EQUATIONAL THEORIES OF UNSTABLE INVOLUTION SEMIGROUPS

EDMOND W. H. LEE

(Communicated by Alexander Olshanskii)

Abstract. It is long known that with respect to the property of having a finitely axiomatizable equational theory, there is no relationship between a general involution semigroup and its semigroup reduct. The present article establishes such a relationship within the class of involution semigroups that are unstable in the sense that the varieties they generate contain semilattices with nontrivial involution. Specifically, it is shown that the equational theory of an unstable involution semigroup is not finitely axiomatizable whenever the equational theory of its semigroup reduct satisfies the same property. Consequently, many results on equational properties of semigroups can be converted into results applicable to involution semigroups.

1. Introduction

An algebra is finitely based if its equational theory is finitely axiomatizable. The question of which algebras are finitely based—the finite basis problem—is one of the most prominent research problems in universal algebra. Apart from being very natural by itself, this problem also has a number of interesting and unexpected connections with other topics of theoretical and practical importance, for example, feasible algorithms for membership in certain classes of formal languages [1] and classical number-theoretic conjectures such as Goldbach's conjecture, the twin prime conjecture, and the odd perfect number conjecture [2].

In the 1960s, Tarski [32] explicitly posed the finite basis problem for finite algebras as a decision problem. For classical algebras such as groups [26], associative rings [13, 22], Lie algebras [4], and lattices [23], all finite members are finitely based. However, there exist algebras with as few as three elements that are non-finitely based [10, 25]. In 1969, Perkins [27] published the first examples of non-finitely based finite semigroups. Since then, the finite basis problem for semigroups, especially finite ones, has been intensely investigated. The same problem for involution semigroups has also been considered shortly after. Recall that a unary operation ∗ on a semigroup S is an involution if S satisfies the identities

\[(x^*)^* \approx x \quad \text{and} \quad (xy)^* \approx y^* x^*.
\]
An **involution semigroup** is a pair \((S, \ast)\) where \(S\) is a semigroup with involution \(\ast\). An **inverse semigroup** is an involution semigroup \((S, \ast)\) that satisfies the additional identities \(xx^\ast x \approx x\) and \(xx^\ast y y^\ast \approx yy^\ast xx^\ast\). Examples of inverse semigroups include any group \((G, \^{-1})\) with inversion \(\^{-1}\). The multiplicative \(n \times n\) matrix semigroup \((M_n, T)\) over any field with transposition \(T\) is an involution semigroup that is not an inverse semigroup. Refer to Volkov [36] for a survey of results on the finite basis problem for both semigroups and involution semigroups established since the 1960s.

The motivation for investigating involution semigroups and inverse semigroups is the same as that for semigroups—to study natural generalizations of groups—except that the involution is intended to describe the intricate symmetries of groups more accurately. Therefore it is counterintuitive that an involution semigroup \((S, \ast)\) and its semigroup **reduct** \(S\) can behave independently with respect to the finite basis property. Infinite counterexamples demonstrating this phenomenon have long been available, as observed by Volkov [36, Section 2], while finite counterexamples have only recently been discovered.

**Example 1** (Lee [17]). The \(J\)-trivial semigroup \(L = \langle e, f \mid e^2 = e, f^2 = f, ef = 0 \rangle\) of order six admits an involution \(\ast\) that fixes the generators \(e\) and \(f\). The involution semigroup \((L, \ast)\) is finitely based while its reduct \(L\) is non-finitely based.

**Example 2** (Jackson and Volkov [9, Proposition 6.3]). The Rees matrix semigroup \(JV = M^0(\{1, 2, 3\}, E, \{1, 2, 3\}; \begin{bmatrix} 0 & e & e \\ e & e & 0 \end{bmatrix})\) over the trivial group \(E = \{e\}\) is a semigroup of order 10 that admits an involution \(\ast\) given by \((i, e, j)^\ast = (j, e, i)\) and \(0^\ast = 0\). The involution semigroup \((JV, \ast)\) is non-finitely based while its reduct \(JV\) is finitely based.

In view of these counterexamples, it is instinctive to consider the **relationship problem**: with respect to the finite basis property, under what conditions will there be some relationship between an involution semigroup and its semigroup reduct? Such a relationship does exist among finite inverse semigroups.

**Theorem 3** (Kleîman [11] and Volkov [34]). Let \((S, \ast)\) be any finite inverse semigroup. Suppose that the reduct \(S\) is finitely based. Then \((S, \ast)\) is also finitely based.

This theorem remained the only positive result on the relationship problem since the 1980s. Even though inverse semigroups can be considered specialized due to their characteristics of being regular and having commuting idempotents, no significant progress on this problem was made on more general involution semigroups. The relationship problem for infinite or non-inverse involution semigroups turns out to be more tantalizing than expected. The present article addresses this problem by providing a solution for a certain wide class of involution semigroups.

### 2. Main result

Recall that a **semilattice** is a semigroup that is commutative and idempotent. An involution semilattice with nontrivial involution is called an **unstable semilattice**. Up to isomorphism, the smallest unstable semilattice is the non-chain semilattice

\[ S\ell_3 = \langle e, f \mid e^2 = e, f^2 = f, ef = fe = 0 \rangle \]
of order three with involution \( * \) that interchanges the nonzero idempotents \( e \) and \( f \). An involution semigroup \((S, *)\) is said to be unstable if the variety \( \text{Var}\{(S, *)\} \) it generates contains an unstable semilattice, otherwise it is said to be stable. Since \((\mathcal{S} \ell_3, *)\) embeds into every unstable semilattice, an involution semigroup \((S, *)\) is unstable if and only if \((\mathcal{S} \ell_3, *) \in \text{Var}\{(S, *)\}\). For instance, the involution semigroup \((\mathcal{J} \mathcal{V}, *)\) in Example 2 is unstable because its involution subsemigroup generated by the idempotent \((1, e, 2)\) is isomorphic to \((\mathcal{S} \ell_3, *)\). On the other hand, any semigroup with idempotent-fixing involution is stable—the involution semigroup \((\mathcal{L}, *)\) being an example. Inverse semigroups are also stable since the regularity axiom \(xx^*x \approx x\) is violated by \((\mathcal{S} \ell_3, *)\).

The main result of the present article is concerned with the converse of Theorem 3 in the class of unstable involution semigroups.

**Theorem 4.** Let \((S, *)\) be any unstable involution semigroup. Suppose that the reduct \( S \) is non-finitely based. Then \((S, *)\) is also non-finitely based.

Now since the variety \( \text{Var}\{(\mathcal{S} \ell_3, *)\} \) is an atom in the lattice of all varieties of involution semigroups [8], the unstableness assumption in Theorem 4 is a weakest assumption—the only way to weaken it is to omit it altogether. On the other hand, unstableness is essential to this theorem since the stable involution semigroup \((\mathcal{L}, *)\) serves as a counterexample. Further, as demonstrated by the involution semigroup \((\mathcal{J} \mathcal{V}, *)\), the converse of Theorem 4 does not hold for all unstable involution semigroups. Consequently, Theorem 4 has reached its full potential.

After some background material is given in Section 3, it is shown in Section 4 that each unstable involution semigroup possesses some basis of identities of very specific form. This result is then used in Section 5 to establish the proof of Theorem 4.

In general, any involution semigroup \((S, *)\) that is not already unstable is embeddable in the unstable involution semigroup \((S, *) \times (\mathcal{S} \ell_3, *)\). Hence unstable involution semigroups constitute a large and general class of involution semigroups, and many results on equational properties of semigroups can be converted by Theorem 4 into results applicable to involution semigroups. It is infeasible to provide full details in all such instances. Therefore only a few representative examples are recorded in Section 6, while other results will be surveyed or developed elsewhere.

### 3. Preliminaries

Acquaintance with rudiments of universal algebra is assumed of the reader. Refer to the monograph of Burris and Sankappanavar [6] for more information.

#### 3.1. Terms and words.

Let \( \mathcal{A} \) be a countably infinite alphabet throughout and let \( \mathcal{A}^* = \{ x^* \mid x \in \mathcal{A} \} \) be a disjoint copy of \( \mathcal{A} \). Elements of \( \mathcal{A} \cup \mathcal{A}^* \) are called variables. The set \( T(\mathcal{A}) \) of terms over \( \mathcal{A} \) is the smallest set containing \( \mathcal{A} \) that is closed under concatenation and \( * \). The subterms of a term \( s \) are defined as follows: \( s \) is a subterm of \( s \); if \( s_1s_2 \) is a subterm of \( s \) where \( s_1, s_2 \in T(\mathcal{A}) \), then \( s_1 \) and \( s_2 \) are subterms of \( s \); if \( t^* \) is a subterm of \( s \) where \( t \in T(\mathcal{A}) \), then \( t \) is a subterm of \( s \).

Elements of the free monoid \( (\mathcal{A} \cup \mathcal{A}^*)^* \cup \{ \emptyset \} \) are called words while words in \( \mathcal{A}^* \cup \{ \emptyset \} \) are called plain words. The proper inclusion \((\mathcal{A} \cup \mathcal{A}^*)^+ \subset T(\mathcal{A}) \) holds and the identities \((0)\) can be used to convert any term \( t \in T(\mathcal{A}) \) into some unique word \( |s| \in (\mathcal{A} \cup \mathcal{A}^*)^+ \cup \{ \emptyset \} \). For instance, \( [x(x^3(yx^*)^*)^*zy^*] = xy(x^3)^4zy^* \).

**Remark 5.** For any subterm \( s \) of a term \( t \), either \( |s| \) or \( [s^*] \) is a subword of \( |t| \).
3.2. Identities and deducibility. In general, an identity is an expression \( s \approx t \) formed by terms \( s, t \in T(\mathcal{A}) \). A word identity is an identity \( u \approx v \) formed by words \( u, v \in (\mathcal{A} \cup \mathcal{A}^*)^+ \), while a plain identity is an identity \( p \approx q \) formed by plain words \( p, q \in \mathcal{A}^+ \). For any involution semigroup \( (S, \ast) \), let \( \text{id}((S, \ast)) \) denote the set of identities satisfied by \( (S, \ast) \), commonly called the \textit{equational theory} of \( (S, \ast) \). Let \( \text{id}_W((S, \ast)) \) and \( \text{id}_P((S, \ast)) \) denote the subsets of \( \text{id}((S, \ast)) \) consisting of word identities and plain identities, respectively. It is clear that the inclusions

\[
\text{id}((S, \ast)) \supset \text{id}_W((S, \ast)) \supset \text{id}_P((S, \ast)) = \text{id}(S)
\]

hold, where \( \text{id}(S) \) is the equational theory of the semigroup \( S \).

In the involution semigroup signature, an identity \( s \approx t \) is \textit{directly deducible} from an identity \( s' \approx t' \) if there exists some substitution \( \varphi : \mathcal{A} \rightarrow T(\mathcal{A}) \) such that \( s' \varphi \) is a subterm of \( s \), and replacing this particular subterm \( s' \varphi \) of \( s \) with \( t' \varphi \) results in the term \( t \). An identity \( s \approx t \) is \textit{deducible} from some set \( \Sigma \) of identities if there exists a sequence \( s = s_1, s_2, \ldots, s_r = t \) of terms such that each identity \( s_i \approx s_{i+1} \) is directly deducible from some identity in \( \Sigma \).

In the semigroup signature, deducibility is defined analogously, but direct deducibility can be given more precisely: a plain identity \( p \approx q \) is \textit{directly deducible} from a plain identity \( p' \approx q' \) if there exist some substitution \( \varphi : \mathcal{A} \rightarrow \mathcal{A}^+ \) and plain words \( a, b \in \mathcal{A}^+ \cup \{ \varepsilon \} \) such that \( p = a(p' \varphi) b \) and \( q = a(q' \varphi) b \).

In general, by Birkhoff’s completeness theorem of equational logic [5], an identity is satisfied by an algebra \( \mathfrak{A} \) if and only if it is deducible from \( \text{id}(\mathfrak{A}) \).

Remark 6. An identity \( s \approx t \) is deducible from \( (0) \) if and only if \( |s| = |t| \).

3.3. Bases and varieties. For any algebra \( \mathfrak{A} \), a subset \( \Sigma \) of \( \text{id}(\mathfrak{A}) \) is a \textit{basis} for \( \mathfrak{A} \) if every identity in \( \text{id}(\mathfrak{A}) \) is deducible from \( \Sigma \). An algebra is \textit{finitely based} if it has some finite basis. A class of algebras of a fixed type is a \textit{variety} if it is closed under the formation of homomorphic images, subalgebras, and arbitrary direct products. The \textit{variety generated} by an algebra \( \mathfrak{A} \), denoted by \( \text{Var}\{\mathfrak{A}\} \), is the smallest variety containing \( \mathfrak{A} \). A variety satisfies an identity if every one of its algebras satisfies the identity. Any algebra \( \mathfrak{A} \) and the variety \( \text{Var}\{\mathfrak{A}\} \) satisfy the same identities.

For any word \( u \in (\mathcal{A} \cup \mathcal{A}^*)^+ \), let \( \overline{u} \) denote the word obtained by writing \( u \) in reverse. For instance, if \( u = x^a y^b (z^c)^d x^e \), then \( \overline{u} = x^e (z^c)^d y^b x^a \). A set \( \Sigma \) of word identities is \textit{symmetric} if \( \overline{u} \approx \overline{v} \in \Sigma \) for all \( u \approx v \in \Sigma \). A basis \( \Sigma \) for an involution semigroup is \textit{symmetric} if \( \Sigma \) is a symmetric set of word identities.

Lemma 7.

(i) Each involution semigroup has some symmetric basis.

(ii) Each finitely based involution semigroup has some finite symmetric basis.

Proof. (i) Let \( \Sigma \) be any basis for an involution semigroup \( (S, \ast) \). Since the identities \( (0) \) can be used to convert any term \( s \) into the unique word \( |s| \), the set \( |\Sigma| = \{|s| \approx t | s \approx t \in \Sigma\} \) of word identities is also a basis for \( (S, \ast) \). Therefore generality is not lost by assuming that all identities in \( \Sigma \) are word identities to begin with, that is, \( |\Sigma| = \Sigma \). Now if \( (S, \ast) \) satisfies a word identity \( u \approx v \), then it satisfies the identity \( u \varphi \approx v \varphi \), where \( \varphi \) is the substitution \( x \varphi = x^\ast \) for all \( x \in \mathcal{A} \), whence it also satisfies \( \overline{u} \approx (u \varphi)^* \approx (v \varphi)^* = \overline{v} \). It follows that \( \Sigma \cup \overline{\Sigma} \), where \( \overline{\Sigma} = \{ u \approx \overline{v} | u \approx v \in \Sigma\} \), is a symmetric basis for \( (S, \ast) \).

(ii) If the basis \( \Sigma \) in part (i) is finite, then so is the symmetric basis \( \Sigma \cup \overline{\Sigma} \). □
4. Organized bases

In the present section, it is shown that each unstable involution semigroup has some basis of identities of very specific form. This result is essential to the proof of Theorem 4 in Section 5.

Let $u \in (\mathcal{A} \cup \mathcal{A}^*)^+$ be any word. The content of $u$, denoted by $\text{con}(u)$, is the set of variables occurring in $u$. If $x, x^* \in \text{con}(u)$ for some $x \in \mathcal{A}$, then $\{x, x^*\}$ is called a mixed pair of $u$. A word is said to be mixed if it has some mixed pair. A word without mixed pairs is said to be bipartite because its content can be partitioned into $\mathcal{X}$ and $\mathcal{Y}^*$ for some disjoint finite sets $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{A}$. A word identity $u \approx v$ is mixed if both $u$ and $v$ are mixed, bipartite if both $u$ and $v$ are bipartite, and homogeneous if $\text{con}(u) = \text{con}(v)$. A word identity that is both homogeneous and bipartite is said to be homogeneously bipartite.

Lemma 8. Suppose that $u \approx v \in \text{id}_W(\{S_{s_3}, *\})$. Then

(i) $u$ is mixed if and only if $v$ is mixed;
(ii) $\text{con}(u) = \text{con}(v)$ whenever $u$ and $v$ are bipartite.

Consequently, each word identity satisfied by an unstable involution semigroup is either mixed or homogeneously bipartite.

Proof. (i) Seeking a contradiction, suppose that the word $u$ is bipartite while the word $v$ is mixed. Then $\text{con}(u) = \mathcal{X} \cup \mathcal{Y}^*$ for some finite sets $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{A}$ such that $\mathcal{X} \cap \mathcal{Y} = \emptyset$. Let $\varphi : \mathcal{A} \to S_{s_3}$ denote the substitution given by $x\varphi = e$ for all $x \in \mathcal{X}$ and $x\varphi = f$ for all $x \in \mathcal{A}\setminus\mathcal{X}$. Then $u\varphi = e$ and $v\varphi = 0$.

(ii) Seeking a contradiction, suppose that the words $u$ and $v$ are bipartite with $z \in \text{con}(v) \setminus \text{con}(u)$. Then $\text{con}(u) = \mathcal{X} \cup \mathcal{Y}^*$ for some finite sets $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{A}$ such that $\mathcal{X} \cap \mathcal{Y} = \emptyset$. Let $\varphi : \mathcal{A} \to S_{s_3}$ denote the substitution given by $x\varphi = e$ for all $x \in \mathcal{X}$, $x\varphi = f$ for all $x \in \mathcal{Y}$, and $x\varphi = 0$ for all $x \in \mathcal{A}\setminus(\mathcal{X} \cup \mathcal{Y})$. Then $u\varphi = e$ and $v\varphi = 0$.

The plain projection of a word $u \in (\mathcal{A} \cup \mathcal{A}^*)^+$, denoted by $\pi(u)$, is the plain word obtained from $u$ by removing all occurrences of the symbol $*$. For instance, if $u = xy^2y^*z^*(x^*)^3z$, then $\pi(u) = xy^3zx^4$. The plain projection of a word identity $u \approx v$ is $\pi(u) \approx \pi(v)$.

Lemma 9. An involution semigroup $(S, *)$ satisfies a homogeneously bipartite identity $u \approx v$ if and only if it satisfies the plain projection $\pi(u) \approx \pi(v)$.

Proof. By assumption, $\text{con}(u) = \text{con}(v) = \mathcal{X} \cup \mathcal{Y}^*$ for some finite sets $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{A}$ such that $\mathcal{X} \cap \mathcal{Y} = \emptyset$. Let $\varphi$ denote the substitution $x\varphi = x^*$ for all $x \in \mathcal{Y}$. If $(S, *)$ satisfies $u \approx v$, then it satisfies $\pi(u) = |u\varphi| \approx |v\varphi| = \pi(v)$. Conversely, if $(S, *)$ satisfies $\pi(u) \approx \pi(v)$, then it satisfies $u = (\pi(u))\varphi \approx (\pi(v))\varphi = v$.

A symmetric basis $\Sigma$ for an involution semigroup is organized if each identity in $\Sigma$ is either mixed or plain. In other words, an organized basis for an involution semigroup $(S, *)$ is a symmetric basis of the form

$$\Sigma = \Sigma_{\text{mix}} \cup \Sigma_p,$$

where $\Sigma_{\text{mix}} \subseteq \text{id}_W(\{(S, *)\})$ consists of mixed identities and $\Sigma_p \subseteq \text{id}_P(\{(S, *)\})$.

Lemma 10. Let $\Sigma$ be any symmetric basis for an unstable involution semigroup $(S, *)$ and let $\Sigma'$ be the set obtained from $\Sigma$ by replacing every bipartite identity with its plain projection. Then $\Sigma'$ is an organized basis for $(S, *)$. 

Proof. It follows from Lemma 8 that $\Sigma = \Sigma_{\text{mix}} \cup \Sigma_{\text{HB}}$, where $\Sigma_{\text{mix}}$ consists of mixed identities in $\Sigma$ and $\Sigma_{\text{HB}}$ consists of homogeneously bipartite identities in $\Sigma$. Let $\Sigma_p = \{ \pi(u) = \pi(v) \mid u \approx v \in \Sigma_{\text{HB}} \}$ so that $\Sigma' = \Sigma_{\text{mix}} \cup \Sigma_p$. Then by Lemma 9, the set $\Sigma'$ is a basis for $(S, \ast)$; this basis is symmetric and thus also organized.

5. PROOF OF THEOREM 4

Lemma 11. Let $(S, \ast)$ be any unstable involution semigroup with organized basis $\Sigma = \Sigma_{\text{mix}} \cup \Sigma_p$. Suppose that $s \approx t$ is any identity directly deducible from some identity in $(0) \cup \Sigma$ with $[s], [t] \in \mathcal{A}^\ast$. Then the plain identity $[s] \approx [t]$ is directly deducible from some identity in $\Sigma_p$.

Proof. By assumption, $s \approx t$ is directly deducible from some $u \approx v \in (0) \cup \Sigma$. If $u \approx v \in (0)$, then as observed in Remark 6, the identity $[s] \approx [t]$ is trivial and so is vacuously directly deducible from some identity in $\Sigma_p$. Therefore assume that $u \approx v \in \Sigma$. Then there exists a substitution $\varphi : \mathcal{A} \to T(\mathcal{A})$ such that $u \varphi$ is a subterm of $s$, and replacing this particular subterm $u \varphi$ of $s$ with $v \varphi$ results in the term $t$. By Remark 5, either $[u \varphi]$ or $[(u \varphi)']$ is a subword of the plain word $[s]$.

Case 1: $[u \varphi]$ is a subword of $[s]$. Then $[s] = a[u(x_1 \varphi, x_2 \varphi, \ldots, x_m \varphi)]b$ for some $a, b \in \mathcal{A}^\ast \cup \{\emptyset\}$. Since $t$ is obtained by replacing $u \varphi$ in $s$ with $v \varphi$, it follows that $[t] = a[v \varphi]b$. If $u$ has a mixed pair, then $[u \varphi]$ has a mixed pair, contradicting the plainness of $[s]$. Therefore $u$ is bipartite, so that $u \approx v \in \Sigma_p$. Hence $\text{con}(u) = \text{con}(v)$ by Lemma 8(ii), say $u = u(x_1, x_2, \ldots, x_m)$ and $v = v(x_1, x_2, \ldots, x_m)$ each involves all of the variables $x_1, x_2, \ldots, x_m \in \mathcal{A}$. Since the words $[s]$ and $u$ are plain with

$$[s] = a[u(x_1 \varphi, x_2 \varphi, \ldots, x_m \varphi)]b,$$

the words $[x_1 \varphi], [x_2 \varphi], \ldots, [x_m \varphi]$ are also plain. Let $\chi : \text{con}(u) \to \mathcal{A}^\ast$ denote the substitution given by $x_i \chi = [x_i \varphi]$ for all $i \in \{1, 2, \ldots, m\}$. Then $[s] = a(u \chi) b$ and $[t] = a(v \chi) b$. Hence the identity $[s] \approx [t]$ is directly deducible from $u \approx v \in \Sigma_p$.

Case 2: $[(u \varphi)']$ is a subword of $[s]$. Then following the argument in Case 1 yields $[s] = a[(u \varphi)'] b$ and $[t] = a(v \varphi) b$ for some $a, b \in \mathcal{A}^\ast \cup \{\emptyset\}$, and $u \approx v \in \Sigma_p$ with $\text{con}(u) = \text{con}(v)$, say $u = u(x_1, x_2, \ldots, x_m)$ and $v = v(x_1, x_2, \ldots, x_m)$ each involves all of the variables $x_1, x_2, \ldots, x_m \in \mathcal{A}$. Since the words $[s]$ and $u$ are plain with

$$[s] = a[(u(x_1 \varphi, x_2 \varphi, \ldots, x_m \varphi)')] b,$$

the words $[(x_1 \varphi)], [(x_2 \varphi)], \ldots, [(x_m \varphi)]$ are also plain. Let $\chi : \text{con}(u) \to \mathcal{A}^\ast$ denote the substitution given by $x_i \chi = [(x_i \varphi)]$ for all $i \in \{1, 2, \ldots, m\}$. Then $[s] = a(\overline{u} \chi) b$ and $[t] = a(\overline{v} \chi) b$. Therefore the identity $[s] \approx [t]$ is directly deducible from $u \approx v \in \Sigma$ and $\Sigma$ is symmetric, $u \approx v \in \Sigma_p$. □

Lemma 12. Suppose that $(S, \ast)$ is any unstable involution semigroup with organized basis $\Sigma = \Sigma_{\text{mix}} \cup \Sigma_p$. Then $\Sigma_p$ is a basis for the reduct $S$.

Proof. Consider any $p \approx q \in \text{id}(\Sigma)$. Since $p \approx q$ is deducible from $(0) \cup \Sigma$, there exists a sequence $p = s_0, s_1, \ldots, s_r = q$ of terms such that each identity $s_i \approx s_{i+1}$ is directly deducible from some identity in $(0) \cup \Sigma$. It is clear that $[s_1] = p \in \mathcal{A}^\ast$.\[\]

EQUATIONAL THEORIES OF UNSTABLE INVOLUTION SEMIGROUPS 15
Suppose that $|s_i| \in \mathcal{A}^+$ for some $i \geq 1$. Then since $|s_i| \approx |s_{i+1}| \in \mathcal{A}^+$, it follows from Lemma 8(ii) that $|s_{i+1}| \in \mathcal{A}^+$. Therefore by Lemma 11, the plain identity $|s_i| \approx |s_{i+1}|$ is directly deducible from some identity in $\Sigma_p$. It follows that $p = [s_1], [s_2], \ldots, [s_r] = q$ is a sequence of plain words such that each identity $|s_i| \approx |s_{i+1}|$ is directly deducible from some identity in $\Sigma_q$. Consequently, the identity $p \approx q$ is deducible from $\Sigma_q$. \hfill \Box

Now suppose that $(S, *)$ is any unstable involution semigroup that is finitely based. Then by Lemma 7(ii), there exists some finite symmetric basis for $(S, *)$, from which a finite organized basis $\Sigma = \Sigma_{\text{mix}} \cup \Sigma_p$ can be constructed, as shown in Lemma 10. By Lemma 12, the set $\Sigma_p$ is a finite basis for the reduct $S$.

**6. Applications of Theorem 4**

6.1. **Stable involution semigroups.** Example 1 demonstrates that Theorem 4 is not applicable to an involution semigroup $(S, *)$ in general. But the theorem is always applicable to the unstable involution semigroup $(S, *) \times (\mathcal{S}_3, *)$.

**Theorem 13.** Let $(S, *)$ be any involution semigroup. Suppose that the reduct $S$ is non-finitely based. Then either $(S, *)$ or $(S, *) \times (\mathcal{S}_3, *)$ is non-finitely based.

In fact, each involution semigroup $(S, *)$ is embeddable in some unstable involution semigroup that is simpler than the direct product $(S, *) \times (\mathcal{S}_3, *)$: the amalgamation $(S \cup \mathcal{S}_3, *)$ of $(S, *)$ and $(\mathcal{S}_3, *)$, where for any $x \in S$ and $y \in \mathcal{S}_3$, the products $xy$ and $yx$ are equal to 0 in $\mathcal{S}_3$. Therefore Theorem 13 also holds when $(S, *) \times (\mathcal{S}_3, *)$ is replaced by $(S \cup \mathcal{S}_3, *)$.

**Corollary 14.** Any involution semigroup with non-finitely based semigroup reduct is embeddable in some non-finitely based involution semigroup with at most three more elements.

6.2. **Inherent non-finite basis property.** A locally finite variety is inherently non-finitely based if every locally finite variety containing it is non-finitely based. An algebra is inherently non-finitely based if it generates an inherently non-finitely based variety. A structural description of inherently non-finitely based finite semigroups, due to Sapir [29], yields an algorithm that decides if a finite semigroup is inherently non-finitely based. However, this description does not generalize to all involution semigroups. The semigroup reduct of any inherently non-finitely based involution semigroup is indeed inherently non-finitely based [2], but the converse of this result is false in general, as demonstrated by the inverse semigroup $(\mathcal{B}_1, -1)$, where $\mathcal{B}_1 = \mathcal{B}_2 \cup \{1\}$ is the monoid obtained from the Brandt semigroup

$$\mathcal{B}_2 = \langle a, b \mid a^2 = b^2 = 0, aba = a, bab = b \rangle$$

of order five, with involution $^{-1}$ that interchanges the generators $a$ and $b$.

**Example 15** (Sapir [30, 31]). The inverse semigroup $(\mathcal{B}_2, -1)$ is not inherently non-finitely based while its reduct $\mathcal{B}_1$ is inherently non-finitely based.

In fact, inherently non-finitely based finite inverse semigroups do not exist [31]. Nevertheless, using a general sufficient condition [3, Theorem 2.3] that required a rather involved proof, Auinger et al. [2] established a result on inherently non-finitely based involution semigroups that parallels Theorem 4.
Theorem 16 (Auinger et al. [2, Theorem 3.1]). Let \((S, \ast)\) be any unstable involution semigroup. Suppose that the reduct \(S\) is inherently non-finitely based. Then \((S, \ast)\) is also inherently non-finitely based.

This theorem turns out to be an easy consequence of Theorem 4. Suppose that \((S, \ast)\) is any unstable involution semigroup that is not inherently non-finitely based. Then \((S, \ast)\) belongs to some finitely based locally finite variety \(\mathsf{Var}\{(U, \ast)\}\). Since \((U, \ast)\) is unstable and finitely based, its reduct \(U\) is finitely based by Theorem 4. Therefore the reduct \(S\) belongs to the finitely based locally finite variety \(\mathsf{Var}\{U\}\) and so is not inherently non-finitely based.

Although the inverse semigroup \(\mathcal{B}_1^2, -1\) is not inherently non-finitely based, it can be used to provide many inherently non-finitely based involution semigroups.

Proposition 17. The direct product of \((\mathcal{B}_1^2, -1)\) with any finite unstable involution semigroup is inherently non-finitely based.

Proof. The direct product \((\mathcal{B}_1^2, -1) \times (S, \ast)\) of \((\mathcal{B}_1^2, -1)\) with any finite unstable involution semigroup \((S, \ast)\) is vacuously unstable. Since the semigroup \(\mathcal{B}_1^2\) is inherently non-finitely based [30], the reduct \(\mathcal{B}_1^2 \times S\) is also inherently non-finitely based. The result then follows from Theorem 16.

Example 18 (Auinger et al. [3, Corollary 2.7]). The involution semigroup \((\mathcal{B}_1^2, \ast)\) is inherently non-finitely based.

6.3. Weak finite basis property. A finite algebra that is not inherently non-finitely based is weakly finitely based. It follows from Sapir [30, Proposition 7] that the class \(\mathcal{W}\) of weakly finitely based semigroups is a pseudovariety, that is, \(\mathcal{W}\) is closed under the formation of homomorphic images, subsemigroups, and finitary direct products. In contrast, this result does not hold for involution semigroups.

Proposition 19. The class of weakly finitely based involution semigroups is not a pseudovariety.

Proof. The involution semigroup \((\mathcal{B}_1^2, -1)\) is weakly finitely based [31] and the unstable semilattice \((\mathcal{S} \ell_3, \ast)\) is even finitely based [7], but by Proposition 17, the direct product \((\mathcal{B}_1^2, -1) \times (\mathcal{S} \ell_3, \ast)\) is not weakly finitely based.

6.4. Non-finite basis property: sufficient conditions. Since the investigation of the finite basis problem for semigroups began in the 1960s, many sufficient conditions have been established under which semigroups are non-finitely based. By Theorem 4, each of these sufficient conditions is also applicable to any unstable involution semigroup \((S, \ast)\): any condition that guarantees the non-finite basis property of the reduct \(S\) also guarantees the same property for \((S, \ast)\). A sufficient condition of Volkov [35], the statement of which requires the 0-simple semigroup

\[\mathcal{A}_2 = \langle a, b \mid a^2 = aba = a, bab = b, b^2 = 0 \rangle\]

of order five, serves as a prime example. Recall that the core of a semigroup \(S\), denoted by \(\mathcal{C}(S)\), is the subsemigroup of \(S\) generated by its idempotents.
Suppose that \( \text{Var}\{S\} \setminus \text{Var}\{\mathcal{C}(S)\} \) contains some group. Then \( S \) is non-finitely based.

This theorem has been applied to establish the non-finite basis property of many semigroups, including semigroups of transformations, semigroups of binary relations, semigroups of matrices, and the direct product of \( A_2 \) with any nontrivial finite group. A version of this theorem that is applicable to finite unstable involution semigroups is easily deduced from Theorem 4.

\section*{Theorem 21.} Let \((S, *)\) be any finite unstable involution semigroup such that \( A_2 \in \text{Var}\{S\} \). Suppose that \( \text{Var}\{S\} \setminus \text{Var}\{\mathcal{C}(S)\} \) contains some group. Then \((S, *)\) is non-finitely based.

Now, the 0-simple semigroup \( A_2 \) admits an involution \(*\) that fixes the generators \( a \) and \( b \). The involution semigroup \((A_2, *)\) is unstable because its involution subsemigroup generated by \( ab \) modulo the ideal \( \{0, b\} \) is isomorphic to \((S \ell_3, *)\).

Therefore the following result is a particular consequence of Theorem 21.

\section*{Proposition 22.} The direct product of \((A_2, *)\) with any nontrivial finite involution group is non-finitely based.

It is of interest to note that both the involution semigroup \((A_2, *)\) [9] and all finite involution groups [26] are finitely based. Similarly, the involution semigroup \((\mathcal{L}, *)\) and any finite cyclic group with nontrivial involution are finitely based, but their direct product is non-finitely based [16, Corollary 1.5(iii)].

\section*{Proposition 23.} The involution semigroup \((\mathcal{L}, *)\) is the unique smallest finitely based involution semigroup with non-finitely based semigroup reduct.

\section*{Proof.} Up to isomorphism, the \( \mathcal{J} \)-trivial semigroup \( \mathcal{L} \), the Brandt monoid \( \mathcal{B}_1^2 \), the monoid \( A_2^1 = \mathcal{A}_2 \cup \{1\} \), and the semigroup \( A_2^2 = \mathcal{A}_2 \cup \{g\} \) with \( gA_2 = A_2g = \{g\} \) and \( g^2 = 0 \) are precisely all non-finitely based semigroups of order six [19, 20, 21].

The only involution semigroups with semigroup reducts from \( \{\mathcal{L}, \mathcal{B}_1^1, A_2^1, A_2^2\} \) are \((\mathcal{L}, *), (\mathcal{B}_1^1, 1), (\mathcal{B}_1^1, *), (A_2^1, *), (A_2^2, *), (A_2^2, 1), (A_2^2, *), (A_2^2, 1), (A_2^2, *)\) with \( 1^* = 1 \), and \((A_2^2, *), (A_2^2, *)\) with \( g^* = g \). The inverse semigroup \((\mathcal{B}_1^1, 1)\) is non-finitely based by Kleiman [12], while the involution semigroups \((\mathcal{B}_1^1, *), (A_2^1, *), (A_2^2, *), (A_2^2, *)\) are non-finitely based by Theorems 16 and 21. Since \((\mathcal{L}, *)\) is a finitely based involution semigroup [17], it is the only one of order six with non-finitely based semigroup reduct. But there are no smaller examples since all semigroups of order up to five are finitely based [14, 33]. \(\square\)

\section*{6.5. Neutral finite basis property.} A finitely based algebra \( \mathfrak{A} \) is \textit{neutral finitely based} if for any algebra \( \mathfrak{B} \) of the same type, the direct product \( \mathfrak{A} \times \mathfrak{B} \) is finitely based if and only if \( \mathfrak{B} \) is finitely based. Every semilattice is a neutral finitely based semigroup [24]. In contrast, even though every involution semilattice is finitely based [7], it is shown below that not every one of them is neutral finitely based.

\section*{Theorem 24 (Lee [18]).} Let \( S \) be any semigroup such that \( \mathcal{L} \in \text{Var}\{S\} \). Suppose that for some fixed \( n \geq 2 \), the semigroup \( S \) satisfies the identities \( x^{2n} \approx x^n \) and \( x^n y_1^n y_2^n \cdots y_k^n x^n \approx x^n y_k^n y_{k-1} \cdots y_1^n x^n \) for all \( k \geq 2 \). Then \( S \) is non-finitely based.

Since this theorem is a sufficient condition for the non-finite basis property of semigroups, it can be converted by Theorem 4, in a manner similar to Theorem 21, into a result applicable to involution semigroups. It is also possible—and more
appropriate for the purpose of the present subsection—to exhibit an interval of non-finitely based varieties of involution semigroups. Let $\mathbf{LS} = \operatorname{Var}\{ (\mathcal{L}, \ast), (\mathcal{S}_{\ell_3}, \ast) \}$ and, for each fixed $n \geq 2$, let $\mathbf{V}_n$ denote the variety of all involution semigroups that satisfy all the identities in Theorem 24. Since these identities are satisfied by the semigroup $\mathcal{L}$ and all semilattices [15], the inclusion $\mathbf{LS} \subseteq \mathbf{V}_n$ holds.

**Proposition 25.** Every variety in the interval $[\mathbf{LS}, \mathbf{V}_n]$ is non-finitely based.

**Proof.** Let $\mathbf{U} = \operatorname{Var}\{ (U, \ast) \}$ be any variety in $[\mathbf{LS}, \mathbf{V}_n]$. Then $(\mathcal{L}, \ast), (\mathcal{S}_{\ell_3}, \ast) \in \mathbf{U}$, so that $(U, \ast)$ is unstable and $\mathcal{L} \in \operatorname{Var}\{U\}$. Further, since $(U, \ast) \in \mathbf{V}_n$, the semigroup $U$ satisfies the identities in Theorem 24 and so is non-finitely based. Hence by Theorem 4, the unstable involution semigroup $(U, \ast)$, and so also the variety $\mathbf{U}$, are non-finitely based. \hfill $\Box$

**Theorem 26.** Every unstable semilattice is not neutrally finitely based.

**Proof.** The involution semigroup $(\mathcal{L}, \ast)$ is finitely based [17]. But for any unstable semilattice $(U, \ast)$, the direct product $(\mathcal{L}, \ast) \times (U, \ast)$ generates a variety in the interval $[\mathbf{LS}, \mathbf{V}_n]$ and so is non-finitely based by Proposition 25. \hfill $\Box$

**References**


DEPARTMENT OF MATHEMATICS, NOVA SOUTHEASTERN UNIVERSITY, 3301 COLLEGE AVENUE, FORT LAUDERDALE, FLORIDA 33314, USA
E-mail address: edmond.lee@nova.edu