POSITIVE VARIETIES and INFINITE WORDS

Jean-Éric Pin

LIAFA, CNRS and Université Paris VII
2 Place Jussieu 75251 Paris Cedex O5
FRANCE
e-mail: `Jean-Eric.Pin@liafa.jussieu.fr`

Abstract. Carrying on the work of Arnold, Pécuchet and Perrin, Wilke has obtained a counterpart of Eilenberg's variety theorem for finite and infinite words. In this paper, we extend this theory for classes of languages that are closed under union and intersection, but not necessarily under complement. As an example, we give a purely algebraic characterization of various classes of recognizable sets defined by topological properties (open, closed, F_{Σ} and G_{δ}) or by combinatorial properties

Carrying on the work of Arnold [1], Pécuchet [7, 8] and Perrin [9–11], Wilke [25, 26] has obtained a counterpart of Eilenberg's variety theorem for finite *and* infinite words. The word "and" is emphasized in the last sentence, because it is really important to work simultaneously with finite and infinite words. The fitness of this approach was corroborated by recent contributions [13, 14, 18, 25, 27].

A variant of the notion of syntactic semigroup was recently proposed by the author [17]. The key idea is to define a partial order on syntactic semigroups, leading to the notion of ordered syntactic semigroups. The resulting extension of Eilenberg's variety theory permits to treat classes of languages that are closed under union and intersection, but not necessarily under complement, a major difference with the original theory.

The aim of this paper is to extend this theory to infinite words, thus completing the table below. In the setting proposed by Wilke, semigroups are not suitable any more. They can be replaced by ω-semigroups, which are, roughly speaking, semigroups equipped with an infinite product [13, 25, 27].

Finite words	Eilenberg [4] Varieties of semigroups $\Leftrightarrow$ +-varieties	Pin [17] Var. of ordered semigroups $\Leftrightarrow$ Positive +-varieties
Finite or infinite words	Wilke [25, 26] Varieties of ω-semigroups $\Leftrightarrow$ ∞-varieties	*This paper* Var. of ordered ω-semigroups $\Leftrightarrow$ Positive ∞-varieties

The "ordered" approach has many interesting consequences but leads to a complete rewriting of the theory. We have selected three examples, of rather different nature, to convince the reader of the power of this new approach.

In Sect. 5, we give a purely algebraic characterization of four classes of recognizable sets defined by topological properties. This includes in particular the class of deterministic ω-languages (i.e. recognized by a deterministic Büchi automaton). Similar results were known only for topological classes closed under complement. Note that all these characterizations are effective, since a simple algorithm to compute the syntactic ω-semigroup of a recognizable ω-language was given in [13].

In Sect. 6 we address a question originally considered by Pécuchet [7, 8]. Since every recognizable ω-language can be written as a finite union of languages of the form XY^ω, where X and Y are recognizable languages, the question arose to know whether this result could be "relativized" to varieties. Theorem 3, which extends the results of Pécuchet, gives such a result for varieties of ordered semigroups. It requires the concept of weak recognizability, introduced by Perrin [11], and the notion of ordered Büchi automaton, which might be interesting in itself.

Our last example, developed in Sect. 7, is a good illustration of the problems that arise when trying to generalize to infinite words a given class of recognizable languages. Indeed, as was observed by Pécuchet, there are at least three natural ways to associate a class of ω-languages to a variety of finite semigroups $\mathbf{V}$. One can consider the sets recognized by a finite ω-semigroup S whose semigroup part belongs to $\mathbf{V}$, or those weakly recognized by a semigroup of $\mathbf{V}$, or finally, inspired by McNaughton's theorem, the boolean combinations of sets of the form $\overrightarrow{L}$ where L is recognized by a semigroup of $\mathbf{V}$. We will see how these three classes relate to each other in our case study, the shuffle ideals.

Due to the lack of place, no proofs are given, but they can be found in [14]. For more details, the reader is referred to [15, 16, 20] for the variety theory for finite words and to [13, 14, 19, 24–26] for the theory of ω-languages.

1 Notations and basic definitions

Let A be a finite alphabet. The free monoid on A is denoted A^* and the free semigroup, A^+. The set of infinite words on A is denoted A^ω. Finally, A^∞ denotes the set of finite or infinite words. In this paper, a subset X of A^∞ will be systematically identified with the pair (X_+, X_ω), where $X_+ = X \cap A^+$ and $X_\omega = X \cap A^\omega$.

We now briefly review the standard definition of a Büchi automaton and introduce the notion of an ordered Büchi automaton.

1.1 Büchi automata

A *Büchi automaton* is a 5-tuple $\mathcal{A} = (Q, A, E, I, F)$ where

(1) (Q, A, E) is a finite (non deterministic) automaton,

(2) I and F are subsets of Q, called the set of *initial* and *final* states.

A finite path in $\mathcal{A}$ is *successful* if its origin is in I and its end is in F. An infinite path p is *successful* if its origin is in I and if some state of F occurs infinitely often in p.

The set of finite (respectively infinite) words *recognized* by $\mathcal{A}$ is the set of the labels of all successful finite (respectively infinite) paths of $\mathcal{A}$. A set of finite (respectively infinite) words is *recognizable* (or *regular*) if it is recognized by a finite Büchi automaton.

A Büchi automaton $\mathcal{A} = (Q, A, E, I, F)$ is said to be *deterministic* if I is a singleton and if E contains no pair of transitions of the form $(q, a, q_1), (q, a, q_2)$ with $q_1 \neq q_2$. In particular, each word u is the label of at most one path starting from the initial state. It follows that an infinite word is accepted by $\mathcal{A}$ if and only if it has infinitely many prefixes accepted by $\mathcal{A}$. Therefore, if L denotes the set of finite words recognized by $\mathcal{A}$, then

$$\overrightarrow{L} = \{u \in A^\omega \mid u \text{ has infinitely many prefixes in } L\}.$$

is the set of infinite words recognized by $\mathcal{A}$.

1.2 Ordered Büchi automata

An *ordered automaton* is a Büchi automaton $\mathcal{A} = (Q, A, E, I, F)$ in which the set of states Q is equipped with a partial order $\leq$ satisfying the two following conditions, for all $p, q, q' \in Q$ and $a \in A$:

(1) if $p \leq q$ and $(q, a, q') \in E$, there exists a state $p' \leq q'$ such that $(p, a, p') \in E$,

(2) The set F of final states is an order ideal: if $q \in F$ and $p \leq q$, then $p \in F$.

Condition (1) can be extended by transitivity as follows:

(3) if $p \leq q$ and if there is a path from q to q' labeled by u, there exists a state $p' \leq q'$ such that u is the label of a path from p to p'.

If $\mathcal{A}$ is a deterministic automaton, condition (1) can be simplified as follows

(4) if $p \leq q$ and if $q\cdot a$ is defined, then $p\cdot a$ is defined and satisfies $p\cdot a \leq q\cdot a$.

2 Ordered algebraic structures

We introduce in this section our main algebraic tools : ordered semigroups, ω-semigroups and syntactic ω-semigroups.

2.1 Ordered semigroups

A *quasi-order* is a reflexive and transitive relation. Given a quasi-order $\mathcal{R}$, the relation $\sim$ defined by $x \sim y$ if and only if $x \mathrel{\mathcal{R}} y$ and $y \mathrel{\mathcal{R}} x$ is an equivalence relation, called the *equivalence relation associated with* $\mathcal{R}$. If this equivalence relation is the equality relation, the relation $\mathcal{R}$ is an *order*.

A relation $\mathcal{R}$ on a semigroup S is *stable* if, for every $x, y, z \in S$, $x \mathrel{\mathcal{R}} y$ implies $xz \mathrel{\mathcal{R}} yz$ and $zx \mathrel{\mathcal{R}} zy$. An *ordered semigroup* is a semigroup S equipped with a stable order relation $\leq$ on S. Note that any semigroup, equipped with the equality relation, becomes an ordered semigroup.

Let S be an ordered semigroup. An *ordered subsemigroup* of S is a subset T of S such that $t, t' \in T$ implies $tt' \in T$. An *order ideal* of S is a subset I of S such that, if $x \le y$ and $y \in I$, then $x \in I$.

A *congruence on an ordered semigroup* S is a stable quasi-order which is coarser than or equal to $\le$. In particular, the order relation $\le$ is itself a congruence. If $\preceq$ is a congruence on S, then the equivalence relation $\sim$ associated with $\preceq$ is a congruence on S. Furthermore, there is a well-defined stable order on the quotient set $S/{\sim}$, given by $[s] \le [t]$ if and only if $s \preceq t$. Thus $(S/{\sim}, \le)$ is an ordered semigroup, also denoted $S/{\preceq}$.

Given two ordered semigroups S and T, their *product* $S \times T$ is the ordered semigroup defined on the set $S \times T$ by the law $(s, t)(s', t') = (ss', tt')$ and the order given by $(s, t) \le (s', t')$ if and only if $s \le s'$ and $t \le t'$.

A *morphism of ordered semigroups* $\varphi : S \to T$ is a semigroup morphism from S into T such that, for every $x, y \in S$, $x \le y$ implies $\varphi(x) \le \varphi(y)$. A morphism of ordered semigroups $\varphi : S \to T$ is an *isomorphism* if and only if φ is a bijective semigroup morphism and, for every $x, y \in S$, $x \le y$ is equivalent with $\varphi(x) \le \varphi(y)$.

Let S and T be two ordered semigroups. Then S is a *quotient* of T if there exists a surjective morphism from T onto S, and S *divides* T if S is a quotient of an ordered subsemigroup of T. Division is a quasi-order on ordered semigroups. Furthermore, one can show that two finite ordered semigroups divide each other if and only if they are isomorphic.

A subset X of A^+ is *recognized* by an ordered semigroup S if there exists an ideal order I of S and a semigroup morphism $\varphi : A^+ \to S$ such that $X = \varphi^{-1}(I)$. It is not difficult to see that a language is recognizable if and only if it is recognized by a finite ordered semigroup.

The *syntactic congruence* of a recognizable subset X of A^+ is the stable quasiorder $\preceq_X$ defined on A^+ by $u \preceq_X v$ if and only if, for every $x, y \in A^*$,

$$xvy \in X \Rightarrow xuy \in X$$

The equivalence relation $\sim_X$ associated with $\preceq_X$ is called the *syntactic equivalence* of X. Thus $u \sim_X v$ if and only if, for every $x, y \in A^*$,

$$xuy \in X \iff xvy \in X$$

The ordered semigroup $S(X) = A^+/{\preceq_X}$ is the *ordered syntactic semigroup* of X and the order relation on $S(X)$ the *syntactic order* of X.

2.2 Transition semigroup of a Büchi automaton

A standard construction associates to each finite automaton a finite semigroup, its transition semigroup. We define in this section the ordered version of this notion.

Let $\mathcal{A} = (Q, A, E, I, F)$ be a Büchi automaton. Denote by $\mathcal{R}(Q)$ the semigroup of relations on Q, under composition of relations. Then the map $\varphi : A^+ \to$

$\mathcal{R}(Q)$, defined by

$$\varphi(u) = \{(q, q') \in Q \times Q \mid \text{there is a path of label } u \text{ from } q \text{ to } q'\}$$

is a morphism of semigroups, and the semigroup $S(\mathcal{A}) = \varphi(A^+)$ is called the *transition semigroup* of $\mathcal{A}$.

If $\mathcal{A}$ is ordered, $S(\mathcal{A})$ is naturally equipped with a relation $\preceq$ defined by $u' \preceq u$ if and only if, for all $(p, q) \in u$, there exists $q' \leq q$ such that $(p, q') \in u'$. Equivalently, $u' \preceq u$ if and only if, for all $(p, q) \in u$ and for all $p' \leq p$, there exists $q' \leq q$ such that $(p', q') \in u$. One can show that this relation $\preceq$ is a congruence of ordered semigroup on $(S(\mathcal{A}), =)$, but it is not necessarily an order. The ordered semigroup $SO(\mathcal{A}) = (S(\mathcal{A})/\preceq)$ is called the *transition ordered semigroup* of $\mathcal{A}$.

2.3 ω-semigroups

Finite semigroups can be viewed as a two-sided algebraic counterpart of finite automata that recognize finite words. In the case of infinite words, they can be replaced by ω-semigroups, which are, roughly speaking, semigroups equipped with an infinite product.

More formally, an *ω-semigroup* is a two-sorted algebra $S = (S_+, S_\omega)$ equipped with the following operations:

(1) A *product* $S_+ \times S_+ \to S_+$, that associates to each pair $(s, t) \in S_+ \times S_+$ an element of S_+ denoted st,

(2) A *mixed product* $S_+ \times S_\omega \to S_\omega$, that associates to each pair $(s, t) \in S_+ \times S_\omega$ an element of S_ω denoted st,

(3) An *infinite product* $S_+^{\mathbb{N}} \to S_\omega$, that associates to each infinite sequence $s_0, s_1, s_2, \ldots$ of elements of S_+ an element of S_ω denoted $s_0 s_1 s_2 \cdots$.

These three operations satisfy all possible types of associativity (a precise definition can be found in [13]). In particular, S_+, equipped with the binary operation, is a semigroup and for all $s, t \in S_+$ and $u \in S_\omega$, $s(tu) = (st)u$. The infinite product of the sequence $s, s, s, \ldots$ is denoted s^ω.

In particular, we denote by A^∞ the ω-semigroup (A^+, A^ω) equipped with the usual concatenation product.

Given two ω-semigroups $S = (S_+, S_\omega)$ and $T = (T_+, T_\omega)$, a *morphism of ω-semigroups* S is a pair $\varphi = (\varphi_+, \varphi_\omega)$ consisting of a semigroup morphism $\varphi_+ : S_+ \to T_+$ and of a mapping $\varphi_\omega : S_\omega \to T_\omega$ preserving the mixed product and the infinite product: for every sequence $(s_n)_{n \in \mathbb{N}}$ of elements of S_+,

$$\varphi_\omega(s_0 s_1 s_2 \cdots) = \varphi_+(s_0)\varphi_+(s_1)\varphi_+(s_2) \cdots$$

and for every $s \in S_+$, $t \in S_\omega$, $\varphi_+(s)\varphi_\omega(t) = \varphi_\omega(st)$. In the sequel, we shall omit the subscripts, and use the simplified notation φ instead of φ_+ and φ_ω.

Algebraic concepts like subsemigroup, quotient, division and product are easily adapted to ω-semigroups.

2.4 Ordered ω-semigroups

An ordered ω-semigroup is an ω-semigroup (S_+, S_ω) equipped with two partial orders on S_+ and S_ω which are stable under the operations of ω-semigroup:

(1) for all $s, s', t \in S_+$, $s \leq s'$ implies $ts \leq ts'$ and $st \leq s't$,

(2) for all $s, s' \in S_+$ and for all $u \in S_\omega$, $s \leq s'$ implies $su \leq s'u$,

(3) if $(s_n)_{n\in\mathbb{N}}$ and $(s'_n)_{n\in\mathbb{N}}$ are two sequences of elements of $S_+^{\mathbb{N}}$ such that $s_n \leq s'_n$ for all n, then $s_0s_1s_2\cdots \leq s'_0s'_1s'_2\cdots$.

A subset $X = (X_+, X_\omega)$ of A^∞ is *recognized* by an ordered ω-semigroup S if there exists an ideal order $I = (I_+, I_\omega)$ of S and a morphism of ordered ω-semigroup $\varphi : A^\infty \to S$ such that $X_+ = \varphi^{-1}(I_+)$ and $X_\omega = \varphi^{-1}(I_\omega)$.

3 Ordered syntactic ω-semigroup

The *syntactic congruence* of a recognizable subset X of A^∞ is the quasiorder $\preceq_X$ defined on A^+ by $u \preceq_X v$ if and only if, for every $x, y \in A^*$ and for every $z \in A^\omega$,

$$xvy \in X_+ \Rightarrow xuy \in X_+ \tag{1}$$

$$xvyz^\omega \in X_\omega \Rightarrow xuyz^\omega \in X_\omega \tag{2}$$

$$x(vy)^\omega \in X_\omega \Rightarrow x(uy)^\omega \in X_\omega \tag{3}$$

and on A^ω by $u \preceq_X v$ if and only if, for every $x \in A^*$,

$$xv \in X_\omega \Rightarrow xu \in X_\omega$$

The *syntactic ordered ω-semigroup* of X, denoted by $S(X)$, is the quotient of A^∞ by the syntactic congruence of X. It is a finite object, that can be effectively constructed, given a Büchi automaton recognizing X (see [13].)

4 The variety theorem

We state in this section our extended version of the variety theorem.

A *variety of ordered semigroups* is a class of finite ordered semigroups closed under division and finite product. If $\mathbf{V}$ is a variety of ordered semigroups, a $\mathbf{V}$-*language* is a language recognized by an ordered semigroup of $\mathbf{V}$, or, equivalently, whose syntactic ordered semigroup belongs to $\mathbf{V}$.

Similarly, a *variety of ordered ω-semigroups* is a class of finite ordered ω-semigroups closed under division and finite product. If $\mathbf{V}$ is a variety of ordered ω-semigroups, denote by $\mathcal{V}(A^\infty)$ the set of recognizable subsets of A^∞ recognized by an ordered ω-semigroup of $\mathbf{V}$. This is also the set of subsets of A^∞ whose syntactic ordered ω-semigroup belongs to $\mathbf{V}$.

An ∞-*class* of recognizable sets is a correspondence which associates a set $\mathcal{C}(A^\infty)$ of recognizable sets of A^∞ with every finite alphabet A. In particular, the correspondence $\mathbf{V} \to \mathcal{V}$ associates an ∞-class of recognizable sets with every variety of ordered ω-semigroups. The variety theorem states in particular that this correspondence is one-to-one. It also gives an abstract description of the classes $\mathcal{V}$ arising in this way.

Given a subset X of A^∞, a word $u \in A^*$ and an infinite word v of A^ω, set

$$\begin{aligned} u^{-1}X &= \{x \in A^\infty \mid ux \in X\} \\ Xu^{-\omega} &= \{x \in A^+ \mid (xu)^\omega \in X\} \\ Xv^{-1} &= \{x \in A^+ \mid xv \in X\} \end{aligned}$$

A *positive* ∞*-variety* is an ∞-class such that

(1) For every alphabet A, $\mathcal{V}(A^\infty)$ is closed under finite union and finite intersection,

(2) for every semigroup morphism $\varphi : A^+ \to B^+$, $X \in \mathcal{V}(B^\infty)$ implies $\varphi^{-1}(X) \in \mathcal{V}(A^\infty)$,

(3) If $X \in \mathcal{V}(A^\infty)$, then, for all $u \in A^*$, $u^{-1}X \in \mathcal{V}(A^\infty)$ and $Xu^{-\omega} \in \mathcal{V}(A^\infty)$ and for all $u \in A^\omega$, $Xu^{-1} \in \mathcal{V}(A^\infty)$.

It is important to remember that the elements of a positive ∞-variety are sets of finite or infinite words. The variety theorem can now be stated.

Theorem 1. *The correspondence* $\mathbf{V} \to \mathcal{V}$ *defines a bijection between varieties of ordered* ω*-semigroups and positive* ∞*-varieties.*

Varieties are conveniently defined by identities. Let (u, v) be a pair of words of A^+. An ordered semigroup S *satisfies the identity* $u \leq v$ if and only if $\varphi(u) \leq \varphi(v)$ for every morphism of ordered ω-semigroups $\varphi : A^+ \to S$.

Similarly, let (u, v) be a pair of words of A^∞. An ordered ω-semigroup S *satisfies the identity* $u \leq v$ if and only if $\varphi(u) \leq \varphi(v)$ for every morphism of ordered ω-semigroups $\varphi : A^\infty \to S$. Given a set Σ of identities, the class of all finite ordered ω-semigroups that satisfy all the identities of Σ is a variety of ordered ω-semigroup, denoted $[\![\Sigma]\!]$.

In a finite semigroup, the subsemigroup generated by an element x contains a unique idempotent, denoted by x^π. We will also adopt this notation for identities (this can be rigorously justified, see [21]). For instance, an ordered ω-semigroup S belongs to the variety $[\![x^\pi y^\omega \leq x^\omega]\!]$ if and only if, for every $x, y \in S_+$, $x^\pi y^\omega \leq x^\omega$. Note that x^π is an idempotent of S_+, while x^ω and $x^\pi y^\omega$ are elements of S_ω.

5 Topological classes

Our ordered version of the variety theorem is perfectly suited to give algebraic characterization of certain topological properties of recognizable ω-languages. Similar results [25, 26, 13] were so far limited to classes closed under complement.

Recall that the topology on A^ω is defined by considering A as a discrete space and by taking the product topology. In particular the open sets of A^ω are the sets of the form XA^ω for some $X \subset A^+$. A set is closed if its complement is open. Open sets form the first level Σ_1 of the Borel hierarchy. The second level

Σ_2 consists of the countable intersection of open sets. There is a dual hierarchy, whose first level is the class Π_1 of closed sets. The second level Π_2 consists of the countable unions of closed sets. One can show that the recognizable sets of Σ_2 are exactly the deterministic ω-languages. These classes have the following algebraic characterization.

Theorem 2. *Let X be a recognizable subset of A^ω and let $S(X)$ be its syntactic ordered ω-semigroup. Then:*

(1) *X is in Σ_1 if and only if $S(X)$ belongs to $[\![x^\pi y z^\omega \leq x^\omega]\!]$,*
(2) *X is in Π_1 if and only if $S(X)$ belongs to $[\![x^\omega \leq x^\pi y z^\omega]\!]$,*
(3) *X is in Σ_2 if and only if $S(X)$ belongs to $[\![(x^\pi y)^\omega \leq (x^\pi y)^\pi x^\omega]\!]$,*
(4) *X is in Π_2 if and only if $S(X)$ belongs to $[\![(x^\pi y)^\pi x^\omega \leq (x^\pi y)^\omega]\!]$.*

6 Weak recognizability

It is a well know fact that every recognizable set of A^ω is a finite union of sets of the form XY^ω, where X and Y are recognizable languages. Pécuchet [7, 8] succeeded to relativize this result to varieties of semigroups. We give in this section an even stronger version, that works for varieties of ordered semigroups.

We first need some auxiliary definitions. Let S be a finite ordered semigroup and let $\varphi : A^+ \to S$ be a morphism of semigroups. A *linked pair* is a pair $(s, e) \in S \times S$ such that $e^2 = e$ and $se = s$. These linked pairs play an important role in the study of finite ω-semigroups (see [13].) Denote by $\downarrow s$ the order ideal generated by an element s of S. Thus, by definition $\downarrow s = \{t \in S \mid t \leq s\}$. An ω-language is called *φ-simple* if it is of the form

$$\varphi^{-1}(\downarrow s)\big(\varphi^{-1}(\downarrow e)\big)^\omega$$

where (s, e) is a linked pair. Note that $\varphi^{-1}(\downarrow s) = \varphi^{-1}(\downarrow s)\big(\varphi^{-1}(\downarrow e)\big)^*$ and $\varphi^{-1}(\downarrow e) = \big(\varphi^{-1}(\downarrow e)\big)^+$. Furthermore, these two languages are, by construction, recognized by the ordered semigroup S. Finally, let us say that an ω-language is *weakly recognized* by φ if it is a finite union of φ-simple ω-languages. Our theorem can now be stated:

Theorem 3. *Let $\mathbf{V}$ be a variety of ordered semigroups, and let Z be a subset of A^ω. The following conditions are equivalent:*

(1) *Z is recognized by an ordered Büchi automaton whose ordered transition semigroup belongs to $\mathbf{V}$,*
(2) *Z is weakly recognized by an ordered semigroup of $\mathbf{V}$,*
(3) *Z is a finite union of sets of the form XY^ω, where XY^* and Y^+ are $\mathbf{V}$-languages.*

The reader may observe that this result does not fit directly into the theory of varieties, as developed in Sect. 4. Actually, the connection between weak recognizability and recognizability is a difficult topic, that cannot be covered here. But in some cases (see for instance Theorems 5, 7 and 9 below), weak recognizability is equivalent with recognizability.

7 Shuffle ideals

As an illustration of our results, we investigate in this section the analog for infinite words of some standard classes of recognizable languages. Recall that a *shuffle ideal* is a finite union of languages of the form $A^*a_1A^*a_2A^* \cdots a_nA^*$, where $n > 0$ and $a_1, \ldots, a_n$ are letters of A. It was shown in [22] that a language is a shuffle ideal if and only if its syntactic ordered monoid satisfies the identity $x \leq 1$. It is equivalent to saying that its syntactic ordered semigroup satisfies the identities $xy \leq x$ and $yx \leq x$.

Several extensions to ω-languages can be proposed. A first possibility is to consider finite unions of sets of the form $A^*a_1A^*a_2 \cdots A^*a_kA^\omega$.

Theorem 4. *Let Z be a recognizable subset of A^ω. The following conditions are equivalent:*

(1) *Z is a finite union of sets of the form $A^*a_1A^*a_2 \cdots A^*a_kA^\omega$, where a_1, $\ldots$, $a_k \in A$,*
(2) *Z is of the form $\overrightarrow{L}$, where L is a shuffle ideal,*
(3) *$S(Z)$ satisfies the identities $xy \leq x$, $yx \leq x$ and $x^\pi y^\omega \leq x^\omega$.*

Another possibility is to consider ω-languages whose syntactic ordered ω-semigroup satisfy the identities $xy \leq x$ and $yx \leq x$.

Theorem 5. *Let Z be a recognizable subset of A^ω. The following conditions are equivalent:*

(1) *Z is a positive boolean combination of sets of the form $A^*a_1A^* \cdots A^*a_kA^\omega$, with $k > 0$, $a_1, \ldots, a_k \in A$, or $(A^*a)^\omega$, with $a \in A$,*
(2) *Z is a finite union of sets of the form XY^ω, where XY^* and Y^+ are shuffle ideals,*
(3) *Z is weakly recognized by an ordered semigroup of $[\![xy \leq x, yx \leq x]\!]$,*
(4) *Z is weakly recognized by an ordered semigroup of $[\![x^\pi y \leq x^\pi, yx^\pi \leq x^\pi]\!]$,*
(5) *$S(Z)$ satisfies the identities $xy \leq x$ and $yx \leq x$,*
(6) *$S(Z)$ satisfies the identities $x^\pi y \leq x^\pi$ and $yx^\pi \leq x^\pi$.*

The equivalence of conditions (5) and (6) seems to contradict Theorem 1 since the varieties of ordered ω-semigroups $[\![xy \leq x, yx \leq x]\!]$ and $[\![x^\pi y \leq x^\pi, yx^\pi \leq x^\pi]\!]$, which are distinct, define the same class of ω-languages. But in fact, they define two different classes of languages, and hence, two different classes of ∞-languages, which explains the paradox.

The previous results lead to consider the variety of ordered semigroups defined by the identity $xy \leq x$. The corresponding languages are finite unions of languages of the form

$$a_0A^*a_1A^*a_2A^* \cdots a_nA^*$$

where $a_0, a_1, \ldots, a_k \in A$. We have again two possible generalizations.

Theorem 6. *Let Z be a recognizable subset of A^ω. The following conditions are equivalent:*

(1) *Z is a finite union of sets of the form $a_0A^*a_1A^* \cdots A^*a_kA^\omega$, where a_0, a_1, $\ldots$, $a_n \in A$,*
(2) *Z is of the form $\overrightarrow{L}$, where $S(L) \in [\![xy \leq x]\!]$,*
(3) *$S(Z)$ satisfies the identities $xy \leq x$ and $x^\pi y^\omega \leq x^\omega$.*

The other possibility is to consider the ω-languages whose syntactic ordered ω-semigroup satisfy the identity $xy \leq x$.

Theorem 7. *Let Z be a recognizable subset of A^ω. The following conditions are equivalent:*

(1) *Z is a positive boolean combination of sets of the form $a_0A^*a_1A^* \cdots\ a_kA^\omega$ or $(A^*a)^\omega$, where a, a_0, $\ldots$, $a_n \in A$,*
(2) *Z is a finite union of sets of the form XY^ω, where $S(XY^*), S(Y^+) \in [\![xy \leq x]\!]$,*
(3) *Z is weakly recognized by an ordered semigroup of $[\![xy \leq x]\!]$,*
(4) *$S(Z)$ satisfies the identity $xy \leq x$.*

One can also consider boolean combination of shuffle ideals. These languages are called *piecewise testable* and have been intensively studied. They admit a simple, but deep algebraic characterization, discovered by Simon [23]. Recall that a finite semigroup is $\mathcal{J}$-trivial if two elements that generate the same ideal are equal. A recognizable language is piecewise testable if and only if its syntactic semigroup is $\mathcal{J}$-trivial. There are again two possible extensions to infinite words, which are more precise versions of results by Pécuchet [7, 8].)

Theorem 8. *Let Z be a recognizable subset of A^ω. The following conditions are equivalent:*

(1) *Z is a boolean combination of sets of the form $A^*a_1A^*a_2 \cdots A^*a_kA^\omega$, where $a_1, \ldots, a_k \in A$,*
(2) *Z is a boolean combination of sets of the form $\overrightarrow{L}$, where L is a shuffle ideal,*
(3) *$Z = \overrightarrow{L}$, where L is piecewise testable,*
(4) *Z is a boolean combination of sets of the form $\overrightarrow{L}$, where L is piecewise testable,*
(5) *$S_+(Z)$ is $\mathcal{J}$-trivial and $S(Z)$ satisfies the identity $(x^\pi y^\pi)^\pi x^\omega = (x^\pi y^\pi)^\pi y^\omega$.*

Theorem 9. *Let Z be a recognizable subset of A^ω. The following conditions are equivalent:*

(1) *Z is a boolean combination of sets of the form $A^*a_1A^* \cdots A^*a_kA^\omega$, with $a_1, \ldots, a_k \in A$, or $(A^*a)^\omega$, with $a \in A$,*
(2) *Z is a boolean combination of sets of the form XY^ω, where XY^* and Y^+ are shuffle ideals,*
(3) *Z is a finite union of sets of the form XY^ω, where XY^* and Y^+ are piecewise testable,*
(4) *Z is a boolean combination of sets of the form XY^ω, where XY^* and Y^+ are piecewise testable,*

(5) *Z is weakly recognized by a $\mathcal{J}$-trivial semigroup,*
(6) *$S_+(Z)$ is $\mathcal{J}$-trivial.*

References

1. A. Arnold, A syntactic congruence for rational ω-languages, *Theoret. Comput. Sci.* **39**, (1985) 333–335.
2. J. R. Büchi, Weak second-order arithmetic and finite automata, *Z. Math. Logik und Grundl. Math.* **6**, (1960) 66–92.
3. J. R. Büchi, On a decision method in restricted second-order arithmetic, in *Proc. 1960 Int. Congr. for Logic, Methodology and Philosophy of Science*, Stanford Univ. Press, Standford, (1962) 1–11.
4. S. Eilenberg, *Automata, languages and machines*, Vol. B, Academic Press, New York (1976).
5. R. McNaughton, Testing and generating infinite sequences by a finite automaton *Information and Control* **9**, (1966) 521–530.
6. D. Muller, Infinite sequences and finite machines, in *Switching Theory and Logical Design*, Proc. Fourth Annual Symp. IEEE, (1963) 3–16.
7. J.-P. Pécuchet, Variétés de semigroupes et mots infinis, in *B. Monien and G. Vidal-Naquet eds., STACS 86, Lecture Notes in Computer Science* **210**, Springer, (1986) 180–191.
8. J.-P. Pécuchet, Étude syntaxique des parties reconnaissables de mots infinis, in *Proc. 13th ICALP, (L. Kott ed.) Lecture Notes in Computer Science* **226**, Springer, Berlin, (1986) 294–303.
9. D. Perrin, Variétés de semigroupes et mots infinis, *C.R. Acad. Sci. Paris* **295**, (1982) 595–598.
10. D. Perrin, Recent results on automata and infinite words, in *Mathematical Foundations of Computer Science, Lecture Notes in Computer Science* **176**, Springer, Berlin, (1984) 134–148.
11. D. Perrin, An introduction to automata on infinite words, in *Automata on Infinite Words (Nivat, M. ed.), Lecture Notes in Computer Science* **192**, Springer, Berlin, (1984) 2–17.
12. D. Perrin and J.-É. Pin, First order logic and star-free sets, *J. Comput. System Sci.* **32**, (1986), 393–406.
13. D. Perrin and J.-É. Pin, Semigroups and automata on infinite words, in NATO Advanced Study Institute *Semigroups, Formal Languages and Groups, J. Fountain (ed.)*, Kluwer academic publishers, (1995), 49–72.
14. D. Perrin and J.-É. Pin, *Mots infinis*, to appear (LITP report 97–04), (1997). Accessible on the web: `http://liafa.jussieu.fr/~jep`.
15. J.-É. Pin, *Variétés de langages formels*, Masson, Paris (1984); English translation: *Varieties of formal languages*, Plenum, New-York (1986).
16. J.-É. Pin, Finite semigroups and recognizable languages: an introduction, in NATO Advanced Study Institute *Semigroups, Formal Languages and Groups, J. Fountain (ed.)*, Kluwer academic publishers, (1995), 1–32.
17. J.-É. Pin, A variety theorem without complementation, *Russian Mathematics (Iz. VUZ)* **39** (1995), 80–90.
18. J.-É. Pin, A negative answer to a question of Wilke on varieties of ω-languages, *Information Processing Letters*, (1995), 197–200.

19. J.-É. Pin, Logic, Semigroups and Automata on Words, *Annals of Mathematics and Artificial Intelligence* **16** (1996), 343–384.
20. J.-É. Pin, Syntactic semigroups, in *Handbook of formal languages*, G. Rozenberg and A. Salomaa (éd.), Springer Verlag, 1997.
21. J.-É. Pin and P. Weil, A Reiterman theorem for pseudovarieties of finite first-order structures, *Algebra Universalis* **35** (1996), 577–595.
22. J.-É. Pin and P. Weil, Polynomial closure and unambiguous product, *Theory Comput. Systems* **30** (1997), 1–39.
23. I. Simon, Piecewise testable events, *Proc. 2nd GI Conf., Lecture Notes in Computer Science* **33**, Springer, Berlin, (1975) 214–222.
24. W. Thomas, Automata on infinite objects, in *Handbook of Theoretical Computer Science, vol B, Formal models and semantics*, Elsevier, (1990) 135–191.
25. T. Wilke, An Eilenberg theorem for ∞-languages, in *Automata, Languages and Programming, Lecture Notes in Computer Science* **510**, Springer Verlag, Berlin, Heidelberg, New York, (1991), 588–599.
26. T. Wilke, An algebraic theory for regular languages of finite and infinite words, *Int. J. Alg. Comput.* **3**, (1993), 447–489.
27. T. Wilke, Locally threshold testable languages of infinite words, in *STACS 93, P. Enjalbert, A. Finkel, K.W. Wagner (Eds.), Lecture Notes in Computer Science* **665**, Springer, Berlin, (1993) 607–616.