

ON RESIDUALLY FINITE VARIETIES OF INVOLUTION SEMIGROUPS

IGOR DOLINKA

(Received 9 March 2010; accepted 24 October 2010)

Communicated by M. G. Jackson

Dedicated to the memory of my father Vojin (1937–2009)

Abstract

We prove that the variety consisting of all involutory inflations of normal bands is the unique maximal residually finite variety consisting of combinatorial semigroups with involution.

2010 *Mathematics subject classification*: primary 20M07; secondary 08B26.

Keywords and phrases: involution semigroup, variety, residually finite, subdirectly irreducible.

1. Introduction and formulation of the main result

A universal algebra $\mathbf{A} = (A, F)$ is *residually finite* if, for all $a, a' \in A$ such that $a \neq a'$, there exist a finite algebra \mathbf{B} (of the same similarity type as \mathbf{A}) and a homomorphism $\varphi : \mathbf{A} \rightarrow \mathbf{B}$ such that $\varphi(a) \neq \varphi(a')$. A variety \mathcal{V} is residually finite if each algebra from \mathcal{V} is residually finite. It is not difficult to see that a variety is residually finite if and only if all of its subdirectly irreducible members are finite. Recall that an algebra \mathbf{A} is *subdirectly irreducible* if $\theta_i \neq \Delta_{\mathbf{A}}$ for all $i \in I$ implies that $\bigcap_{i \in I} \theta_i \neq \Delta_{\mathbf{A}}$, for each family of its congruences $\{\theta_i : i \in I\}$. In other words, \mathbf{A} has a minimum nonidentical congruence μ , which is called the *monolith* of \mathbf{A} . Similarly, \mathbf{A} is *not* residually finite if and only if the intersection of each family of nonidentical congruences of \mathbf{A} of finite index (that is, with finitely many equivalence classes) is not equal to $\Delta_{\mathbf{A}}$.

A closely related notion is that of a *subdirect product* of algebras $\{\mathbf{A}_i : i \in I\}$: this is just a subalgebra of the direct product $\prod_{i \in I} \mathbf{A}_i$ whose i th projection coincides with the whole of \mathbf{A}_i for all $i \in I$. It is now quite easy to see that subdirectly irreducible algebras are precisely those which have, in a sense, only trivial decompositions into subdirect products: that is, for such an algebra \mathbf{A} we must have $\mathbf{A} \cong \mathbf{A}_i$ for some $i \in I$. On the

The research of the author was supported by Grant No. 144011 of the Ministry of Science and Technological Development of the Republic of Serbia.

© 2010 Australian Mathematical Publishing Association Inc. 1446-7887/2010 \$16.00

other hand, by a well-known theorem of Birkhoff, any algebra is isomorphic to a subdirect product of subdirectly irreducible algebras; therefore subdirectly irreducible algebras serve as basic building blocks of general algebras. This is true in varieties, too: any variety is uniquely determined by its class of subdirectly irreducible members. For additional background, consult [3, 18].

From the point of view just presented, residually finite varieties are the simplest; indeed—as residual finiteness is obviously a hereditary varietal property—they constitute the lower regions in the lattice of varieties of a given type. Characterizations of residually finite varieties are known in several classes of algebras. For example, by a well-known result of Ol’shanskiĭ [25], a group variety is residually finite if and only if it is generated by a finite group all of whose Sylow subgroups are Abelian. On the other hand, Golubov and Sapir [17] provide several characterizations of residually finite semigroup varieties. We shall return to some aspects of their result later. Independent descriptions of residually finite varieties of semigroups have been supplied (as corollaries to solutions of some more general questions) by Kublanovskĭĭ [20] and McKenzie [22, 23].

The aim of this paper is to initiate an investigation of residually finite varieties of *involution semigroups*. Recall that an involution semigroup is a natural ‘intermediate’ notion between (plain) semigroups and groups: we have a semigroup $(S, \cdot, *)$ equipped with a unary operation $*$, subject to the identities $(xy)^* \approx y^*x^*$ and $(x^*)^* \approx x$. For example, any inverse semigroup is an involution semigroup satisfying two additional identities $x \approx xx^*x$ and $xx^*yy^* \approx yy^*xx^*$.

To formulate our result, let us quickly introduce several varieties of involution semigroups, along with notes on their subvarieties:

- \mathcal{O}^* , the variety of all null semigroups with involution (whose only nontrivial proper subvariety is \mathcal{O}^{id} , defined by $x \approx x^*$, the variety of null semigroups with trivial involution [5]);
- \mathcal{RB}^* , the variety of all rectangular bands with involution (which is a minimal involution semigroup variety [14]);
- \mathcal{SL}^* , the variety of all involution semilattices (which has three nontrivial proper subvarieties [8, Theorem 3.1]);
- $\mathcal{NB}^* = \mathcal{RB}^* \vee \mathcal{SL}^*$, the variety of all normal bands with involution (whose subvariety lattice was determined in [9]).

Recall that a semigroup S is *combinatorial* if all of its subgroups are trivial. In this paper we set ourselves the task of proving the following result.

THEOREM 1.1. *Let \mathcal{V} be a variety consisting of combinatorial semigroups with involution. Then \mathcal{V} is residually finite if and only if it is contained in $\mathcal{O}^* \vee \mathcal{NB}^*$.*

We briefly justify the backward implication. By [5, Lemmas 1.7.11 and 1.7.12], the variety $\mathcal{O}^* \vee \mathcal{NB}^*$ consists of all subdirect products of null semigroups with involution and normal bands with involution. As remarked in [5], \mathcal{O}^* has two subdirectly irreducible algebras: the two-element null semigroup with an identical involution, and the three-element null semigroup whose involution fixes the zero and exchanges the

other two elements. On the other hand, the subdirectly irreducible members of \mathcal{NB}^* were discussed in [11, Theorem 7.4]: the largest of them has nine elements. Hence $\mathcal{O}^* \vee \mathcal{NB}^*$ is residually finite.

The proof of the forward implication in the theorem above occupies most of the rest of the paper. In what follows, \mathcal{V} will stand for a residually finite involution semigroup variety in which the underlying semigroup of each member is combinatorial. But first let us introduce the rest of the necessary concepts and notation, and then take a look at some examples of involution semigroup varieties that are not residually finite.

2. Further definitions and notation

We say that a semigroup with zero S is a *nil-semigroup* if for each $a \in S$ there exists an integer $k \geq 1$ such that $a^k = 0$. A very special type of nil-semigroup is an *n-nilpotent semigroup*; in such a semigroup, $a_1 \cdots a_n = 0$ for all $a_1, \dots, a_n \in S$; S is said to be *nilpotent* if it is n -nilpotent for some n . A 2-nilpotent semigroup, in which every product is equal to 0, is called *null*.

As usual, \mathcal{O} , \mathcal{SL} , \mathcal{RB} and \mathcal{NB} denote the varieties of all null semigroups, semilattices, rectangular bands and normal bands, respectively. For a semigroup variety \mathcal{U} , we denote by \mathcal{U}^* the class of all involution semigroups whose semigroup reducts belong to \mathcal{U} . It is quickly seen that \mathcal{U}^* is always an involution semigroup variety, defined by the involution laws and the plain semigroup identities of \mathcal{U} , thus making the notation from the introductory section consistent.

In the following we shall frequently encounter a situation where a semigroup S has an ideal I such that S/I is a nil-semigroup; then S is said to be a *nil-extension* of I by (a semigroup isomorphic to) S/I . In the particular case where S/I is null we are concerned with a *null extension* of I , which is equivalent to saying that $S^2 \subseteq I$. Now if S is an involution semigroup and a nil-extension (or a null extension) of its $*$ -ideal $I = I^*$, then S/I is a nil-semigroup (or a null semigroup) with involution, and S is said to be an *involution nil-extension* (or a null extension, respectively) of I .

A special, and in a way the simplest, type of null extension is obtained from the construction of an *inflation* of a semigroup T . Here we construct a semigroup on a set $S \supseteq T$ by starting from a function $f : S \rightarrow T$ such that f_T is the identity mapping on T and defining

$$a \cdot_S b = f(a) \cdot_T f(b)$$

for all $a, b \in T$. It is easy to see that in this way all products from T are preserved and that T is an ideal of S such that $S^2 \subseteq T$. In the case where T is an involution semigroup (with unary operation $*$) and we have an involution $'$ on the set $S \setminus T$ such that f agrees with $'$ in the sense that, for all $a \in S \setminus T$,

$$f(a') = f(a)^*,$$

then S may be turned into an involution semigroup by defining $a^* = a'$ for all $a \in S \setminus T$. In the setting just described, we say that S is an *involution inflation* of T .

As recorded in [27, Proposition 2.6], for a semigroup variety \mathcal{U} , the variety $\mathcal{O} \vee \mathcal{U}$ consists of all inflations of members of \mathcal{U} . This is true in the involutory setting as well: for each involution semigroup variety \mathcal{W} , the members of $\mathcal{O}^* \vee \mathcal{W}$ are characterized as involutory inflations of members of \mathcal{W} . Therefore the variety $\mathcal{O}^* \vee \mathcal{NB}^*$ from our Theorem 1.1 is precisely the variety of all involutory inflations of normal bands. In addition, if $\mathcal{U} \in \{\mathcal{SL}, \mathcal{RB}, \mathcal{NB}\}$ (and in many other cases), then $\mathcal{O}^* \vee \mathcal{U}^* = (\mathcal{O} \vee \mathcal{U})^*$, so we use these two forms interchangeably.

A congruence ρ of a semigroup S is called a *semilattice congruence* if S/ρ is a semilattice; S is *semilattice indecomposable* if its only semilattice congruence is its universal relation. It is well known (consult, for example, [26, Ch. II] or [29]) that each semigroup S has a least semilattice congruence ρ_0 such that all ρ_0 -classes are semilattice indecomposable subsemigroups of S . It is not difficult to show that if S is an involution semigroup, then the least semilattice congruence of its semigroup reduct agrees with $*$, so that it is in fact a congruence of S , the least *involution semilattice congruence* of S . If ρ is an (involution) semilattice congruence of a semigroup S (with involution) such that each ρ -class belongs to a class of semigroups \mathcal{C} , then we refer to S as an (*involution*) *semilattice of semigroups from \mathcal{C}* . Notice that in the involutory case, not every ρ -class needs to be fixed by $*$, so it is in general not expected to be an involution subsemigroup of S . All semilattice decompositions explicitly mentioned in this paper will have semilattice indecomposable classes, so we can take for granted that whenever an involution is involved, it actually defines an involution semilattice decomposition of the involution semigroup considered.

3. Some examples

EXAMPLE 3.1. First of all, recall the well-known five-element combinatorial *Brandt semigroup* B_2 generated by two elements a, b and a zero 0 , subject to the defining relations $a^2 = b^2 = 0$, $aba = a$ and $bab = b$. However, B_2 may also be viewed as a semigroup reduct of an inverse semigroup, also denoted by B_2 , generated by a and 0 such that $a^* = b$, $a^2 = 0$ and $aa^*a = a$. This is just a special case of the following construction. Let I be any nonempty set; define a multiplication on $(I \times I) \cup \{0\}$ by

$$(i_1, j_1)(i_2, j_2) = \begin{cases} (i_1, j_2) & \text{if } j_1 = i_2, \\ 0 & \text{otherwise,} \end{cases}$$

and $(i, j)^* = (j, i)$, for all $i, i_1, i_2, j, j_1, j_2 \in I$. The resulting involution (in fact, inverse) semigroup is denoted by B_I , and when $|I| = 2$ it is isomorphic to B_2 . It is well known [19, 28] that when $|I| \geq 2$ each B_I generates the same variety \mathcal{B} of *combinatorial strict inverse semigroups* [2, 30], defined within inverse semigroups by $xyx^* \approx xy^2x^*$. It is fairly easy to check that B_I is (congruence-)simple for each nonempty set I , and thus subdirectly irreducible; therefore \mathcal{B} is not residually finite.

Before presenting the next example, we consider a general construction of the 0-direct union of a semigroup with its dual. Let S be a semigroup; denote by S^∂

its dual semigroup: $S^\partial = \{\bar{a} : a \in S\}$ with $\bar{a} \cdot \bar{b} = \overline{ba}$. Upon assuming that $S \cap S^\partial = \emptyset$ and $\mathbf{0} \notin S \cup S^\partial$, define a semigroup on the set $S \cup S^\partial \cup \{\mathbf{0}\}$ whose multiplication \circ is given by $a \circ b = ab$ for all $a, b \in S$ and $\bar{a} \circ \bar{b} = \overline{ba}$ for all $\bar{a}, \bar{b} \in S^\partial$, while all the other products are equal to $\mathbf{0}$. The involution is defined by $\mathbf{0}^* = \mathbf{0}$ and $a^* = \bar{a}$ and $\bar{a}^* = a$ for all $a \in S$. In this way we obtain an involution semigroup, denoted by $I_0^*(S)$.

A characterization of the identities satisfied by these involution semigroups is given in [6] and we reproduce it here, as it will be repeatedly used later. Due to the involution semigroup axioms, the elements of the free involution semigroup on the set X can be considered simply as (nonempty) words over the ‘doubled’ alphabet $X \cup X^*$, where $X^* = \{x^* : x \in X\}$. Such words we call *involutionary words*. Hence involution semigroup identities are of the form $\mathbf{u} \approx \mathbf{v}$ for two involutory words \mathbf{u}, \mathbf{v} . To distinguish between the elements of the set $X \cup X^*$, considered as the alphabet for involutory words, and symbols from X (which, for example, may be a part of a term of the form x^*), we call the former *letters* and the latter *variables*. A variable $x \in X$ such that both letters x and x^* occur in \mathbf{u} is called *paired* (in \mathbf{u}); the set of all paired variables in \mathbf{u} is denoted by $\pi(\mathbf{u})$. A (plain) semigroup identity $\mathbf{w}_1 \approx \mathbf{w}_2$ is called *homotypical* if the set of variables occurring in \mathbf{w}_1 coincides with that of \mathbf{w}_2 . If $\overleftarrow{\mathbf{w}}$ is the word obtained from \mathbf{w} by reading it backwards, then the *reverse* of $\mathbf{w}_1 \approx \mathbf{w}_2$ is the identity $\overleftarrow{\mathbf{w}}_1 \approx \overleftarrow{\mathbf{w}}_2$. We now provide the reformulation of [6, Theorem 6] given in [12].

PROPOSITION 3.2. *Let \mathbf{u}, \mathbf{v} be involutory words and let S be a semigroup. Then the identity $\mathbf{u} \approx \mathbf{v}$ holds in $I_0^*(S)$ if and only if either:*

- (a) $\pi(\mathbf{u}) \neq \emptyset$ and $\pi(\mathbf{v}) \neq \emptyset$; or
- (b) $\pi(\mathbf{u}) = \pi(\mathbf{v}) = \emptyset$ and $\mathbf{u} \approx \mathbf{v}$ is obtained from a homotypical semigroup identity satisfied by S whose reverse is also true in S , by replacing some of the variables by their stars.

For a semigroup variety \mathcal{U} , let \mathcal{U}^∂ denote its *dual variety*: this is defined by the reverses of identities of \mathcal{U} , and if S generates \mathcal{U} , then S^∂ generates \mathcal{U}^∂ . A variety \mathcal{U} is *self-dual* if $\mathcal{U} = \mathcal{U}^\partial$. We note the following easy consequence of the proposition above.

LEMMA 3.3. *Let \mathcal{U} be a semigroup variety generated by a semigroup S . If $T \in \mathcal{U} \vee \mathcal{U}^\partial \vee \mathcal{SL}$, then $I_0^*(T)$ belongs to the involution semigroup variety generated by $I_0^*(S)$.*

PROOF. Let $\mathbf{u} \approx \mathbf{v}$ be an identity satisfied by $I_0^*(S)$. If $\pi(\mathbf{u}) \neq \emptyset$ and $\pi(\mathbf{v}) \neq \emptyset$, then $\mathbf{u} \approx \mathbf{v}$ holds in $I_0^*(T)$ as well. Otherwise, assume that $\pi(\mathbf{u}) = \pi(\mathbf{v}) = \emptyset$. Then there is no loss of generality in assuming that $\mathbf{u} \approx \mathbf{v}$ is a semigroup identity. By the proposition above, it is homotypical and holds in both S and S^∂ , which is equivalent to saying that it holds in $\mathcal{U} \vee \mathcal{U}^\partial \vee \mathcal{SL}$. However, note that $\mathcal{U} \vee \mathcal{U}^\partial \vee \mathcal{SL}$ is a self-dual variety, so that $T \in \mathcal{U} \vee \mathcal{U}^\partial \vee \mathcal{SL}$ implies that $T^\partial \in \mathcal{U} \vee \mathcal{U}^\partial \vee \mathcal{SL}$. Hence $\mathbf{u} \approx \mathbf{v}$ holds in both T and T^∂ , and since it is homotypical, we conclude that it is satisfied by $I_0^*(T)$. \square

EXAMPLE 3.4. Let \mathcal{B}_0 be the variety generated by $I_0^*(B_2)$. For an arbitrary I satisfying $|I| > 1$, let θ be the congruence on $I_0^*(B_I)$ collapsing $0, \bar{0}$ and $\mathbf{0}$; write

$S_I = I_0^*(B_I)/\theta$. It is quite straightforward to check that S_I is congruence-free (and so subdirectly irreducible), just like B_I . On the other hand, \mathcal{B} is self-dual, as B_2 is a semigroup reduct of an involution semigroup, so its equational theory is closed under taking reverses. Since $\mathcal{SL} \subseteq \mathcal{B}$, Lemma 3.3 implies that $I_0^*(B_I)$ (and thus S_I) belongs to \mathcal{B}_0 , preventing the latter variety from being residually finite.

EXAMPLE 3.5. Let \mathcal{M}_1 be the variety of 3-nilpotent semigroups defined by identities $xyz \approx 0$, $x^2 \approx 0$ and $xy \approx yx$. As it is commutative, it can be trivially ‘expanded’ to an involution semigroup variety by the identical involution, $x^* \approx x$. The resulting variety is also denoted by \mathcal{M}_1 . As noted in [17], \mathcal{M}_1 is not residually finite.

LEMMA 3.6. *Let \mathcal{W} be an involution semigroup variety such that $\mathcal{M}_1 \not\subseteq \mathcal{W}$.*

- (i) \mathcal{W} satisfies an identity of the form $xy \approx \mathbf{w}$, where $|\mathbf{w}| \geq 3$.
- (ii) If \mathcal{W} is generated by a nil-semigroup with involution N , then N is null.

PROOF. (i) By the given condition, \mathcal{W} satisfies an identity that fails in \mathcal{M}_1 . Then after possible substitutions of variables by their stars, the identity in question is either of the form $x \approx \mathbf{u}$, where $|\mathbf{u}| \geq 2$, or of the form $xy \approx \mathbf{w}$, where either $|\mathbf{w}| = 2$ and \mathbf{w} contains a variable not in $\{x, y\}$, or \mathbf{w} contains exactly one of the variables $\{x, y\}$, or $|\mathbf{w}| \geq 3$. Clearly, an identity of the first type implies an identity of the second type, so we may conclude that \mathcal{W} satisfies a nontrivial identity $xy \approx \mathbf{w}$. Moreover, if $|\mathbf{w}| = 2$ and \mathbf{w} contains a letter z or z^* different from x, x^*, y, y^* , then z may be substituted by a product of any two letters in order to deduce an identity of the form $xy \approx \mathbf{w}_1$, $|\mathbf{w}_1| \geq 3$. Finally, if, for example, $\mathbf{w} \in \{x^2, xx^*, x^*x, (x^*)^2\}$, then $xy \approx \mathbf{w}$ implies that $xyz \approx \mathbf{w}$ and so $xy \approx xyz$. In any case, \mathcal{W} satisfies an identity of the required form.

(ii) First of all note that, from (i), whenever $m \geq 3$, there is an involutory word \mathbf{w}_m such that $|\mathbf{w}_m| \geq m$ and $xy \approx \mathbf{w}$ implies that $xy \approx \mathbf{w}_m$. Also from (i), the involutory nil-semigroup N satisfies an identity of the form $xyx \approx \mathbf{w}'$ for some involutory word \mathbf{w}' not graphically equal to xyx , so by [13, Proposition 2.6] the involution subsemigroup N' of N generated by two arbitrary elements $a, b \in N$ is finite. Since each finite nil-semigroup is nilpotent, N_1 satisfies the identity $x_1 \cdots x_m \approx 0$ for some m . Therefore N_1 satisfies $xy \approx \mathbf{w}_m \approx 0$, that is, N_1 is a null semigroup. In particular, $ab = 0$; but since the choice of a, b was arbitrary, we conclude that N must be null. □

EXAMPLE 3.7. Let S be a 3-nilpotent semigroup generating \mathcal{M}_1 . Define \mathcal{M}_1^0 to be the variety generated by $I_0^*(S)$ (Proposition 3.2 shows that the actual choice of S is immaterial); we claim that it is not residually finite. To see this, we exhibit concrete examples of infinite subdirectly irreducible algebras from the semigroup variety \mathcal{M}_1 . Let κ be any cardinal. Define N_κ to be the semigroup on $\{0, p\} \cup \{x_i : i < \kappa\}$ such that $p^2 = x_i p = p x_i = 0$ for all $i < \kappa$ and

$$x_i x_j = \begin{cases} p & \text{if } i \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

It is easily verified that $N_\kappa \in \mathcal{M}_1$ and that it is subdirectly irreducible, where the monolith is the congruence collapsing 0 and p . Now let $N'_\kappa = I_0^*(N_\kappa)/\theta$, where the only nontrivial block of θ is $\widehat{0} = \{0, \bar{0}, \mathbf{0}\}$. It suffices to see that each nontrivial congruence ρ of N'_κ collapses p and $\bar{p} = p^*$. This is clear when $(a, b) \in \rho$ satisfy $a \neq b$ and $a, b \in (N_\kappa \setminus \{0\}) \cup \{\widehat{0}\}$ or $a, b \in (N_\kappa^\partial \setminus \{\bar{0}\}) \cup \{\widehat{0}\}$, because ρ then contains either $(\widehat{0}, p)$ or $(\bar{0}, \bar{p})$, so it collapses $\widehat{0}, p, \bar{p}$. The remaining case is when $a, c \in N_\kappa \setminus \{0\}$ and $b = \bar{c}$. The case where $a = c = p$ is immediate. If $a = x_i$ for some i , then for all $j \neq i$ we have $(p, \widehat{0}) = (ax_j, \bar{c}x_j) \in \rho$, with the same conclusion as in the previous subcases. The possibility when $c = x_i$ for some i is analogous, so we are done.

LEMMA 3.8. *Let N be a nil-semigroup, and let \mathcal{W} be the variety generated by $I_0^*(N)$. If $\mathcal{M}_1^0 \not\subseteq \mathcal{W}$, then N is null.*

PROOF. Let S be a semigroup generating \mathcal{M}_1 ; then $I_0^*(S) \notin \mathcal{W}$, so there is an identity satisfied by $I_0^*(N)$ that fails in $I_0^*(S)$. Bearing in mind Proposition 3.2, this cannot be an identity of type (a). Therefore there must be a homotypical *semigroup* identity $\mathbf{u} \approx \mathbf{v}$ that holds in N and N^∂ and fails in S or S^∂ . In any case, there is a semigroup identity ($\mathbf{u} \approx \mathbf{v}$ or $\overleftarrow{\mathbf{u}} \approx \overleftarrow{\mathbf{v}}$) holding in N and failing in S . As noted in [17, proof of Lemma 3], the identity in question may be assumed to be of the form $xy \approx \mathbf{w}$, where \mathbf{w} is a word of length at least three. Hence the proof of this lemma can be finished like that of Lemma 3.6, but using [31, Lemma 3.2] instead of [13, Proposition 2.6]. \square

EXAMPLE 3.9. A band (an idempotent semigroup) is *regular* if it satisfies the identity $xyxzx \approx xyzx$. It is *left regular* if it satisfies $xy \approx xyx$. It is well known that the variety of left regular bands coincides with the class of all semilattices of left zero bands. A **-band* is a band with involution satisfying $x \approx xx^*x$. In [34], Scheiblich proved that each regular **-band* B is a *spined product* of a left regular band and its dual with respect to the structure semilattice of B : this is to say that there is a semilattice Y and a left regular band $L = \bigcup_{\alpha \in Y} L_\alpha$ such that $B = \bigcup_{\alpha \in Y} (L_\alpha \times R_\alpha)$, where $R_\alpha = L_\alpha^\partial$, while the involution is just the operation of exchanging pairs: $(a, \bar{b})^* = (b, \bar{a})$. It is a matter of straightforward (if somewhat tedious) verification to see that B is subdirectly irreducible if and only if L is. On the other hand, the variety of left regular bands is not residually finite (see, for example, [15, 16]), so it follows that the variety $\mathcal{ReB}^{\text{reg}}$ of regular **-bands* is not residually finite either.

EXAMPLE 3.10. If L_2 denotes the two-element left zero band, let \mathcal{ReB}^0 be the variety generated by $I_0^*(L_2^1)$. This is the variety defined within regular bands with involution by the identities $xx^*y \approx xyx^* \approx yxx^* \approx xx^*$. By [11, Theorem 6.2], if L is a subdirectly irreducible left regular band without a zero element, then $I_0^*(L)$ is a subdirectly irreducible member of \mathcal{ReB}^0 . By the already mentioned results of [15, 16], we conclude that \mathcal{ReB}^0 is not residually finite.

LEMMA 3.11. *Let \mathcal{U} be a variety of involution bands that does not contain $\mathcal{ReB}^{\text{reg}}$ and \mathcal{ReB}^0 . Then $\mathcal{U} \subseteq \mathcal{NB}^*$.*

PROOF. This is just an immediate consequence of [8, Lemma 2.1]. □

Finally, we set out an elementary fact that will be used several times in the following proofs.

LEMMA 3.12. *Let S be a residually finite involution semigroup. Then all subsemigroups of its semigroup reduct are residually finite. In particular, if $I_0^*(T)$ is residually finite for a semigroup T , so is T .*

PROOF. Assume that S is residually finite. Then in particular, for arbitrary $a, b \in S$, $a \neq b$, there is an involution semigroup homomorphism $\varphi : S \rightarrow F$ into a finite involution semigroup F such that $\varphi(a) \neq \varphi(b)$. However, $\varphi|_T : T \rightarrow \overline{F}$ is then a plain semigroup homomorphism (where \overline{F} is the semigroup reduct of F) separating a and b . □

4. Proof of Theorem 1.1, part 1

Recall that throughout the rest of the proof of Theorem 1.1, \mathcal{V} denotes a residually finite variety of combinatorial semigroups with involution. The first structural feature we discover about members of \mathcal{V} is as follows.

LEMMA 4.1. *\mathcal{V} consists of involution semilattices of null extensions of rectangular bands.*

PROOF. We employ the main result of [12] (namely, Theorem 5), which implies that any involution semigroup variety omitting B_2 and $I_0^*(B_2)$ consists of involution semilattices of Archimedean semigroups. Recall that a semigroup Q is Archimedean if for all $a, b \in Q$ there is an integer $n \in \mathbb{N}$ such that $b^n = xay$ for some $x, y \in Q^1$. By Examples 3.1 and 3.4, this is the case with our variety \mathcal{V} . Therefore each $S \in \mathcal{V}$ can be decomposed as $S = \bigcup_{\alpha \in Y} S_\alpha$ for an involution semilattice Y , where each S_α is Archimedean. On the other hand, $\mathcal{M}_1 \not\subseteq \mathcal{V}$ by Example 3.5, so by Lemma 3.6(i) it follows that \mathcal{V} satisfies an identity of the form $xy \approx \mathbf{w}$ for an involutory word \mathbf{w} such that $|\mathbf{w}| \geq 3$. Our discussion now splits into two cases.

Case 1: \mathbf{w} is a plain semigroup word. By identifying all the variables involved, we conclude that \mathcal{V} satisfies $x^2 \approx x^{p+2}$ for some $p \geq 1$, so each member S of \mathcal{V} is periodic. It is well known that any periodic Archimedean semigroup, including each S_α , is actually a nil-extension of a completely simple semigroup (see, for example, [33]). So S_α has a completely simple ideal I_α such that S_α/I_α is a nil-semigroup for all $\alpha \in Y$. As we are dealing with the combinatorial case, each I_α is a rectangular band.

Case 2: \mathbf{w} contains the unary symbol $$, that is, a ‘starred’ letter.* Now Y is a homomorphic image of $S = \bigcup_{\alpha \in Y} S_\alpha \in \mathcal{V}$, so Y must satisfy the identity $xy \approx \mathbf{w}$, which in this case implies that the involution on Y is the identity mapping. Hence $S_\alpha = S_{\alpha^*} = S_\alpha^*$ is an involution subsemigroup of S for all $\alpha \in Y$. Further, by substituting xx^* for each variable occurring in $xy \approx \mathbf{w}$, we infer that $(xx^*)^2 \approx (xx^*)^{p+2}$ for some $p \geq 1$, which implies that $(xx^*)^{2p} \approx (xx^*)^{4p} = [(xx^*)^{2p}]^2$. Therefore each

Archimedean component S_α contains an idempotent, as for all $a \in S_\alpha$, the element $(aa^*)^{2p}$ belongs to S_α as well. By [29, Theorem 4.1], (the semigroup reduct of) S_α is a nil-extension of a simple semigroup I_α . In addition, I_α is a $*$ -ideal of S_α , and so is an involution subsemigroup of S that contains an idempotent. Now if I_α fails to be *completely* simple for some $\alpha \in Y$, then by the well-known result of Andersen [1] (see also [4, Theorem 2.54]) the semigroup reduct of I_α contains a copy of the *bicyclic semigroup* as a subsemigroup. Recall that the bicyclic semigroup is the inverse monoid presented by $\langle b, c : cb = 1 \rangle$; hence all of its elements are of the form $b^m c^n$, where $m, n \geq 0$, as the multiplication in B works as follows:

$$b^m c^n \cdot b^p c^q = b^{m-n+\max(n,p)} c^{q-p+\max(n,p)}.$$

The semigroup B is known not to be residually finite (see [7, 21]), so by Lemma 3.12 it follows that the involution semigroup $I_\alpha \in \mathcal{V}$ is not residually finite either, a contradiction. We conclude that each I_α is completely simple, and thus a rectangular band, just as in Case 1.

It remains to be seen that all nil-semigroups S_α/I_α are in fact null. Again, we have two cases to consider. If $\alpha = \alpha^*$, then $S_\alpha^* = S_{\alpha^*} = S_\alpha$, and I_α is a $*$ -ideal of the involution subsemigroup S_α of S . Hence S_α/I_α is a nil-semigroup with involution belonging to \mathcal{V} . Now $\mathcal{M}_1 \not\subseteq \mathcal{V}$, so \mathcal{M}_1 does not belong to the variety generated by S_α/I_α , and by Lemma 3.6(ii) the latter nil-semigroup is null. The other case occurs when $\alpha^* \neq \alpha$. Then $P = S_\alpha \cup S_{\alpha^*} \cup S_{\alpha\alpha^*}$ is an involution subsemigroup of S , while the equivalence δ on P whose only nonsingleton blocks are I_α, I_{α^*} and $S_{\alpha\alpha^*}$ is a congruence on P . The involution semigroup P/δ belongs to \mathcal{V} and is of the form $I_0^*(N)$, where $N \cong S_\alpha/I_\alpha$ is a nil-semigroup. Since $\mathcal{M}_1^0 \notin \mathcal{V}$, Lemma 3.8 implies that N must be null, as required. □

LEMMA 4.2. *In involution semigroups, each identity of the form*

$$xy \approx \mathbf{w}(x, y),$$

where \mathbf{w} is an involutory word such that $|\mathbf{w}| \geq 3$ and both variables x and y occur in \mathbf{w} , implies an identity of the form

$$xy \approx \mathbf{u}(x, y)\mathbf{v}(x, y),$$

where both variables x and y occur in both \mathbf{u} and \mathbf{v} . Moreover, if \mathbf{w} is a plain word, then both \mathbf{u} and \mathbf{v} may be assumed to be plain words as well.

PROOF. For an involutory word $\mathbf{w}_0 = \mathbf{w}_0(x, y)$, let $\delta_x(\mathbf{w}_0)$ denote the total number of occurrences of the variable x in \mathbf{w}_0 (both as x and x^*), that is, $\delta_x(\mathbf{w}_0) = |\mathbf{w}_0|_x + |\mathbf{w}_0|_{x^*}$. Also, if $\mathbf{u}_1, \mathbf{u}_2$ are involutory words, then $\mathbf{w}_0[\mathbf{u}_1, \mathbf{u}_2]$ will stand for the word obtained by substituting $x \rightarrow \mathbf{u}_1, y \rightarrow \mathbf{u}_2$ into \mathbf{w}_0 . In cases where \mathbf{w}_0 omits one of the variables x, y , we write only $\mathbf{w}_0[\mathbf{u}_1]$ to indicate that the variable that occurs in \mathbf{w}_0 has been substituted by \mathbf{u}_1 .

Since $|\mathbf{w}| \geq 3$, either $\delta_x(\mathbf{w}) \geq 2$ or $\delta_y(\mathbf{w}) \geq 2$; without loss of generality, assume that the first case occurs. Now write $\mathbf{w} = \mathbf{w}_1\mathbf{w}_2\mathbf{w}_3$, where $\delta_y(\mathbf{w}_1) = \delta_y(\mathbf{w}_3) = 0$ and \mathbf{w}_2

begins and ends with a letter from $\{y, y^*\}$. Then $\mathbf{w}_2 = \eta\mathbf{w}'_2$ for some $\eta \in \{y, y^*\}$, where the involutory word \mathbf{w}'_2 is either empty or is of the form $\mathbf{w}'_2 = \mathbf{w}''_2\eta'$ for some $\eta' \in \{y, y^*\}$.

Consider first the case where both \mathbf{w}_1 and $\mathbf{w}'_2\mathbf{w}_3$ are nonempty; then, in particular, $\mathbf{w}_1 \in \{x, x^*\}^+$. Write $\mathbf{w}_1 = \xi\mathbf{w}'_1$, where $\xi \in \{x, x^*\}$. We apply the substitution $x \rightarrow \mathbf{w}_1\eta, y \rightarrow \mathbf{w}'_2\mathbf{w}_3$ to the identity $xy \approx \mathbf{w}$, yielding

$$\mathbf{w} = (\mathbf{w}_1\eta)(\mathbf{w}'_2\mathbf{w}_3) \approx \mathbf{u} \cdot \mathbf{w}'_1[\mathbf{w}_1\eta]\mathbf{w}_2[\mathbf{w}_1\eta, \mathbf{w}'_2\mathbf{w}_3]\mathbf{w}_3[\mathbf{w}_1\eta, \mathbf{w}'_2\mathbf{w}_3],$$

where $\mathbf{u} \in \{\mathbf{w}_1\eta, \eta^*\mathbf{w}'_1\}$. In any case, $\delta_x(\mathbf{u}), \delta_y(\mathbf{u}) \neq 0$, since both variables x and y are by construction contained in $\mathbf{w}_1\eta$. On the other hand, by our assumptions $\delta_x(\mathbf{w}'_1\mathbf{w}_2\mathbf{w}_3) > 0$, so when

$$\mathbf{v} = \mathbf{w}'_1[\mathbf{w}_1\eta]\mathbf{w}_2[\mathbf{w}_1\eta, \mathbf{w}'_2\mathbf{w}_3]\mathbf{w}_3[\mathbf{w}_1\eta, \mathbf{w}'_2\mathbf{w}_3],$$

$\delta_x(\mathbf{v}), \delta_y(\mathbf{v}) \neq 0$ as well. Hence we infer an identity of the form $xy \approx \mathbf{u}\mathbf{v}$ with the desired properties.

If \mathbf{w}_1 is empty but \mathbf{w}'_2 is not, then $\delta_y(\mathbf{w}_2) \geq 2$ and we may apply essentially the same argument as above with the roles of x, y switched to obtain the same result. Therefore it remains to consider the case where either $\mathbf{w}_1\mathbf{w}'_2$ or $\mathbf{w}'_2\mathbf{w}_3$ is empty. Then our identity $xy \approx \mathbf{w}$ is of one of the forms $xy \approx \mathbf{u}_0\eta, xy \approx \eta\mathbf{u}_0$, where $\mathbf{u}_0 \in \{x, x^*\}^+$ and $|\mathbf{u}_0| \geq 2$. However, in the involution semigroup setting, $xy \approx \mathbf{u}_0\eta$ implies that $yx \approx (x^*y^*)^* \approx \eta\overleftarrow{\mathbf{u}_0}$, and thus $xy \approx \xi\mathbf{v}_0$, where $\xi \in \{x, x^*\}^+$ and $\mathbf{v}_0 = \overleftarrow{\mathbf{u}_0}[y] \in \{y, y^*\}^+$. So if $\mathbf{u}_0 = \mathbf{u}'_0\xi'$ for some $\xi' \in \{x, x^*\}$, then \mathbf{u}'_0 is nonempty, and we deduce the following chain of identities:

$$xy \approx \mathbf{u}_0\eta = \mathbf{u}'_0(\xi'\eta) \approx \mathbf{u}'_0\xi''\mathbf{v}_0[\eta],$$

where $\xi'' \in \{x, x^*\}$. Since \mathbf{u}'_0 is nonempty and $|\mathbf{v}_0| \geq 2$, the identity

$$xy \approx \mathbf{u}'_0\xi''\mathbf{v}_0[\eta]$$

has the form already discussed in the preceding cases. The same conclusion is reached if one starts from an identity of the form $xy \approx \eta\mathbf{u}_0$.

Finally, if \mathbf{w} contains no occurrences of the unary symbol $*$, then it is not difficult to see that all words constructed in the proof above, including \mathbf{u} and \mathbf{v} , fail to contain that symbol as well. Thus the proof of the lemma is complete. □

LEMMA 4.3. *\mathcal{V} consists of involutory null extensions of normal bands.*

PROOF. Since $\mathcal{M}_1 \not\subseteq \mathcal{V}$, Lemma 3.6(i) shows that \mathcal{V} satisfies an identity of the form $xy \approx \mathbf{w}$, where $|\mathbf{w}| \geq 3$. Moreover, it is straightforward to see that we may confine ourselves to the case where both x and y but no other variables occur in \mathbf{w} . By the previous lemma, the identity $xy \approx \mathbf{w}$ implies an identity of the form $xy \approx \mathbf{u}(x, y)\mathbf{v}(x, y)$, with both variables x and y occurring in both \mathbf{u} and \mathbf{v} .

Now let $S \in \mathcal{V}$ and let $a, b \in S$ be arbitrary. We already know that S is a semilattice of null extensions of rectangular bands, $S = \bigcup_{\alpha \in Y} S_\alpha$, and so $a \in S_\alpha$ and $b \in S_\beta$ for some $\alpha, \beta \in Y$. If \mathbf{w} is a plain word, then both \mathbf{u} and \mathbf{v} may be assumed to be plain

words, whence

$$ab = \mathbf{u}(a, b)\mathbf{v}(a, b) \in S_{\alpha\beta}^2.$$

On the other hand, assume that \mathbf{w} contains an occurrence of a ‘starred’ letter. Then Y must have an identical involution, implying that $S_{\gamma^*} = S_{\gamma}$ holds for all $\gamma \in Y$. Thus $a^* \in S_{\alpha}$ and $b^* \in S_{\beta}$, which allows us to deduce the same conclusion as above: $ab = \mathbf{u}(a, b)\mathbf{v}(a, b) \in S_{\alpha\beta}^2$. So in any case, ab is an idempotent of S ; in other words, S^2 is a band. By Lemma 3.11 we must have $S^2 \in \mathcal{NB}^*$, so S is an involutory null extension of a normal band. \square

We finish this section by recording a fact that is already contained in [27]. In the rest of the paper, we denote by \mathcal{B}_1 and \mathcal{C}_1 the varieties of all null extensions of rectangular and normal bands, respectively.

LEMMA 4.4. *Let $S = \bigcup_{\alpha \in Y} S_{\alpha}$, where each S_{α} is a null extension of the rectangular band S_{α}^2 , such that S^2 is a normal band. If \mathcal{U} is the semigroup variety generated by all S_{α} where $\alpha \in Y$, then $S \in \mathcal{U} \vee \mathcal{SL}$.*

PROOF. By [27, Lemma 5.3] we have $\mathcal{C}_1 = \mathcal{B}_1 \vee \mathcal{SL}$. So $S \in \mathcal{C}_1$, whence Proposition 2.9 of the same paper (which is in fact due to Mel’nik [24]) implies that S is a dense semilattice of the S_{α} (see the cited proposition for the definition of the dense semilattice construction). Conversely, since \mathcal{B}_1 satisfies a nonhomotypical identity (for example, $x^2 \approx x^2y^2x^2$), the same proposition yields that if $S_{\alpha} \in \mathcal{U}$ for all $\alpha \in Y$ and S is a dense semilattice of the S_{α} , then $S \in \mathcal{U} \vee \mathcal{SL}$. \square

The previous lemma implies that if we manage to prove, for each $S \in \mathcal{V}$ such that $S = \bigcup_{\alpha \in Y} S_{\alpha}$ (with S_{α} as above), that the semigroup reduct of each S_{α} is an inflation of the rectangular band S_{α}^2 , then we can complete the proof of Theorem 1.1. Indeed, as $\mathcal{O} \vee \mathcal{RB}$ is the variety consisting precisely of all inflations of rectangular bands, $\mathcal{U} \subseteq \mathcal{O} \vee \mathcal{RB}$ (where \mathcal{U} is the semigroup variety from the previous lemma). Hence, the semigroup reduct of S belongs to $\mathcal{O} \vee \mathcal{RB} \vee \mathcal{SL} = \mathcal{O} \vee \mathcal{NB}$, so that $S \in (\mathcal{O} \vee \mathcal{NB})^* = \mathcal{O}^* \vee \mathcal{NB}^*$, as required.

5. Proof of Theorem 1.1, part 2

In this section our goal is to prove the following result.

PROPOSITION 5.1. *Let $S \in \mathcal{V}$ such that $S = \bigcup_{\alpha \in Y} S_{\alpha}$, where each S_{α} is a null extension of the rectangular band S_{α}^2 . Then S_{α} is an inflation of S_{α}^2 for all $\alpha \in Y$.*

As we have just noted, once we prove this proposition, Theorem 1.1 will be proved as well.

We approach the proposition above by contradiction: if S_{α} fails to be an inflation of S_{α}^2 for some $\alpha \in Y$, then S generates a variety that is not residually finite. Two cases emerge: $\alpha^* = \alpha$ and $\alpha^* \neq \alpha$.

First we resolve the latter possibility. We have that $S_{\alpha} \cup S_{\alpha^*} \cup S_{\alpha\alpha^*}$ is an involution subsemigroup of S , in which $S_{\alpha\alpha^*}$ is a $*$ -ideal, and the corresponding Rees quotient

is isomorphic to $I_0^*(S_\alpha)$. Since S_α is not an inflation of S_α^2 , that is, $S_\alpha \notin \mathcal{O} \vee \mathcal{RB}$, from the description of the subvariety lattice of \mathcal{B}_1 (see, for example, [27, Diagrams 1 and 2]), it follows that the variety generated by S_α contains either the variety \mathcal{P} of all null extensions of left zero bands, defined by $xyz \approx xy$, or its dual variety \mathcal{Q} . Since $S_\alpha^\partial \cong S_{\alpha^*}$, there is no loss of generality in assuming that the former case occurs.

LEMMA 5.2. *For all $P \in \mathcal{P}$, $I_0^*(P)$ belongs to the variety generated by $I_0^*(S_\alpha)$.*

PROOF. It suffices to use Proposition 3.2 to see that each identity holding in $I_0^*(S_\alpha)$ holds in $I_0^*(P)$ as well. This is clear for the identities of type (a). On the other hand, for identities of type (b) it is enough to confine ourselves to plain semigroup identities $\mathbf{u} \approx \mathbf{v}$. This identity is homotypical and holds in both S_α and S_{α^*} , that is, in the smallest self-dual semigroup variety containing S_α . However, the latter variety can only be $\mathcal{P} \vee \mathcal{Q} = \mathcal{B}_1$ (see [27, Diagram 2, p. 122]). Therefore $\mathbf{u} \approx \mathbf{v}$ holds in both \mathcal{P} and \mathcal{Q} and thus in both P and P^∂ . This implies that $\mathbf{u} \approx \mathbf{v}$ is satisfied in $I_0^*(P)$. \square

It is proved in [17] that \mathcal{P} (denoted there by \mathcal{M}_2) is not residually finite by showing an explicit example of a null extension A of a left zero band that is not residually finite. By the previous lemma and Lemma 3.12, $I_0^*(A)$ belongs to the variety generated by $I_0^*(S_\alpha)$ (which is, in turn, contained in the variety generated by S), while $I_0^*(A)$ is not residually finite.

Hence it remains to deal with the case $\alpha^* = \alpha$, that is, when S has an involution subsemigroup S_α that is an involutory null extension, but not an inflation, of the rectangular band S_α^2 . In the process of resolving this situation, we will determine the subvariety lattice of \mathcal{B}_1^* .

For $\mathbf{w} \in (X \cup X^*)^+$, let $h_1(\mathbf{w})$ and $t_1(\mathbf{w})$ denote the first and the last letter in \mathbf{w} , respectively. Furthermore, when $|\mathbf{w}| \geq 2$, let $h_2(\mathbf{w})$ and $t_2(\mathbf{w})$ denote the second and the second-last letter of \mathbf{w} , respectively. Since \mathcal{B}_1^* satisfies $abxcd \approx abcd$, the identity

$$\mathbf{w} \approx h_1(\mathbf{w})h_2(\mathbf{w})t_2(\mathbf{w})t_1(\mathbf{w})$$

holds in this variety for each involutory word \mathbf{w} . The following observation makes some further considerations easier.

LEMMA 5.3. *Within \mathcal{B}_1^* , every involution semigroup identity $\mathbf{u} \approx \mathbf{v}$ such that $|\mathbf{u}|, |\mathbf{v}| \geq 2$ is equivalent to the conjunction of two the identities*

$$h_1(\mathbf{u})h_2(\mathbf{u})h_1(\mathbf{u}) \approx h_1(\mathbf{v})h_2(\mathbf{v})h_1(\mathbf{v})$$

and

$$t_1(\mathbf{u})t_2(\mathbf{u})t_1(\mathbf{u}) \approx t_1(\mathbf{v})t_2(\mathbf{v})t_1(\mathbf{v}).$$

PROOF. As already noted, $\mathbf{u} \approx \mathbf{v}$ is equivalent in \mathcal{B}_1^* to

$$h_1(\mathbf{u})h_2(\mathbf{u})t_2(\mathbf{u})t_1(\mathbf{u}) \approx h_1(\mathbf{v})h_2(\mathbf{v})t_2(\mathbf{v})t_1(\mathbf{v}). \tag{5.1}$$

However, due to the involution axioms, by substituting for each variable its star and then starring the whole identity, we deduce the reverse identity

$$t_1(\mathbf{u})t_2(\mathbf{u})h_2(\mathbf{u})h_1(\mathbf{u}) \approx t_1(\mathbf{v})t_2(\mathbf{v})h_2(\mathbf{v})h_1(\mathbf{v}). \tag{5.2}$$

By multiplying (5.1) and (5.2) and using $abxcd \approx abcd$, we obtain

$$h_1(\mathbf{u})(h_2(\mathbf{u}))^2h_1(\mathbf{u}) \approx h_1(\mathbf{v})(h_2(\mathbf{v}))^2h_1(\mathbf{v})$$

and

$$t_1(\mathbf{u})(t_2(\mathbf{u}))^2t_1(\mathbf{u}) \approx t_1(\mathbf{v})(t_2(\mathbf{v}))^2t_1(\mathbf{v}).$$

The lemma now follows, since \mathcal{B}_1^* satisfies $ax^2b \approx (ax)b(xb) = a(xb)^2 \approx axb$. \square

We are now in a position to characterize those involution semigroup identities that, when they hold, are strong enough to impose the structure of an involutory inflation of a rectangular band on members of \mathcal{B}_1^* .

LEMMA 5.4. *Let \mathcal{X} be a subvariety of \mathcal{B}_1^* . Then $\mathcal{X} \subseteq (\mathcal{O} \vee \mathcal{RB})^*$ if and only if \mathcal{X} satisfies an identity $\mathbf{u} \approx \mathbf{v}$ such that either $1 \in \{|\mathbf{u}|, |\mathbf{v}|\}$ or, alternatively, $|\mathbf{u}|, |\mathbf{v}| \geq 2$ and one of the following two cases holds:*

- (i) $h_1(\mathbf{u}) \neq h_1(\mathbf{v})$;
- (ii) $h_1(\mathbf{u}) = h_1(\mathbf{v})$ and $h_2(\mathbf{u}) \neq h_2(\mathbf{v})$, with at least one of $h_2(\mathbf{u}), h_2(\mathbf{v})$ not belonging to $\{h_1(\mathbf{u}), h_1(\mathbf{u})^*\}$.

PROOF. The forward implication is immediate, as $(\mathcal{O} \vee \mathcal{RB})^*$ satisfies $xyz \approx xz$, which complies with (ii). So assume that \mathcal{X} satisfies an identity $\mathbf{u} \approx \mathbf{v}$ of one of the forms specified, and let $S \in \mathcal{X}$.

If, for example, $|\mathbf{u}| = 1$, then the identity $\mathbf{u} \approx \mathbf{v}$ implies that $S = S^2$ is a rectangular band. If $|\mathbf{u}|, |\mathbf{v}| \geq 2$ and $h_1(\mathbf{u}) \neq h_1(\mathbf{v})$, then, bearing in mind the previous lemma, the rectangular involution band S^2 (which also satisfies $\mathbf{u} \approx \mathbf{v}$) must be trivial, so S is a null semigroup. Therefore we may assume that $\mathbf{u} \approx \mathbf{v}$ is of type (ii). Without loss of generality, assume that $h_2(\mathbf{u}) \notin \{h_1(\mathbf{u}), h_1(\mathbf{u})^*\}$. Recall that by the lemma above, S satisfies

$$h_1(\mathbf{u})h_2(\mathbf{u})h_1(\mathbf{u}) \approx h_1(\mathbf{v})h_2(\mathbf{v})h_1(\mathbf{v}).$$

In view of the assumptions made, the latter identity is equivalent to one of the following:

- (1) $xyx \approx xzx$;
- (2) $xyx \approx xy^*x$;
- (3) $xyx \approx xx^*x$;
- (4) $xyx \approx x^3$.

It is quite easy to verify that (3) implies (1) and (1) implies (4) within \mathcal{B}_1^* . Also, (2) yields $xyz^*x \approx xzy^*x$, which in turn implies (1) by Lemma 5.3. However, (4) is equivalent in \mathcal{B}_1^* to $xyx \approx x^2$, since $x^3 \approx x^2$ holds in \mathcal{B}_1^* . By [27, Lemma 3.3], S is then an inflation of S^2 , as S satisfies $(xy)^2 \approx xy$. \square

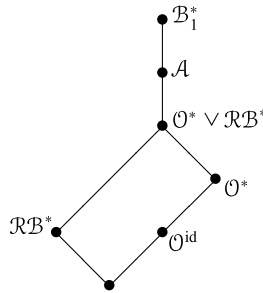


FIGURE 1. Subvarieties of \mathcal{B}_1^* .

PROPOSITION 5.5. *The lattice of all subvarieties of \mathcal{B}_1^* is as given in Figure 1, in which \mathcal{A} denotes the subvariety of \mathcal{B}_1^* determined by the identity $x^2 \approx xx^*x$.*

PROOF. The structure of the subvariety lattice of $(\mathcal{O} \vee \mathcal{RB})^* = \mathcal{O}^* \vee \mathcal{RB}^*$ follows immediately from [5, Theorem 1.7.14]. Now assume that \mathcal{X} is a proper subvariety of \mathcal{B}_1^* such that $\mathcal{X} \not\subseteq (\mathcal{O} \vee \mathcal{RB})^*$. Then \mathcal{X} must not satisfy any of the identities described in the previous lemma. On the other hand, \mathcal{X} satisfies an identity $\mathbf{u} \approx \mathbf{v}$ that does not hold in \mathcal{B}_1^* . Lemmas 5.3 and 5.4 imply that we must have $|\mathbf{u}|, |\mathbf{v}| \geq 2$, $h_1(\mathbf{u}) = h_1(\mathbf{v})$ and $h_2(\mathbf{u}) \neq h_2(\mathbf{v})$ while $h_2(\mathbf{u}), h_2(\mathbf{v}) \in \{h_1(\mathbf{u}), h_1(\mathbf{u})^*\}$ (dually, we also have $t_1(\mathbf{u}) = t_1(\mathbf{v})$ and $t_2(\mathbf{u}) \neq t_2(\mathbf{v})$ while $t_2(\mathbf{u}), t_2(\mathbf{v}) \in \{t_1(\mathbf{u}), t_1(\mathbf{u})^*\}$). By Lemma 5.3 and the fact that \mathcal{B}_1^* satisfies $x^3 \approx x^2$, we conclude that $\mathbf{u} \approx \mathbf{v}$ is equivalent in \mathcal{B}_1^* to $x^2 \approx xx^*x$. Hence $\mathcal{X} = \mathcal{A}$. □

The final piece of our argument is now as follows.

LEMMA 5.6. *The variety \mathcal{A} is not residually finite.*

PROOF. We modify the example of a null extension A that is not residually finite of a left zero band, given in [17, proof of Lemma 1]. This semigroup is defined on the set $\{a, b, c_i, d_i, x : i \in \mathbb{N}\}$ and, setting $C = \{c_i : i \in \mathbb{N}\}$, the subset $L = A \setminus C$ forms a left zero band, $c_i d_i = a$, $c_i d_j = b$ and $c_i t = x$ for all $i, j \in \mathbb{N}$ such that $j \neq i$, and $t \in \{a, b, c_i, x : i \in \mathbb{N}\}$. Now define \widehat{A} , with the multiplication \circ , to be a null extension of the rectangular involution band on $L \times L$ (that is, $(u, v) \circ (w, z) = (u, z)$ and $(u, v)^* = (v, u)$ for all $u, v, w, z \in L$) by the null semigroup on $C \cup \{0\}$ such that $c_i \circ c_j = (x, x)$, $c_i \circ (y, z) = (c_i y, z)$ and $(y, z) \circ c_i = (y, c_i z)$ for all $i, j \in \mathbb{N}$ and $y, z \in L$. It remains to specify an involution on the set C , a function $f : C \rightarrow C$ such that $f(f(c_i)) = c_i$ holds for all $i \in \mathbb{N}$; as we shall see, the way in which this is done is immaterial.

We have $\widehat{A} \in \mathcal{A}$ since $\xi \circ \xi = \xi \circ \xi^* \circ \xi$ obviously holds if $\xi \in L \times L$, while if $\xi = c_i$ for some i , then $c_i \circ c_i = (x, x) = c_i \circ c_j \circ c_i$, where $c_j = c_i^*$. On the other hand, let ρ be a congruence of \widehat{A} of finite index. Then there are $m, n, p, q \in \mathbb{N}$ such that $(d_m, d_n)\rho(d_p, d_q)$, where either $m \neq p$ or $n \neq q$. For example, assume that the

latter occurs. Then

$$(d_m, a) = (d_m, c_n d_n) = (d_m, d_n) \circ c_n \rho(d_p, d_q) \circ c_n = (d_p, c_n d_q) = (d_p, b),$$

and since ρ agrees with the involution, $(a, d_m) \rho(b, d_p)$ as well. Hence

$$(a, a) = (a, d_m) \circ (d_m, a) \rho(b, d_p) \circ (d_p, b) = (b, b),$$

that is, all congruences of finite index of \widehat{A} have a nonidentical intersection. In other words, there is no homomorphism of \widehat{A} onto a finite involution semigroup that separates (a, a) and (b, b) , so \widehat{A} is not residually finite. \square

Thus if $\alpha^* = \alpha$ and S_α is not an involutory inflation of the rectangular band S_α^2 , then S_α generates a variety that contains \mathcal{A} , whence the variety generated by S cannot be residually finite. This completes the proofs of Proposition 5.1 and Theorem 1.1.

6. Discussion and open problems

By the main theorem of Golubov and Sapir [17], there are precisely three maximal combinatorial varieties of residually finite semigroups. One is $\mathcal{O} \vee \mathcal{NB}$, the variety consisting of all inflations of normal bands. For the other two, define P to be the semigroup on $\{0, a, e\}$ such that $e^2 = e$ and $ea = a$, all the other products being equal to 0, and let L_2, R_2 be the two-element left zero and right zero bands, respectively. The remaining varieties in question are the variety \mathcal{U} generated by $L_2 \times P$, and its dual variety \mathcal{U}^∂ , generated by $R_2 \times P^\partial$. We see that in the involutory case the latter two simply ‘disappear’. Roughly speaking, this is because these varieties are not self-dual, so that they cannot be generated by a class of semigroup reducts of some involution semigroup variety; on the other hand, the largest self-dual variety contained in both is their intersection, which is equal to $\mathcal{O} \vee \mathcal{SL}$.

In comparison to the exhaustive result of [17], here we imposed the restriction of dealing with *combinatorial* involution semigroups only. First of all, this is because currently we do not know the solution to the following problem.

PROBLEM 6.1. Describe the residually finite varieties of groups with involution. Are they all periodic?

We draw the reader’s attention to the fact that a given group may admit various involutions; more precisely, any involution on a group G is of the form $g \mapsto \psi(g^{-1})$ for some $\psi \in \text{Aut}(G)$ such that $\psi^2 = \iota_G$. Hence an involution group (in the signature $\{\cdot, {}^{-1}, *\}$) is in fact term-equivalent to a group with a distinguished automorphism of order two. This circumstance implies that it is likely that most of the considerations concerning varieties of involution groups lead to elementary but quite tedious number-theoretic complications, such as in [10]. Nevertheless, it is worth asking whether an analogue of Ol’shanskiĭ’s result holds in the presence of an additional unary operation. The removal of nontrivial subgroups probably reveals the semigroup-theoretical aspects of residually finite varieties in a more transparent way. It is very

likely, however, that the following conjecture holds: a variety of involution semigroups is residually finite if and only if it is of the form $\mathcal{V} \vee \mathcal{X}$, where $\mathcal{V} \subseteq (\mathcal{O} \vee \mathcal{N}\mathcal{B})^*$ and \mathcal{X} is a residually finite variety of periodic groups with involution (since, technically speaking, only periodic groups are semigroups).

Significant progress towards this result may be made along the lines of this paper. Namely, it is evident from the proof of Lemma 4.1 that each residually finite involution semigroup variety \mathcal{V} must consist of involution semilattices of null extensions of completely simple semigroups. The ‘missing ingredient’ is the argument that the completely simple semigroups in question are in fact rectangular groups; this conclusion is true in the plain semigroup case, as any completely simple semigroup that is not a rectangular group fails to generate a residually finite variety; see [17]. This explains the motivation for the following task, which is more general than Problem 6.1.

PROBLEM 6.2. Describe residually finite varieties generated either by a completely simple semigroup with involution, or by $I_0^*(A)$ where A is completely simple.

If we overcame this obstacle, we would obtain a result parallel to Lemma 4.3, that all members of \mathcal{V} must be involutory null extensions of orthodox normal bands of groups. Furthermore, we could conclude that each $S \in \mathcal{V}$ is a dense involution semilattice of null extensions of rectangular groups S_α^2 , that is, $S = \bigcup_{\alpha \in Y} S_\alpha$. Our aim would be then to show that these null extensions must be inflations. At this point, it would be useful to invoke [27, Corollary 2.2] which implies that any null extension Q of a rectangular group $T \times G$ (where T is a rectangular band and G a group) is a subdirect product of G and $Q/\mathcal{H} \in \mathcal{B}_1$. Then by our Proposition 5.1, S_α/\mathcal{H} must be an inflation of a rectangular band for all $\alpha \in Y$, since otherwise it would follow that a homomorphic image of an involution subsemigroup of S generates a variety that is not residually finite, whence S_α is an inflation of $(S_\alpha^2/\mathcal{H}) \times G_\alpha$ for each maximal subgroup G_α of S_α .

Here is where we encounter the second difficulty. The considerations just presented suffice to conclude, by using Lemma 4.4, that if the maximal subgroups of S generate a (periodic) group variety \mathcal{G} , then the semigroup reduct of S belongs to $(\mathcal{O} \vee \mathcal{RB} \vee \mathcal{G}) \vee \mathcal{SL} = \mathcal{O} \vee \mathcal{NB} \vee \mathcal{G}$, so that $S \in \mathcal{O}^* \vee (\mathcal{NB} \vee \mathcal{G})^*$. But what can be said about \mathcal{G} ? If $\alpha^* = \alpha$ then there is a maximal subgroup G_α of S_α such that $G_\alpha^* = G_\alpha$, thus each such involution subgroup must generate a residually finite variety, so we depend on the solution of Problem 6.1. However, if $\alpha^* \neq \alpha$, then we should analyze the property of residual finiteness for a variety generated by an involution semigroup of the form $I_0^*(G_\alpha)$, which returns us to Problem 6.2. Is it true that these conditions are satisfied if and only if each subgroup of the semigroup reduct of S generates a residually finite group variety? If *not*, then the localization of S within $\mathcal{O}^* \vee (\mathcal{NB} \vee \mathcal{G})^*$ might turn out to be not precise enough. In such a case, we would need an ‘involutory version’ of Mel’nik’s dense semilattice theorem (that is, [27, Proposition 2.9]) to answer the question: under what circumstances does a dense involution semilattice of semigroups belong to $\mathcal{W} \vee \mathcal{SL}^*$ for an arbitrary involution semigroup variety \mathcal{W} ?

We finish this paper with two remarks.

REMARK 6.3. The subvariety lattice of $\mathcal{O}^* \vee \mathcal{N}\mathcal{B}^*$ can be easily computed from [5, Theorem 1.7.14]: it has exactly 30 elements, comprising the 14 subvarieties of $\mathcal{N}\mathcal{B}^*$ (see [9]), \mathcal{O}^{id} , $\mathcal{O}^{\text{id}} \vee \mathcal{S}\mathcal{L}^{\text{id}}$, and the varieties of the form $\mathcal{O}^* \vee \mathcal{V}$, where $\mathcal{V} \subseteq \mathcal{N}\mathcal{B}^*$.

REMARK 6.4. A variety \mathcal{V} is *residually less than* λ if each subdirectly irreducible member of \mathcal{V} has cardinality less than λ . A variety \mathcal{V} is *residually small* if \mathcal{V} is residually less than λ for some cardinal λ (in other words, the class of isomorphism types of subdirectly irreducible algebras of \mathcal{V} is a set). Residually small varieties of semigroups were investigated by McKenzie [22, 23] and Sapir and Shevrin [32]. In particular, it follows from their results that a combinatorial semigroup variety is residually small if and only if it is residually finite. By a careful inspection of this paper, one can see that the same holds for combinatorial varieties of involution semigroups: it is possible to replace the words ‘residually finite’ by ‘residually small’ or ‘residually less than λ ’ (for any infinite cardinal λ) in the formulation of our Theorem 1.1. This observation relies upon the fact that each of Examples 3.1, 3.4, 3.5, 3.7, 3.9 and 3.10 actually supplies instances of subdirectly irreducible involution semigroups of arbitrarily large cardinality. Also, for each cardinal λ , an analog of Lemma 3.12 will hold for a semigroup S that is residually less than λ , provided that in its proof, F is selected such that $|F| < \lambda$. Finally, the proof of Lemma 5.6 will work verbatim if we let the indices of c_i, d_i range over λ instead of \mathbb{N} , so that the variety \mathcal{A} is residually large (that is, not residually small). These modifications, in conjunction with the other results here, yield a strengthened version of Theorem 1.1.

Acknowledgement

The author is grateful to the anonymous referee for a thorough reading of the paper, resulting in a number of improvements in the presentation.

References

- [1] O. Andersen, ‘Ein Bericht über die Struktur abstrakter Halbgruppen’, Thesis, Hamburg, 1952.
- [2] K. Auinger, ‘Free products of combinatorial strict inverse semigroups’, *Pacific J. Math.* **164** (1994), 201–227.
- [3] S. Burris and H. P. Sankappanavar, *A Course in Universal Algebra*, Graduate Texts in Mathematics (Springer, New York, 1981).
- [4] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups*, Vol. I (American Mathematical Society, Providence, RI, 1961).
- [5] S. Crvenković and I. Dolinka, ‘Varieties of involution semigroups and involution semirings: a survey’, *Bull. Soc. Math. Banja Luka* **9** (2002), 7–47.
- [6] S. Crvenković, I. Dolinka and M. Vinčić, ‘Equational bases for some 0-direct unions of semigroups’, *Studia Sci. Math. Hungar.* **36** (2000), 423–431.
- [7] L. Descalço and N. Ruškuc, ‘Properties of the subsemigroups of the bicyclic monoid’, *Czechoslovak Math. J.* **58**(133) (2008), 311–330.
- [8] I. Dolinka, ‘Remarks on varieties of involution bands’, *Comm. Algebra* **28** (2000), 2837–2852.
- [9] I. Dolinka, ‘All varieties of normal bands with involution’, *Period. Math. Hungar.* **40** (2000), 109–122.

- [10] I. Dolinka, 'On the lattice of varieties of involution semigroups', *Semigroup Forum* **62** (2001), 438–459.
- [11] I. Dolinka, 'Subdirectly irreducible bands with involution', *Acta Sci. Math. (Szeged)* **67** (2001), 535–554.
- [12] I. Dolinka, 'Varieties of involution semilattices of Archimedean semigroups', *Publ. Math. Debrecen* **66** (2005), 439–447.
- [13] I. Dolinka, 'On identities of finite involution semigroups', *Semigroup Forum* **80** (2010), 105–120.
- [14] S. Fajtlowicz, 'Equationally complete semigroups with involution', *Algebra Universalis* **1** (1972), 355–358.
- [15] J. A. Gerhard, 'Subdirectly irreducible idempotent semigroups', *Pacific J. Math.* **39** (1971), 669–676.
- [16] J. A. Gerhard, 'Some subdirectly irreducible idempotent semigroups', *Semigroup Forum* **5** (1973), 362–369.
- [17] E. A. Golubov and M. V. Sapir, 'Varieties of finitely approximable semigroups', *Izv. Vysš. Učebn. Zaved. Mat.* **11**(246) (1982), 21–29 (in Russian); English transl. *Soviet Math. (Iz. VUZ)* **26** (1982), 25–36.
- [18] J. M. Howie, *Fundamentals of Semigroup Theory* (Oxford University Press, New York, 1995).
- [19] E. I. Kleiman, 'On basis of identities of Brandt semigroups', *Semigroup Forum* **13** (1977), 209–218.
- [20] S. I. Kublanovskii, 'Finite approximability of prevarieties of semigroups with respect to predicates', in: *Modern Algebra* (Gos. Ped. Inst., Leningrad, 1980), pp. 58–88 (in Russian).
- [21] G. Lallement, 'On monoids presented by a single relation', *J. Algebra* **32** (1974), 370–388.
- [22] R. McKenzie, 'Residually small varieties of semigroups', *Algebra Universalis* **13** (1981), 171–201.
- [23] R. McKenzie, 'A note on residually small varieties of semigroups', *Algebra Universalis* **17** (1983), 142–149.
- [24] I. I. Mel'nik, 'Varieties and lattices of varieties of semigroups', in: *Studies in Algebra*, Vol. 2 (ed. V. V. Vagner) (Izdat. Saratov. Univ., Saratov, 1970), pp. 47–57 (in Russian).
- [25] A. Yu. Ol'shanskiĭ, 'Varieties of finitely approximable groups', *Izv. Akad. Nauk SSSR, Ser. Mat.* **33** (1969), 915–927 (in Russian); English transl. *Math. USSR-Izv.* **3** (1969), 867–877.
- [26] M. Petrich, *Introduction to Semigroups* (Merrill, Columbus, OH, 1973).
- [27] M. Petrich, 'All subvarieties of a certain variety of semigroups', *Semigroup Forum* **7** (1974), 104–152.
- [28] M. Petrich, *Inverse Semigroups* (Wiley, New York, 1984).
- [29] M. S. Putcha, 'Semilattice decompositions of semigroups', *Semigroup Forum* **6** (1973), 12–34.
- [30] N. R. Reilly, 'Free combinatorial strict inverse semigroups', *J. Lond. Math. Soc. (2)* **39** (1989), 102–120.
- [31] M. V. Sapir, 'Problems of Burnside type and the finite basis property in varieties of semigroups', *Izv. Akad. Nauk SSSR, Ser. Mat.* **51** (1987), 319–340 (in Russian); English transl. *Math. USSR-Izv.* **30** (1987), 295–314.
- [32] M. V. Sapir and L. N. Shevrin, 'Residually small varieties of semigroups and groups', *Izv. Vysš. Učebn. Zav. Mat.* **10**(317) (1988), 41–49 (in Russian); English transl. *Soviet Math. (Iz. VUZ)* **32** (1988), 57–73.
- [33] M. V. Sapir and E. V. Sukhanov, 'Varieties of periodic semigroups', *Izv. Vysš. Učebn. Zav. Mat.* **4**(227) (1981), 48–55 (in Russian); English transl. *Soviet Math. (Iz. VUZ)* **25** (1981), 53–63.
- [34] H. E. Scheiblich, 'Projective and injective bands with involution', *J. Algebra* **109** (1987), 281–291.

IGOR DOLINKA, Department of Mathematics and Informatics,
 University of Novi Sad, Trg Dositeja Obradovića 4,
 21000 Novi Sad, Serbia
 e-mail: dockie@dmf.uns.ac.rs