Available online at www.sciencedirect.com
ScienceDirect
C. R. Acad. Sci. Paris, Ser. I 346 (2008) 813–818
Comptes Rendus Mathematique
http://france.elsevier.com/direct/CRASS1/

Combinatorics

On the triplex substitution – combinatorial properties ☆

Bo Tan [a], Zhi-Xiong Wen [a], Yiping Zhang [b]

[a] *Department of Mathematics, Huazhong University of Science and Technology, Wuhan 430074, PR China*
[b] *Department of Mathematics, Wuhan University, Wuhan 430072, PR China*

Received 18 February 2008; accepted after revision 18 June 2008

Available online 24 July 2008

Presented by Jean-Pierre Kahane

Abstract

If a substitution τ over a three-letter alphabet has a positively linear complexity, that is, $P_\tau(n) = C_1 n + C_2$ $(n \geqslant 1)$ with $C_1, C_2 \geqslant 0$, there are only 4 possibilities: $P_\tau(n) = 3$, $n+2$, $2n+1$ or $3n$. The first three cases have been studied by many authors, but the case $3n$ remained unclear. This leads us to consider the triplex substitution $\sigma : a \mapsto ab$, $b \mapsto acb$, $c \mapsto acc$. Studying the factor structure of its fixed point, which is quite different from the other cases, we show that it is of complexity $3n$. We remark that the triplex substitution is also a typical example of invertible substitution over a three-letter alphabet. ***To cite this article: B. Tan et al., C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

Résumé

Sur la substitution triplexe – propriétés combinatoires. Si une substitution τ sur un alphabet de trois lettres a une complexité positivement linéaire, c'est-à-dire $P_\tau(n) = C_1 n + C_2$ $(n \geqslant 1)$ où $C_1, C_2 \geqslant 0$, alors il n'y a que quatre possibilités : $P_\tau(n) = 3$, $n+2$, $2n+1$ ou $3n$. Les trois premiers cas ont été étudiés par différents auteurs, mais le cas $3n$ reste non entièrement élucidé. Nous considérons donc la substitution triplexe $\sigma : a \mapsto ab$, $b \mapsto acb$, $c \mapsto acc$. Analysant la structure des facteurs de son point fixe nous montrons que sa complexité est $3n$. La substitution triplexe est un exemple typique de substitution inversible sur un alphabet de trois lettres. ***Pour citer cet article : B. Tan et al., C. R. Acad. Sci. Paris, Ser. I 346 (2008).***

Version française abrégée

Nous considérons la substitution $\sigma = (ab, acb, acc)$ (c'est-à-dire, $a \mapsto ab$, $b \mapsto acb$, $c \mapsto acc$), appelée substitution triplexe. Elle a un point fixe unique $\xi = \xi_1\xi_2\xi_3\cdots = abacbabaccacb\cdots$.

Le but est de démontrer que ξ a pour complexité $3n$. Pour cela on analyse ses facteurs spéciaux :

Lemme 1. *Pour $n \geqslant 1$, les facteurs suivants sont spéciaux : $b\sigma^n(b)b^{-1}$, $c\sigma^n(b)b^{-1}$, $b\sigma^n(ab)b^{-1}$, $c\sigma^n(ab)b^{-1}$, $\sigma^{n-1}(accab)b^{-1}$, $\sigma^n(bab)b^{-1}$.*

☆ Research supported by NSFC No. 10501035, 10631040, 10571140 and 10671150.
E-mail addresses: bo_tan@163.com (B. Tan), wen.gong@263.net (Z.-X. Wen), ypzhang@whu.edu.cn (Y. Zhang).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2008.06.013

On obtient ainsi un arbre de facteurs spéciaux (voir Fig. 1), qui démontre le :

Lemme 2. *Pour $k \geqslant 2$, il existe au moins 3 facteurs spéciaux de longueur k.*

Pour obtenir une égalité, on contrôle le nombre de facteurs spéciaux de longueur $|A_n|$:

Proposition 1. *Le nombre de facteurs spéciaux de longueur $|A_n|$ est au plus $3|A_n|$.*

Par une méthode d'interpolation, la proposition et les lemmes précédents nous permettent de démontrer

Théorème 1. *La complexité de ξ est $P(n) = 3n$ $(n \geqslant 1)$.*

En particulier, l'arbre de la Fig. 1 est complet : il donne tous les facteurs spéciaux de ξ. Par conséquent, le language engendré par ξ ou σ est complètement déterminé.

1. Introduction

The study of substitutions over a finite alphabet plays an important rôle in finite automata, symbolic dynamics, formal languages, number theory, and fractal geometry, and it has various applications to quasi-crystals, computational complexity, information theory, ... (see [2,3,6,12] and the references therein). In addition, substitutions are fundamental objects in combinatorial group theory [8,9].

Given a sequence $\xi = \xi_1\xi_2\xi_3\cdots$ over some finite alphabet $\mathcal{A}$, with $\xi_i \in \mathcal{A}$. We denote by $\mathcal{L}_n(\xi)$ the set $\{\xi_i \cdots \xi_{i+n-1} \mid i \geqslant 1\}$ of factors of ξ with length n $(n \geqslant 1)$. The set $\mathcal{L}_\xi = \bigcup_{n\geqslant 1} \mathcal{L}_n(\xi)$ is called the language of ξ, and the function $P_\xi(n) := \#\mathcal{L}_n(\xi)$ is called the complexity of ξ.

Let $\mathcal{A}^*$ be the free monoid generated by $\mathcal{A}$ (its identity is the empty word ε). A morphism $\sigma : \mathcal{A}^* \to \mathcal{A}^*$ is called a substitution. Denote by ξ_σ any one of the fixed points of σ (that is $\sigma(\xi_\sigma) = \xi_\sigma$), if it exists.

The study of complexity and substitutions has a long history. Here are some classical results:

- $P_\xi(n) \leqslant n$ for some n if and only if ξ is ultimately periodic, and in this case the complexity is bounded [10];
- A sequence ξ over a two-letter alphabet with complexity $P_\xi(n) = n + 1$ is called Sturmian. There are many equivalent characterizations of Sturmian sequences (e.g., see [12,14,16]);
- Rote [13] constructed a class of sequences with complexity $2n$ by using graphs;
- Mossé [11] studied the case of q-automata (substitutions of constant length). A method to compute $P(n)$ with linear recurrence formulas was given under some technical conditions;
- Over a three-letter alphabet, a class of Tribonacci type substitutions with complexity $2n + 1$ was introduced by Arnoux and Rauzy [4].

However, the complexity of substitutions over a three-letter alphabet presents much more complex phenomena. Few examples can be explicitly worked out, even for invertible substitutions. This is because the structure of invertible substitutions over a three-letter alphabet is quite different from the case of substitutions over a two-letter alphabet: In [17] we showed that the set of invertible substitutions over a three-letter alphabet is not finitely generated. Nevertheless in [15] we were able to characterize the structure of invertible substitutions.

Now notice that for a primitive substitution, the corresponding complexity $P(n)$ satisfies a linear inequality $P(n) \leqslant C_1 n + C_2$ $(n \geqslant 1)$ for some positive constants C_1 and C_2. If we confine ourselves to the case of a three-letter alphabet, and if the above inequality becomes equality for all $n \geqslant 1$, there are clearly only four possibilities: (1) $P(n) = 3$; (2) $P(n) = n + 2$; (3) $P(n) = 2n + 1$; (4) $P(n) = 3n$. The first case is trivial (periodic), the second case (Sturmian-like) was studied in [1]. Arnoux and Rauzy discussed the case (3) in [4]. For the case (4), as far as we know, the existence of such substitution is not present in the literature yet! This is one of the motivations of this Note. So we consider the substitution: $\sigma = (ab, acb, acc)$, that is, $a \mapsto ab$, $b \mapsto acb$, $c \mapsto acc$. We call it the *triplex substitution*.

We remark that the triplex substitution, which can be seen as a representative of undecomposable invertible substitutions (remark that the inverse of σ is $a \mapsto ab^{-1}ab^{-1}c$, $b \mapsto c^{-1}ba^{-1}b$, $c \mapsto c^{-1}ba^{-1}c$), plays an important rôle in the study of invertible substitutions over a three-letter alphabet [17,15]. In this Note, we show that its complexity