On the finite basis problem for finite Rees quotients of free monoids

Marcel Jackson

Abstract

We examine the problem of determining when a finite Rees quotient of a free monoid has a finite basis for its identities. In [4] and [20] there is shown to be many difficulties associated with this problem but the main examples and theorems there concern the Rees quotients of free monoids on small numbers of generators. Here we extend these results to arbitrary finite generating sets and provide some considerably more general conditions on when a finite Rees quotient of a free monoid is not finitely based. We also introduce the notion of a strongly not finitely based word and construct some examples.

AMS subject classification (1991): primary 20M07; secondary 08B05.

One of the most investigated properties of varieties (equationally defined classes) generated by finite algebras is that of having finitely axiomatisable identities. For many of the "popular" classes of algebras such as the variety of groups [13], the variety of rings [7], [8], and the variety of lattices [9], the identities of every finite member are finitely axiomatisable. On the other hand there are many known examples of finite structures whose identities cannot be finitely axiomatised, even at the fringes of the "nicely behaved" classes above. These include examples amongst groups with distinguished element [1] (that is, with one particular element taken as an extra nullary operation), amongst finite loops [22] and amongst nonassociative rings [14]. There are also many known examples of finite semigroups with this property; the first two to be found are presented in [15]. In general an algebra (or variety) whose identities can be finitely axiomatised is said to be *finitely based*, otherwise it is said to be *not finitely based.*

One of the most famous problems underlying the study of identities of finite algebras is Tarski's Finite Basis Problem which asks if there is an algorithm to determine when a finite algebra is finitely based. While in general this problem has been shown to be undecidable [10], the solution is unknown when restricted to most "natural" classes containing both finitely based and not finitely based examples (an exception is the semigroup identities of inverse semigroups, where there are many not finitely based examples while a complete description of the finitely based ones can be found by combining results of [18] and [23]). In particular, there has been a significant volume of work concerning this problem within the variety of semigroups. As it happens, the problem does appear to be very complicated for semigroups: for example neither of the classes of finitely based finite semigroups or not finitely based finite semigroups are closed under the taking of subsemigroups, homomorphic images and direct products (see [4], [19], and [21] for various examples).

In this paper we investigate Tarski's Finite Basis Problem for finite semigroups that are Rees quotients of free monoids (recall that the Rees quotient S/I of a semigroup S with respect to an ideal I is the semigroup S factored by the congruence whose congruence classes are exactly I and all singletons in $S\backslash I$). These semigroups are very closely related to the syntactic monoids of finite languages in free monoids as follows. If $\mathbf{S}$ is a monoid with a subset W then we define the *discrete syntactic congruence* ρ_W of W by $(u, v) \in \rho_W$ if and only if for any $w \in W$ and $p, q \in \mathbf{S}$, $puq = w \Leftrightarrow pvq = w$; the *discrete syntactic monoid* of W in $\mathbf{S}$ is defined as $\mathbf{S}/\rho_W$. Evidently ρ_W is the largest congruence on $\mathbf{S}$ which separates W, in the sense that each individual element of W constitutes an entire congruence class of ρ_W. If W is a finite language (that is, a finite subset W of a free monoid X^*) then

$$I(W) = \{w \in X^* : pwq \notin W \text{ for all } p, q \in X^*\}$$

is the ideal of X^* consisting of all words in X^* that are not subwords of a word in W and X^*/ρ_W is easily seen to be the Rees quotient $X^*/I(W)$. Conversely it is easily seen that any finite Rees quotient of a free monoid is of this form. We denote the monoid X^*/ρ_W by $S(W)$. We can also visualise $S(W)$ as being the semigroup whose elements are 0, 1 (with their usual multiplicative properties) along with all subwords w of words in W with the multiplication $w_1 \cdot w_2 = w_1 w_2$ if $w_1 w_2$ is a subword of a word in W, and 0 otherwise.

If the language $W = \{w\}$ consists of a single word then any proper congruence on $S(W)$ collapses the element w onto 0 and ρ_W is the same as the *syntactic congruence* $\sim_W$ (where $u \sim_W v$ if and only if for every pair of words s, t from X^*, $sut \in W \Leftrightarrow svt \in W$; see [16]). In other words, $S(W)$ is the syntactic monoid $X^*/\sim_W$ of W. While in general $S(W)$ is not always the syntactic monoid of the language W, if V is a subset of a monoid $\mathbf{S}$ then V/ρ_V is a subset of $\mathbf{S}/\rho_V$ that consists of singleton congruence classes of $\mathbf{S}$ and the syntactic monoid of V in $\mathbf{S}$ is isomorphic to the syntactic monoid of V/ρ_V in $\mathbf{S}/\rho_V$. In other words, the syntactic monoid of a subset V in a semigroup $\mathbf{S}$ is a homomorphic image of the discrete syntactic monoid of V in $\mathbf{S}$. In keeping with the notation of [4] and [20] we will say that a language W (or a single word w) is finitely based if the discrete syntactic monoid $S(W)$ (or $S(\{w\})$ respectively) is finitely based. Otherwise, W (or w respectively) will be said to be not finitely based.

The identities of discrete syntactic monoids of finite languages have been of interest since P. Perkins [15] showed that $\{abcba, acbab, abab, aab\}$ was not finitely based (as noted above, [15] contains the first known examples of finite semigroups without a finite basis of identities). In two recent papers, by O. Sapir and the author [4], and by O. Sapir [20] it has been shown that such examples turn out to be quite abundant, the simplest example (in terms of the number of elements in the corresponding discrete syntactic monoid) being the word $abab$ (see [4]).

The finite basis problem for finite languages appears to be particularly complicated. For example, it is shown in [4] that for every finite language W one can find a chain $W = V_0 \subseteq V_1 \subseteq V_2 \subseteq \dots$ of finite languages with the property that V_i (with $i > 0$) is finitely based if and only if i is even (recently Mashevitzky has obtained a similar result to this for the class of completely 0-simple semigroups [11]). Furthermore if W contains more than one distinct letter then the alphabet of each V_i can be chosen to be the same as for W. We note that since the syntactic monoid $X^*/\sim_W$ of a finite language W is a homomorphic image of $S(W)$, it follows that for every syntactic monoid of a finite language, there exists an infinite chain of finitely generated varieties containing $X^*/\sim_W$ which alternate between being finitely based and not finitely based.

Another indication of the complicated nature of the identities of these semigroups is that one can find finite languages (even in a two letter alphabet) W_1, W_2, W_3 and W_4 so that W_1 and W_2 are finitely based, W_3 and W_4 are not finitely based while $W_1 \cup W_2$ and the direct product $S(W_1) \times S(W_2)$ are not finitely based and $W_3 \cup W_4$ and $S(W_3) \times S(W_4)$ are finitely based [4].

This complicated behaviour reinforces the motivation for Question 7.1 of [21] (due to M. Sapir; originally posed in 1976 [20]) which essentially asks for a description of the finitely based finite languages. The real interest in Question 7.1 of [21] is clearer when it is rephrased in terms of decidability:

Question 0.1 *(M. Sapir) Is there an algorithm that determines, when given a finite language W, whether or not $S(W)$ has a finite basis of identities?*

The quite fickle nature of the finite basis property for these semigroups coupled with their simple description make them a suitable candidate for studying the potential undecidability of Tarski's Finite Basis Problem for finite semigroups.

For languages consisting of one word in a two letter alphabet a complete solution to Tarski's Finite Basis Problem has been obtained by O. Sapir [20]: a word in such an alphabet is finitely based if and only if it is equivalent up to a change of letter names to one of the words $a^n b^m$ or $a^n b a^m$ for some $n, m \geq 0$.

As examples described above indicate, for languages of more than one word in a two letter alphabet, the problem of finding a finite basis for the corresponding discrete syntactic monoid appears to be significantly more complicated. Furthermore in many cases the discrete syntactic monoid of a language of more than one word may be equationally equivalent to (that is, satisfies the same identities as) the discrete syntactic monoid of a single word language in an enlarged alphabet (examples can be found in [4] and below). Thus increasing the number of distinct letters allowed in a language also appears to increase the complexity of the finite basis problem for the corresponding discrete syntactic monoid.

The methods used in [4] and [20] for investigating when a discrete syntactic monoid of a finite language is not finitely based do not seem to be particularly well suited to languages involving larger numbers of distinct letters. In this paper we address this deficiency by establishing theorems regarding not finitely based monoids which can be applied to the discrete syntactic monoids of many finite languages in arbitrary alphabets. The theorems are then used to construct some special subwords whose presence in suitable languages guarantees that the language is not finitely based. These subwords are in turn used to show that every word is a subword of a not finitely based word whose length is at most four letters longer than the original and it is shown that the variety generated by a syntactic monoid of a finite language has an infinite chain of supervarieties, each generated by a syntactic monoid of a one word language whose identities alternate between being finitely based and not finitely based (a syntactic monoid variant of a result from [4] mentioned above). We also show that there are words w in a three letter alphabet $\{x, y, z\}$ such that any word in the alphabet $\{x, y, z\}$ that contains w as a subword is not finitely based.

Some of the results below have appeared in the author's PhD thesis [3] or are announced in the survey article [24].

1 Preliminaries.

The *free monoid* and *free semigroup* on an alphabet X will be denoted X^* and X^+ respectively. Elements of X^+ will be referred to as *words* and elements of X^* will be referred to as *possibly empty words*. The equality relation on a free monoid will be denoted "$\equiv$" and the *length* of a word w (denoted by $|w|$) will be the number of (not necessarily distinct) letters appearing in w. Likewise, if $W = \{w_1, \dots, w_n\}$ is a finite language then the length of W is the maximum of the lengths of the words $w_1, \dots\ w_n$.

Definition 1.1 *(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) if $m > n$ then a letter x is more than n-occurring *in a word w if $occ(x, w) = m$,*
*(v) a word w is n-*limited *if $occ(x, w) \leq n$ for all letters x,*
(vi) If w is a word with $occ(x, w) \geq n$ then ${}_nx$ denotes the n^{th} occurrence of x in w.

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. Throughout this paper, letters of the form t_i will always be used to denote a letter that is linear in a word. Occasionally letters of the form s_i will also be used in this way.

Several of these definitions may also be extended to finite languages. In particular, if $W = \{w_1, \dots, w_m\}$ is a finite language then $c(W) = \cup_{i=1}^{m} c(w_i)$ and W is said to be n-limited if w_i is n-limited for every $i \leq n$.

An *identity* is a formal expression $u \approx v$ where u and v are words. A semigroup $\mathbf{S}$ will be said to *satisfy* $u \approx v$ (written $\mathbf{S} \models u \approx v$) if for every homomorphism θ from X^+ into $\mathbf{S}$, $\theta(u) = \theta(v)$. A set of identities will be said to be satisfied by a semigroup if every identity in the set is satisfied by the semigroup. The set of all identities in some fixed countably infinite alphabet satisfied by a semigroup $\mathbf{S}$ (or class of semigroups K) will be denoted by $Id(\mathbf{S})$ (or $Id(K)$ respectively).

Some of the definitions regarding the structure of words carry over to identities in a natural way.

Definition 1.2 *A letter is n-occurring in an identity $u \approx v$ if it is n-occurring in both u and v. The identity $u \approx v$ is n-limited if both u and v are n-limited.*

If Σ is a set of identities then we will say that $u \approx v$ can be *derived* from Σ (written $\Sigma \vdash u \approx v$) if there is a sequence of words $u \equiv u_1, u_2, \dots u_{n-1}, u_n \equiv v$ in an alphabet X and homomorphisms $\theta_i : X^+ \to X^+$ so that $u_i \equiv u_i' \theta(p_i) v_i'$ and $u_{i+1} \equiv u_i' \theta(q_i) v_i'$ for some (possibly empty) words u_i' and v_i' and some identity $p_i \approx q_i \in \Sigma$. The homomorphisms θ_i are called *substitutions* and the number $n-1$ is called the *length* of the derivation of $u \approx v$ from Σ.

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 $p' \approx p'$.

Definition 1.3 *If $c(w) = \{x_1, \dots, x_n\}$ and $\{x_{i_1}, x_{i_2}, \dots, x_{i_m}\}$ is a subset of $c(w)$ then $w(x_{i_1}, x_{i_2}, \dots, x_{i_m})$ is the word obtained from w by deleting every occurrence of the letters in $c(w) \backslash \{x_{i_1}, x_{i_2}, \dots, x_{i_m}\}$ from w.*

The identities of a monoid that is not a group are always closed under deletion (however a group of exponent n satisfies $x^n y \approx y$ so its identities are not closed under deletion according to the above definition).

The concept of deletion enables one to define a notion of stability for pairs of letters in an identity.

Definition 1.4 *[4], [20] A pair (x, y) of letters is* stable *in an identity $u \approx v$ containing x and y if $u(x, y) \equiv v(x, y)$. In this case (x, y) is called a* stable pair *in $u \approx v$. Otherwise (x, y) is said to be an* unstable pair *in $u \approx v$. If, for a given semigroup $\mathbf{S}$ and a given word u, a pair (x, y) is stable in every identity $u \approx v$ satisfied by $\mathbf{S}$ then (x, y) is said to be a* stable pair in u with respect to $\mathbf{S}$.

We now define a concept (due to Perkins [15]) that is of fundamental importance in arguments to follow.

Definition 1.5 *A word p is an* isoterm *relative to a set of identities Σ if $\Sigma \vdash p \approx q \Rightarrow p \equiv q$, that is, if the equivalence class of p under the fully invariant congruence corresponding to Σ is simply $\{p\}$. When referring to a specific semigroup $\mathbf{S}$, a word will be said to be an isoterm for $\mathbf{S}$ if it is an isoterm for $Id(\mathbf{S})$, the set of all identities satisfied by $\mathbf{S}$ over some fixed countable alphabet.*

Clearly every subword of an isoterm for a semigroup $\mathbf{S}$ is also an isoterm for $\mathbf{S}$. Furthermore, if $u(x, y)$ is an isoterm for $\mathbf{S}$ then (x, y) is stable in u with respect to $\mathbf{S}$.

2 Not finitely based monoids.

In this section we establish some theorems concerning when a monoid may be not finitely based.

To begin with we need some results concerning isoterms for monoids.

Definition 2.1 *If w is a word then a* block *of w is a maximal subword of w between two consecutive linear letters (if these exist).*

The following is a result from [4]. The proof is quite elementary, involving only the deletion properties of monoids.

Lemma 2.2 *[4] Let $\mathbf{S}$ be a monoid such that xy is an isoterm of $\mathbf{S}$ and let u be an isoterm of $\mathbf{S}$ containing a linear letter t_1. Then*
(i) erasing a prefix (suffix) of a block in u gives a new isoterm for $\mathbf{S}$ and
(ii) the word v obtained by adding a linear letter t_2 immediately to the left (or right) of the occurrence of t_1 in u is also an isoterm.

The next lemma is also useful.

Lemma 2.3 *Let w be an isoterm for a monoid $\mathbf{S}$ and X be a subset of $c(w)$. If we replace all maximal subwords of w not containing a member of X by linear letters, then the resulting word is also an isoterm for $\mathbf{S}$.*

Proof: Let v be the word obtained from w by replacing all maximal subwords of w not containing a member of X by linear letters and let $v \approx v'$ be an identity satisfied by $\mathbf{S}$ (we may assume that the new linear letters are not contained in X). Consider the assignment θ which assigns each letter in X to itself and every other linear letter in v to that part of w which it replaced, that is, so that $\theta(v) \equiv w$. It follows that if $v \not\equiv v'$ then $\theta(v) \not\equiv \theta(v')$. Since $\mathbf{S} \models w \equiv \theta(v) \approx \theta(v')$ this contradicts the fact that w is an isoterm. Therefore $v' \equiv v$ and v is an isoterm for $\mathbf{S}$ as required. □

Definition 2.4 *If w is a word containing the letters a and b then let $\tilde{w}$ be the word obtained from w by replacing all maximal subwords of w not containing a or b by linear letters and replacing all subwords of the form ab by words of the form asb, where s is a linear letter. For example, the word $abcddbbcbababd$ becomes $abt_1bbt_2bababt_3$ and then $as_1bt_1bbt_2bas_2bas_3bt_3$.*

Lemma 2.5 *Let w be a word containing at least two distinct letters a, b. If w is an isoterm for a monoid $\mathbf{S}$ then so is $\tilde{w}$.*

Proof: Let $\mathcal{T}$ be the set of linear letters replacing maximal subwords of w not containing a or b and $\mathcal{S}$ be the set of linear letters introduced when replacing ab by asb. By Lemma 2.3, the pairs (a, t), (b, t) and (a, b) are stable in $\tilde{w}$ with respect to $\mathbf{S}$ if t is from $\mathcal{T}$. Because w contains at least two letters it must contain a subword of the form xy where x and y are distinct letters and therefore xy is an isoterm for the monoid **S**. Thus if t_1 and t_2 are linear letters then the pair (t_1, t_2) is stable in $\tilde{w}$ with respect to **S**. It remains to show that (a, s) and (b, s) are stable pairs in $\tilde{w}$ if s is a linear letter from $\mathcal{S}$.

The pair (a, s) is stable in $\tilde{w}$ because we can assign a to a, 1 to b, maximal subwords of the form b^k to corresponding linear letters s from $\mathcal{S}$ and the remaining subwords of w can be assigned to corresponding linear letters from $\mathcal{T}$. The pair (b, s) is stable in $\tilde{w}$ because we can assign b to b, 1 to a, maximal subwords of the form a^k to corresponding linear letters s from $\mathcal{S}$ and the remaining subwords of w to corresponding linear letters from $\mathcal{T}$. $\square$

One final definition is required before we present the main theorems of this section.

Definition 2.6 *Let* $X = \{x_1, x_2, \dots\}$*. Then* $[Xn]$ *and* $[nX]$ *denote the words* $x_1 x_2 \dots x_n$ *and* $x_n x_{n-1} \dots x_1$ *respectively. Also* $[\mathcal{X}(2n)]$ *denotes the word*

$$x_2 x_4 \dots x_{2n} x_1 x_3 \dots x_{2n-1}.$$

We now obtain our first general theorem concerning not finitely based monoids. The proof is a modified and generalised version of that used by O. Sapir in [20] to describe when a monoid $S(\{w\})$ is not finitely based if w is a word involving just two distinct letters.

Theorem 2.7 *Let*

$$w \equiv w_1 a^{\alpha_1} b^{\beta_1} w_2 a^{\alpha_2} p b^{\beta_2} w_3$$

be a word such that a *and* b *are letters,* p*,* w_1*,* w_2 *and* w_3 *are possibly empty subwords and* α_1*,* β_1*,* α_2 *and* β_2 *are non zero and maximal. If both* w *and* $xytyx$ *are isoterms for a monoid* **S** *and for every* $n \in \mathbb{N}$ *the word*

$$u_n \equiv \tilde{w}_1 a^{\alpha_1} [Xn] b^{\beta_1 - 1} \tilde{w}_2 a^{\alpha_2} t [nX] t b^{\beta_2} \tilde{w}_3$$

is not an isoterm for **S***, then* **S** *is not finitely based.*

Proof: The first part of the proof is essentially the same as a corresponding portion of [20]; we include it here for completeness. Let X be the alphabet $\{x_1, x_2, \dots\}$ and let $u_n \approx v_n$ be a nontrivial identity satisfied by **S**. We will show that within $Id(\mathbf{S})$, identities involving at least $n-6$ distinct letters are required to derive the identity $u_n \approx v_n$ for any given $n \in \mathbb{N}$. Thus no finite basis for **S** can exist, since such a basis would necessarily involve identities with a bounded number of letters. We may assume that $n > 6$.

Let Σ be a set of identities that contain less than $n-6$ distinct letters and let $u_n \equiv p_1 \approx p_2 \approx \dots \approx p_m \equiv v_n$ be a derivation of $u_n \approx v_n$ from Σ (we may assume that $p_i \not\equiv p_{i+1}$). So there is an identity $u \approx v$ and words A and B such that $p_1 \equiv A\theta(u)B$ and $p_2 \equiv A\theta(v)B$ for some substitution θ. Replace the word u in the identity $u \approx v$ by the word $t_1 u t_2$ and v by the word $t_1 v t_2$ where t_1 and t_2 are new linear letters and extend θ by letting $\theta(t_1) \equiv A$ and $\theta(t_2) \equiv B$. So we have a derivation of $u_n \approx v_n$ from $\Sigma \cup \{t_1 u t_2 \approx t_1 v t_2\}$ involving at most $n-4$ letters such that $p_1 \equiv \theta(t_1 u t_2)$ and

$p_2 \equiv \theta(t_1 v t_2)$. For the sake of simplicity, we will write simply u in place of $t_1 u t_2$ and v in place of $t_1 v t_2$.

Now let u' be the smallest subword of u such that $\theta(u)$ contains $[Xn]$ and u'' be the smallest subword of u such that $\theta(u'')$ contains $[nX]$. Let t be the first letter in u'. By the choice of u', $\theta(t)$ must contain x_1, the first letter of $[Xn]$. If $\theta(t)$ also contains the letter to the left of x_1 (in this case the letter a) then t must be linear in u_n, since ax_1 occurs just once in u_n. In this case, say where $\theta(t) \equiv z_1 x_1 z_2$ for some words z_1 and z_2 (with z_1 not empty), we can replace the letter t in u and v by the word $t_3 t_4$ where $\theta(t_3) \equiv z_1$ and $\theta(t_4) \equiv x_1 z_2$. Thus we can find a derivation of $u_n \approx v_n$ involving less than $n-3$ letters and such that $[Xn]$ is an initial segment of $\theta(u')$ (where u' is the smallest subword of u such that $\theta(u')$ contains $[Xn]$). Performing the same procedure for the end of $[Xn]$ and the start and end of $[nX]$, we can find a derivation of $u_n \approx v_n$ involving less than n letters and such that $u_n \equiv p_1 \equiv \theta(u)$, $p_2 \equiv \theta(v)$ and the smallest subword of u whose image under θ contains $[Xn]$ is assigned by θ the value $[Xn]$ and likewise for $[nX]$. We will continue with the convention that u' and v' are the smallest subwords of u so that $\theta(u') \equiv [Xn]$ and $\theta(v') \equiv [nX]$.

Since every letter in the set X occurs exactly twice in u_n or not at all, the letters occurring in u' and v' do not occur elsewhere in the word u. So because $u_n \equiv p_1 \approx p_2$ is a nontrivial identity, the word

$$I \equiv \tilde{w}_1 a^{\alpha_1} u' b^{\beta_1 - 1} \tilde{w}_2 a^{\alpha_2} t_1 v' t_2 b^{\beta_2} \tilde{w}_3$$

is not an isoterm since we can easily apply the identity $u \approx v$ to it. Our goal is to show that this contradicts the claim that w is an isoterm, thereby showing that $\mathbf{S} \not\models \Sigma$.

Firstly, since xy is an isoterm for $\mathbf{S}$, any pair of linear letters (s, t) is a stable pair in I with respect to $\mathbf{S}$. Secondly since $xytyx$ is an isoterm for $\mathbf{S}$, $[Xn]t[nX]$ is an isoterm for $\mathbf{S}$ and therefore by the choice of u' and v', the word $u'tv'$ is also an isoterm for $\mathbf{S}$. Because $u \approx v$ involves fewer than n distinct letters but $[Xn]$ and $[nX]$ each have n distinct letters, both u' and v' must contain letters t_3 and t_4 respectively such that for some $i \leq n-1$, one of $\theta(t_3)$ and $\theta(t_4)$ contains the letters x_i and x_{i+1} and the other contains at least one of the letters x_i and x_{i+1}. However every subword of $[Xn]t[nX]$ with length more than 1 occurs just once in u_n and $[Xn]t[nX]$ is 2-limited. Therefore the letters t_3 and t_4 must be linear in I. That is, both u' and v' contain a linear letter. Now if

$$w \equiv w_1 a^{\alpha_1} b^{\beta_1} w_2 a^{\alpha_2} p b^{\beta_2} w_3$$

is an isoterm for $\mathbf{S}$ then

$$\tilde{w} \equiv \tilde{w}_1 a^{\alpha_1} t_1 b^{\beta_1} \tilde{w}_2 a^{\alpha_2} \tilde{p} b^{\beta_2} \tilde{w}_3$$

is an isoterm for $\mathbf{S}$ by Lemma 2.5. (Here for the sake of simplicity we are assuming that p contains at least one subword of the form ab or a letter other than a or b so that $\tilde{p}$ contains a linear letter. The only other case is when p is of the form $b^j a^k$ for some $j, k \geq 0$ and then we can replace $\tilde{p}$ in the above word by $t_2 b^j a^k t_3' \equiv t_2 p t_3'$ without effecting the arguments to follow.) By Lemma 2.2 (i) the words

$$\tilde{w}_1 a^{\alpha_1} t_1 b^{\beta_1 - 1} \tilde{w}_2 a^{\alpha_2} \tilde{p} b^{\beta_2} \tilde{w}_3$$

and

$$I_1 \equiv \tilde{w}_1 a^{\alpha_1} t_1 b^{\beta_1 - 1} \tilde{w}_2 a^{\alpha_2} t_2 b^{\beta_2} \tilde{w}_3$$

are also isoterms for $\mathbf{S}$. Therefore the pairs (a, b), (a, t), (b, t) are stable in I with respect to $\mathbf{S}$ whenever t is a linear letter. If both u' and v' consist entirely of linear letters then I would be just the word I_1 with some extra linear letters placed next to existing linear letters in I_1 and therefore an isoterm by Lemma 2.2 (ii), a contradiction. So let us assume there is a letter z that is 2-occurring in $u'v'$ (all nonlinear letters in $u'v'$ are 2-occurring). To obtain the desired contradiction, it only remains to show that all linear letters t in I (not just ones that appear in $u'tv'$) and the letters a and b form stable pairs with every non linear letter z, from $u'v'$.

Let z be a 2-occurring letter in $u'v'$. For some linear letter t, I can be deleted to ztz, an isoterm for $\mathbf{S}$. If (z, s) is not stable in I for some linear letter s, then $\mathbf{S}$ must satisfy an identity $I \approx J$ with $I(s, z) \approx J(s, z)$ being the identity $zzs \approx szz$ (since zsz is an isoterm). But then $I(z, s, t) \approx J(z, s, t)$ is the identity $ztzs \approx sztz$ and so $I(s, t) \approx J(s, t)$ is the identity $st \approx ts$. This identity is not satisfied by $\mathbf{S}$ since xy is an isoterm for $\mathbf{S}$. Thus for any linear letter t in I, (z, t) is stable in I with respect to $\mathbf{S}$.

Now there is at least one linear letter in both u' and v' (say t_3 and t_4 respectively) and the linear letter t_1 occurs between u' and v' in I. Since there exists a substitution θ such that $\theta(u') \equiv [Xn]$ and $\theta(v') \equiv [nX]$, we can choose t_3 and t_4 such that $u't_1v'$ deletes to a word of the form $zt_3t_1t_4z$ or $t_3zt_1zt_4$. Thus I can be deleted to either

$$\tilde{w}_1 a^{\alpha_1} z t_3 b^{\beta_1 - 1} \tilde{w}_2 a^{\alpha_2} t_1 t_4 z t_2 b^{\beta_2} \tilde{w}_3$$

or

$$\tilde{w}_1 a^{\alpha_1} t_3 z b^{\beta_1 - 1} \tilde{w}_2 a^{\alpha_2} t_1 z t_4 t_2 b^{\beta_2} \tilde{w}_3$$

The following assignment shows that (a, z) is also a stable pair in the first of the words: $a \to a$, $b \to 1$, $z \to b$, $t_3 \to b^{\beta_1 - 1}$, $t_1 \to p$, $t_4 \to 1$, $t_2 \to b^{\beta_2 - 1}$, and the remaining linear letters are assigned the corresponding unassigned (by the above) subwords of w. This gives the first word a value that is a subword of w and therefore an isoterm.

It now follows that (b, z) is stable in the first of these words also since for any linear letter t, (b, t), (z, t) as well as (a, z) are stable pairs in I and between every occurrence of b and an occurrence of z there is either a linear letter or an occurrence of a.

The following assignment shows that (b, z) is also a stable pair in the second of the words: $a \to 1$, $b \to b$, $z \to a$, $t_3 \to a^{\alpha_1 - 1}$, $t_1 \to a^{\alpha_2 - 1}$ (or $ba^{\alpha_2 - 1}$ if w_2, and therefore $\tilde{w}_2$, is empty), $t_4 \to p$, $t_2 \to 1$, and the remaining linear letters are assigned the corresponding unassigned (by the above) subwords of w. This gives the second word a value that is a subword of w and therefore an isoterm.

It now follows that (a, z) is a stable pair in the second of these words also for similar reasons to the case of (b, z) for the first word.

Since every pair of letters from $c(I)$ is stable in I with respect to $\mathbf{S}$ it must be an isoterm for $\mathbf{S}$. We have reached the desired contradiction and thus no finite basis can exist for the identities satisfied by $\mathbf{S}$. □

Note that the proof of Theorem 2.7 holds equally well if we replace the word apb

in the statement of the theorem with bpa along with the requirement that $\tilde{p}$ contains a linear letter or equivalently, that p contains either the subword ab or a letter other than a and b (we then require that for every n,

$$\tilde{w}_1 a^{\alpha_1}[Xn]b^{\beta_1-1}\tilde{w}_2 b^{\beta_2}t_1[nX]t_2 a^{\alpha_2}\tilde{w}_3$$

is not an isoterm). The proof also holds (after making the obvious adjustments) if the order of appearance of the two subwords ab and apb (or bpa) is reversed in the word w.

We now introduce a further definition in the style of Definition 2.4.

Definition 2.8 *If w is a word then let $\ddot{w}$ be the word obtained from w by replacing every maximal subword not containing the letter a by a linear letter. For example if $w \equiv abcbabb$ then $\ddot{w} \equiv at_1at_2$.*

A second general theorem is the following.

Theorem 2.9 *(a) Let $w \equiv w_1u_1au_2w_2u_2au_1w_3$ be a word where a is a letter, u_1 and u_2 are non empty subwords and w_1, w_2 and w_3 are possibly empty subwords. If w is an isoterm for a monoid* **S** *and for every n the word*

$$r_n \equiv \ddot{w}_1 t_1[Xn]at_2\ddot{w}_2 t_3[nX]at_4\ddot{w}_3$$

is not an isoterm, then **S** *is not finitely based.*
(b) Let $w \equiv w_1u_1au_2w_2u_1au_2w_3$ be a word where a is a letter, u_1 and u_2 are non empty subwords and w_1, w_2 and w_3 are possibly empty subwords. If w is an isoterm for a monoid **S** *and for every n the word*

$$g_n \equiv \ddot{w}_1 t_1[X2n]at_2\ddot{w}_2 t_3 a[\mathcal{X}2n]t_4\ddot{w}_3$$

is not an isoterm, then **S** *is not finitely based.*

Proof: (a) By Lemma 2.3 and Lemma 2.2 (ii), $\ddot{w}_1 t_1 a t_2 \ddot{w}_2 t_3 a t_4 \ddot{w}_3$ is an isoterm for **S**. Therefore for any linear letter t, (a, t) is stable in r_n with respect to **S**. By assigning u_1 to x, a to y and $u_2w_2u_2$ to t in the word $xytyx$ we obtain the word $u_1au_2w_2u_2au_1$, an isoterm for **S**. Thus $xytyx$ and consequently $[Xn]t[nX]$ are isoterms for **S**. This combined with the fact that t_1t_2 is an isoterm shows that for any linear letter t, (x_i, t) is a stable pair in r_n with respect to **S**. Therefore if $r_n \approx r'_n$ is a nontrivial identity satisfied by **S**, then the only unstable pairs in $r_n \approx r'_n$ are of the form (x_i, a).

Now by the choice of u_1 and u_2, we also have that

$$\ddot{w}_1 t_1 x a t_2 \ddot{w}_2 t_3 a x t_4 \ddot{w}_3$$

and

$$\ddot{w}_1 t_1 a y t_2 \ddot{w}_2 t_3 y a t_4 \ddot{w}_3$$

are isoterms for $\mathbf{S}$. So if for some i, (x_i, a) is unstable in $r_n \approx r'_n$ then (x_i, a) is unstable in $r_n \approx r'_n$ for all i and $r'_n(a, x_i, \tau) \equiv \ddot{w}_1 t_1 a x_i t_2 \ddot{w}_2 t_3 a x_i t_4 \ddot{w}_3$, where τ is a list of all linear letters in r'_n. Thus r'_n must be the word

$$\ddot{w}_1 t_1 a[Xn] t_2 \ddot{w}_2 t_3 a[nX] t_4 \ddot{w}_3.$$

Therefore any basis for $Id(\mathbf{S})$ must contain an identity $p \approx q$ with $r_n \equiv A\theta(p)B$ and $r'_n \equiv A\theta(q)B$ for some words A and B and a substitution θ. Since the only unstable pairs in $r_n \approx r'_n$ are of the form (x_i, a), we may assume that $\theta(p)$ contains both $[Xn]$ and $[nX]$. Now $[Xn]$ and $[nX]$ each contain n distinct letters and any subword of these with length more than one occurs just once in r_n. So if $p \approx q$ involves fewer than n letters then θ must assign a linear letter, t_5, in p to some subword of $[Xn]$ and a linear letter t_6 to some subword of $[nX]$. Thus (by possibly deleting some letters in $c(p)$) we find that $\mathbf{S}$ must satisfy the identity

$$\bar{r}_n \equiv \ddot{w}_1 t_1 t_5 a t_2 \ddot{w}_2 t_3 a t_6 t_4 \ddot{w}_3 \approx \ddot{w}_1 t_1 a t_5 t_2 \ddot{w}_2 t_3 t_6 a t_4 \ddot{w}_3 \equiv \bar{r}'_n.$$

However this is not possible because of the following assignment: $a \to a$, $t_5 \to u_1$, $t_6 \to 1$ and all other (linear) letters are assigned maximal unassigned portions of w. This assignment takes the word $\bar{r}_n$ to the word w but assigns $\bar{r}'_n$ the value $w_1 a u_1 u_2 w_2 u_2 a u_1 w_3$, therefore contradicting the claim that w was an isoterm. So the identity $p \approx q$ must contain at least n letters. Since r_n is not an isoterm for every n, any basis for $\mathbf{S}$ must contain infinitely many identities.

(b) As in part (a), the word $\ddot{w}_1 t_1 a t_2 \ddot{w}_2 t_3 a t_4 \ddot{w}_3$ is an isoterm for $\mathbf{S}$ and for any linear letter t, the pair (a, t) is stable in g_n with respect to $\mathbf{S}$. Now say that (x_i, x_j) is unstable in g_n with respect to $\mathbf{S}$. So (x_i, x_j) is unstable in $g_n(x_1, \dots x_n)$ with respect to $\mathbf{S}$. It is shown in [4] that if $xytxy$ is an isoterm for a monoid $\mathbf{T}$ and $[X(2n)]t[\mathcal{X}(2n)]$ is not an isoterm for $\mathbf{T}$ then $\mathbf{T} \models [X(2n)]t[\mathcal{X}(2n)] \approx [\mathcal{X}(2n)]t[X(2n)]$ and any basis of identities of $\mathbf{T}$ contains an identity involving at least n distinct letters. Because the word $u_1 a (u_2 w_2) u_1 a$ is a subword of w, $xytxy$ is an isoterm for $\mathbf{S}$. So if (x_i, x_j) is unstable in $[X(2n)]t[\mathcal{X}(2n)]$ with respect to $\mathbf{S}$ it follows that any basis for the identities of $\mathbf{S}$ contains an identity with at least n distinct letters.

Now assume that (x_i, x_j) is a stable pair in a nontrivial identity $g_n \approx g'_n$ satisfied by $\mathbf{S}$ (this is guaranteed if w contains a subword of the form $xyuyx$ for some possibly empty subword u). So for some i, (a, x_i) must be unstable in $g_n \approx g'_n$. By the choice of u_1 and u_2 the words

$$\ddot{w}_1 t_1 x_i a t_2 \ddot{w}_2 t_3 x_i a t_4 \ddot{w}_3$$

and

$$\ddot{w}_1 t_1 a x_i t_2 \ddot{w}_2 t_3 a x_i t_4 \ddot{w}_3$$

are isoterms. Thus $g'_n(a, x_i, \tau)$ must be the word $\ddot{w}_1 t a x_i t \ddot{w}_2 t x_i a t \ddot{w}_3$ (where τ is a list of all linear letters in g'_n). Therefore since (x_i, x_j) are stable in $g_n \approx g'_n$ for all i, the pair (x_j, a) must be unstable for all j and

$$g_n(a, x_j, \tau) \equiv \ddot{w}_1 t_1 a x_j t_2 \ddot{w}_2 t_3 x_j a t_4 \ddot{w}_3.$$

So g'_n is the word $\ddot{w}_1 t_1 a[X2n]t_2\ddot{w}_2 t_3[\mathcal{X}2n]at_4\ddot{w}_3$.

Therefore if Σ is a basis for the identities of $\mathbf{S}$ then there is an identity $p \approx q \in \Sigma$ so that

$$g_n \equiv A\theta(p)B,\ g'_n \equiv A\theta(q)B$$

for some words A and B and a substitution θ. If the identity $p \approx q$ contained fewer than n letters then there must be letters z, z_1 and z_2 in p so that $\theta(z)$ contains a, $\theta(z_1)$ contains $x_i x_{i+1}$ and $\theta(z_2)$ contains $x_{2j}x_{2j+2}$ for some i, j. Evidently z_1 and z_2 are linear in $p \approx q$ and both (z_1, z) and (z_2, z) are unstable in $p \approx q$. However if we rename z as a, then both p and q are easily seen to be equivalent to a subword of the isoterm

$$\ddot{w}_1 t_1 a t_2 \ddot{w}_2 t_3 a t_4 \ddot{w}_3$$

with possibly some extra linear letters introduced next to existing linear letters. Thus a contradiction has been obtained and therefore no such identity $p \approx q$ can exist. Therefore the basis Σ must contain identities with arbitrarily large numbers of letters and is therefore infinite. □

3 Applications to the Finite Basis Problem for discrete syntactic monoids of finite languages.

It does not follow immediately that the conditions described in the theorems from the previous section are actually satisfied by any monoids. In this section we show that the conditions can be met by relating them specifically to discrete syntactic monoids of finite languages.

If w is a word and a is a letter in $c(w)$ then we may write w as

$$w_1 a^{n_1} w_2 a^{n_2} w_3 \dots w_m a^{n_m} w_{m+1}$$

where for every $i \leq m+1$, n_i is a positive integer, w_1 and w_{m+1} are possibly empty words, w_2, w_3, … , w_m are words and for each $i \leq m+1$, a is not contained in w_i. We may then define the *occurrence vector* of a in w to be the m-tuple $V_w(a) = (n_1, n_2, \dots, n_m)$. Clearly $\sum_{i=1}^{m} n_i = occ(a, w)$. If we replace the condition that a is a *single letter* occurring in w with the condition that a is a *subword* of w then we obtain a notion of an occurrence vector for arbitrary subwords of w. However if a is a subword of w of length more than 1, the notation $V_w(a)$ is no longer well defined since a given subword of w may have several distinct occurrence vectors. For example the word $w \equiv aaaaa$ (where a is a letter) can be written as $(aa)^2a$ or $(aa)a(aa)$ or $a(aa)^2$ and so there are two distinct occurrence vectors for aa in w: they are (2) and $(1,1)$. Our primary concern will be with occurrence vectors of letters in words and for our purposes it will suffice to assume that when v is a subword of w then $V_w(v)$ is any one particular occurrence vector of v in w.

Definition 3.1 *An occurrence vector*

$$v_1 = (n_1, n_2, \dots, n_p)$$

contains *an occurrence vector*

$$v_2 = (m_1, m_2, \dots, m_q)$$

if there is a substitution $\theta : X^* \to X^*$ *with* $\theta(a) \equiv a$ *(for some fixed letter* a*) such that the word*

$$a^{n_1} t_1 a^{n_2} t_2 \dots t_{p-1} a^{n_p}$$

(where the t_i *are letters) contains as a subword the word*

$$\theta(a^{m_1} t_1 a^{m_2} t_2 \dots t_{q-1} a^{m_q}).$$

In this case we will write $v_1 \geq v_2$*.*

For example, take v_1 and v_2 as in the definition and let $h_1, h_2, \dots, h_q$ be a subsequence of $n_1, n_2, \dots, n_p$ such that $m_i \leq h_i$. Consider the word $w \equiv a^{n_1} t_1 a^{n_2} t_2 \dots t_{p-1} a^{n_p}$. Since $h_1, h_2, \dots, h_q$ is a subsequence of $n_1, n_2, \dots, n_p$, w is of the form $w_0 a^{h_1} w_1 a^{h_2} w_2 \dots a^{h_q} w_q$ for some words w_1, w_2, … , w_{q-1} and some possibly empty words w_0 and w_q. Now let θ be the substitution defined by $\theta(a) \equiv a$ and $\theta(t_i) \equiv a^{h_i - m_i} w_i$. Evidently

$$\theta(a^{m_1} t_1 a^{m_2} t_2 \dots t_{q-1} a^{m_q}) \equiv a^{h_1} w_1 a^{h_2} w_2 \dots a^{h_q} w_q,$$

a subword of w and so by Definition 3.1, the occurrence vector v_1 contains the occurrence vector v_2. Also if θ is a substitution that assigns 1 to all linear letters of the form t_i in the word $w_1 \equiv a^{n_1} t_1 a^{n_2} t_2 \dots t_{p-1} a^{n_p}$ and assigns a to itself then $\theta(w_1) \equiv a^n$ where $n = \sum_{i=1}^{p} n_i$. Therefore the singleton occurrence vector $(n + i)$ contains the vector v_1 for any non-negative integer i. An occurrence vector of a subword u in a word w is said to be *maximal* in w if for any occurrence vector $V_v(w)$ of a subword v in w, we have that $V_w(v) \geq V_w(u) \Rightarrow u \equiv v$. Likewise if W is a language containing w then $V_w(u)$ is maximal in W if for every subword v of a word $w' \in W$, $V_{w'}(v) \geq V_w(u) \Rightarrow (u \equiv v \;\&\; w' \equiv w)$. Possibly the simplest way in which an occurrence vector $V_w(a)$ of a letter a in a word w can be maximal in a language W is if a occurs more times in w than any other letter and the remaining words in W are $(occ(a, w) - 1)$-limited. Another simple situation is if there is a power of a in w that is higher than the power of any other subword of a word in W. On the other hand, there need not be a maximal occurrence vector amongst the set of all occurrence vectors of a word. For example in the word *aabbcc*, we have $V_w(a) = V_w(b) = V_w(c) = (2)$ and all other occurrence vectors are the singleton (1).

The following lemma gives some indication about when a word w contains a letter a such that $V_w(a)$ is a maximal occurrence vector in w. Here and below we shall say that a letter a is maximally occurring in a word w if for all letters $b \in c(w)$, $occ(a, w) \leq occ(b, w)$ implies that $occ(a, w) = occ(b, w)$.

Lemma 3.2 *If w is a word then either w has a maximally occurring letter whose occurrence vector is maximal or there are two maximally occurring letters in w with identical occurrence vectors.*

Proof: Let a occur a maximal number of times in w and assume that no pair of distinct maximally occurring letters in w share the same occurrence vector. The occurrence vectors of letters in w are partially ordered by the $\leq$ relation and (as noted above) if a letter c occurs more times in w than a letter b occurs in w, then it cannot be the case that $V_w(a) \leq V_w(b)$. Thus since $c(w)$ is a finite set there is an occurrence vector $V_w(x)$ of a maximally occurring letter x in w such that $V_w(a) \leq V_w(x)$ and $V_w(x)$ is maximal amongst those occurrence vectors of single letters. It remains to show there is no subword u of length more than 1 for which has an occurrence vector $V_w(u)$ such that $V_w(x) \leq V_w(u)$.

If such a subword u exists it must have at least $occ(x, w)$ distinct occurrences in w. However every letter in u occurs at least as many times in w as does u and so because x occurs a maximal number of times in w, u must have exactly $occ(x, w)$ nonoverlapping occurrences of the subword u. Each letter occurring in u must be linear in u since otherwise they would occur at least twice as many times as x does. So let y be the first letter in u and z be the last letter (these are distinct since the length of u is at least 2). Since both y and z cannot have any occurrences outside of the relevant occurrences of the subword u (otherwise they would occur more times in w than x does), the occurrence vector of both y and of z are both equal to the $occ(x, w)$-tuple $(1, 1, 1, \ldots, 1)$, contradicting the condition that no two maximally occurring letters share the same occurrence vector. Thus $V_w(x)$ is a maximal occurrence vector for w. $\square$

A second important lemma associated with occurrence vectors is the following.

Lemma 3.3 *Let $V_w(x) = (n_1, n_2, \ldots, n_p)$ and $V_w(y) = (m_1, m_2, \ldots, m_q)$ be the occurrence vectors of two maximally occurring letters x and y in a word w. If $V_w(x) \leq V_w(y)$ then there are numbers $0 = i_0 < i_1 < i_2 < \ldots < i_{q-1} < i_q = p$ such that for each $0 \leq j \leq q - 1$, $\displaystyle\sum_{i_j+1 \leq i \leq i_{j+1}} n_i = m_{j+1}$.*

Proof: Because x and y occur the same number of times in the word w we must have that $\sum_{i=1}^{p} n_i = \sum_{i=1}^{q} m_i$ and so if θ is substitution such that $\theta(a^{n_1}t_1a^{n_2}t_2 \ldots t_{p-1}a^{n_p})$ is a subword of $a^{m_1}t_1a^{m_2}t_2 \ldots t_{q-1}a^{m_q}$ and $\theta(a) \equiv a$, then $\theta(t_i)$ cannot contain the letter a for any $i \leq p - 1$. Therefore $\theta(t_i)$ is either 1 or t_j for some j. We can now choose i_j to be the subscript of the j^{th} linear letter t_i not assigned the value 1 by θ. $\square$

The relevance of maximal occurrence vectors lies in the following simple lemma.

Lemma 3.4 *Let w_1 and w_2 be words with u a subword of w_1 and v a subword of w_2. Let θ be a substitution and let $V_{w_1}(u)$ and $V_{w_2}(v)$ be occurrence vectors of u in w_1 and of v in w_2 respectively. If $V_{w_1}(u) = V_{w_2}(v)$ and $V_{w_2}(v)$ is a maximal occurrence vector* *in a language W of words containing w_2 then $\theta(w_1)$ is a subword of a word in W only if $\theta(u) \equiv 1$ or both $\theta(u) \equiv v$ and $\theta(w_1)$ is a subword of w_2.*

Proof: This is because if $\theta(u) \not\equiv 1$ then any occurrence vector of $\theta(u)$ in $\theta(w_1)$ contains the occurrence vector $V_{w_1}(u)$ which equals $V_{w_2}(v)$. Since $V_{w_2}(v)$ is maximal in W then $\theta(w_1)$ cannot be a subword of any word in W except for the word w_2 and in this case $\theta(u) \equiv v$. $\square$

Theorem 3.5 *Let W be a language and $w \in W$ be a word containing the letters a and b such that $V_w(a)$ is maximal in W (the language W may of course be simply $\{w\}$ itself). Let β_1 and β_2 be any positive numbers and p be any (possibly empty) word not containing a or b. If w satisfies any one of the following conditions (or their reverse) then W is not finitely based (in each case we will assume that the given subwords of w are not contained within each other though they may overlap):*
(i) w has a subword $ab^{\beta_1}a$ and a subword $apb^{\beta_2}ba$;
(ii) w has a subword $abb^{\beta_1}a$ and a subword $apb^{\beta_2}a$;
(iii) w has subwords of the form aba, $apba$ and ba;
(iv) w contains aba and ends with $apba$. For example, w ends with $ababa$;
(v) w has a subword of the form $abbb^{\beta_1}a$ and of the form apb. For example $abbbab$ is a subword of w;
(vi) w has a subword aba and a subword $apbaa$ and $V_w(a)$ is the only occurrence vector of a letter in a word in W that contains the occurrence vector $V_{w'}(a)$, where w' is obtained by replacing the particular occurrence of $apbaa$ by $apaba$.

Proof: In every case we will construct a set of identities $\{u_n \approx v_n\}$ based on the form of w and apply Theorem 2.7. Both sides of the identities constructed will contain the letter a and in all except the last case the occurrence vectors of a in these words will be identical to that of w. Since $V_w(a)$ is maximal in W, by Lemma 3.4, if θ is a substitution then $\theta(u_n)$ or $\theta(v_n)$ is a subword of a word in W only if $\theta(a) \equiv 1$ or $\theta(a) \equiv a$. Furthermore, if $\theta(a) \equiv a$ then $\theta(u_n)$ (or $\theta(v_n)$) is a subword of the word w. The identities $u_n \approx v_n$ will also be constructed so that if $\theta(a) \equiv 1$ then $\theta(u_n) \equiv \theta(v_n)$. Therefore in the arguments to follow in this proof it will be sufficient to consider the case when $W = \{w\}$ and $\theta(a) \equiv a$. However because the proofs are quite repetitious we will only fully prove cases (i) and (ii) (by applying Theorem 2.7). Full proofs of the remaining cases are contained in the author's PhD thesis [3].

First note that in every case in the theorem, w contains a subword of the form ab and another of the form ba (not intersecting). This is all that is required to establish that $xytyx$ is an isoterm for $S(W)$ (as is needed for Theorem 2.7 to be applied).

(i) Let $w \equiv w_1\ {}_ia b^{\beta_1} a w_2\ {}_ja p b^{\beta_2} b a^{\alpha} w_3$, where ${}_ia$ and ${}_ja$ (as defined in Definition 1.1 (vi)) denote the i^{th} and j^{th} occurrences of a in w respectively and α is maximal.
Claim: $S(\{w\}) \models u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1\ {}_ia[Xn]b^{\beta_1-1}a\tilde{w}_2\ {}_jat_1[nX]t_2b^{\beta_2}ba^{\alpha}\tilde{w}_3$$

and

$$v_n \equiv \tilde{w}_1\ {}_ia[Xn]b^{\beta_1-1}a\tilde{w}_2\ {}_jat_1[nX]t_2b^{\beta_2}a^{\alpha}b\tilde{w}_3.$$

Let θ be a substitution such that $\theta(u_n)$ (or $\theta(v_n)$) is a subword of w. Between ${}_ja$ and ${}_{(j+1)}a$ in w there is the word $b^{\beta_2}b$. So $\theta(t_1[nX]t_2b^{\beta_2}b)$ (or $\theta(t_1[nX]t_2b^{\beta_2})$) must be the word $b^{\beta_2}b$. Now if θ assigns b the value 1, then $\theta(u_n) \equiv \theta(v_n)$ because (a, b) is the only unstable pair in $u_n \approx v_n$. The remaining case is when $\theta(b)$ is the letter b and we will show that this never occurs (note that $\theta(b)$ cannot be a higher power of b since then we would have more than $occ(b, w)$ occurrences of b in $\theta(u_n)$ and $\theta(v_n)$).

If we are considering u_n then $\theta(b) \equiv b$ implies $\theta(t_1[Xn]t_2) \equiv 1$ and then the subword of u_n between ${}_ia$ and ${}_{(i+1)}a$ is effectively b^{β_1-1}. Between ${}_ia$ and ${}_{(i+1)}a$ in w however, there is the word b^{β_1} which contradicts the assumption that $\theta(u_n)$ was a subword of w since $\theta(b^{\beta_1-1})$ cannot be b^{β_1} if $\theta(b) \equiv b$. If we are considering v_n then the fact that $\theta(b) \equiv b$ implies $\theta(t_1[nX]t_2) \equiv b$. If $\theta([nX]) \equiv 1$ then the previous argument applies and a contradiction is obtained. If $\theta(x_k) \equiv b$ for some k, then $occ(b, \theta(v_n)) > occ(b, w)$ since x_k is 2-occurring in v_n; again we have a contradiction. It follows that $S(W)$ is not finitely based by Theorem 2.7.

(ii) Let $w \equiv w_1\ {}_iabb^{\beta_1}a^{\alpha}w_2\ {}_japb^{\beta_2}aw_3$, where α is maximal.
Claim: $S(\{w\}) \models u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1\ {}_ia[Xn]b^{\beta_1}a^{\alpha}\tilde{w}_2\ {}_jat_1[nX]t_2b^{\beta_2}a\tilde{w}_3$$

and

$$v_n \equiv \tilde{w}_1\ {}_ia[Xn]b^{\beta_1-1}a^{\alpha}b\tilde{w}_2\ {}_jat_1[nX]t_2b^{\beta_2}a\tilde{w}_3.$$

Let θ be a substitution such that $\theta(u_n)$ is a subword of w. Between ${}_ia$ and ${}_{(i+1)}a$ in w we have the word b^{β_1+1}. So $\theta([Xn]b^{\beta_1}) \equiv b^{\beta_1+1}$. Now $\theta(b)$ cannot be b^k for any k greater than 1 since then $occ(b, \theta(u_n)) > occ(b, w)$. If $\theta(b) \equiv b$, then we must have $\theta(x_k) \equiv b$ for some k. But then $occ(b, \theta(u_n)) > occ(b, w)$ since x_k is 2-occurring in u_n, again contradicting the choice of θ. Thus $\theta(b) \equiv 1$, and therefore $\theta(u_n) \equiv \theta(v_n)$.

Now let $\theta(v_n)$ be a subword of w. Like before, $\theta([Xn]b^{\beta_1-1})$ must be the word b^{β_1}. If b^{β_1-1} is empty (that is, if $\beta_1 = 1$) then for some k, $\theta(x_k)$ must be the word b (since $\beta_1 = 1$). Also $\theta(t_1[nX]t_2b^{\beta_2}) \equiv pb^{\beta_2}$ since this is the word between ${}_ja$ and ${}_{(j+1)}a$ in w. Because p does not contain b and $\theta(x_k) \equiv b$, $\theta(t_2b^{\beta_2})$ must be b^{β_2-1}. Thus $\theta(b) \equiv 1$ and $\theta(v_n) \equiv \theta(u_n)$. It follows that $S(W)$ is not finitely based by Theorem 2.7.

(iii) For example $w \equiv w_1\ {}_iabaw_2\ {}_japbaw_3ba^{\alpha}w_4$, where α is maximal.
The identities to be used are $u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1\ {}_ia[Xn]a\tilde{w}_2\ {}_jat_1[nX]t_2ba\tilde{w}_3t_3ba^{\alpha}\tilde{w}_4$$

and

$$v_n \equiv \tilde{w}_1\ {}_ia[Xn]a\tilde{w}_2\ {}_jat_1[nX]t_2ba\tilde{w}_3t_3a^{\alpha}bw_4.$$

(iv) For example $w \equiv w_1\ {}_iabaw_2\ {}_japba^{\alpha}$, where α is maximal.
The identities to be used are $u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1\ {}_ia[Xn]a\tilde{w}_2\ {}_jat_1[nX]t_2ba^{\alpha}$$

and

$$v_n \equiv \tilde{w}_1 \; {}_i a[Xn]a\tilde{w}_2 \; {}_j a t_1[nX]t_2 a^{\alpha} b.$$

(v) For example $w \equiv w_1 \; {}_i abbb^{\beta_1} a^{\alpha} w_2 \; {}_j apbw_3$, where α is maximal. The identities to be used are $u_n \approx v_n$, where

$$u_n \equiv \tilde{w}_1 \; {}_i a[Xn]b^{\beta_1} b a^{\alpha} \tilde{w}_2 a t_1[nX]t_2 b\tilde{w}_3$$

and

$$v_n \equiv \tilde{w}_1 \; {}_i a[Xn]b^{\beta_1} a^{\alpha} b \tilde{w}_2 a t_1[nX]t_2 b\tilde{w}_3.$$

(vi) For example, $w \equiv w_1 \; {}_i abaw_2 \; {}_j apbaa^{\alpha} w_3$, where α is maximal. The identities to be used are $u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1 \; {}_i a[Xn]a\tilde{w}_2 \; {}_j a t_1[nX]t_2 baa^{\alpha} \tilde{w}_3$$

and

$$v_n \equiv \tilde{w}_1 \; {}_i a[Xn]a\tilde{w}_2 \; {}_j a t_1[nX]t_2 aba^{\alpha} \tilde{w}_3.$$

The extra condition required for this part is due to the fact that the occurrence vector of a in v_n is no longer identical to that of a in w. Once given this condition however we are still able to make the assumptions indicated at the start of this proof. The extra condition is still held in many cases: for example if a occurs more times in w than any other letter in w and more times than any letter does in any other word in W. □

Theorem 3.5 by no means captures all possible applications of Theorem 2.7. For example in the word $w \equiv (ba)^n$ where $n > 2$, the vector $V_w(a)$ is not maximal (since $V_w(a) = V_w(b)$). Yet for every $n > 2$, $S(\{(ba)^n\})$ still satisfies the identity $u_n \approx v_n$ where u_n is the word $(bt_1abt_2a \ldots bt_{n-3}a)babt_{n-2}[Xn]t_{n-1}a[nX]a$ and v_n is the word $(bt_1abt_2a \ldots bt_{n-3}a)bbat_{n-2}[Xn]t_{n-1}a[nX]a \equiv v_n$ since if θ is an assignment that does not assign a the value 1 and $\theta(u_n)$ is a subword of w then either $\theta(a) \equiv a$, $\theta(a) \equiv b$ or $\theta(a) = (ba)$ (these are the only subwords of w that occur as many times as the letter a does in u_n). If $\theta(a) \equiv ba$ then clearly $\theta(b) \equiv 1$ and $\theta(u_n) \equiv \theta(v_n)$. If $\theta(a) \equiv b$ then the first occurrence of a in u_n must be assigned the first occurrence of b in w. The first letter to appear in u_n is b and yet there is no letter left of the first occurrence of b in w. Therefore $\theta(b) \equiv 1$ and $\theta(u_n) \equiv \theta(v_n)$. The remaining case is when $\theta(a) \equiv a$ and then the proof becomes effectively the same as that of Theorem 3.5 (iv). A similar argument applies when considering v_n. It follows by Theorem 2.7 that $(ba)^n$ is not finitely based (this fact was previously established in [20]).

The arguments just used did not depend on the fact that $(ba)^n$ contained only two distinct letters, only on the fact that to the left of the first occurrence of b there was no proper subword occurring at least $n-1$ times (that is, the number of times that the letter b occurs in the identities used in the example). Thus we can deduce the following theorem.

Theorem 3.6 *Let w be a word which has exactly two maximally occurring letters, a and b, with the first occurrence of b occurring in w before the first occurrence of a and* *with the property that every letter left of the first occurrence of b occurs fewer than $occ(a, w) - 1$ times. If w satisfies one of the conditions (i) to (vi) of Theorem 3.5 and the subwords described in the relevant part of Theorem 3.5 do not involve the first occurrence of a and of b in w then w is not finitely based.*

Definition 3.7 *An occurrence vector $V_w(u)$ of a subword u in a word w is said to be* super maximal *if the deletion of any one particular occurrence of u in w gives a new word v with the property that for any subword u' of w, $V_v(u) \leq V_w(u')$ only if $u' \equiv u$. Likewise $V_w(u)$ is super maximal in a language W containing w if for every subword u' of a word $w' \in W$, $V_v(u) \leq V_{w'}(u')$ only if $u \equiv u'$ and $w \equiv w'$.*

Clearly in this definition $V_v(u)$ can be obtained by subtracting the number 1 from one of the entries of $V_w(u)$ and deleting any zero entries from the resulting list. A simple example of a super maximal occurrence vector is the occurrence vector of a letter in a word that has at least two extra occurrences in the word than any other letter.

We may now extend Lemma 3.4 as follows (the proof is similar to that of Lemma 3.4).

Lemma 3.8 *Let W be a language, $w \in W$ be a word and u be a subword of w for which $V_w(u)$ is super maximal in W. If an occurrence vector $V_{w'}(u')$ of a subword u' in a second word w' (not necessarily in W) can be obtained by subtracting the number 1 from one of the entries of $V_w(u)$ and deleting any zero entries from the resulting list then for any substitution θ, $\theta(w')$ is a subword of a word in W only if $\theta(u') \equiv 1$ or both $\theta(w')$ is a subword of w and $\theta(u') \equiv u$.*

Using this lemma and Theorem 2.7 one can now obtain a variation on Theorem 3.5

Theorem 3.9 *Let W be a language and $w \in W$ be a word containing a letter a and a letter b such that $V_w(a)$ is super maximal in W (the language W may of course be simply $\{w\}$ itself). Let p be any (possibly empty) word not containing the letter a and α_1, α_2, β_1 and β_2 be arbitrary positive integers. If w satisfies any one of the following conditions or their reverse then $S(W)$ is not finitely based (in a similar way to before, we will assume that unless otherwise stated the given subwords may overlap but may not be contained within one another):*
(i) $apb^{\beta_1}aaa^{\alpha_1}b$ is a subword of w and $V_w(a)$ is the only occurrence vector of a subword in W that contains the occurrence vector of the letter a in the word obtained from w by replacing the given occurrence of $apb^{\beta_1}aaa^{\alpha_1}b$ by $apb^{\beta_1-1}aba^{\alpha_1}b$;
(ii) $baa^{\alpha_1}b$, apb and ba are subwords of w and the occurrence of ba in w does not overlap with that of $baa^{\alpha_1}b$. For example $baababa$, $abaabba$ or $baabbab$ is a subword of w;
(iii) $apbb^{\beta_1}aa^{\alpha_1}b$ is a subword of w, where α is maximal. For example $abbaab$ is a subword of w;
(iv) $b^{\beta_1}a^{\alpha_1}b^{\beta_2}a^{\alpha_2}pb$ is a subword of w, $\beta_1 > \beta_2$ and p does not contain b. For example, $bbabab$ is a subword of w.

Proof: The proof of this theorem is similar to that of Theorem 3.5 except that the identities $u_n \approx v_n$ we construct in this case have only $occ(a, w) - 1$ occurrences of a (instead of $occ(a, w)$ occurrences). It is for this reason that we require $V_w(a)$ to be super maximal so that by Lemma 3.8 if θ is a substitution such that $\theta(u_n)$ (or $\theta(v_n)$) is a subword of a word in W, then either both $\theta(a) \equiv a$ and $\theta(u_n)$ (or $\theta(v_n)$) is a subword of w or $\theta(a) \equiv 1$, in which case $\theta(u_n) \equiv \theta(v_n)$. We only fully prove case (i) since this is the only part we use below. Proofs of all other cases are similar and appear in the author's PhD thesis [3].

First note that in every case in the theorem, w contains a subword of the form ab and another of the form ba (not intersecting). As in Theorem 3.5, this is all that is required to establish that $xytyx$ is an isoterm for $S(W)$. In each case, this and the fact that the given identities are satisfied by $S(W)$ is enough to ensure that W is not finitely based by Theorem 2.7.

(i) For example, $w \equiv w_1 apb^{\beta_1}\ {}_iaaa^{\alpha_1}bw_2$.

Claim: $S(\{w\}) \models u_n \approx v_n$, where

$$u_n \equiv \tilde{w}_1 at_1[Xn]t_2b^{\beta_1}\ {}_iaa^{\alpha_1}[nX]b\tilde{w}_2$$

and

$$v_n \equiv \tilde{w}_1 at_1[Xn]t_2b^{\beta_1-1}\ {}_iaba^{\alpha_1}[nX]b\tilde{w}_2$$

If θ is a substitution such that $\theta(u_n)$ is a subword of w then because $occ(a, u_n) = occ(a, w) - 1$, θ must take the i^{th} occurrence of a in u_n to either the i^{th} or the $(i+1)^{th}$ occurrence of a in w (we will write this as $\theta({}_ia) \equiv\ {}_ia$ or ${}_{(i+1)}a$ in w). Now $\theta([nX])$ and $\theta(b)$ cannot contain a else we would have more than $occ(a, w)$ occurrences of a in $\theta(u_n)$. So in the first case (when $\theta({}_ia) \equiv\ {}_ia$) since the word a^{α_1+1} occurs immediately to the right of ${}_ia$ in w but the word $a^{\alpha_1}[nX]b$ occurs immediately to the right of ${}_ia$ in u_n, we must have $\theta([nX]b) \equiv 1$ or $\theta([nX]b)$ contains an occurrence of a. If $\theta([nX]b)$ contains an occurrence of a then $\theta(u_n)$ contains more than $occ(a, w)$ occurrences of a which is not possible. In the second case (when $\theta({}_ia) \equiv\ {}_{(i+1)}a$), there is the letter a immediately to the left of ${}_{(i+1)}a$ in w which implies the next letter left of ${}_ia$ in u_n not assigned the value 1 by θ, must be assigned a word ending in a. However, b cannot be assigned a word containing a. Consequently $\theta(b) \equiv 1$ and therefore $\theta(u_n) \equiv \theta(v_n)$.

Since $V_w(a)$ is the only occurrence vector of a letter in w that contains $V_{v_n}(a)$, we may assume as before that if θ is a substitution with the property that $\theta(v_n)$ is a subword of w then $\theta(a) \equiv a$. So θ must assign the i^{th} occurrence of a in v_n to either the i^{th} or the $(i+1)^{th}$ occurrence of a in w. In the first instance, ${}_{(i+1)}a$ lies immediately to the right of ${}_ia$ in w but in v_n, b lies immediately to the right of ${}_ia$. Since $\theta(b)$ does not contain a, $\theta(b)$ must be 1 and therefore $\theta(v_n) \equiv \theta(u_n)$. The second case follows in a similar way since immediately to the right of ${}_{(i+1)}a$ in w is the $(i+2)^{th}$ occurrence of a but b occurs to the right of ${}_ia$ in v_n.

(ii) For example $w \equiv w_1 b\ {}_iaa^{\alpha_1}bw_2apbw_3ba^{\alpha}w_4$, where α is maximal.
Use the identities $u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1 b\ {}_iaa^{\alpha_1-1}[Xn]b\tilde{w}_2at_1[nX]t_2b\tilde{w}_3ba^{\alpha}\tilde{w}_4$$

and

$$v_n \equiv \tilde{w}_1 b \ {}_i a a^{\alpha_1 - 1}[Xn] b \tilde{w}_2 a t_1 [nX] t_2 b \tilde{w}_3 a^{\alpha} b \tilde{w}_4.$$

(iii) For example, $w \equiv w_1 apbb^{\beta_1} \ {}_i a a^{\alpha_1} b w_2$.
Use the identities $u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1 a t_1 [Xn] t_2 b^{\beta_1} b \ {}_i a a^{\alpha_1 - 1}[nX] b \tilde{w}_2$$

and

$$v_n \equiv \tilde{w}_1 a t_1 [Xn] t_2 b^{\beta_1} \ {}_i a a^{\alpha_1 - 1} b [nX] b \tilde{w}_2.$$

(iv) For example $w \equiv w_1 b^{\beta_1} b a^{\alpha_1} b^{\beta_2} \ {}_i a a^{\alpha_2 - 1} p b w_2$.
Use the identities $u_n \approx v_n$ where

$$u_n \equiv \tilde{w}_1 b^{\beta_1} a^{\alpha_1 - 1}[Xn] b^{\beta_2} \ {}_{(i-1)} a a^{\alpha_2 - 1} t_1 [nX] t_2 b \tilde{w}_2$$

and

$$v_n \equiv \tilde{w}_1 b^{\beta_1} a^{\alpha_1 - 1}[Xn] b^{\beta_2 - 1} \ {}_{(i-1)} a a^{\alpha_2 - 1} b t_1 [nX] t_2 b \tilde{w}_2.$$

□

This theorem and Theorem 3.5 depend on Theorem 2.7. We now present an analogous theorem using Theorem 2.9.

Theorem 3.10 *Let W be a language and $w \in W$ be a word containing letters a, b, c for which $V_w(a)$ is maximal in W and let u and v be any (possibly empty) words with $a, b \notin c(u)$ and $a, c \notin c(v)$ (the language W may of course be simply $\{w\}$ itself). If w satisfies any one of the following properties (or their reverse) then W is not finitely based:*
(i) bac and $aucabva$ are non overlapping subwords of w;
(ii) $bacva$ and $aucab$ are non overlapping subwords of w;
(iii) bac and $avbacua$ are non overlapping subwords of w;
(iv) $avbac$ and $avbac$ are non overlapping subwords of w.

Proof: Parts (i) and (ii) are obtained by an application of part (i) of Theorem 2.9 with $u_1 \equiv b$ and $u_2 \equiv c$. We will only prove (i) here. Likewise parts (iii) and (iv) follow in a very similar manner from part (ii) of Theorem 2.9 and so will also not be proved. Since $V_w(a)$ is maximal in w, Lemma 3.4 implies that if θ is a substitution so that $\theta(r_n)$ (or $\theta(r'_n)$) is a subword of w, then either $\theta(a) \equiv a$ or $\theta(a) \equiv 1$.

(i) For example, $w \equiv w_1 b \ {}_i a c w_2 a u c \ {}_j a b v a w_3$. We will show that $S(\{w\})$ satisfies the identity

$$\dot{w}_1 a t_1 [Xn] \ {}_i a t_2 \dot{w}_2 a t_3 [nX] \ {}_j a t_4 \dot{w}_3 \approx \dot{w}_1 a t_1 \ {}_i a [Xn] t_2 \dot{w}_2 a t_3 \ {}_j a [nX] t_4 \dot{w}_3.$$

For convenience we will denote the left side of this identity by r_n and the right side by r'_n. Firstly if $\theta([Xn]) \equiv 1$ or $\theta(a) \equiv 1$ then $\theta(r_n) \equiv \theta(r'_n)$. Now left of ${}_i a$ in w is the letter b. So if $\theta(r_n)$ is a subword of w and $\theta([Xn]) \not\equiv 1$, then $\theta([Xn])$ contains b. But then $\theta([nX])$ contains b and so contained in the word between ${}_{(j-1)}a$ and ${}_{j}a$ in $\theta(r_n)$ is a letter b. However between ${}_{(j-1)}a$ and ${}_{j}a$ in w there is no letter b, contradicting the assumption that $\theta(r_n)$ was a subword of w. The case when $\theta(r'_n)$ is a subword of w follows by symmetry. □

We will make no further use of this theorem below however in several places it can be used in place of Theorem 3.5.

In [3] it is shown that for any fixed finite alphabet A and fixed positive integer $k \geq 1$, the proportion of all k element sets of words in the alphabet A and of maximum length n, to which the main results (in particular, Theorem 3.5) of this section apply, tends toward 1 as n tends toward infinity. Since the main results of this section concern *not finitely based* monoids, it follows in a natural sense, that "almost all" k element sets of words in a fixed finite alphabet are not finitely based (and that Theorem 3.5 below applies "almost always"). We note in contrast that the proportion of all (not necessarily associative) binary operations definable on an n element set that give rise to *finitely based* groupoids tends toward 1 as n tends toward infinity [12], and the proportion of all associative binary operations (respectively, associative binary operations with identity) definable on an n element set that give rise to 3-nilpotent semigroups (respectively, 3-nilpotent monoids) tends toward 1 as n tends toward infinity [5] (respectively, [6]). (Here a 3-nilpotent semigroup is a semigroup in which the product of any three elements equals 0 and a 3-nilpotent monoid is a 3-nilpotent semigroup with adjoined identity element.) A 3-nilpotent semigroup or monoid always satisfies the identity $xyx \approx xxy$ and therefore is finitely based by one of the central results of [17]. Thus in a natural sense almost all groupoids, semigroups and monoids are finitely based.

4 Subwords of finitely based words and not finitely based words.

In [4] it is shown that every finite language W is a subset of a finite, not finitely based language W' (if the alphabet of W has more than one letter then one can assume that W' has the same alphabet as W). We now prove a variation of this idea.

Theorem 4.1 *Every word w is a subword of a not finitely based word w' whose length is no more than 4 letters longer than w. If $|c(w)| > 1$ then w' can be chosen such that $c(w') = c(w)$.*

Proof: If $|c(w)| = 0$ then w is the empty word and it follows from results obtained in [4] that the shortest not finitely based word containing w is the word *abab* or *abba*.

If $|c(w)| = 1$ then w is of the form a^k for some k. In this case we may choose w' to be the word $a^k bab$ for some new letter b. It follows immediately from the description obtained in [20] of the not finitely based words in a two letter alphabet that w' is not finitely based. Now assume $|c(w)| > 1$.

Case 1. w ends with a letter a that occurs a maximal number of times in w and there is at least one letter b occurring in w fewer times than a. In this case we may take w' to be the word $wbaba$ and apply Theorem 3.5 (iv).

Case 2. Every letter of w occurs an equal number of times. Let b be the last letter in w and a be the next letter left of this that is different to b. So $w \equiv w_1 a b^\beta$ for some $\beta > 0$. Thus we may take w' to be the word $w_1 a b^\beta aaab$ and apply Theorem 3.9 (i).

Case 3. w ends with a letter, b say, not occurring a maximal number of times in w. Let a be the closest letter to the right end of w that does occur a maximal number of times. Then we may choose w' as the word $waaab$ and apply Theorem 3.9 (i). □

Example 4.2 *It is shown in [4] that the word* $abcbadefgef$ *is finitely based while Theorem 4.1 implies that* $abcbadefgefgfgf$ *is not finitely based. Note also that Theorem 3.5 (iv) implies that* $fefabcbadefgef$ *is not finitely based.*

As we now show, all words are also subwords of a finitely based word however the method we present is much less efficient than that obtained above since the alphabet of the new word may need to be many times larger than that of the original word.

We first use the following lemma from [4].

Lemma 4.3 *[4] If* w *is an isoterm for a monoid* $\mathbf{S}$ *then* $Id(\mathbf{S}) \subseteq Id(S(\{w\}))$.

Theorem 4.4 *Every word* w *is a subword of a finitely based word* w'.

Proof: Let k be a number such that w is k-limited. It follows from results in [4] that if W_k is the language of all k-limited words in a two letter alphabet, then $W_k \cup \{w\}$ is finitely based. The proof will be complete if there is a word w' containing w as a subword and for which $S(\{w'\})$ satisfies exactly the same identities as $S(W_k \cup \{w\})$.

First assume (without loss of generality) that each word in $W_k \cup \{w\}$ is written in a distinct alphabet. Let u be the word obtained by concatenating every word in this set. Obviously every word in $W_k \cup \{w\}$ is an isoterm for $S(\{u\})$ since they are subwords of u. On the other hand if (x, y) is an unstable pair in some identity $u \approx v$ satisfied by $S(W_k \cup \{w\})$ then both x and y must be letters which occur in distinct words in $W_k \cup \{w\}$. Thus $u(x, y)$ is of the form $x^n y^m$ or $y^n x^m$ for some $n, m \leq k$. Since these are k-limited words in a two letter alphabet, they are isoterms for $S(W_k)$ and therefore isoterms for $S(W_k \cup \{w\})$, contradicting the claim that (x, y) was an unstable pair. By Lemma 4.3, $S(\{u\})$ and $S(W_k \cup \{w\})$ satisfy the same identities and the proof is complete. □

In [4], the finitely based set of words W_k is used to show that every finite set of words W is contained in a finitely based set of words: simply choose k sufficiently large that $W \cup W_k$ is k-limited. The method of the proof above shows that there is a single (finitely based) word w such that $S(\{w\})$ satisfies the same identities as $S(W \cup W_k)$. That is, for any finite set of words W there is a single finitely based word w such that $\mathbf{V}(S(W)) \subseteq \mathbf{V}(S(\{w\}))$. Combining this theorem with Theorem 4.1 we have

Corollary 4.5 *For every set of words W there are words $w_1, w_2, \ldots$ with w_i a subword of w_{i+1} (so that $\mathbf{V}(S(\{w_i\})) \subset \mathbf{V}(S(\{w_{i+1}\})))$ and such that $S(W) \in \mathbf{V}(S(\{w_1\}))$ and w_i is finitely based if and only if i is even. If $W = \{w\}$ has only one element then w_1 can be chosen such that w is a subword of w_1.*

Of course we have not obtained a bound on the number of distinct letters in the words $w_1, w_2, \ldots$; we will return to this issue in the next section.

As discussed in the introduction, the discrete syntactic monoid $S(\{w_i\})$ is actually the syntactic monoid of the language $\{w_i\}$ and the syntactic monoid of any language is a quotient of the discrete syntactic monoid of that language. Thus we obtain a "syntactic monoid version" of Corollary 4.5 as follows.

Corollary 4.6 *If $\mathbf{S}$ is a syntactic monoid of a finite language then there are syntactic monoids (of finite languages) $\mathbf{M}_1$, $\mathbf{M}_2$, $\ldots$ such that $\mathbf{V}(\mathbf{M}_i) \subset \mathbf{V}(\mathbf{M}_{i+1})$, $\mathbf{S} \in \mathbf{V}(\mathbf{M}_1)$ and $\mathbf{V}(\mathbf{M}_i)$ is finitely based if and only if i is even.*

For syntactic monoids of infinite (recognizable) languages this result does not hold. For example the six element Brandt monoid $\mathbf{B_2^1}$ is the syntactic monoid of a recognizable language, but is inherently nonfinitely based [18], that is, it is not finitely based and every locally finite semigroup $\mathbf{S}$ such that $\mathbf{B_2^1} \in \mathbf{V}(\mathbf{S})$ is not finitely based as well.

The proof of Theorem 4.4 also suggests the following result.

Theorem 4.7 *If W is an n-limited set of words and tx^n, x^nt are isoterms for $S(W)$ then there is a word w such that $S(W)$ is equationally equivalent to $S(\{w\})$.*

Proof: Firstly, since W is n-limited and tx^n, x^nt are isoterms for $S(W)$, $S(W)$ is equationally equivalent to $S(W')$, where W' consists of every word in $v \in W$ replaced by t_1wt_2 (here t_1 and t_2 are new letters not occurring in v). If every word in W' is then written in a distinct alphabet and concatenated to form a word w (as in Theorem 4.4), it easily follows using Lemma 4.3 that $S(\{w\})$ is equationally equivalent to $S(W)$. $\square$

This fact motivates an obvious question.

Question 4.8 *Does there exist a finitely based (not finitely based) set of n-limited words W for which $W \cup \{tx^n, x^nt\}$ is not finitely based (or finitely based respectively).*

While a positive answer to this question would be slightly interesting, it is in a negative answer that the real interest lies since by Theorem 4.7 the finite basis problem for sets of words would then be reduced to the finite basis problem for single words.

5 Strongly not finitely based words.

Theorem 4.4 requires new letters to be introduced to ensure that a word be a subword of a finitely based word. In this section we show why this is, to an extent, a necessary feature. For example, the results of [20] show that if w is a word in a two letter alphabet such that $S(\{w\})$ is not finitely based, then there is *no* word w' in a two letter alphabet that contains w as a subword and is such that $S(\{w'\})$ is finitely based. On the other hand it is possible to show that $abcbadefge$ is a not finitely based word ($S(\{abcbadefge\})$ is easily seen to be equationally equivalent to $S(\{abcba\})$ which is shown to be not finitely based in [4]) but, as discussed in Example 4.2, $abcbadefgef$ is a finitely based word.

This behaviour appears to mimic that of the Finite Basis Problem for actual algebras, a fact that motivates the following definitions.

Definition 5.1 *If A is an alphabet and w is a word from A^* then w is said to be* weakly finitely based in A^* *if it is a subword of a finitely based word in A^*; otherwise w is said to be* strongly not finitely based in A^*. *The word w is* weakly not finitely based in A^* *if it is weakly finitely based in A^* but not finitely based and a word will be said to be* hereditarily finitely based *if it is finitely based and every subword of it is finitely based. A word w is said to be* weakly finitely based *(or* strongly not finitely based *or* weakly not finitely based*) if it is weakly finitely based (or strongly not finitely based or weakly not finitely based respectively) in $c(w)^*$.*

It follows from the main result of [20] for example that every finitely based word in $\{a, b\}^*$ is hereditarily finitely based while every not finitely based word in $\{a, b\}^*$ is strongly not finitely based. On the other hand we have seen that the word $abcbadefge$ is weakly not finitely based (and therefore weakly finitely based) in any free monoid containing $\{a, b, c, d, e, f, g\}^*$. Indeed Theorem 4.4 shows that every word is weakly finitely based in a free monoid generated by a sufficiently enlarged finite alphabet.

Strongly not finitely based words could have an important role in the finite basis problem for the discrete syntactic monoids (or equivalently the syntactic monoids) of single word languages in a fixed finite alphabet since the construction of a single example establishes the strongly not finitely based property in infinitely many others. This is similar to the notion of an inherently nonfinitely based algebra (see comments after Corollary 4.6), or more particularly the notion of a strongly not finitely based algebra introduced in [24].

Lemmas 3.2 and 3.3 provide a key connection between the main theorems of Section 3 and the problem of finding strongly not finitely based words and words with similarly pathological properties. To demonstrate how useful these lemmas are, we briefly examine an amusing "application" of the results to genetics.

The *base sequence* of a DNA molecule can be described as a word in the alphabet $\{a, c, g, t\}$ where each of the four letters correspond to particular molecules (called nucleotides). Figure 41 of [2] is an example of such a word (it is the base sequence of the single chromosome of bacteriophage ϕX174) and has an initial segment

$$ccgtcaggattgacaccctcccaattgtatgttttcatgcctccaaat\ldots$$

If we denote the entire word by w then the occurrence vectors of a, c, g, and t are $V_w(a) = (1, 1, 1, 1, \ldots)$, $V_w(c) = (2, 1, 1, 3, \ldots)$, $V_w(g) = (1, 2, 1, 1, \ldots)$, and $V_w(t) =$

$(1, 2, 1, 2, \ldots)$. Evidently the occurrence vectors of no pair of distinct letters in w are identical and so it follows from Lemma 3.2 that there is an occurrence vector of a letter that is maximal in w. Now for any pair of distinct letters (x, y) in $\{a, c, g, t\}$ it is easy to find distinct subwords in w of the form $xyy^{\beta_1}x$ and xyx, and so Theorem 3.5 implies that w is a not finitely based word.

While this example may not be of interest to geneticists, it does illustrate the ability of Theorem 3.5 and Lemma 3.2 to apply to long and complicated words.

We will now use the basic ideas of this example to show how to construct words w in any given finite alphabet A such that for any word $u \in A^+$, the word wu is not finitely based. We will call a word w with this property *left* strongly not finitely based (the existence of right strongly not finitely based words follows by symmetry).

Proposition 5.2 *Let $A = \{a_1, a_2, \ldots a_n\}$ be a finite alphabet and $w' \in A^+$ be a word with the property that for each letter a there is a letter b and such that w' contains subwords of the one of the forms described in Theorem 3.5 parts (i) to (v). Let w be the word $a_1a_2^2 \ldots a_n^n w'$. Then for any word u the word $a_1a_2^2 \ldots a_n^n w'u$ is not finitely based, that is, w is left strongly not finitely based (and $w'a_1a_2^2 \ldots a_n^n$ is right strongly not finitely based).*

Proof: The initial part $a_1a_2^2 \ldots a_n^n$ of w ensures that the occurrence vectors of each distinct pair of letters in w are distinct. Thus by Lemma 3.2, there is maximal occurrence vector of the form $V_v(a_i)$ for some letter a_i. Combining this with the chosen properties of the subword w', we have that v is not finitely based by Theorem 3.5. $\square$

The word $a_1a_2^2 \ldots a_n^n$ in the above theorem only ensures that the occurrence vector of each letter is distinct from the occurrence vector of every other letter. This can be excluded if the second part, w', of w is already guarantees this. In the alphabet $\{a, b, c\}$ for example, the word

$$w \equiv abacbbacbccb$$

is both left and right strongly not finitely based since the occurrence vectors of a, b and c can obviously never be identical in any word in the alphabet $\{a, b, c\}$ beginning or ending in w. However we have not shown that *every* word in the alphabet $\{a, b, c\}$ containing w as a subword is not finitely based. Despite Proposition 5.2 and the combined power of Theorem 3.5 and Lemmas 3.2 and 3.3, it appears to be difficult to find strongly not finitely based words in alphabets with more than two letters. As a final result we show how to construct some strongly not finitely based words in a three letter alphabet.

Theorem 5.3 *Let w_0 be a word in the alphabet $\{x, y, z\}$ such that the occurrence vectors of each of x, y and z are maximal in w_0 and w_0 contains subwords of the form $ab^\beta a$ for each pair of letters $a, b \in \{x, y, z\}$ and for some positive integer β (not necessarily identical for different choices of a and b). Let w_1 be a word in the alphabet $\{x, y, z\}$ such that for each pair of letters a, b, there are subwords of the form required by Theorem 3.5 (say ababba for example). Then w_0w_1 is strongly not finitely based.*

For example $(xyyyxxzzzxxzyxxyzzyyz)(xyxyyxxyzxzxxzzxyzyzzyyz)$ is a strongly not finitely based word (where the brackets indicate w_0 and w_1).

Proof: Let $v \in \{x, y, z\}^*$ be a word containing w. We need to show that v is not finitely based. For a given pair of letters a, b, there is a subword of w corresponding to a part of Theorem 3.5. We can therefore construct an infinite list of identities $u_n \approx v_n$ corresponding to those used to prove the relevant part of Theorem 3.5. We will often be considering several distinct pairs of letters a, b and it therefore be convenient to denote the n^{th} identity corresponding to the pair a, b by $u_n^{[a,b]} \approx v_n^{[a,b]}$.

If there is a unique maximally occurring letter a in v then the occurrence vector $V_v(a)$ is maximal and v is not finitely based by Theorem 3.5. So we may assume that v contains two maximally occurring letters. Indeed by Theorem 3.6 we may assume that either all three letters occur the same number of times in v or there are two letters occurring the same number of times in v and one letter occurring exactly one less time than the other two letters.

We can assume without loss of generality that x and y are maximally occurring letters in v and of the maximally occurring letters, x is the first to occur in v and that y is the last. We are going to examine the identities $u_n^{[x,y]} \approx v_n^{[x,y]}$ and the ways that these identities can possibly fail on $S(\{v\})$. The form of the word w ensures that there is no subword of v with length greater than 1 in v which has $occ(x, v) - 1$ distinct occurrences in v. Thus if $u_n^{[x,y]} \approx v_n^{[x,y]}$ is to fail on $S(\{v\})$ then the letter a in $u_n^{[x,y]} \approx v_n^{[x,y]}$ must be assigned a single $occ(x, v)$-occurring letter in v (that is, x or y or possibly z if $occ(z, v) = occ(x, v)$) and the letter b must be assigned one of the letters x, y, z. We will say that $u_n^{[x,y]} \approx v_n^{[x,y]}$ *fails at* (c, d) if c and d are two distinct letters in $\{x, y, z\}$ and $u_n^{[x,y]} \approx v_n^{[x,y]}$ fails on $S(\{v\})$ under some assignment that assigns a to c and b to d (and likewise for any identity $u_n^{[e,f]} \approx v_n^{[e,f]}$ where e and f are distinct elements of $\{x, y, z\}$).

There are a number of cases to consider and will address each one in the form of a lemma. In keeping with notation used in the proof of Theorem 3.5, if θ is an assignment which assigns the letters a and b in one of the words $u_n^{[x,y]}$ or $v_n^{[x,y]}$ (or similarly constructed words) to single letters in the word v then we will write $\theta({}_ia) \equiv {}_jc$ and $\theta({}_ib) \equiv {}_kd$ (where c and d are two of the letters x, y, z) if θ assigns the i^{th} occurrence of a to the j^{th} occurrence of c and the i^{th} occurrence of b to the k^{th} occurrence of d in v. It will also be convenient to use the notation ${}_ia^n$ to denote an occurrence of the subword a^n in a word u (where a is any letter) for which the first occurrence of a is the i^{th} occurrence of a in u.

The two words $u_n^{[x,y]}$ and $v_n^{[x,y]}$ are almost identical; they differ only in the part corresponding to w_1. For this reason it is convenient to denote the initial segment of v up to the start of the subword w_1 by V and the corresponding (identical) parts of $u_n^{[x,y]}$ and $v_n^{[x,y]}$ by $U_n^{[x,y]}$.

Lemma 5.4 *Let c, d be two distinct letters from $\{x, y, z\}$ which occur maximally in v and e be the third letter. The identity $u_n^{[c,d]} \approx v_n^{[c,d]}$ does not fail at (c, e).*

Proof: Assume $u_n^{[c,d]} \approx v_n^{[c,d]}$ fails at (c, e) under the assignment θ. Now for all i, $\theta({}_ia) \equiv {}_ic$ and, since w_0 is a subword of V, for some j there is a subword of V of the form ${}_jcd^\beta c$ where $\beta > 1$. This means that $U_n^{[c,d]}$ is of the form $\dots {}_jatb^\beta a \dots$ but since $\theta(b) \equiv e$ we have $\theta(U_n^{[c,d]}) \equiv \dots {}_jc \dots e^\beta \dots {}_{(j+1)}c \dots$, which contradicts the fact that $v \equiv \dots {}_jcd^\beta\, {}_{(j+1)}c$. Therefore the identity $u_n^{[c,d]} \approx v_n^{[c,d]}$ does not fail at (c, e).□

Lemma 5.5 *Let* c, d *be two distinct letters from* $\{x, y, z\}$ *which occur maximally in* v *and* e *be the third letter. The identity* $u_n^{[c,d]} \approx v_n^{[c,d]}$ *does not fail at* (e, d)*.*

Proof: Assume $u_n^{[c,d]} \approx v_n^{[c,d]}$ fails at (e, d) under the assignment θ. We use the same subword ${}_jcd^\beta c$ of the previous lemma. Let k be such that the first occurrence of d after ${}_jc$ is the k^{th} occurrence. It follows that $\theta({}_jatb^\beta a) \equiv {}_je\theta(t)\, {}_kd^\beta e$ or ${}_je\theta(t)\, {}_{(k+1)}d^\beta e$ which is inconsistent with the form of v (since a second d follows the k^{th} occurrence of d in v while the word $d^{\beta-1}c$ follows the $(k+1)^{th}$ occurrence of d in v). Therefore the identity $u_n^{[c,d]} \approx v_n^{[c,d]}$ does not fail at (e, d). □

If z occurs one less time in v than x and y then the above lemmas show that the identity $u_n^{[y,x]} \approx v_n^{[y,x]}$ cannot fail at (y, z) or (z, x) while the argument underlying Theorem 3.6 shows that it cannot fail at (x, y). This leaves only the possibility that $u_n^{[y,x]} \approx v_n^{[y,x]}$ fails at (x, z) (the case (z, y) is impossible if z occurs fewer times than y). Since x occurs before y in v, this is only possible if z occurs before x in v. Thus we may assume below that one of the following holds: z is the first of the three letters to occur in v, x is the second (distinct) letter to occur, y is the third (distinct) letter to occur and z occurs exactly $occ(x, v) - 1 = occ(y, v) - 1$ times in v; or all letters occur the same number of times and x is the first to occur, z is the second (distinct) letter to occur and y is the last distinct letter to occur in v (recall that we have assumed that x is the first maximally occurring letter to occur in v and y is the last maximally occurring letter to occur in v).

Lemma 5.6 *The identity* $u_n^{[x,y]} \approx v_n^{[x,y]}$ *does not fail at* (y, x)*.*

Proof: Assume $u_n^{[x,y]} \approx v_n^{[x,y]}$ fails at (y, x) under the assignment θ. Since x occurs before y in v, a occurs before b in $u_n^{[x,y]}$. There are $occ(x, v) - 1$ occurrences of b in $u_n^{[x,y]}$ and $v_n^{[x,y]}$ and since these all occur after the first occurrence of a, there must be $occ(x, v) - 1$ occurrences of x after the first occurrence of y in v and $\theta({}_ia) \equiv {}_iy$, $\theta({}_ib) \equiv {}_{(i+1)}x$. Thus $v \equiv z^{\gamma_0}xz^{\gamma_1}y^{\beta_1}\dots$ for some $\gamma_0, \gamma_1 \geq 0$ and $\beta_1 > 0$. Therefore $U_n^{[x,y]} \equiv t_0at_1b^{\beta_1}\dots$ (where t_0 is the empty word if $\gamma_0 = 0$). Since $\theta(a) \equiv y$ and $\theta(b) \equiv x$, we must have $\beta_1 = 1$ and $v \equiv z^{\gamma_0}xz^{\gamma_1}yz^{\gamma_2}x\dots$, where $\gamma_2 \geq 0$. Now assume that $V \equiv z^{\gamma_0}xz^{\gamma_1}yz^{\gamma_2}xz^{\gamma_3}y\dots xz^{\gamma_{2m-1}}\dots$ for some m and some $\gamma_i \geq 0$ (so that the subword w_1 in v occurs after the m^{th} occurrence of x in v). Then $U_n^{[x,y]} \equiv t_0at_1bt_2at_3b\dots at_m\dots$ (where t_{2i} is the empty word if $\gamma_{2i} = 0$) and since $\theta({}_ia) \equiv {}_iy$

and $\theta({}_ib) \equiv {}_{(i+1)}x$ we can deduce that $V \equiv z^{\gamma_0}xz^{\gamma_1}yz^{\gamma_2}xz^{\gamma_3}y\dots xz^{\gamma_m}y\dots$ and then that $U_n^{[x,y]} \equiv t_0at_1bt_2at_3b\dots at_mb\dots$. This in turn implies that

$$V \equiv z^{\gamma_0}xz^{\gamma_1}yz^{\gamma_2}xz^{\gamma_3}y\dots xz^{\gamma_{2m-1}}yz^{\gamma_{2m}}x\dots.$$

Thus by induction we arrive at the conclusion that

$$v \equiv z^{\gamma_0}xz^{\gamma_1}yz^{\gamma_2}xz^{\gamma_3}y\dots xz^{\gamma_{n-1}}yz^{\gamma_n}w_1\dots$$

or

$$v \equiv z^{\gamma_0}xz^{\gamma_1}yz^{\gamma_2}xz^{\gamma_3}y\dots xz^{\gamma_{n-1}}yz^{\gamma_n}xz^{\gamma_{n+1}}w_1\dots.$$

However this implies the occurrence vectors of x and y in w_0 are of the form $(1,1,1\dots,1)$ contradicting the fact that the occurrence vectors of each letter in w_0 is maximal. Thus $u_n^{[x,y]} \approx v_n^{[x,y]}$ does not fail at (y,x). □

Lemma 5.7 *The identity* $u_n^{[x,y]} \approx v_n^{[x,y]}$ *does not fail at* (z,x).

Proof: Assume $u_n^{[x,y]} \approx v_n^{[x,y]}$ fails at (z,x) under the assignment θ. In this case z must occur the same number of times in v as x and y do and therefore x is the first letter to occur and then z and then y. It follows that $\theta({}_ib) \equiv {}_{i+1}x$ and for some $m \geq 0$, $v \equiv xz^{\gamma_0}x^{\alpha_1}z^{\gamma_1}\dots x^{\alpha_m}z^{\gamma_m}y^{\beta_1}\dots$ (for some $\gamma_0 > 0$ and $\alpha_i, \gamma_i \geq 0$). But then $U_n^{[x,y]} \equiv at_1a^{\alpha_1}\dots a^mt_{m+1}\,{}_1b^{\beta_1}\dots$ and so m must equal 0 (since $\theta({}_1b) \equiv {}_2x$). That is, $v \equiv xz^{\gamma_0}y^{\beta_1}\dots$. Furthermore, since $\theta({}_ia) \equiv {}_iz$ we must have $\gamma_0 = 1$. We now have that $U_n^{[x,y]} \equiv at_1b^{\beta_1}\dots$ and therefore $v \equiv xzy^{\beta_1}z^{\gamma_0}x^{\beta_1}\dots$ where $\gamma_0 \geq 0$. Now if $\gamma_0 \geq 1$, then there would have to be at least two occurrences of a in $U_n^{[x,y]}$ before the first occurrence of b. This is not the case and therefore $\gamma_0 = 0$ and $v \equiv xzy^{\beta_1}x^{\beta_1}\dots$. It follows that $U_n^{[x,y]} \equiv at_1b^{\beta_1}a^{\beta_1}\dots$ and therefore $v \equiv xzy^{\beta_1}x^{\beta_1}z^{\beta_1}\dots$. We now use induction to show that (for some $\beta_i \geq 1$ and some $n \geq 1$), V is an initial segment of the word $xzy^{\beta_1}x^{\beta_1}z^{\beta_1}y^{\beta_2}x^{\beta_2}z^{\beta_2}\dots y^{\beta_n}x^{\beta_n}z^{\beta_n}$ which contradicts the choice of w_0 in the word w (which we can assume is not a subword of this word).

Assume that $V \equiv xzy^{\beta_1}x^{\beta_1}z^{\beta_1}\dots y^{\beta_k}x^{\beta_k}\,{}_iz^{\beta_k}\dots$ for some k (and some i). So $U_n^{[x,y]} \equiv at_1b^{\beta_1}a^{\beta_1}\dots t_kb^{\beta_k}a^{\beta_k}t_{k+1}\dots$. We first want to show that the value of $\theta(t_{k+1})$ is $y^{\beta_{k+1}}$ (for some $\beta_{k+1} > 0$). That is we want to show that the next letter to occur in V after the subword ${}_iz^{\beta_k}$ is the letter y. Now z cannot be the next letter to occur after the subword ${}_iz^{\beta_k}$ because this would imply that x would have to have occurred immediately after the subword ${}_ix^{\beta_k}$ (which it does not). So let us assume that x occurs after the subword ${}_iz^{\beta_k}$. Then $U_n^{[x,y]} \equiv at_1b^{\beta_1}a^{\beta_1}\dots t_kb^{\beta_k}\,{}_ia^{\beta_k}t_{k+1}\,{}_{(i+\beta_k)}a^{\alpha}\dots$, for some $\alpha > 0$. Now $\theta(t_{k+1})$ cannot contain x or z since all occurrences of a are assigned the value z and all occurrences of b are assigned the value x while there are exactly $occ(a, u_n^{[x,y]})$ occurrences of z in v and $occ(b, u_n^{[x,y]})$ of x after the first occurrence of z in v. Therefore $\theta(t_{k+1})$ must take the value 1, which implies $\theta({}_{(i+\beta_k)}a^{\alpha})$ is assigned a value starting with x, a contradiction since $\theta(a) \equiv z$. Thus the assumption that x follows ${}_iz^{\beta_k}$ in v is false and we can deduce that y is the next letter to follow ${}_iz^{\beta_k}$.

Now if

$$V \equiv xzy^{\beta_1}x^{\beta_1}\dots y^{\beta_k}x^{\beta_k}\ {}_iz^{\beta_k}y^{\beta_{k+1}}\dots$$

then

$$U_n^{[x,y]} \equiv at_1b^{\beta_1}a^{\beta_1}\dots t_kb^{\beta_k}\ {}_ia^{\beta_k}t_{k+1}b^{\beta_{k+1}}\dots$$

and since $\theta({}_ia) \equiv\ {}_iz$ and $\theta({}_ib) \equiv\ {}_{(i+1)}x$, we have that

$$V \equiv xzy^{\beta_1}x^{\beta_1}\dots y^{\beta_k}x^{\beta_k}\ {}_iz^{\beta_k}y^{\beta_{k+1}}x^{\beta_{k+1}}\dots$$

and then

$$U_n^{[x,y]} \equiv at_1b^{\beta_1}a^{\beta_1}\dots t_kb^{\beta_k}\ {}_ia^{\beta_k}t_{k+1}b^{\beta_{k+1}}a^{\beta_{k+1}}\dots$$

and consequently

$$V \equiv xzy^{\beta_1}x^{\beta_1}\dots y^{\beta_k}x^{\beta_k}\ {}_iz^{\beta_k}y^{\beta_{k+1}}x^{\beta_{k+1}}z^{\beta_{k+1}}\dots\ .$$

This completes our proof by induction that for some n, V is an initial segment of the word $xzy^{\beta_1}x^{\beta_1}z^{\beta_1}y^{\beta_2}x^{\beta_2}z^{\beta_2}\dots y^{\beta_n}x^{\beta_n}z^{\beta_n}$. Thus a contradiction has been obtained (since w_0 is not a subword of this word but is a subword of V), and we can conclude that $u_n^{[x,y]} \approx v_n^{[x,y]}$ cannot fail at the pair (z, x). □

To complete the proof of Theorem 5.3 it remains to examine the pair (y, z).

Lemma 5.8 *If for any $n > 0$ the identity $u_n^{[x,y]} \approx v_n^{[x,y]}$ fails at (y, z) then v is not finitely based.*

Proof: Assume $u_n^{[x,y]} \approx v_n^{[x,y]}$ fails at (y, z) under the assignment θ. If $occ(z, v) < occ(x, v)$ then the first letter to appear in v is the letter z and the second distinct letter to appear is the letter x and the third is y. However $\theta({}_1a) \equiv\ {}_1y$ while $\theta({}_1b)$ occurs after $\theta({}_1a) \equiv\ {}_1y$, and so we cannot have $\theta({}_1b) \equiv\ {}_1z$, a contradiction (since there are $occ(z, v)$ occurrences of b in $u_n^{[x,y]}$ and $v_n^{[x,y]}$, but only $occ(z, v) - 1$ occurrences of z in v after the first occurrence of y). Thus we may assume that $occ(z, v) = occ(x, v) = occ(y, v)$ and that x is the first letter to occur in v, z is the second (distinct) letter to occur and y is the last (distinct) letter to occur. Furthermore we may assume there is exactly one occurrence of z before the first occurrence of the letter y in v (since there must be $occ(z, v) - 1$ occurrences of z after the first occurrence of y in v). An initial segment of v is then $x^{\alpha_0}zx^{\alpha_1}y^{\beta_0}$, where $\alpha_0, \beta_0 > 1$ and $\alpha_1 \geq 0$. Thus $U_n^{[x,y]} \equiv a^{\alpha_0}t_0a^{\alpha_1}t_1b^{\beta_0}\dots$ where t_1 is the empty word if $\alpha_1 = 0$. Since $\theta({}_ia) \equiv\ {}_iy$ and $\theta({}_ib) \equiv\ {}_{(i+1)}z$, we may assume that $v \equiv x^{\alpha_0}zx^{\alpha_1}y^{\alpha_0}x^{\alpha_2}y^{\alpha_1}x^{\alpha_3}z^{\alpha_0}\dots$, where $\alpha_2, \alpha_3 \geq 0$.

Case 1: If α_1 and $\alpha_2 > 0$.
We have in this case that $U_n^{[x,y]} \equiv a^{\alpha_0}t_1a^{\alpha_1}t_2b^{\alpha_0}a^{\alpha_2}t_3b^{\alpha_1}\dots$ and therefore that

$$v \equiv x^{\alpha_0}zx^{\alpha_1}y^{\alpha_0}x^{\alpha_2}y^{\alpha_1}x^{\alpha_3}z^{\alpha_0}y^{\alpha_2}x^{\alpha_4}z^{\alpha_1}\dots,$$

where (α_3 and) $\alpha_4 \geq 0$. Note that by Lemma 3.3, either the occurrence vector of z in v is maximal or $\alpha_0 = 1$. It is now easy to verify that the identity $u_n^{[y,x]} \approx v_n^{[y,x]}$ cannot fail on $S(\{v\})$. It cannot fail at the pairs (z,x) or (y,z) by Lemmas 5.4 and 5.5. It cannot fail at the pair (x,y) or (x,z) since ${}_1x$ is the first letter to occur in v while ${}_1a$ has at least two occurrences of b occurring before it in $U_n^{[y,x]}$. Likewise it cannot fail at the pair (z,y) since there are no occurrences of y to the left of the first occurrence of z in v. Thus in any assignment θ of letters to the elements of $S(\{v\})$, either both sides of the identity $u_n^{[y,x]} \approx v_n^{[y,x]}$ are assigned the value 0 or the letter a is assigned the value 1 and therefore both sides are assigned the same value. The case is completed.

Case 2: If $\alpha_1 > 0$ and $\alpha_2 = 0$.
In this case $V \equiv x^{\alpha_0}zx^{\alpha_1}y^{\alpha_0+\alpha_1}x^{\alpha_3}z^{\alpha_0+\alpha_1}\dots$, where $\alpha_3 \geq 0$. In this case it follows from Lemma 3.3 that the occurrence vector of z is maximal in v. The case is complete.

Therefore $v \equiv x^{\alpha_0}zy^{\alpha_0}x^{\alpha_1}z^{\alpha_0}\dots$, $\alpha_1 \geq 0$. However Lemma 3.3 implies that in this case $V_v(z)$ is maximal unless $\alpha_0 = 1$. Thus we may assume that $v \equiv xzyx^{\alpha_1}z\dots$. We now examine two cases; when $\alpha_1 = 0$ and when $\alpha_1 > 0$.

Case 3: When $v \equiv xzyz^{\gamma_0}\dots$ for some $\gamma_0 \geq 1$.
If y occurs immediately after ${}_{(1+\gamma_0)}z$ in v (so that $v \equiv xzyz^{\gamma_0}y^{\beta_0}\dots$ for some $\beta_0 \geq 1$) then for every $n \geq 1$, the identity $u_n^{[y,z]} \approx v_n^{[y,z]}$ cannot fail on $S(\{v\})$ at (y,x) or (x,z) (by Lemmas 5.4 and 5.5) or at (x,y) (since x has no letters occurring left of its first occurrence whereas a in $U_n^{[y,z]}$ does). It cannot fail at (z,y) either since there are no occurrences of y before the first occurrence of z. The remaining case is the pair (z,x). Now there are at least two occurrences of z to the left of the second occurrence of y in v. However there is only the one occurrence of x left of the second occurrence of z in v. Thus $u_n^{[y,z]} \approx v_n^{[y,z]}$ cannot fail at the pair (z,x). Theorem 3.5 now shows that $S(\{v\})$ is not finitely based.

So we may assume that y does not follow after the second occurrence of z. Now if x occurs after the $(\gamma_0+1)^{th}$ occurrence of z then $v \equiv xzyz^{\gamma_0}x^{\alpha_0}\dots$ for some $\alpha_0 \geq 1$ (and $\gamma_0 \geq 1$). Therefore $U_n^{[x,y]} \equiv at_1bt_2a^{\alpha_0}\dots$ and so $V \equiv xzyz^{\gamma_0}x^{\alpha_0}y^{\alpha_0}\dots$ and $\gamma_0 = 1$. But then $U_n^{[x,y]} \equiv at_1bt_2a^{\alpha_0}t_3b^{\alpha_0}\dots$ and so $V \equiv xzyzx^{\alpha_0}y^{\alpha_0}a^{\alpha_1}z^{\alpha_0}\dots$, where $\alpha_1 \geq 0$. By Lemma 3.3, α_0 must be 1; otherwise the occurrence vector of z in v will be maximal and the subcase will be complete. That is, $V \equiv xzyzxyx^{\alpha_1}z\dots$, where $\alpha_1 \geq 0$. We show that for each $n > 1$ the identity $u_n^{[y,z]} \approx v_n^{[y,z]}$ does not fail. As in cases above, it cannot fail at the pairs (x,z) or (y,x) and taking into consideration the number of occurrences of z before the first occurrence of y in v, it cannot fail at the pair (z,y) or (x,y) either. The remaining case is the pair (z,x). However there is an occurrence of z between the first and second occurrences of y in v but no such pattern occurs for the pair (z,x). Theorem 3.5 now implies that $S(\{v\})$ is not finitely based. This completes Case 3 since with γ_0 chosen to be maximal one of x and y must follow the $(\gamma_0+1)^{th}$ occurrence of z in v.

We now examine the final case.

Case 4: When $v \equiv xzyx^{\alpha_0}z^{\gamma_0}\dots$ for some α_0, $\gamma_0 > 0$.
We use an induction argument to show that if $S(\{v\})$ is finitely based and $v \equiv xzyx^{\alpha_0}z^{\gamma_0}\dots$ then V is an initial segment of the word $(xzy)^n$ for some n. This of course contradicts the fact that V contains the word w_0 which will show that $S(\{v\})$

is not finitely based in the case we are considering.

Assume that the word v is finitely based word and that

$$V \equiv xzyxzy \dots \ _n xzyx^{\alpha_0} z^{\gamma_0} \dots ,$$

where α_0, $\gamma_0 \geq 1$. Now

$$U_n^{[x,y]} \equiv at_1bat_2b \dots \ _n at_nba^{\alpha_0}t_{n+1} \dots$$

for some n. Since $\theta({}_ia) \equiv \ {}_iy$ and $\theta({}_ib) \equiv \ {}_{(i+1)}z$, we have that

$$V \equiv xzyxzy \dots \ _n xzyx^{\alpha_0} zy^{\alpha_0} \dots$$

and then

$$U_n^{[x,y]} \equiv at_1bat_2b \dots \ _n at_nba^{\alpha_0}t_{n+1}b^{\alpha_0} \dots .$$

As above this implies that

$$V \equiv xzyxzy \dots \ _n xzyx^{\alpha_0} \ _{n+1}zy^{\alpha_0}x^{\alpha_1}z^{\alpha_0} \dots$$

for some $\alpha_1 \geq 0$. In this case, Lemma 3.3 implies that $\alpha_0 = 1$; otherwise the occurrence vector of z in v is maximal and the case is complete. Now if $\alpha_1 = 0$ we can apply an argument involving the identities $u_n^{[y,z]} \approx v_n^{[y,z]}$ that is almost identical to that used in the final subcase of Case 3 to show that $S(\{v\})$ is not finitely based: the only essential difference is that here one uses the fact that there is no occurrence of x between the $(n+1)^{th}$ occurrence of z and the $(n+2)^{th}$ occurrence of z in v (instead of the first and second occurrence of z in v). Since this contradicts the assumption that v is finitely based, we must have that $\alpha_1 > 0$ and the induction is complete. Thus the desired contradiction has been obtained.

In every case we have shown that if $u_n^{[x,y]} \approx v_n^{[x,y]}$ fails then $S(\{v\})$ is not finitely based and so the proof of the lemma is complete. □

This also completes the proof of Theorem 5.3. □

Note that while the strongly not finitely based words found in this theorem are quite long, there are no examples known to the author which suggest that much simpler words might not also be strongly not finitely based. Even the simplest candidate, the word $xyxy$ (in the alphabet $\{x, y, z\}$) is not ruled out by any of the known results to date. The author also believes that strongly not finitely based words exist in larger alphabets.

Problem 5.9 *(i) Find strongly not finitely based words in alphabets with more than three letters.*
(ii) Describe (if possible) all words w which are weakly finitely based.
(iii) Given an alphabet A, describe (if possible) all words that are weakly finitely based in A^*.

References

[1] R. M. BRYANT, The laws of finite pointed groups, *Bull. London Math. Soc.* **14** (1982), 119–123.

[2] D. R. HOFSTADTER, "Gödel, Escher, Bach: An Eternal Golden Braid," Penguin Books Ltd., England, 1987.

[3] M. JACKSON, "Small semigroup related structures with infinite properties," PhD thesis, University of Tasmania, Hobart, 1999.

[4] M. JACKSON and O. SAPIR, Finitely based sets of words, to appear in *Internat. J. Algebra Comput.*

[5] D. J. KLEITMAN, B. R. ROTHSCHILD, and J.H. SPENCER, The number of semigroups of order n, *Proc. Amer. Math. Soc.* **55** No. 1 (1976), 227–232.

[6] V. KOUBEK and V. RODL, Note on the number of monoids of order n, *Comment. Math. Univ. Carolin.* **26** (1985) 309–314.

[7] R. KRUSE, Identities satisfied in a finite ring, *J. Algebra* **26** (1973), 298–318.

[8] I. V. L'VOV, Varieties of associative rings I, *Algebra i Logika* **12** (1973), 269–297 [Russian].

[9] R. MCKENZIE, Equational bases for lattice theories, *Math. Scand.* **27** (1970), 24–38.

[10] R. MCKENZIE, Tarski's finite basis problem is undecidable, *Internat. J. Algebra Comput.* **6** (1996), 49–104.

[11] G. MASHEVITZKY, On the finite basis problem for completely 0-simple semigroup identities, *Semigroup Forum* **59** (1999), 197–219.

[12] V. L. MURSKII, Concerning the number of k-element algebras with one binary operation which have no finite basis of identities, *Problemy Kibernet* **35** (1979), 5–27.

[13] S. OATES and M. B. POWELL, Identical relations in finite groups, *J. Algebra* **1** (1964), 11–39.

[14] S. OATES-MACDONALD and M. VAUGHAN-LEE, Varieties that make one Cross, *J. Austral Math. Soc. Ser.* (A) **26** (1978), 368–382.

[15] P. PERKINS, Bases for equational theories of semigroups, *J. Algebra* **11** (1969), 298–314.

[16] J. E. PIN, "Varieties of Formal Languages," North Oxford Academic Publishers Ltd (English edition), 1986.

[17] 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.

[18] 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.

[19] M. V. SAPIR, On Cross semigroup varieties and related questions, *Semigroup Forum* **42** (1991), 345–364.

[20] O. SAPIR, Finitely Based Words, *Internat. J. Algebra Comput.* **10** (2000) 457–480.

[21] L. N. SHEVRIN and M. V. VOLKOV, Identities of semigroups, *Izv. Vyssh. Uchebn. Zaved. Mat.* (1985), **11**, 3–47 [Russian].

[22] M. VAUGHAN-LEE, Laws in finite loops, *Algebra Universalis* **9** (1979), 269–280.

[23] M. V. VOLKOV, On the identity bases of Brandt semigroups, in L. N. Shevrin (ed.), "Investigations of Algebraic Systems by Properties of Their Subsystems" (*Ural. Gos. Univ. Mat. Zap.* **14**, No. 1), Ural State University, Sverdlovsk, 1985, 38–42 [Russian].

[24] M. V. VOLKOV, The finite basis problem for finite semigroups: a survey, in E. Giraldes, P. Martin Mendes, M.P. Marques-Smith (eds.), “Semigroups”, World Scientific, Singapore, 2000.

Discipline of Mathematics
University of Tasmania
Hobart 7001
Australia
Now at
Department of Mathematics
La Trobe University
Victoria 3086
Australia