

RESEARCH ARTICLE

Small Inherently Nonfinitely Based Finite Semigroups

Marcel Jackson

Communicated by M. V. Volkov

Abstract

We show how to construct all “forbidden divisors” for the pseudovariety of *not* inherently nonfinitely based finite semigroups. Several other results concerning finite semigroups that generate an inherently nonfinitely based variety that is minimal amongst those generated by finite semigroups are obtained along the way. For example, aside from the variety generated by the well known six element Brandt monoid \mathbf{B}_2^1 , a variety of this type is necessarily generated by a semigroup with at least 56 elements (all such semigroups with 56 elements are described by the main result).

1. Introduction

One of the most investigated problems in the study of finite algebras is determining when a finite set of identities satisfied by a given algebra implies all identities of that algebra. For arbitrary algebras, this problem (known as *Tarski’s finite basis problem*) has been shown by McKenzie [10] to be undecidable however for some natural classes (such as the varieties of groups [12], commutative semigroups [13], rings [7], [8], and lattices [9]) there are no finite algebras without a finite basis for their identities. An algebra \mathbf{A} (or a variety \mathcal{V}) is said to be *inherently nonfinitely based* (abbreviated to INFB) if it is locally finite, not finitely based (that is, has no finite basis for its identities) and has the property that every locally finite variety \mathcal{V}' for which $\mathbf{A} \in \mathcal{V}'$ (or $\mathcal{V} \subseteq \mathcal{V}'$ respectively) is also not finitely based. This concept has been useful in many investigations into Tarski’s finite basis problem (the earliest being [11] and [14]; the term “inherently nonfinitely based” is introduced in the second of these), and McKenzie’s negative solution to Tarski’s finite basis problem in fact shows that the class of finite INFB algebras is not recursively enumerable and that the class of finitely based algebras is not recursive. This makes the situation for semigroups of particular interest since there are many known nonfinitely based finite semigroups and there also exists an aesthetic algorithmic description of those which are INFB. Let the words Z_1, Z_2, \dots be defined by $Z_1 \equiv x_1$, $Z_{n+1} \equiv Z_n x_{n+1} Z_n$ (“ \equiv ” denotes graphic equality of words). The words Z_1, Z_2, \dots are known as *Zimin words* and were introduced in [24].

Theorem 1.1. (M. Sapir [17]) *A finite semigroup is not INFB if and only if for some n it satisfies a nontrivial identity of the form $Z_n \approx W$.*

This gives a description of the finite INFB semigroups in terms of identities. A powerful structural description is given by the following (recall that the *upper hypercentre* $\Gamma(\mathbf{G})$ of a group \mathbf{G} is the last term in the upper central series of that group and that an element x of a semigroup \mathbf{S} *divides* an element y if there are elements a and b in \mathbf{S}^1 so that $axb = y$).

Theorem 1.2. (M. Sapir [18]) *(i) If \mathbf{S} is a finite INFB semigroup then for some idempotent $e \in \mathbf{S}$, $e\mathbf{S}e$ is an INFB submonoid of \mathbf{S} with identity element e .*

(ii) If \mathbf{S} is a finite monoid with period d then \mathbf{S} is INFB if and only if for some element $a \in \mathbf{S}$ dividing an idempotent $e \in \mathbf{S}$ the elements eae and $ea^{d+1}e$ do not lie in the same coset of the maximal subgroup \mathbf{S}_e of \mathbf{S} containing e with respect to the upper hypercentre $\Gamma(\mathbf{S}_e)$.

The convention of writing \mathbf{S}_e for the maximal subgroup containing a given idempotent e of a semigroup \mathbf{S} will be used throughout.

It will frequently be necessary to consider pairs of the form (a, e) where a and e are elements of a semigroup \mathbf{S} , $e^2 = e$ and a divides e in \mathbf{S} . Such a pair will be called a *dividing pair* for \mathbf{S} and we will say *INFB occurs at (a, e)* if this pair satisfies the conditions of Theorem 1.2 (ii).

One interesting corollary of Theorem 1.1 is that the class of finite, not INFB semigroups (known as *weakly finitely based*, or WFB semigroups) is closed under the formation of finite direct products, homomorphic images and subsemigroups and therefore forms a *pseudovariety*.

Consider the following monoid \mathbf{B}_2^1 of matrices with the standard multiplication:

$$\left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

This semigroup was the first finite semigroup to be shown to have no finite basis for its identities [13] and is one of the smallest semigroups with this property since Trahtman [22], [23] has shown that the identities of every semigroup with fewer than six elements are finitely based (see [21] for a discussion of this). As is shown in [17], it also follows very easily from Theorem 1.2 that \mathbf{B}_2^1 is INFB (INFB occurs at (a, e) if we take a to be the second last matrix and e to be the last). Indeed it is shown in [18] that this semigroup generates an INFB variety that is minimal amongst all INFB varieties generated by finite semigroups. Some other results emphasising the importance of this example are: a regular semigroup \mathbf{S} is INFB if and only if $\mathbf{B}_2^1 \in \mathbf{V}(\mathbf{S})$ [5] and a semigroup \mathbf{T} containing only nilpotent subgroups is INFB if and only if $\mathbf{B}_2^1 \in \mathbf{V}(\mathbf{T})$ [18].

Definition 1.3. Let \mathbf{M} be a finite INFB semigroup from a variety \mathcal{V} . If there is a subclass C of \mathcal{V} containing \mathbf{M} such that for every semigroup $\mathbf{S} \in C$, \mathbf{S} is INFB if and only if $\mathbf{M} \in \mathcal{V}(\mathbf{S})$ then \mathbf{M} (or $\mathcal{V}(\mathbf{M})$) will be said to be a *minimal finite INFB semigroup* in \mathcal{V} and C (or *minimal finitely generated INFB variety* in \mathcal{V} respectively) and $\mathcal{V}(\mathbf{M})$ will be denoted the *minimum finitely generated INFB variety* for C .

In the terminology of this definition, \mathbf{B}_2^1 generates a minimal finitely generated INFB variety in the variety of all semigroups and generates a variety that is the minimum INFB variety for the class of regular semigroups.

By constructing a finite INFB monoid \mathbf{S} for which $\mathbf{B}_2^1 \notin \mathbf{V}(\mathbf{S})$ it is shown in [18] that $\mathbf{V}(\mathbf{B}_2^1)$ is not the only minimal finitely generated INFB variety. We will extend these results by providing a construction for all “forbidden divisors” for the pseudovariety of WFB finite semigroups and by proving the following theorem.

Theorem 1.4. *If \mathbf{S} is a finite semigroup from one of the following classes then \mathbf{S} is INFB if and only if $\mathbf{B}_2^1 \in \mathbf{V}(\mathbf{S})$:*

- (i) *the class of (not necessarily regular) semigroups whose idempotents form a subsemigroup;*
- (ii) *the class of all semigroups with at most eight non-nilpotent subgroups;*
- (iii) *the class of all semigroups with at most fifty five elements.*

As well as the abbreviations INFB and WFB (introduced above) we will also use the abbreviations: FB; NFB; and WNFB to denote semigroups (or locally finite varieties) which are: finitely based; not finitely based; and not finitely based but not inherently not finitely based (*weakly not finitely based*) respectively. Each of these classes is known to be nonempty (see [21] for examples). A further property of a semigroup is that of being hereditarily finitely based (abbreviated to HFB): a semigroup or variety of semigroups has this property if it is FB and every subvariety of the variety it generates is also FB.

2. Classes for which \mathbf{B}_2^1 generates the minimum INFB variety

We first recall an extract of a result that is central to the arguments used in [18]. Here \mathbf{A}_2^1 is the monoid given by matrix multiplication on the following set of matrices:

$$\left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right\}.$$

Lemma 2.1. (See [18] or [20].) *Let \mathbf{S} be a finite monoid. If there is no homomorphic image of a submonoid of \mathbf{S} isomorphic to \mathbf{B}_2^1 or \mathbf{A}_2^1 then for every idempotent $e \in \mathbf{S}$ and every element a dividing e in \mathbf{S} the element eae*

belongs to \mathbf{S}_e . Furthermore if for every idempotent $e \in \mathbf{S}$ and every element a dividing e in \mathbf{S} the element eae belongs to \mathbf{S}_e then for every idempotent $f \in \mathbf{S}$ and any element b dividing f in \mathbf{S} , the element b^2 divides f in \mathbf{S} .

As is easily verified, the monoid \mathbf{B}_2^1 is isomorphic to a homomorphic image of a subsemigroup (that is, a *divisor*) of the direct product $\mathbf{A}_2^1 \times \mathbf{A}_2^1$ and so $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{A}_2^1)$. For this reason, \mathbf{B}_2^1 will feature more frequently than \mathbf{A}_2^1 in the arguments to follow. Note that it follows from the definition of being inherently nonfinitely based that \mathbf{A}_2^1 is also an INFB semigroup.

The following result from [5], which we include here for completeness, follows by combining Lemma 2.1 with Theorems 1.1 and 1.2.

Theorem 2.2. [5] *If \mathbf{S} is a finite regular semigroup with period d then the following are equivalent:*

- (i) \mathbf{S} is INFB;
- (ii) $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{S})$;
- (iii) the three element monoid \mathcal{N}^1 consisting of a null semigroup with adjoined identity element is contained in $\mathcal{V}(\mathbf{S})$;
- (iv) $\mathbf{S} \not\approx xyx \approx (xy)^{d+1}x$.

This shows that the WFB regular semigroups have a quite restricted nature, a situation emphasised by the following.

Corollary 2.3. *A finite monoid \mathbf{S} is embeddable in a finitely based finite regular semigroup only if \mathbf{S} is regular.*

Proof. The statement follows immediately because a finite monoid containing a non group element generates a variety containing \mathcal{N}^1 . ■

It would be interesting if the reverse implication also held true for WFB monoids and to describe (if possible) the class of finite semigroups that are subsemigroups of FB regular semigroups.

Any semigroup satisfying $xyx \approx (xy)^{d+1}x$ satisfies $x^3 \approx x^{2d+3}$, so Theorem 2.2 implies that no semigroup with index greater than three can be embedded in a finitely based finite regular semigroup. In fact we can reduce these bounds further.

Proposition 2.4. *If \mathbf{S} is a semigroup with index greater than two, then \mathbf{S} is not embeddable into a finite finitely based regular semigroup.*

Proof. Assume that \mathbf{S} is embedded in a finite regular semigroup \mathbf{R} of period d . There is an element $a \in \mathbf{S} \subseteq \mathbf{R}$ so that $a^2 \neq a^{2+i}$ for any $i > 0$. Since \mathbf{R} is regular there is an idempotent e so that $ea = a$. The element eae cannot lie in a subgroup of \mathbf{R} since then $a^2 = (ea)(ea) = (eae)a = (eae)^{d+1}a = a^{d+2}$. Therefore the monoid $e\mathbf{R}e$ is INFB since $\mathcal{N}^1 \in \mathcal{V}(e\mathbf{R}e)$ and $e\mathbf{R}e$ is regular (see [3] for example). ■

Note that there are many WFB and even FB regular semigroups with index equal to two (the semigroup $\mathbf{B}_2 = \mathbf{B}_2^1 \setminus \{1\}$ was shown to have this property by A. N. Trahtman; see [21]).

An orthodox semigroup is a regular semigroup whose idempotents form a subsemigroup. Orthodox semigroups are a well known generalisation of inverse semigroups. A second result obtained in [5] which relates to INFB semigroups is the following (this theorem follows immediately from Theorem 2.2 and results established by Rasin in [15]).

Corollary 2.5. [5] *A finite orthodox monoid is FB if and only if it is HFB if and only if it is not INFB. A finite orthodox semigroup \mathbf{S} is INFB if and only if $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{S})$.*

There are a large number of of WNFB finite semigroups whose idempotents form a subsemigroup but are not regular (see [6] and [19] for example). Therefore if the condition of regularity is removed from the definition of an orthodox semigroup the first statement of Corollary 2.5 no longer holds. The second statement however does continue to hold.

Theorem 2.6. *If the idempotents of a finite semigroup \mathbf{S} form a subsemigroup of \mathbf{S} then \mathbf{S} is INFB if and only if $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{S})$.*

Proof. If the idempotents of a semigroup \mathbf{S} form a subsemigroup then for every idempotent e , the idempotents of the submonoid $e\mathbf{S}e$ also form a subsemigroup of $e\mathbf{S}e$. Therefore by Theorem 1.2 we need only consider the case when \mathbf{S} is a monoid.

If $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{S})$ then \mathbf{S} is INFB by the definition of being inherently nonfinitely based. Assume that $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$ and \mathbf{S} has period d . Since $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{A}_2^1)$, by Lemma 2.1 for every dividing pair (a, e) , $ea^i e \in \mathbf{S}_e$. Now for any $i > 0$, a^i divides a^{2^i} and by Lemma 2.1, a^{2^i} divides e . Therefore $ea^i e \in \mathbf{S}_e$ for all $i > 0$. We now use induction to show that $ea^i e = (eae)^i$. From this it follows that $ea^{d+1} e = (eae)^{d+1} = eae$, and by Theorem 1.2, \mathbf{S} is not INFB.

Firstly $(eae)^{-1}a$ and $a(eae)^{-1}$ are both idempotent since, for example,

$$(eae)^{-1}a(eae)^{-1}a = (eae)^{-1}eae(eae)^{-1}a = (eae)^{-1}a.$$

Therefore $(eae)^{-1}aa(eae)^{-1} = (eae)^{-1}ea^2e(eae)^{-1}$ is idempotent and since

$$(eae)^{-1}ea^2e(eae)^{-1} \in \mathbf{S}_e,$$

$(eae)^{-1}ea^2e(eae)^{-1} = e$. Therefore $ea^2e = (eae)^2$.

Now assume that $ea^k e = (eae)^k$. Since $(eae)^{-1}a$ and $a^k(ea^k e)^{-1}$ are idempotent, so is the element $(eae)^{-1}aa^k(ea^k e)^{-1}$. Therefore

$$(eae)^{-1}aa^k(ea^k e)^{-1} = (eae)^{-1}ea^{k+1}e(ea^k e)^{-1} = (eae)^{-1}ea^{k+1}e(eae)^{-k} = e.$$

Therefore $ea^{k+1}e = (eae)^{k+1}$ as required. In particular $ea^{d+1}e = (eae)^{d+1} = eae$ since the exponent of \mathbf{S}_e divides the period, d , of \mathbf{S} . ■

By a well known result from [1] the class of all finite semigroups whose idempotents form a subsemigroup is exactly the psuedovariety generated by the class of finite orthodox semigroups.

Theorem 1.2 also provides a way of increasing the power of Theorem 2.6.

Definition 2.7. If P is a property of semigroups then a semigroup \mathbf{S} has the property P locally or \mathbf{S} is locally- P , if for every idempotent e , $e\mathbf{S}e$ has the property P .

Corollary 2.8. *If P is a property such that the finite semigroups with P are INFB if and only if \mathbf{B}_2^1 is contained in the variety they generate then a finite locally- P semigroup is INFB if and only if \mathbf{B}_2^1 is contained in the variety it generates.*

Proof. Let \mathbf{S} be a finite locally- P semigroup. Then by Theorem 1.2, \mathbf{S} is INFB if and only if $e\mathbf{S}e$ is INFB for some idempotent $e \in \mathbf{S}$. The semigroup $e\mathbf{S}e$ has the property P and therefore is INFB if and only if $\mathbf{B}_2^1 \in \mathcal{V}(e\mathbf{S}e)$. Since $e\mathbf{S}e$ is a subsemigroup of \mathbf{S} the result follows. ■

It follows for example that \mathbf{B}_2^1 generates the minimum finitely generated INFB variety in the class of locally regular semigroups.

3. The number of elements in a INFB semigroup

As noted above, every semigroup of order less than six is necessarily FB so an INFB semigroup must have at least six elements. While \mathbf{B}_2^1 and \mathbf{A}_2^1 are INFB semigroups with exactly six elements, the example presented in [18] of a finite INFB semigroup whose variety does not contain \mathbf{B}_2^1 has quite a few elements (in fact it has at least 57 elements, the exact size depending on the choice of a centreless group). We now address the problem of finding the smallest possible size of an INFB semigroup \mathbf{S} with $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$.

It will be convenient to always assume that \mathbf{S} is a finite INFB monoid of period d with $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$ and that INFB occurs at the dividing pair (a, e) . As in the previous section \mathbf{S}_e will denote the largest subgroup of \mathbf{S} containing e and $ea^i e \in \mathbf{S}_e$ for every $i \geq 0$. A number of simple lemmas will lead to the desired lower bound; the last two conditions of Theorem 1.4 will then follow.

Lemma 3.1. *No subgroup of \mathbf{S} contains a .*

Proof. If a were in a subgroup of \mathbf{S} then $a = a^{d+1}$ and $eae = ea^{d+1}e$ contradicting the fact that INFB occurs at (a, e) . ■

This shows, for example, that no power of a can equal 1.

Lemma 3.2. *Let s and t be elements of \mathbf{S}_e and $i \geq 0$.*

(i) *The elements as , sa are not contained in \mathbf{S}_e ,*

(ii) *$sa^i = ta^i \Rightarrow s = t$,*

(iii) *$sa^i \neq at$ and $sa \neq a^i t$.*

Proof. (i) If $sa \in \mathbf{S}_e$ then $rs^{-1}sa = ra \in \mathbf{S}_e$ for any $r \in \mathbf{S}_e$. Say $ea = r$ for some $r \in \mathbf{S}_e$. Then $ea^{d+1}e = rea^d e = r^2ea^{d-1}e = \dots = r^d eae = eae$, contradicting the fact that INFB occurs at (a, e) . By symmetry we also have $as \notin \mathbf{S}_e$.

(ii) Say $sa^i = ta^i$. Then

$$\begin{aligned} s &= s(ea^i e)(ea^i e)^{-1} \\ &= (sa^i)e(ea^i e)^{-1} \\ &= (ta^i)e(ea^i e)^{-1} \\ &= t(ea^i e)(ea^i e)^{-1} \\ &= t. \end{aligned}$$

(iii) Say $sa^i = at$ for some $s, t \in \mathbf{S}_e$. So $eae = eatt^{-1} = sa^i t^{-1} = att^{-1} = ae$. But then $ae = eae \in \mathbf{S}_e$, contradicting (i). The case $sa = a^i t$ follows by symmetry. ■

Lemma 3.3. *Let s and t be arbitrary elements of \mathbf{S}_e . Then $sa^2 \neq ta$ and $a^2 s \neq at$.*

Proof. Say $sa^2 = ta$. Then

$$\begin{aligned} sa^{d+1} &= sa^2 a^{d-1} \\ &= ta a^{d-1} \\ &= ts^{-1} sa^d \\ &= ts^{-1} (sa^2) a^{d-2} \\ &= ts^{-1} ta^{d-1} \\ &= \dots \\ &= (ts^{-1})^{d-1} ta \\ &= (ts^{-1})^{-1} ta \\ &= st^{-1} ta = sa. \end{aligned}$$

Therefore $eae = s^{-1} sae = s^{-1} sa^{d+1} e = ea^{d+1} e$, contradicting the fact that INFB occurs at (a, e) . The case $a^2 s \neq at$ follows by symmetry. ■

Lemma 3.4. *For every $s \in \mathbf{S}_e$, $sa^2 \notin \mathbf{S}_e$ and $a^2 s \notin \mathbf{S}_e$.*

Proof. Assume $sa^2 \in \mathbf{S}_e$. So $s^{-1}sa^2 = ea^2 \in \mathbf{S}_e$ and therefore $ea^2 = ea^2e$. Let i and p be the index and period respectively of the subsemigroup $\langle a \rangle$ of \mathbf{S} generated by a .

Case 1. p is odd.

If i is odd then $a^{i+1} = a^{i+1+p}$ and $i+1$ is even. Let $2j$ be the even element of $\{i, i+1\}$ (that is, j is the integer part of $(i+1)/2$). So $ea^{2j} = ea^2a^{2j-2} = ea^2ea^{2j-2} = \dots = (ea^2)^j \in \mathbf{S}_e$. But since p is odd, $ea^{2j} = ea^{2j+p} = ea^{2j}(a^2)^{(p-1)/2}a = (ea^2)^j(a^2)^{(p-1)/2-1}a = \dots = (ea^2)^{j+(p-1)/2}a$. Now $ea^2 \in \mathbf{S}_e$ and by Lemma 3.2, $sa \notin \mathbf{S}_e$, so therefore $(ea^2)^{j+(p-1)/2}a \notin \mathbf{S}_e$, contradicting the fact that $(ea^2)^{j+(p-1)/2}a = ea^{2j} = (ea^2)^j \in \mathbf{S}_e$.

Case 2. p is even.

Since p divides d and p is even, $\frac{p}{2}$ divides $\frac{d}{2}$ and d is even. Therefore if for some $s \in \mathbf{S}_e$, we have $s^{p/2} = e$ then $s^{d/2} = e$. Now since $ea^2 \in \mathbf{S}_e$,

$$ea^{d+1}e = ea^2a^{d-1}e = ea^2eea^{d-1}e = \dots = (ea^2)^{d/2}eae \quad (\text{since } d \text{ is even})$$

We now show that $(ea^2)^{p/2} = e$ and therefore $ea^{d+1}e = eae$, giving the required contradiction.

Let $2j$ be the even element of $\{i, i+1\}$. So $ea^{2j} = ea^2a^{2j-2} = \dots = (ea^2)^j \in \mathbf{S}_e$. But

$$ea^{2j} = ea^{2j+p} = ea^2(a^2)^{j+p/2-1} = ea^2e(a^2)^{j+p/2-1} = \dots = (ea^2)^{j+p/2}$$

Therefore $(ea^2)^j = (ea^2)^{j+p/2} = (ea^2)^j(ea^2)^{p/2}$ and so $(ea^2)^{p/2} = e$ as required.

Therefore sa^2 is not contained in \mathbf{S}_e . That a^2s is not contained in \mathbf{S}_e follows by symmetry. \blacksquare

Lemma 3.5. For any elements $s, t \in \mathbf{S}_e$, $sa^2 \neq a^2t$.

Proof. If $sa^2 = a^2t$ then $sa^2e = a^2te = a^2t$. But by Lemma 3.4, $sa^2e \in \mathbf{S}_e$ and $a^2t \notin \mathbf{S}_e$, a contradiction. \blacksquare

Lemma 3.6. If $i, j \in \{1, 2\}$ and $s \in \mathbf{S}_e$ then $a^i sa^j \notin \mathbf{S}_e$.

Proof. Say $a^i sa^j \in \mathbf{S}_e$ and let $t = ea^i s \in \mathbf{S}_e$. Then $a^i sa^j = ea^i sa^j = ta^j$, a contradiction since ta^j is not an element of \mathbf{S}_e by Lemma 3.2 (i) and Lemma 3.4. \blacksquare

Lemma 3.7. If $i, j, k, l \in \{1, 2\}$ and $s, t \in \mathbf{S}_e$ then $a^i sa^j = a^k ta^l$ implies $s = t$, $i = k$, $j = l$.

Proof. Say $j \neq l$. Without loss of generality we may assume $j = 1$ and $l = 2$. Then $a^i sa = a^k ta^2$ and so $(ea^i s)a = (ea^k t)a^2$, contradicting Lemma 3.3. Therefore, by symmetry, $i = k$ and $j = l$. So $a^i sa^j = a^i ta^j$ and therefore $ea^i sa^j e = ea^i ta^j e$, giving $s = t$ as required. \blacksquare

Lemma 3.8. If $i, j, k \in \{1, 2\}$ and $s, t \in \mathbf{S}_e$ then $a^i sa^j \neq a^k t$ or ta^k .

Proof. If $a^i sa^j = a^k t$ then $ea^i sa^j = ea^k t$, contradicting Lemmas 3.2 (i) and 3.4. Likewise, $a^i sa^j \neq ta^k$, by symmetry. ■

Lemma 3.9. For any $s \in \mathbf{S}_e$, $a \notin \{s, sa, as, saa, aas, asa, aasa, asaa, aasaa, 1\}$.

Proof. Firstly $a \neq 1$ by Lemma 3.1. Secondly for any $i, j \in \{0, 1, 2\}$, $(a^i sa^j)^{d+1} = a^i (sa^{i+j} e)^d sa^j = a^i sa^j$. Since $ea^i sa^j \neq ea^{d+1} e$, the result follows. ■

Lemma 3.10. For any $i, j \in \{0, 1, 2\}$ and $s \in \mathbf{S}_e$, $1 \neq a^i sa^j$.

Proof. If $i > 0$, $1 \neq a^i sa^j$ since then $e = e1 = a^i sa^j e$, contradicting Lemmas 3.2 (i) and 3.4. By symmetry the only remaining case is when $i = j = 0$, that is, when $1 = s \in \mathbf{S}_e$. This is impossible since $a = a1 \neq as$ by Lemma 3.2 (i). ■

Combining Lemmas 3.2 through 3.10 we have the following.

Theorem 3.11. The sets $\{1\}$, $\{a\}$, $\{a^i sa^j : s \in \mathbf{S}_e, i, j \leq 2\}$ are disjoint in \mathbf{S} .

Corollary 3.12. If \mathbf{T} is a semigroup with $|\mathbf{T}| < 56$ then \mathbf{T} is INFB if and only if $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{T})$.

Proof. If \mathbf{S} is a finite INFB semigroup and $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$ then $ea^i sa^j \in \mathbf{S}_e$ for every dividing pair (a, e) . By Theorem 1.2, for one such dividing pair (a, e) , $ea^i sa^j$ and $ea^{d+1} e$ do not lie in the same coset of \mathbf{S}_e modulo $\Gamma(\mathbf{S}_e)$. Since both these elements are contained in \mathbf{S}_e , we must have $\Gamma(\mathbf{S}_e) \neq \mathbf{S}_e$. If \mathbf{G} is a group then by definition, $\Gamma_{\mathbf{G}} = \mathbf{G}$ exactly when \mathbf{G} is nilpotent. The smallest non nilpotent group \mathbf{G} is the six element centerless group \mathbf{S}_3 with upper hypercenter equal to $\{1\}$. By Theorem 3.11 there is a disjoint copy of \mathbf{S}_e for each pair $\{(i, j) : i, j \in \{0, 1, 2\}\}$ (that is, nine copies of \mathbf{S}_e) as well as an element 1 and the element a . This sets the minimum size for such a semigroup as $9 \times 6 + 1 + 1 = 56$. ■

As will be shown, there do exist quite a few INFB semigroups \mathbf{S} with 56 elements and with $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$, so this bound is the best possible (recall that the smallest example of this type that can be made using the construction of [18] has 57 elements and so is itself nearly a smallest possible example).

A second corollary of Theorem 3.11 is the following

Corollary 3.13. If \mathbf{S} is a semigroup with at most 8 disjoint non-nilpotent subgroups then \mathbf{S} is INFB if and only if $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{S})$.

All statements in Theorem 1.4 have now been proved.

4. Minimal INFB divisors for finite semigroups

We now describe two constructions for making finite INFB monoids generating varieties not containing \mathbf{B}_2^1 . These constructions will be based around finite centreless groups. The importance of centreless groups here lies in the fact that the upper hypercentre of a group \mathbf{G} is a normal subgroup $\Gamma(\mathbf{G})$ such that $\mathbf{G}/\Gamma(\mathbf{G})$ is centreless.

Throughout the remainder of this paper it will be convenient to consider (contrary to the usual convention) the ij^{th} entry of a matrix as the entry in the $(i+1)^{\text{th}}$ row and the $(j+1)^{\text{th}}$ column. For example the first entry in any matrix will be the 00^{th} entry and a Rees matrix semigroup (without 0 element) $\mathcal{M}(\mathbf{G}, m, n, P)$ over a group \mathbf{G} with $n \times m$ sandwich matrix P will be considered as a set of the form

$$\{(i, g, j): g \in \mathbf{G}, 0 \leq i \leq m-1, 0 \leq j \leq n-1\}$$

with multiplication $(i, g, j)(i', g', j') = (i, gP_{j,i'}g', j')$, where $P_{i,j}$ is the ij^{th} entry of the matrix P (according to the altered convention above). If a is a non group element of a monoid \mathbf{S} we will let a^0 denote the identity element 1 of \mathbf{S} .

Definition 4.1. Let \mathbf{G} be a finite centreless group with identity element e and exponent d . Let g and g_1 be (possibly identical) elements of the group \mathbf{G} . Construct a 3×3 matrix with group entries as follows: let g_i denote the element $(g_1g)^{i-1}g_1$; let h be any element of $\mathbf{G} \setminus \{g_2^{-1}g_1g_2^{-1}\}$; and for $i, j \leq 2$ define

$$P_{i,j} = \begin{cases} e, & \text{if } i = j = 0 \\ h, & \text{if } i + j = 1 \\ g_2^{-1}g_{i+j}g_2^{-1}, & \text{if } i + j \geq 2. \end{cases}$$

Then $\Xi_1[\mathbf{G}, g, g_1, h]$ consists of the set $\mathcal{M}(\mathbf{G}, 3, 3, P) \cup \{a, 1\}$ with multiplication $1x = x1 = x$ (for every x), $aa = (2, g_2, 2)$,

$$a(i, k, j) = \begin{cases} (i+1, k, j), & \text{if } i < 2 \\ (2, g_3g_2^{-1}k, j) = (2, g_1gk, j), & \text{if } i = 2 \end{cases}$$

and

$$(i, k, j)a = \begin{cases} (i, k, j+1), & \text{if } j < 2 \\ (i, kg_2^{-1}g_3, 2) = (2, kgg_1, j), & \text{if } j = 2. \end{cases}$$

Multiplication within $\mathcal{M}(\mathbf{G}, 3, 3, P)$ will be as usual.

Remark 4.2. In general,

$$\begin{aligned} g_2^{-1}g_i &= g_1^{-1}g^{-1}g_1^{-1}(g_1g)(g_1g)(g_1g)^{i-3}g_1 \\ &= g(g_1g)^{i-3}g_1 \end{aligned}$$

and likewise, $g_i g_2^{-1} = g_1(gg_1)^{i-3}g$. This means, in particular, that $a(2, g_i, 2) = (2, g_{i+1}, 2) = (2, g_i, 2)a$, and that $P_{2,2} = g$.

Lemma 4.3. *For any centreless group \mathbf{G} , the groupoid $\Xi_1[\mathbf{G}, g, g_1, h]$ as constructed in Definition 4.1 is an INFB semigroup with $\mathbf{B}_2^1 \notin \mathcal{V}(\Xi_1[\mathbf{G}, g, g_1, h])$.*

Proof. The routine though tedious task of showing that $\Xi_1[\mathbf{G}, g, g_1, h]$ is a semigroup (performed in [4]) will be omitted. To show $\Xi_1[\mathbf{G}, g, g_1, h]$ is INFB, first note that $\Xi_1[\mathbf{G}, g, g_1, h]$ is a monoid and that the element a divides the idempotent $(0, e, 0)$ because

$$(0, g_2^{-1}P_{0,2}^{-1}, 0)aa(2, P_{2,2}^{-1}, 0) = (0, g_2^{-1}P_{0,2}^{-1}, 0)(2, g_2, 2)(2, P_{2,2}^{-1}, 0) = (0, e, 0).$$

However the period of $\Xi_1[\mathbf{G}, g, g_1, h]$ is d (the exponent of \mathbf{G}) and $a^{d+1} = (2, g_1, 2)$ so

$$(0, e, 0)a(0, e, 0) = (0, h, 0) \neq (0, g_2^{-1}g_1g_2^{-1}, 0) = (0, e, 0)(2, g_1, 2)(0, e, 0).$$

as required by Theorem 1.2. That is, INFB occurs at $(a, (0, e, 0))$.

Now to show that \mathbf{B}_2^1 is not contained in the variety $\mathcal{V}(\Xi_1[\mathbf{G}, g, g_1, h])$. It is well known and easy to verify that a Rees matrix semigroup over a group of exponent d satisfies the identities $x \approx x^{d+1}$ and $(xyz)^d \approx (xz)^d$. Therefore $\mathcal{M}[\mathbf{G}, 3, 3, P]$ satisfies the identity $(xyx^2y)^d \approx (xy)^d$. Furthermore if we delete all occurrences of a given letter from this identity then the resulting identity is still satisfied by $\mathcal{M}[\mathbf{G}, 3, 3, P]$. Therefore the monoid obtained from $\mathcal{M}[\mathbf{G}, 3, 3, P]$ by adjoining an identity element satisfies $(xyx^2y)^d \approx (xy)^d$. So in order to show that $\Xi_1[\mathbf{G}, g, g_1, h]$ satisfies $(xyx^2y)^d \approx (xy)^d$ we need only check cases where the element a is assigned to at least one of the letters x and y . If a is assigned to both x and y or if a is assigned to just one of these and 1 is assigned to the other then both sides simply equal a^d . If a is assigned to x but (i, s, j) is assigned to y , then xy becomes (i', t, j) for some i' and some $t \in \mathbf{G}$. In this case, both sides of the identity become the idempotent in the subgroup $H_{i',j}$ of all elements of the form (i', r, j) , where $r \in \mathbf{G}$. The case when a is assigned to y and (i, s, j) is assigned to x is similar. Thus

$$\Xi_1[\mathbf{G}, g, g_1, h] \models (xyx^2y)^d \approx (xy)^d.$$

However $\mathbf{B}_2^1 \not\models (xyx^2y)^d \approx (xy)^d$ since \mathbf{B}_2^1 contains two elements, a and b , such that ab is a nonzero idempotent but $a^2 = 0$ (as a and b one can take the third last and last matrices respectively from the definition of \mathbf{B}_2^1 in the introduction). Since the left side of $(xyx^2y)^d \approx (xy)^d$ contains x^2 but the right side is of the form $(xy)^d$, assigning a to x and b to y ensures the left side becomes 0 but the right side becomes the nonzero idempotent ab . ■

We will say that $\Xi_1[\mathbf{G}, g, g_1, h]$ is a *small INFB finite semigroup of the first kind* and denote the set of all such monoids by Ξ_1 .

Remark 4.4. In [18] a finite INFB monoid \mathbf{T} is presented for any centreless group \mathbf{G} with a non identity element g with the property $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{T})$. By letting e be the identity element of \mathbf{G} it is possible to show that the monoid $\Xi_1[\mathbf{G}, g, g^{-1}, e]$ is a (proper) homomorphic image of \mathbf{T} .

Note also that if \mathbf{S}_3 is the six element centreless group then $\Xi_1[\mathbf{S}_3, g, g_1, h]$ has exactly 56 elements for any valid choice of g, g_1 and h from \mathbf{S}_3 . By Corollary 3.12 this is the smallest possible size for such a semigroup.

For integers a, b, r we will use the notation $a + (b \bmod(r))$ to denote the sum of a with the smallest non-negative element of the equivalence class $b \bmod(r)$. We will also use the notation $[a/b]$ to denote the *integer part* of the rational number a/b . For example, for any pair of integers n and m we have $n = m[n/m] + (n \bmod(m))$. We now present a second construction for INFB semigroups whose variety does not contain \mathbf{B}_2^1 .

Definition 4.5. Let \mathbf{G} be a centreless group with exponent d and identity element e and let $\langle a \rangle$ be a finite cyclic semigroup of index 2 and period p generated by an element a . Let q be the lowest common multiple of d and p . Suppose p has two divisors l and r , not both 1, such that there are elements L and R of \mathbf{G} with order p/l and p/r respectively and a mapping $f: \langle a \rangle \rightarrow \mathbf{G}$ satisfying:

- (i) $f(a) \neq f(a^{1+p})$;
- (ii) for all $i, j \geq 0$ with $i + j \leq 1 + p$,

$$f(a^{2+i+j}) = L^{[j/l]} f(a^{2+i+(j \bmod(l))}) = f(a^{2+i+(j \bmod(r))}) R^{[j/r]};$$

- (iii) for any $j \leq p$, if $f(a^{i+j}) = f(a^i)$ for every i with $1 + p \geq i \geq 2$ then $p = j$.

Then $\Xi_2[\mathbf{G}, L, R, f, p]$ is the groupoid

$$\mathcal{M}(\mathbf{G}, 2 + l, 2 + r, P) \cup \{1, a, a^2, \dots, a^{1+p}\}$$

where $P_{i,j} = f(a^{i+j})$ for $i + j \geq 1$, $P_{0,0} = e$ and multiplication is defined by

$$a(i, k, j) = \begin{cases} (i + 1, k, j), & \text{if } i < 2, \\ (2 + ((i - 1) \bmod(r)), R^{[(i-1)/r]} k, j), & \text{if } 2 \leq i \leq 2 + r \end{cases}$$

$$(i, k, j)a = \begin{cases} (i, k, j + 1), & \text{if } j < 2, \\ (i, kL^{[(j-1)/l]}, 2 + ((j - 1) \bmod(l))), & \text{if } 2 \leq j \leq 2 + l \end{cases}$$

and $a^i x = a^{i-1}(ax)$, $xa^i = (xa)a^{i-1}$.

The following lemma is useful when considering this construction.

Lemma 4.6. *In general, $a^n(i, s, j) = (2 + ((i+n-2) \bmod(r)), R^{[(i+n-2)/r]} s, j)$ and $(i, s, j)a^n = (i, sL^{[(j+n-2)/l]}, 2 + ((j+n-2) \bmod(l)))$.*

Proof. We use induction. Firstly $a(i, s, j) = (2 + ((i-1) \bmod(r)), R^{[(i-1)/r]} s, j)$ so the claim is true for $n = 1$. Assume that $a^n(i, s, j) = (2 + ((i+n-$

$2) \bmod(r), R^{[(i+n-2)/r]}_s, j)$. We now show that $a^{n+1}(i, s, j) = (2 + ((i + n - 1) \bmod(r)), R^{[(i+n-1)/r]}_s, j)$.

$$\begin{aligned} & \text{Now } a^{n+1}(i, s, j) \\ &= a(2 + ((i + n - 2) \bmod(r)), R^{[(i+n-2)/r]}_s, j), \text{ (by assumption)} \\ &= (2 + (2 + ((i + n - 2) \bmod(r)) - 1) \bmod(r), \\ &\quad R^{[(2+((i+n-2) \bmod(r))-1)/r]} R^{[(i+n-2)/r]}_s, j) \\ &= (2 + ((i + n - 1) \bmod(r)), R^{[(1+((i+n-2) \bmod(r)))/r]} R^{[(i+n-2)/r]}_s, j). \end{aligned}$$

It remains to show that $R^{[(1+((i+n-2) \bmod(r)))/r]} R^{[(i+n-2)/r]} = R^{[(i+n-1)/r]}$. Let t be the element $R^{[(1+((i+n-2) \bmod(r)))/r]} R^{[(i+n-2)/r]} \in \mathbf{G}$. Now either $1 + ((i + n - 2) \bmod(r)) < r$ or $1 + ((i + n - 2) \bmod(r)) = r$. If $1 + ((i + n - 2) \bmod(r)) < r$ then $((i + n - 1) \bmod(r)) < r$ and so $[\frac{i+n-2}{r}] = [\frac{i+n-1}{r}]$. In this case $t = R^{[(i+n-2)/r]} = R^{[(i+n-1)/r]}$ as required. If $1 + ((i + n - 2) \bmod(r)) = r$ then $R^{[(1+((i+n-2) \bmod(r)))/r]} = R$, $((i + n - 2) \bmod(r)) = r - 1$ and $[\frac{i+n-2}{r}] = [\frac{i+n-1}{r}] - 1$. So $t = RR^{[(i+n-1)/r]-1} = R^{[(i+n-1)/r]}$ as required. Therefore by induction the result is true for all $n \geq 1$.

The corresponding result for multiplication on the right follows by symmetry. \blacksquare

Lemma 4.7. *The groupoid $\Xi_2[\mathbf{G}, L, R, f, p]$ as constructed in Definition 4.5 is an INFB semigroup and $\mathbf{B}_2^1 \notin \mathcal{V}(\Xi_2[\mathbf{G}, L, R, f, p])$.*

Proof. Again we will omit the proof that $\Xi_2[\mathbf{G}, L, R, f, p]$ is a semigroup though it is routine (once given Lemma 4.6 [4]). To show that $\Xi_2[\mathbf{G}, L, R, f, p]$ is INFB, note that if q is the period of $\Xi_2[\mathbf{G}, L, R, f, p]$ (the lowest common multiple of the exponent d of \mathbf{G} and the period p of $\langle a \rangle$) then INFB occurs at the dividing pair $(a, (0, e, 0))$ since

$$(0, e, 0)a(0, e, 0) = (0, f(a), 0)$$

and by Lemma 4.6,

$$(0, e, 0)a^{q+1}(0, e, 0) = (0, e, 0)a^{p+1}(0, e, 0) = (0, f(a^{p+1}), 0)$$

where $f(a) \neq f(a^{p+1})$.

It remains to show that $\mathbf{B}_2^1 \notin \mathcal{V}(\Xi_2[\mathbf{G}, L, R, f, p])$. As in Lemma 4.3, we use the identity $(xyx^2y)^q \approx (xy)^q$. Since both p and d divide q as numbers, it follows that $\Xi_2[\mathbf{G}, L, R, f, p]$ satisfies this identity for essentially the same reasons as in the proof of Lemma 4.3. \blacksquare

We will say that $\Xi_2[\mathbf{G}, L, R, f, p]$ is a *small INFB semigroup of the second kind* and denote the set of all such monoids by Ξ_2 .

We will now construct an example.

Example 4.8. Consider the symmetric group \mathbf{S}_3 of order 6 (and exponent 6) with presentation $\langle L, R; L^3 = R^2 = e, LR = RL^2 \rangle$ and let p be the number 6. The orders of L and R are 3 and 2 respectively so the numbers l and r required by Definition 4.5 are 2 and 3 respectively. Finally, define a mapping f according to the following table:

$f(a)$	$f(a^2)$	$f(a^3)$	$f(a^4)$	$f(a^5)$	$f(a^6)$	$f(a^7)$
L	L	R	L^2	LR	$L^3 = R^2 = e$	L^2R

It is easily verified that f satisfies the requirements of Definition 4.5. So the sandwich matrix P of the completely simple ideal of $\Xi_2[\mathbf{S}_3, L, R, f, p]$ is

$$\begin{pmatrix} e & L & L & R \\ L & L & R & L^2 \\ L & R & L^2 & LR \\ R & L^2 & LR & e \\ L^2 & LR & e & L^2R \end{pmatrix}$$

We now show that the class $\Xi_1 \cup \Xi_2 \cup \mathbf{B}_2^1$ contains all minimal finite INFB semigroups.

Theorem 4.9. *Let \mathbf{S} be a finite semigroup. The \mathbf{S} is INFB if and only if there is a monoid $\mathbf{T} \in \Xi_1 \cup \Xi_2 \cup \{\mathbf{B}_2^1\}$ with $\mathbf{T} \in \mathcal{V}(\mathbf{S})$.*

Proof. The “if” implication follows immediately from the property of being INFB. Now we show that the reverse implication is also true. Firstly by Theorem 1.2 we may assume that \mathbf{S} is a monoid with identity element 1 and that there is an idempotent e and an element a so that INFB occurs at (a, e) . Now assume that $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$. So by Lemma 2.1, $ea^i e \in \mathbf{S}_e$ for every $i \geq 0$. We will take a series of subsemigroups and homomorphic images until we arrive at a semigroup isomorphic to one from $\Xi_1 \cup \Xi_2$. This process is equivalent to taking a single homomorphic image of a subsemigroup of \mathbf{S} (see [2] for example). The small INFB semigroup we arrive at is therefore a *divisor* of \mathbf{S} .

Consider the subsemigroup \mathbf{T} of \mathbf{S} generated by the set $\mathbf{S}_e \cup \{a, 1\}$. Now a still divides e in \mathbf{T} since $(e)a(e(eae)^{-1}) = e$. Let i and p be the index and period respectively of $\langle a \rangle$ (the subsemigroup generated by a). By Lemma 3.1, i is at least 2. Also since p divides the period of \mathbf{T} and the period of \mathbf{T} divides the period of \mathbf{S} (say d), the property $ea^i e \not\equiv ea^{p+1}e$ modulo $\Gamma(\mathbf{S}_e)$ is preserved and therefore \mathbf{T} is an INFB submonoid of \mathbf{S} . Note that \mathbf{T}_e is identical to \mathbf{S}_e .

Since \mathbf{T} is generated by $\mathbf{T}_e \cup \{a, 1\}$ and $ea^k e \in \mathbf{T}_e$ for every $k \geq 0$ (recall $a^0 = 1$), every element in \mathbf{T} except 1 and a can be considered as a word of the form $a^n s a^m$ where n and m are non negative integers and $s \in \mathbf{T}_e$. We now want to replace the non-nilpotent group \mathbf{T}_e with a centreless (and therefore also non-nilpotent) group. Consider the equivalence θ_1 defined as

$$\{(x, y): x = y \text{ or } x = a^n s a^m, y = a^n t a^m, n, m \geq 0 \text{ and } s \equiv t \pmod{\Gamma(\mathbf{T}_e)}\}.$$

This is a congruence since if $a^n sa^m$ and $a^n ta^m$ are equivalent modulo θ_1 then

$$a^{n'} ga^{m'} a^n sa^m = a^{n'} gea^{m'+n} esa^m$$

and

$$a^{n'} ga^{m'} a^n ta^m = a^{n'} gea^{m'+n} eta^m$$

for any non-negative integers n' and m' and $g \in \mathbf{T}_e$. Since $s \equiv t \pmod{\Gamma(\mathbf{T}_e)}$ we must have

$$gea^{m'+n} es \equiv gea^{m'+n} et \pmod{\Gamma(\mathbf{T}_e)}$$

and therefore

$$(a^{n'} gea^{m'+n} esa^m, a^{n'} gea^{m'+n} eta^m) \in \theta_1.$$

So θ_1 is a left congruence and likewise, by symmetry, a right congruence. Let $\bar{\mathbf{T}}$ denote the monoid \mathbf{T}/θ_1 . This is still an INFB monoid since $ea e$ and $ea^{d+1}e$ are not equivalent modulo $\Gamma(\mathbf{T}_e)$ and so $(ea e)/\theta_1 = (ea e)\Gamma(\mathbf{T}_e) \neq (ea^{d+1}e)\Gamma(\mathbf{T}_e) = (ea^{d+1}e)/\theta_1$. To avoid unnecessarily complicated expressions we will relabel the equivalence classes of $\bar{\mathbf{T}}$ so that a/θ_1 becomes a and e/θ_1 becomes e . By the definition of the upper central series, the group $\bar{\mathbf{T}}_e$ is centreless and so $\Gamma(\bar{\mathbf{T}}_e) = \{e\}$. Therefore two elements of $\bar{\mathbf{T}}_e$ are equivalent modulo $\Gamma(\bar{\mathbf{T}}_e)$ if and only if they are equal.

Let j be the smallest positive integer such that $ea^{i-j}e \neq ea^{i+p-j}e$ (recall that i is the index of $\langle a \rangle$). Such a j exists since $ea^{i-(i-1)}e = ea e \neq ea^{p+1}e = ea^{i+p-(i-1)}e$. So by the choice of j , for any $k < j$, $ea^{i-k}e = ea^{i+p-k}e$. Now a divides $a^{i-j-1}(ea^{i-j-1}e)^{-1}$ since $a^{i-j-1}(ea^{i-j-1}e)^{-1} = a^{i-j-1}(ea^{i-j-1}e)^{-1} \times a \times (ea e)^{-1}$. Also $a^{i-j-1}(ea^{i-j-1}e)^{-1}$ is idempotent since

$$\begin{aligned} & a^{i-j-1}(ea^{i-j-1}e)^{-1} a^{i-j-1}(ea^{i-j-1}e)^{-1} \\ &= a^{i-j-1}(ea^{i-j-1}e)^{-1} ea^{i-j-1}e (ea^{i-j-1}e)^{-1} \\ &= a^{i-j-1}(ea^{i-j-1}e)^{-1}. \end{aligned}$$

Therefore $(a, a^{i-j-1}(ea^{i-j-1}e)^{-1})$ is a dividing pair and the set $\{a^{i-j-1}s : s \in \bar{\mathbf{T}}_e\}$ is a subgroup of $\bar{\mathbf{T}}$ isomorphic to $\bar{\mathbf{T}}_e$ (it is easily verified that the map $f: \bar{\mathbf{T}}_e \rightarrow \{a^{i-j-1}s : s \in \bar{\mathbf{T}}_e\}$ given by $f(s) = a^{i-j-1}(ea^{i-j-1}e)^{-1}s$ is an isomorphism). Now

$$(a^{i-j-1}(ea^{i-j-1}e)^{-1})a(a^{i-j-1}(ea^{i-j-1}e)^{-1})$$

is equal to

$$a^{i-j-1}(ea^{i-j-1}e)^{-1} ea^{i-j}e (ea^{i-j-1}e)^{-1}$$

and

$$(a^{i-j-1}(ea^{i-j-1}e)^{-1})a^{p+1}(a^{i-j-1}(ea^{i-j-1}e)^{-1})$$

is equal to

$$a^{i-j-1}(ea^{i-j-1}e)^{-1} ea^{i+p-j}e (ea^{i-j-1}e)^{-1}.$$

But since $ea^{i-j}e$ and $ea^{i+p-j}e$ are not equal, neither can be

$$a^{i-j-1}(ea^{i-j-1}e)^{-1}ea^{i-j}e(ea^{i-j-1}e)^{-1}$$

and

$$a^{i-j-1}(ea^{i-j-1}e)^{-1}ea^{i+p-j}e(ea^{i-j-1}e)^{-1}.$$

Therefore INFB occurs at $(a, a^{i-j-1}(ea^{i-j-1}e)^{-1})$ and for every $k > 1$,

$$(a^{i-j-1}(ea^{i-j-1}e)^{-1})a^k(a^{i-j-1}(ea^{i-j-1}e)^{-1})$$

is equal to

$$(a^{i-j-1}(ea^{i-j-1}e)^{-1})a^{p+k}(a^{i-j-1}(ea^{i-j-1}e)^{-1}).$$

Let the idempotent $a^{i-j-1}(ea^{i-j-1}e)^{-1}$ be denoted by f and let \mathbf{U} be the submonoid of $\bar{\mathbf{T}}$ generated by $\bar{\mathbf{T}}_f \cup \{a, 1\}$. Using the same kind of argument as was used in the case of \mathbf{T} , we have that \mathbf{U} is an INFB submonoid of $\bar{\mathbf{T}}$ and INFB occurs at (a, f) . However, as was noted above, $fa^k f = fa^{k+p} f$ in \mathbf{U} for all $k > 1$.

Now consider the equivalence on the set $\{a, a^2, a^3, \dots, a^{i+p-1}\}$ given by

$$\phi_p = \{(a^j, a^k): j = k \text{ or } j, k > 1 \text{ and } j \equiv k \pmod{p}\}.$$

Since $sfa^k ft = sfa^{k+p} ft$ in \mathbf{U} for all $k > 1$ and any $s, t \in \mathbf{U}_f$, ϕ_p generates a congruence θ_2 on \mathbf{U} equal to

$$\{(x, y): x = y \text{ or } (x, y) \in \phi_p; \\ \text{or } x = a^{j_1} s a^{k_1}, y = a^{j_2} s a^{k_2} \text{ and both } (a^{j_1}, a^{j_2}), (a^{k_1}, a^{k_2}) \in \phi_p\}.$$

Let $\bar{\mathbf{U}}$ be the semigroup \mathbf{U}/θ_2 . For the sake of simplicity we will relabel the equivalence classes so that a/θ_2 becomes a and f/θ_2 becomes e . So (a, e) is a dividing pair, $\bar{\mathbf{U}}_e$ is centreless, $ea^j e \in \bar{\mathbf{U}}_e$ for all $j \geq 0$, and $eae \neq ea^{1+p}e$. Furthermore, the index of $\langle a \rangle$ (the subsemigroup generated by a) is now 2, that is $a^2 = a^{2+p}$.

Consider now the subsemigroup

$$\mathbf{C} = \{a^i s a^j: i, j \geq 0, s \in \bar{\mathbf{U}}_e\}$$

This is an ideal of the semigroup $\bar{\mathbf{U}}$ and every element in \mathbf{C} divides every other element since for any $i, i', j, j' \geq 0, s, t \in \bar{\mathbf{U}}_e$,

$$a^i s a^j = (a^i t^{-1} (ea^{i'} e)^{-1}) (a^{i'} t a^{j'}) ((ea^{j'} e)^{-1} s a^j).$$

Thus \mathbf{C} is a completely simple subsemigroup of $\bar{\mathbf{U}}$. It is clear also that the \mathcal{H} -classes of \mathbf{C} are sets of the form $\{a^i s a^j: s \in \bar{\mathbf{U}}_e\}$ and we will denote such an \mathcal{H} -class by $H_{i,j}$. The proof will now split into two cases. The first is the

situation when a^2 is contained in \mathbf{C} . The corresponding semigroups will be elements of Ξ_1 , the INFB semigroups of the first kind. The second situation is when $a^2 \notin \mathbf{C}$. In this case it is possible that some further reduction may be made.

Case 1. $a^2 \in \mathbf{C}$.

In this case the set $\{a^2, a^3, \dots, a^{1+p}\}$ is a cyclic subgroup of some group $H_{i,j}$. Now by the Rees-Suschkewitz theorem, \mathbf{C} is isomorphic to a Rees Matrix Semigroup over the centreless group $\bar{\mathbf{U}}_e$ with sandwich matrix P . Since every element in \mathbf{C} is of the form $a^i sa^j$ and $a^2 \in \mathbf{C}$, P must be at most a 3×3 matrix. Since Theorem 3.11 shows that the sets $\{1, a\}$ and each $H_{i,j}$ for $i, j \leq 2$ are disjoint, P must be exactly a 3×3 matrix. Now if $a^2 = a^i sa^j$ where $i < 2$ and $j \leq 2$ then every element in \mathbf{C} can be written in the form $a^{i'} ta^{j'}$ for $t \in \bar{\mathbf{U}}_e$, $i' < 2$ and $j' \leq 2$ and then P is only an $i' \times 3$ or $3 \times i'$ matrix, a contradiction. Therefore, by symmetry, $a^2 = a^2 sa^2$ for some $s \in \bar{\mathbf{U}}_e$ and the subgroup $\{a^2, a^3, \dots, a^{1+p}\}$ is a cyclic subgroup of $H_{2,2}$. Note also that $\{a^2, a^3, \dots, a^{1+p}\}$ is generated by a^{1+p} since $(a^{1+p})^n = a^{n+np} = a^n$. For some $g_1 \in \bar{\mathbf{U}}_e$, $a^{1+p} = a^2 g_1 a^2$. Let g be the element $ea^4 e$ and define a map $\iota: \bar{\mathbf{U}} \rightarrow \Xi_1(\bar{\mathbf{U}}_e, g, g_1, eae)$ by

$$\iota(1) = 1, \quad \iota(a) = a, \quad \iota(a^i sa^j) = (i, s, j).$$

We show that ι is an isomorphism. It is certainly a bijection since in the arguments above we have shown that $\bar{\mathbf{U}}$ contains only elements of the form 1 , a , and $a^i sa^j$ (for $i, j \leq 2$ and $s \in \bar{\mathbf{U}}_e$) and Theorem 3.11 shows that these are distinct in any finite INFB semigroup whose variety does not contain \mathbf{B}_2^1 . We need to show that for any elements $x, y \in \bar{\mathbf{U}}$, $\iota(xy) = \iota(x)\iota(y)$. The case when x or y is 1 is trivial.

Consider the case when $x = a^i sa^j$ and $y = a^{i'} ta^{j'}$. Now

$$\iota(a^i sa^j a^{i'} ta^{j'}) = \iota(a^i sea^{i'+j} eta^{j'}) = (i, sea^{i'+j} et, j')$$

and

$$\iota(a^i sa^j)\iota(a^{i'} ta^{j'}) = (i, s, j)(i', t, j') = (i, sP_{j,i'}t, j').$$

As in Definition 4.1, put $g_i = (g_1 g)^{i-1} g_1$. If $i' + j \geq 2$ then

$$a^{j+i'} = (a^{1+p})^{j+i'} = (a^2 g_1 a^2)^{j+i'} = a^2 (g_1 e a^4 e)^{j+i'-1} g_1 a^2 = a^2 g_{j+i'} a^2$$

and therefore $ea^{j+i'} e = ea^2 e g_{j+i'} ea^2 e$. Now $a^2 = a^2 g_2 a^2$ so $ea^2 e = ea^2 g_2 a^2 e$. Therefore $g_2 = (ea^2 e)^{-1}$. This implies that

$$ea^{j+i'} e = ea^2 e g_{j+i'} ea^2 e = g_2^{-1} g_{j+i'} g_2^{-1} = P_{j,i'}$$

as required. If $i' = j = 0$ then $P_{j,i'} = P_{0,0} = e = ea^0 e$ as required. Finally if $i' = 1$ and $j = 0$ (the case when $i' = 0$ and $j = 1$ follows by symmetry) then

$\iota(a^i s a t a^{j'}) = (i, seaet, j') = (i, sP_{0,1}t, j') = (i, s, 0)(1, t, j') = \iota(a^i s)\iota(ata^{j'})$, also as required.

Now consider the case when $x = a$ and $y = a^i s a^j$. Firstly assume $i = 0$. Then $\iota(aa^i s a^j) = \iota(asa^j) = (1, s, j) = a(0, s, j) = \iota(a)\iota(sa^j)$, as required.

Now assume that $i > 0$. Therefore

$$\iota(aa^i s a^j) = \iota(a^{i+1} s a^j) = \iota(a^2 g_{i+1} a^2 s a^j) = (2, g_{i+1} e a^2 e s, j).$$

But $ea^2 e = g_2^{-1}$, so $\iota(aa^i s a^j) = (2, g_{i+1} g_2^{-1} s, j)$. If $i = 1$ then

$$(2, g_{i+1} g_2^{-1} s, j) = (2, s, j) = a(1, s, j) = \iota(a)\iota(asa^j)$$

as required. If $i = 2$ then

$$(2, g_{i+1} g_2^{-1} s, j) = (2, g_3 g_2^{-1} s, j) = a(2, s, j) = \iota(a)\iota(a^2 s a^j),$$

also as required.

Up to symmetry, the only remaining case is when $x = y = a$, and $\iota(aa) = \iota(a)\iota(a)$ follows immediately since $a^2 = a^2 g_2 a^2$ in $\bar{\mathbf{U}}$ and $a^2 = (2, g_2, 2)$ in $\Xi_1(\bar{\mathbf{U}}_e, g, g_1, eae)$. Therefore ι is an isomorphism.

Case 2. $a^2 \notin \mathbf{C}$.

Since $\{a^2, a^3, \dots, a^{1+p}\}$ forms a cyclic subgroup of $\bar{\mathbf{U}}$, it must be that $a^i \notin \mathbf{C}$ for all $i \geq 0$. Recall that if $i > 1$ then $ea^i e = ea^{i+p} e$ and that ϕ_p is the equivalence

$$\{(a^j, a^k): j = k \text{ or } j, k > 1 \text{ and } j \equiv k \pmod{p}\}$$

on the set $\{a, a^2, a^3, \dots\}$. Let q be the smallest number such that for all $i > 1$, $ea^i e = ea^{i+q} e$. It is easily verified that the equivalence θ_3 given by

$$\{(x, y): x = y; \text{ or } (x, y) \in \phi_q; \\ \text{ or } x = a^i s a^j, y = a^{i'} s a^{j'}, \text{ and both } (a^i, a^{i'}), (a^j, a^{j'}) \in \phi_q\}$$

is a congruence that preserves the property of being an INFB monoid. Let the semigroup $\bar{\mathbf{U}}/\theta_3$ be denoted \mathbf{V} and let the equivalence classes a/θ_3 and e/θ_3 be relabeled a and e respectively. Note that the group $\mathbf{V}_e = \bar{\mathbf{U}}_e/\theta_3$ is isomorphic to $\bar{\mathbf{U}}_e$ and that the period of $\langle a \rangle$ (the subsemigroup of \mathbf{V} generated by a) is now the number q .

In \mathbf{V} , $ea^i e = ea^{i+j} e$ for any fixed $j \leq q$ and for all $i > 1$ if and only if $j = q$. If $ea^2 = ea^{2+j}$ then $ea^2 a^{i-2} = ea^i = ea^{2+j} a^{i-2} = ea^{i+j}$ for any $i \geq 2$. Therefore the number j in these equations must be q . Likewise $a^2 e = a^{2+j} e$ if and only if $j = q$. Say there exists $j \geq 2$ such that $ea^j = ga$ for some $g \in \mathbf{V}_e$. Then $ea^j a^q = ea^j$ and $gaa^q = ga^{q+1}$. Therefore $gae = ea^j e = ga^{q+1} e$, a contradiction since $eae \neq ea^{q+1} e$. It follows that no such j exists and likewise that there is no integer $j \geq 2$ such that $a^j e = ag$. This also guarantees

that no $j \geq 1$ exists so that $a^j e = g$ or $ea^j = g$ since then, for example, $a^{j+1}e = ag$. Now let r be the smallest number such that $a^2 e = a^{2+r}R$ for some $R \in \mathbf{V}_e$ and l be the smallest number such that $ea^2 = La^{2+l}$ for some $L \in \mathbf{V}_e$. Now r must divide q since otherwise there are numbers k and k' such that $kr \equiv k' \pmod{q}$ and $k' < r$. In this case $a^{2+kr}e = a^2 R^k$ and $a^{2+k'r}e = a^{2+k'}e$, contradicting the minimality of r . Likewise, l must divide q also: say $q = nr = ml$. Now since $a^{2+q}e = a^2 e$ and $a^{2+q}e = a^{2+nr}e = a^2 R^n$ we must have that $R^n = e$ and therefore the order of R divides q . Let the order of R be k (note that k necessarily divides n). Then $a^{2+rk}e = a^2 R^k = a^2 e$ and therefore $ea^i e = ea^{i+rk}e$ for every $i > 1$. By the choice of \mathbf{V} however, this is true only if $rk = q$. Therefore the order of R is n and, by symmetry, the order of L is m .

If $l = r = 1$ then L and R have the same order and for any integer $i \geq 0$, $L^i(ea^2 e) = ea^{2+i}e = (ea^2 e)R^i$. Furthermore for any $k \geq 1$ and $i, j \geq 0$,

$$a^{2+k} \times a^i sa^j = a^2 R^{i+k} sa^j$$

and

$$\begin{aligned} a^2 R^k (ea^2 e)^{-1} a^2 \times a^i sa^j &= a^2 R^k (ea^2 e)^{-1} ea^2 R^i sa^j \\ &= a^2 R^{i+k} sa^j. \end{aligned}$$

Likewise $a^i sa^j \times a^{2+k} = a^i sa^j \times a^2 (ea^2 e)^{-1} L^k a^2$. But $ea^2 e R^k = L^k ea^2 e$ so it follows that $a^2 (ea^2 e)^{-1} L^k a^2 = a^2 R^k (ea^2 e)^{-1} a^2$. Therefore multiplication on the left of an element of \mathbf{C} by a^{2+k} is the same as multiplication on the left by $a^2 R^k (ea^2 e)^{-1} a^2$ (or equivalently by $a^2 (ea^2 e)^{-1} L^k a^2$) and likewise for multiplication on the right. We also have $a \times a^{2+k} = a^{2+(k+1)} = a^{2+k} \times a$ and

$$a \times a^2 R^k (ea^2 e)^{-1} a^2 = a^2 R^{k+1} (ea^2 e)^{-1} a^2 = a^2 R^k (ea^2 e)^{-1} a^2 \times a.$$

Therefore the equivalence θ_4 given by

$$\{(x, y): x = y; \text{ or } x = a^{2+k} \text{ and } \\ y = a^2 R^k (ea^2 e)^{-1} a^2; \text{ or } x = a^2 R^k (ea^2 e)^{-1} a^2 \text{ and } y = a^{2+k}\}$$

is a congruence. The resulting quotient of \mathbf{V} is an INFB semigroup of the type described in Case 1. Therefore we can assume that not both of l and r are 1.

Now we are ready to compare \mathbf{V} to a semigroup from Ξ_2 . Define a map

$$f: \{a, a^2, \dots, a^{1+p}\} \rightarrow \mathbf{V}_e$$

by $f(a^i) = ea^i e$. For any $i < r$ and any $s \in \mathbf{V}_e$, $a^{2+i}e \neq a^2 s$ and likewise for l , there are at most $(2+r) \times (2+l)$ \mathcal{H} -classes $H_{i,j}$. To see that there are exactly $(2+r) \times (2+l)$ \mathcal{H} -classes of the form $H_{i,j}$, note that if $a^i sa^j = a^{i'} ta^{j'}$ with j and j' less than $2+l$ then $ea^i sa^j = ea^{i'} ta^{j'}$, that is there is an element

$v \in \mathbf{V}_e$ so that $ea^j = va^{j'}$. Say $j' \neq j$. If both j and j' are greater than 1 then because $a^{2+p} = a^2$ we have $ea^2 = va^{2+|j-j'|}$, contradicting the minimality of l . If one of j and j' , say j' , is less than 2 then, we have either $ea = ea^j$ or $e = ea^j$. In either case we obtain $ea = ea^k$ for some $k > 1$. But then $ea^2 = ea^{1+k}$. By the arguments above, $1+k$ must equal $2+p$ and therefore $ea = ea^{1+p}$, contradicting the fact that INFB occurs at (a, e) . Therefore j must equal j' . Likewise by symmetry if i and i' are less than $2+r$ then $i = i'$.

For $i < 2+r$, $j < 2+l$ and $0 \leq k \leq 1+q$ define a map $\iota': \mathbf{V} \rightarrow \Xi_2(\mathbf{V}_e, L, R, f, q)$ by $\iota'(a^i sa^j) = (i, s, j)$, $\iota'(a^k) = a^k$. It is clear that ι' is a bijection. To show it is an isomorphism, up to symmetry there is only one nontrivial case to check that is not already covered by corresponding arguments for the map ι in Case 1. The case that remains is to verify that $\iota'((a^k)(a^i sa^j)) = \iota'(a^k)\iota'(a^i sa^j)$. If $k+i \geq 2$, the left side of this expression equals

$$\begin{aligned} \iota'(a^{2+((k+i-2)\bmod(r))} R^{[(k+i-2)/r]} sa^j) &= (2 + ((k+i-2)\bmod(r)), \\ &\quad R^{[(k+i-2)/r]} s, j) \\ &= a^k(i, s, j) \quad (\text{by Lemma 4.6}) \\ &= \iota'(a^k)\iota'(a^i sa^j) \end{aligned}$$

as required. If $k+i < 2$ we can assume that $k=1$ and $i=0$ (since the cases when $k=0$ are trivial) and the left side becomes

$$\iota'((a)(sa^j)) = (1, s, j) = a(0, s, j) = \iota'(a)\iota'(sa^j)$$

as required. Therefore \mathbf{V} is isomorphic to $\Xi_2[\mathbf{V}_e, L, R, f, q]$ and the proof is complete. \blacksquare

The next result follows from Theorem 4.9 but can also be obtained (using effectively the same proof) as an easy corollary of Sapir's construction in [18] of an INFB finite semigroup not containing \mathbf{B}_2^1 .

Corollary 4.10. *There are infinitely many minimal finitely generated INFB varieties.*

Proof. We will only consider the small INFB semigroups of the first kind. It is well known that if $p > 2$ is a prime number then the dihedral group \mathbf{D}_p given by $\langle a, b: a^p = 1 = b^2, a^{p-1}b = ba \rangle$ is centreless (the proof of this and more general results are popular exercises in many group theory texts; see [16] for example). Let \mathbf{S} and \mathbf{T} be two monoids from Ξ_1 with largest subgroup \mathbf{D}_p and \mathbf{D}_q respectively (p and q distinct primes). For each number $n \geq 1$, \mathbf{D}_n has exponent $2n$ so therefore $\mathbf{S} \models x^2 \approx x^{2+2p}$ and $\mathbf{T} \models x^2 \approx x^{2+2q}$. In this case, any semigroup $\mathbf{U} \in \mathcal{V}(\mathbf{S}) \cap \mathcal{V}(\mathbf{T})$ satisfies $x^2 \approx x^{2+g.c.d.(2p, 2q)} \equiv x^{2+2}$. Any subgroup \mathbf{G} of \mathbf{U} must therefore have exponent 2. So \mathbf{G} satisfies $xy \approx xy(yx) \approx x(yy)xy \approx xxyx \approx yx$. That is, \mathbf{G} is abelian, and therefore nilpotent. Therefore \mathbf{U} is INFB if and only if $\mathbf{B}_2^1 \in \mathcal{V}(\mathbf{U})$. Since $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$ and $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{T})$, \mathbf{U} is WFB. The result now follows since there are infinitely many prime numbers and consequently infinitely many dihedral groups \mathbf{D}_p . \blacksquare

Note that if INFB occurs at a dividing pair (x, y) in a semigroup $\mathbf{S} \in \Xi_1 \cup \Xi_2$ with maximal subgroup \mathbf{G} , then since a is the only non group element in \mathbf{S} , x must equal a by Lemma 3.1. Every idempotent in \mathbf{S} of the form (i, s, j) for $s \in \mathbf{G}$ and $i+j \geq 1$ can be written in the form $a^i(0, t, 0)a^j$ for some element $t \in \mathbf{G}$. But since $i+j \geq 1$ and the index of \mathbf{S} is 2,

$$\begin{aligned} a^i(0, t, 0)a^j a a^i(0, t, 0)a^j &= a^i(0, t, 0)a^{j+i+1}(0, t, 0)a^j \\ &= a^i(0, t, 0)a^{j+i+1+d}(0, t, 0)a^j \\ &= a^i(0, t, 0)a^j a^{d+1} a^i(0, t, 0)a^j \end{aligned}$$

and therefore INFB does not occur at $(a, a^i(0, t, 0)a^j)$. So the idempotent y must be one of 1, $(0, e, 0)$ or possibly a^d if $\mathbf{S} \in \Xi_2$. Since INFB occurs at (a, y) it is easily verified that the only possibility for y is $(0, e, 0)$. That is, there is only one dividing pair in \mathbf{S} where INFB occurs. Since INFB occurs for at least two distinct dividing pairs in both \mathbf{B}_2^1 and \mathbf{A}_2^1 , by Lemma 2.1 we have proved the following theorem.

Theorem 4.11. *If INFB occurs at only one dividing pair in a finite semigroup \mathbf{S} then \mathbf{B}_2^1 is not contained in the variety $\mathcal{V}(\mathbf{S})$ and there is a monoid $\mathbf{T} \in \Xi_1 \cup \Xi_2$ such that $\mathbf{T} \in \mathcal{V}(\mathbf{S})$.*

Note also that since the index of every semigroup \mathbf{S} in $\Xi_1 \cup \Xi_2$ is only 2, for any element x and any $i > 1$, the element x^i lies in a subgroup of \mathbf{S} . From this it is easily verified that every semigroup from Ξ_1 of period d satisfies $(x^2y)^d \approx (x^3y)^d$ and $(yx^2)^d \approx (yx^3)^d$. However if \mathbf{S} is a semigroup from Ξ_2 then the numbers l and r are not both 1, and either $a^2(0, e, 0)$ and $a^3(0, e, 0)$ or $(0, e, 0)a^2$ and $(0, e, 0)a^3$ must lie in different subgroups of \mathbf{S} . In this case one of the identities $(x^2y)^d \approx (x^3y)^d$ and $(yx^2)^d \approx (yx^3)^d$ must fail on \mathbf{S} . Thus a minimal finite INFB semigroup in a variety generated by a semigroup from Ξ_1 must be a semigroup from Ξ_1 . It is unknown if the same is true for the semigroups in Ξ_2 : possibly there are no minimal finite INFB semigroups in Ξ_2 (in which case $\Xi_1 \cup \{\mathbf{B}_2^1\}$ contains all minimal finite INFB semigroups). If however we replace “minimal finite INFB semigroups” with “minimal finite INFB *divisors*” a complete description is possible and indeed, this class contains many small INFB semigroups of the second kind. To prove this we consider the cases of Ξ_1 and Ξ_2 separately.

If \mathbf{G} is a centreless group and a and b are elements of \mathbf{G} then we will say that a and b are Γ -*separate* in \mathbf{G} if for every proper normal subgroup \mathbf{N} , $a\mathbf{N} \not\equiv b\mathbf{N}$ modulo $\Gamma(\mathbf{G}/\mathbf{N})$. In other words, a and b are distinct modulo $\Gamma(\mathbf{G})$ but in every proper quotient of \mathbf{G} , the cosets containing a and b are equivalent modulo the corresponding upper hypercentre. In a semigroup from Ξ_1 we have that $(0, e, 0)a(0, e, 0) = (0, h, 0)$ and $(0, e, 0)a^{d+1}(0, e, 0) = (0, g_2^{-1}g_i g_2^{-1}, 0)$. Since INFB occurs at the pair $(a, (0, e, 0))$ the group elements h and $g_2^{-1}g_i g_2^{-1}$ must be distinct. This motivates the following definition.

Definition 4.12. Let $\bar{\Xi}_1$ be the subset of Ξ_1 that consists of all monoids of the form $\Xi_1[\mathbf{G}, g, g_1, h]$ such that h and $(g_1g)^{-2}g_1^{-1}$ are Γ -separate in \mathbf{G} and equivalent modulo $\Gamma(\mathbf{H})$ for every proper subgroup \mathbf{H} of \mathbf{G} containing h and $g_2^{-1}g_ig_2^{-1}$ (for $1 \leq i \leq p$).

Theorem 4.13. *Every monoid in $\bar{\Xi}_1$ is a minimal INFB divisor for the class of finite semigroups and every minimal INFB divisor for the class of finite semigroups that is contained in Ξ_1 is contained in $\bar{\Xi}_1$.*

Proof. Firstly if $\Xi_1[\mathbf{G}, g, g_1, h]$ is not contained in $\bar{\Xi}_1$ then either $(g_1g)^{-2}g_1^{-1}$ and h are not Γ -separate in \mathbf{G} or there is a subgroup \mathbf{H} of \mathbf{G} containing the entries of the sandwich matrix of $\mathcal{M}[\mathbf{G}, 3, 3, P]$ in which $(g_1g)^{-2}g_1^{-1}$ and h are not equivalent modulo $\Gamma(\mathbf{H})$. In the first case, there is a normal subgroup \mathbf{N} of \mathbf{G} so that $(g_1g)^{-2}g_1^{-1}\mathbf{N}$ and $h\mathbf{N}$ are not equivalent modulo $\Gamma(\mathbf{G}/\mathbf{N})$. The quotient \mathbf{G}/\mathbf{N} induces a congruence θ on $\Xi_1[\mathbf{G}, g, g_1, h]$ defined by

$$\{(x, y): x = y, \text{ or } x = (i, s, j), y = (i, t, j), \text{ and } s\mathbf{N} = t\mathbf{N}\}.$$

Since $(g_1g)^{-2}g_1^{-1}\mathbf{N}$ and $h\mathbf{N}$ are not equivalent modulo $\Gamma(\mathbf{G}/\mathbf{N})$, it must be the case that $\Xi_1[\mathbf{G}/\mathbf{N}, g\mathbf{N}, g_1\mathbf{N}, h\mathbf{N}]$ is an INFB divisor of $\Xi_1[\mathbf{G}, g, g_1, h]$.

In the second case, every entry in the sandwich matrix P of $\mathcal{M}(\mathbf{G}, 3, 3, P)$ is an element of \mathbf{H} , and so there is a proper INFB subsemigroup of $\Xi_1[\mathbf{G}, g, g_1, h]$ generated by \mathbf{H} , 1 and a . That is, $\Xi_1[\mathbf{G}, g, g_1, h]$ is not a minimal INFB divisor.

Now assume that $\mathbf{S} = \Xi_1[\mathbf{G}, g, g_1, h]$ is an element of $\bar{\Xi}_1$. Since $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{S})$, Theorem 3.11 implies that any congruence on \mathbf{S} whose corresponding quotient \mathbf{T} is INFB must only collapse elements within \mathcal{H} -classes. This corresponds to taking a quotient of the group \mathbf{G} in every \mathcal{H} -class of $\mathcal{M}[\mathbf{G}, 3, 3, P]$. But since $(g_1g)^{-2}g_1^{-1}$ and h are Γ -separate and \mathbf{T} is INFB, the normal subgroup of \mathbf{G} must be trivial and so \mathbf{T} is isomorphic to \mathbf{S} . For similar reasons, the definition of $\bar{\Xi}_1$ and Theorem 3.11 imply that there are no proper INFB subsemigroups of \mathbf{S} . Therefore \mathbf{S} is a minimal INFB divisor. ■

We now investigate small INFB semigroups of the second kind.

Definition 4.14. Let $\bar{\Xi}_2$ be the subset of Ξ_2 that consists of all monoids of the form $\Xi_2[\mathbf{G}, L, R, f, p]$ so that:

- (i) the elements $f(a)$ and $f(a^{1+p})$ are Γ -separate in \mathbf{G} and equivalent modulo $\Gamma(\mathbf{H})$ for every subgroup \mathbf{H} of \mathbf{G} containing $f(a^i)$ for all $1 \leq i \leq p+1$;
- (ii) the numbers r and l (that is $p/(\text{order}(R))$ and $p/(\text{order}(L))$) are the smallest choices of i and j respectively with the property that for all $k \leq p$, $f(a^{2+k+i}) = f(a^{2+k})g$ and $f(a^{2+k+j}) = hf(a^{2+k})$ for some elements g and h of \mathbf{G} not dependent on k .

Theorem 4.15. *Every monoid in $\bar{\Xi}_2$ is a minimal INFB divisor for the class of finite semigroups and every minimal INFB divisor for the class of finite semigroups that is contained in $\bar{\Xi}_2$ is contained in $\bar{\Xi}_2$.*

Proof. Let $\mathbf{S} = \Xi_2[\mathbf{G}, L, R, f, p]$ be a semigroup from Ξ_2 for which at least one of the conditions of Definition 4.14 is not satisfied. If the first condition is not satisfied then it follows by only trivial modifications of the argument used in the proof of Theorem 4.13 that there is a proper INFB divisor of \mathbf{S} . So now assume that the first condition holds for \mathbf{S} but the second condition does not. In particular let us assume that there is a smallest number $r' < r$ so that $f(a^{2+k+r'}) = f(a^{2+k})K$ for some $K \in \mathbf{G}$ and for every $k \leq p$. Now $f(a^{2+k})R = f(a^{2+k+r}) = f(a^{2+k+(r) \bmod (r')})K^{[r/r']}$ and since $(r) \bmod (r') < r'$, by the minimality of r' we must have that $(r) \bmod (r') = 0$ and $K^{[r/r']} = R$. Therefore r' divides r and $RK = KR$. We now show that the equivalence θ given by the symmetric closure of

$$\Delta \vee \{(2+k, Kg, j), (2 + ((k+r') \bmod (r)), R^{[(k+r')/r]}g, j)\};$$

$$0 \leq k \leq r-1, 0 \leq j \leq 1+l, g \in \mathbf{G}\},$$

(where Δ denotes the diagonal relation on \mathbf{S}) is a congruence so that \mathbf{S}/θ is INFB. Let $(2+i, Kg, j)$ and $(2 + ((i+r') \bmod (r)), R^{[(i+r')/r]}g, j)$ be two θ equivalent elements. Firstly

$$a(2 + ((i+r') \bmod (r)), R^{[(i+r')/r]}g, j)$$

$$= \begin{cases} (2 + ((1+i+r') \bmod (r)), R^{[(1+i+r')/r]}g, j), & \text{if } i < r-1 \\ (2 + ((r') \bmod (r)), R^{[r'/r]}Rg, j), & \text{if } i = r-1, \end{cases}$$

which is equivalent modulo θ to $(2+1+i, Kg, j) = a(2+i, Kg, j)$ if $i < r-1$ and equivalent modulo θ to $(2, KRg, j) = (2, RKg, j) = a(2+i, Kg, j)$ if $i = r-1$. That $(2+i, Kg, j)a$ and $(2 + ((i+r') \bmod (r)), R^{[(i+r')/r]}g, j)a$ are equivalent modulo θ is trivial. Likewise if $k_1, k_2 \leq 1+l$ and $h \in \mathbf{G}$ then $(2+i, Kg, j)(k_1, h, k_2)$ and $(2 + ((i+r') \bmod (r)), R^{[(i+r')/r]}g, j)(k_1, h, k_2)$ are also trivially equivalent modulo θ . Now

$$(k_1, h, k_2)(2 + ((i+r') \bmod (r)), R^{[(i+r')/r]}g, j)$$

$$= (k_1, hf(a^{k_2+2+((i+r') \bmod (r))})R^{[(i+r')/r]}g, j)$$

$$= (k_1, hf(a^{2+k_2+i+r'})g, j)$$

$$= (k_1, hf(a^{2+k_2+i})Kg, j)$$

$$= (k_1, h, k_2)(2+i, Kg, j)$$

as required. So therefore θ is a congruence on \mathbf{S} . By definition θ does not collapse elements within \mathcal{H} -classes of \mathbf{S} and so therefore $f(a)$ and $f(a^{1+p})$ are still not equivalent modulo $\Gamma(\mathbf{G})$ and \mathbf{S} is INFB. This means that members

of Ξ_2 that are not members of $\bar{\Xi}_2$ are not minimal INFB divisors. We now show that elements of Ξ_2 that are not minimal INFB divisors are not elements of $\bar{\Xi}_2$.

Let θ be a proper congruence on a semigroup $\mathbf{S} = \Xi_2[\mathbf{G}, L, R, f, p]$ from $\bar{\Xi}_2$ so that $\mathbf{T} = \mathbf{S}/\theta$ is INFB.

Case 1. $(a^i, a^j) \in \theta$ where $p+1 \geq i, j > 1$ and $i \neq j$.

Since the set $\{a^2, a^3, \dots, a^{1+p}\}$ is a cyclic subgroup of \mathbf{S} we must have $(a^{k+|i-j|}, a^k) \in \theta$ for all $k \geq 2$. Then $((0, f(a^{k+|i-j|}), 0), (0, f(a^k), 0)) \in \theta$ for all $k \geq 2$. Since for some $k \geq 2$ the elements $f(a^{k+|i-j|})$ and $f(a^k)$ are distinct in \mathbf{G} (by Definition 4.5), θ induces a nontrivial congruence on \mathbf{G} and therefore, $f(a)$ and $f(a^{1+p})$ cannot be Γ -separate in \mathbf{G} (since \mathbf{T} is INFB). That is, $\mathbf{S} \notin \bar{\Xi}_2$.

Case 2. $((j, g, k), (j', h, k')) \in \theta$ where (j, g, k) does not equal (j', h, k') .

If $j = j'$ and $k = k'$ but $g \neq h$ then clearly the restriction of θ to \mathbf{G} is a nontrivial congruence so $f(a)$ and $f(a^{1+p})$ again cannot be Γ -separate and $\mathbf{S} \notin \bar{\Xi}_2$. If $j \neq j'$ (say, $j < j'$) then both j and j' are greater than 1 (see proof of Case 2 of Theorem 4.9). So we can assume that $1 < j < j' \leq r+1$. Now because $(j, g, k) = a^j(0, e, 0)(0, g, k)$ and $(j', h, k') = a^{j'}(0, e, 0)(0, h, k')$ we have that

$$a^{2+p-j}a^j(0, e, 0)(0, g, k)(0, P_{k,0}^{-1}g^{-1}, 0) = a^2(0, e, 0)$$

and

$$a^{2+p-j}a^{j'}(0, e, 0)(0, h, k')(0, P_{k,0}^{-1}g^{-1}, 0) = a^{2+j'-j}(0, hg^{-1}, 0).$$

By multiplying on the left by $(0, e, 0)a^k$ ($0 \leq k \leq 1+p$) we have that

$$((0, e, 0)a^{2+k}(0, e, 0), (0, e, 0)a^{2+k+j'-j}(0, hg^{-1}, 0)) \in \theta$$

and therefore

$$((0, f(a^{2+k}), 0), (0, f(a^{2+k+j'-j})hg^{-1}, 0)) \in \theta$$

for every $k \geq 0$. If $f(a^{2+k}) = f(a^{2+k+j'-j})hg^{-1}$ then condition (iii) implies $\mathbf{S} \notin \bar{\Xi}_2$ (since $j'-j < r$). If $f(a^{2+k}) \neq f(a^{2+k+j'-j})hg^{-1}$ then the congruence θ induces a nontrivial congruence on the group \mathbf{G} . Since we have assumed \mathbf{S}/θ is INFB the group elements $f(a)$ and $f(a^{1+p})$ are not equivalent modulo $\Gamma(\mathbf{G}/\theta)$ and therefore they are also not Γ -separate in \mathbf{S} . So \mathbf{S} is not a semigroup from $\bar{\Xi}_2$.

Case 3. $(a^i, (j, s, k)) \in \theta$ for some $i, j, k \leq p+1$.

By Theorem 3.11, $(a, (j, s, k)) \notin \theta$. Say $(a^i, (j, s, k)) \in \theta$ with $i \geq 2$. Since we have that $\{a^2, a^3, \dots, a^{1+p}\}$ is a subgroup of \mathbf{S} , $(a^{i'}, (j, s, k)) \in \theta$ for every $i' \geq 2$ and some g depending on i' . By the arguments used above in Case 2 of the proof of Theorem 4.9, we can assume that $j = k = 2$. In accordance with Definition 4.5, let l and r be such that p/l and p/r are the orders of L and R respectively. Now at least one of l and r (say r) are greater than 1 and

therefore without loss of generality we may assume that $a^2(1, e, 2) = (3, e, 2)$. But since $(a^2, (2, g, 2)) \in \theta$ for some group element $g \in \mathbf{G}$, we must have that $((2, g, 2)(1, e, 2), (3, e, 2)) \in \theta$, that is $((2, gf(a^3), 2), (3, e, 2)) \in \theta$ and therefore $\mathbf{S} \notin \bar{\Xi}_2$ by Case 2 above.

Finally, if \mathbf{S} contains an INFB subsemigroup then the condition that $f(a)$ and $f(a^{1+p})$ be equivalent modulo $\Gamma(\mathbf{H})$ for every subgroup \mathbf{H} of \mathbf{G} containing $f(a^i)$ for all $1 \leq i \leq p+1$ is easily seen to fail and therefore $\mathbf{S} \notin \bar{\Xi}_2$. ■

Example 4.16. For any integer $p > 1$ and minimal centreless group divisor \mathbf{G} (say \mathbf{S}_3 for example) the semigroup $\Xi_2[\mathbf{G}, e, e, f, p]$ is a minimal INFB divisor if $f(a^i) = e$ for every $i \geq 1$ except when $i = 1+p$.

Thus even given a particular minimal centreless group divisor, there are infinitely many minimal finite INFB divisors containing \mathbf{G} (but also containing cyclic subgroups of arbitrary finite order). A further example that illustrates the large number of minimal finite INFB divisors is the following.

Example 4.17. If $\mathbf{G} = \{e, g_1, g_2, \dots, g_{p+1}\}$ is a finite simple group (where p is an integer) then $\Xi_2[\mathbf{G}, e, e, f, p]$ is a minimal INFB divisor if f is given by $f(a^i) = g_i$.

This follows immediately from Theorem 4.15 since there are no proper nontrivial quotients of \mathbf{G} and no proper subgroups of \mathbf{G} containing $f(a^i)$ for every $i \leq 1+p$. Note that a finite simple group need not be a minimal centreless group divisor (since every finite group \mathbf{G} is embeddable in the finite simple group $\mathbf{A}_{|\mathbf{G}|+2}$; see [16]) and so there may be a subgroup \mathbf{H} of \mathbf{G} containing both $f(a)$ and $f(a^{1+p})$ (but not $f(a^i)$ for some i) and in which $f(a)$ and $f(a^{1+p})$ are not equivalent modulo $\Gamma(\mathbf{H})$. It is for this reason that the definition of Γ -separate depends only on quotients of a group and not the more general notion of divisors.

Combining Lemma 2.1, and Theorems 4.13 and 4.15 we have a description of all minimal finite INFB divisors.

Corollary 4.18. *The class $\bar{\Xi}_1 \cup \bar{\Xi}_2 \cup \{\mathbf{B}_2^1, \mathbf{A}_2^1\}$ is, up to isomorphism, the class of minimal finite INFB divisors.*

Theorem 2.2 shows that even though the semigroups from $\bar{\Xi}_1$ and $\bar{\Xi}_2$ can each be embedded in finite regular semigroups, these semigroups necessarily generate varieties containing \mathbf{B}_2^1 . Another well known embedding theorem is that every (finite) semigroup is embeddable in an idempotent generated (finite) semigroup (see [3] for two alternative constructions). As a final observation we show that there are finite INFB idempotent generated semigroups that do not generate varieties containing \mathbf{B}_2^1 . We use a construction due to T. E. Hall (see [3]). Take an arbitrary semigroup \mathbf{S} from $\bar{\Xi}_1 \cup \bar{\Xi}_2$ with period d and therefore satisfying the identity $(xyx^2y)^d \approx (xy)^d$. Construct a Rees matrix semigroup $\mathcal{M}[\mathbf{S}, |\mathbf{S}|, |\mathbf{S}|, P]$ over \mathbf{S} with sandwich matrix satisfying $P_{0,i} = P_{i,0} = 1$ and

$\mathbf{S} = \{P_{i,j} : i, j \neq 0\}$ (here, for the sake of consistency, we have continued with our convention regarding the labelling of entries of a matrix). Then $\mathcal{M}[\mathbf{S}, |\mathbf{S}|, |\mathbf{S}|, P]$ is idempotent generated.

Proposition 4.19. *The semigroup $\mathbf{M} = \mathcal{M}[\mathbf{S}, |\mathbf{S}|, |\mathbf{S}|, P]$ as constructed above is an INFB idempotent generated finite semigroup not generating a variety containing \mathbf{B}_2^1 .*

Proof. \mathbf{M} is obviously finite and also INFB since \mathbf{S} is embedded in \mathbf{M} (for example $s \mapsto (1, s, 1)$ is an embedding) and \mathbf{S} is INFB. We show that $\mathbf{M} \models (xyx^2y)^d \approx (xy)^d$, where d is the period of \mathbf{S} . Since both sides of the identity $(xyx^2y)^d \approx (xy)^d$ start and finish with the same letter, any value these words assume in \mathbf{M} always lies within the same set $M_{i,j} = \{(i, s, j) : s \in \mathbf{S}\}$. Indeed $(xyx^2y)^d$ lies in the same subgroup of $M_{i,j}$ as $(xy)^d$. Now \mathbf{M} has period d and index 2 since \mathbf{S} has this period and index respectively. Therefore $(xyx^2y)^d$ and $(xy)^d$ are both idempotents (note that $d > 2$ since, as noted in the proof of Corollary 4.10, a group of period 2 is abelian) and therefore equal. Since \mathbf{B}_2^1 does not satisfy $(xyx^2y)^d \approx (xy)^d$ (see proof of Lemma 4.3), $\mathbf{B}_2^1 \notin \mathcal{V}(\mathbf{M})$. ■

References

- [1] Birget, C.J., S. Margolis, and J. Rhodes, *Semigroups whose idempotents form a subsemigroup*, Bull. Austral. Math. Soc. **41** (1990), 161–184.
- [2] Burris S. and H. Sankappanavar, “A Course in Universal Algebra”, Graduate Texts in Mathematics **78**, Springer Verlag, 1981.
- [3] Howie, J.M., “Fundamentals of Semigroup Theory”, London Mathematical Society Monographs, Oxford University Press, 1995.
- [4] Jackson, M., “Small Semigroup Related Structures with Infinite Properties”, PhD thesis, University of Tasmania, 1999.
- [5] Jackson, M., *Finite semigroups whose varieties have uncountably many subvarieties*, J. Algebra **228** (2000), 512–535.
- [6] Jackson, M. and O. Sapir, *Finitely based, finite sets of words*, Internat. J. Algebra Comput. **10** (2000), 683–708.
- [7] Kruse, R., *Identities satisfied in a finite ring*, J. Algebra **26** (1973), 298–318.
- [8] L’vov, I.V., *Varieties of associative rings I*, Algebra i Logika **12** (1973), 269–297.

- [9] McKenzie, R., *Equational bases for lattice theories*, Math. Scand. **27** (1970), 24–38.
- [10] McKenzie, R., *Tarski's finite basis problem is undecidable*, Internat. J. Algebra and Comput. **6** (1996), 49–104.
- [11] Murskii, V.L., *On the number of k -element algebras with one binary operation which have no finite basis of identities*, Problemy Kibernet, **35** (1979), 5–27.
- [12] Oates, S. and M.B. Powell, *Identical relations in finite groups*, J. Algebra **1** (1964), 11–39.
- [13] Perkins, P., *Bases for equational theories of semigroups*, J. Algebra **11** (1969), 298–314.
- [14] Perkins, P., *Basic questions for general algebras*, Algebra Universalis **19** (1984), 16–23.
- [15] Rasin, V.V., *Varieties of orthodox Clifford semigroups*, Russian Math. (Izv. VUZ) **26** (1982), 82–85.
- [16] Rotman, J.J., “An Introduction to the Theory of Groups”, 4th ed., Springer Verlag, 1991.
- [17] Sapir, M.V., *Problems of Burnside type and the finite basis property in varieties of semigroups*, Math. USSR Izv. **30** No. 2 (1988), 295–314.
- [18] Sapir, M.V., *Inherently nonfinitely based finite semigroups*, Math. USSR Sbornik **61** No. 1 (1988), 155–166.
- [19] Sapir, O., *Finitely based words*, Internat. J. Algebra Comput. **10** (2000) 457–480.
- [20] Shevrin, L.N., *To the theory of epigroups I*, Matem. Sbornik **185** No. 8 (1994), 129–160.
- [21] Shevrin, L.N. and M. V. Volkov, *Identities of semigroups*, Izv. Vyssh. Uchebn. Zaved. Mat. **11** (1985), 3–47.
- [22] Trahtman, A.N., *The finite basis problem for semigroups of order less than six*, Semigroup Forum **27**, 387–389.
- [23] Trahtman, A.N., *Finiteness of identity bases of 5-element semigroups*, in E.S. Lyapin (ed.), “Semigroups and Their Homomorphisms”, Russian State Pedagogical Univ., Leningrad, 1991, 76–97.

- [24] Zimin, A.I., *Blocking sets of terms*, Math. USSR Sbornik **47** (1984), 363–375.

School of Mathematics and Physics
The University of Tasmania
Hobart
Australia

now at

Department of Mathematics
La Trobe University
Victoria 3086
Australia
email: m.g.jackson@latrobe.edu.au

Received September 23, 1999
and in final form February 19, 2001
Online publication