ON HIGHLY PALINDROMIC WORDS

ŠTĚPÁN HOLUB AND KALLE SAARI

Abstract. We study some properties of palindromic (scattered) subwords of binary words. In view of the classical problem on subwords, we show that the set of palindromic subwords of a word characterizes the word up to reversal.

Since each word trivially contains a palindromic subword of length at least half of its length—a power of the prevalent letter—we call a word that does not contain any palindromic subword longer than half of its length *minimal palindromic*. We show that every minimal palindromic word is abelian unbordered.

We also propose to measure the degree of palidromicity of a word w by the ratio $|rws|/|w|$, where the word rws is minimal palindromic and rs is as short as possible. We prove that the ratio is always bounded by four, and construct a sequence of words that achieves this bound asymptotically.

1. Introduction

In this paper we propose a property of binary words, which measures how palindromic they are. For that purpose we investigate palindromes that are (scattered) subwords of a given word. A binary word trivially contains a palindromic subword of length at least half of its length: a power of the prevalent letter. Therefore a word that does not contain any palindromic subword longer that half of its length has the lowest degree of palindromicity, and deserves to be called *minimal palindromic.* Table 1 lists the minimal palindromic words up to length 9 starting with the letter 0. In this paper, we give some properties of such words.

It is less clear, on the other hand, which words should be understood as highly palindromic. Our approach is the following: *a word w is the more palindromic the harder it is to construct a word z that is minimal palindromic and w is its factor.* Therefore we study shortest extensions rws of w that are minimal palindromic, and use the fraction $|rws|/|w|$ as the measure of palindromicity. It is not very difficult to see that this measure always exists and it is bounded by 4, see Theorem 3. Rather surprisingly, this bound is also optimal as we show by constructing a sequence of words which reach the bound asymptotically.

We will also consider some other aspects of palindromic subwords to motivate this notion further. We will show that minimal palindromic subwords are unbordered, even in a stronger, abelian, sense. A classical problem regarding subwords is to characterize a word by a collection of its subwords, see [4, 7, 3, 6, 2, 8]. We will show that a word is characterized, up to reversal, by its palindromic subwords.

The content of the paper follows. In Section 2, we fix the notation and present some definitions. In Section 3, we show that minimal palindromic words are abelian unbordered. In Section 4, we show that a word is essentially characterized by its

The first author is supported by the research project MSM 0021620839 financed by MSMT.
The second author is supported by the Finnish Academy under grant 8206039.

Length	**Minimal Palindromic Words**
1	0
2	01
3	011, 001
4	0011,
5	00011, 00101, 00111, 01011
6	001011, 000111
7	0000111, 0001011, 0001101, 0001111, 0010111, 010011
8	00001111, 00010111
9	000001111, 000010111, 000011011, 000011101, 000011111, 000100111, 000101011, 000101111, 000110111, 000111011, 001000111, 001001111, 001010111, 010001111

Table 1. Minimal palindromic words up to length 9.

palindromic subwords. In Section 5, we present the main result of this paper. Namely, we show that the number 4 is the optimal upper bound for the measure of palindromicity of words. Section 6 concludes the paper with a discussion of some further problems for future consideration.

The notion of *palindromic subwords* of a finite word has, to our knowledge, not been much considered before. For a review on subwords, we refer to [1], and for a more throughout treatment, we refer to [4, Chapter 6].

1ex

2. Definitions

We start by fixing the notation and presenting some definitions. For the notions left undefined, the reader should consult, for example, [1].

Let w and u be words over the alphabet $\{0,1\}$. The word u is a *factor* of w if $w = xuy$ for some words $x, y \in \{0,1\}^*$. The word u is a *subword* of w if there exist words $x_1, y_1, \ldots, x_n, y_n \in \{0,1\}^*$ such that

$$u = y_1 y_2 \cdots y_n \qquad \text{and} \qquad w = x_1 y_1 x_2 y_2 \cdots x_n y_n.$$

Note that factor is often called a subword in the literature; and subword is often called a *scattered* subword. If the subword u of w is a palindrome, then we say that u is a *subpalindrome* of w.

The length of the word w is denoted by $|w|$. For a letter $a \in \{0,1\}$, the symbol $|w|_a$ denotes the number of occurrences of a in w. The reversal of w is denoted by $\tilde{w}$.

If w can be written in the form $w = uxv$, where u and v are nonempty words with $|u|_0 = |v|_0$ and $|u|_1 = |v|_1$, we say that w is *abelian bordered.* Otherwise we say that w is *abelian unbordered.*

We say that a letter $a \in \{0,1\}$ is *prevalent* in a word $w \in \{0,1\}^*$ if

$$|w|_a = \max\{|w|_0, |w|_1\}.$$

Then we have

$$|w|_a \geq \left\lceil \frac{|w|}{2} \right\rceil,$$

and $a^{|w|_a}$ is a subpalindrome of w.

We denote the length of a longest palindromic subword of a word $w \in \{0,1\}^*$ by $\kappa(w)$. Since the word w is over a binary alphabet, it follows that

$$\left\lceil \frac{|w|}{2} \right\rceil \leq \kappa(w) \leq |w|.$$

We say that the word w is *minimal palindromic* if $\kappa(w) = \lceil |w|/2 \rceil$. Observe that the number of letters 0 and 1 in a minimal palindromic word w differ by at most 1, that is,

$$\big|\, |w|_0 - |w|_1 \,\big| \leq 1. \tag{2.1}$$

For an arbitrary word $w \in \{0,1\}^*$, if there exists two words $r, s \in \{0,1\}^*$ such that a word rws is minimal palindromic, then we say that the pair (r,s) is an *MP-extension* of w (an MP-extension always exists, see Theorem 3). If, in addition, the word rws is as short as possible, then we call the pair (r,s) a shortest MP-*extension*, or SMP-extension for short, of w. Finally, the rational number $|rws|/|w|$, where (r,s) is an SMP-extension of w, is called the MP-*ratio* of w. The MP-ratio measures how much a word has to be extended to obtain a minimal palindromic word.

Some additional definitions that are needed only locally will be presented later.

3. Minimal Palindromic Words are Abelian Unbordered

Here we establish a strong unborderedness property of minimal palindromic words.

Theorem 1. *Suppose a word $w \in \{0,1\}^*$ is an abelian bordered word. Then w is not minimal palindromic.*

Proof. Let us write $w = xuy$, where x and y satisfy $|x|_0 = |y|_0$ and $|x|_1 = |y|_1$. Suppose further that $|x|$ is the least possible. Let a denote a prevalent letter of u, and denote $m = |u|_a$.

If x is a letter, then $w = xux$ and $xa^m x$ is a subpalindrome of w. We have

$$\left\lceil \frac{|w|}{2} \right\rceil = \left\lceil \frac{|u|}{2} \right\rceil + 1 \leq |u|_a + 1 < |xa^m x| \leq \kappa(w),$$

and hence w is not minimal palindromic.

Now we may suppose that $|x| \geq 2$. Since the length of x (and y) is minimal, both letters 0 and 1 must occur in x (and y). Let b denote a prevalent letter of x (and y), and denote $l = |x|_b = |y|_b$. The last letter of x differs from the first letter of y. Indeed, otherwise w would have an abelian border of length $|x| - 1$, contradicting the minimality of $|x|$. Therefore the letter a is either a suffix of x or a prefix of y, and consequently the word

$$b^l a^{m+1} b^l$$

is a subpalindrome of w. This implies that

$$\left\lceil \frac{|w|}{2} \right\rceil \leq \left\lceil \frac{|x|}{2} \right\rceil + \left\lceil \frac{|u|}{2} \right\rceil + \left\lceil \frac{|y|}{2} \right\rceil \leq 2l + m < |b^l a^{m+1} b^l| \leq \kappa(w),$$

so that w is not minimal palindromic. □

By negating the previous theorem, we get the following corollary.

Corollary 1. *If a word is minimal palindromic, then it is abelian unbordered.*

As an application of this corollary, we show that already a very short prefix of a word can tell if the word is not minimal palindromic.

Corollary 2. *No word with a proper prefix* 011 *is minimal palindromic.*

Proof. If a word w has a proper prefix 011, then Corollary 1 implies that w has a suffix 111. If $|w| \leq 6$, the word w clearly is not minimal palindromic. If $|w| > 6$, we can write $w = 011u111$, and then w has a subpalindrome $11p11$, where p is a longest subpalindrome of u. Consequently,

$$\kappa(w) \geq 4 + \kappa(u) \geq 4 + \left\lceil \frac{|u|}{2} \right\rceil = 1 + \left\lceil \frac{6 + |u|}{2} \right\rceil > \left\lceil \frac{|w|}{2} \right\rceil .$$

Hence w is not minimal palindromic. □

Incidentally, the word 011 itself is minimal palindromic.

4. Determining a Word by Its Subpalindromes

In this section we show that a word is characterized, up to reversal, by the set of its subpalindromes. Let $\mathcal{P}(w)$ denote the set of all subpalindromes of w.

Theorem 2. *If $w, z \in \{0,1\}^*$ with $\mathcal{P}(w) = \mathcal{P}(z)$, then either $w = z$ or $w = \tilde{z}$.*

Proof. We proof the claim by induction on the length of w. The claim is clear if w is empty or a letter; hence we may assume that $|w| \geq 2$.

Denote $P = \mathcal{P}(w) = \mathcal{P}(z)$, and let m_a, where $a \in \{0,1\}$, be the maximal integer such that $a^{m_a} \in P$. Then, clearly, $|w|_0 = |z|_0 = m_0$ and $|w|_1 = |z|_1 = m_1$.

Let us first suppose that $w = aw'a$, where $a \in \{0,1\}$ and $w' \in \{0,1\}^*$. Then $ab^{m_b}a \in P$, where $b = 1 - a$. This implies that $z = az'a$ for some $z' \in \{0,1\}^*$ and, furthermore,

$$\mathcal{P}(w') = \mathcal{P}(z') = \{\, u \mid aua \in P \,\} .$$

By the induction assumption, either $w' = z'$ or $w' = \tilde{z}'$, and so it follows that either $w = z$ or $w = \tilde{z}$.

Let us then suppose that $a = \mathrm{pref}_1(w) \neq \mathrm{suff}_1(w) = b$, and let i, j be the largest integers such that w can be written in the form $w = a^i w' b^j$. The word w' is either empty, or $\mathrm{pref}_1(w') = b$ and $\mathrm{suff}_1(w') = a$. Then $ab^{m_b}a \notin P$, whence either z or $\tilde{z}$ is of the form $a^{i'} z' b^{j'}$, where z' satisfies the same conditions as w'. The definition of w' and z' implies that we have

$$\mathcal{P}(w') = \mathcal{P}(z') = \{\, u \mid \mathrm{pref}_1(u) = \mathrm{suff}_1(u) = c \text{ and } dud \in P,\ d = 1 - c \,\} \cup \{\epsilon\} .$$

The induction assumption now implies that $w' = z'$. Moreover $i = m_a - |w'|_a = m_a - |z'|_a = i'$. Similarly, we obtain $j = j'$, and the proof is complete. □

5. Extending a Word into a Minimal Palindromic Word

As mentioned in the introduction of this paper, we consider a word w highly palindromic if it is difficult to construct a minimal palindromic word z with w as a factor. That is to say, large MP-ratio corresponds to high palindromicity. In this section we investigate how large the MP-ratio of a word can be. The first result gives an upper bound.

Theorem 3. *The MP-ratio of any word $w \in \{0,1\}^*$ is at most 4.*

Proof. Let us denote
$$z = 0^{|w|+|w|_1} w\, 1^{|w|+|w|_0}.$$
Since w is a factor of z and $|z|/|w| = 4$, it suffices to show that z is minimal palindromic. To do that, let p denote a subpalindrome of z. By symmetry, we may suppose that p starts and ends with the letter 1. Consequently p is a subword of $w1^{|w|+|w|_0}$, and hence either $p \in 1^*$ or $p = r\,01^j$, where $r\,0$ is a subword of w. If the former holds, then clearly $|p| \le 2|w| = \lceil |z|/2 \rceil$. If the latter holds, then the palindomicity of p implies $j \le |r|_1 \le |w|_1$. Therefore
$$|p| = |r\,0| + j \le |w| + |w|_1 \le 2|w| = \left\lceil \frac{|z|}{2} \right\rceil .$$
We have shown that $\kappa(z) = \lceil |z|/2 \rceil$, so that z is minimal palindromic. □

Next we show that the constant 4 in the previous theorem is optimal. Next theorem is the main result of this paper. Denote
$$R(n) = \max\{\, \text{MP-ratio of } w \mid w \in \{0,1\}^*,\ |w| = n \,\}.$$

Theorem 4. *We have*
$$\lim_{n\to\infty} R(n) = 4.$$

To prove the claim, we need some auxiliary definitions and Lemmas 1–4 below.

We say that a word $w \in \{0,1\}^*$ is *k-economic* (with respect to the letter 1), where $k \ge 0$ is an integer, if w is a palindrome and the word $w\,1^k$ contains a subpalindrome of length at least $|w|_1 + k + 2$. Such a subpalindrome can be written in the form $1^m q\, 1^m$, where $m \le k$, the word q is a palindrome, and $1^m q$ is a subword of w. We call the pair (q, m) a *k-witness* of w. Note that the word w may have several k-witnesses. Finally, we say that w is *economic*, if it is k-economic for each integer $k = 0, \ldots, |w|_1$.

The importance of an economic word is that it is highly palindromic, provided it has a high density of letter 1. This follows from the next two lemmas.

Lemma 1. *Suppose that a word $w \in \{0,1\}^*$ is economic. If a pair (r, s) is an MP-extension of w, then $|rs|_1 > |w|_1$.*

Proof. Suppose, contrary to what we want to prove, that $|rs|_1 \le |w|_1$. Denote $|r|_1 = i$ and $|s|_1 = j$. Suppose first that $i \le j$. Now, we have $(j - i) \le |w|_1$, and so the word w is $(j - i)$-economic. Let (q, m) be a $(j - i)$-witness of w. Then the word $1^m q$ is a subword of w, and
$$|1^m q\, 1^m| \ge |w|_1 + j - i + 2.$$
On one hand, since $m + i \le j = |s|_1$, we see that the palindrome $1^{m+i} q\, 1^{m+i}$ is a subword of rws. On the other hand, we have
$$|1^{m+i} q\, 1^{m+i}| \ge |w|_1 + j + i + 2 > |rws|_1 + 1 \ge \left\lceil \frac{|rws|}{2} \right\rceil ,$$
where the last inequality holds because rws is minimal palindromic, see Equation (2.1). But then the palindrome $1^{m+i} q\, 1^{m+i}$ cannot be a subword of rws as it is too long, a contradiction. The case $i > j$ can be proved in the same way, we just have to recall that w is a palindrome. This completes the proof. □

Lemma 2. *Suppose that a word $w \in \{0,1\}^*$ is economic. If a pair (r,s) is an MP-extension of w, then $|rws| > 4|w|_1$.*

Proof. Since rws is minimal palindromic, Equation (2.1) and Lemma 1 implies that

$$|rws| = |rws|_0 + |rws|_1 \geq 2|rws|_1 - 1 = 2|w|_1 + 2|rs|_1 - 1 > 4|w|_1.$$

□

The previous lemma implies that if a word w is economic and the density of the letter 1 in w is large, then w has a large MP-ratio. Therefore we want to find a sequence of economic words with large densities.

Let w_0 be an economic word. We define a sequence $(w_i)_{i\geq 0}$ recursively by

$$w_{i+1} = w_i\, 1^{t_i} w_i \qquad \text{for } i \geq 0, \tag{5.1}$$

where t_i is a positive integer. Our sequence is fully defined by the starting word w_0 and by the sequence $(t_i)_{i\geq 0}$. Note that, since w_0 is a palindrome, the word w_i is a palindrome for all $i \geq 0$.

Lemma 3. *If $t_i < |w_i|_0$ for every integer $i \geq 0$, then the words w_i are economic.*

Proof. We prove the claim by induction on i. The word w_0 is economic by assumption. Suppose now that w_i is economic. We show that w_{i+1} is economic, that is, k-economic for all $k = 0, \ldots, |w_{i+1}|_1$. This is done in three parts as follows.

Suppose first that $0 \leq k \leq |w_i|_1$. By the induction assumption, w_i is k-economic. Let (q,m) be a k-witness of w_i, and denote

$$p = 1^m q\, 1^{t_i+m} q\, 1^m.$$

Since $1^m q$ is a subword of w_i and $w_{i+1} = w_i 1^{t_i} w_i$, we see that the word

$$1^m q\, 1^{t_i+m} q$$

is a subword of w_{i+1}, and therefore p is a subpalindrome of $w_{i+1}1^m$. So, to show that w_{i+1} is k-economic, we only have to show that

$$|p| \geq |w_{i+1}|_1 + k + 2.$$

Since (q,m) is a k-witness of w_i, we have

$$|q| + 2m = |1^m q\, 1^m| \geq |w_i|_1 + k + 2,$$

and since $m \leq k$, we obtain

$$|p| = 2|q| + 3m + t_i \geq 2|q| + 4m - k + t_i \geq 2|w_i| + k + 4 + t_i > |w_{i+1}|_1 + k + 2.$$

Hence w_{i+1} is k-economic.

Suppose then that $|w_i|_1 < k \leq |w_i|_1 + t_i$. Now the word $1^k w_i 1^k$ is a subpalindrome of $w_{i+1}1^k$. Since $t_i < |w_i|_0$ and $|w_i|_1 < k$, we have

$$|1^k w_i 1^k| = 2k + |w_i|_1 + |w_i|_0 \geq 2|w_i|_1 + t_i + k + 2 = |w_{i+1}|_1 + k + 2,$$

and it follows that w_{i+1} is k-economic.

Suppose finally that $|w_i|_1 + t_i < k \leq |w_{i+1}|_1$. Denote $j = |w_i|_1 + t_i$ and $l = k - j$. Since $l \leq |w_i|_1$, the word w_i is l-economic by the induction assumption. Let (q,m) denote an l-witness of w_i. Since the word 1^j is a subword of $w_i 1^{t_i}$, the word $1^m q$ is a subword of w_i, and $m + j \leq k$, we see that the palindrome

$$1^{m+j} q\, 1^{m+j}$$

is a subword of $w_{i+1}1^k$. Furthermore,

$$|1^{m+j}q\,1^{m+j}| \geq 2j + |w_i|_1 + l + 2 = |w_{i+1}|_1 + k + 2,$$

and hence w_{i+1} is k-economic. The proof of Lemma 3 is complete. □

Next we will show that there exist economic words of any sufficiently large length with high density of the letter 1. To do that, we denote by $w(t_0, \ldots, t_{j-1})$, where $j \geq 1$, the word w_j defined by (5.1) with $w_0 = 0000$. If, in addition, the integers t_i's satisfy

$$2^i \leq t_i < 2^{i+2}, \tag{5.2}$$

then we have

$$t_i < 2^{i+2} = |w(t_0, \ldots, t_{i-1})|_0$$

for all $0 \leq i \leq j-1$, and therefore the word $w(t_0, \ldots, t_{j-1})$ is economic by Lemma 3.

Lemma 4. *For every integer $k \geq 448$, there exist a word v_k such that $|v_k| = k$, and $v_k = w(t_0, \ldots, t_{n-1})$ for some $n \geq 6$, and some integers $t_0, \ldots, t_{n-1}$ satisfying the inequality* (5.2).

Proof. To enhance readability, let us denote, for integers $i \geq 0$,

$$\alpha_i = 2^i \qquad \text{and} \qquad \beta_i = 2^{i+2} - 1;$$

then (5.2) is equivalent to $\alpha_i \leq t_i \leq \beta_i$. It is easy to verify that, for all $j \geq 1$, we have

$$|w(t_0, t_1, \ldots, t_{j-1})| = 2^{j+2} + 2^{j-1}t_0 + 2^{j-2}t_1 + \cdots + 2t_{j-2} + t_{j-1},$$

and furthermore,

$$|w(\alpha_0, \ldots, \alpha_{j-1})| = 2^{j-1}(8 + j) \qquad \text{and} \qquad |w(\beta_0, \ldots, \beta_{j-1})| = 2^j(3 + 2j) + 1.$$

Now, a straightforward calculation shows that, for each $j \geq 6$, we have

$$|w(\alpha_0, \ldots, \alpha_j)| < |w(\beta_0, \ldots, \beta_{j-1})|.$$

This implies that, for each $k \geq |w(\alpha_0, \ldots, \alpha_5)| = 448$, there exists an integer $n \geq 6$ such that

$$|w(\alpha_0, \ldots, \alpha_{n-1})| \leq k \leq |w(\beta_0, \ldots, \beta_{n-1})|.$$

It is now enough to verify that integers of the form $|w(t_0, \ldots, t_{n-1})|$, where t_i's satisfy (5.2), cover the whole interval between $|w(\alpha_0, \ldots, \alpha_{n-1})|$ and $|w(\beta_0, \ldots, \beta_{n-1})|$. This can be done inductively by noting that if

$$k = |w(t_0, \ldots, t_i, \beta_{i+1}, \ldots, \beta_{n-1})|$$

with $t_i < \beta_i$, then, as in the usual binary numeration system,

$$k + 1 = |w(t_0, \ldots, t_{i-1}, t_i + 1, \beta_{i+1} - 1, \ldots, \beta_{n-1} - 1)|.$$

This concludes the proof of Lemma 4. □

In the previous lemma, the constant 448 is by no means essential—it could possibly be chosen to be smaller. However, since we are interested in the asymptotical behavior of economic words, the constant suffices for our purposes.

Now, it is easy to see that

$$\frac{|w(\alpha_0, \ldots, \alpha_{j-1})|_1}{|w(\alpha_0, \ldots, \alpha_{j-1})|} \leq \frac{|w(t_0, \ldots, t_{j-1})|_1}{|w(t_0, \ldots, t_{j-1})|}$$

for all j, as soon as the t_i's satisfy inequality (5.2). We calculate

$$\frac{|w(\alpha_0,\ldots,\alpha_{j-1})|_1}{|w(\alpha_0,\ldots,\alpha_{j-1})|} = \frac{2^{j-1}j}{2^{j-1}(j+8)} = \frac{j}{j+8}.$$

Therefore, if we consider the density of the letter 1 in the words v_k guaranteed by Lemma 4, we obtain

$$\lim_{k\to\infty} \frac{|v_k|_1}{|v_k|} = 1.$$

We are now ready to finish the proof. Choose a positive real number ε, and let k_0 denote an integer such that

$$\frac{|v_k|_1}{|v_k|} > 1 - \varepsilon/4$$

for all $k \geq k_0$. Let a pair (r, s) be an SMP-extension of v_k. Then Lemma 2 implies that

$$\frac{|rv_ks|}{|v_k|} > \frac{4|v_k|_1}{|v_k|} > 4 - \varepsilon,$$

that is, the MP-ratio of v_k is at least $4 - \varepsilon$ for $k \geq k_0$. It follows that

$$\lim_{n\to\infty} R(n) = 4,$$

and Theorem 4 is now proved.

6. Future Research

It is interesting to study the structure of words with maximal palindromicity, i.e., with MP-ratio $R(|w|)$ (they should be called *maximal palindromic*). It seems that economic words are best candidates for such maximal property. It is, however, an open question whether each maximal palindromic word is economic, and whether maximal palindromic words can be obtained using our construction. Oddly enough, we do not even know whether maximal palindromic words are always palindromes.

We would also like to ask whether in the definition of the MP-extension it is possible to suppose, without loss of generality, that the extending words r and s are powers of a single letter. Again, we cannot answer even a weaker question, whether the maximal extending words r and s can be always elements of 0^*1^* or 1^*0^*.

Finally, since we are mostly dealing with subwords, not factors, it may be more natural to study minimal palindromic extensions as follows. If w is an arbitrary word, its minimal palindromic extension is a word u such that u is minimal palindromic and it contains w as a subword (i.e., w does not have to be factor of u). Finding an optimal upper bound for the length of a shortest MP extension in this sense is open, but computational evidence suggests that the optimal bound may be strictly less than 4.

References

[1] Choffrut, Christian; Karhumäki, Juhani Combinatorics of words. Handbook of formal languages, Vol. 1, 329–438, Springer, Berlin, 1997.

[2] Dudík, Miroslav; Schulman, Leonard J. Reconstruction from subsequences. *J. Combin. Theory Ser.* A 103 (2003), no. 2, 337–348.

[3] Krasikov, Ilia; Roditty, Yehuda On a Reconstruction Problem for Sequences, *J. Comb. Theory, Ser. A* 77(2): 344-348 (1997)

[4] Lothaire, M. Combinatorics on words. Encyclopedia of Mathematics and its Applications, 17. *Addison-Wesley Publishing Co., Reading, Mass.*, 1983.

[5] Lothaire, M. Algebraic combinatorics on words, Encyclopedia of Mathematics and its Applications, 90. *Cambridge University Press*, Cambridge, 2002.

[6] Maňuch, Ján Characterization of a word by its subwords. *Developments in language theory (Aachen, 1999)*, 210–219, *World Sci. Publ., River Edge, NJ,* 2000.

[7] Manvel, B.; Meyerowitz, A.; Schwenk, A.; Smith, K.; Stockmeyer, P. Reconstruction of sequences. *Discrete Math.* 94 (1991), no. 3, 209–219.

[8] Salomaa, Arto Connections between subwords and certain matrix mappings. *Theoret. Comput. Sci.* 340 (2005), no. 2, 188–203.

Department of Algebra, Charles University, Sokolovská 83, 175 86 Praha, Czech Republic

E-mail address: holub@karlin.mff.cuni.cz

Department of Mathematics and Turku Centre for Computer Science, University of Turku, 20014 Turku, Finland

E-mail address: kasaar@utu.fi