Semigroup Forum
Springer-Verlag New York Inc.

Version of November 14, 1994

Finite semigroups as categories, ordered semigroups or compact semigroups

Jean-Eric Pin

Communicated by K. H. Hofmann and M. W. Mislove

I met Professor A.H. Clifford just once, in 1984. It was my first semigroup conference in the USA and I felt very honored when he told me very kindly how he enjoyed the way automata theory and semigroups mixed together. Ten years later, it was again a great honor for me to participate to the Tulane conference in his memory.

1. Introduction

Algebraic semigroup theory has made great strides in the recent years. It is a remarkable fact that these new results did require the introduction of some auxiliary structures. In this article, I would like to emphasize the rôle of three of these tools: topology, partial orders and categories. The three parts are relatively independent.

2. Identities and topology

The results presented in this section are a good illustration of the following quotation of Marshall Stone [34]: 'A cardinal principle of modern mathematical research may be stated as a maxim: "One must always topologize" '. Varieties of finite semigroups are a good example where Stone's principle was applied successfully.

Recall that a *variety of semigroups* is a class of semigroups closed under taking subsemigroups, quotients and direct products. A *variety of finite semigroups*, or *pseudovariety*, is a class of finite semigroups closed under taking subsemigroups, quotients and finite direct products.

It is a well known theorem of Birkhoff that varieties can be defined by identities. Formally, an identity is an equality of the form $u = v$ between two elements of the free semigroup Σ^+ on a countable number of generators Σ. A semigroup S satisfies an identity $u = v$ if and only if $u\varphi = v\varphi$ for every morphism $\varphi : \Sigma^+ \to S$. For instance, the identity $xy = yx$ defines the variety of commutative semigroups and $x = x^2$ defines the variety of bands. Birkhoff's theorem can be summarized by saying that each variety is an *equational class*. It was an interesting question to know whether pseudovarieties could also be defined by identities. The problem was solved by several authors but the most satisfactory answer is due to Reiterman [31]. Reiterman's theorem states that pseudovarieties are also equational classes. The difference with Birkhoff's theorem lies in the definition of the identities. For Reiterman, an identity is also a formal equality of the form $u = v$, but u and v are now elements of a certain completion $\hat{\Sigma}^+$ of the free semigroup Σ^+. The idea is actually quite simple. Let us say that a finite semigroup S *separates* two words $u, v \in \Sigma^+$ if $u\varphi \neq v\varphi$ for some morphism $\varphi : \Sigma^+ \to S$. Now set, for $u, v \in A^+$,

$$r(u, v) = \min\{\, |S| \;\mid\; S \text{ is a finite semigroup separating } u \text{ and } v \,\}$$

and $d(u,v) = 2^{-r(u,v)}$, with the usual conventions $\min \emptyset = +\infty$ and $2^{-\infty} = 0$. One can verify that d is a metric for which two words are close if a large semigroup is required to separate them. Now, (Σ^+, d) is a metric space and $\hat{\Sigma}^+$ is its completion. In particular, every element of $\hat{\Sigma}^+$ is the limit of some Cauchy sequence of (Σ^+, d). An important such limit is the ω-power, which traditionally designates the idempotent power of an element of a finite semigroup. If $x \in \hat{\Sigma}^+$, the sequence $(x^{n!})_{n \geq 0}$ converges in $\hat{\Sigma}^+$ to an idempotent denoted x^ω. Note that if $\pi : \hat{\Sigma}^+ \to S$ is a continuous morphism onto a finite discrete semigroup S, then $x^\omega \pi$ is equal to the unique idempotent power of $x\pi$.

Some more topology is required to fully understand Reiterman's theorem. We now consider only topological semigroups. In particular, as above, every finite semigroup is considered as equipped with the discrete topology. A topological semigroup S satisfies an identity $u = v$ (where $u, v \in \hat{\Sigma}^+$) if and only if $u\varphi = v\varphi$ for every continuous morphism $\varphi : \hat{\Sigma}^+ \to S$. Let $(I, \leq)$ be a partially ordered set and suppose that for every $i, j \in I$, there exists $k \in I$ such that $i \leq k$ and $j \leq k$. Assume that for every $i \in I$, there is a finite semigroup S_i and for every pair $(i,j) \in I \times I$ with $i \leq j$, there is a morphism $\pi_{j,i} : S_j \to S_i$ such that $\pi_{i,i}$ is the identity map for every $i \in I$ and if $i \leq j \leq k$, $\pi_{k,i} = \pi_{k,j}\pi_{j,i}$. Then $(S_i, \pi_{j,i})_{i,j \in I}$ is called a *projective system* and its *projective limit* is the subsemigroup of $\prod_{i \in I} S_i$ consisting of all elements $s = (s_i)_{i \in I}$ such that $s_j \pi_{j,i} = s_i$ for every $i \leq j$. A topological semigroup is *profinite* if it is a projective limit of finite semigroups. One can show that a topological semigroup is profinite if and only if it is compact and 0-dimensional (that is, every connected component is trivial). By extension, a topological semigroup is called *pro-*$\mathbf{V}$ if it is a projective limit of semigroups of $\mathbf{V}$, or, equivalently, if it is profinite and if all its finite continuous homomorphic images are in $\mathbf{V}$. Pro-$\mathbf{V}$ semigroups form a *pro-variety*, that is, a class of profinite semigroups closed under taking closed subsemigroups, continuous homomorphic images and direct products. The definition of a pro-variety allows infinite direct products and thus is very close to the original definition of a variety, with some topological sugar (or spice?) added. Reiterman's theorem can now be completed as follows. Let $\mathbf{V}$ be a pseudovariety and let E be a set of identities (in $\hat{\Sigma}^+$) defining $\mathbf{V}$. Then the class of all profinite semigroups satisfying E is exactly the class of pro-$\mathbf{V}$ semigroups. In other words, pro-varieties are equational classes.

Thus the core of Reiterman's theorem is a topological extension of Birkhoff's theorem. Its application to pseudovarieties just shows the emerging part of the iceberg, since pseudovarieties are then considered as the finite semigroups of a pro-variety.

Reiterman's theorem suggests that most standard results on varieties might be extended in some way to pseudovarieties. For instance, it is well known that varieties have free objects. More precisely, if $\mathbf{V}$ is a variety and A is a finite set, there exists an A-generated semigroup $F_A(\mathbf{V})$ of $\mathbf{V}$, such that every A-generated semigroup of $\mathbf{V}$ is a quotient of $F_A(\mathbf{V})$. This semigroup is unique (up to an isomorphism) and is called the *free semigroup* of the variety $\mathbf{V}$. How to extend this result to pseudovarieties? Again the answer is "topologize!". If $\mathbf{V}$ is a pseudovariety, one can relativize to $\mathbf{V}$ the definition of r and d as follows:

$$r_{\mathbf{V}}(u,v) = \min\{\, |S| \;\mid\; S \in \mathbf{V} \text{ and } S \text{ separates } u \text{ and } v \,\}$$

and $d_{\mathbf{V}}(u,v) = 2^{-r_{\mathbf{V}}(u,v)}$. The function $d_{\mathbf{V}}(u,v)$ still satisfies the triangular inequality and even the stronger inequality

$$d_{\mathbf{V}}(u,v) \leq \max\{\, d_{\mathbf{V}}(u,w), d_{\mathbf{V}}(w,v) \,\}$$

but it is not a metric anymore because one can have $d_{\mathbf{V}}(u,v) = 0$ with $u \neq v$: for instance, if $\mathbf{V}$ is the pseudovariety of commutative finite semigroups, $d_{\mathbf{V}}(xy,yx) = 0$ since xy and yx cannot be separated by a commutative semigroup. However, the relation $\sim_{\mathbf{V}}$ defined on A^+ by $u \sim_{\mathbf{V}} v$ if and only if $d_{\mathbf{V}}(u,v) = 0$ is a congruence and $d_{\mathbf{V}}$ induces a metric on the quotient semigroup $A^+/\sim_{\mathbf{V}}$. The completion of this metric space is a topological compact semigroup $\hat{F}_A(\mathbf{V})$, called the *free pro-*$\mathbf{V}$ *semigroup*. This semigroup is pro-$\mathbf{V}$, is generated by A as a topological semigroup (this just means that $A^+/\sim_{\mathbf{V}}$ is dense in $\hat{F}_A(\mathbf{V})$) and satisfies the following universal properties:

(1) If σ is a map from A into a pro-$\mathbf{V}$ semigroup S, then σ induces a unique continuous morphism $\hat{\sigma} : \hat{F}_A(\mathbf{V}) \to S$

(2) Every A-generated semigroup of $\mathbf{V}$ is a continuous homomorphic image of $\hat{F}_A(\mathbf{V})$.

(3) Let S and T be pro-$\mathbf{V}$ semigroups. If $\sigma : \hat{F}_A(\mathbf{V}) \to S$ and $\varphi : T \to S$ are continuous morphisms and if φ is onto, then there exists a continuous morphism $\pi : \hat{F}_A(\mathbf{V}) \to T$ such that $\pi\varphi = \sigma$.

One can actually extend this construction to the case where A is a *profinite* set, that is a topological set which is a projective limit of finite sets [28]. This more general setting is required in the study of the semidirect product [5, 6].

A more detailed presentation of Reiterman's theorem can be found in [1, 2, 4, 38].

3. Ordered semigroups and Eilenberg's theorem

An *ordered semigroup* $(S, \leq)$ is a semigroup S equipped with a (partial) *stable* order relation $\leq$: for every $u, v, x \in S$, $u \leq v$ implies $ux \leq vx$ and $xu \leq xv$. An *order ideal* of $(S, \leq)$ is a subset I of S such that, if $x \leq y$ and $y \in I$, then $x \in I$. Morphisms, ordered subsemigroups, quotients, direct products and (pseudo)varieties are defined in the natural way. Note that the free ordered semigroup on a set A is just $(A^+, =)$, where A^+ is the free semigroup on A.

The *dual* of an ordered semigroup $(S, \leq)$ is the ordered semigroup S equipped with the dual order $\leq'$ defined by $u \leq' v$ if and only if $v \leq u$. The class of all dual ordered semigroups of members of a (pseudo)variety of ordered semigroups $\mathbf{V}$ is also a (pseudo)variety, called the *dual* of $\mathbf{V}$ and denoted $\breve{\mathbf{V}}$.

Ordered semigroups occurred recently in connection with language theory [26]. The reader is referred to one of the books [12, 13, 18, 22] or to the survey article [24] for an introduction to this theory. Recall that a *variety of languages* is a class of recognizable languages closed under finite union, finite intersection, complement, left and right quotients and inverse morphisms between free semigroups. Eilenberg's variety theorem gives a bijective correspondence between varieties of finite semigroups and varieties of languages. However, certain important classes of recognizable languages occurring in language theory are not closed under complement but are closed under the other operations defining a variety of languages. This observation motivated the following definition: a *positive variety of languages* is a class of recognizable languages closed under finite union, finite intersection, left and right quotients and inverse morphisms between free semigroups. The term "positive" is borrowed from formal logic, where a *positive formula* is a formula without negation.

It turns out that Eilenberg's variety theorem can be extended to positive varieties. On the algebraic side, varieties of finite semigroups are replaced by varieties of finite ordered semigroups. The first thing to do is to extend the

definition of recognizability and the second will be to generalize the notion of syntactic semigroup.

Let $(S, \leq)$ be an finite ordered semigroup and let η be a surjective semigroup morphism from A^+ onto S, which can be considered as a morphism of ordered semigroup from $(A^+, =)$ onto $(S, \leq)$. A language L of A^+ is said to be *recognized* by η if $L = P\eta^{-1}$ for some order ideal P of S. Note that if $(T, \leq)$ recognizes L and $(T, \leq)$ is a quotient of $(S, \leq)$, then $(S, \leq)$ recognizes L. By extension, a language L is said to be *recognized* by $(S, \leq)$ if there exists a surjective morphism from A^+ onto S that recognizes Q.

Let L a language of A^+. One defines a stable quasiorder $\preceq_L$ and a congruence relation $\sim_L$ on A^+ by setting

$$u \preceq_L v \text{ if and only if, for every } x, y \in A^*, xvy \in L \text{ implies } xuy \in L$$

$$u \sim_L v \text{ if and only if } u \preceq_L v \text{ and } v \preceq_L u$$

The congruence $\sim_L$ is called the *syntactic congruence* of L and the quasiorder $\preceq_L$ induces a stable order $\leq_L$ on $S(L) = A^+/\sim_L$. The ordered semigroup $(S(L), \leq_L)$ is called the *syntactic ordered semigroup* of L, the relation $\leq_L$ is called the *syntactic order* of L and the canonical morphism η_L from A^+ onto $S(L)$ is called the *syntactic morphism* of L. The syntactic ordered semigroup is the smallest ordered semigroup that recognizes L. More precisely, an ordered semigroup $(S, \leq)$ recognizes L if and only if $(S(L), \leq_L)$ is a quotient of $(S, \leq)$.

If $\mathbf{V}$ is variety of finite ordered semigroups and A is a finite alphabet, we denote by $A^+\mathcal{V}$ the set of recognizable languages of A^+ which are recognized by an ordered semigroup of $\mathbf{V}$. Equivalently, $A^+\mathcal{V}$ is the set of recognizable languages of A^+ whose ordered syntactic semigroup belongs to $\mathbf{V}$. It is shown in [26] that the correspondence $\mathbf{V} \mapsto \mathcal{V}$ is a bijective correspondence between varieties of finite ordered semigroups and positive varieties of languages. The opposite correspondence is also easy to describe: with each positive variety of languages is associated the variety of finite ordered semigroups generated by all ordered syntactic semigroups of languages of $\mathcal{V}$. A similar result holds if languages are considered as subsets of the free monoid A^*. Then one should consider monoids and varieties of finite ordered monoids instead of semigroups and varieties of finite semigroups.

We would like to illustrate this correspondence on three non-trivial examples. Other examples can be found for instance in [26] or [29]. It is convenient to describe our examples by identities. Indeed, it was proved by Bloom [11] that a Birkhoff's theorem holds for varieties of ordered semigroups, if one considers identities of the form $u \leq v$, with $u, v \in \Sigma^*$. Similarly, a Reiterman's theorem holds for varieties of finite ordered semigroups [30]. Given a set E of identities of the form $u \leq v$, with $u, v \in \hat{\Sigma}^+$, one denotes by $[\![E]\!]$ the variety of finite ordered semigroups which satisfy all the identities of E. Our three examples are the two varieties of ordered monoids $\mathbf{V}_1 = [\![x \leq 1]\!]$, $\mathbf{V}_2 = [\![x^\omega \leq 1]\!]$ and the variety of ordered semigroups $\mathbf{V}_3 = [\![x^\omega y x^\omega \leq x^\omega]\!]$.

Thus $\mathbf{V}_1$ consists of all finite ordered monoids $(M, \leq)$ in which the identity is the top element. Observe that this implies that M is $\mathcal{J}$-trivial. Indeed, if $x \leq_J y$, then $x = ayb$ for some $a, b \in M$. Now $a \leq 1$, $y \leq y$, $b \leq 1$ and thus $x = ayb \leq y$. Thus, if $x \mathcal{J} y$, then $x \leq y$ and $y \leq x$, whence $x = y$.

Next $\mathbf{V}_2$ consists of all finite ordered monoids $(M, \leq)$ in which $e \leq 1$, for every idempotent $e \in E(M)$. This condition too, imposes severe restrictions on the algebraic structure of M. For instance, the argument above shows that the submonoid of M generated by its idempotents is $\mathcal{J}$-trivial. Finite monoids

satisfying this condition are called *block groups*. This terminology is justified in [25], where several equivalent conditions are also given. For instance:

(1) Every $\mathcal{R}$-class and every $\mathcal{L}$-class contains at most one idempotent,

(2) For every regular $\mathcal{D}$-class D of M, D^0 is a Brandt semigroup

Finally $\mathbf{V}_3$ consists of all finite semigroups $(S, \leq)$ in which $ese \leq e$, for every idempotent e and every element s in S: this is equivalent to saying that the "local" ordered monoids $(eSe, \leq)$, for $e \in E(S)$, are in $\mathbf{V}_1$. In particular, the local monoids are $\mathcal{J}$-trivial, but we shall see later on that S satisfies an even stronger condition.

Our three examples are defined by some rather natural conditions on the order and the corresponding positive varieties of languages are also relatively natural families. For every alphabet A, $A^*\mathcal{V}_1$ is the set of languages which are finite union of languages of the form $A^*a_1A^*a_2 \cdots a_kA^*$ where the a_i's are letters of A. The languages of $A^*\mathcal{V}_2$ are finite union of languages of the form $L_0a_1L_1 \cdots a_kL_k$ where the a_i's are letters and the L_i's are group languages (that is, languages whose syntactic monoid is a finite group). There is an obvious similarity between these two descriptions which is discussed in detail in [29]. The languages of $A^+\mathcal{V}_3$ are finite union of languages of the form $u_0A^*u_1A^* \cdots u_{k-1}A^*u_k$, where $k \geq 0$ and $u_0, \ldots, u_k \in A^*$.

These three results are a particular case of a much more general result established in [29] but they will suffice to illustrate a typical back and forth argument between semigroups and languages to obtain results of pure semigroup theory. The idea is to make use of the natural relations between varieties of finite semigroups and varieties of finite ordered semigroups.

If $\mathbf{V}$ is a variety of finite semigroups, the class of all finite ordered semigroups of the form $(S, \leq)$, where $S \in \mathbf{V}$, is a variety of finite ordered semigroups, denoted $\mathbf{V}_\leq$, and called the *variety of finite ordered semigroups associated with* $\mathbf{V}$. Conversely, given a variety of finite ordered semigroups $\mathbf{W}$, the class of all semigroups S such that $(S, \leq) \in \mathbf{W}$ for some stable order $\leq$ on S is a variety of finite semigroups, called the *variety of finite semigroups associated with* $\mathbf{W}$. Now, for every variety of finite semigroups $\mathbf{V}$, $\mathbf{V}$ is the variety of finite semigroups associated with $\mathbf{V}_\leq$. But if $\mathbf{W}$ is a variety of finite ordered semigroups, and if $\mathbf{V}$ is the variety of finite semigroups associated with $\mathbf{W}$, then $\mathbf{V}_\leq$ is not in general equal to $\mathbf{W}$. In fact, $\mathbf{V}_\leq$ is equal to $\mathbf{W} \vee \breve{\mathbf{W}}$, the join of $\mathbf{W}$ and its dual (that is, the smallest variety of finite ordered semigroups containing $\mathbf{W}$ and its dual).

Now, if $\mathbf{V}$ is a variety of finite ordered monoids (resp. semigroups), one can try to compute the associated variety of finite monoids (resp. semigroups) according to the following plan:

Step 1 Characterize the positive variety of languages corresponding to $\mathbf{V}$.

Step 2 Characterize the positive variety of languages corresponding to $\breve{\mathbf{V}}$.

Step 3 Characterize the positive variety of languages $\mathcal{V}$ corresponding to $\mathbf{V} \vee \breve{\mathbf{V}}$.

Step 4 Characterize the variety of finite monoids corresponding to $\mathcal{V}$.

Step 1 was discussed above for our examples $\mathbf{V}_1$, $\mathbf{V}_2$ or $\mathbf{V}_3$. Step 2 is easy: taking the dual of a variety of ordered semigroups $\mathbf{V}$ corresponds to complementation at the language level. More precisely, if $\mathcal{V}$ (resp. $\breve{\mathcal{V}}$) is the positive variety corresponding to $\mathbf{V}$ (resp. to $\breve{\mathbf{V}}$), then, for each alphabet A, $A^+\breve{\mathcal{V}}$ is the class of all complements in A^+ of the languages of $A^+\mathcal{V}$. Step 3 is also simple: $\mathcal{W} = \mathcal{V} \vee \breve{\mathcal{V}}$ is not only a positive variety, but also a variety of languages, and for each alphabet A, $A^+\mathcal{W}$ is the boolean algebra generated by $A^+\mathcal{V}$.

For the last step, we consider our three examples separately. $A^+\mathcal{W}_1$ is the set of languages which are boolean combinations of languages of the form $A^*a_1A^*a_2\cdots a_kA^*$ where the a_i's are letters of A. It is a well-known result of Simon [33] that these languages correspond to the variety **J** of finite $\mathcal{J}$-trivial monoids, which can be defined by the identities

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

The languages of $A^*\mathcal{W}_2$ are boolean combinations of languages of the form $L_0a_1L_1\cdots a_kL_k$ where the a_i's are letters and the L_i's are group languages. These languages correspond to the variety **BG** of block groups [19, 25, 28], defined by the identities

$$(x^\omega y^\omega)^\omega x^\omega = (x^\omega y^\omega)^\omega = y^\omega(x^\omega y^\omega)^\omega$$

Finally, the languages of $A^+\mathcal{W}_3$ are boolean combination of languages of the form $u_0A^*u_1A^*\cdots u_{k-1}A^*u_k$, where $k \geq 0$ and $u_0, \ldots, u_k \in A^*$. These languages are known as languages of "dot-depth one" in the literature [13, 22, 29] and the corresponding variety of finite semigroups is denoted $\mathbf{B_1}$ (the "B" refers to Brzozowski who introduced the dot-depth hierarchy, and the 1 to the first level of this hierarchy). A characterization of $\mathbf{B_1}$ was obtained by Knast [16, 17] and relies on the notion of *graph of a finite semigroup*. Given a semigroup S, form a graph $G(S)$ as follows: the vertices are the idempotents of S and the edges from e to f are the elements of the form esf. Then a finite semigroup is in $\mathbf{B_1}$ if its graph satisfies the following condition : if e and f are two vertices, p and r edges from e to f, and q and s edges from f to e, then $(pq)^\omega ps(rs)^\omega = (pq)^\omega(rs)^\omega$. Thus $\mathbf{B_1}$ is defined by the identities

$$(x^\omega py^\omega qx^\omega)^\omega x^\omega py^\omega sx^\omega(x^\omega ry^\omega sx^\omega)^\omega = (x^\omega py^\omega qx^\omega)^\omega(x^\omega ry^\omega sx^\omega)^\omega$$

Our computation can be summarized into the following three equations

$$\begin{aligned}
[\![x \leq 1]\!] \vee [\![1 \leq x]\!] &= [\![(xy)^\omega x = (xy)^\omega = y(xy)^\omega]\!] \\
[\![x^\omega \leq 1]\!] \vee [\![1 \leq x^\omega]\!] &= [\![(x^\omega y^\omega)^\omega x^\omega = (x^\omega y^\omega)^\omega = y^\omega(x^\omega y^\omega)^\omega]\!] \\
[\![x^\omega yx^\omega \leq x^\omega]\!] \vee [\![x^\omega \leq x^\omega yx^\omega]\!] &= \mathbf{B_1}
\end{aligned}$$

The difficult part is to prove that the right hand side of these equalities is contained in the left hand side. The opposite inclusion follows from simple manipulations of the identities. For instance, here is a proof that the identity $x \leq 1$ implies $x^\omega = x^{\omega+1}$. The identity $x \leq 1$ clearly implies $x^{\omega+1} \leq x^\omega$. It follows that, for all $n > 0$, $x^{\omega+n!} \leq x^{\omega+1} \leq x^\omega$ and by continuity, $x^\omega = \lim_{n\to\infty} x^{\omega+n!} \leq x^{\omega+1} \leq x^\omega$, whence $x^\omega = x^{\omega+1}$.

This former formulation is only appealing to people familiar with identities, but there is a more attractive version of these results, also given in [29].

(1) Every finite $\mathcal{J}$-trivial monoid is a quotient of an ordered monoid satisfying the identity $x \leq 1$.
(2) Every block group is a quotient of an ordered monoid satisfying the identity $x^\omega \leq 1$.
(3) Every semigroup of $\mathbf{B_1}$ is a quotient of an ordered semigroup satisfying the identity $x^\omega yx^\omega \leq x^\omega$.

The first of these results was first proved by Straubing and Thérien by a remarkable induction on the size of the monoid [35]. It would be very interesting to have

a similar proof for the two other results. It is easy, however, to prove directly the third result for powergroups, which are particular cases of block groups. Given a group G, denote by $\mathcal{P}'(G)$ the monoid of all non-empty subsets of G under multiplication. Then $\mathcal{P}'(G)$ is naturally ordered by the relation $\leq$ defined by

$$X \leq Y \text{ if and only if } Y \subseteq X$$

The idempotents of $\mathcal{P}'(G)$ are the subgroups of G and they all contain the trivial subgroup $\{1\}$, which is the identity of $\mathcal{P}'(G)$. Therefore $X^\omega \leq \{1\}$ for every $X \in \mathcal{P}'(G)$ and thus $(\mathcal{P}'(G), \leq)$ satisfies the identity $x^\omega \leq 1$.

4. Decomposition theorems

Let us briefly review some well-known facts of group theory. Kernels and group extensions are two central notions of this theory. Given a group morphism $\varphi : G \to H$, the *kernel* of φ is the subgroup $1\varphi^{-1}$ of G. Given three groups G, H and K, G is said to be an *extension* of K by H (or an *expansion*(*) of H by K) if there exists a group morphism from G onto H whose kernel is K. Group expansions are intimately related to wreath products and to semidirect products. For instance, if G is an expansion of H by K, then G divides the wreath product $K \circ H$. Furthermore, if K and H are finite groups whose order are relatively prime, then G is a semidirect product of K by H: this is the Schur-Zassenhaus lemma.

It is tempting to develop a similar theory for semigroups, but it is a non trivial task. We first consider a simple case, which is half-way between semigroups and groups. Let N be a monoid and let G be group. It is natural to say that a monoid M is an *expansion of G by N* if there exists a surjective morphism $\varphi : M \to G$ such that $1\varphi^{-1} = N$. For instance, one may consider the monoids which are expansions of a group by a band (resp. a semilattice). In this case $1\varphi^{-1}$ is a submonoid of $E(M)$, the set of idempotents of M, and since φ maps any idempotent onto 1, the equality $1\varphi^{-1} = E(M)$ actually holds. In particular $E(M)$ is a submonoid of M. Furthermore, if e and es are idempotent, then $e\varphi = (es)\varphi = 1$, whence $s\varphi = 1$ and $s \in E(M)$. Thus $E(M)$ is an *unitary* submonoid of M. Furthermore, M is *E-dense*: for every $s \in M$, there exists an element $t \in M$ such that $st, ts \in E(M)$ (choose for t an arbitrary element in $g\varphi^{-1}$, where $g = (s\varphi)^{-1}$.) Thus M is *E-unitary dense*. This is actually a characterization: a monoid is an expansion of a group by an idempotent monoid if and only if it is E-unitary dense. This result was proved for monoids with commuting idempotents (or *E-commutative* monoids) in [20] and generalized in [3]. Note that we didn't make any assumption on the regularity of M.

Although their natural definition as expansions makes these monoids worth to be studied, they were originally introduced for another purpose. Define a *covering* to be a surjective morphism which is one-to-one on idempotents. It was conjectured in [20] that every E-dense E-commutative monoid is covered by an E-unitary dense monoid. A version for finite monoids was also proposed: every finite E-commutative monoid is covered by a finite E-unitary dense monoid. Both conjectures received a positive answer : the latter was proved by Ash [7] and the first one by Fountain [14]. Birget, Margolis and Rhodes [10], Ash [8], and Almeida, Pin, Weil [3] gave further extensions. It is convenient to

(*) Although it might be a little bit confusing to have a double terminology, I am taking the risk of introducing the word "expansion", which is more adapted to the generalizations introduced below.

summarize these results into a single statement with some optional properties written between brackets:

First structure theorem Every [finite] [E-commutative] [regular] E-dense monoid is covered by a [finite] [E-commutative] [regular] E-unitary dense monoid.

By an E-dense monoid, we mean of course a monoid in which $E(M)$ is a dense submonoid. Note that the density condition is rather weak: if $E(M)$ is a submonoid of M, then M^0 is always E-dense.

Let us return to the structure of an E-unitary dense monoid M. We already mentioned that M is expansion of a group by a band. Clearly, the band has to be $E(M)$, but what about the group ? The group is the *fundamental group* of M, denoted $\pi_1(M)$, and defined as the quotient of the free group $F(M)$ with basis M by the relations $(s)(t) = (st)$, for every $s, t \in M$. One can show that $\pi_1(M)$ is the maximal quotient group of M. Then if $\pi: M \to \pi_1(M)$ denotes the natural morphism, π is onto and $1\pi^{-1} = E(M)$. Therefore M is an expansion of $\pi_1(M)$ by $E(M)$.

Our last characterization motivates the introduction of categories as a generalization of monoids. We follow the presentation of [3]. Formally, a category C is given by

(a) a set $Ob(C)$ of objects,

(b) for each pair (u, v) of objects, a set $C(u, v)$ of arrows,

(c) for each triple (u, v, w) of objects, a mapping from $C(u, v) \times C(v, w)$ into $C(u, w)$ which associates to each $p \in C(u, v)$ and $q \in C(v, w)$ the composition $p + q \in C(u, w)$.

(d) for each object u, an arrow 0_u such that, for each pair (u, v) of objects, for each $p \in C(u, v)$ and $q \in C(v, u)$, $0_u + p = p$ and $q + 0_u = q$.

The additive notation is used for convenience, but it does not imply commutativity. Composition is assumed to be associative (when defined).

For each object u, $C(u, u)$ is a monoid, called the *local monoid* of u. In particular a monoid can be considered as a category with exactly one object. A category is said to be *locally idempotent* (resp. *locally commutative*, etc.) if all its local monoids are idempotent (resp. commutative, etc.). A category C is *regular* if, for each arrow $p \in C(u, v)$, there exists an arrow $q \in C(v, u)$ such that $p + q + p = p$, and it is *inverse* if, for each arrow $p \in C(u, v)$, there exists a unique arrow $\bar{p} \in C(v, u)$ such that $p + \bar{p} + p = p$ and $\bar{p} + p + \bar{p} = \bar{p}$. It is connected if $C(u, v) \neq \emptyset$ for each pair (u, v) of objects of C. The $\mathcal{J}$ partial order is defined as in a semigroup : given two arrows p and q, $p \leq_{\mathcal{J}} q$ if and only if $p = r + q + s$ for some arrows r and s. The other Green relations and the corresponding equivalence classes are defined analogously.

A *morphism* $\varphi: C \to D$ between two categories C and D is given by a map $\varphi: Ob(C) \to Ob(D)$ and, for every $u, v \in Ob(C)$, a map, also denoted φ from $C(u, v)$ into $D(u\varphi, v\varphi)$ such that, for each pair (p, q) of consecutive arrows, $p\varphi + q\varphi = (p + q)\varphi$ and for each object u, $0_u\varphi = 0_{u\varphi}$. An automorphism φ of a category C is defined as usual. An action of a group G on C is given by a group morphism from G into the group of automorphisms of C. In this case we write gu (resp. gp) the result of the action of g on the object u (resp. the arrow p). Note the following identities:

(1) $g(p + q) = gp + gq$ for all $g \in G$, $p \in C(u, v)$ and $q \in C(v, w)$.

(2) $(gh)p = g(hp)$ for all $g, h \in G$ and $p \in C(u, v)$.

Whenever a group G acts on a category C, a quotient category C/G is defined, with object set $Ob(C)/G$, that is, the set of disjoint subsets of $Ob(C)$ of the

form Gu ($u \in Ob(C)$), and with arrow sets

$$C/G(Gu, Gv) = \{Gp \mid p \in C(u', v'),\ u' \in Gu,\ v' \in Gv\}$$

Composition of consecutive arrows Gp and Gq (that is, $p \in C(u, v)$ and $q \in C(gv, w)$ for some objects u, v and w and some element g of G) is given by

$$Gp + Gq = G(p + g^{-1}q)$$

If a group G acts transitively without fixpoints on a category C, then the category C/G is a monoid and for each object u of C,

$$C_u = \{(p, g) \mid g \in G,\ p \in C(u, gu)\}$$

is a monoid isomorphic to C/G for the multiplication defined by $(p, g)(q, h) = (p + gq, gh)$.

Let us come back to expansions of groups by monoids. Let G be group and let $\varphi : M \to G$ be a surjective morphism. The *derived category* of φ is the category C such that $Ob(C) = G$ and, for $u, v \in G$,

$$C(u, v) = \{(u, s, v) \in G \times M \times G \mid u(s\varphi) = v\}.$$

Composition is given by $(u, s, v)+(v, t, w) = (u, st, w)$. Then G acts transitively without fixpoints on C and M is isomorphic (as a monoid) to C/G. This fact is the key of the second main result of this section, also stated with some optional properties. The first version was proved in [20] and extensions were given in [10], [3] and [36].

Second structure theorem A [finite] [E-commutative] [regular] E-monoid is E-unitary dense if and only if it is isomorphic to C/G where G is a group acting transitively without fixpoints on some [finite] [locally commutative] [regular] connected, locally idempotent category C.

The two structure theorems give a satisfying description of E-dense monoids. They cover important particular cases, especially in the regular case, since a regular E-commutative monoid is nothing else than an inverse monoid and a regular E-dense monoid is an orthodox monoid. One can also derive McAlister's P-theorem. Indeed, if G is a group acting transitively without fixpoints on a connected, locally commutative, inverse category C, then the inverse monoid C/G is isomorphic to a P-semigroup $P(G, F, E)$, where F is the partially ordered set of $\mathcal{J}$-classes of C and E is a subsemilattice of F isomorphic to $E(C/G)$. Actually, if u is an object of C, one can take for E the set $\{J \in F \mid J \cap C(u, u) \neq \emptyset\}$. Then E is a semilattice isomorphic to $C(u, u)$ (and to $E(C/G)$) and an order ideal of F. See [20] for more details.

For finite monoids, one can generalize further on the notion of expansion as follows. Let $\mathbf{V}$ be a variety of finite semigroups and let M and N be monoids. A surjective morphism $\varphi : M \to N$ is a $\mathbf{V}$-morphism if, for every $e \in E(M)$, $e\varphi^{-1} \in \mathbf{V}$. A monoid M is a $\mathbf{V}$-*expansion of* N if there exists a surjective $\mathbf{V}$-morphism from M onto N. Given a variety of finite monoids $\mathbf{W}$, the variety of finite monoids generated by all $\mathbf{V}$-expansions of a monoid of $\mathbf{W}$ is called the *Mal'cev product* of $\mathbf{V}$ and $\mathbf{W}$ and is denoted $\mathbf{V}$ Ⓜ $\mathbf{W}$.

Given $\mathbf{V}$ and $\mathbf{W}$, the computation of $\mathbf{V}$ Ⓜ $\mathbf{W}$ can sometimes be very difficult, but some recent results put some new light on this problem. The first breakthrough was Ash's solution of Rhodes "Type II" conjecture [8]. A detailed

account of Rhodes conjecture and its far reaching consequences can be found in [15], so we just mention the following consequence : if $\mathbf{V}$ is a decidable variety of finite monoids (that is, if there is an algorithm to decide membership in $\mathbf{V}$ for a given finite monoid), then $\mathbf{V}$ Ⓜ $\mathbf{G}$ is also decidable. This result has a number of interesting corollaries (see [15]). For instance:

(1) If $\mathbf{V}$ is the variety of finite semilattices, then $\mathbf{V}$ Ⓜ $\mathbf{G}$ is the variety of finite monoids with commuting idempotents and is also the variety of finite monoids generated by all finite inverse monoids.
(2) If $\mathbf{V}$ is the variety of finite idempotent monoids (bands), then $\mathbf{V}$ Ⓜ $\mathbf{G}$ is the variety of finite monoids generated by all finite orthodox monoids.
(3) If $\mathbf{V}$ is the variety of finite $\mathcal{J}$-trivial monoids, then $\mathbf{V}$ Ⓜ $\mathbf{G}$ is the variety of block groups.

P. Weil and the author recently introduced another powerful technique for computing Mal'cev products [28]. The idea is to find the identities defining $\mathbf{V}$ Ⓜ $\mathbf{W}$, given the identities defining $\mathbf{V}$ and $\mathbf{W}$. Although the general result is a little bit too technical to be stated here in extenso, some consequences are worth mentioning. It is shown in particular that if $\mathbf{V}$ is a decidable variety of finite semigroups, then $\mathbf{V}$ Ⓜ $\mathbf{W}$ is also decidable in the following cases

(1) $\mathbf{W}$ is the variety of finite nilpotent semigroups
(2) $\mathbf{W}$ is the variety of finite semilattices
(3) $\mathbf{W}$ is the variety of finite $\mathcal{J}$-trivial monoids

It is interesting to note that these results rely heavily on the structure of profinite semigroups. In particular, a crucial step of the proof is a compactness argument.

As it is the case for groups, the Mal'cev product is sometimes related to the wreath product or to the semidirect product. For instance, an E-unitary inverse monoid is isomorphic to a submonoid of a semidirect product of a semilattice by a group. More precisely, if φ is a surjective morphism from an inverse monoid M onto a group G such that $1\varphi^{-1} = E(M)$, and if C is the derived category of C, then M is isomorphic to a subsemigroup of a semidirect product $S * G$, where S is the semilattice of ideals of C under intersection [20]. It is not clear whether this type of result can be generalized to the non regular case or to regular expansions of groups by bands of a given variety of bands, although partial results were obtained by P. Jones and M. Szendrei.

Given a variety of finite semigroups $\mathbf{V}$ and a variety of finite monoids $\mathbf{W}$, one denotes by $\mathbf{V} * \mathbf{W}$ the variety of finite monoids generated by all wreath products of the form $M \circ N$ with $M \in \mathbf{V}$ and $N \in \mathbf{W}$. Almeida and Weil give a description of the free profinite semigroup of $\mathbf{V} * \mathbf{W}$ [5] and provide identities defining $\mathbf{V} * \mathbf{W}$, given identities defining $\mathbf{V}$ and $\mathbf{W}$ [6]. For instance, the equality $\mathbf{V}$ Ⓜ $\mathbf{G} = \mathbf{V} * \mathbf{G}$ holds if $\mathbf{V}$ is one of the following variety of finite monoids:

(1) aperiodic monoids.
(2) semilattices
(3) $\mathcal{J}$-trivial monoids

Another important tool for studying the relations between expansions and wreath product decompositions is again the derived category. Indeed, as it was shown by Tilson [37], the use of the derived category is not limited to expansions of groups by monoids. Consider now the more general situation of a surjective morphism $\varphi : M \to N$, where M and N are monoids. One can mimic the construction given above to define a category C such that $Ob(C) = N$ and, for all $u, v \in N$,

$$C(u, v) = \{(u, s, v) \in N \times M \times N \mid u(s\varphi) = v\}.$$

Again, composition is given by $(u, s, v) + (v, t, w) = (u, st, w)$. Now the *derived*

category $D(\varphi)$ of φ is the quotient of C by the congruence $\sim$ defined by

$$(u,s,v) \sim (u,t,v) \text{ if and only if } ms = mt \text{ for all } m \in u\varphi^{-1}$$

Thus the derived category identifies elements with the same action on each fiber $u\varphi^{-1}$.

A few more definitions on categories are in order to state Tilson's theorem precisely. A category C is a *subcategory* of a category D if there exists a morphism $\varphi : C \to D$ which is injective on arrows (that is, for each pair of objects (u,v), the map from $C(u,v)$ into $D(u\varphi, v\varphi)$ is injective). A category C is a *quotient* of a category D if there exists a morphism $D \to C$ which is bijective on objects and surjective on arrows. Finally C *divides* D if C is a quotient of a subcategory of D. Tilson's derived category theorem [37] relates expansions to wreath products as follows

Derived Category Theorem Let M and N be monoids and let $\varphi : M \to N$ be a surjective morphism. Then, for each monoid K such that $D(\varphi)$ divides K, M divides $K \circ N$.

Tilson's theorem can actually be stated in the more general setting of relational morphisms.

5. Conclusion

The three tools presented in this paper (profinite semigroups, ordered semigroups and categories) share a common feature. All three can be viewed as extensions of the purely algebraic structure of semigroup, but none of them was introduced for the purpose of gratuitous generalization. Rather, they were simply *needed* to solve existing problems: this is actually a rather common phenomenon in mathematics and it would not be surprising if even more sophisticated constructions were required in the future.

Acknowledgements

I would like to thank Pascal Weil for a careful reading of a first version of this paper.

References

[1] Almeida, J., *Semigrupos Finitos e Álgebra Universal*, Publicações do Instituto de Matemática e Estatística da Universidade de São Paulo, 1992.

[2] Almeida, J., Equations for pseudovarieties, J.-E. Pin ed. *Formal properties of finite automata and applications*, Lecture Notes in Computer Science 386, 1989, 148–164.

[3] Almeida, J., J.-E. Pin, and P. Weil, Semigroups whose idempotents form a subsemigroup *Math. Proc. Camb. Phil. Soc.* **111**, (1992), 241–253.

[4] Almeida, J., and P. Weil, Relatively free profinite monoids: an introduction and examples, in NATO Advanced Study Institute *Semigroups, Formal Languages and Groups*, J. Fountain et V. Gould (ed.), Kluwer Academic Publishers, to appear.

[5] Almeida, J., and P. Weil, Free profinite semigroups over semidirect products, *Izvestija vuzov. Matematika*, to appear.

[6] Almeida, J., and P. Weil, Profinite categories and semidirect products, LITP Report 94–35, submitted.

[7] Ash, C. J., Finite semigroups with commuting idempotents, *J. Austral. Math. Soc. (Series A)* **43**, (1987) 81–90.

[8] Ash, C. J., Inevitable Graphs: A proof of the type II conjecture and some related decision procedures, *Int. Jour. Alg. and Comp.* **1** (1991) 127–146.

[9] Birget, J.-C., S. Margolis and J. Rhodes. Finite semigroups whose idempotents commute or form a subsemigroup, in S. Goberstein and P. Higgins (eds.) *Semigroups and their applications*, Reidel (Dordrecht, Holland), (1987) 25–35.

[10] Birget, J.-C., S.W. Margolis and J. Rhodes, Finite semigroups whose idempotents form a subsemigroup, *Bull. Austral. Math. Soc.* **41** (1990) 161–184.

[11] Bloom, S. L., Varieties of ordered algebras, *J. Comput. System Sci.* **13**, (1976) 200–212.

[12] Eilenberg, S., *Automata, languages and machines*, Vol. A, Academic Press, New York, 1974.

[13] Eilenberg, S., *Automata, languages and machines*, Vol. B, Academic Press, New York, 1976.

[14] Fountain., J., E-unitary dense covers of E-dense monoids, *Bull. London Math. Soc.* **22**, (1990), 353–358.

[15] Henckell, K., S. W. Margolis, J.-E. Pin and J. Rhodes, Ash's Type II Theorem, Profinite Topology and Malcev Products, *International Journal of Algebra and Computation* **1** (1991) 411–436.

[16] Knast, R., A semigroup characterization of dot-depth one languages, *RAIRO Inform. Théor.* **17**, (1983) 321–330.

[17] Knast, R., Some theorems on graph congruences, *RAIRO Inform. Théor.* **17**, (1983) 331–342.

[18] Lallement, G., *Semigroups and combinatorial applications*, Wiley, New York, 1979.

[19] Margolis, S.W., and J.E. Pin, Product of group languages, *FCT Conference, Lecture Notes in Computer Science* **199** (1985) 285–299..

[20] Margolis, S.W., and J.E. Pin, Inverse semigroups and extensions of groups by semilattices, *Journal of Algebra* **110** (1987) 277–297.

[21] Margolis, S.W., and J.E. Pin, Inverse semigroups and varieties of finite semigroups, *Journal of Algebra* **110** (1987) 306–323.

[22] Pin, J.-E., *Variétés de langages formels*, Masson, Paris, 1984; English translation: *Varieties of formal languages*, Plenum, New-York, 1986.

[23] Pin, J.-E., Topologies for the free monoid, *Journal of Algebra* **137** (1991) 297–337.

[24] Pin, J.-E., Finite semigroups and recognizable languages : an introduction, in NATO Advanced Study Institute *Semigroups, Formal Languages and Groups*, J. Fountain et V. Gould (ed.), Kluwer Academic Publishers, to appear.

[25] Pin, J.-E., **BG = PG**, a success story, in NATO Advanced Study Institute *Semigroups, Formal Languages and Groups*, J. Fountain et V. Gould (ed.), Kluwer Academic Publishers, to appear.

[26] Pin, J.-E., A variety theorem without complementation, *Izvestija vuzov. Matematika*, to appear.

[27] Pin, J.-E., Polynomial closure of group languages and open sets of the Hall topology, *ICALP 1994, Lecture Notes in Computer Science* **820**, (1994) 424–435.

[28] Pin, J.-E., and P. Weil, Free profinite semigroups, Mal'cev products and identities, to appear.

[29] Pin, J.-E., and P. Weil, Polynomial closure and unambiguous product, LITP report 94-60, Paris, (1994), submitted.

[30] Pin, J.-E., and P. Weil, A Reiterman theorem for pseudovarieties of finite first-order structures, LITP Report 94–38, Paris, (1994), submitted.

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

[32] Rhodes, J., and B. Tilson, The kernel of monoid morphisms, *Journal of Pure and Applied Algebra* **62**, (1989) 227–268.

[33] Simon, I., Piecewise testable events, *Proc. 2nd GI Conf., Lecture Notes in Computer Science* **33**, (1975) 214–222..

[34] Stone, M., The representation of boolean algebras, *Bulletin of the AMS* **44**, 807–816, reviewed in Zentralblatt für Mathematik **20**, 342.

[35] Straubing, H., and D. Thérien, Partially ordered finite monoids and a theorem of I. Simon, *J. of Algebra* **119**, (1985) 393–399.

[36] Szendrei, M., A generalization of McAlister's P-theorem for E-unitary regular semigroups, *Acta Sci. Math. (Szeged)* **51**, (1987), 229–249.

[37] Tilson, B., Categories as algebras: an essential ingredient in the theory of finite monoids, *J. Pure Applied Algebra* **48** (1987), 83–198..

[38] Weil, P., Implicit operations on pseudovarieties: an introduction, in J. Rhodes (ed.) *Monoids and semigroups with applications* World Scientific, Singapore, (1991), 89–104.

LITP/IBP, Université Paris VI et CNRS
Tour 55-65, 4 Place Jussieu
75252 Paris Cedex 05, FRANCE
E-mail: pin@litp.ibp.fr