Bull. Aust. Math. Soc. **109** (2024), 350–364
doi:10.1017/S0004972723000989

FINITE BASIS PROBLEM FOR INVOLUTION MONOIDS OF ORDER FIVE

BIN BIN HAN, WEN TING ZHANG ✉ and YAN FENG LUO

(Received 15 March 2023; accepted 19 August 2023; first published online 9 October 2023)

Abstract

An example of a nonfinitely based involution monoid of order five has recently been discovered. We confirm that this example is, up to isomorphism, the unique smallest among all involution monoids.

2020 *Mathematics subject classification*: primary 20M05.

Keywords and phrases: monoid, involution, identity basis, finite basis problem.

1. Introduction

An algebra is *finitely based* if the identities it satisfies are finitely axiomatisable; otherwise, it is *nonfinitely based*. The celebrated theorem of Oates and Powell [20], published in 1964, states that all finite groups are finitely based. In the decade that followed, finite members from other classes of algebras such as lattices [19], associative rings [7, 8] and Lie rings [2] were also shown to be finitely based. However, this is not true in general. In the 1960s, Perkins [21] published the first examples of nonfinitely based finite semigroups, one of which is the well-known Brandt monoid B_2^1 of order six. The discovery of this example focused attention upon the finite basis problem for small semigroups. In particular, is there a nonfinitely based semigroup of order less than six? After several decades of cumulative work, a complete solution has been found for all semigroups of order up to six: every semigroup of order five or less is finitely based [9, 22] and there are only four nonfinitely based semigroups of order six (including B_2^1) up to isomorphism [16–18].

This paper is concerned with *involution semigroups*, that is, unary semigroups $(S, {}^*)$ that satisfy the identities

$$(x^*)^* \approx x \quad \text{and} \quad (xy)^* \approx y^*x^*; \tag{1.1}$$

The authors are partially supported by the National Natural Science Foundation of China (Nos. 12271224, 12171213, 12161062) and the Fundamental Research Funds for the Central University (No. lzujbky-2023-ey06).

the unary operation * is an *involution* of S, and S is the semigroup *reduct* of $(S, ^*)$. Common examples of involution semigroups include groups $(G, ^{-1})$ under inversion $^{-1}$ and multiplicative matrix semigroups $(M_n, ^T)$ over any field under the usual matrix transposition ^{T}.

With respect to the finite basis problem, involution semigroups have not been considered as much as semigroups, perhaps due to the supposition that a finite involution semigroup $(S, ^*)$ and its reduct S are simultaneously finitely based; but this has been refuted by recent examples [6, 10, 12]. An interesting example is the monoid A_0^1, obtained by adjoining a unit element to the semigroup

$$A_0 = \langle e, f \mid e^2 = e,\ f^2 = f,\ ef = 0\rangle$$

of order four, with involution * given by the *transposition* $e \leftrightarrow f$ on A_0^1, that is, * interchanges e and f but fixes every other element. It is long known that the monoid A_0^1 of order five is finitely based [4], but recently, Gao *et al.* [5] have shown that the involution monoid $(A_0^1, ^*)$ is nonfinitely based. This result is surprising given that every semigroup of order up to five is finitely based [9, 22]. As in the case for semigroups, it is natural to ask if there exists a nonfinitely based involution semigroup of order less than five [5, Question 1.3]. The objective of the present article is to provide an answer to this question for involution monoids.

Theorem 1.1. *Up to isomorphism, the involution monoid $(A_0^1, ^*)$ of order five is the unique smallest nonfinitely based algebra in the class of all involution monoids.*

Notation and background information are first given in Section 2. An outline of the proof of Theorem 1.1 is then given in Section 3, while the finer details of the proof are deferred to Sections 4–6.

Question 1.2. Is there a nonfinitely based involution semigroup of order five or less?

Since every involution semigroup of order up to three is finitely based [14], any example that positively answers Question 1.2 is of order four or five. If the answer to Question 1.2 is negative, then $(A_0^1, ^*)$ is also the unique smallest nonfinitely based involution semigroup. Refer to the monograph of Lee [15] for more information on the finite basis problem for involution semigroups.

2. Preliminaries

Most of the notation and background material of this article are given in this section. Refer to the monograph of Burris and Sankappanavar [3] for more information.

2.1. Words. Let $\mathscr{A}$ be a countably infinite alphabet that excludes the symbol 1, and let $\mathscr{A}^* = \{x^* \mid x \in \mathscr{A}\}$ be a disjoint copy of $\mathscr{A}$. Elements of $\mathscr{A} \cup \mathscr{A}^*$ are called *variables*. The *free involution semigroup* over $\mathscr{A}$ is the free semigroup $F_{\mathsf{inv}}(\mathscr{A}) = (\mathscr{A} \cup \mathscr{A}^*)^+$ with unary operation * given by $(x^*)^* = x$ for all $x \in \mathscr{A}$ and

$$(x_1 x_2 \cdots x_n)^* = x_n^* x_{n-1}^* \cdots x_1^*$$

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

for all $x_1, x_2, \ldots, x_n \in \mathcal{A} \cup \mathcal{A}^*$. The *free involution monoid* over $\mathcal{A}$ is $F^1_{\mathsf{inv}}(\mathcal{A}) = F_{\mathsf{inv}}(\mathcal{A}) \cup \{1\}$, where 1 is the empty word with $1^* = 1$. Elements of $F^1_{\mathsf{inv}}(\mathcal{A})$ are called *words* and elements of $\mathcal{A}^+ \cup \{1\}$ are called *plain words*. A word $\mathbf{u}$ is a *factor* of a word $\mathbf{v}$ if $\mathbf{puq} = \mathbf{v}$ for some $\mathbf{p}, \mathbf{q} \in F^1_{\mathsf{inv}}(\mathcal{A})$.

The *plain projection* of a word $\mathbf{u} \in F_{\mathsf{inv}}(\mathcal{A})$, denoted by $\overline{\mathbf{u}}$, is the plain word obtained from $\mathbf{u}$ by removing all occurrences of the symbol *. The *content* of a word $\mathbf{u}$, denoted by $\mathsf{con}(\mathbf{u})$, is the set of variables occurring in $\mathbf{u}$; the number of times that a variable x occurs in $\mathbf{u}$ is denoted by $\mathsf{occ}(x, \mathbf{u})$. A variable $x \in \mathcal{A} \cup \mathcal{A}^*$ is *simple* in $\mathbf{u}$ if $\mathsf{occ}(\overline{x}, \overline{\mathbf{u}}) = 1$; otherwise, it is *nonsimple*. A word $\mathbf{u}$ is *simple* if every variable in $\mathbf{u}$ is simple in $\mathbf{u}$. Let $\mathsf{sim}(\mathbf{u})$ denote the set of simple variables occurring in $\mathbf{u}$ and $\mathsf{non}(\mathbf{u})$ denote the set of nonsimple variables occurring in $\mathbf{u}$.

For any $\mathbf{u} \in F^1_{\mathsf{inv}}(\mathcal{A})$ and $x_1, x_2, \ldots, x_n \in \mathcal{A}$, let $\mathbf{u}[x_1, x_2, \ldots, x_n]$ denote the word obtained from $\mathbf{u}$ by retaining the variables $x_1, x_1^*, x_2, x_2^*, \ldots, x_n, x_n^*$. In particular, $\mathbf{u}[\mathsf{sim}(\mathbf{u})]$ is obtained from $\mathbf{u}$ by retaining its simple variables.

EXAMPLE 2.1. If $\mathbf{u} = x^*xy^*x^2yz^*yx^*$ with $x, y, z \in \mathcal{A}$, then

- $\overline{\mathbf{u}} = x^2yx^2yzyx$;
- $\mathsf{con}(\mathbf{u}) = \{x, x^*, y, y^*, z^*\}$;
- $\mathsf{occ}(x, \mathbf{u}) = 3$, $\mathsf{occ}(x^*, \mathbf{u}) = 2$, $\mathsf{occ}(y, \mathbf{u}) = 2$, $\mathsf{occ}(z^*, \mathbf{u}) = \mathsf{occ}(y^*, \mathbf{u}) = 1$;
- $\mathsf{occ}(x, \overline{\mathbf{u}}) = 5$, $\mathsf{occ}(y, \overline{\mathbf{u}}) = 3$, $\mathsf{occ}(z, \overline{\mathbf{u}}) = 1$;
- $\mathsf{sim}(\mathbf{u}) = \{z^*\}$, $\mathsf{non}(\mathbf{u}) = \{x, x^*, y, y^*\}$;
- $\mathbf{u}[x] = x^*x^3x^*$, $\mathbf{u}[x, y] = x^*xy^*x^2y^2x^*$, $\mathbf{u}[y, z] = y^*yz^*y$.

2.2. Identities. An *identity* is an expression $\mathbf{u} \approx \mathbf{v}$ formed by words $\mathbf{u}, \mathbf{v} \in F^1_{\mathsf{inv}}(\mathcal{A})$. An involution semigroup $(S, ^*)$ *satisfies* an identity $\mathbf{u} \approx \mathbf{v}$ if, for any substitution $\varphi : \mathcal{A} \to S$, the elements $\mathbf{u}\varphi$ and $\mathbf{v}\varphi$ of S coincide; in this case, $\mathbf{s} \approx \mathbf{t}$ is also said to be an *identity of* $(S, ^*)$.

An involution monoid that satisfies an identity $\mathbf{u} \approx \mathbf{v}$ also satisfies the identity $\mathbf{u}[x_1, x_2, \ldots, x_n] \approx \mathbf{v}[x_1, x_2, \ldots, x_n]$ for any $x_1, x_2, \ldots, x_n \in \mathcal{A}$, since assigning the unit element 1 to a variable x in an identity is effectively the same as removing all occurrences of x and x^*.

For any involution semigroup $(S, ^*)$, a set Σ of identities of $(S, ^*)$ is an *identity basis* for $(S, ^*)$ if every identity of $(S, ^*)$ can be deduced from Σ. An involution semigroup is *finitely based* if it has some finite identity basis; otherwise, it is *nonfinitely based*.

2.3. Periodic commutative involution semigroups. Perkins [21] proved that every commutative semigroup is finitely based. In this subsection, a similar result is established for involution semigroups.

PROPOSITION 2.2. *Every periodic commutative involution semigroup is finitely based.*

Recall that a semigroup S is *periodic* if it satisfies the identity $x^i \approx x^{i+j}$ for some $i, j \geq 1$. If $m \geq 1$ is the least such that S satisfies $x^m \approx x^{m+j}$ for some $j \geq 1$, and k is the least such that S satisfies $x^m \approx x^{m+k}$, then S is (m, k)-*periodic*. An involution semigroup $(S, ^*)$ is (m, k)-*periodic* if S is (m, k)-periodic. An identity $\mathbf{u} \approx \mathbf{v}$ of an (m, k)-periodic

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

involution semigroup $(S, {}^*)$ is *reduced* if the words $\mathbf{u}$ and $\mathbf{v}$ belong to the set

$$\{x_1^{e_1}x_2^{e_2}\cdots x_n^{e_n} \mid 0 \le e_1, e_2, \ldots, e_n < m+k\}$$

for some distinct variables $x_1, x_2, \ldots, x_n \in \mathscr{A} \cup \mathscr{A}^*$.

Let $\mathbf{u} \approx \mathbf{v}$ be a reduced identity of an (m,k)-periodic involution semigroup $(S, {}^*)$. For any integers p, q, s, t such that $0 \le p, q, s, t < m+k$, a nonempty set $U_{(p,q,s,t)} = \{x_1, x_2, \ldots, x_n\}$ of variables from $\mathsf{con}(\overline{\mathbf{uv}})$ is called the (p,q,s,t)-*block* of $\mathbf{u} \approx \mathbf{v}$ if $U_{(p,q,s,t)}$ is the maximal subset of $\mathsf{con}(\overline{\mathbf{uv}})$ such that for each variable x_i in $U_{(p,q,s,t)}$,

$$\mathsf{occ}(x_i, \mathbf{u}) = p, \quad \mathsf{occ}(x_i^*, \mathbf{u}) = q, \quad \mathsf{occ}(x_i, \mathbf{v}) = s \quad \text{and} \quad \mathsf{occ}(x_i^*, \mathbf{v}) = t.$$

Note that $(p,q,s,t) \neq (0,0,0,0)$ because $U_{(p,q,s,t)} \neq \emptyset$. The *length* of $U_{(p,q,s,t)}$ is denoted by $|U_{(p,q,s,t)}|$. For instance, if $\mathbf{u} \approx \mathbf{v}$ is an identity with reduced words

$$\mathbf{u} = x_1^2x_2^3x_3^3x_4^3x_5^2x_1^*(x_2^*)^2(x_3^*)^2(x_4^*)^2x_5^*x_6^* \quad \text{and} \quad \mathbf{v} = x_1^6x_5^6(x_2^*)^2(x_3^*)^2(x_4^*)^2,$$

then $\{x_1, x_5\}$, $\{x_2, x_3, x_4\}$ and $\{x_6\}$ are the $(2,1,6,0)$-block, the $(3,2,0,2)$-block and the $(0,1,0,0)$-block of $\mathbf{u} \approx \mathbf{v}$, respectively.

For any reduced identity $\mathbf{u} \approx \mathbf{v}$, since each component of a quadruple (p,q,s,t) is from $\{0, 1, 2, \ldots, m+k-1\}$ and $(p,q,s,t) \neq (0,0,0,0)$, the number of possible quadruples is $r = (m+k)^4 - 1$. Encode these quadruples so that we can refer to the i-block U_i of $\mathbf{u} \approx \mathbf{v}$, where $1 \le i \le r$, instead of the (p,q,s,t)-block $U_{(p,q,s,t)}$ of $\mathbf{u} \approx \mathbf{v}$. An r-dimensional vector $\overrightarrow{\mathbf{l}} \in \mathbb{N}^r$ (where $\mathbb{N} = \{0, 1, 2, \ldots\}$) is called the *length vector* of blocks for a reduced identity $\mathbf{u} \approx \mathbf{v}$ if the ith component

$$\overrightarrow{\mathbf{l}}(i) = \begin{cases} |U_i| & \text{if the } i\text{-block of } \mathbf{u} \approx \mathbf{v} \text{ exists,} \\ 0 & \text{otherwise.} \end{cases}$$

It is routine to check that every reduced identity $\mathbf{u} \approx \mathbf{v}$ of an (m,k)-periodic commutative involution semigroup can be uniquely determined by some r-dimensional vector. Let $\overrightarrow{\mathbf{l}_1}$ and $\overrightarrow{\mathbf{l}_2}$ be r-dimensional length vectors corresponding to reduced identities $\mathbf{u}_1 \approx \mathbf{v}_1$ and $\mathbf{u}_2 \approx \mathbf{v}_2$, respectively. Define a partial order $\preceq$ on $\mathbb{N}^r$ such that $\overrightarrow{\mathbf{l}_1} \preceq \overrightarrow{\mathbf{l}_2}$ if $\overrightarrow{\mathbf{l}_1}(i) \le \overrightarrow{\mathbf{l}_2}(i)$ for all $i \in \{1, 2, \ldots, r\}$. Similar to the argument given by Perkins [21, Section 4] for the case of semigroups, we can deduce a 'long' identity from a 'short' identity using some appropriate substitution, that is, if $\overrightarrow{\mathbf{l}_1} \preceq \overrightarrow{\mathbf{l}_2}$, then $\mathbf{u}_1 \approx \mathbf{v}_1$ implies $\mathbf{u}_2 \approx \mathbf{v}_2$.

PROOF OF PROPOSITION 2.2. Let $(S, {}^*)$ be any periodic commutative involution semigroup, say $(S, {}^*)$ is (m,k)-periodic. Suppose that $(S, {}^*)$ is nonfinitely based. Then there exists an infinite set

$$\Sigma = \{\mathbf{u}_1 \approx \mathbf{v}_1,\ \mathbf{u}_2 \approx \mathbf{v}_2,\ \mathbf{u}_3 \approx \mathbf{v}_3, \ldots\}$$

of identities of $(S, {}^*)$ such that for each $i \ge 1$, the first i identities

$$\Sigma_i = \{\mathbf{u}_1 \approx \mathbf{v}_1,\ \mathbf{u}_2 \approx \mathbf{v}_2,\ \ldots,\ \mathbf{u}_i \approx \mathbf{v}_i\}$$

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

do not imply the $(i+1)$st identity $\mathbf{u}_{i+1} \approx \mathbf{v}_{i+1}$. Since $(S, {}^*)$ is commutative, each identity $\mathbf{u}_i \approx \mathbf{v}_i \in \Sigma$ can be converted into reduced form and is thus uniquely associated with some length vector $\overrightarrow{\mathbf{l}_i} \in \mathbb{N}^r$. Since $\Sigma_i \nvdash \mathbf{u}_{i+1} \approx \mathbf{v}_{i+1}$ for each $i \geq 1$, the length vectors $\overrightarrow{\mathbf{l}_1}, \overrightarrow{\mathbf{l}_2}, \overrightarrow{\mathbf{l}_3}, \ldots$ corresponding to $\mathbf{u}_1 \approx \mathbf{v}_1, \mathbf{u}_2 \approx \mathbf{v}_2, \mathbf{u}_3 \approx \mathbf{v}_3, \ldots$ are distinct.

The set $\{\overrightarrow{\mathbf{l}_1}, \overrightarrow{\mathbf{l}_2}, \overrightarrow{\mathbf{l}_3}, \ldots\}$ of infinitely many pairwise distinct r-dimensional vectors must contain two vectors $\overrightarrow{\mathbf{l}_k}$ and $\overrightarrow{\mathbf{l}_\ell}$ with $k < \ell$ such that $\overrightarrow{\mathbf{l}_k} \preceq \overrightarrow{\mathbf{l}_\ell}$. Indeed, this can be proved by induction on the dimension r. If $r = 1$, the conclusion holds obviously. Suppose that infinitely many pairwise distinct $(r-1)$-dimensional vectors contain two $\preceq$-related vectors. Then we can show that it also holds for the r-dimensional vectors. Let $L_1 = \{\overrightarrow{\mathbf{l}_1}, \overrightarrow{\mathbf{l}_2}, \overrightarrow{\mathbf{l}_3}, \ldots\}$. We can find an infinite set $L_2 = \{\overrightarrow{\mathbf{l}_{p_1}}, \overrightarrow{\mathbf{l}_{p_2}}, \overrightarrow{\mathbf{l}_{p_3}}, \ldots\} \subseteq L_1$ for some $p_1 < p_2 < p_3 < \cdots$ such that for at least one component, say the ith component with $1 \leq i \leq r$, one has $\overrightarrow{\mathbf{l}_{p_1}}(i) \leq \overrightarrow{\mathbf{l}_{p_2}}(i) \leq \overrightarrow{\mathbf{l}_{p_3}}(i) \leq \cdots$. By the inductive hypothesis, we can find two distinct vectors $\overrightarrow{\mathbf{l}_k}, \overrightarrow{\mathbf{l}_\ell} \in L_2$ such that $\overrightarrow{\mathbf{l}_k} \preceq \overrightarrow{\mathbf{l}_\ell}$. Therefore, there exists some j such that $\Sigma_j \vdash \mathbf{u}_{j+1} \approx \mathbf{v}_{j+1}$, which is a contradiction. □

3. Proof of Theorem 1.1

Since a finite involution monoid is finitely based if it is either commutative (Proposition 2.2) or of order at most three [14], it suffices to consider those that are noncommutative and of order four or five. It is routine to check with a computer that every involution monoid of order four is commutative and so is finitely based, and there are only six noncommutative involution monoids of order five:

M_1	1	2	3	4	5
1	1	1	1	1	1
2	1	1	1	1	2
3	1	1	1	1	3
4	1	1	2	1	4
5	1	2	3	4	5
x^*	1	2	4	3	5

M_2	1	2	3	4	5
1	1	1	1	1	1
2	1	1	1	1	2
3	1	1	1	3	3
4	1	2	1	4	4
5	1	2	3	4	5
x^*	1	3	2	4	5

M_3	1	2	3	4	5
1	1	1	1	1	1
2	1	1	1	1	2
3	1	1	2	1	3
4	1	1	2	2	4
5	1	2	3	4	5
x^*	1	2	4	3	5

M_4	1	2	3	4	5
1	1	1	1	1	1
2	1	1	1	2	2
3	1	2	3	1	3
4	1	1	1	4	4
5	1	2	3	4	5
x^*	1	2	4	3	5

M_5	1	2	3	4	5
1	1	1	1	1	1
2	1	1	1	2	2
3	1	2	3	2	3
4	1	1	1	4	4
5	1	2	3	4	5
x^*	1	2	4	3	5

M_6	1	2	3	4	5
1	1	1	1	4	4
2	1	2	3	4	5
3	3	3	3	5	5
4	1	4	1	4	4
5	3	5	3	5	5
x^*	1	2	4	3	5

The involution monoid $(M_1, {}^*)$, which appears in [13] as $\langle \mathsf{Rq}\{xx^*\}, {}^* \rangle$, is finitely based; in particular, its identities are axiomatised by (1.1) and

$$x^3 \approx x^2, \quad xyx \approx x^2y, \quad xyx \approx yx^2, \quad xyx^* \approx xx^*y, \quad xyx^* \approx yxx^*, \quad (x^*)^2 \approx x^2.$$

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

The involution monoids $(M_2, {}^*)$, $(M_3, {}^*)$ and $(M_4, {}^*)$ are shown to be finitely based in Sections 4, 5 and 6, respectively.

The involution monoid $(M_5, {}^*)$ is isomorphic to $(A_0^1, {}^*)$ and so is nonfinitely based [5], while it follows from Adair [1] that the identities of $(M_6, {}^*)$ are axiomatised by (1.1) and

$$x^2 \approx x, \quad xx^*x \approx x, \quad xx^*yxy \approx xy, \quad xyxy^*y \approx xy.$$

4. The involution monoid $(M_2, {}^*)$

If $x, x^* \in \mathsf{con}(\mathbf{u})$ for some $x \in \mathcal{A}$, then $\{x, x^*\}$ is a *mixed pair* of $\mathbf{u}$. A word is *mixed* if it has some mixed pair. A word without mixed pairs is *bipartite*.

LEMMA 4.1 (Lee [11, Lemma 9]). *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be any bipartite words such that* $\mathsf{con}(\mathbf{u}) = \mathsf{con}(\mathbf{v})$. *Then an involution semigroup satisfies* $\mathbf{u} \approx \mathbf{v}$ *if and only if it satisfies* $\overline{\mathbf{u}} \approx \overline{\mathbf{v}}$.

LEMMA 4.2. *Let* $(M, {}^*)$ *be any involution monoid that satisfies the identities*

$$x^*yx \approx xyx, \quad xyx^* \approx xyx, \quad x^*yx^* \approx xyx. \tag{4.1}$$

Suppose that M *is finitely based. Then* $(M, {}^*)$ *is also finitely based.*

PROOF. There exists some set Σ of identities of $(M, {}^*)$ such that $\{(1.1), (4.1)\} \cup \Sigma$ is an identity basis for $(M, {}^*)$. In view of the identities (4.1), the identities in Σ can be assumed to be formed by words whose nonsimple variables are all plain; note that these words are bipartite. If $\mathsf{con}(\mathbf{u}) \neq \mathsf{con}(\mathbf{v})$ for some $\mathbf{u} \approx \mathbf{v} \in \Sigma$, then $(M, {}^*)$ satisfies either $x^a \approx x^*$ or $x^b \approx 1$ for some $a, b \geq 1$. Note that $x \approx (x^*)^a \overset{(4.1)}{\approx} x^a \approx x^*$ if $a \geq 2$ and $x \approx x^{b+1} \overset{(4.1)}{\approx} x^b x^* \approx x^*$. Hence, $(M, {}^*)$ satisfies the identity $x \approx x^*$, and so $(M, {}^*)$ is finitely based since M is finitely based. If $\mathsf{con}(\mathbf{u}) = \mathsf{con}(\mathbf{v})$ for all $\mathbf{u} \approx \mathbf{v} \in \Sigma$, then by Lemma 4.1, the identities in Σ can be chosen to be plain. In other words, Σ is a set of identities of the monoid M. By assumption, there exists a finite identity basis Σ_0 for M, so that Σ_0 implies Σ. Therefore, $\{(1.1), (4.1)\} \cup \Sigma_0$ implies $\{(1.1), (4.1)\} \cup \Sigma$ and so is a finite identity basis for $(M, {}^*)$. □

COROLLARY 4.3. *The involution monoid* $(M_2, {}^*)$ *is finitely based.*

PROOF. It is routine to check that the involution monoid $(M_2, {}^*)$ satisfies the identities (4.1). Since M_2 is finitely based [4], the result holds by Lemma 4.2. □

5. The involution monoid $(M_3, {}^*)$

PROPOSITION 5.1. *The identities (1.1) and*

$$x^4 \approx x^3, \quad xyx \approx x^2y, \quad xyx \approx yx^2,$$
$$xyx^* \approx xx^*y, \quad xyx^* \approx yxx^*, \quad (x^*)^3 \approx x^3, \quad (x^*)^2 \approx x^2$$

constitute an identity basis for $(M_3, {}^*)$.

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

For any variables $x_1, x_2, \dots, x_n \in \mathcal{A}$ in strict alphabetical order, define

- $x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n}$, where $e_1, e_2, \dots, e_n \in \{2, 3\}$, to be a *plain restricted word*;
- $\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_n$, where $\mathbf{x}_i \in \{x_i x_i^*, x_i^* x_i\}$, to be a *mixed restricted word*.

It is easy to show that the identities in Proposition 5.1 can be used to convert every word in $F_{\mathsf{inv}}^1(\mathcal{A})$ into some word of the form $\mathbf{pms}$, where $\mathbf{p} \in \mathcal{A}^+ \cup \{1\}$ is a plain restricted word, $\mathbf{m} \in F_{\mathsf{inv}}^1(\mathcal{A})$ is a mixed restricted word, and $\mathbf{s} \in F_{\mathsf{inv}}^1(\mathcal{A})$ is a simple word such that the sets $\mathsf{con}(\mathbf{p})$, $\mathsf{con}(\overline{\mathbf{m}})$ and $\mathsf{con}(\overline{\mathbf{s}})$ are pairwise disjoint; in this section, such a word $\mathbf{pms}$ is said to be in *canonical form*.

It is routine to check that $(M_3, {}^*)$ satisfies the identities in Proposition 5.1. Let $\mathbf{u}_1 \approx \mathbf{u}_2$ be any identity of $(M_3, {}^*)$, where $\mathbf{u}_i = \mathbf{p}_i \mathbf{m}_i \mathbf{s}_i$ is in canonical form for each $i \in \{1, 2\}$. In the remainder of this section, it is shown that $\mathbf{u}_1 = \mathbf{u}_2$. This completes the proof of Proposition 5.1.

LEMMA 5.2. $\mathbf{p}_1 = \mathbf{p}_2$.

PROOF. Suppose that $\mathbf{p}_1 \neq \mathbf{p}_2$. Then there are two cases.

Case 1: $\mathsf{con}(\mathbf{p}_1) = \mathsf{con}(\mathbf{p}_2)$. Then $\mathsf{occ}(x, \mathbf{p}_1) \neq \mathsf{occ}(x, \mathbf{p}_2)$ for some $x \in \mathcal{A}$, so that $\{\mathbf{p}_1[x], \mathbf{p}_2[x]\} = \{x^2, x^3\}$ by the definition of plain restricted words. It follows that $\mathbf{u}_1[x] \approx \mathbf{u}_2[x]$ is the identity $x^3 \approx x^2$; but this identity is not satisfied by $(M_3, {}^*)$, giving a contradiction.

Case 2: $\mathsf{con}(\mathbf{p}_1) \neq \mathsf{con}(\mathbf{p}_2)$. Generality is not lost by assuming the existence of some $x \in \mathsf{con}(\mathbf{p}_1) \setminus \mathsf{con}(\mathbf{p}_2)$. If $x \in \mathsf{con}(\mathbf{m}_2)$, then $\mathbf{u}_1[x] \approx \mathbf{u}_2[x]$ is either $x^2 \approx xx^*$, $x^2 \approx x^*x$, $x^3 \approx xx^*$ or $x^3 \approx x^*x$. If $x \in \mathsf{con}(\mathbf{s}_2)$, then $\mathbf{u}_1[x] \approx \mathbf{u}_2[x]$ is either $x^2 \approx x$, $x^2 \approx x^*$, $x^3 \approx x$ or $x^3 \approx x^*$. If $x \notin \mathsf{con}(\mathbf{m}_2\mathbf{s}_2)$, then $\mathbf{u}_1[x] \approx \mathbf{u}_2[x]$ is either $x^2 \approx 1$ or $x^3 \approx 1$. However, these ten identities are not satisfied by $(M_3, {}^*)$, giving a contradiction. □

LEMMA 5.3. $\mathbf{m}_1 = \mathbf{m}_2$.

PROOF. Suppose that $\mathbf{m}_1 \neq \mathbf{m}_2$. Then there are two cases.

Case 1: $\mathsf{con}(\mathbf{m}_1) = \mathsf{con}(\mathbf{m}_2) = \{x_1, x_1^*, x_2, x_2^*, \dots, x_n, x_n^*\}$ for some variables $x_1, x_2, \dots, x_n \in \mathcal{A}$ in strict alphabetical order. Then by the definition of mixed restricted words, $\mathbf{m}_1 = \mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_n$ and $\mathbf{m}_2 = \mathbf{x}_1' \mathbf{x}_2' \cdots \mathbf{x}_n'$, where $\mathbf{x}_i, \mathbf{x}_i' \in \{x_i x_i^*, x_i^* x_i\}$ for all i. The assumption $\mathbf{m}_1 \neq \mathbf{m}_2$ implies that $\mathbf{x}_j \neq \mathbf{x}_j'$ for some j. Therefore, $\mathbf{u}_1[x_j] \approx \mathbf{u}_2[x_j]$ is the identity $x_j x_j^* \approx x_j^* x_j$; but this identity is not satisfied by $(M_3, {}^*)$, giving a contradiction.

Case 2: $\mathsf{con}(\mathbf{m}_1) \neq \mathsf{con}(\mathbf{m}_2)$. Generality is not lost by assuming the existence of some $x \in \mathcal{A}$ such that $x, x^* \in \mathsf{con}(\mathbf{m}_1) \setminus \mathsf{con}(\mathbf{m}_2)$. If $x \in \mathsf{con}(\mathbf{p}_2)$, then $x \in \mathsf{con}(\mathbf{p}_1)$ by Lemma 5.2, whence $\mathsf{con}(\mathbf{p}_1)$ and $\mathsf{con}(\overline{\mathbf{m}_1})$ are not disjoint, contradicting the choice of $\mathbf{p}_1$ and $\mathbf{m}_1$. Hence, $x \notin \mathsf{con}(\mathbf{p}_2)$. Clearly, $x^* \notin \mathsf{con}(\mathbf{p}_2)$ because the word $\mathbf{p}_2$ is plain. Therefore, the remaining possibilities are $x \in \mathsf{con}(\overline{\mathbf{s}_2})$ and $x \notin \mathsf{con}(\overline{\mathbf{s}_2})$. If $x \in \mathsf{con}(\overline{\mathbf{s}_2})$, then $\mathbf{u}_1[x] \approx \mathbf{u}_2[x]$ is either $xx^* \approx x$, $xx^* \approx x^*$, $x^*x \approx x$ or $x^*x \approx x^*$. If $x \notin \mathsf{con}(\overline{\mathbf{s}_2})$, then $\mathbf{u}_1[x] \approx \mathbf{u}_2[x]$ is either $xx^* \approx 1$ or $x^*x \approx 1$. However, these six identities are not satisfied by $(M_3, {}^*)$, giving a contradiction. □

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

LEMMA 5.4. $\mathbf{s}_1 = \mathbf{s}_2$.

PROOF. Recall that $(M_3, {}^*)$ satisfies $\mathbf{p}_1\mathbf{m}_1\mathbf{s}_1 \approx \mathbf{p}_2\mathbf{m}_2\mathbf{s}_2$, where for each $i \in \{1, 2\}$, the sets $\mathsf{con}(\mathbf{p}_i)$, $\mathsf{con}(\overline{\mathbf{m}}_i)$ and $\mathsf{con}(\overline{\mathbf{s}}_i)$ are pairwise disjoint. Since $\mathbf{p}_1 = \mathbf{p}_2$ and $\mathbf{m}_1 = \mathbf{m}_2$ by Lemmas 5.2 and 5.3, it follows that $(M_3, {}^*)$ satisfies $\mathbf{s}_1 \approx \mathbf{s}_2$. It is then easy to show that $\mathbf{s}_1 = \mathbf{s}_2$. □

6. A finite basis for $(M_4, {}^*)$

PROPOSITION 6.1. *The identities (1.1) and*

$$xyxzx \approx xyzx, \tag{6.1a}$$

$$x^2y^2 \approx y^2x^2, \tag{6.1b}$$

$$xx^*y \approx yxx^*, \tag{6.1c}$$

$$x^*yxzx \approx x^*yxz, \quad xzxyx^* \approx zxyx^* \tag{6.1d}$$

$$xyx^*zx \approx yzxx^*, \tag{6.1e}$$

$$xx^* \approx x^*x, \tag{6.1f}$$

$$x^{\circledast_1}hxyky^{\circledast_2} \approx x^{\circledast_1}hyxky^{\circledast_2}, \tag{6.1g}$$

$$x^{\circledast_1}hy^{\circledast_2}kxy \approx x^{\circledast_1}hy^{\circledast_2}kyx, \tag{6.1h}$$

$$xyhx^{\circledast_1}ky^{\circledast_2} \approx yxhx^{\circledast_1}ky^{\circledast_2}, \tag{6.1i}$$

where $\circledast_1, \circledast_2 \in \{1, *\}$, *constitute an identity basis for* $(M_4, {}^*)$.

Some basic results are given in Section 6.1. A canonical form for words forming identities of $(M_4, {}^*)$ is given in Section 6.2. Results established in these two subsections are then used to prove Proposition 6.1 in Section 6.3.

REMARK 6.2. The identities (6.1a)–(6.1d) actually imply the latter identities (6.1e)–(6.1i) and so constitute an identity basis for $(M_4, {}^*)$. However, as we will see shortly, the identities (6.1e)–(6.1i) are crucial to the proof of Proposition 6.1.

6.1. Basic results.

REMARK 6.3. It is routine to check that the involution monoid $(M_4, {}^*)$ satisfies the identities (6.1) but not any of the identities

$$\begin{aligned} &xyx \approx x^2y, \quad xyx^* \approx xx^*y, \quad xyx^* \approx x^*xy, \quad xyx^* \approx x^*yx, \\ &xyx \approx yx^2, \quad xyx^* \approx yxx^*, \quad xyx^* \approx yx^*x, \quad x^2y \approx yx^2. \end{aligned} \tag{6.2}$$

A word $\mathbf{u} \in F^1_{\mathsf{inv}}(\mathscr{A})$ is 2-*limited* if for any $x \in \mathscr{A}$, the total number of times x and x^* occur in $\mathbf{u}$ is at most two, that is, $\mathsf{occ}(x, \mathbf{u}) + \mathsf{occ}(x^*, \mathbf{u}) \leq 2$. An identity is 2-*limited* if it is formed by a pair of 2-limited words.

LEMMA 6.4. *Given any word* $\mathbf{u} \in F^1_{\mathsf{inv}}(\mathcal{A})$, *there exists some* 2*-limited word* $\mathbf{u}'$ *such that* $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{u}')$, $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{u}')$ *and (6.1)* $\vdash \mathbf{u} \approx \mathbf{u}'$.

PROOF. It is easy to see that the identities {(6.1a), (6.1d), (6.1e)} can be used to convert any word $\mathbf{u}$ into some 2-limited word $\mathbf{u}'$ satisfying $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{u}')$ and $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{u}')$. □

Define a relation $\sim$ on $F^1_{\mathsf{inv}}(\mathcal{A})$ by $\mathbf{u} \sim \mathbf{v}$ if $\mathbf{u}$ and $\mathbf{v}$ are the same word up to arrangement of their variables. Equivalently, $\mathbf{u} \sim \mathbf{v}$ if and only if $xy \approx yx \vdash \mathbf{u} \approx \mathbf{v}$.

LEMMA 6.5. *Let* $\mathbf{u} \approx \mathbf{v}$ *be any* 2*-limited identity of* $(M_4, {}^*)$. *Then*

(i) $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v})$ *and* $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v})$;
(ii) $\mathbf{u} \sim \mathbf{v}$;
(iii) $\mathbf{u}[\mathsf{sim}(\mathbf{u})] = \mathbf{v}[\mathsf{sim}(\mathbf{v})]$.

PROOF. (i) First, suppose that $x \in \mathsf{con}(\mathbf{u}) \setminus \mathsf{con}(\mathbf{v})$. Generality is not lost by assuming that $x \in \mathcal{A}$. Let $\varphi : \mathcal{A} \to M_4$ denote the substitution that maps x to 3 and every other variable to 5. Then

$$\mathbf{u}\varphi = \begin{cases} 1 & \text{if } x^* \in \mathsf{con}(\mathbf{u}), \\ 3 & \text{otherwise;} \end{cases} \qquad \mathbf{v}\varphi = \begin{cases} 4 & \text{if } x^* \in \mathsf{con}(\mathbf{v}), \\ 5 & \text{otherwise.} \end{cases}$$

Therefore, the contradiction $\mathbf{u}\varphi \neq \mathbf{v}\varphi$ is obtained. Hence, the variable x does not exist, so that $\mathsf{con}(\mathbf{u}) = \mathsf{con}(\mathbf{v})$.

Now suppose that $x \in \mathsf{sim}(\mathbf{u}) \setminus \mathsf{sim}(\mathbf{v})$. Since $x \in \mathsf{con}(\mathbf{u}) = \mathsf{con}(\mathbf{v})$, we have $x \in \mathsf{non}(\mathbf{v})$. Let $\psi : \mathcal{A} \to M_4$ denote the substitution that maps x to 2 and every other variable to 5. Then $\mathbf{u}\psi = 2$ and $\mathbf{v}\psi = 1$, resulting in the contradiction $\mathbf{u}\psi \neq \mathbf{v}\psi$. Therefore, the variable x does not exist, so that $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v})$; this, together with $\mathsf{con}(\mathbf{u}) = \mathsf{con}(\mathbf{v})$, implies that $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v})$.

(ii) This is an easy consequence of part (i) because $\mathbf{u}$ and $\mathbf{v}$ are 2-limited words.

(iii) Suppose that $\mathbf{u}[\mathsf{sim}(\mathbf{u})] \neq \mathbf{v}[\mathsf{sim}(\mathbf{v})]$. Then there exist $x, y \in \mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v})$ such that x precedes y in $\mathbf{u}$ but y precedes x in $\mathbf{v}$. Hence, $\mathbf{u}[x, y] \approx \mathbf{v}[x, y]$ is the identity $xy \approx yx$, which implies that $(M_4, {}^*)$ is commutative, a contradiction. □

LEMMA 6.6. *Let* $\mathbf{u} \approx \mathbf{v}$ *be any* 2*-limited identity of* $(M_4, {}^*)$.

(i) *Suppose that either* $\mathsf{sim}(\mathbf{u}) = \emptyset$ *or* $\mathsf{sim}(\mathbf{v}) = \emptyset$. *Then (6.1)* $\vdash \mathbf{u} \approx \mathbf{v}$.
(ii) *Suppose that either* $\mathsf{non}(\mathbf{u}) = \emptyset$ *or* $\mathsf{non}(\mathbf{v}) = \emptyset$. *Then* $\mathbf{u} = \mathbf{v}$.

PROOF. (i) By Lemma 6.5(i), we have $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v}) = \emptyset$ and $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v})$, so that both $\mathbf{u}$ and $\mathbf{v}$ consist entirely of nonsimple variables. Since $\mathbf{u} \sim \mathbf{v}$ by Lemma 6.5(ii), the identities (6.1g)–(6.1i) can be used to convert $\mathbf{u}$ into $\mathbf{v}$.

(ii) By Lemma 6.5(i), we have $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v})$ and $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v}) = \emptyset$, so that both $\mathbf{u}$ and $\mathbf{v}$ are simple words. Therefore, $\mathbf{u} = \mathbf{u}[\mathsf{sim}(\mathbf{u})] = \mathbf{v}[\mathsf{sim}(\mathbf{v})] = \mathbf{v}$ by Lemma 6.5(iii). □

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

6.2. Canonical form. Any alphabetical order $\prec$ on $\mathcal{A}$ can be extended to a total order $\prec$ on $\mathcal{A} \cup \mathcal{A}^*$ in the following manner: $x \prec x^*$ for all $x \in \mathcal{A}$ and for all $x, y \in \mathcal{A} \cup \mathcal{A}^*$, define $x \prec y$ if $\overline{x} \prec \overline{y}$. An *ordered word* is a word of the form

$$x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n},$$

where $x_1, x_2, \ldots, x_n \in \mathcal{A} \cup \mathcal{A}^*$ with $x_1 \prec x_2 \prec \cdots \prec x_n$ and $e_1, e_2, \ldots, e_n \geq 1$.

In this section, a 2-limited word $\mathbf{u}$ with $\mathsf{sim}(\mathbf{u}) \neq \emptyset$ and $\mathsf{non}(\mathbf{u}) \neq \emptyset$ is said to be in *canonical form* if

$$\mathbf{u} = \mathbf{u}_0 \prod_{i=1}^{m} (\mathbf{s}_i \mathbf{u}_i) \tag{6.3}$$

for some $m \geq 1$, where

(CF1) $\mathbf{u}_0, \mathbf{s}_1, \mathbf{u}_m \in F^1_{\mathsf{inv}}(\mathcal{A})$ and $\mathbf{s}_2, \mathbf{s}_3, \ldots, \mathbf{s}_m, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{m-1} \in F_{\mathsf{inv}}(\mathcal{A})$;
(CF2) $\mathbf{u}[\mathsf{sim}(\mathbf{u})] = \mathbf{s}_1 \mathbf{s}_2 \cdots \mathbf{s}_m$;
(CF3) $\mathbf{u}_0 = x_1 x_1^* x_2 x_2^* \cdots x_r x_r^*$ for some $x_1, x_2, \ldots, x_r \in \mathcal{A}$ with $x_1 \prec x_2 \prec \cdots \prec x_r$ and $r \geq 0$;
(CF4) $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{m-1} \in \mathsf{non}(\mathbf{u})^+$ and $\mathbf{u}_m \in \mathsf{non}(\mathbf{u})^+ \cup \{1\}$ are bipartite ordered words.

LEMMA 6.7. *Let $\mathbf{u}$ be any 2-limited word such that $\mathsf{sim}(\mathbf{u}) \neq \emptyset$ and $\mathsf{non}(\mathbf{u}) \neq \emptyset$. Then the identities (6.1) can be used to convert $\mathbf{u}$ into a word in canonical form.*

PROOF. Write $\mathbf{u} = \prod_{i=1}^{m} (\mathbf{s}_i \mathbf{u}_i)$, where $\mathbf{s}_1 \in F^1_{\mathsf{inv}}(\mathcal{A})$ and $\mathbf{s}_2, \mathbf{s}_3, \ldots, \mathbf{s}_m \in F_{\mathsf{inv}}(\mathcal{A})$ are maximal factors of $\mathbf{u}$ formed by simple variables, and $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{m-1} \in F_{\mathsf{inv}}(\mathcal{A})$ and $\mathbf{u}_m \in F^1_{\mathsf{inv}}(\mathcal{A})$ are maximal factors of $\mathbf{u}$ formed by nonsimple variables.

Suppose that some $\mathbf{u}_i$ contains a mixed pair $\{x, x^*\}$. Then apply the identities (6.1f)–(6.1i) to group x and x^* together as some factor xx^* of $\mathbf{u}_i$, and apply the identity (6.1c) to move xx^* to the left of $\mathbf{s}_1$.

The procedure in the previous paragraph can be repeated on every mixed pair of every $\mathbf{u}_i$, so that every $\mathbf{u}_i$ no longer has a mixed pair and so is bipartite. The factors of the form xx^* that are collected on the left of $\mathbf{s}_1$ can be rearranged by the identity (6.1c) to form the prefix $\mathbf{u}_0$ satisfying (CF3). Note that since $\mathbf{u}$ is 2-limited, the prefix $\mathbf{u}_0$ does not share any variable with the rest of the word.

Therefore, $\mathbf{u} = \mathbf{u}_0 \prod_{i=1}^{m} (\mathbf{s}_i \mathbf{u}'_i)$, where each $\mathbf{u}'_i$ is a bipartite word obtained from $\mathbf{u}_i$ by removing all its mixed pairs. If $\mathbf{u}'_i$ is empty for some $i < m$, then $\mathbf{s}_i$ and $\mathbf{s}_{i+1}$ can be combined into a single maximal factor of $\mathbf{u}$ formed by simple variables:

$$\mathbf{u} = \mathbf{u}_0 \cdots \mathbf{s}_i \mathbf{u}'_i \cdot \mathbf{s}_{i+1} \mathbf{u}'_{i+1} \cdots = \mathbf{u}_0 \cdots (\mathbf{s}_i \cdot \mathbf{s}_{i+1}) \mathbf{u}'_{i+1} \cdots .$$

The resulting word is of the form (6.3) satisfying (CF1). Now apply the identities (6.1g)–(6.1i) to rearrange each $\mathbf{u}_i$ $(1 \leq i \leq m)$ into an ordered word, so that (CF4) is satisfied. It is clear that (CF2) is also satisfied since no simple variable has been introduced or removed, and the order of appearance of the simple variables has not been changed. □

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

6.3. Proof of Proposition 6.1. It suffices to show that any identity $\mathbf{u} \approx \mathbf{v}$ of $(M_4, {}^*)$ is deducible from the identities (6.1). By Lemmas 6.4 and 6.5, we may further assume that

(a) $\mathbf{u}$ and $\mathbf{v}$ are 2-limited;
(b) $\mathbf{u} \sim \mathbf{v}$;
(c) $\mathbf{u}[\mathsf{sim}(\mathbf{u})] = \mathbf{v}[\mathsf{sim}(\mathbf{v})]$.

If either $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{u}) = \emptyset$ or $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v}) = \emptyset$, then (6.1) $\vdash \mathbf{u} \approx \mathbf{v}$ by Lemma 6.6. Therefore, it remains to address the case when $\mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v}) \neq \emptyset$ and $\mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v}) \neq \emptyset$, whence by Lemma 6.7, the words $\mathbf{u}$ and $\mathbf{v}$ can be assumed to be in canonical form, say

$$\mathbf{u} = \mathbf{u}_0 \prod_{i=1}^{m} (\mathbf{s}_i \mathbf{u}_i) \quad \text{and} \quad \mathbf{v} = \mathbf{v}_0 \prod_{i=1}^{n} (\mathbf{t}_i \mathbf{v}_i).$$

It follows from (a) and (CF3) that

(d) $\mathsf{con}(\mathbf{u}_0) \cap \mathsf{con}(\mathbf{s}_i \mathbf{u}_i) = \mathsf{con}(\mathbf{v}_0) \cap \mathsf{con}(\mathbf{t}_i \mathbf{v}_i) = \emptyset$ for all $i \geq 1$.

The results in the remainder of this subsection verify that $\mathbf{u} = \mathbf{v}$. The proof of Proposition 6.1 is therefore complete.

LEMMA 6.8. *$m = n$ and $\mathbf{s}_i = \mathbf{t}_i$ for all i.*

PROOF. Suppose that $y_1, y_2 \in \mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v})$ are such that $y_1 y_2$ is a factor of $\mathbf{u}$ but not of $\mathbf{v}$. Then $y_1 y_2$ is a factor of some $\mathbf{s}_j$ but is not a factor of any $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n$. However, since $\mathbf{s}_1 \mathbf{s}_2 \cdots \mathbf{s}_m = \mathbf{t}_1 \mathbf{t}_2 \cdots \mathbf{t}_n$ by (c), the word $y_1 y_2$ is a factor of $\mathbf{t}_1 \mathbf{t}_2 \cdots \mathbf{t}_n$. It follows that for some j, the last variable of $\mathbf{t}_j$ is y_1 and the first variable of $\mathbf{t}_{j+1}$ is y_2; in other words, $y_1 \mathbf{v}_j y_2$ is a factor of $\mathbf{v}$. By (CF4), the factor $\mathbf{v}_j$ contains some nonsimple variable of $\mathbf{v}$, say $x^{\circledast}$ with $\circledast \in \{1, *\}$. Then by (a) and (CF4),

$$\mathbf{u}[x, y_1, y_2] \in \{x^{\circledast_1} x^{\circledast_2} y_1 y_2,\ x^{\circledast_1} y_1 y_2 x^{\circledast_2},\ y_1 y_2 x^{\circledast_1} x^{\circledast_2}\} \quad \text{and}$$
$$\mathbf{v}[x, y_1, y_2] \in \{x^{\circledast_3} y_1 x^{\circledast_4} y_2,\ y_1 (x^{\circledast_3})^2 y_2,\ y_1 x^{\circledast_3} y_2 x^{\circledast_4}\}$$

for some $\circledast_1, \circledast_2, \circledast_3, \circledast_4 \in \{1, *\}$. Now (b) implies that $\mathbf{u}[x, y_1, y_2] \sim \mathbf{v}[x, y_1, y_2]$, whence it is routine to check that for any $\mathbf{u}[x, y_1, y_2] \approx \mathbf{v}[x, y_1, y_2]$, there exists an appropriate $i \in \{1, 2\}$ such that $\mathbf{u}[x, y_i] \approx \mathbf{v}[x, y_i]$ is one of the following identities:

$$x^2 y_i \approx y_i x^2, \quad (x^*)^2 y_i \approx y_i (x^*)^2, \quad x^{\circledast_1} x^{\circledast_2} y_i \approx x^{\circledast_3} y_i x^{\circledast_4}, \quad y_i x^{\circledast_1} x^{\circledast_2} \approx x^{\circledast_3} y_i x^{\circledast_4},$$

where $\circledast_1, \circledast_2, \circledast_3, \circledast_4 \in \{1, *\}$ are such that $\{\circledast_1, \circledast_2\} = \{\circledast_3, \circledast_4\}$. However, by Remark 6.3, none of these identities is satisfied by $(M_4, {}^*)$, so we have a contradiction.

Therefore, for any $y_1, y_2 \in \mathsf{sim}(\mathbf{u}) = \mathsf{sim}(\mathbf{v})$, the word $y_1 y_2$ is a factor of $\mathbf{u}$ if and only if it is a factor of $\mathbf{v}$. The present lemma thus follows from (c). □

LEMMA 6.9. $\mathbf{u}_0 = \mathbf{v}_0$.

PROOF. Suppose that $\mathsf{con}(\mathbf{u}_0) \neq \mathsf{con}(\mathbf{v}_0)$, say $x, x^* \in \mathsf{con}(\mathbf{u}_0) \backslash \mathsf{con}(\mathbf{v}_0)$. Then since $x, x^* \in \mathsf{non}(\mathbf{v})$ by (b), the variables x, x^* occur in the factors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. However,

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

these factors are bipartite by (CF4), so the variables x, x^* cannot occur in the same $\mathbf{v}_i$, whence their occurrence in $\mathbf{v}$ must sandwich some simple variable y. Then $\mathbf{u}[x,y] = xx^*y$ and $\mathbf{v}[x,y] \in \{xyx^*, x^*yx\}$. It follows that $(M_4, {}^*)$ satisfies an identity from (6.2), which is impossible by Remark 6.3. Therefore, $\mathsf{con}(\mathbf{u}_0) = \mathsf{con}(\mathbf{v}_0)$, whence $\mathbf{u}_0 = \mathbf{v}_0$ by (CF3). □

LEMMA 6.10. $\mathbf{u}_m = \mathbf{v}_m$.

PROOF. Suppose that $\mathsf{con}(\mathbf{u}_m) \neq \mathsf{con}(\mathbf{v}_m)$, say $x \in \mathsf{con}(\mathbf{u}_m) \setminus \mathsf{con}(\mathbf{v}_m)$. Generality is not lost by assuming that $x \in \mathscr{A}$. It follows from (a), (b), (d) and (CF4) that there are three cases. (In each case, let y be any simple variable in $\mathbf{s}_m$.)

Case 1: $x^{\circledast} \in \mathsf{con}(\mathbf{u}_i)$ for some $i \in \{1, 2, \dots, m-1\}$ with $\circledast \in \{1, *\}$ and $x^* \notin \mathsf{con}(\mathbf{v}_m)$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{u}_1 \cdot \mathbf{s}_2\mathbf{u}_2 \cdots \mathbf{s}_{m-1}\mathbf{u}_{m-1}}_{x^{\circledast}} \cdot \mathbf{s}_m \underbrace{\ \mathbf{u}_m\ }_{x},$$

$$\mathbf{v} = \mathbf{v}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{v}_1 \cdot \mathbf{s}_2\mathbf{v}_2 \cdots \mathbf{s}_{m-1}\mathbf{v}_{m-1}}_{x \text{ and } x^{\circledast}} \cdot \mathbf{s}_m\mathbf{v}_m.$$

Hence, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is either $x^{\circledast}yx \approx xx^{\circledast}y$ or $x^{\circledast}yx \approx x^{\circledast}xy$, which contradicts Remark 6.3.

Case 2: $x^* \in \mathsf{con}(\mathbf{u}_i)$ for some $i \in \{1, 2, \dots, m-1\}$ and $x^* \in \mathsf{con}(\mathbf{v}_m)$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{u}_1 \cdot \mathbf{s}_2\mathbf{u}_2 \cdots \mathbf{s}_{m-1}\mathbf{u}_{m-1}}_{x^*} \cdot \mathbf{s}_m \underbrace{\ \mathbf{u}_m\ }_{x},$$

$$\mathbf{v} = \mathbf{v}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{v}_1 \cdot \mathbf{s}_2\mathbf{v}_2 \cdots \mathbf{s}_{m-1}\mathbf{v}_{m-1}}_{x} \cdot \mathbf{s}_m \underbrace{\ \mathbf{v}_m\ }_{x^*}.$$

Hence, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is $x^*yx \approx xyx^*$, which contradicts Remark 6.3.

Case 3: $\mathsf{occ}(x, \mathbf{u}_m) = 2$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1\mathbf{u}_1 \cdot \mathbf{s}_2\mathbf{u}_2 \cdots \mathbf{s}_{m-1}\mathbf{u}_{m-1} \cdot \mathbf{s}_m \underbrace{\ \mathbf{u}_m\ }_{x^2},$$

$$\mathbf{v} = \mathbf{v}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{v}_1 \cdot \mathbf{s}_2\mathbf{v}_2 \cdots \mathbf{s}_{m-1}\mathbf{v}_{m-1}}_{\text{two occurrences of } x} \cdot \mathbf{s}_m\mathbf{v}_m.$$

Hence, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is $yx^2 \approx x^2y$, which contradicts Remark 6.3.

Since all three cases are impossible, we must have $\mathsf{con}(\mathbf{u}_m) = \mathsf{con}(\mathbf{v}_m)$.

Now suppose that $\mathbf{u}_m \neq \mathbf{v}_m$. Then by (CF4), there exists some $x \in \mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v})$ such that $\mathsf{occ}(x, \mathbf{u}_m) \neq \mathsf{occ}(x, \mathbf{v}_m)$. Generality is not lost by assuming that $\mathsf{occ}(x, \mathbf{u}_m) = 2$ and $\mathsf{occ}(x, \mathbf{v}_m) = 1$ with $x \in \mathscr{A}$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1\mathbf{u}_1 \cdots \mathbf{s}_{m-1}\mathbf{u}_{m-1} \cdot \mathbf{s}_m \underbrace{\ \mathbf{u}_m\ }_{x^2},$$

$$\mathbf{v} = \mathbf{v}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{v}_1 \cdots \mathbf{s}_{m-1}\mathbf{v}_{m-1}}_{x} \cdot \mathbf{s}_m \underbrace{\ \mathbf{v}_m\ }_{x}.$$

Let y be any simple variable in $\mathbf{s}_m$. Then, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is $yx^2 \approx xyx$, which contradicts Remark 6.3. Consequently, $\mathbf{u}_m = \mathbf{v}_m$. □

LEMMA 6.11. $\mathbf{u}_i = \mathbf{v}_i$ *for all* $i = 1, 2, \ldots, m-1$.

PROOF. Suppose that $\ell \in \{1, 2, \ldots, m-1\}$ is the least index such that $\mathbf{u}_\ell \neq \mathbf{v}_\ell$. Then $\mathbf{u}_i = \mathbf{v}_i$ for all $i < \ell$. First, suppose that $\mathsf{con}(\mathbf{u}_\ell) \neq \mathsf{con}(\mathbf{v}_\ell)$. Then generality is not lost by assuming the existence of some plain variable $x \in \mathsf{con}(\mathbf{u}_\ell) \backslash \mathsf{con}(\mathbf{v}_\ell)$. It follows from (a), (b), (d) and (CF4) that there are four cases. (In each case, let y be any simple variable in $\mathbf{s}_{\ell+1}$.)

Case 1: $x^{\circledast} \in \mathsf{con}(\mathbf{u}_i)$ for some $i \in \{1, 2, \ldots, \ell-1\}$ with $\circledast \in \{1, *\}$ and $x^* \notin \mathsf{con}(\mathbf{v}_\ell)$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1}}_{x^{\circledast}} \cdot \mathbf{s}_\ell \underbrace{\mathbf{u}_\ell}_{x} \cdot \mathbf{s}_{\ell+1} \mathbf{u}_{\ell+1} \cdots \mathbf{s}_m \mathbf{u}_m,$$

$$\mathbf{v} = \mathbf{u}_0 \cdot \mathbf{s}_1 \underbrace{\mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1}}_{x^{\circledast}} \cdot \mathbf{s}_\ell \mathbf{v}_\ell \cdot \mathbf{s}_{\ell+1} \underbrace{\mathbf{v}_{\ell+1} \cdots \mathbf{s}_m \mathbf{v}_m}_{x}.$$

Hence, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is $x^{\circledast}xy \approx x^{\circledast}yx$, which contradicts Remark 6.3.

Case 2: $\mathsf{occ}(x, \mathbf{u}_\ell) = 2$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \underbrace{\mathbf{u}_\ell}_{x^2} \cdot \mathbf{s}_{\ell+1} \mathbf{u}_{\ell+1} \cdots \mathbf{s}_m \mathbf{u}_m,$$

$$\mathbf{v} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \mathbf{v}_\ell \cdot \mathbf{s}_{\ell+1} \underbrace{\mathbf{v}_{\ell+1} \cdots \mathbf{s}_m \mathbf{v}_m}_{\text{two occurrences of } x}.$$

Hence, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is $x^2y \approx yx^2$, which contradicts Remark 6.3.

Case 3: $x^{\circledast} \in \mathsf{con}(\mathbf{u}_i)$ for some $i \in \{\ell+1, \ell+2, \ldots, m\}$ with $\circledast \in \{1, *\}$ and $x^* \notin \mathsf{con}(\mathbf{v}_\ell)$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \underbrace{\mathbf{u}_\ell}_{x} \cdot \mathbf{s}_{\ell+1} \underbrace{\mathbf{u}_{\ell+1} \cdots \mathbf{s}_m \mathbf{u}_m}_{x^{\circledast}},$$

$$\mathbf{v} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \mathbf{v}_\ell \cdot \mathbf{s}_{\ell+1} \underbrace{\mathbf{v}_{\ell+1} \cdots \mathbf{s}_m \mathbf{v}_m}_{x \text{ and } x^{\circledast}}.$$

Hence, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is either $xyx^{\circledast} \approx yxx^{\circledast}$ or $xyx^{\circledast} \approx yx^{\circledast}x$, which contradicts Remark 6.3.

Case 4: $x^* \in \mathsf{con}(\mathbf{u}_i)$ for some $i \in \{\ell+1, \ell+2, \ldots, m\}$ and $x^* \in \mathsf{con}(\mathbf{v}_\ell)$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \underbrace{\mathbf{u}_\ell}_{x} \cdot \mathbf{s}_{\ell+1} \underbrace{\mathbf{u}_{\ell+1} \cdots \mathbf{s}_m \mathbf{u}_m}_{x^*},$$

$$\mathbf{v} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \underbrace{\mathbf{v}_\ell}_{x^*} \cdot \mathbf{s}_{\ell+1} \underbrace{\mathbf{v}_{\ell+1} \cdots \mathbf{s}_m \mathbf{v}_m}_{x}.$$

Hence, $\mathbf{u}[x,y] \approx \mathbf{v}[x,y]$ is $xyx^* \approx x^*yx$, which contradicts Remark 6.3.

https://doi.org/10.1017/S0004972723000989 Published online by Cambridge University Press

Since all four cases are impossible, we must have $\mathsf{con}(\mathbf{u}_\ell) = \mathsf{con}(\mathbf{v}_\ell)$. Then by (CF4), there exists some $x \in \mathsf{non}(\mathbf{u}) = \mathsf{non}(\mathbf{v})$ such that $\mathsf{occ}(x, \mathbf{u}_\ell) \neq \mathsf{occ}(x, \mathbf{v}_\ell)$. Generality is not lost by assuming $\mathsf{occ}(x, \mathbf{u}_\ell) = 2$ and $\mathsf{occ}(x, \mathbf{v}_\ell) = 1$ with $x \in \mathscr{A}$. Then

$$\mathbf{u} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \underbrace{\mathbf{u}_\ell}_{x^2} \cdot \mathbf{s}_{\ell+1} \mathbf{u}_{\ell+1} \cdots \mathbf{s}_m \mathbf{u}_m,$$

$$\mathbf{v} = \mathbf{u}_0 \cdot \mathbf{s}_1 \mathbf{u}_1 \cdots \mathbf{s}_{\ell-1} \mathbf{u}_{\ell-1} \cdot \mathbf{s}_\ell \underbrace{\mathbf{v}_\ell}_{x} \cdot \mathbf{s}_{\ell+1} \underbrace{\mathbf{v}_{\ell+1} \cdots \mathbf{s}_m \mathbf{v}_m}_{x}.$$

Let y be any simple variable in $\mathbf{s}_{\ell+1}$. Then, $\mathbf{u}[x, y] \approx \mathbf{v}[x, y]$ is $x^2 y \approx xyx$, which contradicts Remark 6.3. Consequently, the index ℓ does not exist and the present lemma is established. □

Acknowledgements

The authors are very grateful to the anonymous referees whose careful reading and helpful suggestions led to the improvement of this paper. They also thank Edmond W. H. Lee for pointing out that every involution monoid of order four is commutative and for his help in checking and revising this paper.

References

[1] C. Adair, 'Bands with an involution', *J. Algebra* **75** (1982), 297–314.
[2] Y. A. Bahturin and A. Y. Ol'šanskiĭ, 'Identical relations in finite Lie rings', *Math. USSR-Sb.* **25** (1975), 507–523; translation of *Mat. Sb. (N.S.)* **96**(138) (1975), 543–559.
[3] S. Burris and H. P. Sankappanavar, *A Course in Universal Algebra* (Springer, New York, 1981).
[4] C. C. Edmunds, 'On certain finitely based varieties of semigroups', *Semigroup Forum* **15** (1977), 21–39.
[5] M. Gao, W. T. Zhang and Y. F. Luo, 'A non-finitely based involution semigroup of order five', *Algebra Universalis* **81** (2020), Paper no. 31.
[6] M. Jackson and M. V. Volkov, 'The algebra of adjacency patterns: Rees matrix semigroups with reversion', in: *Fields of Logic and Computation* (eds. A. Blass, N. Dershowitz and W. Reisig) (Springer, Berlin, 2010), 414–443.
[7] R. Kruse, 'Identities satisfied in a finite ring', *J. Algebra* **26** (1973), 298–318.
[8] I. V. L'vov, 'Varieties of assiciative rings I', *Algebra i Logika* **12** (1973), 269–297.
[9] E. W. H. Lee, 'Finite basis problem for semigroups of order five or less: generalization and revisitation', *Studia Logica* **101** (2013), 95–115.
[10] E. W. H. Lee, 'Finitely based finite involution semigroups with nonfinitely based reducts', *Quaest. Math.* **39** (2016), 217–243.
[11] E. W. H. Lee, 'Equational theories of unstable involution semigroups', *Electron. Res. Announc. Math. Sci.* **24** (2017), 10–20.
[12] E. W. H. Lee, 'Non-finitely based finite involution semigroups with finitely based semigroup reducts', *Korean J. Math.* **27** (2019), 53–62.
[13] E. W. H. Lee, 'Varieties of involution monoids with extreme properties', *Q. J. Math.* **70** (2019), 1157–1180.
[14] E. W. H. Lee, 'Intervals of varieties of involution semigroups with contrasting reduct intervals', *Boll. Unione Mat. Ital.* **15** (2022), 527–540.
[15] E. W. H. Lee, *Advances in the Theory of Varieties of Semigroups* (Birkhäuser, Cham, 2023).

[16] E. W. H. Lee and J. R. Li, 'Minimal non-finitely based monoids', *Dissertationes Math.* **475** (2011), 65 pages.
[17] E. W. H. Lee, J. R. Li and W. T. Zhang, 'Minimal non-finitely based semigroups', *Semigroup Forum* **85** (2012), 577–580.
[18] E. W. H. Lee and W. T. Zhang, 'Finite basis problem for semigroups of order six', *LMS J. Comput. Math.* **18**(1) (2015), 1–129.
[19] R. McKenzie, 'Equational bases for lattice theories', *Math. Scand.* **27** (1970), 24–38.
[20] S. Oates and M. B. Powell, 'Identical relations in finite groups', *J. Algebra* **1** (1964), 11–39.
[21] P. Perkins, 'Bases for equational theories of semigroups', *J. Algebra* **11** (1969), 298–314.
[22] A. N. Trahtman, 'Finiteness of identity bases of five-element semigroups', in: *Semigroups and Their Homomorphisms* (ed. E. S. Lyapin) (Leningrad State Pedagogical Institute, Leningrad, 1991), 76–97 (in Russian).

BIN BIN HAN, School of Mathematics and Statistics,
Lanzhou University, Lanzhou, Gansu 730000, PR China
e-mail: hanbb19@lzu.edu.cn

WEN TING ZHANG, School of Mathematics and Statistics,
Lanzhou University, Lanzhou, Gansu 730000, PR China
e-mail: zhangwt@lzu.edu.cn

YAN FENG LUO, School of Mathematics and Statistics,
Lanzhou University, Lanzhou, Gansu 730000, PR China
e-mail: luoyf@lzu.edu.cn