JEAN BERSTEL

Some Recent Results on Squarefree Words

Publications du Département de Mathématiques de Lyon, 1985, fascicule 2B
« Compte rendu des journées infinitistes », , p. 21-36

<http://www.numdam.org/item?id=PDML_1985___2B_21_0>

SOME RECENT RESULTS ON SQUAREFREE WORDS

by Jean BERSTEL [1]

"Für die Entwicklung der logischen Wissenschaften wird es, ohne Rücksicht auf etwaige Anwendungen, von Bedeutung sein, ausgedehnte Felder für Spekulation über schwierige Probleme zu finden."

Axel Thue, 1912.

1. Introduction.

When Axel Thue wrote these lines in the introduction to his 1912 paper on squarefree words, he certainly did not feel as a theoretical computer scientist. During the past seventy years, there was an increasing interest in squarefree words and more generally in repetitions in words. However, A. Thue's sentence seems still to hold : in some sense, he said that there is no reason to study squarefree words, excepted that it's a difficult question, and that it is of primary importance to investigate new domains. Seventy years later, these questions are no longer new, and one may ask if squarefree words served already.

First, we observe that infinite squarefree, overlap-free or cube-free words indeed served as examples or counter-examples in several, quite different domains. In symbolic dynamics, they were introduced by Morse in 1921 [36] . Another use is in group theory, where an infinite square-free word is one (of the numerous) steps in disproving the Burnside conjecture (see Adjan [2]). Closer to computer science is Morse and Hedlund's interpretation in relation

1) Le présent texte reproduit avec sa permission l'article de l'auteur au "Symposium for theoretical aspect of Computer Science (STACS'84). N.D.E.

with chess [37]. We also mention applications to formal language theory : Brzozowsky, K. Culik II and Gabriellian [7] use squarefree words in connection with noncounting languages, J. Goldstine uses the Morse sequence to show a property of some family of languages [22] . See also Shyr [52], and Reutenauer [43]. All these are cases where repetition-free words served as explicit examples. In other cases, questions about these words led to new insights in other domains, such as for DOL languages and for context-free languages. At the present time, the set of result on repetitions constitutes a topic in combinatorics on words.

This paper gives a survey of some recent results concerning squarefree words and related topics. In the past years, the interest in this topic was indeed growing, and a number of results are now available. An account of basic results may be found in Salomaa [45,46] and in Lothaire [30] . For earlier work, see also Hedlund's paper [25] . The more general concept of unavoidable pattern is introduced in Bean, Ehrenfeucht, McNulty [4] . Part 2 deals with powers and repetitions, part 3 with language-theoretic results, part 4 gives the estimations on growth, part 5 describes results on morphisms.

2. Powers and repetitions.

A square is a word of the form xx, with x a nonempty word. Cubes and k-th powers are defined accordingly. A word is squarefree if none of his factors (in the sense of Lothaire [30] , or subwords) is a square. A word is overlap-free if it contains no factor of the form xuxux, with x nonempty. The concept of k-th power free words where k implicitly is a positive integer, can be extendee to rational numbers as follows : if r = n + s is a rational positive number with n positive integer and $0 < s < 1$, then an r-th power is a word of the form

$$u^n u'$$

with exactly n consecutive u's and one left factor u' of u satisfying $|u'|/|u| = s$.

The The-Morse sequence

m = 01101001100101101001011O...

contains squares and is overlap-free (Thue [54] , Morse [36]) , the word

t = abcacbabcba...

derived from m by the inverse morphism a —> 011, b —> 01 , c —> 0 is square-free (Thue [55]). The Fibonacci word

f = abaababaabaababaabab...

containe cubes but is 4-th power free (see e.g. Karhumaki [27]). Many other special infinite words with some repetition property are known. Usually, they are constructed by iterating morphisms or by tag systems in the sense of Minsky [35] . (See also Pansiot's paper in the STACS'84 proceedings). Let us mention that some words may also be defined by an explicit description of the positions of the letters occuring in them. This holds for the Thue-Morse sequence, since the i-th letter can be shown to be 0 or 1 according to the number of "1" in the binary expansion of i being even or odd. A more systematic treatement of these descriptions is given in Christol, Kamae, Mendès-France, Rauzy [10] . One of the properties of these generalized sequences is given by Cerny [9] . He defines, for a given fixed word w over {0,1} an infinite word by setting the i-th letter to 0 or to 1 when the number of occurences of w in the binary expansion of i is even or is odd. Thus the original Thue-Morse sequence is the special case where w=1. Cerny shows that the infinite word that is obtained in this manner has no factors of the form

$$(xu)^k x$$

where $k=2^{|w|}$, and x is nonempty.

Squares are unavoidable over two letters, and they are avoidable over three letters. Here "unavoidable" means that every long enough word has a square. On the contrary, avoidable means that there are infinite square-free words. So one may ask for the minimal avoidable repetition or (almost) equivalently for the maximal unavoidable repetition over a fixed k letter alphabet. Denote the maximal unavoidable repetition over k letters by s(k).

If s(k)=r, then every long enough word has a r-th power, and there is an infinite word with no factor of the form wa with w an r-th power and a the first letter of w. The Thue-Morse sequence shows that s(2)=2 (since squares are unavoidable over 2 letters). Over three letters, squares are avoidable. So s(3)<2. It has been shown by F. Dejean [15] that s(3)=7/4. (Indeed, every word over three letters of length 39 contains a 7/4th power !) For four letters, the lower bound $s(4) \geqslant 7/5$ has been verified by F. Dejean and her conjecture that this bound is sharp has been proved by Pansiot [40] . For more than four letters, the precise value of s(k) is unknown. F. Dejan shows that $s(k) \geqslant k/(k-1)$ and conjectures that this is the right value.

The interesting question of constructing efficient algorithms for testing whether a word is squarefree was considered by several authors. The naive algorithm derived from the pattern matching algorithm is in time $O(n^2)$ for words of length n. The first significant improvement was made by Main, Lorentz [32] who proposed an O(nlog n) algorithm for testing square-freeness. A linear algorithm, based on the suffix tree of Weiner [56] and McCreight [34] is given by Crochemore [13] . (See also Slisenko [53]).

Another related problem is to determine ALL repetitions in a word. Considering for instance the Fibonacci word (a finite left factor of it), one cas see that there are "many" repetitions, even if one restricts to maximal repetitions, i.e. those which cannot be extended, neither to the left nor to the right : there are O(nlog n) in this left factor of length n.

THEOREM. - *There is an algorithm to compute all powers in a word of length* n *in time* O(n log n).

There are at least three different proofs of this result. By alphabetic order, Apostolico, Preparata [3] give an algorithm using suffix trees. Crochemore [11] uses partitioning in his algorithm. Main and Lorentz' paper [33] contains an extension of the divide-and-conquer method of their previous paper.

There still remain several open problems. First, to give linear algorithms for testing cubefreeness etc. Then, to give efficient algorithms (if they exist) for testing abelian squarefreeness.

3. The language of squarefree words.

The study of the language of squarefree words has produced new insights in formal language theory. Indeed, the special form of words of the complement of this language implies that standard techniques cannot be applied.

More precisely, it is a straightforward consequence of the pumping lemma that the set of squarefree words over a fixed alphabet is not context-free. Consider the complement of this set, that is the language of words containing a square. This set is not rational. However, the standard pumping technique obviously fails in this case, since any strict power of a word is in the language. The question whether the language of words containing powers is context-free was asked by Autebert, Beauquier, Boasson, Nivat [1]. The answer is no :

THEOREM. - *The set of words containing a square is not context-free.*

Two different proofs of this result exist, one by Ehrenfeucht, Rozenberg [20] (also [19]) , and the other one by Ross and Winklmann [44] . The proofs are quite different, the first shows that there is no EOL-language separating squarefree words from the set of words containing squares, the second proof uses an argument on pushdown automata.

The second technique is more developed in the socalled Interchange Lemma for context-free languages of Ogden, Ross and Winklmann [38] . This lemma was used by Main in the proof of the following result.

THEOREM (Main [31]). - *The set of words over an alphabet of at least 16 letters containing an abelian square is not context-free.*

An abelian square is a word uv, such that v is a permutation of u. It is known that there exist infinite words without abelian squares over a five letter alphabet (Pleasants [41]). It is easily seen by inspection that any word of length 8 over 3 letter contains an abelian square. It is open whether there is an infinite word without abelian square over a four letter alphabet.

OPEN PROBLEMS : There are several questions which seem still to be open, concerning the language of words containing squares. A conjecture by Ehrenfeucht, Haussler and Rozenberg [16] says that any context-free language which contains all square-containing words is cofinite. Another question concerns transformations that maintain the separation of square-free and square-containing words : squarefree morphisms have this property, but also the reversal function. Are there other transformations of this kind ? The set of words containing cubes presumably is also noncontext-free. The same should hold for the set of words containing overlapping factors.

4. How many words are squarefree ?

There is an interesting question which remained open for some time and which was solved recently : are there "many" squarefree words ?

Consider a three letter alphabet. Since there is some difficulty involved with constructing squarefree words, one may suppose that there are only "few"" such words. In other terms, the number c(n) of squarefree words of length n on three letters may grow as a polynomial. The following table gives the first values of c(n), they are token from Brandenburg [5] who gives values up to 24 :

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
c(n)	3	6	12	18	30	42	60	78	108	144	204	264	342	456

They seem to grow rather slowly. However, there is a surprising result due to Brandenburg showing that the number c(n) grows exponentially :

THEOREM (Brandenburg [5]). - *Let* c(n) *be the number of squarefree words of length* n *on a three letter alphabet. Then*

$$6 * c_1^n \leqslant c(n) \leqslant 6 * c_2^n$$

where $c_1 \geqslant 1.032$ *and* $c_2 \leqslant 1.38$.

The proof goes approximatively as follows. Take any squarefree word of length k over three letter a,b,c and replace any letter by itself and by a primed copy in all possible ways. This gives exactly 2^k squarefree words of length k over a six letter alphabet a,a',b,b',c,c'. Next map these words back into a three letter alphabet by a morphism that preserves squarefreeness. Such a morphism exists. Each of the six letters is mapped onto a word of length 22. Moreover, the morphism is injective, and consequently there are at least 2^k squarefree words of length 22k over a three letter alphabet. Since

$$2^{(1/22)} = 1.032...$$

this gives the lower bound. The upper bound is obtained by observing that each squarefree word w can be extended to the right by at most c(n) words of length n. Thus $c(n+m) \leqslant c(n) + c(m)$, from which the conclusion follows by taking n=22.

There still remains a gap between the upper and the lower bounds, but the very precise value is not so important. There is also a similar proof of the result by Brinkhuis [6].

An analoguous proof shows that the number of cubefree words over a two letter alphabet also grows exponentially. In contrast, there is a very interesting polynomial bound on the number of overlap-free words :

THEOREM (Restivo, Salemi [42]). - *There is a constant* C *such that the number* p(n) *of overlap-free words of length* n *over a two letter alphabet satisfies*

$$p(n) \leqslant C.n^{\log 15}$$

The proof is based on a clever factorization of overlap-free words into factors which are the initial factors of length 4^n of the two letter Thue-Morse sequence and those obtained by exchanging a and b. Each overlap-free word is shown to have a unique factorization of this kind. A computation of all possible factorizations for words of length n then gives the upper bound.

It remains to investigate the tree of squarefree words in more detail. This tree is obtained by assigning a node to each squarefree word and by connecting the node of a word to the node for each extension by a letter added on the right. Since there are infinitely many squarefree words, this tree is infinite. Therefore, there are infinite paths in it (Konig's lemma). But there are also finite branches in it, as for example abacaba. These correspond to maximal squarefree words which cannot be extended by any of the three letters. These right-maximal squarefree words were described by Li [29] : they have exactly the expected form namely (over three letters) :

wvuabuacvuabua

provided they are squarefree. They are derived from the simplest of them, abacaba, by inserting a word u before the a's, a word v before the uabua's and w in front.

It was shown by Kakutani (see [21]) that there are uncountably many infinite squarefree words over three letters. So one may ask "where" these words are in the tree : more precisely, is the tree uniform in some sense ? One could imagine indeed that there are infinite paths in the tree where all leaving paths are finite, yielding a "sparse" infinite branch. That this cannot happen was proved by Shelton and Soni.

THEOREM (Shelton, Soni [50,51]-. - *The set of infinite squarefree words over three letters is perfect.*

This statement means that if there is an infinite word going through a node of the tree, then this infinite word will eventually split into two (and therefore into infinitely many infinite squarefree words. There is a related result which says that one must not walk too much in the tree to find an infinite path.

THEOREM (Shelton, Soni [51]). - *There is a constant* K *such that if* u *is a squarefree finite word on a three letter alphabet of length* n *and if* u *can be extended to a squarefree word* uv *of length* $n + K*n^{3/2}$, *then* u *can be extended to an infinite squarefree word.*

5. Squarefree morphisms.

The first, and up to now the only technique to construct squarefree words which was systematically investigated are morphisms. The method goes as follows. First, a endomorphism is iterated, giving an infinite set of words (which can also be considered as an infinite word). Then a second morphism is applied to the set (infinite word). If everything is conveniently choosen, the result is squarefree.

This technique was used already by Axel Thue [55] to compute the first infinite squarefree word. Of course, there exist infinite squarefree words which cannot be constructed this way, since there are uncountably many of these words. However, the method still is very useful. The sets of words squarefree or not, obtained by morphism, have interesting combinatorial properties. Among these, their "subword complexity". See Ehrenfeucht et al. [17,18,19] .

One of the basic questions asked in this context is whether a given morphism

$$h : A^* \longrightarrow B^*$$

is squarefree. By definition, h is a squarefree morphism if h preserves squarefree words, i.e. if the image h(w) is a squarefree word whenever w is squarefree.

Examples : The morphism of Thue [55]

$$h(a) = abcab \quad , \quad h(b) = acabcb \quad , \quad h(c) = acbcacb$$

is squarefree. The following morphism (see Hall [23] ,Istrail [26]

$$h(a) = abc \quad , \quad h(b) = ac \quad , \quad h(c) = a$$

is not squarefree since h(abc) = abcacabc.

The last morphism is too "simple" to be squarefree. Indeed, A. Carpi [8] has shown that a squarefree morphism over three letters must have size at least 18. Here the size is the sum of the lengths of the images of the letters. Thue's morphism given above has size 18 , so it is (already) optimal. The second morphism has only size 6.

Several people have investigated squarefreeness of morphisms, and have derived conditions that ensure that they are. The most precise description is that given by Crochemore :

THEOREM (Crochemore [12,13]). - *Let* $h:A^* \longrightarrow B^*$ *be a morphism, with* A *having at least three letters. Then* h *is squarefree iff the two following conditions hold :*

i) h(x) *is squarefree for squarefree words* x *in* A *of length* 3 ;

ii) No h(a) , *for* a *in* A, *contains a internal presquare.*

Roughly speaking, a presquare is a factor u of h(a) such that h(ax) or h(xa) contains the square uu for some word x for which ax resp. xa is itself squarefree. Another condition is given in Ehrenfeucht, Rozenberg [19].

The theorem implies that it is decidable whether a morphism is squarefree : it suffices to test squarefreeness for long enough words. The following bound is derived by Crochemore from his theorem. For an nonerasing morphism $h : A^* \longrightarrow B^*$, define

$$m(h) = \min \{ |h(a)| : a \text{ in } A \}$$
$$M(h) = \max \{ |h(a)| : a \text{ in } A \} \quad .$$

PROPOSITION (Crochemore [ibid.]). - *Let* $h : A^* \longrightarrow B^*$ *be a morphism. Then* h *is squarefree iff* h(x) *is squarefree for all squarefree words* x *of length* $k = \max \{3, \lceil (M(h)-3)/m(h) \rceil + 1\}$.

Examples show that this bound is sharp. For uniform morphisms, i.e. when M(h) = m(h), the bound is 3. The nicest corollary of Crochemore's theorem is perhaps the following

THEOREM (Crochemore [ibid.]). - *Let* $h : A^* \longrightarrow B^*$ *be a morphism, with* A *a three letter alphabet. Then* h *is squarefree iff* h(x) *is squarefree for all words* x *of lengh 5.*

Such an explicit bound which does not depend on the morphism cannot exist for bigger alphabets. Crochemore gives counterexamples. The results for higher powers than 2 are not yet so complete. I quote two of them which are particularly beautiful. The first concerns cube-free words generated by iterating a morphism over a two letter alphabet.

THEOREM (Karhumaki [27]). - *Let* $h : A^* \longrightarrow A^*$ *be a morphism over a two letter alphabet, such that* h(a) *starts with an* a. *Then the infinite word* $h^{\omega}(a)$ *is cube-free iff the tenth power* $h^{10}(a)$ *is cube-free.*

Another result concerns power-free morphisms. A morphism h is called power-free if for all $k \geqslant 2$, h(w) is k-th power free for all k-th power free words w.

THEOREM (Leconte [28]). - *Let* h : $A^* \longrightarrow B^*$ *be a morphism. Then* h *is power-free iff* h *is squarefree and if* h(aa) *is cube-free for each letter* a *in* A.

The situation for overlap-free morphisms is different : there are (essentially) only two such morphisms ! A more general result was proved by Seebold. Recall first that the socalled Morse morphism (rediscovered independently by Morse [36] after Thue [54]) is defined by :

$$m(a) = ab \quad ; \quad m(b) = ba.$$

Pansiot [39] has observed that the only morphisms generating the Thue-Morse word are powers of m. This was extended by Seebold to :

THEOREM (Seebold [47,48,49]). - *Let* x *be an infinite overlap-free word over the alphabet* {a,b} *that is generated by iterating some morphism* h. *Then* h *is a power of* m.

The following is proved by Harju :

THEOREM (Harju [24]). - *If* h : $\{a,b\}^* \longrightarrow \{a,b\}^*$ *is an overlap-free morphism, then either* h *is a power of* m, *or* h *is the product of the morphism that exchanges* a *and* b *and of a power of* m.

This shows that there are only very few non overlapping morphisms over two letters. Harju characterizes also cyclically non overlapping words and morphisms and asks for a similar characterization of cyclically square-free words.

*****

I gratefully acknowledge helpful discussions with M. Crochemore.

REFERENCES.

1 AUTEBERT J. , J. BEAUQUIER, L. BOASSON, M. NIVAT, Quelques problèmes ouverts en théorie des langages algébriques. RAIRO Informatique théorique 13 (1979), 363-379.

2 ADJAN S., The Burnside problem and identities in groups", Ergeb. Math. Grenzgeb. vol. 95, Springer 1979.

3 APOSTOLICO A., F. PREPARATA, Optimal off-line detection of repetitions in a string, Theor. Comp. Sci. 22 (1983), 297-315.

4 BEAN D. , A. EHRENFEUCHT , G. McNULTY, Advoidable patterns in strings of symbols, Pacific J. Math. 85 (1979), 261-294.

5 BRANDEBURG F., Uniformly growing K-th power-free homomorphisms, Theor. Comput. Sci. 23 (1983), 69-82.

6 BRINKHUIS J. Non-repetitive sequences on three symbols, Quart. J. Math. Oxford (2) 34 (1983), 145-149.

7 BRZOZOWSKI J. , K. CULIK II , A. GABRIELIAN, Classification of noncounting events, J. Comp. Syst. Sci. 5 (1971), 41-53.

8 CARPI A. , On the size of a square-free morphism on a three letter alphabet, Inf. Proc. Letters 16 (1983), 231-236.

9 CERNY A., On generalized words of Thue-Morse, Techn. Report, Université Paris VI, L.I.P.T. 83-84.

10 CHRISTOL C., T. KAMAE, M. MENDES-FRANCE, G. RAUZY, Suites algébriques automates et substitutions, Bull. Soc. Math. France 108 (1980), 401-419.

11 CROCHEMORE M., An optimal algorithm for computing the repetitions in a word, Inf. Proc. Letters, 12 (1981), 244-250.

12 CROCHEMORE M. , Sharp characterizations of squarefree morphisms, Theor. Comp. Sci. 18 (1982), 221-226.

13 CROCHEMORE M. Mots et morphismes sans carré, Annals of Discr. Math. 17 (1983), 235-245.

14 CROCHEMORE M., Recherche linéaire d'un carré dans un mot, C. R. Acad. Sci. Paris, 296 (1983) , 781-784.

15 DEJEAN F., Sur un théorème de Thue, J. Combinatorial Theory 13 (1972), 90-99.

16 EHRENFEUCHT A., D. HAUSSLER, G. ROZENBERG, Conditions enforcing regularity of context-free languages, Techn. Report, Boulder University, 1982.

17 Ehrenfeucht A., K. LEE, G. ROZENBERG, Subword complexities of various classes of deterministic developmental languages without iteraction, Theor. Comput. Sci. 1 (1975), 59-75.

18 EHRENFEUCHT A., G. ROZENBERG, On the subword complexity of square-free DOL languages, Theor. Comp. Sci. 16 (1981), 25-32.

19 EHRENFEUCHT A., G. ROZENBERG, Repetitions in homomorphisms and languages, 9th ICALP Symposium, Springer Lecture Notes in Computer Science 1982 192-196.

20 EHRENFEUCHT A., G. ROZENBERG, On the separating power of EOL systems RAIRO Informatique 17 (1983), 13-22.

21 GOTTSCHALK W., G. HEDLUND, "Topological Dynamics", American Math. Soc. Colloq. Pub. Vol. 36, 1955.

22 GOLDSTINE J., Bounded AFLs, J. Comput. Syst. Sci. 12 (1976), 399-419.

23 HALL M. , Generators and relations in groups - the Burnside problem in : T.L. Saaty (ed.), "Lectures in Modern Mathematics", Vol. II, John Wiley & Sons, 1964 , 42-92.

24 HARJU T., Morphisms that avoid overlapping, University of Turku May 1983.

25 HEDLUND G. , Remarks on the work of Axel Thue on sentences, Nord. Mat. Tidskr 16 (1967), 148-1.

26 ISTRAIL S., On irreductible languages and non rational numbers, Bull. Mat. Soc. Sci. Mat. R. S. Roumanie 21 (1977), 301-308.

27 KARHUMAKI J. , On cube-free ω-words generated by binary morphisms, Discr. Appl. Math. 5 (1983), 279-297.

28 LECONTE M., A fine characterization of power-free morphisms, Techn. Report L.I.T.P. 1984.

29 LI S. , Annihilators in non repetitive semi groups, Studies in Appl. Math. 55 (1976), 83-85.

30 LOTHAIRE M., "Combinatorics on Words", Addison-Wesley 1983.

31 MAIN M., Permutations are not context-free : an application of the "Interchange Lemma", Inf. Proc. Letters 15 (1982), 68-71.

32 MAIN M. , R. LORENTZ, An O(n log n) algorithm for recognizing repetition, Washington State University, Techn. Report CS-79-056.

33 MAIN M., R. LORENTZ, An O(n log n) algorithm for finding all repetitions in a string, J. Algorithms, 1983, to appear.

34 McCREIGHT E. A space-economical suffix tree construction algorithm J. Assoc. Mach. Comp. 23 (1976) 262-272.

35 MINSKY S., "Computations : finite and infinite machines" , Prentice-Hall 1967.

36 MORSE M. , Recurrent geodesics on a surface of negative curvature, Trans. Amer. Math. Soc. 22 (1921), 84-100.

37 MORSE M. , G. HEDLUND, Unending chess, symbolic dynamics and a problem in semigroups, Duke Math. J. 11 (1944), 1-7.

38 OGDEN W., R. ROSS, K. WINKLMANN, Ann "Interchange lemma" for context-free languages, Washington State University, Techn. Report CS-81-080.

39 PANSIOT J., The Morse sequence and iterated morphisms, Inf. Proc. Letters 12 (1981), 68-70.

40 PANSIOT J., A propos d'une conjecture de F. Dejean sur les répétitions dans les mots, Discrete Appl. Math.

41 PLEASANTS P., Nonrepetitive sequences, Proc. Cambridge Phil. Soc. 68 (1970) 267-274.

42 RESTIVO A., S. SALEMI, On weakly square free words, Inf. Proc. Letters, to appear.

43 REUTENAUER C., A new characterization of the regular languages, 8th ICALP, Springer Lecture Notes in Computer Science 115, 1981, 175-183.

44 ROSS R., K. WINKLMANN, Repetitive strings are not context-free, RAIRO Informatique théorique 16 (1982), 191-199.

45 SALOMAA A., Morphisms on free monoids and language theory, in : R. Book (ed.) : "Formal language theory : perpectives and open problems", Academic Press 1098, 141-166.

46 SALOMAA A., "Jewels of Formal Language Theory", Pitman 1981.

47 SEEBOLD P., Morphismes itérés, mot de Morse et mot de Fibonacci, C.R. Acad. Sci. Paris, 295 (1982), 439-441.

48 SEEBOLD P., Sur les morphismes qui engendrent des mots infinis ayant des facteurs prescrits, 6. GI Tagung Theoretische Informatik Lecture Notes Comp. Sci. 145, 1983, 301-311.

49 SEEBOLD P., Sequences generated by infinitely iterated morphisms Techn. Report Université de Paris VII, L.I.T.P. 83-13.

50 SHELTON R., Aqperiodic words on three symbols, I, II, J. Reine Angew, Math. 321 (1981), 195-209, 327 (1981), 1-11.

51 SHELTON R., R. SONI, Aperiodic words on three symbols III, J. Reine Angew, Math. 330 (1982), 44-52.

52 SHYR J. , A stongly primitive word of arbitrary length ant its applications, Int. J. Comp. Math. A-6 (1970), 165-170.

53 SLISENKO A., Determination in real time of all the periodicities in a word, Soviet Math. Dokl. 21 (1980), 392-395.

54 THUE A., Uber unendliche Zeichenreihen, Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiania (1906), 1-22.

55 THUE A., Uber die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiania (1912), 1-67.

56 WEINER P., Linear pattern matching algorithms, Proc. 14 th Symp. switching automata theory, 1973. 1-11.

Jean BERSTEL
L I T P
Université de Paris VII
2, place Jussieu
75231 PARIS Cedex 05