Advances in Applied Mathematics 50 (2013) 737–748

Contents lists available at SciVerse ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

On palindromic factorization of words

A.E. Frid [a,*,1], S. Puzynina [a,b,2], L.Q. Zamboni [c,b,3]

[a] *Sobolev Institute of Mathematics, 4 Koptyug av., 630090, Novosibirsk, Russia*
[b] *Department of Mathematics and Turku Centre for Computer Science, University of Turku, 20014 Turku, Finland*
[c] *Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, F69622 Villeurbanne Cedex, France*

ARTICLE INFO

Article history:
Received 12 November 2012
Accepted 4 January 2013
Available online 29 January 2013

MSC:
68R15

Keywords:
Palindrome
Periodicity of words
Complexity of words

ABSTRACT

Given a finite word u, we define its *palindromic length* $|u|_{\text{pal}}$ to be the least number n such that $u = v_1 v_2 \dots v_n$ with each v_i a palindrome. We address the following open question: let P be a positive integer and w an infinite word such that $|u|_{\text{pal}} \leqslant P$ for every factor u of w. Must w be ultimately periodic? We give a partial answer to this question by proving that for each positive integer k, the word w must contain a k-power, i.e., a factor of the form u^k. In particular, w cannot be a fixed point of a primitive morphism. We also prove more: for each pair of positive integers k and l, the word w must contain a position covered by at least l distinct k-powers. In particular, w cannot be a Sierpinski-like word.

1. Introduction

Let A be a finite non-empty set, and let A^+ denote the set of all finite non-empty words in A. A word $u = u_1 u_2 \dots u_n \in A^+$ is called a *palindrome* if $u_i = u_{n-i+1}$ for each $i = 1, \dots, n-1$. In particular each $a \in A$ is a palindrome. We also regard the empty word as a palindrome.

Palindrome factors of finite or infinite words have been studied from different points of view. In particular, Droubay, Justin and Pirillo [4] proved that a word of length n can contain at most $n+1$ distinct palindromes, which gave rise to the theory of *rich* words (see [5]). The number of palindromes of a given length occurring in an infinite word is called its *palindrome complexity* and is bounded by a function of its usual subword complexity [1]. However, in this paper we study palindromes in an infinite word from the point of view of decompositions.

* Corresponding author.
E-mail addresses: anna.e.frid@gmail.com (A.E. Frid), svepuz@utu.fi (S. Puzynina), lupastis@gmail.com (L.Q. Zamboni).
1 Supported in part be RFBR grant 12-01-00089 and by the Presidential grant MK-4075.2012.1.
2 Supported in part by grant No. 251371 from the Academy of Finland and by RFBR grant 12-01-00448.
3 Supported in part by a FiDiPro grant from the Academy of Finland and by ANR grant SUBTILE.

0196-8858/$ – see front matter
http://dx.doi.org/10.1016/j.aam.2013.01.002

For each word $u \in A^+$ we define its *palindromic length*, denoted by $|u|_{\mathrm{pal}}$, to be the least number P such that $u = v_1 v_2 \ldots v_P$ with each v_i a palindrome. As each letter is a palindrome, we have $|u|_{\mathrm{pal}} \leqslant |u|$, where $|u|$ denotes the length of u. For example, $|01001010010|_{\mathrm{pal}} = 1$ while $|010011|_{\mathrm{pal}} = 3$. Note that 010011 may be expressed as a product of 3 palindromes in two different ways: (0)(1001)(1) and (010)(0)(11). In [10], O. Ravsky obtains an intriguing formula for the supremum of the palindromic lengths of all binary words of length n. The question considered in this paper is

Question 1. Do there exist an infinite non-ultimately periodic word w and a positive integer P such that $|u|_{\mathrm{pal}} \leqslant P$ for each factor u of w?

We conjecture that such aperiodic words do not exist, but at the moment we can prove it only partially. Namely, in this paper we prove that if such a word exists, then it is not k-power-free for any k and moreover, for all $k > 1$, $l \geqslant 0$ it does not satisfy the (k, l)-condition defined in Section 4. A discussion what exactly the condition means and which class of words should be studied now to give a complete answer to the question is given in Section 5.

A preliminary version of this paper has been reported at Journées Montoises 2012.

2. The case of *k*-power-free words

Let k be a positive integer. A word $v \in A^+$ is called a *k-power* if $v = u^k$ for some word $u \in A^+$. An infinite word $w = w_1 w_2 \ldots \in A^{\mathbb{N}}$ is said to be *k-power-free* if no factor u of w is a k-power. For instance, the Thue–Morse word 0110100110010110... fixed by the morphism $0 \mapsto 01$, $1 \mapsto 10$ is 3-power-free (see for example [7]).

Theorem 1. *Let k be a positive integer and $w = w_1 w_2 \ldots \in A^{\mathbb{N}}$. If w is k-power-free, then for each positive integer P there exists a prefix u of w with $|u|_{\mathrm{pal}} > P$.*

Recall that a word $u_1 \ldots u_n$ is called t-periodic if $u_i = u_{i+t}$ for all i such that $1 \leqslant i \leqslant n - t$.

The proof of Theorem 1 will make use of the following lemmas.

Lemma 2. *Let u be a palindrome. Then for every palindromic proper prefix v of u, we have that u is $(|u| - |v|)$-periodic.*

Proof. If u and v are palindromes with v a proper prefix of u, then v is also a suffix of u and hence u is $(|u| - |v|)$-periodic. □

In what follows, the notation $w[i..j]$ can mean the factor $w_i w_{i+1} \ldots w_j$ of a word $w = w_1 \ldots w_n \ldots$ as well as its precise occurrence starting at the position numbered i; we always specify it when necessary.

Lemma 3. *Suppose the infinite word w is k-power-free. If $w[i_1..i_2]$ and $w[i_1..i_3]$ are palindromes with $i_3 > i_2$, then*

$$\frac{|w[i_1..i_3]|}{|w[i_1..i_2]|} > 1 + \frac{1}{k-1}.$$

Proof. By Lemma 2, the word $w[i_1..i_3]$ is $(i_3 - i_2)$-periodic; at the same time, it cannot contain a k-power, so, $|w[i_1..i_3]| < k(i_3 - i_2)$. Thus,

$$\frac{|w[i_1..i_3]|}{|w[i_1..i_2]|} = \frac{|w[i_1..i_3]|}{|w[i_1..i_3]| - (i_3 - i_2)} > \frac{|w[i_1..i_3]|}{(1 - \frac{1}{k})(|w[i_1..i_3]|)} = 1 + \frac{1}{k-1}. \qquad \square$$