LUDWIG STAIGER

# On infinitary finite length codes

*Informatique théorique et applications*, tome 20, nº 4 (1986), p. 483-494

<http://www.numdam.org/item?id=ITA_1986__20_4_483_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 20, n° 4, 1986, p. 483 à 494)

# ON INFINITARY FINITE LENGTH CODES (*)

by Ludwig STAIGER [1]

Communicated by J. BERSTEL

Abstract. – *For a code $C \subseteq X^*$ the following four conditions are considered:*

*(i) for all words u, v, in $X^*$ and for every word w in $C^*$, wu, uv, vu in $C^*$ imply u, v in $C^*$.*

*(ii) every one- sided infinite product of words of C is unambiguous.*

*(iii) C has finite decoding delay.*

*(iv) C has bounded decoding delay.*

*It is shown that in general (iv) → (iii) → (ii) → (i), and the reverse implications are not true; whereas in the case of regular codes C we have (i) → (ii) and (iii) → (iv) but not (ii) → (iii)..*

Résumé. – *On démontre que pour un code rationnel X, les conditions suivantes sont équivalentes: (i) tout produit infini à droite de mots de X est non ambigu.*

*(ii) pour tout mot u, v de $A^*$ et pour tout mot x de $X^*$, xu, uv, vu dans $X^*$ entraîne u, v dans $X^*$;*

*(iii) X a un délai de déchiffrage borné.*

## 0. INTRODUCTION

Let $X \supseteq \{0, 1\}$ be a finite alphabet and $X^*$ be the free monoid generated by $X$. The elements of $X^*$ will be called words, and the unit element $e \in X^*$ is called the empty word. A subset $C \subseteq X^*$ is referred to as a code iff every word $w \in X^*$ admits at most one factorization as a product of elements of $C$, i. e. $w$ is uniquely decipherable.

In this paper we study a class of codes $C$ which satisfy the stronger property that the equality

$$w_1 \cdot w_2 \cdot \ldots = v_1 \cdot v_2 \cdot \ldots \tag{1}$$

(*) Received November 1985, revised in April 1986

([1]) ZKJ der AdW/D.D.R.-1086 Berlin/PSF 1298, R.D.A., Allemagne de l'Est.

of the (one-sided) infinite products where $w_i$, $v_i \in C$ implies that

$$w_i = v_i \quad \text{for all } i \in N.$$

We call those codes infinitary finite length codes (ifl-codes); infinitary according to the product $L$ in (1) and finite length according to the condition $C \subseteqq X^*$. A completely different kind of infinitary codes has been considered in [D 3] (*cf.* also [D 1, 2]), where ω-words are admitted as codewords, i. e. $C \subseteqq X^* \cup X^\omega$ ($X^\omega$ being the set of all semiinfinite words $x_1 x_2 \ldots$ on $X$) but only finite products are considered.

In the case of finite codes it is known that ifl-codes are exactly the codes having finite decipherability delay [L, Sc], but for infinite ifl-codes only little is known. In this paper we show that the classes of ifl-codes having bounded, or finite, or infinite decoding delay are pairwise distinct. Moreover, we show that ifl-codes are weakly prefix codes (*cf.* [C 2]) but the converse does not hold for arbitrary codes. However, it appears that in the case of regular codes the notions of bounded and finite decoding delay as well as ifl-code and weakly prefix code coincide. Since ifl-codes are defined via the infinite product property Equation (1), their study requires not only a consideration of the free monoid

$$C^* =_{df} \{ w_1 \cdot w_2 \cdot \ldots \cdot w_n : w_i \in C \wedge n \in N \}$$

generated by $C$ (*cf.* [Sh]) but also of the ω-power

$$C^\omega =_{df} \{ w_1 \cdot w_2 \cdot \ldots : w_i \in C \setminus \{ e \} \} \subseteqq X^\omega.$$

This brings into the play specific results of the theory of ω-languages and topological methods developed there (*cf.* [S 2, LS, W and S 3]).

## 1. IFL-CODES AND WEAKLY PREFIX CODES

As in [C 2] we call a code $C$ weakly prefix provided for all $w \in C^*$, $u$, $v \in X^*$ the condition $w \cdot u$, $u \cdot v$, $v \cdot u \in C^*$ implies $u$, $v \in C^*$. We obtain our first result.

LEMMA 1. — *Every ifl-code is weakly prefix.*

Proof: Let $C$ be a code not weakly prefix. Then there are a $w \in C^*$ and nonempty words $u$, $v \in X^*$ such that $w \cdot u$, $u \cdot v$, $v \cdot u \in C^*$.

Hence $(w \cdot u) \cdot (v \cdot u) \cdot (v \cdot u) \cdot \ldots$ and $w \cdot (u \cdot v) \cdot (u \cdot v) \cdot \ldots$ are two distinct (infinite) factorizations of the same ω-word, which implies that $C$ is not an ifl-code. □

To show that the converse ist not true we derive the following example.

*Example* 1: $C_1 =_{df} \{1\} \cup \{0^i 10^{i+1} 1 : i \geqq 0\}$.

One easily observes that $\beta =_{df} 1010^2 10^3 1 \ldots$ is the unique $\omega$-word having two distinct factorizations in $C_1$. Since $\beta$ is not ultimately periodic, $C_1$ is weakly prefix, but, clearly, $C_1$ is not an ifl-code. □

Moreover, since every word $w \in C$ takes part in one of the factorizations of $\beta$, any proper subcode of $C_1$ is an ifl-code.

This verifies the following.

LEMMA 2: *The set* $\mathbb{C}_{\mathrm{ifl}} =_{df} \{C : C \subseteqq X^* \setminus \{e\}$ *and* $C$ *is an ifl-code*$\}$ *is not inductive.*

Proof: It suffices to consider a chain $C'_0 \subseteqq C'_1 \ldots \subset C_1$ of proper subcodes of the $C_1$ of Example 1 which satisfy $\bigcup_{i \in N} C'_i = C_1$. □

Lemma 2 leads to the following open problem: Is there for any ifl-code C a maximal ifl-code C′ *containing* $C$?

Since Theorem 6 of [C 2] shows that the set $\mathbb{C}_{wp}$ of all weakly prefix codes $C \subseteqq X^* \setminus \{e\}$ is inductive, every weakly prefix code is contained in a maximal weakly code, and this theorem an our Lemma 2 exhibit a principal difference between weakly prefix and ifl-codes.

Finally, we mention a condition on $C^\omega$ being equivalent to the property of $C$ being an ifl-code.

To this end we introduce some notation. Let $wb$ be the concatenation of $w \in X^*$ and $b \in X^* \cup X^\omega$. This in an obvious way defines a product $W.B$ of subsets $W \subseteqq X^*$ and $B \subseteqq X^* \cup X^\omega$. For the sake of brevity we shall write $w \cdot B$ and $W \cdot b$ instead of $\{w\} \cdot B$ and $W \cdot \{b\}$, as well as $w^*$ instead of $\{w\}^*$.

LEMMA3: *A subset* $C \subseteqq X^*$ *is an ifl-code iff* $w$, $v \in C$ *and* $w \cdot C^\omega \cap v \cdot C^\omega \neq \varnothing$ *imply* $w = v$.

*Proof*: Clearly, the condition is necessary. Now, assume $C$ to be not an ifl-code. Then there is a $\beta \in C^\omega$ having two different representations as an infinite product, i. e.

$$\beta = w_1 \cdot w_2 \cdot \ldots = v_1 \cdot v_2 \cdot \ldots$$

where $\quad w_i, v_i C \qquad$ and $\qquad w_j \neq v_j \quad$ for some $j \in N$.

Take the least such $j \in N$ and consider

$$\eta = w_j \cdot w_{j+1} \cdot \ldots = v_j \cdot v_{j+1} \cdot \ldots$$

Obviously

$$\eta \in w_j \cdot C^\omega \cap v_j \cdot C^\omega \quad \text{where} \quad w_j \neq v_j. \quad \square$$

## 2. DECODING DELAY

Well known is the concept of finite decipherability delay (*cf.* [*L*, C 1]). But since this delay is measured in letters of $X$ as units, this concept would not be appropriate in the case of infinite codes, for they have codewords of arbitrarily large lengths.

Therefore, we introduce another concept as follows:

DEFINITION: A code $C$ is said to have finite decoding delay iff the following condition holds true:

$$\bigwedge_{w \in C} \bigvee_{m} \bigwedge_{v \in C^m} \bigwedge_{\beta \in C^\omega} (w \cdot v \sqsubseteq \beta \Rightarrow \beta \in w \cdot C^\omega) \tag{2}$$

Here we abbreviate by $u \sqsubseteq b$ the fact that $u \in X^*$ is an initial word of $b \in X^* \cup X^\omega$, and $C^m =_{df} \{ w_1 \cdot \ldots \cdot w_m : w_i \in C \}$ is the $m$-fold product of the set $C \subseteq X^*$.

The idea here is to measure the delay in units of codewords, i. e. we can be sure that the first factor of $\beta \in C^\omega$ is $w$ whenever we have an initial word $w \cdot v$ of $\beta$ being a product of $m+1$ codewords ($m$ depending on $w \in C$).

If $m$ is independent of the specific $w \in C$ we obtain the following.

DEFINITION: A code $C$ is said to have a decoding delay of $m$ units iff the following condition holds true

$$\bigwedge_{w \in C} \bigwedge_{v \in C^m} \bigwedge_{\beta \in C^\omega} (w \cdot v \sqsubseteq \beta \rightarrow \beta \in w \cdot C^\omega) \tag{3}$$

We say that *a* code $C$ has bounded decoding delay provided $C$ has a delay of $m$ units for some $m \in N$.

Let $\mathbb{C}_{fd}$, $\mathbb{C}_{bd}$ and $\mathbb{C}_m$ denote the sets of codes $C \subseteq X^* \setminus \{ e \}$ having finite, or bounded decoding delay, or a delay of $m$-units resp.

In particular, $\mathbb{C}_0$ is the class of prefix codes.

PROPERTY 4: (i) $\mathbb{C}_0 \subseteqq \mathbb{C}_1 \subseteqq \mathbb{C}_2 \subseteqq \ldots \subseteqq \mathbb{C}_{bd} = \bigcup_{m \in N} \mathbb{C}_m$

(ii) $\mathbb{C}_{bd} \subseteqq \mathbb{C}_{fd} \subseteqq \mathbb{C}_{\text{ifl}}$.

*Proof*: Lemma 3 ensures that $\mathbb{C}_{fd} \subseteqq \mathbb{C}_{\text{ifl}}$, and the other inclusions are trivial. □

We are not going to prove that the inclusions in (i) are all strong. To this end we refer to the case of finite codes with finite decipherability delay which is widely investigated (*cf.* [L, C 2]). We confine to the inequalities of (ii). Before we are going to show that these inequalities are proper we need some auxiliary considerations.

*Lemma* 5: *Let* $C \subseteqq X^*$. $C \in \mathbb{C}_{fd}$ *iff it holds*

$$\bigwedge_{w \in C} \bigvee_{m} \bigwedge_{v \in C^m} \bigwedge_{w' \in C} \bigwedge_{u \in C^*} (w \cdot v \sqsubseteq w' \cdot u \rightarrow w = w'). \tag{4}$$

*Proof*: Let $C \in \mathbb{C}_{fd}$ and $w \in C$. Choose $m$ depending on $w$ according to Equation (2), and let $v \in C^m$, $w' \in C$ and $u \in C^*$ be such that $w \cdot v \sqsubseteq w' \cdot u$. Then for every $\eta \in C^\omega$ we have $\beta = w' \cdot u \cdot \eta \in C^\omega$ and $w \cdot v \sqsubseteq \beta$.

Consequently, $\beta \in w \cdot C^\omega$, and since $C$ is also an ifl-code, we have $w = w'$.

Conversely, let $C$ satisfy Equation (4) and let $w \in C$. Again, choose $m$ depending on $w$ according to Equation (4), and let $v \in C^m$ and $\beta \in C^\omega$ be such that $w \cdot v \sqsubseteq \beta$. Since $\beta \in C^\omega$, there are $w' \in C$, $u \in C^*$ and $\eta \in C^\omega$ such that $\beta = w' u \eta$ and $w \cdot v \sqsubseteq w' \cdot u$. Now Equation (4) implies $w' = w$, i. e. $\beta \in w \cdot C^\omega$. □

The same proof works in the case of codes having bounded decoding delay i. e. the following lemma is also true.

LEMMA 6: *Let* $C \subseteqq X^*$. $C \in \mathbb{C}_m$ *iff it holds*

$$\bigwedge_{w, w' \in C} \bigwedge_{v \in C^m} \bigwedge_{u \in C^*} (wv \sqsubseteq w' u \rightarrow w = w').$$

These characterizations of the classes $\mathbb{C}_{fd}$ and $\mathbb{C}_m$, or $\mathbb{C}_{bd}$ resp. give rise to the following characterizations of $C^\omega$.

Let $C \in \mathbb{C}_{fd}$. To every $w \in C$ define $m(w) \geqq 1$ as one value satisfying Equation (4) for $w$. Extend the function $m$ to $C^*$ via $m(wv) =_{df} m(w) + m(v)$.

Now we define

$$C_i =_{df} \bigcup_{u \in C^i} u \cdot C^{m(u)} \tag{6}$$

By definition, we have

$$C^{\omega} \subseteqq C_i \cdot X^{\omega} \subseteqq C_{i-1} \cdot X^{\omega},$$

and since $m(w) \geqq 1$ for $w \in C$,

$$C_i \cdot X^{\omega} \subseteqq C \cdot C_{i-1} \cdot X^{\omega}.$$

Let $\beta \in \bigcap_{i \in N} C_i \cdot X^{\omega}$. Then $\beta \in C_1 \cdot X^{\omega}$. Hence, $w \cdot v \sqsubseteq \beta$ where $w \in C$ and $v \in C^{m(w)}$.

Let $i > |wv|$ ($|u|$ denotes the length of the word $u \in X^*$) and $w \cdot v \sqsubseteq w_1 \cdot \ldots \cdot w_i$ where $w_i \in C$. Then Equation (4) implies $w_1 = w$.

Consequently, every $v' \in C_i$, $i > |w \cdot v|$ satisfying $v' \sqsubseteq \beta$ has as its first factor the word $w$, and we obtain from Equations (8) and (7) that

$$\beta \in \bigcap_{i \in N} w \cdot C_{i-1} \cdot X^{\omega} = w \cdot \bigcap_{i \in N} C_{i-1} \cdot X^{\omega}.$$

This shows, that $C \in \mathbb{C}_{fd}$ implies

$$\bigcap_{i \in N} C_i \cdot X^{\omega} \subseteqq C \cdot \bigcap_{i \in N} C_i \cdot X^{\omega}.$$

Applying the following property (*cf.* [S 1, LS]):

$$F \subseteqq X^{\omega}, \qquad L \subseteqq X^* \setminus \{e\} \quad \text{and} \quad F \subseteqq L \cdot F \text{ imply } F \subseteqq L^{\omega}, \tag{9}$$

we obtain together with Equation (7) the following.

THEOREM 7: *If* $C \in \mathbb{C}_{fd}$, *then*

$$C^{\omega} = \bigcap_{i \in N} C_i \cdot X^{\omega}. \tag{10}$$

The above construction also applies to codes having a bounded decoding delay. In that case we can choose $m(w) = m$ for all $w \in C$ and a suitable $m \in N$. Hence $C_i$ becomes $C^{i \cdot (m+1)}$, and we get the following theorem.

THEOREM 8: *If* $C \in \mathbb{C}_{bd}$, *then*

$$C^{\omega} = \bigcap_{i \in N} C^i \cdot X^{\omega}. \tag{11}$$

*Remark*: It was widely believed (*cf.* [S 1, BN, DK]) and utilized in [BN, Property 2 (3)] that Equation (11) holds true for all $e$-free languages

$L \subseteqq X^* \setminus \{e\}$. In the following section we shall show that this is not true even for regular ifl-codes. Moreover, we shall use the above properties of $\mathbb{C}_{fd}$ and $\mathbb{C}_{bd}$ to prove that the inclusions in Property 4 (ii) are all proper.

## 3. TOPOLOGICAL RESULTS

We regard in $X^\omega$ the product topology which is induced by the basis $(w \cdot X^\omega)_{w \in X^*}$.

Consequently, a set $E \subseteqq X^\omega$ is open if and only if there is a language $L \subseteqq X^*$ such that $E = L \cdot X^\omega$. The collection of denumerable intersections of open subsets of $X^\omega$ is known as the family of $G_\delta$-subsets of $X^\omega$. The following characterization of $G_\delta$-subsets by languages is due to Davis [Da].

For $L \subseteqq X^*$ we denote by $L^\delta$ the $\delta$-limit of the language $L$ where

$$L =_{df} \{ \beta : \beta \in X^\omega \text{ and } u \sqsubseteq \beta \text{ for infinitely many } u \in L \}. \tag{12}$$

LEMMA 9: *$F \subseteqq X^\omega$ is a $G_\delta$-set iff*

$$F = L^\delta \quad \textit{for some } L \subseteqq X^*.$$

Above, in Theorem 7 we have shown, that for all codes $C$ having finite decoding delay the set

$$C = \bigcap_{i \in N} C_i \cdot X^\omega$$

is a $G_\delta$-set.

As a first example, we shall show that there is a regular ifl-code $C_2$ such that $C_2$ is not a $G_\delta$-set. Hence $C_2 \notin \mathbb{C}_{fd}$. This example was first obtained (but not published) by K. Wagner who constructed it utilizing his DA-reducibility of $\omega$-languages [W].

*Example* 2: The set $C_2 =_{df} \{0, 10\} \cup \{01, 02\}^* \cdot 020$ is an ifl-code, but $C_2^\omega = (\{01, 02\}^* \cdot 0 \cup \{10\})^\omega$ is not a $G_\delta$-set. The latter fact is proved in [S 3].

That $C_2$ is an ifl-code may be verified directly or via the procedure indicated by Properties 11 and 12 and Lemma 15 below. □

The next example (first obtained by K. Wagner and G. Wechsung, but up to now unpublished) shows that $\mathbb{C}_{bd} \subset \mathbb{C}_{fd}$.

*Example* 3: *Lt* $C_3 =_{df} \{w_1, w_2, \ldots\}$ where

$$w_1 =_{df} 1 \qquad \text{and} \qquad w_{i+1} =_{df} w_i^i \cdot 0^i 1.$$

This code has finite delay (take e. g. $m(w_i) =_{df} |w_{i+1}|$).

On the other hand, for $\beta$ with $\{\beta\} = C_3^\delta$ we have $\beta \notin C_3^\omega$ and $w_{i+1} \sqsubseteq \beta$ for all $i$. Hence $w_i^i \sqsubseteq \beta$ for all $i \in N$.

Consequently, $\beta \in \bigcap_{i \in N} C_3^i \cdot X^\omega$, and $C_3 \notin \mathbb{C}_{bd}$. $\square$

We ask now whether the statements of the Theorems 7 and 8 are reversible. To this end we recall some connections between the $\omega$-power and the $\delta$-limit from [S 2, 3].

Let $L \subseteq X^* \setminus \{e\}$.

$$L \subseteq \bigcap_{i \in N} L^i \cdot X^\omega \subseteq (L^*)^\delta = L^\omega \cup L^* \cdot L^\delta. \tag{13}$$

This yields:

$$L^\omega = (L^*)^\delta \quad \text{iff} \quad L^\omega \supseteq L^\delta. \tag{14}$$

*Example* 4: Let $C_4 =_{df} \{010, 20\} \cup 2(001)^*$.

*Then* $C_4^\delta = \{20(010)(010)\ldots\} \subseteq C_4^\omega$. Hence

$$C_4^\omega = \bigcap_{i \in N} C_4^i \cdot X^\omega = (C_4^*)^\delta.$$

but $C_4 \in \mathbb{C}_{\mathrm{ifl}} \setminus \mathbb{C}_{fd}$, since

$$20(010)^i \sqsubseteq 2 \cdot (001)^{i+1}.$$

That $C_4$ is an ifl-code may be verified directly or as described in Example 2. $\square$

From Equation (13) it follows that for prefix codes $C$ (they satisfy e. g. $C^\delta = \varnothing$) we have

$$C^\omega = \bigcap_{i \in N} C_i^i \cdot X^\omega = (C^*)^\delta,$$

but already in the case $C \in \mathbb{C}_1$ we may have $C^\omega \neq (C^*)^\delta$, as the code $C_5 = 1 \cdot 0^*$ shows.

Thus, the conclusion of this section is that topological methods are helpful in the study of ifl-codes, but do not provide a thorough characterization of the classes of codes considered here:

#### 4. $C$-CHAINS

In this section following an idea of Levenshtejn (*cf.* [L] and also [LS, sec. 2.2.1]) we introduce a relation useful in the study of codes. This same idea reappeared as the concept of $L$-sequences in [C 1, 2].

We set $C/w =_{df} \{v : wv \in C\}$ and we call $C/w$ the state (left derivative) of the set $C \subseteq X^*$ derived by the word $w$. As it is well-known, a subset $C \subseteq X^*$ is regular iff the set $\{C/w : w \in X^*\}$ is finite, i. e. $C$ is finite-state.

For a subset $C \subseteq X^* \setminus \{e\}$ we define Levenshtejn's relation $\prec$ on $X^*$ as follows:

$$w \prec_1 v \quad \Leftrightarrow_{df} \quad v \in C/w$$

$$w \prec_2 v \quad \Leftrightarrow_{df} \quad w \in C \cdot v \qquad \text{and} \qquad \prec =_{df} \prec_1 \cup \prec_2.$$

We consider $C$-chains, i. e. sequences of the form

$$u_1 \prec_{k_1} u_2 \prec_{k_2} u_3 \prec \ldots \prec_{k_{n-1}} u_n$$

where $u_1 \in C$.

By induction one easily proves that any $C$-chain of length $n \geqq 2$ is in one-to-one correspondence to a covering relation:

$$w_1 \cdot \ldots \cdot w_{i-1} \sqsubseteq v_1 \cdot \ldots \cdot v_j \sqsubseteq w_1 \cdot \ldots \cdot w_i$$

where $i+j=n$, $w_k$, $v_k \in C$ and $w_1 \neq v_1$.

Moreover, in this case $v_1 \cdot \ldots \cdot v_j \cdot u_n = w_1 \cdot \ldots \cdot w_i$.

Similar investigations can be found in [L] and [C 1,2].

This observation makes the following equivalences obvious:

PROPERTY 10 (The Sardinas-Patterson Theorem). $C$ is a code iff there is no $C$-chain terminating with the empty word $e$.

PROPERTY 11. $C$ is an ifl-code iff there is no infinite $C$-chain.

In the proof of Theorem 4 of [C 2] it is shown the following.

PROPERTY 12. $C$ is a weakly prefix code iff there is no $C$-chain in which a word occurs twice.

Let $l_c(w)$ denote supremum over the lengthes of all $C$-chains starting with $w$. From the connection of the length $n$ of the $C$-chain and the lengthes $i$ and $j$ of the products in the covering relations above one readily sees that for $m \geqq l_c(w)-1$ one has that $w \cdot v \sqsubseteq w' \cdot u$, where $v \in C^m$, $w' \in C$, $u \in C^*$, implies $w = w'$. This yields the following connection to codes having finite decoding delay.

PROPERTY 13: Let $l_c(w) < \infty$ for any $w \in C$. Then $C$ has finite decoding delay.

On the other hand, if $C$ has a decoding delay of $m$ units, then a covering relation

$$w_1 \cdot \ldots \cdot w_{i-1} \sqsubseteq v_1 \cdot \ldots \cdot v_j \sqsubseteq w_1 \cdot \ldots \cdot w_i$$

is possible only if

$$i-2<m \qquad \text{and} \qquad j-1<m, \quad \text{i.e. it implies } i+j<2 \cdot m+3.$$

Together with the above consideration we obtain the following.

PROPERTY 14: If $C \in \mathbb{C}_m$ then $l_c(w) \leqq 2m+2$ for every $w \in C$, and if $l_c(w) \leqq n$ for every $w \in C$ then $C \in \mathbb{C}_{n-1}$.

## 5. REGULAR CODES

Observing that any word in a $C$-chain is a suffix of a word in $C$, we get from the above properties that a finite code $C$ is weakly prefix iff $C$ is an ifl-code iff $C$ has bounded decoding delay. We are now going to investigate what happens if $C$ is regular but infinite. In the Examples 2 and 4 above we have seen that there are regular codes without finite decoding delay.

LEMMA 15: *Let $C$ be regular. Then $C$ is an ifl-code iff it is a weakly prefix code.*

*Proof*: In virtue of Lemma 1 it remains to show that if $C$ is not an ifl-code it is also not a weakly prefix code.

If C is not an ifl-code then $C$ is not a code (and hence not a weakly prefix code) or otherwise there is an infinite $C$-chain. Since $u \prec_2 v$ implies $|v|<|u|$, this infinite C-chain contains infinitely many pairs $u_i \prec_1 u_{i+1}$ $(i \in M \subseteq N)$.

By definition $u_i \prec_1 u_{i+1}$ is equivalent to $u_{i+1} \in C/u_i$. Since $C$ is regular, there are only finitely many distinct $C/u_i$.

Hence, there is a subchain

$$u_1 \prec \ldots \prec u_i \prec_1 u_{i+1} \prec \ldots \prec u_j$$

such that

$$j \geqq i+1 \qquad \text{and} \qquad C/u_i = C/u_j.$$

Therefore we can continue

$$u_1 \prec \ldots \prec u_i \prec u_{i+1} \prec \ldots \prec u_j \prec u_{i+1},$$

and, following Property 12, $C$ is not weakly prefix. □

LEMMA 16: *Let $C$ be regular. Then $C$ has finite decoding delay iff $C$ has bounded decoding delay.*

*Proof*: We show that if $l_c(w)$ is an unbounded function on $C$ then it takes on the value $\infty$.

To this end let $(w_i)_{i \in N}$ be a family of words in $C$ such that there is a $C$-chain starting with $w_i$ and having length $i$. Without loss of generality we may assume the chain to start with the relation $\prec_1$, for otherwise we replace $w_i$ by $w_i'$ where $w_i'$ is defined by

$$w_i = w_i' \cdot u_i \qquad \text{when} \quad w_i \prec_2 u_i.$$

Again $w_i \prec_1 u_i$ iff $u_i \in C/w_i$.

Since $C$ is regular, there are only finitely many $C/w_i$, and we conclude that there is a $w_i$ such that $C/w_i = C/w_j$ for infinitely many $j$. Then there are $C$-chain starting with $w_i$ and having length $j$ for any $j$ such that $C/w_i = C/w_j$, i. e. $l_c(w_i) = \infty$. □

## ACKNOWLEDGMENTS

Im am grateful to Prof. Dr. K. Wagner and Prof. Dr. G. Wechsung for their help in finding suggestive examples, and to Prof. Dr. J. Berstel for his encouraging discussion during the FCT'85 conference in Cottbus.

## REFERENCES

BN. L. BOASSON and M. NIVAT, *Adherences of languages*, J. Comput. System Sci., Vol. 20, 1980, pp. 285-309.

C 1. R. M. CAPOCELLI, *A note on uniquely decipherable codes*, IEEE Trans. Information Theory, vol. IT-25, 1979, pp. 90-94.

C 2. R. M. CAPOCELLI, *On weakly prefix subsemigroups of a free semigroup* in: *Advances in Communications*, D. G. LAINIOTIS and N. S. TSANNES, Eds, D. Reidel, Dordrecht 1980, pp. 123-129.

Da. M. DAVIS, *Infinite games of perfect Information*, in: *Advances in Game Theory*. Ann. of Mathematical Studies, No. 52, 1964, pp. 85-101.

D 1. Do L. V., *Codes avec des mots infinis*. RAIRO Inform. Theor., Vol. 16, 1982, pp. 371-386.

D 2. Do L. V., *Sous-monoides et codes avec des mots infinis*, Semigroup Forum, Vol. 26, 1983, pp. 75-87.

D 3. Do L. V., *Contribution to combinatorics on words*, Diss. B. Humboldt-Univ., Berlin, 1985.

DK. P. DARONDEAU and L. KOTT, *A formal proof system for infinitary rational expressions* in: *Automata on Infinite Words*, M. NIVAT and D. PERRIN Eds, Lecture Notes in Computer Science, No. 192, Springer-Verlag, Berlin 1985, pp. 68-80.

L. V. I. LEVENSHTEIN, *Some properties of coding and self adjusting automata for decoding messages*, Problemy Kibernetiki, Vol. 11, 1964, pp. 63-121.

LS. R. LINDER und L. STAIGER, *Algebraische Codierungstheorie-Theorie der sequentiellen Codierungen, Akademie-Verlag, Berlin* 1977.

Sc. N. P. SCHÜTZENBERGER, *On a question concerning certain free submonoids*, J. Combinatorial Theory, Vol. 1, 1966, pp. 437-442.

Sh. St J. SHYR, *Free Monoids and Languages*, Lecture Notes, Soochow Univ., Taipei, 1979.

S 1. L. STAIGER, *Eine Bemerkung zur Charakterisierung von Folgenmengen durch Wortmengen*, Elektron. Inf. verarb. Kybernetik, Vol. 8, 1972, pp. 589-592.

S 2. L. STAIGER, *Zur Topologie der regulären Mengen*, Diss. A., Friedrich-Schiller-Univ., Jena, 1976.

S 3. L. STAIGER, *A note on connected ω-languages*, Elektron. Inf. verarb. Kybernetik, Vol. 16, 1980, pp. 245-251.

W. K. WAGNER, *On ω-regular sets*, Information and Control, Vol. 43, 1979, pp. 123-177.