Shuffle on positive varieties of languages.

Antonio Cano Gómez* and Jean-Éric Pin†

acano@dsic.upv.es, Jean-Eric.Pin@liafa.jussieu.fr

Abstract

We show there is a unique maximal positive variety of languages which does not contain the language $(ab)^*$. This variety is the unique maximal positive variety satisfying the two following conditions: it is strictly included in the class of rational languages and is closed under the shuffle operation. It is also the largest proper positive variety closed under length preserving morphisms. The ordered monoids of the corresponding variety of ordered monoids are characterized as follows: for every pair (a, b) of mutually inverse elements, and for every element z of the minimal ideal of the submonoid generated by a and b, $(abzab)^\omega \leq ab$. In particular this variety is decidable.

1 Introduction

The shuffle product is a standard tool for modeling process algebras [3]. This motivates the study of "robust" classes of recognizable languages which are closed under shuffle product. By "robust" classes, we mean classes which are closed under standard operations, like boolean operations, morphisms or inverse morphisms, etc. For instance, a complete classification is known for varieties of languages. Recall that a variety of languages is a class of recognizable languages closed under the following operations: union, intersection, complement, inverse morphisms and residuals. It is easy to see that the variety of all recognizable languages is closed under shuffle. Finding the proper varieties (ie. not equal to the variety of all recognizable languages) closed under shuffle proved to be a much more challenging problem. Actually, all these varieties are commutative, a very restrictive condition. In particular, the variety of commutative languages is the largest proper variety closed under shuffle. This result, first conjectured by Perrot in 1978 [7], was finally proved by Ésik and Simon in 1998 [6].

*Departamento de Sistemas Informáticos y Computación, Universidad Politénica de Valencia, Camino de Vera s/n, P.O. Box: 22012, E-46020 - Valencia.

†LIAFA, Université Paris VII and CNRS, Case 7014, 2 Place Jussieu, 75251 Paris Cedex 05, France. Work supported by INTAS project 1224.

In this paper, we are interested in positive varieties closed under shuffle. A positive variety is obtained by relaxing the definition of a variety, in the sense that only positive boolean operations (union and intersection) are allowed — no complement. Again the positive variety of all recognizable languages is closed under shuffle, but the question arises to know whether there is a largest proper positive variety closed under shuffle.

The main result of this paper is a positive solution to this problem. First we show there is a largest positive variety $\mathcal{W}$ which does not contain the language $(ab)^*$. Then we show that this variety $\mathcal{W}$ is the largest proper positive variety closed under shuffle. We also characterize the corresponding variety $\mathbf{W}$ of ordered monoids. An ordered monoid $(M, \leq)$ belongs to $\mathbf{W}$ if, for every pair (a, b) of mutually inverse elements of M, and for every element z of the minimal ideal of the submonoid generated by a and b, $(abzab)^\omega \leq ab$. It follows that the variety $\mathbf{W}$ is decidable, and consequently, there is an algorithm to decide whether or not a given recognizable language belongs to $\mathcal{W}$.

Another important property of $\mathcal{W}$ is proved along the way. We show that $\mathcal{W}$ is the largest proper positive variety closed under length preserving morphisms. This result is proved by first showing that power monoids form the algebraic counterpart of length preserving morphisms. This result is not new, but is adapted here for ordered monoids and positive varieties of languages.

Our proofs require some classical semigroup theory (Green's relations, etc.) but, more surprisingly, make a nontrivial use of profinite techniques, especially in the detailed study of the variety $\mathbf{W}$. It would be interesting to know whether this type of arguments can be avoided.

Our paper is organized as follows. Section 2 gives the basic definitions. Section 3 is devoted to the algebraic study of the variety $\mathbf{W}$. This study is completed by the examples and counterexamples presented in Section 4. Power semigroups are introduced in Section 5. They form the main algebraic tool for the study of the operations on languages considered in Section 6: the length preserving morphisms and the shuffle operation. Section 7 is devoted to our main result, and we conclude the paper in Section 8.

2 Preliminaries

We assume that the reader has a basic background in finite semigroup theory (in particular Green's relations). All semigroups and monoids considered in this paper are either free or finite.

In this section we provide the most important concepts and tools used in this article. Subsections 2.2, 2.4 and 2.6 come from [13, 18]. The reader is referred to [12, 13, 15, 18, 19, 14] for further information about ordered semigroups.

2.1 Semigroups

If S is a semigroup, S^1 denotes the monoid equal to S if S has an identity element and $S \cup \{1\}$ otherwise, with $s1 = 1s = s$ for all $s \in S$. An element $e \in S$ is idempotent if $e^2 = e$. The set of idempotents of a semigroup S is denoted by $E(S)$. Given an element s of a finite semigroup S, s^ω denotes the unique idempotent of the subsemigroup of S generated by s.

Two elements a and b of a semigroup are *mutually inverse* if $aba = a$ and $bab = b$.

A relation $\mathcal{R}$ on a monoid M is *stable on the right* (resp. left) if, for every $x, y, z \in M$, $x \mathcal{R} y$ implies $xz \mathcal{R} yz$ (resp. $x \mathcal{R} y$ implies $zx \mathcal{R} zy$). A relation is *stable* if it is stable on the right and on the left.

A congruence on a semigroup is a stable equivalence relation. If $\sim$ is a congruence on S, there is a well-defined multiplication on the quotient set $S/{\sim}$ given by $[s][t] = [st]$ where [s] denotes the $\sim$-class of $s \in S$.

An *ideal* of a semigroup S is a subset $I \subseteq S$ such that $S^1IS^1 \subseteq I$. A nonempty ideal I of a subsemigroup S is called *minimal* if, for every nonempty ideal J of S, $J \subseteq I$ implies $J = I$. Every finite semigroup S admits a unique minimal ideal, denoted by $I(S)$. In particular, if s is an element of S, the minimal ideal of the subsemigroup generated by s is a group, with identity s^ω. There is a unique element t of this group such that $ts = st = s^\omega$. This element t is denoted by $s^{\omega-1}$.

The next proposition, which applies in particular to minimal ideals, is a particular case of [10, Chapter 2, Proposition 1.2]. For the convenience of the reader, we give a self-contained proof. Recall that if J and K are subsets of a semigroup S, then $K^{-1}J = \{s \in S^1 \mid Ks \cap J \neq \emptyset\}$ and $JK^{-1} = \{s \in S^1 \mid sK \cap J \neq \emptyset\}$.

Proposition 2.1 *Let J be a $\mathcal{J}$-class of a semigroup S which is also a semigroup. Then $J^{-1}J = JJ^{-1}$ and this set is a submonoid of S^1 in which J is the minimal ideal.*

Proof. Let $s \in J^{-1}J$. By definition, there exists an element $t \in J$ such that $ts \in J$. Since J is a semigroup, it follows that tst and $(tst)^\omega$ are elements of J. Furthermore, since two conjugate idempotents are $\mathcal{J}$-equivalent, the idempotent $(stt)^\omega$ is also in J. But since $(stt)^\omega = st(tst)^{\omega-1}t$ and $t(tst)^{\omega-1}t \in J$, one has $s \in JJ^{-1}$. It follows that $J^{-1}J$ is contained in JJ^{-1} and a dual argument would show the opposite inclusion. Thus $J^{-1}J = JJ^{-1}$.

Let now $s_1, s_2 \in J^{-1}J$. Then $t_1s_1 \in J$ for some $t_1 \in J$, and since $J^{-1}J = JJ^{-1}$, $s_2t_2 \in J$ for some $t_2 \in J$. Therefore $t_1s_1s_2t_2$ and hence $(t_1s_1s_2t_2)^\omega$ are in J. By conjugacy, $(s_1s_2t_2t_1)^\omega \in J$ and since $t_2t_1(s_1s_2t_2t_1)^{\omega-1} = t_2(t_1s_1s_2t_2)^{\omega-1}t_1 \in J$, $s_1s_2 \in JJ^{-1}$. Therefore JJ^{-1} is a semigroup and since it clearly contains 1, it is a submonoid of S^1.

We claim that J is an ideal of JJ^{-1}. Let $s \in J^{-1}J$ and let $u \in J$. By definition, there exists an element $t \in J$ such that $ts \in J$. It follows $tsu \in J$ and thus $su \in J^{-1}J$. Symmetrically, since $J^{-1}J = JJ^{-1}$, there exits $r \in J$ such that $sr \in J$. It follows $usr \in J$, and thus $us \in JJ^{-1}$, which proves the claim. Since J is a simple semigroup, it is necessarily equal to the minimal ideal of JJ^{-1}. □

2.2 Ordered monoids

An ordered monoid is a monoid equipped with a stable partial order relation. For instance, U_1^- denotes the ordered monoid $\{0, 1\}$, consisting of an identity 1 and a zero 0, ordered by $1 \leq 0$.

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

The product of a family $(M_i)_{i \in I}$ of ordered monoids is the ordered monoid defined on the set $\prod_{i \in I} M_i$. The multiplication and the order relation are defined componentwise.

A *morphism* from an ordered monoid $(M, \leq)$ into an ordered monoid $(N, \leq)$ is a function $\varphi : M \to N$ such that $\varphi(1) = 1$, $\varphi(s_1 s_2) = \varphi(s_1)\varphi(s_2)$ and such that $s_1 \leq s_2$ implies $\varphi(s_1) \leq \varphi(s_2)$. Ordered submonoids and quotients are defined in the usual way. Complete definitions can be found in [18]. We just mention a special case of one of the so-called "homomorphism theorems". It is stated here in a negative form which is more suitable for our purpose (see the proof of Theorem 3.6).

Proposition 2.2 *Let $\leq_1$ and $\leq_2$ two order relations on a monoid M. If, for all $x, y \in M$, $x \not\leq_2 y$ implies $x \not\leq_1 y$, then $(M, \leq_2)$ is a quotient of $(M, \leq_1)$.*

Ordered monoids are a generalization of monoids. Taking the equality as a stable order relation, we obtain the same definitions as above for the unordered case.

An *order ideal* I of an ordered monoid $(M, \leq)$ is a subset of M such that if $x \in I$ and $y \leq x$ then $y \in I$. Given an element s of M, the set

$$\downarrow s = \{t \in M \mid t \leq s\}$$

is an order ideal, called the order ideal generated by s. More generally, if X is a subset of M, the order ideal generated by X is the set

$$\downarrow X = \bigcup_{s \in X} \downarrow s$$

A *filter* F of an ordered monoid $(M, \leq)$ is a subset of M such that if $x \in F$ and $x \leq y$ then $y \in F$. Note that a set is a filter if and only if its complement is an order ideal. Given an element s of M, the set

$$\uparrow s = \{t \in M \mid s \leq t\}$$

is a filter, called the filter generated by s. More generally, if X is a subset of M, the filter generated by X is the set

$$\uparrow X = \bigcup_{s \in X} \uparrow s$$

2.3 Rees quotient

Let M be a monoid and let I be an ideal of M. The *Rees quotient* of M by I, denoted by M/I, is the monoid defined on the set $(M \setminus I) \cup \{0\}$ by the multiplication (temporarily denoted by $s \cdot t$) defined as follows

$$s \cdot t = \begin{cases} st & \text{if } s,\, t \text{ and } st \text{ are in } M \setminus I \\ 0 & \text{otherwise} \end{cases}$$

The natural morphism π from M onto M/I is defined by

$$\pi(s) = \begin{cases} s & \text{if } s \in M \setminus I \\ 0 & \text{otherwise} \end{cases}$$

If M is an ordered monoid, it is not always possible to order the Rees quotient M/I in such a way that the natural morphism $\pi : M \to M/I$ be a morphism of ordered monoids. The next proposition gives some sufficient conditions that make possible the construction of such an order.

Proposition 2.3 *Let $(M, \leq)$ be an ordered monoid and let I be an ideal of M. Assume that no relations of the form $t_1 \leq s \leq t_2$ hold with $s \in M \setminus I$ and $t_1, t_2 \in I$. Define a relation $\preceq$ on M/I as follows. If $s_1, s_2 \in M \setminus I$, then $s_1 \preceq s_2$ if $s_1 \leq s_2$ or if $s_1 \leq t_1$ and $t_2 \leq s_2$ for some elements $t_1, t_2 \in I$. If $s \in M \setminus I$, $s \preceq 0$ (resp. $0 \preceq s$) if $s \leq t$ (resp. $t \leq s$) for some $t \in I$. Finally, $0 \preceq 0$. Then $\preceq$ is a stable order relation and $(M/I, \preceq)$ is a quotient of $(M, \leq)$.*

Proof. Let us show that the relation $\preceq$ is an order relation. It is clearly reflexive by construction. If $s_1 \preceq s_2$ and $s_2 \preceq s_1$, with $s_1, s_2 \in M \setminus I$, three cases may apparently arise, but two of them are not compatible with the hypothesis. More precisely, if $s_1 \leq t_1$ and $t_2 \leq s_2$ for some elements $t_1, t_2 \in I$, the relation $s_2 \leq s_1$ does not hold, for otherwise $t_2 \leq s_1 \leq t_1$, a contradiction. Similarly, if $s_2 \leq t_3$ and $t_4 \leq s_2$ for some $t_3, t_4 \in I$, one gets $t_2 \leq s_2 \leq t_3$ with a new contradiction. Thus, the only possible case is

$s_1 \leq s_2$ and $s_2 \leq s_1$, and hence $s_1 = s_2$. Finally, if $s \preceq 0 \preceq s$, for some $s \in M \setminus I$, there must exist two elements t_1, t_2 of I such that $t_2 \leq s \leq t_1$, with a new contradiction. Thus $\preceq$ is antisymmetric.

We finally prove that $\preceq$ is transitive. Suppose that $s_1 \preceq s_2$ and $s_2 \preceq s_3$. If $s_1 = 0$, then $t_1 \leq s_2$ for some element $t_1 \in I$. If $s_2 \leq t_2$ for some $t_2 \in I$, the condition of the theorem is violated. Therefore $s_2 \leq s_3$, whence $t_1 \leq s_3$ and $0 \preceq s_3$. The case $s_3 = 0$ is similar. Suppose that $s_2 = 0$. Then $s_1 \leq t_1$ and $t_3 \leq s_3$ for some $t_1, t_3 \in I$ and thus $s_1 \preceq s_3$. Suppose finally that $s_1, s_2, s_3 \in M \setminus I$. If $s_1 \preceq s_2$ and $s_2 \preceq s_3$, then we have $s_1 \leq s_3$ and hence $s_1 \preceq s_3$. If $s_1 \leq t_1$ and $t_2 \leq s_2$ for some $t_1, t_2 \in I$, then necessarily $s_2 \leq s_3$ by the same argument as before. Thus the relation $t_2 \leq s_3$ holds, and gives $s_1 \preceq s_3$.

Let us show that $\preceq$ is right stable. If $s \preceq 0$, then $s \leq t$ for some $t \in I$. Therefore $su \leq tu$ with $tu \in I$ and $su \leq 0$. If $s_1 \preceq s_2$, then either $s_1 \leq s_2$ and hence $s_1 u \leq s_2 u$, whence $s_1 u \preceq s_2 u$, or $s_1 \leq t \leq s_2$ for some $t \in I$ and thus $s_1 u \leq tu \leq s_2 u$, whence $s_1 u \preceq s_2 u$. A similar argument would show that $\preceq$ is left stable.

Let us show that π is a morphism of ordered monoids. If $s \leq t$ and $s, t \in M \setminus I$, then $\pi(s) = s \preceq t = \pi(t)$. If $s \in M \setminus I$ and $t \in I$, then $\pi(s) = s \preceq 0 = \pi(t)$. If $s \in I$ and $t \in M \setminus I$, then $\pi(s) = 0 \preceq t = \pi(t)$. Finally, if $s, t \in I$, then $\pi(s) = \pi(t) = 0$. □

2.4 Syntactic ordered monoids

A language L of A^* is recognized by an ordered monoid $(M, \leq)$ if and only if there exist an order ideal I of M and a monoid morphism η from A^* into M such that $L = \eta^{-1}(I)$.

Let A^* be a free monoid. Given a language P of A^* we define the *syntactic congruence* $\sim_P$ and the *syntactic preorder* $\leq_P$ as follows:

(1) $u \sim_P v$ if and only if for all $x, y \in A^*$, $xvy \in P \Leftrightarrow xuy \in P$,

(2) $u \leq_P v$ if and only if for all $x, y \in A^*$, $xvy \in P \Rightarrow xuy \in P$.

The monoid $A^*/{\sim_P}$ is called the *syntactic monoid* of P, and is denoted by $M(P)$. The monoid $A^*/{\sim_P}$, ordered with the stable order relation induced by $\leq_P$ is called the *ordered syntactic monoid* of P. The syntactic (ordered) monoid of a rational language is finite.

Two ordered monoids play an important role in the sequel. First U_1^-, the ordered syntactic monoid of the language a^* on the alphabet $\{a, b\}$, and B_2^{1-}, the ordered syntactic monoid of the language $(ab)^*$ on the alphabet $\{a, b\}$.

Example 2.1 The structure of the ordered syntactic monoid B_2^{1-} of the language $L_1 = (ab)^*$ on the alphabet $A = \{a, b\}$ is given in Figure 2.1.

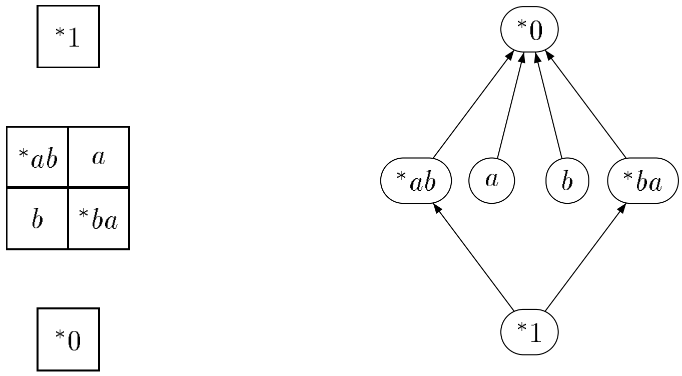

Figure 2.1: $\mathcal{J}$-classes and order of B_2^{1-}.

2.5 Profinite monoids

We briefly recall the definition of a free profinite monoid. More details can be found in [1, 2]. Let A be a finite alphabet. A monoid M *separates* two words u and v of the free monoid A^* if there exists a morphism φ from A^* onto M such that $\varphi(u) \neq \varphi(v)$. We set

$$r(u, v) = \min\{\, |M| \mid M \text{ is a monoid that separates } u \text{ and } v \,\}$$

and $d(u, v) = 2^{-r(u,v)}$, with the usual conventions $\min \emptyset = +\infty$ and $2^{-\infty} = 0$. Then d is an ultrametric on A^*. For the metric d, the closer are two words, the larger is the monoid needed to separate them.

As a metric space, A^* admits a completion, denoted by $\widehat{A^*}$. For instance, it can be shown that, for each $x \in \widehat{A^*}$, the sequence $(x^{n!})_{n\geq 0}$ is a Cauchy sequence. It converges to an idempotent element of $\widehat{A^*}$, denoted by x^ω. The product on A^* is uniformly continuous. Since A^* is dense in $\widehat{A^*}$ by definition, the product can be extended by continuity to $\widehat{A^*}$. The resulting monoid is called the *free profinite monoid* on A. This is a topological compact monoid which admits a unique minimal ideal.

Every monoid morphism from A^* into a finite monoid M (considered as a discrete metric space), can be extended by continuity to a morphism from $\widehat{A^*}$ into M. In particular, the image of x^ω under any morphism $\varphi : \widehat{A^*} \to M$ into a finite monoid M is the unique idempotent of the subsemigroup of M generated by $\varphi(x)$. This fully justifies the natural formulas $\varphi(x^\omega) = (\varphi(x))^\omega$ and $\varphi(x^{\omega-1}) = (\varphi(x))^{\omega-1}$, which are, in practice, the only thing to remember.

2.6 Varieties

A variety of finite (ordered) monoids, or pseudovariety, is a class of finite monoids closed under taking submonoids, quotients and finite direct products. Varieties of (ordered) monoids will be denoted by boldface capital letters (e.g. $\mathbf{V}$, $\mathbf{W}$).

Let $u, v \in \widehat{A^*}$. A finite ordered monoid M *satisfies the identity* $u \leq v$ (resp. $u = v$) if and only if, for each morphism $\varphi : \widehat{A^*} \to M$, $\varphi(u) \leq \varphi(v)$ (resp. $\varphi(u) = \varphi(v)$). Given a set E of identities, it is easy to see that the class of finite ordered monoids satisfying all the identities of E form a variety of finite ordered monoids, denoted by $[\![E]\!]$.

Reiterman's theorem [21] shows that every variety of finite monoids can be defined by a set of identities. Pin and Weil [17] have extended this result to varieties of finite ordered monoids.

For instance the variety $\mathbf{Com}$ of finite commutative monoids is defined by the identity $xy = yx$. The variety $\mathbf{J_1^-} = [\![xy = yx, x^2 = x, 1 \leq x]\!]$ is generated by the ordered monoid U_1^-. It is the variety of semilattices ordered by $x \leq y$ if and only if $xy = y$.
A positive variety of languages is a class of recognizable languages $\mathcal{V}$ such that:

(1) for every alphabet A, $\mathcal{V}(A^*)$ is a positive boolean algebra (closed under union and intersection),

(2) if φ: $A^* \to B^*$ is a morphism of semigroups, $L \in \mathcal{V}(B^*)$ implies that $\varphi^{-1}(L) \in \mathcal{V}(A^*)$,

(3) if $L \in \mathcal{V}(A^*)$ and if $a \in A$, then $a^{-1}L$ and La^{-1} are in $\mathcal{V}(A^*)$.

A variety of languages is a positive variety closed under complement.

Given two positive varieties of languages $\mathcal{V}$ and $\mathcal{W}$, we write $\mathcal{V} \subseteq \mathcal{W}$ if, for each alphabet A, $\mathcal{V}(A^*) \subseteq \mathcal{W}(A^*)$.

There is a one to one correspondence between varieties of finite monoids (resp. varieties of finite ordered monoids) and varieties of recognizable languages (resp. positive varieties of recognizable languages) [5, 12].

For instance, the positive variety of languages corresponding to $\mathbf{Com}$ is the variety $\mathcal{C}om$ of all commutative languages. Recall that a language L is commutative if $a_1a_2 \cdots a_n \in L$ implies $a_{\sigma(1)}a_{\sigma(2)} \cdots a_{\sigma(n)} \in L$ for each permutation σ of $\{1, 2, \ldots, n\}$. Other descriptions of $\mathcal{C}om$ can be found in [5, 11].

The positive variety of languages $\mathcal{J}_1^-$ corresponding to $\mathbf{J_1^-}$ is defined as follows: for each alphabet A, $\mathcal{J}_1^-(A^*)$ is the positive boolean algebra generated by the languages of the form B^* where $B \subseteq A$.

As a preparation to our main theorem, we prove a technical result on varieties containing the language $(ab)^*$. A finite language F of A^* is said to be *multilinear* if, for each letter of A, $\sum_{u \in F} |u|_a \leq 1$. Thus, for instance, the language $\{ab, cde\}$ is multilinear, but the language $\{ab, cae\}$ is not, because the letter a occurs twice: once in ab and another time in cae.

Proposition 2.4 *A positive variety containing the language $(ab)^*$ contains all the languages of the form L^*, where L is a multilinear language.*

Proof. Let a and b be distinct letters, and let $\mathcal{V}$ be a positive variety such that $\mathcal{V}(\{a,b\}^*)$ contains the language $(ab)^*$. Then B_2^{1-} belongs to the variety $\mathbf{V}$ corresponding to $\mathcal{V}$ and since U_1^- is a quotient of B_2^{1-} (obtained by identifying a, b, ab, ba and 0), U_1^- also belongs to $\mathbf{V}$. It follows that $\mathbf{V}$ contains $\mathbf{J_1^-}$, and thus $\mathcal{V}$ contains $\mathcal{J}_1^-$. In particular, for each alphabet A, $\mathcal{V}(A^*)$ contains the languages of the form B^*, where B is a subset of A.

Let $\{a_1,\ldots,a_n\}$ be a set of pairwise distinct letters and let A be an alphabet containing them. We first show that the language $(a_1\cdots a_n)^*$ is in $\mathcal{V}(A^*)$. Indeed, if B is a subset of A, denote by π_B the projection of A onto B defined by

$$\pi_B(a)=\begin{cases} a & \text{if } a\in B\\ 1 & \text{if } a\in A\setminus B\end{cases}$$

Let us show that

$$(a_1\cdots a_n)^* = \{a_1,\ldots,a_n\}^* \cap \Big(\bigcap_{i<j} \pi^{-1}_{\{a_i,a_j\}}(a_ia_j)^*\Big) \tag{1}$$

Let K be the right hand side of (1). It is clear that $(a_1\cdots a_n)^*$ is contained in K. Let $x=a_1\cdots a_n$ and let $u\in K$. We claim that u is a prefix of x^r for some $r>0$. If not, let $u=x^pa_1\cdots a_ka_iv$, with $p\geq 0$, $0\leq k<n$ and $i\neq k+1$. Then, if $i\leq k$, then $\pi_{\{a_i,a_{k+1}\}}(u)$ contains two consecutive occurrences of a_i, a contradiction. If now $i>k+1$, $\pi_{\{a_{k+1},a_i\}}(u)=(a_{k+1}a_i)^pa_i\pi_{\{a_{k+1},a_i\}}(v)$ and thus a_{k+1} has to be the first letter of v. But then $\pi_{\{a_{k+1},a_i\}}(u)=(a_{k+1}a_i)^pa_ia_{k+1}\pi_{\{a_{k+1},a_i\}}(v)\notin(a_{k+1}a_i)^*$, a contradiction again. Hence u is a prefix of x^r for a certain $r>0$. Symmetrically, u is a suffix of some x^s. Since x is multilinear, this implies that $u\in x^*$. Thus (1) holds.

Now, $(ab)^*\in\mathcal{V}(\{a,b\}^*)$, and thus the language $(a_ia_j)^*$ is in $\mathcal{V}(\{a_i,a_j\}^*)$. Since a positive variety is closed under intersection and inverse morphisms, Formula (1) shows that $(a_1\cdots a_n)^*\in\mathcal{V}(A^*)$.

Let now $L=\{u_1,u_2,\ldots,u_n\}$ be a multilinear language of A^*. For $1\leq i\leq n$, let C_i be the set of letters occurring in u_i, let $C=\bigcup_{1\leq i\leq n}C_i$ and let $\pi_i : A^*\to C_i^*$ be the morphism defined by

$$\pi_i(a)=\begin{cases} a & \text{if } a\in C_i\\ u_i & \text{otherwise}\end{cases}$$

We claim that

$$L^* = C^*\cap\Big(\bigcap_{1\leq i\leq n}\pi_i^{-1}(u_i^*)\Big) \tag{2}$$

Let R be the right hand side of (2). We first observe that

$$\pi_i(L) = \{u_i^{|u_1|}, \ldots, u_i^{|u_{i-1}|}, u_i, u_i^{|u_{i+1}|}, \ldots, u_i^{|u_n|}\}$$

and thus $\pi_i(L^*) \subset u_i^*$. It follows that L^* is a subset of R. Let now $v \in R$, and let us show that $v \in L^*$. First observe that, for $1 \leq i \leq n$,

$$\pi_i(v) \in u_i^* \tag{3}$$

Since, by a result of [4], L^* is a local language, it can be expressed as follows

$$L^* = \{1\} \cup \big((PA^* \cap A^*S) \setminus A^*NA^*\big)$$

where P (resp. S) is the set of first (resp. last) letters of the words of L, $N = A^2 \setminus F$ and F is the set of factors of length 2 of the words of L^*. Suppose that v is nonempty, and let a be its first letter. If $a \in C_i$, then $\pi_i(a) = a$, and thus the first letter of $\pi_i(v)$ is a. It follows from (3) that a is the first letter of u_i, and thus belongs to P. A similar argument would show that the last letter of v belongs to S. Consider now two consecutive letters a and b of v. If a and b belong to the same alphabet C_i, then $\pi_i(ab) = ab$ is a factor of $\pi_i(v)$ and thus, by (3), a factor of u_i^2. Thus $ab \in F$ in this case. Suppose now that $a \in C_i$ and $b \in C_j$ for some $i \neq j$. Then again $\pi_i(ab) = au_i$ is a factor of u_i^2 and thus a is the last letter of u_i. A similar argument would show that b is the first letter of u_j and thus ab is a factor of u_iu_j and hence belongs to F. It follows that v belongs to L^*, which proves the claim.

We have seen that if u_i is multilinear, then $u_i^* \in \mathcal{V}(A^*)$. It follows now from (2) that $L^* \in \mathcal{V}(A^*)$. $\square$

3 The variety W

This section is devoted to the algebraic study of a variety of ordered monoids which plays a central role in this article. Indeed, we shall see in Section 7 that the corresponding positive variety of languages is the largest proper positive variety closed under shuffle. This variety is denoted by $\mathbf{W}$ and is defined as follows: a monoid M belongs to $\mathbf{W}$ if and only if, for any pair (a, b) of mutually inverse elements of M, and any element z of the minimal ideal of the submonoid generated by a and b, $(abzab)^\omega \leq ab$.

It is not easy to see directly from the definition of $\mathbf{W}$ that it is a variety of ordered monoids. To overcome this difficulty, we shall give an equivalent definition of $\mathbf{W}$, that relies on an apparently weaker condition on the minimal ideal.

Let us denote by $\hat{F}$ the free profinite monoid generated by x and y. Given an element ρ of $\hat{F}$, and two elements u, v of a monoid M, we denote by $\rho(u,v)$ the image of ρ under the morphism from $\hat{F}$ into M which maps x onto u and y onto v. For instance, if $M = \hat{F}$, $\rho = (xy)^\omega x$, $u = yx$ and $v = xy^\omega x$, then $\rho(u,v) = (yxxy^\omega x)^\omega yx$.

Consider, for each element ρ of $\hat{F}$, the variety

$$\mathbf{W}_\rho = [\![((xy)^\omega \rho((xy)^{\omega-1}x, y(xy)^\omega)(xy)^\omega)^\omega \leq (xy)^\omega]\!]$$

The next proposition gives a simple characterization of these varieties.

Proposition 3.1 *An ordered monoid $(M, \leq)$ belongs to $\mathbf{W}_\rho$ if and only if, for any pair (a,b) of mutually inverse elements of M, $(ab\rho(a,b)ab)^\omega \leq ab$.*

Proof. Let $(M, \leq) \in \mathbf{W}_\rho$ and let (a,b) be a pair of mutually inverse elements of M. Then ab is idempotent, $(ab)^{\omega-1}a = a$, and $b(ab)^\omega = b$. Therefore, the identity defining $\mathbf{W}_\rho$ yields $(ab\rho(a,b)ab)^\omega \leq ab$.

Conversely, suppose that, for any pair (a,b) of mutually inverse elements of M, $(ab\rho(a,b)ab)^\omega \leq ab$. Let u and v be two elements of M. Then $a = (uv)^{\omega-1}u$ and $b = v(uv)^\omega$ are mutually inverse and satisfy $ab = (uv)^\omega$. Therefore, the relation $((uv)^\omega\rho((uv)^{\omega-1}u, v(uv)^\omega)(uv)^\omega)^\omega \leq (uv)^\omega$ holds in M, and hence $M \in \mathbf{W}_\rho$. $\square$

The definition of $\mathbf{W}_\rho$ is quite similar to that of $\mathbf{W}$, but the condition $(abzab)^\omega \leq ab$, which was imposed on any element z of the minimal ideal, is now restricted to only one element, namely $\rho(a,b)$. In particular, it is clear that if ρ is an element of the minimal ideal of $\hat{F}$, then $\mathbf{W} \subseteq \mathbf{W}_\rho$. The main result of this section states that this inclusion is actually an equality.

Theorem 3.2 *For any element ρ of the minimal ideal of $\hat{F}$, $\mathbf{W}_\rho = \mathbf{W}$.*

The proof relies on several lemmas. The first one gives a factorization of the elements of the minimal ideal of $\hat{F}$. Its statement requires an auxiliary notation. We denote by $\tilde{}$ the automorphism of $\hat{F}$ defined by $\tilde{x} = y$ and $\tilde{y} = x$. For instance $\widetilde{xyy} = yxx$.

Lemma 3.3 *Let ρ be an element of the minimal ideal of $\hat{F}$. Then either ρ or $\tilde{\rho}$ can be factorized as $\rho''x^2\rho'$, where ρ' belongs to the closure of the language $(yx)^*\{1,y\}$.*

Proof. Since ρ is in the minimal ideal of $\hat{F}$, $\rho \leq_{\mathcal{J}} x^2$ and thus ρ is the limit of a sequence of words of the form $r_n = r''_n x^2 r'_n$. Furthermore, we may assume that the occurrence of x^2 defined by the context (r''_n, r'_n) is the right-most occurrence of x^2 in r_n. Let I be the subset of $\mathbb{N}$ consisting of all indices n such that y^2 is not a factor of r'_n. If I is infinite, we simply consider the subsequence $(r_n)_{n\in I}$. Then r'_n contains no factor x^2 nor y^2 and its first letter cannot be x. In other words, $r'_n \in (yx)^*\{1,y\}$. Since $\hat{F}$ is compact, one can extract a subsequence from r_n such that r''_n converges to some element ρ'' and r'_n converges to some element ρ'. By the choice of r'_n, ρ' belongs to the closure of the language $(yx)^*\{1, y\}$, and since the multiplication is uniformly continuous, $\rho = \rho'' x^2 \rho'$.

If I is finite, for n large enough, each r'_n can be written as $r'_n = s_n y^2 s'_n$, where the context (s_n, s'_n) defines the right-most occurrence of y^2 in r'_n. Setting $s''_n = r''_n x^2 s_n$, we obtain $r_n = s''_n y^2 s'_n$, with $s'_n \in (xy)^*\{1, x\}$. It follows that $\tilde{r}_n = \widetilde{s''_n} x^2 \widetilde{s'_n}$, where $\widetilde{s'_n} \in (yx)^*\{1, y\}$, and we conclude as in the previous case. □

Let M be an ordered semigroup of $\mathbf{W}_\rho$ and let (a, b) be a pair of mutually inverse elements of M. Let N be the ordered subsemigroup of M generated by a and b. Set $e = (ab\rho(a, b)ab)^\omega$ and $f = (ba\rho(b, a)ba)^\omega$. We first observe that $abe = e = eab$ and $baf = f = fba$ since ab and ba are idempotent. Two other relations require a separate proof.

Lemma 3.4 *The relations $e = (afb)^\omega$ and $f = (bea)^\omega$ hold in N.*

Proof. Let I be the minimal ideal of N. Since ρ is an element of the minimal ideal of $\hat{F}$, $\rho(a, b) \in I$ and thus e and f are in I. Since $M \in \mathbf{W}_\rho$ and N is an ordered subsemigroup of M, Proposition 3.1 shows that the relations $e \leq ab$ and $f \leq ba$ hold in N. From the relation $e \leq ab$ follows $bea \leq baba = ba$ and since ba is idempotent, the relation $(bea)^\omega \leq ba$ holds in N. By multiplying both sides on the right (resp. on the left) by f, we obtain the relations $(bea)^\omega f \leq baf = f$ and $f(bea)^\omega \leq fba = f$. It follows $(bea)^\omega f f (bea)^\omega \leq ff$, that is

$$(bea)^\omega f (bea)^\omega \leq f \tag{4}$$

Now, since f and $(bea)^\omega$ are idempotent elements of I, $((bea)^\omega f (bea)^\omega)^\omega = (bea)^\omega$ and hence

$$(bea)^\omega \leq f \tag{5}$$

Similarly, by multiplying both sides of the relation $f \leq ba$ by $(bea)^\omega$ on the left (resp. right), we obtain $f(bea)^\omega \leq (bea)^\omega$ and $(bea)^\omega f \leq (bea)^\omega$, whence $f(bea)^\omega f = f(bea)^\omega (bea)^\omega f \leq (bea)^\omega$ and by taking the ω-power on both sides

$$f = (f(bea)^\omega f)^\omega \leq (bea)^\omega \tag{6}$$

Relations (5) and (6) together give $f = (bea)^\omega$. It follows that $(afb)^\omega = (a(bea)^\omega b)^\omega = e$. □

It follows $f \leq_{\mathcal{L}} ea$ and thus $ea \mathrel{\mathcal{L}} f$ since ea and f are both in I.

Lemma 3.5 *The element e (resp. f) belongs to the left ideal Na^2b (resp. Na^2).*

Proof. It follows immediately from Lemma 3.4 that the two conditions $e \in Na^2b$ and $f \in Na^2$ are equivalent.

Lemma 3.3 leads to the consideration of two cases. First assume that $\rho = \rho'' x^2 \rho'$, where ρ' belongs to the closure $\bar{L}$ of the language $L = (yx)^*\{1, y\}$. Let $\pi : \hat{F} \to N$ be the continuous morphism defined by $\pi(x) = a$ and $\pi(y) = b$. By a standard result of topology, $\pi(\bar{L}) \subseteq \overline{\pi(L)}$, and since the closure of $\pi(L)$ is computed in the discrete monoid N, it reduces to $(ba)^*\{1, b\} = \{1, b, ba\}$. It follows that

$$\begin{aligned} ab\rho(a,b)ab = ab\rho''(a,b)a^2\rho'(a,b)ab &\in Na^2\{1, b, ba\}ab \\ &= Na^3b \cup Na^2b \cup Na^2ba^2b \subseteq Na^2b \end{aligned}$$

Thus $e = (ab\rho(a,b)ab)^\omega \in Na^2b$.

Next assume that $\tilde{\rho} = \rho'' x^2 \rho'$ (or, equivalently, $\rho = \tilde{\rho''} y^2 \tilde{\rho'}$), where $\rho' \in \bar{L}$. Observing that $L = (yx)^*y \cup (yx)^*$, we consider successively two subcases

(a) $\rho' \in \overline{(yx)^*y}$

(b) $\rho' \in \overline{(yx)^*}$

We claim that in case (a), $\rho'(b,a) = a$. Indeed, $\pi((xy)^*x) = (ab)^*a = \{a\}$, and thus $\rho'(b,a) \in \pi(\overline{(xy)^*x}) \subseteq \overline{\pi((xy)^*x)} = \{a\}$. It follows that

$$ab\rho(a,b)ab = ab\rho''(b,a)b^2\rho'(b,a)ab \in Nb^2a^2b \subseteq Na^2b$$

and again $e = (ab\rho(a,b)ab)^\omega \in Na^2b$. In case (b), we have, by a similar argument, $\rho'(a,b) \in \{1, ba\}$, whence

$$ba\rho(b,a)ba = ba\rho''(a,b)a^2\rho'(a,b)ba \in Na^2\{1, ba\}ba = Na^2$$

It follows that $f = (ba\rho(b,a)ba)^\omega \in Na^2$. □

We can now conclude the proof of Theorem 3.2 by showing that $M \in \mathbf{W}$. We claim that $e = (eaab)^\omega$. By Lemma 3.5, there exists an element w of N such that $e = waab$. The relation $e \leq ab$ gives on one hand $(ew)e(aab) \leq (ew)ab(aab)$, that is $eweaab \leq e$ and on the other hand $(ea)e \leq (ea)ab$, that is $eae \leq eaab$. Taking the ω-power on both sides of these relations gives

$$(eweaab)^\omega \leq e \text{ and } (eae)^\omega \leq (eaab)^\omega \tag{7}$$

But since I is a simple semigroup containing e, $(eweaab)^\omega = (eaab)^\omega$ and $(eae)^\omega = e$. Thus (7) reduces to $(eaab)^\omega \leq e$ and $e \leq (eaab)^\omega$, which proves the claim.

It follows that $e \mathrel{\mathcal{H}} eaab$, and hence $ea \mathrel{\mathcal{L}} eaaba = eaa$. Now since $ea \mathrel{\mathcal{L}} f$ by Lemma 3.4, we have

$$eaa \mathrel{\mathcal{L}} ea \mathrel{\mathcal{L}} f$$

A similar argument can be used to obtain the relations $fbb \mathrel{\mathcal{L}} fb \mathrel{\mathcal{L}} e$, $bbe \mathrel{\mathcal{R}} be \mathrel{\mathcal{R}} f$ and $aaf \mathrel{\mathcal{R}} af \mathrel{\mathcal{R}} e$. Therefore, by Green's lemma, the union of the $\mathcal{L}$-class L_e of e and the $\mathcal{L}$-class L_f of f is stable under right and left multiplication by a and b. More precisely, the right multiplication by a maps L_e onto L_f and L_f onto itself and the right multiplication by b maps L_f onto L_e and L_e onto itself. A similar argument would show that the union of the $\mathcal{R}$-classes of e and f is invariant under left multiplication by a and b. Since a and b generate N, it follows that the minimal ideal I is equal to the union of the $\mathcal{H}$ classes of e, ea, be and f, as represented in Figure 3.1.

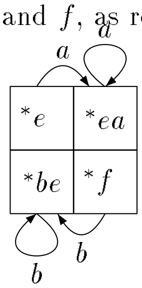

Figure 3.1: The elements a and b acting by right multiplication on I.

Finally, let z be an element of I. By the previous results, $abzab \; \mathcal{H} \; e$ and thus $(abzab)^\omega = e \leq ab$. Thus $M \in \mathbf{W}$. $\square$

The next proposition shows that $\mathbf{W}$ is the largest variety of ordered monoids not containing B_2^{1-}.

Theorem 3.6 *Every variety of ordered monoids not containing B_2^{1-} is contained in* $\mathbf{W}$.

Proof. First, $B_2^{1-} \notin \mathbf{W}$, since the relation $0 \leq ab$ does not hold in B_2^{1-}.

Let $(M, \leq)$ be an ordered monoid not in $\mathbf{W}$. Let $(u_1, v_1), \ldots, (u_n, v_n)$ be the list of pairs of mutually inverse elements of M and let N be the ordered submonoid of $\underbrace{M \times \cdots \times M}_{n \text{ times}}$ generated by $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$. Note that uv is an idempotent of N. The rest of the proof consists in proving that B_2^{1-} divides $N \times N$. As a first step, we exhibit a $\mathcal{J}$-class of N which is not a semigroup.

Suppose that $s \leq uv$ for some $s \in I(N)$. Let ρ be an element of the minimal ideal of $\hat{F}$ such that, in N, $\rho(u, v) = s$. Then, in M, the relations $\rho(u_i, v_i) \leq u_i v_i$ hold for $1 \leq i \leq r$, and hence the relation

$$(ab\rho(a, b)ab)^\omega \leq (ab)^\omega$$

holds for any pair (a, b) of mutually inverse elements of M. It follows by Proposition 3.1 that $M \in \mathbf{W}_\rho$, and thus, by Theorem 3.2, $M \in \mathbf{W}$, a contradiction. Thus, for each element $s \in I(N)$, $s \not\leq uv$. Therefore the set

$$S = \{s \in N \mid s \leq uv\}$$

is a nonempty subsemigroup of N disjoint from $I(N)$. Let h be an idempotent of $I(S)$ and let J be the $\mathcal{J}$-class of h in N.

Lemma 3.7 *The $\mathcal{J}$-class J is not a semigroup.*

Proof. By definition of S, $h \leq uv$ and thus $(uv)h(uv) \leq uv$. Therefore $(uv)h(uv) \in S$, and since $(uv)h(uv) \leq_{\mathcal{J}} h$ and $h \in I(S)$, $(uv)h(uv)\ \mathcal{J}\ h$ in S, and thus also in N. It follows that, in N, the relations $vh\ \mathcal{L}\ h\ \mathcal{R}\ hu$ hold, showing that $u \in J^{-1}J$ and $v \in JJ^{-1}$. In particular, if J is a semigroup, Proposition 2.1 shows that the set $J^{-1}J = JJ^{-1}$ is a submonoid of N whose minimal ideal is J. This submonoid contains u and v and therefore, is equal to N. It follows $J = I(N)$, a contradiction, since S does not meet $I(N)$. $\square$

Since J is not a semigroup, one can find two idempotents e, f in J such that $ef \notin J$. Since $e\ \mathcal{J}\ f$, there exist two elements $a, b \in J$ such that $e = ab$, $f = ba$, $aba = a$ and $bab = b$, as pictured in Figure 3.2, in which a star denotes the presence of an idempotent. Note that b is not idempotent, otherwise $ef = abba = aba = a \in J$. However it is possible that $h = e$ or $h = f$, and that a is idempotent.

$*$	$^*e$	a
	b	$^*f$
$^*h$		

Figure 3.2: The elements e, f and h in J.

Let R be the submonoid of $N \times N$ generated by $x = (a, b)$ and $y = (b, a)$. Then $xyx = x$, $yxy = y$, $xy = (e, f)$ and $yx = (f, e)$. Thus xy and yx are idempotents. Suppose that x or y is idempotent. Then a and b are idempotent, and thus $ef = a$ and $fe = b$ are in J, a contradiction. Thus neither x nor y are idempotent.

We claim that B_2^{1-} is a quotient of R. The monoid R is the disjoint union of the singleton $\{1\}$, the $\mathcal{J}$-class $D = \{x, y, xy, yx\}$ and the ideal $I = \{r \in R \mid r <_{\mathcal{J}} x\}$. As a monoid, R/I is isomorphic to B_2^1. We now carefully analyse the forbidden relations among the elements of R.

Lemma 3.8 *None of the following relations hold in R:*

(a) $1 \leq x$, $1 \leq y$,

(b) $r \leq s$ *for some* $r \in I$, $s \in D \cup \{1\}$.

(c) $r \leq s$ *for some* $r \in D$ *and* $s \in R \setminus \{r, 0\}$,

Proof. Let us first observe that x and y are not $\leq$-comparable. Indeed, if for instance $y \leq x$, then $(b, a) \leq (a, b)$, whence $a = b$, a contradiction.

If the relation $1 \leq x$ holds, then $y \leq xy \leq xyx = x$, and hence $y \leq x$, a forbidden relation. By a similar argument, the relations $1 \leq y$, $x \leq 1$ and $y \leq 1$ cannot hold. In particular, this proves (a).

Suppose that (b) holds with $s = xy$. Then $rx \leq xyx = x$. Similarly, if $r \leq yx$ then $xr \leq xyx = x$ and if $r \leq y$, then $xrx \leq xyx = x$. Therefore, the only remaining case is $r \leq x$, with $r \in I$. Setting $r = (r_1, r_2)$, we obtain $r_1 \leq a$ and $r_2 \leq b$. It follows that $r_1 r_2 \leq (ab)^\omega = e$. Since $e \mathcal{J} h$, $h = ced$ for some elements $c, d \in N$. Therefore $cr_1r_2d \leq ced = h \leq uv$ and thus $cr_1r_2d \in S$. But since $r \in I$, either $r_1 <_{\mathcal{J}} a$ or $r_2 <_{\mathcal{J}} b$. In both cases, it implies $cr_1r_2d <_{\mathcal{J}} h$, a contradiction, since $h \in I(S)$.

We now show that each of the remaining relations of type (c) implies a forbidden relation of type (b). If $xy \leq 1$ (resp. $yx \leq 1$), then $x^2y \leq x$ (resp. $yx^2 \leq x$), a type (b) relation. Similarly, if $r \leq 1$ with $r \in I$, then $rx \leq x$. Finally, if $r \leq s$ for some $r \in D$ and $s \in D \setminus \{r\}$, let $\bar{s}$ be the inverse of s in D (that is, if $s = x$, $\bar{s} = y$, if $s = y$, $\bar{s} = x$ and if $s = xy$ or $s = yx$, $\bar{s} = s$). Then $\bar{s}r\bar{s} \leq \bar{s}s\bar{s} = \bar{s}$ and since $s \neq r$, $\bar{s}r\bar{s} \in I$. Thus the relation $\bar{s}r\bar{s} \leq \bar{s}$ is of type (b). $\square$

In particular, no relation of the form $t \leq s$ hold with $t \in I$ and $s \in D \cup \{1\}$. It follows, by Proposition 2.3, that $(B_2^1, \preceq)$ is a quotient of R, where the order relation $\preceq$ is defined as follows. If $s_1, s_2 \in D \cup \{1\}$, then $s_1 \preceq s_2$ if $s_1 \leq s_2$. If $s \in D \cup \{1\}$, $s \preceq 0$ if $s \leq t$ for some $t \in I$. Finally, $0 \preceq 0$. It remains now to apply Proposition 2.2 to conclude that B_2^{1-} is a quotient of $(B_2^1, \preceq)$. It follows that B_2^{1-} divides $N \times N$. $\square$

Denote by **DS** the variety of ordered monoids whose regular $\mathcal{D}$-classes are subsemigroups. This variety contains in particular the variety of commutative monoids. Since the ordered monoid B_2^{1-} does not belong to **DS**, the following corollary is an immediate consequence of Theorem 3.6.

Corollary 3.9 *The variety* **W** *contains the variety* **DS***.*

4 Some examples and counterexamples

It is tempting to try to simplify the definition of **W**, either by relaxing or by strengthening the defining condition of **W**. This section presents the Hall of Fame of these failed attempts.

The first natural attempt is to require that, for each pair (a, b) of mutually inverse elements, there exists an idempotent e in the minimal ideal of the submonoid generated by a and b such that $abeab \leq ab$. However, the monoid M of Example 4.1 satisfies this condition (and even a stronger condition), but does not belong to **W**.

Example 4.1 Let $(M, \leq)$ be the ordered monoid presented by the relations $a^3 = a^2$, $b^3 = b^2$, $aba = a$, $bab = b$, $ba^2b^2 = a^2b^2$ and $a^2b^2 \leq ab$. There are 21 elements in M, and its $\mathcal{J}$-class structure is represented in Figure 4.1 below.

^{*1}

$^*ab$	a
b	$^*ba$

$^*a^2$	$^*a^2b$
$^*ba^2$	ba^2b

$^*b^2$	$^*b^2a$
$^*ab^2$	ab^2a

b^2a^2	b^2a^2b
ab^2a^2	ab^2a^2b

$^*a^2b^2$	$^*a^2b^2a$	$^*a^2b^2a^2$	$^*a^2b^2a^2b$

Figure 4.1: The $\mathcal{J}$-class structure of M.

The order relation is defined as follows

$$\begin{aligned} a^2b^2 &< b, ab, b^2, a^2b, ab^2, ba^2b, b^2a^2b, ab^2a^2b, a^2b^2a^2b \\ a^2b^2a &< a, a^2, ba, ba^2, b^2a, ab^2a, b^2a^2, ab^2a^2, a^2b^2a^2 \\ a^2b^2a^2 &< a^2, ba^2, b^2a^2, ab^2a^2 \\ a^2b^2a^2b &< a^2b, ba^2b, b^2a^2b, ab^2a^2b \end{aligned}$$

This monoid (without order) is also the transition monoid of the automaton represented in Figure 4.2. However, it is not possible to choose the final states of this automaton in such a way that $(M, \leq)$ is the ordered syntactic monoid of this automaton.

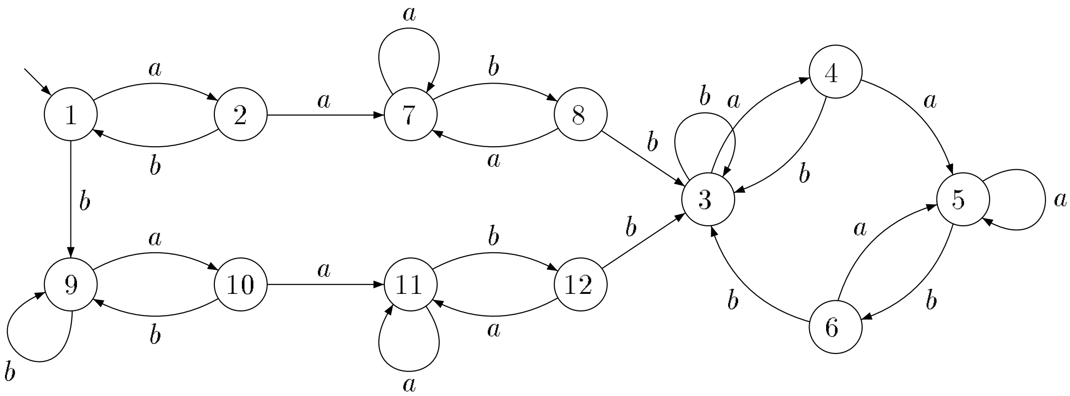

Figure 4.2: An automaton for M.

One can verify that, for each pair (x, y) of elements of M, there exists an idempotent e in the minimal ideal of the submonoid of M generated by x and y such that $e \leq (xy)^\omega$. However, $(M, \leq)$ does not belong to $\mathbf{W}$. Indeed a and b are mutually inverse in M, and generate M. However $z = a^2b^2a^2$ belongs to the minimal ideal of M, but $(abzab)^\omega = (aba^2b^2a^2ab)^\omega = a^2b^2a^2b \not\leq ab$.

Note that the quotient of M under the congruence $a^2b^2 = a^2b^2a^2b$ and $a^2b^2a = a^2b^2a^2$ is an ordered monoid of the variety defined by the identity

$$((xy)^\omega(x^\omega y^\omega)^\omega(xy)^\omega)^\omega \leq (xy)^\omega$$

which is contained in $\mathbf{W}$.

One can also try to strengthen the defining condition of $\mathbf{W}$ by requiring that, for any pair (a, b) of elements of M (not necessarily mutually inverse) and any element z of the minimal ideal of the submonoid generated by a and b, $(abzab)^\omega \leq ab$. However, the monoid M of Example 4.2 is in $\mathbf{W}$ but does not satisfy this condition.

Example 4.2 Let $(M, \leq)$ be the ordered monoid presented by the relations $a^2ba = a^2b$, $a^2b^2 = a^2b$, $aba^2 = ab^2$, $ba^2b = b^3$, $bab^2 = b^2$, $b^2ab = b^3$, $b^3a = b^3$, $b^4 = b^3$, $a^6 = a^5$, $a^5b = a^5$, $ababab = ab$, $ba^4b = ba^3b$, $b^2a^4 = b^2a^3$, $b^2a^3b = b^2a^3$, $ab^3 \leq abab$ and $b^3 \leq baba$. There are 34 elements in M, and its regular $\mathcal{J}$-classes are represented in Figure 4.3 below.

^{*1}

ab $^*abab$	aba $ababa$
bab $babab$	$^*baba$ $bababa$

$^*a^2b$
$^*a^3b$
$^*ab^3$
$^*b^3$
$^*ba^3b$
$^*a^4b$
$^*b^2a^3$
$^*ba^5$
$^*ab^2a^3$
$^*a^5$

Figure 4.3: The regular $\mathcal{J}$-classes of M.

This monoid (without order) is also the transition monoid of the automaton represented in Figure 4.4

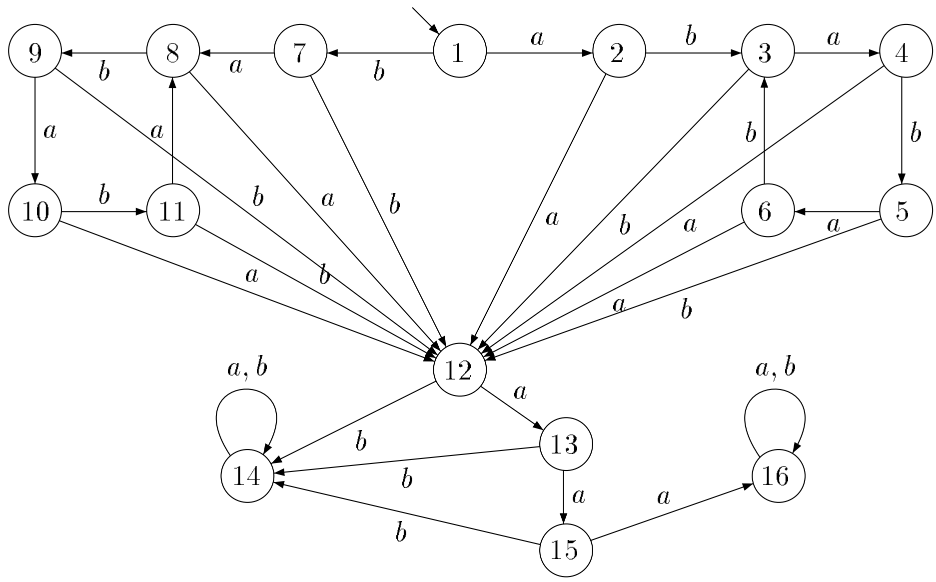

Figure 4.4: An automaton for M.

The order in M is defined by the following relations

$$b^3 < b^2, bab, b^2a, baba, b^2a^2, babab, b^2a^3, bababa$$
$$ab^3 < ab, aba, ab^2, abab, ab^2a, ababa, ab^2a^2, ab^2a^3$$

There are four elements of M which are their own inverse: ab, $abab$, $baba$ and $bababa$. Each of them generates a cyclic group of order 2, and thus, the condition defining $\mathbf{W}$ is trivially verified for these elements. The two other pairs of mutually inverse elements are $\{aba, babab\}$ and $\{bab, ababa\}$. The subsemigroup generated by aba and $babab$ is

$$\{aba, babab, abab, baba, ab^3, b^3\}$$

and the subsemigroup generated by $ababa$ and bab is

$$\{ababa, bab, abab, baba, ab^3, b^3\}$$

It follows that, for each pair of mutually inverse elements (x, y) of M and for any element z in the minimal ideal of the submonoid of M generated by x and y, the relation $((xy)^\omega z (xy)^\omega)^\omega \leq (xy)^\omega$ holds, and hence $(M, \leq) \in \mathbf{W}$. Actually, M even belongs to the variety $[\![(xy)^\omega (yx)^\omega (xy)^\omega \leq (xy)^\omega]\!]$. However $((ab)^\omega a^5 (ab)^\omega)^\omega = ab^2a^3 \not\leq (ab)^\omega = abab$.

Another attempt consisted to compare $\mathbf{W}$ with a variety of the form $\mathbf{W}_\rho$, where ρ is not necessarily in the minimal ideal of $\hat{F}$. By a suitable choice of ρ, we may insure that B_2^{1-} does not belong to $\mathbf{W}$ and thus $\mathbf{W}_\rho \subseteq \mathbf{W}$. One can take for instance $\rho = (yx)^\omega$. However, Example 4.3 shows that the variety $[\![((xy)^\omega (yx)^\omega (xy)^\omega)^\omega \leq (xy)^\omega]\!]$ is strictly contained in $\mathbf{W}$.

Example 4.3 Let $(M, \leq)$ be the ordered monoid with zero presented by the relations $aba = a$, $bab = b$, $ba^2b = ab^2a$, $a^3 = b^3 = 0$, $0 \leq a$ and $0 \leq b$. There are 15 elements in M, and its $\mathcal{J}$-class structure is represented in Figure 4.5 below.

^{*1}

a	$^*ab$
$^*ba$	b

a^2	a^2b	$^*a^2b^2$
ba^2	$^*ab^2a$	ab^2
$^*b^2a^2$	b^2a	b^2

^{*0}

Figure 4.5: The $\mathcal{J}$-class structure of M.

The order relation is defined by $0 \leq x$ for every $x \in M$. This ordered monoid does not belong to the variety $[\![((xy)^\omega(yx)^\omega(xy)^\omega)^\omega \leq (xy)^\omega]\!]$ since in M, ab and ba are idempotent, but $(ab)(ba)(ab) = ab^2a \not\leq ab$. However, M belongs to $\mathbf{W}$.

5 Power ordered monoids

The definitions given in this section were first given by the second author in [14].

Given a monoid M, we denote by $\mathcal{P}(M)$ the set of subsets of M with the multiplication defined, for all $X, Y \in \mathcal{P}(M)$ by

$$XY = \{xy \mid x \in X \text{ and } y \in Y\}$$

It is possible to extend this notion to ordered monoids. Let $(M, \leq)$ be an ordered monoid. Three ordered monoids, denoted respectively by $\mathcal{P}^+(M, \leq)$, $\mathcal{P}^-(M, \leq)$ and $\mathcal{P}(M, \leq)$, can be defined. Let $\leq_+$ be the relation defined on $\mathcal{P}(M)$ by setting $X \leq_+ Y$ if and only if, for all $y \in Y$, there exists $x \in X$ such that $x \leq y$, that is, if the filter generated by Y is included in the filter generated by X.

It is immediate to see that the relation $\leq_+$ is a stable preorder relation on $\mathcal{P}(M)$. Furthermore, if $Y \subseteq X$, then $X \leq_+ Y$. Denote by $\sim_+$ the equivalence relation defined by $X \sim_+ Y$ if $X \leq_+ Y$ and $Y \leq_+ X$. Then again, $\leq_+$ induces a stable order on the monoid $\mathcal{P}(M)/{\sim_+}$. The underlying ordered monoid is denoted by $\mathcal{P}^+(M, \leq)$. The monoid $\mathcal{P}^-(M, \leq)$ is the same monoid, equipped with the dual order.

For technical reasons, it is sometimes useful to use the monoid $\mathcal{P}'^+(M)$, which is the submonoid of $\mathcal{P}^+(M)$ obtained by removing the empty set. The two monoids are related as follows.

Proposition 5.1 *Let $(M, \leq)$ be an ordered monoid. Then $\mathcal{P}'^+(M, \leq)$ is a submonoid of $\mathcal{P}^+(M, \leq)$ and $\mathcal{P}^+(M, \leq)$ is a quotient of $U_1^- \times \mathcal{P}'^+(M, \leq)$.*

Proof. By definition, $\mathcal{P}^+(M, \leq) = \mathcal{P}'^+(M, \leq) \cup \{\{\emptyset\}\}$ and $\emptyset$ is a zero such that, for every $s \in M$, $s \leq \emptyset$. Thus $\mathcal{P}'^+(M, \leq)$ is a submonoid of $\mathcal{P}^+(M, \leq)$. Furthermore the map $\gamma : \mathcal{P}'^+(M, \leq) \times U_1^- \to \mathcal{P}^+(M, \leq)$ defined by $\gamma(X, 0) = \emptyset$ and $\gamma(X, 1) = X$ is a surjective morphism. Therefore, $\mathcal{P}^+(M, \leq)$ is a quotient of $U_1^- \times \mathcal{P}'^+(M, \leq)$. $\square$

To define $\mathcal{P}(M, \leq)$, we introduce another relation on $\mathcal{P}(M)$, denoted by $\leq$, and defined by setting $X \leq Y$ if and only if,

(1) for every $y \in Y$, there exists $x \in X$ such that $x \leq y$,

(2) for every $x \in X$, there exists $y \in Y$ such that $x \leq y$.

It is not difficult to see that $\leq$ is also a stable preorder on the semiring $\mathcal{P}(M)$. The associated semiring congruence $\sim$ is defined by setting $X \sim Y$ if $X \leq Y$ and $Y \leq X$. Then again, $\leq$ induces a stable order on the semiring $\mathcal{P}(M)/{\sim}$ and the underlying ordered semiring (resp. monoid) is denoted by $\mathcal{P}(M, \leq)$.

Looking at the definitions of the orders $\leq_+$, $\leq_-$ and $\leq$ for power monoids, one can describe the equivalence relations $\sim_+$, $\sim_-$ and $\sim$ and the corresponding equivalence classes $[\ ]_+$, $[\ ]_-$ and $[\ ]$.

Proposition 5.2 *Let $(M, \leq)$ be an ordered monoid and let X and Y be subsets of M.*

(1) *$X \sim_+ Y$ if and only if X and Y have the same set of minimal elements.*

(2) *$X \sim_- Y$ if and only if X and Y have the same set of maximal elements.*

(3) *$X \sim Y$ if and only if X and Y have the same set of minimal and maximal elements.*

The following characterization of the equivalence $\sim_+$ was given in [20].

Corollary 5.3 *Let $(M, \leq)$ be an ordered monoid and let X and Y be subsets of M. Then $X \sim_+ Y$ if and only if $\uparrow X = \uparrow Y$.*

In view of these results, there are two natural representations of the ordered power monoids. First, $\mathcal{P}^+(M,\leq)$ and $\mathcal{P}^-(M,\leq)$ can be identified with the set of antichains of M. If X and Y are two antichains of M, their product in $\mathcal{P}^+(M,\leq)$ (resp. in $\mathcal{P}^-(M,\leq)$) is the set of minimal (resp. maximal) elements of the set XY. In particular, if M is totally ordered, then both $\mathcal{P}^+(M,\leq)$ and $\mathcal{P}^-(M,\leq)$ are isomorphic to M, and $\mathcal{P}(M,\leq)$ is isomorphic to the set of intervals of M under the following multiplication: if I and J are two intervals, their product is the interval $[\min(I)\min(J), \max(I)\max(J)]$.

Secondly, as was observed in [20], $P^+(M,\leq)$ can be identified with the *monoid of filters* of M, where the product of two filters F and G is defined as the filter generated by the set FG, and the order relation is $\supseteq$. This is an immediate consequence of Corollary 5.3. Note that the identity of $P^+(M,\leq)$ is the filter generated by 1 and the maximal element is the empty filter.

In the rest of the paper, we shall use this latter approach and consider $P^+(M,\leq)$ as the monoid of filters of M. Let us mention a useful property.

Proposition 5.4 *Let $(M,\leq)$ be an ordered monoid and let X_1 and X_2 be subsets of M. Then in $\mathcal{P}^+(M,\leq)$ holds the formula $(\uparrow X_1)(\uparrow X_2) = \uparrow(X_1X_2)$.*

Proof. If $t \in (\uparrow X_1)(\uparrow X_2)$, then $t_1t_2 \leq t$ for some $t_1 \in \uparrow X_1$ and some $t_2 \in \uparrow X_2$. Thus $s_1 \leq t_1$ and $s_2 \leq t_2$ for some $s_1 \in X_1$ and $s_2 \in X_2$. It follows that $s_1s_2 \leq t_1t_2 \leq t$ and thus $t \in \uparrow(X_1X_2)$.

Conversely, let $t \in \uparrow(X_1X_2)$. Then $s_1s_2 \leq t$ for some $s_1 \in X_1$ and $s_2 \in X_2$. But since $s_1 \in \uparrow X_1$ and $s_2 \in \uparrow X_2$, $t \in (\uparrow X_1)(\uparrow X_2)$. □

Example 5.1 Let $(M,\leq)$ be the ordered monoid $(\{0,a,1\},\leq)$ in which 1 is the identity, 0 is a zero, $a^2 = a$ and $0 \leq a \leq 1$.

First, $\{0,1\} \sim \{0,a,1\}$. Thus in $\mathcal{P}(M)$, $\{0,1\}$ and $\{0,a,1\}$ should be identified. Similarly $\{0\} \sim_+ \{0,1\} \sim_+ \{0,a\} \sim_+ \{0,a,1\}$ and $\{a\} \sim_+ \{a,1\}$. Thus $\mathcal{P}^+(M,\leq) = \{\emptyset, \{0,a,1\}, \{a,1\}, \{1\}\}$. The orders are represented in Figure 5.1.

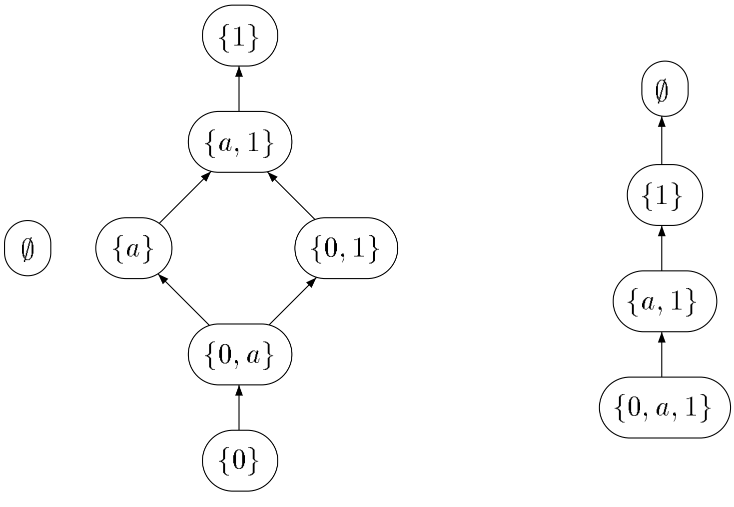

Figure 5.1: The monoids $\mathcal{P}(M)$ (on the left) and $\mathcal{P}^+(M)$ (on the right).

Given a variety of ordered monoids $\mathbf{V}$, we define $\mathbf{P}^+\mathbf{V}$ (resp. $\mathbf{P}'^+\mathbf{V}$) as the variety of ordered monoids generated by the monoids of the form $\mathcal{P}^+(M, \leq)$ (resp. $\mathcal{P}'^+(M, \leq)$) where $(M, \leq) \in \mathbf{V}$. The connection between the operators $\mathbf{P}^+$ and $\mathbf{P}'^+$ is a direct consequence of Proposition 5.1.

Proposition 5.5 *If* $\mathbf{V}$ *is a variety of ordered monoids containing* U_1^-*, then* $\mathbf{P}^+\mathbf{V} = \mathbf{P}'^+\mathbf{V}$.

A general result on power monoids is required to compute the varieties $\mathbf{P}^+\mathbf{W}$.

Proposition 5.6 *Let* $(M, \leq)$ *be an ordered monoid and let* $P \in \mathcal{P}(M)$*, then for each* $x \in P^\omega$ *there exists* $e \in E(P^\omega)$*, such that* $x \leq_{\mathcal{J}} e$ *in the semigroup* P^ω*.*

Proof. Let S be the semigroup P^ω. By [11, Proposition 1.12], there exists $n > 0$ such that $S^n = SE(S)S$. Since $S^2 = S$, it means that for all $x \in S$, there exists an idempotent $e \in S$ such that $x \leq_{\mathcal{J}} e$. □

Proposition 5.7 *The equality* $\mathbf{W} = \mathbf{P}^+\mathbf{W} = \mathbf{P}'^+\mathbf{W}$ *holds.*

Proof. Since $\mathbf{W}$ contains U_1^-, it suffices to show, by Proposition 5.5, that $\mathbf{W} = \mathbf{P}'^+\mathbf{W}$. Let ρ be an element of the minimal ideal of $\hat{F}$. By Theorem 3.2, $\mathbf{W} = \mathbf{W}_\rho$. We use the characterization given in Proposition 3.1. Let $(M, \leq)$ be an ordered monoid of $\mathbf{W}$ and let X and Y be two mutually inverse elements of $\mathcal{P}'^+(M, \leq)$. Then XY is idempotent, and if $a \in XY$, there exists by Proposition 5.6 an idempotent e of XY such that $a \leq_{\mathcal{J}} e$ in the semigroup XY. Thus there exist $a_1, a_2 \in XY$ such that $a = a_1 e a_2$. Furthermore, $e = x'y'$ for some $x' \in X$ and $y' \in Y$. Now, the elements

$x = (x'y')^{\omega-1}x' \in (XY)^{\omega-1}X = X$ and $y = y'(x'y')^{\omega} \in Y(XY)^{\omega} = Y$ are mutually inverse and satisfy $xy = e$, and thus $(e\rho(x,y)e)^{\omega} \leq e$ by Proposition 3.1, whence $a_1(e\rho(x,y)e)^{\omega}a_2 \leq a_1ea_2 = a$. Finally, since $a_1(e\rho(x,y)e)^{\omega}a_2 \in XY(XY\rho(X,Y)XY)^{\omega}XY = (XY\rho(X,Y)XY)^{\omega}$, the relation $(XY\rho(X,Y)XY)^{\omega} \leq_+ XY$ holds. Thus $\mathcal{P}'^+(M,\leq) \in \mathbf{W}_\rho$ and $\mathbf{P}'^+\mathbf{W} = \mathbf{W}$. $\square$

6 Operations on languages

In this section we establish a connection between the shuffle operation, the length preserving morphisms on languages and the operations of power ordered monoids defined in the previous section. These results are the counterpart, for positive varieties, of well-known results on varieties of languages [7, 22, 23].

6.1 Length preserving morphisms

A morphism φ from A^* into B^* is *length preserving* if $\varphi(A) \subseteq B$. Given a positive variety of languages $\mathcal{V}$, we define the positive variety of languages $\Lambda^+\mathcal{V}$ as follows. For each alphabet A, $\Lambda^+\mathcal{V}(A^*)$ consists of the languages which are positive boolean combinations of sets of the form $\varphi(L)$, where $L \in \mathcal{V}(B^*)$ for some finite alphabet B and φ is a length preserving morphism from B^* into A^*.

The next proposition extends to positive varieties a result of Straubing [23] about the relation between power monoids and length preserving morphisms. See also [20, Remark 1, page 414].

Proposition 6.1 *Let $\mathcal{V}$ be a positive variety of languages and let $\mathbf{V}$ be the corresponding variety of ordered monoids. Then $\Lambda^+\mathcal{V}$ is a positive variety of languages and the corresponding variety of ordered monoids is $\mathbf{P}^+\mathbf{V}$.*

Proof. Let $\mathcal{U}$ be the positive variety of languages corresponding to $\mathbf{P}^+\mathbf{V}$. We first show that $\mathcal{U} \subseteq \Lambda^+\mathcal{V}$. Every language of $\mathcal{U}(B^*)$ is a positive boolean combination of languages of the form $\psi^{-1}(\downarrow Z)$ where $\psi : B^* \to \mathcal{P}^+(M,\leq)$ is a morphism, $(M,\leq)$ is an ordered monoid of $\mathbf{V}$ and Z is an element of $\mathcal{P}^+(M,\leq)$, that is, a filter of M. Observe that $\downarrow Z$ is the order ideal of $\mathcal{P}^+(M,\leq)$ generated by Z, and since the order relation is reverse inclusion, $\downarrow Z$ actually denotes the set of filters of M that contain Z.

Set, for each $s \in M$,

$$X_s = \{u \in B^* \mid \psi(u) \cap \downarrow s \neq \emptyset\},$$

We claim that

$$\psi^{-1}(\downarrow Z) = \bigcap_{s \in Z} X_s \tag{8}$$

Indeed, first suppose that $u \in \psi^{-1}(\downarrow Z)$. Then $\psi(u)$ contains Z and thus, for each $s \in Z$, $\psi(u)$ meets $\downarrow s$. Therefore $u \in \bigcap_{s \in Z} X_s$. Conversely, if this property holds, then for each $s \in Z$, $\psi(u)$ meets $\downarrow s$, and since $\psi(u)$ is a filter, it contains s. Therefore $\psi(u)$ contains Z, proving (8).

Since $\Lambda^+\mathcal{V}(B^*)$ is a positive boolean algebra, it now suffices to show that $X_s \in \Lambda^+\mathcal{V}$ for each $s \in M$. Let

$$A = \{(b, x) \mid b \in B, x \text{ is a minimal element of } \psi(b)\}$$

Define a length preserving morphism $\varphi : A^* \to B^*$ by setting $\varphi(b, x) = b$ and a morphism of ordered monoids $\eta : A^* \to M$ by setting $\eta(b, x) = x$. We claim that

$$X_s = \varphi(\eta^{-1}(\downarrow s)) \tag{9}$$

First, if $b_1 \cdots b_n \in X_s$, then by definition there exist some elements $y_1 \in \psi(b_1)$, …, $y_n \in \psi(b_n)$ such that $y_1 \cdots y_n \leq s$. For $1 \leq i \leq n$, let us choose a minimal element $x_i \in \psi(b_i)$ such that $x_i \leq y_i$. Then (b_i, x_i) is a letter of A, $\eta(b_i, x_i) = x_i$ and $\varphi(b_i, x_i) = b_i$. Furthermore, $x_1 \cdots x_n \leq y_1 \cdots y_n \leq s$, and thus $b_1 \cdots b_n \in \varphi(\eta^{-1}(\downarrow s))$. Conversely, if $b_1 \cdots b_n \in \varphi(\eta^{-1}(\downarrow s))$, there exists, for $1 \leq i \leq n$, an element $x_i \in \psi(b_i)$ such that $x_1 \cdots x_n \leq s$. It follows that $b_1 \cdots b_n \in X_s$, proving the claim. Formula (9) shows that $X_s \in \Lambda^+\mathcal{V}$ and thus $\mathcal{U} \subseteq \Lambda^+\mathcal{V}$.

We now prove that $\Lambda^+\mathcal{V} \subseteq \mathcal{U}$. Let $\varphi : A^* \to B^*$ be a length preserving morphism and let $L \in \mathcal{V}(A^*)$. We want to prove that $\varphi(L) \in \mathcal{U}(B^*)$.

By definition, there is an ordered monoid $(M, \leq) \in \mathbf{V}$, a monoid morphism $\eta : A^* \to M$ and an order ideal P of M such that $L = \eta^{-1}(P)$.

Lemma 6.2 *The map $\psi : B^* \to \mathcal{P}^+(M, \leq)$ defined by setting, for each $u \in B^*$,*

$$\psi(u) = \{\uparrow \eta(v) \mid v \in A^*, \varphi(v) = u\},$$

is a morphism.

Proof. Let $u_1, u_2 \in A^*$. Let F be an element of $\psi(u_1)\psi(u_2)$. By definition, $F = (\uparrow \eta(v_1))(\uparrow \eta(v_2))$ for some $v_1, v_2 \in A^*$ such that $\varphi(v_1) = u_1$ and $\varphi(v_2) = u_2$. Now by Proposition 5.4,

$$(\uparrow \eta(v_1))(\uparrow \eta(v_2)) = \uparrow (\eta(v_1)\eta(v_2)) = \uparrow \eta(v_1 v_2) \tag{10}$$

Since $\varphi(v_1 v_2) = u_1 u_2$, it follows that $F \in \psi(u_1 u_2)$.

Conversely, let $F \in \psi(u_1 u_2)$. Then $F = \uparrow \eta(v)$ for some $v \in A^*$ such that $\varphi(v) = u_1 u_2$. Since φ is length preserving, $v = v_1 v_2$ for some v_1, v_2 such that $\varphi(v_1) = u_1$ and $\varphi(v_2) = u_2$. By (10), $F = (\uparrow \eta(v_1))(\uparrow \eta(v_2))$, and thus $F \in \psi(u_1)\psi(u_2)$. Therefore $\psi(u_1)\psi(u_2) = \psi(u_1 u_2)$. ◻

The set

$$\mathcal{F} = \{F \in \mathcal{P}^+(M, \leq) \mid F \cap P \neq \emptyset\}$$

is an order ideal. Furthermore, since P is an order ideal, the conditions $\uparrow s \cap P \neq \emptyset$ and $s \in P$ are equivalent. Therefore

$$\begin{aligned}
\psi^{-1}(\mathcal{F}) &= \{u \in B^* \mid \psi(u) \in \mathcal{F}\} \\
&= \{u \in B^* \mid \text{there exists } v \in A^*, \varphi(v) = u \text{ and } \uparrow \eta(v) \cap P \neq \emptyset\} \\
&= \{u \in B^* \mid \text{there exists } v \in A^*, \varphi(v) = u \text{ and } \eta(v) \in P\} \\
&= \varphi(L).
\end{aligned}$$

Thus $\varphi(L)$ is recognized by $\mathcal{P}^+(M, \leq)$ and hence belongs to $\mathcal{U}(B^*)$. Therefore $\Lambda^+\mathcal{V} \subseteq \mathcal{U}$. □

A slight adjustment in the proof would establish a similar result for surjective length preserving morphisms. More precisely, given a positive variety of languages $\mathcal{V}$, define the positive variety of languages $\Lambda'^+\mathcal{V}$ as follows. For each alphabet A, $\Lambda'^+\mathcal{V}(A^*)$ is the positive boolean closure of the class of sets of the form $\varphi(L)$, where $L \in \mathcal{V}(B^*)$ for some finite alphabet B and φ is a surjective length preserving morphism from B^* into A^*.

Proposition 6.3 *Let* $\mathbf{V}$ *be a variety of ordered monoids and let* $\mathcal{V}$ *be the corresponding variety of languages. Then the positive variety of languages corresponding to* $\mathbf{P}'^+\mathbf{V}$ *is* $\Lambda'^+\mathcal{V}$.

We now apply these results to the positive variety $\mathcal{W}$ corresponding to $\mathbf{W}$. Let us first give an immediate corollary of Theorem 3.6, Proposition 6.1 and Proposition 5.7.

Corollary 6.4 *The positive variety* $\mathcal{W}$ *is the largest variety not containing the language* $(ab)^*$. *It is closed under length-preserving morphisms.*

The next result concerns the varieties containing the language $(ab)^*$.

Theorem 6.5 *If a positive variety of languages* $\mathcal{V}$ *contains the language* $(ab)^*$, *then* $\Lambda^+\mathcal{V}$ *and* $\Lambda'^+\mathcal{V}$ *are both equal to the class of all rational languages.*

Proof. Let $\mathcal{V}$ be a positive variety of languages containing the language $(ab)^*$ and let $\mathbf{V}$ be the corresponding variety of ordered monoids. Then $\mathbf{V}$ contains B_2^{1-}, the ordered syntactic monoid of $(ab)^*$. And since U_1^- divides B_2^{1-}, $\mathbf{V}$ also contains U_1^-. It follows by Proposition 5.5 that $\mathbf{P}^+\mathbf{V} = \mathbf{P}'^+\mathbf{V}$ and hence, by Propositions 6.1 and 6.3, $\Lambda^+\mathcal{V} = \Lambda'^+\mathcal{V}$.

By Proposition 2.4, $\mathcal{V}$ contains all the languages of the form F^*, where F is a finite multilinear language.

The end of the proof is based on a result of [8] (see also [16, Theorem 8.1]), that we now briefly recall. Let L be a rational language of A^* and let $\mathcal{A} = (Q, A, \cdot, 1, F)$ be its minimal automaton, where $Q = \{1, 2, \ldots, n\}$. Let

$B = A \cup \{c\}$, where c is a new letter, and let $\tau : \mathbb{N} \to \mathbb{N}$ be the function defined by $\tau(k) = 2^{k-1} - 1$. Set

$$R = \{c^{\tau(i)} a c^{\tau(n)-\tau(i\cdot a)} \mid a \in A, i \in Q\} \quad \text{and} \quad P = \{c^{\tau(i)} \mid i \in F\}.$$

The aforementioned result states that $L = \psi^{-1}(R^*P)$, where ψ is the morphism from A^* to B^* defined by $\psi(a) = ac^{\tau(n)}$ for every $a \in A$.

Since a positive variety of languages is closed under inverse morphisms, it suffices to show that R^*P is in $\Lambda^+\mathcal{V}(B^*)$. Since $R^*P = \cup_{p \in P} R^*p$, it amounts to showing that, for each $p \in P$, R^*p is in $\Lambda^+\mathcal{V}(B^*)$. Observe that any word of P is the prefix of a word in R. We need a last lemma to conclude.

Lemma 6.6 *Every language of A^* of the form R^*p, where R is finite and p is the prefix of a word of R, can be written as $\varphi(F^*u^{-1})$, where φ is a length preserving morphism, F is a finite multilinear language and u is a word.*

Proof. Let $R = \{u_1, \ldots, u_k\}$, with $u_i = a_{i,1} \cdots a_{i,r_i}$. We may assume that $p = a_{1,1} \cdots a_{1,s}$ for some $s \le r_1$. Define a new alphabet with $r_1 + \ldots + r_k$ letters $B = \{b_{i,j_i} \mid 1 \le i \le k, 1 \le j_i \le r_i\}$ and a multilinear language $F = \{v_1, \ldots, v_p\}$, with $v_i = b_{i,j_1} \cdots b_{i,j_{r_i}}$. Finally let φ be the length preserving morphism from B^* into A^* defined by $\varphi(b_{i,j}) = a_{i,j}$ and let $u = b_{1,s+1} \cdots b_{1,r_1}$. Observe that $F^*u^{-1} = F^*v$, with $v = b_{1,1} \cdots b_{1,s}$. It follows that $\varphi(F^*u^{-1}) = \varphi(F^*v) = R^*p$. $\square$

Let us now conclude the proof of the theorem. We have seen that $\mathcal{V}$ contains the languages of the form F^*, where F is a finite multilinear set. Therefore, it also contains the languages of the form F^*u^{-1}. It follows by Lemma 6.6, that $\Lambda^+\mathcal{V}$ contains the languages of the form R^*p, where R is finite and p is the prefix of a word of R. $\square$

Corollary 6.7 *The variety $\mathcal{W}$ is the largest proper positive variety closed under length-preserving morphisms.*

6.2 The shuffle operator

The *shuffle* of two languages L_1 and L_2 of A^* is the language $L_1 ш L_2$ of A^* defined by:

$$L_1 ш L_2 = \{w \in A^* \mid w = u_1v_1 \cdots u_nv_n \text{ for some } n \ge 0 \text{ such that } u_1 \cdots u_n \in L_1, v_1 \cdots v_n \in L_2\}$$

The next proposition presents a connection between this operation and the operator $\mathcal{P}^+$.

Proposition 6.8 *Let L_1 and L_2 be two languages on A^* and let $(M_1, \leq_1)$ and $(M_2, \leq_2)$ be ordered monoids recognizing L_1 and L_2 respectively. Then $L_1 ш L_2$ is recognized by the ordered monoid $\mathcal{P}^+(M_1 \times M_2, \leq)$.*

Proof. Let, for $1 \leq i \leq 2$, $\eta_i : A^* \to M_i$ be a monoid morphism and let $P_i \subseteq M_i$ be an order ideal of M_i such that $L_i = \eta_i^{-1}(P_i)$. Then $P_1 \times P_2$ is an order ideal of $M_1 \times M_2$.

Define a morphism $\eta : A^* \to \mathcal{P}^+(M_1 \times M_2, \leq)$ by setting, for each $u \in A^*$,

$$\eta(u) = \{\uparrow(\eta_1(u_1), \eta_2(u_2)) \mid u \in u_1 ш u_2\}$$

Let us verify that η is a morphism of monoids. First,

$$\eta(1) = \{\uparrow(\eta_1(1), \eta_2(1))\} = \uparrow(1, 1)$$

which is the identity of $\mathcal{P}^+(M_1 \times M_2, \leq)$. Now, by Proposition 5.4, we have for all $u, v \in A^*$,

$$\begin{aligned}\eta(u)\eta(v) &= \{\uparrow(\eta_1(u_1), \eta_2(u_2)) \mid u \in u_1 ш u_2\} \\ &\qquad \{\uparrow(\eta_1(v_1), \eta_2(v_2)) \mid v \in v_1 ш v_2\} \\ &= \{\uparrow(\eta_1(u_1v_1), \eta_2(u_2v_2)) \mid u \in u_1 ш u_2, v \in v_1 ш v_2\}\end{aligned}$$

Now, since $uv \in x ш y$ if and only if there are factorizations $x = u_1u_2$ and $y = v_1v_2$ such that $u \in u_1 ш u_2$ and $v \in v_1 ш v_2$, one has

$$\eta(u)\eta(v) = \{\uparrow(\eta_1(x), \eta_2(y)) \mid uv \in x ш y\} = \eta(uv)$$

Now the set

$$\mathcal{F} = \{F \in \mathcal{P}^+(M_1 \times M_2, \leq) \mid F \cap (P_1 \times P_2) \neq \emptyset\}$$

is an order ideal of $\mathcal{P}^+(M_1 \times M_2, \leq)$. Furthermore

$$\eta^{-1}(\mathcal{F}) = \{u \in A^* \mid \eta(u) \cap P_1 \times P_2 \neq \emptyset\}$$

and by Proposition 5.2, and the definition of η,

$$\begin{aligned}\eta^{-1}(\mathcal{F}) = \{u \in A^* \mid & \text{ there exist } u_1, u_2 \in A^*, u \in u_1 ш u_2, \\ & \eta_1(u_1) \in P_1, \eta_2(u_2) \in P_2\} \\ = L_1 ш L_2 &\end{aligned}$$

Thus η recognizes $L_1 ш L_2$. □

The shuffle operation can be extended to positive varieties of languages as follows. Given a positive variety of languages $\mathcal{V}$, denote by $ш\mathcal{V}$ the positive variety of languages generated by $\mathcal{V}$ and by the languages of the form $L_1 ш L_2$, where L_1, L_2 are in $\mathcal{V}(A^*)$.

The closure of $\mathcal{V}$ under shuffle is the smallest positive variety containing $\mathcal{V}$ such that, if L_1 and L_2 are in $\mathcal{V}(A^*)$, $L_1 ш L_2$ is also in $\mathcal{V}(A^*)$.

7 Closure under shuffle

The following result was first conjectured in [9] and proved in [6]: "Given a variety of languages, either it is included in $\mathcal{C}om$ and then its closure under shuffle is included in $\mathcal{C}om$, or its closure under shuffle is the class of rational languages". It follows that there is a largest proper variety of languages closed under shuffle. We now establish a similar result for positive varieties of languages. The first step consists in adapting a proposition of [6] to positive varieties of languages.

Proposition 7.1 *If a positive variety of languages $\mathcal{V}$ contains the language $(ab)^*$, then $ш\mathcal{V}$ is the class of all rational languages.*

Proof. Let $\mathbf{V}$ be the variety of ordered monoids corresponding to $\mathcal{V}$. By Theorem 6.5, $\Lambda'^{+}\mathcal{V}$ is the class of all rational languages. Therefore, it suffices to show that $\Lambda'^{+}\mathcal{V} \subseteq ш\mathcal{V}$.

Let $L \in \mathcal{V}(A^*)$, and let $\varphi : A^* \to B^*$ be a surjective length preserving morphism. Note that $\varphi(L) \in \Lambda'^{+}\mathcal{V}$ by definition. Let c be a new letter and let $C = A \cup \{c\}$. Denote by π the projection from C^* onto A^* obtained by erasing all occurrences of c. We claim that the languages

$$L_1 = L \;ш\; c^* \qquad L_2 = (Ac)^* \qquad L_3 = A^*$$

are all in $\mathcal{V}(C^*)$. For L_1, this follows from the equality $L_1 = \pi^{-1}(L)$, since any positive variety is closed under inverse morphic images. Next, we observe that $L_2 = \gamma^{-1}((ab)^*)$, where γ denotes the morphism from C into $\{a, b\}$ mapping c to b and each letter of A to a. The fact that L_3 is in $\mathcal{V}(C^*)$ is a consequence of Proposition 2.4. It follows that the language

$$L_4 = (L_1 \cap L_2) \;ш\; L_3$$

belongs to $ш\mathcal{V}(C^*)$.

To finish the proof, let, for each $b \in B$, u_b be a word containing exactly one occurrence of each letter in $\varphi^{-1}(b)$, and no other letter. Consider the morphism $\eta : B^* \to C^*$ defined, for each $b \in B$, by $\eta(b) = u_b c$. We claim that

$$\varphi(L) = \eta^{-1}(L_4).$$

Indeed, let $u = a_1 \cdots a_n \in L$ and let, for $1 \leq i \leq n$, $b_i = \varphi(a_i)$. Let w_i be the word obtained by deleting the letter a_i in u_{b_i}. Now the word $a_1 c \cdots a_n c$ belongs to $L_1 \cap L_2$ and $w = w_1 \cdots w_n$ belongs to L_3. It follows that $u_{b_1} c \cdots u_{b_n} c$ belongs to L_4 and hence $b_1 \cdots b_n$ is in $\eta^{-1}(L_4)$. Thus $\varphi(L) \subseteq \eta^{-1}(L_4)$. To establish the opposite inclusion, consider a word $b_1 \cdots b_n$ in $\eta^{-1}(L_4)$. Then $w = \eta(b_1 \cdots b_n) = u_{b_1} c \cdots u_{b_n} c$ belongs to L_4. Therefore, there exist two words $u \in L_1 \cap L_2$ and $v \in L_3$ such that $w \in u \;ш\; v$. Setting $u = a_1 c a_2 c \cdots a_n c$, we have necessarily $a_1 \cdots a_n \in L$ since $u \in L \;ш\; c^*$.

Furthermore, for $1 \leq i \leq n$, a_i is a letter of u_{b_i}, since $w \in a_1 c a_2 c \cdots a_n c \text{ Ш } v$. It follows, by the definition of u_{b_i}, that $\varphi(a_i) = b_i$ and thus $b_1 \cdots b_n \in \varphi(L)$. This proves the claim and shows that $\varphi(L) \in \text{Ш}\mathcal{V}(B^*)$. Thus $\Lambda^+\mathcal{V} \subseteq \text{Ш}\mathcal{V}$ and hence $\text{Ш}\mathcal{V}$ is the class of all rational languages. □

Combining the results above (Lemma 6.8, Theorem 6.5, Proposition 5.7 and Proposition 3.6) we arrive to the final theorem.

Theorem 7.2 *The variety $\mathcal{W}$ is the largest proper positive variety which is closed under shuffle.*

8 Conclusion

It was shown that there is a largest positive variety not containing the language $(ab)^*$. This variety is also the largest proper positive variety closed under length-preserving morphisms, and the unique largest positive variety closed under shuffle. The corresponding variety of ordered monoids is defined by the identities

$$[\![((xy)^\omega \rho((xy)^{\omega-1}x, y(xy)^\omega)(xy)^\omega)^\omega (xy)^\omega)^\omega \leq (xy)^\omega]\!]$$

where ρ is any element of the minimal ideal of the free profinite monoid generated by x and y. It is also the class of all finite monoids M such that, for any pair (a, b) of mutually inverse elements of M, and any element z of the minimal ideal of the submonoid generated by a and b, $(abzab)^\omega \leq ab$. In particular, this variety is decidable.

Acknowledgements

The authors would like to thank Jorge Almeida, Zoltan Ésik, Pascal Weil and Marc Zeitoun for their useful comments and suggestions.

References

[1] J. Almeida, *Finite semigroups and universal algebra. Series in Algebra*, vol. 3, World Scientific, Singapore, 1994.

[2] J. Almeida and P. Weil, Relatively free profinite monoids: an introduction and examples, in *NATO Advanced Study Institute Semigroups, Formal Languages and Groups*, J. Fountain (éd.), pp. 73–117, Kluwer Academic Publishers, 1995.

[3] J. Baeten and W. Weijland, *Process algebra*, *Cambridge Tract in Theoretical Computer Science* vol. 18, Cambridge University Press, Cambridge UK, 1990.

[4] J. Berstel and J.-E. Pin, Local languages and the Berry-Sethi algorithm, *Theoret. Comput. Sci.* **155** (1996), 439–446.

[5] S. Eilenberg, *Automata, Languages and Machines*, vol. B, Academic Press, New York, 1976.

[6] Z. Ésik and I. Simon, Modeling Literal Morphisms by Shuffle, *Semigroup Forum* **56** (1998), 225–227.

[7] J.-F. Perrot, Variétés de langages et operations, *Theoret. Comput. Sci.* **7** (1978), 197–210.

[8] J.-E. Pin, Sur le monoïde de L^* lorsque L est un langage fini, *Theoret. Comput. Sci.* **7** (1978), 211–215.

[9] J.-E. Pin, Variétés de langages et monoïde des parties, *Semigroup Forum* **20** (1980), 11–47.

[10] J.-E. Pin, *Variétés de langages et variétés de semigroupes*, Thèse d'état, Université Paris VI, 1981.

[11] J.-E. Pin, *Varieties of formal languages*, North Oxford, London and Plenum, New-York, 1986. (Translation of *Variétés de langages formels*, Masson, 1984).

[12] J.-E. Pin, A variety theorem without complementation, *Russian Mathematics (Izvestija vuzov.Matematika)* **39** (1995), 80–90.

[13] J.-E. Pin, Syntactic semigroups, in *Handbook of formal languages*, G. Rozenberg and A. Salomaa (éd.), vol. 1, ch. 10, pp. 679–746, Springer-Verlag, 1997.

[14] J.-E. Pin, Algebraic tools for the concatenation product, *Theoret. Comput. Sci.* **292** (2003), 317–342.

[15] J.-E. Pin, A. Pinguet and P. Weil, Ordered categories and ordered semigroups, *Comm. Algebra* **30** (2002), 5651–5675.

[16] J.-E. Pin, H. Straubing and D. Thérien, Some results on the generalized star-height problem, *Information and Computation* **101** (1992), 219–250.

[17] J.-E. Pin and P. Weil, A Reiterman theorem for pseudovarieties of of finite first-order structures, *Algebra Universalis* **35** (1996), 577–595.

[18] J.-E. Pin and P. Weil, Semidirect products of ordered semigroups, *Communications in Algebra* **30** (2002), 149–169.

[19] J.-E. Pin and P. Weil, The wreath product principle for ordered semigroups, *Communications in Algebra* **30** (2002), 5677–5713.

[20] L. Polák, Operators on classes of regular languages, in *Semigroups, Algorithms, Automata and Languages*, G. Gomes, J.-E. Pin and P. Silva (éd.), pp. 407–422, World Scientific, 2002.

[21] J. Reiterman, The Birkhoff theorem for finite algebras, *Algebra Universalis* **14** (1982), 1–10.

[22] C. Reutenauer, Sur les variétés de langages et de monoïdes, in *Theoretical computer science (Fourth GI Conf., Aachen)*, vol. 67, pp. 260–265, Springer-Verlag, Berlin, 1979.

[23] H. Straubing, Recognizable Sets and Power Sets of Finite Semigroups, *Semigroup Forum* **18** (1979), 331–340.