Finite Semigroups Whose Varieties Have Uncountably Many Subvarieties

Marcel Jackson

School of Mathematics and Physics, University of Tasmania, Hobart, Tasmania 7001, Australia

Abstract

For several large classes of semigroups we provide a description of all semigroups which generate varieties with uncountably many subvarieties. These include the class of all Rees quotients of free monoids, the class of finite orthodox monoids, the class of monoids of index greater than two and the class of finite inherently not finitely based semigroups. The first example of a finite, finitely based semigroup generating a variety with uncountably many subvarieties is presented and a number of related results are obtained. All varieties found with uncountably many subvarieties contain uncountable chains of subvarieties with the same ordering as the real numbers.

Key words: Finite semigroup, variety, lattice.

1 Introduction

By a famous result of Oates and Powell [16], every finite group generates a semigroup (or equivalently a group) variety $\mathbf{V}$ with the property that $\mathbf{V}$ and every subvariety of $\mathbf{V}$ can be given by finitely many identities. Such a variety is called *hereditarily finitely based.* Since there are only countably many finite sets of identities, a hereditarily finitely based variety has only countably many subvarieties (in fact in [16] it is shown that a variety generated by a finite group has only finitely many subvarieties). This situation does not extend to semigroups in general however. In [33], Trahtman shows that the variety $\mathbf{V}(\mathbf{A_2^1})$ generated by the (regular) semigroup $\mathbf{A_2^1}$ of the matrices

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

under matrix multiplication has uncountably many subvarieties. Since any variety containing $\mathbf{A_2^1}$ also has uncountably many subvarieties, this example immediately provides a host of finite semigroups, each generating uncountably many subvarieties. Results of [24] however easily show that the identities of $\mathbf{A_2^1}$ are not finitely axiomatizable and that this is also true of any locally finite variety containing it. A locally finite semigroup or variety with this property is called *inherently not finitely based* (which we will abbreviate to INFB) and a semigroup or variety with a finite basis for its identities will be called *finitely based.* The following question naturally arises:

Question 1.1 *Does there exist a finitely based, finite semigroup generating a variety with uncountably many subvarieties?*

An important INFB semigroup is the semigroup $\mathbf{B_2^1}$ consisting of the following matrices

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

under multiplication. This semigroup was the first finite semigroup to be shown not to have a finite basis for its identities (see [18]). In [24], $\mathbf{B_2^1}$ was shown to generate a minimal INFB variety amongst those generated by finite semigroups. Indeed the class of INFB semigroups for which $\mathbf{B_2^1}$ is minimal in this sense contains many important classes of semigroups including the class of all finite semigroups with only nilpotent subgroups [24]. In particular, $\mathbf{B_2^1}$ is contained in the variety $\mathbf{V}(\mathbf{A_2^1})$ generated by $\mathbf{A_2^1}$. It is also shown in [24] that there exist finite INFB semigroups whose varieties do not contain $\mathbf{B_2^1}$. These facts make the following questions of definite interest:

Question 1.2 *(i) Does $\mathbf{B_2^1}$ generate a variety with uncountably many subvarieties?*
(ii) Is there a finite INFB semigroup not generating a variety with uncountably many subvarieties?

Note that there does exist a finite semigroup which is not finitely based (though it is not INFB) but generates a variety with only finitely many subvarieties (see [25]).

A subsemigroup or homomorphic image of a semigroup $\mathbf{S}$ generates a subvariety of $\mathbf{V}(\mathbf{S})$. In [25] however it is shown that the class of finite semigroups generating a variety with only finitely many subvarieties is not closed under direct products. Likewise we may ask the following question:

Question 1.3 *Is the class of finite semigroups generating varieties with countably many subvarieties closed under direct products?*

We will prove (using work by M. Sapir and M. Volkov in [26]) a result that can be used to answer Question 1.1 in the positive and Questions 1.2 (ii) and 1.3 in the negative (a positive answer to part (i) of Question 1.2 follows immediately from the solution to the second part of that question). As corollaries, complete descriptions of the semigroups from a number of classes that generate varieties with uncountably many subvarieties are obtained. These classes include the class of monoids with index greater than two (this result can also be obtained using a minor variation of the arguments of [26]), the class of finite orthodox monoids and the class of all Rees quotients of free monoids. It also follows that there exist pairs of finite semigroups generating varieties with only countably many subvarieties but whose direct product has uncountably many subvarieties. In fact one of these finite semigroups can be taken so that the variety it generates has a lattice of subvarieties with only 3 elements. In the final section a number of different examples are examined. In particular it is shown that the direct product of the semigroup $\mathbf{B_2} = \mathbf{B_2^1} \backslash \{1\}$ with a null semigroup with adjoined identity element generates a variety with uncountably many subvarieties.

2 Preliminaries

Let X^+ and X^* respectively be the free semigroup and free monoid over some countable, non empty set X. Elements of X^+ will be called *words* and elements of X will be called *letters*. A *possibly empty word* will mean an element from the monoid X^*. Two possibly empty words, u and v, will be said to be *graphically equal* (in symbols: $u \equiv v$) if $u = v$ in X^*. The *length* of a word w (written $|w|$) will be the number of not necessarily distinct letters appearing in w.

Definition 2.1 *(i) If x is a letter and w is a word, then $occ(x, w)$ is the number of occurrences of x in w,*
(ii) $c(w) = \{x : occ(x, w) > 0\}$ (the content of w),
(iii) a letter x is n-occurring in a word w if $occ(x, w) = n$,
(iv) a word w is n-limited if $occ(x, w) \leq n$ for all letters x.

In the special case when a letter t is 1-occurring in a word w we will say that t is a *linear letter* in w.

An *identity* is an expression of the form $u \approx v$ where u and v are words. The identity $u \approx v$ is said to be *balanced* if for all letters x, $occ(x, u) = occ(x, v)$.

Definition 2.2 *If X is an alphabet of letters then a substitution θ is a homomorphism $\theta : X^+ \to X^+$ defined by its action on the generators X. If Σ is a set of identities then an identity $u \approx v$ can be derived from Σ (in symbols: $\Sigma \vdash u \approx v$) if and only if there is a sequence of words $u \equiv u_1, u_2, \dots u_{n-1}, u_n \equiv v$ and substitutions θ_i so that $u_i \equiv u_i'\theta_i(p_i)v_i'$ and $u_{i+1} \equiv u_i'\theta_i(q_i)v_i'$ for some (possibly empty) words u_i' and v_i' and some identity $p_i \approx q_i \in \Sigma$.*

An identity $u \approx v$ in an alphabet X will be said to be *satisfied* by a semigroup $\mathbf{S}$ (in symbols: $\mathbf{S} \models u \approx v$) if for every homomorphism $\phi : X^+ \to \mathbf{S}$, $\phi(u) = \phi(v)$. Such a homomorphism will be called an *assignment*. Some important kinds of identities are those of the form $x^n \approx x^{n+m}$ for some natural numbers n and m. If n and m are the smallest numbers so that a given semigroup $\mathbf{S}$ satisfies $x^n \approx x^{n+m}$ then $\mathbf{S}$ is said to be *periodic* and to have *index* n and *period* m. In the case when $m = 1$, $\mathbf{S}$ is said to be an *aperiodic* semigroup and in the case when no such n and m exist for $\mathbf{S}$, then $\mathbf{S}$ is said to be *nonperiodic*. Clearly there are no finite nonperiodic semigroups.

By an *isoterm* for a set Σ of identities we mean a word w such that $\Sigma \vdash w \approx v$ if and only if $w \equiv v$; that is, the equivalence class of the fully invariant congruence corresponding to Σ on X^+ containing w is $\{w\}$. The word w is an isoterm for a semigroup $\mathbf{S}$ if it is an isoterm for $Id(\mathbf{S})$, the set of all identities satisfied by $\mathbf{S}$. A set Σ of identities will be said to be *closed under deletion* if both $\Sigma \vdash p \approx q \Rightarrow c(p) = c(q)$ and $\Sigma \vdash p_x \approx q_x$, where $p_x \approx q_x$ is the identity obtained by deleting every occurrence of some letter x from $p \approx q$. We will say that an identity $p \approx q$ *deletes to* or *can be deleted to* $p' \approx q'$ if there is a sequence of such deletions starting at $p \approx q$ and ending at $p' \approx q'$. A word p deletes to a word p' if $p \approx p$ deletes to $p' \approx p'$. If $\mathbf{S}$ is a monoid for which there is no word w taking the value 1 for all possible assignments (such as the word x^n does on a group of exponent n) then the set of semigroup identities satisfied by $\mathbf{S}$ is closed under deletion since assigning the element 1 to a letter in an identity is effectively the same as deleting that letter.

3 Monoids Whose Semigroup Varieties Have Uncountably Many Subvarieties

For each $n > 2$ let L_n be the word

$$y_1x_1x_2x_3x_4y_1y_2x_5y_2y_3x_6y_3 \cdots y_{n-1}x_{n+2}y_{n-1}y_nx_{n+3}x_{n+4}x_{n+5}x_{n+6}y_n.$$

The following result is proved in [26].

Lemma 3.1 *([26]) Assume that $L_n \approx w$ is a balanced identity, that w can be deleted to $y_jx_iy_j$ if and only if L_n can be deleted to $y_jx_iy_j$, and that $1 \leq i < j \leq n+6$ implies w deletes to x_ix_j. If a substitution θ exists so that $\theta(L_m)$ is a subword of w then $m = n$.*

We now use this lemma to show the following.

Theorem 3.2 *Let Σ be a set of identities closed under deletion. If xyx is an isoterm for Σ then the variety defined by Σ has uncountably many subvarieties.*

Proof: A semigroup $\mathbf{S}$ with a zero element in the signature $\{\cdot, 0\}$ satisfies an identity $u \approx 0$ exactly when it satisfies the semigroup identities $ux \approx yu \approx u$ (x and y are letters not occurring in the word u). For this reason it will be convenient to consider semigroups with zero element to be in the signature $\{\cdot, 0\}$ and satisfying the identities $x0 \approx 0x \approx 0$. This is not essential, but simplifies the arguments to be used. Let $\mathbf{V}$ be a variety defined by a set, Σ, of identities closed under deletion and for which xyx is an isoterm. If M is a subset of the natural numbers, $\mathbb{N}$, then we will take Σ_M to be the set of identities $\{L_n \approx 0 : n \in M\}$. We show that for every subset M of $\mathbb{N}$, $\Sigma \cup \Sigma_M \vdash L_n \approx 0$ if and only if $n \in M$. That is for each pair of subsets P, Q of $\mathbb{N}$, the sets of identities $\Sigma \cup \Sigma_P$ and $\Sigma \cup \Sigma_Q$ define the same subvariety of $\mathbf{V}$ if and only if $P = Q$. Since there are uncountably many subsets of the natural numbers, there are uncountably many subvarieties of $\mathbf{V}$.

Fix some set $M \subseteq \mathbb{N}$ and assume that $\Sigma \cup \Sigma_M \vdash L_m \approx 0$ for some $m \in \mathbb{N}$. By Definition 2.2 there are words $u_1, \ldots, u_n$ with $u_1 \equiv L_m$, $u_n \equiv 0$. The set Σ is closed under deletion and xyx is an isoterm for Σ, so $\Sigma \not\vdash L_m \approx 0$. Therefore we may find a smallest number k such that u_{k+1} is obtained from u_k by an application of an identity from Σ_M. Now since xyx is an isoterm for Σ the words x and xy are also isoterms for Σ. So a letter x_i is linear in u_k if and only if it is linear in L_m. Also every 2-occurring letter y_j in L_m occurs either side of a linear letter x_i, that is, L_m deletes to $y_jx_iy_j$. Since xyx is an isoterm, this happens exactly when u_k deletes to $y_jx_iy_j$ and therefore y_j is 2-occurring in u_k also. So $L_m \approx u_k$ satisfies the first two conditions of Lemma 3.1. Finally xy is an isoterm for Σ so u_k deletes to x_ix_j if $1 \leq i < j \leq n+6$ and the third condition also holds. Therefore we may apply Lemma 3.1 to the identity $L_m \approx u_k$

Now u_{k+1} is obtained from u_k by an application of an identity of the form L_i for some $i \in M$. So by Lemma 3.1, i must equal m and therefore $m \in M$ as required. $\square$

Note 3.3 *It is not immediately obvious that a variety $\mathbf{V}$ satisfying the conditions of Theorem 3.2 has an uncountable chain of subvarieties even though it does have uncountably many subvarieties. This is however true: we outline an argument from [17] (page 82). If $\mathbb{Q}$ is the set of rational numbers then for each real number r one can define A_r to be the set $\{q \in \mathbb{Q} : q < r\}$. It is easily seen that $A_{r_1} \subseteq A_{r_2}$ if and only if $r_1 \leq r_2$. Thus there is an* *uncountable chain in the lattice of subsets of* $\mathbb{Q}$ *and therefore also in the lattice of subsets of* $\mathbb{N}$*, the natural numbers. Since the lattice of subvarieties of* $\mathbf{V}$ *contains a copy of the lattice of all subsets of* $\mathbb{N}$*, the result follows. This argument applies to every example of a variety with uncountably many subvarieties to follow.*

We note also that if xx is an isoterm for a monoid $\mathbf{S}$ and $\mathbf{S}$ satisfies a nontrivial identity of the form $xyx \approx w$ then w must be a nontrivial permutation of the letters in xyx. So w is one of the words xxy or yxx. However it is shown in [21] that either of the identities $xyx \approx xxy$ and $xyx \approx yxx$ define hereditarily finitely based varieties and therefore the variety generated by $\mathbf{S}$ can have only countably many subvarieties. Since xx is an isoterm for a monoid if and only if it has index three or more we have proved the following.

Corollary 3.4 *A monoid of index three or more generates a variety with uncountably many subvarieties if and only if it does not satisfy* $xyx \approx xxy$ *or* $xyx \approx yxx$ *or equivalently if and only if it is not hereditarily finitely based.*

This corollary can also be extracted from the proof of Lemma 7 and Proposition 4 of [26]. These two results of [26] explicitly concern only nonperiodic monoids and make extensive use the fact (established elsewhere in [26]) that a nonperiodic hereditarily finitely based semigroup necessarily satisfies the implication $e^2 = e \ \& \ f^2 = f \rightarrow ef = efe$ or its dual. While this implication is not always available in the periodic case (for example the variety $\mathcal{N}$ of normal bands does not satisfy this implication and yet by results of Perkins [18] there exist hereditarily finitely based periodic semigroups of arbitrarily large index generating varieties containing $\mathcal{N}$), it has been pointed out to the author by M. Volkov (private communication) that if the condition of being nonperiodic is replaced by being a monoid of index at least three then this implication is no longer necessary and the corresponding arguments in [26] continue to hold.

The remainder of this section will be concerned with examining the many consequences of Theorem 3.2.

3.1 Small finite examples and Rees quotients of free monoids.

Finite bases for all monoids of less than 6 elements are established in [4], [5] and [32]. By examining bases of identities described in these papers, it is evident that Theorem 3.2 does not apply to any of them: all monoids of order five or less satisfy a nontrivial identity of the form $xyx \approx t(x, y)$. A basic seven element monoid with a finite basis for identities can however be constructed as follows, answering Question 1.1. We first need the following definition.

Definition 3.5 *If* W *is a language (that is, a set of words) in an alphabet* X *then let* $S(W)$ *be the monoid consisting of* 0, 1 *and all subwords of words in* W *with the multiplication* $w_1 \cdot w_2 = w_1w_2$ *if* w_1w_2 *is a subword of a word in* W *and* 0 *otherwise. Equivalently we may regard* $S(W)$ *as the Rees quotient* $X^*/I(W)$ *where* $I(W)$ *is the ideal of* X^* *containing all words that are not subwords of a word in* W*.*

Consider the monoid $S(\{aba\})$. Certainly the identities of $S(\{aba\})$ are closed under deletion and xyx is an isoterm for these identities since if $xyx \approx t(x, y)$ is an identity satisfied by $S(\{aba\})$ then by assigning a to x and b to y we have that $t(a, b) = aba$ and so

$t(x,y) \equiv xyx$. Therefore by Theorem 3.2, $S(\{aba\})$ generates a variety with uncountably many subvarieties. It is reasonably easy to show (see [28] for example) that a finite basis for the semigroup identities of $S(\{aba\})$ is the closure under deleting letters of the set of identities

$$\{xyxzx \approx xxyz,\ xyy \approx yyx,$$
$$xuyvxy \approx xuyvyx, xuyxvy \approx xuxyvy, xyuxvy \approx yxuxvy\}.$$

Let $\mathbf{S}$ be a semigroup such that the set $Id(\mathbf{S})$ of all identities satisfied by $\mathbf{S}$ satisfy the conditions of Theorem 3.2. Since xyx is an isoterm for $\mathbf{S}$, if an identity $u \approx v \in Id(\mathbf{S})$ can be deleted to an identity $u' \approx v'$ where u' is of the form aba (or a subword of this), then $u' \equiv v'$. Therefore $S(\{aba\})$ satisfies the identity $u \approx v$ and so Theorem 3.2 applies to a set Σ of identities only when $S(\{aba\})$ is contained in the variety defined by Σ. We have proved the following theorem.

Theorem 3.6 *A set of identities Σ contains a subset satisfying the conditions of Theorem 3.2 if and only if $S(\{aba\})$ is contained in the variety generated by Σ.*

Note that the semigroup obtained from $S(\{aba\})$ by removing the identity element satisfies $x_1x_2x_3x_4 \approx y_1y_2y_3y_4$ and consequently has only *finitely* many subvarieties.

If a word w contains a subword of the form xyx which is not equivalent to xxy or yxx then $S(\{w\})$ will generate a variety containing $S(\{aba\})$ and therefore have uncountably many subvarieties. This means that monoids of the form $S(W)$ which generate such varieties are likely to be very common. Indeed we have

Theorem 3.7 *Let W be a non-empty set of words. The following are equivalent:*
(i) the variety $\mathbf{V}(S(W))$ has only countably many subvarieties;
(ii) the variety $\mathbf{V}(S(W))$ has infinitely many but not uncountably many subvarieties;
(iii) $S(W) \models xyx \approx yxx$ or $S(W) \models xyx \approx xxy$;
(iv) either every word in W is, for some $n \geq 1$ and $m > 1$, of one of the forms $a_1a_2 \dots a_n$ and $a_1a_2 \dots a_{n-1}a_n^m$ or every word in W is of one of the forms $a_1a_2 \dots a_n$ and $a_1^m a_2 \dots a_{n-1}a_n$ exclusively (the a_i are distinct letters);
(v) $S(W)$ generates a hereditarily finitely based variety;
(vi) $S(\{aba\}) \notin \mathbf{V}(S(W))$.

Proof: (i)⇔(iii). Since both $xyx \approx yxx$ and $xyx \approx xxy$ define hereditarily finitely based varieties, in order to prove the equivalence of the conditions (i) and (iii) we need to show that if $S(W)$ has only countably many subvarieties then it satisfies one of these identities. By the last statement of Theorem 3.2 we need only consider the case when xx is not an isoterm for $S(W)$. If xx is not an isoterm for $S(W)$ then W contains no subwords of the form uu (where u is a word). If it does not contain a subword of the form uvu either (since xx is not an isoterm for $S(W)$, v must be a word distinct from u), then it is a collection of words of the form $a_1a_2 \dots a_n$ (where the a_i are distinct letters) and is easily seen to satisfy $xyx \approx xxy$. If W does contain a subword of the form uvu then xyx is an isoterm for $S(W)$ and so Theorem 3.2 implies $S(W)$ does not have countably many subvarieties.

(iii)⇔(v). Since $xyx \approx yxx$ and $xyx \approx xxy$ define hereditarily finitely based varieties we need only show that condition (v) implies condition (iii). This follows since a hereditarily finitely based variety necessarily satisfies condition (i) and condition (i) implies condition (iii).

(iii)$\Leftrightarrow$(iv). That condition (iv) implies condition (iii) is easily verified. Now assume that $S(W) \models xyx \approx xxy$ or $xyx \approx yxx$. So W cannot have a subword of the form uvu where $uvu \not\equiv uuv$ or vuu, since then xyx would be an isoterm. Similarly W cannot contain both a subword of the form uuv and a subword of the form $v'u'u'$ (where $uuv \not\equiv vuu$ or uvu and $v'u'u' \not\equiv u'v'u'$ or $u'u'v'$) since then both xxy and yxx are isoterms and $S(W)$ would not satisfy condition (iii). Therefore W must satisfy exactly one of the two situations described in (iv).

(i)$\Leftrightarrow$(vi). From the proof of the equivalence of conditions (i) and (iii), the monoid $S(W)$ generates a variety with uncountably many subvarieties if and only if Theorem 3.2 applies to the identities of $S(W)$. The equivalence of conditions (i) and (vi) now follows from Theorem 3.6.

To complete the proof it remains to show that $S(W)$ at least an infinity of subvarieties. Since W is non-empty, $\mathbf{V}(S(W))$ is a supervariety of $\mathbf{V}(S(\{a\}))$ where a is a single letter. It is trivial to establish that this semigroup variety is given by the identities $\{xy \approx yx, xx \approx xxx\}$. For each $n \in I\!N$ the (n-nilpotent) variety given by

$$\{x_1x_2\dots x_n \approx y_1y_2\dots y_n, xx \approx xxx, xy \approx yx\}$$

defines a distinct subvariety of $\mathbf{V}(S(\{a\}))$. The theorem is proved. $\square$

In [11] it is shown that every finite set of words W is a subset of a finite set of words U so that $S(U)$ is finitely based. Thus there are many finitely based finite semigroups generating varieties with uncountably many subvarieties (although it is shown in [10] that in a natural sense almost all monoids of the form $S(W)$ are not finitely based). The set U can also be chosen so that $S(U)$ is not finitely based (but not inherently not finitely based).

3.2 Finite INFB semigroups.

One of the main results of [23] and [24] is the following theorem. Here if e is an idempotent element of a semigroup $\mathbf{S}$ then S_e is the largest subgroup of $\mathbf{S}$ containing e and $\Gamma(S_e)$ is the final term in the upper central series of the group S_e (the *upper hypercenter* of S_e).

Theorem 3.8 *([23], [24]) Let* $\mathbf{S}$ *be a finite semigroup with period* p*. Then the following are equivalent:*
(i) $\mathbf{S}$ *is INFB;*
(ii) for every $n \in I\!N$*, the word* Z_n *is an isoterm for the identities of* $\mathbf{S}$*, where* $Z_1 \equiv x_1$ *and* $Z_n \equiv Z_{n-1}x_nZ_{n-1}$ *(these words are called Zimin words);*
(iii) for some idempotent $e \in \mathbf{S}$ *the monoid* $e\mathbf{S}e$ *is INFB;*
(iv) There is an idempotent $f \in \mathbf{S}$ *and elements* a *and* e *in the monoid* $\mathbf{M} = f\mathbf{S}f$ *such that* eae *and* $ea^{p+1}e$ *do not lie in the same coset of* $M_e/\Gamma(M_e)$*.*

We can now answer Question 1.2.

Corollary 3.9 *If* $\mathbf{S}$ *is a finite INFB semigroup then the variety* $\mathbf{V}(\mathbf{S})$ *generated by* $\mathbf{S}$ *has uncountably many subvarieties.*

Proof: Condition (iii) of Theorem 3.8 shows that an INFB semigroup $\mathbf{S}$ contains an INFB subsemigroup, $\mathbf{T}$, with identity. Condition (ii) shows that $Z_2 \equiv x_1x_2x_1$ is an isoterm for the identities of $\mathbf{T}$. Therefore by Theorem 3.2 $\mathbf{V}(\mathbf{T})$ (and consequently $\mathbf{V}(\mathbf{S})$) has uncountably many subvarieties. $\square$

In [24] it is shown that there is a unique minimum INFB semigroup contained in the variety of any INFB semigroup whose subgroups are not INFB (it is not known if there exists an INFB group). However this semigroup is infinite and so the variety it generates is not covered by Corollary 3.9. Using a modification of Theorem 3.2 the author has shown that this variety also has uncountably many subvarieties however the proof is long and will not be given here (see [10]).

3.3 Regular and orthodox semigroups.

A further class for which we can give a complete description of the finite monoids generating varieties with uncountably many subvarieties is the class of orthodox semigroups. It is convenient to first examine exactly when a finite regular semigroup is INFB.

Using the fact that $\mathbf{B_2^1} \in \mathbf{V}(\mathbf{A_2^1})$ (see comments before Question 1.2) the next lemma is a small part of the main results of [30] and [31]. (see also Lemma 1 of [24]).

Lemma 3.10 *[30] Let* $\mathbf{S}$ *is a finite monoid. If there are elements* a *and* e *of* $\mathbf{S}$ *so that* a *divides* e*,* e *is idempotent and* a^2 *does not divide* e *then* $\mathbf{B_2^1} \in \mathbf{V}(\mathbf{S})$*.*

We can now show the following.

Theorem 3.11 *If* $\mathbf{S}$ *is a finite regular semigroup with period* d *then the following are equivalent:*
(i) $\mathbf{S}$ *is INFB;*
(ii) $\mathbf{B_2^1} \in \mathbf{V}(\mathbf{S})$*;*
(iii) $S(\{a\}) \in \mathbf{V}(\mathbf{S})$*;*
(iv) $\mathbf{S} \not\models xyx \approx (xy)^{d+1}x$*.*

Proof: The implications (ii)⇒(i) and (ii)⇒(iii)⇒(iv) follow immediately since $\mathbf{B_2^1}$ is INFB, $S(\{a\}) \in \mathbf{V}(\mathbf{B_2^1})$ and $S(\{a\}) \not\models xyx \approx (xy)^{d+1}x$ for any $d > 0$. The implication (i)⇒(iv) follows since if $\mathbf{S} \models xyx \approx (xy)^{d+1}x$, the Zimin word Z_2 is not an isoterm for $\mathbf{S}$ and by Theorem 3.8, $\mathbf{S}$ is not INFB.

We now show that condition (iv) implies condition (ii). Say that the identity $xyx \approx (xy)^{d+1}x$ fails on the finite regular semigroup $\mathbf{S}$ (recall that d is the period of $\mathbf{S}$). So there are elements a and b of $\mathbf{S}$ for which $aba \neq (ab)^{d+1}a$. Since $\mathbf{S}$ is regular there is an idempotent e with $e\mathcal{R}a$ and $ea = a$. So $aba = (eabe)a$ and $(ab)^{d+1}a = (eab)^{d+1}ea = (eabe)^{d+1}a$. Now consider the monoid $\mathbf{M} = e\mathbf{S}e$. This is a regular monoid since for any element $exe \in \mathbf{M}$ with inverse x' in $\mathbf{S}$, $exe = (exe)x'(exe) = (exe)(ex'e)(exe)$. If $\mathbf{M}$ is completely regular then it satisfies $x \approx x^{d+1}$. In this case $eabe = (eabe)^{d+1}$ and therefore $aba = eabea = (eabe)^{d+1}a = (ab)^{d+1}a$, a contradiction. Therefore $\mathbf{M}$ is not completely regular and there is an element $c \in \mathbf{M}$ which does not lie in a subgroup of $\mathbf{M}$. Consider the $\mathcal{D}$-class D_c of c in $\mathbf{M}$. The principle factor of D_c is a completely 0-simple semigroup in which $c^2 = 0$ (see [9] or [3] for details). Since D_c is regular there is a non zero idempotent $f \in D_c$ so that c divides f. But $c^2 = 0$ and 0 cannot divide f so therefore by Lemma 3.10,

$\mathbf{B_2^1} \in \mathbf{V}(\mathbf{M}) \subseteq \mathbf{V}(\mathbf{S})$ as required. $\square$

We can now examine the class of finite orthodox monoids.

Corollary 3.12 *Let* $\mathbf{S}$ *be a finite orthodox monoid with period* p *and let* $\mathbf{V}(\mathbf{S})$ *be the (semigroup) variety generated by* $\mathbf{S}$*. The following are equivalent:*
(i) $\mathbf{V}(\mathbf{S})$ *has uncountably many subvarieties;*
(ii) $\mathbf{V}(\mathbf{S})$ *has infinitely many subvarieties;*
(iii) $\mathbf{V}(\mathbf{S})$ *is not hereditarily finitely based;*
(iv) $\mathbf{S}$ *is not finitely based;*
(v) $\mathbf{S}$ *is INFB;*
(vi) $\mathbf{B_2^1} \in \mathbf{V}(\mathbf{S})$*;*
(vii) $S(\{a\}) \in \mathbf{V}(\mathbf{S})$*;*
(viii) $\mathbf{S}$ *is not a union of groups;*
(ix) $\mathbf{S} \not\models x \approx x^{p+1}$*.*

Proof: That (i)⇒(ii) is trivial. The implications (ii)⇒(viii), (iii)⇒(viii) and (iv)⇒(viii) are central results of [22] and the implication (i)⇒(iii) follows since there are only countably many finite sets of identities. That (vi)⇒(i), (ii), (iii), (iv) and (v) follows from Corollary 3.9 and the fact that $\mathbf{B_2^1}$ is INFB. The implications (viii)⇒(vii), (viii)⇔(ix) and (vii)⇒(ix) are also easily established. All remaining implications follow immediately from Theorem 3.11 and the fact that a monoid satisfies $xyx \approx (xy)^{p+1}x$ if and only if it satisfies $x \approx x^{p+1}$. $\square$

So for the class of orthodox monoids the properties of being finitely based and generating a variety with only finitely many subvarieties are in fact equivalent. This result is also of interest in connection with Question 8.2 of [29] which asks if a finite Orthodox semigroup is not finitely based if and only if the variety it generates contains $\mathbf{B_2^1}$.

Note 3.13 *If* $\mathbf{S}$ *is a finite orthodox semigroup (not necessarily a monoid) which is not a union of groups then the variety* $\mathbf{V}(\mathbf{S})$ *generated by* $\mathbf{S}$ *has infinitely many subvarieties.*

This is because $\mathbf{S}$ contains a non group element a which, since $\mathbf{S}$ is regular, lies in an ideal whose principal factor is an orthodox completely 0-simple semigroup which is not a union of groups (see [3] or [9] for details). Consider the two semigroups $\mathbf{B_2}$ and $\mathbf{A_2} = \mathbf{A_2^1}\backslash\{1\}$. If $\mathbf{C}$ is a completely 0-simple semigroup that is not a union of groups then there is a subsemigroup of a quotient of $\mathbf{C}$ that is isomorphic to either $\mathbf{B_2}$ or $\mathbf{A_2}$. Since $\mathbf{B_2} \in \mathbf{V}(\mathbf{A_2})$ (see [29] for example) it must be the case that $\mathbf{B_2} \in \mathbf{V}(\mathbf{S})$ (in fact $\mathbf{A_2}$ is cannot be contained in the variety of $\mathbf{S}$ anyway since its idempotents do not form a subsemigroup). A finite basis for the semigroup identities of $\mathbf{B_2}$ has been found by A. N. Trahtman ([35]; see also [29]): it is the set

$$\{x^2 \approx x^3,\ x^2y^2 \approx y^2x^2,\ xyx \approx xyxyx\}.$$

Since every identity in this set contains a letter that occurs at least twice on both sides, they are never applicable to any identity of the form $x_1x_2\dots x_n \approx y_1y_2\dots y_n$. Thus by adjoining an identity of this form to the above set of identities, a proper subvariety of $\mathbf{V}(\mathbf{B_2})$ is obtained. Since there are infinitely many such identities and each describes a distinct variety it follows that the variety $\mathbf{V}(\mathbf{B_2})$ and consequently $\mathbf{V}(\mathbf{S})$ contains infinitely many subvarieties. Thus a finite orthodox semigroup containing a non group element always generates a variety with infinitely many subvarieties.

3.4 Joins of finitely generated, hereditarily finitely based varieties.

Examples presented in [25] show that the class of semigroups generating varieties with only finitely many subvarieties is not closed under the taking of direct products (or equivalently joins of varieties). Likewise we have the following result:

Theorem 3.14 *The class of finite semigroups each generating a variety with countably many subvarieties is not closed under direct products. Therefore the class of varieties with countably many subvarieties does not form a sublattice of the class of all varieties.*

Proof: Theorem 3.7 shows that $S(\{xyy\})$ and $S(\{xxy\})$ generate varieties with countably many subvarieties. However $S(\{xyy\}) \times S(\{xxy\})$ does not satisfy either of the identities $xxy \approx xyx$ or $yxx \approx xyx$ and the word xx is an isoterm for this monoid. So by Theorem 3.2, $S(\{xyy\}) \times S(\{xxy\})$ generates a variety with uncountably many subvarieties. $\square$

In fact the examples used in this theorem show that the join of two hereditarily finitely based varieties generated by finite semigroups can have uncountably many subvarieties. A more striking example is obtained by considering any finite group not satisfying one of the identities $xxy \approx xyx$ or $yxx \approx xyx$. As mentioned above, the semigroup variety generated by a finite group $\mathbf{G}$ has only finitely many subvarieties. If $\mathbf{G}$ does not satisfy $xxy \approx xyx$, say, then the direct product $\mathbf{G} \times S(\{baa\})$ is a monoid of index three not satisfying either of the identities $xxy \approx xyx$ or $yxx \approx xyx$. By Theorem 3.2 $\mathbf{G} \times S(\{baa\})$ generates a variety with uncountably many subvarieties (clearly if $\mathbf{G}$ did not satisfy either of the described identities then instead of $S(\{baa\})$ one may take the semigroup $S(\{aa\})$). The smallest group with this property is the symmetric group $\mathbf{S_3}$ with six elements.

In terms of subvarieties however, a quite surprisingly small example is possible. Let $\mathbf{G_3}$ be the 27 element group with presentation

$$\langle a, b, c : a^3 = b^3 = 1, cb = bc, ac = ca, ab = bac \rangle$$

([2], page 145). This group satisfies neither of the identities $xyx \approx xxy$ or $xyx \approx yxx$ since $aba = baca = baac$, $aab = abac = bacac = baacc$, and baa represent different elements of $\mathbf{G_3}$. It is also easy to establish that $\mathbf{G_3}$ can be generated by just the two elements a, b, that it is of exponent 3 and that it is nilpotent of class 2. Indeed it is the only nonabelian group of order dividing 27 that has exponent 3 (see [2]) and is in fact the free Burnside group of exponent 3 on two generators (see [8] for example). Thus every two generated group in the variety of $\mathbf{G_3}$ has order dividing 27 and therefore is either isomorphic to $\mathbf{G_3}$ or is abelian. However, since the identity $xy \approx yx$ involves just two letters, any nonabelian variety must contain a two generated nonabelian group. Therefore there are no nonabelian proper subvarieties of $\mathbf{V}(\mathbf{G_3})$. Since the only abelian variety of exponent 3 is that generated by the additive group of integers modulo 3, the lattice of subvarieties of $\mathbf{V}(\mathbf{G_3})$ is a three element chain (note that every group variety with fewer than three subvarieties is abelian). The group $\mathbf{G_3}$ also plays an important role in the examples constructed in [25].

A small, aperiodic (that is, with only trivial subgroups) example of a pair of semigroups generating hereditarily finitely based varieties whose join has uncountably many subvarieties is also possible. Let $\mathbf{L^1}$ be the left zero semigroup with adjoined identity element. The lattice of band varieties has been completely described in [1], [6] and [7], and it follows that this semigroup generates a variety with only three proper, nontrivial subvarieties (the variety of semilattices, the variety of left zero semigroups and the variety of left normal bands). Since $\mathbf{L^1}$ contains a left zero semigroup it does not satisfy the identity $xyx \approx yxx$. So the direct product $S(\{aab\}) \times \mathbf{L^1}$ is a monoid of index three not satisfying either of the identities $xxy \approx xyx$ or $yxx \approx xyx$ and therefore by Theorem 3.2 it generates a variety with uncountably many subvarieties. As seen above, the monoid $S(\{aab\})$ generates a hereditarily finitely based variety. These examples suggest the following question.

Question 3.15 *Do there exist two (finite) semigroups each generating a variety with only finitely many subvarieties whose direct product has uncountably many subvarieties?*

Note that the direct product of the semigroup $\mathbf{L^1}$ above with any finite band generates a variety with still only finitely many subvarieties (in fact from results of [22], it follows that the direct product of $\mathbf{L^1}$ with any finite group also generates such a variety; see Corollary 3.12 above).

Combining the ideas above we obtain the following theorem.

Theorem 3.16 *(i) For any semigroup* $\mathbf{S}_1$ *(finite or otherwise) there are finite semigroups* $\mathbf{S}_2$ *and* $\mathbf{S}_3$ *generating hereditarily finitely based varieties so that* $\mathbf{S}_1 \times \mathbf{S}_2 \times \mathbf{S}_3$ *generates a variety with uncountably many subvarieties.*
(ii) If $\mathbf{M}$ *is a monoid of index more than two then there is a finite group* $\mathbf{G}$ *generating a hereditarily finitely based variety with only* 3 *subvarieties so that* $\mathbf{M} \times \mathbf{G}$ *generates a variety with uncountably many subvarieties.*
(iii) If $\mathbf{M}$ *is a monoid of index less than or equal to two then either* $\mathbf{M}$ *satisfies both* $xyx \approx xxy$ *and* $xyx \approx yxx$ *or there is a finite semigroup* $\mathbf{S}$ *generating a hereditarily finitely based variety so that* $\mathbf{M} \times \mathbf{S}$ *generates a variety with uncountably many subvarieties.*

Proof of (i): For $\mathbf{S}_2$ and $\mathbf{S}_3$ one can take, for example, the semigroups $\mathbf{L^1}$ and $S(\{aab\})$ or the semigroups $\mathbf{G_3}$ and $S(\{aa\})$.
Proof of (ii): The monoid $S(\{aa\})$ is contained in the variety generated by $\mathbf{M}$ and therefore the claim follows by taking $G\mathbf{G}$ to be the group $\mathbf{G_3}$ above. To obtain an aperiodic example one may replace the group $\mathbf{G_3}$ in this argument by the direct product of $\mathbf{L^1}$ with its right dual $\mathbf{R^1}$ and obtain a similar result. The semigroup $\mathbf{L^1} \times \mathbf{R^1}$ generates a band variety with a lattice of subvarieties consisting of 13 elements.
Proof of (iii): If $\mathbf{M}$ does not satisfy one of the described identities then one of the semigroups $\mathbf{M} \times S(\{aab\})$ or $\mathbf{M} \times S(\{abb\})$ generates a variety whose identities are closed under deletion, have index three and do not contain either of the identities $xyx \approx xxy$ and $xyx \approx yxx$. By the last part of Theorem 3.2, one of these semigroups generates a variety with uncountably many subvarieties. □

Note that by a famous theorem of Clifford (see [3] or [9]) a semigroup of index one is a semilattice of completely simple semigroups. A completely simple semigroup that is not merely a group cannot not satisfy both the identities $xyx \approx xxy$ and $xyx \approx yxx$ since it contains a divisor that is isomorphic to either a left or a right zero semigroup. Therefore if

M is a monoid of index one satisfying both of these identities it is a semilattice of groups (a Clifford semigroup), each satisfying these identities.

3.5 A final example.

Let $\mathcal{S}_n$ be the semigroup variety generated by all semigroups of order n and $\mathcal{M}_n$ be the semigroup variety generated by all monoids of order n. Naturally, $\mathrm{M}_n \subseteq \mathcal{S}_n$.

Corollary 3.17 *The semigroup $\mathcal{M}_n$, and consequently $\mathcal{S}_n$, has uncountably many subvarieties for $n > 3$. For $n < 3$, $\mathcal{M}_n$ and $\mathcal{S}_n$ have at most countably many subvarieties.*

Proof: If $n \geq 4$, $\mathcal{M}_n$ contains the following: the three element monoid $\mathbf{L^1}$ consisting of the two element left zero semigroup with adjoined identity element; its right zero counterpart $\mathbf{R^1}$; and the four element monoid $S(\{aa\})$. Therefore $\mathcal{M}_n$ contains the direct product of these. Since xx is an isoterm for $S(\{aa\})$, $\mathbf{L} \not\models xyx \approx yxx$, and $\mathbf{R} \not\models xyx \approx xxy$, Theorem 3.2 now applies. Up to isomorphism there are only two, two element monoids (the two element group and the two element semilattice) and these are both commutative. There are five, two element semigroups (the two previously mentioned along with the two element null semigroup and the two element left and right zero semigroups) and it is trivial to verify that these all satisfy the identities $xyzw \approx xzyw$ and $x^2 \approx x^4$. Therefore both $\mathcal{M}_n$ and $\mathcal{S}_n$ generate hereditarily finitely based varieties and consequently have countably many subvarieties (see [18]). □

The following question remains unanswered.

Question 3.18 *Do $\mathcal{M}_3$ and $\mathcal{S}_3$ have uncountably many subvarieties?*

It can be checked that $xyx \approx xyx^7$ is an identity for both of these varieties (for a list of all semigroups of order three the reader is referred to [19]). We note that S_3 is finitely based (see [29]).

4 Further Varieties with Uncountably Many Subvarieties

By Theorem 3.6 every example of a semigroup variety with uncountably many subvarieties that was found in the previous section contains the semigroup $S(\{aba\})$. We now find two semigroup varieties with uncountably many subvarieties which do not contain this semigroup.

4.1 Varieties containing $\mathbf{B_2}$

Theorem 4.1 *If **V** is a variety containing the semigroups $\mathbf{B_2}$ and $S(\{a\})$ then **V** has uncountably many subvarieties.*

Proof: Let S be the semigroup $\mathbf{B_2} \times S(\{a\})$. Since $\mathbf{B_2}$ and $S(\{a\})$ are (up to isomorphism) subsemigroups of S, a variety **V** contains S if and only if it contains both $\mathbf{B_2}$ and $S(\{a\})$ and the semigroup S satisfies an identity $p \approx q$ exactly when both the subsemigroups $\mathbf{B_2}$ and $S(\{a\})$ satisfy $p \approx q$. It is easily verified that the semigroup $S(\{a\}) \models p \approx q$

if and only if $c(p) = c(q)$ and $occ(x,p) = 1 \Leftrightarrow occ(x,q) = 1$. As noted above, a basis for the identities of $\mathbf{B_2}$ is the set $\{x^2y^2 \approx y^2x^2, xyx \approx xyxyx, x^2 \approx x^3\}$. It is clear that $\mathbf{B_2} \models p \approx q$ implies $c(p) \approx c(q)$ and therefore $S \models p \approx q$ if and only if both $\mathbf{B_2} \models p \approx q$ and there is no letter t that is linear on one side of $p \approx q$ but nonlinear on the other. We now show that for every odd number $n > 0$ the word

$$L_n \equiv (z_1t_1t_2z_1)(x_1y_1x_1)(x_2y_2x_2)\dots(x_ny_nx_n)(z_2t_3t_4z_2)$$

is an isoterm for S (the condition of being odd here merely serves to reduce in what follows the number of cases necessary to consider). It is shown in [20] that these words are independent in the sense that for any distinct odd natural numbers n and m there is no substitution θ so that L_m contains $\theta(L_n)$ as a subword. Thus if the word L_i is an isoterm for S for all odd numbers $i > 0$ then for any two distinct subsets P and Q of the odd natural numbers, the sets $Id(S) \cup \{L_n \approx 0 : n \in P\}$ and $Id(S) \cup \{L_n \approx 0 : n \in Q\}$ define distinct varieties. This shows that $\mathbf{V}(S)$ has uncountably many subvarieties.

It will be convenient to consider the semigroup $\mathbf{B_2}$ as the semigroup on the set $\{a, b, ab, ba, 0\}$ with presentation $\langle a, b : aba = a, bab = b, aa = bb = 0\rangle$. It is clear that any word in the alphabet $\{a, b\}$ that starts with the letter a represents in $\mathbf{B_2}$ one of the words a, ab or 0 and likewise words starting with b represent one of the words b, ba or 0. The following two lemmas establish the structure of possible words r for which $\mathbf{B_2} \models L_n \approx r$.

Lemma 4.2 *If* $\mathbf{B_2} \models L_n \approx r$ *then* r *begins with the letter* z_1 *and ends with the letter* z_2.

Proof: For every number i less than n assign a to the letters x_{2i-1}, y_{2i}, t_1 and t_3, b to the letters x_{2i}, y_{2i-1} and z_i, and ba to t_2 and t_4. Call this assignment θ_1. Under θ_1, L_n takes the value $[(b)(a)(ba)(b)](aba)(bab)\dots(aba)[(b)(a)(ba)(b)] = b$. Since $\mathbf{B_2} \models L_n \approx r$, the word r must also be assigned the value b under θ_1. This shows that r cannot start with any of the letters x_{2i-1}, y_{2i}, t_1 or t_3. Let θ_2 be the same as θ_1 except with ab assigned to z_1, and b assigned to t_2. This gives L_n the value ab. This shows that r cannot start with any of the letters x_{2i}, y_{2i-1}, z_2, t_2 and t_4. Thus r starts with the letter z_1. By the symmetry of the word L_n and of the semigroup $\mathbf{B_2}$ there are dual assignments to the above that show that r must finish with the letter z_2. $\square$

Lemma 4.3 *If* $\mathbf{B_2} \models L_n \approx r$ *and* u *is a two letter subword of* r *then either* u *is a two letter subword of* L_n *or* u *is contained in the set* $\{y_1t_1,\ y_iy_{i-1},\ t_4y_n : 0 < i \leq n\}$.

Proof: Since bb and aa equal zero in the semigroup $\mathbf{B_2}$, the assignments θ_1, θ_2 and their duals above show that the only possible two letter subwords involving letters of the form x_i and y_j are $x_{2i}y_{2j}$ or reverse, $x_{2i-1}y_{2j-1}$ or reverse, $x_{2i}x_{2j-1}$ or reverse, and $y_{2i-1}y_{2j}$ or reverse. Assume that r contains the subword of the form $x_{2i}y_{2j}$ or reverse. Say $i \leq j$ and define an assignment ϕ_{2i} as follows. Assign a to all letters $x_{2i'}$ and $y_{2i'-1}$ with $i' \leq i$ and b to all letters $x_{2i'-1}$ and $y_{2i'}$ with $i' \leq i$. Assign a to all letters $x_{2j'-1}$ and $y_{2j'}$ for $j' > i+1$ and b to all letters $x_{2j'}$ and $y_{2j'-1}$ for $j' > i+1$. Assign ba to x_{2i+1} and y_{2i+1}. Since $2i$ is even and n is odd, ϕ_{2i} assigns the word

$$(x_1y_1x_1)(x_2y_2x_2)\dots(x_{2i}y_{2i}x_{2i})[x_{2i+1}y_{2i+1}x_{2i+1}](x_{2i+2}y_{2i+2}x_{2i+2})\dots(x_ny_nx_n)$$

the value

$$(bab)(aba)\dots(aba)[(ba)(ba)(ba)](bab)\dots(aba) = ba.$$

To complete the definition of ϕ_{2i}, let ϕ_{2i} assign ba to z_1 and z_2, b to t_1 and t_3 and a to t_2 and t_4. An analogous assignment for odd numbers $2i-1$ exists and we will denote this by ϕ_{2i-1}. Now ϕ_{2i} gives L_n the value ba on the semigroup $\mathbf{B_2}$. However it also assigns any word $x_{2i}y_{2j'}$ (or reverse) the value $aa = 0$ if $i < j'$. Since it has been assumed that r contains the subword $x_{2i}y_{2j}$ for $i \leq j$ and $\mathbf{B_2} \models L_n \approx r$ it must be that $i = j$. In the case when $j \leq i$ the same arguments using the substitution ϕ_{2j} instead of ϕ_{2i} again show that $i = j$. These assignments also show that if $x_{2i}x_{2j+1}$ is a subword of r then $j = i$ and that if $y_{2j+1}y_{2i}$ is a subword of r then $j = i$ (note however that there are no such subwords in L_n). Similarly, using ϕ_{2i-1} one can show that if $x_{2i-1}y_{2j-1}$ (or reverse), $x_{2i-1}x_{2j}$, and $y_{2i}y_{2j-1}$ are subwords of r then $i = j$.

Thus the only possible two letter subwords of r in the alphabet $\{x_i, y_j : 0 < i, j \leq n\}$ are those already occurring in L_n and subwords of the form y_iy_{i-1}. The arguments above are easily extended to the two letter subwords of r containing any of the letters x_i, y_i, z_i or t_i. It is routine to verify in this case that the only possible two letter subwords of r that do not already occur in L_n are those found above and the words y_1t_1 and t_4y_n. The Lemma is proved. □

Denote the set of all possible two letter subwords of r by $\mathbf{R}$ (note that not all of these subwords need occur in any particular choice of the word r). We now complete the proof of Theorem 4.1 by showing that if $S \models L_n \approx r$ then $L_n \equiv r$.

We associate with the word r a sequence of consecutive edges, or a pathway, in a directed graph $G(r)$ with vertex set $V(G(r)) = c(r) \cup \{0\}$ and edge set $E(G(r)) = \{(u, v) : uv \in \mathbf{R}\} \cup \{(0, z_1), (z_2, 0)\}$ (no duplicate edges are allowed). This graph is shown in Figure 1 (here the dotted lines represent edges corresponding to the two letter subwords contained in $\mathbf{R}$ but not occurring in the word L_n). The first edge in the pathway corresponding to r

Figure 1: The directed graph constructed for the word r.

is the edge $(0, z_1)$ and successive edges correspond to successive two letter subwords in r. That is, the i^{th} edge in this pathway corresponds to the $(i-1)^{th}$ two letter subword to occur in r. Finally, the last edge in the pathway is the edge $(z_2, 0)$. Naturally for some choices of r the corresponding pathway does not contain every edge. For example, the word L_n (which is a possible choice for r since $S \models L_n \approx L_n$ trivially) corresponds to the (unique) pathway passing every non dotted edge exactly once. If the semigroup $S \models L_n \approx r$ then all linear letters in L_n are linear in r also. Therefore for every linear letter, say t, the pathway corresponding to r contains only one edge leaving the vertex t and one entering the vertex t. We will assume that this pathway contains a dotted edge (that is, r contains a two letter subword not contained in L_n) and show that a contradiction arises.

Assume that the edge (y_i, y_{i-1}) is contained in the pathway corresponding to r and that i is the largest number with this property. Thus either the edge immediately preceding (y_i, y_{i-1}) is (x_i, y_i) or $i = n$ and the edge immediately preceding (y_i, y_{i-1}) is (t_4, y_n). Let j be the smallest number for which (y_{j+1}, y_j) is an edge succeeding (y_i, y_{i-1}) in the pathway. Therefore either the edge immediately following (y_{j+1}, y_j) is (y_j, x_j) or $j = 1$ and the edge immediately following (y_{j+1}, y_j) is (y_1, t_1). For the sake of simplicity we will only consider the cases when i does not equal n and j does not equal 1. The remaining cases follow in the same manner essentially by using z_1 and z_2 instead of x_j and x_i respectively (aside from simple arguments regarding t_2 and t_3). So r contains the subword $x_i y_i y_{i-1} y_{i-2} \dots y_{j+1} y_j x_j$. The only edges pointing left in the graph are of the form (y_k, y_{k-1}), (y_1, t_1) and (t_4, y_n). Thus if an edge of the form (y_k, x_k) is contained in the pathway corresponding to r then, since y_k is linear in r, every edge to follow can never finish at the vertex x_k. Therefore r must be of the form

$$[A] x_j x_{j+1} \dots x_{i-1} x_i y_i y_{i-1} \dots y_{j+1} y_j x_j x_{j+1} \dots x_{i-1} x_i x_{i+1} [B]$$

where A does not contain x_k or y_k for any $k \geq j$ or the letters z_2, t_3 and t_4 and B does not contain $x_{k'}$ or $y_{k'}$ for any $k' < i+1$ or the letters z_1, t_1 and t_2. Assign ab to all letters in r up to (but not including) the first occurrence of the letter x_i, assign a to x_i, ba to y_i and b to y_{i-1}. Assign ab to the letters y_k for $i-1 < k \leq j$ and ba to all other letters. Clearly (since ab and ba are idempotent in $\mathbf{B_2}$) these rules assign A the value ab and B the value ba. Thus r is assigned the value

$$[ab](ab)(ab)\dots(ab)(a)(ba)(b)(ab)(ab)\dots(ab)(ab)(ab)(ab)\dots(ab)(a)(ba)[ba] = a.$$

However L_n contains the subword $x_{i-1} y_{i-1}$ which takes the value $abb = 0$ under this assignment. Thus we have reached a contradiction.

So the pathway corresponding to r does not pass along any of the dotted edges but does pass through every vertex. Since the vertices $t_1, \dots, t_4$ and $y_1, \dots, y_n$ can be passed only once, it is easily verified that the pathway corresponding to r must be identical to that of L_n. Thus $r \equiv L_n$ as required. □

It is a routine exercise to verify that both $\mathbf{B_2}$ and $S(\{a\})$ satisfy the identity $xyxzx \approx xzxyx$ but $S(\{aba\})$ does not and therefore $S(\{aba\}) \notin \mathbf{V}(\mathbf{B_2} \times S(\{a\}))$. Similarly $\mathbf{B_2} \times S(\{a\}) \notin \mathbf{V}(S(\{aba\}))$ since $S(\{xyx\}) \models xyxy \approx yxyx$ but $\mathbf{B_2} \times S(\{a\}) \not\models xyxy \approx yxyx$.

Proposition 3 of [26] shows that if $\mathbf{V}$ is a nonperiodic variety then $\mathbf{V}$ is hereditarily finitely based only if the regular elements of every semigroup $\mathbf{S}$ in $\mathbf{V}$ lie in subgroups of

S (and that a nonperiodic semigroup containing a nongroup, regular element generates a variety with uncountably many subvarieties). The condition of nonperiodicity here serves only to ensure that certain identities are balanced. The identities in question will also be balanced if (as in the comments after Corollary 3.4) the condition of being nonperiodic is replaced by the condition of containing a monoid of index more than three. Thus the regular elements of a semigroup in a hereditarily finitely based variety containing a monoid of index at least four are all group elements. By combining the arguments used to prove Proposition 4 of [26] and the result of Theorem 4.1 above we obtain the following improvement on these results.

Corollary 4.4 *If* $\mathbf{V}$ *is a hereditarily finitely based variety containing a monoid of index greater than one (that is, a monoid that is not completely regular) then the regular elements of any semigroup* $\mathbf{S}$ *in* $\mathbf{V}$ *lie in subgroups of* $\mathbf{S}$*. On the other hand if a variety* $\mathbf{V}$ *contains a semigroup with a nongroup, regular element and* $\mathbf{V}$ *also contains a monoid of index greater than one then* $\mathbf{V}$ *has uncountably many subvarieties.*

Theorem 4.1 also provides an example of a seven element, not INFB semigroup whose identities are not closed under deletion. We may think of $\mathbf{B_2}$ and $S(\{c\})$ as sharing a single common element, the zero element (here we use the letter c in the semigroup $S(\{c\})$ to avoid confusion between elements of $S(\{c\})$ and elements of $\mathbf{B_2}$) and define a semigroup multiplication on the set $\mathbf{B_2} \cup (S(\{c\}))$ to coincide with that on the subsemigroups $\mathbf{B_2}$ and $S(\{c\})$ and to equal zero elsewhere (this construction is called the *zero direct join* of $\mathbf{B_2}$ and $S(\{c\})$).

Example 4.5 *The seven element semigroup* $\mathbf{B_2} \cup (S(\{c\}))$ *with the described multiplication generates a variety with uncountably many subvarieties.*

It is trivial to verify that this semigroup has seven elements and generates a variety satisfying the conditions of Theorem 4.1 (it generates the same variety as $\mathbf{B_2} \times S(\{c\})$). This semigroup also satisfies $xyxzx \approx xzxyx$ and so is not INFB by results of [23] since this identity implies that the word Z_3 is not an isoterm (see proof of Corollary 3.9 above). The identities of this semigroup are not closed under deletion since $xyxzx \approx xzxyx$ deletes to $yz \approx zy$ and $\mathbf{B_2}$ is not commutative. Indeed since the identity $xy \approx yx$ defines a hereditarily finitely based variety (see [18]), this argument shows that any subvariety of $\mathbf{V}(\mathbf{B_2} \times S(\{c\}))$ whose identities are closed under deletion has only countably many subvarieties.

A more extreme example is that found by Jezek in [12]. There it is shown that the variety $\mathbf{V}'$ defined by $x^2y \approx yx^2 \approx x^2$ has uncountably many subvarieties (this is the variety of all semigroups where the square of any element is the zero element). Since if 1 is the identity element of a monoid then $1^2 = 1$, it follows that if s is an element of a monoid $\mathbf{S}$ from $\mathbf{V}'$, then $s = s1^2 = 1^2 = 1$. That is, all monoids in $\mathbf{V}'$ are trivial! The variety $\mathbf{V}'$ however is not generated by a finite semigroup, indeed it contains the famous three generated infinite semigroup constructed by Morse and Hedlund [15] and so is not even locally finite.

Note also that Theorem 4.1 shows that the direct product $\mathbf{B_2}$ with any monoid of index greater than one generates a variety with uncountably many subvarieties. If $\mathbf{B_2}$ generates a hereditarily finitely based variety then an improvement of Theorem 3.16 would be obtained. Furthermore since $\mathbf{B_2}$ contains a submonoid of index one, the index described in Corollary 4.4 would be shown to be minimal. As an inverse semigroup in the signature $\{\cdot, ^{-1}\}$, $\mathbf{B_2}$ does generate such a variety [14]. On the other hand if $\mathbf{V}(\mathbf{B_2^1})$ is not hereditarily finitely based then neither can be the variety generated by any finite orthodox semigroup containing a non group element and so a complete description of the hereditarily finitely based varieties generated by finite orthodox semigroups is obtained (extending Corollary 3.12). If $\mathbf{V}(\mathbf{B_2^1})$ has uncountably many subvarieties then Corollary 3.12 would be extended further to include a complete description of the finite orthodox semigroups whose varieties have uncountably many subvarieties. This motivates the following questions.

Question 4.6 *Does* $\mathbf{B_2}$ *generate a hereditarily finitely based semigroup variety? Does* $\mathbf{B_2}$ *generate a semigroup variety with uncountably many subvarieties?*

4.2 Varieties for which xyx is not an isoterm.

The word xyx is an isoterm for both $S(\{aba\})$ and $\mathbf{B_2} \times S(\{a\})$. On the other hand a recent result of J. Kadourek [13] shows that the semigroup variety defined by the powerful identity $x^2y \approx xy$ has uncountably many subvarieties. Clearly xyx is not an isoterm for this variety. We now present a second example with this property which permits a proof along similar lines to others in this paper. However it is not known whether the example in [13] or the example below can be modified to imply the existence of *finite* semigroups whose varieties have uncountably many subvarieties. For instance, the variety defined by $x^2y \approx xy$ contains the variety of all bands and therefore by a result from [27], cannot be generated by any finite semigroup.

For every $k > 0$ let $\mathbf{V_k}$ be the variety defined by $\{xyx \approx xy^{k+1}x, xyxy \approx yxyx$. Note that while $x^2y \approx xy \vdash xyx \approx xy^{k+1}x$, the variety defined by $\{x^2y \approx xy, xyxy \approx yxyx\}$ has only countably many subvarieties since these identities imply $xyx \approx xyxyx \approx yxyxx \approx yxx$.

Proposition 4.7 *For every* $k > 0$, $\mathbf{V_k}$ *has uncountably many subvarieties.*

Proof: For every $n > 0$ let L_n be the word

$$x_1x_2x_1y_1^2y_2^2\dots y_n^2x_3x_4x_3.$$

and R_n be the word

$$x_1x_2x_1y_n^2y_{n-1}^2\dots y_1^2x_3x_4x_3.$$

Fix a subset M of $I\!N$ and let n be any number in $I\!N$. We will show that $\{xyx \approx xy^{k+1}x, xyxy \approx yxyx, L_i \approx R_i : i \in M\} \vdash L_n \approx R_n$ only if $n \in M$. As in previous proofs, this implies that the variety $\mathbf{V_k}$ has uncountably many subvarieties.

Let the set $\{xyx \approx xy^{k+1}x, xyxy \approx yxyx, L_i \approx R_i : i \in M\}$ be denoted by Σ_M and assume that $\Sigma_M \vdash L_n \approx R_n$. By Definition 2.2 we can select a number m and pairwise distinct words $u_1, u_2, \dots, u_m$ with $u_1 \equiv L_n$, $u_m \equiv R_n$ so that for each $i \leq m$, there is a substitution θ_i and an identity $p_i \approx q_i \in \Sigma_M$ so that u_{i+1} is obtained from u_i by replacing a subword of the form $\theta_i(p_i)$ in u_i with the subword $\theta_i(q_i)$. Let j be the largest number so that $\{xyx \approx xy^{k+1}x, xyxy \approx yxyx\} \vdash u_1 \approx u_j$. There are only two subwords of L_n of the form xyx and none of the form $xyxy$. Since $\{xyx \approx xy^{k+1}x, xyxy \approx yxyx\} \vdash L_n \approx u_j$ it is easily established by induction on j that for some integers $p, q \geq 0$,

$$u_j \equiv x_1x_2^{pk+1}x_1y_1^2y_2^2\dots y_n^2x_3x_4^{qk+1}x_3.$$

Because this word is not R_n it follows that $\{xyx \approx xy^{k+1}x, xyxy \approx yxyx\} \not\vdash L_n \approx R_n$ and so there exists a number $h \in M$ and a substitution θ so that $u_j \equiv r\theta(L_h)s$ and $u_{j+1} \equiv r\theta(R_h)s$. The first letter of L_h is x_1. Since x_1 is 2-occurring in L_h and $x_1x_2x_1$ is a subword of L_h, there must be a subword of u_j of the form uvu for some words u and v. By inspection, the pair (u, v) is one of the following: (x_1, x_2^{pk+1}), $(x_2^{e_1}, x_2^{f_1})$, (x_3, x_4^{qk+1}), $(x_4^{e_2}, x_4^{f_2})$ (where e_i and f_i are natural numbers satisfying $e_1 + f_1 \leq pk + 1$ and $e_2 + f_2 \leq qk + 1$). The second last of these is obviously impossible since then uvu would be a final segment of u_j but uvu must be followed in u_j by $\theta(y_1)$ since this follows $x_1x_2x_1$ in L_h. The last of the possibilities is also impossible since the only letter that occurs twice to the right of x_4 in u_j is x_4 itself. This enforces $\theta(x) = x_4^i$ for every letter $x \in c(L_h)$ (for some i depending on x) and therefore $\theta(L_h) \equiv \theta(R_h)$. In this case $u_j \equiv u_{j+1}$, contradicting both the choice of j as the largest such that $\{xyx \approx xy^{k+1}x, xyxy \approx yxyx\} \vdash L_n \approx u_j$ and the fact that u_j and u_{j+1} are distinct words. A similar argument applies for the second of the possibilities unless for some $x \in c(L_h) \backslash \{x_1, x_2\}$ the letter x_1 appears in $\theta(x)$. In this case however, there is only one occurrence of x_1 to the right of x_2 in u_j and so x must be 1-occurring in L_h. The only remaining 1-occurring letter in L_h is x_4. However then for every $i \leq h$ there is an i' so that $\theta(y_i) \equiv x_2^{i'}$. Therefore $\theta(L_h) \equiv \theta(R_h)$, once again contradicting the fact that $u_j \not\equiv u_{j+1}$. So the only remaining possibility is that $\theta(x_1) \equiv x_1$ and $\theta(x_2) \equiv x_2^{pk+1}$. The same arguments show that $\theta(x_3) \equiv x_3$ and $\theta(x_4) \equiv x_4^{qk+1}$. In this case it is easily verified that $h = n$ and $\theta(y_i) \equiv y_i$ for all $i \leq n$. Thus $n \in M$ as required.□

We finish with a number of questions.

Question 4.8 *Does* $\mathbf{A_2}$ *generate a variety with uncountably many subvarieties (see also Question 4.6)?*

Question 4.9 *Is there a finite regular semigroup generating uncountably many subvarieties that is not INFB?*

Note that a negative answer to this question would imply a negative answer to both parts of Question 4.8 and enable a generalisation of Corollary 3.12 to regular semigroups.

Question 4.10 *Is there a finite monoid generating a variety with uncountably many subvarieties for which* xyx *is not an isoterm?*

A negative answer to this question would help improve the bounds for a solution to Question 4.11 as well as provide a partial solution to Question 3.18. In connection with this question and Question 4.9 we note that there exist finitely based semigroup varieties generated by *locally finite* regular monoids for which xyx is not an isoterm; the free group in the variety of groups (considered in the semigroup signature) of exponent 16 in [36] suffices.

Question 4.11 *(i) What is the smallest (element-wise) finite semigroup (or monoid) generating a variety with uncountably many subvarieties?*
(ii) What is the smallest finitely based finite semigroup (or monoid) generating a variety with uncountably many subvarieties?

Question 4.12 *Is the membership problem for the class of finite semigroups generating varieties with uncountably many subvarieties decidable?*

Acknowledgement

The author is indebted to M. V. Volkov for suggesting the investigation of finite semigroups whose varieties have uncountably many subvarieties, for many comments and for a potential example that inspired several of the examples in this paper.

References

[1] A. P. Birjukov, Varieties of idempotent semigroups, *Algebra i Logika* **9** (1970), 255-273 (Russian).

[2] W. Burnside, “Theory of Groups of Finite Order”, Cambridge Univ. Press, 2nd ed., 1911.

[3] A. H. Clifford and G. B. Preston, “The Algebraic Theory of Semigroups,” Volumes 1,2, Amer. Math. Soc., Providence, RI, 1961, 1967.

[4] C. C. Edmunds, On certain finitely based varieties of semigroups, *Semigroup Forum* **15** (1977), 21-39.

[5] C. C. Edmunds, Varieties generated by semigroups of order four, *Semigroup Forum* **21** (1980), 67-81.

[6] C. F. Fennemore, All varieties of bands I, II, *Math. Nachr.* **48** (1971), 237-252.

[7] J. A. Gerhard, The lattice of equational classes of idempotent semigroups, *J. Algebra* **15** (1970), 195-224.

[8] M. Hall Jr., “Theory of Groups”, Chelsea Publishing Company, New York, 2nd ed., 1976.

[9] J.M. Howie, “Fundamentals of Semigroup Theory”, London Mathematical Society Monographs, Oxford University Press, 1995.

[10] M. Jackson, “Small semigroup related structures with infinite properties”, PhD thesis, University of Tasmania, Hobart (1999).

[11] M. Jackson and O. Sapir, Finitely based sets of words, to appear in *Internat. J. Algebra Comput.*.

[12] J. Jezek, Intervals in the lattice of varieties, *Algebra Universalis*, **6** (1976), 147-158.

[13] J. Kadourek, Uncountably many subvarieties of semigroups satisfying $x^2y \approx xy$, to appear in *Semigroup Forum.*

[14] E. I. Kleiman, On bases of identities of varieties of inverse semigroups, *Sibirsk. Mat. Zh.* **20** No 4 (1979), 760-777.

[15] M. Morse and G. Hedlund, Unending chess, symbolic dynamics, and a problem in semigroups, *Duke Math J.* **11** (1944), 1-7.

[16] S. Oates and M. B. Powell, Identical relations in finite groups, *J. Algebra* **1** (1964), 11-39.

[17] A. Ju. Ol'shanskii, On some infinite systems of identities, *Amer. Math. Soc. Transl.* Ser. 2 **119** (1983), 81-88.

[18] P. Perkins, Bases for equational theories of semigroups, *J. Algebra* **11** (1969), 298-314.

[19] M. Petrich, "Introduction to semigroups", Charles E. Merrill Publishing Co. (1973).

[20] G. Pollak, On identities that define hereditarily finitely based varieties of semigroups, *in* "Algebraic Theory of Semigroups", Proc. Conf. Szeged (1976), Colloq. Math. Soc. János Bolyai, North-Holland Amsterdam, **20** (1979), 447-452.

[21] G. Pollak and M. V. Volkov, On almost simple semigroup identities, *in* "Semigroups", Proc. Conf. Szeged (1981), Colloq. Math. Soc. János Bolyai, North-Holland Amsterdam, **39** (1985) 287-323.

[22] V. V. Rasin, Varieties of orthodox Clifford semigroups, *Russian Math. (Izv. VUZ)* **26** No. 7 (1982), 82-85.

[23] M. V. Sapir, Problems of Burnside type and the finite basis property in varieties of semigroups, *Math. USSR Izv.*, **30** No. 2 (1988), 295-314.

[24] M. V. Sapir, Inherently nonfinitely based finite semigroups, *Russian Acad. Sci. Sb. Math.*, **61**, No.1 (1988), 155-166.

[25] M. V. Sapir, On cross semigroup varieties and related questions, *Semigroup Forum*, **42**, (1991), 345-364.

[26] M. V. Sapir and M. V. Volkov, HFB property and structure of semigroups, *Contributions to General Algebra*, Vol.6, 303-310.

[27] O. Sapir, "Identities of finite semigroups and related questions", PhD thesis, University of Nebraska (1997).

[28] O. Sapir, Finitely based words, to appear in *Internat. J. Algebra Comput.*.

[29] L. N. Shevrin and M. V. Volkov, Identities of semigroups, *Izv. Vyssh. Uchebn. Zaved. Mat.* (1985), No. 11, 3-47 (Russian).

[30] L. N. Shevrin, To the theory of epigroups I, *Matem. Sbornik* **185**, No. 8 (1994), 129-160 (Russian).

[31] L. N. Shevrin, To the theory of epigroups II, *Matem. Sbornik* **185**, No. 9 (1994), 153-176 (Russian).

[32] A. V. Tishchenko, The finiteness of a base of identities for five element monoids, *Semigroup Forum* **20** (1980), 171-186.

[33] A. N. Trahtman, Six-element semigroup generates a variety with uncountably many subvarieties, *in* “Alg. Systems and their varieties”, Sverdlovsk (1988), 138-143 (Russian).

[34] A. N. Trahtman, Identities of a five-element 0-simple semigroup, *Semigroup Forum* **48** (1994), 385-387.

[35] A. N. Trahtman, A basis of identities of a five-element Brandt semigroup, *in* “Research in Modern Algebra”, Sverdlovsk (1981), 147-149 (Russian).

[36] M. R. Vaughan-Lee, Uncountably many varieties of groups, *Bull. London Math. Soc.* **2** (1970), 280-286.

Address:

Marcel Jackson
School of Mathematics and Physics,
University of Tasmania,
GPO Box 252c-37
Hobart 7001
Australia.
Email: marcel_j@hilbert.maths.utas.edu.au