

GÉRARD JACOB

**Generalized cancellation-and-permutation properties,  
regular languages and supports of rational series**

*RAIRO. Informatique théorique et applications*, tome 22, n° 4 (1988),  
p. 461-486

[http://www.numdam.org/item?id=ITA\\_1988\\_\\_22\\_4\\_461\\_0](http://www.numdam.org/item?id=ITA_1988__22_4_461_0)

© AFCET, 1988, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Informatique théorique et applications » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

## GENERALIZED CANCELLATION-AND-PERMUTATION PROPERTIES, REGULAR LANGUAGES AND SUPPORTS OF RATIONAL SERIES (\*)

by Gérard JACOB <sup>(1)</sup>

Communicated by J. BERSTEL

---

*Abstract.* – *In this paper, we prove two results.*

*The first one is a characterization of regular languages by generalized stability property under cancellation-and-permutation of factors. This result includes as particular cases the characterization of regular languages, of Ehrenfeucht and al. By the "block cancellation property", as well as the characterization of regular languages in the class of periodic languages by the "transposition property", or by the " $\sigma$ -property" given by Restivo and Reutenauer.*

*Secondly, solving an open question of Restivo and Reutenauer, we prove that supports of rational power series satisfy very strong stability property by any finite set of cancellations.*

*Résumé.* – *Dans cet article, nous prouvons deux résultats.*

*Le premier est une caractérisation des langages réguliers par des propriétés généralisées de stabilité par les effacements-avec-permutations de facteurs. Ce résultat contient comme cas particuliers la caractérisation des langages réguliers par la « block cancellation property » due à Ehrenfeucht et al., aussi bien que la caractérisation des langages réguliers dans la classe des langages périodiques par la « transposition property » ou par la «  $\sigma$ -property » donnée par Restivo et Reutenauer.*

*Le second résultat répond à une question ouverte de Restivo et Reutenauer : nous prouvons que les supports des séries formelles rationnelles satisfont une très forte propriété de stabilité pour tout ensemble fini d'effacements.*

### INTRODUCTION

Regular languages satisfy some very strong stability properties with respect to cancellation, or pumping, or permutation of factors in a word.

Indeed, let  $\mathcal{A}$  be a finite deterministic automaton that recognizes a regular language  $L$ . Then, for each  $k$  integer, it is clear that any "long enough"

---

(\*) Received June 1986, revised in March 1988.

(<sup>1</sup>) Université des Sciences et Techniques de Lille Flandres-Artois, U.F.R. d'I.E.E.A., Bât. M3, 59655 Villeneuve-d'Ascq.

transition in  $\mathcal{A}$  passes  $(k+1)$  times through the same state. Thus this implies the “pumping lemma” for regular sets. More generally, this implies that for any word  $w = x u_1 u_2 \dots u_n y$  (where the  $u_i$  are nonempty words, and  $n$  some great enough integer), we can obtain an other factorization  $w = x' v_1 v_2 \dots v_k y'$  where each  $v_j$  is a concatenation of consecutive factors  $u_i$ , and operates a loop on the state  $q_0 * x'$  obtained from initial state  $q_0$  by the transition labelled by the word  $x'$ ). Therefore we obtain that any transformation  $\theta$  of  $w$  by cancellation, or pumping, or permutation of the factors  $v_j$  gives a transformed word  $w \circ \theta$  that satisfies the two symmetric conditions:

$$w \in L \Rightarrow w \circ \theta \in L$$

and

$$w \in CL \Rightarrow w \circ \theta \in CL.$$

Converse results attempt to determine which kind of stability by cancellation, or pumping, or permutation implies regularity. The first converse result has been proved by Ehenfeucht, Parikh and Rozenberg [2] stating that the “block cancellation property”, i. e. stability by cancellation of a single block, implies regularity. The main argument of the proof is a well known theorem of Ramsey [14] (see also [5] theorem 1.7.1). A similar result of Restivo and Reutenauer [18] state that stability by transposition of two consecutive blocks, and more generally by a single permutation, implies also regularity, but only for the *periodic* languages, i. e. for languages having a torsion syntactic monoid. The proof uses conjointly theorem of Ramsey, and a theorem of Shirshov [21, 11].

In this paper, we obtain a general theorem including as particular cases the previous results, in the following way. First we replace a single cancellation or permutation by a “transformation scheme”, i. e. any finite set of transformations. Secondly, instead of the use of one cancellation *or* permutation, we permit any transformation by cancellation *and* permutation, in other words any non increasing transformation.

For a transformation scheme without permutation (that is strictly decreasing) we obtain a characterization of regularity that generalizes the block cancellation property [2].

For a transformation scheme eventually with permutations (but not increasing!) we obtain *for periodic languages* a characterization of regularity that generalizes the transposition property, or the  $\sigma$ -property [18].

Our arguments are, essentially, the same as that of the previous authors: Ramsey's theorem and Shirshov's theorem. It must be noted also that our characterizations are *not symmetric*, and we require a strong property on  $L$  and a weak property on  $CL$  (or the converse). But they become symmetric again in case of transformation scheme reduced to a single transformation.

Finally, we present – as an illustration – a theorem concerning supports of rational power series on a commutative field. We prove that these languages satisfy the same strong stability property by *cancellation* as regular languages. As a particular case, we solve an open question of Restivo and Reutenauer [18]. Consequently, using our main theorem, we obtain: in order to prove that some support  $L$  of rational series is a regular language, it would suffice to prove for  $CL$  a very weak stability property by cancellation.

1. FACTORISATIONS, RAMSEY, SHIRSHOV

Let  $A$  be a fixed finite alphabet. Denote by  $A^*$  the free monoid over  $A$ . The elements of  $A^*$ , also called words over  $A$ , are the finite sequences of letters, i. e. of elements  $A$ . We note  $|w|$  the length of a word  $w$  (i. e. of the sequence defining  $w$ ). The empty word (i. e. the empty sequence) is noted  $\epsilon$ , we note  $A^+$  the set of all the nonempty words over  $A$ .

DEFINITION 1.1: We call  $n$ -factorization ( $n \in \mathbb{N}$ ) over  $A$  any finite sequence  $F = (u, f_1, f_2, \dots, f_n, v)$  of words of  $A^*$ , such that each  $f_i$  is not empty.

Then we set

$$\text{word}(F) = u f_1 f_2 \dots f_n v.$$

Finally, we note  $\Phi_n$  the set of all  $n$ -factorizations.

DEFINITION 1.2: Soit  $F = (u, f_1, f_2, \dots, f_n, v) \in \Phi_n$ .

A  $k$ -factorization  $G$  is said to be *compatible with  $F$*  if and only if there is a sequence of integers

$$1 \leq s_1 < s_2 < \dots < s_{k+1} \leq n + 1$$

that satisfies:

$$u' = u f_1 \dots f_{s_1-1}, \quad v' = f_{s_k+1} \dots f_{n-1} f_n v$$

and

$$g_j = f_{s_j} f_{s_j+1} \dots f_{s_{j+1}-1} \quad \text{for any } j \in \{1, 2, \dots, k\}.$$

Finally, we note  $\Phi_k(F)$  the set of all  $k$ -factorizations compatible with  $F$ . (see Fig. 1).

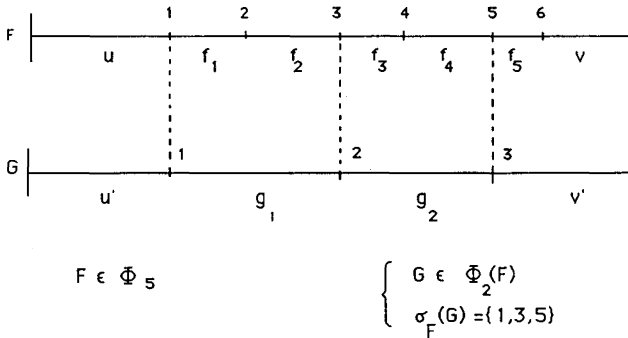


Figure 1. — A factorization  $G$  compatible with  $F$ .

(We adopt the evident generalizations in case of  $1 = s_1$  or  $s_{k+1} = n + 1$ : see example below.)

Observe that  $G \in \Phi_k(F)$  implies:

$$\text{word}(G) = \text{word}(F).$$

So, with the notations of the last definition,  $G \in \Phi_k(F)$  if and only if:

- $\text{word}(F) = \text{word}(G)$ ;
- $u$  is a prefix of  $u'$ ;
- $v$  is a suffix of  $v'$ ;

for any  $j \in \{1, 2, \dots, k\}$ ,  $g_j$  is a concatenation of consecutive factors  $f_i$  of  $F$ . Thus,  $G \in \Phi_k(F)$  if and only if  $w(G) = w(F)$ , and  $F$  is a refinement of  $G$ .

To any word  $w$  of length  $n$

$$w = x_{i_1} x_{i_2} \dots x_{i_n}$$

there is exactly one factorization  $l_w \in \Phi_n$  such that

$$w = \text{word}(l_w).$$

We call it the *literal factorization* of  $w$ :

$$l_w = (\epsilon, x_{i_1}, x_{i_2}, \dots, x_{i_n}, \epsilon).$$

Now, let  $F \in \Phi_n$ , and  $k \leq n$  a positive integer.

Let  $X$  be a finite set. We note  $X[k+1]$  the set of all subsets of  $X$  whose cardinality is  $k+1$ .

Now, according with notations of definition 1.2, we note  $\sigma_F(G) = \{s_1, s_2, \dots, s_{k+1}\} \subset X_F$  the sequence of  $k+1$  integers  $s_j$  that characterizes  $G \in \Phi_n(F)$ .

LEMMA 1.1: *Let  $F \in \Phi_n$ . The map  $\sigma_F$  is a bijection*

$$\sigma_F: \Phi_k(F) \rightarrow X_F[n+1].$$

Furthermore (transitivity of factorizations), if  $H$  is any  $m$ -factorization compatible with  $F$ , and if  $G \in \Phi_k(F)$  for some integer  $k \leq m$ , we have

$$G \in \Phi_k(H) \quad \text{iff} \quad \sigma_F(G) \subset \sigma_F(H).$$

Consequently,  $\sigma_F$  induces a bijection of  $\Phi_k(H)$  onto the set  $\sigma_F(H)[k+1]$ .

*Proof:* Straightforward from the definitions (see Fig. 2).

These remarks allow to translate a combinatorial theorem of Ramsey in the terminology of factorizations.

THEOREM 1 (of Ramsey [14], equivalent translation in factorization's terminology):

*For any natural integers  $k$  and  $N$ , one can compute an integer  $\text{Ram}(k, N)$  that satisfies:*

*For any  $\text{Ram}(k, N)$ -factorization  $F$ , and for any partition of  $\Phi_k(F)$  in two subsets  $B$  and  $C$ , there is a factorization  $G \in \Phi_N(F)$  such that*

- either  $\Phi_k(G) \subset B$ ;*
- or  $\Phi_k(G) \subset C$ .*

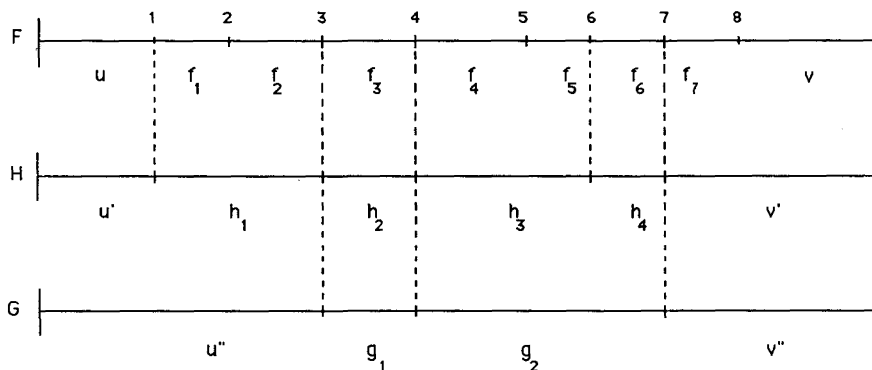
We recall now also, concerning factorizations, a combinatorial theorem of Shirshov [21]. See also Lothaire [11].

DEFINITION 1.3: We call  **$p$ -power of a word**  $h \in A^+$  any  $p$ -factorization of the form

$$H = (x, \underset{\substack{h \\ p \text{ times}}}{h}, h, \dots, h, y).$$

Otherwise, let  $<$  be a total order on  $A$ . We call **lexicographical-by-length order** the order  $<$  over  $A^*$  defined as follows:

- $w < w'$  if and only if
- either  $|w| < |w'|$ ;
- or  $|w| = |w'|$  and  $w$  is smaller than  $w'$  for the lexicographical order.



$$\begin{aligned}
 F \in \Phi_7 & \quad \left\{ \begin{array}{l} H \in \Phi_4(F) \\ \sigma_F(H) = \{1, 3, 4, 6, 7\} \end{array} \right. & \quad \left\{ \begin{array}{l} G \in \Phi_2(F) \\ \sigma_F(G) = \{3, 4, 7\} \end{array} \right. \\
 & & \quad G \in \Phi_2(H) \quad \text{iff} \quad \sigma_F(G) \subset \sigma_F(H)
 \end{aligned}$$

Figure 2. — “Transitivity of factorizations”.

DEFINITION 1.4: We call *k*-division any factorization

$$H = (x, h_1, h_2, \dots, h_k, y)$$

such that for any permutation  $\sigma$  of  $\{1, 2, \dots, k\}$ , we have:

$$x h_{\sigma(1)} h_{\sigma(2)} \dots h_{\sigma(k)} y < x h_1 h_2 \dots h_k y$$

[that will be written in the following section:  $\text{word}(H \circ \sigma) < \text{word}(H)$ ].

THEOREM 2 (Shirshov 1957): Let  $d, p, N \in \mathbb{N} - \{0\}$ . One can compute an integer  $\text{Sh}(d, p, N)$  that satisfies:

For any totally ordered alphabet  $A$ , of cardinality  $d$ , and any word  $w \in A^*$  of length  $|w| = \text{Sh}(d, p, N)$  we have:

- either  $w = \text{word}(H)$  for a  $p$ -power  $H$  of a word  $h$ , with  $1 \leq |h| < N$ ,
- or  $w = \text{word}(G)$  for an  $N$ -division  $G$ .

2. TRANSFORMATION SCHEMES

**Definition and notations**

We call *k*-transformation a couple  $(q, \theta)$  where  $q \in \mathbb{N}$  and  $\theta$  is an injective application not equal to the identity map:

$$\theta: \{1, 2, \dots, q\} \rightarrow \{1, 2, \dots, k\}.$$

We call *k*-transformation-scheme any finite set  $\Theta$  of *k*-transformations.

If  $(q, \theta)$  is a *k*-transformation and  $G = (x, g_1, g_2, \dots, g_k, y)$  any *k*-factorization, we set

$$G \circ \theta = (x, g_{\theta(1)}, g_{\theta(2)}, \dots, g_{\theta(q)}, y)$$

$$\text{word}(G \circ \theta) = x g_{\theta(1)} g_{\theta(2)} \dots g_{\theta(q)} y.$$

*Example 1* (Cancellation of one factor):  $\Theta$  is the unique map  $\theta: \{ \ } \rightarrow \{1\}$

$$g = (u, g_1, v), \quad g \circ \theta = (u, v)$$

$$\text{word}(G) = u g_1 v, \quad \text{word}(G \circ \theta) = u v.$$

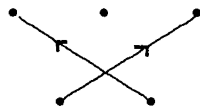
*Example 2* (Cancellation of one factor and permutation of the two others):

$$\theta: \{1, 2\} \rightarrow \{1, 2, 3\} \text{ is defined by } \theta(1) = 3 \text{ and } \theta(2) = 1$$

$$G = (u, g_1, g_2, g_3, v), \quad G \circ \theta = (u, g_3, g_1, v)$$

$$\text{word}(G) = u g_1 g_2 g_3 v, \quad \text{word}(G \circ \theta) = u g_3 g_1 v.$$

(see Fig. 3).



$$G = (u, g_1, g_2, g_3, v)$$

$$\theta = \{3, 3, 1\}$$

$$G \circ \theta = (u, g_3, g_1, v)$$

where, for  $\theta: \{1, 2, \dots, q\} \rightarrow \{1, 2, \dots, k\}$ , we adopt the notation:

$$\theta = \{k; \theta(1), \theta(2), \dots, \theta(q)\}$$

Figure 3. — The 3-transformation  $\theta$  (example 2).

*Example 3* (The set of all  $k$ -cancellations): In the last part of this paper, we shall use the  $k$ -transformation scheme  $C_k$  where  $C_k$  is the set of all strictly increasing map (not surjective) into  $\{1, 2, \dots, k\}$ .

Then  $C_k$  describes all effective cancellations of factors of the  $k$ -factorization. For example, if  $k=2$ , then  $C_2 = \{\theta_1, \theta_2, \theta_3\}$  such that

$$\text{if } G = (u, g_1, g_2, v)$$

$$G \circ \theta_1 = (u, g_1, v)$$

$$G \circ \theta_2 = (u, g_2, v)$$

$$G \circ \theta_3 = (u, v).$$

Observe that  $C_{k+1} \cup \{\text{Id}_{k+1}\}$  is the disjoint union of the two following subsets

$$\{\theta \in C_{k+1} \cup \{\text{Id}_{k+1}\} \mid k+1 \in \text{Im } \theta\}$$

and

$$\{\theta \in C_{k+1} \mid k+1 \notin \text{Im } \theta\}$$

and there is a bijection of each one with  $C_k \cup \{\text{Id}_k\}$ .

*Remark:* Let  $G \in \Theta_k$  and let  $\theta$  be some  $k$ -transformation.

(a) If  $\theta$  is not surjective, then we have

$$|\text{word}(G \circ \theta)| < |\text{word}(G)|.$$

(b) We suppose  $A$  totally ordered, and suppose that  $G$  is compatible with a  $p$ -division  $D$ . Then for any  $\theta$  eventually surjective (but not equal to the identity), we have:

$$\text{word}(G \circ \theta) < \text{word}(G)$$

*Notations 2.2:* Let  $\Theta$  be some  $k$ -transformation scheme. For any  $L \subset A^*$ , let  $E_\Theta(L)$  and  $U_\Theta(L)$  be the two subsets of  $\Phi_k$  defined as follows:

$$G \in E_\Theta(L) \Leftrightarrow \exists \theta \in \Theta, \text{ word}(G \circ \theta) \in L.$$

$$G \in U_\Theta(L) \Leftrightarrow \forall \theta \in \Theta, \text{ word}(G \circ \theta) \in L.$$

It is clear that these two sets coincide if  $\Theta$  is reduced to a single transformation.

Let  $L$  be a language

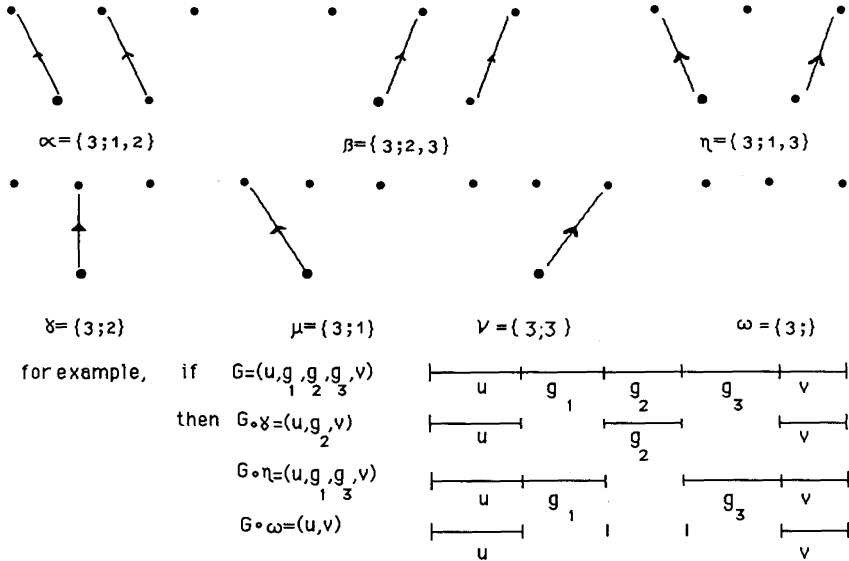


Figure 4. — The 3-transformation scheme  $C_3$  (example 3).

LEMMA 2.1: We have two partitions of  $\Phi$  defined by

$$\Phi_k = E_{\Theta}(L) \sqcup U_{\Theta}(CL)$$

and

$$\Phi_k = U_{\Theta}(L) \sqcup E_{\Theta}(CL).$$

These two partitions will be use later in order to apply the combinatorial Ramsey Theorem.

DEFINITION 2.2: Let  $\Theta$  be some  $k$ -transformation scheme, and let  $N \in \mathbb{N} - \{0\}$ . We say that a language  $L$  is strictly  $\Theta$ -transformable at the order  $N$  is any  $N$ -factorization  $F$  satisfies:

$$\Phi_k(F) \subset E_{\Theta}(CL) \Rightarrow \text{word}(F) \in CL.$$

DEFINITION 2.3: Let  $\Theta$  be some  $k$ -transformation scheme, we say that a language  $L$  is weakly  $\Theta$ -transformable at the order  $N$  if any  $N$ -transformation  $F$  satisfies:

$$\Phi_k(F) \subset U_{\Theta}(CL) \Rightarrow \text{word}(F) \in CL.$$

These two definitions are given in another equivalent form by Ehrenfeucht and al., and Restivo and Reutenauer, as it can be showed by the two following lemmas.

Thus, let  $L$  be a language,  $N$  a natural integer, and  $\Theta$  a  $k$ -transformation scheme.

LEMMA 2.2 (strictly transformable factorizations): *For an  $N$ -factorization  $F$ , the following statements are equivalent*

- (i)  $\Phi_k(F) \subset E_\Theta(CL) \Rightarrow \text{word}(F) \in CL$ ;
- (ii)  $\text{word}(F) \in L \Rightarrow \exists G \in \Phi_k(F), G \in U_\Theta(L)$ ;
- (iii)  $\exists G \in \Phi_k(F), \text{word}(F) \in L \Rightarrow G \in U_\Theta(L)$ .

LEMMA 2.3 (Weakly transformable factorization): *For an  $N$ -factorization  $F$ , the following statements are equivalent*

- (i)  $\Phi_k(F) \subset U_\Theta(CL) \Rightarrow \text{word}(F) \in CL$ ;
- (ii)  $\text{word}(F) \in L \Rightarrow \exists G \in \Phi_k(F), G \in E_\Theta(L)$ ;
- (iii)  $\exists G \in \Phi_k(F), \text{word}(F) \in L \Rightarrow G \in E_\Theta(L)$ .

*Proof:* By the lemma 2.1, these two lemmas are only logical equivalences.

THEOREM 3 (The main result): *Let  $\Theta$  be a  $k$ -transformation scheme, and  $L \subset A^*$ . We suppose satisfied one of the two hypothesis:*

- **either**  $\Theta$  does not contain any surjection;
- **or**  $L$  is a periodic language (see below).

*Then the following two statements are equivalent:*

- (i)  $L$  is a regular language.
- (ii) *There exists an integer  $N \in \mathbb{N} - \{0\}$  such that  $L$  is strictly  $k$ -transformable at the order  $N$ , and such that  $CL$  is weakly  $k$ -transformable at the order  $N$ .*

*Remark:* We can permute  $L$  and  $CL$  in the theorem, because a language  $L$  is regular (resp. periodic) if and only if  $CL$  is regular (resp. periodic).

*Particular cases of theorem 3:* (1) If  $\Theta$  is the cancellation of one factor (example 1), we obtain the characterization of regular languages by the “block cancellation property” of Ehrenfeucht *and al.*, [2].

(2) If  $\Theta = \{\sigma\}$  where  $\sigma$  is a permutation of  $\{1, 2, \dots, k\}$ , we obtain the characterization of the regularity of periodic languages by the “ $\sigma$ -property” of Restivo and Reutenauer [18].

Supports of rational series will give examples where  $\Theta$  is the set  $C_k$  of all the  $k$ -cancellations (Fig. 4) (see example 3). Of course, for  $k \geq 2$ ,  $C_k$  is not reduced to a single transformation. We shall prove that for any  $k \in \mathbb{N}$ , the supports of rational series are strictly  $C_k$ -transformable.

3. PROOF OF THE MAIN THEOREM

(A) *Proof of the direct part ((i) ⇒ (ii))*

Let  $k, N \in \mathbb{N} - \{0\}$ , and  $\Theta$  a  $k$ -transformation scheme, and let  $L$  be a regular language, recognized by a finite deterministic automaton, with  $d$  states.

Let us say that an  $N$ -factorization  $F = (x, f_1, f_2, \dots, f_N, y)$  crosses  $k$  times the same state if there is a state  $q_1$ , and  $k$  integers  $j_i$ , with  $1 \leq j_1 \leq j_2 \leq \dots \leq j_k \leq N$  such that for any  $i \in \{1, 2, \dots, k\}$  we have:

$$q_0 * f_1 f_2 \dots f_{j_i} = q_1$$

(where  $q_0$  is the initial state, and  $*$  denote the right action of the words on the states). Clearly, we have:

LEMMA 3.1: *If  $L$  is recognized by some deterministic automaton with  $d$  states, then any  $k$ .  $d$ -factorization crosses  $k + 1$  times the same state.*

Consequently, if  $F$  is a  $k$ .  $d$ -factorization, the lemma gives some  $k$ -factorization  $G \in \Phi_k(F)$ , each factor of which being the transition word of a loop on the same state. Hence, for any map

$$\theta: \{1, 2, \dots, q\} \rightarrow \{1, 2, \dots, k\}$$

(not necessarily injective), one has:

$$\text{word}(F) \in L \Leftrightarrow \text{word}(G \circ \theta) \in L$$

and that achieves the proof of the direct part.

(B) *Proof of the converse ((ii) ⇒ (i))*

**First case** ( $\Theta$  does not contain any surjection).

Let  $\Theta$  be a  $k$ -transformation scheme.

Denote by  $L_N(\Theta)$  the set of languages  $L \subset A^*$  that are strictly  $\Theta$ -transformable at the order  $N$ , and whose the complementary is weakly  $\Theta$ -transformable at the order  $N$ .

Recall that the residual of  $L \subset A^*$  by a word  $u \in A^*$  is the language

$$L \circ u = \{w \in A^* \mid uw \in L\}$$

clearly, we have

$$C(L \circ u) = (CL) \circ u.$$

Recall also the “Nerode’s criterium” [13]: a language is regular if and only if it has only a finite set of distinct residuals.

Thus the theorem will be proved, in this case, by the following two lemmas.

LEMMA 3.2:  $L_N(\Theta)$  is closed by residuals.

*Proof:* For any  $m$ -factorization  $F=(x, f_1, f_2, \dots, f_m, y)$ , we set

$$uF=(ux, f_1, f_2, \dots, f_m, y).$$

We have clearly:

$$\text{word}(F) \in L \circ u \Leftrightarrow u \text{word}(F) \in L \Leftrightarrow \text{word}(uF) \in L.$$

Furthermore, a  $k$ -factorization  $G$  is in  $\Phi_k(F)$  if and only if  $uG$  is in  $\Phi_k(uF)$ .

Now, if we suppose for any  $F \in \Phi_N$

$$\Phi_k(F) \subset E_{\Theta}(CL) \Rightarrow \text{word}(F) \in CL$$

we deduce in particular, for any  $F \in \Phi_N$ , and  $u \in A^*$

$$\Phi_k(uF) \subset E_{\Theta}(CL) \Rightarrow \text{word}(uF) \in CL$$

and finally:

$$\Phi_k(F) \subset E_{\Theta}(C(L \circ u)) \Rightarrow \text{word}(F) \in C(L \circ u).$$

Hence,  $L$  strictly  $\Theta$ -transformable at the order  $N$  implies  $L \circ u$  strictly  $\Theta$ -transformable at the order  $N$ .

The same argument agrees in the weakly  $\Theta$ -transformable case, and that proves the lemma.

LEMMA 3.3: Let  $N \in \mathbb{N} - \{0\}$ , and  $\Theta$  a  $k$ -transformation scheme. If  $\Theta$  does not contain any surjection, then  $L_N(\Theta)$  is finite.

*Proof:* Let  $L \in L_N(\Theta)$ .

Let  $\rho = \text{Ram}(k, N)$ , and let  $w_1 \in A^*$  such that  $|w_1| = m \geq \rho$ .

Let  $A^{<m} = \{w \in A^* \mid |w| < m\}$ , and let

$$E = (\varepsilon, l_1, l_2, \dots, l_m, \varepsilon)$$

the litteral factorization of  $w_1$ .

(a) If  $G \in \Phi_k(E)$ , recall that we have

$$G \in U_{\Theta}(L) \quad \text{iff} \quad \forall \theta \in \Theta, \quad \text{word}(G \circ \theta) \in L.$$

Now, because  $\Theta$  does not contain any surjection we have:

$$\forall \theta \in \Theta, \quad |\text{word}(G \circ \theta)| < |\text{word}(G)| = m.$$

Hence the set  $L \cap A^{<m}$  suffices to determine if  $G$  belongs to  $U_\Theta(L)$ , or not.

Thus the set  $L \cap A^{<m}$  completely determines the partition of  $\Phi_k(E)$  induced by:

$$\Phi_k = U_\Theta(L) \sqcup E_\Theta(CL).$$

(b) Otherwise, since  $m \geq \rho = \text{Ram}(k, N)$ , and by the Ramsey's theorem, there is an  $N$ -factorization  $F$  compatible with  $E$ , such that

$$\begin{aligned} &\text{either } \Phi_k(F) \subset U_\Theta(L) \\ &\text{or } \Phi_k(F) \subset E_\Theta(CL). \end{aligned}$$

In the second case, since  $L$  is strictly  $\Theta$ -transformable at the order  $N$ , we have:

$$w_1 = \text{word}(F) \in CL.$$

In the first case, since  $CL$  is weakly  $\Theta$ -transformable at the order  $N$ , we obtain

$$w_1 = \text{word}(F) \in L.$$

Thus, we see that the set  $L \cap A^{<m}$  completely determines whether or not  $w_1 \in A^m$  belongs to  $L$ .

(c) Finally, the set  $L \cap A^{<m}$  completely determines  $L \cap A^m$ , if  $m > \rho$ . By induction, we obtain that  $L \cap A^{<\rho}$  completely determines any language  $L \subset L_N(\Theta)$ . Thus we have, if  $d = \text{Card } A$ :

$$\text{Card}(L_N(\Theta)) \leq \text{Card}(\mathcal{P}(A^{<\rho})) = 2^s,$$

with

$$s = (d^{\rho-1}) / (d-1).$$

That achieves the proof of the theorem, in the case "without surjection".

(c) *Second case ( $\Theta$  contains at least one surjection, but  $L$  is a periodic language)*

Recall first that a language is called *periodic* if for any  $v \in A^*$ , the image of  $v$  in the syntactic monoid of  $L$  is a periodic element. In other words,  $L$  is

periodic if and only if

$$\begin{aligned} &\forall v \in A^*, \exists p_v, q_v \in \mathbb{N}; \\ &\forall x, y \in A^*; \\ & xv^{p_v+q_v} y \in L \Leftrightarrow xv^{q_v} y \in L. \end{aligned}$$

We would like to replace the families of integers  $p_v$  and  $q_v$  by a single  $(p, q)$  independent of  $v$ . That is generally not possible, however:

**DEFINITION:** Let  $R, p, q \in \mathbb{N} - \{0\}$ . A language  $L$  will be called *R-periodic of type  $(p, q)$*  if and only if

$$\begin{aligned} &\forall v \in A^{<R}, \\ &\forall x, y \in A^*; \\ & xv^{p+q} y \in L \Leftrightarrow xv^q y \in L \end{aligned}$$

and we note  $\text{Per}_R(p, q)$  the set of languages  $L \subset A^*$  that are *R-periodic of type  $(p, q)$* .

**LEMMA 3.4 (Restivo and Reutenauer):** *For any  $R \in \mathbb{N} - \{0\}$  and any periodic language  $P$ , one can compute two integers  $p = p(R, P)$  and  $q = q(R, P)$  such that*

$$P \in \text{Per}_R(p, q).$$

*Proof:* Let  $p$  be a common multiple of all the  $p_v$  such that  $|v| < R$ , and let  $q$  any integer greater than all the  $q_v$  such that  $|v| < R$ . Clearly,  $p$  and  $q$  satisfy the lemma.

Now, as in the first case, the theorem will be prove if we establish the two following lemmas:

**LEMMA 3.5:** *For any  $R, p, q \in \mathbb{N}$ , the set  $\text{Per}_R(p, q)$  is closed by residuals, and by complementation.*

*Proof:* The stability by complementation is a direct consequence of the definition of  $\text{Per}_R(p, q)$ .

Furthermore, suppose  $L \in \text{Per}_R(p, q)$ . Thus for any words  $u, x, y \in A^*$  and  $v \in A^{<R}$ , we have:

$$uxv^{p+q} y \in L \Leftrightarrow uxv^q y \in L$$

and equivalently:

$$xv^{p+q} y \in L \circ u \Leftrightarrow xv^q y \in L \circ u$$

and consequently

$$L \circ u \in \text{Per}_R(p, q).$$

LEMMA 3. 6: Let  $N \in \mathbb{N} - \{0\}$  and let  $\Theta$  be a  $k$ -transformation scheme. Then for any  $R, p, q \in \mathbb{N} - \{0\}$ , the set  $L_N(\Theta) \cap \text{Per}_R(p, q)$  is finite (restricted to a fixed alphabet  $A$ ).

*Proof:* For any word  $w_1 \in A^*$ , we set:

$$A^{<w_1} = \{w \in A^* \mid w < w_1\}.$$

Let  $d = \text{Card}(A)$ . We set:

$$\rho = \text{Ram}(k, N), \quad \text{and} \quad \sigma = \text{Sh}(d, p + q, \rho).$$

Let  $L \subset A^*$  be any language belonging to

$$L_N(\Theta) \cap \text{Per}_\rho(p, q).$$

We shall prove that for any word  $w_1 \in A^*$  such that  $|w_1| \geq \sigma$ , the subset

$$L \cap A^{<w_1}$$

completely determines if  $w_1$  belongs to  $L$  or not.

Indeed:

By the Shirshov's theorem, and since  $|w_1| > \sigma = \text{Sh}(d, p + q, \rho)$ , we have:

either  $w_1 = \text{word}(H)$  for some  $(p + q)$ -power  $H$  of a word  $h$  such that  $1 \leq |h| < \rho$ ,

or  $w_1 = \text{word}(E)$  for some  $\rho$ -division  $E$ .

In the first case, since  $|h| < \rho$  and  $L \in \text{Per}_\rho(p, q)$ , we have:

$$w_1 = xh^{p+q}y \in L \iff xh^qy \in L.$$

Now,  $|xh^qy| < |w_1|$ , and consequently  $xh^qy < w_1$ . That implies, of course, that  $L \cap A^{<w_1}$  completely determines whether  $w_1$  belongs to  $L$  or not.

In the last case, we proceed as in the proof of lemma 3. 3:

(a) Recall that for any  $G \in \Phi_k(E)$

$$G \in U_\Theta(L) \quad \text{iff} \quad \forall \theta \in \Theta, \quad \text{word}(G \circ \theta) \in L.$$

Now, for any  $G \in \Phi_k(E)$ , because  $E$  is a  $\rho$ -division, we have:

$$\forall \theta \in \Theta, \quad \text{word}(G \circ \theta) < \text{word}(G) = w_1$$

(recall that the identity map does not belong to  $\Theta$ ). Hence the subset  $L \cap A^{<w_1}$  completely determines whether  $G$  belongs to  $U_\Theta(L)$ , or not.

Thus the set  $L \cap A^{<w_1}$  completely determines the partition of  $\Phi_k(E)$  induced by:

$$\Phi_k = U_\Theta(L) \sqcup E_\Theta(CL).$$

(b) Otherwise, by the Ramsey's theorem, and because  $E \in \Phi_\rho$  and  $\rho = \text{Ram}(k, N)$ , there exists an  $N$ -factorization  $F \in \Phi_N(E)$  such that:

either  $\Phi_k(F) \subset U_\Theta(L)$ ;

or  $\Phi_k(F) \subset E_\Theta(CL)$ .

Because  $L$  is strictly  $\Theta$ -transformable, and  $CL$  is weakly  $\Theta$ -transformable, at the order  $N$ , we deduce in the first case  $w_1 \in L$ , and in the second case  $w_1 \in CL$ .

Thus in any case,  $L \cap A^{<w_1}$  completely determines whether  $w_1$  belongs to  $L$ , or not.

(c) Let  $u_0$  be the smallest word of length  $\sigma$  for the lexicographical-by-length ordering. By induction on that order, we deduce that  $L \cap A^{<u_0} = L \cap A^{<\sigma}$  completely determines any language  $L$  on  $A$  that belongs to  $L_N(\Theta) \cap \text{Per}_\rho(p, q)$ . An finally:

$$\begin{aligned} \text{Card}(\mathcal{P}(A^*) \cap L_N(\Theta) \cap \text{Per}_\rho(p, q)) \\ \leq \text{Card}(\mathcal{P}(A^{<\sigma})) = 2^t \\ \text{with } t = (d^\sigma - 1)/(d - 1). \end{aligned}$$

That achieves the proof of the theorem in the case of periodic languages.

#### 4. SUPPORTS AND CANCELLATIONS: THE MAIN THEOREM

##### (A) Rational power series

For definitions and general developments on rational power series, see for example [1].

DEFINITION 4. 1: We call *formal power series* (in noncommuting variables) on  $A$  with coefficients in a field  $K$  any map  $S : A^* \rightarrow K$ , and we denote it by a formal sum:

$$S = \sum_{w \in A^*} \langle S | w \rangle w$$

We call *support of S* the language

$$\text{supp}(S) = \{ w \in A^* \mid \langle S \mid w \rangle = 0 \}.$$

In all the sequel, we identify any vector  $V \in K^N$  with the linear application:

$$\begin{aligned} V: K &\rightarrow K^N \\ 1 &\mapsto V \end{aligned}$$

In particular, any  $k \in K$  will be identified with the linear endomorphism of  $K$ , "product by  $k$ ".

DEFINITION 4.2: A power series  $S$  will be called *recognizable*, or *rational*, if there are:

- some finite dimensional vector space  $E = K^d$  ( $d$  interger) and
- a linear map  $\gamma: K \rightarrow E$  (i. e. an element of  $E$ );
- a monoid homomorphism  $\mu: A^* \rightarrow \text{End}(E)$ ;
- a linear form  $\lambda: A \rightarrow K$ ,

such that for any  $w \in A^*$ , we have:

$$\langle S \mid w \rangle = \lambda \circ \mu \circ w \circ \gamma.$$

The quadruplet  $(E, \lambda, \mu, \gamma)$  will be called a *linear representation* of the series  $S$ .

*Remark:* The sign  $\circ$  can be interpreted as a matricial product,  $\lambda$  (resp.  $\gamma$ ) as an arrow matrix (resp. column matrix) and each  $\mu w$  as a square matrix.)

DEFINITION 4.3: Following Restivo and Reutenauer [17], we shall call *support* any language that is support of some rational power series.

It is well known that *any regular language is a support*. Indeed, since any rational language is non-ambiguously rational, its characteristic power series  $\chi L = \sum_{w \in L} 1 \cdot w$  is a rational series.

But *the converse is false*: for example, it is well known that

$$D_1^* = \{ w \in \{ a, b \}^* \mid \text{length}_a(w) = \text{length}_b(w) \}$$

is not a regular language nor its complement:

$$T = \{ w \in \{ a, b \}^* \mid \text{length}_a(w) \neq \text{length}_b(w) \}.$$

In spite of that,  $T$  is the support of the following rational power series

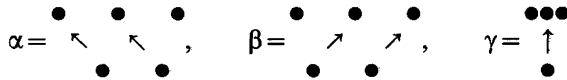
$$S = \sum_{w \in \{a, b\}^*} [\text{length}_a(w) - \text{length}_b(w)] w.$$

Indeed,  $S$  can be defined by the linear representation of the series  $S$ :

$$\begin{matrix} \lambda = (1 & 0) \\ \mu a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, & \mu b = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, & \gamma = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{matrix}$$

Restivo and Reutenauer have proved that each support has “the weak cancellation property”. In other words, each support is  $\{\theta\}$ -transformable, where  $\theta$  is the unique application of the empty set into  $\{1\}$  (see example 1), and they have deduced that if  $L$  and  $CL$  are supports, then  $L$  is a regular language (particular case of our theorem 3).

They have proved also (in the their lemma 4.2) that any support is strictly  $\Gamma$ -transformable, where  $\Gamma$  is the following 3-transformation scheme:



that is, in case  $G = (f, g_1, g_2, g_3, h)$ :

- word  $[g \circ \alpha] = fg_1g_2h$ ;
- word  $[g \circ \beta] = fg_2g_3h$ ;
- word  $[g \circ \gamma] = fg_2h$ .

And they ask whether there is a more general result (p. 258, remark).

Now we state our main theorem on supports.

**THEOREM 4:** *For any  $k \in \mathbb{N} - \{0\}$ , and any support  $L$ , one can compute an integer  $m$  such that  $L$  is strictly  $k$ -cancellable at the order  $m$ .*

**COROLLARY:** *Let  $L$  be a support. In order  $L$  to be regular, it suffices that there are two integers  $k$  and  $m$  such that  $CL$  is weakly  $k$ -cancellable at the order  $m$ .*

*Proof:* This corollary is a direct consequence of our Theorem 3.

Thus, in a certain sense, supports are very close to regular languages, because a very weak condition on their complementary implies regularity.

The proof of theorem 4 requires a combinatorial result on sequences of vectors in a finite dimensional vector space, that we present in the following section.

(B) sequences of vectors

If  $\mathcal{V} = (V_1, V_2, \dots, V_m)$  is a sequence of vectors, we denote by  $\text{span}(\mathcal{V})$  the vector space spanned by the vectors  $V_j (j = 1, 2, \dots, m)$ .

DEFINITION 4.4: A sequence  $\mathcal{V} = (V_1, V_2, \dots, V_m)$  of nonnul vectors of  $K^d$  ( $d$  integer) will be called a *festoon of size  $q$*  if it is the concatenation of  $q$  subsequences  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_q$  that satisfy:

$$\text{span}(\mathcal{V}_1) = \text{span}(\mathcal{V}_2) = \dots = \text{span}(\mathcal{V}_q) = \text{span}(\mathcal{V}).$$

In other words, for  $i = 1, 2, \dots, q$  there is a sequence

$$\mathcal{V}_i = (V_{j_i+1}, V_{j_i+2}, \dots, V_{j_{i+1}})$$

with

$$0 \leq j_1 < j_2 < \dots < j_q < j_{q+1} = m$$

such that for any  $i \in \{1, 2, \dots, q\}$

$$\text{span}(\mathcal{V}_i) = \text{span}(\mathcal{V}).$$

The integer  $\dim(\text{span}(\mathcal{V}))$  is called *dimension of the festoon*. Each subsequence  $(\mathcal{V}_i)$  is called a *mesh of the festoon*.

THEOREM 5: For any  $q, d \in \mathbb{N} - \{0\}$  and any sequence  $\mathcal{W}$  of  $q^d$  nonnul vectors of  $K^d$ , there exists a subsequence of consecutive vectors of  $\mathcal{W}$  which is a festoon of size  $q$ .

*Proof.* We could obtain this theorem as an application of Ramsey's theorem, or of the "théorème du rang constant" [7]. We give here a direct and economical proof.

The theorem is trivially true for  $d = 1$ .

Suppose now the theorem already proved for  $d \leq d_0$ , and let  $\mathcal{W}$  be a sequence of  $q^{d_0+1}$  vectors of  $K^{d_0+1}$ . We distinguish two cases.

(a) Suppose that there is a subsequence  $\mathcal{V}$  of  $q^{d_0}$  consecutive vectors of  $\mathcal{W}$  such that  $\dim(\text{span}(\mathcal{V})) = d_0$ . Then, by induction hypothesis, we can find in  $\mathcal{V}$  the required festoon.

(b) In other case, any sequence of  $q^{d_0}$  consecutive vectors of  $\mathcal{W}$  generates  $\text{span}(\mathcal{W})$ , because it is the only  $(d_0 + 1)$ -dimensional subspace of  $\text{span}(\mathcal{W})$ .

Hence, if  $s = q^{d_0}$ , the  $q$  following sequences:

$$\mathcal{W}_1 = (V_1, V_2, \dots, V_s);$$

$$\mathcal{W}_2 = (V_{s+1}, V_{s+2}, \dots, V_{2s});$$

$$\mathcal{W}_q = (V_{(q-1)s+1}, V_{(q-1)s+2}, \dots, V_{qs}).$$

satisfy:

$\text{Span}(\mathcal{W}_1) = \text{span}(\mathcal{W}_2) = \dots = \text{Span}(\mathcal{W}_s) = \text{span}(\mathcal{W})$ , and consequently, the  $\mathcal{W}_i$  are the meshes of a festoon of dimension  $d_0 + 1$  and of size  $q$ .

Finally, the theorem 5 is proved by induction on  $d_0$ .

**5. SUPPORTS ARE STRICTLY  $k$ -CANCELLEABLE: THE PROOF**

(a) Let  $f, g, f', g'$  be four linear applications such that  $f \circ g$  and  $f' \circ g'$  are defined.

$$\text{Then } (f \circ g) \otimes (f' \circ g') = (f \otimes f') \circ (g \otimes g').$$

Recall also that the tensor product of two elements of the field  $K$  is nothing else as their product in  $K$ .

Thus, if  $V$  and  $W \in E = K^d$  and if  $f$  and  $g$  are two linear forms on  $K^d$ , then we have

$$(f \otimes g) \circ (V \otimes W) \neq 0$$

if and only if

$$f \circ V \neq 0 \quad \text{and} \quad g \circ W \neq 0.$$

Thus the tensor product allows to “simultaneously control” two inequalities.

(b) Now, let  $(E \cong K^d, \lambda, \mu, \gamma)$  be a linear representation of some rational series  $S$ , and let  $k$  be any positive integer. In order to prove that  $\text{supp}(S)$  is strictly  $k$ -cancellable at some order  $m$ , we shall have to “simultaneously control”  $2^{k-1}$  inequalities, and thus to compute in the  $2^{k-1}$ -th tensor power of  $E$ .

Thus we set

$$\hat{E} = E \otimes E \otimes \dots \otimes E$$

$2^{k-1}$  times

hence  $E$  is a vector space of dimension

$$\dim(\hat{E}) = \tau, \quad \tau = d^{2^{k-1}}$$

Now let  $M = (k + 1)\tau$ . We shall prove that the support of  $S$  is strictly  $k$ -cancellable at the order  $M$ , that is [using lemma 2.1, (iii)]: for any  $F \in \Phi_M$  such that  $\text{word}(F) \in \text{supp}(S)$  we shall prove the following hypothesis.

(H) one can find some  $G \in \Phi_k(F)$  such that

$$G \in U_{C_k}(\text{supp } S)$$

(c) In fact, we shall prove a stronger result. For any  $F = (x, u_1, u_2, \dots, u_M, y) \in \Phi_M$  such that  $\text{word}(F) \in \text{supp}(S)$  we set

$$\begin{aligned} s_j &= u_j u_{j+1} \dots u_M y \\ S_j &= \mu(u_j u_{j+1} \dots u_M y) \circ \gamma \\ V_j &= S_j \otimes S_j \otimes \dots \otimes S_j \\ &\quad \text{\scriptsize } 2^{k-1} \text{ times} \end{aligned}$$

Then  $\mathcal{V} = (V_1, V_2, \dots, V_M)$  is a sequence of  $M = (k + 1)^\tau$  vectors in the space  $\hat{E}$  of dimension  $\tau$ . Thus let  $(F)$  be a festoon of size  $k + 1$ ,

$$F = (\mathcal{W}_1, \mathcal{W}_2, \dots, \mathcal{W}_{k+1})$$

that can be found in  $\mathcal{V}$  by theorem 4.

In fact we shall prove the following hypothesis, for  $q = 1, 2, \dots, k + 1$ .

(H<sub>q</sub>) One can find some factorization  $G_q \in \Phi_q(F)$ , defined by a sequence of integers

$$0 \leq j_1 < j_2 < \dots < j_{q+1} \leq M$$

that satisfies

- (i)  $G_q \in U_{C_{j_q}}(\text{supp } S)$ .
- (ii)  $V_{j_i} \in \mathcal{W}_i$  for  $i = 1, 2, \dots, q + 1$

(end of H<sub>q</sub>) (Fig. 5).

Clearly (H<sub>k</sub>) implies (H), with  $G = G_k$ .

Clearly also, (H<sub>0</sub>) is trivially true. Indeed, C<sub>0</sub> being the empty set (because Id<sub>( )</sub> is not a transformation), so (H<sub>0</sub>) only asserts the existence of some  $V_{j_1} \in \mathcal{W}_1$ .

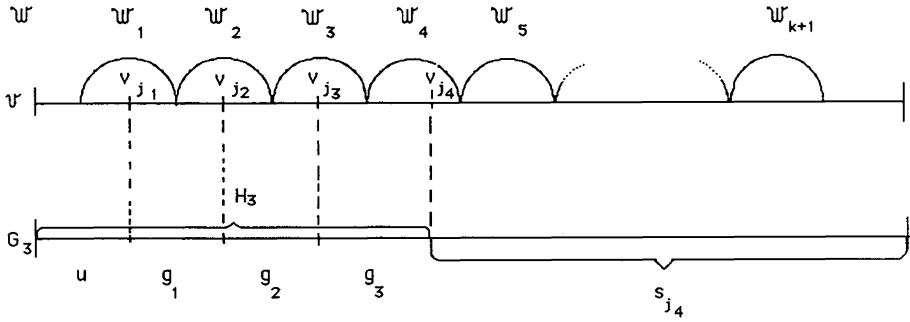
(d) Now, we suppose (H<sub>q</sub>) satisfied for  $q < k$ , and we prove (H<sub>q+1</sub>).

Introduce first some notations. The factorizations G is of the form

$$G_q = (f, g_1, g_2, \dots, g_q, s_{j_{q+1}}).$$

We set

$$H_q = (f, g_1, g_2, \dots, g_q, \epsilon)$$



- (i)  $\forall \theta \in C_3, \text{word}(G_3 \circ \theta) \in \text{Supp}(S)$
- (ii)  $W$  is a festoon of size  $(k+1)$ , and  $\forall j_4 \in W_4$

Figure 5. —  $(H_3)$ -hypothesis.

and so we have, for any  $\theta \in C_q \cup \{Id_q\}$

$$\text{word}(G_q \circ \theta) = \text{word}(H_q \circ \theta) \cdot s_{j_{q+1}}$$

[where  $\text{word}(G_q \circ Id) = \text{word}(G_q)$ ].

But  $\text{word}(G_q \circ \theta)$  belongs to  $\text{supp } L$  if and only if

$$\langle S \mid \text{word}(G_q \circ \theta) \rangle \neq 0$$

in other words, if and only if

$$\begin{aligned} \lambda \circ \mu(\text{word}(G_q \circ \theta)) \circ \gamma &\neq 0 \\ \lambda \circ \mu(\text{word}(H_q \circ \theta)) \circ \mu(s_{j_{q+1}}) \circ \gamma &\neq 0 \end{aligned}$$

and that can be write

$$P_\theta \circ S_{j_{q+1}} \neq 0$$

where we have set:

$$P_\theta = \lambda \circ \mu[\text{word}(H_q \circ \theta)].$$

Hence we have, by  $(H_q)$  and because  $\text{word}(G_q) \in \text{supp } S$ :

$$(E_q) \quad \forall \theta \in C_q \cup \{ \text{Id}_q \}, \quad P_\theta \circ S_{j_{q+1}} \neq 0.$$

We have  $\text{Card}(C_q \cup \{ \text{Id}_q \}) = 2^q \leq 2^{k-1}$ . Thus we can “simultaneously control” these set of  $2^q$  equations as follows. We set

$$\hat{P} = \left( \bigotimes_{\Theta \in C_q \cup \{ \text{Id}_q \}} P_\Theta \right) \otimes (P_{\text{Id}}^{\otimes (2^{k-1} - 2^q)})$$

Recall that  $V_{j_{q+1}} = S_{j_{q+1}} \otimes S_{j_{q+1}} \otimes \dots \otimes S_{j_{q+1}}$   
 $2^{k-1} \text{ times}$

Thus the  $2^q$  equations  $(E_q)$  are equivalent to

$$\hat{P} \circ V_{j_{q+1}} \neq 0.$$

(e) Now,  $V_{j_{q+1}} \in W_{q+1}$ , and because  $q+1 < k+1$ , there is a mesh  $\mathcal{W}_{q+2}$  in the festoon  $F$ , and we have:

$$V_{j_{q+1}} \in \text{span}(\mathcal{W}_{q+2}).$$

Thus we have successively, for some  $\alpha_s \in k$ ,

$$V_{j_{q+1}} = \sum_{V_s \in \mathcal{W}_{q+2}} \alpha_s V_s$$

$$\hat{P} \circ V_{j_{q+1}} = \sum_{V_s \in \mathcal{W}_{q+2}} \alpha_s (\hat{P} \circ V_s) \neq 0.$$

Consequently, we can find  $V_{j_{q+2}} \in \mathcal{W}_{q+2}$  such that

$$\hat{P} \circ V_{j_{q+2}} \neq 0$$

and that is equivalent to

$$(E'_q) \quad \forall \theta \in C_q \cup \{ \text{Id}_q \}, \quad P_\theta \circ S_{j_{q+2}} \neq 0$$

and in other words:

$$\lambda \circ \mu [\text{word}(H_q \circ \theta)] \circ \mu (s_{j_{q+2}}) \circ \gamma \neq 0$$

and finally we obtain

$$\forall \Theta \in C_q \cup \{ \text{Id}_q \}$$

$$\text{word}(H_q \circ \Theta) \cdot s_{j_{q+2}} \in \text{supp } L.$$

(see Fig. 6).

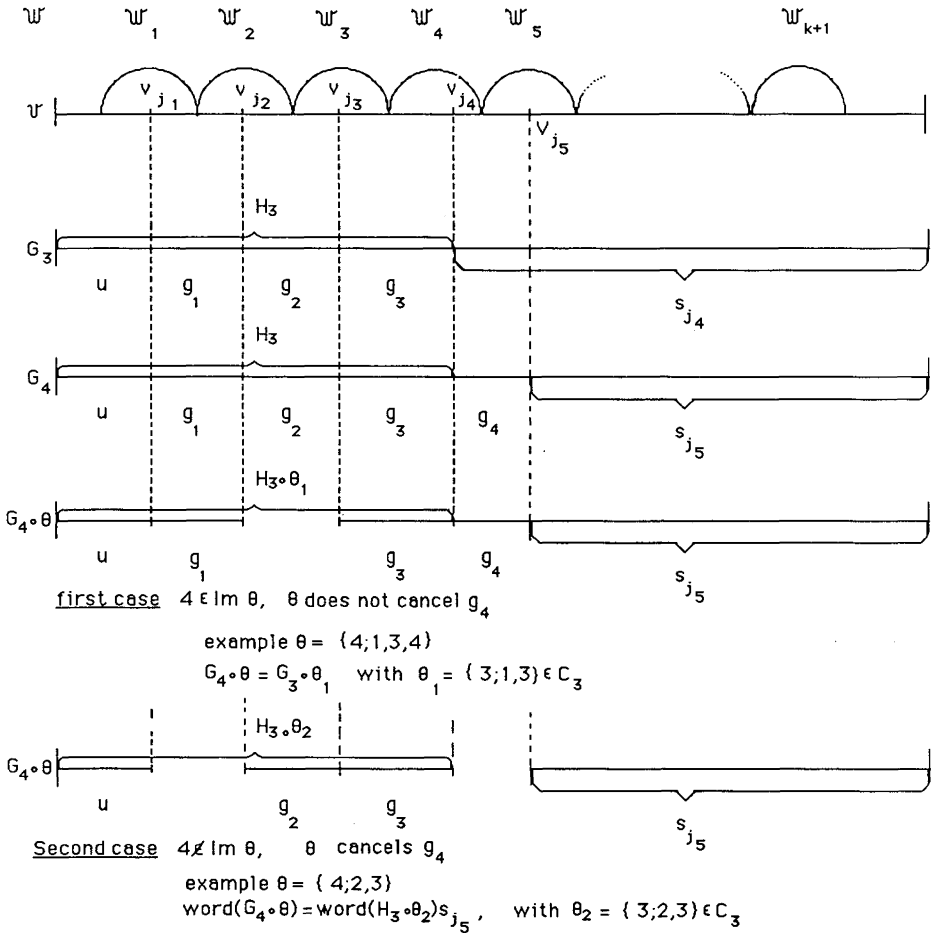


Figure 6. — hypothesis  $(H_3)$  implies hypothesis  $(H_4)$ .

(f) Now we can define  $G_{q+1} \in \Phi_{q+1}(F)$  by:

$$G_{q+1} = (f, g_1, g_2, \dots, g_{q+1}, s_{j_{q+2}})$$

and we claim that

$$\forall \theta \in C_{q+1}, \text{ word}(G_{q+1} \circ \theta) \in \text{supp } L.$$

Indeed, if  $\theta \in C_{q+1}$  (see example)

either  $q + 1$  belongs to  $\text{Im } \theta$ , and then

$$\begin{aligned} \text{word } (G_{q+1} \circ \theta) &= \text{word } (G_q \circ \theta_1) \\ &= \text{word } (H_q \circ \theta_1) g_{q+1} s_{j_{q+2}} \\ &\text{for some } \theta_1 \in C_q \end{aligned}$$

and  $\text{word } (G_q \circ \theta_1)$  belongs to  $\text{supp } L$  by hypothesis  $(H_q)$

or  $q + 1$  does not belong to  $\text{Im } \theta$ , and then

$$\begin{aligned} \text{word } (G_{q+1} \circ \theta) &= \text{word } (H_q \circ \theta_2) s_{j_{q+2}} \\ &\text{for some } \theta_2 \in C_q \cup \{ \text{Id}_q \} \end{aligned}$$

and  $\text{word } (G_{q+\theta})$  belongs to  $\text{supp } L$  by  $(E'_q)$ .

Thus we have proved  $(H_{q+1})$ , and that achieves the proof of the theorem 5.

CONCLUSION

We recall only the corollary:

Let  $L$  be a support. In order  $L$  to be regular, it suffices that  $CL$  is weakly  $k$ -cancellable at the order  $m$ , for some  $k, m$  strictly positive integers.

Perhaps that will be a useful tool in order to study the following conjecture of Restivo and Reutenauer [18], p. 26 of "rational separability of disjoint supports". If  $L_1$  and  $L_2$  are two disjoint supports of rational power series, there exists a regular language  $K$  containing  $L_1$  and not intersecting  $L_2$ .

REFERENCES

1. J. BERSTEL and C. REUTENAUER, *Les séries rationnelles et leurs langages*, Masson, Paris, 1984.
2. A. EHRENFEUCHT, R. PARIKH and G. ROZENBERG, *Pumping Lemmas for Regular Sets*, SIAM J. Comput., Vol. 10, 1981, pp. 536-541.
3. A. EHRENFEUCHT, D. HAUSSLER and G. ROZENBERG, *On Regularity of Context-free Languages*, Theoret. Comput. Sci. Vol. 27, 1983, pp. 311-332.
4. S. EILENGERG, *Automata, Languages and Machines*, Volumes A, B, Academic Press, New York, 1974.
5. M. HARRISON, *Introduction to Formal Language Theory*, Addison-Wesley, Reading, MA 1978.
6. K. HASHIGUSHI, *Positive Pumping Conditions and Cancellation Conditions for Regular Languages*, to be published.
7. G. JACOB, *Un théorème de factorisation des produits d'endomorphismes de  $K^N$* , J. of Algebra, Vol. 63, No. 2, 1980, pp. 389-412.
8. G. JACOB, *La finitude des représentations linéaires des semi-groupes est décidable*, J. of Algebra, Vol. 52, No. 2, 1978, pp. 437-459.
9. J. JAFFE, *A Necessary and Sufficient Pumping Lemma for Regular Languages*, Sigact News, 1948, pp. 48-49.

10. G. LALLEMENT, *Semigroups and combinatorial applications*, Wiley, New York, 1979.
11. M. LOTHAIRE, *Combinatorics of Words*, Addison-Wesley, Reading, MA 1983.
12. J. MYHILL, *Finite Automata and the Representation of Events*, WADD Technical Report 57-624, 1957, Wright-Patterson Air Force Base, Ohio, 45433.
13. A. NERODE, *Linear Automata Transformations*, Proc. Amer. Math. Soc., Vol. 9, 1958, pp. 541-544.
14. F. P. RAMSEY, *On a Problem of Formal Logic*, The Foundations of Mathematics, Routledge and Kegan Paul, 1954, pp. 82-111 and reprinted from Proc. London Math. Soc., Ser 2, Vol. 30, 1928, pp. 338-384.
15. A. RESTIVO and C. REUTENAUER, *Some Applications of a Theorem of Shirshov to Language Theory*, Inform. and Control, Vol. 57, 1983, pp. 205-213.
16. A. RESTIVO and C. REUTENAUER, *On the Burnside Problem for Semigroups*, J. Algebra, Vol. 89, 1984, pp. 102-104.
17. A. RESTIVO and C. REUTENAUER, *On Cancellation Properties of Languages which are Supports of Rational Power Series*, J. Comput. System Sci., Vol. 29, 1984, pp. 153-159.
18. A. RESTIVO and C. REUTENAUER, *Rational Languages and the Burnside Problem*, Theoret. Comput. Sci., Vol. 40, 1985, pp. 13-30.
19. A. SALOMAA and M. SOITTOLA, *Automata Theoretic Aspects on Fromal Power Series*, Springer, Berlin, 1978.
20. M. P. SCHUTZENBERGER, *Finite Counting Automata*, Inform. Control, Vol. 5, 1962, pp. 91-107.
21. A. I. SHIRSHOV, *On rings with identity relations*, Mat. Sb., Vol. 43, 1957, pp. 277-283 (in russian).
22. H. STRAUBING, *The burnside problem for semigroups of matrices*, Combinatorics on Words, Progress and Perspectives, L. J. CUMMINGS, Ed., Academic Press, New York, 1983, pp. 279-295.