Pro-Aperiodic Monoids via Saturated Models*

Samuel J. v. Gool[1] and Benjamin Steinberg[2]

1 **Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam, The Netherlands; and Department of Mathematics, City College of New York, New York, NY, USA**
`samvangool@me.com`

2 **Department of Mathematics, City College of New York, New York, NY, USA**
`bsteinberg@ccny.cuny.edu`

Abstract

We apply Stone duality and model theory to study the structure theory of free pro-aperiodic monoids. Stone duality implies that elements of the free pro-aperiodic monoid may be viewed as elementary equivalence classes of pseudofinite words. Model theory provides us with saturated words in each such class, i.e., words in which all possible factorizations are realized. We give several applications of this new approach, including a solution to the word problem for ω-terms that avoids using McCammond's normal forms, as well as new proofs and extensions of other structural results concerning free pro-aperiodic monoids.

1998 ACM Subject Classification F.4.1 Model Theory, F.4.3 Algebraic Language Theory

Keywords and phrases aperiodic monoids, profinite monoids, Stone duality, saturated models

Digital Object Identifier 10.4230/LIPIcs.STACS.2017.39

1 Introduction

The pseudovariety of aperiodic monoids has long played a fundamental role in finite semigroup theory and automata theory. The famous Schützenberger theorem [32] proved that the aperiodic monoids recognize precisely the star-free languages. This class, which also coincides with the class of languages recognizable by counter-free automata, was later shown by McNaughton and Papert [27] to be the class of first-order definable languages, also see Straubing's book [35]. Algorithmic questions about aperiodic languages lead to challenges that the algebraic approach can often help resolve. In particular, aperiodic monoids play a prime role in two of the oldest and most difficult open problems in automata theory: the dot-depth problem [10, 29, 28] and the Krohn-Rhodes complexity problem [23].

Within this algebraic approach to aperiodic languages, *free pro-aperiodic monoids* are a useful tool. The importance of profinite monoids in automata theory and finite semigroup theory was first highlighted, starting in the late eighties, by Almeida [1]; also see the more recent monograph by Rhodes and the second-named author [30], or Straubing and Weil's handbook article [36]. The structure of free pro-aperiodic monoids has been studied recently by several authors, e.g., [4, 20, 22, 6], but many difficult questions remain open. Many of the existing results about free pro-aperiodic monoids are about the submonoid of elements definable by *ω-terms* and rely on an ingenuous normal form algorithm due to

* The first-named author was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant #655941; the second-named author was supported by Simons Foundation #245268, United States – Israel Binational Science Foundation #2012080 and NSA MSP #H98230-16-1-0047.

34th Symposium on Theoretical Aspects of Computer Science (STACS 2017).
Editors: Heribert Vollmer and Brigitte Vallée; Article No. 39; pp. 39:1–39:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

McCammond [25], which solves the word problem for ω-terms. In our new approach to pro-aperiodic monoids we replace the use of normal forms by the model-theoretic concept of saturated word.

Contributions of this paper

Our contributions can be separated into useful theoretical results about pro-aperiodic monoids on the one hand, and applications of this theory to structural and decidability results for pro-aperiodic monoids on the other hand.

The main theoretical results (Section 4) are the following:

- Identification of the free pro-aperiodic monoid with a monoid of elementary equivalence classes of pseudofinite words (Theorem 10), and of homomorphisms with substitutions (Theorem 12).
- Substitution-invariance of ω-saturated words (Theorem 15).
- A classification of the factors of the image of a substitution (Theorem 18).

The main applications of these results (Sections 5–7) are the following:

- Decidability of the word problem for aperiodic ω-terms that avoids McCammond's normal forms (Section 5).
- New, easier proofs of known facts about factors of ω-terms, notably that factors of ω-terms are given by ω-terms (Theorem 22).
- Having a well-quasi-order of factors, prefixes, or suffixes is stable under substitutions (Theorem 23); in particular, ω-terms have these properties (Corollary 26).
- Having a regular language of factors, prefixes, or suffixes is stable under non-erasing substitutions (Theorem 28); in particular, ω-terms have these properties (Corollary 29).

Methodology

Early on in the study of profinite monoids, Almeida [1] made the important observation that the Boolean algebra of clopen subsets of the free pro-**V** monoid can be identified with the Boolean algebra of **V**-recognizable languages; in other words, the free pro-**V** monoid is the *Stone dual* of the Boolean algebra of **V**-recognizable languages. In recent years, a number of authors have made explicit use of Stone duality theory to redevelop and expand the foundations of the profinite approach to studying varieties of languages in the sense of Eilenberg [14]; most closely related to our work here is the work of Gehrke, Pin et al. [16, 17, 18, 15], Bojańczyk [9], and Rhodes and the second-named author [30, Chapter 8].

In this paper we apply Stone duality in the case where **V** is the variety of aperiodic monoids to study the structure theory of free pro-aperiodic monoids. The crucial idea is that we can view elements of the free pro-aperiodic monoid as complete first-order theories containing the theory of finite words. By the completeness theorem of first-order logic, these in turn can be identified with elementary equivalence classes of so-called pseudofinite words, models of the theory of finite words. These models can be concatenated in a natural way, allowing us to recover the algebraic structure as well as the topological structure from this approach. See Section 3 for details.

In addition to Stone duality, the correspondence between aperiodic monoids and first-order logic allows us to import techniques from *first-order model theory*, in particular *Ehrenfeucht-Fraïssé games*, which are commonly used in logic on words, and more importantly *ω-saturated models*. In our context, an ω-saturated model is a word where every possible factorization is realized in the word itself; see Section 4 for details. These ω-saturated words allow us to perform an explicit combinatorial analysis of the factors of elements of pro-aperiodic monoids.

We leave model-theoretic prerequisites to a minimum; for the precise correspondence to the general model-theoretic concepts and other technical proofs, we refer the reader to our extended technical report [19], where we provide more background information than space permits here.

2 Logic on words

We briefly recall the necessary preliminaries on logic on words. We mostly follow the standard terminology as in, e.g., [35].

Words. In what follows, $A, B, \dots$ denote non-empty finite alphabets. In this paper, by a *word* over A, or A-*word*, we mean a tuple $W = (|W|, <^W, (P_a^W)_{a\in A})$, where

- $|W|$ is a set,
- the relation $<^W$ is a discrete linear order with endpoints on $|W|$, i.e., there are a first and a last element, every element except the last has a unique immediate $<^W$-successor and every element except the first has a unique immediate $<^W$-predecessor,
- $(P_a^W)_{a\in A}$ is a partition of $|W|$.

An A-word is called *finite* if $|W|$ is finite. We denote by ε the unique A-word with $|W| = \emptyset$.

Let W be an A-word. For any $i \in |W|$, we write $W(i)$ for the unique letter a such that $i \in P_a^W$; we write $W(<i)$ for the A-word obtained by restricting W to the set of positions strictly less than i (the *ray* left of i), and similarly we define $W(>i)$ (the *ray* right of i). For $i, j \in |W|$, we write $W(i, j)$ for the *open interval*, $W[i, j]$ for the *closed interval* of W between these positions, and *half-open intervals* $W[i, j)$ and $W(i, j]$. Since the order on an A-word is discrete with endpoints, any (half-)open interval or ray can be written as a closed interval; we use this fact without mention in what follows.

Logic. An *atomic formula* is an expression of the form $x < y$ or $P_a(x)$, where, here and in what follows, x and y denote first-order variables and a denotes a letter from the alphabet. A *first-order formula* is an expression built up from atomic formulas by inductively applying the connectives $\wedge$, $\vee$, $\neg$, $\to$, $\exists x$, and $\forall x$. The notation $\varphi(\overline{x})$ indicates that all the variables that occur freely in φ lie in $\overline{x}$. A *sentence* is a formula without free occurrences of variables.

For a first-order formula $\varphi(\overline{x})$, a word W and an assignment $\overline{x}^W$ of the variables $\overline{x}$ in $|W|$, one defines $W, \overline{x}^W \models \varphi(\overline{x})$ if $\varphi(\overline{x})$ is true in W under the assignment $\overline{x}^W$. The *language defined by an FO-sentence* φ is the set L_φ of *finite* A-words W such that $W \models \varphi$.

The first-order definable languages form a strict subclass of the regular languages. Indeed, recall [32, 27] that *a language L of finite A-words is first-order definable if and only if the syntactic monoid of L is finite and aperiodic.* This fundamental result is the starting point for our perspective on the free pro-aperiodic monoid, cf. Section 4 below.

Quantifier depth and games. The *quantifier depth* of a formula φ is the maximum nesting depth of quantifiers in φ. If U and V are A-words, we write $U \equiv_k V$ if U and V satisfy exactly the same first-order sentences φ of quantifier depth at most k; in this case, we say that U and V are *elementarily equivalent up to quantifier depth k* or simply *k-equivalent.* We write $U \equiv V$ if $U \equiv_k V$ for all k, i.e., U and V satisfy exactly the same first-order sentences; in this case, we say that U and V are *elementarily equivalent.* Importantly, for each $k \geq 0$, there are only *finitely* many k-equivalence classes, and each such class is definable by a first-order sentence [21, Thm. 3.3.2], see also [35, Sec. IV.1].

One can alternatively describe k-equivalence using Ehrenfeucht-Fraïssé games, cf., e.g., [21, Ch. 3] or [35, Sec. IV.1]. These games allow us to prove the following lemma. If U and V are A-words, $i \in |U|$ and $j \in |V|$, we say that the position i in U *k-corresponds* to the position j in V provided that $U(<i) \equiv_k V(<j)$, $U(i) = V(j)$, and $U(>i) \equiv_k V(>j)$.

▶ **Lemma 1.** *Let U and V be A-words. For all $k \geq 0$, $U \equiv_{k+1} V$ if and only if for every $i \in |U|$, there exists $j \in |V|$ that k-corresponds to i, and for every $j \in |V|$, there exists $i \in |U|$ that k-corresponds to j.*

3 Pro-aperiodic monoids as Stone dual spaces

In this section, we define free pro-aperiodic monoids and a new object $\Lambda(A)$, and identify them with spaces of elementary equivalence classes of models.

Pro-aperiodic monoids. A *profinite monoid* is an inverse limit of finite discrete monoids in the category of topological monoids (i.e., monoids whose underlying set is equipped with a topology in which the monoid operation is continuous). Equivalently, a profinite monoid is a topological monoid whose underlying space is a *Boolean* or *Stone space*, i.e., compact, Hausdorff and zero-dimensional. Profinite monoids inherit many properties of finite monoids. Of particular interest to us is the fact that any element x in a profinite monoid has a unique idempotent, denoted x^ω, in its orbit-closure $\overline{\{x^n \mid n \geq 1\}}$. A *pro-aperiodic* monoid M is a profinite monoid in which $x^\omega = x^\omega x$ for all $x \in M$. Equivalently, a pro-aperiodic monoid is an inverse limit of finite aperiodic monoids; here, finite monoids are equipped with the discrete topology, and the inverse limit is taken in the category of topological monoids.

The *free pro-aperiodic monoid* generated by a finite set A is a pro-aperiodic monoid $\widehat{F}_{\mathbf{A}}(A)$ containing A such that any function $f\colon A \to M$, with M a finite aperiodic monoid, extends uniquely to a continuous homomorphism $\overline{f}\colon \widehat{F}_{\mathbf{A}}(A) \to M$, where M is given the discrete topology. The free pro-aperiodic monoid is unique up to topological isomorphism, and the same extension property still holds if in the previous sentence M is replaced by an arbitrary pro-aperiodic monoid.

Let M be a monoid and u, v elements of M. Then $u \leq_{\mathcal{J}} v$ means that there exist x and y such that $u = xvy$, and in this case v is called a *factor* of u; $u \leq_{\mathcal{L}} v$ means that there exists x such that $u = xv$, and in this case v is called a *suffix* of u; $u \leq_{\mathcal{R}} v$ means that there exists y such that $u = vy$, and in this case v is called a *prefix* of u. Each of these relations is a quasi-order on M; the equivalence relations they induce are denoted $\mathcal{J}$, $\mathcal{L}$ and $\mathcal{R}$, e.g., $u \mathrel{\mathcal{J}} v$ means that $u \leq_{\mathcal{J}} v$ and $v \leq_{\mathcal{J}} u$.

Stone duality. We make use of *Stone duality*, which, we briefly recall, is the dual equivalence between the categories of Boolean algebras and Boolean spaces that takes a Boolean space X to its algebra $\mathcal{K}(X)$ of clopen sets, and a Boolean algebra B to the set of ultrafilters $\mathrm{Spec}(B)$ of B, which is given a Boolean topology by declaring, for each $L \in B$, the set $\widehat{L} := \{x \in \mathrm{Spec}(B) \mid L \in x\}$ to be open. Stone's duality theorem [34] says that the assignment $L \mapsto \widehat{L}$ is an isomorphism from B to $\mathcal{K}(\mathrm{Spec}(B))$, and that moreover Boolean algebra homomorphisms $B_1 \to B_2$ are in natural bijection with continuous functions $\mathrm{Spec}(B_2) \to \mathrm{Spec}(B_1)$.

Spaces of first-order theories. Recall that, for T a set of sentences in first-order logic, the *Lindenbaum-Tarski algebra* $\mathrm{LT}(T)$ of T is the Boolean algebra of T-equivalence classes of first-order sentences. Here, two first-order sentences φ and ψ are *T-equivalent* if, in any model W where all sentences in T are true, φ and ψ are either both true or both false. Stone dual spaces of Lindenbaum-Tarski algebras are well understood in logic.

▶ **Proposition 2.** *Let T be a set of first-order sentences. The Stone dual space of* $\mathrm{LT}(T)$ *is homeomorphic to the Boolean space X whose points are elementary equivalence classes of models of T, in which the clopen sets are exactly the* truth sets
$\widehat{\varphi} := \{x \in X \mid \varphi$ *is true in the models in the class* $x\}$*, for φ any first-order sentence.*

Theories of A-words. For a finite alphabet A, we denote by T_A the (finitely axiomatized) *theory of A-words*, that is, the set of first-order sentences deducible from axioms expressing that the order is a discrete linear order with endpoints, and that exactly one letter predicate holds at each position. We further let T_A^{fin} denote the *theory of finite A-words*, i.e., the set of first-order sentences that are true in all finite A-words. A model of the theory T_A^{fin} is called a *pseudofinite A-word*.[1] The theories T_A and T_A^{fin} do not coincide in general; in fact, they coincide only if the alphabet A contains a single letter. In this case, both theories T_A and T_A^{fin} are the theory of discrete linear orders with endpoints, with a unary predicate that is true everywhere.

As soon as A contains at least two letters, the situation is very different. In particular, there are A-words that are not pseudofinite.

▶ **Example 3.** Let W be the word $aaaa\ldots\ \ldots bbbb$, i.e., W is the word over the alphabet $\{a, b\}$ with underlying order $\mathbb{N} + \mathbb{N}^{\mathrm{op}}$, where $W(i) = a$ for all $i \in \mathbb{N}$ and $W(i) = b$ for all $i \in \mathbb{N}^{\mathrm{op}}$. The sentence $\exists x P_a(x) \to \exists x(P_a(x) \wedge \forall y(y > x \to \neg P_a(y))$, expressing 'if there exists an a-position, then there exists a last such' is true in every finite A-word, and therefore lies in T_A^{fin}, but it fails to hold in W. Thus, W is not pseudofinite.

▶ Remark. In [19, Sec. 4], we discuss the axiomatizability of the theory T_A^{fin} of finite A-words. In particular, we show that the theory is not finitely axiomatizable, and we give an axiomatization of it using an axiom scheme similar to the one given by Doets [13].

Pro-aperiodic monoids as spaces of theories. We will denote by $\Lambda(A)$ the Stone dual space of $\mathrm{LT}(T_A)$. By Proposition 2, $\Lambda(A)$ is the space of elementary equivalence classes of A-words. Combining this with the characterization theorem [32, 27] that first-order languages are exactly the aperiodic-recognizable languages, we obtain the following lemma and proposition.

▶ **Lemma 4.** *The Lindenbaum-Tarski algebra* $\mathrm{LT}(T_A^{\mathrm{fin}})$ *of the theory* T_A^{fin} *is isomorphic to the algebra* $\mathrm{Rec}_{\mathbf{A}}(A)$ *of aperiodic-recognizable languages of finite A-words.*

▶ **Proposition 5.** *The subspace of* $\Lambda(A)$ *consisting of the elementary equivalence classes of pseudofinite A-words is homeomorphic to the space underlying the free pro-aperiodic monoid over A.*

We will see in the next section that there is a natural multiplication on $\Lambda(A)$ which makes it into a pro-aperiodic monoid, and makes the homeomorphism of Proposition 5 into a topological isomorphism of pro-aperiodic monoids.[2]

[1] This is an instance of the general model-theoretic use of the term 'pseudofinite', cf., e.g., [37].
[2] Here and in what follows, we use the term 'homeomorphism' to indicate an isomorphism in the category of topological spaces, and 'topological isomorphism' for an isomorphism in the category of topological monoids.

4 Substitutions, saturated words, and factorizations

This is the main theoretical section of the paper. We introduce substitutions and saturated words and state our main results about them.

Substitutions

Suppose that V is a word over a finite alphabet B, and that for each $b \in B$, U_b is a word over a finite alphabet A. The *substitution* of the A-words $(U_b)_{b\in B}$ into the B-word V is the A-word $W = V[b/U_b]$ defined as follows.

- The underlying order of W is the *lexicographic order* on the disjoint union $|W| := \bigsqcup_{i\in|V|} |U_{V(i)}|$, i.e., $(i,j) <^W (i',j') \overset{\text{def}}{\iff} i <^V i'$, or $i = i'$ and $j <^{U_{V(i)}} j'$.
- The letter at position (i,j) in W is the letter at position j in $U_{V(i)}$.

There are two important special cases of substitution. If U_0 and U_1 are A-words, then the *concatenation* $U_0 \cdot U_1$ of U_0 and U_1 is defined as the substitution of $(U_b)_{b\in\{0,1\}}$ into the $\{0,1\}$-word 01. If U is an A-word and λ is a discrete linear order with endpoints, then the *λ-power* U^λ of U is defined as the substitution of $U_b = U$ into the unique $\{b\}$-word with underlying order λ.

Importantly, the operation of substitution respects the equivalence relations $\equiv_k$.

▶ **Proposition 6.** *Let $k \geq 0$. For any finite alphabets A and B and any B-indexed collections of A-words $(U_b)_{b\in B}$ and $(U'_b)_{b\in B}$ such that $U_b \equiv_k U'_b$ for each $b \in B$, if V and V' are A-words such that $V \equiv_k V'$, then $V[b/U_b] \equiv_k V'[b/U'_b]$.*

▶ **Corollary 7.** *If $V \equiv V'$ and $U_b \equiv U'_b$ for each $b \in B$, then $V[b/U_b] \equiv V'[b/U'_b]$.*

Pro-aperiodic monoids of elementary equivalence classes

Corollary 7 in particular implies that there is a well-defined binary operation of concatenation on the set $\Lambda(A)$ of elementary equivalence classes of A-words.

▶ **Theorem 8.** *The Stone space $\Lambda(A)$, equipped with the operation of concatenation up to elementary equivalence, is a pro-aperiodic monoid. The ω-power of an element $[U]_\equiv$ is $[U^\lambda]_\equiv$, where λ is any infinite discrete linear order with endpoints.*

We showed in Proposition 5 that the space underlying the free pro-aperiodic monoid, $\widehat{F}_{\mathbf{A}}(A)$, is homeomorphic to the subspace of $\Lambda(A)$ consisting of the elementary equivalence classes of pseudofinite A-words. Indeed, we can prove more.

▶ **Proposition 9.** *The set of elementary equivalence classes of pseudofinite A-words is a topologically closed submonoid of $\Lambda(A)$ which is closed under taking factors.*

We now establish our first main theoretical contribution.

▶ **Theorem 10.** *Let A be a finite alphabet. The free pro-aperiodic monoid over A is topologically isomorphic to the pro-aperiodic monoid of elementary equivalence classes of pseudofinite A-words.*

▶ **Convention.** In view of Theorem 10, we henceforth identify the free pro-aperiodic monoid $\widehat{F}_{\mathbf{A}}(A)$ with the closed submonoid $\mathrm{PF}(A)$ of $\Lambda(A)$ consisting of the elementary equivalence classes of pseudofinite A-words.

▶ Remark. The pro-aperiodic monoids $\Lambda(A)$ and $\widehat{F}_{\mathbf{A}}(A)$ can also be usefully described as inverse limits of chains of finite monoids of k-equivalence classes, cf. [19, Sec. 3].

▶ **Proposition 11.** *Let $(U_b)_{b \in B}$ be a B-indexed collection of A-words. The function $f \colon \Lambda(B) \to \Lambda(A)$, which sends an element $[V]_\equiv$ of $\Lambda(B)$ to $[V[b/U_b]]_\equiv$, is a well-defined continuous homomorphism.*

We call a continuous homomorphism $f \colon \Lambda(B) \to \Lambda(A)$ a *substitution* if it arises as in Proposition 11. Restricting attention to the free pro-aperiodic monoids, we obtain the following theorem.

▶ **Theorem 12.** *The continuous homomorphisms from $\widehat{F}_{\mathbf{A}}(B)$ to $\widehat{F}_{\mathbf{A}}(A)$ are exactly the substitutions of pseudofinite A-words into pseudofinite B-words.*

Saturated words

Let A be a finite alphabet and let U be an A-word. For any position $i \in |U|$, define the triple $t^U(i) := ([U({<}i)]_\equiv, U(i), [U({>}i)]_\equiv)$, an element of the Cartesian product $\Lambda(A) \times A \times \Lambda(A)$. We call $t^U(i)$ the (complete 1-)*type* of i in U (see [19, Prop. A.3] for the correspondence with types in model theory). We refer to the product space $\Lambda(A) \times A \times \Lambda(A)$, where the middle component A has the discrete topology, as the *type space*.

We write $\mathrm{RT}(U)$ for the *set of types realized in* U, i.e., $\mathrm{RT}(U)$ is the subset $\{t^U(i) \mid i \in |U|\}$ of the type space. If V is an A-word elementarily equivalent to U and $j \in |V|$, then we say the type $t^V(j)$ is *consistent with* U. We write $\mathrm{CT}(U)$ for the *set of types consistent with* U, i.e., $\mathrm{CT}(U)$ is the subset $\{t^V(j) \mid V \equiv U, j \in |V|\}$ of the type space.

▶ **Definition 13.** Let U be an A-word. We say that U is

- *weakly saturated* if every type consistent with U is realized in U, i.e., $\mathrm{CT}(U) = \mathrm{RT}(U)$;
- *ω-saturated* if every closed interval $U[i,j]$ in U is weakly saturated;
- *countably saturated* if the underlying set $|U|$ is countable and U is ω-saturated.

It is well-known in model theory that any elementary equivalence class contains an ω-saturated model, which typically has an uncountable underlying set; also see [19, Prop. A.5].

▶ **Example 14.** Consider the case of a one-letter alphabet, $\{a\}$. All infinite $\{a\}$-words are elementarily equivalent and pseudofinite. Hence, $\widehat{F}_{\mathbf{A}}(\{a\}) = \Lambda(\{a\})$ is topologically isomorphic to the topological monoid $\mathbb{N} \cup \{\omega\}$, i.e., the one-point compactification of $\mathbb{N}$ with the usual addition, where ω is an absorbing element. The space of types of $\{a\}$-words is $\widehat{F}_{\mathbf{A}}(\{a\}) \times \{a\} \times \widehat{F}_{\mathbf{A}}(\{a\})$. Concretely, types of $\{a\}$-words are of the following four forms:

- (a^n, a, a^m) for $n, m \in \mathbb{N}$;
- (a^n, a, a^ω) for $n \in \mathbb{N}$;
- (a^ω, a, a^m) for $m \in \mathbb{N}$;
- (a^ω, a, a^ω).

Consider the following infinite $\{a\}$-words:

1. $W_1 := a^{\mathbb{N} + \mathbb{N}^{\mathrm{op}}}$,
2. $W_2 := a^{\mathbb{N} + \mathbb{Z} + \mathbb{N}^{\mathrm{op}}}$,
3. $W_3 := a^{\mathbb{N} + \mathbb{Q} \times \mathbb{Z} + \mathbb{N}^{\mathrm{op}}}$.

The word W_1 is not weakly saturated, because the elementarily equivalent word W_2 realizes the type (a^ω, a, a^ω), which is not realized in W_1, that is, $(a^\omega, a, a^\omega) \in \mathrm{CT}(W_1) \setminus \mathrm{RT}(W_1)$. The word W_2 is weakly saturated, because it realizes all the types. However, W_2 is not ω-saturated, because the ray to the left of i, where i is any point in the summand $\mathbb{Z}$, is isomorphic to W_1, and not weakly saturated. Notice that any closed interval in the word in W_3 is either finite or isomorphic to W_3, using the well-known fact that any open interval in the order $\mathbb{Q}$ is isomorphic to $\mathbb{Q}$ (cf. e.g., [21, p. 100]). Since finite words and W_3 are weakly saturated, the word W_3 is in fact ω-saturated. Since $|W_3|$ is countable, W_3 is countably saturated.

The key technical result about saturated words is the following.

▶ **Theorem 15.** *If V is an ω-saturated B-word and $(U_b)_{b\in B}$ is a B-indexed collection of A-words that are ω-saturated, then $V[b/U_b]$ is ω-saturated.*

Note that Theorem 15 has the following immediate consequence.

▶ **Corollary 16.** *The concatenation of two ω-saturated words is ω-saturated. The power of an ω-saturated word by an ω-saturated linear order is ω-saturated.*

Factorizations

A useful fact about ω-saturated words is that they realize any finite factorization of their elementary equivalence class.

▶ **Lemma 17.** *Let W be an ω-saturated A-word and suppose that $W \equiv W_1 \cdots W_n$ with $W_1, \dots, W_n$ non-empty A-words. Then we can find positions $i_1 < i_2 < \cdots < i_{n-1}$ in $|W|$ such that $W(<i_1) \equiv W_1$, $W(\geq i_{n-1}) \equiv W_n$ and $W[i_j, i_{j+1}) \equiv W_{j+1}$ for $1 \leq j \leq n-2$.*

As a first application of Theorem 15, we will now analyze what factors of the result of a substitution can look like in $\Lambda(A)$ (and hence, by Proposition 9, in $\widehat{F}_{\mathbf{A}}(A)$).

▶ **Theorem 18.** *Let $f\colon \Lambda(B) \to \Lambda(A)$ be a substitution. Let $v \in \Lambda(B)$ and $w \in \Lambda(A)$.*

1. *$f(v) \leq_{\mathcal{J}} w$ if and only if one of the following holds:*
   (a) *$w = \varepsilon$, or*
   (b) *there exists $b \in B$ such that $v \leq_{\mathcal{J}} b$ and $f(b) \leq_{\mathcal{J}} w$, or*
   (c) *there exist $b_1, b_2 \in B$, $x, y \in \Lambda(A)$ and $z \in \Lambda(B)$ such that $v \leq_{\mathcal{J}} b_1 z b_2$, $f(b_1) \leq_{\mathcal{L}} x$, $f(b_2) \leq_{\mathcal{R}} y$, and $w = x f(z) y$.*
2. *$f(v) \leq_{\mathcal{R}} w$ if and only if $w = \varepsilon$ or there exist $x \in \Lambda(A)$, $b \in B$, and $z \in \Lambda(B)$ such that $f(b) \leq_{\mathcal{R}} x$, $v \leq_{\mathcal{R}} zb$, and $w = f(z)x$.›*
3. *$f(v) \leq_{\mathcal{L}} w$ if and only if $w = \varepsilon$ or there exist $y \in \Lambda(A)$, $c \in B$, and $z \in \Lambda(B)$ such that $f(c) \leq_{\mathcal{L}} y$, $v \leq_{\mathcal{L}} cz$, and $w = y f(z)$.*

Theorem 18 can be viewed as extension of [7, Lemma 8.2], where the special case that v is a finite word is handled in the context of the free profinite monoid.

We may usefully summarize Theorem 18 by recalling some more notation. For any $u \in \Lambda(A)$, write $\uparrow_{\mathcal{J}} u := \{w \in \Lambda(A) \mid w \geq_{\mathcal{J}} u\}$ for the set of factors of u, and similarly $\uparrow_{\mathcal{R}} u$ for the set of prefixes of u, and $\uparrow_{\mathcal{L}} u$ for the set of suffixes of u. Also, for any subset L of $\Lambda(A)$ and $a, b \in A$, write $a^{-1}L := \{u \in \Lambda(A) \mid au \in L\}$ and $Lb^{-1} := \{u \in \Lambda(A) \mid ub \in L\}$, and $a^{-1}Lb^{-1} := \{u \in \Lambda(A) \mid aub \in L\}$. Finally, if $u \in \Lambda(A)$, write $\mathcal{C}(u) := \{a \in A \mid u \leq_{\mathcal{J}} a\}$ for the *content* of u. The following restates Theorem 18 in this notation.

▶ **Corollary 19.** *Let $f\colon \Lambda(B) \to \Lambda(A)$ be a substitution. For any $v \in \Lambda(B)$, we have:*

$$\uparrow_{\mathcal{J}} f(v) = \{\varepsilon\} \cup \bigcup_{b \in \mathcal{C}(v)} (\uparrow_{\mathcal{J}} f(b)) \cup \bigcup_{b_1, b_2 \in B} \left[\uparrow_{\mathcal{L}} f(b_1) \cdot f\left(b_1^{-1} \left(\uparrow_{\mathcal{J}} v\right) b_2^{-1}\right) \cdot \uparrow_{\mathcal{R}} f(b_2)\right],$$

$$\uparrow_{\mathcal{R}} f(v) = \{\varepsilon\} \cup \bigcup_{b \in B} f\left(\left(\uparrow_{\mathcal{R}} v\right) b^{-1}\right) \cdot \left(\uparrow_{\mathcal{R}} f(b)\right),$$

$$\uparrow_{\mathcal{L}} f(v) = \{\varepsilon\} \cup \bigcup_{c \in B} \left(\uparrow_{\mathcal{L}} f(c)\right) \cdot f\left(c^{-1}\left(\uparrow_{\mathcal{L}} v\right)\right).$$

▶ Remark. Several known structural results about $\widehat{F}_{\mathbf{A}}(A)$ are easy to prove using the saturated-models approach, as one can apply identical arguments to usual combinatorics on words, also see [19, Section 6] or the appendix to this paper. In particular, we can recover the result [20, 3] that $\widehat{F}_{\mathbf{A}}(A)$ is equidivisible; a monoid M is called *equidivisible* if for all $u, v, u', v' \in M$, if $uv = u'v'$ then there exists $x \in M$ such that either $ux = u'$ and $xv' = v$, or $u'x = u$ and $xv = v'$. The pro-aperiodic monoid $\Lambda(A)$ is equidivisible as well. We also re-prove in this way that the $\mathcal{L}$- and $\mathcal{R}$-orders on $\Lambda(A)$ and $\widehat{F}_{\mathbf{A}}(A)$ are unambiguous, and are total on stabilizers. This was used in [20] to establish the decidability of membership in a number of semidirect products of the form $\mathbf{V} * \mathbf{A}$.

5 The word problem for omega-terms

In this section, we use our techniques to give an improved proof of the decidability of the word problem for ω-terms in $\widehat{F}_{\mathbf{A}}(A)$. We begin with a few definitions. An *ω-term* over A is a term built up from finite words by using concatenation and ω-power. If M is a profinite monoid containing the alphabet A, then any ω-term t has a natural interpretation $[\![t]\!]_M$ in M, which can be defined inductively using the multiplication and the unary operation $()^\omega$ on the profinite monoid M. In the case $M = \widehat{F}_{\mathbf{A}}(A)$, we will now inductively define, for any ω-term t, a particular A-word U_t in the class $[\![t]\!]_{\widehat{F}_{\mathbf{A}}(A)}$. Let ρ denote the linear order $\mathbb{N} + \mathbb{Q} \times \mathbb{Z} + \mathbb{N}^{\mathrm{op}}$, which is countably saturated (cf. Example 14).

- If t is a term representing a finite word, let U_t be that finite word.
- If $t = t_1 \cdot t_2$, let U_t be the A-word $U_{t_1} \cdot U_{t_2}$.
- If $t = s^\omega$, let U_t be the A-word $(U_s)^\rho$.

▶ **Proposition 20.** *For any ω-term t, the A-word U_t is a countably saturated A-word in the elementary equivalence class $[\![t]\!]_{\widehat{F}_{\mathbf{A}}(A)}$.*

Proof. Finite words are countably saturated. Concatenations and ρ-powers of countably saturated A-words are clearly countable, and, by Corollary 16, ω-saturated. An easy induction, using Theorem 8 for the step involving ω-power, shows that U_t lies in $[\![t]\!]_{\widehat{F}_{\mathbf{A}}(A)}$. ◀

▶ **Theorem 21.** *For any ω-terms t_1, t_2, the following are equivalent:*
1. $[\![t_1]\!]_{\widehat{F}_{\mathbf{A}}(A)} = [\![t_2]\!]_{\widehat{F}_{\mathbf{A}}(A)}$,
2. *U_{t_1} is isomorphic to U_{t_2}.*

Proof. (2) ⇒ (1) is clear, since isomorphic A-words are elementarily equivalent. (1) ⇒ (2). By Proposition 20, both U_{t_1} and U_{t_2} are countably saturated models in the same elementary equivalence class. By the uniqueness of countably saturated models (cf., e.g., [12, Thm. 2.3.9]), U_{t_1} and U_{t_2} are isomorphic. ◀

In order to decide the word problem for ω-terms in $\widehat{F}_{\mathbf{A}}(A)$, one can now proceed as in [22] and use a decidability procedure for isomorphism of regular words (cf. [8] or [24]) to decide isomorphism of the countably saturated A-words interpreting the ω-terms.

▶ Remark. The original proof of decidability was due to McCammond [26], who introduced normal forms based on an inductively defined notion of rank. The separation of normal forms makes use of his solution to the word problem for free Burnside semigroups of sufficiently large exponent [25], which inspired the definition of the normal forms in the first place. Recently, Almeida, Costa and Zeitoun [6] have provided a new proof that distinct McCammond normal forms represent distinct ω-terms in $\widehat{F}_{\mathbf{A}}(A)$ by introducing star-free languages associated to normal forms whose closures can be used to separate them. Huschenbett and Kufleitner developed in [22] a new algorithm to solve the word problem for ω-terms in $\widehat{F}_{\mathbf{A}}(A)$ using model-theoretic ideas. They assign the same A-word that we did above to each ω-term. They then prove that two such interpretations of ω-terms are elementarily equivalent to each other if and only if they are isomorphic by making use of the non-trivial direction McCammond's normal form theorem. They use work of Bloom and Esik [8] to prove that isomorphism can be decided in exponential time in the size of the expression as an ω-term; in the conference presentation of this result they announced that this can be improved to polynomial time using recent work of Lohrey and Mathissen [24]. Here, countably saturated models allow us to give a direct proof of the correctness of the Huschenbett and Kufleitner algorithm, circumventing McCammond's results entirely.

6 Factors of omega-terms and well-quasi-orders

In this section and the next, we exploit our model-theoretic view of $\Lambda(A)$ and $\widehat{F}_{\mathbf{A}}(A)$ to analyze the $\mathcal{J}$-orderings on sets of *factors* of elements in $\Lambda(A)$. As special cases of our more general results, we recover several of the structural results on factors of ω-terms that were obtained in [4, 5] using McCammond's normal forms. Again, our proofs avoid such normal forms altogether; instead, we use Theorem 18 on factors of a substitution.

We first show how our results in Section 4 give a simple proof of the following result that was proved in an entirely different way by Almeida, Costa, and Zeitoun [5, Theorem 7.4].

▶ **Theorem 22.** *Prefixes, suffixes and factors of elements in $\widehat{F}_{\mathbf{A}}(A)$ that are interpretations of ω-terms are again interpretations of ω-terms.*

Proof. We prove the statement for prefixes. We write $[\![t]\!]$ as shorthand for $[\![t]\!]_{\widehat{F}_{\mathbf{A}}(A)}$. The statement for suffixes then follows by symmetry, and the statement for factors, in turn, then follows because any factor is a suffix of a prefix. We will prove by induction on the complexity of an ω-term t that, for any $w \in \widehat{F}_{\mathbf{A}}(A) \setminus \{\varepsilon\}$ such that $[\![t]\!] \leq_{\mathcal{R}} w$, we have $w = [\![s]\!]$ for some ω-term s. Prefixes of a finite word are finite words. If $t = t_0 \cdot t_1$ for some ω-terms t_0 and t_1, and $[\![t]\!] \leq_{\mathcal{R}} w$, we apply Theorem 18.2 in the special case of a concatenation, which we recall is the substitution of $u_0 := [\![t_0]\!]$ and $u_1 := [\![t_1]\!]$ into the $\{0,1\}$-word $v := 01$. If $[\![t_0]\!] \leq_{\mathcal{R}} w$, then we are done immediately by induction. Otherwise, we have $w = [\![t_0]\!] \cdot v$ for some v with $[\![t_1]\!] \leq_{\mathcal{R}} v$. By induction, pick an ω-term s' such that $[\![s']\!] = v$. Then for the ω-term $s := t_0 s'$, we have $[\![s]\!] = w$. If $t = r^\omega$ for some ω-term r, and $[\![t]\!] \leq_{\mathcal{R}} w$, we apply Theorem 18.2 in the special case of an ω-power, which we recall is the substitution of $u := [\![r]\!]$ into the $\{b\}$-word $v := b^\omega$. There exist $z \in \widehat{F}_{\mathbf{A}}(1) = \mathbb{N} \cup \{\omega\}$ and $x \in \widehat{F}_{\mathbf{A}}(A)$ with $[\![r]\!] \leq_{\mathcal{R}} x$ such that $w = r^z x$. By the induction hypothesis, pick an ω-term s' such that $[\![s']\!] = x$. The ω-term $s := r^z s'$ gives $[\![s]\!] = w$. ◀

▶ Remark. Almeida, Costa and Zeitoun also proved that interpretations of ω-terms have only finitely many regular $\mathcal{J}$-classes above them. In [19, Section 8], we easily prove this result using our new method; we omit this here for reasons of space.

Recall that a *well-quasi-order* (wqo) is a quasi-order that does not contain infinite antichains or infinite descending chains; equivalently, for every infinite sequence $(q_i)_{i \in \omega}$ in Q, there exist $i < j$ such that $q_i \preceq q_j$ (see e.g. [31, §10.3]). We will call an element $u \in \Lambda(A)$ *well-factor-ordered* (wfo) if the reverse $\mathcal{J}$-order, $\geq_{\mathcal{J}}$, is a well-quasi-order on the set $\uparrow_{\mathcal{J}} u$ of factors of u. Similarly, we call $u \in \Lambda(A)$ *well-prefix-ordered* (wpo) if $\geq_{\mathcal{R}}$ is a wqo on $\uparrow_{\mathcal{R}} u$ and *well-suffix-ordered* (wso) if $\geq_{\mathcal{L}}$ is a wqo on $\uparrow_{\mathcal{L}} u$.

▶ Remark. Note that, by Proposition 9, $\widehat{F}_{\mathbf{A}}(A)$ is upward closed in $\Lambda(A)$ with respect to the $\mathcal{J}$-, $\mathcal{R}$- and $\mathcal{L}$-orders, so that the following results apply immediately to $\widehat{F}_{\mathbf{A}}(A)$, as well.

The proof of the following theorem uses Theorem 18 to analyze the factors of the result of substituting wfo elements into wfo elements.

▶ **Theorem 23.** *Let $f\colon \Lambda(B) \to \Lambda(A)$ be a substitution. If $f(b)$ is well-factor-ordered for each $b \in B$, and $v \in \Lambda(B)$ is well-factor-ordered, then $f(v)$ is well-factor-ordered.*

Analogous results, with simpler proofs, hold for well-prefix-ordered and well-suffix-ordered elements.

▶ **Theorem 24.** *Let $f\colon \Lambda(B) \to \Lambda(A)$ be a substitution. If $f(b)$ is well-prefix-ordered for each $b \in B$, and $v \in \Lambda(B)$ is well-prefix-ordered, then $f(v)$ is well-prefix-ordered.*

By symmetry, we have the following corollary.

▶ **Corollary 25.** *Let $f\colon \Lambda(B) \to \Lambda(A)$ be a substitution. If $f(b)$ is well-suffix-ordered for each $b \in B$, and $v \in \Lambda(B)$ is well-suffix-ordered, then $f(v)$ is well-suffix-ordered.*

The following special case of the above three results recovers [5, Corollary 5.6] and [5, Theorem 7.3].

▶ **Corollary 26.** *The sets of well-factor-ordered, well-prefix-ordered, and well-suffix-ordered elements in $\Lambda(A)$ are closed under concatenation and ω-power. In particular, interpretations of ω-terms in $\widehat{F}_{\mathbf{A}}(A)$ are well-factor-ordered.*

We end this section by showing that there are many more well-factor-ordered elements in $\widehat{F}_{\mathbf{A}}(A)$ and $\Lambda(A)$ than just the interpretations of ω-terms. Almeida showed [2, Theorem 2.6] that $\widehat{F}_{\mathbf{A}}(A) \setminus A^*$ contains maximal elements with respect to the $\mathcal{J}$-ordering and that they correspond in a sense that can be made precise to uniformly recurrent words. Moreover, it is shown in [5, Proposition 3.2] that every element of $\widehat{F}_{\mathbf{A}}(A) \setminus A^*$ is $\mathcal{J}$-below a maximal element. Notice that ω is the unique maximal element of $\widehat{F}_{\mathbf{A}}(1) \setminus \mathbb{N}$. The $\mathcal{J}$-maximal elements of $\widehat{F}_{\mathbf{A}}(A) \setminus A^*$ are in particular $\mathcal{J}$-maximal in $\Lambda(A) \setminus A^*$, and all of the latter elements are well-factor-ordered:

▶ **Proposition 27.** *Let w be a maximal element of $\Lambda(A) \setminus A^*$ in the $\mathcal{J}$-order. Then w is well-factor-ordered.*

It follows that the smallest submonoid of $\widehat{F}_{\mathbf{A}}(A)$ containing all finite words and $\mathcal{J}$-maximal elements that is closed under the ω-power consists of well-factor-ordered elements. Also, substituting ω-terms into $\mathcal{J}$-maximal elements will provide new example of well-factor-ordered elements.

7 Regular languages of factors

We call an element $u \in \Lambda(A)$ *factor-regular* if the set $F(u) := {\uparrow_{\mathcal{J}}} u \cap A^*$ of finite factors of u is a regular language. We call u *prefix-regular* if $P(u) := {\uparrow_{\mathcal{R}}} u \cap A^*$ is a regular language, and *suffix-regular* if $S(u) := {\uparrow_{\mathcal{L}}} u \cap A^*$ is a regular language. If $u \in \Lambda(A) \setminus A^*$ is prefix-regular, then so is any element of $u\Lambda(A)$, and dually for suffix-regular elements. We call a substitution $f\colon \Lambda(B) \to \Lambda(A)$ *non-erasing* if $f(b) \neq \varepsilon$ for every $b \in B$. Notice that a non-erasing substitution f sends the ideal $\Lambda(B) \setminus B^*$ to the ideal $\Lambda(A) \setminus A^*$.

In the following theorem, which is another application of Theorem 18, we prove that prefix-regularity and suffix-regularity are stable under non-erasing substitutions, and so is factor-regularity under certain additional assumptions.

▶ **Theorem 28.** *Let $f\colon \Lambda(B) \to \Lambda(A)$ be a non-erasing substitution and $v \in \Lambda(B)$.*

1. *If $f(b)$ is prefix-regular, for each $b \in B$, and v is prefix-regular, then $f(v)$ is prefix-regular.*
2. *If $f(b)$ is suffix-regular, for each $b \in B$, and v is suffix-regular, then $f(v)$ is suffix-regular.*
3. *If $f(b)$ is factor-regular, prefix-regular and suffix-regular, for each $b \in B$, and v is factor-regular, then $f(v)$ is factor-regular.*

Proof. We only prove (3), referring to the appendix for the proofs of (1) and (2). Put $C = \{b \in B^* \mid f(b) \in A^*\}$. Using Corollary 19, it is straightforward to verify that

$$F(f(v)) = \bigcup_{b \in B \cap F(v)} F(f(b)) \cup \bigcup_{b_1, b_2 \in B \cup \{\varepsilon\}} S(f(b_1)) f(C^* \cap b_1^{-1} F(w) b_2^{-1}) P(f(b_1))$$

and so the desired result follows from closure of regular languages under boolean operations, product, left and right quotients and homomorphic image. ◀

For a finite alphabet A and $b \notin A$, let $f\colon \Lambda(A \cup \{b\}) \to \Lambda(A)$ be the substitution erasing b and fixing A. Note that if $v \in \Lambda(A)$ is not prefix-regular, then $b^\omega v$ is prefix-regular, but $f(b^\omega v) = v$ is not. Thus, the assumption in Theorem 28 that f is non-erasing is necessary.

The next corollary recovers [5, Corollary 7.6].

▶ **Corollary 29.** *Interpretations of ω-terms are prefix-, suffix- and factor-regular.*

▶ Remark. There are more factor-regular elements than just interpretations of ω-terms. In particular, the minimal ideal I of $\widehat{F}_{\mathbf{A}}(A)$ consists of those elements containing every finite word as a factor, so every element of I is factor-regular. Thus if we substitute ω-terms over B into an element of the minimal ideal of $\widehat{F}_{\mathbf{A}}(A)$, then we obtain factor-regular elements of $\widehat{F}_{\mathbf{A}}(B)$. It is not the case that every element of the minimal ideal of $\widehat{F}_{\mathbf{A}}(A)$ is prefix-regular or suffix-regular. In fact, it is easy to see using the pumping lemma that $w \in \Lambda(A) \setminus A^*$ is prefix-regular if and only if $w = uv^\omega z$ with u, v finite and $v \neq \varepsilon$. A dual description holds for suffix-regular elements.

8 Conclusion

In this paper we gave a new approach to the free pro-aperiodic monoid by viewing its elements as elementary equivalence classes of pseudofinite words. This view led us to consider $\widehat{F}_{\mathbf{A}}(A)$ as a closed submonoid of the larger monoid $\Lambda(A)$ consisting of elementary equivalence classes of arbitrary A-words. The model-theoretic fact that each such class contains an ω-saturated model enabled us to analyze factors in $\Lambda(A)$ combinatorially. Thus, we substantiate the claim made in [20] that one may "transfer arguments from Combinatorics on Words to the profinite context": our approach using saturated models makes this idea precise.

The newly identified pro-aperiodic monoid $\Lambda(A)$ poses several interesting questions for future work. In particular, it would be interesting to study the algebraic structure of $\Lambda(A)$ in more detail. Here, connections with the work of Carton, Colcombet and Puppis on algebras for words over countable linear orderings [11] are to be expected.

We also plan to explore in future work how this method might be extended to other profinite monoids. In particular, one could try to analyze the absolutely free profinite monoid in this way, but this would require replacing first-order model theory by monadic second-order model theory. In this direction, we foresee connections to Shelah's seminal work [33].

In a different direction, we hope that our approach could be useful for more easily analyzing aperiodic pointlike sets and related notions, in particular because a logical approach has recently proved useful for deciding problems in the first-order quantifier alternation hierarchy [29].

References

1 J. Almeida. *Finite semigroups and universal algebra*, volume 3 of *Series in Algebra*. World Scientific Publishing Co. Inc., River Edge, NJ, 1994. Translated from the 1992 Portuguese original and revised by the author.

2 J. Almeida. Profinite groups associated with weakly primitive substitutions. *Fundam. Prikl. Mat.*, 11(3):13–48, 2005. doi:10.1007/s10958-007-0242-y.

3 J. Almeida and A. Costa. Infinite-vertex free profinite semigroupoids and symbolic dynamics. *J. Pure Appl. Algebra*, 213(5):605–631, 2009. doi:10.1016/j.jpaa.2008.08.009.

4 J. Almeida, J. C. Costa, and M. Zeitoun. Some structural properties of the free profinite aperiodic semigroup. In *2nd AutoMathA Conference*, 2009.

5 J. Almeida, J. C. Costa, and M. Zeitoun. Iterated periodicity over finite aperiodic semigroups. *European J. Combin.*, 37:115–149, 2014. doi:10.1016/j.ejc.2013.07.011.

6 J. Almeida, J. C. Costa, and M. Zeitoun. McCammond's normal forms for free aperiodic semigroups revisited. *LMS J. Comput. Math.*, 18(1):130–147, 2015. doi:10.1112/S1461157014000448.

7 J. Almeida and M. V. Volkov. Subword complexity of profinite words and subgroups of free profinite semigroups. *Internat. J. Algebra Comput.*, 16(2):221–258, 2006. doi:10.1142/S0218196706002883.

8 S. L. Bloom and Z. Ésik. The equational theory of regular words. *Information and Computation*, 197:55–89, 2005.

9 M. Bojańczyk. Recognisable Languages over Monads. arXiv:1502.04898v1, 2015.

10 J. A. Brzozowski. Hierarchies of aperiodic languages. *Theoret. Informatics Appl.*, 10(2):33–49, 1976.

11 O. Carton, T. Colcombet, and G. Puppis. Regular languages of words over countable linear orderings. In *Proceedings of the 38th International Colloquium on Automata, Languages and Programming (ICALP)*, volume 6756 of *LNCS*, pages 125–136. Springer, 2011.

12 C.-C. Chang and J. H. Keisler. *Model Theory*. North-Holland, Amsterdam-London, third edition, 1990.

13 H. C. Doets. *Completeness And Definability : Applications Of The Ehrenfeucht Game In Second-Order And Intensional Logic*. PhD thesis, Universiteit van Amsterdam, May 1987. URL: http://oai.cwi.nl/oai/asset/22742/22742A.pdf.

14 S. Eilenberg. *Automata, languages, and machines. Vol. B*. Academic Press, New York, 1976. With two chapters ("Depth decomposition theorem" and "Complexity of semigroups and morphisms") by Bret Tilson, Pure and Applied Mathematics, Vol. 59.

15 M. Gehrke. Stone duality, topological algebra, and recognition. *J. Pure Appl. Algebra*, 220(7):2711–2747, 2016.

16 M. Gehrke, S. Grigorieff, and J.-É. Pin. Duality and equational theory of regular languages. In L. Aceto et al., editor, *ICALP 2008, Part II*, number 5126 in Lect. Notes Comput. Sci., pages 246–257. Springer, 2008.

17 M. Gehrke, A. Krebs, and J.-É. Pin. Ultrafilters on words for a fragment of logic. *Theoretical Computer Science*, 610 A:37–58, 2016.

18 M. Gehrke, D. Petrisan, and L. Reggio. The Schützenberger product for syntactic spaces. In *43rd International Colloquium on Automata, Languages and Programming (ICALP 2016)*, 2016.

19 S. J. v. Gool and B. Steinberg. Pro-aperiodic monoids and model theory. Technical report, 2016. URL: http://www.samvangool.net/papers/goolsteinberg2016TR.pdf.

20 K. Henckell, J. Rhodes, and B. Steinberg. A profinite approach to stable pairs. *Internat. J. Algebra Comput.*, 20(2):269–285, 2010. doi:10.1142/S0218196710005650.

21 W. Hodges. *Model theory*, volume 42 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 1993.

22 M. Huschenbett and M. Kufleitner. Ehrenfeucht-Fraisse Games on Omega-Terms. In *STACS*, pages 374–385, 2014.

23 K. Krohn and J. Rhodes. Complexity of finite semigroups. *Ann. of Math. (2)*, 88:128–160, 1968.

24 M. Lohrey and C. Mathissen. Isomorphism of regular trees and words. *Inform. and Comput.*, 224:71–105, 2013. doi:10.1016/j.ic.2013.01.002.

25 J. P. McCammond. The solution to the word problem for the relatively free semigroups satisfying $T^a = T^{a+b}$ with $a \geq 6$. *Internat. J. Algebra Comput.*, 1(1):1–32, 1991. doi:10.1142/S021819679100002X.

26 J. P. McCammond. Normal forms for free aperiodic semigroups. *Int. J. Algebra Comput.*, 11(5):581–625, 2001.

27 R. McNaughton and S. Papert. *Counter-free automata.* Number 65 in M.I.T. Research Monograph. M.I.T. Press, 1971.

28 J.-É. Pin. The dot-depth hierarchy, 45 years later. Preprint available from the author's homepage.

29 T. Place and M. Zeitoun. Going Higher in the First-Order Quantifier Alternation Hierarchy on Words. In *Proceedings ICALP*, 2014.

30 J. Rhodes and B. Steinberg. *The q-theory of Finite Semigroups.* Springer, 2009.

31 J. G. Rosenstein. *Linear orderings.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], 1982.

32 M.-P. Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8:190–194, 1965.

33 S. Shelah. The monadic theory of order. *Ann. of Math. (2)*, 102(3):379–419, 1975.

34 M. H. Stone. The Theory of Representation for Boolean Algebras. *Trans. Amer. Math. Soc.*, 74(1):37–111, 1936.

35 H. Straubing. *Finite automata, formal logic, and circuit complexity.* Progress in Theoretical Computer Science. Birkhäuser Boston Inc., Boston, 1994.

36 H. Straubing and P. Weil. Varieties. In J.-É. Pin, editor, *Handbook of automata theory.* European Mathematical Society Publishing House, to appear. ArXiv:1502.03951.

37 J. Väänänen. Pseudo-finite model theory. *Mat. Contemp*, 24:169–183, 2003.