, a \notin \{a^k \mid k \geqslant 2\}$ as the index of $a$ is at least 2, so factoring $\mathbf{A}$ by the ideal $\{a^k \mid k \geqslant 2\}$ yields a monoid isomorphic to $\mathbf{N}^1$. Hence, $\mathbf{N}^1 \in \mathbb{V}(\mathbf{M})$. □

Proposition 3.6 shows that $\mathbf{N}^1$ can be thought of as the minimal failure of a monoid to be completely regular.

We can now easily derive the following sufficient condition for a monoid to generate a residually large semigroup variety. The proof is from Jackson's article [**40**, Theorem 8.1(1)].

**Lemma 3.7.** *The variety of* $\mathbf{N}^1$ *contains* $\mathbf{N}_4$*. Consequently, if* $\mathbf{M}$ *is a monoid that is not completely regular, then the semigroup variety of* $\mathbf{M}$ *contains* $\mathbf{N}_4$ *and is therefore residually large.*

PROOF. Write the underlying set of $\mathbf{N}$ as $\{0, \emptyset\}$. Let $J$ be the ideal of pairs in $\mathbf{N}^1 \times \mathbf{N}^1$ with at least one zero coordinate. The table below shows that $S := \{(1,\emptyset), (\emptyset,1), (\emptyset,\emptyset)\} \cup J$ forms a subsemigroup $\mathbf{S}$ of $\mathbf{N}^1 \times \mathbf{N}^1$.

| $\cdot$ | $(1,\emptyset)$ | $(\emptyset,1)$ | $(\emptyset,\emptyset)$ |
|---|---|---|---|
| $(1,\emptyset)$ | $J$ | $(\emptyset,\emptyset)$ | $J$ |
| $(\emptyset,1)$ | $(\emptyset,\emptyset)$ | $J$ | $J$ |
| $(\emptyset,\emptyset)$ | $J$ | $J$ | $J$ |

In the table, a $J$ entry indicates that the corresponding product lies in $J$. Clearly $\mathbf{S}/J \cong \mathbf{N}_4$, so $\mathbf{N}_4 \in \mathbb{V}(\mathbf{N}^1)$. Now, if $\mathbf{M}$ is a monoid that is not completely regular, then its variety contains $\mathbf{N}_4$ by Proposition 3.6, so the result follows from Theorem 3.4. □

To apply Lemma 3.7, we are lead to introducing the two semigroups $\mathbf{P}$ and $\mathbf{Q}$ in Figure 3.2. These are important semigroups in general, but they will be especially important for this thesis.

The next result, due to the author, shows how $\mathbf{P}$ (and by duality, $\mathbf{Q}$) can arise in a variety under somewhat broad conditions.

| $\cdot$ | $e$ | $u$ | $0$ |
|---|---|---|---|
| $e$ | $e$ | $u$ | $0$ |
| $u$ | $0$ | $0$ | $0$ |
| $0$ | $0$ | $0$ | $0$ |

**P**

| $\cdot$ | $e$ | $u$ | $0$ |
|---|---|---|---|
| $e$ | $e$ | $0$ | $0$ |
| $u$ | $u$ | $0$ | $0$ |
| $0$ | $0$ | $0$ | $0$ |

**Q**

FIGURE 3.2. The semigroups $\mathbf{P}$ and $\mathbf{Q}$.

**Lemma 3.8.** *Let* $\mathbf{S}$ *be a semigroup, let* $u \in S$ *be a non-group element, and let* $e \in S$ *be an idempotent with* $eu = u$*. Then* $\mathbb{V}(\mathbf{S})$ *contains either* $\mathbf{P}$ *or* $\mathbf{N}_4$*.*

Proof. Without loss of generality, we can assume that $\mathbf{S}$ is generated by $\{e, u\}$. We will separately consider the cases $u \in \langle ue \rangle$ and $u \notin \langle ue \rangle$.

Assume that $u \in \langle ue \rangle$, so that $u$ can be written as a word $w$ in $u$ and $e$ containing $ue$ as a subword. If $w$ ends in $u$, then $w$ contains at least two occurrences of $u$ ($ue$ is a subword), and we can remove all occurrences of $e$ in $w$ using $eu = u$; this gives $u = w = u^k$ for some $k \geqslant 2$. But $u^k = u$ implies that $u$ is a group element (Theorem 2.34), which is a contradiction. It follows that $w$ ends in $e$, and hence $u \in Se$. This then gives $ue = u$, so $\mathbf{S}$ is a monoid that is not completely regular, and the result follows from Lemma 3.7.

Now assume that $u \notin \langle ue \rangle$. Then $e \notin \langle ue \rangle$ as $u = eu \preccurlyeq_J e$. Since $uu = ueu \in \langle ue \rangle$, it follows that $\{\{e\}, \{u\}, \langle ue \rangle\}$ forms a subsemigroup of $\mathbf{S}/\langle ue \rangle$ isomorphic to $\mathbf{P}$. □

Since $\mathbf{P}$ and $\mathbf{Q}$ are dual to each other, they are in a sense equally important semigroups, but, due to the next result, we will usually be dealing with only one of them. Again, the proof is from Jackson's article [**40**, Theorem 9.1].

**Lemma 3.9.** *The variety* $\mathbb{V}(\mathbf{P}, \mathbf{Q})$ *contains* $\mathbf{N}_4$ *and is therefore residually large.*

Proof. Let $J$ be the set of elements of $\mathbf{P} \times \mathbf{Q}$ with at least one zero coordinate. Then the set $S := \{(u, u), (e, u), (u, e)\} \cup J$ forms a subsemigroup $\mathbf{S}$ of $\mathbf{P} \times \mathbf{Q}$, as can be seen from the following table of products.

| $\cdot$ | $(e,u)$ | $(u,e)$ | $(u,u)$ |
|---|---|---|---|
| $(e,u)$ | $J$ | $(u,u)$ | $J$ |
| $(u,e)$ | $J$ | $J$ | $J$ |
| $(u,u)$ | $J$ | $J$ | $J$ |

Now $\mathbf{S}/J \in \mathbb{V}(\mathbf{P}, \mathbf{Q})$ is nilpotent of class 3, so $\mathbf{N}_4 \in \mathbb{V}(\mathbf{P}, \mathbf{Q})$ by Lemma 3.2. □

We now come to a simple condition for a variety to contain $\mathbf{N}_4$, due to the author.

**Theorem 3.10.** *Let* $\mathbf{S}$ *be a semigroup with idempotent elements* $e, f \in S$ *and a non-group element* $u \in S$ *such that* $eu = uf = u$*. Then* $\mathbb{V}(\mathbf{S})$ *contains* $\mathbf{N}_4$ *and so is residually large.*

Proof. By Lemma 3.8 and its dual, either $\mathbf{P}, \mathbf{Q} \in \mathbb{V}(\mathbf{S})$ or $\mathbf{N}_4 \in \mathbb{V}(\mathbf{S})$. But in the former case, we have $\mathbf{N}_4 \in \mathbb{V}(\mathbf{S})$ by Lemma 3.9. □

The semigroups $\mathbf{P}$ and $\mathbf{Q}$ will now fade into the background; we will encounter them again in Section 3.5. In the remainder of this section, we will deduce some important results that are now obtained as easy consequences of Theorem 3.10.

In [**40**, Theorem 4.3], it is shown that a regular semigroup that is not completely regular has $\mathbf{N}_4$ in its variety. Theorem 3.10 gives us the following extension of this result.

**Theorem 3.11.** *Let* $\mathbf{S}$ *be a semigroup with a regular element* $u \in S$ *that is not a group element. Then* $\mathbb{V}(\mathbf{S})$ *contains* $\mathbf{N}_4$ *and is therefore residually large.*

Proof. Let $x \in S$ with $uxu = u$. Then $e := ux$ and $f := xu$ are idempotent and $eu = uf = u$, so the result follows from Theorem 3.10. □

Theorem 3.10 also allows us to strengthen Theorem 2.39 to the following result.

**Theorem 3.12.** *Let* $\mathbf{S}$ *be a semigroup with* $\mathbf{N}_4 \notin \mathbb{V}(\mathbf{S})$. *Then the following hold:*

(i) *if* $\mathbf{S}$ *is simple, then* $\mathbf{S}$ *is completely simple;*

(ii) *if* $\mathbf{S}$ *is zero-simple, then* $\mathbf{S} \cong \mathbf{T}^0$ *for some completely simple semigroup* $\mathbf{T}$.

PROOF. It suffices to prove (ii). Assume that $\mathbf{S}$ is zero-simple. By Lemma 3.5, $\mathbf{S}$ is periodic, so $\mathbf{S}$ is completely zero-simple by Theorem 2.39. Thus, $\mathbf{S}$ is regular by Theorem 2.25(i), and is therefore completely regular by Theorem 3.11. Since the $\mathcal{J}$-classes of $\mathbf{S}$ are $S\backslash\{0\}$ and $\{0\}$, the result now follows from Theorem 2.44. □

In Theorem 2.11, which applies to any semigroup, we saw that principal factors must be either simple, zero-simple, or null. The next result shows that, in a variety not containing $\mathbf{N}_4$, we are reduced to the dichotomy we hinted at in Example 2.5: a $\mathcal{J}$-class is either a completely simple subsemigroup, or its principal factor is a null semigroup.

**Theorem 3.13.** *Let* $\mathbf{S}$ *be a semigroup with* $\mathbf{N}_4 \notin \mathbb{V}(\mathbf{S})$, *and let* $a \in S$. *If there are* $x, y \in J_a$ *with* $xy \in J_a$, *then* $J_a$ *forms a completely simple subsemigroup of* $\mathbf{S}$, *and so every element of* $J_a$ *is a group element of* $\mathbf{S}$.

PROOF. If $J_a$ is the minimum ideal of $\mathbf{S}$, then Theorems 2.11, 2.43, and 3.12(i) give the result. Thus, we can assume that $I := \{x \in S \mid J_x < J_a\}$ is non-empty. Now, the principal factor $\langle a\rangle/I$ is, by assumption, not a null semigroup, so it is zero-simple by Theorem 2.11. By Theorem 3.12(ii), the non-zero elements of $\langle a\rangle/I$ form a completely simple subsemigroup of $\langle a\rangle/I$, which implies that $J_a$ forms a completely simple subsemigroup of $\mathbf{S}$. □

Our final result shows that in a variety $\boldsymbol{\mathcal{V}}$ not containing $\mathbf{N}_4$, the group elements of any $\mathbf{S} \in \boldsymbol{\mathcal{V}}$ must form a subsemigroup of $\mathbf{S}$, as hinted at in Section 2.6. This can be obtained as a corollary of [**62**, Theorem 2], but Theorem 3.10 again yields a simple direct argument.

**Theorem 3.14.** *Let* $\mathbf{S}$ *be a semigroup such that the set of group elements of* $\mathbf{S}$ *does not form a subsemigroup of* $\mathbf{S}$. *Then* $\mathbb{V}(\mathbf{S})$ *contains* $\mathbf{N}_4$ *and is therefore residually large.*

PROOF. Since the set of group elements is not a subsemigroup, it is either empty or is not closed under multiplication. If $\mathbf{S}$ has no group elements, then $\mathbf{S}$ has no idempotents, so $\mathbf{S}$ is not periodic by Theorem 2.36. Thus, by Lemma 3.5, $\mathbb{V}(\mathbf{S})$ contains $\mathbf{N}_4$. We can therefore assume $\mathbf{S}$ has a non-empty set of group elements.

Let $a, b \in S$ be group elements of $\mathbf{S}$ such that $ab$ is not a group element of $\mathbf{S}$, and let $e$ and $f$ be the identity elements of $H_a$ and $H_b$, respectively. By assumption, $u := ab$ is not a group element of $\mathbf{S}$, and $eu = u = uf$. Thus, $\mathbb{V}(\mathbf{S})$ contains $\mathbf{N}_4$ by Theorem 3.10. □

## 3.2. Two identities for residually small varieties

From Section 3.1, we know that a residually small semigroup variety $\boldsymbol{\mathcal{V}}$ satisfies $x^2 \approx x^{n+2}$ for some $n \geqslant 2$. The goal of this section is to show that $\boldsymbol{\mathcal{V}}$ also satisfies at least one of two powerful identities: $x^{n+1}y \approx xy$ or $xy^{n+1} \approx xy$. First, we will show that $\mathbf{N}_4 \notin \boldsymbol{\mathcal{V}}$ leads to three possible identities. We will then introduce two semigroups $\mathbf{L}^+$ and $\mathbf{R}^+$ whose varieties are residually large and use their exclusion from $\boldsymbol{\mathcal{V}}$ to get our two desired identities.

In Theorem 3.14, we saw that a semigroup $\mathbf{S}$ with $\mathbf{N}_4 \notin \mathbb{V}(\mathbf{S})$ has a subsemigroup consisting of all of its group elements. This subsemigroup is evidently equal to the union of all subgroups of $\mathbf{S}$, so it is completely regular, and we will therefore be able to utilise the results of Section 2.6.

The subsemigroup of group elements will be an important object, so it will be convenient to introduce notation for it. If $\mathbf{S}$ is a semigroup, then we denote by $\mathrm{G}(\mathbf{S})$ the set of group elements of $\mathbf{S}$. We also denote by $\mathrm{E}(\mathbf{S})$ the set of idempotent elements of $\mathbf{S}$, which is contained in $\mathrm{G}(\mathbf{S})$ as a subset. We will see in Section 3.3 that $\mathrm{E}(\mathbf{S})$ is also a subsemigroup of $\mathbf{S}$ when $\mathbb{V}(\mathbf{S})$ is residually small.

Due to Lemma 3.5, we will be dealing exclusively with semigroups of finite exponent in this chapter. The next result with therefore be of fundamental importance. We will state and prove it for clarity, but we will frequently use this result without mention.

**Proposition 3.15.** *Let $\mathbf{S}$ be a semigroup of exponent $n \geqslant 1$. Then the following hold:*

(i) $\mathrm{E}(\mathbf{S}) = \{x^n \mid x \in S\}$;
(ii) *if $a \in \mathrm{G}(\mathbf{S})$, then $a^n$ is the identity element of $H_a$;*
(iii) $\mathrm{G}(\mathbf{S}) = \{x \in S \mid x^{n+1} = x\} = \{x^{n+1} \mid x \in S\}$.

Proof. Statement (i) follows immediatly from the definition of exponent.

To prove (ii), let $a \in \mathrm{G}(\mathbf{S})$, so $H_a$ is a subgroup of $\mathbf{S}$. Then $H_a$ has exponent $n$, so $a^n$ is the identity element of $H_a$.

To prove (iii), let $a \in S$. Since $a^n$ is idempotent, the index of $a$ is at most $n$, so by Theorem 2.34 we have $a^{n+1} \in \mathrm{G}(\mathbf{S})$. This gives $\{x^{n+1} \mid x \in S\} \subseteq \mathrm{G}(\mathbf{S})$. Now, if $a \in \mathrm{G}(\mathbf{S})$, then $a^{n+1} = a$ by (ii), which gives $\mathrm{G}(\mathbf{S}) \subseteq \{x \in S \mid x^{n+1} = x\} \subseteq \{x^{n+1} \mid x \in S\}$. $\square$

The following result gives us three possible identities for residually small varieties, using the results of Section 3.1. The proof is based on the original proofs of Golubov, Sapir, and McKenzie, but the number of cases has been reduced to the number of identities.

**Lemma 3.16.** *Let $\mathcal{V}$ be a variety with $\mathbf{N}_4 \notin \mathcal{V}$. Then $\mathcal{V} \models x^2 \approx x^{n+2}$ for some $n \geqslant 2$, and for every such choice of $n$, at least one of the following identities holds in $\mathcal{V}$:*

$$(xy)^{n+1} \approx xy, \qquad x^{n+1}y \approx xy, \qquad xy^{n+1} \approx xy.$$

Proof. By Lemma 3.5, we have $\mathcal{V} \models x^2 \approx x^{n+2}$ for some $n \geqslant 2$. Let $\mathbf{F} := \mathbf{F}_{\mathcal{V}}(\{x, y\})$. The semigroup $\mathbf{F}/FFF \in \mathcal{V}$ is 3-nilpotent, so it is null by Lemma 3.2. Thus, $xy \in FFF$, and so in $\mathbf{F}$ we have $xy = w$ for some word $w$ in $x, y$ of length at least 3.

Consider first the case where $yx$ is a subword of $w$, so $w = uyxv$ for some words $u$, $v$, at least one of which is non-empty. Let $z \mapsto \overline{z}$ denote the automorphism of $\mathbf{F}$ interchanging $x$ and $y$. Then $xy = uyxv$ implies $yx = \overline{u}xy\overline{v}$, so $xy = uyxv = u\overline{u}xy\overline{v}v$.

We aim to show that there are words $s$, $t$, *both* non-empty, such that $xy = sxyt$ in $\mathbf{F}$. If $u$ and $v$ are both non-empty, we are done, so by symmetry we can assume that $v$ is empty and $u$ is not; thus, $xy = u\overline{u}xy$. The Pumping Lemma 2.37 with $k = 3$ gives $xy = u\overline{u}(u\overline{u})^2xy$. Now, $u\overline{u}$ contains $x$ and $y$, so $(u\overline{u})^2$ contains the subword $xy$. Thus, $xy = u\overline{u}(u\overline{u})^2xy = sxyt$ for some non-empty words $s$, $t$, as required.

The Pumping Lemma 2.37 now gives $xy = s^n xyt^n$. Letting $e := s^n$ and $f := t^n$, we have $e, f \in \mathrm{E}(\mathbf{F})$ and $exy = xyf = xy$. Now $xy \in \mathrm{G}(\mathbf{F})$ by Theorem 3.10, so $(xy)^{n+1} = xy$, and hence $\boldsymbol{\mathcal{V}} \models (xy)^{n+1} \approx xy$ by Lemma 1.8.

We are left with the case where $w$ does not contain the subword $yx$, so $xy = x^k y^\ell$ for some $k, \ell \geqslant 0$ with $k + \ell \geqslant 3$. Then either $k \geqslant 2$ or $\ell \geqslant 2$. If $k \geqslant 2$, then $x^k \approx x^{n+k}$ follows from $x^2 \approx x^{n+2}$, so $x^n xy = x^n x^k y^\ell = x^{n+k} y^\ell = x^k y^\ell = xy$, and we get $\boldsymbol{\mathcal{V}} \models x^{n+1} y \approx xy$ by Lemma 1.8. Dually, if $\ell \geqslant 2$, we get $\boldsymbol{\mathcal{V}} \models xy^{n+1} \approx xy$. $\square$

Now, we want to further narrow down the equations that hold in a residually small variety. The next three lemmas will take us as far as we can go by omitting only $\mathbf{N}_4$. After this, we will introduce the semigroups $\mathbf{L}^+$ and $\mathbf{R}^+$.

Theorem 3.14 is used implicitly in the statement of the next result, which shows that the ordered set $\mathrm{G}(\mathbf{S})/\mathcal{J}$ sits inside the ordered set $S/\mathcal{J}$ when $\mathbb{V}(\mathbf{S})$ is residually small.

**Lemma 3.17.** *Let $\mathbf{S}$ be a semigroup with $\mathbf{N}_4 \notin \mathbb{V}(\mathbf{S})$, and define $\mathbf{T} := \mathrm{G}(\mathbf{S})$. Then every $\mathcal{J}$-class of $\mathbf{T}$ is a $\mathcal{J}$-class of $\mathbf{S}$, and $\preccurlyeq_J^{\mathbf{S}}$ agrees with $\preccurlyeq_J^{\mathbf{T}}$ on $T$.*

Proof. By Theorem 3.13, we have $J_a^{\mathbf{S}} \subseteq T$ for all $a \in T$, which is to say that if $a \in T$ is $\mathcal{J}^{\mathbf{S}}$-related to $b \in S$, then $b \in T$. To show that every $\mathcal{J}$-class of $\mathbf{T}$ is a $\mathcal{J}$-class of $\mathbf{S}$, it suffices to show that $a \mathrel{\mathcal{J}^{\mathbf{S}}} b \Rightarrow a \mathrel{\mathcal{J}^{\mathbf{T}}} b$ for all $a, b \in T$. For this, in turn, it will suffice to show that $a \preccurlyeq_J^{\mathbf{S}} b \Rightarrow a \preccurlyeq_J^{\mathbf{T}} b$ for all $a, b \in T$, and this will also prove the entire result.

Assume that $a = ubv$ for some $u, v \in S^1$. Since $a = aa^{-1}aa^{-1}a = aa^{-1}ubva^{-1}a$, we can assume that $u, v \in S$ and that $u, v \preccurlyeq_J a$. But then $a = ubv \preccurlyeq_J u, v$, so $u, v \in J_a^{\mathbf{S}} \subseteq T$. Thus, we have $a = ubv$ for some $u, v \in T$, as required. $\square$

The following result is a special case of a result that apparently appears in reference 13 of [**62**], though there does not appear to be an English translation of the referenced paper, so the following proof is due to the author. Note the implicit use of Theorem 2.44.

**Lemma 3.18.** *Let $\mathbf{S}$ be a semigroup such that $\mathbf{N}_4 \notin \mathbb{V}(\mathbf{S})$, and let $\mathbf{T} := \mathrm{G}(\mathbf{S})$. Then the map $\varphi\colon \mathbf{S} \to \mathbf{T}/\mathcal{J}$ given by $x \mapsto J_{x^2}$ is a homomorphism, and $\ker(\varphi{\restriction_T}) = \mathcal{J}^{\mathbf{T}}$.*

Proof. By Lemma 3.17, we have $J_a^{\mathbf{T}} = J_a^{\mathbf{S}}$ for all $a \in T$, so we can omit superscripts on $\mathcal{J}$-classes without confusion. By Lemma 3.5 and Theorem 2.34, we also have $a^2 \in T$ for all $a \in S$, so $\varphi$ is well defined. Since $a \mathrel{\mathcal{J}} a^2$ for all $a \in T$, we have for all $a, b \in T$ that

$$\varphi(a) = \varphi(b) \iff J_{a^2} = J_{b^2} \iff a^2 \mathrel{\mathcal{J}} b^2 \iff a \mathrel{\mathcal{J}} b,$$

so $\ker(\varphi{\restriction_T}) = \mathcal{J}^{\mathbf{T}}$. It remains to show that $\varphi$ is a homomorphism.

Let $a, b \in S$. Then $\varphi(ab) = J_{(ab)^2}$, while $\varphi(a)\varphi(b) = J_{a^2} J_{b^2} = J_{a^2 b^2}$ (using the multiplication of $\mathbf{T}/\mathcal{J}$), so showing that $\varphi(ab) = \varphi(a)\varphi(b)$ amounts to showing that

$$(ab)^2 \mathrel{\mathcal{J}} a^2 b^2.$$

By Lemma 3.16, we can assume by symmetry that $\mathbf{S}$ satisfies $(xy)^{n+1} \approx xy$ or $x^{n+1} y \approx xy$ for some $n \geqslant 2$. Using either of these identities, we get $a^2 b^2 = aabb = a(ab)^{n+1} b \preccurlyeq_J (ab)^2$, so it remains to show that $(ab)^2 \preccurlyeq_J a^2 b^2$. Now, in the case that $\mathbf{S}$ satisfies $x^{n+1} y \approx xy$, we have $abab = a^{n+1} b^{n+1} ab \preccurlyeq_J a^2 b^2$, so we can assume that $\mathbf{S} \models (xy)^{n+1} \approx xy$. We then have

$$(ab)^2 = abab \preccurlyeq_J aba = (aba)(aba)^n \preccurlyeq_J a^2,$$

and $(ab)^2 \preccurlyeq_J b^2$ by symmetry. By Lemma 3.17 and Theorem 2.44, $\preccurlyeq_J$ is compatible with $\mathbf{T}$, so $(ab)^4 \preccurlyeq_J a^2b^2$. Finally, since $(ab)^2 \in T$, we have $(ab)^2 \mathrel{\mathcal{J}} (ab)^4$, hence $(ab)^2 \preccurlyeq_J a^2b^2$. $\square$

Our only use of Lemma 3.18 will be to prove Lemma 3.19 below. Recall from Theorem 2.41 that a completely simple semigroup of exponent $n$ satisfies $(xy)^n x \approx x$. The shadow of this equation appears in the next result.

**Lemma 3.19.** *Let $\mathcal{V}$ be a variety not containing $\mathbf{N}_4$, and assume that $\mathcal{V} \models x^2 \approx x^{n+2}$ for some $n \geqslant 2$. If $w$ is a word in the variables $x$ and $y$ that contains both $x$ and $y$, then $\mathcal{V}$ satisfies $((xy)^n w)^n (xy)^n \approx (xy)^n$.*

Proof. Let $\mathbf{F} := \mathbf{F}_{\mathcal{V}}(\{x, y\})$. Using any one of the three equations from Lemma 3.16, we get $\mathcal{V} \models ((xy)^n w)^n (xy)^n \approx ((xy)^n w^{n+1})^n (xy)^n$, so it suffices to prove the desired identity with $w^{n+1}$ in place of $w$. Thus, we may assume without loss of generality that $w \in \mathrm{G}(\mathbf{F})$.

Let $\varphi\colon \mathbf{F} \to \mathrm{G}(\mathbf{F})/\mathcal{J}$ be the homomorphism from Lemma 3.18. Then $\varphi(\mathbf{F})$ is a semilattice by Theorem 2.44, so because $w$ contains $x$ and $y$, we have $\varphi(w) = \varphi(x)\varphi(y) = \varphi((xy)^n)$. But $w, (xy)^n \in \mathrm{G}(\mathbf{F})$, so by Lemma 3.18 we have $(xy)^n \mathrel{\mathcal{J}} w$ in $\mathrm{G}(\mathbf{F})$. By Theorem 2.44, it follows that $(xy)^n$ and $w$ lie in some completely simple subsemigroup of $\mathbf{F}$. Hence, by Theorem 2.41, we have $((xy)^n w)^n (xy)^n = (xy)^n$. Lemma 1.8 gives the result. $\square$

For the remainder of this section we aim to show that if $\mathcal{V}$ is a residually small semigroup variety, then there is some $n \geqslant 2$ such that $\mathcal{V}$ satisfies $x^{n+1}y \approx xy$ or $xy \approx xy^{n+1}$, thus strengthening Lemma 3.16. To obtain our desired result, we introduce two more minimal residually large varieties. We denote by $\mathbf{L}^+$ the semigroup in Figure 3.3, and the dual semigroup by $\mathbf{R}^+$.

| $\cdot$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $a$ | $a$ | $a$ | $a$ | $a$ |
| $b$ | $b$ | $b$ | $b$ | $b$ |
| $c$ | $c$ | $c$ | $c$ | $c$ |
| $d$ | $a$ | $a$ | $b$ | $a$ |

Figure 3.3. The semigroup $\mathbf{L}^+$.

We may represent $\mathbf{L}^+$ as a semigroup of self-maps on $\{0, 1, 2\}$. Take $a$, $b$, and $c$ to be the constant maps with values 0, 1, and 2, respectively, and take $d$ to be the following map:

$$\circlearrowleft 0 \leftarrow 1 \leftarrow 2$$

This shows that $\mathbf{L}^+$ is indeed a semigroup. We will show that $\mathbf{L}^+$ generates a residually large variety (and therefore so does $\mathbf{R}^+$). For this, it will be useful to give an equational basis for its identities.

**Proposition 3.20.** *The variety $\mathbb{V}(\mathbf{L}^+)$ is axiomatised by the equation $xyz \approx xy$. Moreover, a non-trivial equation $u \approx v$ holds in $\mathbf{L}^+$ if and only if $u$ and $v$ have length at least 2 and agree on the two left-most variables.*

Proof. Since all products in $\mathbf{L}^+$ are left-zero elements, we have $\mathbf{L}^+ \models xyz \approx xy$. Now, if two words $u$, $v$ agree on the two left-most variables, it is clear that $u \approx v$ can be deduced from $xyz \approx xy$. Thus, it suffices to prove the 'moreover' part in the left-to-right direction.

Let $u$ and $v$ be distinct words with $\mathbf{L}^+ \models u \approx v$. Note that $\{a, d\}$ forms a subsemigroup of $\mathbf{L}^+$ isomorphic to $\mathbf{N}$, so $\mathbf{N} \models u \approx v$, from which it is clear that $u$ and $v$ must have length at least 2. Also, $\mathbf{L}^+$ contains a non-trivial left-zero subsemigroup, so the left-most variables of $u$ and $v$ must be the same. Thus, $u = xyu'$ and $v = xzv'$ for some variables $x, y, z$ and some possibly empty words $u'$, $v'$. We must show that $y = z$. Suppose not; then $x$ differs from $y$ or $z$, so we can assume by symmetry that $x \neq y$. Now, if we set $(x, y, z) = (d, c, d)$ in $\mathbf{L}^+$, and assign arbitrary values to the remaining variables, then we get $u = dc = b \neq a = dd = v$, which contradicts $\mathbf{L}^+ \models u \approx v$. Thus, $y = z$, so $u$ and $v$ agree on the two left-most variables, as required. $\square$

The construction in the following proof was originally based on that of Golubov and Sapir [**32**, Lemma 1], but it evolved into a simpler form, partly by using elements of Gerhard's paper [**30**]. Golubov and Sapir defined their semigroups via a table, so associativity needed to be checked, and they used the maximal congruence argument to infer the existence of large subdirectly irreducibles. We construct our subdirectly irreducibles directly as semigroups of maps.

**Theorem 3.21.** *The semigroups $\mathbf{L}^+$ and $\mathbf{R}^+$ generate residually large varieties.*

Proof. It suffices to prove the result for $\mathbf{L}^+$. Let $\kappa > 1$ be a cardinal. For each $\alpha \in \kappa$, let $\underline{\alpha}$ denote the constant map $\kappa \to \kappa$ with value $\alpha$, and define

$$0^1_\alpha \colon \kappa \to \kappa, \qquad 0^1_\alpha(x) := \begin{cases} 1 & \text{if } x = \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

Now, let

$$S := \{\underline{\alpha} \mid \alpha \in \kappa\} \cup \{0^1_\alpha \mid \alpha \in \kappa\backslash\{0, 1\}\},$$

and let $x, y \in S$. If $x$ or $y$ is constant, then $x \circ y$ is constant, and if $x, y \in \{0^1_\alpha \mid \alpha \in \kappa\backslash\{0, 1\}\}$ then $x \circ y = \underline{0}$. In either case, $x \circ y$ is constant, so $x \circ y \in S$. Thus, $\mathbf{S} := \langle S; \circ\rangle$ is a semigroup in which every product is constant, so $\mathbf{S} \models xyz \approx xy$. Now $\mathbf{S} \in \mathbb{V}(\mathbf{L}^+)$ by Proposition 3.20, and clearly $|S| \geqslant \kappa$.

We will show that $\mathbf{S}$ is subdirectly irreducible by showing that $\underline{0}$ and $\underline{1}$ are identified by every non-trivial congruence on $\mathbf{S}$. Let $\theta \neq \Delta_S$ be a congruence on $\mathbf{S}$, and choose $x, y \in S$ with $x \neq y$ and $x \mathrel{\theta} y$. Then there is a constant map $z \in S$ such that $xz \neq yz$, and so $xz$ and $yz$ are distinct constant maps and $xz \mathrel{\theta} yz$; in other words, there are distinct $\alpha, \beta \in \kappa$ with $\underline{\alpha} \mathrel{\theta} \underline{\beta}$. Now, if $\{\alpha, \beta\} = \{0, 1\}$, then we are done, so we may assume by symmetry that $\alpha \notin \{0, 1\}$. Then $\underline{\alpha} \mathrel{\theta} \underline{\beta}$ implies $\underline{1} = 0^1_\alpha\,\underline{\alpha} \mathrel{\theta} 0^1_\alpha\,\underline{\beta} = \underline{0}$, as required. $\square$

**Remark.** Taking $\kappa = 3$ in the above proof gives $\mathbf{S} = \mathbf{L}^+$, so we see incidentally that $\mathbf{L}^+$ is subdirectly irreducible, with its monolith generated by $(a, b)$.

**Remark.** To provide some intuition for the role of $\mathbf{L}^+$ and $\mathbf{R}^+$, we will give a brief historical note. In [**52**, §3], Petrich determines the subvariety lattice of $[uvxyz \approx uvyz,\ (xy)^2 \approx xy]$, which is the variety of semigroups $\mathbf{S}$ such that $SS$ is a rectangular band. We show this subvariety lattice in Figure 3.2, though we will not attempt to prove its correctness, since we will not need to use it here.

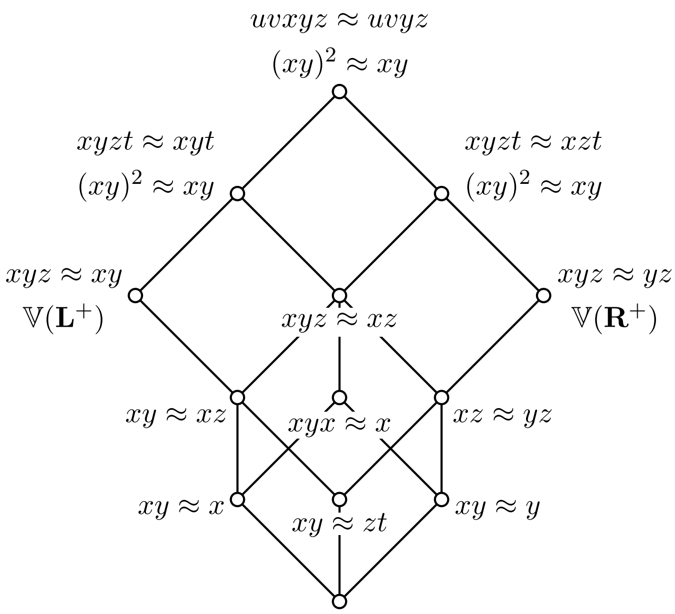



FIGURE 3.4. The lattice of subvarieties of $[uvxyz \approx uvyz,\ (xy)^2 \approx xy]$.

From Figure 3.2, it can be seen that any subvariety of $[uvxyz \approx uvyz,\ (xy)^2 \approx xy]$ not containing $\mathbf{L}^+$ or $\mathbf{R}^+$ is a subvariety of $[xyz \approx xz]$, which turns out to be the variety of subdirect products of rectangular bands and null semigroups. This echoes the role of $\mathbf{L}^+$ and $\mathbf{R}^+$ in the context of this chapter, though the general picture is more complicated.

This role was recognised by Golubov and Sapir [**32**]. In part of their proof, Golubov and Sapir drew on Petrich's description of various lattices similar to that in Figure 3.2, though we were able to circumvent the dependence on Petrich's results.

It is not observed in [**52**] or [**32**] that $[xyz \approx xy]$ and $[xyz \approx yz]$ have the finite generators $\mathbf{L}^+$ and $\mathbf{R}^+$, which is most likely because finite generators were not of interest. Petrich's arguments in [**52**] are almost exclusively syntactic. As far we are aware, the semigroups $\mathbf{L}^+$ and $\mathbf{R}^+$ have appeared (in the context of residual character) only in McKenzie's paper [**47**, Lemma 21]. The names $\mathbf{L}^+$ and $\mathbf{R}^+$ were chosen by the present author.

The remainder of this section follows the approach of Golubov and Sapir [**32**], so it does not make explicit use of the generators $\mathbf{L}^+$ and $\mathbf{R}^+$. However, the fact that finite generators exist will be extremely useful for our dualisability results in later chapters.

**Lemma 3.22.** *Let $\mathcal{V}$ be a variety with $\mathbf{N}_4, \mathbf{R}^+ \notin \mathcal{V}$, and assume that $\mathcal{V} \models x^2 \approx x^{n+2}$ for some $n \geqslant 2$. Then $\mathcal{V} \models (xy)^{n+1} \approx (xy)^{n+1}y^n$. Consequently, the following hold:*

(i) *if $\mathcal{V} \models (xy)^{n+1} \approx xy$, then $\mathcal{V} \models xy^{n+1} \approx xy$;*
(ii) *if $\mathcal{V} \models x^{n+1}y \approx xy$, then $\mathcal{V} \models (xy)^{n+1} \approx x^{n+1}y^{n+1}$.*

PROOF. Let $u \approx v$ be a non-trivial identity that holds in $\mathcal{V}$ but not in $\mathbf{R}^+$. Then, by Proposition 3.20, either

(1) the right-most variables of $u$ and $v$ are different, or
(2) $u = u'y$ and $v = v'y$ for some variable $y$ and some non-empty words $u'$, $v'$ with different right-most variables, or
(3) one of $u$, $v$ is a single variable while the other is not.

Case (3) implies that $\mathcal{V} \models x^k \approx x$ for some $k \geqslant 2$, so every member of $\mathcal{V}$ is completely regular, and $(xy)^{n+1} \approx (xy)^{n+1}y^n$ then follows from $y \approx y^{n+1}$. In case (1), we can multiply the identity $u \approx v$ on the right by some variable to bring us to case (2). So, we may assume that (2) holds. Now, $u'$ and $v'$ cannot both end in $y$, so we can assume that $u'$ ends in some variable $x \neq y$. By setting all other variables equal to $y$, we get $\mathcal{V} \models sxy \approx ty^2$ for some words $s$, $t$ in the variables $x$, $y$. We deduce that

$$\mathcal{V} \models sxyy^n \approx ty^2y^n \approx ty^{2+n} \approx ty^2 \approx sxy.$$

Now, substituting $x \mapsto (xy)^n x$ and $y \mapsto y$ in the identity $sxy^{n+1} \approx sxy$, we get an identity of the form $w(xy)^n xy^{n+1} \approx w(xy)^n xy$ for some word $w$. By multiplying on the left by $x$ and $y$, we can assume that $w$ contains both variables. We then have $\mathcal{V} \models ((xy)^n w)^n (xy)^n \approx (xy)^n$ by Lemma 3.19, and so we deduce

$$\mathcal{V} \models (xy)^n xy \approx ((xy)^n w)^n (xy)^n xy \approx ((xy)^n w)^n (xy)^n xy^{n+1} \approx (xy)^n xy^{n+1} \approx (xy)^{n+1}y^n,$$

with the second '$\approx$' following from $w(xy)^n xy \approx w(xy)^n xy^{n+1}$.

Now, if $\mathcal{V} \models (xy)^{n+1} \approx xy$, then $\mathcal{V} \models xy \approx (xy)^{n+1} \approx (xy)^{n+1}y^n \approx xy^{n+1}$, proving (i). If $\mathcal{V} \models x^{n+1}y \approx xy$, then $\mathcal{V} \models (xy)^{n+1} \approx (xy)^{n+1}y^n \approx xyy^n \approx x^{n+1}y^{n+1}$, proving (ii). $\square$

Finally, combining Lemma 3.22(i) with Lemma 3.16, we get our desired result.

**Theorem 3.23.** *Let $\mathcal{V}$ be a variety with $\mathbf{N}_4, \mathbf{R}^+, \mathbf{L}^+ \notin \mathcal{V}$. Then $\mathcal{V}$ satisfies $x^{n+1}y \approx xy$ or $xy^{n+1} \approx xy$ for some $n \geqslant 1$.*

To lead us into the next section, we will prove one further result. This result tells us in particular that a subdirectly irreducible semigroup satisfying both identities of Theorem 3.23 is either completely regular or null.

**Lemma 3.24.** *Let $\mathbf{S}$ be a semigroup such that $\mathbf{S} \models x^{n+1}y \approx xy \approx xy^{n+1}$ for some $n \geqslant 1$. Then $\mathbf{S} \models (xy)^{n+1} \approx x^{n+1}y^{n+1} \approx xy$, and $\mathbf{S}$ is a subdirect product of a completely regular semigroup and a null semigroup.*

Proof. We have $\mathbf{S} \models (xy)^{n+1} \approx (xy)^n xy \approx (xy)^n xyy^n \approx (xy)^{n+1}y^n \approx xyy^n \approx xy$, and clearly $\mathbf{S} \models x^{n+1}y^{n+1} \approx xy$.

To obtain the subdirect decomposition, we can assume that $n \geqslant 2$ (since $x^2y \approx xy \approx xy^2$ implies $x^3y \approx xy \approx xy^3$), so that $n$ is an exponent of $\mathbf{S}$. Since $\mathbf{S}$ satisfies $(xy)^{n+1} \approx xy$, we have $SS \subseteq \mathrm{G}(\mathbf{S})$, and the reverse inclusion is obvious, so $\mathrm{G}(\mathbf{S}) = SS$. Thus, $\mathbf{S}/\mathrm{G}(\mathbf{S})$ is a null semigroup, and clearly $\mathrm{G}(\mathbf{S})$ is completely regular.

Since $\mathbf{S}$ satisfies $(xy)^{n+1} \approx x^{n+1}y^{n+1}$, the mapping $x \mapsto x^{n+1}$ is a homomorphism from $\mathbf{S}$ onto $\mathrm{G}(\mathbf{S})$ fixing $\mathrm{G}(\mathbf{S})$ pointwise. This homomorphism and the quotient map $\mathbf{S} \to \mathbf{S}/\mathrm{G}(\mathbf{S})$ together separate the points of $\mathbf{S}$, and so $\mathbf{S}$ is a subdirect product of $\mathrm{G}(\mathbf{S})$ and $\mathbf{S}/\mathrm{G}(\mathbf{S})$. $\square$

## 3.3. Completely regular semigroups

In this section, we will study completely regular semigroups in residually small varieties, with a particular aim of describing the subdirectly irreducible completely regular semigroups in such varieties. In view of Lemma 3.24, this will do most of the work in describing residually small varieties satisfying both $x^{n+1}y \approx xy$ and $xy^{n+1} \approx xy$, but it will also be important in the case where we have only one of these identities.

Just as $\mathrm{G}(\mathbf{S})$ must be a subsemigroup of $\mathbf{S}$ when $\mathbb{V}(\mathbf{S})$ is residually small, $\mathrm{E}(\mathbf{S})$ must also be a subsemigroup for such $\mathbf{S}$, as we will show via the next two results. To this end, we will introduce a family of completely simple semigroups, which can be thought of as minimal failures of a completely regular semigroup to be orthodox.

Let $p$ be a prime, let $\mathbf{G}$ be the (cyclic) group of order $p$, and let $g$ be a generator of $\mathbf{G}$. We define $\mathbf{M}_p$ to be the Rees matrix semigroup $\mathcal{M}[\mathbf{G}, P]$, where

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

Thus, $\mathbf{M}_p$ is a completely simple semigroup of exponent $p$.

We will first show that $\mathbf{M}_p$ generates a residually large variety. For the proof, we will call on a paper of Rasin [**58**], whose objects of study are completely simple semigroups with Abelian subgroups and varieties thereof. The results in question are those in Section 2.2 of [**58**], amounting to about two pages, which show how certain Rees matrix semigroups can be obtained from simpler ones (resembling $\mathbf{M}_p$) via direct products, subsemigroups, and quotients. The proofs are conceptually straightforward, but are cumbersome due to the nature of Rees matrix semigroups, so we will refer the reader to [**58**] for the proofs. On the other hand, it is not so straightforward to extract from the literature a tangible construction of large subdirectly irreducibles in $\mathbb{V}(\mathbf{M}_p)$, so we will detail this part of the argument.

**Theorem 3.25.** *For each prime $p$, the variety $\mathbb{V}(\mathbf{M}_p)$ is residually large.*

Proof. Let $p$ be a prime, and use the symbols $\mathbf{M}_p$, $\mathbf{G}$, $g$ as in the definition of $\mathbf{M}_p$ above. To show that $\mathbb{V}(\mathbf{M}_p)$ is residually large, let $\kappa > 0$ be a cardinal.

Define the $\kappa \times \kappa$ matrix $Q$ over $G$ by setting $q_{ij} := g$ if $i = j \neq 0$ and $q_{ij} := 1$ otherwise, and let $\mathbf{S} := \mathcal{M}[\mathbf{G}, Q]$. Then $\mathbf{S}$ is completely simple with exponent $p$. By [**58**, §2.2], we have $\mathbf{S} \in \mathbb{V}(\mathbf{M}_p)$, and clearly $|S| \geqslant \kappa$.

By Theorem 2.26(ii), $\mathcal{H}$ is a congruence on $\mathbf{S}$. We will show that $\mathbf{S}$ is subdirectly irreducible by showing that $\mathcal{H}$ is the least non-trivial congruence on $\mathbf{S}$. First, we will prove that for every congruence $\theta$ on $\mathbf{S}$, we have

$$\theta \neq \Delta_S \implies \theta \cap \mathcal{H} \neq \Delta_S.$$

Let $\theta$ be a non-trivial congruence on $\mathbf{S}$, so there are $a, b \in S$ with $a \neq b$ and $a \mathrel{\theta} b$. We will show that $\theta$ identifies two distinct $\mathcal{H}$-related elements of $\mathbf{S}$. If $a \mathrel{\mathcal{H}} b$, then we are done, so we can assume that $a \mathrel{\not\mathcal{H}} b$. By symmetry, we can assume further that $a \mathrel{\not\mathcal{R}} b$.

Define $e := (0, 0, 1)$, and let $c := (ae)^p$ and $d := (be)^p$. Then $c \mathrel{\theta} d$ because $a \mathrel{\theta} b$, and by Lemma 2.24, we have $c \in R_a \cap L_e$ and $d \in R_b \cap L_e$; in particular, $c \mathrel{\not\mathcal{R}} d$. Thus, $c$ and $d$ are idempotents in $L_e$ with $c \mathrel{\not\mathcal{R}} d$, so there are distinct $i, j \in \kappa$ such that $c \in \{i\} \times \{0\} \times G$

and $d \in \{j\} \times \{0\} \times G$. From the definition of $Q$, one easily verifies that $(i, 0, 1)$ and $(j, 0, 1)$ are the unique idempotents of $\{i\} \times \{0\} \times G$ and $\{j\} \times \{0\} \times G$, respectively, so we in fact have $c = (i, 0, 1)$ and $d = (j, 0, 1)$. Now, we can assume by symmetry that $j \neq 0$, so, using $c\,\theta\,d$, we have

$$e = (0, 0, 1) = (0, 0, q_{ji}) = (0, j, 1)(i, 0, 1)\ \theta\ (0, j, 1)(j, 0, 1) = (0, 0, q_{jj}) = (0, 0, g).$$

Thus, $e = (0, 0, 1)$ and $(0, 0, g)$ are distinct elements of $H_e$ that are $\theta$-related, as required.

Next, we will prove that for every congruence $\theta$ on $\mathbf{S}$, we have

$$\theta \cap \mathcal{H} \neq \Delta_S \implies \mathcal{H} \subseteq \theta.$$

Let $\theta$ be a congruence on $\mathbf{S}$, and assume there are distinct $a, b \in S$ identified by both $\mathcal{H}$ and $\theta$; we will show that $\mathcal{H} \subseteq \theta$. Note that $\theta{\restriction}_{H_a}$ is a non-trivial congruence on $H_a$, which is a $p$-element cyclic group, so $\theta$ identifies all elements of $H_a$. Now, let $H$ be any $\mathcal{H}$-class of $\mathbf{S}$. Then, as $\mathcal{D} = S^2$ by Theorem 2.26(i), Green's Lemma 2.17 gives a bijection $H_a \to H$ obtained by composing translations of $\mathbf{S}$, so the congruence property implies that all elements of $H$ are identified by $\theta$. Thus, $\mathcal{H} \subseteq \theta$, as required. We have shown that $\mathcal{H}$ is the least non-trivial congruence on $\mathbf{S}$, so $\mathbf{S}$ is subdirectly irreducible. $\square$

**Theorem 3.26.** *Let $\mathbf{S}$ be a non-orthodox completely regular semigroup of finite exponent. Then $\mathbb{V}(\mathbf{S})$ contains $\mathbf{M}_p$ for some prime $p$, and so $\mathbb{V}(\mathbf{S})$ is residually large.*

Proof. By Lemma 2.45 and Theorem 2.44, $\mathbf{S}$ has a non-orthodox completely simple subsemigroup, so without loss of generality, we can assume that $\mathbf{S}$ is completely simple.

Choose idempotents $e, f \in S$ such that $ef$ is not idempotent. By Theorem 2.41, every subsemigroup of $\mathbf{S}$ is completely simple, so we can assume that $\mathbf{S}$ is generated by $\{e, f\}$.

Since $ef$ is not idempotent, we have $ef \notin \{e, f\}$, so we cannot have $e\ \mathcal{R}\ f$ or $e\ \mathcal{L}\ f$ by Proposition 2.29. Thus, by Theorem 2.26(ii), the four $\mathcal{H}$-classes $H_e$, $H_{ef}$, $H_{fe}$, $H_f$ are pairwise distinct, and their union forms a subsemigroup of $\mathbf{S}$. Because $\{e, f\}$ generates $\mathbf{S}$, these must be the only $\mathcal{H}$-classes of $\mathbf{S}$.

| | | |
|---|---|---|
| $e$ | | $ef$ |
| | | |
| $fe$ | | $f$ |

By Theorem 2.28, we may represent $\mathbf{S}$ as $\mathcal{M}[\mathbf{G}, P]$, where $P$ is a $\{0, 1\} \times \{0, 1\}$ matrix over the group $\mathbf{G} = \langle H_e; \cdot \rangle$. Moreover, we may assume that all entries of $P$ equal $e$ except for $a := p_{11}$. Thus, $e$ is represented as $(0, 0, e)$, while $f$ is represented as $(1, 1, a^{-1})$. Now, in any product of $(0, 0, e)$ and $(1, 1, a^{-1})$, the third coordinate is always an integer power of $a$, so because $\{e, f\}$ generates $\mathbf{S}$, it follows that $a$ generates $\mathbf{G}$.

Now, since $\mathbf{G}$ is a non-trivial cyclic group, there is a prime $p$, a cyclic group $\mathbf{H}$ of order $p$, and an onto homomorphism $\varphi\colon \mathbf{G} \to \mathbf{H}$. If we define the $\{0, 1\} \times \{0, 1\}$ matrix $Q$ by setting $q_{\lambda i} := \varphi(p_{\lambda i})$, we get an onto homomorphism $\psi\colon \mathcal{M}[\mathbf{G}, P] \to \mathcal{M}[\mathbf{H}, Q]$, given by $\psi(i, \lambda, x) := (i, \lambda, \varphi(x))$. Since all entries of $Q$ are the identity except for $q_{11} = \varphi(a)$, which generates $\mathbf{H}$, we have that $\mathcal{M}[\mathbf{H}, Q]$ is isomorphic to $\mathbf{M}_p$, so $\mathbf{M}_p \in \mathbb{V}(\mathbf{S})$, as required. $\square$

Thus, a semigroup in a residually small variety must be orthodox. To study the subsemigroup of idempotents, we will now consider bands in particular.

The characterisation of residually small band varieties was completed (somewhat indirectly) by Gerhard [**29**] [**30**], about a decade before Golubov, Sapir, and McKenzie's results were published. As we will see by the end of this section, there is a single variety containing all residually small varieties of bands; its members are called normal bands. A band $\mathbf{S}$ is **normal** if $\mathbf{S} \models xyzt \approx xzyt$. Our immediate goal is to give a characterisation of normal bands that will enable us to show that all bands in a residually small variety are normal.

To state this characterisation, we introduce some important bands. We denote by $\mathbf{L}$ and $\mathbf{R}$ the two-element left-zero and right-zero semigroups, respectively. These are of course normal bands, as are $\mathbf{L}^0$ and $\mathbf{R}^0$, which we will encounter again later. On the other hand, the bands $\mathbf{L}^1$ and $\mathbf{R}^1$ are not normal. The next result characterises normal bands in terms of embedding $\mathbf{L}^1$ and $\mathbf{R}^1$.

Our proof of the following theorem is based on ideas from Howie [**37**]. The important point for us is the equivalence of (i) and (ii) (which is a well-known consequence of [**29**], but we give a more direct proof). We include conditions (iii) and (iv) mainly because the proof of (ii) ⇒ (i) passes through (iii) and (iv). Note that (iii) references the idempotent order from Section 2.4.

**Theorem 3.27.** *Let* $\mathbf{S}$ *be a band. The following are equivalent:*

(i) $\mathbf{S}$ *is normal;*
(ii) $\mathbf{L}^1, \mathbf{R}^1 \notin \mathbb{S}(\mathbf{S})$*;*
(iii) $(\forall e, f, g \in S)\ \ f, g \leqslant e\ \ \&\ \ f \mathrel{\mathcal{J}} g \implies f = g$*;*
(iv) $\mathbf{S} \models xyzx \approx xzyx$.

Proof. (i) ⇒ (ii): Any monoid satisfying $xyzt \approx xzyt$ must be commutative, so $\mathbf{L}^1$ and $\mathbf{R}^1$ are not normal, and therefore any band containing them is not normal.

(ii) ⇒ (iii): To show the contrapositive, assume that (iii) fails; we will show that either $\mathbf{L}^1$ or $\mathbf{R}^1$ embeds into $\mathbf{S}$. Let $c, d, e \in S$ with $c, d$ distinct, $c \mathrel{\mathcal{J}} d$, and $c, d \leqslant e$. Since $e$ is an identity for $c$ and $d$, it is also an identity for the subsemigroup $\mathbf{B}$ generated by $\{c, d\}$. By Theorem 2.47, the subsemigroup on $J_c$ is rectangular band, and $J_c$ contains $B$, so $\mathbf{B}$ is a rectangular band. Since $\mathbf{B}$ is a band, we have $\mathcal{H}^{\mathbf{B}} = \Delta_B$, and $\mathcal{D}^{\mathbf{B}} = B^2$ by Lemma 2.26. Since $\mathbf{B}$ is non-trivial, $\mathbf{B}$ has a non-trivial $\mathcal{R}$ or $\mathcal{L}$ relation, so by Proposition 2.29, $\mathbf{B}$ contains a subsemigroup $\{a, b\}$ isomorphic to $\mathbf{R}$ or $\mathbf{L}$. Since $e$ is an identity for $B$, we have $e \notin \{a, b\}$, so $\mathbf{R}^1$ or $\mathbf{L}^1$ embeds into $\mathbf{S}$ on $\{e, a, b\}$.

(iii) ⇒ (iv): Assume that (iii) holds. To show that $\mathbf{S} \models xyzx \approx xzyx$, let $a, b, c \in S$. By Theorem 2.47, we have $abca \mathrel{\mathcal{J}} acba$ because $\mathbf{S}/\mathcal{J}$ is commutative, and clearly $a$ is an identity for $abca$ and $acba$, so by (iii) we have $abca = acba$. Thus, $\mathbf{S} \models xyzx \approx xzyx$.

(iv) ⇒ (i): Assume that $\mathbf{S} \models xyzx \approx xzyx$. Below is a derivation of $xyzt \approx xzyt$ from $xyzx \approx xzyx$ and $x^2 \approx x$, where the square brackets indicate which two substrings are to be swapped using $xyzx \approx xzyx$.

$$xyzt \approx (xyzt)^2 \approx x[y][zt]xyzt \approx xz[tyx][y]zt \approx xzyt[y][xz]t \approx xzytxzyt \approx xzyt.$$

Hence, $\mathbf{S}$ is normal. □

We will now show that $\mathbf{L}^1$ and $\mathbf{R}^1$ generate residually large varieties. It will then follow immediately from Theorem 3.27 that all bands in a residually small variety are normal.

As is now routine, we will start by finding equational bases for these varieties.

**Proposition 3.28.** *The variety of $\mathbf{L}^1$ is axiomatised by the equations $x^2 \approx x$ and $xyx \approx xy$.*

Proof. It is easily seen that $\mathbf{L}^1 \models x^2 \approx x,\ xyx \approx xy$. Let $u \approx v$ be an identity holding in $\mathbf{L}^1$; we must show that $u \approx v$ can be deduced from $x^2 \approx x$ and $xyx \approx xy$.

First, note that the two-element semilattice $\mathbf{I}$ embeds into $\mathbf{L}^1$, so $\mathbf{I} \models u \approx v$, which implies that $u$ and $v$ must contain exactly the same variables. We will prove the result by induction on the number of variables occurring in $u$ and $v$.

Since $\mathbf{L} \models u \approx v$, the left-most variables of $u$ and $v$ must be the same; let $x$ denote this variable. Using $x^2 \approx x$, we can reduce $u$ to a word with no consecutive occurrences of $x$ (but still beginning with $x$). Now, suppose $x$ occurs elsewhere in the reduced word, so we have a word of the form $xsxt$ for some words $s$, $t$ with $s$ non-empty. Then $xsxt$ can be reduced to $xst$ using $xyx \approx xy$. This process can be repeated to remove other instances of $x$.

By symmetry, we can deduce $u \approx xu'$ and $v \approx xv'$ from $x^2 \approx x$ and $xyx \approx xy$, for some words $u'$, $v'$ not containing $x$. Now, it is clear that $\mathbf{L}^1 \models xu' \approx xv'$; in fact, $\mathbf{L}^1 \models u' \approx v'$, since $x$ does not occur in $u'$, $v'$ and $\mathbf{L}^1$ has an identity element. Now $u'$ and $v'$ contain one fewer distinct variable than $u$, $v$, so inductively we can assume that $u' \approx v'$ follows from the identities $x^2 \approx x$ and $xyx \approx xy$. Thus, $u \approx xu' \approx xv' \approx v$ follows from these identities. $\square$

The proof of the following result is adapted from Gerhard's final example in [**30**].

**Theorem 3.29.** *The semigroups $\mathbf{L}^1$ and $\mathbf{R}^1$ generate residually large varieties.*

Proof. It suffices to prove the result for $\mathbf{L}^1$. Let $\kappa > 1$ be a cardinal. For each $\alpha \in \kappa$, let $\underline{\alpha}$ denote the constant map $\kappa \to \kappa$ with value $\alpha$, and for each $c \in \{0, 1\}$, define

$$c^\alpha_\alpha \colon \kappa \to \kappa, \qquad c^\alpha_\alpha(x) := \begin{cases} \alpha & \text{if } x = \alpha, \\ c & \text{otherwise.} \end{cases}$$

Now, let

$$S := \{\underline{\alpha} \mid \alpha \in \kappa\} \cup \left\{c^\alpha_\alpha \mid c \in \{0,1\},\ \alpha \in \kappa\backslash\{0,1\}\right\}.$$

Let $\alpha, \beta \in \kappa\backslash\{0,1\}$ with $\alpha \neq \beta$, and let $\{c, d\} = \{0, 1\}$. Note that if one of $x, y \in S$ is constant, then $x \circ y$ is constant (hence in $S$), so the following table shows that $S$ is closed under composition. The table also shows that $\mathbf{S} := \langle S; \circ\rangle$ is a band.

| $\circ$ | $c^\alpha_\alpha$ | $c^\beta_\beta$ | $d^\alpha_\alpha$ | $d^\beta_\beta$ |
|---|---|---|---|---|
| $c^\alpha_\alpha$ | $c^\alpha_\alpha$ | $\underline{c}$ | $c^\alpha_\alpha$ | $\underline{c}$ |

Let $x, y \in S$. Then, from the table, either $xy$ is a constant map, in which case $xyx = xy$, or else $xy = x$, in which case $xyx = xx = x = xy$. This shows that $\mathbf{S} \models xyx \approx xy$, and therefore $\mathbf{S} \in \mathbb{V}(\mathbf{L}^1)$ by Proposition 3.28. Clearly, $|S| \geqslant \kappa$.

We will show that $\mathbf{S}$ is subdirectly irreducible by showing that $\underline{0}$ and $\underline{1}$ are identified by every non-trivial congruence on $\mathbf{S}$. Let $\theta \neq \Delta_S$ be a congruence on $\mathbf{S}$. Arguing as in the last paragraph of the proof of Theorem 3.21, we can assume that there are distinct $\alpha, \beta \in \kappa$ with $\underline{\alpha}\ \theta\ \underline{\beta}$ and $\alpha \notin \{0, 1\}$. Thus, $\underline{\alpha} = 0^{\alpha}_{\alpha}\, \underline{\alpha}\ \theta\ 0^{\alpha}_{\alpha}\, \underline{\beta} = \underline{0}$, and similarly $\underline{\alpha}\ \theta\ \underline{1}$, so $\underline{0}\ \theta\ \underline{1}$. □

We conclude that all bands in a residually small variety are normal. Since any semigroup $\mathbf{S}$ in a residually small variety is orthodox, it follows that $\mathrm{E}(\mathbf{S})$ is a normal band. We will characterise the completely regular semigroups with this property in Lemma 3.31. First, we record the following simple but incredibly important result.

**Lemma 3.30.** *Let $\mathbf{S}$ be a semigroup in $[x^{n+1}y \approx xy,\ xy^n z^n t \approx xz^n y^n t]$ for some $n \geqslant 1$. Then, for every $e \in \mathrm{E}(\mathbf{S})$, the left translation $x \mapsto ex$ is an endomorphism of $\mathbf{S}$.*

Proof. Let $e \in \mathrm{E}(\mathbf{S})$, and let $x, y \in S$. Then

$$exy = ex(ex)^n y = exe^n(ex)^n y = ex(ex)^n e^n y = (ex)^{n+1} ey = exey. \qquad \square$$

**Lemma 3.31.** *Let $\mathbf{S}$ be a completely regular semigroup of exponent $n \geqslant 1$. The following are equivalent:*

(i) *$\mathbf{S}$ is orthodox and $\mathrm{E}(\mathbf{S})$ is a normal band;*

(ii) *$\mathbf{S} \models xy^n z^n t \approx xz^n y^n t$;*

(iii) *for each $e \in \mathrm{E}(\mathbf{S})$, the maps $x \mapsto ex$ and $x \mapsto xe$ are endomorphisms of $\mathbf{S}$.*

*Moreover, if $\mathbf{S}$ satisfies (i)–(iii), then $\mathcal{H}$ is a congruence on $\mathbf{S}$.*

Proof. (i) ⇒ (ii): Clearly $\mathbf{S} \models x^{n+1} \approx x$, and by (i), we have $\mathbf{S} \models x^n y^n z^n t^n \approx x^n z^n y^n t^n$. Thus, $\mathbf{S} \models xy^n z^n t \approx xx^n y^n z^n t^n t \approx xx^n z^n y^n t^n t \approx xz^n y^n t$.

(ii) ⇒ (iii): Since $\mathbf{S} \models x^{n+1} \approx x$, this is true by Lemma 3.30 and its dual.

(iii) ⇒ (i): Let $e, f, g, h \in S$. Then, since $x \mapsto fx$ and $x \mapsto xf$ are endomorphisms, we have $efgh = efgfh = egfh$, and this implies $efef = eeff = ef$. This shows that $\mathbf{S}$ is orthodox and $\mathrm{E}(\mathbf{S})$ is a normal band.

We have shown that (i)–(iii) are equivalent. Now assume that $\mathbf{S}$ satisfies (i)–(iii); we will show that $\mathcal{H}$ is a congruence on $\mathbf{S}$.

Let $a, b \in S$, and let $e := a^n \in H_a \subseteq J_a$ and $f := b^n \in H_b \subseteq J_b$. Then by Theorem 2.44, we have that $ab\ \mathcal{J}\ ef$ and that $ab$ and $ef$ lie in some completely simple subsemigroup $\mathbf{J}$ of $\mathbf{S}$. Now, since $x \mapsto fx$ and $x \mapsto xf$ are endomorphisms of $\mathbf{S}$ by (iii), we have

$$efab = efafb = eafb = a^n ab^n b = ab,$$

so $ab \in efS \subseteq R^{\mathbf{J}}_{ef}$ by Lemma 2.24, which implies $ab\ \mathcal{R}\ ef$ in $\mathbf{S}$. We then have $ab\ \mathcal{H}\ ef$ by symmetry. Now $ef \in H_{ab}$ is idempotent by (i), so $(ab)^n = ef = a^n b^n$. The mapping $x \mapsto x^n$ therefore defines an endomorphism of $\mathbf{S}$ whose kernel is $\mathcal{H}$; thus, $\mathcal{H}$ is a congruence on $\mathbf{S}$. □

**Remark.** Lemma 3.31 holds without the assumption of finite exponent, provided one interprets $x^n$ as $xx^{-1}$. However, (ii) is not expressible in the semigroup signature when $\mathbf{S}$ does not have finite exponent. The result itself was rather painstakingly extracted from McKenzie's proofs; it does not seem to appear in any convenient place in the English literature.

Now, we would like to characterise the subdirectly irreducible completely regular semigroups satisfying the conditions of Lemma 3.31. We will prove this using two more preliminary lemmas.

**Lemma 3.32.** *Let* $\mathbf{S}$ *be a semigroup without zero. Then* $\mathbf{S}$ *is subdirectly irreducible if and only if* $\mathbf{S}^0$ *is subdirectly irreducible.*

Proof. Assume that $\mathbf{S}$ is subdirectly irreducible, let $(a, b)$ generate the monolith of $\mathbf{S}$, and let $\theta$ be a non-trivial congruence on $\mathbf{S}^0$; we will show that $a \mathrel{\theta} b$. For this, it suffices to show that $\theta{\restriction_S}$ is a non-trivial congruence on $\mathbf{S}$. Choose distinct $x, y \in S^0$ with $x \mathrel{\theta} y$. If $x, y \in S$, then we are done, so we can assume that $y = 0$. Since $x$ is not a zero element of $\mathbf{S}$, we can assume by duality that $cx \neq x$ for some $c \in S$, which gives $x \mathrel{\theta} 0 \Rightarrow cx \mathrel{\theta} 0 \Rightarrow x \mathrel{\theta} cx$, so $\theta{\restriction_S}$ is non-trivial, as required. It follows that $\mathbf{S}^0$ is subdirectly irreducible.

Conversely, assume that $\mathbf{S}^0$ is subdirectly irreducible, and let $(a, b)$ generate the monolith of $\mathbf{S}^0$. Since $S^2 \cup \{(0, 0)\}$ is a non-trivial congruence on $\mathbf{S}^0$, we have $a, b \in S$. Now, if $\theta \neq \Delta_S$ is a congruence on $\mathbf{S}$, then $\theta \cup \{(0, 0)\}$ is a congruence on $\mathbf{S}^0$, so $a \mathrel{\theta} b$. □

In the following result, the hypothesis on $a$, $b$ is satisfied if $\mathbf{S}$ is subdirectly irreducible and $(a, b)$ generates the monolith of $\mathbf{S}$. For later use, we have expressed it more generally.

**Lemma 3.33.** *Let* $\mathbf{S} \in [x^{n+1}y \approx xy,\ xy^nz^nt \approx xz^ny^nt]$ *for some* $n \geqslant 1$. *Suppose there are distinct* $a, b \in S$ *such that* $\varphi(a) = \varphi(b)$ *for every endomorphism* $\varphi \colon \mathbf{S} \to \mathbf{S}$ *that is not one-to-one. If* $e \in \mathrm{E}(\mathbf{S})$ *and* $ea \neq eb$, *then* $e$ *is a left identity of* $\mathbf{S}$.

Proof. Let $e \in \mathrm{E}(\mathbf{S})$ with $ea \neq eb$, and let $\lambda \colon S \to S$ denote the left translation $x \mapsto ex$. By Lemma 3.30, $\lambda$ is an endomorphism of $\mathbf{S}$, which must be one-to-one since $\lambda(a) \neq \lambda(b)$. As $\lambda$ is also an idempotent map (i.e., $\lambda \circ \lambda = \lambda$), it must be the identity map. □

We now come to the main result of this section.

**Theorem 3.34.** *Let* $\mathbf{S} \in [xy^nz^nt \approx xz^ny^nt]$ *be completely regular with exponent* $n \geqslant 1$. *Then* $\mathbf{S}$ *is subdirectly irreducible if and only if* $\mathbf{S} \in \mathbb{I}(\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0, \mathbf{I})$ *or* $\mathbf{S} \in \mathbb{I}(\mathbf{G}, \mathbf{G}^0)$ *for some subdirectly irreducible group* $\mathbf{G}$.

Proof. The right-to-left implication follows immediately from Lemma 3.32. Assume that $\mathbf{S}$ is subdirectly irreducible. If $\mathcal{J}$ is trivial, then $\mathbf{S}$ is a semilattice by Theorem 2.44, so $\mathbf{S} \cong \mathbf{I}$ by Theorem 1.12. Thus, we can assume that $\mathcal{J}$ is non-trivial. Let $(a, b)$ generate the monolith of $\mathbf{S}$. By Theorem 2.44, $\mathcal{J}$ is a congruence on $\mathbf{S}$, so $a \mathrel{\mathcal{J}} b$. Let $\mathbf{J}$ be the subsemigroup of $\mathbf{S}$ on $J_a = J_b$, so $\mathbf{J}$ is a non-trivial completely simple semigroup by Theorem 2.44.

Let $e$ be the idempotent element of $H_a$. Then $ea = a$, so if $eb \neq a$, then $e$ is a left identity of $\mathbf{S}$ by Lemma 3.33. Dually, if $be \neq a$, then $e$ is a right identity of $\mathbf{S}$.

Suppose $\mathcal{H}$ is trivial. Then $e = a$, and either $a \mathrel{\not\mathcal{R}^{\mathbf{J}}} b$ or $a \mathrel{\not\mathcal{L}^{\mathbf{J}}} b$, so by Proposition 2.29 we have either $ba \neq a$ or $ab \neq a$. Thus, $e$ is either a left or a right identity of $\mathbf{S}$ in this case.

Now assume that $\mathcal{H}$ is non-trivial. By Lemma 3.31, $\mathcal{H}$ is a congruence on $\mathbf{S}$, so $a \mathrel{\mathcal{H}} b$, and therefore $eb = be = b \neq a$. Thus, $e$ is a two-sided identity for $\mathbf{S}$ in this case.

In both cases, $e$ is at least a one-sided identity of $\mathbf{S}$, so we have $\langle e \rangle = S$. Since $a \mathrel{\mathcal{H}} e$, we also have $\langle a \rangle = \langle e \rangle = S$. Now, if $K$ is a non-trivial ideal of $\mathbf{S}$, then $\Delta_S \cup K^2$ is a non-trivial congruence, so $K$ must contain $a$ (and $b$) and therefore $\langle a \rangle$; that is, $K = S$. Thus, $\mathbf{S}$ is either simple or zero-simple. Since a zero-simple semigroup has exactly two $\mathcal{J}$-classes, we have either $\mathbf{S} = \mathbf{J}$ or $\mathbf{S} \cong \mathbf{J}^0$. Now, $\mathbf{J}$ is simple and non-trivial and therefore cannot have a zero, so if $\mathbf{S} \cong \mathbf{J}^0$, then $\mathbf{J}$ is subdirectly irreducible by Lemma 3.32.

In either case $\mathbf{S} = \mathbf{J}$ or $\mathbf{S} \cong \mathbf{J}^0$, we have that $\mathbf{J}$ is a subdirectly irreducible completely simple semigroup. Now, by Lemma 3.31, $\mathbf{J}$ is orthodox, so by Theorem 2.30, we can assume by symmetry that $\mathbf{J}$ is a group or a right-zero semigroup. In the former case, we are done. For the latter case, note that every equivalence relation on a right-zero semigroup is a congruence, so if $\mathbf{J}$ is right-zero, then $\mathbf{J} \cong \mathbf{R}$. □

We close this section with two important corollaries, describing the subdirectly irreducible normal bands and characterising the residually small varieties of bands.

**Theorem 3.35.** *The class of subdirectly irreducible normal bands is* $\mathbb{I}(\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0, \mathbf{I})$. *Consequently, the variety of normal bands equals* $\mathbb{SP}(\mathbf{L}^0, \mathbf{R}^0)$ *and* $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I})$.

Proof. It is clear that $\mathbf{L}$, $\mathbf{R}$, and $\mathbf{I}$ are normal bands. Any variety containing $\mathbf{I}$ is closed under the construction $\mathbf{S} \mapsto \mathbf{S}^0$ (since $\mathbf{S}^0 \in \mathbb{H}(\mathbf{S} \times \mathbf{I})$), so $\mathbf{L}^0$ and $\mathbf{R}^0$ are also normal bands. The result now follows from Theorem 3.34 with $n = 1$ (and Birkhoff's Theorem 1.6). □

**Corollary 3.36.** *Let $\mathcal{V}$ be a variety of bands. The following are equivalent:*

(i) $\mathcal{V}$ *is residually small;*
(ii) $\mathcal{V}$ *is residually very finite;*
(iii) $\mathbf{si}(\mathcal{V}) \subseteq \mathbb{I}(\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0, \mathbf{I})$;
(iv) $\mathcal{V}$ *is a variety of normal bands;*
(v) $\mathbf{L}^1, \mathbf{R}^1 \notin \mathcal{V}$.

Proof. In the sequence (i) ⇒ (v) ⇒ (iv) ⇒ (iii) ⇒ (ii) ⇒ (i), the first three implications follow from Theorems 3.29, 3.27, and 3.35, respectively, while the last two are trivial. □

## 3.4. Varieties satisfying both identities

We now return to the main thread of the chapter. In Theorem 3.23, we saw that a residually small variety satisfies $x^{n+1}y \approx xy$ or $xy^{n+1} \approx xy$ for some $n \geqslant 1$. In this section, we will deal with the case where both identities are satisfied. Given the results of Section 3.3, there is little work left to do in the general case; all we need is some notation.

For each $n \geqslant 1$, we denote by $\mathcal{G}_n$ the (semigroup) variety of all groups of exponent $n$, and we denote by $\mathcal{N}_n$ the variety $[x^{n+1}y \approx xy \approx xy^{n+1},\ xy^n z^n t \approx xz^n y^n t]$. As the next result shows, $\mathcal{N}_n$ is equal to the join of the variety of groups of exponent $n$ with the varieties of **n**ormal bands and **n**ull semigroups (hence the letter 'N').

**Theorem 3.37.** *For each $n \geqslant 1$, we have*

$$\mathcal{N}_n = \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N}) \vee \mathcal{G}_n, \qquad \mathbf{si}(\mathcal{N}_n) = \mathbb{I}(\{\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}, \mathbf{G}, \mathbf{G}^0 \mid \mathbf{G} \in \mathbf{si}(\mathcal{G}_n)\}).$$

Proof. Let $\mathcal{S}_n := \mathbb{I}(\{\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}, \mathbf{G}, \mathbf{G}^0 \mid \mathbf{G} \in \mathbf{si}(\mathcal{G}_n)\})$. It is clear that $\mathcal{S}_n \subseteq \mathcal{N}_n$, and we have $\mathcal{S}_n \subseteq \mathbf{si}(\mathcal{N}_n)$ by Lemma 3.32. Now, by Lemma 3.24, each $\mathbf{S} \in \mathbf{si}(\mathcal{N}_n)$ is either null or completely regular, and since $\mathcal{N}_n \models x^{n+2} \approx x^2$, every completely regular semigroup in $\mathcal{N}_n$ has exponent $n$. Thus, by Proposition 3.1 and Theorem 3.34, we have $\mathbf{si}(\mathcal{N}_n) = \mathcal{S}_n$.

Now, clearly $\mathcal{N}_n \supseteq \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N}) \vee \mathcal{G}_n$, and since $\mathbf{S}^0 \in \mathbb{H}(\mathbf{S} \times \mathbf{I})$ for any semigroup $\mathbf{S}$, we have $\mathcal{S}_n \subseteq \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N}) \vee \mathcal{G}_n$. Hence, $\mathcal{N}_n = \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N}) \vee \mathcal{G}_n$. □

**Lemma 3.38.** *Let $\mathcal{V}$ be a variety satisfying $x^{n+1}y \approx xy \approx xy^{n+1}$ for some $n \geqslant 1$, and assume that $\mathbf{L}^1, \mathbf{R}^1 \notin \mathcal{V}$ and $\mathbf{M}_p \notin \mathcal{V}$ for all primes $p$. Then $\mathcal{V} \subseteq \mathcal{N}_n$.*

Proof. Let $\mathbf{S} \in \mathcal{V}$. Since $\mathbf{N}_4 \not\models x^{n+1}y \approx xy$, it follows from Theorems 3.14 and 3.26 that $\mathrm{G}(\mathbf{S})$ is an orthodox subsemigroup of $\mathbf{S}$, and therefore $\mathbf{S}$ is also orthodox. By Corollary 3.36, $\mathrm{E}(\mathbf{S})$ is a normal band, so $\mathbf{S} \models x^m y^m z^m t^m \approx x^m z^m y^m t^m$, where $m := \max\{2, n\}$, which is an exponent of $\mathbf{S}$. Since $m \leqslant n+1$, we deduce

$$\mathbf{S} \models xy^m z^m t \approx x^{n+1} y^m z^m t^{n+1} \approx x^{n+1} z^m y^m t^{n+1} \approx xz^m y^m t.$$

If $n \neq m$ (i.e., $n = 1$), then $xyzt \approx xzyt$ follows from $x^2 y \approx xy$ and $xy^2 z^2 t \approx xz^2 y^2 t$. □

**Theorem 3.39.** *Let $\mathcal{V}$ be a variety satisfying $x^{n+1}y \approx xy \approx xy^{n+1}$ for some $n \geqslant 1$. Then the following are equivalent:*

(i) *$\mathcal{V}$ is residually small;*
(ii) *$\mathbf{L}^1, \mathbf{R}^1 \notin \mathcal{V}$ and $\mathbf{M}_p \notin \mathcal{V}$ for all primes $p$, and $\mathcal{V} \cap \mathcal{G}_n$ is residually small;*
(iii) *$\mathcal{V} \subseteq \mathcal{N}_n$ and $\mathcal{V} \cap \mathcal{G}_n$ is residually small.*

Proof. Theorems 3.25 and 3.29 give (i) ⇒ (ii), while Lemma 3.38 gives (ii) ⇒ (iii) and Theorem 3.37 gives (iii) ⇒ (i). □

As we remarked in the chapter introduction, we do not know exactly when a subvariety of $\mathcal{G}_n$ is residually small, so Theorem 3.39 is incomplete in this regard. For the remainder of this section, we will strengthen Theorem 3.39 with the added assumption that $\mathcal{V}$ is generated by a finite semigroup. For this, we will call on Ol'šanskiĭ's results in [**50**], where the residually *finite* varieties of groups are characterised. Naturally, to state these results, we require some group-theoretic notions.

The following common notions may be found in Macdonald [**45**], for example. We define the **centre** of a group $\mathbf{G}$ to be the normal subgroup $Z(\mathbf{G}) := \{g \in G \mid (\forall x \in G)\ xg = gx\}$. Now, define $\mathbf{G}_n$ inductively for all $n \geqslant 0$ by $\mathbf{G}_0 := \mathbf{G}$ and $\mathbf{G}_{n+1} := \mathbf{G}_n / Z(\mathbf{G}_n)$. If $n \geqslant 0$ and $\mathbf{G}_n$ is trivial, then $\mathbf{G}$ is called **$n$-nilpotent**, and we say that $\mathbf{G}$ is **nilpotent** if $\mathbf{G}$ is $n$-nilpotent for some $n \geqslant 0$. Thus, a group is nilpotent if iteratively factoring by the centre results in the trivial group after a finite number of steps. Any finite nilpotent group is isomorphic to the direct product of its Sylow subgroups [**45**, Theorem 9.08].

The next concept is not quite as common. Let $\mathbf{G}$ be a group. Then $\mathbf{G}$ is called an **A-group** if $\mathbf{G}$ is finite and every nilpotent subgroup of $\mathbf{G}$ is Abelian; equivalently, if $\mathbf{G}$ is finite and every Sylow subgroup of $\mathbf{G}$ is Abelian. Ol'šanskiĭ's main result in [**50**] can now be stated as follows: a variety $\mathcal{V}$ of groups (in the group signature) is residually finite if and only if $\mathcal{V} = \mathbb{V}(\mathbf{G})$ for some (finite) A-group $\mathbf{G}$.

Although there is no complete characterisation of the residually small group varieties, we can say when a finite group generates a residually small variety. As discussed in the chapter introduction, Freese and McKenzie showed in [**27**] that the RS conjecture holds for finite groups. In other words, a finite group that generates a residually small variety in fact generates a residually very finite variety. Combining this with Ol'šanskiĭ's results [**50**] gives the following result (which holds in both the group and the semigroup signature).

**Theorem 3.40.** *Let* $\mathbf{G}$ *be a finite group. The following are equivalent:*

(i) $\mathbb{V}(\mathbf{G})$ *is residually small;*
(ii) $\mathbb{V}(\mathbf{G})$ *is residually very finite;*
(iii) $\mathbf{G}$ *is an A-group;*
(iv) $\mathbb{V}(\mathbf{G}) = \mathbb{SP}(\mathbf{H})$ *for some (finite) A-group* $\mathbf{H}$.

Proof. Freese and McKenzie [**27**, Theorem 8] gives (i) $\Leftrightarrow$ (ii), while Ol'šanskiĭ [**50**, Theorem 1, Theorem 2, Lemma 1] gives (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv). $\square$

To apply Theorem 3.39 to the variety of a finite semigroup $\mathbf{S}$, we will first prove the well-known result that the subgroups of $\mathbf{S}$ generate the variety of groups lying in $\mathbb{V}(\mathbf{S})$.

**Lemma 3.41.** *Let* $\mathbf{S}$ *be a finite semigroup, and let* $\mathbf{G} \in \mathbb{H}(\mathbf{S})$ *be a group. Then* $\mathbf{G} \in \mathbb{H}(\mathbf{H})$ *for some subgroup* $\mathbf{H}$ *of* $\mathbf{S}$.

Proof. Fix a surjective homomorphism $\varphi\colon \mathbf{S} \to \mathbf{G}$. By Theorem 2.4, $\mathbf{S}$ has a minimum ideal $\mathbf{M}$, which is a completely simple subsemigroup of $\mathbf{S}$ by Theorem 2.40. Now $\varphi(M)$ is an ideal of $\mathbf{G}$ as $\varphi$ is surjective, so $\varphi(M) = G$ because $\mathbf{G}$ is simple as a semigroup.

By Theorem 2.36, we may choose some $e \in \mathrm{E}(\mathbf{M})$. Then $\varphi(e) \in \mathrm{E}(\mathbf{G})$, so $\varphi(e) = 1$. Now, let $H := H_e^{\mathbf{M}}$. Then, by Lemma 2.24, we have $eMe \subseteq H$, and clearly $H = eHe \subseteq eMe$, so $eMe = H$ is a subgroup of $\mathbf{S}$, and $\varphi(eMe) = \varphi(e)\varphi(M)\varphi(e) = G$. $\square$

**Theorem 3.42.** *Let* $\mathbf{S}$ *be a finite semigroup, and let* $\mathbf{G}$ *denote the direct product of the subgroups of* $\mathbf{S}$*. Then* $\mathbb{V}(\mathbf{G})$ *is precisely the class of all groups in* $\mathbb{V}(\mathbf{S})$.

Proof. Let $\mathbf{H} \in \mathbb{V}(\mathbf{S})$ be a finite group. By Lemma 1.11, we have $\mathbf{H} \in \mathbb{H}(\mathbf{T})$ for some finite semigroup $\mathbf{T} \in \mathbb{SP}(\mathbf{S})$, and by Lemma 3.41, we have $\mathbf{H} \in \mathbb{H}(\mathbf{K})$ for some subgroup $\mathbf{K}$ of $\mathbf{T}$. Now $\mathbf{K} \in \mathbb{SP}(\mathbf{S})$, so the homomorphisms $\mathbf{K} \to \mathbf{S}$ separate the points of $\mathbf{K}$. But the image of $\mathbf{K}$ under a homomorphism $\mathbf{K} \to \mathbf{S}$ is isomorphic to a subgroup of $\mathbf{S}$ and therefore of $\mathbf{G}$, so the homomorphisms $\mathbf{K} \to \mathbf{G}$ also separate the points of $\mathbf{K}$; that is, $\mathbf{K} \in \mathbb{SP}(\mathbf{G})$. This shows that $\mathbf{H} \in \mathbb{V}(\mathbf{G})$. Thus, $\mathbb{V}(\mathbf{G})$ contains every finite group in $\mathbb{V}(\mathbf{S})$.

Now let $\mathbf{A}$ be any group in $\mathbb{V}(\mathbf{S})$. Then, by Theorems 1.9 and 1.10, $\mathbf{A}$ is in the variety generated by its finite subgroups, which we have shown are in $\mathbb{V}(\mathbf{G})$. Thus, $\mathbf{A} \in \mathbb{V}(\mathbf{G})$. $\square$

We close this section with the following specialisation of Theorem 3.39 to finitely generated varieties.

**Theorem 3.43.** *Let* $\mathbf{S}$ *be a finite semigroup satisfying* $x^{n+1}y \approx xy \approx xy^{n+1}$ *for some* $n \geqslant 1$*. Then the following are equivalent:*

(i) $\mathbb{V}(\mathbf{S})$ *is residually small;*
(ii) $\mathbb{V}(\mathbf{S})$ *is residually very finite;*
(iii) *every finite group in* $\mathbb{V}(\mathbf{S})$ *is an A-group,* $\mathbf{L}^1, \mathbf{R}^1 \notin \mathbb{V}(\mathbf{S})$*, and* $\mathbf{M}_p \notin \mathbb{V}(\mathbf{S})$ *for all primes* $p$*;*
(iv) $\mathbf{S} \in \mathcal{N}_n$ *and every subgroup of* $\mathbf{S}$ *is an A-group;*

Proof. Note that a finite direct product of A-groups is an A-group, so by Theorems 3.40 and 3.42, every subgroup of $\mathbf{S}$ is an A-group if and only if every finite group in $\mathbb{V}(\mathbf{S})$ is an A-group if and only if $\mathbb{V}(\mathbf{S}) \cap \mathcal{G}_n$ is residually small. Theorem 3.39 now gives the result. $\square$

## 3.5. Groups in the presence of P

In Section 3.4, we considered residually small varieties satisfying both $x^{n+1}y \approx xy$ and $xy^{n+1} \approx xy$, and saw that there is no restriction on their subvarieties of groups, as long as they are residually small. It is perhaps surprising, then, that if a residually small variety $\mathcal{V}$ satisfies only one of these two identities, there is a significant restriction on the groups in $\mathcal{V}$: they must all be Abelian. The broader goal of this section is to prove this fact.

Of course, if a variety satisfies precisely one of the identities $x^{n+1}y \approx xy$, $xy^{n+1} \approx xy$, we can assume by symmetry that the first identity holds but the second does not. The remainder of this chapter will predominantly be concerned with such varieties.

Lemma 3.8 showed us that **P** manifests somewhat mysteriously in varieties under certain conditions. Lemma 3.45 below shows that these conditions are met in a residually small variety satisfying $x^{n+1}y \approx xy$ but not $xy^{n+1} \approx xy$. First, we prove a simple result concerning group elements.

**Lemma 3.44.** *Let* $\mathbf{S}$ *be a semigroup satisfying* $x^{n+1}y \approx xy$ *for some* $n \geqslant 2$. *Then* $\mathrm{G}(\mathbf{S})$ *is a left ideal of* $\mathbf{S}$.

PROOF. If $a \in S$ and $b \in \mathrm{G}(\mathbf{S})$, then $(ab)^{n+1} = (ab)^n ab = (ab)^n abb^n = abb^n = ab$. $\square$

The following result was proved by Golubov and Sapir [**32**, Lemma 6], but we give a new proof using Lemma 3.8. In fact, Lemma 3.8 was discovered while refining the next proof.

**Lemma 3.45.** *Let* $\mathcal{V}$ *be a semigroup variety satisfying* $x^{n+1}y \approx xy$ *but not* $xy^{n+1} \approx xy$ *for some* $n \geqslant 1$, *and assume that* $\mathbf{R}^+ \notin \mathcal{V}$. *Then* $\mathbf{P} \in \mathcal{V}$.

PROOF. Suppose $n = 1$. Then $\mathcal{V} \models x^3 \approx x^2$, so $\mathcal{V} \models x^3y \approx xy$; however $\mathcal{V} \models xy^3 \approx xy$ would imply $\mathcal{V} \models xy^2 \approx xy$, so $\mathcal{V} \not\models xy^3 \approx xy$. Thus, we can assume that $n \geqslant 2$.

Now, note that $\mathbf{N}_4 \not\models x^{n+1}y \approx xy$, so $\mathbf{N}_4 \notin \mathcal{V}$. Let $\mathbf{F} := \mathbf{F}_{\mathcal{V}}(\{x, y\})$, and consider the elements $e := x^n$ and $u := x^n y = ey$ of $\mathbf{F}$. Then $e \in \mathrm{E}(\mathbf{F})$ and $eu = u$. By Lemma 3.8, it suffices to show that $u \notin \mathrm{G}(\mathbf{F})$.

Suppose $u \in \mathrm{G}(\mathbf{F})$. Then $xy = x^{n+1}y = xu \in \mathrm{G}(\mathbf{F})$ by Lemma 3.44, so $(xy)^{n+1} = xy$, and therefore $\mathcal{V} \models (xy)^{n+1} \approx xy$ by Lemma 1.8. But we then have $\mathcal{V} \models xy \approx xy^{n+1}$ by Lemma 3.22(i), which is a contradiction. $\square$

Lemma 3.45 tells us that **P** is always present in the varieties we are interested in; so, effectively, our task for the remainder of this chapter is to characterise the residually small varieties containing **P**. In this section, we will focus on groups in varieties containing **P**. We will start by introducing a construction that will be of fundamental importance for the remainder of the chapter, as well as the second half of the thesis.

To introduce this construction, it would perhaps be prudent to clarify the notion of composition of partial maps. Given partial maps $f\colon U \to V$ and $g\colon V \to W$, we define the composite map $g \circ f\colon U \to W$ by setting $\mathrm{dom}(g \circ f) := \{x \in U \mid g(f(x)) \text{ is defined}\}$ and defining $(g \circ f)(x) := g(f(x))$ for all $x \in U$.

Now, let **T** be a semigroup, and let $U$ be a non-empty set. An **action** of **T** on $U$ is a homomorphism $\Phi$ from **T** to the semigroup of *partial* maps $U \to U$. If $\Phi(a)\colon U \to U$ is total for all $a \in T$, then $\Phi$ is called **total**. If $\Phi$ is one-to-one, then $\Phi$ is called **faithful**.

If the action of $\mathbf{T}$ on $U$ is understood, we will typically write $au$ to mean $\Phi(a)(u)$, for all $a \in T$ and all $u \in U$. We will often refer to a semigroup $\mathbf{T}$ acting on a set $U$ without specifying the action; it will be assumed that there is some fixed action involved. The action will either be clear from context, or the exact nature will not be important. Additionally, we will always assume that semigroups act on non-empty sets.

Let $\Phi$ be an action of a semigroup $\mathbf{T}$ on a set $U$. We define the semigroup $\mathbf{T} \circlearrowleft U$ to have underlying set $T \mathbin{\dot\cup} U \mathbin{\dot\cup} \{0\}$, with multiplication defined as follows:

- if $a, b \in T$, we take $ab$ to be the product in $\mathbf{T}$;
- if $a \in T$ and $u \in U$, we define $au$ to be $\Phi(a)(u)$ if the latter is defined;
- all other products are defined to be 0.

Of course, the multiplication of $\mathbf{T} \circlearrowleft U$ depends on $\Phi$, but again, the action will usually be understood. It is straightforward to verify that $\mathbf{T} \circlearrowleft U$ is a semigroup.

The construction $\mathbf{T} \circlearrowleft U$ was used by McKenzie in [**47**]. One of the strong points of McKenzie's approach is his use of this construction, as it makes certain aspects of the main proof more lucid. For total actions, we can roughly summarise the definition of $\mathbf{T} \circlearrowleft U$ in the following table:

| $\cdot$ | $T$ | $U$ | $0$ |
|---|---|---|---|
| $T$ | $T$ | $U$ | $0$ |
| $U$ | $0$ | $0$ | $0$ |
| $0$ | $0$ | $0$ | $0$ |

We now begin to see the connection between $\mathbf{P}$ and semigroup actions: the above table is a copy of the table for $\mathbf{P}$ (Figure 3.2). This gives us a surjective homomorphism $\mathbf{T} \circlearrowleft U \to \mathbf{P}$. In fact, $\mathbf{P}$ itself can be represented by the action of the trivial semigroup, $\mathbf{1}$, on a one-element set; that is, $\mathbf{P} \cong \mathbf{1} \circlearrowleft \{u\}$. This hints that we have more than just a neat construction, and explains to some extent what the semigroup $\mathbf{P}$ really is.

In this section, we will restrict our attention to group actions, but we will encounter more general semigroup actions in the next section. Let $\Phi$ be an action of a group $\mathbf{G}$ on a set $U$. We say that $\Phi$ is a **group action** if $\Phi(g)$ is a total permutation of $U$ for every $g \in G$. If $\Phi$ is an action of a group $\mathbf{G}$ on a set $U$ (not necessarily total), then a sufficient condition for $\Phi$ to be a group action is that $\Phi(1)$ is the (total) identity map on $U$.

By means of the next two results, we will show that if $\mathbf{G}$ is a group in some variety $\boldsymbol{\mathcal{V}}$ containing $\mathbf{P}$, then $\boldsymbol{\mathcal{V}}$ contains *every* semigroup $\mathbf{G} \circlearrowleft U$, where $U$ is a set and $\mathbf{G}$ acts on $U$ (via a group action). Since we want to know when a variety containing $\mathbf{P}$ is residually small, this naturally leads to the question of when $\mathbf{G} \circlearrowleft U$ is subdirectly irreducible, as well as the question of when there is a cardinal bound on these subdirectly irreducible semigroups.

Let us recall some well-known properties of group actions. Fix some group action of $\mathbf{G}$ on a set $U$. The **orbit** of $x \in U$ is the set $\{gx \mid g \in G\}$; the set of orbits is a partition of $U$. A group action is **transitive** if it has precisely one orbit.

If $X$ is an orbit, then the action of each $g \in G$ restricts to a permutation of $X$, and this induces a transitive action of $\mathbf{G}$ on $X$. This idea gives us the following subdirect decomposition of $\mathbf{G} \circlearrowleft U$.

**Proposition 3.46.** *Let* $\mathbf{G}$ *be a group acting on a set* $U$ *via a group action, and let* $\mathcal{X}$ *be the set of orbits. Then* $\mathbf{G} \circlearrowright U$ *is a subdirect product of the semigroups* $\{\mathbf{G} \circlearrowright X \mid X \in \mathcal{X}\}$*.*

PROOF. Let $\mathbf{S} := \mathbf{G} \circlearrowright U$. For each $X \in \mathcal{X}$, the set $J_X := (U \backslash X) \cup \{0\}$ is an ideal of $\mathbf{S}$, and $\theta_X := \Delta_S \cup J_X^2$ is a congruence on $\mathbf{S}$ with $\mathbf{S}/\theta_X \cong \mathbf{G} \circlearrowright X$. Since $\bigcap\{\theta_X \mid X \in \mathcal{X}\} = \Delta_S$, the result follows. □

Proposition 3.46 tells us in particular that if $\mathbf{G} \circlearrowright U$ is subdirectly irreducible, the action must be transitive. This enables us to call on the well-known representation theorem for group actions, usually called the Orbit-Stabiliser Theorem or the Stabiliser-Orbit Theorem.

Recall that if $H$ is a subgroup of a group $\mathbf{G}$, then there is a natural group action of $\mathbf{G}$ on the set $G/H := \{aH \mid a \in G\}$ of left cosets of $H$ in $\mathbf{G}$, which is given by left multiplication. This action is transitive. The Orbit-Stabiliser Theorem says in effect that every transitive group action can be represented in this way.

For each $x \in U$, the stabiliser of $x$ is defined as the subgroup $G_x := \{g \in G \mid gx = x\}$ of $\mathbf{G}$. The precise content of the Orbit-Stabiliser Theorem is that, for a fixed $x \in U$, the mapping $gG_x \mapsto gx$ is a well-defined bijection $G/G_x \to X$, where $X$ is the orbit of $x$. It is easily seen that this bijection induces an isomorphism $\mathbf{G} \circlearrowright G/G_x \to \mathbf{G} \circlearrowright X$, and so $\mathbf{G} \circlearrowright X$ is representable via a left coset action. By combining Proposition 3.46 with the next result, it will follow immediately that $\mathbf{G} \circlearrowright U \in \mathbb{V}(\mathbf{G}, \mathbf{P})$.

**Proposition 3.47.** *Let* $\mathbf{G}$ *be a group and* $H$ *a subgroup of* $\mathbf{G}$*. Then* $\mathbf{G} \circlearrowright G/H \in \mathbb{V}(\mathbf{G}, \mathbf{P})$*.*

PROOF. For simplicity, assume that the union $G \cup G/H \cup \{0\}$ is disjoint. Let $\mathbf{S} := \mathbf{G} \circlearrowright G/H$, and define the map $\varphi \colon G \times P \to S$ by

$$\varphi(g, x) := \begin{cases} g & \text{if } x = e; \\ gH & \text{if } x = u; \\ 0 & \text{if } x = 0. \end{cases}$$

for all $(g, x) \in G \times P$. Then $\varphi$ is easily seen to be a surjective homomorphism $\mathbf{G} \times \mathbf{P} \to \mathbf{S}$. □

In view of these results, we would like to know when a semigroup $\mathbf{G} \circlearrowright U$ arising from a transitive group action is subdirectly irreducible. We can rather easily show that the action must also be faithful. Consider the following equivalence relation $\gamma$ on $\mathbf{G} \circlearrowright U$ identifying pairs of elements of $\mathbf{G}$ that act as the same map on $U$:

$$x \mathrel{\gamma} y \iff x = y \text{ or } \big(x, y \in G \;\;\&\;\; (\forall w \in U)\;\; xw = yw\big).$$

One easily verifies that $\gamma$ is a congruence on $\mathbf{G} \circlearrowright U$. Now, observe that the congruence corresponding to the ideal $U \cup \{0\}$ is non-trivial, but intersects $\gamma$ trivially. Thus, if $\mathbf{G} \circlearrowright U$ is subdirectly irreducible, $\gamma$ must be trivial, and so $\mathbf{G}$ acts faithfully on $U$.

Now, we would like to translate this condition to left coset actions. First, let us agree to use the term **group kernel** to refer to the preimage of the identity element under a group homomorphism (as opposed to the 'map' kernel). If $H$ is a subgroup of a group $\mathbf{G}$, then the group kernel of the natural action of $\mathbf{G}$ on $G/H$ is called the **normal core** of $H$ in $\mathbf{G}$, or simply the **core** of $H$ in $\mathbf{G}$. The name reflects a basic fact: the normal core of $H$ in $\mathbf{G}$ is the largest normal subgroup of $\mathbf{G}$ contained in $H$.

Indeed, if $C$ is the core of $H$, then we may write $C = \{g \in G \mid (\forall a \in G) \;\; gaH = aH\}$. Each $g \in C$ then satisfies $gH = H$ in particular, so $C$ is contained in $H$. Now, being a group kernel, $C$ is of course is a normal subgroup of $\mathbf{G}$. Moreover, an elementary manipulation shows that $C = \bigcap\{aHa^{-1} \mid a \in G\}$, the intersection of all conjugates of $H$, from which it is clear that any normal subgroup of $\mathbf{G}$ contained in $H$ is also contained in $C$.

As is well known, a group homomorphism is one-to-one if and only if its group kernel is trivial. Thus, the action of $\mathbf{G}$ on $G/H$ is faithful if and only if the core of $H$ in $\mathbf{G}$ is trivial. If $H$ has a trivial core in $\mathbf{G}$, we say that $H$ is **core-free** in $\mathbf{G}$. Equivalently, $H$ is core-free in $\mathbf{G}$ if $H$ contains no non-trivial normal subgroup of $\mathbf{G}$.

Thus, if $\mathbf{G}\circlearrowright G/H$ is subdirectly irreducible, then $H$ must be core-free in $\mathbf{G}$. However, $H$ being core-free in $\mathbf{G}$ is not a sufficient condition for $\mathbf{G} \circlearrowright G/H$ to be subdirectly irreducible. As we will establish in Proposition 3.49 below, $H$ must also be completely meet irreducible in the subgroup lattice of $\mathbf{G}$. To reframe this property, let $a \in G$. We say that the subgroup $H$ is **$a$-maximal** in $\mathbf{G}$ if $a \notin H$ and every subgroup of $\mathbf{G}$ that properly contains $H$ also contains $a$. It is easy to see that a subgroup of $\mathbf{G}$ is completely meet irreducible in the subgroup lattice of $\mathbf{G}$ if and only if it is $a$-maximal for some $a \in G$.

We are now ready to describe the subdirectly irreducibles of the form $\mathbf{G} \circlearrowright G/H$. We first prove the following lemma, which we will use several times in this chapter to verify that a semigroup of the form $\mathbf{T} \circlearrowright U$ is subdirectly irreducible.

**Lemma 3.48.** *Let $\mathbf{T}$ be a semigroup acting faithfully on a set $U$ such that no element of $\mathbf{T}$ acts as the empty map on $U$, and let $\theta$ be a non-trivial congruence on $\mathbf{T} \circlearrowright U$. Then there are distinct $x, y \in U \cup \{0\}$ with $x \;\theta\; y$. Furthermore, if for every $z \in U$ there is some $a \in T$ with $az \in U\backslash\{z\}$, then there are distinct $x, y \in U$ with $x \;\theta\; y$.*

Proof. We will first show that $x \;\theta\; y$ for some distinct $x, y \in U \cup \{0\}$. Fix $a, b \in T \cup U \cup \{0\}$ with $a \neq b$ and $a \;\theta\; b$. If $a, b \in U \cup \{0\}$, then we are done, so we can assume by symmetry that $a \in T$.

If $b \in U \cup \{0\}$, then because $a$ does not act as the empty map on $U$, there is some $u \in U$ with $au \neq 0$. Then $0 = bu \;\theta\; au \in U$, so the required result holds in this case.

If $b \in T$, then because $\mathbf{T}$ acts faithfully on $U$, there is some $u \in U$ with $au \neq bu$. If we define $x := au$ and $y := bu$, we have $a \;\theta\; b \Rightarrow x \;\theta\; y$ and $x, y \in U \cup \{0\}$, as required.

Finally, assume that for every $z \in U$, there is some $a \in T$ with $az \in U\backslash\{z\}$. We have shown there are distinct $x, y \in U \cup \{0\}$ with $x \;\theta\; y$. Suppose one of $x$ or $y$ is 0; by symmetry, we can assume it is $y$. Then, by the assumption, there exists $a \in T$ with $ax \in U\backslash\{x\}$, which gives $x \;\theta\; 0 \Rightarrow ax \;\theta\; 0$, and therefore $x \;\theta\; ax$. $\square$

**Proposition 3.49.** *Let $\mathbf{G}$ be a non-trivial group and $H$ a subgroup of $\mathbf{G}$. Then $\mathbf{G} \circlearrowright G/H$ is subdirectly irreducible if and only if $H$ is core-free in $\mathbf{G}$ and is completely meet irreducible in the lattice of subgroups of $\mathbf{G}$.*

Proof. First, assume that $\mathbf{G} \circlearrowright G/H$ is subdirectly irreducible with monolith $\mu$. Then $H$ is core-free, and is therefore a proper subgroup of $\mathbf{G}$ since $\mathbf{G}$ is non-trivial, so $|G/H| > 1$. Thus, the 'furthermore' part of Lemma 3.48 applies, so $\mu$ identifies a distinct pair in $G/H$. By translation, we have $aH \;\mu\; H$ for some $a \in G\backslash H$. We will show that $H$ is $a$-maximal.

Let $K$ be a subgroup of $\mathbf{G}$ with $K \supset H$, and let $\alpha$ be the congruence on $\mathbf{G} \circlearrowright G/H$ identifying pairs of elements of $G/H$ that are contained in the same left coset of $K$:

$$x \;\alpha\; y \iff x = y \ \text{ or } \ \big(x, y \in G/H \text{ and } x, y \subseteq gK \text{ for some } g \in G\big).$$

Then $\alpha$ is non-trivial as $K \supset H$, so $aH \;\mu\; H \Rightarrow aH \;\alpha\; H$, and hence $aH$ and $H$ are contained in the same left coset of $K$. Since $H \subseteq K$, this implies that $a \in K$, so $H$ is $a$-maximal.

Conversely, assume that $H$ is core-free and completely meet irreducible. Then $H$ is $a$-maximal for some $a \in G \backslash H$. To prove that $\mathbf{G} \circlearrowright G/H$ is subdirectly irreducible, let $\theta$ be a non-trivial congruence on $\mathbf{G} \circlearrowright G/H$; we will show that $aH \;\theta\; H$. By Lemma 3.48, $\theta$ identifies some pair in $G/H$, so by translation, we may obtain some $g \in G$ with $gH \neq H$ and $gH \;\theta\; H$. Now, consider the subgroup $K := \{x \in G \mid xH \;\theta\; H\}$ of $\mathbf{G}$. Then $K \supseteq H$, and the inclusion is strict as $g \in K \backslash H$. Since $H$ is $a$-maximal, we have $a \in K$, and so $aH \;\theta\; H$. $\square$

Proposition 3.49 is essentially McKenzie's Proposition 3 in [**47**]. In [**32**], on the other hand, Golubov and Sapir do not study semigroups of the form $\mathbf{G} \circlearrowright U$ in any detail; instead, they make use of a notion known as $\kappa$-separability. There are advantages to both approaches, so we will combine them to make the best of both worlds.

Let $\kappa$ be a cardinal, let $\mathbf{G}$ be a group, let $a \in G$, and let $H$ be a subgroup of $\mathbf{G}$. We say that the pair $(H, a)$ is **$\kappa$-inseparable** if $a \notin H$ and, for every congruence $\theta$ on $\mathbf{G}$ such that $|G/\theta| < \kappa$, there exists $h \in H$ with $a \;\theta\; h$. We call a group **$\kappa$-separable** if it does not have a $\kappa$-inseparable pair.

To connect the ideas of McKenzie with those of Golubov and Sapir, we introduce one further definition. A $\kappa$-inseparable pair $(H, a)$ is called **irreducible** if $H$ is $a$-maximal and core-free in $\mathbf{G}$. The significance of this concept is encapsulated in the following result.

**Proposition 3.50.** *Let $\mathbf{G}$ be a group, let $\kappa$ be a cardinal, and let $(H, a)$ be an irreducible $\kappa$-inseparable pair of $\mathbf{G}$. Then $|G| \geqslant \kappa$ and $\mathbf{G} \circlearrowright G/H \in \mathbf{si}(\mathbb{V}(\mathbf{G}, \mathbf{P}))$.*

Proof. If $|G| < \kappa$, then $|G/\Delta_G| < \kappa$ contradicts the fact that $(H, a)$ is $\kappa$-inseparable, so we must have $|G| \geqslant \kappa$. By assumption, $H$ is completely meet irreducible and core-free, so by Propositions 3.47 and 3.49, we have $\mathbf{G} \circlearrowright G/H \in \mathbf{si}(\mathbb{V}(\mathbf{G}, \mathbf{P}))$. $\square$

We now have the following sufficient condition for a variety $\mathcal{V}$ containing $\mathbf{P}$ to be residually large, expressed completely in terms of groups: for every cardinal $\kappa$, there is a group in $\mathcal{V}$ with an irreducible $\kappa$-inseparable pair. The next result will show that we can omit the word 'irreducible'. The proof uses another basic fact about the core: if $H$ is a subgroup of $\mathbf{G}$ and $C$ is the normal core of $H$ in $\mathbf{G}$, then $H/C$ is core-free in $\mathbf{G}/C$. To see this, let $\varphi$ denote the canonical map $\mathbf{G} \to \mathbf{G}/C$, and let $C'$ be the core of $H/C$ in $\mathbf{G}/C$. Then $\varphi^{-1}(C') \subseteq H$ and therefore $\varphi^{-1}(C') \subseteq C$, so $C'$ is contained in $\varphi(C)$, which is trivial.

**Lemma 3.51.** *Let $\kappa$ be a cardinal, and let $\mathbf{G}$ be a group which is not $\kappa$-separable. Then some homomorphic image of $\mathbf{G}$ has an irreducible $\kappa$-inseparable pair. Consequently, there is a subdirectly irreducible semigroup in $\mathbb{V}(\mathbf{G}, \mathbf{P})$ of cardinality at least $\kappa$.*

Proof. Let $(H, a)$ be a $\kappa$-inseparable pair of $\mathbf{G}$. By Zorn's Lemma, $H$ is contained in some $a$-maximal subgroup $H'$ of $\mathbf{G}$, and it is easily verified that $(H', a)$ is also $\kappa$-inseparable. So we may assume without loss of generality that $H$ is $a$-maximal.

Now, let $C$ be the core of $H$ in $\mathbf{G}$, let $\mathbf{G}' := \mathbf{G}/C$, and let $\varphi\colon \mathbf{G} \to \mathbf{G}'$ be the canonical quotient map. It is straightforward to check that $\varphi(H)$ is $\varphi(a)$-maximal and core-free in $\mathbf{G}'$. To see that $(\varphi(H), \varphi(a))$ is $\kappa$-inseparable, let $\theta$ be a congruence on $\mathbf{G}'$ with $|G'/\theta| < \kappa$, let $\eta$ be the canonical map $\mathbf{G}' \to \mathbf{G}'/\theta$, and define $\alpha := \ker(\eta \circ \varphi)$. Then $|G/\alpha| = |G'/\theta| < \kappa$, so there is some $h \in H$ with $a \mathrel{\alpha} h$. By definition of $\alpha$, this gives $\varphi(a) \mathrel{\theta} \varphi(h)$, so $(\varphi(H), \varphi(a))$ is $\kappa$-inseparable. The 'consequently' part follows from Proposition 3.50. $\square$

For the following discussion, let us call a variety $\mathcal{V}$ of groups **uniformly separable** if there is a cardinal $\kappa$ such that every group in $\mathcal{V}$ is $\kappa$-separable, and **finitely separable** if we can choose $\kappa = \omega$. Lemma 3.51 says that if a residually small variety contains $\mathbf{P}$, the subvariety of groups must be uniformly separable. Thus, in characterising the residually small varieties containing $\mathbf{P}$, we are now faced with the following question: which subvarieties of $\mathcal{G}_n$ are uniformly separable?

This question has a somewhat interesting history. Golubov and Sapir proved in [**32**] that every *finitely* separable variety must be Abelian; as they were interested only in residually *finite* semigroup varieties, this was enough to prove their main result. On the other hand, McKenzie did not address this question at all in [**47**], since it was not needed for his proof of the RS conjecture for semigroups (which was McKenzie's main motivation in [**47**]).

It was not long before McKenzie became aware of Golubov and Sapir's work, and so, in a short sequel to [**47**], McKenzie discussed in [**48**] the extent of the overlap of the two papers [**47**], [**32**], and showed that every *finite* group in a uniformly separable subvariety of $\mathcal{G}_n$ is Abelian. He also raised the more general question of whether all uniformly separable subvarieties of $\mathcal{G}_n$ are Abelian, but was unable to answer this.

Several years later, Sapir revisited the problem with Shevrin in [**61**], after becoming aware of McKenzie's work. Sapir and Shevrin finally proved in [**61**] that all uniformly separable subvarieties of $\mathcal{G}_n$ are Abelian, which completed the characterisation of the residually small varieties containing $\mathbf{P}$. Interestingly, part of the proof came from a paper of Zamyatin concerning first-order theories of group varieties [**69**].

At this point, the proof that all uniformly separable subvarieties of $\mathcal{G}_n$ are Abelian was scattered throughout at least six papers, since Golubov, Sapir, and Shevrin called on some fairly non-trivial results on group varieties. McKenzie's weaker version of this result used more elementary group-theoretic techniques, which led the present author to search for a simpler proof of the general result.

Although McKenzie's weaker result suffices for the locally finite case, which in turn suffices for the purposes of this thesis, the general result is, in my opinion, rather interesting in its own right, and after several years has accumulated some sentimental value. In the end, the broader view afforded by the various approaches revealed the core argument. It is my great pleasure to present a simplified, self-contained proof over the next three pages.

The key trigger for non-uniform separability is identified in the following result. Recall that the commutator term $[x, y]$ is defined as $xyx^{-1}y^{-1}$ (or $x^{-1}y^{-1}xy$, to the same effect), and that a normal subgroup $K$ of a group $\mathbf{G}$ is 'superclosed' under the commutator, in the sense that $[g, k], [k, g] \in K$ for all $g \in G$ and $k \in K$.

**Lemma 3.52.** *Let* $\mathbf{G}$ *be a subdirectly irreducible group with trivial centre. Then, for every cardinal* $\kappa$*, there is a group in* $\mathbb{V}(\mathbf{G})$ *that is not* $\kappa$*-separable.*

PROOF. Let $\kappa > 0$ be a cardinal, and let $\mathbf{F} \in \mathbb{V}(\mathbf{G})$ denote the subgroup of $\mathbf{G}^\kappa$ formed by the elements of finite support; i.e., $F = \{x \in G^\kappa \mid \{\alpha \in \kappa \mid x(\alpha) \neq 1\} \text{ is finite}\}$. We will show that $\mathbf{F}$ is not $\kappa$-separable by constructing a $\kappa$-inseparable pair.

Choose some $m \in G$ such that $(m, 1)$ generates the monolith of $\mathbf{G}$, let $\mathbf{A}$ denote the cyclic subgroup of $\mathbf{G}$ generated by $m$, and define

$$H := \left\{ x \in F \cap A^\kappa \;\middle|\; \prod_{\alpha \in \kappa} x(\alpha) = 1 \right\}.$$

From the definition of $\mathbf{F}$ and the fact that $\mathbf{A}$ is Abelian, the above product over $\kappa$ is well defined, and $H$ is a subgroup of $\mathbf{F}$. Now, for each $\alpha \in \kappa$ and each $g \in G$, let $1_\alpha^g$ denote the element of $G^\kappa$ with $1_\alpha^g(\alpha) = g$ and $1_\alpha^g(\beta) = 1$ for all $\beta \in \kappa \backslash \{\alpha\}$. We will show that $(H, 1_0^m)$ is $\kappa$-inseparable in $\mathbf{F}$.

Let $\theta$ be a congruence on $\mathbf{F}$ with $|F/\theta| < \kappa$, and let $K$ be the corresponding normal subgroup of $\mathbf{F}$. Since $|F| \geqslant \kappa$, there is a non-identity element $u \in K$. Fix some $\alpha \in \kappa$ such that $u(\alpha) \neq 1$, and let $a := u(\alpha)$. Because $Z(\mathbf{G}) = \{1\}$, there exists $g \in G$ with $[g, a] \neq 1$, so as $K$ is a normal subgroup of $\mathbf{F}$, we have $1_\alpha^{[g,a]} = [1_\alpha^g, u] \in K$. Now, the normal subgroup of $\mathbf{G}$ generated by $[g, a]$ contains $m$, so $1_\alpha^{[g,a]} \in K$ implies $1_\alpha^m \in K$. Finally, since $1_0^m \cdot (1_\alpha^m)^{-1} \in H$, we have $1_0^m \in HK$, so $1_0^m \;\theta\; h$ for some $h \in H$. Thus, $(H, 1_0^m)$ is $\kappa$-inseparable in $\mathbf{F}$. □

Now, our main goal for this section is, in effect, to prove that a non-Abelian subvariety $\mathcal{V}$ of $\mathcal{G}_n$ is not uniformly separable. This is clear if $\mathcal{V}$ is residually large, since a subdirectly irreducible group of cardinality at least $\kappa$ is not $\kappa$-separable, as is easily verified. It therefore suffices to show that if $\mathcal{V}$ is residually small, then $\mathcal{V}$ contains a subdirectly irreducible group with a trivial centre.

In the case that $\mathcal{V}$ contains a *finite* non-Abelian group, this is easily done (the argument is given in the last two paragraphs of the proof of Theorem 3.58, if the reader wishes to skip ahead), so the locally finite case is covered. The case where every finite group in $\mathcal{V}$ is Abelian requires some extra work. (It would of course be natural to ask if there actually exists a non-Abelian variety of groups whose finite members are Abelian. An example of such a variety was constructed by Ol'šanskiĭ in [**51**].)

Let $\mathbf{G}$ be a group. Recall that the **commutator subgroup** $[G, G]$ is defined as the normal subgroup of $\mathbf{G}$ generated by the set $\{[x, y] \mid x, y \in G\}$ of all commutators in $\mathbf{G}$. This also equals the subgroup of $\mathbf{G}$ generated by $\{[x, y] \mid x, y \in G\}$ (in general, if $X \subseteq G$ is closed under conjugation, then the subgroup generated by $X$ is normal).

The fundamental property of the commutator subgroup is that it is the least normal subgroup of $\mathbf{G}$ for which the resulting factor group is Abelian. We record this in the following proposition.

**Proposition 3.53.** *Let* $\mathbf{G}$ *be a group. Then the following hold:*

(i) *if* $N$ *is a normal subgroup of* $\mathbf{G}$*, then* $\mathbf{G}/N$ *is Abelian if and only if* $[G, G] \subseteq N$*;*
(ii) *if* $X$ *is a generating set for* $\mathbf{G}$*, then* $[G, G]$ *is the normal subgroup of* $\mathbf{G}$ *generated by* $\{[x, y] \mid x, y \in X\}$*.*

PROOF. Since $G$ is a generating set for $\mathbf{G}$, it suffices for (i) and (ii) to prove that if $X$ is a generating set for $\mathbf{G}$ and $N$ is a normal subgroup of $\mathbf{G}$, then $\mathbf{G}/N$ is Abelian if and only if $[x, y] \in N$ for all $x, y \in X$.

Let $\mathbf{H} := \mathbf{G}/N$, and let $\varphi\colon \mathbf{G} \to \mathbf{H}$ be the quotient map. Then $[\varphi(x), \varphi(y)] = \varphi([x, y])$ for all $x, y \in X$, so we have $[x, y] \in N$ for all $x, y \in X$ if and only if all elements of $\varphi(X)$ commute. But $\mathbf{H}$ is generated by $\varphi(X)$, so the latter is true if and only if $\mathbf{H}$ is Abelian. $\square$

Next, we recall the notion of a solvable group. Let $\mathbf{G}$ be a group. For $n \geqslant 0$, we define the $n$**th derived subgroup** $G^{(n)}$ of $\mathbf{G}$ inductively by $G^{(0)} := G$ and $G^{(n+1)} := [G^{(n)}, G^{(n)}]$. For each $n \geqslant 0$, we say that $\mathbf{G}$ is $n$**-solvable** if $G^{(n)} = \{1\}$. Finally, $\mathbf{G}$ called is **solvable** if it is $n$-solvable for some $n \geqslant 0$.

We will use the following basic facts about the derived subgroups.

**Proposition 3.54.** *Let $\mathbf{G}$ be a group, and let $n \geqslant 0$. Then the following hold:*

(i) *if $\mathbf{H} \leqslant \mathbf{G}$, then $H^{(n)} \subseteq G^{(n)}$;*
(ii) *if $\varphi\colon \mathbf{G} \to \mathbf{H}$ is an onto homomorphism, then $H^{(n)} = \varphi(G^{(n)})$;*
(iii) *$G^{(n)}$ is a normal subgroup of $\mathbf{G}$, and $\mathbf{G}/G^{(n)}$ is $n$-solvable.*

PROOF. Statements (i) and (ii) follow from the $n = 1$ case by easy induction arguments. For (i), the $n = 1$ case is obvious. To prove (ii) in the $n = 1$ case, let $\varphi\colon \mathbf{G} \to \mathbf{H}$ be onto. Then $\varphi(G^{(1)})$ is generated as a subgroup by $\{[\varphi(x), \varphi(y)] \mid x, y \in G\}$, but since $\varphi$ is onto, this is exactly the set $\{[x, y] \mid x, y \in H\}$.

To prove (iii), first apply (ii) to an arbitrary conjugation map $\mathbf{G} \to \mathbf{G}$ to see that $G^{(n)}$ is normal. Now, let $\mathbf{H} := \mathbf{G}/G^{(n)}$, and let $\varphi\colon \mathbf{G} \to \mathbf{H}$ be the quotient map. Then $G^{(n)}$ is the group kernel of $\varphi$, so by (ii) we have that $H^{(n)} = \varphi(G^{(n)})$ is trivial. Thus, $\mathbf{H}$ is $n$-solvable. $\square$

In addition to the basic results of Proposition 3.54, we require the result that a solvable group of finite exponent is locally finite. Our proof of this is from Dixon's text on Sylow theory [**23**, Proposition 1.1.5], and uses the following well-known result.

**Lemma 3.55** (Schreier's Lemma)**.** *Let $\mathbf{G}$ be a finitely generated group. Then every finite-index subgroup of $\mathbf{G}$ is finitely generated.*

PROOF. Let $X$ be a finite generating set for $\mathbf{G}$, let $H$ be a finite-index subgroup of $\mathbf{G}$, and let $S$ be a set of representatives of the right cosets of $H$ in $\mathbf{G}$ with $1 \in S$. For each $g \in G$, let $\overline{g}$ denote the unique element of $S \cap Hg$. We claim that the finite set

$$Y := \{sx(\overline{sx})^{-1} \mid s \in S,\ x \in X\}$$

is a generating set for $H$. Since $g\overline{g}^{-1} \in H$ for all $g \in G$, we have $Y \subseteq H$. Let $h \in H$, and write $h = x_1 \cdots x_n$ for some $x_1, \dots, x_n \in X$. For each $i \in \{0, \dots, n\}$, define $h_i := 1x_1 \cdots x_i$ and $u_i := \overline{h_i}$ (so $h_0 = u_0 = 1$). Since $u_n = \overline{h} = 1$, we may write

$$h = u_0 x_1 u_1^{-1} u_1 x_2 u_2^{-1} \cdots u_{n-1} x_n u_n^{-1},$$

so it suffices to show that $u_{i-1} x_i u_i^{-1} \in Y$ for all $i \in \{1, \dots, n\}$.

Let $i \in \{1, \dots, n\}$. Then $Hu_{i-1} = Hh_{i-1}$ by definition of $u_{i-1}$, so $Hu_{i-1}x_i = Hh_{i-1}x_i$. Thus, $\overline{u_{i-1}x_i} = \overline{h_{i-1}x_i} = \overline{h_i} = u_i$, and hence $u_{i-1}x_i u_i^{-1} = u_{i-1}x_i(\overline{u_{i-1}x_i})^{-1} \in Y$. $\square$

**Lemma 3.56.** *Let* $\mathbf{G}$ *be a solvable group of finite exponent. Then* $\mathbf{G}$ *is locally finite.*

Proof. Since every subgroup of a solvable group is solvable (Proposition 3.54(i)), it suffices to prove the following statement by induction on $n \geqslant 0$:

*If* $\mathbf{G}$ *is a finitely generated* $n$*-solvable group of finite exponent, then* $\mathbf{G}$ *is finite.*

The case $n = 0$ is trivial. Assume that the statement holds for some $n \geqslant 0$, and let $\mathbf{G}$ be an $(n+1)$-solvable finitely generated group of finite exponent. Then $\mathbf{G}/G^{(1)}$ is Abelian by Proposition 3.53(i), so as it is finitely generated with finite exponent, it is finite. Thus, $G^{(1)}$ is finite index in $\mathbf{G}$, and is therefore finitely generated by Schreier's Lemma 3.55. Now $G^{(1)}$ is $n$-solvable, so $G^{(1)}$ is finite by the inductive hypothesis. As $\mathbf{G}/G^{(1)}$ and $G^{(1)}$ are finite, it follows that $\mathbf{G}$ is finite. □

The last ingredient we need for the main result of this section is the following argument from Zamyatin [**69**, p. 17].

**Lemma 3.57.** *Let* $\mathcal{V}$ *be a non-Abelian group variety of finite exponent such that every finite group in* $\mathcal{V}$ *is Abelian. Then there is a subdirectly irreducible group in* $\mathcal{V}$ *with a trivial centre.*

Proof. If $\mathcal{V}$ has a non-Abelian solvable group, then by Lemma 3.56, there is a finite non-Abelian group in $\mathcal{V}$, which is a contradiction. Thus, every solvable group in $\mathcal{V}$ is Abelian.

Let $\mathbf{G}$ be a non-Abelian group in $\mathcal{V}$. Then there are $a, b \in G$ with $[a, b] \neq 1$, and we can assume $\mathbf{G}$ is generated by $\{a, b\}$. By Theorem 1.6, there is a congruence $\theta$ on $\mathbf{G}$ such that $\mathbf{G}/\theta$ is subdirectly irreducible and $([a, b]/\theta, 1/\theta)$ generates the monolith of $\mathbf{G}/\theta$. Thus, we can assume that $\mathbf{G}$ is subdirectly irreducible with monolith generated by $([a, b], 1)$.

Now, by Proposition 3.53(ii), the normal subgroup of $\mathbf{G}$ generated by $[a, b]$ equals $G^{(1)}$ (since $[b, a] = [a, b]^{-1}$), so $G^{(1)}$ is the least non-trivial normal subgroup of $\mathbf{G}$. By Proposition 3.54(iii), $\mathbf{G}/G^{(2)} \in \mathcal{V}$ is solvable, so it is Abelian. Thus, by Proposition 3.53(i), we have $G^{(1)} = G^{(2)}$, which implies that $G^{(1)}$ is non-Abelian, and therefore every non-trivial normal subgroup of $\mathbf{G}$ is non-Abelian. Since $Z(\mathbf{G})$ is Abelian, it must be trivial. □

**Theorem 3.58.** *Let* $\mathcal{V}$ *be a semigroup variety of finite exponent with* $\mathbf{P} \in \mathcal{V}$*. If* $\mathcal{V}$ *contains a non-Abelian group, then* $\mathcal{V}$ *is residually large.*

Proof. By Theorem 3.40, we can assume that every finite group in $\mathcal{V}$ is an A-group.

Assume that $\mathcal{V}$ contains a non-Abelian group. By Lemma 3.51 and Lemma 3.52, it suffices to show that there is a subdirectly irreducible group in $\mathcal{V}$ with trivial centre. By Lemma 3.57, we may assume that $\mathcal{V}$ contains a finite non-Abelian group.

Choose a finite non-Abelian group $\mathbf{G} \in \mathcal{V}$ of least cardinality. If all subdirectly irreducible homomorphic images of $\mathbf{G}$ were Abelian, then $\mathbf{G}$ would be a subdirect product of Abelian groups and would therefore be Abelian. Therefore, $\mathbf{G}$ must have a non-Abelian subdirectly irreducible homomorphic image, which cannot be smaller than $\mathbf{G}$, and so must be isomorphic to $\mathbf{G}$. That is, $\mathbf{G}$ is subdirectly irreducible.

Suppose that the centre of $\mathbf{G}$ is non-trivial. Then $\mathbf{G}/Z(\mathbf{G})$, being smaller than $\mathbf{G}$, is Abelian, and therefore $\mathbf{G}$ is 2-nilpotent and non-Abelian, which contradicts the fact that $\mathbf{G}$ is an A-group. Thus, $\mathbf{G}$ has trivial centre. □

## 3.6. Actions of bands

We have shown that all groups in a residually small variety containing $\mathbf{P}$ are Abelian. This already is quite a restriction, but there are still other semigroups that interact badly with $\mathbf{P}$. Much like in Section 3.5, we will show in this section that certain varieties containing $\mathbf{P}$ must contain arbitrarily large subdirectly irreducibles of the form $\mathbf{T} \circlearrowleft U$. Here, the semigroups $\mathbf{T}$ will be bands rather than groups.

To understand such semigroups, we will need to understand the identities they satisfy, and this calls for some compact notation. Let $s$ be a (non-empty) word. We define $c(s)$ (the **content** of $s$) to be the set of all variables that occur in $s$; for example, $c(xyxy) = \{x, y\}$. Importantly, if $s$ and $t$ are words, then $c(s) = c(t)$ if and only if $s \approx t$ holds in the two-element semilattice $\mathbf{I}$.

We define $\ell(s)$ and $r(s)$, respectively, to be the left-most and right-most variables occuring in $s$. Thus, if $s$ and $t$ are words, then $\mathbf{L} \models s \approx t$ if and only if $\ell(s) = \ell(t)$, and dually, we have $\mathbf{R} \models s \approx t$ if and only if $r(s) = r(t)$.

One further concept we require is that of a $u$-term. The name is based on the labelling of the elements of $\mathbf{P}$ given in Figure 3.2. If $s$ is a word, then $s$ is called a ***u*-term** if $r(s)$ occurs only once in $s$. The significance of this concept is that the $u$-terms are precisely the words that are capable of taking the value $u$ when evaluated in $\mathbf{P}$. Specifically, $s$ takes the value $u$ in $\mathbf{P}$ if and only if $s$ is a $u$-term, $r(s)$ takes the value $u$, and all other variables in $c(s)$ take the value $e$. This is easily seen from the table for $\mathbf{P}$ in Figure 3.2.

Now, let $s \approx t$ be a non-trivial identity holding in $\mathbf{P}$. Then $c(s) = c(t)$ since $\mathbf{I}$ embeds into $\mathbf{P}$. Obviously, $s$ and $t$ take the same values under all evaluations of the variables, so $s$ is a $u$-term if and only if $t$ is a $u$-term. Assume that $s$ and $t$ are both $u$-terms, and fix some evaluation of the variables so that $s$ and $t$ take the value $u$. Then $r(s)$ and $r(t)$ must take the value $u$, so because $c(s) = c(t)$, this is possible only if $r(s) = r(t)$. This proves the left-to-right implication of the following result.

**Proposition 3.59.** *Let $s$ and $t$ be distinct words. Then $\mathbf{P} \models s \approx t$ if and only if*

(i) *$c(s) = c(t)$, and*
(ii) *neither of $s$, $t$ are $u$-terms, or $s$, $t$ are both $u$-terms with $r(s) = r(t)$.*

The right-to-left implication follows from the next lemma, which is a preservation result of sorts; it shows how certain identities holding in $\mathbf{T}$ can be translated to identities in $\mathbf{T} \circlearrowleft U$. Since $\mathbf{P}$ is of the form $\mathbf{T} \circlearrowleft U$ (with $\mathbf{T}$ trivial), the result is applicable to $\mathbf{P}$ in particular.

**Lemma 3.60.** *Let $\mathbf{T}$ be a semigroup acting on a set $U$, and let $s \approx t$ be an identity holding in $\mathbf{T}$ with $c(s) = c(t)$. Then the following hold:*

(i) *if $z$ is a variable not in $s$, $t$, then $\mathbf{T} \circlearrowleft U \models sz \approx tz$;*
(ii) *if neither of $s$, $t$ are $u$-terms, then $\mathbf{T} \circlearrowleft U \models s \approx t$.*

Proof. For (i), note that $sz$ and $tz$ will evaluate to 0 in $\mathbf{T} \circlearrowleft U$ unless all variables in $s$, $t$ take values in $T$, in which case $sz$ and $tz$ take the same value because $\mathbf{T} \models s \approx t$.

In case (ii), $s$ and $t$ will evaluate to 0 in $\mathbf{T} \circlearrowleft U$ unless their variables take values in $T$, in which case $s$ and $t$ take the same value because $\mathbf{T} \models s \approx t$. $\square$

We will now introduce the final two minimal residually large varieties in this chapter. One of these is the variety $\mathbb{V}(\mathbf{L}, \mathbf{P})$. We will show that this variety contains all semigroups of the form $\mathbf{T} \circlearrowright U$, where $\mathbf{T}$ is a left-normal band.

A semigroup is called **left-normal** if it satisfies $xyz \approx xzy$. Evidently every left-normal band is normal. Thus, by Theorem 3.35, the only subdirectly irreducible left-normal bands are $\mathbf{L}$, $\mathbf{L}^0$, and $\mathbf{I}$, so the variety of left-normal bands equals $\mathbb{V}(\mathbf{L}, \mathbf{I})$. Clearly, an identity $s \approx t$ holds in all left-normal bands if and only if $c(s) = s(t)$ and $\ell(s) = \ell(t)$; this can also be seen directly from the axiomatisation $x^2 \approx x$, $xyz \approx xzy$. Of course, **right-normal** semigroups are defined dually, and $\mathbb{V}(\mathbf{R}, \mathbf{I}) = [x^2 \approx x,\ xyz \approx yxz]$ is the class of right-normal bands.

Although it will turn out that $\mathbb{V}(\mathbf{L}, \mathbf{P})$ contains all semigroups $\mathbf{T} \circlearrowright U$ where $\mathbf{T}$ is a left-normal band, the variety $\mathbb{V}(\mathbf{R}, \mathbf{P})$ does not necessarily contain $\mathbf{T} \circlearrowright U$ if $\mathbf{T}$ is a right-normal band. In fact, $\mathbb{V}(\mathbf{R}, \mathbf{P})$ will turn out to be residually small. The reason for these contrasting facts remains somewhat mysterious, though perhaps the best intuition is that the identities of $\mathbb{V}(\mathbf{R}, \mathbf{P})$ are too restrictive to facilitate faithful actions of right-normal bands. Note that $\mathbf{R}$ and $\mathbf{P}$ satisfy $xyz \approx yxz$ (i.e., they are right-normal); however, it is easily shown that if $\mathbf{T}$ is a band acting faithfully on a set $U$ and $\mathbf{T} \circlearrowright U$ is right-normal, then $\mathbf{T}$ is a semilattice.

It turns out that if a variety contains $\mathbf{R} \circlearrowright U$ for a certain action of $\mathbf{R}$ on a two-element set, then this is enough to facilitate all actions of right-normal bands. We denote by $\mathbf{R} \circlearrowright 2$ the semigroup in Figure 3.5. The element $b$ acts on $\{u, v\}$ as the total constant map with value $u$, while $a$ acts as the restriction of this map to $\{u\}$. Of course, the notation $\mathbf{R} \circlearrowright 2$ is ambiguous, but it will always refer to this particular semigroup. Note that $\mathbf{P} \in \mathbb{S}(\mathbf{R} \circlearrowright 2)$.

| $\cdot$ | $a$ | $b$ | $u$ | $v$ | $0$ |
|---|---|---|---|---|---|
| $a$ | $a$ | $b$ | $u$ | $0$ | $0$ |
| $b$ | $a$ | $b$ | $u$ | $u$ | $0$ |
| $u$ | $0$ | $0$ | $0$ | $0$ | $0$ |
| $v$ | $0$ | $0$ | $0$ | $0$ | $0$ |
| $0$ | $0$ | $0$ | $0$ | $0$ | $0$ |

FIGURE 3.5. The semigroup $\mathbf{R} \circlearrowright 2$.

The variety of $\mathbf{R} \circlearrowright 2$ is the final residually large variety that we will introduce. The aim for the remainder of this section is to prove that $\mathbb{V}(\mathbf{L}, \mathbf{P})$ and $\mathbb{V}(\mathbf{R} \circlearrowright 2)$ are residually large. As we will need to verify membership in these varieties, we will study their identities.

The identities of $\mathbb{V}(\mathbf{L}, \mathbf{P})$ are simply understood in terms of the identities of $\mathbf{L}$ and $\mathbf{P}$. To describe the identities of $\mathbf{R} \circlearrowright 2$, we will use the following notation. Given a word $s$ whose length is at least 2, we denote by $r'(s)$ the variable in $s$ occurring immediately to the left of $r(s)$; for example, $r'(xyzx) = z$.

**Proposition 3.61.** *Let $s$ and $t$ be distinct words. Then $\mathbf{R} \circlearrowright 2 \models s \approx t$ if and only if*

(i) $\mathbf{R}, \mathbf{P} \models s \approx t$, *and*
(ii) *if $s$, $t$ are both $u$-terms, then $r'(s) = r'(t)$.*

Proof. Assume that $\mathbf{R} \circlearrowleft 2 \models s \approx t$. Then (i) clearly holds; in particular, $s$, $t$ satisfy the conditions of Proposition 3.59. To prove (ii), assume that $s$, $t$ are both $u$-terms; we must show that $r'(s) = r'(t)$. That $r'(s)$ and $r'(t)$ exist is because $s$ and $t$ are distinct $u$-terms with $c(s) = c(t)$ and $r(s) = r(t)$. Now, suppose that $r'(s) \neq r'(t)$, so $s = s'xz$ and $t = t'yz$, where $x$, $y$, $z$ are pairwise distinct variables with $z \notin c(s't')$. If we evaluate the variables in $\mathbf{R} \circlearrowleft 2$ so that $(x, y, z) = (a, b, v)$ and all other variables in $s't'$ lie in $\{a, b\}$, then $s$ and $t$ will take the values $av = 0$ and $bv = v$, respectively. This contradicts $\mathbf{R} \circlearrowleft 2 \models s \approx t$.

Conversely, assume (i) and (ii) hold. Then $r(s) = r(t)$, and $s$, $t$ satisfy the conditions of Proposition 3.59; in particular, $c(s) = c(t)$. If neither of $s$, $t$ are $u$-terms, then Lemma 3.60(ii) implies $\mathbf{R} \circlearrowleft 2 \models s \approx t$, because $s \approx t$ holds in all right-normal bands. Assume that $s$ and $t$ are both $u$-terms, so $s = s'z$ and $t = t'z$ for some variable $z \notin c(s't')$. Then by (i) and (ii) we have $c(s') = c(t')$ and $r(s') = r(t')$, so $s' \approx t'$ holds in all right-normal bands, and we therefore get $\mathbf{R} \circlearrowleft 2 \models s \approx t$ by Lemma 3.60(i). $\square$

Our residual largeness proofs will also require the following result.

**Lemma 3.62.** *Let $\mathbf{T}$ be a band acting on a set $U$. Then the following hold:*

(i) *if $\mathbf{T}$ is left-normal, then $\mathbf{T} \circlearrowleft U \in \mathbb{V}(\mathbf{L}, \mathbf{P})$;*
(ii) *if $\mathbf{T}$ is right-normal, then $\mathbf{T} \circlearrowleft U \in \mathbb{V}(\mathbf{R} \circlearrowleft 2)$.*

Proof. To prove (i), assume that $\mathbf{T}$ is left-normal, and let $s \approx t$ be a non-trivial identity holding in $\mathbf{L}$ and $\mathbf{P}$. Then $\ell(s) = \ell(t)$, and $s$, $t$ satisfy the conditions of Proposition 3.59; in particular, $c(s) = c(t)$. If neither of $s$, $t$ are $u$-terms, then $\mathbf{T} \circlearrowleft U \models s \approx t$ by Lemma 3.60(ii). Assume that $s$ and $t$ are both $u$-terms, so $s = s'z$ and $t = t'z$ for some variable $z \notin c(s't')$. Then $c(s') = c(t')$ and $\ell(s') = \ell(t')$, so $\mathbf{T} \models s' \approx t'$; hence $\mathbf{T} \circlearrowleft U \models s \approx t$ by Lemma 3.60(i).

For (ii), let $s \approx t$ be a non-trivial identity holding in $\mathbf{R} \circlearrowleft 2$, so $s$, $t$ satisfy the conditions of Proposition 3.61. The argument that $\mathbf{T} \circlearrowleft U \models s \approx t$ is exactly the same as in the last paragraph of the proof of Proposition 3.61, with $\mathbf{T} \circlearrowleft U$ in place of $\mathbf{R} \circlearrowleft 2$. $\square$

**Remark.** Despite the asymmetry between the varieties $\mathbb{V}(\mathbf{L}, \mathbf{P})$ and $\mathbb{V}(\mathbf{R} \circlearrowleft 2)$, the proofs of the two cases in Lemma 3.62 are almost identical in form. It is the somewhat strange condition (ii) of Proposition 3.61 that allows for the symmetric proofs.

For our residual largeness results, all that remains now is to construct large subdirectly irreducibles.

**Theorem 3.63.** *The variety $\mathbb{V}(\mathbf{L}, \mathbf{P})$ is residually large.*

Proof. Let $\kappa$ be a cardinal, and let $X$ be a set of cardinality $\kappa$. Adjoin to $X$ two more distinct points $u, v \notin X$ to form $U := X \cup \{u, v\}$. For each finite $A \subseteq X$ and each $s \in \{u, v\}$, define the total map $f^s_A \colon U \to U$ by

$$(\forall x \in U) \;\; f^s_A(x) := \begin{cases} x & \text{if } x \in X \backslash A, \\ s & \text{if } x \in A \cup \{u, v\}. \end{cases}$$

Then, for all $s, t \in \{u, v\}$ and all finite $A, B \subseteq X$, we have $f^s_A \circ f^t_B = f^s_{A \cup B}$. This shows that the set $T := \{f^s_A \mid s \in \{u, v\},\ A \subseteq X \text{ is finite}\}$ forms a subsemigroup, $\mathbf{T}$, of the semigroup of maps $U \to U$, and that $\mathbf{T}$ is isomorphic to the direct product of $\mathbf{L}$ and the $\cup$-semilattice of finite subsets of $X$. Thus, $\mathbf{T}$ is a left-normal band.

Since $\mathbf{T}$ is a semigroup of maps on $U$, there is a natural faithful action of $\mathbf{T}$ on $U$. It is clear that $\mathbf{T} \circlearrowright U$ has cardinality at least $\kappa$, and by Lemma 3.62(i), we have $\mathbf{T} \circlearrowright U \in \mathbb{V}(\mathbf{L}, \mathbf{P})$. To see that $\mathbf{T} \circlearrowright U$ is subdirectly irreducible, let $\theta$ be a non-trivial congruence on $\mathbf{T} \circlearrowright U$. We will show that $u \mathrel{\theta} v$.

For all $x \in U$, there are maps in $T$ sending $x$ to $u$ and $v$; evidently, one of $u$, $v$ must differ from $x$. Thus, by Lemma 3.48, there are distinct $x, y \in U$ with $x \mathrel{\theta} y$. If $\{x, y\} = \{u, v\}$, then we are done, so we can assume by symmetry that $x \in X$. Choose $f \in T$ with $fx = x$ and $fy = u$. Then $x = fx \mathrel{\theta} fy = u$, and similarly $x \mathrel{\theta} v$, so $u \mathrel{\theta} v$ by transitivity. $\square$

**Theorem 3.64.** *The variety* $\mathbb{V}(\mathbf{R} \circlearrowright 2)$ *is residually large.*

Proof. Let $\kappa$ be a cardinal, and let $X$ be a set of cardinality $\kappa$. Adjoin to $X$ some $u \notin X$ to form $U := X \cup \{u\}$. For each $x \in U$, define the partial map $f_x \colon U \to U$ by

$$\operatorname{dom}(f_x) := \{u, x\}, \quad f(x) = f(u) = u.$$

Then $f_x \circ f_y = f_y$ for all $x, y \in U$, so $\mathbf{T} := \langle \{f_x \mid x \in U\}; \circ \rangle$ is a right-zero (hence right-normal) band with a natural faithful action on $U$. Clearly $\mathbf{T} \circlearrowright U$ has cardinality at least $\kappa$, and $\mathbf{T} \circlearrowright U \in \mathbb{V}(\mathbf{R} \circlearrowright 2)$ by Lemma 3.62(ii). To show that $\mathbf{T} \circlearrowright U$ is subdirectly irreducible, let $\theta$ be a non-trivial congruence on $\mathbf{T} \circlearrowright U$; we will show that $u \mathrel{\theta} 0$.

By Lemma 3.48, there are distinct $x, y \in U \cup \{0\}$ with $x \mathrel{\theta} y$. If $\{x, y\} = \{u, 0\}$, then we are done, so we can assume by symmetry that $x \in X$. If $y \in U$, then $u = f_y y \mathrel{\theta} f_y x = 0$, and if $y = 0$, then $u = f_x x \mathrel{\theta} f_x y = 0$. $\square$

**Remark.** Using Propositions 3.59 and 3.61, it is straightforward to show that

$$\mathbb{V}(\mathbf{P}) = [x^2 y \approx xy,\ xy^2 \approx yx^2], \quad \mathbb{V}(\mathbf{L}, \mathbf{P}) = [x^2 y \approx xy,\ xyz^2 \approx xzy^2],$$

$$\mathbb{V}(\mathbf{R} \circlearrowright 2) = [x^2 y \approx xy,\ xyz^2 \approx yxz^2,\ xyx \approx yx^2].$$

We note that the semigroup $\mathbf{R} \circlearrowright 2$ did not appear in Golubov and Sapir's proofs in [**32**]; they worked directly with the above equational basis. On the other hand, McKenzie worked with $\mathbb{V}(\mathbf{R} \circlearrowright U)$ for a certain total action of $\mathbf{R}$ on a three-element set [**47**, Definition 21], which is the same variety. McKenzie also allowed only total actions in the construction $\mathbf{T} \circlearrowright U$. Allowing for partial maps turned out make the residual largeness proof quite a lot simpler, in that we were able to construct our subdirectly irreducibles directly. Both of the original papers used the maximal congruence argument on more complicated semigroups, as they did in all of their residual largeness proofs.

## 3.7. Varieties satisfying one identity

In Sections 3.5 and 3.6, we found that a residually small variety satisfying $x^{n+1} y \approx xy$ but not $xy^{n+1} \approx xy$ cannot contain $\mathbf{L}$, nor $\mathbf{R} \circlearrowright 2$, nor a non-Abelian group. In this section, we establish the converse; that is, omitting these semigroups guarantees residual smallness. This will essentially complete the proof of the main result of this chapter.

The following two results will use the results of Section 3.6 to derive a powerful identity. This identity will then be used to narrow down the subdirectly irreducibles.

**Lemma 3.65.** *Let $\mathcal{V}$ be a semigroup variety satisfying $x^{n+1}y \approx xy$ for some $n \geqslant 2$. If $\mathcal{V}$ contains $\mathbf{P}$ but not $\mathbf{R}\circlearrowleft 2$, then $\mathcal{V} \models x^ny^nz \approx x^ny^nx^nz$, and therefore $\mathcal{V} \models (x^ny^n)^2 \approx x^ny^n$.*

PROOF. We claim that $\mathcal{V}$ satisfies an identity of the form $s'xz \approx t'yz$, where $s'$, $t'$ are words and $x$, $y$, $z$ are pairwise distinct variables with $z \notin c(s't')$. To prove this, let $s \approx t$ be an identity holding in $\mathcal{V}$ but not in $\mathbf{R}\circlearrowleft 2$. If $r(s) \neq r(t)$, then we can multiply $s \approx t$ on the right by some variable to get an identity of the required form. Assume that $r(s) = r(t)$. Then $s$, $t$ satisfy Proposition 3.61(i) (since $\mathbf{P} \in \mathcal{V}$), and so must fail (ii); that is, $s$, $t$ are $u$-terms with $r'(s) \neq r'(t)$. The identity $s \approx t$ now has the required form, which proves the claim.

By multiplying the identity $s'xz \approx t'yz$ on the left by $xy$ and substituting out any variables not in $\{x, y, z\}$, we can assume that $c(s') = c(t') = \{x, y\}$ and $\ell(s') = \ell(t') = x$. Now, substitute $x \mapsto x^n$ and $y \mapsto y^n$ in $s'xz \approx t'yz$. Using $x^{2n} \approx x^n$, the resulting equation can be reduced to $(x^ny^n)^ix^nz \approx (x^ny^n)^jz$ for some $i, j \geqslant 1$. By multiplying on the left by an appropriate power of $x^ny^n$ and then reducing the right-hand side using $x^{n+1}y \approx xy$, we can assume that $j = 1$, so that $\mathcal{V} \models (x^ny^n)^ix^nz \approx x^ny^nz$. If we now substitute $z \mapsto x^n$ in $(x^ny^n)^ix^nz \approx x^ny^nz$, we get $(x^ny^n)^ix^n \approx x^ny^nx^n$, and so $(x^ny^n)^ix^nz \approx x^ny^nz$ reduces to $x^ny^nx^nz \approx x^ny^nz$. From this, we get $(x^ny^n)^2 \approx x^ny^n$ by substituting $z \mapsto y^n$. □

For the next proof, we will use a simple fact about bands that we have not yet had occasion to mention: any band $\mathbf{B}$ not containing $\mathbf{L}$ satisfies $xyx \approx yx$ (so $\mathbf{B} \in \mathbb{V}(\mathbf{R}^1)$, by the dual of Proposition 3.28). To see this, let $\mathbf{B}$ be a band with $\mathbf{L} \notin \mathbb{S}(\mathbf{B})$, and let $x, y \in B$. Then $\{xyx, yx\}$ is a left-zero subsemigroup of $\mathbf{B}$, so $xyx = yx$.

**Lemma 3.66.** *Let $\mathcal{V}$ be a semigroup variety satisfying $x^{n+1}y \approx xy$ but not $xy^{n+1} \approx xy$ for some $n \geqslant 1$. If $\mathbf{R}^+, \mathbf{R}\circlearrowleft 2, \mathbf{L} \times \mathbf{P} \notin \mathcal{V}$, then $\mathcal{V} \models x^ny^nz \approx y^nx^nz$.*

PROOF. We have $\mathbf{P} \in \mathcal{V}$ by Lemma 3.45, so $\mathbf{L} \notin \mathcal{V}$; thus, every band in $\mathcal{V}$ satisfies $xyx \approx yx$. Let $m := \max\{2, n\}$, so $m$ is an exponent of $\mathcal{V}$ and $\mathcal{V} \models x^{m+1}y \approx xy$. By Lemma 3.65 with $m$ in place of $n$, every member of $\mathcal{V}$ is orthodox, so $\mathcal{V} \models x^my^mx^m \approx y^mx^m$. Now, by Lemma 3.65, we also have $\mathcal{V} \models x^my^mz \approx x^my^mx^mz$, so $\mathcal{V} \models x^my^mz \approx x^my^mx^mz \approx y^mx^mz$. If $n \neq m$ (i.e., $n = 1$), then $xyz \approx yxz$ follows from $x^2y \approx xy$ and $x^2y^2z \approx y^2x^2z$. □

Using the identity $x^ny^nz \approx y^nx^nz$, we can now narrow down the subdirectly irreducibles. The assumption that all groups are Abelian will allow us to obtain an even stronger identity.

**Lemma 3.67.** *Let $\mathcal{V}$ be a subvariety of $[x^{n+1}y \approx xy,\ x^ny^nz \approx y^nx^nz]$ for some $n \geqslant 1$ such that all groups in $\mathcal{V}$ are Abelian, and let $\mathbf{S} \in \mathcal{V}$ be subdirectly irreducible. Then $\mathbf{S}$ is isomorphic to a member of $\{\mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}, \mathbf{P}, \mathbf{G}, \mathbf{G}^0, \mathbf{G}\circlearrowleft G\}$ for some subdirectly irreducible Abelian group $\mathbf{G}$ acting on itself by left translation. Consequently, $\mathcal{V} \models xyz \approx yxz$.*

PROOF. Note that $[x^2y \approx xy,\ xyz \approx yxz] \subseteq [x^3y \approx xy,\ x^2y^2z \approx y^2x^2z]$, so we can assume that $n \geqslant 2$; thus, $n$ is an exponent of $\mathbf{S}$. Let $(a, b)$ generate the monolith of $\mathbf{S}$.

First, assume that $a, b \in \mathrm{G}(\mathbf{S})$. Since $\mathbf{N}_4 \not\models x^{n+1}y \approx xy$ and $\mathbf{R}^+ \not\models x^ny^nz \approx y^nx^nz$ (the latter by Proposition 3.20), we have $\mathbf{N}_4, \mathbf{R}^+ \notin \mathbb{V}(\mathbf{S})$, and therefore $\mathbf{S} \models (xy)^{n+1} \approx x^{n+1}y^{n+1}$ by Lemma 3.22(ii). Thus, the mapping $x \mapsto x^{n+1}$ is an endomorphism of $\mathbf{S}$, and it is one-to-one as it separates $a$ and $b$. Since a one-to-one idempotent map is the identity map, $\mathbf{S}$ is completely regular. As $\mathbf{L}, \mathbf{L}^0 \not\models x^ny^nz \approx y^nx^nz$, Theorem 3.34 gives the result in this case.

Thus, we may assume by symmetry that $a \notin \mathrm{G}(\mathbf{S})$. We will first prove that

$$(\forall e \in \mathrm{E}(\mathbf{S}))\ \ e \text{ is a left identity of } \langle e \rangle. \tag{$\circ$}$$

Let $e \in \mathrm{E}(\mathbf{S})$, and let $x \in \langle e \rangle$, so $x = uev$ for some $u, v \in S^1$. Then

$$x = uev = ueev = (ue)^n ueev = (ue)^n e^n ueev = e^n (ue)^n ueev \in eS,$$

so $ex = x$, which proves ($\circ$).

Now, we claim that $a^n$ is a zero element of $\mathbf{S}$. Suppose not; then $\langle a^n \rangle$ has more than one element, and so the corresponding ideal congruence is non-trivial. As $(a, b)$ generates the monolith of $\mathbf{S}$, we have $a \in \langle a^n \rangle$, and so $a^{n+1} = a^n a = a$ by ($\circ$), contradicting $a \notin \mathrm{G}(\mathbf{S})$. We conclude that $a^n$ is a zero element of $\mathbf{S}$, which we will denote by 0.

Suppose that $\mathrm{E}(\mathbf{S}) = \{0\}$. Then, for all $x, y \in S$, we have $xy = x^n xy = 0$ because $x^n$ is idempotent; thus, $\mathbf{S}$ is null, and we get $\mathbf{S} \cong \mathbf{N}$ by Proposition 3.1. We can therefore assume that $\mathbf{S}$ has a non-zero idempotent. We will show in this case that $\mathbf{S}$ has the form $\mathbf{G} \circlearrowright U$.

First, we will show that each non-zero idempotent is a left identity of $\mathbf{S}$. Let $e \in \mathrm{E}(\mathbf{S})$ be non-zero. Then $\langle e \rangle$ is a non-zero ideal, so $a, b \in \langle e \rangle$. By ($\circ$), we have $ea = a$ and $eb = b$, so by Lemma 3.33, $e$ is a left identity of $\mathbf{S}$.

Consider the congruence $\gamma$ on $\mathbf{S}$ identifying pairs in $\mathrm{G}(\mathbf{S})$ with the same left translation:

$$x \,\gamma\, y \iff x = y \text{ or } \big(x, y \in \mathrm{G}(\mathbf{S})\ \&\ (\forall w \in S)\ \ xw = yw\big).$$

Using Lemma 3.44, it is straightforward to verify that $\gamma$ is a congruence on $\mathbf{S}$. As $a \notin \mathrm{G}(\mathbf{S})$, we have $a/\gamma = \{a\}$, so it follows that $\gamma = \Delta_S$. Now, if $e, f \in \mathrm{E}(\mathbf{S})\backslash\{0\}$, then both $e$ and $f$ are left identities for $\mathbf{S}$, and so $e \,\gamma\, f$. Since $\gamma$ is trivial, we have $e = f$, so $\mathbf{S}$ has precisely one non-zero idempotent, which we will denote by $e$. Thus, $\mathrm{E}(\mathbf{S}) = \{0, e\}$.

Let $\mathbf{G}$ be the subgroup of $\mathbf{S}$ formed by $H_e$, and let $J := S\backslash G$. Then $x^n \in \mathrm{E}(\mathbf{S}) = \{0, e\}$ for all $x \in S$. If $x^n = e$, then $x^{n+1} = ex = x$, so $x \in \mathrm{G}(\mathbf{S})$, and therefore $x \in G$ as $x^n = e$. Conversely, if $x \in G$, then $x^n = e$. Thus, we have

$$G = \{x \in S \mid x^n = e\}, \qquad J = \{x \in S \mid x^n = 0\}.$$

Let $U := J\backslash\{0\}$. Then $U \neq \varnothing$ because $a \in S\backslash\mathrm{G}(\mathbf{S}) \subseteq U$, so $\{G, U, \{0\}\}$ is a partition of $S$.

If $x \in U$ and $y \in S$, then $xy = x^n xy = 0$, so $US = \{0\}$. If $x \in GU$, then since $US = \{0\}$, we have $x^2 = 0$, so $x \notin G$; thus, $GU \subseteq J$. To show that $0 \notin GU$, let $g \in G$ and $x \in U$. Then $0 \neq x = ex = g^{n-1}gx$, so $gx \neq 0$. Thus, $GU \subseteq U$. Hence, $\mathbf{S} \cong \mathbf{G} \circlearrowright U$, where $\mathbf{G}$ acts on $U$ by left translation. Since $e$ is a left identity of $\mathbf{S}$, this is a group action. The action is transitive by Proposition 3.46, so $\mathbf{S} \cong \mathbf{G} \circlearrowright G/H$ for some subgroup $H$ of $\mathbf{G}$.

If $\mathbf{G}$ is trivial, then $\mathbf{S} \cong \mathbf{P}$, so assume that $\mathbf{G}$ is non-trivial. Then $H$ is core-free and completely meet irreducible by Proposition 3.49. But $\mathbf{G}$ is Abelian; all subgroups are normal, so $H$ coincides with its core and must therefore be trivial. Now clearly $\mathbf{G} \circlearrowright G/H \cong \mathbf{G} \circlearrowright G$, where $\mathbf{G}$ acts on $G$ by left translation. Moreover, the trivial subgroup $H$ is completely meet irreducible in the lattice of *normal* subgroups; that is, $\mathbf{G}$ itself is subdirectly irreducible.

For the 'consequently' part, we simply note that each $\mathbf{S} \in \mathbf{si}(\mathcal{V})$ satisfies $xyz \approx yxz$; this is obvious if $\mathbf{S} \in \mathbb{I}(\mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}, \mathbf{G}, \mathbf{G}^0)$ for some Abelian group $\mathbf{G}$, and by Lemma 3.60(i) we have $\mathbf{G} \circlearrowright U \models xyz \approx yxz$ whenever $\mathbf{G}$ is an Abelian group. ☐

For each $n \geqslant 1$, we define $\mathcal{P}_n$ to be the variety $[x^{n+1}y \approx xy,\ xyz \approx yxz]$. As we have essentially shown, a variety $\mathcal{V}$ containing $\mathbf{P}$ is residually small if and only if $\mathcal{V} \subseteq \mathcal{P}_n$ for some $n \geqslant 1$ (hence the letter 'P'). The remainder of this section will just be filling in the details of this argument as well as tying the threads of the last three sections.

The next theorem describes the subdirectly irreducibles of $\mathcal{P}_n$. To state the result neatly, denote by $\mathcal{A}_n$ the variety of all Abelian groups of exponent $n$, for each $n \geqslant 1$. The members of $\mathbf{si}(\mathcal{P}_n)$ will be described in terms of the members of $\mathbf{si}(\mathcal{A}_n)$. Recall that a finite Abelian group is **primary** if it is of prime-power order. It is well known that an Abelian group of finite exponent is subdirectly irreducible if and only if it is a non-trivial primary cyclic group; see [**28**, Theorem 3.1], for example. Thus, $\mathbf{si}(\mathcal{A}_n)$ can be read as the class of all non-trivial primary cyclic groups of order dividing $n$.

**Theorem 3.68.** *For each $n \geqslant 1$, we have*

$$\mathcal{P}_n = \mathbb{V}(\mathbf{R}, \mathbf{P}) \vee \mathcal{A}_n, \qquad \mathbf{si}(\mathcal{P}_n) = \mathbb{I}(\{\mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}, \mathbf{P}, \mathbf{G}, \mathbf{G}^0, \mathbf{G} \circlearrowright G \mid \mathbf{G} \in \mathbf{si}(\mathcal{A}_n)\}).$$

*Consequently, $\mathcal{P}_n$ is residually very finite.*

Proof. Let $\mathcal{S}_n := \mathbb{I}(\{\mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}, \mathbf{P}, \mathbf{G}, \mathbf{G}^0, \mathbf{G} \circlearrowright G \mid \mathbf{G} \in \mathbf{si}(\mathcal{A}_n)\})$. First, we will show that $\mathcal{S}_n = \mathbf{si}(\mathcal{P}_n)$, starting with the inclusion $\mathcal{S}_n \subseteq \mathbf{si}(\mathcal{P}_n)$.

It follows easily from Lemma 3.60(i) that $\mathcal{S}_n \subseteq \mathcal{P}_n$. Certainly $\mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}$ are subdirectly irreducible, and it is easily seen that $\mathbf{P}$ is subdirectly irreducible (directly, or by letting $\kappa = 0$ in the construction of Theorem 3.64). Let $\mathbf{G} \in \mathbf{si}(\mathcal{A}_n)$. Then $\mathbf{G}^0$ is subdirectly irreducible by Lemma 3.32. Now, clearly $\mathbf{G} \circlearrowright G \cong \mathbf{G} \circlearrowright G/\{1\}$ and $\{1\}$ is core-free in $\mathbf{G}$. Since $\mathbf{G}$ is subdirectly irreducible and Abelian, $\{1\}$ is completely meet irreducible in the subgroup lattice of $\mathbf{G}$, so $\mathbf{G} \circlearrowright G$ is subdirectly irreducible by Proposition 3.49. Thus, $\mathcal{S}_n \subseteq \mathbf{si}(\mathcal{P}_n)$.

Now, it is clear that $\mathcal{P}_n \models x^n y^n z \approx y^n x^n z$ and that every group in $\mathcal{P}_n$ is Abelian. Moreover, every group in $\mathcal{P}_n$ has exponent $n$, as $\mathcal{P}_n \models x^{n+2} \approx x^2$. We then get $\mathbf{si}(\mathcal{P}_n) \subseteq \mathcal{S}_n$ from Lemma 3.67, so $\mathcal{S}_n = \mathbf{si}(\mathcal{P}_n)$.

Let $\mathcal{V}_n := \mathbb{V}(\mathbf{R}, \mathbf{P}) \vee \mathcal{A}_n$. Clearly $\mathcal{P}_n \supseteq \mathcal{V}_n$. Now, $\mathbf{I}, \mathbf{N} \in \mathbb{S}(\mathbf{P})$, and since $\mathbf{S}^0 \in \mathbb{H}(\mathbf{S} \times \mathbf{I})$ for any $\mathbf{S}$, we have $\mathbf{R}^0, \mathbf{G}^0 \in \mathcal{V}_n$ for any $\mathbf{G} \in \mathcal{A}_n$. We also have $\mathbf{G} \circlearrowright G \in \mathcal{V}_n$ for any $\mathbf{G} \in \mathcal{A}_n$ by Proposition 3.47, so $\mathcal{S}_n \subseteq \mathcal{V}_n$. Thus, $\mathcal{P}_n = \mathcal{V}_n$. Finally, each $\mathbf{G} \in \mathbf{si}(\mathcal{A}_n)$ is a primary cyclic group of order dividing $n$, so $\mathcal{P}_n$ is residually very finite. $\square$

We collect the results of the last three sections in the following theorem.

**Theorem 3.69.** *Let $\mathcal{V}$ be a semigroup variety satisfying $x^{n+1}y \approx xy$ but not $xy^{n+1} \approx xy$ for some $n \geqslant 1$. The following are equivalent:*

(i) *$\mathcal{V}$ is residually small;*
(ii) *$\mathcal{V}$ is residually very finite;*
(iii) *$\mathbf{R}^+$, $\mathbf{L} \times \mathbf{P}$, $\mathbf{R} \circlearrowright 2 \notin \mathcal{V}$, and $\mathbf{G} \times \mathbf{P} \notin \mathcal{V}$ for every non-Abelian group $\mathbf{G}$;*
(iv) *$\mathcal{V} \subseteq \mathcal{P}_n$.*

Proof. (i) $\Rightarrow$ (iii) is true by Theorems 3.21, 3.58, 3.63, and 3.64, while (iv) $\Rightarrow$ (ii) by Theorem 3.68, and (ii) $\Rightarrow$ (i) is trivial. It suffices to show that (iii) $\Rightarrow$ (iv).

Assume that (iii) holds. We have $\mathbf{P} \in \mathcal{V}$ by Lemma 3.45, so all groups in $\mathcal{V}$ are Abelian, and $\mathcal{V} \models x^n y^n z \approx y^n x^n z$ by Lemma 3.66. Hence, $\mathcal{V} \models xyz \approx yxz$ by Lemma 3.67. $\square$

## 3.8. The main theorem

With this short section, we come to the end of the chapter and the first half of the thesis. At this point, all we need is some notation to compactify the statements of the results.

Define $\mathcal{F}$ to be the class of all semigroups that are either isomorphic or anti-isomorphic to a member of the following class:

$$\{\mathbf{N}_4, \mathbf{L}^+, \mathbf{L}^1, \mathbf{L}\times\mathbf{P}, \mathbf{R}\circlearrowleft 2, \mathbf{M}_p, \mathbf{G}\times\mathbf{P} \mid p \text{ is prime, } \mathbf{G} \text{ is a non-Abelian group}\}.$$

The semigroups in $\mathcal{F}$ are the 'forbidden semigroups': a semigroup variety is residually small if and only if it does not contain a member of $\mathcal{F}$. (To recall the definitions of these semigroups, the reader may find useful the notation index on page 157.)

The varieties $\mathcal{N}_n$ and $\mathcal{P}_n$, respectively, were defined in Sections 3.4 and 3.7, and their subdirectly irreducibles described in Theorems 3.37 and 3.68. For each $n \geqslant 1$, let $\mathcal{Q}_n$ denote the variety $[xy^{n+1} \approx xy,\ xyz \approx xzy]$, which is the dual variety of $\mathcal{P}_n$.

Combining Theorems 3.23, 3.39, and 3.69 now gives the following.

**Theorem 3.70.** *Let $\mathcal{V}$ be a semigroup variety, and let $\mathcal{G}$ be the class of all groups in $\mathcal{V}$. Then the following are equivalent:*

(i) *$\mathcal{V}$ is residually small;*
(ii) *$\mathcal{V}$ does not contain a semigroup in $\mathcal{F}$, and there is a cardinal bound on $\mathbf{si}(\mathcal{G})$;*
(iii) *$\mathcal{V}$ is contained in $\mathcal{N}_n$, $\mathcal{P}_n$, or $\mathcal{Q}_n$ for some $n \geqslant 1$, and $\mathcal{G}$ is residually small.*

**Remark.** In Theorem 3.70, the equivalence of condition (ii) with the others has not been explicitly shown in the literature, but is of course embedded in the proofs of Golubov, Sapir, and McKenzie. For the next half of the thesis, it was important for us to extract this 'forbidden semigroup' condition.

Finally, we give a strengthened version of Theorem 3.70 for finitely generated varieties. Define the following class of finite semigroups:

$$\mathcal{F}_\omega := \{\mathbf{S} \in \mathcal{F} \mid \mathbf{S} \text{ is finite}\} \cup \{\mathbf{G} \mid \mathbf{G} \text{ is a finite group that is not an A-group}\}.$$

Note that if $\mathbf{S}$ is a finite semigroup, then by Theorem 1.10, $\mathbb{V}(\mathbf{S})$ contains a member of $\mathcal{F}_\omega$ if and only if it contains a finite member of $\mathcal{F}$.

Theorem 3.70 now combines with Theorems 3.43 and 3.68 to give the following result.

**Theorem 3.71.** *Let $\mathbf{S}$ be a finite semigroup. Then the following are equivalent:*

(i) *$\mathbb{V}(\mathbf{S})$ is residually small;*
(ii) *$\mathbb{V}(\mathbf{S})$ is residually very finite;*
(iii) *$\mathbb{V}(\mathbf{S})$ does not contain a semigroup in $\mathcal{F}_\omega$;*
(iv) *$\mathbf{S} \in \mathcal{N}_n \cup \mathcal{P}_n \cup \mathcal{Q}_n$ for some $n \geqslant 1$, and every subgroup of $\mathbf{S}$ is an A-group.*

# Part 2

# The Dualisability of Finite Aperiodic Semigroups

# Preface to Part 2

This brief discussion is for the benefit of two roughly-defined audiences: those who are inclined towards semigroup theory but are not familiar with the theory of natural dualities for algebraic structures (as developed by Davey and Werner in [**21**]), as well as those who are familiar with duality theory but require a refresher on the known results for semigroups. Dualisability will be defined precisely in Chapter 4; for now, we can roughly define a finite algebra $\mathbf{A}$ to be dualisable if there is a dual categorical equivalence between $\mathbb{SP}(\mathbf{A})$ and another category (of a certain kind).

The **dualisability problem** for a class $\mathfrak{K}$ is the problem of determining which algebras in $\mathfrak{K}$ have the property of being dualisable. Prior to the author's work, the dualisability problem for finite semigroups had been solved for several well-known subclasses. A finite group is dualisable if and only if it is an A-group [**57**, Theorem 4.1] [**41**, Corollary 6.14]; a finite band is dualisable if and only if it is normal [**2**, Corollary 4.9]; a finite nilpotent semigroup is dualisable if and only if it is null [**5**, Theorem 10.7.1] [**40**, Corollary 5.3]; a finite simple or zero-simple semigroup is dualisable if and only if its non-zero $\mathcal{J}$-class is the direct product of a rectangular band and an A-group [**40**, Theorem 7.1, Corollary 7.3].

In all of these classes, the dualisable finite members are precisely those that generate a residually small variety, as can be readily deduced from the results of Chapter 3. Although the dualisability problem is unresolved for classes such as finite inverse semigroups and finite monoids, it was at least known that the each of the dualisable members must generate residually small varieties [**39**, Theorem 15, Theorem 16]. Furthermore, there were no known finite semigroups $\mathbf{S}$ failing the equivalence

$$\text{‘}\mathbb{V}(\mathbf{S}) \text{ is residually small'} \iff \text{‘}\mathbf{S} \text{ is dualisable'}. \qquad (\heartsuit)$$

Despite the known results for semigroups, it would still have been somewhat rash to conjecture that ($\heartsuit$) holds for all finite semigroups $\mathbf{S}$, because the equivalence was known to be false in other classes; indeed, counterexamples existed in both directions. There were many known examples of non-dualisable algebras that generate a residually small variety, such as the two-element implication algebra [**5**, Theorem 10.7.1]. In the other direction, the only known example of a dualisable algebra generating a residually large variety was a certain term reduct of a four-element unital ring [**16**, §4], which was constructed specifically as a counterexample.

Even without these counterexamples, it would be strange if the equivalence held in general, because the dualisability of an algebra is fundamentally a quasivariety property. In the first place, the very notion of a natural duality based on a finite algebra $\mathbf{A}$ is a dual equivalence between $\mathbb{SP}(\mathbf{A})$ and another category. Furthermore, it is known that if two finite algebras generate the same quasivariety, then one is dualisable if and only if the other is [**22**] [**63**]. There is no obvious theoretical reason that the dualisability of $\mathbf{A}$ ought to depend only on $\mathbb{V}(\mathbf{A})$; it just happens to be the case for the above-mentioned classes.

The combination of all of these facts and observations is what drew the author into the depths of the dualisability problem for finite semigroups. In the end, this tantalising connection between dualisability and residual smallness perhaps became more interesting than the dualisability problem itself. If ($\heartsuit$) holds for all finite semigroups $\mathbf{S}$, is there some underlying reason for this? If not, then what exactly is the class of dualisable finite semigroups, and is there nonetheless some connection with residual smallness?

In this half of the thesis, we will answer these questions in some part as we tackle the dualisability problem for finite aperiodic semigroups. A semigroup $\mathbf{S}$ is called **aperiodic** if every subgroup of $\mathbf{S}$ is trivial, or equivalently, if $\mathrm{G}(\mathbf{S}) = \mathrm{E}(\mathbf{S})$. The choice to focus on such semigroups was not made at the outset of the author's graduate studies; in fact, the initial focus was on the dualisability problem for Clifford semigroups (completely regular semigroups whose idempotents commute). The shift to aperiodic semigroups was partly due to the difficulty of working with groups, but the main reason was that the aperiodic case turned out to be far more interesting than we could have imagined.

The main results of this thesis will be presented in Chapters 5 and 6. Chapter 4 will be devoted to introducing the concept of dualisability, as well as covering the relevant known results and some basics of the theory of quasivarieties. Once dualisability has actually been defined, it will be quite easy to state the main results of this thesis; however, I will deliberately hold off on doing so. The reader can always skip to the end, but I think the story is more interesting without knowing the ending.

CHAPTER 4

# Natural Dualities and Quasivarieties

The overall goal of this chapter is to cover the background necessary to understand and prove the novel results in Chapters 5 and 6. We begin in Section 4.1 with a brief introduction to the theory of natural dualities, as developed by Davey and Werner in [**21**]. For understanding the remainder of the thesis, the contents of Section 4.1 should be sufficient, but there is much more that can and should be said to those wishing to understand the inner workings of this theory; for such readers, consulting the standard text by Clark and Davey [**5**] is recommended.

The discussion of natural dualities will assume basic knowledge of common category-theoretic notions (such as the general definition of a dual equivalence of categories), but the definition of dualisability will be comprehensible without a precise understanding of the category theory, and we will not use any category theory beyond Section 4.1.

The kinds of dual equivalences we construct will always involve topology; however, the general results in Section 4.1 were developed so that one can obtain dualisability and non-dualisability results without engaging in any topology. To apply these results, one must verify conditions that, while somewhat technical, are easily formulated and understood in purely algebraic terms.

The (IC) Duality Theorem, for example, gives a sufficient condition for a finite algebra $\mathbf{A}$ to be dualisable. The theorem requires one to show that certain partial maps $A^n \to A$ that preserve certain operations and relations can always be extended to term functions of $\mathbf{A}$. The only potentially new concept that must be understood by a universal algebraist or semigroup theorist is that of a finitary partial operation, and particularly, what it means for such operations to be preserved. This merely requires some careful definitions.

As mentioned in the preface to Part 2, the dualisability of a finite algebra depends only on its quasivariety; that is, if finite algebras $\mathbf{A}$, $\mathbf{A}'$ generate the same quasivariety, then $\mathbf{A}$ is dualisable if and only if $\mathbf{A}'$ is dualisable. This fundamental result, known as the Independence of Generator Theorem, can be extremely useful for proving that all members of a certain class are dualisable. For example, AlDhamri's proof that all normal bands are dualisable [**2**] made use of Gerhard's result that there are only finitely many quasivarieties of normal bands [**31**]. Given the Independence of Generator Theorem, it was sufficient to choose a generator for each quasivariety and show it is dualisable, reducing the dualisability proofs to a finite number of semigroups.

In order to appeal to the Independence of Generator Theorem in this manner, we will cover some of the basics of the theory of quasivarieties, with particular emphasis on the notion of a quasicritical algebra. All locally finite quasivarieties are generated by their quasicritical members, just as varieties are generated by their subdirectly irreducible members. In particular, a locally finite quasivariety with only finitely many quasicriticals (up to isomorphism) has only finitely many subquasivarieties. For a semigroup quasivariety, this situation implies that every subquasivariety is generated by a finite direct product of quasicriticals. Consequently, understanding the quasicritical members of the classes we consider will be crucial to deriving our general dualisability results.

In Section 4.2, we will explain the known results concerning the dualisability of semigroups, and in Section 4.5, we will give a new proof that all normal bands are dualisable, which slightly strengthens the original result of AlDhamri. Since every band is aperiodic, our dualisability results for aperiodic semigroups will naturally build on the results for bands, so an understanding of our dualisability proofs for normal bands should prepare the reader for the rather technical dualisability proofs in Chapter 5.

## 4.1. Natural dualities and dualisability

The theory of natural dualities, developed by Davey and Werner in [**21**], provides a general framework for studying and constructing dualities for quasivarieties generated by a finite algebra. In this section, we will cover the basics of this theory, to the extent required to prove our main results. Although we will provide some exposition, we direct the reader to Clark and Davey [**5**] for a more thorough development and a wide range of examples.

Fix a finite algebra $\mathbf{A}$, and let $\boldsymbol{\mathcal{A}} := \mathbb{SP}(\mathbf{A})$. One of the main goals within the theory of natural dualities is to determine whether there exists a category $\boldsymbol{\mathfrak{X}}$ (of a certain kind) that is dually equivalent to $\boldsymbol{\mathcal{A}}$. A desirable aspect of such a duality is that it gives a representation for the members of $\boldsymbol{\mathcal{A}}$, in a manner generalising Stone's representation for Boolean algebras, for example.

We consider a uniform method for constructing dualities. The basis of the construction is to take a topological structure $\underset{\sim}{\mathbf{A}}$ with the same underlying set as $\mathbf{A}$, generate a topological quasivariety $\boldsymbol{\mathfrak{X}}$ from $\underset{\sim}{\mathbf{A}}$, and define functors $D\colon \boldsymbol{\mathcal{A}}^{\mathrm{op}} \to \boldsymbol{\mathfrak{X}}$ and $E\colon \boldsymbol{\mathfrak{X}}^{\mathrm{op}} \to \boldsymbol{\mathcal{A}}$ which constitute a 'weak duality', formally called a dual adjunction. The goal then is to study these functors and determine whether we have a true duality between $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathfrak{X}}$—or at least a subcategory of $\boldsymbol{\mathfrak{X}}$, as this is enough to obtain a representation for the members of $\boldsymbol{\mathcal{A}}$.

The functors $D$ and $E$ are defined from $\mathbf{A}$ and $\underset{\sim}{\mathbf{A}}$ in a specific way (by 'homming into' $\mathbf{A}$ and $\underset{\sim}{\mathbf{A}}$, respectively), so the construction of this potential duality for $\boldsymbol{\mathcal{A}}$ depends entirely on the choice of structure $\underset{\sim}{\mathbf{A}}$. All of the category theory is taken care of by a few general results, so determining whether $D$ and $E$ give us a duality is reduced to a careful study of the structures $\mathbf{A}$ and $\underset{\sim}{\mathbf{A}}$ using familiar algebraic and combinatorial techniques.

Since $\underset{\sim}{\mathbf{A}}$ is the core of the duality construction, we must carefully specify the structure that we will allow on $\underset{\sim}{\mathbf{A}}$. To have any hope of obtaining a duality, it is necessary to put a topology on $\underset{\sim}{\mathbf{A}}$, and for us, this will always be the discrete topology. Other than the topology, the structure of $\underset{\sim}{\mathbf{A}}$ will be limited to finitary relations and finitary partial operations (including total operations). Thus, $\underset{\sim}{\mathbf{A}}$ will have the form $\langle A; G, R, \mathfrak{T}\rangle$, where $\mathfrak{T}$ is the discrete topology, $G$ is a set of finitary partial operations on $A$, and $R$ is a set of finitary relations on $A$. The potential dual category $\boldsymbol{\mathfrak{X}}$ will consist of structures of the same type as $\underset{\sim}{\mathbf{A}}$. To define $\boldsymbol{\mathfrak{X}}$ precisely, we will now introduce the necessary notions and constructs.

Extending the notion of an algebraic signature or type from universal algebra, we will think of a **partial algebraic type** as a set $G$ of partial operation symbols with a function $G \to \omega$ specifying the arity of each symbol. A **partial algebra** $\mathbf{X} = \langle X; G^{\mathbf{X}}\rangle$ of type $G$ generalises the notion of an algebra (though we allow the set $X$ to be empty): the set $G^{\mathbf{X}} = \{g^{\mathbf{X}} \mid g \in G\}$ interprets the symbols in $G$ as operations on $X$, so for each $g \in G$ with arity $n \in \omega$, we have a partial map $g^{\mathbf{X}} \colon X^n \to X$.

There are several natural ways to define the preservation of partial operations, so we require a careful definition. Let $\mathbf{X}$ and $\mathbf{Y}$ be partial algebras of type $G$, and let $\varphi \colon X \to Y$ be a total map. If $g \in G$ and $n \in \omega$ is the arity of $g$, we say that $\varphi$ **preserves** $g$ if for all $(x_1, \dots, x_n) \in \operatorname{dom}(g^{\mathbf{X}})$, we have

$$(\varphi(x_1), \dots, \varphi(x_n)) \in \operatorname{dom}(g^{\mathbf{Y}}) \;\;\&\;\; \varphi(g^{\mathbf{X}}(x_1, \dots, x_n)) = g^{\mathbf{Y}}(\varphi(x_1), \dots, \varphi(x_n)).$$

The preservation of relations is perhaps a more familiar concept. A **relational type** is a set $R$ of relation symbols with an arity function $R \to \omega \backslash \{0\}$. A **relational structure** of type $R$ is a structure $\mathbf{X} := \langle X; R^{\mathbf{X}}\rangle$, where $X$ is a set and $R^{\mathbf{X}} = \{r^{\mathbf{X}} \mid r \in R\}$ is an interpretation of $R$ on the set $X$; i.e., for each $r \in R$ with arity $n$, we have $r^{\mathbf{X}} \subseteq X^n$.

Let $\mathbf{X}$ and $\mathbf{Y}$ be relational structures of type $R$, and let $\varphi \colon X \to Y$ be a (total) map. If $r \in R$ and $n$ is the arity of $r$, we say that $\varphi$ **preserves** $r$ if for all $(x_1, \dots, x_n) \in r^{\mathbf{X}}$, we have $(\varphi(x_1), \dots, \varphi(x_n)) \in r^{\mathbf{Y}}$.

Now, let $G$ be a partial algebraic type and $R$ a relational type. A **topological structure** of type $(G, R)$ is a structure of the form $\mathbf{X} = \langle X; G^{\mathbf{X}}, R^{\mathbf{X}}, \mathcal{T}\rangle$, where $\mathcal{T}$ is a topology on $X$, while $\langle X; G^{\mathbf{X}}\rangle$ is a partial algebra of type $G$ and $\langle X; R^{\mathbf{X}}\rangle$ is a relational structure of type $R$. Finally, let $\mathbf{X}$ and $\mathbf{Y}$ be topological structures of type $(G, R)$, and let $\varphi \colon X \to Y$ be a map. We say that $\varphi$ is a **morphism** if $\varphi$ is continuous and preserves every $g \in G$ and every $r \in R$. With this notion of morphism, the class of all topological structures of type $(G, R)$ is easily seen to be a category; isomorphisms are thus defined categorically.

To construct our potential dual category $\mathcal{X}$ from a given topological structure $\underset{\sim}{\mathbf{A}}$, we require notions of product and substructure. Let $\{\mathbf{X}_i \mid i \in I\}$ be a set of topological structures of type $(G, R)$. The **product** of $\{\mathbf{X}_i \mid i \in I\}$ is the topological structure $\mathbf{X}$ of type $(G, R)$ defined as follows. The underlying set of $\mathbf{X}$ is the Cartesian product $\prod_{i \in I} X_i$, and the topology of $\mathbf{X}$ is the product topology. The structure of $\mathbf{X}$ is defined pointwise; spelling this out, for each $r \in R$ of arity $n$, we define

$$r^{\mathbf{X}} := \{(x_1, \dots, x_n) \in X^n \mid (\forall i \in I) \;\; (x_1(i), \dots, x_n(i)) \in r^{\mathbf{X}_i}\},$$

while for each $g \in G$ of arity $n$, the partial operation $g^{\mathbf{X}} \colon X^n \to X$ is defined by

$$\operatorname{dom}(g^{\mathbf{X}}) := \{(x_1, \dots, x_n) \in X^n \mid (\forall i \in I) \;\; (x_1(i), \dots, x_n(i)) \in \operatorname{dom}(g^{\mathbf{X}_i})\},$$

$$(\forall (x_1, \dots, x_n) \in \operatorname{dom}(g^{\mathbf{X}}))(\forall i \in I) \;\; g^{\mathbf{X}}(x_1, \dots, x_n)(i) := g^{\mathbf{X}_i}(x_1(i), \dots, x_n(i)).$$

The product $\mathbf{X}$ can be shown to be a categorical product of $\{\mathbf{X}_i \mid i \in I\}$ (in the category of all topological structures of type $(G, R)$). If $\mathcal{K}$ is a class of topological structures, we denote by $\mathbb{P}^+(\mathcal{K})$ the class of all topological structures that are isomorphic to a product of members of $\mathcal{K}$, where the product is taken over a non-empty index set. (The empty product is excluded to account for the fact that we do not allow the algebras in $\mathcal{A}$ to be empty.)

Finally, let $\mathbf{X}$ be a topological structure of type $(G, R)$, and let $Y \subseteq X$. Assume that $Y$ is closed under the operations in $G^{\mathbf{X}}$, in the sense that for every $g \in G$ with arity $n$ and for all $(x_1, \dots, x_n) \in \mathrm{dom}(g^{\mathbf{X}}) \cap Y^n$, we have $g^{\mathbf{X}}(x_1, \dots, x_n) \in Y$. Then we may define a structure on $Y$ in the obvious way: restrict each operation and relation of $\mathbf{X}$ to $Y$, and put the subspace topology on $Y$. Again, to spell this out, we define $r^{\mathbf{Y}} := r^{\mathbf{X}} \cap Y^n$ for each $n$-ary $r \in R$, and for each $n$-ary $g \in G$, we define $g^{\mathbf{Y}} : Y^n \to Y$ by

$$\mathrm{dom}(g^{\mathbf{Y}}) := \mathrm{dom}(g^{\mathbf{X}}) \cap Y^n, \quad (\forall (x_1, \dots, x_n) \in \mathrm{dom}(g^{\mathbf{Y}}))\ \ g^{\mathbf{Y}}(x_1, \dots, x_n) := g^{\mathbf{X}}(x_1, \dots, x_n).$$

The resulting structure $\mathbf{Y}$ is a topological structure of type $(G, R)$. Any structure arising in this fashion is called a **substructure** of $\mathbf{X}$. If, further, $Y$ is a closed subset of the underlying topological space of $\mathbf{X}$, then we say that $\mathbf{Y}$ is a **closed substructure** of $\mathbf{X}$, and write $\mathbf{Y} \leqslant \mathbf{X}$. If $\mathcal{K}$ is a class of topological structures of type $(G, R)$, we denote by $\mathbb{S}_c(\mathcal{K})$ the class of all topological structures of type $(G, R)$ that are isomorphic to a closed substructure of a member of $\mathcal{K}$.

We can now construct our potential dual category $\mathcal{X}$. Assume that we have fixed a discrete topological structure $\underset{\sim}{\mathbf{A}}$ of type $(G, R)$ with the same underlying set as our finite algebra $\mathbf{A}$. We define $\mathcal{X}$ to be the category $\mathbb{S}_c\mathbb{P}^+(\underset{\sim}{\mathbf{A}})$ of topological structures of type $(G, R)$.

Next, we aim to define functors $D\colon \mathcal{A}^{\mathrm{op}} \to \mathcal{X}$ and $E\colon \mathcal{X}^{\mathrm{op}} \to \mathcal{A}$ that give us a dual adjunction between $\mathcal{A}$ and $\mathcal{X}$. These will be based on contravariant hom-functors, which are important objects in category theory. Recalling the general construction, let $\mathcal{C}$ be a (locally small) category, and let **Set** denote the category of sets and maps. For a fixed object $\mathbf{A}$ in $\mathcal{C}$, we define the **contravariant hom-functor** $h^{\mathbf{A}}\colon \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$ as follows. An object $\mathbf{B}$ in $\mathcal{C}$ is mapped to the set $h^{\mathbf{A}}(\mathbf{B}) := \mathcal{C}(\mathbf{B}, \mathbf{A})$ of all $\mathcal{C}$-morphisms $\mathbf{B} \to \mathbf{A}$, while a $\mathcal{C}$-morphism $u\colon \mathbf{B} \to \mathbf{C}$ is mapped to the function

$$h^{\mathbf{A}}(u)\colon \mathcal{C}(\mathbf{C}, \mathbf{A}) \to \mathcal{C}(\mathbf{B}, \mathbf{A}), \quad (\forall x\colon \mathbf{C} \to \mathbf{A})\ \ h^{\mathbf{A}}(u)(x) := x \circ u.$$

It is straightforward to check that $h^{\mathbf{A}}$ is a functor $\mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$.

The idea now is to take the hom-functors $h^{\mathbf{A}}\colon \mathcal{A}^{\mathrm{op}} \to \mathbf{Set}$ and $h^{\underset{\sim}{\mathbf{A}}}\colon \mathcal{X}^{\mathrm{op}} \to \mathbf{Set}$ and lift them to functors $D\colon \mathcal{A}^{\mathrm{op}} \to \mathcal{X}$ and $E\colon \mathcal{X}^{\mathrm{op}} \to \mathcal{A}$. More precisely, let $\mathbf{B}$ be an object in $\mathcal{A}$. The hom-set $h^{\mathbf{A}}(\mathbf{B}) = \mathcal{A}(\mathbf{B}, \mathbf{A})$ is a set of maps $B \to A$ and can therefore be regarded as a subset of the topological structure $\underset{\sim}{\mathbf{A}}^B$. If it happens that $\mathcal{A}(\mathbf{B}, \mathbf{A})$ is a closed substructure of $\underset{\sim}{\mathbf{A}}^B$, then we can put this structure on $\mathcal{A}(\mathbf{B}, \mathbf{A})$ to obtain an object in $\mathcal{X}$. This will give us an object mapping for a functor $D\colon \mathcal{A}^{\mathrm{op}} \to \mathcal{X}$.

Whether or not $\mathcal{A}(\mathbf{B}, \mathbf{A}) \leqslant \underset{\sim}{\mathbf{A}}^B$ obviously depends on the choice of $\underset{\sim}{\mathbf{A}}$, and there turns out to be a simple condition on $\underset{\sim}{\mathbf{A}}$ that ensures that we do get a substructure. In fact, this condition will also guarantee that $\mathcal{X}(\mathbf{X}, \underset{\sim}{\mathbf{A}})$ is a subalgebra of $\mathbf{A}^X$ for every object $\mathbf{X}$ in $\mathcal{X}$, and will be sufficient for us to obtain a dual adjunction from the hom-functors. The compatibility condition is given in the following definition.

**Definition 4.1.** Let $\mathbf{A}$ be a finite algebra.

- Let $r \subseteq A^n$ for some $n \in \omega\backslash\{0\}$. The relation $r$ is **compatible with A** if $r$ forms a subalgebra of $\mathbf{A}^n$. (Since subalgebras are non-empty, we then have $r \neq \varnothing$.)
- Let $g\colon A^n \to A$ be a partial map for some $n \in \omega$. The operation $g$ is **compatible with A** if $\mathrm{dom}(g)$ forms a subalgebra of $\mathbf{A}^n$ and $g$ is a homomorphism $\mathrm{dom}(g) \to \mathbf{A}$.

- Let $\underset{\sim}{\mathbf{A}}$ be a discrete topological structure (of some type $(G, R)$) with the same underlying set as $\mathbf{A}$. Then $\underset{\sim}{\mathbf{A}}$ is called an **alter ego** of $\mathbf{A}$ if every operation and relation of $\underset{\sim}{\mathbf{A}}$ is compatible with $\mathbf{A}$.

We can now define our functors precisely.

**Proposition 4.2.** *Let $\mathbf{A}$ be a finite algebra, let $\underset{\sim}{\mathbf{A}}$ be an alter ego of $\mathbf{A}$, and define $\boldsymbol{\mathcal{A}}{:=}\mathbb{SP}(\mathbf{A})$ and $\boldsymbol{\mathfrak{X}} := \mathbb{S}_c\mathbb{P}^+(\underset{\sim}{\mathbf{A}})$. Define the functors $D\colon \boldsymbol{\mathcal{A}}^{\mathrm{op}} \to \boldsymbol{\mathfrak{X}}$ and $E\colon \boldsymbol{\mathfrak{X}}^{\mathrm{op}} \to \boldsymbol{\mathcal{A}}$ as follows. For each object $\mathbf{B}$ in $\boldsymbol{\mathcal{A}}$ and each object $\mathbf{X}$ in $\boldsymbol{\mathfrak{X}}$, define*

$$D(\mathbf{B}) := \boldsymbol{\mathcal{A}}(\mathbf{B}, \mathbf{A}) \leqslant \underset{\sim}{\mathbf{A}}^B, \qquad\qquad E(\mathbf{X}) := \boldsymbol{\mathfrak{X}}(\mathbf{X}, \underset{\sim}{\mathbf{A}}) \leqslant \mathbf{A}^X.$$

*For each $\boldsymbol{\mathcal{A}}$-morphism $u\colon \mathbf{B} \to \mathbf{C}$ and each $\boldsymbol{\mathfrak{X}}$-morphism $\varphi\colon \mathbf{X} \to \mathbf{Y}$, define*

$$D(u)\colon D(\mathbf{C}) \to D(\mathbf{B}), \qquad\qquad E(\varphi)\colon E(\mathbf{Y}) \to E(\mathbf{X}),$$

$$(\forall x\colon \mathbf{C} \to \mathbf{A})\ \ D(u)(x) := x \circ u, \qquad (\forall \alpha\colon \mathbf{Y} \to \underset{\sim}{\mathbf{A}})\ \ E(\varphi)(\alpha) := \alpha \circ \varphi.$$

*Then $D$ and $E$ are well-defined functors constituting a dual adjunction between $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathfrak{X}}$.*

Proof. See Clark and Davey [**5**, Theorem 1.5.3]. □

**Remark.** Our constructions in Propositions 4.2 may seem somewhat arbitrary to those seeing them for the first time, but they are quite natural. As shown by Porst and Tholen [**55**], a very broad class of dual equivalences between concrete categories can be induced by contravariant hom-functors based on two objects with the same underlying set (so-called concrete dualities). Even those that cannot (such as Esakia duality and other restricted Priestley dualities) are typically built off of concrete dualities.

Now, the functors $D$ and $E$ will have the stronger property of being a dual equivalence provided that the composite functors $ED\colon \boldsymbol{\mathcal{A}} \to \boldsymbol{\mathcal{A}}$ and $DE\colon \boldsymbol{\mathfrak{X}} \to \boldsymbol{\mathfrak{X}}$ are naturally isomorphic to the respective identity functors $\mathrm{id}_{\boldsymbol{\mathcal{A}}}\colon \boldsymbol{\mathcal{A}} \to \boldsymbol{\mathcal{A}}$ and $\mathrm{id}_{\boldsymbol{\mathfrak{X}}}\colon \boldsymbol{\mathfrak{X}} \to \boldsymbol{\mathfrak{X}}$. At the object level, this implies that each object $\mathbf{B}$ in $\boldsymbol{\mathcal{A}}$ is isomorphic to its double dual $ED(\mathbf{B})$ (and symmetrically for objects in $\boldsymbol{\mathfrak{X}}$). In particular, the isomorphism $\mathbf{B} \cong ED(\mathbf{B})$ gives a representation of $\mathbf{B}$ as an algebra of $\boldsymbol{\mathfrak{X}}$-morphisms with operations defined pointwise.

To see how we might obtain an isomorphism $\mathbf{B} \to ED(\mathbf{B})$, let us unpack the definition of $ED(\mathbf{B})$. First, $ED(\mathbf{B})$ is defined as the set of all $\boldsymbol{\mathfrak{X}}$-morphisms $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$. Thus, we must associate each $b \in B$ with an $\boldsymbol{\mathfrak{X}}$-morphism $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$. Now, $D(\mathbf{B})$ is a substructure of $\underset{\sim}{\mathbf{A}}^B$, so an $\boldsymbol{\mathfrak{X}}$-morphism $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$ is a morphism from a substructure of $\underset{\sim}{\mathbf{A}}^B$ into $\underset{\sim}{\mathbf{A}}$.

Our definition of products in $\boldsymbol{\mathfrak{X}}$ ensures that the projection maps $\underset{\sim}{\mathbf{A}}^B \to \underset{\sim}{\mathbf{A}}$ are morphisms, as are the restricted projections $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$. Thus, given $b \in B$, the most natural morphism $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$ that can be associated with $b$ is the restricted $b$th projection map $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$. This is how we will define a mapping $B \to ED(\mathbf{B})$, which, it turns out, is always an embedding.

The same idea will work in the other direction, giving us an embedding $\mathbf{X} \to DE(\mathbf{X})$ for each object $\mathbf{X}$ in $\boldsymbol{\mathfrak{X}}$. Making sense of this requires us to define embeddings in $\boldsymbol{\mathfrak{X}}$ (though we will not be working with them at all). An $\boldsymbol{\mathfrak{X}}$-morphism $\varphi\colon \mathbf{X} \to \mathbf{Y}$ is called an **embedding** if $\varphi(X)$ forms a substructure $\varphi(\mathbf{X})$ of $\mathbf{Y}$ and $\varphi$ is an isomorphism $\mathbf{X} \to \varphi(\mathbf{X})$.

**Proposition 4.3.** *Let* $\mathbf{A}$ *be a finite algebra, let* $\underset{\sim}{\mathbf{A}}$ *be an alter ego of* $\mathbf{A}$*, and define* $\boldsymbol{\mathcal{A}}$*,* $\boldsymbol{\mathfrak{X}}$*,* $D$*, and* $E$ *as in Proposition 4.2. For each object* $\mathbf{B}$ *in* $\boldsymbol{\mathcal{A}}$ *and each object* $\mathbf{X}$ *in* $\boldsymbol{\mathfrak{X}}$*, define*

$$e_{\mathbf{B}} \colon \mathbf{B} \to ED(\mathbf{B}), \qquad \varepsilon_{\mathbf{X}} \colon \mathbf{X} \to DE(\mathbf{X}),$$

$$(\forall b \in B)(\forall x \in D(\mathbf{B}))\ \ e_{\mathbf{B}}(b)(x) := x(b), \qquad (\forall x \in X)(\forall \alpha \in E(\mathbf{X}))\ \ \varepsilon_{\mathbf{X}}(x)(\alpha) := \alpha(x).$$

*Then* $e_{\mathbf{B}}$ *and* $\varepsilon_{\mathbf{X}}$ *are embeddings and* $e\colon \mathrm{id}_{\boldsymbol{\mathcal{A}}} \to ED$ *and* $\varepsilon\colon \mathrm{id}_{\boldsymbol{\mathfrak{X}}} \to DE$ *are natural transformations.*

Proof. See Clark and Davey [**5**, Theorem 1.5.3]. □

From the results presented so far, we see that we get 'most' of the way to a dual equivalence simply by choosing $\underset{\sim}{\mathbf{A}}$ to be an alter ego of $\mathbf{A}$. All we require for a duality is that the morphisms $e_{\mathbf{B}}$ and $\varepsilon_{\mathbf{X}}$ are surjective for all $\mathbf{B}$ and $\mathbf{X}$. Of course, the surjectivity of every $e_{\mathbf{B}}$ and $\varepsilon_{\mathbf{X}}$ is not guaranteed by assuming only that $\underset{\sim}{\mathbf{A}}$ is an alter ego of $\mathbf{A}$; that would be too easy. All of the work in establishing a dual equivalence lies in choosing an appropriate alter ego $\underset{\sim}{\mathbf{A}}$ so that the maps $e_{\mathbf{B}}$ and $\varepsilon_{\mathbf{X}}$ are always surjective.

Finally, we come to the definition of dualisability. Let $\mathbf{A}$ be a finite algebra. If $\underset{\sim}{\mathbf{A}}$ is an alter ego of $\mathbf{A}$, we say that $\underset{\sim}{\mathbf{A}}$ **dualises** $\mathbf{A}$ if $e_{\mathbf{B}} \colon \mathbf{B} \to ED(\mathbf{B})$ (as defined in Proposition 4.3) is surjective for every object $\mathbf{B} \in \mathbb{SP}(\mathbf{A})$. If there is some choice of alter ego $\underset{\sim}{\mathbf{A}}$ that dualises $\mathbf{A}$, then we say that $\mathbf{A}$ is (**naturally**) **dualisable**.

If $\underset{\sim}{\mathbf{A}}$ dualises $\mathbf{A}$, we do not necessarily have a dual equivalence between $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathfrak{X}}$, but we still have a representation for the members of $\boldsymbol{\mathcal{A}}$. If there is some choice of $\underset{\sim}{\mathbf{A}}$ dualising $\mathbf{A}$ with the further property that $\varepsilon_{\mathbf{X}} \colon \mathbf{X} \to DE(\mathbf{X})$ is surjective for all objects $\mathbf{X}$ in $\boldsymbol{\mathfrak{X}}$, then we say that $\mathbf{A}$ is **fully dualisable**. The property of full dualisability is certainly well studied, but we will not consider it any further in this thesis.

The aim of this thesis, then, is to determine which finite aperiodic semigroups are dualisable. To this end, we will now record the general results that will be used to establish dualisability and non-dualisability.

Showing that $\underset{\sim}{\mathbf{A}}$ dualises $\mathbf{A}$ amounts to showing that, for every $\mathbf{B} \in \mathbb{SP}(\mathbf{A})$, every morphism $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$ is equal to $e_{\mathbf{B}}(b)$ for some $b \in B$ (which is really just a restriction of the $b$th projection map $\underset{\sim}{\mathbf{A}}^B \to \underset{\sim}{\mathbf{A}}$). This does not require any further consideration of category-theoretic concepts; Propositions 4.2 and 4.3 take care of all of the category theory for us. However, we may still need to consider topology, since a morphism $D(\mathbf{B}) \to \underset{\sim}{\mathbf{A}}$ is in particular a continuous map. But since all finite structures in $\boldsymbol{\mathfrak{X}}$ are discrete, topology comes into play only when $\mathbf{B}$ is infinite. For these reasons, one of the early endeavours of duality theorists was to find conditions on $\underset{\sim}{\mathbf{A}}$ so that consideration of finite $\mathbf{B}$ is sufficient to determine dualisability, thus eliminating the need to use topology. These efforts resulted in the next theorem.

Let $\underset{\sim}{\mathbf{A}}$ be an alter ego of the finite algebra $\mathbf{A}$. We say that $\underset{\sim}{\mathbf{A}}$ **dualises** $\mathbf{A}$ **at the finite level** if $e_{\mathbf{B}} \colon \mathbf{B} \to ED(\mathbf{B})$ is surjective for every *finite* $\mathbf{B} \in \mathbb{SP}(\mathbf{A})$. The following result was obtained independently by Willard and Zádori. Willard's proof may be found in [**5**, Theorem 2.2.11], [**67**, Theorem 1.3], or [**7**, Theorem 4.3], while Zádori's proof may be found in [**68**, Corollary 3.5] or in [**5**, Theorem 10.6.4].

**Theorem 4.4** (Duality Compactness Theorem)**.** *Let* $\mathbf{A}$ *be a finite algebra, and let* $\underset{\sim}{\mathbf{A}}$ *be an alter ego of* $\mathbf{A}$ *of finite type. If* $\underset{\sim}{\mathbf{A}}$ *dualises* $\mathbf{A}$ *at the finite level, then* $\underset{\sim}{\mathbf{A}}$ *dualises* $\mathbf{A}$*.*

All of our dualisability results will effectively be established using the Duality Compactness Theorem, so we will not need to consider topology. Our proofs will proceed by establishing the following **interpolation condition** on $\underset{\sim}{\mathbf{A}}$:

(IC) *For every* $n \in \omega \backslash \{0\}$ *and every substructure* $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$, *every morphism* $\varphi \colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ *extends to an* $n$*-ary term function of* $\mathbf{A}$.

The following corollary of the Duality Compactness Theorem (see [**5**, Corollary 2.2.12]) allows us to establish dualities via (IC).

**Theorem 4.5** ((IC) Duality Theorem)**.** *Let* $\mathbf{A}$ *be a finite algebra, and let* $\underset{\sim}{\mathbf{A}}$ *be an alter ego of* $\mathbf{A}$ *of finite type satisfying* (IC) *with respect to* $\mathbf{A}$*. Then* $\underset{\sim}{\mathbf{A}}$ *dualises* $\mathbf{A}$*.*

In view of Theorem 4.5, we say that $\mathbf{A}$ is **dualisable via (IC)** if $\mathbf{A}$ has an alter ago $\underset{\sim}{\mathbf{A}}$ of finite type satisfying (IC) with respect to $\mathbf{A}$.

Once we have established the dualisability of $\mathbf{A}$, the following theorem allows us to deduce that every finite algebra generating the same quasivariety as $\mathbf{A}$ is also dualisable. This result was proved independently by Davey and Willard [**22**] and Saramago [**63**].

**Theorem 4.6** (Independence of Generator Theorem)**.** *Let* $\mathbf{A}$ *and* $\mathbf{A}'$ *be finite algebras such that* $\mathbb{SP}(\mathbf{A}) = \mathbb{SP}(\mathbf{A}')$*. Then* $\mathbf{A}$ *is dualisable if and only if* $\mathbf{A}'$ *is dualisable.*

In [**1**, Chapter 5], AlDhamri proves the following modified version of Theorem 4.6.

**Theorem 4.7** ((IC) Independence of Generator Theorem)**.** *Let* $\mathbf{A}$ *and* $\mathbf{A}'$ *be finite algebras with* $\mathbb{SP}(\mathbf{A}) = \mathbb{SP}(\mathbf{A}')$*. Then* $\mathbf{A}$ *is dualisable via* (IC) *if and only if* $\mathbf{A}'$ *is dualisable via* (IC)*.*

Now, we also require results to establish non-dualisability. In all such cases, we will be establishing a stronger property. We say that the finite algebra $\mathbf{A}$ is **inherently non-dualisable** (**IND**) if $\mathbf{A}$ does not lie in $\mathbb{SP}(\mathbf{B})$ for any dualisable finite algebra $\mathbf{B}$. Theorem 4.6 implies that $\mathbf{A}$ is IND if and only if $\mathbf{A}$ does not embed into any dualisable finite algebra. To see this, observe that if $\mathbf{A} \in \mathbb{SP}(\mathbf{B})$ for some dualisable $\mathbf{B}$, then $\mathbf{A}$ (if non-trivial) embeds into $\mathbf{B}^n$ for some $n \geqslant 1$, which is dualisable by Theorem 4.6 as $\mathbb{SP}(\mathbf{B}) = \mathbb{SP}(\mathbf{B}^n)$.

The following result was first published by Davey, Idziak, Lampe, and McNulty in [**12**, Theorem 3] (see also [**5**, Theorem 10.5.5]). To date, all known examples of IND algebras can be shown to be IND using this result.

**Theorem 4.8** (Inherent Non-Dualisability Theorem)**.** *Let* $\mathbf{A}$ *be a finite algebra. Suppose there is a set* $Z$*, a subalgebra* $\mathbf{B} \leqslant \mathbf{A}^Z$*, an infinite subset* $B_0 \subseteq B$*, and a function* $\varphi \colon \mathbb{N} \to \mathbb{N}$ *such that the following hold:*

(i) *for every* $n \in \mathbb{N}$ *and every congruence* $\theta$ *on* $\mathbf{B}$ *with* $|B/\theta| \leqslant n$*, there is a unique block of* $\theta{\restriction}_{B_0}$ *with more than* $\varphi(n)$ *elements;*

(ii) $\mathbf{B}$ *does not contain the element* $g \in A^Z$ *defined by* $g(i) := \rho_i(b_i)$*, where* $\rho_i \colon \mathbf{B} \to \mathbf{A}$ *is the ith projection and* $b_i$ *is any element of the infinite block of* $\ker(\rho_i){\restriction}_{B_0}$*.*

*Then* $\mathbf{A}$ *is inherently non-dualisable.*

Our final result in this section is Theorem 4.10 below, due to Jackson [**40**, Theorem 3.1]. This rather innovative theorem, making use of projective planes, was originally developed to prove that a finite nilpotent semigroup of class 3 is inherently non-dualisable. In [**40**], Jackson showed that many of the non-dualisability results for semigroups and groups that were known at the time can be obtained from Theorem 4.10, and also used it to prove that $\mathbf{M}_p$ is inherently non-dualisable for all primes $p$. We will later use Theorem 4.10 to obtain the new results that $\mathbf{L}^+$ and $\mathbf{R}\circlearrowright 2$ are inherently non-dualisable.

Recall that a projective plane is a triple $\mathsf{P} = (\mathsf{P}, \mathsf{L}, \mathsf{I})$, where $\mathsf{P}$, $\mathsf{L}$ are sets and $\mathsf{I} \subseteq \mathsf{P} \times \mathsf{L}$, satisfying certain conditions (to be stated shortly). The elements of $\mathsf{P}$ and $\mathsf{L}$ are respectively called **points** and **lines**, and $\mathsf{I}$ is called the **incidence relation**. If $p \in \mathsf{P}$ and $\ell \in \mathsf{L}$, the relation $p \mathrel{\mathsf{I}} \ell$ can be read as "the point $p$ lies on the line $\ell$", or "the line $\ell$ contains the point $p$", or "$p$ is incident to $\ell$", or "$\ell$ is incident to $p$". Now, $\mathsf{P} = (\mathsf{P}, \mathsf{L}, \mathsf{I})$ is called a **projective plane** if the following three conditions hold:

- for every distinct pair $p, q \in \mathsf{P}$, there exactly one line in $\mathsf{L}$ incident to $p$ and $q$;
- for every distinct pair $\ell, k \in \mathsf{L}$, there is exactly one point in $\mathsf{P}$ incident to $\ell$ and $k$;
- there exist pairwise distinct $p, q, r, s \in \mathsf{P}$ such that every line in $\mathsf{L}$ is incident to at most two of the points $p$, $q$, $r$, $s$.

Although these axioms are used in the proof of Theorem 4.10, the only property of projective planes that we will use in our applications is the fact that every pair of lines is incident to a common point. (Also, we will need the fact that infinite projective planes actually exist; the real projective plane is one such example.)

Since a line is determined by the set of points incident to it, we can think of lines as subsets of $\mathsf{P}$, so that the incidence relation $\mathsf{I}$ coincides with $\in$. In the statement and use of the next theorem, it will be important for us to think of lines as subsets of $\mathsf{P}$ in this way. For the statement, we also require some notation, which we will use in applications of Theorems 4.8 and 4.10. We used this notation in Theorems 3.21 and 3.29 and Lemma 3.52.

**Notation 4.9.** Let $A$ and $Z$ be sets; typically, $A$ will be the underlying set of a finite algebra. For each $a \in A$, we denote by $\underline{a}$ the constant map in $A^Z$ with value $a$. Now, let $I_1, \dots, I_n$ be non-empty pairwise disjoint subsets of $Z$, and let $a, b_1, \dots, b_n \in A$ (not necessarily distinct). We denote by $a^{b_1,\dots,b_n}_{I_1,\dots,I_n}$ the element of $A^Z$ given by

$$(\forall i \in Z)\ \ a^{b_1,\dots,b_n}_{I_1,\dots,I_n}(i) := \begin{cases} b_k & \text{if } i \in I_k, \\ a & \text{otherwise.} \end{cases}$$

If $I_k$ is a singleton set $\{j\}$, we will write $j$ in the subscript rather than $\{j\}$.

**Theorem 4.10** (Projective Plane Theorem)**.** *Let $\mathbf{A}$ be a finite algebra with a binary term function $\cdot$ and elements $\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}, \mathsf{f} \in A$ satisfying the following table of products:*

| $\cdot$ | c | d |
|---|---|---|
| a | e | f |
| b | f | f |

*Let $\mathsf{P} = (\mathsf{P}, \mathsf{L}, \mathsf{I})$ be an infinite projective plane, and let $\mathsf{P}_\infty := \mathsf{P} \mathbin{\dot\cup} \{\infty\}$. If the subalgebra of $\mathbf{A}^{\mathsf{P}_\infty}$ generated by the set $\{\mathsf{b}^{\mathsf{a},\mathsf{a}}_{\ell,\infty}, \mathsf{d}^{\mathsf{c},\mathsf{c}}_{\ell,\infty} \mid \ell \in \mathsf{L}\}$ does not contain $\mathsf{f}^{\mathsf{e}}_{\infty}$, then $\mathbf{A}$ is inherently non-dualisable.*

## 4.2. The known dualisability results for semigroups

The dualisability problem for finite semigroups was first explicitly addressed in [**39**], where Jackson collected the known dualisability results for finite semigroups and proved several non-dualisability results. One of the main results in [**39**] characterised the IND bands, showing that a finite band is normal if and only if it is not IND (with the hope that 'not IND' could eventually be replaced by 'dualisable'). One direction of this equivalence was established by showing that $\mathbf{L}^1$ and $\mathbf{R}^1$ are IND [**39**, Proposition 4].

**Proposition 4.11.** *The semigroups* $\mathbf{L}^1$ *and* $\mathbf{R}^1$ *are inherently non-dualisable.*

By Theorem 3.27, this immediately implies that every finite non-normal band is IND. In [**39**], Jackson also considered finite monoids and finite inverse semigroups, showing that all such semigroups that generate residually large varieties are IND.

The next major dualisability result for semigroups came when AlDhamri completed the proof that all finite normal bands are dualisable [**1**] [**2**], strengthening Jackson's result that they are not IND. We give a new proof that all finite normal bands are dualisable in Section 4.5.

Jackson returned to the dualisability problem for semigroups in [**40**]. The main result of this paper was the Projective Plane Theorem 4.10, which was used to derive several new and old non-dualisability results. One of the re-proved results was Quackenbush and Szabó's result that any finite non-Abelian nilpotent group is IND [**57**]. Among the new results were the two following theorems (respectively Theorem 5.1 and Theorem 7.6 in [**40**]).

**Theorem 4.12.** *Let* $\mathbf{S}$ *be a finite semigroup. If* $\mathbf{N}_4 \in \mathbb{V}(\mathbf{S})$*, then* $\mathbf{S}$ *is inherently non-dualisable.*

**Theorem 4.13.** *Let* $\mathbf{S}$ *be a finite semigroup. If* $\mathbf{M}_p \in \mathbb{V}(\mathbf{S})$ *for some prime* $p$*, then* $\mathbf{S}$ *is inherently non-dualisable.*

In [**40**, Theorem 8.1(2)], Jackson extended Proposition 4.11 to show that a finite semigroup $\mathbf{S}$ is inherently non-dualisable whenever $\mathbb{V}(\mathbf{S})$ contains $\mathbf{L}^1$ or $\mathbf{R}^1$. The essence of the proof is the following lemma, which we record for later use.

**Lemma 4.14.** *Let* $\mathbf{S}$ *be a finite semigroup, and let* $\varphi\colon \mathbf{S} \to \mathbf{R}^1$ *be an onto homomorphism. Then there is an embedding* $\psi\colon \mathbf{R}^1 \to \mathbf{S}$ *such that* $\varphi \circ \psi$ *is the identity on* $\mathbf{R}^1$*.*

**Theorem 4.15.** *Let* $\mathbf{S}$ *be a finite semigroup. If* $\mathbb{V}(\mathbf{S})$ *contains* $\mathbf{L}^1$ *or* $\mathbf{R}^1$*, then* $\mathbf{S}$ *is inherently non-dualisable.*

The results presented in this rather short section constitute all of the previously-known IND results for semigroups. We remark that all of these semigroups were known to generate residually large varieties; see Section 3.8.

In the other direction, Kearnes and Szendrei proved the substantial result that all A-groups are dualisable in [**41**, Corollary 6.14]; this is a consequence of their much more general main result, though the only semigroups to which it can be applied are groups. Apart from normal bands, A-groups, and some simple combinations of these (see [**1**, Corollary 4.2.7, Chapter 6]), there were relatively few dualisability results for finite semigroups prior to the present work; thus, the dualisability problem for semigroups was ripe for exploration.

## 4.3. Quasicriticals

Let $\mathcal{K}$ be a class of finite algebras, and suppose we wish to show that all members of $\mathcal{K}$ are dualisable. Rather than giving a completely general proof, it is often easier to determine all quasivarieties that can be generated by members of $\mathcal{K}$, and then show that each of these quasivarieties has at least one dualisable generator. By the Independence of Generator Theorem 4.6, this suffices to show that all members of $\mathcal{K}$ are dualisable. In this section, we collect some basic facts about quasivarieties so that we may take this approach to dualisability.

Let $\mathcal{Q}$ be a quasivariety. Our goal is to determine all subquasivarieties of $\mathcal{Q}$ by studying the lattice of subquasivarieties of $\mathcal{Q}$, which we will denote by $\mathrm{L_q}(\mathcal{Q})$. First, we will use the fact that all subquasivarieties can be decomposed into completely join irreducibles according to the following result.

**Theorem 4.16.** *Let $\mathcal{Q}$ be a quasivariety. Then every quasivariety in $\mathrm{L_q}(\mathcal{Q})$ is a join of completely join irreducible quasivarieties in $\mathrm{L_q}(\mathcal{Q})$.*

Proof. Any subquasivariety lattice is dually algebraic. (For a proof, see [**4**, Theorem 5.22]; the proof there is carried out in the more general setting of elementary classes, but the proof in the restricted setting of quasivarieties is identical.) The result therefore follows from Lemma 1.3. $\square$

Thus, to determine the subquasivarieties of $\mathcal{Q}$, it will be extremely useful to describe the completely join irreducibles of $\mathrm{L_q}(\mathcal{Q})$. In the general case, this is quite difficult, but it will suffice for our purposes to consider locally finite quasivarieties. The next definition is the key to describing completely join irreducible subquasivarieties in the locally finite case.

An algebra is called **quasicritical** if it is finite and it is not in the quasivariety generated by the set of its proper subalgebras. If $\mathcal{K}$ is a class of algebras, we denote by $\mathbf{qc}(\mathcal{K})$ the class of all quasicritical algebras in $\mathcal{K}$.

We will often use the following formulation of quasicriticality. Let $\mathbf{A}$ be an algebra, and let $c, d \in A$ with $c \neq d$. We say that $(c, d)$ is a **critical pair of A** if for every $\mathbf{B} \in \mathbb{S}(\mathbf{A})$ such that there is homomorphism $\mathbf{A} \to \mathbf{B}$ separating $c$ and $d$, we have $\mathbf{A} \cong \mathbf{B}$. It is easily shown that a finite algebra is quasicritical if and only if it has a critical pair.

The following result and its corollary explain the role of quasicriticals in determining the subquasivarieties of $\mathcal{Q}$.

**Theorem 4.17.** *Let $\mathcal{Q}$ be a locally finite quasivariety, and let $\mathcal{K} \in \mathrm{L_q}(\mathcal{Q})$. If $\mathcal{K}$ is completely join irreducible in $\mathrm{L_q}(\mathcal{Q})$, then $\mathcal{K} = \mathbb{SP}(\mathbf{A})$ for some $\mathbf{A} \in \mathbf{qc}(\mathcal{Q})$.*

Proof. See Hyndman and Nation [**38**, Theorem 2.8]. It is also shown that the converse holds if $\mathcal{Q}$ has finite type, but we will not need this fact. $\square$

**Corollary 4.18.** *Let $\mathcal{Q}$ be a locally finite quasivariety. Then every subquasivariety of $\mathcal{Q}$ is generated by a set of quasicritical algebras from $\mathcal{Q}$. Consequently, if $\mathcal{Q}$ has finitely many quasicriticals up to isomorphism, then $\mathcal{Q}$ has finitely many subquasivarieties.*

Proof. Combine Theorems 4.16 and 4.17. $\square$

We will close this section with a simple method of constructing subquasivariety lattices. The next theorem, due to the author, shows that certain subquasivariety lattices can be naturally represented as down-set lattices (see Section 1.1).

**Theorem 4.19.** *Let $\mathbf{\mathcal{Q}}$ be a locally finite quasivariety with finitely many quasicriticals up to isomorphism, and let $\mathbf{\mathcal{S}}$ be a set of representatives of the isomorphism classes of $\mathbf{qc}(\mathbf{\mathcal{Q}})$. Assume that every member of $\mathbf{\mathcal{S}}$ is subdirectly irreducible, and order $\mathbf{\mathcal{S}}$ by containment up to isomorphism. Then the mapping $\mathcal{O}(\mathbf{\mathcal{S}}) \to \mathrm{L_q}(\mathbf{\mathcal{Q}})$ given by $\mathbf{\mathcal{U}} \mapsto \mathbb{SP}(\mathbf{\mathcal{U}})$ is a lattice isomorphism. In particular, $\mathrm{L_q}(\mathbf{\mathcal{Q}})$ is a distributive lattice.*

Proof. By Corollary 4.18, every subquasivariety of $\mathbf{\mathcal{Q}}$ is generated by a subset of $\mathbf{\mathcal{S}}$. But since a quasivariety is closed under $\mathbb{S}$, every subquasivariety of $\mathbf{\mathcal{Q}}$ is of the form $\mathbb{SP}(\mathbf{\mathcal{U}})$ for some down-set $\mathbf{\mathcal{U}}$ of $\mathbf{\mathcal{S}}$. This shows that the mapping in the statement is onto.

Let $\mathbf{\mathcal{U}}$ and $\mathbf{\mathcal{V}}$ be down-sets of $\mathbf{\mathcal{S}}$. Clearly $\mathbf{\mathcal{U}} \subseteq \mathbf{\mathcal{V}}$ implies $\mathbb{SP}(\mathbf{\mathcal{U}}) \subseteq \mathbb{SP}(\mathbf{\mathcal{V}})$. Conversely, assume that $\mathbb{SP}(\mathbf{\mathcal{U}}) \subseteq \mathbb{SP}(\mathbf{\mathcal{V}})$, and let $\mathbf{A} \in \mathbf{\mathcal{U}}$. Then $\mathbf{A} \in \mathbb{SP}(\mathbf{\mathcal{U}}) \subseteq \mathbb{SP}(\mathbf{\mathcal{V}})$, so because $\mathbf{A}$ is subdirectly irreducible, we have $\mathbf{A} \in \mathbb{S}(\mathbf{\mathcal{V}})$. Thus $\mathbf{A} \in \mathbf{\mathcal{V}}$ as $\mathbf{\mathcal{V}}$ is a down-set, which shows that $\mathbf{\mathcal{U}} \subseteq \mathbf{\mathcal{V}}$. The mapping $\mathbf{\mathcal{U}} \mapsto \mathbb{SP}(\mathbf{\mathcal{U}})$ is therefore an order-isomorphism $\mathcal{O}(\mathbf{\mathcal{S}}) \to \mathrm{L_q}(\mathbf{\mathcal{Q}})$, so the result follows from Proposition 1.2. □

## 4.4. Quasivarieties of normal bands

The remainder of this chapter will be devoted to re-proving AlDhamri's result that all finite normal bands are dualisable [**2**]. In this section, we will use Theorem 4.19 to derive Gerhard's description of the lattice of quasivarieties of normal bands [**31**]. Our approach will be completely different from the original, and the results proved along the way will be used in our dualisability proofs in Section 4.5.

In this section, we introduce the important concept of a retraction. Let $\mathbf{A}$ and $\mathbf{B}$ be algebras with $\mathbf{A} \leqslant \mathbf{B}$. A homomorphism $\varphi\colon \mathbf{B} \to \mathbf{A}$ is called a **retraction of B onto A** if $\varphi$ is surjective and $\varphi(x) = x$ for all $x \in A$. We write $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{B}$, and say **B retracts onto A**, if there exists a retraction of $\mathbf{B}$ onto $\mathbf{A}$. We note that there are other definitions in standard use; our definition is, of course, the most appropriate for our purposes.

We have encountered retractions several times in Chapter 3. If $\mathbf{S}$ is a semigroup of exponent $n$ with $\mathbf{N}_4, \mathbf{L}^+, \mathbf{R}^+ \notin \mathbb{V}(\mathbf{S})$, then the mapping $x \mapsto x^{n+1}$ is a retraction of $\mathbf{S}$ onto $\mathrm{G}(\mathbf{S})$ (Lemma 3.22(ii), Theorem 3.23). Also, in the situation of Lemma 3.30, the left translations by idempotents are retractions (onto their images).

To obtain the lattice of quasivarieties of normal bands, we can directly apply Theorem 4.19 once we know the quasicritical normal bands are subdirectly irreducible. We will prove this in Theorem 4.25 using the following five lemmas.

**Lemma 4.20.** *Let $\mathbf{S}$ be a rectangular band. Then $\mathbf{S}$ retracts onto every right-zero subsemigroup of $\mathbf{S}$.*

Proof. Let $\mathbf{A} \leqslant \mathbf{S}$ be a right-zero semigroup. Then $A$ is contained in an $\mathcal{R}$-class of $\mathbf{S}$, which forms a right-zero subsemigroup $\mathbf{B}$ of $\mathbf{S}$ by Proposition 2.29. Fix any $b \in B$, and let $\lambda\colon S \to S$ denote left translation by $b$. By Lemma 3.30, $\lambda$ is an endomorphism of $\mathbf{S}$, and by Lemma 2.24, the image of $\lambda$ is contained in $R_b = B$. Moreover, $\lambda$ fixes every point of $B$;

thus, $\lambda$ is a retraction of $\mathbf{S}$ onto $\mathbf{B}$. Now, since any map between right-zero semigroups is a homomorphism, it is trivial that $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{B}$, so $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$ by transitivity. □

**Lemma 4.21.** *Let $\mathbf{S}$ be a band. If $\mathbf{S}$ is not a rectangular band, then $\mathbf{I}$ embeds into $\mathbf{S}$.*

Proof. Assume that $\mathbf{I}$ does not embed into $\mathbf{S}$, and let $x, y \in S$. Then the subsemigroup of $\mathbf{S}$ on $\{xyx, x\}$ is semilattice, so $xyx = x$. Thus, $\mathbf{S}$ is a rectangular band. □

**Lemma 4.22.** *Let $\mathbf{S}$ be a normal band, and let $\mathbf{A} \leqslant \mathbf{S}$ with $\mathbf{A} \cong \mathbf{I}$. Then $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$.*

Proof. Write $A = \{a, b\}$. By Theorem 2.47, the $\mathcal{J}$-classes of $\mathbf{S}$ are rectangular bands, so $A$ is not contained in any $\mathcal{J}$-class; thus, $J_a \neq J_b$. Now, by Theorem 1.12, there is a homomorphism $\mathbf{S}/\mathcal{J} \to \mathbf{A}$ separating $J_a$ and $J_b$, and composing with the quotient map $\mathbf{S} \to \mathbf{S}/\mathcal{J}$ gives an onto homomorphism $r\colon \mathbf{S} \to \mathbf{A}$ separating $a$ and $b$. Since $r{\restriction}_A\colon \mathbf{A} \to \mathbf{A}$ preserves the semilattice order, it is the identity map on $A$, so $r$ is a retraction onto $\mathbf{A}$. □

**Lemma 4.23.** *Let $\mathbf{S}$ be a normal band, and let $\mathbf{A} \leqslant \mathbf{S}$ with $\mathbf{A} \cong \mathbf{L}^0$. Then $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$.*

Proof. Let $\{a, b\}$ be the non-trivial $\mathcal{L}$-class of $\mathbf{A}$. We have $\mathbf{S} \in \mathbb{SP}(\mathbf{R}^0, \mathbf{L}^0) = \mathbb{SP}(\mathbf{R}^0, \mathbf{A})$ by Theorem 3.35, but every homomorphism $\mathbf{S} \to \mathbf{R}^0$ identifies $a$ and $b$ since $\mathbf{R}^0$ has a trivial $\mathcal{L}$ relation, so there is a homomorphism $r\colon \mathbf{S} \to \mathbf{A}$ with $r(a) \neq r(b)$. Since $r(a) \mathrel{\mathcal{L}} r(b)$, we must have $\{r(a), r(b)\} = \{a, b\}$; thus, we can assume by symmetry that $r(a) = a$ and $r(b) = b$. Now, $r{\restriction}_A$ is an endomorphism of $\mathbf{A}$ separating $a$ and $b$, so it must be an automorphism because $(a, b)$ generates the monolith of $\mathbf{A}$. Hence, $r{\restriction}_A$ is the identity on $A$. □

**Lemma 4.24.** *Let $\mathbf{S}$ be a finite normal band, and let $\mathbf{A} \leqslant \mathbf{S}$ with $\mathbf{A} \cong \mathbf{L}$. If $\mathbf{L}^0$ does not embed into $\mathbf{S}$, then $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$.*

Proof. Write $A = \{a, b\}$. Let $M$ be the minimum ideal of $\mathbf{S}$ (Theorem 2.4), let $w \in M$, and let $\mathbf{T}$ be the subsemigroup of $\mathbf{S}$ on $T := Sw$. Then $T \subseteq M$. Since the subsemigroup on $M$ is a rectangular band by Theorem 2.47, $\mathbf{T}$ is also a rectangular band.

Now, we have $wa = waba = wbaa = wb$. If we also had $aw = bw$, then $z := awa = bwb$ would be a zero for $\{a, b\}$, contradicting the fact that $\mathbf{L}^0$ does not embed into $\mathbf{S}$. We must therefore have $aw \neq bw$.

Let $c := aw$ and $d := bw$. By Lemma 3.30, the translation $x \mapsto xw$ is a retraction of $\mathbf{S}$ onto $\mathbf{T}$, so $\{c, d\}$ is a homomorphic image of $\mathbf{A}$ and is therefore a left-zero subsemigroup of $\mathbf{T}$. Now, by Lemma 4.20, there is a retraction of $\mathbf{T}$ onto $\{c, d\}$. Thus, there is a retraction, $r$, of $\mathbf{S}$ onto $\{c, d\}$ with $r(a) = c$ and $r(b) = d$. Now let $f$ be the bijection $\{c, d\} \to \{a, b\}$ with $f(c) = a$. Then $f \circ r$ is a retraction of $\mathbf{S}$ onto $\mathbf{A}$. □

Now we can describe the quasicritical normal bands. The proof is based on that of [**60**, Lemma 3.2], where Sapir describes the quasicriticals in a more general class.

**Theorem 4.25.** *Let $\mathbf{S}$ be a quasicritical normal band. Then $\mathbf{S}$ is isomorphic to a member of $\{\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0, \mathbf{I}\}$, and so $\mathbf{S}$ is subdirectly irreducible.*

Proof. Let $(c, d)$ be a critical pair of $\mathbf{S}$. First, assume that $c \mathrel{\not\mathcal{J}} d$. Then, by Theorem 2.47 and Theorem 1.12, there is a homomorphism $\mathbf{S}/\mathcal{J} \to \mathbf{I}$ separating $J_c$ and $J_d$, so composing with the quotient map $\mathbf{S} \to \mathbf{S}/\mathcal{J}$ gives a homomorphism $\mathbf{S} \to \mathbf{I}$ separating $c$ and $d$. But $\mathbf{S}$ has more than one $\mathcal{J}$-class and is therefore not rectangular, so $\mathbf{I} \in \mathbb{S}(\mathbf{S})$ by Lemma 4.21, and hence $\mathbf{S} \cong \mathbf{I}$ (by the choice of $(c, d)$).

Now assume that $c \mathrel{\mathcal{J}} d$. Let $\mathbf{J}$ be the subsemigroup on $J_c$, which is a rectangular band by Theorem 2.47. Since $\mathcal{H}^{\mathbf{J}} = \Delta_J$, we may assume by symmetry that $c \mathrel{\not\mathcal{L}^{\mathbf{J}}} d$, so by Proposition 2.29, we have $cd \neq c = cc$. By Lemma 3.33, $c$ is a left identity for $\mathbf{S}$, and by symmetry, $d$ is also a left identity. In particular, $\{c, d\}$ is a right-zero subsemigroup of $\mathbf{S}$. Now, if $\mathbf{R}^0$ does not embed into $\mathbf{S}$, then by Lemma 4.24, there is a retraction of $\mathbf{S}$ onto $\{c, d\}$, and so $\mathbf{S} \cong \mathbf{R}$. Assume that $\mathbf{R}^0$ embeds into $\mathbf{S}$. Since $c \mathrel{\mathcal{R}} d$, every homomorphism $\mathbf{S} \to \mathbf{L}^0$ identifies $c$ and $d$, so by Theorem 3.35, there is a homomorphism $\mathbf{S} \to \mathbf{R}^0$ separating $c$ and $d$. Since $\mathbf{R}^0$ embeds into $\mathbf{S}$, this gives $\mathbf{S} \cong \mathbf{R}^0$. □

Now that we have described the quasicritical normal bands, we can apply Theorem 4.19. This requires us to order the quasicriticals by embedding; the resulting ordered set is shown in Figure 4.1.

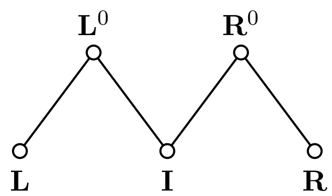



FIGURE 4.1. The quasicritical normal bands ordered by embedding.

Each quasivariety of normal bands is generated by a down-set of the ordered set in Figure 4.1. Of course, rather than using the entire down-set, we can generate the same quasivariety by taking the maximal elements. For example, the down-set $\{\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{I}\}$ generates the same quasivariety as $\{\mathbf{L}^0, \mathbf{R}\}$.

To tighten the description even further, we can use the fact that $\mathbb{SP}(\mathbf{S}, \mathbf{T}) = \mathbb{SP}(\mathbf{S} \times \mathbf{T})$ for finite semigroups $\mathbf{S}$, $\mathbf{T}$, so each quasivariety can be generated by a single semigroup. The equality $\mathbb{SP}(\mathbf{A}, \mathbf{B}) = \mathbb{SP}(\mathbf{A} \times \mathbf{B})$ does not hold for all algebras $\mathbf{A}$, $\mathbf{B}$, but it is true provided that there exist homomorphisms $\mathbf{A} \to \mathbf{B}$ and $\mathbf{B} \to \mathbf{A}$. This is guaranteed for finite semigroups because all finite semigroups have idempotent elements.

For later use, the following proof that $\mathbb{SP}(\mathbf{S}, \mathbf{T}) = \mathbb{SP}(\mathbf{S} \times \mathbf{T})$ will also show that $\mathbf{S} \times \mathbf{T}$ can be retracted onto subsemigroups isomorphic to the direct factors.

**Lemma 4.26.** *Let* $\mathbf{S}$ *and* $\mathbf{T}$ *be periodic semigroups. Then there are* $\mathbf{S}', \mathbf{T}' \leqslant_{\mathrm{r}} \mathbf{S} \times \mathbf{T}$ *such that* $\mathbf{S} \cong \mathbf{S}'$ *and* $\mathbf{T} \cong \mathbf{T}'$*. Consequently, we have* $\mathbb{SP}(\mathbf{S}) \vee \mathbb{SP}(\mathbf{T}) = \mathbb{SP}(\mathbf{S}, \mathbf{T}) = \mathbb{SP}(\mathbf{S} \times \mathbf{T})$.

PROOF. By Theorem 2.36, $\mathbf{T}$ has an idempotent element, say $e$. Let $\mathbf{S}'$ be the subsemigroup of $\mathbf{S} \times \mathbf{T}$ on $S \times \{e\}$. Evidently $\mathbf{S}' \cong \mathbf{S}$, and the maping $(x, y) \mapsto (x, e)$ is a retraction of $\mathbf{S} \times \mathbf{T}$ onto $\mathbf{S}'$. By symmetry, there is some $\mathbf{T}' \leqslant_{\mathrm{r}} \mathbf{S} \times \mathbf{T}$ with $\mathbf{T} \cong \mathbf{T}'$.

Now, $\mathbb{SP}(\mathbf{S}) \vee \mathbb{SP}(\mathbf{T}) = \mathbb{SP}(\mathbf{S}, \mathbf{T})$ holds in general, as does $\mathbb{SP}(\mathbf{S} \times \mathbf{T}) \subseteq \mathbb{SP}(\mathbf{S}, \mathbf{T})$. We have shown that $\mathbf{S}, \mathbf{T} \in \mathbb{S}(\mathbf{S} \times \mathbf{T})$, so $\mathbb{SP}(\mathbf{S}, \mathbf{T}) \subseteq \mathbb{SP}(\mathbf{S} \times \mathbf{T})$. □

**Theorem 4.27.** *The lattice of quasivarieties of normal bands is the* 13*-element distributive lattice shown in Figure 4.2. The varieties are indicated by filled circles.*

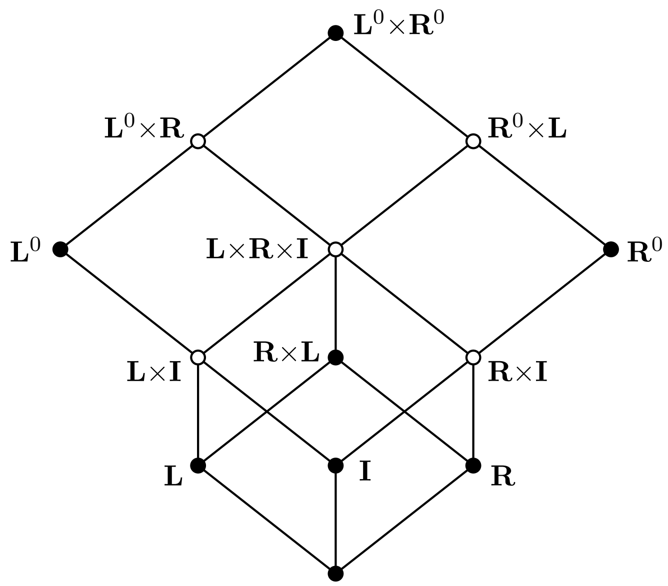



FIGURE 4.2. The lattice of quasivarieties of normal bands, labelled by generators.

PROOF. By Theorem 3.35 and Theorem 1.10, the (quasi)variety of normal bands is locally finite. The lattice of subquasivarieties can be directly constructed from Theorem 4.19 as follows: first, list all down-sets of the ordered set in Figure 4.1, then order the down-sets by containment, and then replace each down-set by the quasivariety it generates. Finally, by Lemma 4.26, each down-set generates the same quasivariety as the direct product of its maximal elements.

Using Corollary 1.7, the varieties can be easily identified. By Theorem 1.12, $\mathbb{SP}(\mathbf{I})$ is the variety of semilattices, and by Theorem 3.35, $\mathbf{L}$ and $\mathbf{R}$ are the only subdirectly irreducible rectangular bands, so $\mathbb{SP}(\mathbf{L})$, $\mathbb{SP}(\mathbf{R})$, and $\mathbb{SP}(\mathbf{L}, \mathbf{R})$ are the varieties of left-zero, right-zero, and rectangular bands, respectively. By Theorem 3.35, the subdirectly irreducible left-normal bands are $\mathbf{I}$, $\mathbf{L}$, $\mathbf{L}^0$ (Section 3.6), so $\mathbb{SP}(\mathbf{L}^0)$ is the variety of left-normal bands. Dually, $\mathbb{SP}(\mathbf{R}^0)$ is a variety. Now, $\mathbf{L}^0 \in \mathbb{H}(\mathbf{I} \times \mathbf{L}) \subseteq \mathbb{V}(\mathbf{I}, \mathbf{L})$, but since $\mathbf{L}^0$ is subdirectly irreducible, we have $\mathbf{L}^0 \notin \mathbb{SP}(\mathbf{I}, \mathbf{L})$; dually, $\mathbf{R}^0 \in \mathbb{V}(\mathbf{I}, \mathbf{R}) \setminus \mathbb{SP}(\mathbf{I}, \mathbf{R})$. Thus, the remaining subquasivarieties are not varieties. □

### 4.5. The dualisability of normal bands revisited

In this section, we will give a new proof that all finite normal bands are dualisable, a result which was first proved by AlDhamri [**2**]. The notation introduced here will be used in Chapter 5, and the techniques built upon, so this section will serve as a warm-up for our dualisability proofs in Chapter 5.

Our proof of the normal band result will also give a stronger statement than the original. We will show that a finite normal band can be dualised by an alter ago with only binary total operations and binary relations (and topology). This will imply that every finite normal band has relational degree at most three (Corollary 4.43).

First, we will introduce some notation related to Definition 4.1. Let $\mathbf{A}$ be a finite algebra, and let $n \in \omega$. For $n > 0$, we denote by $\mathcal{R}_n(\mathbf{A})$ the set of all $n$-ary compatible relations on $\mathbf{A}$ (i.e., the set of all subalgebras of $\mathbf{A}^n$). We also denote by $\mathcal{P}_n(\mathbf{A})$ the set of all $n$-ary compatible partial operations on $\mathbf{A}$ (i.e., all homomorphisms from a subalgebra of $\mathbf{A}^n$ into $\mathbf{A}$), and we denote by $\mathcal{T}_n(\mathbf{A})$ the set of all total compatible operations on $\mathbf{A}$ (i.e., all homomorphisms $\mathbf{A}^n \to \mathbf{A}$). Note that $\mathcal{T}_n(\mathbf{A}) \subseteq \mathcal{P}_n(\mathbf{A})$. We put

$$\mathcal{R}_\omega(\mathbf{A}) := \bigcup_{n\in\omega\backslash\{0\}} \mathcal{R}_n(\mathbf{A}), \qquad \mathcal{P}_\omega(\mathbf{A}) := \bigcup_{n\in\omega} \mathcal{P}_n(\mathbf{A}), \qquad \mathcal{T}_\omega(\mathbf{A}) := \bigcup_{n\in\omega} \mathcal{T}_n(\mathbf{A}).$$

An alter ego of $\mathbf{A}$ thus has the form $\langle A; G, R, \mathfrak{T}\rangle$ for some $G \subseteq \mathcal{P}_\omega(\mathbf{A})$ and some $R \subseteq \mathcal{R}_\omega(\mathbf{A})$.

Constructing an alter ego $\underset{\sim}{\mathbf{A}}$ of $\mathbf{A}$ can sometimes require the inclusion of many operations or relations in the type of $\underset{\sim}{\mathbf{A}}$, and this will certainly be the case for the alter egos we construct. In this situation, it is somewhat inconvenient to list all of the operations and relations explicitly in the type of $\underset{\sim}{\mathbf{A}}$. Notation-wise, a more convenient alternative is to equip $\underset{\sim}{\mathbf{A}}$ with *all* compatible operations or relations of a suitable arity. If desired, one can recover a more optimal set of operations and relations from the relevant proofs.

To obtain dualities, it will suffice in all of our proofs to use compatible partial binary operations. Thus, we will usually take our alter ego of $\mathbf{A}$ to be $\underset{\sim}{\mathbf{A}} = \langle A; \mathcal{P}_2(\mathbf{A}), \mathfrak{T}\rangle$. This will certainly include more partial operations than necessary to achieve dualisability. In [**17**] and [**18**], Davey and Priestley developed methods of achieving dualities with minimal structure on $\underset{\sim}{\mathbf{A}}$ (called optimal dualities), but in this thesis, we will be interested only in whether it is possible to achieve a duality, so the notational efficiency of including all operations or relations of a given arity will be preferred over an optimal dualising structure.

For normal bands in particular, we will be able to achieve (IC) using only total operations, provided that we also allow compatible binary relations. That is, we will show that if $\mathbf{S}$ is a finite normal band, then $\underset{\sim}{\mathbf{S}} = \langle S; \mathcal{T}_2(\mathbf{S}), \mathcal{R}_2(\mathbf{S}), \mathfrak{T}\rangle$ dualises $\mathbf{S}$ via (IC).

Whenever we include a partial operation or relation in the structure of $\underset{\sim}{\mathbf{A}}$, this will allow us to use other operations and relations without necessarily including them in the type. For example, if we include a meet operation in $\underset{\sim}{\mathbf{A}}$, we will also be able to use the associated order relation when working with morphisms. We make this idea precise with the notion of entailment. Our definition, suited to (IC), is a combination of hom-entailment ([**5**, §9.4]) and a finitary version of structural entailment as defined by Davey, Haviar, and Willard [**11**] (cf. the more widely-used notion of duality entailment [**5**, § 2.4]).

Let $\mathbf{A}$ be a finite algebra and $\underset{\sim}{\mathbf{A}}$ an alter ego of $\mathbf{A}$.

- If $r \in \mathcal{R}_\omega(\mathbf{A})$, we say that $\underset{\sim}{\mathbf{A}}$ **entails** $r$, and write $\underset{\sim}{\mathbf{A}} \Vdash r$, if for every $n \in \omega\backslash\{0\}$ and every $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$, every morphism $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ preserves $r$.
- If $h \in \mathcal{P}_\omega(\mathbf{A})$, we say that $\underset{\sim}{\mathbf{A}}$ **entails** $h$, and write $\underset{\sim}{\mathbf{A}} \Vdash h$, if for every $n \in \omega\backslash\{0\}$ and every non-empty substructure $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$, we have that $X$ is closed under $h$ and every morphism $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ preserves $h$.
- If $H \subseteq \mathcal{P}_\omega(\mathbf{A}) \cup \mathcal{R}_\omega(\mathbf{A})$, we say that $\underset{\sim}{\mathbf{A}}$ **entails** $H$, and write $\underset{\sim}{\mathbf{A}} \Vdash H$, if $\underset{\sim}{\mathbf{A}} \Vdash h$ for every $h \in H$.

To understand these definitions properly, there is a subtle point regarding types that needs to be addressed. When constructing an alter ego $\underset{\sim}{\mathbf{A}}$, we start by taking sets $G \subseteq \mathcal{P}_\omega(\mathbf{A})$

and $R \subseteq \mathcal{R}_\omega(\mathbf{A})$ to form the topological structure $\underset{\sim}{\mathbf{A}} = \langle A; G, R, \mathfrak{T}\rangle$. We then take the type of $\underset{\sim}{\mathbf{A}}$ to be the actual pair of $(G, R)$ of sets, with the arities as defined in the obvious way. In other words, we use the operations and relations in $G, R$ as symbols for the type of $\underset{\sim}{\mathbf{A}}$. The operations and relations on $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$ therefore have the natural coordinatewise definitions.

For example, suppose $\underset{\sim}{\mathbf{A}}$ has a meet operation $\wedge \in \mathcal{T}_2(\mathbf{A})$ (i.e., $\wedge$ is associative, idempotent, and commutative). If $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$ for some $n \in \omega\backslash\{0\}$, then the interpretation of $\wedge$ in $\mathbf{X}$, which we denote by $\wedge^{\mathbf{X}}$, can be described as follows: for each $(x_1, \dots, x_n), (y_1, \dots, y_n) \in X$, we have

$$(x_1, \dots, x_n) \wedge^{\mathbf{X}} (y_1, \dots, y_n) = (x_1 \wedge y_1, \dots, x_n \wedge y_n).$$

We will often omit the superscript $\mathbf{X}$, but it will be important to use it in certain situations.

When interpreting the definition of entailment, we assume the relations and operations being entailed are also defined pointwise. For example, suppose $\leqslant$ is the order relation on $A$ associated with the meet operation $\wedge$ (i.e., $a \leqslant b \Leftrightarrow a \wedge b = a$, for all $a, b \in A$). Then $\leqslant$ is interpreted on $\mathbf{X}$ pointwise:

$$(x_1, \dots, x_n) \leqslant^{\mathbf{X}} (y_1, \dots, y_n) \iff x_1 \leqslant y_1 \ \ \& \ \ \cdots \ \ \& \ \ x_n \leqslant y_n.$$

It is easily verified that if $\underset{\sim}{\mathbf{A}}$ includes $\wedge \in \mathcal{T}_2(\mathbf{A})$ in its type, then any morphism $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ preserves $\leqslant$; that is, $\underset{\sim}{\mathbf{A}} \Vdash \leqslant$.

The following basic result, which we will use without mention, shows that the compatible operations of a fixed arity entail all compatible operations of lower arities (and similarly for relations). Also, the compatible partial operations of a fixed arity entail all compatible relations of the same or lower arity.

**Proposition 4.28.** *Let $\mathbf{A}$ be a finite algebra, and let $m \in \omega\backslash\{0\}$. Then the following hold:*

(i) *for $0 \leqslant k \leqslant m$, we have $\langle A; \mathcal{T}_m(\mathbf{A}), \mathfrak{T}\rangle \Vdash \mathcal{T}_k(\mathbf{A})$ and $\langle A; \mathcal{P}_m(\mathbf{A}), \mathfrak{T}\rangle \Vdash \mathcal{P}_k(\mathbf{A})$;*

(ii) *for $1 \leqslant k \leqslant m$, we have $\langle A; \mathcal{R}_m(\mathbf{A}), \mathfrak{T}\rangle \Vdash \mathcal{R}_k(\mathbf{A})$ and $\langle A; \mathcal{P}_m(\mathbf{A}), \mathfrak{T}\rangle \Vdash \mathcal{R}_k(\mathbf{A})$.*

Proof. Let $\underset{\sim}{\mathbf{A}} := \langle A; \mathcal{P}_m(\mathbf{A}), \mathfrak{T}\rangle$, and let $g \in \mathcal{P}_k(\mathbf{A})$ with $0 \leqslant k \leqslant m$, so $\mathrm{dom}(g)$ is a subalgebra of $\mathbf{A}^k$; we will show that $\underset{\sim}{\mathbf{A}} \Vdash g$. Define the partial operation $g'\colon A^m \to A$ by

$$\mathrm{dom}(g') := \mathrm{dom}(g) \times A^{m-k} \subseteq A^m, \qquad g'(x_1, \dots, x_m) := g(x_1, \dots, x_k).$$

Clearly $g' \in \mathcal{P}_m(\mathbf{A})$. Now, if $\mathbf{X}$ is a non-empty substructure of $\underset{\sim}{\mathbf{A}}^n$ for some $n \in \omega\backslash\{0\}$ and $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ is a morphism, then $\varphi$ preserves $g'$ by assumption, from which one easily verifies that $X$ is closed under $g$ and $\varphi$ preserves $g$. Thus, $\underset{\sim}{\mathbf{A}} \Vdash g$, showing that $\underset{\sim}{\mathbf{A}} \Vdash \mathcal{P}_k(\mathbf{A})$.

The proofs that $\langle A; \mathcal{T}_m(\mathbf{A}), \mathfrak{T}\rangle \Vdash \mathcal{T}_k(\mathbf{A})$ and $\langle A; \mathcal{R}_m(\mathbf{A}), \mathfrak{T}\rangle \Vdash \mathcal{R}_k(\mathbf{A})$ are analogous. It remains to show that $\underset{\sim}{\mathbf{A}} \Vdash \mathcal{R}_m(\mathbf{A})$ (where $\underset{\sim}{\mathbf{A}} = \langle A; \mathcal{P}_m(\mathbf{A}), \mathfrak{T}\rangle$). Let $r \in \mathcal{R}_m(\mathbf{A})$, and let the partial map $g\colon A^m \to A$ be the restriction of some projection map to $r$. Clearly $g \in \mathcal{P}_m(\mathbf{A})$, and it is easily shown that a morphism $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ for some $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$ preserves $r$. $\square$

Our duality proofs for normal bands will employ the following simple entailment result, which allows us to use the compatible total operations on certain substructures.

**Lemma 4.29.** *Let $\mathbf{S}$ be a finite algebra, and let $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$. Then, for every $k \in \omega$, every operation in $\mathcal{T}_k(\mathbf{A})$ extends to an operation in $\mathcal{T}_k(\mathbf{S})$. In particular, $\langle S; \mathcal{T}_k(\mathbf{S}), \mathfrak{T}\rangle \Vdash \mathcal{T}_k(\mathbf{A})$.*

PROOF. Let $r$ be a retraction of $\mathbf{S}$ onto $\mathbf{A}$. Given $g \in \mathcal{T}_k(\mathbf{A})$, define the extension $g' \colon \mathbf{S}^k \to \mathbf{S}$ as the natural product map $r^k \colon \mathbf{S}^k \to \mathbf{A}^k$ followed by $g \colon \mathbf{A}^k \to \mathbf{A}$. Clearly $g' \in \mathcal{T}_k(\mathbf{S})$ and agrees with $g$ on $A^k$, and it is easily shown that $\langle S; g', \mathcal{T}\rangle \Vdash g$. $\square$

Now, there are two important cases that need special attention: unary relations and nullary operations. Let us first consider unary relations. As usual, we identify $A$ with its first Cartesian power, so that a unary relation on $A$ is a subset of $A$. A compatible unary relation on $\mathbf{A}$ is then a subalgebra of $\mathbf{A}$. If $B \in \mathcal{R}_1(\mathbf{A})$, then the interpretation of $B$ on $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$ is the subset $B^{\mathbf{X}} = B^n \cap X$ of $X$. That is, for all $(x_1, \dots, x_n) \in X$, we have $(x_1, \dots, x_n) \in B^{\mathbf{X}}$ if and only if $x_1, \dots, x_n \in B$. Concerning unary relations, we have the following simple but important result.

**Lemma 4.30.** *Let $\mathbf{S}$ be a finite algebra, let $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$, let $r$ be a retraction of $\mathbf{S}$ onto $\mathbf{A}$, and let $\underset{\sim}{\mathbf{S}}$ be an alter ego of $\mathbf{S}$ with $\underset{\sim}{\mathbf{S}} \Vdash r$. If $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ for some $n \in \omega\backslash\{0\}$ and $\varphi \colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ is a morphism, then $\varphi(A^{\mathbf{X}}) \subseteq A$. (That is, $\underset{\sim}{\mathbf{S}} \Vdash A$.)*

PROOF. Let $x \in A^{\mathbf{X}}$. Then $r(x) = x$, and so $\varphi(x) = \varphi(r(x)) = r(\varphi(x)) \in A$. $\square$

Next, we address nullary operations. As subalgebras are non-empty and $\mathbf{A}^0$ is a one-element algebra, we have $\mathcal{P}_0(\mathbf{A}) = \mathcal{T}_0(\mathbf{A})$, so there are no compatible *proper* partial nullary operations. As usual, a nullary operation on $A$ will be identified with the element of $A$ that it distinguishes. From this point of view, an element $a \in A$ is a compatible nullary operation on $\mathbf{A}$ precisely if $\{a\}$ is a one-element subalgebra of $\mathbf{A}$. Thus, for a semigroup, the compatible nullary operations are precisely the idempotents. Now, if $a \in \mathcal{T}_0(\mathbf{A})$ and $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$, then the interpretation of $a$ as a nullary operation on $\mathbf{X}$ is the constant $n$-tuple with value $a$, which we denote by $\underline{a}$. Thus, if $\underset{\sim}{\mathbf{A}} \Vdash a$, then any morphism $\varphi \colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ will satisfy $\varphi(\underline{a}) = a$.

Before getting into the (IC) proofs, there is one minor point that needs to be addressed. When verifying (IC), the case where the substructure $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$ is empty does not need to be considered, since the empty map $\varnothing \to A$ obviously extends to an $n$-ary term function (any projection map $A^n \to A$ will do). Since we would like $\mathbf{X}$ to be closed under any entailed nullary operations, it will be convenient to assume that $X \neq \varnothing$. We will therefore assume without further mention that all substructures of powers of alter egos are non-empty.

The following lemma (which is a modified version of [**2**, Lemma 4.2]) will be crucial for obtaining our general dualisability results; it will allow us to obtain an (IC) duality for a given algebra from (IC) dualities on certain subalgebras.

**Lemma 4.31.** *Let $\mathbf{S}$ be a finite algebra with a set $\mathfrak{U}$ of subalgebras such that $\mathbf{S} \in \mathbb{SP}(\mathfrak{U})$. Let $\underset{\sim}{\mathbf{S}}$ be an alter ego of $\mathbf{S}$ with $\underset{\sim}{\mathbf{S}} \Vdash \mathcal{T}_1(\mathbf{S})$, let $\varphi \colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ be a morphism for some $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ and some $n \in \omega\backslash\{0\}$, and let $t$ be an $n$-ary term function of $\mathbf{S}$ that agrees with $\varphi$ on the set $\bigcup\{A^{\mathbf{X}} \mid \mathbf{A} \in \mathfrak{U}\}$. Then $t{\restriction}_X = \varphi$.*

PROOF. Let $H \subseteq \bigcup\{\hom(\mathbf{S}, \mathbf{A}) \mid \mathbf{A} \in \mathfrak{U}\}$ separate the points of $S$. Then $H \subseteq \mathcal{T}_1(\mathbf{S})$, so $\underset{\sim}{\mathbf{S}} \Vdash H$ by assumption. Now, let $x \in X$, and let $u \in H$. Then $u(x) \in A^{\mathbf{X}}$ for some $\mathbf{A} \in \mathfrak{U}$, so $\varphi(u(x)) = t(u(x))$ by the choice of $t$, and therefore $u(\varphi(x)) = u(t(x))$. Since $u \in H$ was arbitrary, we have $\varphi(x) = t(x)$. $\square$

In a similar vein, we have the following result, which was based on [**1**, Lemma 5.1.2].

**Lemma 4.32.** *Let* $\mathbf{S}$ *be a finite algebra, and let* $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$. *Suppose there is some* $k \in \omega\backslash\{0\}$ *such that* $\underset{\sim}{\mathbf{A}} := \langle A; \mathcal{T}_k(\mathbf{A}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{A}$, *and let* $\underset{\sim}{\mathbf{S}}$ *be an alter ego of* $\mathbf{S}$ *with* $\underset{\sim}{\mathbf{S}} \Vdash \mathcal{T}_k(\mathbf{S})$. *If* $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ *for some* $n \in \omega\backslash\{0\}$ *and* $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ *is a morphism, then there is an* $n$*-ary term* $t$ *such that* $t^{\mathbf{A}}$ *agrees with* $\varphi$ *on* $A^{\mathbf{X}}$. *Consequently, if* $\mathbf{S} \in \mathbb{SP}(\mathbf{A})$, *then* $\underset{\sim}{\mathbf{S}}$ *satisfies* (IC) *with respect to* $\mathbf{S}$.

Proof. Let $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ be a morphism for some $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ and some $n \in \omega\backslash\{0\}$. Then, by Lemma 4.30, we have $\varphi(A^{\mathbf{X}}) \subseteq A$. Now, by Lemma 4.29, we have $\underset{\sim}{\mathbf{S}} \Vdash \mathcal{T}_k(\mathbf{A})$, so $A^{\mathbf{X}}$ forms a substructure, $\mathbf{X}_A$, of $\underset{\sim}{\mathbf{A}}^n$, and $\varphi{\restriction}_{A^{\mathbf{X}}}$ is a morphism $\mathbf{X}_A \to \underset{\sim}{\mathbf{A}}$. Since $\underset{\sim}{\mathbf{A}}$ satisfies (IC), there is an $n$-ary term $t$ such that $t^{\mathbf{A}}$ agrees with $\varphi$ on $A^{\mathbf{X}}$. Finally, if $\mathbf{S} \in \mathbb{SP}(\mathbf{A})$, then applying Lemma 4.31 with $\mathfrak{U} = \{\mathbf{A}\}$ gives $t^{\mathbf{S}}{\restriction}_X = \varphi$. ☐

Since all of our compatible operations on semigroups will be at most binary, we record in the following lemma a useful equation for verifying the compatibility of a binary operation.

**Lemma 4.33.** *Let* $\mathbf{S}$ *be a semigroup, and let* $*\colon S^2 \to S$ *be a partial binary operation on* $S$. *Then* $* \in \mathcal{P}_2(\mathbf{S})$ *if and only if* $\mathrm{dom}(*)$ *forms a subsemigroup of* $\mathbf{S}^2$ *and*

$$(\forall (x,y),(z,t) \in \mathrm{dom}(*))\ \ (x*y)(z*t) = xz * yt.$$

*In particular, the multiplication of* $\mathbf{S}$ *is in* $\mathcal{T}_2(\mathbf{S})$ *if and only if* $\mathbf{S} \models xyzt \approx xzyt$.

Lemma 4.33 shows in particular that the multiplication of a normal band is compatible, so we will always have it available to us when verifying (IC). Interestingly, however, we will very rarely need it.

Finally, we come to the specific results for normal bands. The first of these results gives us compatible meet operations on the subdirectly irreducible normal bands. Meet operations will play a crucial role in all of our (IC) proofs.

**Lemma 4.34.** *The semigroups* $\mathbf{I}$, $\mathbf{L}$, *and* $\mathbf{L}^0$ *have compatible meet operations. Specifically:*

(i) *the multiplication of* $\mathbf{I}$ *is a meet operation in* $\mathcal{T}_2(\mathbf{I})$;
(ii) *any linear ordering on* $L$ *gives a meet operation in* $\mathcal{T}_2(\mathbf{L})$;
(iii) *any linear ordering on* $L^0$ *where* $0$ *is the least element gives a meet operation in* $\mathcal{T}_2(\mathbf{L}^0)$.

Proof. Lemma 4.33 gives (i), while (ii) follows from the fact that all maps between left-zero semigroups are homomorphisms (hence all binary operations are compatible). To prove (iii) from (ii), let $\wedge$ be a meet operation on $L^0$ induced by a linear order with least element 0, and let $x, z, y, t \in L^0$ to show that $(x \wedge y)(z \wedge t) = xz \wedge yt$. If $0 \in \{x, y, z, t\}$, then the equality clearly holds, and otherwise, we get equality from (ii). ☐

We will now establish the dualisability of all finite normal bands over the following six lemmas. In the proofs, it will be important to interpret the elements of $\omega$ as ordinals; that is, each $n \in \omega\backslash\{0\}$ is identified with the set $\{0, \ldots, n-1\}$.

**Lemma 4.35.** *If* $\mathbf{A} \in \mathbb{I}(\mathbf{L}, \mathbf{R})$, *then* $\underset{\sim}{\mathbf{A}} := \langle A; \mathcal{T}_2(\mathbf{A}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{A}$.

Proof. Let $\mathbf{X} \leqslant \underset{\sim}{\mathbf{A}}^n$ for some $n \in \omega\backslash\{0\}$, and let $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{A}}$ be a morphism.

Write $A = \{a, b\}$. Let $\wedge$ be the meet operation on $A$ with $a > b$, which is compatible with $\mathbf{A}$ by Lemma 4.34. Since $\varphi(\underline{a}) = a$, the set $\varphi^{-1}(a)$ is non-empty, so it is a subsemilattice of $\langle X; \wedge\rangle$. Let $\widehat{a}$ denote the least element of $\varphi^{-1}(a)$. Since $\varphi(\underline{b}) = b$, we must have $\widehat{a} \neq \underline{b}$, so $\widehat{a}(j) = a$ for some $j \in n$. We will show that $\varphi(x) = x(j)$ for all $x \in X$.

Let $x \in X$. If $\varphi(x) = a$, then we have $x \geqslant \widehat{a}$ by the choice of $\widehat{a}$, so $x(j) = a$ because $a$ is maximal in $\langle A; \wedge\rangle$. Assume that $\varphi(x) = b$, and let $f$ be the non-trivial automorphism of $\mathbf{A}$. Then $\varphi(f(x)) = f(\varphi(x)) = f(b) = a$, so the previous argument gives $f(x)(j) = a$, and hence $x(j) = b$. It follows that $\varphi$ is extended by the $j$th projection $A^n \to A$. $\square$

**Lemma 4.36.** *Let* $\mathbf{S}$ *be a finite rectangular band. Then* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{T}_2(\mathbf{S}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$.

Proof. If $\mathbf{S}$ is left-zero or right-zero, then applying Lemmas 4.20, 4.32, and 4.35 gives the result, so we may assume that there are $\mathbf{A}, \mathbf{B} \leqslant \mathbf{S}$ with $\mathbf{A} \cong \mathbf{L}$ and $\mathbf{B} \cong \mathbf{R}$. By Lemma 4.20, we have $\mathbf{A}, \mathbf{B} \leqslant_{\mathrm{r}} \mathbf{S}$.

Let $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ for some $n \in \omega\backslash\{0\}$, and let $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ be a morphism. By Lemma 4.32 and Lemma 4.35, there is an $n$-ary term function of $\mathbf{A}$ that agrees with $\varphi$ on $A^{\mathbf{X}}$. But a term function of $\mathbf{A}$ is a projection, so there exists $j \in n$ such that $\varphi(x) = x(j)$ for all $x \in A^{\mathbf{X}}$. Similarly, there exists $k \in n$ such that $\varphi(x) = x(k)$ for all $x \in B^{\mathbf{X}}$.

Define $t(x) := x(j)x(k)$ for all $x \in S^n$, so $t$ is an $n$-ary term function of $\mathbf{S}$. Then, for all $x \in A^{\mathbf{X}}$, we have $t(x) = x(j) = \varphi(x)$, and for all $x \in B^{\mathbf{X}}$, we have $t(x) = x(k) = \varphi(x)$, so $t$ agrees with $\varphi$ on $A^{\mathbf{X}} \cup B^{\mathbf{X}}$. Since $\mathbf{S} \in \mathbb{SP}(\mathbf{A}, \mathbf{B})$, we have $t{\restriction}_X = \varphi$ by Lemma 4.31. $\square$

The next result effectively establishes (IC) dualities for $\mathbf{I}$ and $\mathbf{L}^0$. We record some technical details in the statement for use in later proofs. Note that $v_i$ denotes the $i$th variable, for each $i \in n$.

**Lemma 4.37.** *Let* $\mathbf{S}$ *be a finite semigroup, let* $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$ *with* $\mathbf{A} \in \mathbb{I}(\mathbf{I}, \mathbf{L}^0)$, *and let* $0$ *denote the zero of* $\mathbf{A}$. *Let* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{T}_2(\mathbf{S}), \mathfrak{T}\rangle$, *let* $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ *for some* $n \in \omega\backslash\{0\}$, *and let* $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ *be a morphism. Then there exist* $z \in A^{\mathbf{X}}$ *and* $j \in J := \{i \in n \mid z(i) \neq 0\}$ *such that*

$$(\forall x \in A^{\mathbf{X}})\ \big(\varphi(x) \neq 0 \iff (\forall i \in J)\ x(i) \neq 0\big)\ \&\ \big(\varphi(x) \neq 0 \implies x(j) = \varphi(x)\big).$$

*Consequently, the term function of* $\mathbf{S}$ *induced by the* $n$*-ary term*

$$v_j\Big(\prod_{i \in J} v_i\Big)$$

*agrees with* $\varphi$ *on* $A^{\mathbf{X}}$.

Proof. Write $\{a, b\}$ for the set $A\backslash\{0\}$, so that $a = b \Leftrightarrow \mathbf{A} \cong \mathbf{I} \Leftrightarrow \mathbf{A} \ncong \mathbf{L}^0$.

Let $\wedge$ denote the meet operation on $A$ such that $\langle A; \wedge\rangle$ is a chain with greatest element $a$ and least element $0$. Then $\wedge \in \mathcal{T}_2(\mathbf{A})$ by Lemma 4.34, so $\underset{\sim}{\mathbf{S}} \Vdash \wedge$ by Lemma 4.29. By Lemma 4.30, we have $\varphi(A^{\mathbf{X}}) \subseteq A$, so $\varphi{\restriction}_{A^{\mathbf{X}}}$ is a morphism $\langle A^{\mathbf{X}}; \wedge\rangle \to \langle A; \wedge\rangle$.

Since $\varphi(\underline{a}) = a$, the set $\varphi^{-1}(\{a, b\}) \cap A^{\mathbf{X}}$ is non-empty, so it is a subsemilattice of $\langle A^{\mathbf{X}}; \wedge\rangle$. Let $\widehat{b}$ denote the least element of $\varphi^{-1}(\{a, b\}) \cap A^{\mathbf{X}}$; we will show that $\widehat{b}$ has the required properties of the element $z$. Since $\varphi(\underline{b}) = b$, we have $\widehat{b} \leqslant \underline{b}$, so $\widehat{b} \in \{0, b\}^n$, and because $\varphi$ is order-preserving, $\widehat{b} \leqslant \underline{b}$ implies $\varphi(\widehat{b}) \leqslant b$, so $\varphi(\widehat{b}) = b$ since $\varphi(\widehat{b}) \in \{a, b\}$. Define

$$J := \{i \in n \mid \widehat{b}(i) = b\} = \{i \in n \mid \widehat{b}(i) \neq 0\}.$$

From the definition of $\widehat{b}$ and the fact that $\varphi$ is order-preserving, we have for all $x \in A^{\mathbf{X}}$ that

$$\varphi(x) \neq 0 \iff x \geqslant \widehat{b} \iff (\forall i \in J)\ \ x(i) \neq 0.$$

To complete the proof, it remains to show that there is some $j \in J$ such that $\varphi(x) = x(j)$ for all $x \in A^{\mathbf{X}}$. Since $\varphi(\underline{0}) = 0 \notin \{a, b\}$, we cannot have $\widehat{b} = \underline{0}$, so $J \neq \varnothing$. Thus, if $a = b$, then choosing any $j \in J$ gives the required result, so we may assume for the remainder of the proof that $a \neq b$.

Let $\widehat{a}$ denote the least element of $\varphi^{-1}(a) \cap A^{\mathbf{X}}$. We claim that

$$J = \{i \in n \mid \widehat{a}(i) \neq 0\}.$$

That is, $\widehat{b}(i) \neq 0 \Leftrightarrow \widehat{a}(i) \neq 0$, for all $i \in n$. Clearly $\widehat{a} \geqslant \widehat{b}$ from the definition of $\widehat{b}$, which implies that $J \subseteq \{i \in n \mid \widehat{a}(i) \neq 0\}$. For the reverse inclusion, let $f$ be the non-trivial automorphism of $\mathbf{A}$, so $\underset{\sim}{\mathbf{S}} \Vdash f$ by Lemma 4.29. Then $\varphi(f(\widehat{b})) = f(\varphi(\widehat{b})) = a$, so $f(\widehat{b}) \geqslant \widehat{a}$ by the definition of $\widehat{a}$. Thus, for all $i \in n$, we have $\widehat{a}(i) \neq 0 \Rightarrow f(\widehat{b})(i) \neq 0 \Rightarrow i \in J$, giving the reverse inclusion. Hence, $J = \{i \in n \mid \widehat{a}(i) \neq 0\}$.

Now clearly $\widehat{a} \neq \widehat{b}$ since $\varphi(\widehat{a}) \neq \varphi(\widehat{b})$, so because $J = \{i \in n \mid \widehat{a}(i) \neq 0\}$, there must be some $j \in J$ with $\widehat{a}(j) = a$. Thus, if $x \in A^{\mathbf{X}}$ with $\varphi(x) = a$, then $x \geqslant \widehat{a}$, so $x(j) = a = \varphi(x)$. Also, if $x \in A^{\mathbf{X}}$ with $\varphi(x) = b$, then $\varphi(f(x)) = f(\varphi(x)) = a$, so $f(x)(j) = a$ by the previous argument, and therefore $x(j) = b = \varphi(x)$. $\square$

**Lemma 4.38.** *Let* $\mathbf{S}$ *be a finite normal band with* $\mathbb{SP}(\mathbf{S}) = \mathbb{SP}(\mathbf{A})$*, where* $\mathbf{A} \in \{\mathbf{I}, \mathbf{L}^0, \mathbf{R}^0\}$*. Then* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{T}_2(\mathbf{S}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$*.*

Proof. Note that $\mathbf{A}$ embeds into $\mathbf{S}$ since $\mathbf{A}$ is subdirectly irreducible, so by Lemma 4.22 or Lemma 4.23, we can assume that $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$. By Lemma 4.37, $\langle A; \mathcal{T}_2(\mathbf{A}), \mathfrak{T}\rangle$ satisfies (IC) with respect to $\mathbf{A}$, so Lemma 4.32 gives the result. $\square$

**Lemma 4.39.** *Let* $\mathbf{S}$ *be a finite normal band such that* $\mathbb{SP}(\mathbf{S})$ *is the variety* $\mathbb{SP}(\mathbf{L}^0, \mathbf{R}^0)$ *of all normal bands. Then* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{T}_2(\mathbf{S}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$*.*

Proof. Lemma 4.23, combined with the fact that $\mathbf{L}^0$ and $\mathbf{R}^0$ are subdirectly irreducible, implies that there are $\mathbf{A}, \mathbf{B} \leqslant_{\mathrm{r}} \mathbf{S}$ with $\mathbf{A} \cong \mathbf{L}^0$ and $\mathbf{B} \cong \mathbf{R}^0$.

Let $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ for some $n \in \omega\backslash\{0\}$, and let $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ be a morphism. By Lemma 4.37, there are $n$-ary term functions $t_\ell$ and $t_r$ of $\mathbf{S}$ that agree with $\varphi$ on $A^{\mathbf{X}}$ and $B^{\mathbf{X}}$, respectively. Define the $n$-ary term function $t$ of $\mathbf{S}$ by $t(x) := t_\ell(x)t_r(x)$ for all $x \in S^n$. We claim that $t$ extends $\varphi$. By Lemma 4.31, it suffices to show that $t$ agrees with $\varphi$ on $A^{\mathbf{X}} \cup B^{\mathbf{X}}$, and by symmetry, it suffices to show that $t$ agrees with $\varphi$ on $A^{\mathbf{X}}$.

Let 0 and $z$ respectively denote the zeroes of $\mathbf{A}$ and $\mathbf{B}$, let $f\colon \mathbf{A} \to \mathbf{B}$ be a homomorphism with $f^{-1}(z) = \{0\}$, and let $g := f \circ r$ for some retraction $r$ of $\mathbf{S}$ onto $\mathbf{A}$. Then $\varphi$ preserves $g$.

Let $x \in A^{\mathbf{X}}$. By Lemma 4.30, we have $\varphi(x) \in A$. If $\varphi(x) = 0$, then $t_\ell(x) = \varphi(x) = 0$, so $t(x) = 0t_r(x) = 0$. Assume that $\varphi(x) \neq 0$. Since $g(x) \in B^{\mathbf{X}}$, we have $t_r(g(x)) = \varphi(g(x))$, and hence $g(t_r(x)) = g(\varphi(x))$. Now, $\varphi(x) \neq 0 \Rightarrow g(\varphi(x)) \neq z \Rightarrow g(t_r(x)) \neq z \Rightarrow t_r(x) \neq 0$; thus, $t_r(x)$ is a right identity element of $\mathbf{A}$, so $t(x) = t_\ell(x)t_r(x) = \varphi(x)t_r(x) = \varphi(x)$. $\square$

**Lemma 4.40.** *Let* $\mathbf{S}$ *be a finite normal band, let* $\mathbf{A}, \mathbf{B} \leqslant \mathbf{S}$ *with* $\mathbf{A} \in \mathbb{I}(\mathbf{I}, \mathbf{L}^0)$ *and* $\mathbf{B} \cong \mathbf{R}$*, and assume that* $\mathbf{R}^0$ *does not embed into* $\mathbf{S}$*. Let* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{T}_2(\mathbf{S}), \mathcal{R}_2(\mathbf{S}), \mathfrak{T}\rangle$*, let* $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ *for some* $n \in \omega\backslash\{0\}$*, and let* $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ *be a morphism. Then there is an* $n$*-ary term function* $t$ *of* $\mathbf{S}$ *that agrees with* $\varphi$ *on* $A^{\mathbf{X}} \cup B^{\mathbf{X}}$*. Consequently, if* $\mathbf{S} \in \mathbb{SP}(\mathbf{A}, \mathbf{B})$*, then* $\underset{\sim}{\mathbf{S}}$ *satisfies* (IC) *with respect to* $\mathbf{S}$*.*

Proof. By Lemma 4.24, we have $\mathbf{B} \leqslant_{\mathrm{r}} \mathbf{S}$. Write $B = \{a, b\}$, and let $\wedge$ denote the meet operation on $B$ with $a > b$. Then $\wedge \in \mathcal{T}_2(\mathbf{B})$ by Lemma 4.34, so $\underset{\sim}{\mathbf{S}} \Vdash \wedge$ by Lemma 4.29. Since $\varphi(B^{\mathbf{X}}) \subseteq B$ by Lemma 4.30, it follows that $\varphi{\restriction}_{B^{\mathbf{X}}}$ is a morphism $\langle B^{\mathbf{X}}; \wedge\rangle \to \langle B; \wedge\rangle$, so $\varphi^{-1}(a) \cap B^{\mathbf{X}}$ is a subsemilattice of $\langle B^{\mathbf{X}}; \wedge\rangle$. Let $\widehat{a}$ denote the least element of $\varphi^{-1}(a) \cap B^{\mathbf{X}}$.

By Lemma 4.22 or Lemma 4.23, we have $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$. Let 0 denote the zero of $\mathbf{A}$, and use Lemma 4.37 to choose $z \in A^{\mathbf{X}}$ and $j \in J := \{i \in n \mid z(i) \neq 0\}$ so that the term function, $t$, of $\mathbf{S}$ induced by $v_j\big(\prod_{i \in J} v_i\big)$ agrees with $\varphi$ on $A^{\mathbf{X}}$. In particular, we have $\varphi(z) \neq 0$.

Define the binary relation

$$s := \{(x, y) \in A \times B \mid x \neq 0 \implies y = b\}.$$

If $(x, y), (x', y') \in s$, then $xx' \neq 0 \Rightarrow x, x' \neq 0 \Rightarrow y = y' = b \Rightarrow yy' = b$, so $(xx', yy') \in s$. Thus, $s \in \mathcal{R}_2(\mathbf{S})$, so $\varphi$ preserves $s$. Since $(\varphi(z), \varphi(\widehat{a})) \notin s$, we have $(z, \widehat{a}) \notin s^{\mathbf{X}}$, so there exists $k \in J$ with $\widehat{a}(k) = a$. Now, if $x \in \varphi^{-1}(a) \cap B^{\mathbf{X}}$, then $x \geqslant \widehat{a}$, so $x(k) = a$, and applying the non-trivial automorphism of $\mathbf{B}$ shows that $x(k) = b$ for all $x \in \varphi^{-1}(b) \cap B^{\mathbf{X}}$. Hence, we have $\varphi(x) = x(k)$ for all $x \in B^{\mathbf{X}}$.

Define the $n$-ary term function $t$ of $\mathbf{S}$ by $t(x) := t_\ell(x)x(k)$ for all $x \in S^n$. Since $k \in J$, we have $t(x) = t_\ell(x) = \varphi(x)$ for all $x \in A^{\mathbf{X}}$, and since $\mathbf{B} \cong \mathbf{R}$, we have $t(x) = x(k) = \varphi(x)$ for all $x \in B^{\mathbf{X}}$. Hence, $t$ agrees with $\varphi$ on $A^{\mathbf{X}} \cup B^{\mathbf{X}}$. The last claim is by Lemma 4.31. $\square$

We now obtain the desired dualisability result for normal bands.

**Theorem 4.41.** *Let* $\mathbf{S}$ *be a finite normal band. Then* $\langle S; \mathcal{T}_2(\mathbf{S}), \mathcal{R}_2(\mathbf{S}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$*. Moreover, if* $\mathbb{SP}(\mathbf{S})$ *is a variety, then* $\langle S; \mathcal{T}_2(\mathbf{S}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$*. Consequently, every finite normal band is dualisable via* (IC)*.*

Proof. By Theorem 4.27, $\mathbb{SP}(\mathbf{S})$ is one of the quasivarieties in Figure 4.2. If $\mathbb{SP}(\mathbf{S})$ is a variety, then the result follows from Lemmas 4.36, 4.38, and 4.39. Assume that $\mathbb{SP}(\mathbf{S})$ is not a variety, so by symmetry, $\mathbb{SP}(\mathbf{S})$ is either $\mathbb{SP}(\mathbf{I}, \mathbf{L})$, $\mathbb{SP}(\mathbf{L}^0, \mathbf{R})$, or $\mathbb{SP}(\mathbf{L}, \mathbf{R}, \mathbf{I})$. In the former two cases, we have $\mathbf{R}^0 \notin \mathbb{S}(\mathbf{S})$, and the result then follows from Lemma 4.40. Thus, we may assume that $\mathbb{SP}(\mathbf{S}) = \mathbb{SP}(\mathbf{I}, \mathbf{L}, \mathbf{R})$. We then have $\mathbf{I}, \mathbf{L}, \mathbf{R} \in \mathbb{S}(\mathbf{S})$ and $\mathbf{L}^0, \mathbf{R}^0 \notin \mathbb{S}(\mathbf{S})$, so by Lemma 4.22 and Lemma 4.24, there are $\mathbf{A}, \mathbf{B}, \mathbf{C} \leqslant_{\mathrm{r}} \mathbf{S}$ isomorphic to $\mathbf{I}$, $\mathbf{L}$, $\mathbf{R}$, respectively.

Let $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{T}_2(\mathbf{S}), \mathcal{R}_2(\mathbf{S}), \mathfrak{T}\rangle$, let $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ for some $n \in \omega\backslash\{0\}$, and let $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ be a morphism. By Lemma 4.40, there are term functions $t_\ell$, $t_r$ of $\mathbf{S}$ that agree with $\varphi$ on $A^{\mathbf{X}} \cup B^{\mathbf{X}}$ and $A^{\mathbf{X}} \cup C^{\mathbf{X}}$, respectively. Define the $n$-ary term function $t$ of $\mathbf{S}$ by $t(x) := t_\ell(x)t_r(x)$ for all $x \in S^n$. Clearly $t(x) = t_\ell(x) = \varphi(x)$ for all $x \in B^{\mathbf{X}}$, so $t$ agrees with $\varphi$ on $B^{\mathbf{X}}$, and by symmetry $t$ agrees with $\varphi$ on $C^{\mathbf{X}}$. Now, let 0 denote the zero of $\mathbf{A}$. Since $t_\ell$ and $t_r$ agree with $\varphi$ on $A^{\mathbf{X}}$, we have for all $x \in A^{\mathbf{X}}$ that $\varphi(x) \neq 0 \Leftrightarrow t_\ell(x), t_r(x) \neq 0 \Leftrightarrow t(x) \neq 0$, so $t$ also agrees with $\varphi$ on $A^{\mathbf{X}}$. Lemma 4.31 now implies that $t$ extends $\varphi$. $\square$

Theorem 4.41 improves on the original result [**2**] in two aspects: we have removed the need to use alter egos with proper partial operations, and we have also achieved a uniform bound on the arities of the operations and relations used.

Although we do not need proper partial operations to obtain dualities for normal bands, it will be convenient in later proofs to simply include all partial operations. We will therefore record the following consequence of Theorem 4.41 and Proposition 4.28(ii).

**Theorem 4.42.** *Let* $\mathbf{S}$ *be a finite normal band. Then* $\langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$.

To close the chapter, we will briefly discuss a property related to dualisability and state a corollary of Theorem 4.42 concerning this property.

Let $\mathbf{A}$ be an algebra, and let $R \subseteq \mathcal{R}_\omega(\mathbf{A})$. We say that $R$ **determines the clone of A** if the following condition holds:

(CLO) *For every* $n \in \omega\backslash\{0\}$, *every map* $\varphi\colon A^n \to A$ *preserving all of the relations in* $R$ *is an* $n$*-ary term function of* $\mathbf{A}$.

The relations are of course interpreted pointwise on $A^n$. If the clone of $\mathbf{A}$ is determined by a finite set of finitary relations, then we say that $\mathbf{A}$ has **finite degree**, or that $\mathbf{A}$ is **finitely related**. If $\mathbf{A}$ has finite degree, we define the **relational degree** of $\mathbf{A}$ to be the least $d \in \omega\backslash\{0\}$ such that $\mathcal{R}_d(\mathbf{A})$ determines the clone of $\mathbf{A}$. (If $\varnothing$ determines the clone of $\mathbf{A}$—that is, if $\mathbf{A}$ is primal—the relational degree of $\mathbf{A}$ is defined to be 0.)

For alter egos with no operations, (IC) is effectively a strengthening of (CLO), since the topology plays no role. In fact, we can obtain (CLO) from any dualising alter ego, even if there are operations present. The trick, which is standard in duality theory, is to replace each operation with its graph. More precisely, if $g\colon A^m \to A$ is a partial operation for some $m \in \omega$, then the graph of $g$ is defined to be the $(m+1)$-ary relation

$$\operatorname{graph}(g) := \{(x_1, \dots, x_m, x_{m+1}) \in A^{m+1} \mid (x_1, \dots, x_m) \in \operatorname{dom}(g) \ \ \&\ \ x_{m+1} = g(x_1, \dots, x_m)\}.$$

By [**5**, Lemma 2.1.2], if $\langle A; G, R, \mathcal{T}\rangle$ dualises the finite algebra $\mathbf{A}$, then we can replace each operation $g \in G$ with the relation $\operatorname{graph}(g)$ without destroying the duality (though we may lose (IC)). We then obtain a dualising alter ago $\underset{\sim}{\mathbf{A}}$ with no operations, and by [**5**, Theorem 2.2.2], the relations of $\underset{\sim}{\mathbf{A}}$ satisfy (CLO) with respect to $\mathbf{A}$. Given these facts, the following result is an immediate consequence of Theorem 4.41.

**Corollary 4.43.** *A finite normal band has relational degree at most 3.*

In [**24**], Dolinka showed that every finite band $\mathbf{S}$ has finite degree $d$, where $d$ is an unbounded function of $|S|$. We have improved this result for the subclass of finite normal bands by giving a uniform bound on the relational degree.

CHAPTER 5

# Natural Dualities for Aperiodic Semigroups

At long last, we come to the new results of this thesis. In this chapter and the next, we will present a proof of the classification theorem for dualisable finite aperiodic semigroups. This chapter contains a few non-dualisability results, but the bulk of the chapter will be dedicated to deriving dualisability results needed for the main proof. Across the two chapters, the constituent results will be presented more or less in order of discovery.

## 5.1. A roadmap

In Section 4.2, we gave an overview of the known dualisability results for finite semigroups, all of which are consistent with the tentative conjecture that a finite semigroup is dualisable if and only if it generates a residually small variety. Since the latter property has been completely characterised (Theorem 3.71), this conjecture has served as a useful guide for researchers in obtaining new dualisability and non-dualisability results. For example, since $\mathbf{P}$ generates a residually small variety, it was thought likely to be dualisable, and this was eventually proved by Jackson [**40**, Theorem 9.2].

The starting point for the present author was to continue along these lines, with the overall aim of pushing the frontiers of the dualisability problem for finite semigroups as far as possible. Many of the simplest semigroups whose dualisability had not yet been determined were aperiodic, so the author's attention was quickly drawn to these examples. However, it did not seem at the time that classifying the dualisable aperiodic semigroups would merit an entire thesis. To explain why, let us give an explicit statement of Theorem 3.71 restricted to aperiodic semigroups. This is easily done, but there are a few details that need to be checked. First, we have the following fundamental fact about aperiodic semigroups.

**Proposition 5.1.** *Let $\mathbf{S}$ be an aperiodic semigroup, and suppose that every element of $\mathbf{S}$ has index at most $k$ for some $k \geqslant 1$. Then $\mathbf{S} \models x^{k+1} \approx x^k$.*

Proof. Let $a \in S$, and let $i$ be the index of $a$. By Theorem 2.34, the set $\{a^\ell \mid \ell \geqslant i\}$ is a subgroup of $\mathbf{S}$, so it must equal $\{a^i\}$. Since $k \geqslant i$, it follows that $a^{k+1} = a^k$. □

Before stating the specialisation of Theorem 3.71, we will take the opportunity to introduce a convenient piece of notation. Given a class $\mathcal{K}$ of semigroups, we define $\mathbb{A}(\mathcal{K})$ to be the class of all semigroups that are either isomorphic or anti-isomorphic to a member of $\mathcal{K}$; thus, $\mathbb{A}(\mathcal{K})$ is the smallest class containing $\mathcal{K}$ that is closed under anti-isomorphism. For example, $\mathbb{A}(\mathbf{L}) = \mathbb{I}(\mathbf{L}, \mathbf{R})$.

Now, define

$$\mathcal{F}_1 := \mathbb{A}(\mathbf{N}_4, \mathbf{L}^+, \mathbf{L}^1, \mathbf{L} \times \mathbf{P}, \mathbf{R} \circlearrowleft 2),$$

which is precisely the class of aperiodic members of the class $\mathcal{F}_\omega$ defined in Section 3.8. We then have the following result.

**Theorem 5.2.** *Let* $\mathbf{S}$ *be a finite aperiodic semigroup. Then the following are equivalent:*

(i) $\mathbb{V}(\mathbf{S})$ *is residually small;*
(ii) $\mathbb{V}(\mathbf{S})$ *is residually very finite;*
(iii) $\mathbb{V}(\mathbf{S})$ *does not contain a semigroup in* $\mathcal{F}_1$*;*
(iv) $\mathbf{S} \in \mathcal{N}_1 \cup \mathcal{P}_1 \cup \mathcal{Q}_1 = \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N}) \cup \mathbb{V}(\mathbf{R}, \mathbf{P}) \cup \mathbb{V}(\mathbf{L}, \mathbf{Q})$.

PROOF. We will show that (iii) and (iv) are equivalent to the corresponding conditions in Theorem 3.71. Because $\mathbf{S}$ is finite, there is some $k \geqslant 1$ with $\mathbf{S} \models x^{k+1} \approx x^k$ by Proposition 5.1, which implies that all groups in $\mathbb{V}(\mathbf{S})$ are trivial. Thus, condition (iii) of the present theorem is equivalent to condition (iii) of Theorem 3.71.

Assume that Theorem 3.71(iv) holds, so $\mathbf{S} \in \mathcal{N}_n \cup \mathcal{P}_n \cup \mathcal{Q}_n$ for some $n \geqslant 1$. We then have $\mathbf{S} \models x^{n+2} \approx x^2$, so all elements of $\mathbf{S}$ have index at most 2. Thus, $\mathbf{S} \models x^3 \approx x^2$ by Proposition 5.1, and therefore $\mathbf{S} \models x^{n+1} \approx x^2$. It is now straightforward to verify from the definitions of $\mathcal{N}_n$, $\mathcal{P}_n$, and $\mathcal{Q}_n$ that $\mathbf{S} \in \mathcal{N}_1 \cup \mathcal{P}_1 \cup \mathcal{Q}_1$. $\square$

This result gave us the following working conjecture.

**Conjecture 5.3.** *Let* $\mathbf{S}$ *be a finite aperiodic semigroup. Then* $\mathbf{S}$ *is dualisable if and only if* $\mathbf{S}$ *satisfies the conditions of Theorem 5.2.*

The author first entertained the idea of proving this conjecture after a thorough study of the results that now constitute Chapter 3. Condition (iii) of Theorem 3.71 had never been explicitly stated in the literature, so extracting this condition illuminated a roadmap to proving Conjecture 5.3. The general idea was to show that Theorem 5.2(iv) implies dualisability, while a failure of Theorem 5.2(iii) implies non-dualisability.

To understand what is involved in the former implication, let us look more closely at the varieties $\mathcal{N}_1$ and $\mathcal{P}_1$. First, it can be shown that each semigroup in $\mathcal{N}_1$ is an inflation of a normal band. (Inflations will be defined precisely in Section 5.3, but for now they can be understood as rather trivial expansions of semigroups.) As it was known that all finite normal bands are dualisable, proving that all finite members of $\mathcal{N}_1$ are dualisable was expected to be a simple task (which, in the end, it was).

For the variety $\mathcal{P}_1$, the idea was to appeal to the Independence of Generator Theorem 4.6, and this required a determination of the subquasivarieties of $\mathcal{P}_1$. It was shown by Sapir [**60**] that $\mathcal{P}_1$ has finitely many subquasivarieties; in particular, it could be deduced from [**60**, Lemma 3.8] that the only quasicriticals in $\mathcal{P}_1$ up to isomorphism are $\mathbf{N}$, $\mathbf{I}$, $\mathbf{R}$, $\mathbf{R}^0$, and $\mathbf{P}$. Since all of these semigroups are subdirectly irreducible, the lattice $\mathrm{L_q}(\mathcal{P}_1)$ was then easily constructed from Theorem 4.19. The subquasivarieties of $\mathcal{P}_1$ that are not already subquasivarieties of $\mathcal{N}_1$ were found to be those generated by $\mathbf{P}$, $\mathbf{R} \times \mathbf{P}$, and $\mathbf{R}^0 \times \mathbf{P}$. Since $\mathbf{P}$ was known to be dualisable, this left the author to show that $\mathbf{R} \times \mathbf{P}$ and $\mathbf{R}^0 \times \mathbf{P}$ are dualisable. The dualisability of the finite members of $\mathcal{Q}_1$ would then follow by symmetry.

Showing that a failure of Theorem 5.2(iii) implies non-dualisability seemed relatively easy once we were actually aware of condition (iii). Proving this implication amounted to showing that if $\mathbb{V}(\mathbf{S})$ contains one of $\mathbf{N}_4$, $\mathbf{L}^+$, $\mathbf{L}^1$, $\mathbf{L} \times \mathbf{P}$, $\mathbf{R} \circlearrowleft 2$, then $\mathbf{S}$ is non-dualisable. Theorems 4.12 and 4.15 took care of two of these cases, leaving three to the author. Thus, we began with the following roadmap to a proof of Conjecture 5.3.

(1) Show that a finite inflation of a normal band is dualisable (presumably by showing that dualisability is preserved by taking inflations).
(2) Show that $\mathbf{R} \times \mathbf{P}$ and $\mathbf{R}^0 \times \mathbf{P}$ are dualisable.
(3) Show that if $\mathbb{V}(\mathbf{S})$ contains $\mathbf{L}^+$, $\mathbf{L} \times \mathbf{P}$, or $\mathbf{R} \circlearrowleft 2$, then $\mathbf{S}$ is non-dualisable.

Based on this outline, it did not seem (to the author, at least) that a proof of Conjecture 5.3 would suffice for a PhD thesis; proving (2) and (3) requires the consideration of five small semigroups, and (1) was expected to be obtained by a simple general result. But, of course, here we are. To quote Luke Skywalker, this is not going to go the way you think.

## 5.2. Two non-dualisability results

In this section, we will show that if $\mathbf{S}$ is a finite semigroup such that $\mathbb{V}(\mathbf{S})$ contains $\mathbf{L}^+$ or $\mathbf{R} \circlearrowleft 2$, then $\mathbf{S}$ is inherently non-dualisable. At the outset of the author's graduate studies, these two low-hanging pieces of fruit were overripe and simply needed to be nudged off of their stems. The only reason that these results had not been obtained earlier was due to the relative obscurity of the semigroups $\mathbf{L}^+$ and $\mathbf{R} \circlearrowleft 2$. These semigroups did not appear at all in Golubov and Sapir's proofs in [**32**], and they were very deeply embedded in McKenzie's proofs in [**47**]. Once the role of these two semigroups was clarified by the author, it was a simple matter to demonstrate their non-dualisability.

**Theorem 5.4.** *Let* $\mathbf{S}$ *be a finite semigroup such that* $\mathbf{L}^+ \in \mathbb{V}(\mathbf{S})$. *Then* $\mathbf{S}$ *is inherently non-dualisable.*

Proof. To show that $\mathbf{S}$ is inherently non-dualisable, it suffices to show that some finite member of $\mathbb{SP}(\mathbf{S})$ is inherently non-dualisable. By Lemma 1.11, $\mathbf{L}^+$ is a homomorphic image of some finite member of $\mathbb{SP}(\mathbf{S})$, so we can assume without loss of generality that $\mathbf{L}^+$ is in fact a homomorphic image of $\mathbf{S}$. We will apply Theorem 4.10.

Let $\varphi\colon \mathbf{S} \to \mathbf{L}^+$ be onto. Label the elements of $\mathbf{L}^+$ as in Figure 3.3 (page 49), and define $S_t := \varphi^{-1}(t)$ for all $t \in \{a, b, c, d\}$. Let $x$ be any element of $S_d$. Since $c$ is idempotent, $S_c$ is a subsemigroup of $\mathbf{S}$, so by Theorem 2.36, we may choose an idempotent $y \in S_c$. Define

$$\mathsf{a} := x, \quad \mathsf{b} := xy, \quad \mathsf{c} := x, \quad \mathsf{d} := yx, \quad \mathsf{e} := xx, \quad \mathsf{f} := xyx.$$

As $y$ is idempotent, we have

$$\mathsf{ac} = xx = \mathsf{e}, \quad \mathsf{ad} = xyx = \mathsf{f}, \quad \mathsf{bc} = xyx = \mathsf{f}, \quad \mathsf{bd} = xyyx = xyx = \mathsf{f}.$$

Now, let $\mathbf{P} = (\mathsf{P}, \mathsf{L}, \mathsf{I})$ be an infinite projective plane, and let $\mathsf{P}_\infty := \mathsf{P} \,\dot\cup\, \{\infty\}$. Define

$$B := \{\mathsf{b}^{\mathsf{a},\mathsf{a}}_{\infty,\ell} \mid \ell \in \mathsf{L}\} \subseteq S^{\mathsf{P}_\infty}, \qquad D := \{\mathsf{d}^{\mathsf{c},\mathsf{c}}_{\infty,\ell} \mid \ell \in \mathsf{L}\} \subseteq S^{\mathsf{P}_\infty},$$

and let $\mathbf{T}$ be the subsemigroup of $\mathbf{S}^{\mathsf{P}_\infty}$ generated by $B \cup D$.

Suppose by way of contradiction that $\mathsf{f}^{\mathsf{e}}_\infty \in T$. Since $\varphi(\mathsf{a}) = \varphi(\mathsf{c}) = d$ and $\varphi(\mathsf{f}) = b$, we have $\mathsf{a}, \mathsf{c} \neq \mathsf{f}$, so $\mathsf{f}^{\mathsf{e}}_\infty \notin B \cup D$, and we must therefore have $\mathsf{f}^{\mathsf{e}}_\infty = x_1 \cdots x_n$ for some $n \geqslant 2$ and some $x_1, \dots, x_n \in B \cup D$. Since $\mathsf{a} = \mathsf{c} = x$, there is some $p \in \mathsf{P}$ such that $x_1(p) = x_2(p) = x$, so $(x_1 \dots x_n)(p) \in xxS^1 \subseteq S_a$. Since $\mathsf{f} \notin S_a$, this contradicts $\mathsf{f}^{\mathsf{e}}_\infty = x_1 \cdots x_n$. Thus, $\mathsf{f}^{\mathsf{e}}_\infty \notin T$. Theorem 4.10 now gives the result. ☐

For the non-dualisability of $\mathbf{R} \circlearrowleft 2$, we will use the following result.

**Lemma 5.5.** *Let* $\mathbf{S}$ *be a finite semigroup, and let* $\varphi\colon \mathbf{S} \to \mathbf{R}$ *be an onto homomorphism. Then there is an embedding* $\psi\colon \mathbf{R} \to \mathbf{S}$ *such that* $\varphi \circ \psi$ *is the identity on* $\mathbf{R}$.

Proof. Extend $\varphi$ to a monoid homomorphism $\varphi^{+}\colon \mathbf{S}^1 \to \mathbf{R}^1$. By Lemma 4.14, there is a semigroup embedding $\psi^{+}\colon \mathbf{R}^1 \to \mathbf{S}^1$ such that $\varphi^{+} \circ \psi^{+}$ is the identity on $\mathbf{R}^1$. The restriction of $\psi^{+}$ to $\mathbf{R}$ is the required embedding. $\square$

**Theorem 5.6.** *Let* $\mathbf{S}$ *be a finite semigroup such that* $\mathbf{R} \circlearrowleft 2 \in \mathbb{V}(\mathbf{S})$. *Then* $\mathbf{S}$ *is inherently non-dualisable.*

Proof. We will apply Theorem 4.10. As in the proof of Theorem 5.4, we can assume there is an onto homomorphism $\varphi\colon \mathbf{S} \to \mathbf{R} \circlearrowleft 2$. Label the elements of $\mathbf{R} \circlearrowleft 2$ as in Figure 3.5 (page 72), and define the sets $S_t := \varphi^{-1}(t)$ for all $t \in \{0, a, b, u, v\}$.

By Lemma 5.5 applied to the restriction of $\varphi$ to $\varphi^{-1}(\{a, b\})$, there exist $\mathsf{a} \in S_a$ and $\mathsf{b} \in S_b$ such that $\{\mathsf{a}, \mathsf{b}\}$ is a right-zero subsemigroup of $\mathbf{S}$. Select some $\mathsf{c} \in S_v$, and define

$$\mathsf{d} := \mathsf{bc}, \quad \mathsf{e} := \mathsf{ac}, \quad \mathsf{f} := \mathsf{bc}.$$

Note that $\mathsf{e} \in S_0$ while $\mathsf{f} \in S_u$, so in particular $\mathsf{a}, \mathsf{c}, \mathsf{e} \neq \mathsf{f}$. We also have

$$\mathsf{ac} = \mathsf{e}, \quad \mathsf{ad} = \mathsf{abc} = \mathsf{bc} = \mathsf{f}, \quad \mathsf{bd} = \mathsf{bbc} = \mathsf{bc} = \mathsf{f}.$$

Now, let $\mathbf{P} = (\mathsf{P}, \mathsf{L}, \mathsf{I})$ be an infinite projective plane, and let $\mathsf{P}_\infty := \mathsf{P} \mathbin{\dot\cup} \{\infty\}$. Define

$$B := \{\mathsf{b}^{\mathsf{a},\mathsf{a}}_{\infty,\ell} \mid \ell \in \mathsf{L}\} \subseteq S^{\mathsf{P}_\infty}, \qquad D := \{\mathsf{d}^{\mathsf{c},\mathsf{c}}_{\infty,\ell} \mid \ell \in \mathsf{L}\} \subseteq S^{\mathsf{P}_\infty},$$

and let $\mathbf{T}$ be the subsemigroup of $\mathbf{S}^{\mathsf{P}_\infty}$ generated by $B \cup D$.

Suppose by way of contradiction that $\mathsf{f}^{\mathsf{e}}_{\infty} \in T$. Clearly $\mathsf{f}^{\mathsf{e}}_{\infty} \notin B \cup D$ since $\mathsf{a}, \mathsf{c} \neq \mathsf{f}$, so we may write $\mathsf{f}^{\mathsf{e}}_{\infty} = x_1 \cdots x_n$ for some $n \geqslant 2$ and some $x_1, \dots, x_n \in B \cup D$. Now, if we had $D \cap \{x_1, \dots, x_{n-1}\} \neq \varnothing$, then $x_1 \cdots x_n$ would lie in $S_0^{\mathsf{P}_\infty}$; thus, $x_1, \dots, x_{n-1} \in B$. But $B$ is a right-zero semigroup, so we can assume that $n = 2$. Thus, we have $\mathsf{f}^{\mathsf{e}}_{\infty} = x_1 x_2$, where $x_1 \in B$ and $x_2 \in B \cup D$. However, $x_2 \in B$ would imply $\mathsf{f}^{\mathsf{e}}_{\infty} = x_2 \in B$; thus, $x_2 \in D$. Now, write $x_1 = \mathsf{b}^{\mathsf{a},\mathsf{a}}_{\infty,\ell}$ and $x_2 = \mathsf{d}^{\mathsf{c},\mathsf{c}}_{\infty,k}$ for some $\ell, k \in \mathsf{L}$, and let $p$ be a point in $\ell \cap k$. Then $(x_1 x_2)(p) = \mathsf{ac} \neq \mathsf{f}$, contradicting $\mathsf{f}^{\mathsf{e}}_{\infty} = x_1 x_2$. Theorem 4.10 now gives the result. $\square$

One of the advantages of Theorem 4.10 is that it takes care of much of the creativity needed to apply the more general Theorem 4.8. However, Theorem 4.10 is not as widely applicable as Theorem 4.8. For example, it cannot be used to show that $\mathbf{R}^1$ is IND, since none of the possible ways of choosing $\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}, \mathsf{e}, \mathsf{f}$ satisfy the assumptions of Theorem 4.10.

After obtaining the results in this section, it was quickly seen that Theorem 4.10 is not applicable to $\mathbf{L} \times \mathbf{P}$, so this semigroup would have to be dealt with ‘from scratch’. This proved to be rather difficult, so, after several unsuccessful attempts to show that $\mathbf{L} \times \mathbf{P}$ is non-dualisable, the author moved on to dualisability.

## 5.3. Dualities for inflations

In this section, we will show that all finite semigroups in $\mathcal{N}_1 = \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N})$ are dualisable. The main result of this section, Theorem 5.9, will be applicable to algebras of any type. We will therefore define inflations in this general setting, although the concept does not seem to be particularly useful outside of semigroup theory.

Let $\mathbf{B}$ be an algebra of type $F$, let $\mathbf{A} \leqslant \mathbf{B}$, and let $^{-}\colon \mathbf{B} \to \mathbf{A}$ be a retraction onto $\mathbf{A}$ (see Section 4.4). Then $\mathbf{B}$ is called an **inflation of A via** $^{-}$ if for every $n \in \omega$ and every $n$-ary $f \in F$, we have $f^{\mathbf{B}}(x_0, \dots, x_{n-1}) = f^{\mathbf{A}}(\overline{x_0}, \dots, \overline{x_{n-1}})$ for all $(x_0, \dots, x_{n-1}) \in B^n$. One may easily show that if $t$ is an $n$-ary term involving at least one operation symbol, then we have $t^{\mathbf{B}}(x_0, \dots, x_{n-1}) = t^{\mathbf{A}}(\overline{x_0}, \dots, \overline{x_{n-1}})$ for all $(x_0, \dots, x_{n-1}) \in B^n$.

Since $\mathcal{N}_1$ satisfies $x^4 \approx x^2$, we have $\mathrm{E}(\mathbf{S}) = \{x^2 \mid x \in S\}$ for every $\mathbf{S} \in \mathcal{N}_1$. As $\mathcal{N}_1$ also satisfies $(xy)^2 \approx x^2y^2 \approx xy$, we have the following result.

**Proposition 5.7.** *Let* $\mathbf{S} \in \mathcal{N}_1$*. Then* $\mathbf{S}$ *is an inflation of the normal band* $\mathrm{E}(\mathbf{S})$ *via the mapping* $x \mapsto x^2$*.*

Theorem 5.9 will show that if $\mathbf{A}$ is a finite algebra possessing a certain kind of unary term, and if $\mathbf{B}$ is a finite inflation of $\mathbf{A}$, then an (IC) duality for $\mathbf{A}$ can be lifted to an (IC) duality for $\mathbf{B}$. Given that all normal bands are dualisable via (IC), it will follow immediately that all finite semigroups in $\mathcal{N}_1$ are dualisable.

The following result will allow us to utilise existing (IC) dualities. It is adapted from a result of Clark, Davey, and Pitkethly [**6**, Lemma 24] (see also Pitkethly and Davey [**54**, Lemma 2.3.2]). Note that the symbol $\Vdash$ was introduced in Section 4.5.

**Lemma 5.8** (Term Retract Lemma)**.** *Let* $\mathbf{S}$ *be a finite algebra, let* $\mathbf{A} \leqslant_{\mathrm{r}} \mathbf{S}$*, and let* $r\colon \mathbf{S} \to \mathbf{A}$ *be a retraction onto* $\mathbf{A}$ *that is also a unary term function of* $\mathbf{S}$*. Let* $\underset{\sim}{\mathbf{S}}$ *and* $\underset{\sim}{\mathbf{A}} = \langle A; G, R, \mathcal{T}\rangle$ *be alter egos of* $\mathbf{S}$ *and* $\mathbf{A}$*, respectively, and assume that* $\underset{\sim}{\mathbf{A}}$ *satisfies* (IC) *with respect to* $\mathbf{A}$ *and that* $\underset{\sim}{\mathbf{S}} \Vdash G \cup R \cup \{r\}$*. If* $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ *for some* $n \in \omega\backslash\{0\}$ *and* $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ *is a morphism such that* $\varphi(X) \subseteq A$*, then* $\varphi$ *extends to an* $n$*-ary term function of* $\mathbf{S}$*.*

Proof. Note that $A^{\mathbf{X}} = X \cap A^n$ forms a substructure, $\mathbf{X}_A$, of $\underset{\sim}{\mathbf{A}}^n$. Since $\varphi(A^{\mathbf{X}}) \subseteq A$ by assumption, it follows that $\varphi{\restriction}_{A^{\mathbf{X}}}$ is a morphism $\mathbf{X}_A \to \underset{\sim}{\mathbf{A}}$. As $\underset{\sim}{\mathbf{A}}$ satisfies (IC) with respect to $\mathbf{A}$, there is a term $t$ in the signature of $\mathbf{A}$ and $\mathbf{S}$ such that $t^{\mathbf{A}}$ extends $\varphi{\restriction}_{A^{\mathbf{X}}}$.

Let $x \in X$. Then $r(x) \in A^{\mathbf{X}}$, so because $\varphi(X) \subseteq A$ and $\varphi$ preserves $r$, we have

$$\varphi(x) = r(\varphi(x)) = \varphi(r(x)) = t^{\mathbf{A}}(r(x)) = t^{\mathbf{S}}(r(x)).$$

Thus, $\varphi$ is extended by the term function $(x_0, \dots, x_{n-1}) \mapsto t^{\mathbf{S}}(r(x_0), \dots, r(x_{n-1}))$ of $\mathbf{S}$. $\square$

The author's original proof of the dualisability result for $\mathcal{N}_1$ was carried out by showing that each of its subquasivarieties has a dualisable generator; this required eight separate dualisability proofs taking up some 20 to 30 pages. It was hoped that this 'brute force' approach would lead to a general proof that a finite inflation of a normal band is dualisable, but such a proof eluded the author for some time.

Nearly a year later, the author visited Ross Willard for a research collaboration, and as a warm-up exercise of sorts, we attempted to show that the two possible one-point inflations of the two-element group are dualisable (incidentally, these were the final two semigroups of order three whose dualisability had not yet been determined). This successful attempt finally led the author to Theorem 5.9, which brings the dualisability proof for $\mathcal{N}_1$ down to less than a page. The author wishes to acknowledge Ross Willard's involvement in the discovery of the following proof.

**Theorem 5.9.** *Let* $\mathbf{A}$ *be a finite algebra of type* $F$ *with a unary term* $\iota$*, involving at least one operation symbol, such that* $\mathbf{A} \models \iota(x) \approx x$*. Suppose that* $\underset{\sim}{\mathbf{A}} := \langle A; \mathcal{P}_m(\mathbf{A}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{A}$ *for some* $m \in \omega$*. If* $\mathbf{B}$ *is a finite inflation of* $\mathbf{A}$*, then* $\underset{\sim}{\mathbf{B}} := \langle B; \mathcal{P}_{m'}(\mathbf{B}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{B}$*, where* $m' = \max\{m, 2\}$*.*

Proof. Assume $\mathbf{B}$ is an inflation of $\mathbf{A}$ via $^{-}\colon \mathbf{B} \to \mathbf{A}$. Note that $\overline{x} = \iota^{\mathbf{B}}(x)$ for all $x \in B$, since $\iota$ involves at least one operation symbol; thus, $^{-}$ is a unary term function of $\mathbf{B}$.

Let $n \in \omega \backslash \{0\}$, let $\mathbf{X} \leqslant \underset{\sim}{\mathbf{B}}^n$, and let $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{B}}$ be a morphism; we must show that $\varphi$ extends to an $n$-ary term function of $\mathbf{B}$. If $\varphi(X) \subseteq A$, the result follows by taking $r$ to be $^{-}$ in the Term Retract Lemma 5.8. Thus, we may assume that there is some $c \in \varphi(X)\backslash A$. We will show that $\varphi$ is extended by a projection term.

For every $b \in B$, define the total binary operation $*_b$ on $B$ by

$$(\forall x, y \in B)\ \ x *_b y := \begin{cases} c & \text{if } (x, y) = (b, c) \\ \overline{y} & \text{otherwise.} \end{cases}$$

The operation $*_b$ is "almost the second projection"; we have $\overline{x *_b y} = \overline{y}$ for all $x, y \in B$. To show that $*_b \in \mathcal{T}_2(\mathbf{B})$, let $f \in F$ be $k$-ary, and let $x, y \in B^k$. Then $f^{\mathbf{B}}(y) \in A$, so

$$f^{\mathbf{B}}(x) *_b f^{\mathbf{B}}(y) = f^{\mathbf{B}}(y) = f^{\mathbf{B}}(\overline{y}) = f^{\mathbf{B}}\big(\overline{x *_b y}\big) = f^{\mathbf{B}}(x *_b y).$$

Thus, $*_b \in \mathcal{T}_2(\mathbf{B})$, and therefore $\underset{\sim}{\mathbf{B}} \Vdash *_b$. Also, since $B\backslash\{c\} \in \mathcal{R}_1(\mathbf{B})$, we have $\underset{\sim}{\mathbf{B}} \Vdash B\backslash\{c\}$.

Let $\widehat{c}$ be a $(*_c)$-product of all of the elements of $\varphi^{-1}(c)$, in any order and any bracketing. Then $\varphi(\widehat{c}) = c$. Since $\varphi$ preserves the unary relation $B\backslash\{c\}$, there exists $k \in n$ such that $\widehat{c}(k) = c$. By the definitions of $\widehat{c}$ and $*_c$, this implies that $x(k) = c$ for all $x \in \varphi^{-1}(c)$.

Let $x \in X$, and let $b := \varphi(x)$. Then $\varphi(x *_b \widehat{c}) = b *_b c = c$, so $(x *_b \widehat{c})(k) = c$ by the choice of $k$. This is possible only if $x(k) = b$. Thus, $\varphi$ is extended by the $k$th projection. □

To apply Theorem 5.9 to a finite normal band $\mathbf{A}$, we may choose $\iota$ to be the term $x^2$ and use the (IC) duality for $\mathbf{A}$ from Theorem 4.42 to conclude that every finite inflation of $\mathbf{A}$ is dualisable. Combining this with Proposition 5.7, we obtain the following result.

**Corollary 5.10.** *Let* $\mathbf{S}$ *be a finite semigroup in* $\mathbf{N}_1 = \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N})$*. Then* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathfrak{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$*, and consequently* $\mathbf{S}$ *is dualisable.*

In fact, Theorem 5.9 can also be applied to finite completely regular semigroups. This more general case will be discussed in the epilogue (Section 7.1).

## 5.4. Critical discoveries

After obtaining the results of Sections 5.2 and 5.3, there were only three semigroups left to consider, according the roadmap in Section 5.1. The non-dualisability result for $\mathbf{L} \times \mathbf{P}$ remained elusive, so the author's efforts were focused on showing that the semigroup $\mathbf{R}^0 \times \mathbf{P}$ is dualisable. To simplify matters, the author considered the semigroup $\mathbf{R} \circlearrowright \{u\}$, where $\mathbf{R}$ acts totally on the one-element set $\{u\}$, as this four-element semigroup generates the same quasivariety as the nine-element semigroup $\mathbf{R}^0 \times \mathbf{P}$. Even this four-element semigroup proved to be quite difficult to deal with, but after at least a month of calculations, the author showed that $\mathbf{R} \circlearrowright \{u\}$, and therefore $\mathbf{R}^0 \times \mathbf{P}$, is dualisable.

Of course, dualisability is not in general preserved by taking direct products (for example, by Theorem 4.12 and Lemma 3.9, $\mathbf{P} \times \mathbf{Q}$ is non-dualisable, though $\mathbf{P}$ and $\mathbf{Q}$ are). The dualisability of $\mathbf{R}^0 \times \mathbf{P}$ seems to come about from a certain link between $\mathbf{R}^0$ and $\mathbf{P}$, though, even now, it is not clear to the author what this means in precise terms. Nonetheless, having discovered the 'trick' for dealing with this direct product, the dualisability of $\mathbf{R} \times \mathbf{P}$ was obtained with minor modifications to the proof for $\mathbf{R}^0 \times \mathbf{P}$.

With these dualisability proofs, it appeared that we had obtained one converse half of Conjecture 5.3 (generating a residually small variety implies dualisability). However, the result depended on having a complete enumeration of the quasicriticals of $\mathcal{P}_1$. As mentioned in Section 5.1, this could be derived from Sapir's result [**60**, Lemma 3.8], which described the quasicriticals in $\mathcal{P}_n$. Since Sapir did not include a proof of this lemma, the author attempted to prove it for $\mathcal{P}_1$. In the course of this attempted proof, the table in Figure 5.1 arose quite naturally; the goal from there was to show that $a = b$, yielding $\mathbf{P}$. But, to the author's surprise, the table in Figure 5.1 already yielded a quasicritical semigroup in $\mathcal{P}_1$.

| $\cdot$ | $e$ | $a$ | $u$ | $b$ |
|---|---|---|---|---|
| $e$ | $e$ | $a$ | $u$ | $b$ |
| $a$ | $a$ | $a$ | $b$ | $b$ |
| $u$ | $a$ | $a$ | $b$ | $b$ |
| $b$ | $a$ | $a$ | $b$ | $b$ |

FIGURE 5.1. The semigroup $\mathbf{C}$.

According to the statement of [**60**, Lemma 3.8], the minimum ideal of a quasicritical in $\mathcal{P}_n$ other than $\mathbf{R}$ is always an Abelian group, which implies in particular that all quasicriticals in $\mathcal{P}_1$ other than $\mathbf{R}$ have zero elements. However, the semigroup in Figure 5.1 does not have a zero element, so it is not of the form described in [**60**, Lemma 3.8], which shows that this result is false as stated. The author named this counterexample $\mathbf{C}$.

Of course, this meant that we did not have a complete enumeration of the quasicriticals in $\mathcal{P}_1$. Worse, there was no longer any guarantee that there are only finitely many quasicriticals in $\mathcal{P}_1$; the result implying this depended on [**60**, Lemma 3.8]! This prompted a frantic search for more quasicriticals. It was seen almost immediately that $\mathbf{C}^0 \in \mathcal{P}_1$ is quasicritical, but a computer search ruled out the existence of quasicriticals of size six or seven, so it seemed plausible that $\mathbf{C}$ and $\mathbf{C}^0$ were the only new quasicriticals. The author eventually obtained a proof of this, to be published in [**42**]. We will defer the proof until later.

**Theorem 5.11.** *Up to isomorphism, the only quasicriticals in* $\mathcal{P}_1$ *are* $\mathbf{R}$, $\mathbf{R}^0$, $\mathbf{I}$, $\mathbf{N}$, $\mathbf{P}$, $\mathbf{C}$, *and* $\mathbf{C}^0$.

The structure of the lattice $\mathrm{L_q}(\mathcal{P}_1)$ is also determined in [**42**]; it is shown here in Figure 5.2. The results of [**42**] were presented at SandGAL 2019 in Cremona, Italy.

In the aperiodic case, Theorem 5.11 patched up the main semigroup-theoretic result of Sapir in [**60**], which was intended to characterise the locally finite semigroup varieties that have finitely many subquasivarieties. It is now an open problem to repair the proof in the non-aperiodic case, but we will not address this problem here.

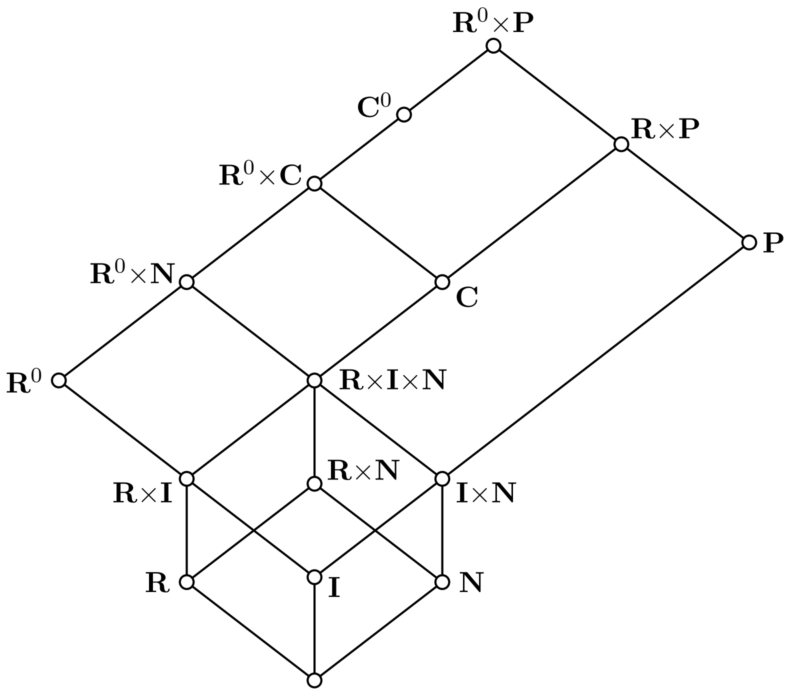



FIGURE 5.2. The lattice $\mathrm{L_q}(\mathcal{P}_1)$, labelled by generators.

The addition of the two quasicriticals $\mathbf{C}$ and $\mathbf{C}^0$ added only three more quasivarieties not already considered; namely, those generated by $\mathbf{C}$, $\mathbf{C}^0$, and $\mathbf{R}^0 \times \mathbf{C}$. The dualisability proofs required some new ideas, but drawing from the proof for $\mathbf{R}^0 \times \mathbf{P}$ allowed these three new dualisability proofs to be obtained without too much difficulty. Thus, we finally had a proof of one converse half of Conjecture 5.3. This result was presented at the 98th AAA conference in Dresden, Germany. Again, we defer the proof.

**Theorem 5.12.** *Let* $\mathbf{S}$ *be a finite aperiodic semigroup. If* $\mathbb{V}(\mathbf{S})$ *is residually small, then* $\mathbf{S}$ *is dualisable.*

In the months leading up these conferences, the author had also been working on the semigroup $\mathbf{L} \times \mathbf{P}$ with the first supervisor. We proceeded via a standard approach: attempt to prove that $\mathbf{L} \times \mathbf{P}$ is dualisable, and then try to understand why it is not dualisable. Many of the techniques used for $\mathbf{R} \times \mathbf{P}$ were applicable to $\mathbf{L} \times \mathbf{P}$, but the essential difference with the $\mathbf{R} \times \mathbf{P}$ case quickly became clear. It seemed that we had identified the obstruction to dualisability, so it was only a matter of time before we obtained a non-dualisability proof.

However, upon returning home from the conferences, the author received some surprising news from the first supervisor: he had shown that $\mathbf{L} \times \mathbf{P}$ is in fact dualisable. It turned out that the apparent obstruction to dualisability could easily be tempered with a particularly simple partial operation. In fact, the overall dualisability proof for $\mathbf{L} \times \mathbf{P}$ was much simpler than the proof for $\mathbf{R} \times \mathbf{P}$, and there was no doubt that it was correct.

Not only did this disprove Conjecture 5.3, but it significantly expanded the scope of the dualisability problem. The first step from here was to figure out exactly what we knew about the class of dualisable finite aperiodic semigroups (which we will refer to as the **dualisable class**). Based on Theorems 5.2 and 5.12, we knew that the dualisable class contains all finite memebers of $\mathbf{N}_1 \cup \boldsymbol{\mathcal{P}}_1 \cup \boldsymbol{\mathfrak{Q}}_1$, so this provided us with a lower boundary. Determining an upper boundary required the non-dualisability results we already had, as well as an in-depth knowledge of the results of Chapter 3.

For possible future generalisations, we present the next result without the assumption of aperiodicity. Note that, by Theorems 4.12, 4.13, 4.15, 5.4, and 5.6, the assumptions of the next theorem are satisfied by a dualisable finite semigroup (more generally, a non-IND finite semigroup).

**Theorem 5.13.** *Let* $\mathbf{S}$ *be a semigroup such that* $\mathbb{V}(\mathbf{S})$ *does not contain any member of the class* $\mathbb{A}(\{\mathbf{N}_4, \mathbf{L}^+, \mathbf{L}^1, \mathbf{R} \circlearrowleft 2, \mathbf{M}_p \mid p \text{ prime}\})$. *Then* $\mathbf{S} \in \mathbb{A}([x^{n+1}y \approx xy,\ xy^nz^nt \approx xz^ny^nt])$ *for some* $n \geqslant 1$.

Proof. By Theorem 3.23, we can assume by symmetry that $\mathbf{S}$ satisfies $x^{n+1}y \approx xy$ for some $n \geqslant 1$. Since $x^2y \approx xy$ implies $x^3y \approx xy$, we can assume further that $n \geqslant 2$, so $n$ is an exponent of $\mathbf{S}$. Now, if $\mathbf{S}$ also satisfies $xy^{n+1} \approx xy$, then $\mathbf{S} \models xy^nz^nt \approx xz^ny^nt$ by Lemma 3.38 and we are done. We may therefore assume that $\mathbf{S} \not\models xy^{n+1} \approx xy$.

We have $\mathbf{P} \in \mathbb{V}(\mathbf{S})$ by Lemma 3.45, so $\mathbf{S} \models x^ny^nz \approx x^ny^nx^nz$ by Lemma 3.65. Now, as $\mathbf{N}_4 \not\models x^{n+1}y \approx xy$, Theorems 3.14,3.26, and 3.27 imply that $\mathrm{E}(\mathbf{S})$ is a normal band, so $\mathbf{S} \models x^ny^nz^nt^n \approx x^nz^ny^nt^n$. Combining this with $x^{n+1}y \approx xy$ and $x^ny^nz \approx x^ny^nx^nz$, we deduce $\mathbf{S} \models xy^nz^nt \approx xx^ny^nz^nt \approx xx^ny^nz^ny^nt \approx xx^nz^ny^ny^nt \approx xx^nz^ny^nt \approx xz^ny^nt$. □

Restricting to the aperiodic case, we get the following result.

**Theorem 5.14.** *Let* $\mathbf{S}$ *be an aperiodic semigroup with* $\mathbb{V}(\mathbf{S}) \cap \mathbb{A}(\mathbf{N}_4, \mathbf{L}^+, \mathbf{L}^1, \mathbf{R} \circlearrowleft 2) = \varnothing$. *Then* $\mathbf{S} \in \mathbb{A}([x^2y \approx xy,\ xyzt \approx xzyt])$.

Proof. By Theorem 5.13, we have $\mathbf{S} \in \mathbb{A}([x^{n+1}y \approx xy,\ xy^nz^nt \approx xz^ny^nt])$ for some $n \geqslant 1$. Thus, $\mathbf{S} \models x^{n+2} \approx x^2$, so by Proposition 5.1 we have $\mathbf{S} \models x^3 \approx x^2$. Using this latter identity, $x^{n+1}y \approx xy$ ($xy^{n+1} \approx xy$) can be reduced to $x^2y \approx xy$ ($xy^2 \approx xy$), and using either of $x^2y \approx xy$, $xy^2 \approx xy$, we may reduce $xy^nz^nt \approx xz^ny^nt$ to $xyzt \approx xzyt$. □

Theorem 5.14 provided us with an upper bound for the dualisable class. We were now reduced to identifying the dualisable finite members of the variety $[x^2y \approx xy,\ xyzt \approx xzyt]$. This variety had not been previously studied, but it turned out to be generated by a set of familiar semigroups.

Using Proposition 3.59, it is easily seen that a semigroup identity follows from $x^2y \approx xy$ and $xyzt \approx xzyt$ if and only if it holds in $\mathbf{L}$, $\mathbf{R}$, and $\mathbf{P}$, so we have the following result.

**Proposition 5.15.** *The variety* $[x^2y \approx xy,\ xyzt \approx xzyt]$ *is precisely* $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$.

Shortly after obtaining the dualisability result for $\mathbf{L}\times\mathbf{P}$, the author showed that $\mathbf{L}\times\mathbf{R}\times\mathbf{P}$ is also dualisable, implying that there is no proper subvariety of $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ containing all of its dualisable members. Of course, there is no reason that the dualisable members ought to constitute the finite part of some subvariety, but since there were no obvious candidates for non-dualisable members of $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, our best conjecture was the following.

**Conjecture 5.16.** *Let* $\mathbf{S}$ *be a finite aperiodic semigroup. Then* $\mathbf{S}$ *is dualisable if and only if* $\mathbf{S} \in \mathbb{A}\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$.

By Theorem 3.63, the variety $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ is residually large, and the proof shows that it has infinitely many finite subdirectly irreducibles (though this latter fact follows from a more general result [**56**]). These finite subdirectly irreducibles generate pairwise distinct quasivarieties, so $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ has infinitely many finitely generated subquasivarieties. Thus, in contrast to Theorem 5.12, there was no hope of proving Conjecture 5.16 by dualising only a finite number of semigroups. Considering the effort it took to dualise five small semigroups for the proof of Theorem 5.12, solving the dualisability problem for aperiodic semigroups now seemed like an insurmountable task.

Of course, the only way forward in this situation is to consider the easiest unknown examples. Since quasicriticals are the fundamental building blocks for locally finite quasivarieties, it made sense to search for the smallest quasicriticals whose dualisability was not yet known. The semigroup $\mathbf{L} \circlearrowleft L$ in Figure 5.3 came to mind almost immediately; it is obtained from the action of $\mathbf{L}$ by left translation on its underlying set.

| $\cdot$ | $a$ | $b$ | $u$ | $v$ | $0$ |
|---|---|---|---|---|---|
| $a$ | $a$ | $a$ | $u$ | $u$ | $0$ |
| $b$ | $b$ | $b$ | $v$ | $v$ | $0$ |
| $u$ | $0$ | $0$ | $0$ | $0$ | $0$ |
| $v$ | $0$ | $0$ | $0$ | $0$ | $0$ |
| $0$ | $0$ | $0$ | $0$ | $0$ | $0$ |

FIGURE 5.3. The semigroup $\mathbf{L} \circlearrowleft L$.

The semigroup $\mathbf{L} \circlearrowleft L$ arises from the construction in Theorem 3.63 when $\kappa = 0$, which shows that $\mathbf{L} \circlearrowleft L$ is a subdirectly irreducible (hence quasicritical) semigroup in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$.

Considering the relationship between $\mathbf{P}$ and $\mathbf{C}$, it was clear how to construct a quasicritical of size six. Note that $\mathbf{P}$ can be obtained from $\mathbf{C}$ as the Rees quotient $\mathbf{C}/\{a, b\}$, and that factoring $\mathbf{C}$ by the partition $\{\{e, a\}, \{u, b\}\}$ yields the two-element right-zero semigroup. Although $\mathbf{C}$ was discovered organically, it can in hindsight be constructed with these properties in mind, and applying this same construction to $\mathbf{L} \circlearrowleft L$ turns out to yield another quasicritical.

| $\cdot$ | $a$ | $b$ | $c$ | $u$ | $v$ | $d$ |
|---|---|---|---|---|---|---|
| $a$ | $a$ | $a$ | $c$ | $u$ | $u$ | $d$ |
| $b$ | $b$ | $b$ | $c$ | $v$ | $v$ | $d$ |
| $c$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |
| $u$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |
| $v$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |
| $d$ | $c$ | $c$ | $c$ | $d$ | $d$ | $d$ |

FIGURE 5.4. The semigroup $\mathbf{E}$.

The idea is to split the zero element of $\mathbf{L} \circlearrowleft L$ into a two-element ideal $\{c, d\}$, so that the quotient by this ideal is isomorphic to $\mathbf{L} \circlearrowleft L$. If we require also that $c$ and $d$ act respectively as zeroes for $\{a, b\}$ and $\{u, v\}$, and that the quotient by $\{\{a, b, c\}, \{u, v, d\}\}$ is a right-zero semigroup, this uniquely determines a multiplication on $\{a, b, u, v, c, d\}$, given by Figure 5.4. The resulting semigroup will be called $\mathbf{E}$.

The structure in Figure 5.4 can be factored by the partitions $\{\{a\}, \{b\}, \{u\}, \{v\}, \{c, d\}\}$ and $\{\{a, b, c\}, \{u, v, d\}\}$ to yield $\mathbf{L} \circlearrowright L$ and $\mathbf{R}$, so it is a subdirect product of $\mathbf{L} \circlearrowright L$ and $\mathbf{R}$. This confirms that $\mathbf{E}$ is a semigroup, and shows that $\mathbf{E} \in \mathbb{SP}(\mathbf{R}, \mathbf{L} \circlearrowright L) \subseteq \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$. (A similar argument shows that $\mathbf{C}$ is a semigroup in $\mathbb{SP}(\mathbf{R}, \mathbf{P})$.)

Just as adding a zero to $\mathbf{C}$ yields another quasicritical, it turns out that $\mathbf{E}^0$ is quasicritical. In general, if $\mathbf{S}$, $\mathbf{T}$, $\mathbf{U}$ are semigroups with $\mathbf{S} \in \mathbb{SP}(\mathbf{T}, \mathbf{U})$, then $\mathbf{S}^0 \in \mathbb{SP}(\mathbf{T}^0, \mathbf{U}^0)$. Since $\mathbf{P}$ and $\mathbf{L} \circlearrowright L$ already have zero elements, we have $\mathbf{C}^0 \in \mathbb{SP}(\mathbf{R}^0, \mathbf{P})$ and $\mathbf{E}^0 \in \mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowright L)$.

**Proposition 5.17.** *The semigroups* $\mathbf{C}$, $\mathbf{C}^0$, $\mathbf{E}$, $\mathbf{E}^0$ *are quasicritical, and*

$$\mathbf{C} \in \mathbb{SP}(\mathbf{R}, \mathbf{P}), \quad \mathbf{C}^0 \in \mathbb{SP}(\mathbf{R}^0, \mathbf{P}), \quad \mathbf{E} \in \mathbb{SP}(\mathbf{R}, \mathbf{L} \circlearrowright L), \quad \mathbf{E}^0 \in \mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowright L).$$

Proof. We have already justified the containments, so it remains to show that the four semigroups are quasicritical. We will prove only that $\mathbf{E}^0$ is quasicritical; the other cases are analogous but simpler.

We will show that $(u, v)$ is a critical pair of $\mathbf{E}^0$. Let $\varphi\colon \mathbf{E}^0 \to \mathbf{E}^0$ be a homomorphism with $\varphi(u) \neq \varphi(v)$; it will suffice to show that $\varphi$ is an automorphism.

Note that partition of $\mathbf{E}^0$ into $\mathcal{L}$-classes is $\{\{a, b\}, \{u, v\}, \{c\}, \{d\}, \{0\}\}$ (see Example 2.3). We therefore have $\varphi(u) \mathrel{\mathcal{L}} \varphi(v)$, as homomorphisms preserve Green's relations. By inspecting the $\mathcal{L}$-classes, we see that $\varphi(u)$ and $\varphi(v)$ are either both idempotent or $\{\varphi(u), \varphi(v)\} = \{u, v\}$. However, if $\varphi(u)$ and $\varphi(v)$ are both idempotent, then

$$\varphi(u) = \varphi(u)^2 = \varphi(u^2) = \varphi(d) = \varphi(v^2) = \varphi(v)^2 = \varphi(v),$$

contradicting the choice of $\varphi$. Thus, $\{\varphi(u), \varphi(v)\} = \{u, v\}$. By composing $\varphi$ with the automorphism of $\mathbf{E}^0$ interchanging $(a, u)$ and $(b, v)$, we can assume that $\varphi(u) = u$ and $\varphi(v) = v$.

Now, we have $\varphi(a)u = \varphi(a)\varphi(u) = \varphi(au) = \varphi(u) = u$, so $\varphi(a)$ is a left identity for $u$, implying that $\varphi(a) = a$. By symmetry, $\varphi(b) = b$. Since $\{a, b, u, v\}$ generates $\mathbf{E}$, it follows that $\varphi{\restriction}_E$ is the identity map on $E$. Finally, since $\varphi(0)$ is a zero for $\varphi(E^0)$ and $\mathbf{E}$ does not have a zero, we must have $\varphi(E^0) \neq E$, and therefore $\varphi(0) = 0$.

By deleting parts of the above argument, one easily shows that $\mathbf{E}$ is quasicritical. For $\mathbf{C}$ and $\mathbf{C}^0$, use a similar proof, taking $(u, b)$ as the critical pair. □

The quasicritical normal bands, along with $\mathbf{N}$, $\mathbf{P}$, $\mathbf{C}$, $\mathbf{C}^0$, $\mathbf{L} \circlearrowright L$, $\mathbf{E}$, and $\mathbf{E}^0$, are the only quasicriticals in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ of size at most seven (though we will not need to prove this fact just now). Finding dualities for these semigroups, and various direct products thereof, was the author's starting point for proving the newly formulated Conjecture 5.16.

We can now state the result that will occupy us for the rest of the chapter: All finite direct products of these 12 quasicriticals are dualisable. Although some of these products are quite complicated, many of the techniques used to dualise $\mathbf{L} \times \mathbf{P}$ and $\mathbf{R} \times \mathbf{P}$ will be applicable, so proving the main will result will primarily be a matter of perseverance. The next largest quasicritical (obtained by taking $\kappa = 1$ in the proof of Theorem 3.63) introduces a new configuration, present in all larger quasicriticals; this situation will be dealt with in the next chapter.

## 5.5. Meet operations

As explained in Section 5.4, our goal for the remainder of this chapter is to show that all finite direct products of members of the set

$$\mathcal{C} := \{\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0, \mathbf{I}, \mathbf{N}, \mathbf{P}, \mathbf{C}, \mathbf{C}^0, \mathbf{L} \circlearrowright L, \mathbf{E}, \mathbf{E}^0\}$$

are dualisable. When working with semigroups from $\mathcal{C}$, we will often use without mention the identities $x^2y \approx xy$ and $xyzt \approx xzyt$ (see Proposition 5.15).

We define the subset $\mathcal{P}$ of $\mathcal{C}$ by

$$\mathcal{P} := \{\mathbf{P}, \mathbf{C}, \mathbf{C}^0, \mathbf{L} \circlearrowright L, \mathbf{E}, \mathbf{E}^0\}.$$

The members of $\mathcal{P}$ will be called **P-like** semigroups (because they have $\mathbf{P}$ as a homomorphic image). By Corollary 5.10, we know that all finite products of semigroups in $\mathcal{C} \backslash \mathcal{P}$ are dualisable, so the $\mathbf{P}$-like semigroups will be at the heart of our dualisability proofs.

In this section, we will construct compatible operations based on the $\mathbf{P}$-like semigroups; these will of course be used in our dualisability proofs in later sections. The proofs will often require the reader to perform case-checks over the $\mathbf{P}$-like semigroups; for this reason, we recommend that the reader first familiarises themself with the members of $\mathcal{P}$, for example by determining Green's relations in each of these semigroups and filling in the details of the proof of Proposition 5.17. Determining the embedding order on $\mathcal{P}$ would also be worthwhile.

**Theorem 5.18.** *Let $\mathbf{S} \in \mathbb{I}(\mathcal{P})$, let $w \in S \backslash \mathrm{E}(\mathbf{S})$, and let $f \in S$ be the left identity for $w$. Then there is a meet operation $\wedge \in \mathcal{T}_2(\mathbf{S})$ such that $f$ and $w$ are maximal in $\langle S; \wedge \rangle$, the $\mathcal{L}$-classes $L_f$ and $L_w$ are up-sets of $\langle S; \wedge \rangle$, and left translation by $f$ is an endomorphism of $\langle S; \wedge \rangle$.*

Proof. Label the elements of $\mathbf{P}$, $\mathbf{C}$, $\mathbf{L} \circlearrowright L$, and $\mathbf{E}$ as in Figures 3.2, 5.1, 5.3, and 5.4.

If $\mathbf{S} \in \{\mathbf{P}, \mathbf{C}, \mathbf{C}^0\}$, then the only choice for $(f, w)$ is $(e, u)$, while if $\mathbf{S} \in \{\mathbf{L} \circlearrowright L, \mathbf{E}, \mathbf{E}^0\}$, then $(f, w)$ can be either $(a, u)$ or $(b, v)$. In the latter case, however, there is an automorphism of $\mathbf{S}$ interchanging $(a, u)$ with $(b, v)$, so we can assume in this case that $(f, w) = (a, u)$. Thus, for the six possible choice of $\mathbf{S} \in \mathcal{P}$, we will define the meet operation $\wedge$ on $S$ as in Figures 5.5 and 5.6. In each case, it is clear that $f$ and $w$ are maximal in $\langle S; \wedge \rangle$ and their $\mathcal{L}$-classes are up-sets, and it is easily seen that the left translation $x \mapsto fx$ preserves $\wedge$. To complete the proof, then, it suffices to show that the six possible meet operations given in Figures 5.5 and 5.6 are compatible with the respective choice of $\mathbf{S} \in \mathcal{P}$.

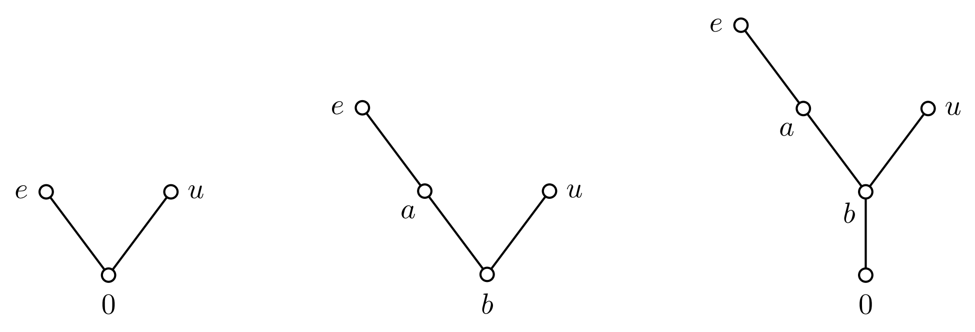



Figure 5.5. From left to right: meet operations on $\mathbf{P}$, $\mathbf{C}$, and $\mathbf{C}^0$.

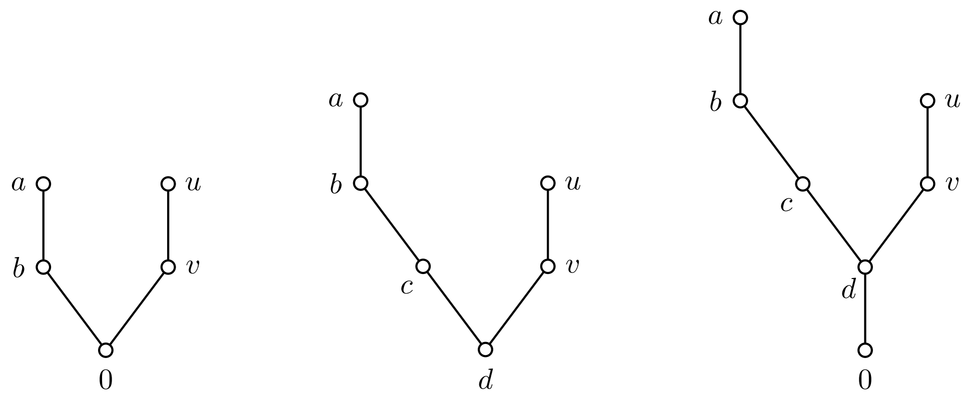



FIGURE 5.6. From left to right: meet operations on $\mathbf{L} \circlearrowleft L$, $\mathbf{E}$, and $\mathbf{E}^0$.

We will show that $\wedge \in \mathcal{T}_2(\mathbf{S})$ using Lemma 4.33. Let $x, y, z, t \in S$ to show that

$$(x \wedge y)(z \wedge t) = xz \wedge yt. \tag{$\triangle$}$$

Let $s\colon \mathbf{S} \to \mathrm{E}(\mathbf{S})$ denote the square map $x \mapsto x^2$ (which is homomorphism).

First, consider the case $\mathbf{S} = \mathbf{P}$. Note that the subsemigroup $\langle \{e, 0\}; \cdot \rangle$ of $\mathbf{P}$ is a semilattice, and by Lemma 4.34(i), the restriction of $\wedge$ to $\{e, 0\}^2$ is compatible with $\langle \{e, 0\}; \cdot \rangle$. Since $\{e, 0\}$ is the image of $s$, we have

$$(s(x) \wedge s(y))(s(z) \wedge s(t)) = s(x)s(z) \wedge s(y)s(t)$$

by Lemma 4.33. But $s$ is an endomorphism of $\langle P; \wedge \rangle$ (and of $\mathbf{P}$), so this implies that

$$s((x \wedge y)(z \wedge t)) = s(xz \wedge yt)$$

That is, the two sides of ($\triangle$) are related by $\ker(s)$. Now, the only non-trivial block of $\ker(s)$ is $\{u, 0\}$, so to obtain ($\triangle$) it suffices to show that $(x \wedge y)(z \wedge t) = u \Leftrightarrow xz \wedge yt = u$. We have

$$\begin{aligned} (x \wedge y)(z \wedge t) = u &\iff x \wedge y = e \;\;\&\;\; z \wedge t = u \\ &\iff x = y = e \;\;\&\;\; z = t = u \\ &\iff xz = yt = u \\ &\iff xz \wedge yt = u. \end{aligned}$$

Thus, ($\triangle$) holds for $\mathbf{S} = \mathbf{P}$.

Next, consider the case $\mathbf{S} = \mathbf{C}$. Let $h$ be the unique retraction of $\mathbf{C}$ onto the right-zero subsemigroup $\langle \{a, b\}; \cdot \rangle$ (so $h(e) = a$ and $h(u) = b$). Then $h$ preserves $\wedge$, and $\wedge$ is compatible with $\langle \{a, b\}; \cdot \rangle$ by Lemma 4.34(ii), so, as in the argument for $\mathbf{P}$, we see that the two sides of ($\triangle$) are related by $\ker(h)$. Now, let $q$ denote the unique surjective homomorphism $\mathbf{C} \to \mathbf{P}$ (so $q^{-1}(0) = \{a, b\}$). Then $q$ is also a semilattice homomorphism $\langle C; \wedge \rangle \to \langle P; \wedge \rangle$. As $\wedge$ is compatible with $\mathbf{P}$, it follows that the two sides of ($\triangle$) are also related by $\ker(q)$. But the kernels of $h$ and $q$ intersect trivially, so ($\triangle$) holds for $\mathbf{S} = \mathbf{C}$.

For $\mathbf{C}^0$, note that $0$ is a zero element for both multiplication and meet. Thus, the equality ($\triangle$) is clear if $0 \in \{x, y, z, t\}$, and otherwise, we get equality from the case $\mathbf{S} = \mathbf{C}$.

Next, consider the case $\mathbf{S} = \mathbf{L} \circlearrowleft L$. The square map $s$ retracts $\mathbf{L} \circlearrowleft L$ onto the subsemigroup $\langle \{a, b, 0\}; \cdot \rangle \cong \mathbf{L}^0$, and by Lemma 4.34(iii), the restriction of $\wedge$ to $\{a, b, 0\}$ is compatible with this subsemigroup. Thus, as in the $\mathbf{S} = \mathbf{P}$ case, we find that the two sides of $(\triangle)$ are related by $\ker(s)$. Now, let $h\colon \mathbf{L}\circlearrowright L \to \mathbf{L}\circlearrowright L$ denote left translation by $a$, which is a retraction onto the subsemigroup $\langle\{a,u,0\};\cdot\rangle \cong \mathbf{P}$. Then $h$ is a semilattice homomorphism onto $\langle P;\wedge\rangle$, so, using the $\mathbf{S} = \mathbf{P}$ case, we find that the two sides of $(\triangle)$ are related by $\ker(h)$. The intersection of $\ker(s)$ and $\ker(h)$ has $\{u,v\}$ as the only non-trivial block, so it now suffices to show that $(x\wedge y)(z\wedge t) = u \Leftrightarrow xz\wedge yt = u$. We have

$$\begin{aligned} (x\wedge y)(z\wedge t) = u &\iff x\wedge y = a \ \ \&\ \ z\wedge t \in \{u,v\} \\ &\iff x = y = a \ \ \&\ \ z,t \in \{u,v\} \\ &\iff xz = yt = u \\ &\iff xz\wedge yt = u. \end{aligned}$$

The case $\mathbf{S} = \mathbf{E}$ is proved analogously to the $\mathbf{S} = \mathbf{C}$ case, using the unique retraction of $\mathbf{E}$ onto $\langle\{c,d\};\cdot\rangle \cong \mathbf{R}$ and the unique surjective homomorphism $q\colon \mathbf{E} \to \mathbf{L}\circlearrowright L$ such that $q(u) = u$. Finally, the case $\mathbf{S} = \mathbf{E}^0$ is proved analogously to the $\mathbf{S} = \mathbf{C}^0$ case. $\square$

Now, we have three more operations to introduce (the word 'meet' in the section title was actually meant in the sense of making acquaintance for the first time). These operations integrate perhaps dozens of precursory operations constructed in specific cases. At this point, they appear to perform genuinely different functions in the proofs, and it is unlikely that any of them can be combined further.

**Lemma 5.19.** *Let $\mathbf{S}$ be a semigroup, and let $\mathbf{A}$ and $\mathbf{B}$ be subsemigroups of $\mathbf{S}$ with $\mathbf{B} \in \mathbb{I}(\mathcal{P})$. Let $F \subseteq A$ be a set of left identities of $\mathbf{A}$, let $U$ be an arbitrary subset of $A$, let $f$ be in the maximum $\mathcal{J}$-class of $\mathbf{B}$, and define the partial binary operation $\rtimes$ on $S$ by*

$$\operatorname{dom}(\rtimes) := \{(x,y) \in A\times B \mid y\ \mathcal{L}^{\mathbf{B}}\ f \implies x \in F\}, \qquad x\rtimes y := \begin{cases} fy & \text{if } x \in U, \\ fy^2 & \text{otherwise.} \end{cases}$$

*Then $\rtimes \in \mathcal{P}_2(\mathbf{S})$.*

Proof. By checking cases over $\mathbf{B}$, it is easily verified that $\operatorname{dom}(\rtimes)$ is a subsemigroup of $\mathbf{S}^2$, using the fact that $yy'\ \mathcal{L}\ f$ implies $y\ \mathcal{L}\ y'\ \mathcal{L}\ f$ for all $y,y' \in B$. Note that $(x\rtimes y)^2 = fy^2$ for all $(x,y) \in \operatorname{dom}(\rtimes)$; this follows from the equations $x^2y \approx xy$ and $xyzt \approx xzyt$, which hold in $\mathbf{B}$.

Using Lemma 4.33, let $(x,y),(z,t) \in \operatorname{dom}(\rtimes)$ to show that

$$(x\rtimes y)(z\rtimes t) = xz\rtimes yt. \qquad (\triangle)$$

Note that the image of $\rtimes$ is contained in $B$, so we have $((x\rtimes y)(z\rtimes t))^2 = (x\rtimes y)^2(z\rtimes t)^2$ as $\mathbf{B} \models (xy)^2 \approx x^2y^2$. Also, $fy^2ft^2 = f(yt)^2$ follows from $x^2y \approx xy$ and $xyzt \approx xzyt$, so

$$((x\rtimes y)(z\rtimes t))^2 = (x\rtimes y)^2(z\rtimes t)^2 = fy^2ft^2 = f(yt)^2 = (xz\rtimes yt)^2.$$

Thus, the two sides of $(\triangle)$ are related by the kernel of the square map $s\colon \mathbf{B} \to \mathbf{B}$. Now, let $w$ be the non-idempotent element of $\mathbf{B}$ with $fw = w$. The image of $\rtimes$ is contained in $fB$, and the kernel of $s$ restricted to $fB$ has $\{w,w^2\}$ as the only non-trivial block, so to obtain $(\triangle)$, it suffices to show that $(x \rtimes y)(z \rtimes t) = w \Leftrightarrow xz \rtimes yt = w$. To prove this, we will use the following properties of $\rtimes$, which can be verified by case-checking over $\mathbf{B}$ again.

$$(\forall (x,y) \in \operatorname{dom}(\rtimes)) \ \ (x \rtimes y = f \iff y \,\mathcal{L}^{\mathbf{B}}\, f) \ \&\ (x \rtimes y = w \iff x \in U \ \&\ y \,\mathcal{L}^{\mathbf{B}}\, w). \quad (\times)$$

Now, we have

$$\begin{aligned}
(x \rtimes y)(z \rtimes t) = w &\iff x \rtimes y = f \ \&\ z \rtimes t = w && \text{(as the image of } \rtimes \text{ is in } fB) \\
&\iff y \,\mathcal{L}^{\mathbf{B}}\, f \ \&\ z \in U \ \&\ t \,\mathcal{L}^{\mathbf{B}}\, w && \text{(by } (\times)) \\
&\iff y \,\mathcal{L}^{\mathbf{B}}\, f \ \&\ xz \in U \ \&\ t \,\mathcal{L}^{\mathbf{B}}\, w && \text{(as } y \,\mathcal{L}^{\mathbf{B}}\, f \Rightarrow x \in F) \\
&\iff xz \in U \ \&\ yt \,\mathcal{L}^{\mathbf{B}}\, w \\
&\iff xz \rtimes yt = w && \text{(by } (\times)). \qquad \square
\end{aligned}$$

The prototype for the following operation was discovered by the first supervisor in the original dualisability proof for $\mathbf{L} \times \mathbf{P}$.

**Lemma 5.20.** *Let $\mathbf{S}$ be a semigroup, and let $\mathbf{A}$ and $\mathbf{B}$ be subsemigroups of $\mathbf{S}$ with $\mathbf{A} \in \mathbb{I}(\mathcal{C})$ and $\mathbf{B} \in \mathbb{I}(\mathcal{P})$. Assume that the maximum $\mathcal{J}$-class of $\mathbf{A}$ is a (possibly trivial) left-zero semigroup. Let $a$ be an element of the maximum $\mathcal{J}$-class of $\mathbf{A}$, let $u \in B\backslash \mathrm{E}(\mathbf{B})$, let $e$ be the left identity for $u$ in $\mathbf{B}$, and define the partial binary operation $\ltimes$ on $S$ by*

$$\operatorname{dom}(\ltimes) := (A \times B)\backslash\{(a,e)\}, \qquad x \ltimes y := \begin{cases} u & \text{if } (x,y) = (a,u), \\ u^2 & \text{otherwise.} \end{cases}$$

*Then $\ltimes \in \mathcal{P}_2(\mathbf{S})$.*

Proof. By assumption, we have $\mathbf{A} \in \mathbb{I}(\mathcal{C}\backslash\{\mathbf{R}, \mathbf{R}^0, \mathbf{N}\})$. Using also $\mathbf{B} \in \mathbb{I}(\mathcal{P})$, we have

$$(\forall x, x' \in A) \ \ xx' = a \implies x = a, \qquad (\forall y, y' \in B) \ \ yy' \in \{e, u\} \implies y = e.$$

This shows that $\operatorname{dom}(\ltimes)$ is a subsemigroup of $\mathbf{A} \times \mathbf{B} \leqslant \mathbf{S}^2$. It also shows that $(a,u)$ cannot arise as a product in $\operatorname{dom}(\ltimes)$, for if $(x,y)(x',y') = (a,u)$ for some $(x,y), (x',y') \in A \times B$, then we must have $(x,y) = (a,e)$. This implies that $J := \operatorname{dom}(\ltimes)\backslash\{(a,u)\}$ is an ideal of $\operatorname{dom}(\ltimes)$ and that $\operatorname{dom}(\ltimes)/J \cong \mathbf{N}$. Since $\{u, u^2\}$ is a two-element null subsemigroup of $\mathbf{B}$, it follows that $\ltimes$ is a homomorphism. $\square$

The last operation we will need is the following.

**Lemma 5.21.** *Let $\mathbf{S}$ be a semigroup, and let $\mathbf{A}$ and $\mathbf{B}$ be subsemigroups of $\mathbf{S}$ with*

$$\mathbf{A} \in \mathbb{I}(\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0), \qquad \mathbf{B} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0).$$

*Let $u \in A\backslash \mathrm{E}(\mathbf{A})$, let $a$ be the left identity for $u$ in $\mathbf{A}$, and let $r\colon \mathbf{A} \to \mathbf{A}$ denote the retraction of $\mathbf{A}$ onto the ideal $A\backslash(L_a^{\mathbf{A}} \cup L_u^{\mathbf{A}})$ of $\mathbf{A}$. Let $w$ be the non-idempotent element of $\mathbf{B}$, let $f$ be the left identity of $\mathbf{B}$, and define the partial binary operation $\lhd$ on $S$ by*

$$\operatorname{dom}(\lhd) := \{(x,y) \in A \times B \mid (y = f \implies x \,\mathcal{L}^{\mathbf{A}}\, a) \ \&\ (y = w \implies x \,\mathcal{L}^{\mathbf{A}}\, u)\},$$

$$x \lhd y := \begin{cases} x & \text{if } y \in \{f, w\}, \\ r(x) & \text{otherwise.} \end{cases}$$

*Then $\lhd \in \mathcal{P}_2(\mathbf{S})$.*

Proof. To see that $\mathrm{dom}(\lhd)$ is a subsemigroup of $\mathbf{S}^2$, let $(x, y), (x', y') \in \mathrm{dom}(\lhd)$. Then

$$yy' = f \implies y = y' = f \implies x, x \in L_a^{\mathbf{A}} \implies xx' \in L_a^{\mathbf{A}},$$
$$yy' = w \implies y = f \ \& \ y' = w \implies x \in L_a^{\mathbf{A}} \ \& \ x' \in L_u^{\mathbf{A}} \implies xx' \in L_u^{\mathbf{A}},$$

so $(xx', yy') \in \mathrm{dom}(\lhd)$, as required.

Using Lemma 4.33, let $(x, y), (z, t) \in \mathrm{dom}(\lhd)$ to show that

$$(x \lhd y)(z \lhd t) = xz \lhd yt. \tag{$\triangle$}$$

First, fix some $c \in L_a^{\mathbf{A}} \cup L_u^{\mathbf{A}}$. We will show that $(x \lhd y)(z \lhd t) = c \Leftrightarrow xz \lhd yt = c$. Note that since $(x, y) \in \mathrm{dom}(\lhd)$, we have the implication $xc = c \Rightarrow y \neq w$, because $xc = c \Rightarrow x \in L_a^{\mathbf{A}}$. This gives the equivalence

$$xc = c \ \& \ y = f \iff xc = c \ \& \ y \in \{f, w\}. \tag{$\heartsuit$}$$

Now, let $k$ be the unique left identity for $c$ in $\mathbf{A}$, so $k \in L_a^{\mathbf{A}}$. Then

$$(\forall x, y \in A) \ \ xy = c \iff x = k \ \& \ y \,\mathcal{L}^{\mathbf{A}}\, c. \tag{$\diamondsuit$}$$

We have

$$\begin{aligned}
(x \lhd y)(z \lhd t) = c &\iff x \lhd y = k \ \& \ z \lhd t \,\mathcal{L}^{\mathbf{A}}\, c && \text{(by } (\diamondsuit))\\
&\iff x = k \ \& \ y \in \{f, w\} \ \& \ z \,\mathcal{L}^{\mathbf{A}}\, c \ \& \ t \in \{f, w\} \\
&\iff x = k \ \& \ z \,\mathcal{L}^{\mathbf{A}}\, c \ \& \ y = f \ \& \ t \in \{f, w\} && \text{(by } (\heartsuit))\\
&\iff xz = c \ \& \ yt \in \{f, w\} && \text{(by } (\diamondsuit))\\
&\iff xz \lhd yt = c.
\end{aligned}$$

This shows that either the two sides of $(\triangle)$ both lie in $L_a^{\mathbf{A}} \cup L_u^{\mathbf{A}}$, in which case they are equal, or else they both lie in the ideal $A \backslash (L_a^{\mathbf{A}} \cup L_u^{\mathbf{A}})$ of $\mathbf{A}$. Assume that we are in the latter situation. Then $r$ fixes both $(x \lhd y)(z \lhd t)$ and $xz \lhd yt$. Using the fact that $r(x \lhd y) = r(x)$ for all $(x, y) \in \mathrm{dom}(\lhd)$, we have

$$(x \lhd y)(z \lhd t) = r((x \lhd y)(z \lhd t)) = r(x \lhd y)r(z \lhd t) = r(x)r(z) = r(xz) = r(xz \lhd yt) = xz \lhd yt,$$

as required. $\square$

## 5.6. Three lemmas

Having introduced all of the necessary operations, we are now ready to prove our dualisability results. An overview of the main proof is as follows. Given a finite product $\mathbf{S}$ of members of $\mathfrak{C}$, we will consider the alter ago $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathfrak{T} \rangle$ and a morphism $\varphi \colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ for some $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ and some $n \in \omega \backslash \{0\}$; the goal is to extend $\varphi$ to an $n$-ary term function of $\mathbf{S}$, thus establishing (IC). The quasicritical direct factors of $\mathbf{S}$ can be viewed as subsemigroups, so we will aim to use Lemma 4.31, which requires us to construct term functions that agree with $\varphi$ on certain subsets of $S^n$.

In this section, we will establish several results concerning the existence of such term functions. The proofs of these are certainly the most intricate in this thesis. Almost all of the discussion in Section 4.5 up to Lemma 4.35 will be relevant here, and many of the ideas in the remainder of Section 4.5 will be expanded upon.

We first introduce the kinds of terms that we will construct in this chapter. Let $n \in \omega$; as in Section 4.5, it will be important here to identify $n$ with the set $\{0, \dots, n-1\}$. Let $v_i$ denote the $i$th variable for each $i \in n$, and let $J, K \subseteq n$ with $J \cap K = \varnothing$ and $K \neq \varnothing$. A **$(J, K)$-term** is an $n$-ary semigroup term of the form

$$\Big(\prod_{j \in J} v_j\Big) v_k,$$

where $k \in K$ and the product over $J$ is taken in any order. A **$(J, K)$-term function** of a semigroup $\mathbf{S}$ is a term function of $\mathbf{S}$ induced by a $(J, K)$-term. The term functions we construct in the remainder of this chapter will always be $(J, K)$-term functions for suitable $J$ and $K$, so this notion will be of fundamental importance.

For a semigroup satisfying $xyz \approx yxz$ (which holds in $\mathbb{V}(\mathbf{R}, \mathbf{P})$), the function induced by a $(J, K)$-term does not depend on the ordering of the $J$-indexed variables. For a semigroup satisfying the weaker identity $xyzt \approx xzyt$, the function induced by a $(J, K)$-term, for a fixed $k \in K$, may depend on which variable occurs first in the product over $J$, but the variables between the first and last can be permuted without changing the induced function. We allow $J = \varnothing$ in the definition; in this case, a $(J, K)$-term is a projection term.

We will now prove three technical lemmas to be used in the main proof. These results were abstracted from close to 30 separate dualisability proofs. As a result, they are quite lengthy and involve many cases, but they still make our overall proof much shorter than the collection of individual proofs. The reader will be rewarded for their perseverance in Chapter 6, as the author was.

**Lemma 5.22.** *Let $\mathbf{S}$ be a finite semigroup, and let $\mathbf{A}$ and $\mathbf{B}$ be subsemigroups of $\mathbf{S}$ such that $\mathbf{B} \in \mathbb{I}(\mathcal{P})$. Let $w \in B \backslash \mathrm{E}(\mathbf{B})$, let $f \in B$ be the left identity for $w$, and let $\wedge \in \mathcal{T}_2(\mathbf{B})$ be a meet operation such that $L^{\mathbf{B}}_f$ and $L^{\mathbf{B}}_w$ are up-sets of $\langle B; \wedge \rangle$. Let $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T} \rangle$, let $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ for some $n \in \omega \backslash \{0\}$, and let $\varphi \colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ be a morphism. Assume that $w \in \varphi(B^{\mathbf{X}})$, and let $\widehat{w}$ denote the least element of $\varphi^{-1}(w) \cap B^{\mathbf{X}}$ with respect to $\wedge$. Define*

$$J := \{ i \in n \mid \widehat{w}(i) \mathrel{\mathcal{L}^{\mathbf{B}}} f \}, \qquad K := \{ i \in n \mid \widehat{w}(i) \mathrel{\mathcal{L}^{\mathbf{B}}} w \}.$$

*Then the following hold:*

(i) *if $F$ is the set of all left identities of $\mathbf{A}$ and $A \backslash F$ is either an ideal of $\mathbf{A}$ or empty, then for every $(J, K)$-term function $t$ of $\mathbf{S}$ and every $x \in A^{\mathbf{X}}$, we have*

$$\varphi(x) \in F \iff t(x) \in F \implies \varphi(x) = t(x);$$

(ii) *if $\mathbf{A} \in \mathbb{I}(\mathbf{R}, \mathbf{R}^0, \mathbf{I})$, then every $(J, K)$-term function of $\mathbf{S}$ agrees with $\varphi$ on $A^{\mathbf{X}}$.*

Proof. Since $A, B \in \mathcal{R}_1(\mathbf{S})$, we have $\underset{\sim}{\mathbf{S}} \Vdash A, B$, so $\varphi(A^{\mathbf{X}}) \subseteq A$ and $\varphi(B^{\mathbf{X}}) \subseteq B$. Thus, $\varphi$ restricts to a morphism $\langle B^{\mathbf{X}}; \wedge \rangle \to \langle B; \wedge \rangle$, so the set $\varphi^{-1}(w) \cap B^{\mathbf{X}}$ is closed under $\wedge$, which shows that $\widehat{w}$ exists.

We first prove (i). If $F$ is empty then (i) vacuously holds, so assume that $F \neq \varnothing$, and let $F' := A \backslash F$. If $F' \neq \varnothing$, then the partition $\{F, F'\}$ of $A$ corresponds to a congruence on $\mathbf{A}$, and the resulting quotient is isomorphic to $\mathbf{I}$. Now, note that $\{f, wf\}$ is a two-element subsemilattice of $\mathbf{B}$, so, whether or not $F' = \varnothing$, we may define a homomorphism $r \colon \mathbf{A} \to \mathbf{B}$ by $r(F) = \{f\}$ and $r(F') \subseteq \{wf\}$. Since $r \in \mathcal{P}_1(\mathbf{S})$, we have $\underset{\sim}{\mathbf{S}} \Vdash r$.

We claim that

*For every $(J,K)$-term function $t$ of* $\mathbf{S}$, *we have* $(\forall x \in A^{\mathbf{X}})\ \varphi(x) \in F \implies t(x) \in F$. $(\triangleright)$

To prove $(\triangleright)$, let $t$ be a $(J,K)$-term function of $\mathbf{S}$, and let $x \in A^{\mathbf{X}}$ with $\varphi(x) \in F$. Then $\varphi$ preserves the multiplication of $\mathbf{B}$ by Lemma 4.33, so $\varphi(r(x)\widehat{w}) = r(\varphi(x))\varphi(\widehat{w}) = fw = w$, and by the definition of $\widehat{w}$ we then have $r(x)\widehat{w} \geqslant \widehat{w}$ in $\langle B^{\mathbf{X}}; \wedge\rangle$. Now, since $L_f^{\mathbf{B}}$ and $L_w^{\mathbf{B}}$ are up-sets of $\langle B; \wedge\rangle$, we have for all $i \in J \cup K$ that $(r(x)\widehat{w})(i) \in L_f^{\mathbf{B}} \cup L_w^{\mathbf{B}}$, so $r(x)(i) \in L_f^{\mathbf{B}}$; thus, $x(i) \in F$ for all $i \in J \cup K$. As $F$ forms a subsemigroup of $\mathbf{A}$, it follows that $t(x) \in F$, which proves $(\triangleright)$.

For each $c \in A$, define the following partial binary operation on $S$, which is compatible with $\mathbf{S}$ by Lemma 5.19:

$$\operatorname{dom}(\rtimes_c) := \{(x,y) \in A \times B \mid y\ \mathcal{L}^{\mathbf{B}}\ f \implies x \in F\}, \qquad x \rtimes_c y := \begin{cases} fy & \text{if } x = c, \\ fy^2 & \text{otherwise.} \end{cases}$$

Here, the crucial property of $\rtimes_c$ is that whenever $(x,y) \in \operatorname{dom}(\rtimes_c)$ with $x \rtimes_c y\ \mathcal{L}^{\mathbf{B}}\ w$, we must have $x = c$, since otherwise $x \rtimes_c y$ is idempotent. This will be used to prove the claim:

$$(\forall c \in A)(\forall x \in \varphi^{-1}(c) \cap A^{\mathbf{X}})\ \ x \rtimes_c \widehat{w} \textit{ is defined} \implies x(i) = c \textit{ for all } i \in K. \qquad (\triangleleft)$$

To prove $(\triangleleft)$, let $c \in A$ and $x \in \varphi^{-1}(c) \cap A^{\mathbf{X}}$, and assume $x \rtimes_c \widehat{w}$ is defined, in which case it must lie in $B^{\mathbf{X}}$. Then $\varphi(x \rtimes_c \widehat{w}) = c \rtimes_c w = fw = w$, so $x \rtimes_c \widehat{w} \geqslant \widehat{w}$ by definition of $\widehat{w}$. As $L_w^{\mathbf{B}}$ is an up-set, we have for all $i \in K$ that $(x \rtimes_c \widehat{w})(i)\ \mathcal{L}^{\mathbf{B}}\ w$ , so $x(i) = c$, proving $(\triangleleft)$.

Let $t$ be a $(J,K)$-term function of $\mathbf{S}$. To prove (i), it will suffice to prove that the following implications hold for all $x \in A^{\mathbf{X}}$ (recall that $F' = A \backslash F$):

$$\varphi(x) \in F \implies t(x) = \varphi(x), \qquad \varphi(x) \in F' \implies t(x) \in F'.$$

Let $x \in A^{\mathbf{X}}$, and define $c := \varphi(x)$. First, assume that $c \in F$; we must show that $t(x) = c$. By $(\triangleright)$ we have $t(x) \in F$, so $x(i) \in F$ for every $i \in J$ because $t$ is a $(J,K)$-term function and $F'$ is an ideal. It follows that $x \rtimes_c \widehat{w}$ is defined, so by $(\triangleleft)$ we have $x(i) = c$ for all $i \in K$. Since $t$ is a $(J,K)$-term function and $F$ is a right-zero semigroup, we have $t(x) = c = \varphi(x)$.

Now assume that $c \in F'$; we must prove that $t(x) \in F'$. If $x(i) \in F'$ for some $i \in J$, then clearly $t(x) \in F'$, so assume that $x(i) \in F$ for all $i \in J$. Then $x \rtimes_c \widehat{w}$ is defined, so by $(\triangleleft)$ we have $x(i) = c \in F'$ for all $i \in K$. Since $F'$ is an ideal, it follows that $t(x) \in F'$, and this completes the proof of (i).

To prove (ii), we apply (i) to $\mathbf{A}$. Note that $F$ is the set of non-zero elements of $\mathbf{A}$, so (i) implies that $\varphi$ and $t$ must agree on $A^{\mathbf{X}}$, for every $(J,K)$-term function $t$. □

**Lemma 5.23.** *Let* $\mathbf{S}$ *be a finite semigroup, and let* $\mathbf{A}$ *and* $\mathbf{B}$ *be subsemigroups of* $\mathbf{S}$ *such that* $\mathbf{B} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$. *Let* $f$ *be the left identity of* $\mathbf{B}$, *let* $w$ *be the non-idempotent element of* $\mathbf{B}$, *and let* $\wedge \in \mathcal{T}_2(\mathbf{B})$ *be a meet operation such that* $f$ *and* $w$ *are maximal in* $\langle B; \wedge\rangle$. *Let* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T}\rangle$, *let* $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ *for some* $n \in \omega\backslash\{0\}$, *and let* $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ *be a morphism. Assume that* $w \in \varphi(B^{\mathbf{X}})$, *and let* $\widehat{w}$ *denote the least element of* $\varphi^{-1}(w) \cap B^{\mathbf{X}}$ *with respect to* $\wedge$. *Define*

$$J := \{i \in n \mid \widehat{w}(i) = f\}, \qquad K := \{i \in n \mid \widehat{w}(i) = w\}.$$

*Then* $K \neq \varnothing$, *and the following hold:*

(i) *if* $\mathbf{A} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$ *and there exists a homomorphism* $h\colon \mathbf{A} \to \mathbf{B}$ *with* $w \in h(A)$, *then every* $(J, K)$*-term function of* $\mathbf{S}$ *agrees with* $\varphi$ *on* $A^{\mathbf{X}}$*;*

(ii) *if* $\mathbf{A} \in \mathbb{I}(\mathbf{L}, \mathbf{L}^0)$, *then there exists a* $(J, K)$*-term function of* $\mathbf{S}$ *that agrees with* $\varphi$ *on* $A^{\mathbf{X}}$*;*

(iii) *if* $\mathbf{A} \in \mathbb{I}(\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0)$ *and there exists a homomorphism* $h\colon \mathbf{A} \to \mathbf{B}$ *with* $w \in h(A)$, *then there exists a* $(J, K)$*-term function of* $\mathbf{S}$ *that agrees with* $\varphi$ *on* $A^{\mathbf{X}}$.

Proof. The existence of $\widehat{w}$ is proved as in Lemma 5.23. To see that $K \neq \varnothing$, suppose to the contrary that $K = \varnothing$. Then $\widehat{w}^2 = \widehat{w}$, so because $\varphi$ preserves the multiplication of $\mathbf{B}$ by Lemma 4.33, we have $\varphi(\widehat{w}) = \varphi(\widehat{w}^2) = \varphi(\widehat{w})^2 = w^2 \neq w$, contradicting $\widehat{w} \in \varphi^{-1}(w)$.

To prove (i), assume that $\mathbf{A} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$. Let $e$ denote the left identity of $\mathbf{A}$ and $u$ the non-idempotent element of $\mathbf{A}$. Also, let $a := ue$ and $b := u^2$, and denote the zero of $\mathbf{A}$ by 0 if it exists. Then $a = b$ if and only if $\mathbf{A} \cong \mathbf{P}$, in which case $0 = a = b$.

Note that a homomorphism $h\colon \mathbf{A} \to \mathbf{B}$ with $w \in h(A)$ exists if and only if the following implications hold:

$$(\mathbf{A} \cong \mathbf{P} \implies \mathbf{B} \cong \mathbf{P}) \ \& \ (\mathbf{A} \cong \mathbf{C}^0 \implies \mathbf{B} \ncong \mathbf{C}).$$

Fix a homomorphism $h\colon \mathbf{A} \to \mathbf{B}$ with $w \in h(A)$. One easily verifies that there is unique choice for $h$, and it satisfies $h^{-1}(f) = \{e\}$ and $h^{-1}(w) = \{u\}$.

Let $t$ be a $(J, K)$-term function of $\mathbf{S}$. We must show that $t$ agrees with $\varphi$ on $A^{\mathbf{X}}$. Since $\varphi(A^{\mathbf{X}}) \subseteq A$, this is equivalent to showing that $\varphi(x) = c$ implies $t(x) = c$ for all $c \in A$ and all $x \in A^{\mathbf{X}}$.

Let $x \in A^{\mathbf{X}}$ with $\varphi(x) = u$. We will show that $t(x) = u$. We have $h(x) \in B^{\mathbf{X}}$, and since $\underset{\sim}{\mathbf{S}} \Vdash h$, we have $\varphi(h(x)) = h(\varphi(x)) = h(u) = w$, so $h(x) \geqslant \widehat{w}$ with respect to the $\wedge$-order. Now, $f$ and $w$ are maximal in $\langle B; \wedge\rangle$, so $h(x)(i) = f$ for all $i \in J$ and $h(x)(i) = w$ for all $i \in K$. Since $h^{-1}(f) = \{e\}$ and $h^{-1}(w) = \{u\}$, this implies that $x(i) = e$ for all $i \in J$ and $x(i) = u$ for all $i \in K$. Thus, $t(x) = u$, by the definition of a $(J, K)$-term function. Combining this with Lemma 5.22(i) (where $F = \{e\}$), we have

$$(\forall x \in A^{\mathbf{X}}) \ \ \varphi(x) \in \{e, u\} \implies t(x) = \varphi(x).$$

Now, let $\langle a\rangle$ denote the ideal of $\mathbf{A}$ generated by $a$. Then $\langle a\rangle = A\backslash\{e, u\}$, so to prove (i), it remains to show that $\varphi(x) \in \langle a\rangle \Rightarrow t(x) = \varphi(x)$ for all $x \in A^{\mathbf{X}}$. First, we will show that

$$(\forall x \in A^{\mathbf{X}}) \ \ \varphi(x) \in \langle a\rangle \implies t(x) \in \langle a\rangle. \tag{$\heartsuit$}$$

By Lemma 5.19, the following operation is in $\mathcal{P}_2(\mathbf{S})$:

$$\operatorname{dom}(\rtimes) := \{(x, y) \in A \times B \mid y = f \implies x = e\}, \qquad x \rtimes y := \begin{cases} y & \text{if } x \in \langle a\rangle, \\ y^2 & \text{otherwise.} \end{cases}$$

To prove ($\heartsuit$), let $x \in A^{\mathbf{X}}$ with $\varphi(x) \in \langle a\rangle$. We will show that $t(x) \in \langle a\rangle$. This is immediate if $x(i) \neq e$ for some $i \in J$ (by choice of $t$ and $\mathbf{A}$), so we can assume that $x(i) = e$ for all $i \in J$. Then $x \rtimes \widehat{w}$ is defined, $x \rtimes \widehat{w} \in B^{\mathbf{X}}$, and since $\varphi(x) \in \langle a\rangle$, we have $\varphi(x \rtimes \widehat{w}) = \varphi(x) \rtimes w = w$. Now the definition of $\widehat{w}$ gives $x \rtimes \widehat{w} \geqslant \widehat{w}$. By maximality of $w$, we have $(x \rtimes \widehat{w})(k) = w$ for all $k \in K$, so $x(k) \in \langle a\rangle$ for all $k \in K$, and hence $t(x) \in \langle a\rangle$. We have shown that ($\heartsuit$) holds.

In the case $a = b$, we are done, so assume that $a \neq b$. We must show that $\varphi(x) \in \langle a \rangle$ implies $t(x) = \varphi(x)$ for all $x \in A^{\mathbf{X}}$. Let $\mathbf{A}_0$ be the subsemigroup of $\mathbf{A}$ on $A_0 := \langle a \rangle$. Then $\mathbf{A}_0$ is isomorphic to $\mathbf{R}$ or $\mathbf{R}^0$, so by Lemma 5.22(ii), we have that $t$ agrees with $\varphi$ on $A_0^{\mathbf{X}}$. Now, let $x \in A^{\mathbf{X}}$ with $\varphi(x) \in \langle a \rangle$. By ($\heartsuit$), we have $t(x) \in \langle a \rangle$, so $a$ is a left identity for $\varphi(x)$ and $t(x)$. Since $\underline{a}x \in A_0^{\mathbf{X}}$, we have $\varphi(\underline{a}x) = t(\underline{a}x)$, so as $\varphi$ preserves the multiplication of $\mathbf{A}$, we have

$$\varphi(x) = a\varphi(x) = \varphi(\underline{a}x) = t(\underline{a}x) = at(x) = t(x).$$

This completes the proof of (i).

To prove (ii), assume that $\mathbf{A} \in \mathbb{I}(\mathbf{L}, \mathbf{L}^0)$. Let $\{a, b\}$ denote the non-trivial $\mathcal{L}$-class of $\mathbf{A}$, and let 0 denote the zero of $\mathbf{A}$, if it exists. Thus, $A$ is either $\{a, b\}$ or $\{a, b, 0\}$.

First, we will prove that for every $(J, K)$-term function $t$ of $\mathbf{S}$, we have

$$(\forall x \in A^{\mathbf{X}})\ \ \varphi(x) \in \{a, b\} \iff t(x) \in \{a, b\}. \tag{$\diamondsuit$}$$

This is trivial when $A = \{a, b\}$, so assume for the proof of ($\diamondsuit$) that $A = \{a, b, 0\}$. Let $t$ be a $(J, K)$-term function of $\mathbf{S}$, and let $\mathbf{A}_0$ be the subsemigroup of $\mathbf{A}$ on $A_0 := \{a, 0\}$, which is a two-element semilattice. By Lemma 5.22(ii) applied to $\mathbf{A}_0$, we have $\varphi(x) = a \Leftrightarrow t(x) = a$ for all $x \in A_0^{\mathbf{X}}$. Now, if $x \in A^{\mathbf{X}}$, then $\underline{a}x \in A_0^{\mathbf{X}}$, so $\varphi(\underline{a}x) = a \Leftrightarrow t(\underline{a}x) = a$. Hence, for all $x \in A^{\mathbf{X}}$, we have

$$\varphi(x) \in \{a, b\} \Leftrightarrow a\varphi(x) = a \Leftrightarrow \varphi(\underline{a}x) = a \Leftrightarrow t(\underline{a}x) = a \Leftrightarrow at(x) = a \Leftrightarrow t(x) \in \{a, b\}.$$

Thus, ($\diamondsuit$) holds.

Let $\sqcap$ be the meet operation of the three-element chain $a > b > 0$ restricted to $A^2$. Then $\sqcap \in \mathcal{T}_2(\mathbf{A})$ by Lemma 4.34, so $\underset{\sim}{\mathbf{S}} \Vdash \sqcap$. Now, since $\underline{a} \in \varphi^{-1}(a) \cap A^{\mathbf{X}}$, the set $\varphi^{-1}(a) \cap A^{\mathbf{X}}$ is non-empty, so it has a least element, $\widehat{a}$, with respect to $\sqcap$. We will prove that

$$\textit{There exists a } (J, K)\textit{-term function } t \textit{ of } \mathbf{S} \textit{ with } t(\widehat{a}) = a. \tag{$\clubsuit$}$$

We consider two cases. First, consider the case where $\widehat{a}(i) \neq a$ for all $i \in J$. By Lemma 5.20, the following operation is compatible with $\mathbf{S}$:

$$\operatorname{dom}(\ltimes) = (A \times B) \backslash \{(a, f)\}, \qquad x \ltimes y := \begin{cases} w & \text{if } (x, y) = (a, w), \\ w^2 & \text{otherwise.} \end{cases}$$

Now, $\widehat{a} \ltimes \widehat{w} \in B^{\mathbf{X}}$ is defined and $\varphi(\widehat{a} \ltimes \widehat{w}) = a \ltimes w = w$, so $\widehat{a} \ltimes \widehat{w} \geqslant \widehat{w}$ with respect to $\wedge$. Since $f \nleqslant w, w^2$ with respect to $\wedge$ and $\widehat{a} \ltimes \widehat{w} \in \{w, w^2\}^n$, the inequality $\widehat{a} \ltimes \widehat{w} \geqslant \widehat{w}$ implies that $J = \varnothing$. Further, for all $i \in K$, we have $(\widehat{a} \ltimes \widehat{w})(i) = w$ by maximality of $w$, so $\widehat{a}(i) = a$. Now, let $t$ be any $(J, K)$-term function of $\mathbf{S}$ (there is at least one, as $K \neq \varnothing$). Since $J$ is empty, $t$ is a projection onto some coordinate in $K$, so $t(\widehat{a}) = a$.

To complete the proof of ($\clubsuit$), it remains to consider the case where there exists $j \in J$ with $\widehat{a}(j) = a$. In this case, choose any $k \in K$, and define $t$ to be the $(J, K)$-term function of $\mathbf{S}$ induced by the $n$-ary term

$$v_j \Bigg( \prod_{i \in J} v_i \Bigg) v_k.$$

Then, since $t(\widehat{a}) \in \{a, b\}$ by ($\diamondsuit$), we have $t(\widehat{a}) = a$ by the choice of $j$. This completes the proof of ($\clubsuit$).

Given ($\diamondsuit$), to complete the proof of (ii) it suffices to show that

*There is a* $(J,K)$*-term function* $t$ *of* $\mathbf{S}$ *with* $(\forall x \in A^{\mathbf{X}})\ \ \varphi(x) \in \{a,b\} \implies t(x) = \varphi(x)$.

Choose, via ($\clubsuit$), a $(J,K)$-term function $t$ of $\mathbf{S}$ with $t(\widehat{a}) = a$. Let $j \in n$ be the index of the left-most variable appearing in some term inducing $t$. Then $\widehat{a}(j) = a$ as $t(\widehat{a}) = a$.

Let $x \in A^{\mathbf{X}}$ with $\varphi(x) = a$. Then $x \geqslant \widehat{a}$ with respect to $\sqcap$. As $a$ is maximal in $\langle A; \sqcap\rangle$, we have $x(j) = a$, so $t(x) = a$ by ($\diamondsuit$) and the choice of $j$. Now let $x \in A^{\mathbf{X}}$ with $\varphi(x) = b$, and let $g$ be the non-trivial automorphism of $\mathbf{A}$. Then $\varphi(g(x)) = a$, so $t(g(x)) = a$ by the previous argument, and hence $t(x) = b$. This completes the proof of (ii).

To prove (iii), assume that $\mathbf{A} \in \mathbb{I}(\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0)$, and let $h\colon \mathbf{A} \to \mathbf{B}$ with $w \in h(A)$. Let $u$ and $v$ denote the non-idempotent elements of $\mathbf{A}$, and let $a \in A$ and $b \in A$ be the left identities for $u$ and $v$, respectively. Thus, $\{a,b\}$ and $\{u,v\}$ are the non-trivial $\mathcal{L}$-classes of $\mathbf{A}$. Note that $h$ must satisfy $h^{-1}(f) = \{a,b\}$ and $h^{-1}(w) = \{u,v\}$. We claim that

*For every* $(J,K)$*-term function* $t$ *of* $\mathbf{S}$*, we have* $(\forall x \in A^{\mathbf{X}})\ \ \varphi(x)\ \mathcal{L}^{\mathbf{A}}\ t(x)$. ($\spadesuit$)

For the proof of ($\spadesuit$), let $\mathbf{A}_0$ be the subsemigroup of $\mathbf{A}$ on $A_0 := aA$. Then $\mathbf{A}_0 \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$. As $\mathcal{L}^{\mathbf{B}}$ is trivial, the homomorphism $h$ factors through $\mathbf{A}/\mathcal{L}^{\mathbf{A}} \cong \mathbf{A}_0$; hence, there is a homomorphism $h'\colon \mathbf{A}_0 \to \mathbf{B}$ with $w \in h'(A_0)$. Thus, we can apply (i) to $\mathbf{A}_0$ and conclude that every $(J,K)$-term function of $\mathbf{S}$ agrees with $\varphi$ on $A_0^{\mathbf{X}}$. Now, if $t$ is a $(J,K)$-term function of $\mathbf{S}$ and $x \in A^{\mathbf{X}}$, then $\underline{a}x \in A_0^{\mathbf{X}}$, so $\varphi(\underline{a}x) = t(\underline{a}x)$, and hence $a\varphi(x) = at(x)$. Since left translation by $a$ on $\mathbf{A}$ has kernel $\mathcal{L}^{\mathbf{A}}$, this shows that ($\spadesuit$) holds.

Now, $\varphi(A^{\mathbf{X}})$ contains $\mathrm{E}(\mathbf{A})$ and is closed under the automorphisms of $\mathbf{A}$, so $\varphi(A^{\mathbf{X}})$ equals either $A$ or $\mathrm{E}(\mathbf{A})$. Thus, we have $u \in \varphi(A^{\mathbf{X}})$ if and only if $\varphi(A^{\mathbf{X}}) = A$. We will consider separately the cases $u \notin \varphi(A^{\mathbf{X}})$ and $u \in \varphi(A^{\mathbf{X}})$.

First, consider the case where $u \notin \varphi(A^{\mathbf{X}})$, so $v \notin \varphi(A^{\mathbf{X}})$ also. Let $\mathbf{A}_0$ be the subsemigroup of $\mathbf{A}$ on $A_0 := \{a, b, ua\}$, which is isomorphic to $\mathbf{L}^0$. Let $r\colon \mathbf{A} \to \mathbf{A}_0$ be the homomorphism fixing $a$ and $b$ and sending the remaining elements of $A$ to $ua$. By (ii) applied to $\mathbf{A}_0$, there is a $(J,K)$-term function $t$ of $\mathbf{S}$ that agrees with $\varphi$ on $A_0^{\mathbf{X}}$. Now, let $x \in A^{\mathbf{X}}$ with $\varphi(x) \in \{a,b\}$. Then $r(x) \in A_0^{\mathbf{X}}$, and so $\varphi(r(x)) = t(r(x))$. By ($\spadesuit$), we have $t(x) \in \{a,b\}$, and so $r$ fixes $\varphi(x)$ and $t(x)$. Thus,

$$\varphi(x) = r(\varphi(x)) = \varphi(r(x)) = t(r(x)) = r(t(x)) = t(x).$$

Now ($\spadesuit$) implies that $t$ agrees with $\varphi$ on $A^{\mathbf{X}}$.

It remains to consider the case where $u, v \in \varphi(A^{\mathbf{X}})$. Let $\sqcap \in \mathcal{T}_2(\mathbf{A})$ be a meet operation such that $a$ and $u$ are maximal in $\langle A; \sqcap\rangle$ and $\{a,b\}$, $\{u,v\}$ are up-sets of $\langle A; \sqcap\rangle$; such a meet operation exists by Theorem 5.18. Let $\widehat{u}$ denote the least element of $\varphi^{-1}(u) \cap A^{\mathbf{X}}$ with respect to $\sqcap$. Define

$$J_u := \{i \in n \mid \widehat{u}(i) \in \{a,b\}\}, \qquad K_u := \{i \in n \mid \widehat{u}(i) \in \{u,v\}\}.$$

We claim that $J_u = J$ and $K_u = K$. First, we will show that $J \subseteq J_u$ and $K \subseteq K_u$. Using the homomorphism $h\colon \mathbf{A} \to \mathbf{B}$ with $w \in h(A)$, we have $h(u) = w$, so $\varphi(h(\widehat{u})) = h(u) = w$, and therefore $h(\widehat{u}) \geqslant \widehat{w}$ with respect to $\wedge$. Now, since $f$ and $w$ are maximal in $\langle B; \wedge\rangle$, we have $h(\widehat{u})(i) = f$ for all $i \in J$ and $h(\widehat{u})(i) = w$ for all $i \in K$. Since $h^{-1}(f) = \{a,b\}$ and $h^{-1}(w) = \{u,v\}$, this gives $J \subseteq J_u$ and $K \subseteq K_u$.

For the reverse inclusions, let $r\colon \mathbf{A} \to \mathbf{A}$ be the retraction of $\mathbf{A}$ onto $A\backslash\{a,b,u,v\}$. By Lemma 5.21, the following operation is compatible with $\mathbf{S}$:

$$\operatorname{dom}(\lhd) := \{(x,y) \in A \times B \mid (y = f \implies x \in \{a,b\}) \ \& \ (y = w \implies x \in \{u,v\})\},$$

$$x \lhd y := \begin{cases} x & \text{if } y \in \{f,w\},\\ r(x) & \text{otherwise.}\end{cases}$$

From the inclusions $J \subseteq J_u$ and $K \subseteq K_u$, it follows that $\widehat{u} \lhd \widehat{w}$ is defined and lies in $A^{\mathbf{X}}$. Now, $\varphi(\widehat{u} \lhd \widehat{w}) = u \lhd w = u$, so $\widehat{u} \lhd \widehat{w} \geqslant \widehat{u}$ with respect to $\sqcap$. Using that $\{a,b\}$ and $\{u,v\}$ are up-sets of $\langle A;\sqcap\rangle$, we have $(\widehat{u} \lhd \widehat{w})(i) \in \{a,b\}$ for all $i \in J_u$ and $(\widehat{u} \lhd \widehat{w})(i) \in \{u,v\}$ for all $i \in K_u$. We must then have $\widehat{w}(i) \in \{f,w\}$ for all $i \in J_u \cup K_u$. Taking into account the inclusions $J \subseteq J_u$ and $K \subseteq K_u$, we have further that $\widehat{w}(i) = f$ for all $i \in J_u$ and $\widehat{w}(i) = w$ for all $i \in K_u$. Thus, $J_u = J$ and $K_u = K$.

We will now prove the following claim.

$$\textit{There exists a } (J,K)\textit{-term function } t \textit{ of } \mathbf{S} \textit{ with } t(\widehat{u}) = u. \tag{$\dagger$}$$

To prove this, we split once more into two cases. First, suppose that $\widehat{u}(i) \neq a$ for all $i \in n$. By Lemma 5.20, the following operation is compatible with $\mathbf{S}$:

$$\operatorname{dom}(\ltimes) = (A \times A)\backslash\{(a,a)\}, \qquad x \ltimes y := \begin{cases} u & \text{if } (x,y) = (a,u),\\ u^2 & \text{otherwise.}\end{cases}$$

Now, $\underline{a} \ltimes \widehat{u} \in A^{\mathbf{X}}$ is defined and $\varphi(\underline{a} \ltimes \widehat{u}) = a \ltimes u = u$, so $\underline{a} \ltimes \widehat{u} \geqslant \widehat{u}$ with respect to $\sqcap$. The tuple $\underline{a} \ltimes \widehat{u}$ lies in $\{u,u^2\}^n$, so because $\{a,b\}$ is an up-set of $\langle A;\sqcap\rangle$ and $\underline{a} \ltimes \widehat{u} \geqslant \widehat{u}$, the tuple $\widehat{u}$ cannot have any coordinates in $\{a,b\}$. Thus, $J = J_u = \varnothing$. Further, for all $i \in K$, we have $(\underline{a} \ltimes \widehat{u})(i) = u$ by maximality of $u$ in $\langle A;\sqcap\rangle$, so $\widehat{u}(i) = u$. Now, let $t$ be a projection onto some coordinate in $K$ (which is a $(J,K)$-term function of $\mathbf{S}$, since $J$ is empty). Then $t(\widehat{u}) = u$.

Now consider the case where $\widehat{u}(j) = a$ for some $j \in n$. Then $j \in J$. In this case, choose any $k \in K = K_u$, and let $t$ be the $(J,K)$-term function of $\mathbf{S}$ induced by the $n$-ary term

$$v_j\Bigl(\prod_{i\in J} v_i\Bigr)v_k.$$

Then, since $t(\widehat{u}) \in \{u,v\}$ by (♠), we have $t(\widehat{u}) = u$ by the choice of $j$. This proves ($\dagger$).

To complete the proof of (iii), it suffices by (♠) to show that for all $c \in \{a,b,u,v\}$, we have $\varphi(x) = c \Rightarrow t(x) = c$ for all $x \in A^{\mathbf{X}}$. Choose, via ($\dagger$), a $(J,K)$-term function $t$ of $\mathbf{S}$ such that $t(\widehat{u}) = u$.

Let $x \in A^{\mathbf{X}}$ with $\varphi(x) = u$. Then $x \geqslant \widehat{u}$ with respect to $\sqcap$, so because $t$ preserves the $\sqcap$-order, we have $t(x) \geqslant t(\widehat{u}) = u$. By maximality of $u$ in $\langle A;\sqcap\rangle$, this implies that $t(x) = u$.

Let $x \in A^{\mathbf{X}}$ with $\varphi(x) = a$. Then $\varphi(x\widehat{u}) = \varphi(x)\varphi(\widehat{u}) = au = u$, so by the previous paragraph we have $t(x\widehat{u}) = u$. Now $t(x)u = t(x)t(\widehat{u}) = t(x\widehat{u}) = u$, so $t(x)$ is a left identity for $u$, implying that $t(x) = a$.

Now let $x \in A^{\mathbf{X}}$ with $\varphi(x) = b$, and let $g$ be the non-trivial automorphism of $\mathbf{A}$. Then $\varphi(g(x)) = a$, so $t(g(x)) = a$ by the previous argument, and hence $t(x) = b$. Similarly, if $x \in A^{\mathbf{X}}$ with $\varphi(x) = v$, then $t(x) = v$. By (♠), it follows that $t$ agrees with $\varphi$ on $A^{\mathbf{X}}$. $\square$

**Lemma 5.24.** *Let* $\mathbf{S}$ *be a finite semigroup, and let* $\mathbf{T}$ *be a subsemigroup of* $\mathbf{S}$ *with* $\mathbf{T} \in \mathbb{I}(\mathcal{P})$. *Let* $u \in T \backslash \mathrm{E}(\mathbf{T})$, *and let* $a \in T$ *be the left identity for* $u$. *Let* $\wedge \in \mathcal{T}_2(\mathbf{T})$ *be a meet operation such that* $a$ *and* $u$ *are maximal in* $\langle T; \wedge\rangle$ *and left translation by* $a$ *preserves* $\wedge$. *Let* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T}\rangle$, *let* $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ *for some* $n \in \omega\backslash\{0\}$, *and let* $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ *be a morphism. Assume that* $u \in \varphi(T^{\mathbf{X}})$, *let* $\widehat{u}$ *denote the least element of* $\varphi^{-1}(u) \cap T^{\mathbf{X}}$ *with respect to* $\wedge$, *and define*

$$J := \{i \in n \mid \widehat{u}(i) \mathrel{\mathcal{L}^{\mathbf{T}}} a\}, \qquad K := \{i \in n \mid \widehat{u}(i) \mathrel{\mathcal{L}^{\mathbf{T}}} u\}.$$

*Then the following hold:*

(i) *if* $\mathbf{T} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$, *then every* $(J,K)$*-term function of* $\mathbf{S}$ *agrees with* $\varphi$ *on* $T^{\mathbf{X}}$*;*

(ii) *if* $\mathbf{T} \in \mathbb{I}(\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0)$, *then there exists a* $(J,K)$*-term function of* $\mathbf{S}$ *that agrees with* $\varphi$ *on* $T^{\mathbf{X}}$.

Proof. For (i), take $\mathbf{B} = \mathbf{A} = \mathbf{T}$ and $h = \mathrm{id}_{\mathbf{T}}$ in Lemma 5.23(i).

To prove (ii), let $\mathbf{T}_0$ be the subsemigroup of $\mathbf{T}$ on $T_0 := aT$. Then $\mathbf{T}_0 \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$, and left translation by $a$ is a homomorphism $h\colon \mathbf{T} \to \mathbf{T}_0$ that has $u$ in its image.

To apply Lemma 5.23(iii) with $\mathbf{A} = \mathbf{T}$ and $\mathbf{B} = \mathbf{T}_0$, take the restriction of $\wedge$ to $T_0^2$. We claim that $\underline{a}\widehat{u}$ is the least element of $\varphi^{-1}(u) \cap T_0^{\mathbf{X}}$, with respect to the $\wedge$-order. First, we have $\varphi(\underline{a}\widehat{u}) = au = u$, so $\underline{a}\widehat{u} \in \varphi^{-1}(u) \cap T_0^{\mathbf{X}}$. Let $x \in \varphi^{-1}(u) \cap T_0^{\mathbf{X}}$. Then $x \in \varphi^{-1}(u) \cap T^{\mathbf{X}}$, so $x \geqslant \widehat{u}$ with respect to $\wedge$. Now, left translation by $a$ is $\wedge$-preserving and fixes $T_0$ pointwise, so $x \geqslant \widehat{u}$ implies $x = \underline{a}x \geqslant \underline{a}\widehat{u}$. Thus, $\underline{a}\widehat{u}$ is the least element of $\varphi^{-1}(u) \cap T_0^{\mathbf{X}}$. Finally, we have for all $i \in n$ that $\widehat{u}(i) \mathrel{\mathcal{L}^{\mathbf{T}}} a \Leftrightarrow (\underline{a}\widehat{u})(i) = a$ and $\widehat{u}(i) \mathrel{\mathcal{L}^{\mathbf{T}}} u \Leftrightarrow (\underline{a}\widehat{u})(i) = u$. Hence,

$$J = \{i \in n \mid \underline{a}\widehat{u}(i) = a\}, \qquad K := \{i \in n \mid \underline{a}\widehat{u}(i) = u\}.$$

Now Lemma 5.23(iii) gives the result. □

## 5.7. A dualisability theorem

Using the technical results of Section 5.6, we can now prove that certain direct products of members of $\mathcal{C}$ are dualisable. We will then show that every finite direct product of members of $\mathcal{C}$ generates the same quasivariety as one of the specific products covered by Lemma 5.25. By the Independence of Generator Theorem 4.6, it will follow that all finite direct products of members of $\mathcal{C}$ are dualisable.

**Lemma 5.25.** *Let* $\mathbf{1}$ *denote the trivial semigroup, let*

$$\mathbf{T} \in \mathbb{I}(\mathcal{P}), \quad \mathbf{U} \in \mathbb{I}(\mathbf{1}, \mathbf{L}, \mathbf{L}^0, \mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0), \quad \mathbf{V} \in \mathbb{I}(\mathbf{1}, \mathbf{R}, \mathbf{R}^0),$$

*and let* $\mathbf{S}$ *be isomorphic to* $\mathbf{T} \times \mathbf{U} \times \mathbf{V}$. *Assume that the following implications hold:*

(i) *if* $\mathbf{U}$ *is non-trivial, then* $\mathbf{T} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$*;*

(ii) *if, furthermore,* $\mathbf{U} \in \mathbb{I}(\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0)$, *then there is a homomorphism* $h\colon \mathbf{U} \to \mathbf{T}$ *whose image contains the non-idempotent element of* $\mathbf{T}$.

*Then* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$.

Proof. Let $\mathbf{X} \leqslant \underset{\sim}{\mathbf{S}}^n$ for some $n \in \omega\backslash\{0\}$, and let $\varphi\colon \mathbf{X} \to \underset{\sim}{\mathbf{S}}$ be a morphism. We must show that $\varphi$ extends to a term function of $\mathbf{S}$. Note that $\mathrm{E}(\mathbf{T})$, $\mathrm{E}(\mathbf{U})$, and $\mathbf{V}$ are normal bands, so $\mathrm{E}(\mathbf{S})$ is a normal band.

First, assume that $\varphi(X) \subseteq \mathrm{E}(\mathbf{S})$. Applying the Term Retract Lemma 5.8, take $\mathbf{A}$ to be the normal band $\mathrm{E}(\mathbf{S})$, take $\underset{\sim}{\mathbf{A}} := \langle A; \mathcal{P}_2(\mathbf{A}), \mathcal{T}\rangle$, and take $r$ to be the square map $\mathbf{S} \to \mathbf{A}$. Then $\varphi$ extends to a term function of $\mathbf{S}$, by Theorem 4.42 and the Term Retract Lemma 5.8. Thus, for the remainder of the proof, we can assume that $\varphi(X)$ contains a non-idempotent element of $\mathbf{S}$, say $v$.

Let $u \in T\backslash\mathrm{E}(\mathbf{T})$, and let $a$ be the left identity for $u$ in $\mathbf{T}$. We claim that there is an endomorphism of $\mathbf{S}$ taking $v$ to $u$. Since $\mathbf{S} \in \mathbb{SP}(\mathbf{T}, \mathbf{U}, \mathbf{V})$ and $\mathbf{V}$ is a band, there is a homomorphism, $k$, from $\mathbf{S}$ into either $\mathbf{T}$ or $\mathbf{U}$ with $k(v) \neq k(v^2)$. In the former case (i.e., $k\colon \mathbf{S} \to \mathbf{T}$), we can follow $k$ by an automorphism of $\mathbf{T}$ if necessary to get the required endomorphism sending $v$ to $u$. Otherwise, if $k\colon \mathbf{S} \to \mathbf{U}$, then $\mathbf{U}$ is not a band, so the assumptions of (i) and (ii) hold, and then $h \circ k$ must send $v$ to $u$. As $\varphi(X)$ contains $v$ and is closed under the endomorphisms of $\mathbf{S}$, it follows that $u \in \varphi(X)$.

By Lemma 4.26, we can assume that $\mathbf{T}, \mathbf{U}, \mathbf{V} \leqslant_{\mathrm{r}} \mathbf{S}$. Then $\underset{\sim}{\mathbf{S}} \Vdash T, U, V$, so $\varphi$ maps $T^{\mathbf{X}}$, $U^{\mathbf{X}}$, and $V^{\mathbf{X}}$ respectively into $T$, $U$, and $V$. Let $\wedge \in \mathcal{T}_2(\mathbf{T})$ be a meet operation such that $a$ and $u$ are maximal in $\langle T; \wedge\rangle$, left translation by $a$ preserves $\wedge$, and $L_a^{\mathbf{T}}$, $L_u^{\mathbf{T}}$ are up-sets of $\langle T; \wedge\rangle$; such an operation exists by Theorem 5.18. As $u \in \varphi(X)$, we have $u = \varphi(x)$ for some $x \in X$. But there is a retraction $r$ of $\mathbf{S}$ onto $\mathbf{T}$, so $u = r(u) = r(\varphi(x)) = \varphi(r(x)) \in \varphi(T^{\mathbf{X}})$. Thus, we may let $\widehat{u}$ be the least element of $\varphi^{-1}(u) \cap T^{\mathbf{X}}$ with respect to $\wedge$, and also define

$$J := \{i \in n \mid \widehat{u}(i) \mathrel{\mathcal{L}^{\mathbf{T}}} a\}, \qquad K := \{i \in n \mid \widehat{u}(i) \mathrel{\mathcal{L}^{\mathbf{T}}} u\}.$$

By Lemma 4.31, it suffices to show that there is an $n$-ary term function of $\mathbf{S}$ that agrees with $\varphi$ on $T^{\mathbf{X}} \cup U^{\mathbf{X}} \cup V^{\mathbf{X}}$. We will first show that there is a $(J, K)$-term function of $\mathbf{S}$ that agrees with $\varphi$ on $T^{\mathbf{X}} \cup U^{\mathbf{X}}$. We consider three cases:

(1) $\mathbf{U}$ is trivial;
(2) $\mathbf{U} \in \mathbb{I}(\mathbf{L}, \mathbf{L}^0)$;
(3) $\mathbf{U} \in \mathbb{I}(\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0)$.

Case (1): Assume that $\mathbf{U}$ is trivial. We claim that there exists a $(J, K)$-term function of $\mathbf{S}$ that agrees with $\varphi$ on $T^{\mathbf{X}}$. We will apply Lemma 5.24 to $\mathbf{T}$. In case (ii) of Lemma 5.24, we are done; in case (i), note that $K \neq \varnothing$ by Lemma 5.23, so there is at least one $(J, K)$-term function. We conclude in both cases that there exists a $(J, K)$-term function of $\mathbf{S}$ that agrees with $\varphi$ on $T^{\mathbf{X}}$, and also on $U^{\mathbf{X}}$, because $\mathbf{U}$ is trivial.

Case (2): Since $\mathbf{T} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$ by (i), we may apply Lemma 5.23(ii) with $\mathbf{B} = \mathbf{T}$ and $\mathbf{A} = \mathbf{U}$. We conclude that there is a $(J, K)$-term function $t$ of $\mathbf{S}$ that agrees with $\varphi$ on $U^{\mathbf{X}}$. By Lemma 5.24(i) applied to $\mathbf{T}$, the term function $t$ also agrees with $\varphi$ on $T^{\mathbf{X}}$.

Case (3): Again we have $\mathbf{T} \in \mathbb{I}(\mathbf{P}, \mathbf{C}, \mathbf{C}^0)$ by (i). Applying Lemma 5.23(iii) with $\mathbf{A} = \mathbf{U}$ and $\mathbf{B} = \mathbf{T}$, there is a $(J, K)$-term function $t$ of $\mathbf{S}$ that agrees with $\varphi$ on $U^{\mathbf{X}}$, and by applying Lemma 5.24(i) to $\mathbf{T}$, we find that the term function $t$ also agrees with $\varphi$ on $T^{\mathbf{X}}$.

Thus, there is a $(J, K)$-term function $t$ of $\mathbf{S}$ that agrees with $\varphi$ on $T^{\mathbf{X}} \cup U^{\mathbf{X}}$. To complete the proof, we need only see that $t$ also agrees with $\varphi$ on $V^{\mathbf{X}}$. This is obvious if $\mathbf{V}$ is trivial, and if $\mathbf{V}$ is non-trivial, then we may apply Lemma 5.22(ii) with $\mathbf{B} = \mathbf{T}$ and $\mathbf{A} = \mathbf{V}$. $\square$

The following three short lemmas are now all that we need for the main result.

**Lemma 5.26.** *Let $\mathfrak{U} \subseteq \mathfrak{C}$ contain precisely one* $\mathbf{P}$*-like semigroup. Then there is a finite semigroup* $\mathbf{S}$ *such that* $\mathbb{SP}(\mathbf{S}) = \mathbb{SP}(\mathfrak{U})$ *and* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$.

Proof. Let $\mathbf{T}$ be the $\mathbf{P}$-like semigroup in $\mathfrak{U}$. Since $\mathbf{N}$ and $\mathbf{I}$ embed into $\mathbf{T}$, we may assume that $\mathfrak{U}\backslash\{\mathbf{T}\} \subseteq \{\mathbf{L}, \mathbf{L}^0, \mathbf{R}, \mathbf{R}^0\}$. Thus, $\mathbb{SP}(\mathfrak{U})$ can be generated by $\mathbf{S} \cong \mathbf{T} \times \mathbf{U} \times \mathbf{V}$ for some choice of $\mathbf{U} \in \{\mathbf{1}, \mathbf{L}, \mathbf{L}^0\}$ and $\mathbf{V} \in \{\mathbf{1}, \mathbf{R}, \mathbf{R}^0\}$. Now, if $\mathbf{T} \in \mathbb{I}(\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0)$, then $\mathbf{U}$ embeds into $\mathbf{T}$, in which case we can replace $\mathbf{U}$ with the trivial semigroup without changing $\mathbb{SP}(\mathbf{S})$. Thus, we can assume without loss of generality that condition (i) of Lemma 5.25 holds. Since (ii) holds vacuously, the desired result follows from Lemma 5.25. ☐

**Lemma 5.27.** *Let $\mathfrak{U} \subseteq \mathfrak{C}$, and assume that every $\mathfrak{U}' \subseteq \mathfrak{C}$ with $\mathbb{SP}(\mathfrak{U}') = \mathbb{SP}(\mathfrak{U})$ contains at least two* $\mathbf{P}$*-like semigroups. Then $\mathbb{SP}(\mathfrak{U})$ can be generated by one of the following sets:*

$$\{\mathbf{P}, \mathbf{E}^0\}, \{\mathbf{C}^0, \mathbf{E}\}, \{\mathbf{P}, \mathbf{E}\}, \{\mathbf{P}, \mathbf{E}, \mathbf{R}^0\}.$$

Proof. We can assume that $\mathfrak{U}$ is an antichain with respect to the embedding order; i.e., if $\mathbf{U}, \mathbf{V} \in \mathfrak{U}$ with $\mathbf{U} \in \mathbb{S}(\mathbf{V})$, then $\mathbf{U} = \mathbf{V}$.

We have $\mathbb{SP}(\mathbf{L} \circlearrowleft L, \mathbf{E}) = \mathbb{SP}(\mathbf{L} \circlearrowleft L, \mathbf{R})$ and $\mathbb{SP}(\mathbf{L} \circlearrowleft L, \mathbf{E}^0) = \mathbb{SP}(\mathbf{L} \circlearrowleft L, \mathbf{R}^0)$ by Proposition 5.17. Thus, if $\mathfrak{U}$ contains $\mathbf{L} \circlearrowleft L$ and $\mathbf{E}$ ($\mathbf{E}^0$), then $\mathbf{E}$ ($\mathbf{E}^0$) can be replaced by $\mathbf{R}$ ($\mathbf{R}^0$) without changing $\mathbb{SP}(\mathfrak{U})$. We can therefore assume that $\mathfrak{U}$ contains at most one member of $\{\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0\}$. Similarly, we can assume $\mathfrak{U}$ contains at most one member of $\{\mathbf{P}, \mathbf{C}, \mathbf{C}^0\}$. Thus, $\mathfrak{U} \cap \mathcal{P} = \{\mathbf{T}, \mathbf{U}\}$ for some $\mathbf{T} \in \{\mathbf{P}, \mathbf{C}, \mathbf{C}^0\}$ and some $\mathbf{U} \in \{\mathbf{L} \circlearrowleft L, \mathbf{E}, \mathbf{E}^0\}$.

Suppose that $\mathbf{U} = \mathbf{L} \circlearrowleft L$. Then we can assume that $\mathbf{T} = \mathbf{P}$ (otherwise, if $\mathbf{T} \in \{\mathbf{C}, \mathbf{C}^0\}$, then $\mathbf{T}$ can be replaced by $\mathbf{R}$ or $\mathbf{R}^0$). But then $\mathbf{T}$ embeds into $\mathbf{U}$, which is impossible as $\mathfrak{U}$ is an antichain. So the case $\mathbf{U} = \mathbf{L} \circlearrowleft L$ cannot occur; thus, $\mathbf{U} \in \{\mathbf{E}, \mathbf{E}^0\}$. Now, noting that $\mathbf{C} \in \mathbb{S}(\mathbf{E})$ and $\mathbf{C}, \mathbf{C}^0 \in \mathbb{S}(\mathbf{E}^0)$, we see that $\mathbf{U} = \mathbf{E}$ implies $\mathbf{T} \in \{\mathbf{P}, \mathbf{C}^0\}$, while $\mathbf{U} = \mathbf{E}^0$ implies $\mathbf{T} = \mathbf{P}$. This narrows $\mathfrak{U} \cap \mathcal{P}$ down to the three sets $\{\mathbf{P}, \mathbf{E}^0\}$, $\{\mathbf{C}^0, \mathbf{E}\}$, and $\{\mathbf{P}, \mathbf{E}\}$.

If $\mathfrak{U} \cap \mathcal{P} = \{\mathbf{P}, \mathbf{E}^0\}$, then since $\mathfrak{C}\backslash\mathcal{P} \subseteq \mathbb{S}(\mathbf{E}^0)$, we have $\mathfrak{U} = \{\mathbf{P}, \mathbf{E}^0\}$. If $\mathfrak{U} \cap \mathcal{P} = \{\mathbf{C}^0, \mathbf{E}\}$, then since $\mathbf{R}^0 \in \mathbb{S}(\mathbf{C}^0)$ and $\mathbf{L}^0 \in \mathbb{S}(\mathbf{E})$, we have $\mathfrak{C}\backslash\mathcal{P} \subseteq \mathbb{S}(\mathbf{C}^0, \mathbf{E})$, and so $\mathfrak{U} = \{\mathbf{C}^0, \mathbf{E}\}$. Finally, if $\mathfrak{U} \cap \mathcal{P} = \{\mathbf{P}, \mathbf{E}\}$, then $\mathfrak{C}\backslash(\mathcal{P} \cup \{\mathbf{R}^0\}) \subseteq \mathbb{S}(\mathbf{E})$, so $\mathfrak{U}\backslash\mathcal{P}$ is either $\varnothing$ or $\{\mathbf{R}^0\}$. ☐

**Lemma 5.28.** *Let $\mathfrak{U}$ be one of the following subsets of $\mathfrak{C}$:*

$$\{\mathbf{P}, \mathbf{E}^0\}, \{\mathbf{C}^0, \mathbf{E}\}, \{\mathbf{P}, \mathbf{E}\}, \{\mathbf{P}, \mathbf{E}, \mathbf{R}^0\}.$$

*Then there is a finite semigroup* $\mathbf{S}$ *such that* $\mathbb{SP}(\mathbf{S}) = \mathbb{SP}(\mathfrak{U})$ *and* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$.

Proof. Note that $\mathbb{SP}(\mathfrak{U})$ may be generated by $\mathbf{S} \cong \mathbf{T} \times \mathbf{U} \times \mathbf{V}$, where $(\mathbf{T}, \mathbf{U})$ is one of $(\mathbf{P}, \mathbf{E}^0)$, $(\mathbf{C}^0, \mathbf{E})$, or $(\mathbf{P}, \mathbf{E})$ and $\mathbf{V}$ is either $\mathbf{R}^0$ or trivial. By considering the possible choices of $\mathbf{T}$ and $\mathbf{U}$, it is easy to see that there is a homomorphism $h\colon \mathbf{U} \to \mathbf{T}$ containing the non-idempotent element of $\mathbf{T}$ in its image. (The key ideas are: $\mathbf{P}$ is a homomorphic image of $\mathbf{E}$ and $\mathbf{E}^0$, while factoring $\mathbf{E}$ by its $\mathcal{L}$ relation yields $\mathbf{C}$, which embeds into $\mathbf{C}^0$.) Thus, $\mathbf{T}$, $\mathbf{U}$, and $\mathbf{V}$ satisfy conditions (i) and (ii) of Lemma 5.25, so the result follows. ☐

Finally, we can assemble the main result of this chapter.

**Theorem 5.29.** *Let $\mathfrak{U} \subseteq \mathfrak{C}$. Then there is a finite semigroup* $\mathbf{S}$ *such that* $\mathbb{SP}(\mathbf{S}) = \mathbb{SP}(\mathfrak{U})$ *and* $\underset{\sim}{\mathbf{S}} := \langle S; \mathcal{P}_2(\mathbf{S}), \mathcal{T}\rangle$ *satisfies* (IC) *with respect to* $\mathbf{S}$. *Consequently, every finite direct product of members of* $\mathfrak{C}$ *is dualisable.*

Proof. If $\mathfrak{U}$ does not contain a $\mathbf{P}$-like semigroup, then $\mathfrak{U} \subseteq \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N})$ and the result follows from Corollary 5.10. Thus, we can assume that $\mathfrak{U}$ contains a $\mathbf{P}$-like semigroup, and therefore every $\mathfrak{U}' \subseteq \mathfrak{C}$ with $\mathbb{SP}(\mathfrak{U}') = \mathbb{SP}(\mathfrak{U})$ contains a $\mathbf{P}$-like semigroup. If there is some $\mathfrak{U}' \subseteq \mathfrak{C}$ such that $\mathbb{SP}(\mathfrak{U}') = \mathbb{SP}(\mathfrak{U})$ and $\mathfrak{U}'$ contains exactly one $\mathbf{P}$-like semigroup, then the result follows from Lemma 5.26 applied to $\mathfrak{U}'$. Otherwise, the result follows from Lemma 5.27 and Lemma 5.28. The 'consequently' part follows from the (IC) Duality Theorem 4.5 and the Independence of Generator Theorem 4.6. □

CHAPTER 6

# The Dualisable Class

Although Theorem 5.29 covers a lot of ground, significantly extending the dualisability result for normal bands (Theorem 4.41), it deals with only finitely many subquasivarieties of $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$. Theorem 5.29 was the culmination of about two years of dualisability proofs, but in terms of proving Conjecture 5.16, there was still quite a long road ahead.

Not too long after proving Theorem 5.29, the author visited Ross Willard at the University of Waterloo for a research collaboration, where we spent two weeks on problems relating to the dualisability problem for aperiodic semigroups. In all, we looked at only three specific semigroups, but they proved to be quite important examples.

The first two semigroups we looked at were the two possible one-point inflations of the two-element cyclic group; as mentioned in Section 5.3, this led the author almost immediately to Theorem 5.9. Not only did this greatly simplify the proof of Corollary 5.10, it inspired many simplifications of the proofs leading up to Theorem 5.29, which was originally proved by separately considering each possible direct product of quasicriticals. Although the proof of Theorem 5.29 in its current form is fairly complicated, it could easily have been two or three times its current length.

The simplifications to Theorem 5.29 gave the author some hope that a proof of Conjecture 5.16 would actually be possible (if not during the author's candidature, then during their lifetime). If the structure of the quasicriticals in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ could be narrowed down to a sufficiently transparent form, then perhaps, by generalising the techniques of Chapter 5, it could be proved that all finite direct products of these quasicriticals are dualisable.

This was essentially the plan of attack that the author posed to Ross Willard, but he suggested that we look at another example, so we did (the author continued to study the quasicriticals on the side, with limited success). After the quasicriticals covered in Theorem 5.29, the next largest quasicritical in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ was the semigroup in Figure 6.1. This semigroup arises from the construction in Theorem 3.63 by taking $\kappa = 1$.

It was clear almost immediately that there was something different about this semigroup. The proof of Theorem 5.29 relied on the existence of compatible meet operations for the quasicriticals (Theorem 5.18), but it turned out that the semigroup in Figure 6.1 does not have a compatible meet operation. Still, progress was made by using meet operations on certain subsemigroups (such as $\mathbf{L} \circlearrowright L$), and by the end of the first week, we came quite close to a dualisability proof.

In the meantime, while studying quasicriticals, the author found that all subdirectly irreducibles in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ of size greater than three are of the form $\mathbf{T} \circlearrowright U$ for some left-normal band $\mathbf{T}$ acting totally on a set $U$. The author then suspected that all quasicriticals in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ may be constructed from the finite subdirectly irreducibles in the same manner that $\mathbf{E}$ and $\mathbf{E}^0$ were constructed from $\mathbf{L} \circlearrowright L$. Unfortunately, working with quasicriticals in general was much more difficult, and there did not seem to be any humanly way to prove this conjecture. However, we were unknowingly closing in on the dualisable class.

| $\cdot$ | $a$ | $b$ | $c$ | $d$ | $u$ | $v$ | $p$ | $0$ |
|---|---|---|---|---|---|---|---|---|
| $a$ | $a$ | $a$ | $a$ | $a$ | $u$ | $u$ | $u$ | $0$ |
| $b$ | $b$ | $b$ | $b$ | $b$ | $v$ | $v$ | $v$ | $0$ |
| $c$ | $a$ | $a$ | $c$ | $c$ | $u$ | $u$ | $p$ | $0$ |
| $d$ | $b$ | $b$ | $d$ | $d$ | $v$ | $v$ | $p$ | $0$ |
| $u$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ |
| $v$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ |
| $p$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ |
| $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ | $0$ |

FIGURE 6.1. A subdirectly irreducible in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$.

### 6.1. Preliminaries

In preparation for the main result, we will devote this short section to a study of the variety $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$. In this chapter, we will be working exclusively in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$; thus, the equations $x^2y \approx xy$ and $xyzt \approx xzyt$ will be used frequently without mention (see Proposition 5.15). We will also use freely the identities in the following result.

**Proposition 6.1.** *Let $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$. Then the following identities hold in $\mathbf{S}$:*

$$x^n \approx x^2 \ \ (n \geqslant 2), \qquad (xy)^2 \approx x^2y^2 \approx xy^2, \qquad xyz \approx xyxz.$$

*Consequently,* $\mathrm{E}(\mathbf{S})$ *is a normal band and all left translations in* $\mathbf{S}$ *are endomorphisms of* $\mathbf{S}$.

The next three results provide us with useful consequences of the theory developed in Chapters 2 and 3.

**Proposition 6.2.** *Every semigroup in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ has $\mathcal{J} = \mathcal{D}$.*

PROOF. Since $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P}) \models x^3 \approx x^2$, this follows from Theorem 2.38. □

**Proposition 6.3.** *Let $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, and let $a \in S$. Then $a \in \mathrm{E}(\mathbf{S})$ if and only if $a \preccurlyeq_J a^2$.*

PROOF. Since $\mathbf{S}$ is both periodic and aperiodic, this follows from Theorem 2.42. □

**Proposition 6.4.** *Let $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, and let $a \in S$. If $a$ is idempotent, then $J_a$ is a rectangular band. Therefore, if $a^2 \neq a$, then $J_a$ contains no idempotents.*

PROOF. Assume that $a$ is idempotent. Every nilpotent semigroup satisfying $x^2y \approx xy$ is null, so by Theorem 3.13, $J_a$ is a completely simple subsemigroup of $\mathbf{S}$. A completely simple semigroup is completely regular by Theorem 2.43, so $J_a$ is a rectangular band by aperiodicity. This shows in particular that the $\mathcal{J}$-class of an idempotent contains only idempotents, so the last claim is immediate. □

The remaining results in this section reveal some important special features of Green's relations in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$.

**Proposition 6.5.** *Let* $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$*, and let* $a \in S \backslash \mathrm{E}(\mathbf{S})$*. Then* $R_a = \{a\}$ *and* $J_a = L_a$*.*

Proof. Let $b \in R_a$, so $b = au$ for some $u \in S^1$. If $u \neq 1$, then $a \mathrel{\mathcal{J}} b = au = a^2u \preccurlyeq_J a^2$, contradicting Proposition 6.3. Therefore, $u = 1$, so $b = a$.

This shows that $R_a = \{a\}$. Thus, if $c \in J_a$, then by Propositions 2.16 and 6.2 the set $R_a \cap L_c = \{a\} \cap L_c$ is non-empty, so $c \in L_a$. □

**Proposition 6.6.** *Let* $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$*. Then* $\preccurlyeq_J$ *is compatible with* $\mathbf{S}$*.*

Proof. It suffices to show that $\preccurlyeq_J$ is left- and right-compatible. Let $a, b, c \in S$ with $a \preccurlyeq_J b$, so $a = xby$ for some $x, y \in S^1$. Then $ca = cxby = cxcby \in \langle cb \rangle$ and $ac = xbyc = xbybc \in \langle bc \rangle$, so $ca \preccurlyeq_J cb$ and $ac \preccurlyeq_J bc$, as required. □

**Proposition 6.7.** *Let* $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$*, and let* $a, b \in \mathrm{E}(\mathbf{S})$*. Then* $a \preccurlyeq_J b \Leftrightarrow aba = a$*.*

Proof. Let $a, b \in S$. Assume that $a \preccurlyeq_J b$. Then $a = aaa \preccurlyeq_J aba$ by Proposition 6.6, and clearly $aba \preccurlyeq_J a$, so $a \mathrel{\mathcal{J}} aba$. By Proposition 6.4, $J_a$ is a rectangular band containing $a$ and $aba$, so $aba = a(aba)a = a$. Conversely, if $aba = a$, then $a = aba \preccurlyeq_J b$. □

For $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, it can be shown that $\preccurlyeq_L$ is also compatible, though we will not use this fact. However, the quasi-order $\preccurlyeq_R$ is not compatible in general; indeed, $\preccurlyeq_R$ can fail to be compatible even in $\mathbb{V}(\mathbf{P})$. The reader may verify, for example, that $\preccurlyeq_R$ is not compatible with the semigroup $\mathbf{S} := \mathbf{I} \times \mathbf{P}$. Nonetheless, $\mathcal{R}$ is a congruence on all semigroups in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, as the final result in this section shows. (In fact, it can easily be shown from the results in this section that *all* of Green's relations are congruences in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, though we will not need this result.)

**Proposition 6.8.** *Let* $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$*. Then for all* $a, b, c \in S$*, we have*

$$\big(a \mathrel{\mathcal{L}} b \implies ca = cb\big) \;\&\; \big(a \mathrel{\mathcal{R}} b \implies ac = bc\big).$$

*Consequently,* $\mathcal{L}$ *and* $\mathcal{R}$ *are congruences on* $\mathbf{S}$*.*

Proof. We will use only the fact that $\mathbf{S} \models \{xyzt \approx xzyt,\ xy^2z \approx xyz\}$, so it suffices by symmetry to show that $a \mathrel{\mathcal{L}} b \Rightarrow ca = cb$ for all $a, b, c \in S$ and that $\mathcal{L}$ is a congruence.

Let $a, b \in S$ with $a \neq b$ and $a \mathrel{\mathcal{L}} b$, and let $x, y \in S$ with $xb = a$ and $ya = b$. Then $xya = a$ and $yxya = b$, so for all $c \in S$, we have $ca = cxya = cxyya = cyxya = cb$. This shows in particular that $\mathcal{L}$ is left-compatible, so $\mathcal{L}$ is a congruence by Proposition 2.15. □

## 6.2. A quasiequation

During the final week of the author's collaboration with Ross Willard, we continued searching for a dualisability proof for the semigroup in Figure 6.1, which we will denote by $\mathbf{S}_1$ for the moment. Though we thought we were close to a proof at the start of the week, it kept fighting back against any new partial operations we introduced. So, we soon shifted course and tried to prove that $\mathbf{S}_1$ is non-dualisable. By the end of the collaboration, we had established that it is not *finitely* dualisable (i.e., not dualisable by an alter ego of finite type). This was quite strong evidence that $\mathbf{S}_1$ is non-dualisable. The author typed up the proof on the way home from Canada and began dissecting it.

The original non-finite-dualisability proof for $\mathbf{S}_1$ was a somewhat technical proof taking up several pages, but the author eventually stumbled across a one-line calculation that captured the non-dualisability of $\mathbf{S}_1$. This calculation showed that $\mathbf{S}_1$ is in fact inherently non-dualisable, and the proof was then easily generalised so that it no longer relied on $\mathbf{S}_1$ having a zero element; an arbitrary ideal would work in place of the zero.

This generalisation had huge implications for the dualisability problem. The author had already shown that $\mathbf{S}_1$ embeds into every larger finite subdirectly irreducible in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, so this result implied that all of these larger subdirectly irreducibles are inherently non-dualisable as well. Furthermore, the removal of the dependence on the zero element suggested that all quasicriticals of size greater than seven are inherently non-dualisable, but some more work was needed to confirm this suspicion.

In contrast to the rather unwelcome discovery that Conjecture 5.3 is false, the disproof of Conjecture 5.16 gave the author a profound sense of relief. Theorem 5.29, which was thought to be a very special case of a dualisability theorem for $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, now appeared to be all that was needed on the dualisability side. On the non-dualisability side, it seemed that the proof for $\mathbf{S}_1$ would imply the non-dualisability of all of the remaining semigroups. The only thing left to do was some semigroup theory.

Although no more progress had been made on describing the quasicriticals in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, the author pondered the problem from a different point of view after obtaining the inherent non-dualisability proof for $\mathbf{S}_1$. It would be quite auspicious if there were some quasiequations differentiating the quasicriticals in $\mathfrak{C}$ (Section 5.5) from all of the larger quasicriticals. And lo, within minutes of having this thought, the following quasiequation came to mind:

$$xw \approx yw \rightarrow xzw \approx yzw.$$

This was, in effect, the solution to the dualisability problem for aperiodic semigroups. After the several weeks spent on $\mathbf{S}_1$, it was clear that this quasiequation captured the essential difference between $\mathbf{S}_1$ and the quasicriticals in $\mathfrak{C}$; in particular, it is satisfied by the semigroups in $\mathfrak{C}$, but not by $\mathbf{S}_1$. The next obvious step was to show that it fails in all larger quasicriticals. However, a much more convenient result was quickly discovered: a failure of $xw \approx yw \rightarrow xzw \approx yzw$ for a finite $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ implies the existence of an inherently non-dualisable subsemigroup. This is the content of the next two results.

**Lemma 6.9.** *Let* $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ *such that* $\mathbf{S} \not\models xw \approx yw \rightarrow xzw \approx yzw$. *Then there are* $a, b, c, d \in \mathrm{E}(\mathbf{S})$ *and* $u, v, p \in S \backslash \mathrm{E}(\mathbf{S})$ *such that* $|\{a, b, c, d, u, v, p\}| = 7$ *and the following products hold in* $\mathbf{S}$ (*cf. Figure 6.1*).

| $\cdot$ | $a$ | $b$ | $c$ | $d$ | $u$ | $v$ | $p$ |
|---|---|---|---|---|---|---|---|
| $a$ | $a$ | $a$ | $a$ | $a$ | $u$ | $u$ | $u$ |
| $b$ | $b$ | $b$ | $b$ | $b$ | $v$ | $v$ | $v$ |
| $c$ | $a$ | $a$ | $c$ | $c$ | $u$ | $u$ | $p$ |
| $d$ | $b$ | $b$ | $d$ | $d$ | $v$ | $v$ | $p$ |

*Moreover, in the subsemigroup* $\mathbf{T}$ *of* $\mathbf{S}$ *generated by* $\{a, b, c, d, u, v, p\}$, *the ideal* $\langle p^2 \rangle$ *of* $\mathbf{T}$ *coincides with* $T \backslash \{a, b, c, d, u, v, p\}$.

Proof. Choose $w, x, y, z \in S$ such that $xw = yw$ and $xzw \neq yzw$. Since $\mathbf{S} \models x^2y \approx xy$, we may assume that $x, y, z \in \mathrm{E}(\mathbf{S})$, so that the following four elements are also idempotent:

$$c := xy, \quad d := yxy, \quad a := czc, \quad b := dzd.$$

It is easily verified that $\{c, d\}$ is a left-zero subsemigroup of $\mathbf{S}$. Since $c \mathrel{\mathcal{L}} d$, we have $a \mathrel{\mathcal{L}} b$ by Proposition 6.8, so $\{a, b\}$ is also a left-zero subsemigroup of $\mathbf{S}$. By construction, $a$ and $b$ are zeroes for $c$ and $d$, respectively; thus, by Proposition 6.8, we have

$$ad = ac = a, \qquad bc = bd = b, \qquad cb = ca = a, \qquad da = db = b.$$

This suffices to verify the multiplication on $\{a, b, c, d\}$. Next, define

$$p := xw = yw, \quad u := ap, \quad v := bp.$$

Then $xp = yp = p$, so by the definitions of $c$ and $d$ we have $cp = dp = p$. Using the table of products for $\{a, b, c, d\}$, the remainder of the table in the statement is easily verified. We must show that $|\{a, b, c, d, u, v, p\}| = 7$ and that $u, v, p \notin \mathrm{E}(\mathbf{S})$. We have

$$u = ap = czcp = czp = xyzp = xzyp = xzyyw = xzyw = xzxw = xzw,$$

and similarly $v = yzw$, so $u \neq v$. Since $cp = p$ and $cv \neq v$, we have $p \neq v$, and by symmetry we have $p \neq u$. Thus, $|\{u, v, p\}| = 3$.

Now, if $u \in \mathrm{E}(\mathbf{S})$, then $v^2 = (bu)^2 = bu^2 = bu = v$, so $v \in \mathrm{E}(\mathbf{S})$. By symmetry, $u \in \mathrm{E}(\mathbf{S})$ if and only if $v \in \mathrm{E}(\mathbf{S})$. Suppose that $u, v \in \mathrm{E}(\mathbf{S})$. Then $cap^2 = (cap)^2 = u^2 = u$, and we similarly have $dap^2 = v$. But $cp = dp$, so $u = capp = cpap = dpap = dapp = v$, which is a contradiction. Thus, $u, v \notin \mathrm{E}(\mathbf{S})$. Since $ap = u \neq u^2 = (ap)^2 = ap^2$, we have $p^2 \neq p$, and so $u, v, p \notin \mathrm{E}(\mathbf{S})$.

Now, $\{a, b, c, d\} \cap \{u, v, p\} = \varnothing$ as $a, b, c, d \in \mathrm{E}(\mathbf{S})$. Since $|\{u, v, p\}| = 3$ and $a, b, c, d$ have pairwise distinct actions on $\{u, v, p\}$ by left translation, we must have $|\{a, b, c, d\}| = 4$, so $|\{a, b, c, d, u, v, p\}| = 7$.

To prove the 'moreover' part, let $X := \{a, b, c, d, u, v, p\}$; we must show that $T \backslash X = \langle p^2 \rangle$. Note that $u = ap = cbdv$, so $u \preccurlyeq_J x$ for all $x \in X$. Suppose there is some $x \in X \cap \langle p^2 \rangle$. Then $u \preccurlyeq_J x \preccurlyeq_J p^2$, so $u \preccurlyeq_J p^2$, and Proposition 6.6 gives $u = au \preccurlyeq_J ap^2 = (ap)^2 = u^2$, contradicting Proposition 6.3. Thus, $\langle p^2 \rangle \subseteq T \backslash X$.

Conversely, let $x \in T \backslash X$. Since $X$ generates $\mathbf{T}$, there exist $n \geqslant 1$ and $x_1, \ldots, x_n, y \in X$ such that $x = x_1 \cdots x_n y$. Now, if $x_1, \ldots, x_n \in \{a, b, c, d\}$, then $x \in X$, contradicting the choice of $x$, so we must have $\{u, v, p\} \cap \{x_1, \ldots, x_n\} \neq \varnothing$. Therefore, since $u, v \in Tp$, we have $x \in T^1 pT \subseteq \langle p^2 \rangle$. Thus, $T \backslash X = \langle p^2 \rangle$. $\square$

Since a failure of $xw \approx yw \to xzw \approx yzw$ for a semigroup in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ implies the existence of at least eight elements, we have the following consequence of Lemma 6.9.

**Corollary 6.10.** *Let $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$. If $|S| \leqslant 7$, then $\mathbf{S} \models xw \approx yw \to xzw \approx yzw$.*

The next result is a generalised proof that $\mathbf{S}_1$ is inherently non-dualisable. The author is indebted to Ross Willard for his many contributions to the original version of the proof, which are hardly reflected in the almost disappointingly short final version.

**Theorem 6.11.** *Let* $\mathbf{S} \in \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ *be finite with* $\mathbf{S} \not\models xw \approx yw \to xzw \approx yzw$*. Then* $\mathbf{S}$ *is inherently non-dualisable.*

Proof. Choose $a, b, c, d, u, v, p \in S$ as in Lemma 6.9. To prove that $\mathbf{S}$ is IND, it suffices to show that $\mathbf{S}$ has an IND subsemigroup, so we may assume without loss of generality that $\mathbf{S}$ is generated by $X := \{a, b, c, d, u, v, p\}$. By Lemma 6.9, the ideal $I := \langle p^2 \rangle$ is equal to $S \backslash X$.

We claim that $S\backslash\{a, c, u\} = \{b, d, v, p\} \cup I$ is a right ideal of $\mathbf{S}$. Since $I = S\backslash X$ is an ideal, it suffices to show that $\{v, p\}S \subseteq I$ and $\{b, d\}X \subseteq \{b, d, v, p\}$. First, if $x \in \{v, p\}$, then for all $y \in S$, we have $xy \preccurlyeq_J py = p^2y \preccurlyeq_J p^2$, so $xy \in I$. Also, if $x \in \{b, d\}$ and $y \in X$, then $xy \in \{b, d, v, p\}$ from the table in Lemma 6.9. Thus, $\{b, d, v, p\} \cup I$ is a right ideal.

We will apply Theorem 4.8. Let $Z$ be any infinite set. Since $S\backslash\{a, c, u\}$ is a right ideal of $\mathbf{S}$, we may let $\mathbf{B}$ be the subsemigroup of $\mathbf{S}^Z$ on the set

$$B := \{z \in S^Z \mid (\exists i \in Z) \;\; z(i) \notin \{a, c, u\}\}.$$

Recalling Notation 4.9, let $B_0 := \{u_i^v \mid i \in Z\}$, and define $\varphi\colon \mathbb{N} \to \mathbb{N}$ by $\varphi(n) := n$.

Let $n \in \mathbb{N}$, and let $\theta$ be a congruence on $\mathbf{B}$ with $|B/\theta| \leqslant n$. We will show that $\theta{\restriction}_{B_0}$ has a unique block of size greater than $\varphi(n) = n$. Let $J$ and $K$ be disjoint subsets of $Z$ such that $|J| = |K| = n + 1$ and the sets $\{u_i^v \mid i \in J\}$, $\{u_i^v \mid i \in K\}$ are each contained in a block of $\theta{\restriction}_{B_0}$. We will show that $u_j^v \;\theta\; u_k^v$ for some $j \in J$ and some $k \in K$.

Write $J = \{j_0, \dots, j_n\}$ and $K = \{k_0, \dots, k_n\}$. Then the set $\{u_{j_s k_s}^{pp} \mid s \in \{0, \dots, n\}\}$ has size $n+1$, so it must contain a pair of distinct $\theta$-related elements, since $|B/\theta| \leqslant n$. In other words, there exist distinct $i, j \in J$ and distinct $k, l \in K$ such that $u_{ik}^{pp} \;\theta\; u_{jl}^{pp}$. Now, we have

$$u_{ik}^{pp} = a_{ik}^{dc} u_{ik}^{pp} \;\theta\; a_{ik}^{dc} u_{jl}^{pp} = u_i^v,$$

and by symmetry we have $u_{ik}^{pp} \;\theta\; u_k^v$, so $u_i^v \;\theta\; u_k^v$. We have shown that Theorem 4.8(i) holds. Finally, the element $g \in S^Z$ defined in Theorem 4.8(ii) is the constant tuple $\underline{u} \notin B$. $\square$

The first aperiodic semigroup considered by the author in the context of the dualisability problem was $\mathbf{L} \times \mathbf{P}$; this was suggested by the first supervisor as the simplest unresolved example of a semigroup that ought to be non-dualisable. As we now know, any attempt to show that $\mathbf{L} \times \mathbf{P}$ is non-dualisable was doomed from the start, so the author was eventually forced to move on to other examples. It was not until after the author obtained Theorem 5.12 that we found that $\mathbf{L} \times \mathbf{P}$ is dualisable.

After this discovery, it looked to be the case that all finite semigroups in $\mathbb{V}(\mathbf{L}, \mathbf{P})$ (and the larger variety $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$) are dualisable, mainly due to the absence of any obvious candidates for non-dualisable members. This would have been a remarkable result if true, since dualisable algebras generating residually large varieties are quite rare ($\mathbf{L} \times \mathbf{P}$ was only the second example discovered, after Ross Willard's example in [**16**, §4]).

Theorem 6.11, quite poetically, was the last (non-)dualisability result that the author needed for the classification theorem, and it finally restored some sanity to the situation. Residual smallness and dualisability are not so drastically different after all; the trigger for non-dualisability in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ simply manifests at a slightly higher level than the trigger for residual largeness. Certainly, the connection between these properties is still far from being understood, but now, perhaps, there is more of a reason for us to try to understand it.

## 6.3. The dualisable quasicriticals

After discovering Theorem 6.11, the finish line was now visible. It ought to be the case that the only quasicriticals in $\mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$ satisfying $xw \approx yw \to xzw \approx yzw$ are the members of the set

$$\mathcal{C} = \{\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{R}^0,\mathbf{I},\mathbf{N},\mathbf{P},\mathbf{C},\mathbf{C}^0,\mathbf{L}\circlearrowright L,\mathbf{E},\mathbf{E}^0\}.$$

Proving this would complete the classification of dualisable finite aperiodic semigroups. After months of unsuccessful attempts to describe the quasicriticals in $\mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$, the quasiequation $xw \approx yw \to xzw \approx yzw$ released all of the built-up knowledge into a proof of the following result.

**Theorem 6.12.** *Let $\mathbf{S} \in \mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$ be quasicritical. Then $\mathbf{S} \models xw \approx yw \to xzw \approx yzw$ if and only if $\mathbf{S} \in \mathbb{I}(\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{R}^0,\mathbf{I},\mathbf{N},\mathbf{P},\mathbf{C},\mathbf{C}^0,\mathbf{L}\circlearrowright L,\mathbf{E},\mathbf{E}^0)$.*

The goal of this section is to prove Theorem 6.12. We remark that Theorem 5.11 follows from Theorem 6.12, as $xw \approx yw \to xzw \approx yzw$ holds in $\mathcal{P}_1$. The proof of Theorem 6.12 presented here greatly simplifies the proof of Theorem 5.11 published in [**42**].

First, we will identify and dispense with the case where $\mathbf{S}$ is a band or a null semigroup.

**Lemma 6.13.** *Let $\mathbf{S} \in \mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$ be quasicritical, and let $(u,v)$ be a critical pair of $\mathbf{S}$. Then the following hold:*

(i) *if $u^2 \neq v^2$, then $\mathbf{S}$ is isomorphic to a member of $\{\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{R}^0,\mathbf{I}\}$;*
(ii) *if one of $u,v$ does not have a left identity in $\mathbf{S}$, then $\mathbf{S} \cong \mathbf{N}$.*

Proof. (i): By Proposition 6.1, the square map $x \mapsto x^2$ is a retraction of $\mathbf{S}$ onto $\mathrm{E}(\mathbf{S})$. If it separates $(u,v)$, then $\mathbf{S} \cong \mathrm{E}(\mathbf{S})$ is a band and the result follows from Theorem 4.25.

(ii): We may assume by symmetry that $u$ does not have a left identity. Then $u$ is not idempotent, so by Proposition 6.1, $\{u,u^2\}$ is a subsemigroup of $\mathbf{S}$ isomorphic to $\mathbf{N}$. Now, note that if $u \in SS$, so $u = xy$ for some $x,y \in S$, then $x$ is a left identity for $u$, which is a contradiction. Thus, $u \notin SS$. It follows that $J := S\backslash\{u\}$ is an ideal of $\mathbf{S}$ and $\mathbf{S}/J \cong \mathbf{N}$. The quotient map $\mathbf{S} \to \mathbf{S}/J$ evidently separates $u$ and $v$, and therefore $\mathbf{S} \cong \mathbf{N}$. $\square$

For the remainder of this section, we fix a quasicritical $\mathbf{S} \in \mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$, not isomorphic to a member of $\{\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{R}^0,\mathbf{I},\mathbf{N}\}$, such that $\mathbf{S} \models xw \approx yw \to xzw \approx yzw$. The aim is to show that $\mathbf{S}$ is isomorphic to a member of

$$\mathcal{P} = \{\mathbf{P},\mathbf{C},\mathbf{C}^0,\mathbf{L}\circlearrowright L,\mathbf{E},\mathbf{E}^0\}.$$

The proof will be split into two main cases, which will be split further into three sub-cases.

Let $(u,v)$ be a critical pair of $\mathbf{S}$. By Lemma 6.13, we have $u^2 = v^2$, and $u$ and $v$ both have left identities in $\mathbf{S}$. If $u$ and $v$ were both idempotent, we would evidently have $u^2 \neq v^2$, so we must have either $u^2 \neq u$ or $v^2 \neq v$.

The two main cases we will consider are $u \not\mathrel{\mathcal{L}} v$ and $u \mathrel{\mathcal{L}} v$. In the case that $u \not\mathrel{\mathcal{L}} v$, we will show that $\mathbf{S}$ is isomorphic to $\mathbf{P}$, $\mathbf{C}$, or $\mathbf{C}^0$. In this case, we have $u \mathrel{\mathcal{J}} v$ by Proposition 6.5, since $u$ and $v$ are not both idempotent. As the assumptions on $u$ and $v$ so far are symmetric, we will assume in the case $u \not\mathrel{\mathcal{L}} v$ that $u \not\preccurlyeq_J v$. It is then necessary that $u^2 \neq u$, for otherwise we would have $u = u^2 = v^2 \preccurlyeq_J v$.

In the case that $u \mathrel{\mathcal{L}} v$, we will show that $\mathbf{S}$ is isomorphic to $\mathbf{L}\circlearrowleft L$, $\mathbf{E}$, or $\mathbf{E}^0$. Since $u \mathrel{\mathcal{J}} v$, we must have $u^2 \neq u$ and $v^2 \neq v$ by Proposition 6.4.

To summarise, it will suffice by symmetry to consider the following two cases:

(L1) $u \not\preccurlyeq_J v$, which implies $u^2 \neq u$;

(L2) $u \mathrel{\mathcal{L}} v$, which implies $u^2 \neq u$ and $v^2 \neq v$.

In both cases, we have $u^2 = v^2$, and $u$, $v$ both have left identities in $\mathbf{S}$. In Case (L1), it will turn out that $L_u$ has one element, while in Case (L2), it will turn out that $L_u$ has two elements; hence the labels (L1), (L2).

A common feature of the semigroups in $\mathcal{P}$ is that they all have $\mathbf{P}$ as a homomorphic image. Our first step will be to show that our quasicritical $\mathbf{S}$ has $\mathbf{P}$ as a homomorphic image by means of a partition. First, we have the following lemma.

**Lemma 6.14.** *Let $w \in S\backslash \mathrm{E}(\mathbf{S})$, and let $a \in S$. Then $w \preccurlyeq_J a^2$ if and only if $aw \mathrel{\mathcal{L}} w$. Consequently, the set $\{x \in S \mid w \preccurlyeq_J x^2\}$ forms a subsemigroup of $\mathbf{S}$ if it is non-empty.*

Proof. If $aw \mathrel{\mathcal{L}} w$, then $w \mathrel{\mathcal{J}} aw = a^2 w \preccurlyeq_J a^2$. Conversely, assume that $w \preccurlyeq_J a^2$, so we have $w = xa^2y$ for some $x, y \in S^1$. If $y = 1$, so $w = xa^2$, then $w$ is idempotent, which is a contradiction. Thus, $y \neq 1$, which gives

$$w = xa^2y = xay = (xa)^2y = xaxay = xaw \in Saw,$$

and clearly $aw \in Sw$, so $aw \mathrel{\mathcal{L}} w$. For the last claim, it suffices to show that $\{x \in S \mid xw \mathrel{\mathcal{L}} w\}$ is closed under multiplication. Let $a, b \in S$ with $aw \mathrel{\mathcal{L}} bw \mathrel{\mathcal{L}} w$. Then by Proposition 6.8 we have $bw \mathrel{\mathcal{L}} w \Rightarrow abw = aw \mathrel{\mathcal{L}} w$. □

Now, define the following subsets of $S$:

$$\begin{aligned} T &:= \{x \in S \mid xu \mathrel{\mathcal{L}} u\} = \{x \in S \mid u \preccurlyeq_J x^2\}, \\ W &:= \{x \in S \mid u \preccurlyeq_J x \ \ \& \ \ u \not\preccurlyeq_J x^2\}, \\ I &:= S\backslash(T \cup W) = \{x \in S \mid u \not\preccurlyeq_J x\}. \end{aligned}$$

These sets are non-empty, since $T$ contains all left identities of $u$, while $u \in W$ and $u^2 \in I$. Since $u \preccurlyeq_J x^2 \Rightarrow u \preccurlyeq_J x$, it follows that $\{T, W, I\}$ is a partition of $S$.

**Lemma 6.15.** *The partition $\{T, W, I\}$ corresponds to a congruence $\pi$ on $\mathbf{S}$ with $\mathbf{S}/\pi \cong \mathbf{P}$. The left identity of $\mathbf{S}/\pi$ is $T$ and the zero element is $I$.*

| $\cdot$ | $T$ | $W$ | $I$ |
|---|---|---|---|
| $T$ | $T$ | $W$ | $I$ |
| $W$ | $I$ | $I$ | $I$ |
| $I$ | $I$ | $I$ | $I$ |

Proof. It is easily shown that $I$ is an ideal of $\mathbf{S}$, and by Lemma 6.14, $T$ is a subsemigroup of $\mathbf{S}$. To complete the proof, we must show that $TW \subseteq W$ and $WS \subseteq I$.

Let $x \in T$ and $y \in W$. Then $u \preccurlyeq_J y$, so by Proposition 6.6 we have $xu \preccurlyeq_J xy$. By definition of $T$, we have $u \mathrel{\mathcal{J}} xu$, so $u \preccurlyeq_J xy$, and hence $xy \in T \cup W$. But $(xy)^2 = x^2y^2 \preccurlyeq_J y^2$, so because $u \not\preccurlyeq_J y^2$, we have $u \not\preccurlyeq_J (xy)^2$. Hence, $xy \in W$, which shows that $TW \subseteq W$.

Finally, let $x \in W$ and $y \in S$. Then $xy = x^2y \preccurlyeq_J x^2$, so as $u \not\preccurlyeq_J x^2$, we have $u \not\preccurlyeq_J xy$. Hence, $xy \in I$, showing that $WS \subseteq I$. □

We will proceed by further refining the partition $\{T, W, I\}$ according to the various cases. Recall that in Case (L1), we aim to show that $\mathbf{S}$ is isomorphic to $\mathbf{P}$, $\mathbf{C}$, or $\mathbf{C}^0$. Since $u \in W$ and $v \in I$ in this case, we will immediately have $\mathbf{S} \cong \mathbf{P}$ if $\mathbf{P}$ embeds into $\mathbf{S}$. Otherwise, we will see that $\mathbf{C}$ embeds into $\mathbf{S}$, and so we must further refine $I$ into two or three blocks, depending on whether $\mathbf{C}^0$ embeds into $\mathbf{S}$. In Case (L2), we must split both $T$ and $W$ into two blocks, and then split the ideal $I$ depending on cases, analogously to Case (L1).

**Lemma 6.16.** *In Case* (L1)*, every left identity for $u$ is a left identity of $\mathbf{S}$, and $J_u = \{u\}$.*

PROOF. Let $a$ be a left identity for $u$. If $au = av$, then $u = au = av \preccurlyeq_J v$, contradicting (L1). Thus, $au \neq av$. Since left translation by $a$ is an endomorphism of $\mathbf{S}$ separating the critical pair $(u, v)$, it is an automorphism. As all left translations in $\mathbf{S}$ are idempotent maps, $a$ is a left identity of $\mathbf{S}$. Now, if $x, y \in S$ with $x \mathrel{\mathcal{L}} y$, then we have $x = ax = ay = y$ by Proposition 6.8, so the $\mathcal{L}$ relation on $\mathbf{S}$ is trivial. Proposition 6.5 gives $J_u = L_u = \{u\}$. $\square$

Define the following subsets of $T$ and $W$ (which will be interesting only in the (L2) case):

$$\begin{aligned} A &:= \{x \in S \mid xu = u\} \neq \varnothing, \\ B &:= T \backslash A = \{x \in S \mid xu \mathrel{\mathcal{L}} u \ \ \& \ \ xu \neq u\}, \\ U &:= AW, \\ V &:= W \backslash U. \end{aligned}$$

Clearly, if $B$ is non-empty, then $\{A, B\}$ partitions $T$. Since $u \in W$ and $u$ has a left identity, we have $u \in U$; in particular, $U \neq \varnothing$. Since $A \subseteq T$ and $TW \subseteq W$, it follows that $U \subseteq W$, so if $V \neq \varnothing$, then $\{U, V\}$ is a partition of $W$.

In Case (L1), we have $T = A$ since $L_u$ is trivial by Lemma 6.16, and by the same result, $A$ is the set of left identities of $\mathbf{S}$, so $U = AW = W$. Thus, $B = V = \varnothing$ in this case. In Case (L2), we will find that $B$ and $V$ are both non-empty.

**Lemma 6.17.** *Let $w \in W$. Then $u = ew$ for some $e \in A$.*

PROOF. By definition of $W$, we have $u \preccurlyeq_J w$, so $u = ewt$ for some $e, t \in S^1$. If $t \neq 1$, then we would have $u = ew^2t \preccurlyeq_J w^2$, contradicting $w \in W$. Thus, $t = 1$, so $u = ew$. Since $u$ has a left identity in $\mathbf{S}$, we can assume that $e \neq 1$. Then $eu = eew = ew = u$, so $e \in A$. $\square$

**Lemma 6.18.** *In Case* (L2)*, we have that $v \in V$, all left identities of $v$ lie in $B$, and*

$$AT \subseteq A, \quad BT \subseteq B, \quad AW \subseteq U, \quad BW \subseteq V.$$

*In table form:*

| $\cdot$ | $A$ | $B$ | $U$ | $V$ |
|---|---|---|---|---|
| $A$ | $A$ | $A$ | $U$ | $U$ |
| $B$ | $B$ | $B$ | $V$ | $V$ |

*Consequently, $\{A, B, U, V, I\}$ corresponds to a congruence $\lambda$ on $\mathbf{S}$ with $\mathbf{S}/\lambda \cong \mathbf{L} \circlearrowleft L$.*

PROOF. We have $u \preccurlyeq_J v$ and $u \not\preccurlyeq_J u^2 = v^2$, so $v \in W$. To show that $v \in V$, suppose to the contrary that $v \in U$. Then $v = aw$ for some $a \in A$ and some $w \in W$, so $a$ is a left identity for both $u$ and $v$. But since $u \mathrel{\mathcal{L}} v$, Proposition 6.8 gives $u = au = av = v$, which is a contradiction. Hence $v \in V$. Now let $b$ be a left identity for $v$. Then, by Proposition 6.8, we have $bu = bv = v \mathrel{\mathcal{L}} u$, so $b \in T$. But $bu = v \neq u$, so $b \in B$.

Now, $T$ is a subsemigroup and $\{A, B\}$ is a partition of $T$, so to show that $AT \subseteq A$ and $BT \subseteq B$, it suffices to show that $xy \in A \Leftrightarrow x \in A$ for all $x, y \in T$. Let $x, y \in T$. Then $yu \mathrel{\mathcal{L}} u$, so by Proposition 6.8 we have $xyu = xu$. Thus, $xyu = u \Leftrightarrow xy = u$, and this gives $xy \in A \Leftrightarrow x \in A$, as required.

Now, $AW \subseteq U$ is true by the definition of $U$, so it remains to show that $BW \subseteq V$. This will be the only point in this section where we use the quasiequation $xw \approx yw \to xzw \approx yzw$. Note that the statement $BW \subseteq V$ can be expressed as

$$(\forall c \in T)(\forall x \in W) \;\; c \in B \implies cx \in V.$$

We will show this in the contrapositive. Let $c \in T$ and $x \in W$ with $cx \in U$ to show $c \in A$. Since $cx \in U$, there are $d \in A$ and $y \in W$ with $cx = dy$. Let $w := cx = dy$, so $cw = dw = w$. Since $w \in W$, Lemma 6.17 gives $u = ew$ for some $e \in S$. Now, $cw = dw$ implies $cew = dew$ as $\mathbf{S} \models xw \approx yw \to xzw \approx yzw$, so $cu = cew = dew = du = u$. Hence, $c \in A$. □

**Lemma 6.19.** *Assume that* (L2) *holds, and let $e \in T$. Then there are $\mathcal{L}$-related idempotents $a, b \in T$ with $au = av = u$ and $bu = bv = v$ and $ua = ub = ue$.*

Proof. By definition of $T$, we have $eu \mathrel{\mathcal{L}} u$, so by Proposition 6.8, we have $ueu = u^2$, and hence $ue = u^2e = ueue = uee$; thus, $ue = ue^2$, so by replacing $e$ with $e^2$, we can assume that $e$ is idempotent.

Choose $c, d \in S$ with $cv = u$ and $du = v$. Since $\mathbf{S} \models x^2y \approx xy$, we can assume that $c$ and $d$ are idempotent. Let $a' := cd$ and $b' := dcd$. Then $a'$ and $b'$ are idempotent with

$$a'b' = a', \quad b'a' = b', \quad a'u = cdu = u, \quad b'u = dcdu = v.$$

Proposition 6.8 then gives $a'v = a'u = u$ and $b'v = b'u = v$.

Now, by Lemma 6.18, $a := a'e$ is a left identity for $u$ and $b := b'e$ is a left identity for $v$. Clearly $a$ and $b$ are idempotent and lie in $T$. We have $a' \mathrel{\mathcal{L}} b' \Rightarrow a \mathrel{\mathcal{L}} b$ by Proposition 6.8, and the same result gives $au = av$ and $bu = bv$. Finally, we have $ua = ua'e = ua'ue = u^2e = ue$, and $ub = ua$ by Proposition 6.8 again. □

**Lemma 6.20.** *Assume that $ue = u^2$ for some $e \in T$. Then $\mathbf{S}$ is isomorphic to $\mathbf{P}$ or $\mathbf{L} \circlearrowleft L$.*

Proof. In Case (L1), Lemma 6.16 implies that $T = A$ and $e$ is a left identity of $\mathbf{S}$. It is then easily checked that $\{e, u, u^2\}$ is a subsemigroup isomorphic to $\mathbf{P}$. The congruence $\pi$ from Lemma 6.15 separates the critical pair $(u, v)$, so it follows that $\mathbf{S} \cong \mathbf{P}$.

In Case (L2), use Lemma 6.19 to choose $\mathcal{L}$-related idempotents $a, b \in T$ with $au = av = u$ and $bu = bv = v$ and $ua = ub = ue = u^2$. Then $\{a, b, u, v, u^2\}$ is isomorphic to $\mathbf{L} \circlearrowleft L$. The congruence $\lambda$ from Lemma 6.18 separates $u$ and $v$, so $\mathbf{S} \cong \mathbf{L} \circlearrowleft L$. □

**Lemma 6.21.** *Let $m \in S$ be idempotent with $mu^2 = m$, and assume that $me \neq m$ for all $e \in T$. Let $\mathbf{S}_m$ denote the subsemigroup of $\mathbf{S}$ on $S_m := \{x \in S \mid m \preccurlyeq_J x^2\}$, and define*

$$Y := \{x \in S_m \mid u \not\preccurlyeq_J x \;\;\&\;\; mx = m\}, \quad X := \{x \in S_m \mid u \not\preccurlyeq_J x \;\;\&\;\; mx \neq m\}.$$

*Then $\{T, W, X, Y\}$ is a partition of $S_m$, and the corresponding equivalence relation $\psi$ is a congruence on $\mathbf{S}_m$ with $\mathbf{S}_m/\psi \cong \mathbf{C}$.*

| $\cdot$ | $T$ | $X$ | $W$ | $Y$ |
|---|---|---|---|---|
| $T$ | $T$ | $X$ | $W$ | $Y$ |
| $X$ | $X$ | $X$ | $Y$ | $Y$ |
| $W$ | $X$ | $X$ | $Y$ | $Y$ |
| $Y$ | $X$ | $X$ | $Y$ | $Y$ |

PROOF. That $S_m$ is a subsemigroup follows from the compatibility of $\preccurlyeq_J$. As $m \preccurlyeq_J u^2$, we have $T \cup W \subseteq S_m$, and also $u \not\preccurlyeq_J m$, so $m \in Y \neq \varnothing$. To see that $X \neq \varnothing$, choose any $e \in T$. Then $me \in S_m$ because $S_m$ is a subsemigroup. Now, we have $me \preccurlyeq_J m \preccurlyeq_J u^2$, so $u \not\preccurlyeq_J me$, and by assumption we have $me \neq m$, so $me \in X$. Thus, both $X$ and $Y$ are non-empty. We also have $X \cup Y = I \cap S_m$, so $X \cup Y$ is an ideal of $\mathbf{S}_m$, and $\{T, W, X, Y\}$ is indeed a partition of $S_m$. Given Lemma 6.15, it remains to show that

$$(\forall x, y \in S_m)\ \ xy \in W \cup Y \iff y \in W \cup Y.$$

We will first prove the following claim.

$$(\forall y \in S_m)\ \ y \in W \cup Y \iff my = m. \tag{$\heartsuit$}$$

Let $y \in S_m$. First, assume that $y \in W$. Then $u = ey$ for some $e \in A$ by Lemma 6.17. Thus, $u^2 = uey = ueuy = u^2y$, showing that $u^2$ is a left zero for $y$. Now $m$ is a left zero for $u^2$ by assumption, so $my = mu^2y = mu^2 = m$. Thus, $y \in W$ implies $my = m$. If $y \in Y$, then $my = m$ by definition of $Y$. Thus, $y \in W \cup Y$ implies $my = m$. For the converse, assume that $y \in S_m \backslash (W \cup Y) = T \cup X$. If $y \in T$, then by assumption we have $my \neq m$, and if $y \in X$, then $my \neq m$ by definition of $X$. This proves ($\heartsuit$).

Now, let $x, y \in S_m$. Then $m \preccurlyeq_J x^2$, so we have $mxm = mx^2m = m$ by Proposition 6.7. Thus, $mxy = mxmy = my$, so by ($\heartsuit$) we have $xy \in W \cup Y \Leftrightarrow y \in W \cup Y$. □

Let $M$ denote the minimum ideal of $\mathbf{S}$, which exists because $\mathbf{S}$ is a finite semigroup (Theorem 2.4). By Theorems 2.40 and 2.43 and the aperiodicity of $\mathbf{S}$, the subsemigroup on $M$ is a rectangular band.

**Lemma 6.22.** *Assume that $ue \neq u^2$ for all $e \in T$, and that*

$$(\forall m \in M)(\forall e \in T)\ \ me \neq m \text{ or } mu^2 \neq m.$$

*Then $\mathbf{S}$ is isomorphic to $\mathbf{C}$ or $\mathbf{E}$.*

PROOF. In Case (L1), let $e$ be any left identity for $u$. Then $e$ is a left identity of $\mathbf{S}$ by Lemma 6.16. Now, $e$, $u$, and $u^2$ lie in $T$, $W$, and $I$, respectively, while $ue = u^2e \Rightarrow ue \in I$, and $ue \neq u^2$ by assumption. Thus, $|\{e, u, ue, u^2\}| = 4$, and it is now easily checked that the subsemigroup $\{e, u, ue, u^2\}$ is isomorphic to $\mathbf{C}$.

In Case (L2), use Lemma 6.19 to choose $\mathcal{L}$-related idempotents $a, b \in T$ with $au = av = u$ and $bu = bv = v$ (which is possible as there is at least one $e \in T$). Then $\{a, b, u, v, ua, u^2\}$ is isomorphic to $\mathbf{E}$.

Now consider simultaneously the cases (L1) and (L2). Let $m$ be any element of $Mu^2$, so $mu^2 = m$. Then $me \neq m$ for all $e \in T$ by assumption. Define the following subsets of $I$:

$$Y := \{x \in I \mid mx = m\}, \quad X := \{x \in I \mid mx \neq m\}.$$

Note that $m \preccurlyeq_J x^2$ for all $x \in S$, so by Lemma 6.21, we have that $\{T, W, X, Y\}$ partitions $S$ and that the quotient of $\mathbf{S}$ by this partition is isomorphic to $\mathbf{C}$.

In Case (L1), we have $u \in W$ and $v \in I$, so there is a homomorphism from $\mathbf{S}$ onto $\mathbf{C}$ separating $(u, v)$. Since $\mathbf{C}$ embeds into $\mathbf{S}$, we have $\mathbf{S} \cong \mathbf{C}$.

In Case (L2), we have $u \in U$ and $v \in V$, and Lemma 6.18 implies that the refined partition $\{A, B, U, V, X, Y\}$ gives a quotient isomorphic to $\mathbf{E}$, so $\mathbf{S} \cong \mathbf{E}$. $\square$

The final piece of the proof of Theorem 6.12 is the following.

**Lemma 6.23.** *Assume that $ue \neq u^2$ for all $e \in T$, and that*

$$(\exists m \in M)(\exists e \in T)\ \ me = m \text{ and } mu^2 = m.$$

*Then $\mathbf{S}$ is isomorphic to $\mathbf{C}^0$ or $\mathbf{E}^0$.*

Proof. Fix $m \in M$ and $e \in T$ such that $me = m$ and $mu^2 = m$. Define $z := um \in M$. Then $z$ is idempotent as $M$ is a (rectangular) band. Clearly $z$ is a left zero for $ue$ and $u^2$, since $m$ is, and since $zu \in M$ is idempotent, we have $zu = (zu)^2 = zu^2 = z$. This shows that $z$ is a left zero for $\{u, ue, u^2\}$. We will now split into cases to show that $\mathbf{S}$ has a subsemigroup isomorphic to $\mathbf{C}^0$ or $\mathbf{E}^0$.

First, consider Case (L1). Then $e$ is a left identity for $u$. As in Lemma 6.22, $e$ is a left identity of $\mathbf{S}$ and $\{e, u, ue, u^2\}$ is isomorphic to $\mathbf{C}$. We have $ze = z$ as $me = m$, so $z$ is a left zero for $\{e, u, ue, u^2\}$. Since $u^2$ is a right zero for $\{e, u, ue, u^2\}$, so is $z = u^2 m$. That is, $z$ is a zero for $\{e, u, ue, u^2\}$, so $\{e, u, ue, u^2, z\}$ is isomorphic to $\mathbf{C}^0$.

Now consider Case (L2). By Lemma 6.19, we may choose $\mathcal{L}$-related idempotents $a, b \in T$ with $au = av = u$ and $bu = bv = v$ and $ua = ub = ue$. Then $\{a, b, u, v, ua, u^2\}$ is isomorphic to $\mathbf{E}$. Since $ua = ue$, we know that $z$ is a left zero for $\{u, ua, u^2\}$, and by Proposition 6.8, we have $zv = zu = z$. Now, $zu = z$ and $zua = z$ imply $za = zua = z$, and so $zb = za = z$ by Proposition 6.8. Thus, $z$ is a left zero for $\{a, b, u, v, ua, u^2\}$. Since $u^2$ is a right zero for $\{a, b, u, v, ua, u^2\}$, so is $z = u^2 m$. Hence, $\{a, b, u, v, ua, u^2, z\}$ is isomorphic to $\mathbf{E}^0$.

Now consider simultaneously the cases (L1) and (L2). Define the following subsets of $I$:

$$Y := \{x \in I \mid u^2 \preccurlyeq_J x^2 \text{ and } u^2 x = u^2\}, \quad X := \{x \in I \mid u^2 \preccurlyeq_J x^2 \text{ and } u^2 x \neq u^2\}.$$

By Lemma 6.21, $\{T, W, X, Y\}$ partitions the subsemigroup $\mathbf{S}_{u^2}$ on $S_{u^2} = \{x \in S \mid u^2 \preccurlyeq_J x^2\}$, and the quotient of $\mathbf{S}_{u^2}$ by this partition is isomorphic to $\mathbf{C}$.

Let $Z := I \backslash (X \cup Y) = S \backslash (T \cup W \cup X \cup Y)$, so $Z = \{x \in S \mid u^2 \not\preccurlyeq_J x^2\}$. We claim that $z \in Z$, from which it will follows that $Z$ is non-empty and is therefore an ideal of $\mathbf{S}$. Suppose that $u^2 \preccurlyeq_J z^2 = z$. Then $u^2 = u^2 z u^2$ by Proposition 6.7, so $u^2 = u^2 z u^2 = z$ since $z$ is a zero for $u^2$. But then $u^2 = z$ is a left zero for $ue$, implying that $ue = u^2 ue = u^2$, which is a contradiction. Hence, $u^2 \not\preccurlyeq_J z^2$, so $Z$ is an ideal of $\mathbf{S}$. It is now clear that the quotient of $\mathbf{S}$ by the partition $\{T, W, X, Y, Z\}$ is isomorphic to $\mathbf{C}^0$.

In Case (L1), we have $u \in W$ and $v \in I$, and so $\mathbf{S} \cong \mathbf{C}^0$. In Case (L2), we have $u \in U$ and $v \in V$, and Lemma 6.18 implies that the refined partition $\{A, B, U, V, X, Y, Z\}$ gives a quotient isomorphic to $\mathbf{E}^0$, so $\mathbf{S} \cong \mathbf{E}^0$. $\square$

Theorem 6.12 now follows from Lemmas 6.20, 6.22, 6.23 and Corollary 6.10.

### 6.4. The classification theorem and its corollaries

We have effectively shown via Theorems 5.29, 6.11, and 6.12 that the dualisable members of $\mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$ are precisely those finite members that satisfy $xw \approx yw \to xzw \approx yzw$. This completes the classification of dualisable finite aperiodic semigroups. Before stating our main result formally, we will give a neater description of the dualisable class.

**Theorem 6.24.** *The class of semigroups in* $\mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$ *satisfying* $xw \approx yw \to xzw \approx yzw$ *is precisely* $\mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)$*. That is,*

$$\mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L) = [x^2y \approx xy,\ xyzt \approx xzyt,\ xw \approx yw \to xzw \approx yzw].$$

*The class of quasicriticals in* $\mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)$ *is* $\mathbb{I}(\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{R}^0,\mathbf{I},\mathbf{N},\mathbf{P},\mathbf{C},\mathbf{C}^0,\mathbf{L} \circlearrowleft L,\mathbf{E},\mathbf{E}^0)$.

Proof. Let $\mathfrak{Q} := [x^2y \approx xy,\ xyzt \approx xzyt,\ xw \approx yw \to xzw \approx yzw]$. Then $\mathfrak{Q} \subseteq \mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{P})$ by Proposition 5.15, so $\mathfrak{Q}$ is locally finite by Theorem 1.10. By Corollary 4.18 and Theorem 6.12, we therefore have $\mathfrak{Q} = \mathbb{SP}(\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{R}^0,\mathbf{I},\mathbf{N},\mathbf{P},\mathbf{C},\mathbf{C}^0,\mathbf{L} \circlearrowleft L,\mathbf{E},\mathbf{E}^0)$. Now, it is clear that $\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{I},\mathbf{N},\mathbf{P} \in \mathbb{S}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)$, and we have $\mathbf{C},\mathbf{C}^0,\mathbf{E},\mathbf{E}^0 \in \mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)$ by Proposition 5.17. Thus, $\mathfrak{Q} = \mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)$. □

**Theorem 6.25** (The Classification of Dualisable Aperiodic Semigroups)**.** *Let* $\mathbf{S}$ *be a finite aperiodic semigroup. The following are equivalent:*

(i) $\mathbf{S}$ *is dualisable;*
(ii) $\mathbf{S}$ *is dualisable via* (IC)*;*
(iii) $\mathbf{S}$ *is not inherently non-dualisable;*
(iv) $\mathbf{S} \in \mathbb{A}([x^2y \approx xy,\ xyzt \approx xzyt,\ xw \approx yw \to xzw \approx yzw])$*;*
(v) $\mathbf{S} \in \mathbb{ASP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)$*.*

*Moreover, conditions* (i)–(v) *are implied by*

(vi) $\mathbb{V}(\mathbf{S})$ *is residually small.*

Proof. The implications (ii) ⇒ (i) ⇒ (iii) are trivial, and the equivalence of (iv) and (v) is by Theorem 6.24. We will show that (iii) ⇒ (iv) ⇒ (ii).

(iii) ⇒ (iv): Assume that $\mathbf{S}$ is not IND. By Theorems 4.12, 4.15, 5.4, 5.6, and 5.14, we have $\mathbf{S} \in \mathbb{A}([x^2y \approx xy,\ xyzt \approx xzyt]) = \mathbb{AV}(\mathbf{L},\mathbf{R},\mathbf{P})$, so (iv) follows from Theorem 6.11.

(iv) ⇒ (ii): Let $\mathbf{S} \in [x^2y \approx xy,\ xyzt \approx xzyt,\ xw \approx yw \to xzw \approx yzw]$. Then, by Theorem 6.24 and Corollary 4.18, there is some $\mathfrak{U} \subseteq \{\mathbf{L},\mathbf{L}^0,\mathbf{R},\mathbf{R}^0,\mathbf{I},\mathbf{N},\mathbf{P},\mathbf{C},\mathbf{C}^0,\mathbf{L} \circlearrowleft L,\mathbf{E},\mathbf{E}^0\}$ with $\mathbb{SP}(\mathbf{S}) = \mathbb{SP}(\mathfrak{U})$. By Theorem 5.29 and the (IC) Independence of Generator Theorem 4.7, $\mathbf{S}$ is dualisable via (IC). (To obtain (i) from (iv) directly, use Theorem 4.6 instead of Theorem 4.7.)

Finally, (vi) ⇒ (iv) follows from the implication (i) ⇒ (iv) in Theorem 5.2. □

The following corollary of Theorems 6.25 and 5.2 gives us many examples of dualisable algebras generating residually large varieties (though there are only finitely many up to the equivalence $\mathbf{A} \sim \mathbf{B} \Leftrightarrow \mathbb{SP}(\mathbf{A}) = \mathbb{SP}(\mathbf{B})$).

**Corollary 6.26.** *Every finite semigroup in* $\mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)\backslash(\mathbb{V}(\mathbf{L},\mathbf{R},\mathbf{I},\mathbf{N}) \cup \mathbb{V}(\mathbf{R},\mathbf{P}))$ *is dualisable and generates a residually large variety.*

We will now discuss two properties in relation to Theorem 6.25. The first of these is finite relateness. As we argued at the end of Section 4.5, every finite algebra that is dualisable via (IC) is also finitely related. Certainly, then, every dualisable aperiodic semigroup is finitely related. However, we can say more than this. By [**13**, Theorem 2.11], the property of being finitely related is a variety property, in the following sense.

**Theorem 6.27.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be finite algebras with* $\mathbb{V}(\mathbf{A}) = \mathbb{V}(\mathbf{B})$. *Then* $\mathbf{A}$ *is finitely related if and only if* $\mathbf{B}$ *is finitely related.*

Using this fact, we can show that every finite semigroup in $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ is finitely related.

**Lemma 6.28.** *Let* $\mathcal{V}$ *be a subvariety of* $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$. *If* $\mathcal{V}$ *is residually large, then* $\mathcal{V}$ *is equal to either* $\mathbb{V}(\mathbf{L}, \mathbf{P})$ *or* $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$.

Proof. Assume that $\mathcal{V}$ is residually large and does not equal $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$. If $\mathcal{V} \models xy^2 \approx xy$, then $\mathcal{V} \subseteq \mathcal{N}_1$, which is residually small by Theorem 5.2; thus, by Lemma 3.45, we can assume that $\mathbf{P} \in \mathcal{V}$. Now, since $\mathcal{V} \neq \mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$, we must have either $\mathbf{L} \notin \mathcal{V}$ or $\mathbf{R} \notin \mathcal{V}$. If $\mathbf{L} \notin \mathcal{V}$, then $\mathcal{V} \subseteq \mathcal{P}_1$ by Lemma 3.66, which is residually small by Theorem 5.2, so we must have $\mathbf{L} \in \mathcal{V}$ and $\mathbf{R} \notin \mathcal{V}$.

Now, arguing dually to the proof of Lemma 3.66, we have $\mathcal{V} \models x^2y^2x^2 \approx x^2y^2$. Using this equation, along with $x^2y \approx xy$ and $xyzt \approx xzyt$, one easily deduces that $\mathcal{V} \models xyz^2 \approx xzy^2$. Since $\mathbb{V}(\mathbf{L}, \mathbf{P}) = [x^2y \approx xy,\ xyz^2 \approx xzy^2]$ (see Remark 3.6), we have $\mathcal{V} = \mathbb{V}(\mathbf{L}, \mathbf{P})$. ☐

By Lemma 6.28 and Theorem 6.25, every subvariety of $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ can be generated by a finite semigroup that is dualisable via (IC) (hence finitely related). By Theorem 6.27, we have the following.

**Corollary 6.29.** *Every finite semigroup in* $\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{P})$ *is finitely related.*

Next, we will present a corollary concerning the finite basis property for quasivarieties. We say that an algebra $\mathbf{A}$ is **finitely q-based** if $\mathbb{SP}(\mathbf{A})$ has a finite quasiequational basis.

**Lemma 6.30.** *Let* $\mathcal{Q}$ *be a locally finite quasivariety. If* $\mathcal{Q}$ *has a finite quasiequational basis and has only finitely many quasicriticals up to isomorphism, then every algebra in* $\mathcal{Q}$ *is finitely q-based.*

Proof. Let $\mathcal{S}$ be a set of representatives of the isomorphism classes of $\mathbf{qc}(\mathcal{Q})$ (so $\mathcal{S}$ is finite), and let $\mathbf{A} \in \mathcal{Q}$. By Corollary 4.18, we have $\mathbb{SP}(\mathbf{A}) = \mathbb{SP}(\mathcal{U})$, where $\mathcal{U} = \mathcal{S} \cap \mathbb{SP}(\mathbf{A})$. Now, for each $\mathbf{T} \in \mathcal{S} \backslash \mathcal{U}$, there is a quasiequation $\varepsilon_{\mathbf{T}}$ that holds in $\mathcal{U}$ but not $\mathbf{T}$. Thus, if $\Sigma$ is a finite quasiequational basis for $\mathcal{Q}$, then $\Sigma \cup \{\varepsilon_{\mathbf{T}} \mid \mathbf{T} \in \mathcal{S} \backslash \mathcal{U}\}$ is a finite quasiequational basis for $\mathbb{SP}(\mathcal{U}) = \mathbb{SP}(\mathbf{A})$. ☐

By Theorem 6.24 and Lemma 6.30, every $\mathbf{S} \in \mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L)$ is finitely q-based. Thus, by Theorem 6.25 and Corollary 6.26, we have the following results.

**Corollary 6.31.** *Let* $\mathbf{S}$ *be a dualisable finite aperiodic semigroup. Then* $\mathbf{S}$ *is finitely q-based.*

**Corollary 6.32.** *Every finite semigroup in* $\mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L) \backslash (\mathbb{V}(\mathbf{L}, \mathbf{R}, \mathbf{I}, \mathbf{N}) \cup \mathbb{V}(\mathbf{R}, \mathbf{P}))$ *is finitely q-based, dualisable, and generates a residually large variety.*

To the author's (limited) knowledge, there have been no published examples of finite semigroups that are finitely q-based and generate a residually large variety. Corollary 6.32 provides us with many such examples. While we obtained these in passing from our dualisability theorem, the connection between being finitely q-based and generating a residually large variety ostensibly has nothing to do with dualisability. Corollary 6.31 now raises the question of whether being finitely q-based is equivalent to being dualisable, at least for finite aperiodic semigroups. We will discuss this further the epilogue (Section 7.5).

## 6.5. A lattice

To close the main part of the thesis, we leave the reader with the following picture.

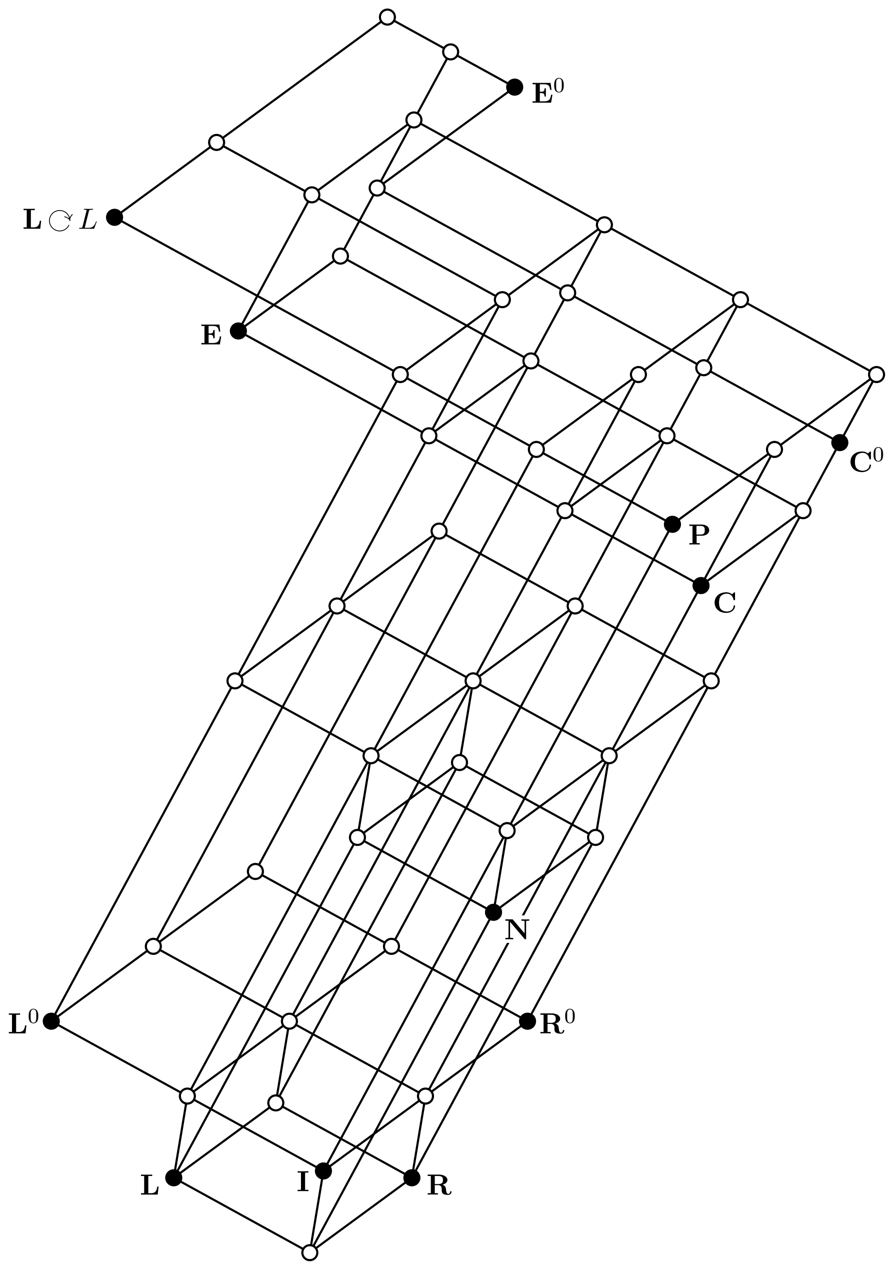



FIGURE 6.2. The lattice $\mathrm{L_q}(\mathbb{SP}(\mathbf{R}^0, \mathbf{L} \circlearrowleft L))$, constructed from a computer output and verified by hand. The join irreducible subquasivarieties are indicated by filled circles and are labelled by their quasicritical generators.

# Epilogue

CHAPTER 7

# The Future

While Theorem 6.25 brings one story to a close, it is only a small part of the solution to the dualisability problem for finite semigroups. We will discuss here a number of possible directions for future work on this broader dualisability problem.

## 7.1. Completely regular semigroups and Clifford semigroups

It is known that any finite completely regular semigroup generating a residually large variety is IND (by [**57**] and Theorems 3.43, 4.13, and 4.15). We conjecture that a finite completely regular semigroup generating a residually small variety is dualisable.

An important subclass for which the problem remains unresolved is the class of Clifford semigroups. A completely regular semigroup $\mathbf{S}$ is called a **Clifford semigroup** if the idempotents of $\mathbf{S}$ commute (so $\mathrm{E}(\mathbf{S})$ is a semilattice). A basic property of Clifford semigroups is that, further to Theorem 2.44, all $\mathscr{J}$-classes are groups [**37**, §4.2].

Beyond the characterisation of dualisable finite groups, there has been relatively little progress on the dualisability problem for finite Clifford semigroups. AlDhamri's work in [**1**, Chapter 6] covers all of the known results to date. Based on our conjecture for completely regular semigroups, we expect that every finite Clifford semigroup whose $\mathscr{J}$-classes are all A-groups is dualisable. For a proof of this conjecture, it is not too difficult to show that a proof of the following restricted version will suffice.

**Problem 7.1.** Let $\mathbf{S}$ be a finite Clifford semigroup with precisely two $\mathscr{J}$-classes, each of which is an A-group. Show that $\mathbf{S}$ is dualisable.

The best starting point for Problem 7.1 would be to show that $\mathbf{G}^1$ is dualisable, where $\mathbf{G}$ is the six-element non-Abelian group. For this, it may be useful to study the duality for $\mathbf{G}$ in [**20**]. I suspect that resolving this case will go a long way towards solving Problem 7.1.

A positive solution to Problem 7.1 would have significant implications for the broader dualisability problem for completely regular semigroups. While editing Chapter 3, I obtained a proof that a positive solution would imply that every completely regular semigroup in a residually small variety is dualisable. Since the proof is not particularly short, I will save it until we have (most of) a solution to Problem 7.1.

Further to this, it is possible to modify the proof of Theorem 5.9 to show that if $\mathbf{A}$ is a finitely dualisable finite algebra with a non-trivial identity term, then every finite inflation of $\mathbf{A}$ is dualisable (this removes all references to (IC)). Consequently, a positive solution to Problem 7.1 would imply that every finite $\mathbf{S}$ satisfying Theorem 3.43(i)–(iv) is dualisable.

My work on this thesis began with Problem 7.1, but I made very little progress. It is somewhat amusing that I was able to bring this full circle at the end of my candidature by showing that a positive solution to Problem 7.1 would solve a much broader problem.

## 7.2. Four-element semigroups

The new results in Chapter 5 complete the classification of dualisable three-element semigroups and cover many of the unresolved cases for the four-element semigroups. Figure 7.1 shows all of the four-element semigroups (up to anti-isomorphism) whose dualisability is not yet known (given also the unpublished results mentioned in Section 7.1).

| $\cdot$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 2 | 3 |
| 3 | 0 | 1 | 3 | 2 |

| $\cdot$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 |
| 1 | 1 | 0 | 1 | 1 |
| 2 | 0 | 1 | 0 | 0 |
| 3 | 0 | 1 | 2 | 3 |

| $\cdot$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 0 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 | 1 |
| 2 | 1 | 0 | 0 | 1 |
| 3 | 0 | 1 | 2 | 3 |

FIGURE 7.1. From left to right: the semigroups $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$.

Let $\mathbf{Z}_2$ denote the two-element cyclic group. The semigroups $\mathbf{S}_1$ and $\mathbf{S}_2$ generate the same quasivarieties as $\mathbf{Z}_2^0 \times \mathbf{P}$ and $\mathbf{Z}_2 \times \mathbf{P}$, respectively, while $\mathbf{S}_3$ lies in $\mathbb{SP}(\mathbf{Z}_2, \mathbf{P})$ but generates a strictly smaller quasivariety. The semigroup $\mathbf{S}_3$ is quasicritical (all proper subsemigroups of $\mathbf{S}_3$ are commutative, but $\mathbf{S}_3$ is not). These three semigroups generate residually small varieties, so they are most likely dualisable.

**Problem 7.2.** Show that $\mathbf{S}_1$, $\mathbf{S}_2$, and $\mathbf{S}_3$ are dualisable.

The best case to start with would be $\mathbf{S}_1$, as this will also contribute to the problem given in the next section.

## 7.3. Residually small implies not IND

We now know that, for finite semigroups, generating a residually small variety is not equivalent to being dualisable. However, it may well be the case that a finite semigroup generating a residually small variety is dualisable. Proving this will most likely be a difficult endeavour. For the time being, a more manageable problem would be to show that a finite semigroup generating a residually small variety embeds into a dualisable semigroup.

Based on Theorems 3.37, 3.68, and 3.71, it will suffice to show that every semigroup of the form $\mathbf{N} \times \mathbf{L}^0 \times \mathbf{R}^0 \times \mathbf{G}^0$, where $\mathbf{G}$ is an A-group, or $\mathbf{R}^0 \times \mathbf{P} \times \mathbf{G}$, where $\mathbf{G}$ is a finite cyclic group, is dualisable. The former case is dealt with by the modified version of Theorem 5.9 mentioned in Section 7.1 along with [**41**, Theorem 1.1], [**1**, Corollary 4.2.7], and [**14**, Theorem 3.6]. For the latter case, it will suffice to prove the following.

**Problem 7.3.** Let $\mathbf{G}$ be a non-trivial finite cyclic group, and let $\mathbf{G} \times \mathbf{R}$ act on $G$ via left translation by the first coordinate. Show that $\mathbf{G} \times \mathbf{R} \circlearrowright G$ is dualisable.

A positive solution to Problem 7.3 would imply that every finite semigroup generating a residually small variety embeds into a dualisable finite semigroup and is therefore not IND.

It would be best to start with a simpler version of the problem: show that $\mathbf{G} \circlearrowright \{u\}$ is dualisable whenever $\mathbf{G}$ is a non-trivial cyclic group (acting totally on $\{u\}$). The semigroup $\mathbf{S}_1$ from Section 7.2 is the smallest example of such a semigroup.

### 7.4. Generalisations

I am not sure what the generalisation of Theorem 6.25 to the class of all finite semigroups ought to be, but it is worth discussing some possibilities. By Theorem 5.13, the dualisability problem for finite semigroups is reduced to determining the dualisable finite members of the variety $\mathfrak{D}_n := [x^{n+1}y \approx xy,\ xy^n z^n t \approx xz^n y^n t]$ for each $n \geqslant 1$.

In the $n = 1$ case, we found that dualisability is equivalent to satisfying the quasiequation $xw \approx yw \to xzw \approx yzw$. There is likely to be an analogous quasiequation for $n > 1$, but at this point it is difficult to say what that may be. It could turn out that dualisability within $\mathfrak{D}_n$ cannot be captured by a single quasiequation, or even by a set of quasiequations. For the time being, it would be best to focus on specific examples until the picture becomes clearer.

Since $\mathbf{G} \times \mathbf{P}$ generates a residually large variety whenever $\mathbf{G}$ is non-Abelian, there is a good chance that such semigroups are non-dualisable, though it would not be too surprising if there were dualisable semigroups of this kind. In any case, the smallest example to consider here would be the following.

**Problem 7.4.** Let $\mathbf{G}$ be the six-element non-Abelian group, and let $\mathbf{G}$ act totally on $\{u\}$. Decide whether $\mathbf{G} \circlearrowright \{u\}$ is dualisable.

The semigroup $\mathbf{S}_1$ from Section 7.2 will probably need to be dealt with before tackling this problem, and solving Problem 7.3 would likely help as well.

Depending on the solution to Problem 7.3, it may be that a finite $\mathbf{S}$ is non-dualisable whenever $\mathbb{V}(\mathbf{S})$ contains $\mathbf{P}$ and a non-Abelian group. However, since $\mathfrak{D}_n \supseteq \mathfrak{G}_n$, we will not be able to eliminate non-Abelian groups altogether. If there is indeed a single quasiequation $\varepsilon_n$ capturing dualisability in $\mathfrak{D}_n$, and if a dualisable $\mathbf{S}$ cannot have both $\mathbf{P}$ and a non-Abelian group in its variety, then $\varepsilon_n$ would somehow have to force all subgroups of $\mathbf{S}$ to be Abelian only if $\mathbb{V}(\mathbf{S})$ contains $\mathbf{P}$. It probably goes without saying that the general dualisability problem for $\mathfrak{D}_n$ is likely to be vastly more complicated than the $n = 1$ case.

### 7.5. The finite q-basis problem

We briefly discussed the finite q-basis property in Section 6.4. Given Corollary 6.31, it is natural to ask whether the converse holds. To my knowledge, there are no published results suggesting otherwise, though Mark Sapir privately communicated to the first supervisor some 18 years ago that $\mathbf{P} \times \mathbf{Q}$ is finitely q-based (it seems the proof is unpublished). If true, this would imply that the converse of Corollary 6.31 is false (as $\mathbf{P} \times \mathbf{Q}$ is non-dualisable by Theorem 4.12 and Lemma 3.9). Though this is mostly wishful thinking on my part, it may turn out that $\mathbf{P} \times \mathbf{Q}$ is not finitely q-based.

**Problem 7.5.** Decide whether $\mathbf{P} \times \mathbf{Q}$ is finitely q-based.

This is not a dualisability problem per se, but it would be interesting to know if the converse to Corollary 6.31 holds. If no counterexamples arise, there are many directions from which one might approach a proof of the converse. For the reader wishing to learn more on the finite q-basis problem for finite semigroups, Margolis, Saliola, and Steinberg give a thorough survey of the known results in [**46**].

### 7.6. Concluding remarks

As much as I have said, I nonetheless feel that a lot was left unsaid. There are many results that did not make the final cut, as they either did not contribute to the main results or were superseded. In the end, what I really wanted was to tell a story, and I have come to believe that in order to craft a good story, it is necessary to delete most of it. (If only that were sufficient.) Being as blessed as I was to obtain my classification result, I have to conclude that the most natural way to approach the thesis was to tell my story (in Part 2) and set up the beginning of someone else's story (via Part 1 and this chapter).

Even now, I think writing Part 1 at the level of depth that I did might seem questionable, but I never wavered from the decision even during the hardest parts of the writing process. While I maintain that any solution to the dualisability problem for semigroups will require most of the knowledge recorded in Part 1, there were other driving factors. For whatever reason, I have always had the urge to re-tell an old story and make it as good as it can be. Arguably, I could have invested this energy into my own results, but they are still fresh, and they make up only one piece of a larger puzzle. There is no doubt that the proofs in Chapter 5 are far from optimal; my feeling is that I have missed many fundamental insights that will be revealed when my results are inevitably generalised. If there ever comes a day when the story of Part 2 can be re-told, I hope that my work in Part 1 inspires the writer to put a piece of their soul into it.

Composer John Williams once said the following: "I spend more time on those little bits of musical grammar, to get them just right, so that they seem inevitable—seem like they've always been there". It is a wonderful approach to creating anything, and perhaps applies in a special way to mathematics, because in some sense it has always been there. This little quote guided me every day that I spent writing this thesis; all things considered, I think there was no other way to have written it.

# Bibliography


[1] AlDhamri, N.: Natural Dualities for Quasi-varieties of Semigroups. PhD Thesis, La Trobe University, Melbourne (2013)

[2] AlDhamri, N.: Dualities for quasi-varieties of bands. Semigroup Forum **88**(2), 417–432 (2014)

[3] Baldwin, J.T., Berman, J.: The number of subdirectly irreducible algebras in a variety. Algebra Universalis **5**, 379–389 (1975)

[4] Bergman, C.: Universal Algebra. Fundamentals and Selected Topics. Pure and Applied Mathematics (Boca Raton) **301**, CRC Press, Boca Raton, Florida (2012)

[5] Clark, D.M., Davey, B.A.: Natural Dualities for the Working Algebraist. Cambridge Studies in Advanced Mathematics **57**, Cambridge University Press, Cambridge (1998)

[6] Clark, D.M., Davey, B.A., Pitkethly, J.G.: Binary homomorphisms and natural dualities. J. Pure Appl. Algebra **169**, 1–28 (2002)

[7] Clark, D.M., Idziak, P.M., Sabourin, L.R., Szabó, C., Willard, R.: Natural dualities for quasivarieties generated by a finite commutative ring. Algebra Universalis **46**, 285–320 (2001)

[8] Clifford, A.H.: The structure of orthodox unions of groups. Semigroup Forum **3**(4), 283–337 (1972)

[9] Clifford, A.H., Preston, G.B.: The Algebraic Theory of Semigroups, Vol. I. Mathematical Surveys, No. 7, American Mathematical Society, Providence, Rhode Island (1961)

[10] Clifford, A.H., Preston, G.B.: The Algebraic Theory of Semigroups, Vol. II. Mathematical Surveys, No. 7, American Mathematical Society, Providence, Rhode Island (1967)

[11] Davey, B.A., Haviar, M., Willard, R.: Structural entailment. Algebra Universalis **54**, 397–416 (2005)

[12] Davey, B.A., Idziak, P.M., Lampe, W.A., McNulty, G.F.: Dualizability and graph algebras. Discrete mathematics **214**(1), 145–172 (2000)

[13] Davey, B.A., Jackson, M.G., Pitkethly, J.G., Szabó, C.: Finite degree: algebras in general and semigroups in particular. Semigroup Forum **83**(1), 89–110 (2011)

[14] Davey, B.A., Knox, B.J.: Regularising natural dualities. Acta Math. Univ. Comenianae **68**(2), 295–318 (1999)

[15] Davey, B.A., Knox, B.J.: From rectangular bands to $k$-primal algebras. Semigroup Forum **64**(1), 29–54 (2002)

[16] Davey, B.A., Pitkethly, J.G., Willard, R.: Dualisability versus residual character: a theorem and a counterexample. J. Pure Appl. Algebra **210**(2), 423–435 (2007)

[17] Davey, B.A., Priestley, H.A.: Optimal natural dualities. Trans. Amer. Math. Soc. **338**(2), 655-677 (1993)

[18] Davey, B.A., Priestley, H.A.: Optimal natural dualities. II: General theory. Trans. Amer. Math. Soc. **348**(9), 3673–3711 (1996)

[19] Davey, B.A., Priestley, H.A.: Introduction to Lattices and Order. Second edition. Cambridge University Press, New York (2002)

[20] Davey, B.A., Quackenbush, R.W.: Natural dualities for dihedral varieties. J. Austral. Math. Soc. (Ser. A) **61**, 216–228 (1996).

[21] Davey, B.A., Werner, H.: Dualities and equivalences for varieties of algebras. Contributions to lattice theory (Szeged, 1980), 101–275, Colloq. Math. Soc. János Bolyai 33, North-Holland, Amsterdam (1983)

[22] Davey, B.A., Willard, R.: The dualisability of a quasi-variety is independent of the generating algebra. Algebra Universalis **45**, 103–106 (2001)

[23] Dixon, M.R.: Sylow theory, formations and fitting classes in locally finite groups. Series in Algebra **2**, World Scientific Publishing Co., Inc., River Edge, New Jersey (1994)

[24] Dolinka, I.: Finite bands are finitely related. Semigroup Forum **97**(1), 115–130 (2018)

[25] Dugundji, J: Topology. Allyn and Bacon, Inc., Boston, Massachusetts (1966)

[26] Evans, T.: The lattice of semigroup varieties. Semigroup Forum **2**(1), 1–43 (1971)

[27] Freese, R., McKenzie, R.: Residually small varieties with modular congruence lattices. Trans. Amer. Math. Soc. **264**(2), 419–430 (1981)

[28] Fuchs, L.: Infinite Abelian Groups, Vol. I. Pure and Applied Mathematics **36**, Academic Press, New York-London (1970)

[29] Gerhard, J.A.: The lattice of equational classes of idempotent semigroups. J. Algebra **15**, 195–224 (1970)

[30] Gerhard, J.A.: Subdirectly irreducible idempotent semigroups. Pacific J. Math. **39**, 669–676 (1971)

[31] Gerhard, J.A., Shafaat, A.: Semivarieties of idempotent semigroups. Proc. London Math. Soc. (3) **22**, 667–680 (1971)

[32] Golubov, È.A., Sapir, M.V.: Varieties of finitely approximable semigroups. Soviet Math. (Iz. VUZ) **26**(11), 25–36 (1982) [Translated from Russian]

[33] Gorbunov, V.A.: Algebraic Theory of Quasivarieties. Siberian School of Algebra and Logic, Consultants Bureau, New York (1998) [Translated from Russian]

[34] Grätzer, G.: Universal Algebra. Second edition. Springer-Verlag, New York-Heidelberg (1979)

[35] Green, J.A.: On the structure of semigroups. Annals of Math. (2) **54**, 163–172 (1951)

[36] Gruenberg, K.W.: Residual properties of infinite soluble groups. Proc. London Math. Soc. (3) **7**, 29–62 (1957)

[37] Howie, J.M.: Fundamentals of Semigroup Theory. London Mathematical Society Monographs. New Series **12**, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York (1995)

[38] Hyndman, J., Nation, J.B.: The Lattice of Subquasivarieties of a Locally Finite Quasivariety. CMS Books in Mathematics, Springer International Publishing AG (2018)

[39] Jackson, M.G.: Dualisability of finite semigroups. Int. J. Algebra Comput. **13**(4), 481–497 (2003)

[40] Jackson, M.G.: Natural dualities, nilpotence and projective planes. Algebra Universalis **74**, 65–85 (2015)

[41] Kearnes, K.A., Szendrei, Á.: Dualizable algebras with parallelogram terms. Algebra Universalis **76**, 497–539 (2016)

[42] Koussas, T.: Semigroup quasivarieties: two lattices and a reopened problem. To appear in Int. J. Algebra Comput.

[43] Lesohin, M.M. [Lesokhin, M.M.], Golubov, È.A.: Finitary approximability of commutative semigroups. Ural. Gos. Univ. Mat. Zap. **5**(3), 82–90 (1966) [Russian]

[44] Mac Lane, S.: Categories for the working mathematician. Second edition. Graduate Texts in Mathematics **5**, Springer-Verlag, New York (1998)

[45] Macdonald, I.D.: The Theory of Groups. Clarendon Press, Oxford (1968)

[46] Margolis, S., Saliola, F., Steinberg, B.: Semigroups embeddable in hyperplane face monoids. Semigroup Forum **89**, 236–248 (2014)

[47] McKenzie, R.: Residually small varieties of semigroups. Algebra Universalis **13**, 171–201 (1981)

[48] McKenzie, R.: A note on residually small varieties of semigroups. Algebra Universalis **17**, 143–149 (1983)

[49] McKenzie, R.: The residual bounds of finite algebras. Int. J. Algebra Comput. **6**, 1–28 (1996)

[50] Ol'šanskiĭ, A.Ju.: Varieties of finitely approximable groups. Math. USSR Izv. **3**(4), 867–877 (1969) [Translated from Russian]

[51] Ol'šanskiĭ, A.Ju.: Varieties in which all finite groups are Abelian. Math. USSR Sb. **54**(1), 57–80 (1986) [Translated from Russian]

[52] Petrich, M.: All subvarieties of a certain variety of semigroups. Semigroup Forum **7**, 104–152 (1974)

[53] Pitkethly, J.G.: Uncountably many dualisable algebras. Int. J. Algebra Comput. **21**(5), 825–839 (2011)

[54] Pitkethly, J.G., Davey, B.A.: Dualisability. Unary Algebras and Beyond. Advances in Mathematics **9**, Springer, New York (2005)

[55] Porst, H.-E., Tholen, W.: Concrete dualities. Category theory at work (Bremen, 1990), Res. Exp. Math. **18**, 111–136 (1991)

[56] Quackenbush, R.W.: Equational classes generated by finite algebras. Algebra Universalis **1**, 265–266 (1971)

[57] Quackenbush, R.W., Szabó, C.: Nilpotent groups are not dualizable. J. Austral. Math. Soc. **72**, 173–179 (2002)

[58] Rasin, V.V.: On the lattice of varieties of completely simple semigroups. Semigroup Forum **17**(2), 113–122 (1979)

[59] Rees, D.: On semi-groups. Proc. Cambridge Philos. Soc. **36**, 387–400 (1940)

[60] Sapir, M.V.: Varieties with a finite number of subquasivarieties. Siberian Math. J. **22**(6), 934–949 (1981) [Translated from Russian]

[61] Sapir, M.V., Shevrin, L.N.: Residually small varieties of semigroups and groups. Soviet Math. (Iz. VUZ) **32**(10), 57–73 (1988) [Translated from Russian]

[62] Sapir, M.V., Sukhanov, E.V.: On manifolds of periodic semigroups. Soviet Math. (Iz. VUZ) **25**(4), 53–63 (1981) [Translated from Russian; note that 'manifolds' are varieties]

[63] Saramago, M.J.: Some remarks on dualisability and endodualisability. Algebra Universalis **43**(2), 197–212 (2000)

[64] Suschkewitsch, A.: Über die endlichen Gruppen ohne das Gesetz der eindeutigen Umkehrbarkeit. Math. Ann. **99**(1), 30–50 (1928) [German]

[65] Taylor, W.: Residually small varieties. Algebra Universalis **2**(1), 33–53 (1972)

[66] Taylor, W.: Subdirectly irreducible algebras in regular, permutable varieties. Proc. Amer. Math. Soc. **75**, 196-200 (1979)

[67] Willard, R.: New tools for proving dualizability. Dualities, Interpretability and Ordered Structures (Lisbon, 1997), eds. J. Vaz de Carvalho and I. Ferreirim, Centro do Álgebra da Universidade de Lisboa 69–74 (1999)

[68] Zádori, L.: Natural duality via a finite set of relations. Bull. Austral. Math. Soc. **51**(3), 469–478 (1995)

[69] Zamyatin, A.P.: A non-Abelian variety of groups has an undecidable elementary theory. Algebra and Logic **17**(1), 13–17 (1978)

# Notation

# Index