

## ATOMS AND PARTIAL ORDERS OF INFINITE LANGUAGES <sup>\*,\*\*</sup>

WERNER KUICH<sup>1</sup> AND N.W. SAUER<sup>2</sup>

**Abstract.** We determine minimal elements, *i.e.*, atoms, in certain partial orders of factor closed languages under  $\subseteq$ . This is in analogy to structural Ramsey theory which determines minimal structures in partial orders under embedding.

**Mathematics Subject Classification.** 68R15, 05C55.

### 1. INTRODUCTION AND PRELIMINARIES

Let  $\Sigma$  be a finite alphabet. The word  $w \in \Sigma^*$  is a *factor* of the word  $v \in \Sigma^*$  if there are words  $u, r \in \Sigma^*$  so that  $v = uwr$ ; symbolically  $w|v$  with negation  $w \not|v$ . We denote by  $|w|$  the *length* of the word  $w \in \Sigma^*$  and by  $\mathbb{N}$  the set  $\{0, 1, 2, 3, 4, \dots\}$ .

A *language*  $L$  is a subset of  $\Sigma^*$  and  $\text{Fact}(L)$  is the set of factors of words in  $L$ . That is  $\text{Fact}(L) := \{w \mid \exists v \in L(w|v)\}$ . The language  $L$  is *factor closed* if  $L = \text{Fact}(L)$ . Also  $\text{Del}(L, w) := \{v \in L \mid w \not|v\}$ . Hence  $\text{Del}(L, w)$  is the set of all words of  $L$  that do not contain  $w$  as a factor. If  $L$  is factor closed then  $\text{Del}(L, w)$  is factor closed.

Let  $L \subseteq \Sigma^*$  be a language. The word  $w \in L$  is *recurrent* in  $L$  if there are infinitely many words  $v \in L$  with  $w|v$  otherwise  $w$  is *rare* in  $L$ . The word  $w$  is *syndetic* in  $L$  if there is  $k \in \mathbb{N}$  so that  $w|v$  for every word  $v \in L$  with  $|v| \geq k$ . The language  $L$  is *recurrent* if every word in  $L$  is recurrent in  $L$ . For example  $\Sigma^*$  is recurrent. The language  $L$  is *uniformly recurrent* if every word in  $L$  is syndetic. Hence, the language  $L$  is recurrent if every word of  $L$  is factor of infinitely many

---

*Keywords and phrases:* Combinatorics of words, structural Ramsey theory.

\* *Partially supported by the "Stiftung Aktion Österreich-Ungarn".*

\*\* *Supported by NSERC of Canada Grant # 691325.*

<sup>1</sup> Technische Universität Wien, Wiedner Hauptstraße 8-10, 1040 Wien, Austria;  
e-mail: [kuich@tuwien.ac.at](mailto:kuich@tuwien.ac.at)

<sup>2</sup> University of Calgary, Department of Mathematics and Statistics, 2500 University Dr. NW,  
Calgary Alberta Canada T2N1N4; e-mail: [nsauer@math.ucalgary.ca](mailto:nsauer@math.ucalgary.ca)

© EDP Sciences 2001

words of  $L$ . The language  $L$  is uniformly recurrent if every word of  $L$  is factor of almost all words of  $L$ . Note that a language which is uniformly recurrent is also recurrent. But  $\Sigma^*$  is recurrent and not uniformly recurrent. Note also that the language  $L$  is recurrent if and only if  $\text{Fact}(L)$  is recurrent and  $L$  is uniformly recurrent if and only if  $\text{Fact}(L)$  is uniformly recurrent.

Let  $L$  and  $K$  be languages. The function  $\lambda : L \mapsto K$  is *admissible* if  $w$  is a factor of  $\lambda(w)$  for all  $w \in L$ . Let  $\mathbf{F}_\Sigma$  be the category with the set of infinite languages over the alphabet  $\Sigma$  as set of objects and morphisms the admissible functions. Let  $\mathbf{I}_\Sigma$  be the category with the set of infinite languages over the alphabet  $\Sigma$  as set of objects and morphisms the admissible injections (admissible one to one functions.)

Let  $\Gamma$  be an alphabet. An *increasing list of words in  $\Gamma^*$*  is a sequence of the form  $R := (r_i; i \in I \subseteq \mathbb{N})$  with  $|r_i| \leq |r_j|$  for all  $i \leq j$  in  $I$ ; we require also that if  $i \leq k \leq j$  and  $i, k, j \in I$  and  $r_i = r_j$  then  $r_i = r_k = r_j$ . The list  $R := (r_i; i \in I \subseteq \mathbb{N})$  is *recurrent* if the language  $\{r_i \mid i \in I\}$  is recurrent and it is *uniformly recurrent* if the language  $\{r_i \mid i \in \mathbb{N}\}$  is uniformly recurrent. Also  $\text{Fact}(R) := \text{Fact}(\{r_i \mid i \in I\})$ .

We are really only interested in the case that the index set  $I$  is infinite because we are only interested in infinite languages. Of course then we might as well replace  $I$  by  $\mathbb{N}$ . Unfortunately the case that the index set  $I$  is finite will be needed in one of the proofs.

Let  $R = (r_i; i \in I \subseteq \mathbb{N})$  be an increasing list of words. Let  $\Gamma$  and  $\Sigma$  be finite alphabets and  $f$  a function of  $\Sigma$  to  $\Gamma$ . We extend  $f$  to words in  $\Sigma^*$  by stipulating that  $f$  is a monoid morphism, i.e.  $f(uv) = f(u)f(v)$  for all words  $u, v \in \Sigma^*$  and  $f(\varepsilon) = \varepsilon$ .

We denote by  $\mathbf{B}(R, f)$  the category with objects all languages  $L \subseteq \Sigma^*$  for which there is a bijection  $\alpha$  of  $L$  to  $I$  so that

$$f(w) = r_{\alpha(w)} \quad \text{for all } w \in L.$$

It follows that if for example  $R = (r_i; i \in \mathbb{N})$  then  $L \in \mathbf{B}(R, f)$  if and only if there is an enumeration

$$w_0, w_1, w_2, w_3 \dots, w_{i-1}, w_i, w_{i+1} \dots$$

of  $L$  so that  $f(w_i) = r_i$  for all  $i \in \mathbb{N}$ . That is

$$R = f(w_0), f(w_1), f(w_2), f(w_3), \dots, f(w_{i-1}), f(w_i), f(w_{i+1}), \dots$$

The morphisms of  $\mathbf{B}(R, f)$  are the admissible injections. It follows that  $\mathbf{B}(R, f)$  is a subcategory of  $\mathbf{I}_\Sigma$  if the index set  $I$  is infinite.

If  $S \subseteq \Gamma^*$  and  $L \subseteq \Sigma^*$  and the monoid morphism  $f$  of  $L$  to  $S$  is a bijection and if  $S$  is ordered into the list of words  $R$  then  $L \in \mathbf{B}(R, f)$ . Conversely, if under these assumptions  $L \in \mathbf{B}(R, f)$  then  $f[L] := \{f(w) \mid w \in L\} = S$ . The purpose of using the device ‘‘increasing list of words’’ is to also consider the case in which  $f$  is not necessarily one to one and then to control exactly the size of the sets of the form  $f^{-1}(v)$ .

Let  $\mathbf{P} = (P; \leq)$  be a partial order. The element  $M \in P$  is *minimal* in  $\mathbf{P}$  if  $X \leq M$  for  $X \in P$  implies  $X = M$ . Such a minimal element of  $\mathbf{P}$  is called *atom* of  $\mathbf{P}$ . The set  $S \subseteq P$  is a *chain* of  $\mathbf{P}$  if  $X \leq Y$  or  $Y \leq X$  for any two elements  $X, Y \in S$ . The partial order  $\mathbf{P}$  is *atomic* if for any element  $X \in P$  there is a minimal element  $M \in P$  with  $M \leq X$ . The partial order  $\mathbf{P}$  is *Zorns* if for every nonempty chain  $S$  of  $\mathbf{P}$  there is an element  $B \in P$  with  $B \leq X$  for every element  $X \in S$ . Such an element  $B$  is a *lower bound* of  $S$ . Zorns lemma says that if the partial order  $\mathbf{P}$  is Zorns then it is atomic.

It is easy to find atomic partial orders which are not Zorns. Let  $P := \{\{n\} \mid n \in \mathbb{N}\} \cup \{\{i \mid i \geq n\} \mid n \in \mathbb{N}\}$  and  $\mathbf{P} := (P; \subseteq)$ . Then each element of the form  $\{n\}$  is minimal in  $\mathbf{P}$  and  $\{\{i \mid i \geq n\} \mid n \in \mathbb{N}\}$  is a chain of  $\mathbf{P}$  which does not have a lower bound in  $\mathbf{P}$ . That is, being Zorns is a stronger condition than being atomic.

Every category  $\mathbf{C}$  can be viewed as a partial quasiorder. The object  $L$  is less than or equal to the object  $K$  if there is a morphism of  $L$  to  $K$ . The objects  $L$  and  $K$  are *equivalent* if there is a morphism of  $L$  to  $K$  and a morphism of  $K$  to  $L$ . Factoring out this equivalence relation one obtains a partial order, the *skeleton* of the category. Lifting the notions above to categories we have:

The object  $M$  is *minimal* in a category  $\mathbf{C}$  if whenever there is a morphism of an object  $L$  in  $\mathbf{C}$  to  $M$  then there is a morphism of  $M$  to  $L$ . The category  $\mathbf{C}$  is *atomic* if for every object  $L$  of  $\mathbf{C}$  there is a minimal object  $M$  of  $\mathbf{C}$  and a morphism of  $M$  to  $L$ . The set  $S$  of objects of  $\mathbf{C}$  is a *chain* if for every two objects  $X, Y \in S$  there is a morphism of  $X$  to  $Y$  or a morphism of  $Y$  to  $X$ . The category  $\mathbf{C}$  is *Zorns* if for every nonempty chain  $S$  of objects in  $\mathbf{C}$  there is an object  $Z$  of  $\mathbf{C}$  which has a morphism to every object in  $S$ ; the object  $Z$  is a *lower bound* of  $S$ .

Note that  $M$  is minimal in the category  $\mathbf{C}$  if and only if the equivalence class of  $\mathbf{C}$  containing  $M$  is minimal in the the partial order which is the skeleton of the category  $\mathbf{C}$ . The category  $\mathbf{C}$  is atomic if and only if the skeleton of  $\mathbf{C}$  is atomic. The category  $\mathbf{C}$  is *Zorns* if [?] and only if [?] the skeleton of  $\mathbf{C}$  is Zorns. It follows that if  $\mathbf{C}$  is Zorns then it is atomic.

It is known that the category  $\mathbf{F}_\Sigma$  is atomic [2] and that a language is minimal in the category  $\mathbf{F}_\Sigma$  if and only if it is uniformly recurrent ([1]; Th. 2.3.1). The fact that the category  $\mathbf{F}_\Sigma$  is atomic and that a language is minimal in  $\mathbf{F}_\Sigma$  if and only if it is uniformly recurrent follows readily from Theorem 2.3.1 of [1] about infinite words. [1] contains an elementary proof. The original argument [2] used methods of symbolic dynamics. Because we use this result in the case of languages and for completeness sake we include a short proof; Theorem 3.2. We actually prove that  $\mathbf{F}_\Sigma$  is Zorns using an even shorter argument than the one in [1].

The category  $\mathbf{I}_\Sigma$  is Zorns and the set of minimal elements of  $\mathbf{I}_\Sigma$  is equal to the set of minimal elements of the category  $\mathbf{F}_\Sigma$ ; Theorem 3.3. We will show, our main results Theorem 5.1 and Theorem 5.2, that categories of the form  $\mathbf{B}(R, f)$  are Zorns and hence atomic and that if  $R$  is uniformly recurrent then the minimal languages of  $\mathbf{B}(R, f)$  are also uniformly recurrent, hence minimal elements of  $\mathbf{F}_\Sigma$ .

The definition of admissible function implies that there is an admissible function of  $L$  to  $K$  if and only if  $L \subseteq \text{Fact}(K)$  if and only if  $\text{Fact}(L) \subseteq \text{Fact}(K)$ . It follows

that there is an admissible function of the language  $L$  to the language  $K$  and an admissible function of  $K$  to  $L$  if and only if  $\text{Fact}(L) = \text{Fact}(K)$ . In this case the languages  $L$  and  $K$  are *factorequal*.

Hence every equivalence class of  $\mathbf{F}_\Sigma$ , that is of factorequal languages, contains exactly one factor closed language. Furthermore, the partial order which is the skeleton of the category  $\mathbf{F}_\Sigma$  is isomorphic to the partial order  $(\Sigma_{\text{cl}}; \subseteq)$  where  $\Sigma_{\text{cl}}$  is the set of all infinite factor closed languages with alphabet  $\Sigma$ . The language  $L \subseteq \Sigma^*$  is minimal in the category  $\mathbf{F}_\Sigma$  if and only if the language  $\text{Fact}(L)$  is a minimal element of the partial order  $(\Sigma_{\text{cl}}; \subseteq)$ . The category  $\mathbf{F}_\Sigma$  is Zorns if and only if the partial order  $(\Sigma_{\text{cl}}; \subseteq)$  is Zorns.

The partial order  $(\Sigma_{\text{cl}}; \subseteq)$  has infinite antichains and chains if  $\Sigma$  contains at least two elements. Let  $\Sigma := \{a, b\}$  and  $w_0 := aa$ ,  $w_1 := aba$ ,  $w_2 := abba$ ,  $w_3 := abbaa$ ,  $\dots$  and  $B^n := \text{Del}(\Sigma^*, w_n)$  for  $n \in \mathbb{N}$ . The set  $\{B^n \mid n \in \mathbb{N}\}$  is an infinite antichain of  $\Sigma_{\text{cl}}$ . Hence  $\{\Sigma^* - \bigcup_{i \in \mathbb{N}} B^i \mid n \in \mathbb{N}\}$  is an infinite descending chain of  $\Sigma_{\text{cl}}$ .

## 2. RAMSEY THEORY AND ATOMS IN QUASIORDERS

A theorem of Ramsey [4] says that if  $S$  is a countable infinite set and  $n \in \mathbb{N}$  and  $A \cup B = [S]^n$  then there is an infinite subset  $T$  of  $S$  so that  $[T]^n \subseteq A$  or  $[T]^n \subseteq B$ . ( $[S]^n$  is the set of all  $n$ -element subsets of  $S$ .) If  $n = 1$  then this theorem is just a special case of the pigeon hole principle.

In order to explain a different point of view of this theorem we will use the case  $n = 2$ . Then we can reformulate the theorem as follows: let  $\mathbf{Q}$  be the quasiorder of all countable infinite graphs with  $G \leq H$  if there is an embedding of  $G$  into  $H$ . The quasiorder  $\mathbf{Q}$  contains two atoms, the complete graph on countably many vertices and the graph without edges on countably many vertices. These atoms have the property that below every element of the quasiorder there is an atom.

This theorem in turn can be formulated as: let  $\mathbf{K}$  be the countable infinite complete graph. Let  $\mathbf{Q}$  be the quasiorder of all two-edge colored copies of  $\mathbf{K}$  with  $G \leq H$  if there is a color preserving embedding of  $G$  to  $H$ . Let the two colors be red and blue. This quasiorder  $\mathbf{Q}$  has two atoms, the complete graph  $\mathbf{K}$  all of whose edges are blue and the complete graph  $\mathbf{K}$  all of whose edges are red.

More generally, given a structure  $\mathbf{R}$  and an embedded substructure  $\mathbf{C}$ , consider the quasiorder of all copies of  $\mathbf{R}$  in which every one of the copies of  $\mathbf{C}$  in  $\mathbf{R}$  are colored with one of  $r$  given colors. For two such colored structures  $\mathbf{G}$  and  $\mathbf{H}$  let  $\mathbf{G} \leq \mathbf{H}$  if there is a color preserving embedding of  $\mathbf{G}$  into  $\mathbf{H}$ . (Of course for each such instance the notions of copy, embedding, substructure etc. have to be defined.)

If such a quasiorder has the set  $\mathcal{A}$  as set of atoms then one has a theorem of the following form: however the copies of  $\mathbf{C}$  in  $\mathbf{R}$  are colored with  $r$  colors one of the atoms in  $\mathcal{A}$  will always appear as an induced color preserving substructure. (See [3] and [5] for a starting point to the literature on this.)

Note that a coloring of the copies of  $\mathbf{C}$  in the structure  $\mathbf{R}$  is a function from the set of copies of  $\mathbf{C}$  in  $\mathbf{R}$  to the set of colors. If  $\Gamma$  and  $\Sigma$  are alphabets and  $f$  a function

of  $\Sigma$  to  $\Gamma$  then  $f$  can be viewed as a coloring of the one element substructures of the “structure”  $\Sigma^*$ . If  $R = (r_i; i \in \mathbb{N})$  is an increasing list of words in  $\Gamma^*$  then every language  $L \in \mathbf{B}(R, f)$  can be viewed as a colored copy of  $R$ . In this case there are more than two atoms. Nevertheless we show in Theorem 5.2 that those atoms have the property that for each element  $L \in \mathbf{B}(R, f)$  there is at least one of the atoms which has an admissible function into  $L$ .

### 3. THE CATEGORIES $\mathbf{F}_\Sigma$ AND $\mathbf{I}_\Sigma$

**Theorem 3.1.** *The language  $L$  is minimal in  $\mathbf{F}_\Sigma$  if and only if it is uniformly recurrent.*

*Proof.* It suffices to prove that  $A$  is minimal in the partial order  $(\Sigma_{\text{cl}}, \subseteq)$  if and only if  $A$  is uniformly recurrent.

Assume  $A \in \Sigma_{\text{cl}}$  is uniformly recurrent and  $L \subseteq A$  is a language in  $\Sigma_{\text{cl}}$ . Let  $w \in A$  and  $k$  be the number so that when  $x \in A$  with  $|x| \geq k$  then  $w|x$ . Let  $v \in L$  with  $|v| \geq k$ . Then  $v \in A$  because  $L \subseteq A$ . Hence  $w|v$  and hence  $w \in L$ . Therefore  $A \subseteq L$ .

Assume  $A \in \Sigma_{\text{cl}}$  is minimal. Let  $w \in A$ . If for every  $k \in \mathbb{N}$  there is a word  $v \in A$  with  $|v| \geq k$  and  $w \not|v$  then the set  $\text{Del}(A, w)$  is infinite and hence an element of  $\Sigma_{\text{cl}}$ . Because  $\text{Del}(A, w) \subset A$  but  $A \not\subseteq \text{Del}(A, w)$  we arrived at a contradiction to  $A$  being minimal. Hence there is a  $k \in \mathbb{N}$  so that  $w|v$  for all  $v \in A$  with  $|v| \geq k$ . It follows that  $A$  is uniformly recurrent. □

**Lemma 3.1.** *Let  $A$  be a minimal element in the category  $\mathbf{F}_\Sigma$  and  $u, w \in \text{Fact}(A)$ . Then there is  $x \in \text{Fact}(A)$  so that  $uxw \in \text{Fact}(A)$ .*

*Proof.* Let  $u, w \in \text{Fact}(A)$ . It follows from Theorem 3.1 that  $A$  is uniformly recurrent and hence there is a number  $n \in \mathbb{N}$  so that whenever  $v \in \text{Fact}(A)$  and  $|v| \geq n$  then  $u \leq v$  and  $w \leq v$ . Let  $v_1v_2 \in \text{Fact}(A)$  with  $|v_1| = |v_2| = n$ . Then  $u \leq v_1$  and  $w \leq v_2$ . Hence  $v_1v_2 = x_0ux_1x_2wx_3$  and we obtain for  $x = x_1x_2$  that  $uxw \in \text{Fact}(A)$ . □

**Theorem 3.2.** *Let  $\Sigma$  be a finite alphabet. The category  $\mathbf{F}_\Sigma$  of all infinite languages with admissible functions as morphisms is Zorns.*

*Proof.* It suffices to prove that the partial order  $(\Sigma_{\text{cl}}; \subseteq)$  is Zorns. Let  $S$  be a chain in  $\Sigma_{\text{cl}}$ . For every  $L \in S$  and  $n \in \mathbb{N}$  let  $L_n$  be the set of words of length  $n$  in  $L$ . Because  $L$  is factor closed  $L_n \neq \emptyset$  for every  $n \in \mathbb{N}$ . Because  $L_n$  is finite there is for every  $n \in \mathbb{N}$  an element  $F(n) \in S$  so that  $K_n = F(n)_n$  for all  $K \in S$  with  $K \subseteq F(n)$ . It follows that

$$\bigcup_{n \in \mathbb{N}} F(n)_n$$

is a lower bound of the chain  $S$  which is infinite because  $F(n)_n \neq \emptyset$  for every  $n \in \mathbb{N}$ . □

**Lemma 3.2.** *Let  $L$  and  $K$  be two infinite languages over the alphabet  $\Sigma$  and  $K$  recurrent. If there is an admissible function of  $L$  to  $K$  then there is an admissible injection of  $L$  to  $K$ . If there is an admissible function of  $K$  to  $L$  then there is an admissible injection of  $K$  to  $L$ .*

*Proof.* Let  $\lambda$  be an admissible function of  $L$  to  $K$  and  $L := \{w_i \mid i \in \mathbb{N}\}$  be an enumeration of  $L$  and  $L_n := \{w_i \mid i \in n\}$ . We will show that for every  $n \in \mathbb{N}$  and every admissible injection  $\phi_n$  of  $L_n$  into  $K$  there is an extension  $\phi_{n+1}$  of  $\phi_n$  which is an admissible injection of  $L_{n+1}$  into  $K$ . Clearly then  $\phi := \bigcup_{n \in \mathbb{N}} \phi_n$  is an admissible injection of  $L$  into  $K$ .

Let  $\phi_n$  be an admissible injection of  $L_n$  into  $K$  and  $B := K - \{\phi_n(w_i) \mid i \in n\}$ . Because  $K$  is recurrent there are infinitely many words in  $K$  which have  $\lambda(w_n)$  as a factor. Hence there is  $v \in B$  so that  $w_n|v$ . Because  $w_n|\lambda(w_n)|v$  it follows that  $w_n|v \in B$ . We extend  $\phi_n$  to  $\phi_{n+1}$  by stipulating that  $\phi_{n+1}(w_n) := v$ . Clearly  $\phi_{n+1}$  is an admissible injection of  $L_{n+1}$  to  $K$ .

Let  $\lambda$  be an admissible function of  $K$  to  $L$  and  $K := \{w_i \mid i \in \mathbb{N}\}$  be an enumeration of  $K$  and  $K_n := \{w_i \mid i \in n\}$ . We will show that for every  $n \in \mathbb{N}$  and every admissible injection  $\phi_n$  of  $K_n$  into  $L$  there is an extension  $\phi_{n+1}$  of  $\phi_n$  which is an admissible injection of  $K_{n+1}$  into  $L$ . Then  $\phi := \bigcup_{n \in \mathbb{N}} \phi_n$  is an admissible injection of  $K$  into  $L$ .

Let  $\phi_n$  be an admissible injection of  $K_n$  into  $L$  and  $B := \{\phi_n(w_i) \mid i \in n\}$ . Because  $K$  is recurrent there are infinitely many words in  $K$  which have  $\lambda(w_n)$  as a factor. Hence there is  $v \in K$  with  $|v| > |\phi_n(w_i)|$  for all  $i \in n$  and  $w_n|v$ . It follows that  $\lambda(v) \notin B$ . We extend  $\phi_n$  to  $\phi_{n+1}$  by stipulating that  $\phi_{n+1}(w_n) := \lambda(v)$ . Clearly  $\phi_{n+1}$  is an admissible injection of  $K_{n+1}$  to  $L$ . □

**Theorem 3.3.** *The category  $\mathbf{I}_\Sigma$  of languages with admissible injections as morphisms is Zorns. The set of minimal elements of  $\mathbf{I}_\Sigma$  is equal to the set of minimal elements of  $\mathbf{F}_\Sigma$ .*

*Proof.* Let  $S$  be a set of infinite languages over the alphabet  $\Sigma$  so that for any two elements  $L, K \in S$  there is an admissible injection of  $L$  to  $K$  or an admissible injection of  $K$  to  $L$ . The category  $\mathbf{F}_\Sigma$  is Zorns according to Theorem 3.2 and hence there is a language  $B$  over the alphabet  $\Sigma$  so that there is an admissible function of  $B$  to  $L$  for every language  $L \in S$ . Because  $\mathbf{F}_\Sigma$  is atomic there is a minimal element  $M$  of  $\mathbf{F}_\Sigma$  and an admissible function of  $M$  to  $B$ . Hence there is an admissible function of  $M$  to every element in  $S$ .

The language  $M$  is uniformly recurrent according to Theorem 3.1 and hence recurrent. Hence there is an admissible injection of  $M$  to every element in  $S$  according to Lemma 3.2. We conclude that the category  $\mathbf{I}_\Sigma$  is Zorns.

Let  $M$  be a minimal element of  $\mathbf{F}_\Sigma$  and  $L$  an infinite language in the alphabet  $\Sigma$  which has an admissible injection to  $M$ . Then there is an admissible function

of  $M$  to  $L$  and therefore because of Lemma 3.2 an admissible injection of  $M$  to  $L$ . It follows that every minimal element of  $\mathbf{F}_\Sigma$  is a minimal element of  $\mathbf{I}_\Sigma$ .

Let  $A$  be a minimal element of  $\mathbf{I}_\Sigma$  and  $M$  a minimal element of  $\mathbf{F}_\Sigma$  which has an admissible function to  $A$ . Then, according to Lemma 3.2 there is an admissible injection of  $M$  to  $A$ . Because  $A$  is minimal in  $\mathbf{I}_\Sigma$  there is an admissible injection of  $A$  to  $M$ . It follows that  $\text{Fact}(A) = \text{Fact}(M)$  and hence  $A$  is uniformly recurrent and therefore a minimal element of  $\mathbf{F}_\Sigma$ .  $\square$

#### 4. ATOMS AND INFINITE WORDS

Let  $\Sigma$  be a finite alphabet. We denote by  $\Sigma^\omega$  the set of *infinite words* with elements in  $\Sigma$ , that is the set of infinite sequences with entries in  $\Sigma$ . If  $v \in \Sigma^\omega$  and  $w \in \Sigma^*$  then  $w|v$  if there is a  $u \in \Sigma^*$  and an  $x \in \Sigma^\omega$  so that  $v = uwx$ . If  $v \in \Sigma^\omega$  then  $\text{Fact}(v) := \{w \in \Sigma^* \mid w|v\}$ .

Let  $A \subseteq \Sigma^*$  be an atom of  $\mathbf{F}_\Sigma$ . We denote by  $\text{Inf}(A)$  the set of all words  $w \in \Sigma^\omega$  so that  $\text{Fact}(A) = \text{Fact}(w)$ .

**Theorem 4.1.** *If  $A \subseteq \Sigma^*$  is an atom of  $\mathbf{F}_\Sigma$  then  $\text{Inf}(A)$  is not empty.*

*Proof.* Let  $A := \{w_i \mid i \in \mathbb{N}\}$ . We use Lemma 3.1 and Theorem 3.3 to successively obtain words

$$w_0x_0w_1, w_0x_0w_1x_1w_2, w_0x_0w_1x_1w_2x_2w_3, w_0x_0w_1x_1w_2x_2w_3x_3w_4, \dots$$

in  $\text{Fact}(A)$ . It follows that if

$$w = w_0x_0w_1x_1w_2x_2w_3x_3w_4x_4 \dots$$

then  $w \in \Sigma^\omega$  and  $\text{Fact}(w) = \text{Fact}(A)$ .  $\square$

The word  $w \in \Sigma^\omega$  is *uniformly recurrent* if for every  $u \in \text{Fact}(w)$  there exists an  $n \in \mathbb{N}$  such that for all  $v \in \text{Fact}(w)$  with  $|v| \geq n$  the word  $u$  is a factor of  $v$ . (See Def. 2.3.2 of [1].) That is  $w \in \Sigma^\omega$  is uniformly recurrent if and only if the language  $\text{Fact}(w)$  is uniformly recurrent if and only if  $\text{Fact}(w)$  is an atom (Th. 3.1.). It follows that  $A$  is an atom if and only if each of the infinite words in  $\text{Inf}(A)$  is uniformly recurrent.

The following corollary supplements Theorem 2.3.1 of [1].

**Corollary 4.1.** *Let  $A$  be an atom in  $\mathbf{F}_\Sigma$ . There exists a uniformly recurrent word  $w \in \Sigma^\omega$  such that  $\text{Fact}(w) = \text{Fact}(A)$ .*

An infinite word  $w \in \Sigma^\omega$  is *periodic* if there exists  $v \in \Sigma^+$  such that  $w = v^\omega$ , i.e.,  $w = vvv \dots v \dots$ . An infinite word  $w \in \Sigma^\omega$  is  *$\omega$ -power-free* if for any  $u \in \text{Fact}(w)$ ,  $u \neq \varepsilon$ , there exists  $p > 1$  such that  $u^p \notin \text{Fact}(w)$ .

For a word  $v \in \Sigma^+$  we define  $\text{Pref}(v) = \{p \mid v = pu, u \neq \varepsilon\}$  and  $\text{Suff}(v) = \{s \mid v = us, u \neq \varepsilon\}$ . Observe that, for  $w = v^\omega$ ,  $v \in \Sigma^+$ ,  $\text{Fact}(w) = \text{Suff}(v)\{v\}^*\text{Pref}(v)$ .

**Theorem 4.2.** *Let  $A$  be an atom in  $\mathbf{F}_\Sigma$ . Then  $\text{Fact}(A)$  is either not context-free or there exists  $v \in \Sigma^+$  such that  $\text{Fact}(A) = \text{Suff}(v)\{v\}^*\text{Pref}(v)$ .*

*Proof.* If  $A$  is an atom there exists, by Theorem 4.1 and Corollary 4.1, a uniformly recurrent  $w \in \Sigma^\omega$  with  $\text{Fact}(w) = \text{Fact}(A)$ . By Lemma 2.6.2 of [1],  $w$  is either  $\omega$ -power-free or periodic.

If  $w$  is  $\omega$ -power-free then, by any pumping lemma,  $\text{Fact}(w)$  is not context-free. If  $w = v^\omega$  for some  $v \in \Sigma^+$  then  $\text{Fact}(w) = \text{Suff}(v)\{v\}^*\text{Pref}(v)$ . □

**Lemma 4.1.** *If  $L$  is context-free then  $\text{Fact}(L)$  is context-free.*

*Proof.* Consider the generalized sequential machine (gsm)  $T_{\text{inf}}$  constructed in the proof of Theorem 4.3 of [6]. This gsm  $T_{\text{inf}}$  realizes the gsm mapping  $\tau : \mathfrak{P}(\Sigma^*) \rightarrow \mathfrak{P}(\Sigma^*)$  given by  $\tau(L) = \text{Fact}(L)$ ,  $L \in \mathfrak{P}(\Sigma^*)$ . Since context-free languages are closed under gsm mappings,  $\text{Fact}(L)$  is context-free. □

**Theorem 4.3.** *Let  $A$  be an atom in  $\mathbf{F}_\Sigma$ . Then  $A$  is not context-free or  $A$  is an infinite subset of  $\text{Suff}(v)\{v\}^*\text{Pref}(v)$  for some  $v \in \Sigma^+$ .*

*Proof.* Assume that  $A$  is context-free. Then by Lemma 4.1  $\text{Fact}(A)$  is again context-free. Hence, by Theorem 4.2 there exists a  $v \in \Sigma^+$  such that  $\text{Fact}(A) = \text{Fact}(v^\omega)$ . Hence  $A \subseteq \text{Fact}(A)$  proves our theorem. □

**Corollary 4.2.** *Let  $A$  be an atom in  $\mathbf{F}_\Sigma$ . Then  $A$  is either not context-free or  $A$  is a regular subset of  $\text{Suff}(v)\{v\}^*\text{Pref}(v)$  for some  $v \in \Sigma^+$ .*

*Proof.* Any context-free subset of  $\text{Suff}(v)\{v\}^*\text{Pref}(v)$ ,  $v \in \Sigma^+$ , is regular. □

**Example.** An infinite word  $w \in \{a, b\}^\omega$  is called *Sturmian* if  $|\text{Fact}(w) \cap \{a, b\}^n| = n + 1$  for all  $n \in \mathbb{N}$ . Any Sturmian word is uniformly recurrent (see [1], Prop. 2.3.2). The most famous Sturmian word is the Fibonacci word

$$f = abaababaabaab \dots$$

which is the limit of the sequence  $(f_n)_{n \in \mathbb{N}}$  of words inductively defined as

$$f_0 = a, \quad f_1 = ab, \quad f_{n+2} = f_{n+1}f_n, \quad n \in \mathbb{N}.$$

The Fibonacci word  $f$  gives rise to the atom  $\text{Fact}(f)$  in  $F_{\{a,b\}}$ . All infinite subsets of  $\text{Fact}(f)$  are atoms in  $F_{\{a,b\}}$ , as well. They are not context-free. Observe that if  $L$  is an atom of  $\mathbf{F}_\Sigma$  and  $K$  is an infinite subset of  $L$  then  $\text{Fact}(K) = \text{Fact}(L)$ . Hence,  $K$  is also an atom.

Another famous infinite word is the Thue–Morse word

$$t = abbabaabbaabba \dots$$

which is the limit of the sequence  $(t_n)_{n \in \mathbb{N}}$  of words inductively defined as

$$t_0 = a, \quad t_{n+1} = t_n \varphi(t_n), \quad n \in \mathbb{N}.$$



Here,  $\varphi : \{a, b\}^* \rightarrow \{a, b\}^*$  is the morphism defined by  $\varphi(a) = b, \varphi(b) = a$ .

The Thue–Morse word  $t$  gives rise to the atom  $\text{Fact}(t)$  in  $F_{\{a,b\}}$ .

The Thue–Morse word  $m$  on three symbols

$$m = abcacbabcbac \dots$$

can be introduced as the limit of the sequence  $(m_n)_{n \in \mathbb{N}}$  of words inductively defined as

$$m_0 = a, \quad m_{n+1} = \mu(m_n), \quad n \in \mathbb{N}.$$

Here,  $\mu : \{a, b, c\}^* \rightarrow \{a, b, c\}^*$  is the morphism defined by  $\mu(a) = abc, \mu(b) = ac, \mu(c) = b$ .

The Thue–Morse word  $m$  gives rise to the atom  $\text{Fact}(m)$  in  $F_{\{a,b,c\}}$ .

The periodic word

$$p = abababab \dots$$

gives rise to the atom  $\text{Fact}(p) = \{\varepsilon, b\}\{ab\}^*\{\varepsilon, a\}$ .

### 5. ATOMS OF SUBSETS

**Lemma 5.1.** *Let  $f$  be a function of the finite alphabet  $\Sigma$  to the finite alphabet  $\Gamma$  and  $R$  a uniformly recurrent and increasing list of words in  $\Gamma^*$ . If  $K \subseteq \text{Fact}(L)$  and  $L \in \mathbf{B}(R, f)$  and  $K$  infinite then there is  $H \subseteq \text{Fact}(K)$  with  $H \in \mathbf{B}(R, f)$  and an admissible injection of  $H$  to  $L$ .*

*Proof.* Because  $R$  is uniformly recurrent it is infinite and hence we may assume that  $R := (r_i; i \in \mathbb{N})$  and that  $L = \{l_i \mid i \in \mathbb{N}\}$  with  $f(l_i) = r_i$ . Because  $L$  is infinite the lengths of the entries of  $R$  are unbounded. We will construct the language  $H = \{h_i \mid i \in \mathbb{N}\} \subseteq \text{Fact}(K)$  so that  $f(h_i) = r_i$  for all  $i \in \mathbb{N}$ . In addition we will construct an injection  $\lambda$  of  $H$  to  $L$  so that  $h_i \mid \lambda(h_i)$  for all  $i \in \mathbb{N}$ .

Assume we have already constructed the set  $H_n := \{h_i \mid i \in n \in \mathbb{N}\} \subseteq \text{Fact}(K)$  so that  $f(h_i) = r_i$  for all  $i \in n$  and an injection  $\lambda_n$  of  $H_n$  to  $L$  so that  $h_i \mid \lambda(h_i)$  for all  $i \in n$ . We have to construct  $r_n$ .

Because  $R$  is uniformly recurrent there is an  $m' \in \mathbb{N}$  so that  $r_n \mid r \in \text{Fact}(R)$  if  $m' \leq |r|$ . Let  $m''$  be the maximum of the set  $\{|\lambda(k_i)| \mid i \in n\}$  and  $m$  the maximum of  $m'$  and  $m''$ . The language  $K$  is infinite and hence there is  $s \in \mathbb{N}$  so that  $|k_s| \geq m$ . There is  $t \in \mathbb{N}$  so that  $k_s \mid l_t$  and hence words  $u$  and  $v$  so that  $l_t = uk_s v$  and hence  $r_t = f(l_t) = f(u)f(k_s)f(v)$ .

Because  $m' \leq |k_s| = |f(k_s)| \in \text{Fact}(R)$  there are words  $u_1$  and  $v_1$  so that  $f(k_s) = u_1 r_n v_1$  and hence a word  $k \in \text{Fact}(K)$  so that  $f(k) = r_n$ . Let  $h_n := k$  and  $H_{n+1} = H_n \cup \{h_n\}$  and  $\lambda_{n+1}$  the extension of  $\lambda_n$  to  $H_{n+1}$  for which  $\lambda_{n+1}(h_n) = l_t$ . The function  $\lambda_{n+1}$  is an injection because  $l_t$  is longer than any of the words in the image of  $\lambda_n$ . □

**Theorem 5.1.** *Let  $f$  be a function of the finite alphabet  $\Sigma$  to the finite alphabet  $\Gamma$  and  $R$  a uniformly recurrent and increasing list of words in  $\Gamma^*$ . Then the category  $\mathbf{B}(R, f)$  is Zorns and every minimal element of  $\mathbf{B}(R, f)$  is a minimal element of  $\mathbf{F}_\Sigma$ .*

*Proof.* Let  $S$  be a chain of infinite languages in  $\mathbf{B}(R, f)$ . According to Theorem 3.2 there is a minimal element  $K$  of  $\mathbf{F}_\Sigma$  so that there is an admissible injection of  $K$  to  $L$  for every  $L$  in  $S$ . The language  $K$  is uniformly recurrent (Th. 3.1).

According to Lemma 5.1 there is for every  $L \in S$  a language  $H_L \subseteq \text{Fact}(K)$  with  $H_L \in \mathbf{B}(R, f)$  and an admissible injection of  $H_L$  to  $L$ . Because  $K$  is uniformly recurrent  $\text{Fact}(H_L) = \text{Fact}(K)$ . Hence each of the languages  $H_L$  is uniformly recurrent. If  $L, L' \in S$  then  $\text{Fact}(H_L) = \text{Fact}(K) = \text{Fact}(H_{L'})$  and hence according to Lemma 3.2 there is an admissible injection of  $H_L$  to  $H_{L'}$ . Because there is an admissible injection of  $H_{L'}$  to  $L'$  there is an admissible injection of  $H_L$  to  $L'$ .

Fix  $L \in S$ . Then for every  $L' \in S$  there is an admissible injection of  $H_L$  to  $L'$  and  $H' \in \mathbf{B}(R, f)$ . It follows that  $\mathbf{B}(R, f)$  is Zorns.

Let  $L$  be a minimal element of  $\mathbf{B}(R, f)$  and  $K$  a minimal element of  $\mathbf{F}_\Sigma$  with  $K \subseteq \text{Fact}(L)$  (Th. 3.2). Using Lemma 3.2 there is a language  $H \subseteq \text{Fact}(K)$  with  $H \in \mathbf{B}(R, f)$  and an admissible injection of  $H$  to  $L$ . Because  $L$  is minimal in  $\mathbf{B}(R, f)$  there is an admissible injection of  $L$  to  $H$ . Hence  $\text{Fact}(H) \subseteq \text{Fact}(K) \subseteq \text{Fact}(L) \subseteq \text{Fact}(H)$  which in turn implies  $\text{Fact}(H) = \text{Fact}(K) = \text{Fact}(L)$ . Hence  $L$  is also uniformly recurrent and according to Theorem 3.1 a minimal element of  $\mathbf{F}_\Sigma$ .  $\square$

Let  $\Sigma = \{a, b\}$  and  $\Gamma = \{0, 1\}$  and  $f(a) = 0$  and  $f(b) = 1$  and  $R$  an enumeration of  $\Gamma^*$  in non descending order of lengths. Then  $\mathbf{B}(R, f) = \{\Sigma^*\}$  and the only minimal element of  $\mathbf{B}(R, f)$  is  $\Sigma^*$ . We conclude, because  $\Sigma^*$  is not uniformly recurrent, that the condition “uniformly recurrent” on  $R$  in Theorem 5.1 can not be omitted and still obtain that the minimal elements of  $\mathbf{B}(R, f)$  are uniformly recurrent. Nevertheless we will show that  $\mathbf{B}(R, f)$  is always Zorns.

**Lemma 5.2.** *Let  $f$  be a function of the finite alphabet  $\Sigma$  to the finite alphabet  $\Gamma$  and  $R$  a recurrent and increasing list of words in  $\Gamma^*$ . Then  $\mathbf{B}(R, f)$  is Zorns.*

*Proof.* Because  $R$  is recurrent we may assume that  $R := (r_i; i \in \mathbb{N})$ . Let  $S$  be a chain of infinite languages of  $\mathbf{B}(R, f)$ . For every  $L \in S$  let  $(w_i^L; i \in \mathbb{N})$  be an enumeration of  $L$  so that  $f(w_i^L) = r_i$  for every  $i \in \mathbb{N}$ .

For every  $i \in \mathbb{N}$  there are only finitely many words  $w \in \Sigma^*$  so that  $f(w) = r_i$ . For every  $i \in \mathbb{N}$  there are infinitely many  $j \in \mathbb{N}$  so that  $r_i$  is a factor of  $r_j$  because  $R$  is recurrent. Hence for every  $L \in S$  and every  $i \in \mathbb{N}$  there are infinitely many  $j \in \mathbb{N}$  so that  $w_j^L$  contains a factor  $w$  with  $f(w) = r_i$ . It follows that the set

$$T_i(L) := \{w \in \text{Fact}(L) \mid f(w) = r_i \text{ and } w \text{ is a factor of } w_j^L \text{ for infinitely many } j \in \mathbb{N}\}$$

is not empty for any  $L \in S$ . Because  $L \supseteq K$  implies  $T_i(L) \supseteq T_i(K)$  and  $T_i(L)$  is finite for every  $L \in S$  it follows that there is for every  $i \in \mathbb{N}$  an element  $L^i \in S$  so that  $T_i(L^i) = T_i(K)$  for all  $K \in S$  for which there is an admissible injection of  $K$  to  $L^i$ . Select an element  $v_i \in T_i(L^i)$  for every  $i \in \mathbb{N}$  and let  $B := \{v_i \mid i \in \mathbb{N}\}$ .

It follows from the definition of  $T_i(L)$  that  $f(v_i) = r_i$  for every  $i \in \mathbb{N}$  and hence  $B \in \mathbf{B}(R, f)$ . In order to prove that  $B$  is a lower bound of  $S$  it suffices to show that for every  $L \in S$  there is an injection  $\alpha$  of  $\mathbb{N}$  to  $\mathbb{N}$  so that  $v_i | w_{\alpha(i)}^L$  for all  $i \in \mathbb{N}$ .

Let  $L \in S$ . We construct step by step a sequence of functions  $\emptyset = \alpha_0 \subseteq \alpha_1 \subseteq \alpha_2 \subseteq \alpha_3 \subseteq \dots$  so that  $\alpha_n$  is an injection of  $n$  to  $\mathbb{N}$  with  $v_i|w_{\alpha(i)}^L$  for all  $i \in n$ .

Let  $\alpha_n$  be given. There are infinitely many  $j \in \mathbb{N}$  so that  $v_n|w_j^L$ . Hence there is a  $j \in \mathbb{N}$  which is not in the range of  $\alpha_n$  and with  $v_n|w_j^L$ . Let  $\alpha_{n+1}$  be given by  $\alpha_n \subseteq \alpha_{n+1}$  and  $\alpha_{n+1}(n) = j$ .

The function  $\alpha := \bigcup_{n \in \mathbb{N}} \alpha_n$  is an injection of  $\mathbb{N}$  to  $\mathbb{N}$  so that  $v_i|w_{\alpha(i)}^L$  for all  $i \in \mathbb{N}$ . □

Let  $S = (s_i; i \in I \subseteq \mathbb{N})$  be a finite or infinite increasing list of words in  $\Gamma^*$ . The list  $S$  is recurrent if and only if for every  $i \in I$  there is  $j > i$  in  $I$  so that  $s_i|s_j$ . If  $S$  is not recurrent let  $N(S)$  be the smallest number in  $I$  so that  $s_{N(S)} \not|s_j$  for all  $j > N(S)$  and let  $M(S)$  be the smallest number in  $I$  so that  $s_{M(S)} = s_{N(S)}$ . Note that according to the definition increasing list of words if  $M(S) \leq i \leq N(S)$  and  $i \in I$  then  $s_{M(S)} = s_i = s_{N(S)}$ .

If  $I = \emptyset$  or if  $S$  is recurrent let  $S' := S$ . If  $S$  is not empty and not recurrent let  $I' := I - \{i \in I \mid M(S) \leq i \leq N(S)\}$  and  $S' = (s_i; i \in I')$  that is the list obtained from the list  $S$  by removing all entries  $r_i$  with  $M(S) \leq i \leq N(S)$ . Let  $\Delta(S) := (r_i; M(S) \leq i \leq N(S) \text{ and } i \in I)$ . If  $L \in \mathbf{B}(S, f)$  with  $L = \{l_i \mid i \in I\}$  and  $f(l_i) = s_i$  then  $\Delta_S(L) := \{l_i \mid M(S) \leq i \leq N(S) \text{ and } i \in I\}$ .

**Lemma 5.3.** *Let  $f$  be a function of the finite alphabet  $\Sigma$  to the finite alphabet  $\Gamma$  and  $S = (s_i; i \in I \subseteq \mathbb{N})$  an increasing and non-recurrent list of words in  $\Gamma^*$ . Let  $L = \{l_i \mid i \in I\}$  and  $K = \{k_i \mid i \in I\}$  with  $f(l_i) = f(k_i) = s_i$  for all  $i \in I$  and  $\lambda$  an admissible injection of  $L$  to  $K$ . Then there is a permutation  $\pi$  of the set  $\{i \in I \mid M(S) \leq i \leq N(S)\}$  so that  $s_i = k_{\pi(i)} = f(s_i)$  for all  $i \in I$  with  $M(S) \leq i \leq N(S)$ .*

*Proof.* Let  $M(S) \leq i \leq N(S)$  and  $i \in I$  and  $\lambda(l_i) = k_j$ . If  $j > N(S)$  then  $l_i|k_j$  hence  $f(l_i)|f(k_j)$  hence  $s_i|s_j$ . But  $s_i = s_{N(S)}$  according to the definition of  $M(S)$  hence  $s_{N(S)}|s_j$  in contradiction to the definition of  $N(S)$ .

If  $j < M(S)$  then as above  $s_i|s_j$ . Because  $s_i = s_{M(S)}$  we get  $s_{M(S)}|s_j$ . Because  $|s_j| \leq |s_{M(S)}|$  it follows that  $s_j = s_{M(S)}$  in contradiction to the definition of  $M(S)$ .

Hence  $M(S) \leq j \leq N(S)$ . Because  $f(l_i) = s_i = s_j = f(k_j)$  it follows that  $|l_i| = |k_j|$  and because  $l_i|k_j$  that  $l_i = k_j$ . Because  $\lambda$  is one-to-one there is a permutation  $\pi$  of the set  $\{i \in I \mid M(S) \leq i \leq N(S)\}$  so that  $s_i = k_{\pi(i)} = f(s_i)$  for all  $i \in I$  with  $M(S) \leq i \leq N(S)$ . □

Under the assumptions of the previous lemma we denote the restriction of the function  $\lambda$  to  $\{l_i \mid i \in I - \{j \in I \mid M(S) \leq j \leq N(S)\}\}$  by  $\lambda_{S,L,K}$ . It follows from Lemma 5.3 that  $\lambda_{S,L,K}$  is an admissible injection of  $\{l_i \mid i \in I - \{j \in I \mid M(S) \leq j \leq N(S)\}\}$  to  $\{k_i \mid i \in I - \{j \in I \mid M(S) \leq j \leq N(S)\}\}$ . Let the restriction of the function  $\lambda$  to  $\{l_i \mid i \in I \text{ and } M(S) \leq i \leq N(S)\}$  be  $\lambda^{S,L,K}$ . It follows from Lemma 5.3 that  $\lambda^{S,L,K}$  is an admissible injection of  $\{l_i \mid i \in I \text{ and } M(S) \leq i \leq N(S)\}$  to  $\{k_i \mid i \in I \text{ and } M(S) \leq i \leq N(S)\}$  and that  $\lambda = \lambda_{S,L,K} \cup \lambda^{S,L,K}$ . Note that  $\lambda^{S,L,K}$  is a bijection; that is an admissible bijection.

**Theorem 5.2.** *Let  $f$  be a function of the finite alphabet  $\Sigma$  to the finite alphabet  $\Gamma$  and  $R$  an increasing list of words in  $\Gamma^*$ . Then  $\mathbf{B}(R, f)$  is Zorns.*

*Proof.* If the lengths of the words in  $R$  are bounded then  $\mathbf{B}(R, f)$  is empty. We assume that the lengths of the words in  $R$  are unbounded. It follows that  $R$  is infinite. Therefore we can write  $R$  as  $R = (r_i; i \in \mathbb{N})$ . Let  $S$  be a chain of languages in  $\mathbf{B}(R, f)$ . For  $L \in S$  let  $(w_i^L; i \in \mathbb{N})$  be an enumeration of  $L$  so that  $f(w_i^L) = r_i$  for every  $i \in \mathbb{N}$ .

We define recursively the list  $R^\alpha$  and the set  $I^\alpha$  for every ordinal number  $\alpha$  so that  $R^\alpha = (r_i; i \in I^\alpha)$ . Let  $R^0 := R$  and  $I^0 = I$  and  $R^{\alpha+1} := (R^\alpha)'$  and  $I^{\alpha+1} := (I^\alpha)'$ . If  $\alpha$  is a limit ordinal then  $I^\alpha = \bigcap_{\beta \in \alpha} I^\beta$  and  $R^\alpha := (r_i; i \in I^\alpha)$ . Note that there is an ordinal  $\alpha$  so that  $R^\alpha = R^\beta$  for all  $\beta > \alpha$ . We call the smallest such ordinal  $\alpha$  the *index* of the list  $R$ . Let  $\nu$  be the index of  $R$ . Then  $R^\nu$  is empty or  $R^\nu$  is recurrent.

For  $\alpha$  an ordinal and  $L \in S$  let  $L^\alpha := \{w_i^L \mid i \in I^\alpha\}$ . □

**Claim 1.** *Let  $L, K \in S$  and  $\lambda$  an admissible injection of  $L$  to  $K$ . Then the restriction  $\lambda_\alpha$  of  $\lambda$  to  $L^\alpha$  is an admissible injection of  $L^\alpha$  to  $K^\alpha$  for all  $\alpha \leq \nu$ .*

*Proof.* By induction on  $\alpha$ . If the claim holds at  $\alpha$  it follows from Lemma 5.3 that  $\lambda_{R^\alpha, L^\alpha, K^\alpha}$  is an admissible injection of  $L^{\alpha+1}$  to  $K^{\alpha+1}$ . If  $\alpha$  is a limit then  $\bigcap_{\beta \in \alpha} \lambda_\beta := \lambda_\alpha$  is an admissible injection of  $L^\alpha$  to  $K^\alpha$ . □

**Claim 2.** *Let  $L, K \in S$  and  $\lambda$  an admissible injection of  $L$  to  $K$ . Then the restriction  $\lambda^\alpha$  of  $\lambda$  to  $L - L^\alpha$  is an admissible bijection of  $L - L^\alpha$  to  $K - K^\alpha$  for all  $\alpha \leq \nu$ .*

*Proof.* By induction on  $\alpha$ . It follows from Lemma 5.3 that  $\lambda^{R^\alpha, L^\alpha, K^\alpha}$  is an admissible bijection of  $L^\alpha - L^{\alpha+1}$  to  $K^\alpha - K^{\alpha+1}$ . If the claim holds at  $\alpha$  then  $\lambda^\alpha$  is an admissible bijection of  $L - L^\alpha$  to  $K - K^\alpha$ . Because  $L - L^{\alpha+1} = L - L^\alpha \cup (L^\alpha - L^{\alpha+1})$  and  $K - K^{\alpha+1} = K - K^\alpha \cup (K^\alpha - K^{\alpha+1})$  it follows that  $\lambda^{\alpha+1}$  is an admissible bijection of  $L - L^{\alpha+1}$  to  $K - K^{\alpha+1}$ .

If  $\alpha$  is a limit then  $\bigcup_{\beta \in \alpha} \lambda_\beta := \lambda_\alpha$  is an admissible bijection of  $L - L^\alpha$  to  $K - K^\alpha$ . □

If  $L, K \in S$  and  $\lambda$  is an admissible injection of  $L$  to  $K$  denote by  $\lambda_0$  be the restriction of  $\lambda$  to  $L - L^\nu$  and by  $\lambda_1$  the restriction of  $\lambda$  to  $L^\nu$ . It follows from Claim 2 that  $\lambda_0$  is an admissible bijection of  $L - L^\nu$  to  $K - K^\nu$  and  $\lambda_1$  is an admissible injection of  $L^\nu$  to  $K^\nu$ .

If  $R^\nu$  is empty then  $L^\nu$  is empty for every  $L \in S$ . It follows that there is an admissible bijection between any two elements  $L$  and  $K$  of  $S$ . Hence any language  $L \in S$  is a lower bound of  $S$ .

If  $R^\nu$  is recurrent then  $S_1 := \{L^\nu \mid L \in S\}$  is a chain in  $\mathbf{B}(R^\nu, g)$  where  $g$  is the restriction of  $f$  to  $I^\nu$ . We obtain from Lemma 5.2 a lower bound, say  $M$ , of the chain  $S_1$ . Then the language  $L - L^\nu \cup M$  is a lower bound of the chain  $S$  for any  $L \in S$ .

**Example.** Let  $J$  be the set of all sequences  $j = (j_i; i \in \mathbb{N})$  which satisfy the following conditions:

- (i)  $j_i \in \{0, 1\}, i \in \mathbb{N}$ ,
- (ii) if, for some  $n \in \mathbb{N}, j_n = 0$  (resp. 1) then there is  $m > n$  such that  $j_m = 1$  (resp. 0).

Hence,  $J$  contains all 0,1-sequences which have infinitely many 0 and 1.

Let  $f$  be the Fibonacci word and  $R = (r_i; i \in \mathbb{N})$  be the sequence of prefixes of  $f$  ordered by length, *i.e.*,

$$r_0 = \varepsilon, \quad r_1 = a, \quad r_2 = ab, \quad r_3 = aba, \quad r_4 = abaa, \quad \dots$$

Let  $\Gamma = \{a, b\}, \Sigma = \{a_0, a_1, b_0, b_1\}$  and  $f(a_0) = f(a_1) = a, f(b_0) = f(b_1) = b$ .

For  $j \in J, j = (j_0 j_1 j_2 \dots)$ , define  $L_j \subseteq \Sigma^*$  by

$$L_j = \{\varepsilon, a_{j_1}, a_{j_2} b_{j_2}, a_{j_3} b_{j_3} a_{j_3}, a_{j_4} b_{j_4} a_{j_4} a_{j_4}, \dots\},$$

*i.e.*, the  $i$ -th word of  $L_j$  is  $h_{j_i}(r_i)$ , where  $r_i$  is the  $i$ -th prefix of  $f, i \geq 0$ , and  $h_t : \Gamma^* \rightarrow \Sigma^*$  is the monoid morphism defined by  $h_t(a) = a_t, h_t(b) = b_t, t = 0, 1$ . Observe that  $L_j$  is a recurrent language. It is clear that  $L_j, j \in J$  is an infinite language in  $\mathbf{B}(R, f)$ . Consider now  $S = \{L_j \mid j \in J\}$ . We claim that  $S$  is a chain in  $\mathbf{B}(R, f)$ . Since all languages in  $S$  are recurrent, we infer by Lemma 3.2 that there is an admissible injection of  $L_{j_1}$  to  $L_{j_2}, j_1, j_2 \in J$ , if  $\text{Fact}(L_{j_1}) \subseteq \text{Fact}(L_{j_2})$ . Consider an arbitrary word in  $L_{j_1}$ , say  $h_t(r_n), t \in \{0, 1\}, n \in \mathbb{N}$ . By construction there exists  $m \geq n$  such that the word  $h_t(r_m)$  is in  $L_{j_2}$ . Hence,  $L_{j_1} \subseteq \text{Fact}(L_{j_2})$  and  $\text{Fact}(L_{j_1}) \subseteq \text{Fact}(L_{j_2})$ .

In fact, we have proven more: given  $j \in J, \text{Fact}(L_j) = \text{Fact}(L_{j'})$  for all  $j' \in J$ . Hence, each  $L_j, j \in J$ , is a lower bound of  $S$ .

Consider the languages  $M_t = \{h_t(r_i) \mid i \in \mathbb{N}\}, t = 0, 1$ , in  $\mathbf{B}(R, f)$ . They are minimal in  $\mathbf{B}(R, f)$ . Hence, by Theorem 5.1, they are also minimal in  $\mathbf{F}_\Sigma$ , and, in accordance with Theorem 2.3.1 of [1] (see also Cor. 4.1),  $M_0$  and  $M_1$  are uniformly recurrent.

### REFERENCES

- [1] A. de Luca and St. Varrichio, *Finiteness and Regularity in Semigroups and Formal Languages*. Springer (1999).
- [2] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton University Press, Princeton (1981).
- [3] M. Pouzet and N. Sauer, Edge partitions of the Rado graph. *Combinatorica* **16** (1996) 1-16.
- [4] F.P. Ramsey, *On a problem of formal logic*. *Proc. London Math. Soc.* **30** (1930) 264-286.
- [5] N. Sauer, *Coloring finite substructures of countable structures*. The Mathematics of Paul Erdős, X. Bolyai Mathematical Society (to appear).
- [6] S. Yu, *Regular Languages*. In: Handbook of Formal Languages, edited by G. Rozenberg and A. Salomaa, Springer (1997).

Communicated by J. Berstel.

Received February 1, 2001. Accepted September 27, 2001.