CLELIA DE FELICE
ANTONIO RESTIVO

Some results on finite maximal codes

<http://www.numdam.org/item?id=ITA_1985__19_4_383_0>

SOME RESULTS
ON FINITE MAXIMAL CODES (*)

by Clelia DE FELICE ([1]) and Antonio RESTIVO ([2])

Communicated by J.-E. PIN

Abstract. — *In the free monoid $\{a, b\}^*$, we give a definition of maximality and completness of a code with respect to the set T_k of all words containing at most k occurrences of b. We show that the intersection of a finite maximal code with T_k is maximal with respect to T_k, for all k. We derive some recessary conditions for a finite code to have a finite completion end we prove for these codes a "local" version of a theorem of Schützenberger.*

Résumé. — *Nous donnons une notion de maximalité et de complétion d'un code dans l'ensemble T_k des mots sur $\{a, b\}^*$ ayant au plus k occurrences de la lettre b. Ces définitions permettent de montrer que l'intersection de tout code maximal fini avec T_k est T_k-maximal pour tout k. De plus nous obtenons une condition suffisante pour qu'un code fini n'ait pas une complétion finie. Nous montrons également que, pour les codes ayant une complétion finie, on a une version locale d'un théorème de Schützenberger.*

INTRODUCTION

The theory of (variable length) codes, born in the framework of the theory of information transmission with the early works of Shannon, has been developed in an algebraic direction by M.P. Schützenberger and his school since 1956 (*see* [10]) in connection with automata and language theory, combinatorics on words and other related topics in computer science. A complete treatment of the theory until very recent developments may be found in [1].

An important role in this theory is played by the notion of maximal code. A code is maximal if it is not a proper subset of any other code on the same alphabet. A fundamental result of Schützenberger states the equivalence, in particular for finite codes, of the algebraic notion of maximality and the combinatorial notion of completness.

In this paper we introduce the notions of maximality and completness of a code with respect to some particular subsets T_k of the free monoid $\{a, b\}^*$. T_k is defined, for any positive integer k, as the set of words in which the number of occurrences of b, is less thar or equal to k. Codes in T_k generalize codes in the "triangle" considered in [2], [4], [7], [8], [9] and [11]: the latest indeed corresponds to the case $k = 1$.

Then we prove a "local" version of the Schützenberger's theorem (see [3]) which states that, for a finite code in T_k, completness implies "locally" maximality and that the converse is true only if the code is contained in a finite maximal code, i. e. if it has a finite completion. This result gives some useful informations on the structure of the words of a finite maximal code. In particular we show that the intersection of a finite maximal code with T_k is maximal with respect to T_k, for all k.

From another point of view, the previous result gives some necessary conditions for a finite code to have a finite completion. From this, one can derives some useful ideas to approach the following open problem: find a procedure to decide whether a finite code has a finite completion. In particular it is not yet known whether the code, recently constructed by Peter Shor (see [11]), as a counterexample to the "triangle conjecture", has a finite completion: the relevance of this question is related to the validity of Schützenberger's conjecture on the commutative equivalence of a finite maximal code to a prefix one (see [6]). This problem was the starting motivation of our investigation.

In the last section the case $k = 1$ of codes in the "triangle" is considered. Some of the theorems of previous section are strengthened, new methods to construct codes having no finite completion are investigated and some result related to the "triangle conjecture" are proved. Finally some unanswered questions are proposed.

1. THE CASE T_k

Let A be a finite alphabet and A^* the free monoid generated by A. For any word $w \in A^*$ denote by $|w|$ the length of w and, for any letter $a \in A$, denote by $|w|_a$ the number of occurrences of the letter a in w. A subset X of A^* is a *code* if X^* is a free submonoid of A^* of base X.

For any $X \subseteq A^*$ set:

$$R_0 = \{ w \in A^+ : Xw \cap X \neq \varnothing \},$$

$$\forall i \geq 1, \quad R_i = \{ w \in A^+ : R_{i-1} w \cap X \neq \varnothing \text{ or } Xw \cap R_{i-1} \neq \varnothing \}.$$

Then X is a code if and only if for any $i \geq 1$ we have that $X \cap R_i = \varnothing$ [1].

If T is a subset of A^* and $X \subseteq T$ is a code, X is a *T-maximal code* if X is not a proper subset of another code $Y \subseteq T$. If $T = A^*$, X is called a *maximal code*.

Let π be the uniform distribution on A^*, i.e. the application defined as follows:

$$\forall w \in A^*, \quad \pi(w) = \frac{1}{|A|^{|w|}}.$$

One has that:

PROPOSITION 1.1 [1]: *Let X be a recognizable code. X is a maximal code if and only if:*

$$\pi(X) = \sum_{x \in X} \pi(x) = 1.$$

For any pair of subsets P, T of A^* such that $P \subseteq T$, P is dense in T if for any word w of T one has that:

$$T w T \cap P \neq \varnothing.$$

If $T = A^*$ we say that P is dense.

The following foundamental result of M. P. Schützenberger will be useful in the sequel (*see* [3]):

THEOREM 1.1. (Schützenberger): *Let B^* be a free submonoid of A^* and let $X \subseteq B^*$ be a recognizable code.*

If X is a B^*-maximal code then X^* is dense in B^*.

If B^* is finitely generated and if X^* is dense in B^* then X is a B^*-maximal code.

In the sequel the alphabet we consider is the binary alphabet $A = \{a, b\}$. For any positive integer k, we introduce the following subset of A^*:

$$T_k = \{w \in A^* \mid |w|_b \leq k\}$$

and consider codes X which are subsets of T_k. For $k = 1$ we obtain the "triangular" codes studied in [2], [4], [7], [8], [9] and [11].

PROPOSITION 1.2: *Any code $X \subseteq T_k$ is contained in a T_k-maximal code.*

Proof: It suffices to apply Zorn's lemma to the family $\mathscr{F}$:

$$\mathscr{F} = \{Y \subseteq T_k \mid Y \text{ code and } Y \supseteq X\}. \quad \square$$

REMARK 1: However it is not generally true that any *finite* code $X \subseteq T_k$ is contained in a *finite* T_k-maximal code, as shown by the following example (*see* prop. 2.3.):

$$X = \{a^5, a^2 b, ba, b, ba^4 b\} \subseteq T_2.$$

In the special case $k = 1$, we shall prove (see cor. 2.2) that any finite code $X \subseteq T_1$ is contained in a finite T_1-maximal code.

PROPOSITION 1.3: *If $X \subseteq T_k$ is a finite code, then there exists a T_k-maximal code Y such that:*

$$Y \supseteq X, \qquad Y \cap a^* \neq \varnothing.$$

Proof: First of all we prove that there is a finite code $X' \subseteq T_k$ such that:

$$X' \supseteq X, \qquad X' \cap a^* \neq \varnothing.$$

If $X \cap a^* \neq \varnothing$ the result follows. Otherwise let:

$$d = \max \{ |x| \mid x \in X \}.$$

We prove that $X \cup \{ a^{2d-1} \}$ is a code. Suppose that this is not the case.

Then there is a word w of $X \cup \{ a^{2d-1} \}$ with two factorizations in terms of elements of $X \cup \{ a^{2d-1} \}$ and of minimal length:

$$w = x_1 \ldots x_k = y_1 \ldots y_h, \qquad x_i, y_j \in X \cup \{ a^{2d-1} \}.$$

Let $x_t = a^{2d-1}$ be an occurrence of a^{2d-1} in the left side of above equation.

Since $X \cap a^* = \varnothing$, by minimality of w and by definition of d there are $i \in \{ 2, \ldots, h-1 \}$, $z', z'', w', w'' \in A^+$ such that:

$$\left. \begin{array}{cc} x_1 \ldots x_{t-2} w' = y_1 \ldots y_{i-1}, & \\ z' x_t z'' = y_i y_{i+1}, & w' z' = x_{t-1}, \\ w'' x_{t+2} \ldots x_k = y_{i+2} \ldots y_h, & z'' w' = x_{t+1}. \end{array} \right\} \qquad (1)$$

(1) shows that either:

$$y_i = a^{2d-1}$$

or:

$$y_{i+1} = a^{2d-1}$$

[otherwise $y_i, y_{i+1} \in X$ implies $|y_i y_{i+1}| \leq 2 d < |z' a^{2d-1} z''|$ against (1)].

Then $y_i = a^{2d-1}$ (resp. $y_{i+1} = a^{2d-1}$) implies $z' z'' = y_{i+1}$ (resp. $z' z'' = y_i$). Since $z', z'' \in A^+$ and by (1) we have that:

$$y_1 \ldots y_{i-1} y_{i+1} y_{i+2} \ldots y_h$$

$$(\text{resp. } y_1 y_2 \ldots y_{i-1} y_i y_{i+2} \ldots y_h)$$

is a word with two factorizations in terms of elements of $X \cup a^{2d-1}$ in contradiction with the minimality of $|w|$.

Then $X \cup a^{2d-1}$ is a code and the result follows by proposition 1.2. □

By proposition 1.3 one has the following corollary:

COROLLARY 1.1: *If $X \subseteq T_k$ is a finite T_k-maximal code then there exists a positive integer n such that $a^n \in X$.*

Remark 2: Proposition 1.3 is no more true for infinite codes, as shown by the following example:

$$X = \{ a^i b a^i \mid i \geq 0 \}.$$

Since:

$$(a^i b a^j)(a^i b a^j) = (a^i b a^i)(a^j b a^j),$$

X is a T_1-*maximal code but* $X \cap a^* = \varnothing$.

For $k = 1$ a partial converse of this proposition holds:

if $X \subset T_1$ is a T_1-maximal code and $X \cap a^* \neq \varnothing$ then X is a finite code.

For $k > 1$ this is not true. For example the code:

$$X = \{ a^5, a^2 b, ba, b \}$$

is contained in an infinite T_2-maximal code Y (*cf.* prop. 1.2, prop. 2.3) and $a^5 \in Y$.

Let us introduce the following notation: for any pair u, v of words of A^*, if v is factor of u we write $v \leq u$.

For any pair of positive integers k, d consider the following subset of A^*:

$$F_{k,d} = \{ w \in A^* \mid \forall u \leq w, \, |u|_b > k \Rightarrow a^{2d} \leq u \}.$$

Let $M_{k,d} = a^d F_{k,d} a^d \cup \{a\}^*$. It is easy to see that $M_{k,d}$ is a submonoid of A^* and that the set of factors of $M_{k,d}$ coincides with $F_{k,d}$.

For k, d positive integers introduce now the following subset of A^*:

$$B_{k,d} = a \cup a^d \left(\bigcup_{i=0}^{k-1} b (a^{<d} b)^i \right) a^d,$$

where $a^{<d} = 1 \cup a \cup a^2 \dots \cup a^{d-1}$.

LEMMA 1.1: *$M_{k,d}$ is a free submonoid of A^* with base $B_{k,d}$.*

Proof: Since $B_{k,d} \subseteq M_{k,d}$ we have that $B_{k,d}^* \subseteq M_{k,d}$. Let us prove that $M_{k,d} \subseteq B_{k,d}^*$. Any word w of $M_{k,d}$ can be uniquely factorized as follows:

$$w = a^{n_1} v_1 a^{n_2} v_2 a^{n_3} v_3 \dots a^{n_r} v_r a^{n_{r+1}},$$

with:

$$n_1, n_{r+1} \geqq d, \qquad n_2, n_3, \ldots, n_r \geqq 2\,d,$$

$$v_i \in \bigcup_{i=0}^{k-1} b\,(a^{<d}\,b)^i.$$

In other words n_1 and n_{r+1} indicate the numbers of consecutive a's at the beginning and at the end of w respectively; the elements a^{n_i} $(2 \leqq i \leqq r)$ indicate the occurrences in w of factors belonging to $a^{2d}\,a^*$ between two consecutive occurrences of the letter b.

By this factorization we easily obtain an unique factorization of w in elements of $B_{k,\,d}$. $\square$

Consider now a finite code $X \subset T_k$ and let $d = \max\{\,|x|\,|\,x \in X\,\}$.

Consider the intersection:

$$Y^* = X^* \cap M_{k,\,d}.$$

Y^* is a free submonoid and Y is a recognizable code (indeed X^* and $M_{k,\,d}$ are recognizable subsets of A^*).

REMARK 3: We have that $Y \neq \varnothing$ if and only if $X \cap a^* \neq \varnothing$.

Indeed suppose that $Y \neq \varnothing$. Since:

$$Y \subseteq Y^* \subseteq M_{k,\,d},$$

if $w \in Y$ either $w \in a^*$ or $w = a^d\,w'\,a^d$, $w' \in F_{k,\,d}$.

Since:

$$Y \subseteq Y^+ \subseteq X^+ \tag{1}$$

and by definition of d we have that $X \cap a^* \neq \varnothing$.

Vice versa, since $a^* \subseteq M_{k,\,d}$, if $X \cap a^* \neq \varnothing$ we have that $X^+ \cap M_{k,\,d} = Y^+ \neq \varnothing$. Moreover, if $Y \neq \varnothing$, we have that:

$$X \cap a^* = Y \cap a^*.$$

Indeed if $Y \neq \varnothing$ let n be the integer such that $a^n \in X$.

Since $a^n \in M_{k,\,d}$ we have $a^n \in Y^+$.

Then there is $t \in N$, $t \leqq n$ such that $a^t \in Y$.

By (1) we have that $a^t \in X^+$.

Then:

$$a^n \in X, \qquad a^t \in X^+, \qquad t \leqq n \quad \Rightarrow \quad t = n.$$

Suppose that $a^n = X \cap a^* = Y \cap a^*$. We have:

LEMMA 1.2: *If Y is a $M_{k,d}$-maximal code, then X is a T_k-maximal code.*

Proof: The proof is by contradiction. Suppose that X is not a T_k-maximal code. Then there exists a word $x \in T_k$ such that $X \cup \{x\}$ is a code.

Let:

$$Z^* = (X \cup \{x\})^* \cap M_{k,d}$$

and let Z be the base of Z^*.

Since $a^{nd} x a^{nd} \in Z^* \setminus Y^*$ one has that:

$$Y^* \subseteq Z^* \quad \text{and} \quad Y^* \neq Z^* \tag{1}$$

If we prove that:

$$Y \subseteq Z,$$

then by (1), it follow:

$$Y \subset Z,$$

which is a contradiction.

Since $Y^* \subseteq Z^*$, if $y \in Y$ then $y \in Z^* \setminus 1$. Then there are $r \in N$, $z_1, \ldots, z_r \in Z \setminus 1$ such that:

$$y = z_1 \cdot \ldots \cdot z_r. \tag{2}$$

By definition of Z^* and Y^* one has that:

$$y \in X^+, \quad z_i \in (X \cup \{x\})^+, \quad z_i \in M_{k,d}, \quad i \in \{1, \ldots, r\}.$$

Then there are $x_1, \ldots, x_n \in X$; $x_1^1, \ldots, x_{k_1}^1, \ldots, x_1^r, \ldots, x_{k_r}^r \in X \cup \{x\}$, such that:

$$x_1 \cdot \ldots \cdot x_n = y,$$

$$\forall i \in \{1, \ldots, r\} \quad x_1^i \cdot \ldots \cdot x_{k_i}^i = z_i.$$

Then:

$$x_1 \cdot \ldots \cdot x_n = x_1^1 \cdot \ldots \cdot x_{k_1}^1 \cdot \ldots \cdot x_1^r \cdot \ldots \cdot x_{k_r}^r. \tag{3}$$

Since $X \cup \{u\}$ is a code and in the left of (3) there is no occurrence of x one has that:

$$\forall i \in \{1, \ldots, r\}, \quad \forall j \in \{1, \ldots, k_i\}, \quad x_j^i \in X \implies \forall i \in \{1, \ldots, r\},$$

$$z_i \in X^+ \cap M_{k,d} \implies \forall i \in \{1, \ldots, r\}, \quad z_i \in Y^* \setminus 1.$$

Then by (2) we have that:

$$y = z_1 \ldots \ldots z_r \in Y^+ \cap Y,$$

which implies $r = 1$ i. e. $y \in Z$. □

Introduce the following definition. A code $X \subset T_k$ is T_k-complete if $Y^* = X^* \cap M_{k, d}$ is dense in $M_{k, d}$, i. e. if:

$$\forall \, w \in M_{k, d}, \qquad M_{k, d} \, w \, M_{k, d} \cap Y^* \neq \varnothing.$$

By theorem 1.1 and lemma 1.2, we obtain the following theorem:

THEOREM 1.2: *Let $X \subset T_k$ be a finite code. If X is T_k-complete then X is T_k-maximal.*

REMARK 4: The converse of theorem 1.2 does not generally hold, as shown by the following example:

$$X = \{ a^5, \, a^2 \, ba, \, a^2 \, b, \, ba, \, b \},$$

X is a T_1-maximal code, but it is not T_1-complete (*see* prop. 2.2).

By remark 4 we see that the relationship between the notions of maximality and completness, as stated by theorem 1.1 for finite (and recognizable) codes in A^*, does not hold "locally" in T_k. As we shall see this is a consequence of the fact that a finite code is not generally contained in a finite maximal code.

REMARK 5: Let X be a T_k-complete finite code.

Then, by theorem 1.2 and corollary 1.1, there is $n \in N$ such that $a^n \in X$.

Moreover for any $w = a^{n_1} ba^{n_2} b \ldots \ldots ba^{n_s} \in M_{k, d}$ there are $w_1, w_2 \in M_{k, d}$ such that:

$$w_1 \, w w_2 \in X^*.$$

Since $n_1 \geq d$, $n_s \geq d$, by definition of d there are $t \leq n_1$, $q \leq n_s$ such that:

$$a^t ba^{n_2} b \ldots \ldots ba^{n_s - 1} ba^q \in X^*.$$

This equation implies $a^* \, wa^* \cap X^* \neq \varnothing$.

Vice versa, since $a^* \subseteq M_{k, d}$, if this condition holds for any $w \in M_{k, d}$, we have that X is a T_k-complete code.

Then:

$$X \text{ is a } T_k\text{-complete code} \quad \Leftrightarrow \quad \forall \, w \in M_{k, d},$$
$$a^* \, wa^* \cap X^* \neq \varnothing.$$

Consider now a finite maximal code C in A^*. We prove the following lemma:

LEMMA 1.3: *Let $C \subset A^*$ be a finite maximal code. Then $X = C \cap T_k$ is T_k-complete.*

Proof: Let d be the maximal length of words in C and let w be a word of $M_{k,d}$. By remark 5 it suffices to prove that:

$$a^* w a^* \cap X^* \neq \varnothing. \tag{1}$$

Since $X \cap a^* = C \cap a^*$, if $w \in a^*$ then (1) is verified. Suppose that:

$$w = a^d w' a^d, \qquad w' \in F_{k,d}.$$

By theorem 1.1 there are $u_1, u_2 \in A^*$ such that:

$$u_1 w u_2 \in C^*$$

and, by the definition of d, there are $r, s \in N$, $r, s < d$, such that:

$$a^r w' a^s \in C^*.$$

By definition of d and of $F_{k,d}$ we have that:

$$a^r w' a^s \in X^*. \tag{2}$$

Let t be an integer such that $tn > d$. By (2) we have:

$$a^{tn} a^r w' a^s a^{tn} \in X^* \cap a^* w a^*. \quad \square$$

As a consequence of theorem 1.2 and lemma 1.3 we obtain the following theorem:

THEOREM 1.3: *If $C \subset A^*$ is a finite maximal code, then $C \cap T_k$ is a T_k-maximal code.*

REMARK 6: The condition stated in theorem 1.3 does not hold in general if one consider, instead of T_k, an arbitrary subset T of A^*. This is shown by the following elementary example. Let T be the set of words of A^* of length pair:

$$T = \{ w \in A^* \mid |w| = 2n, \ n \geq 0 \}.$$

Let $C = \{ aa, ab, b \}$. C is a finite maximal code, but $C \cap T = \{ aa, ab \}$ is not a T-maximal code.

By theorem 1.3 we may deduce some interesting consequences.

COROLLARY 1.2: *If C and C' are two finite maximal codes on the alphabet A which differ for only one word, then C and C' are commutatively equivalent.*

Proof: Let u and u' be words of C and C' respectively such that $C' = (C \setminus \{u\}) \cup \{u'\}$. By well known arguments concerning the mesure of a maximal code (*see* prop. 1.1) we have:

$$|u| = |u'|.$$

Let $k = |u|_b$. By theorem 1.3, $C \cap T_{k-1}$ is a T_{k-1}-maximal code. Hence:

$$|u'|_b \geq k = |u|_b.$$

By changing a for b, the same argument gives $|u'|_a \geq |u|_a$. By these two inequalities and by the condition $|u| = |u'|$, we decuce that $|u'|_a = |u|_a$ and $|u'|_b = |u|_b$, i.e. u and u' are commutatively equivalent. $\square$

COROLLARY 1.3: *Let $X \subset T_k$ be a code that is contained in a finite maximal code in A^*. Then X is T_k-maximal if and only if X is T_k-complete.*

Proof: If C is a finite maximal code which contains X, one has $X \subset C \cap T_k$.

If X is T_k-maximal, then $X = C \cap T_k$. The proof is then obtained as a consequence of lemma 1.3. $\square$

Corollary 1.3 gives a general condition under which the relationship between the notions of maximality and completness of a code, as stated in theorem 1.1 holds "locally" in T_k.

Moreover it gives a necessary condition for a finite code to have finite completion. Indeed it can be formulated in the following way:

COROLLARY 1.3: *Let X be a T_k-maximal code. If X is not T_k-complete then it has no finite completions.*

This result allows us to construct in a simple way finite codes having no finite completions.

2. THE CASE T_1

In this section we consider codes $X \subset T_1$ i.e. such that $X \subset a^* \cup a^* b a^*$. We prove some results which are not true for codes contained in T_k [prop. 2.1, cor. 2.1, cor. 2.2] and we construct finite codes having no finite completions [prop. 2.2, prop. 2.5, cor. 2.3].

Finally proposition 2.6 and corollary 2.4 concern the triangle conjecture formulated by Perrin and Schützenberger [7]:

Triangle conjecture: Let $X \subset T_1$ be a code having finite completions. Set:

$$d = \max \{ |x| \mid x \in X \},$$

then card $(X) \leq d$.

Some partial results on this conjecture can be found in [2], [4], [7], [8] and [11].

PROPOSITION 2.1: *Let* $X \subset T_1$ *be such that* $X \cup a^n$ *is a code. Then card* $(X) \leq n$.

Proof: For all $r \in N$ and $x \in X^r$:

$$x = a^{i_1} ba^{j_1} a^{i_2} ba^{j_2} \ldots \ldots ba^{j_{r-1}} a^{i_r} ba^{j_r},$$

let us consider a map φ_r defined as follows:

$$\varphi_r: \quad t(x) = ((i_1, j_1), \ldots, (i_r, j_r)) \to (i_1, j_1 + i_2, \ldots, i_r + j_{r-1}, j_r) \pmod{n}.$$

One has that if X is a code then φ_r is an injective map. Then one has that:

$$[\operatorname{card}(X)]^r = \operatorname{card}(\{ t(x) \mid x \in X^r \}) \leq n^{r+1}.$$

By the foregoing inequality the result follows. $\quad \square$

COROLLARY 2.1: *Let* $Y \subset T_1$ *be a maximal code. Y has finite cardinality if and only if* $Y \cap a^* \neq \varnothing$.

Proof: Sufficiency of the statement follows by proposition 2.1, necessity by corollary 1.1. $\quad \square$

COROLLARY 2.2: *Every finite code* $X \subset T_1$ *is contained in a finite* T_1-*maximal code.*

Proof: X is contained in a T_1-maximal code Y such that $Y \cap a^* \neq \varnothing$ by proposition 1.3. The results follows by corollary 2.1.

Let p be a number greater than 3 (H, K) a factorization of $\{ 0, 1, \ldots, p-2 \}$ [i.e. for any element z of $\{ 0, 1, \ldots, p-2 \}$ there is an unique pair (h, k) of $H \times K$ such that $h + k = z$] with $H, K \neq \{ 0 \}$. Let X be the code:

$$X = a^p + a^H ba^K = \{ a^p \} \cup \{ a^h ba^k \mid (h, k) \in H \times K \}.$$

If p is a prime number then X belongs to the family of codes of Restivo [9].

REMARK 7 [5]: Since (H, K) is a factorization of $\{ 0, 1, \ldots, p-2 \}$ with $p \geq 3$ and $H, K \neq \{ 0 \}$ then there is $t \in \{ 2, \ldots, p-2 \}$ such that either:

$$H \supseteq \{ 0, 1, \ldots, t-1 \}; \quad \{ 0, t \} \subseteq K, \tag{1}$$

or:

$$K \supseteq \{ 0, 1, \ldots, t-1 \}; \quad \{ 0, t \} \subseteq H \tag{2}$$

PROPOSITION 2.2: *X is a finite T_1-maximal code which is not T_1-complete.* (*Therefore X has not finite completions.*)

Proof: First of all X is not a T_1-complete code. Otherwise, by remark 5, for any integer q such that $q \equiv p - 1 \pmod{p}$ we have that:

$$a^* ba^q ba^* \cap X^* \neq \varnothing.$$

Then there are i, j, t_1, $t_2 \in N$ such that $i + j \equiv p - 1 \pmod{p}$ and $a^{t_1} ba^i$, $a^j ba^{t_2} \in X$, against the definition of X.

Let us prove that X is a maximal code.

For all, $(i, j) \in N^2$ let i_1, i_2 be the integers such that:

$$i = mp + i_1,$$

$$j = np + i_2 \qquad 0 \leq i_1, i_2 < p.$$

One has the following four cases:

(I) $i_1, i_2 < p - 1$;

(II) $i_2 = i_2 = p - 1$;

(III) $i_1 < p - 1, i_2 = p - 1$;

(IV) $i_1 = p - 1, i_2 < p - 1$

and, by remark 7, we have either:

$$H \supseteq \{0, \ldots, t - 1\}; \qquad \{0, t\} \subseteq K, \tag{1}$$

or:

$$K \supseteq \{0, 1, \ldots, t - 1\}; \qquad \{0, t\} \subseteq H. \tag{2}$$

We prove that in each case, if $a^i ba^j \notin X$ then $X \cup a^i ba^j$ is not a code. Indeed there are $h_1, h_2 \in H$, $k_1, k_2 \in K$ such that:
(Case (I) and (III)):

$$i_1 = h_1 + k_1.$$

(Case (I) and (IV)):

$$i_2 = h_2 + k_2$$

and we have that:
(Case I):

$$b(a^i ba^j) b = (ba^{k_1})(a^p)^m a^{h_1} ba^{k_2} (a^p)^n (a^{h_2} b).$$

(Case II + case 1)):

$$ba^t (a^{p-1} ba^{p-1}) ab = ba^p (a^{t-1} b) a^p (b).$$

(Case II + case 2)):

$$ba (a^{p-1} ba^{p-1}) a^t b = b (a^p) (ba^{t-1}) a^p (b).$$

(Case III + case 1)):

$$b (a^i ba^{p-1}) ab = ba^{k_1} (a^{h_1} b) a^p (b).$$

(Case III + case 2)):

$$b (a^i ba^{p-1}) a^t b = ba^{k_1} (a^{h_1} ba^{t-1}) a^p (b).$$

(Case IV + case 1)):

$$ba^t (a^{p-1} ba^j) b = b (a^p) (a^{t-1} ba^{k_2}) (a^{h_2} b).$$

(Case IV + case 2)):

$$ba (a^{p-1} ba^j) b = b (a^p) (ba^{k_2}) (a^{h_2} b).$$

Then $X \cup a^i ba^j$ is not a code. $\quad \square$

Moreover we can prove that X is not contained in any finite T_2-maximal code. In order to prove this statement we need the following preliminar lemma:

LEMMA 2.1: *For all* $i, j, k \in N$ *such that* $j \not\equiv p-1$ *(mod. p)*, $X \cup a^i ba^j ba^k$ *is not a code.*

Proof: If $j \not\equiv p-1$ (mod. p) then there are $h_2 \in H$, $k_2 \in K$, $t_2 \in N$ such that:

$$j = h_2 = k_2 + t_2 p$$

and one of the following conditions is verified:

   (I) $i \not\equiv p-1$ (mod. p), $k \not\equiv p-1$ (mod. p);
   (II) $i' p-1$ (mod. p), $k \equiv p-1$ (mod. p);
   (III) $i \equiv p-1$ (mod. p), $k \not\equiv p-1$ (mod. p);
   (IV) $i \not\equiv p-1$ (mod. p), $k \equiv p-1$ (mod. p).

Moreover, by remark 7, there is $2t \in \{2, \ldots, p-2\}$ such that either:

$$H \supseteq \{0, 1, \ldots, t-1\}, \qquad \{0, t\} \subseteq K, \tag{1}$$

or:

$$K \supseteq \{0, 1, \ldots, t-1\}, \qquad \{0, t\} \subseteq H \tag{2}$$

and there are $h_1, h_3 \in H, k_1, k_3 \in K, t_1, t_3 \in N$ such that:

Case (I) and case (IV):

$$i = h_1 + k_1 + t_1 p,$$

Case (I) and case (III):

$$k = h_3 + k_3 + t_3 p.$$

Then we have:

(Case (I)):

$$b(a^i b a^j b a^k) b = b a^{k_1} (a^p)^{t_1} a^{h_1} b a^{k_2} (a^p)^{t_2} a^{h_2} b a^{k_3} (a^p)^{t_3} (a^{h_3} b).$$

(Case II + case 1)):

$$b a^t (a^{p-1+pt_1} b a^j b a^{p-1+pt_3}) ab = b (a^p)^{1+t_1} a^{t-1} b a^{k_2} (a^p)^{t_2} (a^{h_2} b) (a^p)^{1+t_3} b.$$

(Case (III) + case 1)):

$$b a^t (a^{p-1+pt_1} b a^j b a^k) b = b (a^p)^{1+t_1} a^{t-1} b a^{k_2} (a^p) (a^{h_2} b a^{k_3}) (a^{h_3} b).$$

(Case IV + case 1)):

$$b (a^i b a^j b a^{p-1+pt_3}) ab = b a^{k_1} (a^p)^{t_1} (a^{h_1} b a^{k_2}) (a^p)^{t_2} a^{h_2} b (a^p)^{1+t_3} b.$$

Then in case (1) $X \cup a^i b a^j b a^k$ is not a code.

In case (2) the result follows in a similar way. $\square$

PROPOSITION 2.3: X is not contained in any finite T_2-maximal code.

Proof: The proof is by contradiction. Let Y be a subset of T_2:

$$Y = \{ a^{i_1} b a^{j_1} b a^{k_1}, \ldots, a^{i_n} b a^{j_n} b a^{k_n} \},$$

such that $X \cup Y$ is a T_2-maximal code.

Set:

$$w = a^i b a^j b a^k$$

$$i, j, k \in N; \qquad i, j, k > \{ j_1, \ldots, j_n \}; \qquad j < i,$$

$$i \equiv i_1 \quad (\text{mod. } p),$$

$$j \equiv p - 1 \quad (\text{mod. } p),$$

$$k \equiv k_1 \quad (\text{mod. } p).$$

Since $X \cup Y \cup w$ is not a code there is a word $z \in (X \cup Y \cup w)^* \setminus 1$ with two factorizations in terms of the elements of $X \cup Y \cup w$ and of minimal length:

$$z = x_1 . \ldots . x_k = y_1 . \ldots . y_h, \qquad x_i, y_i \in X \cup Y \cup w.$$

Let x_t be an occurence of w in the left side of the above equation. We have the following cases:

1. There is $q \in \{1, \ldots, h\}$ such that:

$$y_q = x_t$$

and, if $(t, q) \neq (1, 1)$:

$$x_1 \cdot \ldots \cdot x_{t-1} = y_1 \cdot \ldots \cdot y_{q-1},$$

if $(t, q) \neq (k, h)$:

$$x_{t+1} \cdot \ldots \cdot x_k = y_{q+1} \cdot \ldots \cdot y_h.$$

2. There is $i \in \{1, \ldots, h\}$ such that $y_i \in Y$ and $ba^j b$ is a factor of y_i.
3. There are $q_1, q_2 \in N$ such that:

$$a^{q_1} x_t a^{q_2} \in a^* X^2 a^*.$$

4. There are $q_1, q_2, q_3 \in N$ such that:

$$a^{q_1} x_t a^{q_2} ba^{q_3} \in a^* XY a^* \cup a^* YX a^*.$$

5. There are q_1, q_2, q_3, q_4 in N such that:

$$a^{q_1} ba^{q_4} a^i ba^j ba^k a^{q_2} ba^{q_3} \in a^* Y^2 a^*.$$

6. There are $u_1, u_1' \in (X \cup Y)^+$, $v_1, w_1' \in (X \cup Y \cup w)^*$ and an occurrence of w in the right side of above equation such that:

$$\left. \begin{array}{l} z = u_1 w v_1 = u_1' w v_1', \\ |u_1| < |u_1'| < |u_1| + |w|. \end{array} \right\} \tag{1}$$

We prove that, in each of these cases, we have a contradiction. Indeed the first case contradicts minimality of z and the second case contradicts the hypothesis $j > \{j_i, \ldots, j_n\}$.

Moreover we can not have case (3) since it implies that there are r, s, u in N such that:

$$a^* ba^r \cap X \neq \varnothing, \qquad a^s ba^* \cap X \neq \varnothing$$

and:

$$r + s \equiv j \equiv p - 1 \quad (\text{mod. } p),$$

against the definition of X.

We can not have case (4) and (5) because otherwise there are $r \in N$, $r \leq j$ and $q'_3 q_3$ such that:

$$a^r ba^k a^{q_2} ba^{q'_3} \in Y,$$

with $k > \{j_1, \ldots, j_n\}$ and this is a contradiction.

In case (6) we have that there are λ, λ', $\mu \in A^+$, $|\lambda| < |w|$ such that:

$$u'_1 = u_1 \lambda, \qquad \lambda = a^i ba^j ba^s,$$

$$\lambda\mu = \mu\lambda' = w,$$

$$v_1 = \lambda' v'_1.$$

Then one has that:

$$u'_1 = u_1 \lambda = u_1 a^i ba^j ba^s \in (X \cup Y)^+.$$

Since for any $q \in \{1, \ldots, n\}$ we have that i, $j > j_q$, there are h, m, r, $e \in N$ such that:

$$m \equiv s \quad (\text{mod. } p), \qquad h \equiv i \quad (\text{mod. } p),$$

$$r + e \equiv p - 1 \quad (\text{mod. } p),$$

$$a^h ba^e, \ a^r ba^m \in X,$$

against the definition of X.

Then the result follows. $\square$

One can ask the question whether a finite code which is contained in a finite T_k-complete code, for all k, has a finite completion. We have not an answer to this question, however for the code of next proposition, we have not found a finite completion (which probably does not exist): it is then a candidate for giving a negative answer to the question.

PROPOSITION 2.4: *The code:*

$$X = \{ a^4, b, ba^3, a^2 b, a^2 ba^3 \}$$

is T_1-complete and is contained in a finite T_k-complete code for all $k \in N$. The code:

$$Y = a^4 + a^{(0,2)} b (ab)^* a^{(0,3)}$$

is a completion of X.

Proof: It is straightforward to prove that X is a T_1-complete code. Let R_i be l'ensemble of the i-residues of Y and π the uniform distribution on A^*. We have:

$$R_1(Y) = R_3(Y) = \{ a^3, (ab)^+, (ab)^+ a^3 \},$$

$$\pi(Y) = 1,$$

i.e. Y is a recognizable maximal code (*see* prop. 1.1).

Moreover for all $r \in N$ and for all $h, k \in H \times K = (0,2) \times (0,3)$, $s_1, \ldots, s_r \in N$ one has that:

$$w = a^h b a^{s_1} b \ldots \ldots b a^{s_r} b a^k \in Y^* \tag{1}$$

The proof is by induction. Suppose $r = 1$. If $s_1 = 1$ then $w = a^h b a b a^k \in Y$. Otherwise, there are $(h_1, k_1) \in H \times K$, $q \in N$ such that:

$$s_1 = h_1 + k_1 + 4q.$$

Then:

$$w = a^h b a^{k_1} (a^4)^q a^{h_1} b a^k \in Y^*.$$

Let us suppose that (1) is true for all intergers $r' < r$. If $s_1 = \ldots = s_r = 1$ then $w \in Y$.

Otherwise, set:

$$t = \min \{ j \in \{ 1, \ldots, r \} \mid s_j \neq 1 \},$$

there are $h_t, k_t \in H \times K$ and $q \in N$ such that:

$$s_t = h_t + k_t + 4q.$$

We have:

$$a^h b a^{s_1} b \ldots \ldots b a^{s_t - 1} b a^{k_t}, \, a^{h_t} b a^{s_t + 1} b \ldots \ldots b a^{s_r} b a^k \in Y^*,$$

by the hypothesis of induction. Then $w \in Y^*$.

Set, for all $m \in N$:

$$Y_m = a^4 + \sum_{q=0}^{m-1} a^{(0,2)} b (ab)^q a^{(0,3)},$$

Y_m is a T_m-complete code which contains X. In facts, let $n_1, \ldots, n_s \in N$ such that:

$$n_1, n_s \geq d$$

and there are at most m consecutive $n_i' s$ such that $n_i < 2d$.

Let w be the word:

$$w = a^{n_1} b a^{n_2} b \ldots \ldots b a^{n_s}.$$

By remark 5 we have to prove that:

$$a^* w a^* \cap Y_m^* \neq \varnothing. \tag{2}$$

By (1) we have that:

$$a^h ba^{n_2} b. \ldots . ba^{n_s-1} ba^k \in Y_m^*,$$

that implies (2). $\quad\square$

Next propositions give other methods to construct codes having no finite completions.

PROPOSITION 2.5: *Let X be a code such that $a^n \in X$, and let $T = a^* b \cup ba^* \cup a^*$, $Z = X \cap T$. If Z satisfies the following conditions:*

1. $b \in Z$;

2. Z *is a T-maximal code;*

3. $\forall k > 0$, $\exists r \geq k$ *such that $ba^r b \notin Z^*$.*

Then X has not finite completions.

REMARK: Set $T' = b^* a^* \cup a^* b^* \cup a^*$, by 1 and 2 it follows that Z is a T'-maximal code.

Proof: The proof is by contradiction. Let Y be a finite completion of X and let λ be the maximal length of the words of Y.

By (3) there is $r \geq k = \lambda + 1$ such that:

$$ba^r b \notin Z^*. \tag{I}$$

Moreover, since the word:

$$b^\lambda a^r b^\lambda$$

is factor of some word of Y^*, there are $t, s, k, h \geq 0$ such that:

$$b^k a^t, a^s b^h \in Y, \qquad t + s \equiv r \quad (\text{mod. } n)$$

and, by (I), one has that:

$$\{ b^k a^t, a^s b^h \} \notin Z.$$

(Otherwise, since $Z \subseteq T$, we have that $k = h = 1$ and $ba^{t+s} b \in Z^2$).

Since Z is a T'-maximal code the result follows. $\quad\square$

COROLLARY 2.3: *Let X be a finite code such that $a^n \in X$ and let $T = a^* \cup ba^* \cup a^* b$, $Z = X \cap T$. If the following conditions are verified:*

1. *There exists $t \in N$, $0 < t \leq n - 1$ such that:*

$$\{ r + s \,|\, ba^r, a^s b \in X \} = \{ 0, 1, \ldots, n-1 \} \setminus t.$$

2. Z *is neither a prefix code nor a suffix code then X has not finite completions.*

Proof: For all $k > 0$ let q be an integer such that:

$$r = t + nq > \{ k, l(X) \}.$$

By hypothesis one has that:

$$b \in Z, \qquad ba^r b \notin Z^*.$$

If we prove that X is a T-maximal code, then the result follows by proposition 2.5.

For all i, $q \in N$, $q < n$ such that:

$$ba^{q + in} \notin Z,$$

one has two cases:

$$q \neq t,$$
$$q = t.$$

In the first case there are r, $s \in N$ such that:

$$ba^r, \ a^s b \in X; \qquad r + s = q.$$

Then:

$$(ba^{q + in}) b = (ba^r)(a^n)^i (a^s b). \tag{I}$$

In the second case we remark that:

$$\forall v \in N \setminus 0, \qquad a^v b \notin Z \quad \Rightarrow \quad Z \subseteq ba^*.$$

Then by 2 there is $v \in N$, $v \neq 0$ such that $a^v b \in Z$.

By 1, since $Z \subseteq X$, we have that $0 < v < n$ and:

$$q + v = t + v \neq t.$$

Then there are r, $s \in N$ such that:

$$ba^r, \ a^s b \in X$$

and either:

$$r + s = t + v,$$

or:

$$r + s + n = t + v.$$

By these equalities one has either:

$$(ba^{q+in})\, a^v\, b = (ba^r)\, (a^n)^i\, (a^s\, b), \tag{II}$$

or:

$$(ba^{q+in})\, a^v\, b = (ba^r)\, (a^n)^{i+1}\, (a^s\, b). \tag{II'}$$

By (I), (II), (II') one has that $Z \cup ba^{q+in}$ is not a code. Similarly if $a^{q+in}\, b \notin Z$ then $Z \cup a^{q+in}\, b$ is not a code. $\square$

We come now to the triangle conjecture. By proposition 2.1 one has that:

PROPOSITION 2.6: *Let $X \subset T_1$ be a code such that $a^d \in X$. Then X verifies the triangle conjecture.*

PROPOSITION 2.7: *Let $X \subset T_1$ be a code T_1-complete such that $X \cup a^n$ is a code. Then $d \geq n$.*

Proof: By contradiction let $d < n$. Let $t \in N$ be such that:

$$r = n - 1 + tn > d.$$

Set $w = (ba^r)^{n+1}$, since X is a code T_1-complete, by remark 5 there are $e, s \in N$ such that:

$$a^e\, wa^s \in X^*.$$

Then there are $(i_1, j_1), \ldots, (i_n, j_n)$ such that:

$$\left.\begin{array}{l} j_q + i_{q+1} = n - 1, \\ a^{i_q}\, ba^{j_q} \in X, \qquad q = 1, \ldots, n-1. \end{array}\right\} \tag{1}$$

Moreover by the hypothesis:

$$i_q + j_q < n - 1. \tag{2}$$

By (1) and (2) it follows:

$$i_{q+1} \geq i_q + 1.$$

By this inequality, since $i_2 \geq 1$, it follows:

$$i_n \geq n - 1,$$

against the hypothesis $d < n$. $\square$

By proposition 2.1 and 2.7 one has that:

COROLLARY 2.4: *Let $X \subset T_1$ be a code T_1-complete. Then $\operatorname{card}(X) \leq d$.*

We conclude with some open problems. A first question is the one posed before proposition 2.4, i.e. whether a finite code, which for all k is contained in a finite T_k-complete code, has a finite completion. In order to answer negatively to this question one has to prove that the code in proposition 2.4 has no finite completions.

This answer should give also a negative solution to the problem whether the necessary condition of corollary 1.2 is also a sufficient one. If this were the case, one should have two different « types » of finitely uncompletable codes: codes finitely T_k-completable for any k and codes having no finite T_k-completions for some k.

Another open problem is whether there exist codes $X \subset T_1$, with $X \cap a^* = \varnothing$, which are not contained in a finite maximal code. The examples reported in this paper are such that $X \cap a^* \neq \varnothing$. Moreover in these examples if $a^n \in X$, one can always find $n' \neq n$ such that $(X \setminus a^n) \cup \{a^{n'}\}$ has a finite completion.

REFERENCES

1. J. Berstel and D. Perrin, *Théorie des Codes*, L.I.T.P., Université VI et VII, 1981.
2. C. De Felice, *On the Triangle Conjecture*, Inf. Proc. Letters, Vol. 14, 1982, pp. 197-200.
3. S. Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press, New York, 1974.
4. G. Hansel, *Baïonnettes et cardinaux*, Discr. Math., Vol. 39, 1982, pp. 331-335.
5. M. Krasner and B. Ranulac, *Sur une propriété des polynômes de la division du cercle*, C.R. Acad. Sc., Paris, T. 240, 1937, pp. 397-399.
6. D. Perrin and M. P. Schützenberger, *Un problème élémentaire de la théorie de l'information*, Coll. Internat. du C.N.R.S., No. 276, Théorie de l'Information, 1977, pp. 249-260.
7. D. Perrin and M. P. Schützenberger, *A Conjecture on Sets of Differences of Integer Pairs*, Bull. E.A.T.C.S. Vol. 5, 1978, pp. 27-29; J. Comb. Theory, Ser B, Vol. 30, 1981, pp. 91-93.
8. J. E. Pin and I. Simon, *A Note on the Triangle Conjecture*, J. Comb. Theory, Ser. A, Vol. 32, 1982, pp. 106-109.
9. A. Restivo, *On Codes Having no Finite Completions*, Discr. Math., Vol. 17, 1977, pp. 309-316.
10. M. P. Schützenberger, *Une théorie algébrique du codage*, C.R. Acad. Sc., Paris, T. 242, 1956, pp. 862-864.
11. P. Shor, *A Counterexample to the Triangle Conjecture*, J. Comb. Theory, Ser. A, Vol. 38, 1985, pp. 110-112.