MCCAMMOND'S NORMAL FORMS FOR FREE APERIODIC SEMIGROUPS REVISITED†

J. ALMEIDA, J. C. COSTA, AND M. ZEITOUN

Abstract. This paper revisits the normal forms introduced by J. McCammond to solve the word problem for ω-words over the pseudovariety A of aperiodic semigroups. The proof of the uniqueness of these normal forms for ω-words, given by McCammond, is based on his solution of the word problem for certain Burnside semigroups. In this paper, we describe a new proof of correctness of McCammond's algorithm which is based on properties of certain languages associated with the normal forms. This method leads to several new applications beyond ω-words.

Contents

1. Introduction

An ω-term is a formal expression obtained from letters of an alphabet X using the operations of concatenation and ω-power. These expressions naturally define implicit operations on finite semigroups S: the concatenation is viewed as the semigroup multiplication while the ω-power is interpreted as the unary operation which sends each element of S to its unique idempotent power. An implicit operation defined by an ω-term in the elements of a pseudovariety of semigroups V is called an ω-implicit operation over V. The ω-semigroup formed by these implicit operations, denoted by $\Omega^\omega_X\mathsf{V}$, is the V-free ω-semigroup generated by X.

The ω-word problem for a pseudovariety V is the problem of determining whether two ω-terms represent the same operation on the elements of V, that is, is the word problem for $\Omega^\omega_X\mathsf{V}$. This problem has received some attention lately. The case of the pseudovariety J of all $\mathcal{J}$-trivial semigroups, solved by the first author in [1], constitutes a classical example. Another remarkable example,

Date: May 20, 2009.
† TEMPORARY TITLE!!

achieved by McCammond [23] and which plays an important role in the Krohn-Rhodes complexity problem, is given by the pseudovariety A of all aperiodic semigroups. A much simpler case, obtained by the second author [16], is the one associated with the pseudovariety LSl of local semilattices. More recently, the first and third authors [14] solved the word problem for $\Omega_X^\omega\mathsf{R}$, where R is the pseudovariety of $\mathcal{R}$-trivial semigroups.

The ω-word problem for a pseudovariety V is an instance of a more general problem: the word problem for σ-semigroups $\Omega_X^\sigma\mathsf{V}$, where σ is an arbitrary implicit signature (a set of implicit operations containing multiplication). This problem is mainly motivated by recent research [3, 12, 13, 4, 24] that has shown that detailed knowledge of the σ-semigroups $\Omega_X^\sigma\mathsf{V}$ yields important information about V. Indeed, through the seminal work of Ash [15] and its generalizations found by the first author and Steinberg [3, 13, 12], an important property that a pseudovariety V may enjoy has emerged: the so-called *tameness*. Roughly speaking, it serves as a strong form of decidability, entailing that it is decidable whether a finite system of equations with rational constraints over a finite alphabet X admits a solution in every X-generated semigroup from V. By a standard compactness argument, the existence of such solutions can be reduced to the existence of solutions in $\overline{\Omega}_X\mathsf{V}$, the pro-$\mathsf{V}$ semigroup freely generated by X (i.e., the semigroup of all implicit operations over V on the alphabet X). Basically, the idea of tameness is to reduce the search of solutions in such a semigroup, which is often uncountable, to a countable subsemigroup, namely a subalgebra $\Omega_X^\sigma\mathsf{V}$ generated by the same set X with respect to a suitable signature σ, which turns out to be itself a relatively free algebra. The signature in question should be made up of "natural" operations on profinite semigroups, including the multiplication. Among such operations, the most encountered is the pseudoinversion, or $(\omega-1)$-power. In the group case, this operation is the group inversion and the corresponding countable subsemigroup of the free profinite group is just the free group on the same generating set. In the aperiodic case, the $(\omega-1)$-power reduces to the ω-power, in which case an algebra in our signature is simply called an *ω-semigroup*.

To prove tameness, one must establish that the word problem for the chosen relatively free algebra $\Omega_X^\sigma\mathsf{V}$ is decidable and, since the existence of solutions of certain systems in the relatively free profinite semigroup must be shown to entail the existence of solutions in the chosen subalgebra, it is important to understand well how the subalgebra fits in the profinite semigroup. For the case of the pseudovariety of all finite groups, a restricted form of tameness was first proved within the framework of semigroup theory [15], but also independently a special case was proved using methods from profinite group theory [25, 26], and later was rediscovered as a model-theoretic result [19, 10, 11].

Unlike the cases of LSl [18, 17] and R [8, 9], tameness of A has not yet been published, but the above mentioned solution of the word problem for $\Omega_X^\omega\mathsf{A}$, which has been obtained by McCammond in 2001, is a step forward in that direction. McCammond's solution [23] consists in the reduction of arbitrary ω-terms to a certain canonical form. McCammond then goes on to show that different ω-terms in canonical form cannot represent the same implicit operation over A, which he does by invoking his results on free Burnside semigroups [22].

In this paper, we describe an alternative proof of the uniqueness of McCammond's normal forms for ω-terms over A which is independent of the theory of free Burnside semigroups. Our approach consists in associating to each ω-term α and positive integer n a certain rational language $L_n[\alpha]$, whose key property is that, if α is in McCammond normal form and n is large enough then $L_n[\alpha]$ is a star-free language. Another crucial step in our proof is the fact that if α and β are both ω-terms in normal form and n is sufficiently large such that $L_n[\alpha] \cap L_n[\beta] \neq \emptyset$, then $\alpha = \beta$. This new approach, and most particularly the star-freeness of the languages $L_n[\alpha]$, permits also to obtain new important applications. We show for instance that, over A, every factor of an ω-implicit operation is also an ω-implicit operation. In turn this result is a central piece in [6], whose main result provides a characterization of the implicit operations which are given by ω-terms thus contributing to a deeper understanding of how the ω-subsemigroup generated by a finite alphabet X fits in the free pro-A semigroup on X.

The paper is organized as follows. After a section of preliminaries, where we review most of the basic background material including the description of McCammond's normal form, we introduce term expansions and the languages $L_n[\alpha]$ in Section 3 and prove some of its basic properties. Section 4 contains some combinatorial lemmas concerning factors of terms of rank 1. In Section 5, we present the main properties of the languages $L_n[\alpha]$ and the alternative proof of uniqueness of McCammond's normal forms for ω-terms over A. Section 6 is devoted to establish the star-freeness of the languages $L_n[\alpha]$ for α in normal form and n large enough. Finally, we investigate in Section 7 other properties of the languages $L_n[\alpha]$ and derive some applications.

2. Preliminaries

In this section we briefly recall the basic definitions and results that will be used throughout the paper. The reader is referred to [2, 24] for general background, and to [5] for a quick introduction, about the classical theories of pseudovarieties, rational languages and profinite semigroups. For further details about combinatorics on words see [21].

In the following, X always denotes a finite non-empty set called an *alphabet*. The *free semigroup* generated by X, viewed as the set of all finite non-empty words in the letters of X, is denoted by X^+. The *free monoid* on X, denoted by X^*, is obtained from X^+ by adding the empty word. The length of a word $w \in X^*$ is denoted by $|w|$. The following result is known as Fine and Wilf's Theorem (see [21]).

Proposition 2.1. *Let $u, v \in X^+$. If two powers u^k and v^ℓ of u and v have a common prefix of length at least $|u| + |v| - \gcd(|u|, |v|)$, then u and v are powers of the same word.*

A *primitive word* is a word that cannot be written in the form u^n with $n > 1$. Two words w and z are said to be *conjugate* if there exist words $u, v \in X^*$ such that $w = uv$ and $z = vu$. We notice that, if w is a primitive word and z is a conjugate of w, then z is also primitive. Let an order be fixed for the letters of the alphabet X. A *Lyndon word* is a primitive word that is minimal, with respect to the lexicographic ordering, in its conjugacy class. The following proposition (see [21]) gives an alternative definition of Lyndon words.

Proposition 2.2. *A word $w \in X^+$ is a Lyndon word if and only if w is strictly less than any of its proper non-empty suffixes.*

A word $v \in X^*$ is said to be a *border* of another word $w \in X^+$ when $v \neq w$ and v is both a prefix and a suffix of w. If a word $w \in X^+$ admits a non-empty border then w is called *bordered*, otherwise it is called *unbordered*. An immediate consequence of Proposition 2.2 is that every Lyndon word is an unbordered word.

Let V be a pseudovariety of semigroups. A *pro*-V semigroup is a projective limit of elements of V. Equivalently, we may say that a topological semigroup is *pro*-V if it is compact and residually in V. In particular, all semigroups of V, endowed with the discrete topology, are pro-V. A pro-S semigroup, where S is the pseudovariety of all finite semigroups, is also called *profinite*. The projective limit of all X-generated semigroups from V is denoted by $\overline{\Omega}_X\mathsf{V}$. For each mapping $\varphi : X \to S$ into a pro-V semigroup S there is a unique continuous homomorphism $\overline{\varphi} : \overline{\Omega}_X\mathsf{V} \to S$ extending φ, that is $\overline{\Omega}_X\mathsf{V}$ is the pro-V semigroup freely generated by X. The elements of $\overline{\Omega}_X\mathsf{V}$ are called *pseudowords* over V and they are naturally interpreted as *(X-ary implicit)* operations on pro-V semigroups S: the interpretation of $w \in \overline{\Omega}_X\mathsf{V}$ on S is $w_S : S^X \to S$ mapping each function $\varphi \in S^X$ to $\overline{\varphi}(w)$. For instance, for $X = \{x, y\}$, the pseudoword $w = xy$ is interpreted as the semigroup multiplication from $S \times S$ into S. If $X = \{x\}$, the interpretation of the ω-power x^ω is the mapping which associates each element $s \in S$ to s^ω, the unique idempotent of the closed subsemigroup generated by s.

Given $u, v \in \overline{\Omega}_X\mathsf{S}$, we call the formal equality $u = v$ a *pseudoidentity*. For a profinite semigroup S, we then say that S *satisfies* the pseudoidentity $u = v$, and write $S \models u = v$, when $u_S = v_S$. A pseudovariety V *satisfies* a pseudoidentity $u = v$, denoted $\mathsf{V} \models u = v$, if every semigroup from V satisfies $u = v$. Given $u, v \in \overline{\Omega}_X\mathsf{S}$, it is well known that $\mathsf{V} \models u = v$ if and only if $p_\mathsf{V}(u) = p_\mathsf{V}(v)$, where $p_\mathsf{V} : \overline{\Omega}_X\mathsf{S} \to \overline{\Omega}_X\mathsf{V}$ is the only continuous homomorphism which sends each free generator to itself.

An *implicit signature* is a set of pseudowords containing the multiplication xy. In this paper we will be interested in the signature $\{xy, x^\omega\}$, which will also be denoted ω. Every pro-V semigroup S has a natural structure of an ω-semigroup, via the interpretation of the elements of ω as operations on S. We denote by $\Omega^\omega_X\mathsf{V}$ the ω-subsemigroup of $\overline{\Omega}_X\mathsf{V}$ generated by X, whose elements are called *ω-implicit operations* over V. Each ω-implicit operation has a representation by a formal term over X in the signature ω. These terms are called *ω-terms* and they are obtained from the letters of X using multiplication and ω-power. More conveniently, since we are interested in semigroups, we will consider equivalence classes of ω-terms resulting by collecting together ω-terms that can be obtained from each other by applying the associative law of multiplication. The resulting equivalence classes of ω-terms on a set X are called *ω-words*. They constitute the unary semigroup U_X freely generated by X.

We now recall the definition of McCammond's normal form for ω-words. To simplify the notation, McCammond [23] represents ω-words over an alphabet X as correctly parenthesized words in the alphabet $Y = X \cup \{(,)\}$, for which the parentheses are thus viewed as letters. The ω-word associated with such a word is obtained by replacing each matching pair of parentheses $(*)$ by $(*)^\omega$. Conversely, every ω-word over X determines a unique correctly parenthesized word over Y. We define the *length* of an ω-word α to be the length of the word over Y which it determines, and we denote it $|\alpha|$. From hereon, in the absence of mention to the contrary, we will refer to an ω-word meaning its associated word over Y. Note that the words in the extended alphabet that represent ω-words are precisely those for which, by removing all other letters we obtain a Dyck word. In fact, it is easy to check that the ω-subsemigroup of the free semigroup Y^+ generated by X, where the ω-power is interpreted as the operation $w \mapsto (w)$, is freely generated by X as a unary semigroup. Thus, we identify U_X with the set of well-parenthesized words over the alphabet X. In particular, there is a natural homomorphism of ω-semigroups $\epsilon : U_X \to \Omega_X^\omega \mathsf{A}$ that fixes each $x \in X$ when we view X as a subset of U_X and $\Omega_X^\omega \mathsf{A}$ in the natural way. To avoid ambiguities in the meaning of the parentheses, we will write $\epsilon[w]$ for the image of $w \in U_X$ under ϵ. The elements of $\Omega_X^\omega \mathsf{A}$ will sometimes be called *ω-words*.

The ω-word problem for A (over X), consists in deciding when two elements of U_X have the same image under ϵ. This problem was solved by McCammond by showing that it is possible to transform any ω-word into a certain *normal form* with the same image under ϵ, and by proving that two ω-words in normal form with the same image under ϵ are necessarily equal. For the description of the normal form, fix a total ordering of the alphabet X, and extend it to $Y = X \cup \{(,)\}$ by letting $(< x <)$ for all $x \in X$. The *rank* of a word of Y^+ is the maximum number of nested parentheses in it.

McCammond's normal form is defined recursively. The *rank 0 normal forms* are the words from X^+. Assuming that rank i normal form terms have been defined, a *rank $i+1$ normal form term* is a word from Y^+ of the form

$$\alpha_0(\beta_1)\alpha_1(\beta_2)\cdots\alpha_{n-1}(\beta_n)\alpha_n,$$

where the α_j and β_k are ω-words such that

(a) each β_k is a Lyndon word of rank i;

(b) no intermediate α_j is a prefix of a power of β_j or a suffix of a power of β_{j+1};

(c) replacing each subterm (β_k) by $\beta_k\beta_k$, we obtain a rank i normal form ω-word;

(d) at least one of the properties (b) and (c) fails if we remove from α_j a prefix β_j (for $0 < j$) or a suffix β_{j+1} (for $j < n$).

McCammond's procedure of transformation of an arbitrary ω-word into one in normal form consists in applying elementary changes determined by the following *rewriting rules*:

1. $((\alpha)) = (\alpha)$
2. $(\alpha^k) = (\alpha)$
3. $(\alpha)(\alpha) = (\alpha)$
4. $(\alpha)\alpha = (\alpha)$, $\alpha(\alpha) = (\alpha)$
5. $(\alpha\beta)\alpha = \alpha(\beta\alpha)$

If an ω-subword given by the left side of a rule of type 1–4 is replaced in an ω-word by the right side of the rule, then we say there is a *contraction* of that type. If the replacement is done in the opposite direction than we say that there is an *expansion* of that type. For the rules of type 4, we may add an index L or R to indicate on which side of the ω-power the base was added or deleted.

Since all the rules are based on identities of ω-semigroups that are valid in A (in fact, all but those of type 4 are valid in S), it follows that the elementary changes preserve the value of the ω-word under ϵ. Hence McCammond's algorithm does indeed transform an arbitrary ω-word into one in normal form with the same image under ϵ. We don't describe here McCammond's procedure because usually we will work with ω-words already in normal form (or almost). The reader interested in more details of the algorithm is referred to the original paper [23] or to [6] for a more condensed discription of its steps.

To prove that distinct ω-words in normal form have different images in $\Omega^{\omega}_{X}\mathsf{A}$, McCammond used his solution of the word problem for certain free Burnside semigroups [22]. We have obtained a direct combinatorial proof of the same result, which will be described in Sections 3 to 6, which leads to other applications presented in Section 7.

3. Word expansions

The main differences between the following definition and McCammond's "rank i expansions" [23, Definition 10.5] are that we require the exponents to be beyond a fixed threshold and we do not require that the ω-words be in normal form.

Definition 3.1 (Word expansions)**.** Let n be a positive integer. For a word $\alpha \in X^*$, we let $E_n[\alpha] = \{\alpha\}$. For an ω-word

(3.1) $\qquad \alpha = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$ where all the γ_j and δ_k are ω-words, the δ_k have the same rank i and all the γ_j have rank at most i,

we let

$$E_n[\alpha] = \{\gamma_0\delta_1^{n_1}\gamma_1 \cdots \delta_r^{n_r}\gamma_r : \ n_1, \dots, n_r \geq n\}.$$

For a set W of ω-words, we let $E_n[W] = \bigcup_{\alpha\in W} E_n[\alpha]$. For an ω-word α we let

$$L_n[\alpha] = E_n^{\operatorname{rank}\alpha}[\alpha],$$

and, for a set W of ω-words , we let $L_n[W] = \bigcup_{\alpha\in W} L_n[\alpha]$.

We recall now terminology introduced by McCammond for an ω-word α as in (3.1) that will be used later. The ω-subwords of α of the form $(\delta_j)\gamma_j(\delta_{j+1})$ are called *crucial portions* of α, whereas the prefix $\gamma_0(\delta_1)$ and the suffix $(\delta_r)\gamma_r$ are called respectively its *initial portion* and its *final portion.*

Lemma 3.2. *The following formulas hold:*

(*a*) *for ω-words α and β,*

$$E_n[\alpha\beta] = \begin{cases} E_n[\alpha]\, E_n[\beta] & \textit{if } \operatorname{rank}\alpha = \operatorname{rank}\beta \\ \alpha\, E_n[\beta] & \textit{if } \operatorname{rank}\alpha < \operatorname{rank}\beta \\ E_n[\alpha]\, \beta & \textit{if } \operatorname{rank}\alpha > \operatorname{rank}\beta. \end{cases}$$

(*b*) *for an ω-word α, $L_n[\alpha] = L_n[E_n[\alpha]]$;*

(*c*) *for sets U and V of ω-words of bounded rank, $L_n[UV] = L_n[U]\, L_n[V]$.*

(*d*) *if $\alpha = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$ is a factorization of an ω-word as in* (3.1)*, then*

$$L_n[\alpha] = L_n[\gamma_0]\, L_n[(\delta_1)]\, L_n[\gamma_1] \cdots L_n[(\delta_r)]\, L_n[\gamma_r];$$

(*e*) *for an ω-word α, $L_n[(\alpha)] = L_n[\alpha]^n L_n[\alpha]^*$.*

Proof. (a) is immediate from the definition of the operator E_n. For (b), since $E_n[\alpha]$ is a set of ω-words whose rank is $\operatorname{rank}\alpha - 1$, we have

$$L_n[E_n[\alpha]] = E_n^{\operatorname{rank}\alpha-1}[E_n[\alpha]] = E_n^{\operatorname{rank}\alpha}[\alpha] = L_n[\alpha].$$

For (c), we proceed by induction on the maximum rank m of the elements of $U \cup V$. In case $m = 0$, both sides of the equation are equal to UV. Assuming $m \geq 1$ and the result holds for sets involving terms of rank less than m, we first note that, if u and v are ω-words of rank at most m, then

$$L_n[uv] = L_n[E_n[uv]] = \left\{ \begin{array}{ll} L_n[E_n[u]E_n[v]] & \text{if } \operatorname{rank} u = \operatorname{rank} v \\ L_n[uE_n[v]] & \text{if } \operatorname{rank} u < \operatorname{rank} v \\ L_n[E_n[u]v] & \text{if } \operatorname{rank} u > \operatorname{rank} v \end{array} \right\} = L_n[u]\, L_n[v]$$

where the last equality follows from (b) and the induction hypothesis. To conclude the induction step, we note that

$$L_n[UV] = \bigcup_{u\in U,\, v\in V} L_n[uv] = \bigcup_{u\in U,\, v\in V} L_n[u]L_n[v] = L_n[U]L_n[V].$$

Property (d) follows from (c) by induction on the number of factors.

For (e), we have

$$L_n[(\alpha)] \underset{(b)}{=} L_n[E_n[(\alpha)]] = \bigcup_{m\geq n} L_n[\alpha^m] \underset{(c)}{=} \bigcup_{m\geq n} L_n[\alpha]^m = L_n[\alpha]^n L_n[\alpha]^*. \quad \square$$

In case α is a rank $i+1$ ω-word in normal form, the elements of $E_1[\alpha]$ are precisely the McCammond's "rank i expansions of α". Since Lemma 10.7 of [23] states that every such rank i expansion of α is an ω-word in normal form and since $E_1[\alpha] \supseteq E_2[\alpha] \supseteq E_3[\alpha] \supseteq \cdots$, we obtain the following result which will be very useful in the sequel.

Lemma 3.3. *Let n be a positive integer. If α is an ω-word in normal form and $w \in E_n^*[\alpha]$, then w is also an ω-word in normal form.* $\square$

We now associate to each term a parameter which plays an important role in this paper.

Definition 3.4 (Parameter $\mu[\alpha]$)**.** Let $\alpha = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$ be an ω-word as in (3.1). Let $\mu[\alpha]$ denote the integer

$$\mu[\alpha] = 4 + 2\max\{|\delta_j\gamma_j\delta_{j+1}|, |\delta_r\gamma_r|, |\gamma_0\delta_1| : j = 1, \ldots, r-1\}.$$

In case α is a word, we let $\mu[\alpha] = |\alpha|$.

It is easy to check that, if the above expression for α is its normal form, then $\mu[\alpha] \geq \max\{\mu[\gamma_j], \mu[\delta_j]\}$. It is also important to point out the following simple observation.

Lemma 3.5. *If α is an ω-word with $\operatorname{rank}\alpha > 1$ and $\alpha' \in E_n[\alpha]$, then $\mu[\alpha'] \leq \mu[\alpha]$.* $\square$

4. Some combinatorial lemmas

Given words u and v, write $u \preceq v$ if u is a prefix of v and $u \prec v$ if u is a proper prefix of v. The following is a simple consequence of a well-known combinatorial result on words [20, Proposition 1.3.4].

Lemma 4.1. *Suppose that u, v, w are words such that u is nonempty and $uv = vw$. Then v is a prefix of a power of u.* ☐

We will use repeatedly the following consequence of Fine and Wilf's Theorem on the relationship between the periods of a sufficiently long word and their synchronization.

Lemma 4.2. *Let u and v be Lyndon words and suppose that w is a word such that $|w| \geq |u|+|v|$ and w is a factor of both a power of u and a power of v. Then $u = v$. Moreover, for all factorizations $u^m = xwy$ and $v^n = zwt$, there is a factorization $w = w_1w_2$ such that $xw_1, zw_1 \in u^*$.*

Proof. By Fine and Wilf's Theorem (Proposition 2.1), since both u and v are primitive words, and they are periods of w, with $|w| \geq |u| + |v| - \gcd\{|u|, |v|\}$, we conclude that they are conjugates of each other. But, since they are also Lyndon words, they must be equal.

Since xwy and zwt are powers of u, there are factorizations $w = w_1w_2 = w_3w_4$, $x = u^kx'$ and $z = u^\ell z'$ such that $x'w_1 = z'w_3 = u$, and w_2, w_4 are prefixes of some power of u. Furthermore, since $|w| \geq 2|u|$ and $|w_1| \leq |u|$, we have $|w_2| \geq |u|$, and therefore $u \preceq w_2$. Similarly $u \preceq w_4$. Therefore, $w_1u \preceq w_1w_2 = w$ and $w_3u \preceq w_3w_4 = w$ so that w_1u and w_3u are $\preceq$-comparable.

If $w_1 \neq w_3$, then we may as well assume that $|w_1| < |w_3|$ so that $w_1u \prec w_3u$ and $w_1 \prec w_3$. Therefore, there exist t_1, t_2, t_3 such that $w_3 = w_1t_1$, $u = t_1t_2 = t_2t_3$. Now, since Lyndon words are unbordered and t_2 is a border of the Lyndon word u, it follows that $t_2 = 1$. Hence $u = t_1$, so $w_3 = w_1u$. Since $|w_3| \leq |u|$, this implies that $w_1 = 1$ and $w_3 = u$. Therefore, $x' = u$ and $z' = 1$, so $xw_1, zw_1 \in u^*$.

If $w_1 = w_3$, then $x'w_1 = z'w_1 = u$ and again $xw_1, zw_1 \in u^*$, which completes the proof. ☐

The next two lemmas establish important properties of factors of certain ω-words of rank 1.

Lemma 4.3. *Let $\alpha = u_0(v_1)u_1 \cdots (v_r)u_r$ be an ω-word of rank 1, where each u_i and each v_j is a word of X^*, such that each crucial portion $(v_i)u_i(v_{i+1})$ is in normal form and let z be a Lyndon word. Let $w = u_0v_1^{n_1}u_1 \cdots v_r^{n_r}u_r \in E_n[\alpha]$ with $n \geq \max\{\mu[\alpha], |z|+1\}$ and suppose that $w = sz^nt$ with*

$$|u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}| \leq |s| < |u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^{n_i}|.$$

Let x be the word given by $u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^{n_i} = sx$. Then $v_i = z$ and exactly one of the following cases holds:

(a) *$|z^n| \leq |x|$ and there is a factorization $u_{i-1} = pv_i^k$ such that $s = u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}p$;*
(b) *$|z^n| \leq |x|$ and there is a factorization $s = u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^k$ such that $k \leq n_i - n$;*
(c) *$|v_iz| \leq |x| < |z^n|$ and $s = u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^k$ for some k such that $0 \leq k \leq n_i - n + \lfloor\frac{|u_i|}{|v_i|}\rfloor$.*

Proof. Suppose first that $|z^n| \leq |x|$. Since

$$|z^n| \geq |z^{|u_{i-1}v_i|+1}| = (|u_{i-1}v_i| + 1)|z| \geq |u_{i-1}| + |v_i| + |z| = |u_{i-1}| + |v_iz|,$$

and z^n is a prefix of x, which in turn is a suffix of $u_{i-1}v_i^{n_i}$, we deduce that there is a common factor of z^n and $v_i^{n_i}$ of length $|v_i| + |z|$. Lemma 4.2 then implies that $v_i = z$ and the conditions in (a) or (b) hold, depending on whether or not the inequality $|s| < |u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}|$ is verified.

Consider next the case $|v_iz| \leq |x| < |z^n|$. Since

$$|v_i^n| \geq |v_i^{|z|+1}| = (|z| + 1)|v_i| \geq |v_iz|,$$

x is a prefix of z^n and x and $v_i^{n_i}$ are suffixes of the same word, we deduce that the suffix of $v_i^{n_i}$ of length $|v_iz|$ is a factor of z^n. By Lemma 4.2, we conclude that $v_i = z$ and there is an exponent ℓ such that $0 \leq \ell \leq n$ and $sv_i^\ell = u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^{n_i}$. Let $k = n_i - \ell \geq 0$. Thus, to show that (c) holds, it remains to prove that $k \leq n_i - n + \lfloor\frac{|u_i|}{|v_i|}\rfloor$. Note that u_i is not a prefix of a power of v_i because the crucial portion $v_i^\omega u_i v_{i+1}^\omega$ is in normal form. Since u_i and $z^{n-\ell}$ are prefixes of the same word, it follows that the largest possible exponent of a power of z which is a prefix of u_i is $\lfloor\frac{|u_i|}{|v_i|}\rfloor$. Hence $n - \ell \leq \lfloor\frac{|u_i|}{|v_i|}\rfloor$ which, together with $\ell = n_i - k$, implies the desired inequality.

It remains to consider the case where $1 \leq |x| < |v_iz|$. Since

$$|z^n| \geq |z^{|v_iu_i|+1}| = (|v_iu_i| + 1)|z| \geq |u_i| + |v_i| + |z| = |u_i| + |v_iz| > |xu_i|,$$

we conclude that $i < r$. Hence

$$\begin{aligned} |z^n| - |xu_i| &\geq |z^{|v_iu_iv_{i+1}|+2}| - |xu_i| \\ &= (|v_iu_iv_{i+1}| + 2)|z| - |xu_i| \\ &\geq |v_iu_iv_{i+1}| + 2|z| - |xu_i| \\ &= |v_iv_{i+1}| + 2|z| - |x| \\ &> |v_{i+1}z|. \end{aligned}$$

By Lemma 4.2, we have $z = v_{i+1}$ and $sz^k = u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^{n_i}u_i$ for some k such that $1 \leq k \leq n$. Since $|s| < |u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^{n_i}|$, it follows that u_i is a suffix of $z^k = v_{i+1}^k$, contradicting the hypothesis that $v_i^\omega u_i v_{i+1}^\omega$ is a crucial portion in normal form. This shows that at least one of the cases (a), (b) and (c) holds. □

Lemma 4.4. *Let $\alpha = u_0(v_1)u_1 \cdots (v_r)u_r$ be an ω-word of rank 1, where each u_i and each v_j is a word of X^*, such that each crucial portion $(v_i)u_i(v_{i+1})$ is in normal form and let $\beta = (z_1)y(z_2)$ be a crucial portion of rank 1, also in normal form. Let $w = u_0v_1^{n_1}u_1 \cdots v_r^{n_r}u_r \in E_n[\alpha]$ with $n \geq \max\{\mu[\alpha], \mu[\beta]\}$ and suppose that $w = sz_1^nyz_2^nt$ with*

$$|u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}| \leq |s| < |u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^{n_i}|.$$

Then $z_1 = v_i$, $y = u_i$, $z_2 = v_{i+1}$, and $u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_iv_i^{n_i} = sz_1^n$.

Proof. The equality $z_1 = v_i$ follows directly from Lemma 4.3. As in that lemma, let x be the word given by $u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}u_{i-1}v_i^{n_i} = sx$. Note that the last equality in the statement of the lemma is equivalent to $x = z_1^n$. We now consider the three cases of Lemma 4.3.

(a) The inequality $|z_1^n| \leq |x|$ holds and there is a factorization $u_{i-1} = pv_i^k$ such that $s = u_0v_1^{n_1}u_1 \cdots v_{i-1}^{n_{i-1}}p$. Since y is not a prefix of a power of $z_1 = v_i$, because β is in normal form, we must have

$$sv_i^n \prec u_0 v_1^{n_1} u_1 \cdots v_{i-1}^{n_{i-1}} u_{i-1} v_i^{n_i} = sx \prec sv_i^n y$$

i.e., $v_i^n \prec x \prec v_i^n y$. Since $n \geq \mu[\beta] > |y|$, the factor y after the prefix sv_i^n must end somewhere

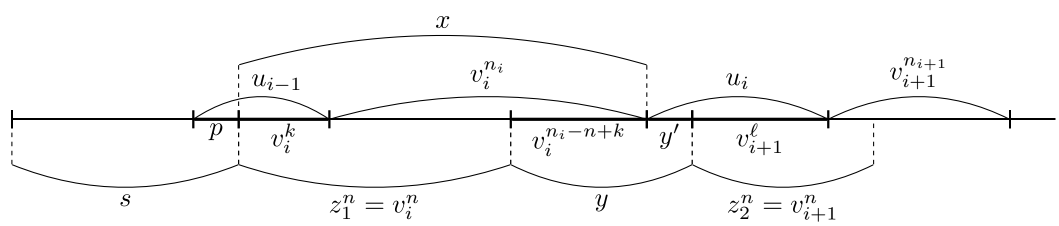


FIGURE 1. Case (a) of Lemma 4.4

within the factor $u_i v_{i+1}^{n_{i+1}}$ or, more precisely,

$$|u_0 v_1^{n_1} u_1 \cdots v_{i-1}^{n_{i-1}} u_{i-1} v_i^{n_i}| = |sx| < |sv_i^n y| < |u_0 v_1^{n_1} u_1 \cdots v_{i-1}^{n_{i-1}} u_{i-1} v_i^{n_i} u_i v_{i+1}^{n_{i+1}}|. \tag{4.1}$$

Hence, by Lemma 4.3, $z_2 = v_{i+1}$. Moreover, we have $u_0 v_1^{n_1} u_1 \cdots v_{i-1}^{n_{i-1}} u_{i-1} v_i^{n_i-k} = sv_i^n$, $y = v_i^{n_i-n+k} y'$ and $u_i = y' v_{i+1}^\ell$. Hence y' is not a prefix of a power of v_i (since otherwise y would also have that property) nor is it a suffix of a power of v_{i+1} (for that would entail that u_i would have the latter property). If u_i is not synchronized with y, that is $n_i - n + k > 0$, or both $n_i - n + k = 0$ and $u_i \neq y$, then neither $(v_i)u_i(v_{i+1})$ nor $(v_i)y(v_{i+1}) = (z_1)y(z_2)$ would be crucial portions in normal form, their normal form being $(v_i)y'(v_{i+1})$. Hence $u_i = y$ and they are synchronized in the two factorizations of w, which proves that $x = z_1^n$.

In case (b), the inequality $|z_1^n| \leq |x|$ holds and there is a factorization $s = u_0 v_1^{n_1} u_1 \cdots v_{i-1}^{n_{i-1}} u_{i-1} v_i^k$ such that $k \leq n_i - n$. In this case, $xu_i v_{i+1}^{n_{i+1}}$ and $v_i^n y z_2^n$ are prefixes of the same word and $x = v_i^{n_i-k}$ is a power of v_i. If $v_i^n y$ were a prefix of x, then y would be a prefix of a power of $z_1 = v_i$,

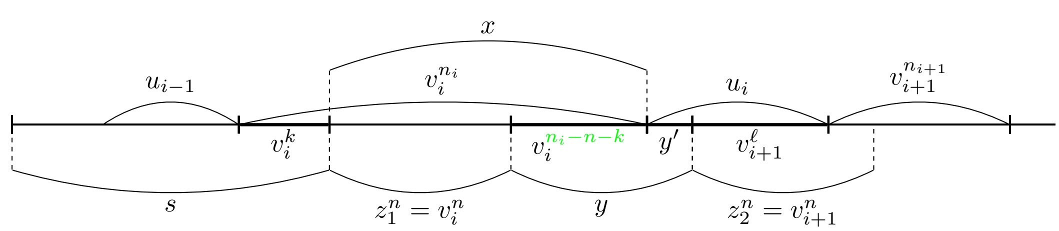


FIGURE 2. Case (b) of Lemma 4.4

which is impossible since $(v_i)u_i(v_{i+1})$ is in normal form. Hence $x \prec v_i^n y$. On the other hand, $|y| < \mu[\beta] \leq n \leq n_{i+1} \leq |u_i v_{i+1}^{n_{i+1}}|$. This shows that the inequalities (4.1) hold. From hereon, in the present case, the argument can be concluded as in the previous case (it suffices to substitute $+k$ by $-k$).

In case (c) the inequalities $|v_i z_1| \leq |x| < |z_1^n|$ hold and $s = u_0 v_1^{n_1} u_1 \cdots v_{i-1}^{n_{i-1}} u_{i-1} v_i^k$ for some k such that $0 \leq k \leq n_i - n + \lfloor \frac{|u_i|}{|v_i|} \rfloor$. As in the previous case, $xu_i v_{i+1}^{n_{i+1}}$ and $v_i^n y z_2^n$ are prefixes of the same word and x is a power of v_i. Since $(v_i)u_i(v_{i+1})$ is in normal form, u_i cannot be a

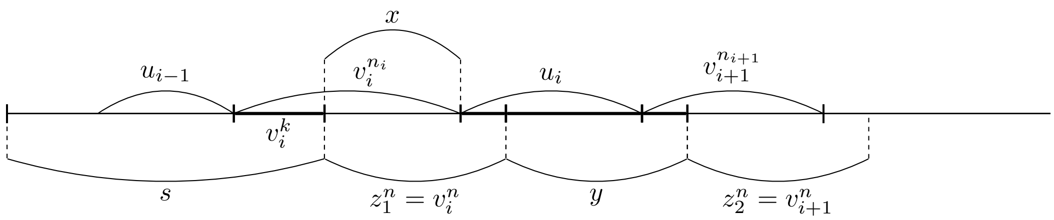


FIGURE 3. Case (c) of Lemma 4.4

prefix of a power of v_i, which implies that $v_i^n \prec xu_i$. Hence the factor y that follows the prefix sv_i^n of w overlaps with the factor u_i in the other given factorization of w. As in the other cases, since $|y| < n \leq |u_i v_{i+1}^{n_{i+1}}|$, we deduce that the inequalities (4.1) hold. A simple adaptation of the argument of case (a) then yields that u_i and y are synchronized in the two factorizations of w, so that $y = u_i$, $z_2 = v_{i+1}$, and $x = v_i^n$, which means that this case cannot actually occur. □

5. THE ω-WORD PROBLEM OVER A

In this section we reveal how the languages $L_n[\alpha]$ can be used to obtain an alternative proof of McCammond's solution of the word problem for ω-words over A.

The fundamental property of the languages $L_n[\alpha]$, whose proof will be left to the next section, is the following.

Theorem 5.1. *Let α be a term in normal form and let $n \geq \mu[\alpha]$. Then the language $L_n[\alpha]$ is star-free.*

A simpler but also important property is that if two ω-words in normal form have a common expansion then one of them is an expansion of the other.

Lemma 5.2. *Let α and β be two ω-words in normal form with $\operatorname{rank}\alpha \leq \operatorname{rank}\beta$, and let $n > \max\{\mu[\alpha], \mu[\beta]\}$. If $L_n[\alpha] \cap L_n[\beta] \neq \emptyset$, then $\alpha \in E_n^*[\beta]$.*

Proof. Let $w \in L_n[\alpha] \cap L_n[\beta]$. We proceed by induction on $\operatorname{rank}\alpha = i$. Suppose first that $i = 0$, that is to say $\alpha \in A^+$, so that $w = \alpha$. Hence $\alpha \in L_n[\beta] = E_n^{\operatorname{rank}\beta}[\beta]$. Assume next that $i \geq 1$ and that the result holds for pairs of ω-words in normal form whose minimum rank is less than i.

By definition of L_n and the choice of w, there exist $\alpha_1 \in E_n^{\operatorname{rank}\alpha-1}[\alpha]$ and $\beta_1 \in E_n^{\operatorname{rank}\beta-1}[\beta]$ such that $w \in L_n[\alpha_1] \cap L_n[\beta_1]$. By Lemma 3.3, α_1 and β_1 are rank 1 ω-words in normal form.

Let $u_0(v_1)u_1 \cdots (v_r)u_r$ and $z_0(t_1)z_1 \cdots (t_s)z_s$ be the normal form expressions of α_1 and β_1, respectively. Then there exist exponents $n_i, m_j \geq n$ such that

$$w = u_0 v_1^{n_1} u_1 \cdots v_r^{n_r} u_r = z_0 t_1^{m_1} z_1 \cdots t_s^{m_s} z_s.$$

Taking into account Lemma 3.5, by Lemma 4.4 it follows that $r = s$, $v_i = t_i$ and $n_i = m_i$ $(i = 1, \ldots, r)$, and $u_i = z_i$ $(i = 0, \ldots, r)$. Hence α_1 and β_1 are the same term. This shows that

$$E_n^{\operatorname{rank}\alpha-1}[\alpha] \cap E_n^{\operatorname{rank}\beta-1}[\beta] \neq \emptyset. \tag{5.1}$$

If $i = 1$, then $\alpha_1 = \alpha$ so that $\alpha \in E_n^{\operatorname{rank}\beta-1}[\beta]$. We now assume that $i > 1$ and, to apply induction, we distinguish the inner, rank 1, parentheses as new letters, namely with the extended ordering $[\![< x <]\!]$ $(x \in X)$ so that, under this interpretation, α and β can be viewed as ω-words $\bar\alpha$ and $\bar\beta$, respectively, over the enlarged alphabet $X \cup \{[\![,]\!]\}$. Note that $\bar\alpha$ and $\bar\beta$ are ω-words in normal form, $1 \leq \operatorname{rank}\bar\alpha = \operatorname{rank}\alpha - 1 \leq \operatorname{rank}\beta - 1 = \operatorname{rank}\bar\beta$, $\mu[\bar\alpha] = \mu[\alpha]$, and $\mu[\bar\beta] = \mu[\beta]$, while $L_n[\bar\alpha] \cap L_n[\bar\beta] \neq \emptyset$ by (5.1). By the induction hypothesis, we deduce that $\bar\alpha \in E_n^*[\bar\beta]$ and, therefore, that $\alpha \in E_n^*[\beta]$, which completes the induction step and the proof of the lemma. □

By raising the lower bound for the integer n, we can present a more precise result.

Theorem 5.3. *Let α and β be two ω-words in normal form and let $n > \max\{|\alpha|, |\beta|, \mu[\alpha], \mu[\beta]\}$. If $L_n[\alpha] \cap L_n[\beta] \neq \emptyset$, then $\alpha = \beta$.*

Proof. Suppose that $\operatorname{rank}\alpha \leq \operatorname{rank}\beta$, so that, by Lemma 5.2, $\alpha \in E_n^*[\beta]$. If $\operatorname{rank}\beta > \operatorname{rank}\alpha$, it follows that $|\alpha| \geq n$, which contradicts the assumption on n. Hence we must have $\operatorname{rank}\beta = \operatorname{rank}\alpha$ and so $\alpha = \beta$. □

For a subset L of $\overline{\Omega}_X\mathsf{S}$, denote its topological closure in $\overline{\Omega}_X\mathsf{S}$ by $\operatorname{cl}(L)$. If $L \subseteq X^+$, then the closure of L in $\overline{\Omega}_X\mathsf{A}$ is $p_\mathsf{A}\big(\operatorname{cl}(L)\big)$ which will be denoted by $\operatorname{cl}_\mathsf{A}(L)$. Combining Theorems 5.1 and 5.3, we obtain a new proof of uniqueness of McCammond's normal form for elements of $\Omega_X^\omega\mathsf{A}$.

Corollary 5.4 (McCammond's solution of the ω-word problem over A [23])**.** *If α and β are ω-words in normal form which define the same implicit operation over A, then $\alpha = \beta$.*

Proof. Let n be any integer greater than $\max\{|\alpha|, |\beta|, \mu[\alpha], \mu[\beta]\}$. By Theorem 5.3, it suffices to show that $L_n[\alpha] \cap L_n[\beta] \neq \emptyset$. Suppose to the contrary that the intersection is empty. Since $L_n[\alpha]$ and $L_n[\beta]$ are star-free languages by Theorem 5.1, their closures in $\overline{\Omega}_X\mathsf{A}$, which are respectively $\operatorname{cl}_\mathsf{A}(L_n[\alpha])$ and $\operatorname{cl}_\mathsf{A}(L_n[\beta])$, are clopen subsets and therefore they are also disjoint. Since the latter sets contain respectively $\epsilon[\alpha]$ and $\epsilon[\beta]$, it follows that $\epsilon[\alpha] \neq \epsilon[\beta]$, that is α and β define different implicit operations over A, and this proves the corollary. □

Given an ω-implicit operation $w \in \Omega_X^\omega\mathsf{A}$, the unique ω-word in normal form that represents it will be called the *normal form representation* or simply the *normal form* of w. For instance, if $a, b \in X$ are such that $a < b$, then

- $(a)ab(b)$ is the normal form of $(a)(b)$ and of $a(a^5)a(a^2)b^8(b)$;
- $b(ab)abaa(a)aaab(aab)$ is the normal form of $(ba)(a)ba(aba)ab$;
- $((a)ab(b)ba)(a)ab(b)$ is the normal form of $((a)(b))$.

6. Star-freeness of the languages $L_n[\alpha]$

This section is dedicated to the proof that the languages $L_n[\alpha]$ are star-free when α is an ω-word in normal form and $n > \mu[\alpha]$.

We say that an ω-word α is in *circular normal form* if either:

- α is a word;
- α is an ω-word of positive rank of the form $\alpha = (\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$, where each δ_i and each γ_i is an ω-word of smaller rank, such that each ω-word $(\delta_k)\gamma_k(\delta_{k+1})$ with $k \in \{1, \ldots, r\}$ is in normal form, where we let $\gamma_{r+1} = \gamma_1$.

A consequence of Lemma 3.3, is that the property of being in circular normal form is preserved by expansions.

Lemma 6.1. *Let α be an ω-word in circular normal form and let $w \in E_n[\alpha]$. Then w is also in circular normal form.*

Proof. If α is a word, then $w = \alpha$ is certainly in circular normal form. Otherwise α and w are of the form $\alpha = (\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$ and $w = \delta_1^{n_1}\gamma_1 \cdots \delta_r^{n_r}\gamma_r$ with each $n_k \geq n$. Note that, according to Lemma 3.2 (a), $w^2 \in L_n[\alpha]^2 = L_n[\alpha^2]$. By Lemma 3.3, applied to the crucial portions $(\delta_k)\gamma_k(\delta_{k+1})$ of α^2, we conclude that each factor $\delta_k^{n_k}\gamma_k\delta_{k+1}^{n_{k+1}}$ is in normal form. Now, each crucial portion of w^2 is a crucial portion of some factor of the form $\delta_k^{n_k}\gamma_k\delta_{k+1}^{n_{k+1}}$ and, therefore it is in normal form. Hence w is in circular normal form. $\square$

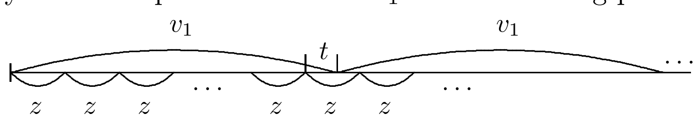
We now prove that if an expansion of an ω-word of rank 1 α in circular normal form has an ℓth root z, then α itself admits an ℓth root $\bar{\alpha}$ and z is an expansion of $\bar{\alpha}$.

Lemma 6.2. *Let α be an ω-word of rank 1 in circular normal form and let $n \geq \mu[\alpha]$. If $z^\ell \in L_n[\alpha]$ then there exists an ω-word of rank 1 in circular normal form $\bar{\alpha}$ such that $\alpha = \bar{\alpha}^\ell$ and $z \in L_n[\bar{\alpha}]$.*

Proof. Let $\alpha = (v_1)u_1 \cdots (v_r)u_r$ be a circular normal form expression for α and let $w = z^\ell$. Then there is a factorization of the form $w = v_1^{n_1}u_1 \cdots v_r^{n_r}u_r$, with $n_1, \ldots, n_r \geq n$. By hypothesis, each u_i is neither a prefix of a power of v_i nor a suffix of a power of v_{i+1}.

If $\ell = 1$ then we can just take $\bar{\alpha} = \alpha$ and there is nothing to prove, and so we assume that $\ell > 1$. We start by eliminating the cases in which the word z is a prefix of a power of v_1.

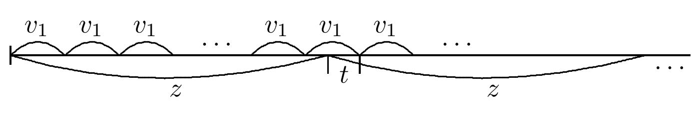
Case 1: $z \prec v_1$. If $v_1 = z^k$ is a power of z, then $k \geq 2$ and we set $t = z$. Otherwise, let t be the prefix of z defined by the overlap between z and v_1 in the following picture:



Since $z \prec v_1$, it follows that t is both a proper suffix and a proper prefix of v_1, which is in contradiction with the assumption on α, which implies that each v_i is a Lyndon word.

Case 2: $z = v_1^k$. Since w admits a factorization of the form $w = v_1^{n_1}u_1w'$, it follows that $u_1w' = v_1^{k\ell - n_1}$ and, therefore, u_1 is a prefix of a power of v_1, which is in contradiction with the hypothesis on α.

Case 3: $v_1^{k-1} \prec z \prec v_1^k$ for some $k \geq 2$. Let t be the prefix of z defined by the overlap between z and v_1 in the following picture:



Note that, since v_1 is a prefix of z, t is both a proper prefix and a proper suffix of v_1. Since v_1 is a Lyndon word, it cannot admit such a border. Hence Case 3 does not occur.

The above shows that $v_1^{n_1} \prec z$ and that z is not a prefix of any power of v_1. In particular, $|z| > n \geq \mu[\alpha] > |u_r|$. It follows that, for every $i \in \{1, \ldots, \ell - 1\}$ there exists $m_i \in \{2, \ldots, r\}$ such that

$$|v_1^{n_1}u_1\cdots v_{m_i-1}^{n_{m_i-1}}| \le |z^i| < |v_1^{n_1}u_1\cdots v_{m_i-1}^{n_{m_i-1}}u_{m_i-1}v_{m_i}^{n_{m_i}}|. \tag{6.1}$$

Taking into account that $v_1^n \prec z$, and applying Lemma 4.3 with $s = z^i$, we conclude that $v_1 = v_{m_i}$. Moreover, by Lemma 4.4, we deduce that $v_1 = v_{m_i}$, $u_1 = u_{m_i}$ and $v_2 = v_{m_i+1}$. Inductively, it follows that $v_j = v_{m_i+j-1}$, $u_j = u_{m_i+j-1}$ and $v_{j+1} = v_{m_i+j}$ for all $j \le r - m_i$. By considering $w^2 \in L_n[\alpha^2]$, where α^2 is in circular normal form, we may also apply Lemma 4.4 to establish that $u_{m_i-1} = u_r$. Hence the word $\bar\alpha = (v_1)u_1\cdots(v_{m_1-1})u_{m_1-1}$ is a period of the word α. Moreover, $z \in L_n(\bar\alpha)$ by Lemma 4.4, which completes the proof. □

We will call *primitive* an ω-word which is primitive when represented as a parenthesized word. An immediate consequence of Lemma 6.2 is the following observation.

Corollary 6.3. *Let α be an ω-word of rank 1 in circular normal form and let $n \ge \mu[\alpha]$. If α is primitive and $w \in L_n[\alpha]$ then w is also primitive.* □

The next result may be regarded as a sort of analog of Corollary 6.3 for ω-words of larger rank.

Corollary 6.4. *Let α be an ω-word of rank $i \ge 1$ in circular normal form and let $n \ge \mu[\alpha]$. If α is primitive and $w \in E_n[\alpha]$, then w is also primitive.*

Proof. We distinguish two types of parentheses in the ω-word α: write $($, $)$ for the parentheses corresponding to the outmost ω-powers of rank i and $[\![,]\!]$ for the remaining parentheses. Consider the alphabet $Y = X \cup \{[\![,]\!]\}$, with the extended ordering $[\![\le x \le]\!]$ $(x \in X)$. Then w may be viewed as a word w_Y over Y and α as an ω-word of rank 1 α_Y over the same alphabet such that $w_Y \in L_n[\alpha]$. Moreover $\mu[w_Y] = \mu[w]$ and it is clear by McCammond's definition of rank i normal form that α_Y is a primitive ω-word in circular normal form (over Y), whence α_Y and w_Y satisfy the hypotheses of Corollary 6.3. To conclude the proof, it suffices to invoke Corollary 6.3. □

Iterating the application of Corollary 6.4, we obtain the following extension of Corollary 6.3 to ω-words of any rank.

Proposition 6.5. *Let α be an ω-word of rank $i \ge 1$ in circular normal form and let $n \ge \mu[\alpha]$. If α is a primitive ω-word and $w \in L_n[\alpha]$ then w is a primitive word.*

Proof. Let $\alpha = (v_1)u_1\cdots(v_r)u_r$ be a circular normal form expression for α. We proceed by induction on i. The case $i = 1$ is given by Corollary 6.3. Assume next that $i > 1$ and that the result holds for ω-words of rank $i-1$. By definition of $L_n[\alpha]$, there is an ω-word α' of the form $\alpha' = v_1^{n_1}u_1\cdots v_r^{n_r}u_r$, with $n_1, \ldots, n_r \ge n$, such that $w \in L_n[\alpha']$. By Corollary 6.4, α' is a primitive ω-word. Moreover, note that $\mu[\alpha'] \le \mu[\alpha]$, by Lemma 3.5, and that, by Lemma 3.3, α' is in circular normal form. Hence, by induction hypothesis, w is primitive, which completes the induction step. □

The following result generalizes Lemma 6.2 in case α is a primitive ω-word.

Lemma 6.6. *Let α be a primitive ω-word of rank $i \ge 0$ in circular normal form and let $n \ge \mu[\alpha]$. If $z^\ell \in L_n[\alpha]^k$ then $z \in L_n[\alpha]^m$ for some m such that $1 \le m \le k$.*

Proof. Note that we may as well assume that z is a primitive word. Recall that the equality $L_n[\alpha]^k = L_n[\alpha^k]$ holds by Lemma 3.2(c). To prove the lemma, we proceed by induction on $i = \operatorname{rank}\alpha$.

In case $i = 0$, α is a primitive word and $L_n[\alpha] = \{\alpha\}$. Since z is also assumed to be primitive, $z^\ell = \alpha^k$ is a power of the two primitive words z and α. By Fine and Wilf's Theorem it follows that $z = \alpha$ and so we must take $m = 1$. So we may assume that $i \geq 1$ and that the result holds for ω-words of rank less than i.

Since $z^\ell \in L_n[\alpha^k] = E_n[E_n^{\operatorname{rank}\alpha - 1}[\alpha^k]]$, there is some ω-word $\alpha' \in E_n^{\operatorname{rank}\alpha-1}[\alpha^k]$ of rank 1 such that $z^\ell \in E_n[\alpha'] = L_n[\alpha']$. By Lemma 6.1, α' is in circular normal form. By Lemma 6.2, there exists an ω-word β of rank 1 such that $z \in L_n[\beta]$ and $\alpha' = \beta^\ell$. In particular, we have $\beta^\ell \in E_n^{\operatorname{rank}\alpha-1}[\alpha^k]$.

Consider first the case where $i = 1$. Then $\beta^\ell \in E_n^{\operatorname{rank}\alpha-1}[\alpha^k]$ means simply that $\beta^\ell = \alpha^k$. Since α is primitive, it follows that $\beta = \alpha^m$, for some m. Hence $z \in L_n[\beta] = L_n[\alpha]^m$, thus completing the proof in case $i = 1$. From hereon, we assume that $i > 1$.

For a word $w \in U_X$, let $\hat{w}$ be the word which is obtained by replacing all the subexpressions of the form (z), where $z \in X^+$, by $[\![z]\!]$. Let $Y = X \cup \{[\![,]\!]\}$ and note that the correspondence $w \mapsto \hat{w}$ is an isomorphism from U_X onto U_Y which sends circular normal forms to circular normal forms. Hence we have $\hat{\beta}^\ell \in L_n[\hat{\alpha}^k]$. Since $\operatorname{rank}\hat{\alpha} = \operatorname{rank}\alpha - 1 \geq 1$, note that the definition of μ implies that $n \geq \mu[\alpha] = \mu[\hat{\alpha}]$. By the induction hypothesis, it follows that there exists m such that $1 \leq m \leq k$ and $\hat{\beta} \in L_n[\hat{\alpha}^m]$. Since $w \mapsto \hat{w}$ is invertible, we deduce that $\beta \in E_n^{\operatorname{rank}\alpha-1}[\alpha^m]$. Hence, we have $z \in L_n[\beta] = E_n[\alpha] \subseteq E_n^{\operatorname{rank}\alpha}[\alpha^m] = L_n[\alpha^m]$. $\square$

JA: this is writt rather different st the usual encodin perhaps it would b to isolate the prop the encoding in a lemma.

We say that an ω-word α is in *quasi-normal form* if one of the following conditions holds:

- α is a word;
- α is of rank 1, there is an ω-word factorization of the form $\alpha = \gamma_0\delta_1^k(\delta_1)\gamma_1(\delta_2)\cdots(\delta_{r-1})\gamma_{r-1}(\delta_r)\delta_r^\ell\gamma_r$ such that $\gamma_0(\delta_1)\gamma_1(\delta_2)\cdots(\delta_{r-1})\gamma_{r-1}(\delta_r)\gamma_r$ is in normal form;
- $\operatorname{rank}\alpha > 1$ and there is an ω-word factorization $\alpha = \gamma_0(\delta_1)\gamma_1\cdots(\delta_r)\gamma_r$ such that $\gamma_0\delta_1^2\gamma_1\cdots\delta_r^2\gamma_r$ is in quasi-normal form.

Note that, if $\gamma_0(\delta_1)\gamma_1\cdots(\delta_r)\gamma_r$ is in normal form then each of the ω-words γ_i and δ_j is in quasi-normal form. This follows easily from condition (c) on the definition of normal form. Also, if α is in circular normal form, then α is in quasi-normal form.

The next lemma is a key result to our objectives. It presents a sort of factorization scheme, for an ω-word α in quasi-normal form, induced by factorizations of elements of $L_n[\alpha]$.

Lemma 6.7. *Let α be an ω-word of rank i in quasi-normal form, $n \geq \mu[\alpha]$, and $m = \lfloor \frac{n}{2} \rfloor$. Then there exists a finite set $P_{\alpha,n}$ of pairs of ω-words and a function $\pi_{\alpha,n}$ associating to each pair of words (w_1, w_2), with $w_1w_2 \in L_n[\alpha]$, an element $(\bar{\alpha}_1, \bar{\alpha}_2)$ of $P_{\alpha,n}$ such that the following conditions hold:*

(a) *$w_k \in L_m[\bar{\alpha}_k]$ and $|\bar{\alpha}_k| \leq (i+1)\cdot|\alpha|\cdot(m+1)$ ($k = 1, 2$);*

(b) *each $\bar{\alpha}_k$ is in quasi-normal form;*

(c) *it is possible to apply McCammond's expansions of types 3 and 4 to α to obtain the ω-word $\bar{\alpha}_1\bar{\alpha}_2$;*

(*d*) *if two factorizations* $w_1w_2, w_1'w_2' \in L_n[\alpha]$ *are such that* $\pi_{\alpha,n}(w_1, w_2) = \pi_{\alpha,n}(w_1', w_2')$*, then the crossed products* $w_1w_2', w_1'w_2$ *belong to* $L_n[\alpha]$ *and* $\pi_{\alpha,n}(w_1, w_2) = \pi_{\alpha,n}(w_1', w_2) = \pi_{\alpha,n}(w_1, w_2')$.

Proof. We proceed by induction on i, the case $i = 0$ being obvious. Suppose that $i \geq 1$ and the result holds for ω-words of rank less than i. Write $\alpha = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$, where the crucial portions $(\delta_j)\gamma_j(\delta_{j+1})$ are in normal form. Let w_1, w_2 be words such that $w_1w_2 \in L_n[\alpha]$. By Lemma 3.2, there exist integers $n_1, \dots, n_r \geq n$ and words $v_{j,h} \in L_n[\delta_j]$, $u_j \in L_n[\gamma_j]$ such that $w_1w_2 = u_0v_1'u_1 \cdots v_r'u_r$, where $v_j' = v_{j,1} \cdots v_{j,n_j}$. We distinguish two cases according to where w_1 ends in the second expression.

Suppose that $w_1 = u_0v_1'u_1 \cdots v_{\ell-1}'u_{\ell-1}v_\ell'z_1$, $u_\ell = z_1z_2$, $w_2 = z_2v_{\ell+1}'u_{\ell+1} \cdots v_r'u_r$. Then, by the induction hypothesis, applied to the relation $z_1z_2 \in L_n[\gamma_\ell]$, we conclude that there is a pair $(\bar\gamma_{\ell,1}, \bar\gamma_{\ell,2}) = \pi_{\gamma_\ell,n}(z_1, z_2) \in P_{\gamma_\ell,n}$. It suffices to take

$$\bar\alpha_1 = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_{\ell-1})\gamma_{\ell-1}(\delta_\ell)\bar\gamma_{\ell,1} \qquad \text{and} \qquad \bar\alpha_2 = \bar\gamma_{\ell,2}(\delta_{\ell+1})\gamma_{\ell+1} \cdots (\delta_r)\gamma_r$$

for it is easy to check that $|\bar\alpha_k| \leq i\cdot|\alpha|\cdot(m+1)$ while, by Lemma 3.2(*c*), $w_k \in L_m[\bar\alpha_k]$ for $k = 1, 2$. We put $\pi_{\alpha,n}(w_1, w_2) = (\bar\alpha_1, \bar\alpha_2)$ in $P_{\alpha,n}$ and we observe that condition (*d*) holds for every additional pair (w_1', w_2') of words. Indeed, if such pair verifies the hypotheses of condition (*d*), then $w_1'w_2' \in L_n[\alpha]$ and $w_1' \in L_m[\bar\alpha_1]$. Therefore, we deduce from Lemma 4.4 that w_1' ends in the factor of $w_1'w_2'$ which corresponds to γ_ℓ (as it happens with w_1 in w_1w_2), and the observation follows from the induction hypothesis. Finally, note that conditions (*b*) and (*c*) are immediate consequences of the hypothesis that the corresponding conditions hold for γ_ℓ.

ed this blue para- o explain the ob- n. However, to Lemma 4.4 (is sary?) I had to the value of n, my opinion is not m. (JC)

understand what on, this is OK, argument is rather

Suppose next that there are factorizations $w_1 = u_0v_1'u_1 \cdots v_{\ell-1}'u_{\ell-1}v_{\ell,1} \cdots v_{\ell,p}z_1$, $v_{\ell,p+1} = z_1z_2$, $w_2 = z_2v_{\ell,p+2} \cdots v_{\ell,n_\ell}u_\ell v_{\ell+1}'u_{\ell+1} \cdots v_r'u_r$. By the induction hypothesis applied to the relation $z_1z_2 \in L_n[\delta_\ell]$, we conclude that there is a pair $(\bar\delta_{\ell,1}, \bar\delta_{\ell,2}) = \pi_{\delta_\ell,n}(z_1, z_2) \in P_{\delta_\ell,n}$. We let

$$\bar\alpha_1 = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_{\ell-1})\gamma_{\ell-1}\delta_\ell^{\bar p}\bar\delta_{\ell,1} \qquad \text{and} \qquad \bar\alpha_2 = \bar\delta_{\ell,2}\delta_\ell^{\bar q}\gamma_\ell(\delta_{\ell+1})\gamma_{\ell+1} \cdots (\delta_r)\gamma_r,$$

where $\bar p = p$ if $p < m$ and $\delta_\ell^{\bar p} = (\delta_\ell)$ otherwise, and similarly for $\bar q$ and $\delta_\ell^{\bar q}$, where $q = n_\ell - p - 1$. Note that, since $p + q + 1 \geq n$, we cannot have both $p, q < m$ and so at least one of the ω-words $\delta_\ell^{\bar p}$ and $\delta_\ell^{\bar q}$ is equal to (δ_ℓ). Hence $\bar\alpha_1\bar\alpha_2$ can be obtained from α by applying expansions of types 3 and 4. We can also easily check that $|\bar\alpha_k| \leq (i + 1) \cdot |\alpha| \cdot (m + 1)$ while, by Lemma 3.2(*c*), $w_k \in L_m[\bar\alpha_k]$ for $k = 1, 2$. We put $(\bar\alpha_1, \bar\alpha_2) = \pi_{\alpha,n}(w_1, w_2)$ in $P_{\alpha,n}$ and we note that condition (*d*) holds for every additional pair (w_1', w_2') of words. Finally, note that conditions (*b*) and (*c*) are immediate consequences of the hypothesis that the corresponding conditions hold for δ_ℓ.

To conclude the proof of the lemma it suffices to notice that, since each $P_{\gamma_\ell,n}$ and $P_{\delta_\ell,n}$ is finite by hypothesis and $\bar p$ and $\bar q$ can only take a finite number of values, $P_{\alpha,n}$ is finite. □

In the situation of the last but one paragraph of the proof of the preceding lemma, we say that the factorization w_1w_2 of an element of $L_n[\alpha]$ *induces a cut of* α *at maximal rank*. More generally, we say that $\pi_{v,n}(w_1, w_2)$ is the *cut induced by* the factorization w_1w_2 of an element of $L_n[\alpha]$. Moreover, in the notation of the proof of Lemma 6.7, we say that *the cut takes place at* γ_ℓ or *at* δ_ℓ according to whether the first or the second situations in that proof hold.

With the help of Lemma 6.7, we can establish the following important property of the languages $L_n[\alpha]$ for primitive ω-words α. In its proof, we apply in both directions Schützenberger's Theorem

[27], stating that a language is star-free if and only if its syntactic semigroup is finite and satisfies the pseudoidentity $x^{\omega+1} = x^{\omega}$.

Lemma 6.8. *Let α be a primitive ω-word in circular normal form and let $n \geq \mu[\alpha]$. If $L_n[\alpha]$ is a star-free language, then so is $L_n[\alpha]^*$.*

Proof. Let M be an integer such that the syntactic semigroup of $L_n[\alpha]$ satisfies the identity $x^M = x^{M+1}$ and let K be a positive integer to be identified later. Let $N > MK$ be an integer and suppose that x, y, z are words such that $xy^N z \in L_n[\alpha]^*$. The result follows from the claim that, for sufficiently large K, depending only on α and n, $xy^{N+1}z$ belongs to $L_n[\alpha]^*$.

To prove the claim, we start with a factorization $xy^N z = w_1 \cdots w_m$ where each $w_j \in L_n[\alpha]$. Consider each product of M consecutive y's within the factor y^N. If at least one of the factors appears completely within one of the w_j, then we have a factorization $w_j = x'y^M z'$ as indicated in Figure 4. In particular, the word $x'y^M z'$ belongs to the star-free language $L_n[\alpha]$. By the choice

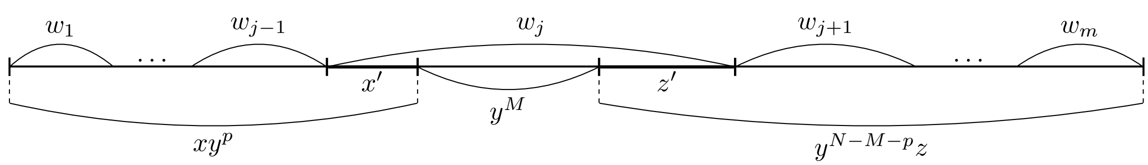


FIGURE 4. Case where some y^M falls within some w_j

of M, we deduce that $w_j' = x'y^{M+1}z' \in L_n[\alpha]$. Hence, for p as in Figure 4,

$$xy^{N+1}z = xy^p \cdot y^{M+1} \cdot y^{N-M-p}z = w_1 \cdots w_{j-1}w_j' w_{j+1} \cdots w_m$$

is again a word from $L_n[\alpha]^m$, independently of the value of $K \geq 1$.

We may therefore assume that no factor y^M appears completely within some factor w_j. Thus, the first $K < N/M$ consecutive factors y^M which form a prefix of y^N as well as the product $y^{N-KM}z$, each starts in a different w_j, say in $w_{j_1}, \ldots, w_{j_{K+1}}$, with $j_1 < \cdots < j_{K+1}$. This determines factorizations

$$w_{j_s} = w_{j_s,1}w_{j_s,2} \tag{6.2}$$

$$y^M = w_{j_s,2}x_s w_{j_{s+1},1} \ (s = 1, \ldots, K) \tag{6.3}$$

$$x = x' w_{j_1,1}$$

$$y^{N-KM}z = w_{j_{K+1},2}z'$$

where each x_s, x', and z' is a word from $L_n[\alpha]^*$, as represented in Figure 5.

Let $P_{\alpha,n}$ and $\pi_{\alpha,n}$ be as in Lemma 6.7 and choose K to be $|P_{\alpha,n}|+1$. By the pigeonhole principle, there exist two indices p, q such that $1 \leq p < q \leq K$ and the pairs $(w_{j_p,1}, w_{j_p,2})$ and $(w_{j_q,1}, w_{j_q,2})$ have the same image under $\pi_{\alpha,n}$. By property (d) of Lemma 6.7, we deduce that $w_{j_q,1}w_{j_p,2}$ belongs to $L_n[\alpha]$. Hence the word

$$w_{j_{p+1},1}y^{M(q-p-2)}w_{j_{q-1},2}x_{q-1}w_{j_q,1} \cdot w_{j_p,2}x_p = w_{j_{p+1},1}y^{M(q-p-1)}w_{j_p,2}x_p = (w_{j_{p+1},1}w_{j_p,2}x_p)^{q-p}$$

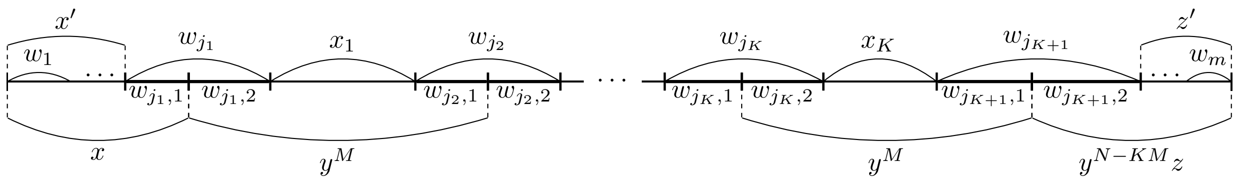


FIGURE 5. Case where each y^M overlaps several w_j

belongs to $L_n[\alpha]^*$ where, for the second equality, we use the factorization (6.3) with $s = p$ for each y^M. Now, $w_{j_{p+1},1}w_{j_p,2}x_p$ is a conjugate of y^M again by (6.3) and, therefore, it is of the form t^M, where t is a conjugate of y. By Lemma 6.6, t belongs to $L_n[\alpha]^*$. On the other hand, note that

$$xy^N z = x'w_{j_1}x_1 \cdots w_{j_{p-1}}x_{p-1}w_{j_p}x_p \cdot w_{j_{p+1}}x_{p+1} \cdots w_{j_K}x_K w_{j_{K+1}}z'$$
$$xy^{N+1}z = x'w_{j_1}x_1 \cdots w_{j_{p-1}}x_{p-1}w_{j_p}x_p \cdot t \cdot w_{j_{p+1}}x_{p+1} \cdots w_{j_K}x_K w_{j_{K+1}}z'$$

where each of the factors separated by the $\cdot$'s belongs to $L_n[\alpha]^*$. Hence $xy^{N+1}z \in L_n[\alpha]^*$. $\square$

We are now ready to complete the proof of aperiodicity of the languages $L_n[\alpha]$ with α in normal form and $n \geq \mu[\alpha]$.

Proof of Theorem 5.1. Let $i = \operatorname{rank}\alpha$. If $i = 0$, then $L_n[\alpha] = \{\alpha\}$ is certainly a star-free language. We will therefore assume that $i \geq 1$. Let $\alpha = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$ be the normal form expression of α.

We claim that each of the languages $L_n[\gamma_0]$, $L_n[\delta_j]$ and $L_n[\delta_j\gamma_j]$ $(j = 1, \ldots, r)$ is star-free. It will then follow, by Lemma 6.8, Lemma 3.2(c), and since the set of star-free languages is closed under concatenation, that each language $L_n[(\delta_j)\gamma_j] = L_n[\delta_j]^* L_n[\delta_j]^{n-1} L_n[\delta_j\gamma_j]$ is also star-free. Taking also into account Lemma 3.2(d), we conclude that the product

$$L_n[\alpha] = L_n[\gamma_0]L_n[(\delta_1)\gamma_1] \cdots L_n[(\delta_r)\gamma_r]$$

is star-free, as stated in the theorem.

To prove the claim, we proceed by induction on $i \geq 1$. The case $i = 1$ is immediate since then all the γ_j and δ_j are words in X^*. Suppose that $i \geq 2$ and assume inductively that the claim holds for ω-words of rank less than i. Consider the ω-word

$$\alpha' = \gamma_0\delta_1\delta_1\gamma_1 \cdots \delta_r\delta_r\gamma_r.$$

By Lemma 3.3, α' is in normal form. By Lemma 3.5, since $\operatorname{rank}\alpha > 1$ and $\alpha' \in E_2[\alpha]$, we have $\mu[\alpha] \geq \mu[\alpha']$. Hence $n \geq \mu[\alpha']$ and we may apply the induction hypothesis to the ω-word α' of rank $i - 1 \geq 1$. Since α is in normal form and the ω-words δ_j are Lyndon words of positive rank, the first letter of each δ_j is the opening parenthesis of an ω-subword of highest (and positive) rank. Hence, if $\alpha' = u_0(v_1)u_1 \cdots (v_s)u_s$ is the normal form expression of α', then each factor $\gamma_0, \delta_j, \delta_j\gamma_j$ $(j = 1, \ldots, r)$ must be a product of some of the factors $u_0, (v_k), (v_k)u_k$ $(k = 1, \ldots, s)$. By the induction hypothesis, each of the languages $L_n[u_0]$, $L_n[v_k]$, and $L_n[v_ku_k]$ $(k = 1, \ldots, s)$ is star-free. By the above argument, it follows that so are the languages $L_n[u_0]$, $L_n[(v_k)]$, and $L_n[(v_k)u_k]$ $(k = 1, \ldots, s)$. Finally, by Lemma 3.2(c), we deduce that each of the languages $L_n[\gamma_0], L_n[\delta_j], L_n[\delta_j\gamma_j]$

$(j = 1, \ldots, r)$ is star-free, thus proving the induction step. This proves the claim and completes the proof of Theorem 5.1. $\square$

We do not know whether the bound $n \geq \mu[\alpha]$ is optimal but we do know that some bound is required, that is that $L_n[\alpha]$ may not be star-free for α in normal form. An example is obtained by taking $\alpha = ((a)ab(b)a^2b^2)$, where a and b are letters. Then $L_1[\alpha] \cap [a^2b^2]^* = [a^2b^2a^2b^2]^+$ so that $L_1[\alpha]$ is not star-free since $[a^2b^2]^*$ is star-free and $[a^2b^2a^2b^2]^+$ is not.

7. Factors of ω-words over A

In this section we present further properties of the languages $L_n[\alpha]$ and derive some applications. The main result of this section is that every factor of an ω-implicit operation in $\overline{\Omega}_X\mathsf{A}$ is also an ω-implicit operation.

For a finite semigroup S let $\mathrm{ind}(S)$ be the smallest $\ell \geq 1$ such that, for some $k \geq 1$ and every $s \in S$, $s^{\ell+k} = s^\ell$. Equivalently, $\mathrm{ind}(S)$ is the minimum positive integer ℓ such that S satisfies the pseudoidentity $x^{\omega+\ell} = x^\ell$. We begin by proving that finite aperiodic semigroups do not separate an ω-word from its expansions of sufficiently large exponent.

Proposition 7.1. *Let $u \in U_X$ be an ω-word and $S \in \mathsf{A}$. If $n \geq \mathrm{ind}(S)$ and $\varphi : \overline{\Omega}_X\mathsf{S} \to S$ is a continuous homomorphism, then $|\varphi(L_n[u])| = 1$. In particular, if $n \geq |S|$ and $w \in L_n[u]$, then $S \models \epsilon[u] = w$.*

Proof. Since $n \geq \mathrm{ind}(S)$, for every $m \geq n$, S satisfies the identity $x^m = x^n$. Hence, for every word $w \in L_n[u]$, $\varphi(w) = \varphi(u')$, where u' is the word which is obtained from u by replacing all occurrences of the ω exponent by n. For the second part of the statement of the proposition, it suffices to observe that $\mathrm{ind}(S) \leq |S|$. $\square$

The following consequence of Proposition 7.1 will be useful.

Corollary 7.2. *If $u \in U_X$ is an arbitrary ω-word, then*

$$p_{\mathsf{A}}\Big(\bigcap_n \mathrm{cl}(L_n[u])\Big) = \{\epsilon[u]\} = \bigcap_n p_{\mathsf{A}}\big(\mathrm{cl}(L_n[u])\big).$$

Proof. Denote by ∂ the unique homomorphism of ω-semigroups $U_X \to \overline{\Omega}_X\mathsf{S}$ extending the identity mapping on X so that $\epsilon = p_{\mathsf{A}} \circ \partial$. First note that, since $\partial[u] \in \mathrm{cl}(L_n[u])$ for every n, certainly

$$\{\epsilon[u]\} \subseteq p_{\mathsf{A}}\Big(\bigcap_n \mathrm{cl}(L_n[u])\Big) \subseteq \bigcap_n p_{\mathsf{A}}\big(\mathrm{cl}(L_n[u])\big).$$

Let $v \in \bigcap_n p_{\mathsf{A}}\big(\mathrm{cl}(L_n[u])\big)$. For a continuous homomorphism $\psi : \overline{\Omega}_X\mathsf{A} \to S$ onto a finite aperiodic semigroup S, let $\varphi = \psi \circ p_{\mathsf{A}}$ and choose any $n \geq \mathrm{ind}(S)$. Then

$$\psi(v) \in \varphi\big(\mathrm{cl}(L_n[u])\big) = \varphi(L_n[u]) = \{\varphi(\partial[u])\}$$

where the first equality follows from the continuity of φ and the finiteness of S, and the second equality is a consequence of Proposition 7.1. Since $\overline{\Omega}_X\mathsf{A}$ is residually in A, it follows that $v = \epsilon[u]$. $\square$

We also have the following somewhat more precise result for ω-words in normal form.

Theorem 7.3. *Let $w \in \Omega_X^\omega \mathsf{A}$ and let α be the normal form representation of w. Then*

$$p_{\mathsf{A}}^{-1}(w) = \bigcap_n \mathrm{cl}(L_n[\alpha]).$$

Proof. The inclusion $\bigcap_n \mathrm{cl}(L_n[\alpha]) \subseteq p_{\mathsf{A}}^{-1}(w)$ follows from Corollary 7.2. For the reverse inclusion, assuming that $v \in \overline{\Omega}_X \mathsf{S}$ is such that $p_{\mathsf{A}}(v) = w$, we have $p_{\mathsf{A}}(v) \in p_{\mathsf{A}}\big(\mathrm{cl}(L_n[\alpha])\big)$ for all n. Let $(v_n)_n$ be a sequence of words converging to v in $\overline{\Omega}_X \mathsf{S}$. Then $\lim v_n = w$ in $\overline{\Omega}_X \mathsf{A}$ and so, since by Theorem 5.1, the set $p_{\mathsf{A}}\big(\mathrm{cl}(L_n[\alpha])\big)$ is open and contains w, by taking a suitable subsequence we may assume that $v_n \in p_{\mathsf{A}}\big(\mathrm{cl}(L_n[\alpha])\big) \cap X^+ = L_n[\alpha]$. Since $(L_n[\alpha])_n$ is a decreasing sequence of languages, it follows that $v \in \mathrm{cl}(L_n[\alpha])$ for all n. □

We now prove the announced main result of this section which does not apparently follow easily from McCammond's results.

Theorem 7.4. *If $v \in \Omega_X^\omega \mathsf{A}$ and $u \in \overline{\Omega}_X \mathsf{A}$ is a factor of v, then $u \in \Omega_X^\omega \mathsf{A}$.*

Proof. By symmetry, it suffices to prove the result when u is a prefix of v, that is, when there exists $w \in \overline{\Omega}_X \mathsf{A}$ such that $uw = v$. Let α be the normal form representation of v. We proceed by induction on $\operatorname{rank}\alpha$. We assume inductively that the result holds for all elements of $\Omega_X^\omega \mathsf{A}$ with rank strictly smaller than $\operatorname{rank}\alpha$.

Since, by Theorem 5.1, $L_n[\alpha]$ is a star-free language for every $n \geq \mu[\alpha]$, its closure $\mathrm{cl}_{\mathsf{A}}(L_n[\alpha])$ is an open subset of $\overline{\Omega}_X \mathsf{A}$. Hence, there exist sequences $(u_m)_m$ and $(w_m)_m$ converging respectively to u and w such that $u_n w_n \in L_n[\alpha]$.

As an ω-word, α admits a unique factorization in U_X of the form $\alpha = x_0 x_1 x_2 \cdots x_{2p-1} x_{2p}$ where each x_{2i} is a finite word and each x_{2i-1} is an ω-word of the form $x_{2i-1} = (y_{2i-1})$. Note that we include here the case where α is a word, for which $p = 0$. Since α is in normal form, each y_{2i-1} is an ω-word of rank less than $\operatorname{rank}\alpha$ (although not necessarily of rank $\alpha - 1$). In view of Lemma 3.2 and each relation $u_n w_n \in L_n[\alpha]$, there is a "cutting" index $c_n \in \{0, \ldots, 2p\}$ and there are factorizations $u_n = u_n' u_n''$ and $w_n = w_n' w_n''$ such that

$$u_n' \in L_n[x_0 \cdots x_{c_n - 1}],\ u_n'' w_n' \in L_n[x_{c_n}],\ w_n'' \in L_n[x_{c_n+1} \cdots x_{2p}].$$

Since the number of possible cutting indices depends only on α and not on n, there is a strictly increasing sequence of indices $(n_k)_k$ whose corresponding cutting indices are all equal to a certain fixed c. By compactness of $\overline{\Omega}_X \mathsf{A}$, one may further assume that the sequences $(u'_{n_k})_k$, $(u''_{n_k})_k$, $(w'_{n_k})_k$, and $(w''_{n_k})_k$ converge, say respectively to u', u'', w', w''. By continuity of multiplication, and since $(L_n[\beta])_n$ is a decreasing sequence of languages for every ω-word β, it follows that

$$u' \in \bigcap_n \mathrm{cl}_{\mathsf{A}}(L_n[x_0 \cdots x_{c-1}]),\ u'' w' \in \bigcap_n \mathrm{cl}_{\mathsf{A}}(L_n[x_c]),\ w'' \in \bigcap_n \mathrm{cl}_{\mathsf{A}}(L_n[x_{c+1} \cdots x_{2p}]).$$

By Corollary 7.2, the preceding intersections are reduced respectively to the ω-words $x_0 \cdots x_{c-1}$, x_c, and $x_{c+1} \cdots x_{2p}$. Hence $u', w'' \in \Omega_X^\omega \mathsf{A}$ and $u'' w' = \epsilon[x_c]$. If c is even, then u'' is a prefix of the word x_c and hence $u = u' u'' \in \Omega_X^\omega \mathsf{A}$, as required. Hence we may as well assume that α is of the form $\alpha = (y)$.

By Lemma 3.2(e), we have $L_n[\alpha] = L_n[y^n] L_n[y]^*$. Thus, in view of the relation $u_n w_n \in L_n[\alpha]$, there exist factorizations $u_n = u_n' u_n''$ and $w_n = w_n' w_n''$ such that $u_n' \in L_n[y^{r_n}]$, $u_n'' w_n' \in L_n[y]$, and

$w''_n \in L_n[y^{s_n}]$, with $r_n + s_n + 1 \geq n$. If there is a strictly increasing sequence of indices $(n_k)_k$ such that $r_{n_k} = r$ is constant, then we may assume that the sequences $(u'_{n_k})_k$, $(u''_{n_k})_k$, $(w'_{n_k})_k$, and $(w''_{n_k})_k$ converge, say respectively to u', u'', w', w''. As above, it follows that $u', w'' \in \Omega^\omega_X \mathsf{A}$ and $u''w' = \epsilon[y]$. Since $\operatorname{rank} y < \operatorname{rank} \alpha$, the induction hypothesis then implies that u'' is an ω-word and, therefore so is $u = u'u''$.

Hence we may assume that $r_n \to \infty$ as $n \to \infty$. This implies that $y^{r_n} \to (y)$ in $\Omega^\omega_X \mathsf{A}$. Assuming again that $(u'_{n_k})_k$, $(u''_{n_k})_k$, $(w'_{n_k})_k$, and $(w''_{n_k})_k$ converge respectively to u', u'', w', w'', we conclude that $u' = \epsilon[(y)] \in \Omega^\omega_X \mathsf{A}$ and $u''w' = \epsilon[y]$. Invoking once more the induction hypothesis as above, the induction step is finally achieved, which proves the theorem. □

Some applications of Theorem 7.4 can be found in [6]. It plays, in particular, an important role in establishing the main result of that paper, namely a characterization of pseudowords over A which are given by ω-words. Other applications of Theorem 7.4 and of properties of the languages $L_n[\alpha]$, such as an algorithm to compute the closure $\mathrm{cl}_{\mathsf{A}}(L)$ of a rational language L, will appear in a paper which is under preparation [7].

Acknowledgements. This work was partly supported by the Pessoa French-Portuguese project Egide-Grices 11113YM *Automata, profinite semigroups and symbolic dynamics*. The work of the first and second authors was supported, in part, by *Fundação para a Ciência e a Tecnologia* (FCT) respectively through *Centro de Matemática da Universidade do Porto* and *Centro de Matemática da Universidade do Minho*. Both the first and second authors were also supported by FCT through the project PTDC/MAT/65481/2006, which is partly funded by the European Community Fund FEDER. The work leading to this paper has also been carried out within the framework of the ESF programme "Automata: from Mathematics to Applications (AutoMathA)", whose support is gratefully acknowledged.

References

1. J. Almeida, *Implicit operations on finite $\mathcal{J}$–trivial semigroups and a conjecture of I. Simon*, J. Pure Appl. Algebra **69** (1990), 205–218.
2. ______, *Finite semigroups and universal algebra*, World Scientific, Singapore, 1995, English translation.
3. ______, *Hyperdecidable pseudovarieties and the calculation of semidirect products*, Int. J. Algebra Comput. **9** (1999), 241–261.
4. ______, *Finite semigroups: an introduction to a unified theory of pseudovarieties*, Semigroups, Algorithms, Automata and Languages (Singapore) (G. M. S. Gomes, J.-É. Pin, and P. V. Silva, eds.), World Scientific, 2002, pp. 3–64.
5. ______, *Profinite semigroups and applications*, Structural Theory of Automata, Semigroups, and Universal Algebra (New York) (Valery B. Kudryavtsev and Ivo G. Rosenberg, eds.), NATO Science Series II: Mathematics, Physics and Chemistry, vol. 207, Springer, 2005, Proceedings of the NATO Advanced Study Institute on Structural Theory of Automata, Semigroups and Universal Algebra, Montréal, Québec, Canada, 7-18 July 2003.
6. J. Almeida, J. C. Costa, and M. Zeitoun, *Iterated periodicity over finite aperiodic semigroups*, Tech. Report CMUP 2009-3, Univ. Porto.
7. ______, *The pin-reutenauer algorithm for aperiodic semigroups*, under preparation.
8. ______, *Tameness of pseudovariety joins involving R*, Monatsh. Math. **146** (2005), no. 2, 89–111.
9. ______, *Complete reducibility of systems of equations with respect to R*, Portugal. Math. **64** (2007), 445–508.

10. J. Almeida and M. Delgado, *Sur certains systèmes d'équations avec contraintes dans un groupe libre*, Portugal. Math. **56** (1999), 409–417.
11. ———, *Sur certains systèmes d'équations avec contraintes dans un groupe libre—addenda*, Portugal. Math. **58** (2001), 379–387.
12. J. Almeida and B. Steinberg, *On the decidability of iterated semidirect products and applications to complexity*, Proc. London Math. Soc. **80** (2000), 50–74.
13. ———, *Syntactic and global semigroup theory, a synthesis approach*, Algorithmic Problems in Groups and Semigroups (J. C. Birget, S. W. Margolis, J. Meakin, and M. V. Sapir, eds.), Birkhäuser, 2000, pp. 1–23.
14. J. Almeida and M. Zeitoun, *An automata-theoretic approach to the word problem for ω-terms over R*, Theoret. Comp. Sci. **370** (2007), 131–169.
15. C. J. Ash, *Inevitable graphs: a proof of the type II conjecture and some related decision procedures*, Int. J. Algebra Comput. **1** (1991), 127–146.
16. J. C. Costa, *Free profinite locally idempotent and locally commutative semigroups*, J. Pure Appl. Algebra **163** (2001), 19–47.
17. J. C. Costa and C. Nogueira, *Complete reducibility of the pseudovariety LSl*, Int. J. Algebra Comput. **19** (2009), 1–36.
18. J. C. Costa and M. L. Teixeira, *Tameness of the pseudovariety LSl*, Int. J. Algebra Comput. **14** (2004), 627–654.
19. B. Herwig and D. Lascar, *Extending partial automorphisms and the profinite topology on free groups*, Trans. Amer. Math. Soc. **352** (2000), 1985–2021.
20. M. Lothaire, *Combinatorics on words*, Addison-Wesley, Reading, Mass., 1983.
21. ———, *Algebraic combinatorics on words*, Cambridge University Press, Cambridge, UK, 2002.
22. J. McCammond, *The solution to the word problem for the relatively free semigroups satisfying $t^a = t^{a+b}$ with $a \geq 6$*, Int. J. Algebra Comput. **1** (1991), 1–32.
23. ———, *Normal forms for free aperiodic semigroups*, Int. J. Algebra Comput. **11** (2001), 581–625.
24. J. Rhodes and B. Steinberg, *The q-theory of finite semigroups*, Springer Monographs in Mathematics, Springer, 2009.
25. L. Ribes and P. A. Zalesskiĭ, *On the profinite topology on a free group*, Bull. London Math. Soc. **25** (1993), 37–43.
26. ———, *Profinite groups*, Ergeb. Math. Grenzgebiete 3, no. 40, Springer, Berlin, 2000.
27. M.P. Schützenberger, *On finite monoids having only trivial subgroups*, Inform. and Control **8** (1965), 190–194.