ITERATED PERIODICITY OVER FINITE APERIODIC SEMIGROUPS

JORGE ALMEIDA, JOSÉ CARLOS COSTA, AND MARC ZEITOUN

Abstract. This paper provides a characterization of pseudowords over the pseudovariety of all finite aperiodic semigroups that can be described from the free generators using only the operations of multiplication and ω-power. A necessary and sufficient condition for this property to hold turns out to be given by the conjunction of two rather simple finiteness conditions: the nonexistence of infinite anti-chains of factors and the rationality of the language of McCammond normal forms of ω-terms that define factors of the given pseudoword. The relationship between pseudowords with this property and arbitrary pseudowords is also investigated.

1. Introduction

Since the mid nineteen seventies, the theory of finite semigroups has seen a significant boost thanks to its connections and applications in computer science. The now classical framework for the relationships between the two areas is provided by Eilenberg's correspondence between pseudovarieties of finite semigroups and varieties of rational languages [16], whose importance is also corroborated by the fact that it has been the object of many extensions [17]. The typical application consists in the solution (positive or negative) of an algorithmic problem about rational languages or finite automata by translating it to a membership problem in a suitable pseudovariety of semigroups. Concomitantly, such applications have served as a guide in the theory of finite semigroups, pointing out the most interesting pseudovarieties to study.

Unlike varieties of algebras, pseudovarieties do not in general have free algebras. In the following decade, a substitute for relatively free algebras emerged and found many applications in the context of pseudovarieties of semigroups: the relatively free profinite algebras. The common algebraic and combinatorial properties of the semigroups in a pseudovariety become encoded in the algebraic and topological properties of their free profinite semigroups which, over a given finite alphabet, can be viewed as the Stone dual for the Boolean algebra of languages over that alphabet recognized by semigroups in the pseudovariety [1, 17, 27]. The difficulty, of course, is to obtain sufficient structural information about these profinite semigroups for the intended applications.

Through the seminal work of Ash [14] and its generalizations found by the first author and Steinberg [3, 11, 10], 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 the pro-V semigroup freely generated by 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 generated by the same set X with respect to a suitable signature, 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 perhaps more familiar ω-power which, in a finite semigroup, gives the unique idempotent power of an element. In this case, let us call an algebra in our signature simply an *ω-semigroup*.

2000 *Mathematics Subject Classification.* 20M05, 20M07, 20M35, 68Q45, 68Q70.

Key words and phrases. Aperiodic semigroup. Free profinite semigroup. Pseudoword. ω-term. McCammond normal form. Rational language. Anti-chain. Semigroupoid.

To prove tameness, one must establish that the word problem for the chosen relatively free algebra 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 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 [14], but also independently a special case was proved using methods from profinite group theory [28, 29], and later was rediscovered as a model-theoretic result [18, 8, 9].

For the pseudovariety A of all finite aperiodic semigroups, tameness has been announced by John Rhodes at the International Conference on Algorithmic Problems in Groups and Semigroups (Lincoln, Nebraska, U.S.A., May 1998) but no published proof has yet appeared except for the solution of the word problem for the relatively free ω-semigroup in the variety of ω-semigroups generated by A, which has been obtained by McCammond [24]. This seems to be explained, at least in part, by the fact that a gap was discovered in the proof of a key result from [13] used in [10] to reduce the decidability of the Krohn-Rhodes complexity of finite semigroups [19] (see [27]) to the tameness of the pseudovariety A. McCammond's solution of the word problem is obtained by constructing a normal form, to which every ω-term may be reduced, preserving its action on semigroups from A, and then showing that two distinct ω-terms in normal form act differently on some finite aperiodic semigroup by using his own solution of the word problem for certain free Burnside semigroups [23].

The present paper is intended as a contribution to a deeper understanding of how the ω-subsemigroup generated by X fits in the free pro-A semigroup F on the finite alphabet X. The main result provides a characterization of the elements u of the ω-subsemigroup in terms of simple finiteness conditions: in any infinite set of factors of u in F, at least one of them is a factor of another one; the language of ω-terms in normal form that determine elements of F that are factors of u is rational. This may be viewed as a sort of iterated periodicity result, by interpreting the operation of taking the ω-power as the infinite iteration of its argument.

The paper is organized as follows. Section 2 gathers most of the preliminary material including some remarks on uniformly recurrent right infinite words, a description of McCammond's normal form and algorithm to compute it, as well as some important properties of the pseudovariety A. Section 3 gives a characterization of pseudovarieties V such that the set of all finite elements of the free pro-V semigroup $\overline{\Omega}_X\mathsf{V}$ over a finite set of free generators X is always open and some related complementary observations to results of [4] on uniformly recurrent pseudowords. In Section 4, we introduce a notion of rank for elements of a semigroup inspired by McCammond's rank of ω-terms. Section 5 investigates some properties of Green's relations in the semigroup of ω-words over A. Several characterizations of periodicity for uniformly recurrent pseudowords over pseudovarieties containing all finite local semilattices are the theme of Section 6 which brings forth the relevance of properties like the rationality of the set of finite factors of a pseudoword as well as various chain conditions.

Sections 7 to 9 contain the main technicalities necessary for the main theorem, which is proved in Section 10. Section 7 establishes necessary conditions for a pseudoword to be describable by an ω-term, from which emerges the notion of slim pseudoword, which is investigated in Section 8. Section 9 gives a way to encode slim pseudowords over A as pseudopaths in suitable profinite categories, tools whose introduction is postponed to this point, rather than included in Section 2 for they play no role earlier in the paper. It is the iterative application of this encoding procedure that allows us to show how to construct an ω-term description of a pseudoword over A satisfying suitable finiteness conditions.

2. Preliminaries

Throughout this paper, we assume that the reader is familiar with the general basic theory of pseudovarieties and specifically with the central role played by relatively free profinite semigroups. A quick introduction, both to the theory and to the applications, is found in [5]. For more comprehensive treatments, see [1, 27].

We adopt the following notation. A general finite alphabet is denoted X. For a pseudovariety V of (finite) semigroups, the pro-V semigroup freely generated by X is represented by $\overline{\Omega}_X\mathsf{V}$. It may be constructed for example as the projective limit of all X-generated semigroups from V or as the completion of the free semigroup X^+ on X with respect to the uniform structure generated by all congruences θ such that $X^+/\theta \in \mathsf{V}$.

The above construction of $\overline{\Omega}_X\mathsf{V}$ provides a natural mapping $\iota_{\mathsf{V}} : X^+ \to \overline{\Omega}_X\mathsf{V}$. While this may not be injective, it is injective for many pseudovarieties of interest. Elements of $\overline{\Omega}_X\mathsf{V}$ are called *implicit operations* or *pseudowords* over V; if no reference to a pseudovariety is made, it is assumed to be S, the pseudovariety of all finite semigroups. The elements of $\iota_{\mathsf{V}}(X^+)$ are said to be *finite* while the remaining elements of $\overline{\Omega}_X\mathsf{V}$ are called *infinite pseudowords*.

The letter A denotes the pseudovariety of all finite aperiodic semigroups, consisting of all finite semigroups all of whose subsemigroups that are groups are trivial. The pseudovariety LSl consists of all finite semigroups S such that, for every idempotent $e \in S$, the monoid eSe is a semilattice, that is eSe is a commutative semigroup in which every element is idempotent.

By a *right infinite word* over an alphabet X we mean an infinite sequence $\mathbf{s} = a_1a_2\cdots$ of letters of X. Its *finite factors* are the words of the form $u = a_i \cdots a_j$ with $i \le j$; we then also say that u *occurs in* $\mathbf{s}$ *starting at position* i. We denote by $F(\mathbf{s})$ the set of all finite factors of $\mathbf{s}$.

We say that the right infinite word $\mathbf{s} = a_1a_2\cdots$ is *periodic* if there exists n such that $a_{i+kn} = a_i$ for all $k \ge 0$ and all i. We *concatenate* a finite word $w = b_1\cdots b_m$ with a right infinite word $\mathbf{s}$ to obtain the right infinite word $w\mathbf{s} = b_1\cdots b_m a_1 a_2 \cdots$. We say that the right infinite word $\mathbf{t}$ is *ultimately periodic* if it is of the form $w\mathbf{s}$ for some finite word w and some periodic right infinite word $\mathbf{s}$.

By K and D we denote the pseudovarieties consisting of all finite semigroups in which idempotents are left zeros, respectively right zeros. Let K_n and D_n be the pseudovarieties consisting of all finite semigroups satisfying respectively the identities $x_1\cdots x_n y = x_1 \cdots x_n$ and $xy_1\cdots y_n = y_1\cdots y_n$.

The structure of $\overline{\Omega}_X\mathsf{K}$ is easy to describe. The mapping $\iota_{\mathsf{K}} : X^+ \to \overline{\Omega}_X\mathsf{K}$ is injective and its image is a discrete subspace, which we identify with X^+. The elements of $\overline{\Omega}_X\mathsf{K} \setminus X^+$ have unique prefixes in X^+ of any given length and they are completely determined by these prefixes, and so they may be identified with right infinite words, that is with elements of $X^{\mathbb{N}}$, where $\mathbb{N}$ denotes the set of all natural numbers. The infinite elements of $\overline{\Omega}_X\mathsf{K}$ are left zeros and multiplication on the left by a finite word is obtained by concatenation. The topology of the subspace of infinite pseudowords, identified with infinite words, is precisely the product topology of $X^{\mathbb{N}}$, so that this space is homeomorphic to the Cantor set of reals. A sequence of finite words converges to an infinite word if every finite prefix of the latter is a prefix of all but finitely many terms of the sequence. The structure of $\overline{\Omega}_X\mathsf{D}$ is dual.

By a *quasi-order* on a set we mean a reflexive and transitive binary relation $\le$. The associated strict order $<$ is defined by $x < y$ if $x \le y$ and $y \not\le x$. A quasi-order is a *well-quasi-order* (*wqo* for short) if it admits no infinite descending chains $x_1 > x_2 > \cdots$ and no infinite anti-chains, that is infinite sets in which, for any two distinct elements x and y, $x \not\le y$. Equivalently, given any sequence $(x_k)_k$, there exist m, n such that $m < n$ and $x_m \le x_n$.

Given two elements s and t of a semigroup S, we write $s \ge_{\mathcal{J}} t$ and say that s lies $\mathcal{J}$*-above* t if t may be written as a product in which s appears as a factor. Theorem 2.4 below may be expressed by saying that $\Omega_X^{\kappa}\mathsf{A}$ forms a filter in $\overline{\Omega}_X\mathsf{A}$ with respect the quasi-order relation $\le_{\mathcal{J}}$. We write $\mathcal{J}$ for the equivalence relation $\le_{\mathcal{J}} \cap \ge_{\mathcal{J}}$. In general, when we talk about anti-chains in this paper, we mean anti-chains for the factor ordering $\ge_{\mathcal{J}}$.

Similarly, we write $s \ge_{\mathcal{R}} t$ if s is a left factor (or *prefix*) of t in a given semigroup S and $s \ge_{\mathcal{L}} t$ if s is a right factor (or *suffix*) of t. The corresponding equivalence relations are $\mathcal{R} = \le_{\mathcal{R}} \cap \ge_{\mathcal{R}}$ and $\mathcal{L} = \le_{\mathcal{L}} \cap \ge_{\mathcal{L}}$. The remaining Green relations are $\mathcal{H} = \mathcal{R} \cap \mathcal{L}$ and $\mathcal{D} = \mathcal{R} \circ \mathcal{L} = \mathcal{L} \circ \mathcal{R}$, where $\circ$ denotes the composition of binary relations. It is well known that the maximal subgroups of a semigroup S are precisely those $\mathcal{H}$-classes that contain idempotents. In the case of a finite aperiodic semigroup, $\mathcal{H}$ is the equality relation.

A language $L \subseteq X^+$ is *factorial* if it is closed under taking factors. If L is infinite, we say that L is *uniformly recurrent* if, for every $u \in L$ there is an integer n such that, for all $w \in L$ with $|w| \ge n$, u is a factor of w. The right infinite word $\mathbf{s}$ is *uniformly recurrent* if the language $F(\mathbf{s})$ is uniformly recurrent. The following combinatorial properties of right infinite words will be useful in the sequel. We do not know if they have been observed elsewhere.

Lemma 2.1. *Let* $\mathbf{s}$ *be a right infinite word such that* $F(\mathbf{s})$ *has no infinite anti-chains. Then* $\mathbf{s}$ *is of the form* $u\mathbf{t}$*, where* u *is finite and* $\mathbf{t}$ *is uniformly recurrent.*

Proof. Suppose that every factorization $\mathbf{s} = u\mathbf{t}$, where u is a finite word, is such that $\mathbf{t}$ is not uniformly recurrent. We construct a sequence $(f_n)_n$ of factors of $\mathbf{s}$ which is an anti-chain, which produces a contradiction and therefore proves the lemma.

Since $\mathbf{s}$ itself is not uniformly recurrent, there exists an arbitrarily long factor f_1 of $\mathbf{s}$ that is not a factor of infinitely many factors of $\mathbf{s}$. If f_1 occurs infinitely often as a factor of $\mathbf{s}$, then there are infinitely many factors of $\mathbf{s}$ of the form f_1uf_1 in which f_1 appears only as prefix and suffix. Since such factors constitute an anti-chain, it follows that there is a factorization $\mathbf{s} = u_1\mathbf{s}_1$ such that $f_1 \notin F(\mathbf{s}_1)$.

Applying the same argument to $\mathbf{s}_1$, we deduce that there is some factorization $\mathbf{s}_1 = u_2\mathbf{s}_2$ and some $f_2 \in F(\mathbf{s}_1) \setminus F(\mathbf{s}_2)$ such that $|f_2| > |f_1|$. Iterating this procedure, we obtain the announced anti-chain $(f_n)_n$, which completes the proof of the lemma. □

Lemma 2.2. *Let* $\mathbf{s}$ *be a uniformly recurrent right infinite word such that the language* $F(\mathbf{s})$ *is rational. Then* $\mathbf{s}$ *is periodic.*

Proof. This could be deduced from Theorem 6.2 below, but we give a direct proof. Since the language $F(\mathbf{s})$ is rational, infinite, and factorial, there is some word $u \in A^+$ such that every finite power of u is a factor of $\mathbf{s}$. Let $\mathbf{s} = u_1u_2\cdots u_n\cdots$ be the factorization of $\mathbf{s}$ in factors of length $|u|$. Since $\mathbf{s}$ is uniformly recurrent, for each n, $u_1\cdots u_n$ is a factor of some power of u. Hence, since all u_i have the same length as $|u|$, it follows that they are all equal (to some conjugate of u). Since this is true for all n, it follows that $\mathbf{s}$ is periodic. □

Combining the two preceding lemmas, we obtain the following result.

Proposition 2.3. *Let* $\mathbf{s}$ *be a right infinite word. Then* $\mathbf{s}$ *is ultimately periodic if and only if the language* $F(\mathbf{s})$ *is rational and it contains no infinite* $\mathcal{J}$*-anti-chains.*

Proof. Suppose first that $\mathbf{s}$ is ultimately periodic, say $\mathbf{s} = uv^\infty$, where v^∞ stands for the right infinite periodic word with period v. Then $F(\mathbf{s})$ is the language of factors of uv^* and, therefore, it is a rational language. More precisely, the finite factors of $\mathbf{s}$ are the factors of u together with the words of the form xv^ky, where x is a suffix of either u or v, $k \geq 0$, and y is a prefix of v, and it is immediate to deduce that there are no infinite $\mathcal{J}$-anti-chains of such words.

Conversely, if there are no infinite anti-chains of finite factors of $\mathbf{s}$ then, by Lemma 2.1, $\mathbf{s}$ has a factorization of the form $\mathbf{s} = u\mathbf{t}$, where u is a finite word and $\mathbf{t}$ is a uniformly recurrent infinite word. We claim that the language $F(\mathbf{t})$ is rational which, by Lemma 2.2, entails that $\mathbf{t}$ is periodic and, therefore, that $\mathbf{s}$ is ultimately periodic.

To prove the claim, note that, since $\mathbf{t}$ is uniformly recurrent, the factors of $\mathbf{t}$ are precisely the factors v of $\mathbf{s}$ such that, for every n, there exists $x \in X^+$ of length n with $xv \in F(\mathbf{s})$. By hypothesis, $F(\mathbf{s})$ is a rational language, whence there is a finite deterministic automaton $\mathscr{A}$ recognizing the reverse L of this language, consisting of the reversed words. In this context, we consider automata as recognizing devices only for nonempty words. Since L is factorial, all states from which a terminal state is accessible are also terminal, that is, either there is only one state or we may as well assume that the only state which is not terminal is a sink. If we change to nonterminal those states from which no strongly connected component of terminal states is accessible, we obtain an automaton $\mathscr{A}'$ that recognizes the set of all words $v \in L$ which can be extended arbitrarily to the right to other words from L. Hence $F(\mathbf{t})$ is rational. □

Going back to the profinite world, we recall that implicit operations over V can be naturally interpreted in pro-V semigroups as follows: given $w \in \overline{\Omega}_X\mathsf{V}$ and a profinite semigroup S, the interpreted operation $w_S : S^X \to S$ associates to each function $\varphi : X \to S$ the element $\hat{\varphi}(w) \in S$, where $\hat{\varphi}$ is the unique extension of φ to a continuous homomorphism $\overline{\Omega}_X\mathsf{V} \to S$.

Given $u, v \in \overline{\Omega}_X\mathsf{V}$, we call the formal equality $u = v$ a *pseudoidentity (over* V*)*. For a pro-V semigroup S, we then say that S *satisfies* the pseudoidentity $u = v$ and write $S \models u = v$ if $u_S = v_S$. For a set Σ of pseudoidentities, we let $[\![\Sigma]\!]$ denote the class of all finite semigroups that satisfy all pseudoidentities from Σ. It is easy to check that $[\![\Sigma]\!]$ is a pseudovariety and Reiterman [26] showed that every pseudovariety is of this form.

For a set σ of implicit operations, one may in particular use natural interpretation of the elements of σ to define the σ-subalgebra of $\overline{\Omega}_X\mathsf{V}$ generated by X, which is denoted $\Omega_X^\sigma\mathsf{V}$. Formal terms over a finite alphabet

X in the signature σ are called *σ-terms.* Since our multiplication is always associative, without further reference, we identify terms that only differ by the order in which multiplications are to be carried out.

Following [10], we denote by κ the set consisting of the operations of multiplication and pseudo-inversion $x \mapsto x^{\omega-1}$, where, for an element s of a finite semigroup, $s^{\omega-1}$ stands for the inverse of ss^{ω} in the maximal subgroup of the subsemigroup generated by s, whose idempotent is denoted s^{ω}. In general, for a profinite semigroup S, one may define $s^{\omega+k} = \lim_n s^{n!+k}$ for $k \in \mathbb{Z}$ and $s^{m^{\omega}} = \lim_n s^{m^{n!}}$ for $m \geq 1$ [5]. For $k, \ell \in \mathbb{Z}$ with $\ell > 0$, we have $s^{\omega+(k+\ell)} = s^{\omega+k}s^{\ell}$.

Note that, for pseudovarieties contained in A, since the subgroup in question is trivial, the operations $x^{\omega-1}$ and x^{ω} coincide. Since this paper is concerned mainly with the pseudovariety A, our κ-terms will use the operation x^{ω} rather than $x^{\omega-1}$ and such terms are also, abusively, called *ω-terms.* More formally, an ω-term on a set X is an element of the unary semigroup U_X freely generated by X.

We will sometimes adopt the simplified notation of McCammond [24] for ω-terms under which the curved parenthesis (α) stands for α^{ω}. This allows us to refer to ω-terms as words over an extended alphabet $X \cup \{(,)\}$, which is particularly useful for McCammond's solution of the word problem for $\Omega_X^{\kappa}\mathsf{A}$ and is also instrumental in the formulation of our main result. Note that the words in the extended alphabet that represent ω-terms are precisely those for which the opening and closing parentheses match, that is, 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 $(X \cup \{(,)\})^+$ 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 $\epsilon : U_X \to \Omega_X^{\kappa}\mathsf{A}$ that fixes each $x \in X$ when we view X as a subset of U_X and $\Omega_X^{\kappa}\mathsf{A}$ 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 $\Omega_X^{\kappa}\mathsf{A}$ will sometimes be called *ω-words.* Whenever we say that an ω-term over the alphabet X is a factor of another, we mean that that is the case in the free semigroup $(X \cup \{(,)\})^+$.

The ω-word problem for A (over X), consists in deciding when two elements of U_X have the same image under ϵ. To solve this problem, McCammond described a normal form for ω-terms over A that we will use extensively in this paper. For its description, a total order is fixed on the underlying alphabet X; on the extended alphabet, we set $(< x <)$ for every $x \in X$. A *primitive word* is a word that cannot be written in the form u^n with $n > 1$. Two words u and v are said to be *conjugate* if there are factorizations of the form $u = xy$ and $v = yx$, with the words x and y possibly empty. A *Lyndon word* is a primitive word that is lexicographically minimum in its conjugacy class. The *rank* of a word in the extended alphabet is the maximum number of nested parentheses in it.

A *rank 0 normal form* ω-term is simply a finite word. Assuming that rank i normal form terms have been defined, a *rank* $i+1$ *normal form term* is a term 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-terms such that

- (a) each β_k is a Lyndon word;
- (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 term;
- (d) at least one of the properties (b) and (c) is lost by canceling from α_j a prefix β_j (in case $j > 0$) or a suffix β_{j+1} (in case $j < n$).

McCammond also described a method to transform an arbitrary ω-term into one in normal form with the same image under ϵ. Moreover, he proved that if two ω-terms in normal form have the same image under ϵ, then they are equal.

Since we will need to refer to McCammond's procedure to transform an arbitrary ω-term into one in normal form, we proceed to describe its steps. The procedure consists in applying elementary changes that obviously retain the value of the ω-term under ϵ. The types of changes are given by the following 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 a subterm given by the left side of a rule of type 1–4 is replaced in a term 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.

The steps in McCammond's normal form algorithm may now be described as follows.

(1) In case the given term is a word, do nothing and stop.
(2) Apply all possible contractions of type 1.
(3) Apply all possible contractions of type 2. The resulting term may be written as a product of the form

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

where each subterm α_j and β_k has strictly smaller rank.
(4) Apply recursively the algorithm to each subterm α_j and β_k, to put it in normal form.
(5) If some β_k is an idempotent (that is, it has the same normal form as its square), then remove the parentheses around it in the expression (1), join it with the adjacent α subterms and apply again the previous item to the new α subterms.
(6) By means of an expansion of type 4_L and an expansion of type 4_R and a shift of parentheses by application of the rule 5, write each remaining subterm of the form (β_k) in the form $\varepsilon_1(\gamma)\varepsilon_2$ where:
   (*i*) γ is a Lyndon word in the extended alphabet;
   (*ii*) $\varepsilon_1\gamma\varepsilon_2$ is the normal form of β_k^3;
   (*iii*) $\varepsilon_1\varepsilon_2$ is the normal form of β_k^2;
   (*iv*) $\varepsilon_2\varepsilon_1$ has the same normal form as γ^2;
   (*v*) the rank of each of ε_1 and ε_2 is $\mathrm{rank}[\beta_k]$.
(7) Using contractions of types 3 and 4, replace maximum rank subterms of the form $(\gamma)\gamma^\ell(\gamma)$ by (γ). The remaining maximum rank subterms of the form $(\gamma)\varepsilon(\delta)$, where $\mathrm{rank}[\gamma] = \mathrm{rank}[\delta] \geq \mathrm{rank}[\varepsilon]$, are called *crucial portions*, whereas the prefix $\varepsilon(\delta)$ and the suffix $(\gamma)\varepsilon$, with similar assumptions on $\gamma, \delta, \varepsilon$, are called respectively the *initial portion* and the *final portion*. An initial portion $\varepsilon(\delta)$ with $\varepsilon = 1$ (the empty word) is said to be *trivial* and similarly for final portions.
(8) Standardize each crucial portion $(\gamma)\varepsilon(\delta)$ by using expansions of type 4_R on (γ) and of type 4_L on (δ) until the intermediate term ε' is no longer a prefix of a power of γ nor a suffix of a power of δ, as a word in the extended alphabet, and apply contractions of type 4_R on the left and of type 4_L on the right while that property holds.
(9) Reduce the initial portion $\varepsilon(\delta)$ by dropping the largest suffix of ε that is a power of δ as a word in the extended alphabet; proceed dually with the final portion.

It is not very hard to check that the above normal form algorithm does indeed transform an arbitrary ω-term into one in normal form. Since all the rewriting rules are based on identities of ω-semigroups that are valid in A, every ω-term over X has the same image under ϵ as its normal form. To prove that distinct ω-terms in normal form have different images in $\Omega^\kappa_X\mathsf{A}$, McCammond used his solution of the word problem for certain free Burnside semigroups [23]. We have obtained a direct combinatorial proof of the same result which leads to many other applications, including the following two theorems, the first of which plays an important role in this paper and which does not apparently follow easily from McCammond's results. The paper containing our proof is under preparation and will appear elsewhere.

Theorem 2.4. *If $v \in \Omega^\kappa_X\mathsf{A}$ and $u \in \overline{\Omega}_X\mathsf{A}$ is a factor of v, then $u \in \Omega^\kappa_X\mathsf{A}$.*

It is well know that the Green relations $\mathcal{J}$ and $\mathcal{D}$ coincide in every compact semigroup, and so in particular they coincide in every finite semigroup. In fact, this property holds more generally in so-called *stable semigroups* (cf. [20]), that is in semigroups S such that, $s \geq_{\mathcal{R}} t$ (respectively $s \geq_{\mathcal{L}} t$) and $s \mathrel{\mathcal{J}} t$ implies $s \mathrel{\mathcal{R}} t$ (respectively $s \mathrel{\mathcal{L}} t$).

Corollary 2.5. *The semigroup $\Omega^\kappa_X\mathsf{A}$ is stable. Hence the Green relations $\mathcal{J}$ and $\mathcal{D}$ of $\Omega^\kappa_X\mathsf{A}$ coincide, while $\mathcal{H}$ is the equality relation.*

Proof. If the relations $s \geq_{\mathcal{R}} t \mathrel{\mathcal{J}} s$ hold in the subsemigroup $\Omega^\kappa_X\mathsf{A}$ then they also hold in $\overline{\Omega}_X\mathsf{A}$. Since $\overline{\Omega}_X\mathsf{A}$, like any compact semigroup, is stable, we deduce that there exists $u \in (\overline{\Omega}_X\mathsf{A})^1$ such that $su = t$. If $u = 1$, then $s = t$. Otherwise, $u \in \Omega^\kappa_X\mathsf{A}$ by Theorem 2.4, and so $s \mathrel{\mathcal{R}} t$ holds in $\Omega^\kappa_X\mathsf{A}$ in any case.

It was already pointed out that stability implies $\mathcal{J} = \mathcal{D}$. That $\mathcal{H}$ is trivial in $\Omega_X^\kappa \mathsf{A}$ is actually true for every subsemigroup of $\overline{\Omega}_X \mathsf{A}$ since, if two elements are $\mathcal{H}$-equivalent in a subsemigroup of $\overline{\Omega}_X \mathsf{A}$, then they are also $\mathcal{H}$-equivalent in $\overline{\Omega}_X \mathsf{A}$. It is well known that $\mathcal{H}$ is trivial in $\overline{\Omega}_X \mathsf{A}$ (cf. [1, Corollary 5.6.2]): if $s \mathrel{\mathcal{H}} t$ and $s \neq t$, then there exists a continuous homomorphism $\varphi : \overline{\Omega}_X \mathsf{A} \to S$ onto a finite aperiodic semigroup S such that $\varphi(s) \neq \varphi(t)$. But, since the Green relations are preserved by applying homomorphisms, $\varphi(s) \mathrel{\mathcal{H}} \varphi(t)$. Since the Green relation $\mathcal{H}$ is trivial in finite aperiodic semigroups, we reach a contradiction. Hence $\mathcal{H}$ is trivial in $\Omega_X^\kappa \mathsf{A}$. □

Following [25, 20], we say that a semigroup S is *equidivisible* if, for any two factorizations $xy = zt$ of the same element of S, there exists $u \in S^1$ such that, either $x = zu$ and $uy = t$, or $xu = z$ and $y = ut$. It follows that any two factorizations of the same element of S may be refined by further factorizating some of the factors so as to reach the same factorization.

We also say that a pseudovariety V is *equidivisible* if $\overline{\Omega}_X \mathsf{V}$ is equidivisible for every finite set X. The following result, which can be easily derived using results from [6, Section 2] is very useful. Details will be given in a forthcoming paper where equidivisibilty is extensively explored.

Theorem 2.6. *The pseudovariety A is equidivisible.*

3. Set of all finite words open

Let $\mathsf{N} = [\![x^\omega = 0]\!]$ be the pseudovariety of all finite nilpotent semigroups. The following result completes well-known properties of N.

Theorem 3.1. *The following conditions are equivalent for a pseudovariety V:*

(*a*) $\mathsf{V} \supseteq \mathsf{N}$;
(*b*) *if V satisfies a pseudoidentity of the form $u = v$, where u is a finite word, then $u = v$;*
(*c*) *if V satisfies a two-letter pseudoidentity of the form $u = v$, where u is a finite word, then $u = v$;*
(*d*) *for every finite alphabet A, the natural homomorphism $\iota_A : A^+ \to \overline{\Omega}_A \mathsf{V}$ is injective and its image is an open discrete subset;*
(*e*) *for the two-letter alphabet $A = \{a, b\}$, the natural homomorphism $\iota_A : A^+ \to \overline{\Omega}_A \mathsf{V}$ is injective and the subset $\{x^n : n \geq 1\}$ is open in $\overline{\Omega}_{\{x\}} \mathsf{V}$.*

Proof. $(a) \Rightarrow (b)$ This is well-known. Simply observe that, if the word u over the finite alphabet A has length n, then the Rees quotient S of A^+ by the ideal $A^n A^+$ is a nilpotent semigroup which satisfies no non-trivial pseudoidentity of the form $u = v$.

$(b) \Rightarrow (a)$ This is also well-known. Indeed, by Reiterman's Theorem [26], it suffices to show that N satisfies every non-trivial pseudoidentity $u = v$ which holds in V. By (b), for such a pseudoidentity, u and v are both infinite pseudowords, whence they have factors of the form w^ω (cf. [1, Corollary 5.6.2]). Hence, N satisfies the pseudoidentities $u = 0 = v$.

$(b) \Rightarrow (c)$ and $(d) \Rightarrow (e)$ are obvious.

$(b) \Rightarrow (d)$ That V satisfies no non-trivial identities $u = v$ with $u, v \in A^+$ is clearly equivalent to the statement that ι_A is injective. Suppose that $\operatorname{Im} \iota_A$ is not an open discrete subset. Then there is an injective sequence $(w_n)_n$ which converges to a word u in $\overline{\Omega}_A \mathsf{V}$. Since the sequence is injective, if it has infinitely many terms which are finite words (meaning that they belong to $\operatorname{Im} \iota_A$), then it contains arbitrarily long words, which implies that the limit u is of the form $u = xy^\omega z$ for some $x, y, z \in \overline{\Omega}_A \mathsf{V}$. Hence V satisfies the non-trivial pseudoidentity $u = xy^\omega z$, in contradiction with (b). Note that in this argument, we worked with a fixed finite alphabet. $(c) \Rightarrow (e)$ follows from the particular case of a two-letter alphabet.

$(e) \Rightarrow (b)$ Let $A = \{a, b\}$ be a two-letter alphabet and let $C = \{c_1, \ldots, c_n\}$ be an arbitrary finite alphabet with n letters. Then the homomorphism $\varphi : C^+ \to A^+$ defined by $\varphi(c_i) = b^i a$ is injective. Let $\hat{\varphi} : \overline{\Omega}_C \mathsf{V} \to \overline{\Omega}_A \mathsf{V}$ be the continuous homomorphism defined by $\hat{\varphi}(\iota_C(c_i)) = \iota_A(b^i a)$, that is so that the following diagram commutes:

$$
\begin{tikzcd}
C^+ \arrow[r, "\iota_C"] \arrow[d, "\varphi"'] & \overline{\Omega}_C \mathsf{V} \arrow[d, "\hat{\varphi}"] \\
A^+ \arrow[r, "\iota_A"] & \overline{\Omega}_A \mathsf{V}
\end{tikzcd}
$$

Since $\iota_A \circ \varphi$ is injective, it follows that ι_C is also injective, that is V satisfies no non-trivial identity on the alphabet C.

Note that, if V satisfies some non-trivial pseudoidentity $u = v$, with $u \in C^+$, then V satisfies some pseudoidentity of the form $x^k = x^{\omega+k}$ and hence $\lim_n x^{(2\cdot3\cdots p_n)^\omega+k} = x^k$, where p_n denotes the nth prime. The assumption that x^+ is an open subset of $\overline{\Omega}_{\{x\}}\mathsf{V}$ entails that V satisfies $x^{(2\cdot3\cdots p_n)^\omega+k} = x^{m_n}$, with m_n finite, for all sufficiently large n. If $\ell \in \{1,\dots,n\}$, n is sufficiently large, and r is so large that p_ℓ^r does not divide $m_n - k$, then the pseudoidentity $x^{(2\cdot3\cdots p_n)^\omega+k} = x^{m_n}$ fails in the cyclic group $\mathbb{Z}/p_\ell^r\mathbb{Z}$, whence this group does not belong to V. It follows that, for every prime p, there is the largest exponent $r = r_p \geq 0$ such that $\mathbb{Z}/p^r\mathbb{Z} \in \mathsf{V}$.

Next, for each $n \geq 1$, let w_n denote an accumulation point of the sequence $\left(x^{(p_n p_{n+1}\cdots p_m)^\omega}\right)_m$. We claim $\lim_n w_n = x^{\omega+1}$ in $\overline{\Omega}_{\{x\}}\mathsf{S}$, from which it follows that $\lim_n w_n = x^{k+1}$ in $\overline{\Omega}_{\{x\}}\mathsf{V}$. Before establishing the claim, note that the assumption that x^+ is open now implies that, for all sufficiently large n, $\mathsf{V} \models w_n = x^{\ell_n}$ for some finite exponent $\ell_n \geq 1$. It follows that, for some n, ℓ_n is divisible by all $p_s^{r_s}$ with $s \geq n$, whence $r_s = 0$ for all s greater than or equal to some n. We deduce that $\mathsf{V} \models x^k = x^{k+\prod_{i<n} p_i^{r_i}}$, which contradicts the assumption that ι_A is injective.

It remains to establish the claim that $\lim_n w_n = x^{\omega+1}$, which means that every monogenic finite semigroup S satisfies the pseudoidentity $w_n = x^{\omega+1}$ for all sufficiently large n. This amounts to an elementary exercise which we solve for the sake of completeness. Since both w_n and $x^{\omega+1}$ are infinite, the substitution of x by $x^{\omega+1}$ leaves the pseudoidentity $w_n = x^{\omega+1}$ unchanged. Hence it suffices to consider the case where S is a cyclic group, say of order r. Then, for every k relatively prime to r, the fact that k is invertible in the multiplicative monoid of $\mathbb{Z}/r\mathbb{Z}$ means that $S \models x^{k^\omega} = x$. Since, for all sufficiently large n and all $m \geq n$, $p_n \cdots p_m$ is relatively prime to r, this proves the claim. □

Just by itself, the property that the set of all finite words is open in $\overline{\Omega}_X\mathsf{V}$ has some interesting consequences.

Proposition 3.2. *For every pseudovariety V such that the set of all finite words is open in $\overline{\Omega}_X\mathsf{V}$, every infinite pseudoword in $\overline{\Omega}_X\mathsf{V}$ is $\mathcal{J}$-below some pseudoword that is $\mathcal{J}$-maximal among all infinite pseudowords.*

Proof. By Zorn's Lemma, it suffices to show that every $\mathcal{J}$-chain C of infinite pseudowords has an upper bound in the set of infinite pseudowords. We consider C as a net, indexed by itself. Since the space $\overline{\Omega}_X\mathsf{V}$ is compact and the set of finite words is open, the subspace of infinite pseudowords is compact as well, whence there is some convergent subnet $(d_i)_{i\in I}$, given by $d_i = \varphi(i)$, where $\varphi : I \to C$ is an order-preserving mapping which is cofinal in the sense that, for every $c \in C$, there exists $i \in I$ with $c \leq_{\mathcal{J}} \varphi(i)$. Note that cofinality implies that C and $\{d_i : i \in I\}$ have the same upper bounds. Hence, it suffices to show that the limit $d = \lim d_i$ is such an upper bound. Since the subnet $(d_j)_{j\in I,\, j\geq i}$ also converges to d and it consists of elements that are $\mathcal{J}$-above d_i, it follows that indeed $d \geq_{\mathcal{J}} d_i$. □

Let now $w \in \overline{\Omega}_X\mathsf{V}$ be an arbitrary pseudoword. We denote by $F(w)$ the set of all $u \in X^+$ such that $\iota_{\mathsf{V}}(u)$ is a factor of w. We say that $w \in \overline{\Omega}_X\mathsf{V}$ is *uniformly recurrent* if the language $F(w)$ is uniformly recurrent. By [4, Theorem 2.6], if V contains LSl then a pseudoword $w \in \overline{\Omega}_X\mathsf{V}$ is uniformly recurrent if and only if w is $\mathcal{J}$-maximal among all infinite pseudowords. We will use this fact from hereon without further reference.

Proposition 3.3. *Let V be a pseudovariety containing LSl and let $v, w \in \overline{\Omega}_X\mathsf{V}$ be such that $F(v) \subseteq F(w)$, w is uniformly recurrent, and $v \notin X^+$. Then v is $\mathcal{J}$-equivalent to w and, in particular, it is uniformly recurrent.*

Proof. Since LSl contains N, by Theorem 3.1 and Proposition 3.2, there exists some uniformly recurrent pseudoword $y \in \overline{\Omega}_X\mathsf{V}$ such that $y \geq_{\mathcal{J}} v$. Hence $F(y) \subseteq F(v) \subseteq F(w)$. By [4, Corollary 2.10], it follows that $y \mathrel{\mathcal{J}} w$, whence $F(y) = F(w)$. Hence $F(v) = F(w)$, so that v is also uniformly recurrent. Applying [4, Corollary 2.10] again, we now conclude that $v \mathrel{\mathcal{J}} w$. □

4. The rank

Given an ω-word $w \in \Omega_X^\kappa\mathsf{A}$, there are in general many ω-terms that represent it. But, by McCammond's results, there is only one such ω-term in normal form, which we call the *normal form representation* or simply the *normal form* of w.

The *rank* of $w \in \Omega_X^\kappa \mathsf{A}$ is i if the normal form of w is a rank i normal form term or, equivalently, an ω-term of rank i. It is denoted $\operatorname{rank}[w]$. In this section, we introduce structural, rather than combinatorial, notions of rank for elements of $\Omega_X^\kappa \mathsf{A}$ and, more generally, for $\overline{\Omega}_X \mathsf{A}$.

Given $w \in \Omega_X^\kappa \mathsf{A}$, define $r(w)$ to be the supremum of the cardinalities of strict $\mathcal{J}$-chains of idempotents of $\Omega_X^\kappa \mathsf{A}$ which lie $\mathcal{J}$-above w. Similarly, for $w \in \overline{\Omega}_X \mathsf{A}$, define $\bar{r}(w)$ to be the supremum of the cardinalities of strict $\mathcal{J}$-chains of idempotents of $\overline{\Omega}_X \mathsf{A}$ which lie $\mathcal{J}$-above w. Note that the definition of r and $\bar{r}$ immediately implies that the value of any of these functions at a product is at least its value at each of the factors.

Lemma 4.1. *Let α and β be ω-terms in normal form and let γ be the normal form of $\alpha\beta$. Then* $\operatorname{rank}[\gamma] \geq \max\{\operatorname{rank}[\alpha], \operatorname{rank}[\beta]\}$.

Proof. According to the normal form algorithm, the only steps that change the depth of nested parentheses are the application of a rule of type 1 $((\delta)) \mapsto (\delta)$ or of the step that replaces (γ) by γ in case γ is idempotent. Now, such changes on the concatenation of two expressions which represent ω-terms (that is, which are properly parenthesized), can only take place entirely within one of the factors, and this is not possible since α and β are in normal form. □

Lemma 4.2. *Let (α) and (β) be ω-terms in normal form.*

(*a*) *If (α) is a subterm of β then $\epsilon[(\alpha)]$ lies strictly $\mathcal{J}$-above $\epsilon[(\beta)]$ in $\Omega_X^\kappa \mathsf{A}$.*
(*b*) *If $\epsilon[(\alpha)]$ lies strictly $\mathcal{J}$-above $\epsilon[(\beta)]$ in $\Omega_X^\kappa \mathsf{A}$ then* $\operatorname{rank}[\alpha] < \operatorname{rank}[\beta]$.

Proof. (*a*) Since β is in normal form, the hypothesis that (α) is a subterm of β implies that $\operatorname{rank}[(\alpha)] \leq \operatorname{rank}[\beta]$ and so that $\operatorname{rank}[(\alpha)] < \operatorname{rank}[(\beta)]$. Moreover, the same hypothesis yields that $\epsilon[(\alpha)]$ lies $\mathcal{J}$-above $\epsilon[(\beta)]$ in $\Omega_X^\kappa \mathsf{A}$. On the other hand, if $\epsilon[(\beta)]$ lies $\mathcal{J}$-above $\epsilon[(\alpha)]$ in $\Omega_X^\kappa \mathsf{A}$ then, by Lemma 4.1, $\operatorname{rank}[(\beta)] \leq \operatorname{rank}[(\alpha)]$, which contradicts an earlier inequality.

(*b*) Assume that indeed $\epsilon[(\alpha)]$ lies strictly $\mathcal{J}$-above $\epsilon[(\beta)]$. By Lemma 4.1, we have $\operatorname{rank}[(\alpha)] \leq \operatorname{rank}[(\beta)]$. Since (α) and (β) are in normal form, it follows that $\operatorname{rank}[\alpha] \leq \operatorname{rank}[\beta]$. Suppose that $\operatorname{rank}[\alpha] = \operatorname{rank}[\beta]$. By hypothesis, there exist possibly empty ω-terms in normal form γ and δ such that $\epsilon[\gamma(\alpha)\delta] = \epsilon[(\beta)]$. We further assume that $|\gamma\delta|$ is minimal for this property.

By McCammond's results, the normal form of $\gamma(\alpha)\delta$ is (β). Again, the only way that the depth of nested parentheses might go down in the computation of the normal form for $\gamma(\alpha)\delta$ would be within one of the three factors, which is impossible since they are assumed to be in normal form. Hence $\operatorname{rank}[\gamma]$ and $\operatorname{rank}[\delta]$ are at most $\operatorname{rank}[(\alpha)]$. If, next, there are any contractions of type $(\varepsilon)(\varepsilon) \mapsto (\varepsilon)$, then they take place at the junction of either γ and (α) or (α) and δ, and so there is no need for such contractions by the minimality assumption on $|\gamma\delta|$. The remaining phases of the normal form algorithm concern therefore the standardization of subterms of the form (ε) of highest rank and of the crucial portions. The first of these steps produces no changes since all subterms of the form (ε) of highest rank are already in normal form. For the second step, since we know that the end result is (β), the minimality assumption on $|\gamma\delta|$ implies that $\gamma = \delta = 1$, which contradicts the hypothesis that $\epsilon[(\alpha)]$ lies strictly $\mathcal{J}$-above $\epsilon[(\beta)]$ in $\Omega_X^\kappa \mathsf{A}$. □

Lemma 4.3. *Let w be an idempotent in $\Omega_X^\kappa \mathsf{A}$. Then there exists an ω-term (α), in normal form, such that $w \mathrel{\mathcal{J}} \epsilon[(\alpha)]$.*

Proof. Let β be the normal form of w. Then β is the normal form of $\beta\beta$ and so β is an *idempotent in* $\operatorname{rank}[\beta]$ in the sense of [24, Definition 5.9]. By [24, Lemma 5.10], it follows that the normal form expression for β is of the form $\varepsilon_0(\alpha)\varepsilon_1$ and that the expression $(\alpha)\varepsilon_1\varepsilon_0(\alpha)$ is $\operatorname{rank}[\alpha]$-equivalent to (α). Hence $w = \epsilon[\beta] \mathrel{\mathcal{J}} \epsilon[(\alpha)]$. □

Proposition 4.4. *Let $w \in \Omega_X^\kappa \mathsf{A}$ and let α be its normal form. Then* $r(w) = \operatorname{rank}[\alpha]$.

Proof. By definition, the rank of an ω-term in normal form is the largest number of nested parentheses. Hence, for $n = \operatorname{rank}[\alpha]$, there are ω-terms in normal form $(\beta_1), \ldots, (\beta_n)$ such that each (β_i) is a subterm of β_{i+1} $(i < n)$ and of α $(i \leq n)$. By Lemma 4.2(*a*), $\{\epsilon[(\beta_1)], \ldots, \epsilon[(\beta_n)]\}$ constitutes a strict $\mathcal{J}$-chain of idempotents $\mathcal{J}$-above w. Hence $r(w) \geq n = \operatorname{rank}[\alpha]$.

Let $e_1, \ldots, e_m$ be a strictly descending $\mathcal{J}$-chain of idempotents of $\Omega_X^\kappa \mathsf{A}$ which lie $\mathcal{J}$-above w. By Lemma 4.3, we may assume that each e_i has a normal form of the form (β_i). By Lemma 4.2(*b*), since $e_i = \epsilon[(\beta_i)]$ is strictly $\mathcal{J}$-above $e_{i+1} = \epsilon[(\beta_{i+1})]$, we have $\operatorname{rank}[(\beta_i)] < \operatorname{rank}[(\beta_{i+1})]$ $(i < m)$. On the other hand, by Lemma 4.1, since $\epsilon[(\beta_m)]$ is $\mathcal{J}$-above $w = \epsilon[\alpha]$, we have $\operatorname{rank}[(\beta_m)] \leq \operatorname{rank}[\alpha]$. Hence $m \leq \operatorname{rank}[\alpha]$, which shows that $r(w) \leq \operatorname{rank}[\alpha]$. $\square$

Since, by Theorem 2.4, the factors of an element of $\Omega_X^\kappa \mathsf{A}$ are the same in $\Omega_X^\kappa \mathsf{A}$ and in $\overline{\Omega}_X \mathsf{A}$, we also have the following result.

Proposition 4.5. *For every element $w \in \Omega_X^\kappa \mathsf{A}$, we have $\bar{r}(w) = r(w)$.* $\square$

Propositions 4.4 and 4.5 imply that, for an ω-term in normal form α, $\operatorname{rank}[\alpha] = r(\epsilon[\alpha]) = \bar{r}(\epsilon[\alpha])$. For an arbitrary element w of $\overline{\Omega}_X \mathsf{A}$, we also call $\bar{r}(w)$ the *rank* of w.

5. Properties of the Green relations on $\Omega_X^\kappa \mathsf{A}$

In this section, we examine some properties of the Green relations on the semigroup $\Omega_X^\kappa \mathsf{A}$.

Lemma 5.1. *In every regular $\mathcal{J}$-class of $\Omega_X^\kappa \mathsf{A}$ there is a unique element which, in normal form, is of the form (α).*

Proof. Let $\beta = \gamma_0(\delta_1)\gamma_1 \cdots (\delta_r)\gamma_r$ be an ω-term in normal form and suppose that $u = \epsilon[\beta]$ is a regular element of $\Omega_X^\kappa \mathsf{A}$. Then there exists $w \in \Omega_X^\kappa \mathsf{A}$ such that $uwu = u$ and $wuw = w$. Let $u_i = \epsilon[\gamma_i]$ and $v_j = \epsilon[\delta_j]$. By the normal form algorithm, $r = 1$ and $v_1^\omega u_1 \cdot w \cdot u_0 v_1^\omega = v_1^\omega$, so that $v_1^\omega = v_1^\omega u_1 \cdot wuw \cdot u_0 v_1^\omega$. Hence $u \mathrel{\mathcal{J}} v_1^\omega$. Suppose that v^ω is another ω-term in normal form in the $\mathcal{J}$-class of u. Then $\operatorname{rank}[v] = \operatorname{rank}[v_1]$ and there exist $x, y \in \Omega_X^\kappa \mathsf{A}$ such that $v_1^\omega = xv^\omega y$. Hence the normal form of $v_1^\omega x v^\omega y$ is v_1^ω. This means that the application of the normal form algorithm must eliminate the crucial portion $v_1^\omega x v^\omega$, since v_1^ω has no crucial portions. The only step where this can happen is Step (7), which implies that $v_1 = v$. $\square$

Proposition 5.2. *In every regular $\mathcal{R}$-class of $\Omega_X^\kappa \mathsf{A}$ there is a unique element which, in normal form, is an initial portion.*

Proof. Let u be a regular element of $\Omega_X^\kappa \mathsf{A}$. By Lemma 5.1, there is in the $\mathcal{J}$-class of u a unique element e which, in normal form, is of the form (α). By Corollary 2.5, there exists a unique element of the form xe such that $u \mathrel{\mathcal{R}} xe \mathrel{\mathcal{L}} e$. Let $y \in \Omega_X^\kappa \mathsf{A}$ be such that $e = yxe$ and let β and γ be the normal forms of x and y, respectively. Then the normal form algorithm reduces $\gamma\beta(\alpha)$ to (α), which implies that the normal form of $\beta(\alpha)$ is an initial portion.

Suppose next that there is another element in the $\mathcal{R}$-class of u whose normal form is an initial portion $\delta(\varepsilon)$. Let $t = \epsilon[\delta]$ and $f = \epsilon[(\varepsilon)]$. Since $e \mathrel{\mathcal{J}} tf$ and $\delta(\varepsilon)$ is an initial portion, which in particular means that $\operatorname{rank}[\delta] \leq \operatorname{rank}[\varepsilon] = \operatorname{rank}[\delta(\varepsilon)] - 1$, by Proposition 4.4 it follows that $r(f) = r(e)$ and so $f \mathrel{\mathcal{J}} e$ since $f \geq_{\mathcal{J}} e$. By Lemma 5.1, we deduce that $\varepsilon = \alpha$. Hence all terms in the $\mathcal{R}$-class of u which, in normal form, are initial portions, lie in the same $\mathcal{H}$-class. Hence there is a unique such term since the Green relation $\mathcal{H}$ is trivial in $\Omega_X^\kappa \mathsf{A}$ by Corollary 2.5. $\square$

Of course, one has a dual result to Proposition 5.2 concerning $\mathcal{L}$-classes and final portions.

Lemma 5.3. *Let α and (β) be ω-terms in normal form and suppose that $\epsilon[(\beta)] \geq_{\mathcal{J}} \epsilon[\alpha]$ in $\Omega_X^\kappa \mathsf{A}$. Then (β) is a factor of α.*

Proof. Let $w = \epsilon[\alpha]$ and $e = \epsilon[(\beta)]$. We proceed by induction on $\operatorname{rank}[\alpha]$. We assume inductively that the result holds for all ω-terms α' with $\operatorname{rank}[\alpha'] < \operatorname{rank}[\alpha]$. Let x and y be ω-words such that $w = xey$. If $\operatorname{rank}[\alpha] = \operatorname{rank}[(\beta)]$, then the result follows from the normal form algorithm since it does not affect factors of the form (ξ) which are already in normal form. Otherwise, at least one of the factors x and y has the same rank as w. As the other cases are similar and simpler, we consider the case where $r(x) = r(y) = r(w) = \operatorname{rank}[\alpha]$.

Let $(\gamma)\delta$ and $\lambda(\mu)$ be, respectively, the final and initial portions of the normal forms of x and y. Then the normal form of the ω-term $(\gamma)\delta(\beta)\lambda(\mu)$ is a factor of α. Let ε be the normal form of $\delta(\beta)\lambda$. There are two cases to consider.

First, Step (7) of the normal form algorithm may apply to the crucial portion $(\gamma)\delta(\beta)\lambda(\mu)$, in which case $\gamma = \mu$ and $\varepsilon = \gamma^n$ is a finite power of γ, which in turn is a factor of α. Since $\operatorname{rank}[\gamma^n] < \operatorname{rank}[\alpha]$ and $\epsilon[\varepsilon] = \epsilon[\delta(\beta)\lambda]$, by the induction hypothesis we obtain that (β) is a factor of γ^n, and therefore also of α.

Second, Step (7) does not apply to the crucial portion $(\gamma)\varepsilon(\mu)$, and applying the standardization procedure of Step (8), $(\gamma)\varepsilon(\mu)$ reduces to its normal form $(\gamma)\varepsilon'(\mu)$ by expansions and contractions of the forms $(\gamma) \to$

$(\gamma)\gamma$, $(\mu) \to \mu(\mu)$, $(\gamma)\gamma \to (\gamma)$, and $\mu(\mu) \to (\mu)$. It follows that there are some finite exponents k, ℓ such that ε is a factor of the ω-term $\gamma^k \varepsilon' \mu^\ell$, which has strictly smaller rank than α. Hence e is a factor of $\epsilon[\gamma^k \varepsilon' \mu^\ell]$ in $\Omega_X^\kappa \mathsf{A}$. By the induction hypothesis, it follows that (β) is a factor of $\gamma^k \varepsilon' \mu^\ell$ and, therefore, also of at least one of the factors γ, ε' and μ, whence of α. □

Combining Lemmas 5.1 and 5.3, we obtain the following result.

Theorem 5.4. *There are only finitely many regular $\mathcal{J}$-classes $\geq_{\mathcal{J}}$ a given $w \in \Omega_X^\kappa \mathsf{A}$.*

Proof. By Lemma 5.1, all regular $\mathcal{J}$-classes $\geq_{\mathcal{J}}$ a given ω-word w contain elements whose normal forms are of the form (α). By Lemma 5.3, all such elements are factors of the normal form of w. Hence there are only finitely many such $\mathcal{J}$-classes. □

Note that a regular $\mathcal{J}$-class of $\Omega_X^\kappa \mathsf{A}$ may be infinite. For instance, all the elements $a^n b(a^\omega b)^\omega$, with $n \geq 1$ lie in the $\mathcal{J}$-class of the idempotent $(a^\omega b)^\omega$ while they are all distinct by McCammond's Normal Form Theorem, since their given descriptions are already in normal form.

The following result is both a consequence of Theorem 2.6 and a special case of a much more general result [7, Lemma 4.7].

Proposition 5.5. *Let $w \in \overline{\Omega}_X \mathsf{A}$ and suppose that u and v are two prefixes of w. Then one of u and v is a prefix of the other.* □

In the terminology of [15], Proposition 5.5 is expressed by stating that the semigroup $\overline{\Omega}_X \mathsf{A}$ has *unambiguous $\mathcal{R}$-order*. Dually, $\overline{\Omega}_X \mathsf{A}$ also has unambiguous $\mathcal{L}$-order. Taking into account Theorem 2.4, we immediately deduce that the semigroup $\Omega_X^\kappa \mathsf{A}$ also has unambiguous $\mathcal{R}$-order and $\mathcal{L}$-order. More precisely, we have the following result.

Corollary 5.6. *The set of prefixes of a given $w \in \Omega_X^\kappa \mathsf{A}$ is wqo under the prefix order.*

Proof. By Proposition 5.5, we already know that the prefixes of w form a chain. Let $[x_n]_n$ be a sequence of prefixes of w and suppose that each x_{n+1} is a strict prefix of the preceding x_n. In particular, the ranks of the x_n constitute a non-increasing sequence of integers. Therefore, without loss of generality, we may assume that all x_n have the same rank. By the normal form algorithm, we may assume that all x_n have a common prefix u, where $w = uv$, and are of the form $x_n = uy_n$ where the y_n have smaller rank and are prefixes of v. An induction argument then shows that there exist indices $m < n$ such that y_m is a prefix of y_n. Hence x_m is a prefix of x_n, which contradicts the hypothesis. This completes the proof of the corollary. □

6. The prefix order on factors

The following proposition shows that the set of factors of a given $w \in \overline{\Omega}_X \mathsf{A}$ may not be wqo under the prefix order, even under the restriction $\bar{r}(w) < \infty$. It concerns an example which is taken from [1, Chapter 12] and satisfies $\bar{r}(w) = 2$. In rank 1, we show that the existence of an infinite ascending $\mathcal{R}$-chain converging to an idempotent is equivalent to periodicity of the idempotent.

Let $(n_{1,k})_k$ be a strictly increasing sequence of integers such that the sequence $(b^{n_{1,k}} a \cdots b^2 aba)_k$ converges in $\overline{\Omega}_X \mathsf{A}$ and call x_1 its limit. Assuming $(n_{i,k})_k$ is a subsequence of $(n_{i-1,k})_k$ for $i = 2, \ldots, \ell$ such that the sequence $(b^{n_{i,k}} a \cdots b^{i+1} ab^i a)_k$ converges to x_i in $\overline{\Omega}_X \mathsf{A}$ for $i = 1, \ldots, \ell$, then there also exists a subsequence $(n_{\ell+1,k})_k$ of $(n_{\ell,k})_k$ such that $(b^{n_{\ell+1,k}} a \cdots b^{\ell+2} ab^{\ell+1} a)_k$ converges in $\overline{\Omega}_X \mathsf{A}$ and we denote by $x_{\ell+1}$ its limit. This recursive definition yields a sequence $(x_n)_n$. The following proposition states some of its properties.

Proposition 6.1. *The following properties hold for the above sequence $(x_n)_n$:*

(*a*) *$x_{n+1} >_{\mathcal{R}} x_n$ for every n;*
(*b*) *$\lim_n x_n = (b^\omega a)^\omega$ in $\overline{\Omega}_X \mathsf{A}$;*
(*c*) *the only regular $\mathcal{J}$-classes $\mathcal{J}$-above any given x_n are those of b^ω and $(b^\omega a)^\omega$;*
(*d*) *$\bar{r}(x_n) = 2$.*

Proof. Multiplying the terms of the sequence $(b^{n_{i,k}} a \cdots b^{i+1} ab^i a)_k$ on the right by $b^{i-1}a$, we obtain a subsequence of $(b^{n_{i-1,k}} a \cdots b^i ab^{i-1} a)_k$. Hence, taking limits, we conclude that $x_{i-1} = x_i b^{i-1} a$. This proves (*a*) since $ab^{i-1}a$ is a factor of x_{i-1} but not a factor of x_i.

Given a finite aperiodic semigroup S and a continuous homomorphism $\varphi : \overline{\Omega}_X\mathsf{A} \to S$, let N be such that $S \models x^\omega = x^N$. Then, for $n \geq N$ and $m \geq n + N$, we have

$$\varphi(b^m a \cdots b^{n+1} a b^n a) = \varphi\big((b^\omega a)^\omega\big),$$

which proves (b).

Since (d) is an immediate consequence of (c), it remains to prove (c). For this purpose, consider an idempotent e $\mathcal{J}$-above x_n. We claim that the finite factors of e are factors of $b^\omega a b^\omega$. A first remark in this direction is that a^2 is not a factor of x_n, and therefore it cannot be a factor of e. Moreover, the set of finite factors of e is prolongable in both directions, meaning that every finite factor is both a proper prefix and a proper suffix of another finite factor. But the only finite factors of x_n that can be prolonged indefinitely on the right are the finite factors of $b^\omega a b^\omega$: any other finite factor has at least a factor of the form $ab^k a$, which can only be found once as a factor of x_n, within fixed finite distance from its right end. Hence all finite factors of e are factors of $b^\omega a b^\omega$, which proves the claim.

If e does not contain the factor a, then it must be b^ω. Otherwise, by the claim, every finite factor of e is a factor of $b^\omega a b^\omega$, so that e has no finite factor in which the letter a occurs more than once. We claim that e is $\mathcal{J}$-equivalent to $(b^\omega a)^\omega$ in $\overline{\Omega}_X\mathsf{A}$. Let $(u_m)_m$ be a sequence of words converging to e and let S be a finite aperiodic semigroup. Let N be such that $S \models x^\omega = x^N$. Then, since $\lim_m u_m^2 = e$, there exists m_0 such that, for all $m \geq m_0$, none of the words $ab^k a$ is a factor of u_m^2 for $k = 0, 1, \ldots, N-1$. Hence u_m is of the form $b^\alpha a b^{\beta_1} a b^{\beta_2} \cdots a b^{\beta_p} a b^\gamma$, for some α, β_i, γ with $\beta_i \geq N$ and $\alpha + \gamma \geq N$. It follows that $S \models u_m = b^\alpha (ab^N)^p ab^\gamma$. Hence, for all m sufficiently large,

$$S \models e = e^\omega = u_m^\omega = b^\alpha (ab^\omega)^\omega ab^\gamma.$$

Since $b^\alpha (ab^\omega)^\omega ab^\gamma \mathrel{\mathcal{J}} (b^\omega a)^\omega$, this proves (c). $\square$

We now turn to the rank 1 case. We say that $w \in \overline{\Omega}_X\mathsf{V}$ is *periodic* if there is some finite $x \in \overline{\Omega}_X\mathsf{V}$ such that $x >_{\mathcal{J}} w \geq_{\mathcal{J}} x^\omega$. In this case, we also say that x is a *period* of w. We define the set of factors of a pseudoword w as

$$\mathcal{F}(w) = \{v \in \overline{\Omega}_X\mathsf{V} : v \geq_{\mathcal{J}} w\}.$$

Theorem 6.2. *Let* V *be any pseudovariety containing* LSl *and let* $w \in \overline{\Omega}_X\mathsf{V}$ *be uniformly recurrent. Then the following conditions are equivalent:*

(a) *w is periodic;*
(b) *the language $F(w)$ is rational;*
(c) *the set $\mathcal{F}(w)$ is open;*
(d) *w is not the limit of a sequence of words w_n, none of which is a factor of w;*
(e) *there is no infinite strictly ascending $\mathcal{J}$-chain in $\overline{\Omega}_X\mathsf{V}$ converging to w;*
(f) *there is no infinite strictly ascending $\mathcal{J}$-chain $(w_n)_n$ in $\overline{\Omega}_X\mathsf{V}$ converging to w such that, for all n, all idempotents $\mathcal{J}$-above w_n are $\mathcal{J}$-equivalent to w;*
(g) *there is no infinite strictly ascending $\mathcal{R}$-chain in $\overline{\Omega}_X\mathsf{V}$ converging to w;*
(h) *there is no infinite strictly ascending $\mathcal{R}$-chain $(w_n)_n$ in $\overline{\Omega}_X\mathsf{V}$ converging to w such that, for all n, all idempotents $\mathcal{J}$-above w_n are $\mathcal{J}$-equivalent to w;*
(i) *there is no sequence $(u_n)_n$ of words such that each u_n is not a factor of w, its factors of length n are factors of w, and the suffix of length n of u_n is also a suffix of w.*
(j) *there is no sequence $(u_n)_n$ of words such that each u_n is not a factor of w and its factors of length n are factors of w.*

Proof. Since the set of factors of an element of a compact semigroup is always closed, the equivalence of conditions (a), (b), and (c) is the statement of [4, Theorem 2.11] for $\mathsf{V} = \mathsf{S}$ and one can check that the results of [4, Section 2], with the same proofs, apply to any pseudovariety V containing LSl (see also [6, Section 6] for an alternative approach). We proceed to establish the equivalence with the remaining conditions.

Note that, if the set $\mathcal{F}(w)$ is open then, whenever $w_n \to w$, all terms in the sequence with sufficiently large index are factors of w. This trivially shows that $(c)\Rightarrow(d)$ and $(c)\Rightarrow(e)$. The implications $(e)\Rightarrow(f)\Rightarrow(h)$, $(e)\Rightarrow(g)\Rightarrow(h)$, and $(i)\Rightarrow(j)$ are also straightforward.

$(d)\Rightarrow(c)$ Suppose that $(w_n)_n$ is a sequence converging to w in which no term belongs to $\mathcal{F}(w)$. By (d), we may assume that all pseudowords w_n are infinite. If every finite factor of w_n is a factor of w then, since

w is uniformly recurrent, w_n has exactly the same finite factors as w and, by [4, Corollary 2.8], for uniformly recurrent pseudowords this condition implies that w_n and w are $\mathcal{J}$-equivalent and, therefore, $w_n \in \mathcal{F}(w)$, in contradiction with the initial assumption. Hence, for each n, w_n has some finite factor $u_n \notin \mathcal{F}(w)$. By compactness, we may assume that the sequence u_n converges to a limit u; continuity of multiplication yields $u \in \mathcal{F}(w)$. Let $x, y \in \overline{\Omega}_X\mathsf{V}$ be such that $w = xuy$ and let $(x_n)_n$ and $(y_n)_n$ be sequences of words converging respectively to x and y. Since u_n is not a factor of w, neither is $x_n u_n y_n$. By construction and continuity of multiplication again, we have $x_n u_n y_n \to w$, which contradicts (d). Hence there is no such sequence of non-factors of w converging to w which is equivalent to (c).

$(h)\Rightarrow(i)$ Suppose that $(u_n)_n$ is a sequence satisfying the conditions of (i). We show how to construct a sequence satisfying the conditions of (h). Denote by p_n and s_n respectively the prefix and the suffix of u_n of length n. Since w is uniformly recurrent and s_{n+1} and p_n are factors of w, for every n there exists a word v_n such that $s_{n+1}v_n p_n$ is a factor of w, and we may assume that $|v_n| \to \infty$. Consider the sequence $\mathcal{S}_m = (v_n u_n \cdots v_{m+1}u_{m+1}v_m u_m)_{n\geq m}$. There is a strictly increasing sequence of indices $(n_{1,k})_k$ whose corresponding subsequence of $\mathcal{S}_1$ converges, and we call x_1 its limit. There is in turn a subsequence $(n_{2,k})_k$ of $(n_{1,k})_k$ such that the subsequence of $\mathcal{S}_2$ which it determines converges, and we call x_2 its limit. And so on. Recursively, we obtain a sequence $(x_n)_n$ such that, for every $n \geq 1$,

$$x_n = x_{n+1}v_n u_n, \tag{2}$$

so that, in particular, $x_{n+1} \geq_{\mathcal{R}} x_n$. Let x be the limit of a subsequence $(x_{n_k})_k$ for which we may assume that the sequences $(s_{n_k})_k$ and $(v_{n_k})_k$ both converge, to the respective limits s and v. Since s_n is a suffix of both x_n and w, s is a suffix of x and w. Since v_n is a prefix of $v_n u_n \cdots v_{m+1}u_{m+1}v_m u_m$, it follows that v is a prefix of x_m, for all m, hence v is a prefix of x. Since $\overline{\Omega}_X\mathsf{V}$ is stable, s and v are infinite pseudowords, and w is a $\mathcal{J}$-maximal infinite pseudoword, it follows that $s \mathrel{\mathcal{L}} w$, whence $w \geq_{\mathcal{J}} x$.

Let u be a finite factor of x. Then u is a factor of x_n for all sufficiently large n, hence u is a factor of $u_{n+1}v_n u_n$ for arbitrarily large n. It follows that there exists $n \geq |u|$ such that u is a factor of $u_{n+1}v_n u_n$. Hence, u is a factor of at least one of the words u_{n+1}, $s_{n+1}v_n p_n$, or u_n. By hypothesis and the choice of v_n, this implies that u is a factor of w. Since $w \geq_{\mathcal{J}} x$, we conclude that x and w have the same finite factors, so that x is also uniformly recurrent. By [4, Corollary 2.10], it follows that $x \mathrel{\mathcal{J}} w$. Since s is a common infinite suffix, we must have $x \mathrel{\mathcal{L}} w$. Let $t \in \overline{\Omega}_X\mathsf{V}$ be an infinite pseudoword such that $w = tx$ and note that $w \geq_{\mathcal{R}} tx_n$ for all n.

Since u_n is a factor of tx_n but not of w, by stability of $\overline{\Omega}_X\mathsf{V}$ it follows that $w >_{\mathcal{R}} tx_n$ for every n, whence some subsequence of $(tx_{n_k})_k$ is a strictly ascending $\mathcal{R}$-chain converging to w.

To complete the proof of $(h)\Rightarrow(i)$, it suffices to show that every idempotent e that lies $\mathcal{J}$-above tx_1 is $\mathcal{J}$-equivalent to w. Let e be such an idempotent. Then the set of finite factors of e is a prolongable language. We claim that every finite factor u of e is a factor of w. Let $(t_n)_n$ be a sequence of finite words converging to t. Since u is a factor $tx_1 = \lim t_k v_{n_{1,k}} u_{n_{1,k}} \cdots v_1 u_1$ and $|v_n| \to \infty$, we conclude that, for all sufficiently large k, at least one of the following conditions holds:

(1) u is a factor of $t_k v_{n_{1,k}}$;
(2) u is a factor of $v_{n_{1,k}} u_{n_{1,k}} \cdots v_1 u_1$.

Suppose first that there are infinitely many values of k for which condition (1) holds. Then u is also a factor of the limit tv. Since $v \mathrel{\mathcal{R}} x$ and $tx = w$, we have $tv \mathrel{\mathcal{J}} w$, whence u is a factor of w. Otherwise, (2) holds for all sufficiently large values of n. Since u is prolongable on the right to a factor of e, it follows that, if we assume that u is not a factor of w, then, for all n, there exists a word z_n of length n such that uz_n is a factor of some $v_{n_{1,k}} u_{n_{1,k}} \cdots v_1 u_1$. If we take $n \geq |v_{|u|-1}u_{|u|-1} \cdots v_1 u_1|$ then we deduce that u is a factor of $v_{n_{1,k}} u_{n_{1,k}} \cdots v_{|u|}u_{|u|}$. By the hypothesis on the sequence $(u_k)_k$, this implies that u is a factor of some factor of the form u_m or $s_{m+1}v_m p_m$, for some $m \geq |u|$. In both cases, the hypotheses imply that u is a factor of w, contrary to the assumption. This proves the claim that u is a factor of w. By Proposition 3.3, we finally conclude that $e \mathrel{\mathcal{J}} w$, thereby establishing that $(h)\Rightarrow(i)$.

$(j)\Rightarrow(a)$ Let G_n be the *Rauzy graph* of order n of w, whose vertices are the factors of w of length n and for which there is an edge $ax \to xb$, where $a, b \in X$, if axb is a factor (of length $n+1$) of w; we label such an edge with the letter b. The label of a path is the concatenation of the labels of its edges. Note that, since w is uniformly recurrent, G_n is strongly connected for every n. Hence G_n has some cycle C. Let u be the label of C. Then u^k is the label of the cycle obtained by going k times around C. If all such labels were factors of w, then u^ω would also be a factor of w which, by [4, Corollary 2.10], implies that $w \mathcal{J} u^\omega$, whence w is periodic. Hence, if w is not periodic then, for every n, there is some non-trivial path in G_n whose label is a word u_n in the conditions of (j), namely such that u_n is not a factor of w but all its factors of length n are. This completes the proof. □

Note that condition (h) of Theorem 6.2 implies that every uniformly recurrent pseudoword over a pseudovariety V containing LSI is the limit of a strictly ascending $\mathcal{R}$-chain of rank 1 pseudowords.

7. Key properties of factors of ω-words over A

Given an ω-term $\alpha = \gamma_0(\beta_1)\gamma_1 \cdots (\beta_n)\gamma_n$ with $\mathrm{rank}[\gamma_i], \mathrm{rank}[\beta_j] < \mathrm{rank}[\alpha]$, we say that an ω-term of the form $\gamma_0\beta_1^m\gamma_1 \cdots \beta_n^m\gamma_n$ with $m \geq 1$ is an *expansion* of α.

Lemma 7.1. *Suppose that $u, w \in \Omega^\kappa_X\mathsf{A}$ are such that u is a factor of w such that $r(u) < r(w)$ and let α, β be the normal forms respectively of w, u. Then there is a portion ε of α such that β is a factor of an expansion of ε.*

Proof. By hypothesis, there exist $x, y \in (\Omega^\kappa_X\mathsf{A})^1$ such that $w = xuy$. Since multiplication does not raise the rank, at least one of the factors x and y must have the same rank as w. Accordingly, there are three similar cases to consider. We only treat the case in which both x and y have the same rank as w, as the other cases are similar.

Let γ, δ be the normal forms respectively of x, y. Since $w = xuy$, the normal form algorithm reduces $\gamma\beta\delta$ to α. The hypotheses on the ranks of the factors x, u, y imply that the only standardization which may take place is in the product $\gamma'\beta\delta'$, where γ' is the final portion of γ and δ' is the initial portion of δ. Such a standardization will produce either a crucial portion of α or a factor of the form (ε) of maximal rank. Moreover, the maximal rank idempotents in γ' and δ' determine bounding idempotents in the crucial portion of α in question, or they are equal to (ε). Since the normalization depends on applying the algorithm at a lower rank in the section of $\gamma'\beta\delta'$ between those two idempotents and using expansions and contractions of type 4, by replacing those idempotents by some finite powers of their bases, we conclude that β is a factor of the expansion of some portion of α. □

The following marginal observation is an immediate consequence of the normal form algorithm.

Remark 7.2. Suppose that u and v are ω-words of the same rank, whose normal forms have the same number of crucial portions, and trivial initial and final portions. If $u \geq_\mathcal{J} v$ then $u = v$.

Given $w \in \Omega^\kappa_X\mathsf{A}$, recall that $\mathcal{F}(w)$ is the set of all factors of w in $\Omega^\kappa_X\mathsf{A}$. We view $\mathcal{F}(w)$ as a quasi-ordered set under the factor ordering.

Theorem 7.3. *Let $w \in \Omega^\kappa_X\mathsf{A}$. Then $\mathcal{F}(w)$ is wqo.*

Proof. Let $(u_n)_n$ be a sequence of factors of w. We claim that there exist indices m and n such that $m < n$ and $u_m \geq_\mathcal{J} u_n$. Let α and β_n be the normal forms respectively of w and u_n.

Suppose first that there is a subsequence of the given sequence all of whose elements have the same rank as w. Then we may as well assume that, for all n, $r(u_n) = r(w)$. By the normal form algorithm, for each n there are factorizations $\alpha = \gamma_n\delta_n\varepsilon_n$ and $\beta_n = \gamma_n'\delta_n\varepsilon_n'$ where γ_n' is a suffix of γ_n, ε_n' is a prefix of ε_n, and both γ_n' and ε_n' have smaller rank than α. Since there are only finitely many such segments of w, we may as well assume that each of the sequences $\gamma_n, \delta_n, \varepsilon_n$ is constant. It follows that all γ_n' are suffixes of the same element of $\Omega^\kappa_X\mathsf{A}$ and all ε_n' are prefixes of the same element of $\Omega^\kappa_X\mathsf{A}$. By Corollary 5.6 and its dual, by taking subsequences, we may assume that each $\epsilon[\gamma_n']$ is a suffix of $\epsilon[\gamma_{n+1}']$ and each $\epsilon[\varepsilon_n']$ is a prefix of $\epsilon[\varepsilon_{n+1}']$. It follows that $u_n = \epsilon[\gamma_n'\delta_n\varepsilon_n']$ is a factor of u_{n+1}.

Hence we may assume that $r(u_n) < r(w)$ for all n. By Lemma 7.1, each β_n is a factor of an expansion of some portion of α. Since there are only finitely many such portions, we may as well assume that the portion in question is the same for every n. One should now consider several cases according to the type of portion, initial, crucial, or final. Since the other cases are similar but somehwat simpler, we treat only the case where the portion is crucial, say $(\gamma)\delta(\varepsilon)$. For each n, let (k_n, ℓ_n) be a lexicographically minimal pair of non-negative integers such that β_n is a factor of $\gamma^{k_n}\delta\varepsilon^{\ell_n}$. By taking subsequences, we may assume that both sequences $(k_n)_n$ and $(\ell_n)_n$ are non-decreasing.

If $(k_n)_n$ is bounded by some k and $(\ell_n)_n$ is bounded by some ℓ, then the β_n are factors of an element of $\Omega_X^\kappa \mathsf{A}$ of rank smaller than $r(w)$, namely $\gamma^k\delta\varepsilon^\ell$. An induction argument then shows that there are indices $m < n$ such that u_m is a factor of u_n, as desired.

There remain three cases to consider, depending on the boundedness of the sequences $(k_n)_n$ and $(\ell_n)_n$. We treat here only the case where $(k_n)_n$ is bounded and $(\ell_n)_n$ is not. By taking subsequences, we may then assume that $k_n = k$ is constant and that $(\ell_n)_n$ is strictly increasing. Then β_n is a factor of $\gamma^k\delta\varepsilon^{\ell_n}$ but not a factor of $\gamma^k\delta\varepsilon^{\ell_n-1}$. We distinguish two subcases, according to whether k is zero or positive.

Assume first that $k = 0$. By the normal form algorithm, there is a factorization $\beta_n = \beta'_n\varepsilon^{m_n}\beta''_n$, where β'_n is a suffix of $\delta\varepsilon$, β''_n is a prefix of ε, and $m_n < \ell_n$. By Corollary 5.6 and its dual, we may assume that each β'_n is a suffix of β'_{n+1}, each β''_n is a prefix of β''_{n+1}, and $m_n \leq m_{n+1}$. But then β_n is a factor of β_{n+1} and, therefore, u_n is a factor of u_{n+1}. The remaining cases are treated similarly. □

Given $w \in \overline{\Omega}_X\mathsf{V}$, denote by $F_\mathsf{V}(w)$ the set of all $v \in X^+$ such that $v >_\mathcal{J} w$. For $L \subseteq X^+$, denote by $\mathrm{cl}_\mathsf{V}(L)$ the closure of L in $\overline{\Omega}_X\mathsf{V}$.

For a pseudovariety V and a fixed alphabet X, denote by p_V the natural continuous homomorphism $\overline{\Omega}_X\mathsf{S} \to \overline{\Omega}_X\mathsf{V}$ determined by sending each generator $x \in X$ to itself.

Lemma 7.4. *If* V *contains* LSI *and* $w \in \overline{\Omega}_X\mathsf{S}$*, then* $F_\mathsf{V}(p_\mathsf{V}(w)) = F_\mathsf{LSI}(p_\mathsf{LSI}(w))$.

Proof. Note that, if $\mathsf{W} \subseteq \mathsf{V}$ then $F_\mathsf{V}(p_\mathsf{V}(w)) \subseteq F_\mathsf{W}(p_\mathsf{W}(w))$. Hence $F_\mathsf{V}(p_\mathsf{V}(w)) \subseteq F_\mathsf{LSI}(p_\mathsf{LSI}(w))$. For the converse, consider a sequence of words $w_n \in X^+$ converging to w and suppose that $v \in F_\mathsf{LSI}(p_\mathsf{LSI}(w))$. Since the set $\mathrm{cl}_\mathsf{LSI}(X^*vX^*)$ is open in $\overline{\Omega}_X\mathsf{LSI}$, we may as well assume that all words w_n belong to $\mathrm{cl}_\mathsf{LSI}(X^*vX^*) \cap X^+ = X^*vX^*$. Hence, we may factorize w_n as $w_n = x_nvy_n$, with $x_n, y_n \in X^*$. By taking subsequences, we may further assume that $x_n \to x$ and $y_n \to y$ in $(\overline{\Omega}_X\mathsf{V})^1$. Then $p_\mathsf{V}(w) = xvy$, which shows that $v \in F_\mathsf{V}(w)$. □

Lemma 7.5. *If* $w \in \Omega_X^\kappa\mathsf{A}$ *then* $F_\mathsf{A}(w)$ *is a rational language.*

Proof. By Lemma 7.4, it suffices to show that $F_\mathsf{LSI}(w)$ is rational whenever $w \in \Omega_X^\kappa\mathsf{LSI}$. Note that, if $v \in \Omega_X^\kappa\mathsf{A}$ has rank at least 1, then $\mathsf{LSI} \models v^\omega = v^2$ since two elements of $\overline{\Omega}_X\mathsf{LSI}$ are equal if and only if they have the same finite factors, the same finite prefixes, and the same finite suffixes, which is the case for v^ω and v^2 since v is infinite. By induction on the rank, it follows that w is either a finite word, for which the result is obvious, or it is given by an ω-term of rank 1, say $w = u_0v_1^\omega u_1 \cdots v_r^\omega u_r$, where the $u_i, v_j \in X^+$.

Let $w_n = u_0v_1^nu_1 \cdots v_r^nu_r$. Then a word $v \in X^+$ is a factor of w if and only if v is a factor of w_n for all sufficiently large values of n. It follows that the finite factors of w are precisely the finite factors of a subterm of one of the forms $u_0v_1^\omega$, $v_i^\omega u_iv_{i+1}^\omega$, or $v_r^\omega u_r$. Now, for instance, $F_\mathsf{LSI}(v_i^\omega u_iv_{i+1}^\omega)$ is the set of all factors of the rational language $v_i^*u_iv_{i+1}^*$ and, therefore, it is rational. The other cases are similar. □

From hereon, for a pseudovariety V containing LSI and $w \in \overline{\Omega}_X\mathsf{S}$, we will write $F(w)$ or $F(p_\mathsf{V}(w))$ for $F_\mathsf{S}(w)$.

Corollary 7.6. *If* $w \in \Omega_X^\kappa\mathsf{A}$ *then the following conditions hold:*
(a) *for every* $v \in \mathcal{F}(w)$*,* $F(v)$ *is a rational language;*
(b) $\mathcal{F}(w)$ *is wqo.*

Proof. (a) follows from Theorem 2.4 and Lemma 7.5, while (b) is Theorem 7.3. □

8. Slim pseudowords over A

For a non-negative integer n, we say that a pseudoword $w \in \overline{\Omega}_X\mathsf{A}$ is *n-slim* if it has no infinite anti-chains of factors of rank at most n and the language $F(w) \subseteq X^+$ is rational. We say that w is *slim* if it is slim for all $n \geq 0$. For example, since every uniformly recurrent pseudoword by its own definition has no infinite anti-chains of finite factors, and all its infinite factors are $\mathcal{J}$-equivalent (cf. [4]), in view of Theorem 6.2 it is slim if and only if it is periodic. Thus, for a uniformly recurrent pseudoword, slimness is equivalent to being an ω-word. More generally, by Corollary 7.6, every ω-word is slim.

This section is dedicated to establishing some useful features of pseudowords that enjoy one or both defining properties of slim pseudowords. We start with a further preliminary observation involving uniformly recurrent pseudowords.

Lemma 8.1. *Let* V *be a pseudovariety containing* LSl *and let* $w_1, \dots, w_n \in \overline{\Omega}_X\mathsf{V}$ *be uniformly recurrent pseudowords. If* $\bigcup_{i=1}^n F(w_i)$ *is a rational language then every* w_i *is periodic.*

Proof. By the pumping lemma and the fact that each $F(w_i)$ is an infinite factorial language, at least one of the languages $F(w_i)$ must contain all words of the form u^k for a fixed word u. Without loss of generality, we may assume that $i = 1$. Since w_1 is $\mathcal{J}$-maximal as an infinite pseudoword, we deduce that w_1 is periodic. Note also that, by a compactness argument and Proposition 3.3, if $x, y \in \overline{\Omega}_X\mathsf{V}$ are uniformly recurrent but not $\mathcal{J}$-equivalent, then $F(x) \cap F(y)$ is a finite language. It follows that $\bigcup_{i=2}^n F(w_i)$ is still a rational language. An induction argument completes the proof of the lemma. □

In analogy with the definition of reduced rank 1 crucial portions of ω-terms, we call a *bridge factorization* of a pseudoword from $\overline{\Omega}_X\mathsf{A}$ a factorization of the form xyz such that:

(a) x and z are uniformly recurrent;
(b) y is a finite word;
(c) y is a factor of neither x nor z;
(d) if $y = av$ and $xa \mathrel{\mathcal{R}} x$, then v is a factor of at least one of x and z;
(e) if $y = vb$ and $bz \mathrel{\mathcal{L}} z$, then v is a factor of at least one of x and z.
(f) y has minimum length among all elements y' for which there exist x', z' such that the preceding conditions hold for the triple (x', y', z') and $x'y'z' = xyz$.

We then also say that the product xyz is a *bridge*. The word y is called a *middle* of the bridge.

Lemma 8.2. *Let* $x, z \in \overline{\Omega}_X\mathsf{A}$ *be uniformly recurrent and* $y \in X^*$*. Then the product* xyz *is either uniformly recurrent or it is a bridge.*

Proof. Suppose that xyz is not uniformly recurrent. Since it is infinite, it follows from Proposition 3.3 that there is some finite factor u of xyz that is not a factor of x. Either u overlaps with y or it is a factor of z (cf. [12, Lemma 8.2]). In either case, we conclude that there exist factorizations $x = x_1x_2$ and $z = z_1z_2$ such that x_2yz_1 is a finite word that is not a factor of x. By symmetry and by further extending the suffix x_2 of x and the prefix z_1 of z, we may assume that x_2yz_1 is a finite word that is a factor of neither x nor z, so that the factorization $x_1 \cdot x_2yz_1 \cdot z_2$ satisfies conditions (a)–(c). Conditions (d) and (e) are obtained by first passing to x_1 all the beginning letters of x_2yz_1 which fail (d) and then proceeding dually with suffix letters. Observe that in this procedure part or all of y may be removed. The resulting factorization $x'y'z'$ of xyz satisfies conditions (a)–(e), and therefore there is such a factorization in which the length of y' is minimum, that is a bridge factorization of xyz. □

We give a couple of examples to further help to understand our notion of bridge. Without condition (f), we could have two middles of the same bridge of different lengths. For instance, the two factorizations $a^\omega \cdot ab \cdot (cdbc)^\omega = a^\omega \cdot bcd \cdot bc(cdbc)^\omega$ both satisfy conditions (a)–(e). According to our definition of bridge, only ab is a middle of the bridge. On the other hand, further assuming condition (f), as in our definition of bridge, even the middles of bridges may not be unique as the following two bridge factorizations show: $(abc)^\omega ab \cdot cb \cdot (abc)^\omega = (abc)^\omega \cdot ba \cdot bc(abc)^\omega$.

Lemma 8.3. *If* $w \in \overline{\Omega}_X\mathsf{A}$ *is a bridge then* $\bar{r}(w) = 1$.

Proof. Let $w = xyz$ be a bridge factorization and suppose that e is an idempotent factor of w. Let $p, q \in (\overline{\Omega}_X\mathsf{A})^1$ be such that $w = peeeq$. Then, by equidivisibility (cf. Theorem 2.6), e must be a factor of at least one of x and z. Since x and z are uniformly recurrent, whence $\mathcal{J}$-maximal as infinite pseudowords, there are no idempotents strictly $\mathcal{J}$-above e, which proves that $\bar{r}(w) = 1$. □

The previous result allows us to show that relatively weak hypotheses force the finiteness of the number bridge factors of a pseudoword.

Lemma 8.4. *Suppose that* $w \in \overline{\Omega}_X\mathsf{A}$ *has no infinite anti-chains of factors of rank at most 1. Then there are only finitely many middles of bridge factors of* w.

Proof. Let tuv and xyz be two bridge factorizations and suppose that $xyz = ptuvq$ for some $p, q \in (\overline{\Omega}_X\mathsf{A})^1$. Since $\overline{\Omega}_X\mathsf{A}$ is equidivisible by Theorem 2.6, the two factorizations have a common refinement. We claim that in such refinements, u and y must overlap. Indeed, otherwise, one of the following two symmetric cases must hold: ptu is a prefix of x or uvq is a suffix of z. In the first case, since t is infinite and x is uniformly recurrent, we have $u \geq_{\mathcal{J}} x \mathrel{\mathcal{J}} t$, which contradicts the hypothesis that tuv is a bridge factorization. The symmetric case is similarly shown to be impossible.

By symmetry, we may as well assume that there are $t', y_1, y_2 \in (\overline{\Omega}_X\mathsf{A})^1$ such that $x = pt'$, $t = t'y_1$, $y = y_1y_2$, and $y_2z = uvq$. Since y is finite, so is y_1 and, therefore, since t is infinite, so is t'. As pt' is a factor of the uniformly recurrent pseudowords x and t, we deduce that x, t, t' are all $\mathcal{J}$-equivalent. Considering the equality $y_2z = uvq$, we conclude similarly that z, v, vq are all $\mathcal{J}$-equivalent. It follows that xyz and $pt \cdot u \cdot vq$ are two bridge factorizations of the same pseudoword. By condition (f) of the definition of bridge, we deduce that y and u have the same length.

Now, if there are infinitely many middles of bridge factors of w, then there is a sequence of such bridges whose middles have strictly increasing length and the above shows that no element in such a sequence is a factor of another. Since we assumed that there are no infinite anti-chains of factors of w of rank at most 1, in view of Lemma 8.3 we conclude that there are only finitely many middles of bridge factors of w. $\square$

We next prove some properties of general factors of 1-slim pseudowords.

Proposition 8.5. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be a 1-slim pseudoword. Then the following conditions hold.*

- (a) *If w' is a factor of w then every element of $F(w')$ is a factor of at least one of the following: a finite prefix of w', a finite suffix of w', a uniformly recurrent factor of w', or a bridge factor of w'.*
- (b) *Every uniformly recurrent factor of w is periodic and the number of $\mathcal{J}$-classes of such factors is finite.*
- (c) *If u is an infinite factor of w then $p_{\mathsf{K}}(u)$ and $p_{\mathsf{D}}(u)$ are ultimately periodic.*

Proof. Suppose that u is an infinite factor of w. Since $F(p_{\mathsf{K}}(u)) \subseteq F(u) \subseteq \mathcal{F}(w)$, by Lemma 2.1 there is a factorization $p_{\mathsf{K}}(u) = u_1\mathbf{s}$, where u_1 is a finite word and $\mathbf{s}$ is a uniformly recurrent right infinite word. Let p_n be the prefix of length n of $\mathbf{s}$. Then $u = u_1p_nr_n$ for some pseudoword r_n. Let (u_2, u_3) be an accumulation point in $\overline{\Omega}_X\mathsf{A} \times \overline{\Omega}_X\mathsf{A}$ of the sequence $\big((p_n, r_n)\big)_n$. Then $F(u_2) = F(\mathbf{s})$ and so u_2 is uniformly recurrent. By continuity of multiplication, we conclude that there is a factorization $u = u_1u_2u_3$, where u_1 is finite and u_2 is uniformly recurrent. Dually one could require instead of u_1 being finite that u_3 be finite.

Suppose next that uv is a factor of w. If u and v are both infinite then, by the preceding paragraph there are factorizations $u = u_1u_2u_3$ and $v = v_1v_2v_3$ such that u_2 and v_2 are uniformly recurrent and $u_3, v_1 \in X^*$. By Lemma 8.2, the product $u_2u_3v_1v_2$ is either uniformly recurrent or it is a bridge.

Given a factorization xyz of some $w' \in \mathcal{F}(w)$ with $y \in X^+$, if one of x or z is finite then y is respectively a factor of $p_{\mathsf{K}}(w')$ or $p_{\mathsf{D}}(w')$, while otherwise we may apply the preceding paragraph say to the factor $x \cdot yz$ of w. It follows that the finite factors of $w' \in \mathcal{F}(w)$ are those of $p_{\mathsf{K}}(w')$, of $p_{\mathsf{D}}(w')$, and of the uniformly recurrent and bridge factors of w', from which (a) follows.

Since there are only finitely many middles of bridge factors of w by Lemma 8.4 and only finitely many $\mathcal{J}$-classes of uniformly recurrent factors of w, as they are all $\leq_{\mathcal{J}}$-incomparable rank 1 pseudowords and w is 1-slim (which includes the rationality of the language $F(w)$), and also taking into account Lemma 2.1 and its dual, we deduce that the (finitary) union of the languages of factors of the uniformly recurrent factors of w is a rational language. By Lemma 8.1, the uniformly recurrent factors of w must be periodic, which proves (b).

In the notation of the first paragraph, we conclude that u_2 is periodic and, therefore, so is $\mathbf{s}$. This proves (c) and completes the proof of the proposition. $\square$

The following is an immediate application of Proposition 8.5.

Corollary 8.6. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be an infinite 1-slim pseudoword. Then the set of rank 1 idempotent factors of w is finite and nonempty, and consists of ω-words of the form u^ω, where $u \in X^+$.*

Proof. It has already been observed that the rank 1 idempotents are the uniformly recurrent idempotents and whence periodic by Proposition 8.5. To complete the proof, it remains to observe that the $\mathcal{J}$-class of a pseudoword of the form u^ω with $u \in X^+$ contains only finitely idempotents. Namely, they are of the form v^ω, where v is an arbitrary conjugate of u. $\square$

We are now ready to prove the following result.

Theorem 8.7. *For $n \geq 1$, every factor of an (n-)slim pseudoword is also (n-)slim.*

Proof. Let $w \in \overline{\Omega}_X\mathsf{A}$ be an n-slim pseudoword and let w' be a factor of w. Since, in case w is finite, the result is obvious, we assume that w is infinite. By Proposition 8.5, the finite factors of w' are factors of factors of w of the form u^ω with $u \in X^+$, of which there are only finitely many by Corollary 8.6, or of factors of one of the forms pu^ω, $u^\omega zt^\omega$, or $u^\omega s$, with $p, s, t, u, z \in X^+$, and where p and s are fixed and we may restrict the choices of z to a finite set of words. It follows that $F(w')$ is a rational language, which shows that w' is n-slim. $\square$

The following couple of examples describe some non-slim pseudowords.

Example 8.8. Let w be an accumulation point of the infinite expression

$$(((ab)^\omega ab^2)^\omega ab^3)^\omega \cdots,$$

meaning an accumulation point of the sequence of truncated expressions

$$(\cdots((ab)^\omega ab^2)^\omega \cdots ab^n)^\omega.$$

We claim that the rank 1 idempotent factors of w are the following: $(ab)^\omega$, $(ba)^\omega$, b^ω. Indeed, every finite factor of w is a factor of some pseudoword of the form $(ab)^\omega ab^2ab^3 \cdots ab^n(ab)^\omega$. Hence no uniformly recurrent factor can have a factor of any the forms ab^k and b^ka with $k > 1$. This implies that the only possibilities for the sets of finite factors of idempotent uniformly recurrent factors of w are the finite factors of $(ab)^\omega$ and b^ω.

Note also that $F(w)$ is not rational by the Pumping Lemma since it contains arbitrarily long words of the form $ab^2ab^3 \cdots ab^na$, none of which admits a factorization of the form xyz such that $xy^mz \in F(w)$ with $|x|, |y|, |z|$ bounded. Since the pseudowords of the form $(ab)^\omega ab^2ab^3 \cdots ab^n(ab)^\omega$ constitute an anti-chain of factors of w, we deduce that w fails both conditions for being 1-slim. $\square$

Example 8.9. Let w be an accumulation point of the infinite expression

$$(((a^\omega b)^\omega a^\omega b^2)^\omega a^\omega b^3)^\omega \cdots.$$

Then the factors of the form $a^\omega b^n a^\omega$ constitute an anti-chain. Note that the set $F(w) = a^*b^*a^*$ is rational.

We claim that, like in the preceding example, there are only finitely many rank 1 idempotent factors of w. More precisely, we claim that the only rank 1 idempotent factors of w are a^ω and b^ω. Indeed, all finite factors are factors of a pseudoword of the form $a^\omega b^n a^\omega$. Hence there are no finite factors of the form ba^rb with $r > 0$ and so a uniformly recurrent factor cannot contain both a and b as factors, whence it must be either a^ω or b^ω. On the other hand, while a^ω is obviously a factor, being even a prefix, b^ω is a factor because every finite power of b is a factor. $\square$

We say that $v \in \mathcal{F}(w)$ is a *special ω-factor* of w if $v = u^\omega$ for some Lyndon word u. In this case, the Lyndon word u is called a *special base* of w. A factor v of w is called a *special factor* if it is of the form $u_1^\omega z u_2^\omega$, where $z \in X^+$ and u_1^ω, u_2^ω are special ω-factors and v is not itself a special ω-factor. Note that every special factor is a bridge. We say that a prefix v of w is *special* if v is of the form zu^ω, where $z \in X^+$ and u^ω is a special ω-factor. The definition of *special suffix* of w is dual.

Corollary 8.10. *Every infinite 1-slim pseudoword has at least one special ω-factor.*

Proof. This is an immediate consequence of Corollary 8.6. $\square$

For special factors, we have the following result.

Lemma 8.11. *If w is 1-slim, then it has only finitely many special ω-factors, finitely many special factors, and precisely one special prefix and one special suffix.*

Proof. We claim that two special factors are either equal or $\mathcal{J}$-incomparable. By the hypothesis that $\mathcal{F}(w)$ has no infinite $\mathcal{J}$-anti-chains, it follows that w has only finitely many special factors. To prove the claim, let $u_1^\omega z u_2^\omega$ and $v_1^\omega t v_2^\omega$ be two special factors and suppose that the latter is a factor the former. Without loss of generality, we may assume that the given expressions are actually the normal forms of these ω-words. Let $x, y \in (\overline{\Omega}_X\mathsf{A})^1$ be such that $u_1^\omega z u_2^\omega = x v_1^\omega t v_2^\omega y$. Then, by Theorem 2.4, we may assume that x, y are rank 1 ω-words. Then, by the normal form algorithm, we must have $(u_1)z(u_2) = (v_1)t(v_2)$, since $(v_1)t(v_2)$ is a crucial portion in normal form of the normal form of the product $xv_1^\omega t v_2^\omega y$, which proves the claim. The proof that w has only finitely many special ω-factors is similar.

That w has some special prefix follows from Proposition 8.5(c). Suppose that xy^ω and uv^ω are both special prefixes of w. Then the right infinite words $xyyy\cdots$ and $uvvv\cdots$ coincide. It is an exercise in combinatorics on words to show that, for two Lyndon words such that the square of each is a factor of a power of the other, the two words coincide. Hence $v = y$ and, because a Lyndon word is lexicographically minimum in its conjugacy class, we must have $xy^m = uv^n$ for some $m, n \geq 0$, which implies that $xy^\omega = uv^\omega$. The argument for suffixes is dual. □

We say that an infinite 1-slim pseudoword is *periodic at the ends* if its special prefix and suffix are special ω-factors.

Corollary 8.12. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be 1-slim. Then there is a factorization $w = xw'y$, where $x, y \in X^*$ and $w' \in \overline{\Omega}_X\mathsf{A}$ is periodic at the ends.*

Proof. Let xu^ω and $v^\omega y$ be respectively the special prefix and suffix of w. Then there are factorizations $w = xu^\omega s = tv^\omega y$. By equidivisibility (cf. Theorem 2.6), and since powers of u and v can only overlap if they do so in a synchronized way, in the sense explained in the proof of Lemma 8.11, taking into account that u^ω is an idempotent, we conclude that there is a factorization of the form $w = xu^\omega zv^\omega y$. Set $w' = u^\omega zv^\omega$ to obtain, by Theorem 8.7, a 1-slim pseudoword which is periodic at the ends and therefore satisfies the requirements of the corollary. □

Proposition 8.5 (a) may be reformulated in terms of special factors as follows.

Lemma 8.13. *Let the pseudoword $w \in \overline{\Omega}_X\mathsf{A}$ be 1-slim and periodic at the ends. Then every finite factor of w is a factor of some special factor or of some special ω-factor of w.*

Proof. Taking into account Proposition 8.5(a) and (b), it suffices to observe that every bridge factor of w is a factor of some special factor of w. Suppose then that u is a bridge factor of w. By Proposition 8.5(b), u is of the form $x^\omega yz^\omega$, with $x, y, z \in X^+$ and x and z primitive, although this is not necessarily a bridge factorization. By the normal form algorithm and since u is not uniformly recurrent, the normal form of u is of the form $\alpha(\beta)\gamma(\delta)\varepsilon$ where β is a conjugate of x, α a suffix of β, δ a conjugate of z, and ε a prefix of δ. It follows that u is a factor of $\beta^\omega\gamma\delta^\omega$, which establishes our claim. □

In the absence of special factors, we can say much more.

Lemma 8.14. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be 1-slim and periodic at the ends. If w has no special factor, then $w = u^\omega$ for some Lyndon word $u \in X^+$.*

Proof. Since w is periodic at the ends, it has a special prefix, whence there is a factorization $w = u^\omega v$ where u is a Lyndon word. Suppose that $w \neq u^\omega$. Let $(v_n)_n$ be a sequence of finite words converging to v and let $v_n = u^{i_n}v'_n$, where $i_n \geq 0$ is maximum. By compactness, we may assume that the sequence $(v'_n)_n$ converges to some pseudoword v' that does not admit u as a prefix. Then we also have $w = u^\omega v'$ and so we may as well assume that u is not a prefix of v. Since the special suffix of w is also periodic, v must be infinite.

By Proposition 8.5(c), v has a factorization of the form $v = xy^\omega z$, where xy^ω is the special prefix. Since u is not a prefix of xy^ω and u and y are Lyndon words, the factor $u^\omega xy^\omega$ of w is a special factor, which contradicts the hypothesis that w has no such factors. Hence we must have $w = u^\omega$. □

9. Encoding 1-slim pseudowords

Let $w \in \overline{\Omega}_X\mathsf{A}$. We denote by $\Gamma(w)$ the directed multigraph whose vertices are the special ω-factors of w and which has an edge $ezf : e \to f$ if $z \in X^+$ is such that ezf is a special factor of w. By Lemma 8.11, $\Gamma(w)$ is finite if w is an infinite 1-slim pseudoword.

In this section, we show how to encode 1-slim pseudowords w as "pseudopaths" in their graphs $\Gamma(w)$. For this purpose, we need to introduce pseudovarieties of categories and their relatively free profinite associated categories, which we do as succintly as possible. See [31] for the definition of pseudovarieties of categories and their relevance for the theory of pseudovarieties of finite semigroups. See also [6] for caveats and pitfalls that one should be aware of when handling pseudovarieties of categories, particularly when profinite techniques are involved.

We view small categories as directed graphs with an associative partial composition of edges which has local identities. We say that a functor $f : C \to D$ is a *quotient* if it is surjective and its restriction to vertices is injective. We say that the functor $f : C \to D$ is *faithful* if its restriction to each hom-set $C(p, q)$ is injective. A category C *divides* a category D if there is a pair of functors $f : E \to C$ and $g : E \to D$ such that f is a quotient and g is faithful. A *pseudovariety of categories* is a class of finite categories that contains the trivial, one-vertex one-edge, category and is closed under taking finite products and divisors.

The hom-set $C(p, p)$ of loops at a vertex p of a category C forms a monoid, called the *local monoid at* p. Conversely, a monoid is viewed as a category by adding a virtual vertex and viewing the elements of the monoid as loops at that vertex, which are composed as they are multiplied in the monoid. Given a pseudovariety V of semigroups, the class of all finite categories which divide some monoid in V is a pseudovariety of categories which is denoted gV. The class of all finite categories whose local monoids belong to V is also a pseudovariety of categories which is denoted $\ell\mathsf{V}$.

By a *congruence* on a category C we mean an equivalence relation $\equiv$ on its set of edges which only identifies edges with the same ends and such that, if $u \equiv u'$ and $v \equiv v'$ and the composite uv is defined, then $uv \equiv u'v'$. Then $C/\equiv$ is the quotient category, with the same vertex set, and edges the $\equiv$-classes of edges of C.

Given a finite graph Γ, the *category freely generated by* Γ is denoted Γ^* and has the same set of vertices as Γ while its edges are the finite paths in Γ. Given a pseudovariety W of categories, by a *pro-*W *category* we mean a finite-vertex compact category C whose edges are separated by continuous quotient functors into members of W. The congruences $\equiv$ on Γ^* for which the quotient $\Gamma^*/\equiv$ belongs to W generate a uniform structure on Γ^*. The completion of Γ^* with respect to this uniform structure is the *pro-*W *category freely generated by* Γ and is denoted $\overline{\Omega}_\Gamma\mathsf{W}$. The subcategory of $\overline{\Omega}_\Gamma\mathsf{W}$ generated by Γ is denoted $\Omega_\Gamma\mathsf{W}$. Note that the ω-power $x \mapsto x^\omega$ is a well-defined operation on loops of $\overline{\Omega}_\Gamma\mathsf{W}$. The subcategory closed under this operation generated by Γ is denoted $\Omega^\kappa_\Gamma\mathsf{W}$.

Let ι_W denote the natural mapping $\Gamma \to \overline{\Omega}_\Gamma\mathsf{W}$. In analogy with the terminology for the semigroup case, and taking into account that the elements of Γ^* are effectively viewed as paths, the elements of $\overline{\Omega}_\Gamma\mathsf{W}$ are called *pseudopaths* while those of $\overline{\Omega}_\Gamma\mathsf{W} \setminus \Gamma^*$ are said to be *infinite pseudopaths*. Elements of $\Omega^\kappa_\Gamma\mathsf{W}$ may be called *ω-paths*.

A pseudovariety of semigroups V is said to be *monoidal* if it is generated by (its) monoids. Equivalently, $S \in \mathsf{V}$ implies $S^1 \in \mathsf{V}$ and so, certainly A is monoidal. For a monoidal pseudovariety V, the unique continuous homomorphism

$$\gamma_\mathsf{V} : \overline{\Omega}_\Gamma\mathsf{gV} \to (\overline{\Omega}_{E(\Gamma)}\mathsf{V})^1$$

that sends each edge to itself (and all vertices to the virtual vertex) is faithful [2].

In case $w \in \overline{\Omega}_X\mathsf{A}$ is infinite and 1-slim, we also consider the unique continuous homomorphism

$$\lambda_w : \overline{\Omega}_{\Gamma(w)}\mathsf{gA} \to \overline{\Omega}_X\mathsf{A}$$

that sends each edge $ezf : e \to f$ to the pseudoword ezf.

Given a semigroupoid S, that is a category without the requirement for local identities, denote by S^c the category which is obtained from S by adding the missing local identities. Recall that

$$\mathsf{Sl} * \mathsf{D}_1 = [\![xyxzx = xzxyx, x^3 = x^2, (xy)^2x = xyx]\!]$$

is the pseudovariety of two-testable semigroups, which is generated by the languages whose membership is characterized by the first letter, the last letter and the two-letter factors (cf. [1, Chapter 10]). It is also the pseudovariety generated by the real matrix semigroup

$$A_2 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}$$

under matrix multiplication.

The following lemma can probably be considered folklore.

Lemma 9.1. *Let Γ be a finite graph and V a pseudovariety containing A_2. Suppose that $v \in \overline{\Omega}_\Gamma\mathsf{gV}$ and $u \in \overline{\Omega}_{E(\Gamma)}\mathsf{V}$ are such that u is a factor of $\gamma_\mathsf{V}(v)$. Then there exists $u' \in \overline{\Omega}_\Gamma\mathsf{gV}$ such that $\gamma_\mathsf{V}(u') = u$ and u' is a factor of v.*

Proof. Suppose that $w \in \overline{\Omega}_{E(\Gamma)}\mathsf{V}$ is such that all its two-letter factors can be read in the graph Γ. In case $w \in \big(E(\Gamma)\big)^+$, w itself can be read in the graph Γ as a path w' such that $\gamma_\mathsf{V}(w') = w$. In general, let $(w_n)_n$

be a sequence in $\bigl(E(\Gamma)\bigr)^+$ converging to w. The hypothesis on V implies that we may assume that all w_n have the same two-letter factors as w. Hence, for each n there exists a path w'_n in Γ such that $\gamma_{\mathsf{V}}(w'_n) = w_n$. If $w' \in \overline{\Omega}_\Gamma \mathsf{gV}$ is an accumulation point of the sequence $(w'_n)_n$, then $w = \gamma_{\mathsf{V}}(w')$ by continuity of γ_{V}.

Now, assuming that $v \in \overline{\Omega}_\Gamma \mathsf{gV}$, $u \in \overline{\Omega}_{E(\Gamma)}\mathsf{V}$, and $\gamma_{\mathsf{V}}(v) = xuy$, then the two-letter factors of each of x, u, y can be read in Γ. By the above, it follows that there exist $x', u', y' \in (\overline{\Omega}_\Gamma \mathsf{gV})^c$ such that $\gamma_{\mathsf{V}}(x') = x$, $\gamma_{\mathsf{V}}(u') = u$, and $\gamma_{\mathsf{V}}(y') = y$. In case $x \neq 1$, then the last letter of x and the first letter of u are consecutive edges in Γ, since their product is a two-letter factor of $\gamma_{\mathsf{V}}(v)$. Hence $x'u'$ is an edge in $\overline{\Omega}_\Gamma \mathsf{gV}$ and, similarly, so is $u'y'$, and, therefore, also $x'u'y'$. Since $\gamma_{\mathsf{V}}(v) = xuy = \gamma_{\mathsf{V}}(x'u'y')$, the pseudowords $\gamma_{\mathsf{V}}(v)$ and $\gamma_{\mathsf{V}}(x'u'y')$ have the same first letter and the same last letter in $E(\Gamma)$ and therefore the edges v and $x'u'y'$ begin in the same vertex and end in the same vertex. Since γ_{V} is faithful, it follows that $v = x'u'y'$, which completes the proof of the lemma. ☐

The following elementary observation from combinatorics on words will be useful in the sequel.

Lemma 9.2. *Suppose that $u, v \in X^+$ and $k > |v|$ are such that v is primitive and u^k is a factor of v^ω. Then u is a conjugate of a power of v.*

Proof. The hypothesis that u^k is a factor of v^ω implies that u^k is a prefix of z^ω for some conjugate z of v. On the other hand, as $k > |v|$, we have $|u^k| = k|u| \geq (|v|+1)|u| \geq |u| + |v| = |u| + |z|$. Therefore, by Fine and Wilf's Theorem, it follows that u and z are both powers of some word y. Now, z is primitive since it is a conjugate of a primitive word. Hence $y = z$ and u is a power of z, say $u = z^\ell$ for some $\ell \geq 1$. Therefore u is a conjugate of v^ℓ, thus proving the lemma. ☐

Given two pseudovarieties of semigroups V and W, $\mathsf{V} * \mathsf{W}$ denotes the pseudovariety generated by all semidirect products $S * T$ with $S \in \mathsf{V}$ and $T \in \mathsf{W}$. The semidirect products of the form $\mathsf{V} * \mathsf{D}_n$ have been extensively studied in [1, Chapter 10] as an approach to semidirect products of the form $\mathsf{V} * \mathsf{D}$, which in turn have received even more attention [16, 31, 30]. We need to recall here some technology from [1, Chapter 10].

Let V be a pseudovariety and n a positive integer. Let B_n denote the de Bruijn graph of order n, whose set of vertices is X^n and whose set of edges is X^{n+1}, where $axb : ax \to xb$ whenever $a, b \in X$ and $x \in X^{n-1}$. Recall that the Rauzy graph $G_n(w)$ of a pseudoword w, which was introduced in the proof of Theorem 6.2, consists of the subgraph of B_n whose vertices and edges are factors of w.

Denote by $X^{\leq n}$ the set of all words $u \in X^+$ such that $|u| \leq n$. We define a continuous mapping $\bar{\Phi}_n^{\mathsf{V}}$ using the following diagram

$$\begin{tikzcd}
\overline{\Omega}_X(\mathsf{V} * \mathsf{D}_n) \setminus X^{\leq n} \arrow[rr, "\bar{\Phi}_n^{\mathsf{V}}"] \arrow[dr, "\iota_{\mathsf{V}}"'] & & E(\overline{\Omega}_{B_n}\mathsf{gV}) \arrow[r, hook] \arrow[dl, "\eta_{\mathsf{V}}"] & \overline{\Omega}_{B_n}\mathsf{gV} \\
& S_n & &
\end{tikzcd} \tag{3}$$

where

- S_n is the subsemigroup of the semigroup $M_n(\overline{\Omega}_{X^{n+1}}\mathsf{V}, \Phi_n^{\mathsf{V}})$ of [1, Section 10.6] whose universe is $X^n \times \overline{\Omega}_{X^{n+1}}\mathsf{V} \times X^n$ and whose operation is given by
$$(u_1, w_1, v_1)\,(u_2, w_2, v_2) = (u_1, w_1 \Phi_n^{\mathsf{V}}(v_1 u_2) w_2, v_2),$$
where, for a word t of length at least $n+1$, $\Phi_n^{\mathsf{V}}(t)$ is the value in $\overline{\Omega}_{X^{n+1}}\mathsf{V}$ of the word over the alphabet X^{n+1} which reads the successive factors of length $n+1$ of t;
- the arrow η_{V} is the continuous map that sends each edge $u : x \to y$ of the category $\overline{\Omega}_{B_n}\mathsf{gV}$ to the triple $(x, \gamma_{\mathsf{V}}(u), y)$;
- the arrow ι_{V} is the continuous homomorphism given by [1, Theorem 10.6.12].

Since η_{V} is continuous and the image of the subsemigroupoid generated by B_n is dense in $\operatorname{Im}\iota_{\mathsf{V}}$, we have $\operatorname{Im}\eta_{\mathsf{V}} = \operatorname{Im}\iota_{\mathsf{V}}$. On the other hand, since γ_{V} is faithful, η_{V} is injective. Hence η_{V} is a homeomorphism of $E(\overline{\Omega}_{B_n}\mathsf{gV})$ with $\operatorname{Im}\iota_{\mathsf{V}}$, and we may define the continuous mapping $\bar{\Phi}_n^{\mathsf{V}}$ to be the composite $\eta_{\mathsf{V}}^{-1} \circ \iota_{\mathsf{V}}$. It is therefore just a reinterpretation of the mapping ι_{V}. Note that each finite word w of length at least $n+1$ is mapped by $\bar{\Phi}_n^{\mathsf{V}}$ to the path which starts at the prefix $i_n(w)$ of length n, ends at the suffix $t_n(w)$ of length n, and goes through the edges given by the successive factors of length $n+1$ of w.

Denote by A_k the pseudovariety $[\![x^{k+1} = x^k]\!]$. Note that $\mathsf{A}_\omega = \mathsf{A} * \mathsf{D}_n = \mathsf{A}$.

Lemma 9.3. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be 1-slim, periodic at the ends, and let n be a common multiple of the lengths of all special bases of w. Let k be a positive integer or ω. Then there is a unique mapping τ_k such that the following diagram commutes:*

$$\begin{tikzcd}
\overline{\Omega}_{\Gamma(w)}\mathsf{gA} \arrow[rr, "\lambda_w"] \arrow[d, "\tau_k"'] & & \overline{\Omega}_X\mathsf{A} \setminus X^{\leq n} \arrow[r, hook] \arrow[d] & \overline{\Omega}_X\mathsf{A} \arrow[d, "p_k"] \\
\overline{\Omega}_{G_n(w)}\mathsf{gA}_k \arrow[r, hook] & \overline{\Omega}_{B_n}\mathsf{gA}_k & \overline{\Omega}_X(\mathsf{A}_k * \mathsf{D}_n) \setminus X^{\leq n} \arrow[l, "\bar{\Phi}_n^{\mathsf{A}_k}"] \arrow[r, hook] & \overline{\Omega}_X(\mathsf{A}_k * \mathsf{D}_n)
\end{tikzcd} \tag{4}$$

where p_k is the natural continuous homomorphism. Moreover the mapping τ_k is a continuous homomorphism.

Proof. Let $\tau_k : \overline{\Omega}_{\Gamma(w)}\mathsf{gA} \to \overline{\Omega}_{G_n(w)}\mathsf{gA}_k$ denote the continuous graph homomorphism which maps each vertex x^ω to the vertex $i_n(x^\omega) = t_n(x^\omega)$ and each edge $x^\omega yz^\omega$ to the "pseudopath reading the successive factors of length $n+1$". More precisely, τ_k is defined so that the diagram (4) commutes. The uniqueness, existence, and continuity of the mapping τ_k is justified by the observation that the composite mapping $\bar{\Phi}_n^{\mathsf{A}_k} \circ p_k \circ \lambda_w$ takes its values in $\overline{\Omega}_{G_n(w)}\mathsf{gA}_k$.

It remains to show that τ_k is a semigroupoid homomorphism. For this purpose, it suffices to show that it is a homomorphism on the subsemigroupoid generated by $\Gamma(w)$, since this subsemigroupoid is dense in $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$. A finite path in $\Gamma(w)$ is a sequence of edges of the form $x_{i-1}^\omega y_i x_i^\omega : x_{i-1}^\omega \to x_i^\omega$ $(i = 1, \ldots, r)$. Its image under λ_w is the product $x_0^\omega y_1 x_1^\omega \cdots y_r x_r^\omega$. Hence it suffices to observe that, for every finite word y whose length divides n and $x, z \in \overline{\Omega}_X(\mathsf{A}_k * \mathsf{D}_n)$, the following equality holds:

$$\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega z) = \bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\bar{\Phi}_n^{\mathsf{A}_k}(y^\omega z). \tag{5}$$

By the definition of $\bar{\Phi}_n^{\mathsf{A}_k}$, to prove (5) it suffices to show that η_{A_k} maps both sides to the same element of $X^n \times \overline{\Omega}_{X^{n+1}}\mathsf{A}_k \times X^n$. The left side is mapped to

$$\iota_{\mathsf{A}_k}(xy^\omega z) = \eta_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega z)\big) = \Big(i_n(xy^\omega), \gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega z)\big), t_n(y^\omega z)\Big) \tag{6}$$

while, since γ_{A_k} is a homomorphism, the right side is mapped to

$$\eta_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\bar{\Phi}_n^{\mathsf{A}_k}(y^\omega z)\big) = \Big(i_n(xy^\omega), \gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\big)\, \gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(y^\omega z)\big), t_n(y^\omega z)\Big). \tag{7}$$

Since ι_{A_k} is a homomorphism, we have $\iota_{\mathsf{A}_k}(xy^\omega z) = \iota_{\mathsf{A}_k}(xy^\omega)\iota_{\mathsf{A}_k}(y^\omega z)$ which shows that the triple in (6) may be rewritten as the product

$$\Big(i_n(xy^\omega), \gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\big)\, \bar{\Phi}_n^{A_k}\big(t_n(xy^\omega)\, i_n(y^\omega z)\big)\, \gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(y^\omega z)\big), t_n(y^\omega z)\Big) \tag{8}$$

Hence, to prove (5) it suffices to establish the equality of the middle components of the triples in (7) and (8), which in turn follows from the stronger equality

$$\gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\big) = \gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\big)\, \bar{\Phi}_n^{\mathsf{A}_k}\big(t_n(y^\omega)\, i_n(y^\omega)\big). \tag{9}$$

Since n is a multiple of the length of y, $t_n(y^\omega)\, i_n(y^\omega) = y^s$ for some integer $s \geq 2$. On the other hand, since γ_{A_k} and $\bar{\Phi}_n^{\mathsf{A}_k}$ are continuous,

$$\gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\big) = \lim_\ell \gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\ell)\big). \tag{10}$$

Now, for $\ell \geq ms$, $\gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\ell)\big)$ is a word in the alphabet X^{n+1} which admits $\big(\Phi_n^{\mathsf{A}_k}(y^s)\big)^m$ as a suffix. By compactness, it follows that $\big(\Phi_n^{\mathsf{A}_k}(y^s)\big)^m$ is a suffix of $\gamma_{\mathsf{A}_k}\big(\bar{\Phi}_n^{\mathsf{A}_k}(xy^\omega)\big)$ for all m, and therefore so is $\big(\Phi_n^{\mathsf{A}_k}(y^s)\big)^\omega$. Aperiodicity now implies the equality (9), which completes the verification of the equality (5) and the proof of the lemma. □

As a consequence of Lemma 8.11, we may associate to each infinite 1-slim pseudoword w a positive integer $\nu(w) = m'm''$ where m' is a multiple of the length of each special base of w and $m'' \geq 8$ is such that $m'' > |xyz|$ for every special factor $x^\omega yz^\omega$ of w in normal form.

Lemma 9.4. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be a 1-slim pseudoword that is periodic at the ends, let $n = \nu(w)$, and let $u = u'au''$ be a factor of w of length $3n$ where $|u'| = n$ and $a \in X$. If u' occurs in a special ω-factor x^ω of w and $u'a$ does not occur in x^ω, then u occurs in a unique special factor of w of the form $x^\omega yz^\omega$. Moreover, u may be written uniquely in the form $x'yz'$ where x' is a suffix of x^ω and z' is a prefix of z^ω with $|x'| \leq n$ and $n \leq |z'| < 2n$.*

Proof. Since w is 1-slim, we know from Lemma 8.11 that it has only finitely many special ω-factors and finitely many special factors. If w would have no special factors, then by Lemma 8.14 it would be a special ω-factor v^ω and the Lyndon word v would be the only special base of w. In that case, since by hypothesis x^ω is a special ω-factor of w, we would have $v = x$ which contradicts the hypotheses that u is a factor of w and that $u'a$ does not occur in x^ω. Therefore w has special factors and, by Lemma 8.13, u is a factor of one of them, say of $t_1^\omega yt_2^\omega$ written in normal form.

Now, notice that since x is a special base of w, the definition of n implies the existence of a positive integer $k > |t_1yt_2|$ such that $n = k|x|$. Hence, from the hypotheses that u' is a factor of x^ω and $|u'| = n$, we deduce that $u' = v^k$ for some conjugate v of x. Since $k > |t_1yt_2|$ and u' is a factor of $t_1^\omega yt_2^\omega$, it follows that v^{k-2} occurs in t_1^ω or in t_2^ω. As $k - 2 > |t_1yt_2| - 2 \geq |t_1|$, if v^{k-2} occurs in t_1^ω, then v is a conjugate of t_1^ℓ for some $\ell \geq 1$ by Lemma 9.2. Therefore v and t_1 are conjugates because v is a primitive word, whence x is a conjugate of t_1. As x and t_1 are Lyndon words, this shows that they are equal. In the second case, we would deduce analogously that $x = t_2$. However, since v is a conjugate of x, if the second case holds, then va does not occur in t_2^ω since otherwise a would be the first letter of v and so $u'a(= v^ka)$ would occur in $x^\omega(= t_2^\omega)$. Therefore, the first case holds necessarily, whence $x = t_1$ and u occurs in the special factor $x^\omega yz^\omega$, where $z = t_2$. Moreover, since $u'a$ is not a factor of x^ω and taking into account that $x^\omega yz^\omega$ is in normal form, the occurrence of u in $x^\omega yz^\omega$ must be such that the a in question is found within y. Hence there is a factorization $u = x'yz'$ with $|x'| \leq n$ a suffix of x^ω and z' a prefix of z^ω with $|z'| < 2n$, and so $|z'| = |u| - |x'| - |y| \geq 3n - n - n = n$.

To prove uniqueness, suppose that u occurs in special factors $x^\omega y_1z_1^\omega$ and $x^\omega y_2z_2^\omega$ in normal form. Then one deduces from the above that u may be written in the forms $u = x'y_1z_1' = x''y_2z_2'$, where x' and x'' are suffixes of x^ω and z_i' is a prefix of z_i^ω $(i = 1, 2)$ with $|x'|, |x''| \geq 2|x|$ and $|z_i'| \geq |z_1| + |z_2|$. As x is a Lyndon word, it follows that $x' = tx^{\ell'}$ and $x'' = tx^{\ell''}$ for some prefix t of x and positive integers ℓ' and ℓ''. On the other hand, since z_1 and z_2 are Lyndon words which have powers with a common factor of length $|z_1| + |z_2|$, we deduce as above that $z_1 = z_2$ and that $z_1' = z_1^{m'}v$ and $z_2' = z_1^{m''}v$ for some suffix v of z_1 and positive integers m' and m''. We thus conclude that $x^\omega y_1z_1^\omega$ and $x^\omega y_2z_2^\omega$ are equal. Since they are written in normal form, we finally deduce from McCammond's normal form algorithm that $y_1 = y_2$. To conclude the proof, we notice that $\ell' = \ell''$ (i.e., $x' = x''$) and that $m' = m''$ (i.e., $z_1' = z_2''$) since otherwise y_1 would be a prefix of x^ω or a suffix of z_1^ω, thus contradicting the definition of normal form. $\square$

Given two pseudowords $u, v \in \overline{\Omega}_X\mathsf{A}$ and a positive integer m, write $u \simeq_m v$ if u and v have precisely the same prefix, the same suffix and the same factors of length m.

Lemma 9.5. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be a 1-slim pseudoword that is periodic at the ends, let $n = \nu(w)$, and let $k \geq 1$ be an integer. Suppose that v is a finite word such that $w \simeq_{kn+1} v$. Then every occurrence of a factor of v of length n which is a factor of some special ω-factor of w is found within a factor of v of length kn which occurs in the same special ω-factor.*

Proof. Let $v = v_1uv_2$ be a factorization such that $|u| = n$ and u is a factor of a special ω-factor x^ω of w.

Suppose first that $|uv_2| \leq kn + 1$. Then u is a factor of the suffix s of length $kn + 1$ of v. Since, by hypothesis, s is also a suffix of w and w is periodic at the ends, we deduce that s is a factor of the special suffix z^ω of w. But $n = \nu(w)$ is at least twice the length of both x and z. Since u is a common factor of x^ω and z^ω and x and z are both Lyndon words, it follows from Fine and Wilf's Theorem (cf. [21, Proposition 1.3.5]) that $x = z$, which concludes the proof in this case. Thus, we may assume that $|uv_2| \geq kn + 1$. Similarly, we may also assume that $|v_1u| \geq kn + 1$.

Let up be the longest prefix of uv_2 that is a factor of x^ω and suppose that $|up| < kn$ for, otherwise, we are done. Let a be the first letter of v_2 and let s be the suffix of v_1 of length $kn + 1 - |upa|$. By hypothesis, the factor $supa$ of v is also a factor of w. By Lemma 8.13, either (a) $supa$ is a factor of a special ω-factor z^ω or (b) $supa$ is a factor of a special factor $\alpha^\omega\beta\gamma^\omega$ of w.

If some factor of length $2m'-1$ of up is a factor of a special ω-factor t^ω of w then, again by Fine and Wilf's Theorem, since x^ω and t^ω have a common factor of length $|x|+|t|-\gcd(|x|,|t|) \leq 2m'-1$, we conclude that $x = t$. Thus case (a) is impossible as upa in particular would be a factor of x^ω, contrary to our assumption and so it remains to treat case (b).

Since $m'' \geq 8$, an elementary calculation shows that $4m' + m'' \leq m'm'' = n$. Let $sup = y_1y_2y_3y_4$ be the factorization such that y_2 and y_4 both have length $2m'-1$ and $|y_3| = m''$. Since $supa$ is a factor of $\alpha^\omega\beta\gamma^\omega$ and $|y_3| = m'' > |\beta|$, we deduce that either y_1y_2 is a factor of α^ω or y_4a is a factor of γ^ω. By the preceding paragraph, it follows that, in the former case, $\alpha = x$ and y_1y_2 is a factor of x^ω while, in the latter case, $\gamma = x$ and y_4a is a factor of x^ω. Since the overlap of both y_1y_2 and y_4a with up has length at least $|x|$ and x is primitive, [22, Proposition 12.1.3] implies that, respectively, sup or upa is a factor of x^ω, the latter being excluded by the choice of p. Hence the given occurrence of u as a factor v is found within the factor sup, of length kn, which is still a factor of x^ω. $\square$

Lemma 9.6. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be a 1-slim pseudoword that is periodic at the ends, let $n = \nu(w)$, and let $k \geq 3$ be an integer. Let w_k be a finite word such that $w_k \simeq_{kn+1} w$. Then there is some path w'_k in $\Gamma(w)$, starting and ending respectively at the special prefix and special suffix of w, such that*

$$\tau_k(w'_k) = \bar{\Phi}_n^{\mathsf{A}_k}(p_k(w_k)). \tag{11}$$

Proof. We first note that $\bar{\Phi}_n^{\mathsf{A}_k}(p_k(w_k))$ is a path in $G_n(w)$ which starts at the vertex $i_n(w_k) = i_n(w)$, ends at $t_n(w_k) = t_n(w)$, and goes through the edges determined by the successive occurrences of factors of w_k of length $n+1$ of w_k. Since w_k and w have the same factors of length $3n \leq kn+1$, Lemma 9.4 shows that each such occurrence is either an occurrence in some special ω-factor, or it determines uniquely a special factor $x^\omega yz^\omega$ in which it occurs. Lemma 9.4 shows moreover that the path $\bar{\Phi}_n^{\mathsf{A}_k}(p_k(w_k))$ determines uniquely an alternated sequence of special ω-factors and special factors of w, where the special ω-factors that come before and after the special factor $x^\omega yz^\omega$ are respectively x^ω and z^ω. In other words, we have an associated path w'_k in the graph $\Gamma(w)$.

It remains to show that (11) holds. From Lemma 9.5, it follows that the path $\bar{\Phi}_n^{\mathsf{A}_k}(p_k(w_k))$, whenever it enters an edge which is determined by a factor of a special ω-factor v^ω, it must go at least k times around the cycle which reads the successive factors of v^ω of length $n+1$. The equality (11) now follows from the definition of τ_k. $\square$

The following lemma is the core of our encoding of 1-slim pseudowords w as pseudopaths in the graph $\Gamma(w)$.

Lemma 9.7. *If $w \in \overline{\Omega}_X\mathsf{A}$ is 1-slim and has at least one special factor, then there are some edge $w' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ and some words $x, y \in X^*$ such that $w = x\lambda_w(w')y$.*

Proof. By Corollary 8.12, we may as well assume that w is periodic at the ends. We will show that $w = \lambda_w(w')$ for a suitable $w' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$.

Let n be as in Lemma 9.4 and for this value of n, consider the continuous homomorphisms τ_k ($k \in \{1, 2, \ldots, \omega\}$) of Lemma 9.3. In particular, the diagram (4) commutes. Combining the diagrams for positive integers $k \leq \ell$ and ω, we obtain the following diagram, where p_k, $p_{\ell,k}$, q_k, and $q_{\ell,k}$ are the natural continuous homomorphisms:

$$\begin{tikzcd}
\overline{\Omega}_{\Gamma(w)}\mathsf{gA} \arrow[drrr, bend left, "\lambda_w"] \arrow[dr, "\tau_\omega"] \arrow[ddr, bend right, "\tau_\ell"'] \arrow[dddr, bend right, "\tau_k"'] & & & & \\
& \overline{\Omega}_{G_n(w)}\mathsf{gA} \arrow[r, hook] \arrow[d] & \overline{\Omega}_{B_n}\mathsf{gA} \arrow[d, "q_\ell"'] \arrow[dd, bend left, "q_k"] & \overline{\Omega}_X\mathsf{A} \setminus X^{\leq n} \arrow[l, "\bar{\Phi}_n^{\mathsf{A}}"] \arrow[r, hook] \arrow[d] \arrow[dd, bend left] & \overline{\Omega}_X\mathsf{A} \arrow[d, "p_\ell"'] \arrow[dd, bend left, "p_k"] \\
& \overline{\Omega}_{G_n(w)}\mathsf{gA}_\ell \arrow[r, hook] \arrow[d] & \overline{\Omega}_{B_n}\mathsf{gA}_\ell \arrow[d, "q_{\ell,k}"'] & \overline{\Omega}_X(\mathsf{A}_\ell * \mathsf{D}_n) \setminus X^{\leq n} \arrow[l, "\bar{\Phi}_n^{\mathsf{A}_\ell}"] \arrow[r, hook] \arrow[d] & \overline{\Omega}_X(\mathsf{A}_\ell * \mathsf{D}_n) \arrow[d, "p_{\ell,k}"'] \\
& \overline{\Omega}_{G_n(w)}\mathsf{gA}_k \arrow[r, hook] & \overline{\Omega}_{B_n}\mathsf{gA}_k & \overline{\Omega}_X(\mathsf{A}_k * \mathsf{D}_n) \setminus X^{\leq n} \arrow[l, "\bar{\Phi}_n^{\mathsf{A}_k}"] \arrow[r, hook] & \overline{\Omega}_X(\mathsf{A}_k * \mathsf{D}_n)
\end{tikzcd} \tag{12}$$

The commutativity of the leftmost and rightmost rectangles is trivial, while for the rectangles in the middle it follows from the continuity of the mappings involved and the way the horizontal sides transform finite words. The commutativity of the leftmost triangles is now a consequence of the uniqueness of the mappings τ_k, cf. Lemma 9.3. Hence the diagram (12) commutes.

Let $(w_k)_k$ be a sequence of finite words converging to w such that, for all k, w_k and w have the same factors, prefixes and suffixes of length $kn+1$. Note that, by Lemma 9.5, if v^ω is a special ω-factor of w, then every occurrence of a factor of v^ω of length n as a factor of w_k can be found within an occurrence of v^k as a factor of w_k.

The reason to consider τ_k for finite k is that the image under τ_k of each edge $x^\omega yz^\omega \in \Gamma(w)$ is equal over gA_k to a finite path. Hence, for finite k, τ_k maps finite paths of $\Gamma(w)$ to finite paths of $G_n(w)$.

By Lemma 9.6, there is some path w'_k in $\Gamma(w)$, starting and ending respectively at the special prefix and special suffix of w such that $\tau_k(w'_k) = \bar{\Phi}_n^{\mathsf{A}_k}(p_k(w_k)) \in \Omega_{G_n(w)}\mathsf{gA}_k$. By compactness, without loss of generality, we may assume that the sequence $(w'_k)_k$ converges to some w' in $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$. Then, in view of the commutativity of the diagram (12) and the fact that the homomorphisms q_k ($k = 1, 2, \ldots$) suffice to separate points of $\overline{\Omega}_{B_n}\mathsf{gA}$, we obtain the following implications:

$$\begin{aligned}
&\tau_\ell(w'_\ell) = \bar{\Phi}_n^{\mathsf{A}_\ell}(p_\ell(w_\ell)) \quad \forall \ell \\
&\quad \Rightarrow \tau_k(w'_\ell) = q_{\ell,k}\big(\tau_\ell(w'_\ell)\big) = q_{\ell,k}\big(\bar{\Phi}_n^{\mathsf{A}_\ell}(p_\ell(w_\ell))\big) = \bar{\Phi}_n^{\mathsf{A}_k}(p_k(w_\ell)) \quad \forall k\ \forall \ell \geq k \\
&\quad \Rightarrow q_k(\tau_\omega(w')) = \tau_k(w') = \bar{\Phi}_n^{\mathsf{A}_k}(p_k(w)) = q_k(\bar{\Phi}_n^{\mathsf{A}}(w)) \quad \forall k \\
&\quad \Leftrightarrow \tau_\omega(w') = \bar{\Phi}_n^{\mathsf{A}}(w)
\end{aligned}$$

From the equality $\tau_\omega(w') = \bar{\Phi}_n^{\mathsf{A}}(w)$ we deduce that $\bar{\Phi}_n^{\mathsf{A}}(\lambda_w(w')) = \tau_\omega(w') = \bar{\Phi}_n^{\mathsf{A}}(w)$. Since $\bar{\Phi}_n^{\mathsf{A}}$ is injective, because the homomorphism ι_{V} of diagram (3) is an embedding, it follows that $\lambda_w(w') = w$, which completes the proof of the lemma. □

Recall that $\gamma_{\mathsf{A}} : \overline{\Omega}_{\Gamma(w)}\mathsf{gA} \to (\overline{\Omega}_{E(\Gamma(w))}\mathsf{A})^1$ is a faithful homomorphism of categories. It maps edges of $\Gamma(w)$ to themselves, viewed as letters. The following result shows how we can carry the theory of ω-words from $\overline{\Omega}_{E(\Gamma(w))}\mathsf{A}$ to $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$.

Proposition 9.8. *Let Γ be a finite graph and let u and v be two edges of $\overline{\Omega}_\Gamma\mathsf{gA}$. Then $\gamma_{\mathsf{A}}(u)$ is a factor of $\gamma_{\mathsf{A}}(v)$ in $\overline{\Omega}_{E(\Gamma)}\mathsf{A}$ if and only if u is a factor of v in $\overline{\Omega}_\Gamma\mathsf{gA}$. In particular, all factors of ω-paths in $\overline{\Omega}_\Gamma\mathsf{gA}$ are again ω-paths.*

Proof. The forward implication follows from Lemma 9.1 and the reverse implication is a consequence of γ_{A} being a homomorphism. □

We also define the *rank* of a pseudopath of $\overline{\Omega}_\Gamma\mathsf{gA}$ to be the rank of its image under γ_{A} and the *normal form* of a pseudopath to be normal form of its image under γ_{A}.

To proceed, we need a rather technical but unsurprising lemma which requires some further notation to state.

We denote by π_k the natural continuous homomorphism $\overline{\Omega}_X\mathsf{A} \to \overline{\Omega}_X\mathsf{A}_k$. Note that $\Omega_X^\kappa\mathsf{A}_k = \Omega_X\mathsf{A}_k$. For $u, v \in \Omega_X^\kappa\mathsf{A}$, we write $u \sim_{w,k} v$ if it is possible to transform u into v by changing the exponents of factors of the form z^p, keeping them at least k, where z is a primitive word such that z^ω is a factor of w. Note that $\sim_{w,k}$ is a congruence on $\Omega_X^\kappa\mathsf{A}$ such that $u \sim_{w,k} v$ implies $\pi_k(u) = \pi_k(v)$ but the converse is false. For instance, the converse fails for $w = a^\omega$, $u = b^k$, and $v = b^{k+1}$, in case X has at least two distinct letters a and b.

Noting that A is local, that is $\mathsf{gA} = \ell\mathsf{A}$ (cf. [31]), there is a natural continuous homomorphism, ϖ_k : $\overline{\Omega}_{\Gamma(w)}\mathsf{gA} \to \overline{\Omega}_{\Gamma(w)}\ell\mathsf{A}_k$.

For a finite semigroup S, we denote by $\operatorname{ind}(S)$ the smallest positive integer m such that S satisfies the pseudoidentity $x^{\omega+m} = x^m$.

Lemma 9.9. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be an infinite 1-slim pseudoword and let $\varphi : \overline{\Omega}_{\Gamma(w)}\mathsf{gA} \to S$ be a continuous homomorphism onto some $S \in \mathsf{A}$ such that the natural continuous homomorphisms $\overline{\Omega}_{\Gamma(w)}\mathsf{gA} \to \overline{\Omega}_{E(\Gamma(w))}\mathsf{K}_1$ and $\overline{\Omega}_{\Gamma(w)}\mathsf{gA} \to \overline{\Omega}_{E(\Gamma(w))}\mathsf{D}_1$ factor through φ. Let $n = \nu(w)$ and let $k = \max\{\operatorname{ind}(S) + 1, n\}$. Denote by μ_w*

the restriction of λ_w to $\Omega_{\Gamma(w)}\mathsf{gA}$. Let m be the supremum of the lengths of the special bases of w. Then the following properties hold:

(a) *μ_w is injective on edges;*

(b) *if $u \in \Omega_{\Gamma(w)}\mathsf{gA}$ is an edge and $v \in \Omega^{\kappa}_{X}\mathsf{A}$ is such that $\pi_k\big(\lambda_w(u)\big) = \pi_k(v)$, then there exists an edge $u' \in \Omega_{\Gamma(w)}\mathsf{gA}$ such that $\lambda_w(u') \sim_{w,k} v$ and $\varpi_{k-1}(u) = \varpi_{k-1}(u')$;*

(c) *$\ker(\pi_k \circ \mu_w) \subseteq \ker\varphi$ so that φ induces a homomorphism of partial semigroups $\psi : \pi_k(\operatorname{Im}\mu_w) \to S$ such that $\psi \circ \pi_k \circ \mu_w = \varphi|_{\Omega_{\Gamma(w)}\mathsf{gA}}$;*

(d) *let θ denote the congruence on $\Omega_X\mathsf{A}_k$ generated by the binary relation $\ker\psi$; then $\pi_k(\operatorname{Im}\mu_w)$ is a union of θ-classes;*

(e) *the restriction of θ to $\pi_k(\operatorname{Im}\mu_w)$ coincides with $\ker\psi$;*

(f) *denote by $\hat\psi$ the natural homomorphism $\Omega_X\mathsf{A}_k \to \Omega_X\mathsf{A}_k/\theta$; then, whenever $u \in \Omega_X\mathsf{A}_k$, $\hat\psi(u)$ is a factor of some element of $\hat\psi \circ \pi_k(\operatorname{Im}\mu_w)$ if and only if u is a factor of some element of $\pi_k(\operatorname{Im}\mu_w)$;*

(g) *if $u \in \Omega^{\kappa}_{X}\mathsf{A}$ and $x, y \in X^*$ are such that $\pi_k(xuy) \in \pi_k(\operatorname{Im}\mu_w)$ and $\pi_k(x'uy') \notin \pi_k(\operatorname{Im}\mu_w)$ whenever x' is a suffix of x and y' is a prefix of y such that $|x'y'| < |xy|$, then $(k+1)m > \max\{|x|, |y|\}$; in particular, there are only finitely many such pairs (x, y).*

Proof. To prove (a), it suffices to observe that, since λ_w maps edges of $\Gamma(w)$ to elements of $\Omega^{\kappa}_{X}\mathsf{A}$, McCammond's solution of the word problem for $\Omega^{\kappa}_{X}\mathsf{A}$ allows us to recover a path in $\Gamma(w)$ from its image under λ_w.

Since $S \in \mathsf{A}_{k-1}$ and $\mathsf{gA} = \ell\mathsf{A} \supseteq \ell\mathsf{A}_{k-1}$, φ factors through a continuous homomorphism $\xi : \overline{\Omega}_{\Gamma(w)}\ell\mathsf{A}_{k-1} \to S$. For the remainder of the proof, it will be useful to keep in mind the following commutative diagram, where the existence of the mapping ψ is asserted in (c) and $\hat\psi$ is defined in (f):

$$\tag{13}\begin{tikzcd}
\overline{\Omega}_{\Gamma(w)}\mathsf{gA} \arrow[rrrrr, "\lambda_w"] \arrow[d, "\varpi_{k-1}"'] \arrow[dddrr, bend left, "\varphi"] & & & & & \overline{\Omega}_X\mathsf{A} \arrow[d, "\pi_k"] \\
\overline{\Omega}_{\Gamma(w)}\ell\mathsf{A}_{k-1} \arrow[ddrr, bend right, "\xi"'] & \Omega^{\kappa}_{\Gamma(w)}\mathsf{gA} \arrow[ul, hook] \arrow[rrr] & & & \Omega^{\kappa}_{X}\mathsf{A} \arrow[ur, hook] \arrow[d] & \overline{\Omega}_X\mathsf{A}_k \\
& & \Omega_{\Gamma(w)}\mathsf{gA} \arrow[ul, hook] \arrow[r, "\mu_w"] \arrow[d] & \operatorname{Im}\mu_w \arrow[ur, hook] \arrow[d] & \Omega_X\mathsf{A}_k \arrow[ur, hook] \arrow[d, "\hat\psi"] & \\
& & S & \pi_k(\operatorname{Im}\mu_w) \arrow[l, dashed, "\psi"'] \arrow[ur, hook] & \Omega_X\mathsf{A}_k/\theta &
\end{tikzcd}$$

Suppose that u and v are as in (b). Since $\Omega_X\mathsf{A}_k$ is the quotient of the ω-semigroup $\Omega^{\kappa}_{X}\mathsf{A}$ by the least congruence which identifies x^k, x^{k+1}, and x^ω for every $x \in \Omega^{\kappa}_{X}\mathsf{A}$, there are $v_0, v_1, \ldots, v_p \in \Omega^{\kappa}_{X}\mathsf{A}$ such that $v_0 = \lambda_w(u)$, $v_p = v$ and, for each $i \in \{1, \ldots, p\}$, there are factorizations $v_{i-1} = x_i y_i^{\alpha_i} z_i$ and $v_i = x_i y_i^{\beta_i} z_i$ with $x_i, z_i \in (\Omega^{\kappa}_{X}\mathsf{A})^1$, $y_i \in \Omega^{\kappa}_{X}\mathsf{A}$, and $\alpha_i, \beta_i \in \{k, k+1, \ldots, \omega\}$. Proceeding inductively to prove (b), we may as well assume that $p = 1$, and so drop subscripts, thus assuming that there are factorizations $\lambda_w(u) = xy^\alpha z$ and $v = xy^\beta z$, with $x, z \in (\Omega^{\kappa}_{X}\mathsf{A})^1$, $y \in \Omega^{\kappa}_{X}\mathsf{A}$, and $\alpha, \beta \in \{k, k+1, \ldots, \omega\}$.

Suppose that $y \in X^+$. Then the choice of k implies that the $\alpha \geq k$ occurrences of y which are distinguished in the factorization $\lambda_w(u) = xy^\alpha z$ are found within the same factor of the form t^ω, where t is a primitive word. In this case, we may simply take $u' = u$. Thus we may assume that $y \notin X^+$. It follows that y is a rank 1 ω-word. Let $y = y_0 t_1^\omega y_1 \cdots t_r^\omega y_r$ be its normal form. Then $xy_0t_1^\omega, t_r^\omega y_r z \in \operatorname{Im}\lambda_w$ so that $v = \lambda_w(u')$ for some u' which is obtained from u by replacing the exponent $\alpha - 1$ by $\beta - 1$ for the cycle

$$t_1^\omega \xrightarrow{t_1^\omega y_1 t_2^\omega} t_2^\omega \xrightarrow{t_2^\omega y_2 t_3^\omega} t_3^\omega \to \cdots \to t_{r-1}^\omega \xrightarrow{t_{r-1}^\omega y_{r-1} t_r^\omega} t_r^\omega \xrightarrow{t_r^\omega y_r y_0 t_1^\omega} t_1^\omega,$$

in case $t_r^\omega y_r y_0 t_1^\omega \neq t_1^\omega$, or otherwise the cycle

$$t_1^\omega \xrightarrow{t_1^\omega y_1 t_2^\omega} t_2^\omega \xrightarrow{t_2^\omega y_2 t_3^\omega} t_3^\omega \to \cdots \to t_{r-1}^\omega \xrightarrow{t_{r-1}^\omega y_{r-1} t_r^\omega} t_r^\omega = t_1^\omega,$$

which proves (b).

To establish (c), we prove a slightly stronger property. Let u, v be two edges in $\Omega_{\Gamma(w)}\mathsf{gA}$ and suppose that $\pi_k \circ \lambda_w(u) = \pi_k \circ \lambda_w(v)$. As above, there are $v_0, v_1, \ldots, v_p \in \Omega^{\kappa}_{X}\mathsf{A}$ such that $v_0 = \lambda_w(u)$, $v_p = \lambda_w(v)$

and, for each $i \in \{1, \dots, p\}$, there are factorizations $v_{i-1} = x_i y_i^{\alpha_i} z_i$ and $v_i = x_i y_i^{\beta_i} z_i$ with $x_i, z_i \in (\Omega_X^\kappa \mathsf{A})^1$, $y_i \in \Omega_X^\kappa \mathsf{A}$, and $\alpha_i, \beta_i \in \{k, k+1, \dots, \omega\}$. Applying (b) inductively, for each $i \in \{1, \dots, p-1\}$ there exists an edge $v_i' \in \Omega_{\Gamma(w)} \mathsf{gA}$ such that $v_0' = u$, $v_p' = v$, $\lambda_w(v_i') \sim_{w,k} v_i$, and $\varpi_{k-1}(v_{i-1}') = \varpi_{k-1}(v_i')$. Hence

$$\varphi(v_{i-1}') = \xi \circ \varpi_{k-1}(v_{i-1}') = \xi \circ \varpi_{k-1}(v_i') = \varphi(v_i')$$

for every i, from which it follows that $\varphi(u) = \varphi(v)$. This proves (c) and establishes the existence of the homomorphism ψ of partial semigroups. Note that $\pi_k(\operatorname{Im} \mu_w)$ may not be a subsemigroup of $\Omega_X^\kappa \mathsf{A}_k$.

To prove (d) and (e), suppose that $u \in \Omega_{\Gamma(w)} \mathsf{gA}$ and $v \in \Omega_X^\kappa \mathsf{A}$ are such that $\pi_k \circ \mu_w(u)\ \theta\ \pi_k(v)$. Then there exist $v_0, v_1, \dots, v_p \in \Omega_X^\kappa \mathsf{A}$ such that $v_0 = \lambda_w(u)$, $v_p = v$, and, for each $i \in \{1, \dots, p\}$, there exist factorizations $\pi_k(v_{i-1}) = \pi_k(x_i y_i z_i)$ and $\pi_k(v_i) = \pi_k(x_i t_i z_i)$ such that $x_i, z_i \in (\Omega_X^\kappa \mathsf{A})^1$, $y_i, t_i \in \operatorname{Im} \mu_w$, and $\psi \circ \pi_k(y_i) = \psi \circ \pi_k(t_i)$. Proceeding inductively, it suffices to show that $\pi_k(v_1) \in \pi_k(\operatorname{Im} \mu_w)$ and $\psi \circ \pi_k(v_1) = \psi \circ \pi_k \circ \lambda_w(u)$ and so we assume that $p = 1$ and we drop the indices, so that $\pi_k \circ \lambda_w(u) = \pi_k(xyz)$ and $\pi_k(v) = \pi_k(xtz)$ with $x, z \in (\Omega_X^\kappa \mathsf{A})^1$, $y, t \in \operatorname{Im} \mu_w$, and $\psi \circ \pi_k(y) = \psi \circ \pi_k(t)$.

Let y' and t' be edges in $\Omega_{\Gamma(w)} \mathsf{gA}$ such that $\lambda_w(y') = y$ and $\lambda_w(t') = t$. Since $\varphi(y') = \psi \circ \pi_k(y) = \psi \circ \pi_k(t) = \varphi(t')$, the hypothesis on the factorizability of φ through the natural continuous homomorphisms $\overline{\Omega}_{\Gamma(w)} \mathsf{gA} \to \overline{\Omega}_{E(\Gamma(w))} \mathsf{K}_1$ and $\overline{\Omega}_{\Gamma(w)} \mathsf{gA} \to \overline{\Omega}_{E(\Gamma(w))} \mathsf{D}_1$ implies that y and t are edges with the same ends.

By (b), there exists $u' \in \Omega_{\Gamma(w)} \mathsf{gA}$ such that $\lambda_w(u') \sim_{w,k} xyz$. Since $y \in \operatorname{Im} \mu_w$, the definition of $\sim_{w,k}$ and (a) imply that there exist $x', z' \in (\Omega_{\Gamma(w)} \mathsf{gA})^c$ such that $x'y'z' = u'$, $\lambda_w(x') \sim_{w,k} x$, and $\lambda_w(z') \sim_{w,k} z$, where λ_w is extended so as to map local identities to 1. Hence $v' = x't'z'$ is an edge in $\Omega_{\Gamma(w)} \mathsf{gA}$ such that $\lambda_w(v') \sim_{w,k} xtz$, whence $\pi_k(v) = \pi_k(xtz) = \pi_k \circ \lambda_w(v')$, which shows that $\pi_k(v) \in \pi_k(\operatorname{Im} \mu_w)$. To conclude the proof of (d) and (e), it now suffices to observe that

$$\psi \circ \pi_k \circ \lambda_w(v') = \varphi(v') = \varphi(x't'z') = \varphi(x'y'z') = \psi \circ \pi_k \circ \lambda_w(u).$$

To prove (f), suppose that $u \in \Omega_X^\kappa \mathsf{A}$ and there exist $v \in \operatorname{Im} \mu_w$ and $x, y \in (\Omega_X^\kappa \mathsf{A})^1$ such that $\hat{\psi} \circ \pi_k(xuy) = \hat{\psi} \circ \pi_k(v)$, that is $\pi_k(xuy)$ is θ-equivalent to some element of $\pi_k(\operatorname{Im} \mu_w)$. Then, by (d), $\pi_k(xuy) \in \pi_k(\operatorname{Im} \mu_w)$, that is $\pi_k(u)$ is a factor of some element of $\pi_k(\operatorname{Im} \mu_w)$, which establishes (f) since the converse is trivial.

For (g), we need to bound the length of the words x and y. Since $\pi_k(xuy) \in \pi_k(\operatorname{Im} \mu_w)$, by (b) there exists $v \in \Omega_{\Gamma(w)} \mathsf{gA}$ such that $\lambda_w(v) \sim_{w,k} xuy$. By the definitions of λ_w and $\sim_{w,k}$, there is some special ω-factor z^ω of w and z^ℓ is a prefix of xuy for some $\ell \geq k$. By the minimality hypothesis on the pair (x, y), it follows that $|x| < (k+1)|z|$ for, otherwise, writing $x = zx'$, we would still have $\lambda_w(v) \sim_{w,k} x'uy$. This shows that $|x| < (k+1)m$ and the dual argument shows that we also have $|y| < (k+1)m$. Note that m is an integer by Lemma 8.11. Hence the set of all pairs (x, y) in (g) for a given u is indeed finite. □

The preceding lemma provides the technicalities required to prove the following rather natural result but for which we do not know of any simpler proof.

Lemma 9.10. *Let $w \in \overline{\Omega}_X \mathsf{A}$ be an infinite 1-slim pseudoword. Then the homomorphism λ_w is injective on edges of $\overline{\Omega}_{\Gamma(w)} \mathsf{gA}$.*

Proof. Let u and v be distinct edges in $\overline{\Omega}_{\Gamma(w)} \mathsf{gA}$. We need to show that $\lambda_w(u) \neq \lambda_w(v)$. There exist some monoid $S \in \mathsf{A}$ and some onto continuous homomorphism $\varphi : \overline{\Omega}_{\Gamma(w)} \mathsf{gA} \to S$ such that $\varphi(u) \neq \varphi(v)$. Since $\mathsf{K}_1 \cup \mathsf{D}_1 \subseteq \mathsf{A}$, by taking the product mapping induced by φ and the natural continuous homomorphisms $\varrho : \overline{\Omega}_{\Gamma(w)} \mathsf{gA} \to \overline{\Omega}_{E(\Gamma(w))} \mathsf{K}_1$ and $\varsigma : \overline{\Omega}_{\Gamma(w)} \mathsf{gA} \to \overline{\Omega}_{E(\Gamma(w))} \mathsf{D}_1$, we obtain a continuous homomorphism into a finite aperiodic semigroup such that $\varphi, \varrho, \varsigma$ factor through it. Hence, we may as well assume that ϱ, ς factor through φ. Let $n = \nu(w)$ and let $k = \max\{\operatorname{ind}(S) + 1, n\}$. Then the hypotheses of Lemma 9.9 hold and we proceed to apply it. We obtain a commutative diagram (13) from which we extract the relevant part for our present purposes in diagram (14) and add the nodes T and R and the arrows involving them which are described below.

(14)
$$\begin{tikzcd}
& & & \operatorname{Im}\lambda_w \arrow[dr, hook] & \\
\overline{\Omega}_{\Gamma(w)}\mathsf{gA} \arrow[urrr, bend left, "\lambda_w"] \arrow[ddrr, bend right, "\varphi"'] & \Omega_{\Gamma(w)}\mathsf{gA} \arrow[l, hook'] \arrow[r, "\mu_w"] \arrow[ddr, bend right] & \operatorname{Im}\mu_w \arrow[r, hook] \arrow[ur, hook] \arrow[d] & \Omega^\kappa_X\mathsf{A} \arrow[r, hook] \arrow[d] & \overline{\Omega}_X\mathsf{A} \arrow[d, "\pi_k"] \\
& & \pi_k(\operatorname{Im}\mu_w) \arrow[r, hook] \arrow[d, "\psi"] & \Omega_X\mathsf{A}_k \arrow[r, hook, dashed] \arrow[d, dashed, "\hat\psi"] & \overline{\Omega}_X\mathsf{A}_k \arrow[d, dashed, "\chi"] \\
& & S \arrow[r, hook, dashed] & T \arrow[r, dashed, "\rho"] & R
\end{tikzcd}$$

As in Lemma 9.9(d) let θ denote the congruence on $\Omega_X\mathsf{A}_k$ generated by $\ker\psi$, so that θ saturates $\pi_k(\operatorname{Im}\mu_w)$ and its restriction to that subset coincides with $\ker\psi$. Hence, if we denote by T the quotient semigroup $\Omega_X\mathsf{A}_k/\theta$, then we may assume that S is a subsemigroup of T in such a way that ψ is the restriction of the natural homomorphism $\hat\psi : \Omega_X\mathsf{A}_k \to T$ to $\pi_k(\operatorname{Im}\mu_w)$. At this point, we need to overcome an essential difficulty: the semigroup T may be infinite. We do so by taking a Rees quotient.

Let $I = \{t \in T : S \cap T^1\, t\, T^1 = \emptyset\}$. Then I is an ideal of T such that $I \cap S = \emptyset$. Denote by R the Rees quotient T/I and let $\rho : T \to R$ be the natural homomorphism so that $\rho \circ \varphi$ still distinguishes the edges u and v. We claim that R is finite. Since then, by construction, $R \in \mathsf{A}_k$, the homomorphism $\rho \circ \hat\psi$ factorizes through a continuous homomorphism $\chi : \overline{\Omega}_X\mathsf{A}_k \to R$. This completes the commutative diagram (14). We thus obtain
$$\chi \circ \pi_k \circ \lambda_w(u) = \rho \circ \varphi(u) \neq \rho \circ \varphi(v) = \chi \circ \pi_k \circ \lambda_w(v)$$
which establishes that $\lambda_w(u) \neq \lambda_w(v)$.

We proceed to establish that R is finite, which is equivalent to showing that $T \setminus I$ is a finite set. By Lemma 9.9(f), $T \setminus I$ consists of all elements of the form $\hat\psi \circ \pi_k(x)$ such that $x \in \Omega^\kappa_X\mathsf{A}$ and $\pi_k(x)$ is a factor of some element of $\pi_k(\operatorname{Im}\mu_w)$. By Lemma 9.9($g$), the number of such elements is bounded by
$$(|S|+1)\left(\sum_{i=0}^{(k+1)m-1} |X|^i\right)^2,$$
where m is given by Lemma 9.9, which shows that R is finite. ☐

The following is an immediate but useful observation.

Lemma 9.11. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be an infinite 1-slim pseudoword and let u be a factor of w. Then $\overline{\Omega}_{\Gamma(u)}\mathsf{gA}$ is a closed subcategory of $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ and λ_u is the restriction to $\overline{\Omega}_{\Gamma(u)}\mathsf{gA}$ of λ_w.*

Proof. By Theorem 8.7, since u is a factor of the 1-slim pseudoword w, u is also 1-slim. On the other hand, from the definitions, we see that $\Gamma(u)$ is a subgraph of $\Gamma(w)$ and so $\overline{\Omega}_{\Gamma(u)}\mathsf{gA}$ is a closed subcategory of $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ and it is immediate that λ_u is the restriction to $\overline{\Omega}_{\Gamma(u)}\mathsf{gA}$ of λ_w. ☐

A further important property of the homomorphism λ_w is that it preserves the factor order in a strong sense.

Lemma 9.12. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be an infinite 1-slim pseudoword and suppose that there is an edge w' in the category $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ such that $\lambda_w(w') = w$. Let u and v be edges in $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ that are factors of w'. Then $u <_{\mathcal{J}} v$ if and only if $\lambda_w(u) <_{\mathcal{J}} \lambda_w(v)$.*

Proof. Note that, since $x <_{\mathcal{J}} y$ means $x \leq_{\mathcal{J}} y$ and $y \not\leq_{\mathcal{J}} x$, it suffices to show that

(15) $$u \leq_{\mathcal{J}} v \iff \lambda_w(u) \leq_{\mathcal{J}} \lambda_w(v)$$

The direct implication in (15) follows immediately from the fact that λ_w is a homomorphism. Conversely, assume that $\lambda_w(u) \leq_{\mathcal{J}} \lambda_w(v)$. Let $s, t \in \overline{\Omega}_X\mathsf{A}$ be such that $\lambda_w(u) = s\lambda_w(v)t$. Since $\lambda_w(u)$ and $\lambda_w(v)$ are periodic at the ends, we may assume that so are s and t, with the infinite suffix of s having the same period as the infinite prefix of $\lambda_w(v)$ and the infinite prefix of t having the same period as the infinite suffix of $\lambda_w(v)$.

By Theorem 8.7, s and t are 1-slim. Hence, by Lemmas 9.11 and 9.7, there exist $s', t' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ such that $\lambda_w(s') = \lambda_s(s') = s$ and $\lambda_w(t') = \lambda_t(t') = t$. Note that the matching of periods at the ends guarantees that $s'vt'$ is an edge in the category $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$. Since λ_w is a homomorphism, it follows that $\lambda_w(s'vt') = \lambda_w(u)$. By Lemma 9.10, we deduce that $s'vt' = u$, which shows that $u \leq_{\mathcal{J}} v$. $\square$

The next lemma describes in detail the encoding procedure on a pseudoword given by an ω-term.

Lemma 9.13. *Let $w \in \Omega^{\kappa}_X\mathsf{A}$ be periodic at the ends and let $w' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ be such that $\lambda_w(w') = w$. Then $w' \in \Omega^{\kappa}_{\Gamma(w)}\mathsf{gA}$. Moreover, if w' is an infinite pseudopath then w' has trivial initial and final portions if and only if so does w.*

Proof. We assume w to be given by its normal form and we transform it according to the application, in order, of each of the following rules once to the parenthesized form of w:

(a) for each well-parenthesized factor of the form (u), where $u \notin X^+$, replace it by $(u)u$;
(b) for each maximal well-parenthesized occurrence of a factor of the form $v(u)t$, where $u \in X^+$ and $v \in \{(\}^*$, replace it by $(u)v(u)t$, unless the occurrence in question is as a prefix or as a suffix of w;
(c) for each factor of the form $)(z)$, where $z \in X^+$, replace it by $(z))$, unless the original factor is preceded by (z).

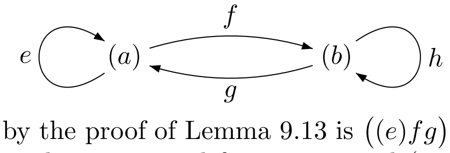
Once a rule has been applied to a given factor, it should not be applied again to that factor, although it may be applied to its factors. This guarantees that the procedure consisting in such application of the rules terminates.

One can verify that each elementary step in the above procedure does not change the pseudoword that the ω-term represents in $\Omega^{\kappa}_X\mathsf{A}$, so that the output remains a representation of w. For condition (c), this is only true because we started with a κ-term in normal form and we first applied (a) and (b). Moreover, by induction on the rank of w, it is easy to show that, replacing in the output each factor of the form $(x)y(z)$, where $x, y, z \in X^+$ by the corresponding edge $(x)y(z) : (x) \to (z)$ of the graph $\Gamma(w)$, we obtain an ω-path $w' \in \Omega^{\kappa}_{\Gamma(w)}\mathsf{gA}$. Hence $\lambda_w(w') = w$, which proves the lemma, in view of Lemma 9.10. $\square$

For example, for the ω-term $w = \Big(((a)b)a(b)c\Big)a(b)$, applying exhaustively each of the rules produces successively the following ω-terms where, in each step, we underline the new factors which come from or are changed by the application of the rules:

$$\Big(((a)b)\underline{(a)b}a(b)c\Big)\underline{((a)b)\underline{(a)b}a(b)c}a(b)$$
$$\Big(((a)b)\underline{(a)}(a)ba\underline{(b)}(b)c\Big)\underline{(a)}((a)b)\underline{(a)}(a)ba\underline{(b)}(b)ca(b)$$
$$\Big(((a)b\underline{(a))}(a)ba(b)(b)c\underline{(a)\Big)}((a)b\underline{(a))}(a)ba(b)(b)ca(b).$$

The graph $\Gamma(w)$ is given by the following diagram, where $e = (a)b(a)$, $f = (a)ba(b)$, $g = (b)c(a)$, and $h = (b)ca(b)$:



The ω-path over this graph given by the proof of Lemma 9.13 is $\big((e)fg\big)(e)fh$.

Note that, given an ω-word of rank 1 in normal form, its initial (respectively final) portion is trivial if and only if it is periodic at the corresponding end. For $w \in \overline{\Omega}_X\mathsf{A}$, let $P(w)$ be the set of all normal forms, as well-parenthesized words $u \in (X \cup \{(,)\})^+$, which are either of rank 0 or have trivial initial and final portions, such that $\epsilon(u) \in \mathcal{F}(w)$. Note that, since $F(w) = P(w) \cap X^+$, if $P(w)$ is rational then so is $F(w)$.

Lemma 9.14. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be a 1-slim pseudoword that is periodic at the ends and let $w' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ be such that $\lambda_w(w') = w$. Let L be the set of all special ω-factors of w. Then the following equalities hold:*

(a) *$\lambda_w\big(P(w')\big) = P(w) \setminus (L \cup X^+)$;*

(b) $\lambda_w^{-1}\big(P(w) \setminus (L \cup X^+)\big) = P(w')$.

Proof. (a) Let $u \in P(w')$. Then u (as an element of $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$) is a factor of w'. Since λ_w is a homomorphism, it follows that $\lambda_w(u)$ is a factor of $\lambda_w(w') = w$, and since it is not a finite word, having rank at least 1, and it contains at least one special factor, it belongs to $P(w) \setminus (L \cup X^+)$.

Conversely, suppose that $u \in P(w) \setminus (L \cup X^+)$. Then u (as an element of $\overline{\Omega}_X\mathsf{A}$), is a factor of w, that is $w = xuy$ for some $x, y \in (\overline{\Omega}_X\mathsf{A})^1$. By Theorem 8.7, each of the pseudowords x, u, y is 1-slim and so the graphs $\Gamma(x)$, $\Gamma(u)$, and $\Gamma(y)$ are well defined contained in $\Gamma(w)$. Moreover, by definition of $P(w)$, u is periodic at the ends and so, in case x or y are not empty factors, we may assume that they are also periodic at the ends, with the special suffix of x coinciding with the special prefix of u and the special suffix of u coinciding with the special prefix of y. By Lemma 9.7 and since $u \notin L$, there exist $x', y' \in (\overline{\Omega}_{\Gamma(w)}\mathsf{gA})^c$ and $u' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ such that $\lambda_w(x') = x$, $\lambda_w(u') = u$, and $\lambda_w(y') = y$. The matching of special prefixes and suffixes guarantees that $x'u'y'$ is an edge in $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$. Since $\lambda_w(w') = w = xuy = \lambda_w(x'u'y')$ and λ_w is injective on edges, we deduce that $w' = x'u'y'$. By Lemma 9.13, $u' \in \Omega^\kappa_{\Gamma(w)}\mathsf{gA}$. Hence the canonical form of u' belongs to $P(w')$, which shows that $u \in \lambda_w\big(P(w)\big)$.

(b) This follows from (a) and Lemma 9.10. $\square$

10. Main theorem

Putting together the main conclusions of the technical lemmas in the preceding section, we are now ready for the following result.

Proposition 10.1. *Let $w \in \overline{\Omega}_X\mathsf{A}$ be n-slim, where $n \geq 1$, periodic at the ends, and not the ω-power of a word. Then there is a unique $w' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ such that $\lambda_w(w') = w$. Moreover, w' enjoys the following properties:*

(a) if $v' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ is such that $\lambda_w(v') = v \in \mathcal{F}(w)$ and $P(v)$ is a rational language, then so is $P(v')$;
(b) $\bar{r}(w') \leq \bar{r}(w)$;
(c) if $\bar{r}(w)$ is finite then $\bar{r}(w') = \bar{r}(w) - 1$;
(d) if $\bar{r}(w)$ is infinite then so is $\bar{r}(w')$;
(e) there are no infinite anti-chains of factors of w' of rank at most $n-1$.

In particular, w' is $(n-1)$-slim.

Proof. The existence and uniqueness of w' is given, respectively, by Lemmas 9.7 and 9.10.

Part (a) follows from Lemma 9.14(b) for the case where $v' = w'$ since the language L of that lemma is finite. The general case is obtained taking into account Theorem 8.7 and Lemma 9.11.

Since λ_w is a homomorphism, the same result implies that a strict $\mathcal{J}$-chain of idempotent factors of w' maps under λ_w to a strict $\mathcal{J}$-chain of idempotent factors of w, which proves (b). Moreover, we can always add at the top of such a chain an idempotent of the form z^ω, where z is a Lyndon word, for instance by choosing the initial vertex of the topmost idempotent in the given chain.

Conversely, suppose that $e_1 >_{\mathcal{J}} e_2 >_{\mathcal{J}} \cdots >_{\mathcal{J}} e_n \geq_{\mathcal{J}} w$, where the e_i are idempotents. By Theorem 8.7, each e_i is slim. By Lemmas 9.7 and 9.11, there exist factors e_i' of w' in $\overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ and finite words $x_i, y_i \in X^*$ such that $e_i = x_i\lambda_w(e_i')y_i$ for all i except possibly for $i = 1$ in case e_1 is the ω-power of a Lyndon word. Since e_i is an idempotent, we deduce that

$$x_i\lambda_w(e_i')y_i = e_i = e_i^3 = x_i\lambda_w(e_i')y_ie_ix_i\lambda_w(e_i')y_i.$$

Say by Theorem 2.6, we may cancel the finite words x_i and y_i at the ends to deduce that $\lambda_w(e_i') = \lambda_w(e_i')y_ie_ix_i\lambda_w(e_i')$, which show that $\lambda_w(e_i')$ lies in the same $\mathcal{J}$-class as e_i. Moreover, since $z^\omega y_ie_ix_it^\omega$ is still a factor of w, where e_i' is an edge $t^\omega \to z^\omega$, there is some $v_i' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ such that $\lambda_w(e_i'v_i'e_i') = \lambda_w(e_i')$. By Lemma 9.10, it follows that $e_i'v_i'e_i' = e_i'$, and so the idempotent $e_i'v_i'$ is $\mathcal{J}$-equivalent to e_i'. By Lemma 9.12, we have a chain $\big(e_1' >_{\mathcal{J}}\big)\, e_2' >_{\mathcal{J}} \cdots >_{\mathcal{J}} e_n' \geq_{\mathcal{J}} w'$, which proves both (c) and (d). Part (e) follows taking also into account Lemma 9.12. $\square$

The following is an elementary observation about the Dyck language which we need in the sequel.

Lemma 10.2. *Suppose that L is a rational language contained in the Dyck language $\bigcup_n D_n$, where $D_0 = \{1\}$, $D_{n+1} = (aD_nb)^*$ $(n \geq 0)$. Then $L \subseteq D_n$ for some n.*

Proof. Since L is rational, L is recognized by a finite deterministic trim automaton, say with m states and initial state q_0. We claim that, given any state q, the difference $|w|_a - |w|_b$ is constant for w in the language $\{w \in \{a,b\}^* : q_0 w = q\}$. Indeed, if w_1 and w_2 are two words in this language and v is any word such that qv is a terminal state, then $w_1 v$ and $w_2 v$ both belong to L and so $|w_i|_a - |w_i|_b = |v|_b - |v|_a$ $(i = 1,2)$. Since every state is reachable from q_0 by a path of length at most $m-1$, it follows that $L \subseteq D_{m-1}$. □

Iterating the application of Proposition 10.1, we obtain the following characterization of the elements of $\overline{\Omega}_X\mathsf{A}$ which are given by ω-terms in terms of properties of their sets of factors, which is the main result of this paper.

Theorem 10.3. *Let $w \in \overline{\Omega}_X\mathsf{A}$. Then $w \in \Omega^\kappa_X\mathsf{A}$ if and only if (a) there are no infinite anti-chains of factors of w and (b) $P(w)$ is a rational language.*

Proof. Suppose first that $w \in \Omega^\kappa_X\mathsf{A}$. Then w is slim by Corollary 7.6 and $\bar{r}(w) = r(w) < \infty$ by Propositions 4.4 and 4.5. By Lemmas 9.7 and 9.13, there exist $w' \in \Omega^\kappa_{\Gamma(w)}\mathsf{gA}$ and $x, y \in X^*$ such that $w = x\lambda_w(w')y$. Moreover, $\bar{r}(w') < \bar{r}(w)$ by Proposition 10.1(c), so that, taking into account Lemma 9.1 and Proposition 9.8, we may view w' as an ω-word of smaller rank over a finite alphabet. Proceeding by induction on the rank, we may assume that $P(w')$ is a rational language. By Lemmas 8.11, 7.5, and 9.14(a), it follows that $P(w)$ is also a rational language.

To establish the converse, we first observe that, since $P(w)$ is rational, so is the language which is obtained from it by erasing all letters of X. By Lemma 10.2, it follows that the elements of $\mathcal{F}(w) \cap \Omega^\kappa_X\mathsf{A}$ have bounded rank. On the other hand, since the two finiteness hypotheses imply in particular that w is slim, by iterated application of Proposition 10.1 taking into account Corollary 8.12 Proposition 9.8, we obtain a sequence starting with the pseudoword $w_0 = w$, and continuing with edges $w_{k+1} \in \overline{\Omega}_{\Gamma(w_k)}\mathsf{gA}$, and words $x_0, y_0 \in X^*$ and paths x_k, y_k in $\Gamma(w_{k-1})$ such that

$$w_k = x_k \lambda_{w_k}(w_{k+1}) y_k, \tag{16}$$

and this sequence can be constructed while $k < \bar{r}(w)$, each term being slim. Now, for w_n we can determine factors of w which are ω-words of rank n: for $n > 1$, if it is not the last term in the sequence, then simply take any edge in $\Gamma(w_n)$ and successively apply to it $\lambda_{w_n}, \lambda_{w_{n-1}}, \ldots, \lambda_{w_0}$ to obtain an ω-term of rank $n+1$ which is a factor of w. Now, by the hypothesis that $P(w)$ is rational and the argument at the beginning of this paragraph, the sequence must stop at some point, that is the graph is trivial, meaning that w_n has no special factors. Since w_n is slim, by Lemma 8.14 and Corollary 8.12, we deduce that either w_n is finite or w_n is a rank 1 ω-word of the form $tu^\omega v$, and therefore trivially an ω-word. Applying successively equation (16) for $k = n-1, n-2, \ldots, 0$, it follows that w is also an ω-word. □

Note that, for a singleton alphabet, both conditions (a) and (b) of Theorem 10.3 are trivially verified. For an alphabet X with at least two letters, there are uniformly recurrent non-periodic pseudowords $w \in \overline{\Omega}_X\mathsf{A}$. Since w is then $\mathcal{J}$-maximal among infinite pseudowords, the very definition of uniformly recurrent pseudoword entails condition (a), whereas condition (b) fails by Theorem 6.2. We conjecture that condition (a) is always a consequence of condition (b). The following result proposes a considerable reduction in the proof of this conjecture.

Proposition 10.4. *Suppose that, for every finite alphabet X and every pseudoword $w \in \overline{\Omega}_X\mathsf{A}$ such that $P(w)$ is a rational language, there are no infinite anti-chains of factors of w of rank at most 1. Then, every pseudoword $w \in \overline{\Omega}_X\mathsf{A}$ such that $P(w)$ is a rational language is an ω-word.*

Proof. Let $w \in \overline{\Omega}_X\mathsf{A}$ be a pseudoword such that $P(w)$ is a rational language. It has already been observed that the rationality of $P(w)$ implies that the language $F(w)$ is also rational. Note that, by Lemma 10.2, there is an upper bound on the rank of the members of $P(w)$. Let $M(w)$ be the least such upper bound. We proceed by induction on $M(w)$ to show that w is an ω-word. If w is finite, then it is in particular an ω-word. Assume that w is infinite and, inductively, that, for every pseudoword v such that $P(v)$ is rational and $M(v) < M(w)$, v is an ω-word.

By hypothesis, w is 1-slim and infinite. Hence, by Corollary 8.12 and Proposition 10.1, w may be encoded as $w = x\lambda_w(w')y$ for a unique edge $w' \in \overline{\Omega}_{\Gamma(w)}\mathsf{gA}$ and some words $x, y \in X^*$ where w' is such that $P(w')$ is rational. Moreover, by Proposition 10.1(c) and taking into account Lemma 9.11, the ranks of the elements of $P(w')$ are bounded above by $M-1$. Hence, by the induction hypothesis, w' is and ω-word and, therefore, so is w. This completes the induction step and the proof of the proposition. □

With a little extra effort, tracing through sections 8–10 where the anti-chain hypothesis has been used, one can actually weaken the hypothesis in Proposition 10.4 to the following condition: if $P(w)$ is rational, then there are no infinite anti-chains of factors of w which are finite, idempotents or bridges.

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 author was supported, in part, by *Fundação para a Ciência e a Tecnologia* (FCT) through the *Centro de Matemática da Universidade do Porto*, by FCT. The work of the second author was supported, in part, by FCT through the *Centro de Matemática da Universidade do Minho*. Both the first and second authors were also supported by 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, *Finite semigroups and universal algebra*, World Scientific, Singapore, 1995, English translation.
2. ———, *A syntactical proof of locality of DA*, Int. J. Algebra Comput. **6** (1996), 165–177.
3. ———, *Hyperdecidable pseudovarieties and the calculation of semidirect products*, Int. J. Algebra Comput. **9** (1999), 241–261.
4. ———, *Profinite groups associated with weakly primitive substitutions*, Fundamentalnaya i Prikladnaya Matematika (Fundamental and Applied Mathematics) **11** (2005), no. 3, 13–48, In Russian. English version in J. Math. Sciences **144**, No. 2 (2007) 3881–3903.
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 and A. Costa, *Infinite-vertex free profinite semigroupoids and symbolic dynamics*, Tech. Report CMUP 2007-20, Univ. Porto, 2007.
7. J. Almeida, J. C. Costa, and M. Zeitoun, *Complete reducibility of systems of equations with respect to R*, Portugal. Math. **64** (2007), 445–508.
8. J. Almeida and M. Delgado, *Sur certains systèmes d'équations avec contraintes dans un groupe libre*, Portugal. Math. **56** (1999), 409–417.
9. ———, *Sur certains systèmes d'équations avec contraintes dans un groupe libre—addenda*, Portugal. Math. **58** (2001), 379–387.
10. J. Almeida and B. Steinberg, *On the decidability of iterated semidirect products and applications to complexity*, Proc. London Math. Soc. **80** (2000), 50–74.
11. ———, *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.
12. J. Almeida and M. V. Volkov, *Subword complexity of profinite words and subgroups of free profinite semigroups*, Int. J. Algebra Comput. **16** (2006), 221–258.
13. J. Almeida and P. Weil, *Profinite categories and semidirect products*, J. Pure Appl. Algebra **123** (1998), 1–50.
14. 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.
15. J.-C. Birget, *Iteration of expansions—unambiguous semigroups*, J. Pure Appl. Algebra **34** (1984), 1–55.
16. S. Eilenberg, *Automata, languages and machines*, vol. B, Academic Press, New York, 1976.
17. M. Gehrke, S. Grigorieff, and J.-E. Pin, *Duality and equational theory of regular languages*, ICALP 2008, Part II, Lecture Notes in Computer Science, no. 5126, Springer Verlag, 2008, pp. 246–257.
18. B. Herwig and D. Lascar, *Extending partial automorphisms and the profinite topology on free groups*, Trans. Amer. Math. Soc. **352** (2000), 1985–2021.
19. K. Krohn and J. Rhodes, *Complexity of finite semigroups*, Ann. of Math. (2) **88** (1968), 128–160.
20. G. Lallement, *Semigroups and combinatorial applications*, Wiley-Interscience, J. Wiley & Sons, Inc., New York, 1979.
21. M. Lothaire, *Combinatorics on words*, Addison-Wesley, Reading, Mass., 1983.
22. ———, *Algebraic combinatorics on words*, Cambridge University Press, Cambridge, UK, 2002.
23. 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.
24. ———, *Normal forms for free aperiodic semigroups*, Int. J. Algebra Comput. **11** (2001), 581–625.
25. J. D. McKnight, Jr. and A. J. Storey, *Equidivisible semigroups*, J. Algebra **12** (1969), 24–48.
26. J. Reiterman, *The Birkhoff theorem for finite algebras*, Algebra Universalis **14** (1982), 1–10.

27. J. Rhodes and B. Steinberg, *The q-theory of finite semigroups*, Springer Monographs in Mathematics, Springer, 2009, To appear (ISBN: 978-0-387-09780-0).
28. L. Ribes and P. A. Zalesskiĭ, *On the profinite topology on a free group*, Bull. London Math. Soc. **25** (1993), 37–43.
29. ______, *Profinite groups*, Ergeb. Math. Grenzgebiete 3, no. 40, Springer, Berlin, 2000.
30. H. Straubing, *Finite semigroup varieties of the form* $V * D$, J. Pure Appl. Algebra **36** (1985), 53–94.
31. B. Tilson, *Categories as algebra: an essential ingredient in the theory of monoids*, J. Pure Appl. Algebra **48** (1987), 83–198.

(Jorge Almeida) Centro de Matemática e Departamento de Matemática Pura, Faculdade de Ciências,, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal.
E-mail address: jalmeida@fc.up.pt

(José Carlos Costa) Centro de Matemática, Universidade do Minho, Campus de Gualtar, 4700-320 Braga, Portugal.
E-mail address: jcosta@math.uminho.pt

(Marc Zeitoun) LaBRI, Université Bordeaux & CNRS UMR 5800. 351 cours de la Libération, 33405 Talence Cedex, France.
E-mail address: mz@labri.fr