Discrete Mathematics and Theoretical Computer Science **4**, 2000, 001–010

Sums of Digits, Overlaps, and Palindromes

Jean-Paul Allouche[1] and Jeffrey Shallit[2†]

[1] *CNRS, Laboratoire de Recherche en Informatique, Bâtiment 490, F-91405 Orsay Cedex, France*
[2] *Department of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

received August 3, 1999, *revised February 8, 2000*, *accepted February 16, 2000*.

Let $s_k(n)$ denote the sum of the digits in the base-k representation of n. In a celebrated paper, Thue showed that the infinite word $(s_2(n) \bmod 2)_{n\geq 0}$ is *overlap-free*, i.e., contains no subword of the form $axaxa$, where x is any finite word and a is a single symbol. Let k, m be integers with $k \geq 2$, $m \geq 1$. In this paper, generalizing Thue's result, we prove that the infinite word $\mathbf{t}_{k,m} := (s_k(n) \bmod m)_{n\geq 0}$ is overlap-free if and only if $m \geq k$. We also prove that $\mathbf{t}_{k,m}$ contains arbitrarily long squares (i.e., subwords of the form xx where x is nonempty), and contains arbitrarily long palindromes if and only if $m \leq 2$.

Keywords: sum of digits, overlap-free sequence, palindrome

1 Introduction

At the beginning of the 20th century, the Norwegian mathematician Axel Thue initiated the study of what is now called combinatorics on words with his results on repetitions in words [18, 19, 6, 8]. We say a nonempty word w is a *square* if it can be written in the form xx for some word x. Examples include the words `chercher` in French and `murmur` in English. We say that w is an *overlap* if it can be written in the form $axaxa$ for some word x and single symbol a. Examples include the words `entente` in French and `alfalfa` in English. Thue explicitly constructed an infinite word on two symbols that is *overlap-free*, that is, contains no subword that is an overlap. He also constructed an infinite word on three symbols that is *square-free*, that is, contains no subword that is a square.

Thue's constructions are based on what is now called the *Thue-Morse sequence*

$$\mathbf{t} = t(0)\, t(1)\, t(2)\ \cdots = 0110100110010110\cdots.$$

There are many alternative ways to define this sequence (see, for example, [3]), one being as the fixed point, starting with 0, of the morphism $h(0) = 01$, $h(1) = 10$. One can also define $\mathbf{t}$ in terms of sums of digits. We define $s_k(n)$ to be the sum of the digits in the base-k representation of n. Then $t(n) = s_2(n) \bmod 2$ for all $n \geq 0$. Thue proved that $\mathbf{t}$ is overlap-free.

Since Thue's pioneering work, many other investigators have studied overlap-free words and their generalizations. For example, Thue's construction was rediscovered by Morse, who used it in a construction

[†] Research supported in part by a grant from NSERC.

in differential geometry [14]. The Dutch chess master Max Euwe rediscovered Thue's construction in connection with a problem about infinite chess games [10].

Fife [12] described all infinite overlap-free binary sequences; also see [7]. Séébold proved the beautiful and remarkable result that $\mathbf{t}$ is essentially the only infinite overlap-free binary sequence which is generated by iterating a morphism [17].

It is natural to wonder if Thue's overlap-free construction is either unique in some sense, or a particular case of a more general construction. In this note we show that $\mathbf{t}$ is a particular case of a more general construction involving sums of digits. We will prove

Theorem 1 *Let $k \geq 2$, $m \geq 1$ be integers. Then the sequence $\mathbf{t}_{k,m} := (s_k(n) \bmod m)_{n \geq 0}$ over the alphabet $\Sigma_m = \{0, 1, \ldots, m-1\}$ is overlap-free if and only if $m \geq k$.*

In contrast to Theorem 1 we also show that $\mathbf{t}_{k,m}$ always contains arbitarily long squares.

We also consider the occurrence of palindromes in $\mathbf{t}_{k,m}$. A palindrome is a word (such as `kayak` or `radar`) that is equal to its reversal. We prove that $\mathbf{t}_{k,m}$ contains arbitrarily long palindromes if and only if $m \leq 2$.

Some combinatorial properties of $\mathbf{t}_{k,m}$ were previously studied by Morton and Mourant [15], who proved among other things that $\mathbf{t}_{k,m}$ is ultimately periodic if and only if $m \mid (k-1)$.

We observe that overlaps, squares, and palindromes in sequences have several applications. For example, in number theory they aid in proving the transcendence of real numbers whose base b expansion or continued fraction expansion have "repetitions" [11, 4, 16, 2], while in statistical physics they are useful for studying the spectrum of certain discrete Schrödinger operators [9, 13, 1, 5].

2 Some useful lemmas

In this section we introduce some notation and prove some useful lemmas.

Lemma 2 *For any $k \geq 2, m \geq 1$, the sequence $\mathbf{t}_{k,m}$ is the fixed point, starting with 0, of the morphism $\varphi_{k,m}$ defined by $\varphi_{k,m}(a) = a\ (a+1)\ (a+2)\ \cdots\ (a+k-1)$ where the sums are taken modulo m.*

Proof. Left to the reader. ∎

Remark. Lemma 2 shows that $\mathbf{t}_{k,m}$ is a k-automatic sequence. For $k = m = 2$, we get the well-known fact that the Thue-Morse infinite word is the fixed point, starting with 0, of the morphism defined by $0 \to 01, 1 \to 10$.

Let $k \geq 2, n \geq 1$ be integers. Then we define $\nu_k(n)$ to be the exponent of the highest power of k which divides n. More precisely, $\nu_k(n) = a$ if $k^a \mid n$ but $k^{a+1} \not\mid n$.

Lemma 3 *For all integers $k \geq 2$ and $n, n' \geq 1$ we have*

$$\nu_k(n+n') \begin{cases} = \min(\nu_k(n), \nu_k(n')), & \text{if } \nu_k(n) \neq \nu_k(n'); \\ \geq \nu_k(n), & \text{if } \nu_k(n) = \nu_k(n'). \end{cases} \tag{1}$$

Proof. Left to the reader. ■

Remark. Note that if k is a prime number, then we have $\nu_k(nn') = \nu_k(n) + \nu_k(n')$, but this is not necessarily true if k is not prime.

Lemma 4 *Any block of $2k$ consecutive values of the sequence $(\nu_k(n))_{n\geq 1}$ contains an occurrence of the value 1.*

Proof. Any k consecutive numbers contains some multiple of k, say g. We have $\nu_k(g) \geq 1$. If $\nu_k(g) = 1$, we are done. Otherwise $\nu_k(g + k) = 1$ by Lemma 3. ■

The link between $\nu_k(n)$ and $s_k(n)$ is given by the following lemma.

Lemma 5 *Let $k \geq 2$, $n \geq 1$ be integers. Then*

$$s_k(n) - s_k(n-1) = 1 - (k-1)\nu_k(n).$$

Proof. Let $\nu_k(n) = a$. Then the base-k expansion of n can be written in the form $w\ c\ \overbrace{0\cdots 0}^{a}$ for some word w and some single digit c, where $c \neq 0$. Then the base-k expansion of $n-1$ is $w\ (c-1)\ \overbrace{(k-1)\ \cdots\ (k-1)}^{a}$. It follows that $s_k(n) - s_k(n-1) = 1 - (k-1)a$. ■

We now turn to the properties of overlaps. If $axaxa$ is an overlap, then we call $p = |ax|$ its *period*.

If $w = a_1a_2\cdots a_t$ is a word over $\Sigma_m = \{0, 1, \ldots, m-1\}$, then we define $S(w) = (\sum_{1\leq i\leq t} a_i) \bmod m$. Given a word $w = a_1a_2\cdots a_t$, we define its word of first differences

$$\Delta w = (a_2 - a_1)\ \cdots\ (a_t - a_{t-1}),$$

where the differences are taken mod m. We can also extend Δ to infinite words.

Lemma 6 *Let* $\mathbf{s}$ *be a finite or infinite word over $\{0, 1, \ldots, m-1\}$. The word* $\mathbf{s}$ *contains an overlap $axaxa$ if and only if the word $\Delta\mathbf{s}$ contains a square yy such that $S(y) \equiv 0 \pmod m$.*

Proof. Suppose $\mathbf{s} = s_0s_1s_2\cdots$.

If $\mathbf{s}$ contains an overlap of period p beginning at position r of $\mathbf{s}$. Let $x = x_0x_1\cdots x_{2p} = s_r\cdots s_{r+2p}$ be this overlap. Then $x_i = x_{i+p}$ for $0 \leq i \leq p$. Then $x_i - x_{i-1} = x_{i+p} - x_{i+p-1}$ for $1 \leq i \leq p$, so letting $y = (x_1 - x_0)\ (x_2 - x_1)\ \cdots\ (x_p - x_{p-1})$, we get $\Delta x = yy$. Furthermore, $S(y) = x_p - x_0 = 0$.

For the converse, suppose $\Delta\mathbf{s}$ contains a square yy. Then there exist integers $r \geq 0$, $p \geq 1$, such that $s_{r+i+1} - s_{r+i} = s_{r+i+1+p} - s_{r+i+p} \pmod m$ for $0 \leq i < p$. By telescoping cancellation, it follows that

$$s_{r+j+1} - s_r = s_{r+j+1+p} - s_{r+p} \pmod m \tag{2}$$

for $0 \leq j < p$. If, as the hypothesis states, we have $S(y) \equiv 0 \pmod m$, then

$$s_{r+p} - s_r \equiv 0 \pmod m. \tag{3}$$

Together Eqs. (2) and (3) imply that $s_{r+j} \equiv s_{r+j+p} \pmod m$ for $0 \leq j \leq p$. But this implies that $\mathbf{s}$ contains at overlap of period $p+1$, beginning at index r. ■

3 Proof of the main theorem

We are now ready to prove Theorem 1.

Proof. Fix integers $k \geq 2$ and $m \geq 1$, and let $\mathbf{t}_{k,m} = t(0)\,t(1)\,t(2)\cdots$. Note that

$$t(kn+c) \equiv s_k(kn+c) \equiv s_k(n) + c \pmod{m} \tag{4}$$

for $0 \leq c < k$. If $m \geq k$, it then follows that

$$\text{Every symbol contained in } z = t(kn)\,t(kn+1)\,\cdots\,t(kn+k-1) \\ \text{appears exactly once in } z. \tag{5}$$

We call such a subword z (of length k, starting at a position in $\mathbf{t}_{k,m}$ which is congruent to 0, modulo k) a *k-aligned subword.* It follows from Eq. (4) that every symbol in a k-aligned subword is completely determined once the value of a single such symbol is known.

$\Longrightarrow$: Assume $m < k$. We will prove that the sequence $\mathbf{t}_{k,m}$ contains an overlap of period m. In fact, the subword

$$t(k^m - (m+1))\,t(k^m - m)\,\cdots\,t(k^m + m - 1)$$

is the overlap $0\;1\,\cdots\,(m-1)\;0\;1\,\cdots\,(m-1)\;0$.

Since $k \geq m+1$, the base-k expansion of $k^m - i$, with $m+1 \geq i \geq 1$, is of the form

$$\overbrace{(k-1)\cdots(k-1)}^{m-1}\,(k-i).$$

Then

$$s_k(k^m - i) = (m-1)(k-1) + k - i \equiv 1 - i \pmod{m}$$

for $m+1 \geq i \geq 1$. Thus $t(k^m - (m+1))\,\cdots\,t(k^m - 1) = 0\;1\,\cdots\,(m-1)\;0$.

Similarly, the base-k expansion of $k^m + i$, $0 \leq i \leq m-1$, is of the form

$$1\;\overbrace{0\cdots0}^{m}\;i.$$

Thus $s_k(k^m + i) = i + 1$ for $0 \leq i \leq m-1$. Thus $t(k^m)\,\cdots\,t(k^m + m - 1) = 1\;2\,\cdots\,(m-1)\;0$. The result follows.

$\Longleftarrow$: If $\mathbf{t}_{k,m}$ has an overlap, then it has an overlap of shortest period p. Let

$$t(r)\,t(r+1)\,\cdots\,t(r+p-1)\,t(r+p)\,\cdots\,t(r+2p-1)\,t(r+2p)$$

be such an overlap. Note $p \geq 1$. By the definition of overlap, we have $t(r+i) = t(r+p+i)$ for $0 \leq i \leq p$.

Case 1: $k \mid p$. In this case, write $p = kp'$, where p' is a positive integer, and, using the division theorem, write $r = kr' + c$ where $0 \leq c < k$.

Since $t(r+i) = t(r+p+i)$ for $0 \le i \le p$, we have, by considering only those i that are multiples of k, that $t(r+kj') = t(r+p+kj')$ for $0 \le j' \le p/k$. Hence $t(kr'+c+kj') = t(kr'+c+kp'+kj')$ for $0 \le j' \le p'$. Hence $t(k(r'+j')+c) = t(k(r'+p'+j')+c)$ for $0 \le j' \le p'$. Hence $t(r'+j')+c \equiv t(r'+p'+j')+c \pmod{m}$, and so $t(r'+j') \equiv t(r'+p'+j') \pmod{m}$ for $0 \le j' \le p'$. But then $t(r') \cdots t(r'+2p')$ is an overlap of period $p' = p/k < p$, contradicting our assumption that p was minimal.

Case 2: $k \not| \, p$. In this case there are three subcases to consider, based on the size of p: (a) $p < k$; (b) $k < p < 2k$; (c) $p > 2k$.

Case 2(a): $p < k$. Let $j = \lceil \frac{r}{k} \rceil$. Then $kj = r + I$ where $0 \le I < k$. There are two cases to consider, (i) $I \le p$ and (ii) $I > p$.

Case 2(a)(i): If $I \le p$, then $w = t(r+I) \cdots t(r+I+p)$ is a subword of $t(kj) \cdots t(kj+k-1)$, and w contains two identical symbols, namely $t(r+I)$ and $t(r+I+p)$, which contradicts observation (5).

Case 2(a)(ii): If $I > p$, then $w = t(r) \cdots t(r+p)$ is a subword of $t(k(j-1)) \cdots t(kj-1)$, and w contains two identical symbols, namely $t(r)$ and $t(r+p)$, again contradicting observation (5).

In both cases we get a contradiction, so there cannot be an overlap with $p < k$.

Case 2(b): $k < p < 2k$. As in Case 2(a), let $j = \lceil \frac{r}{k} \rceil$. Suppose the overlap is

$$t(r) \cdots t(r+2p).$$

Define $x_i = t(r+i)$ for $0 \le i \le 2p$, and note

$$x_i = x_{p+i} \text{ for } 0 \le i \le p. \tag{6}$$

Set $I = kj - r$, so that $A_1 = x_I \cdots x_{I+k-1}$ is a k-aligned subword of $\mathbf{t}_{k,m}$ and $0 \le I < k$. There are two cases to consider: (i) $0 \le I \le p-k$; and (ii) $I > p-k$.

Case 2(b)(i): $0 \le I \le p-k$. If further $I \le 2p-3k+1$, then define

$$X = \overbrace{x_0 \cdots x_{I-1}}^{A_0} \; \overbrace{x_I \cdots x_{I+k-1}}^{A_1} \; \overbrace{x_{I+k} \cdots x_{I+p} \cdots x_{I+2k-1}}^{A_2} \; \overbrace{x_{I+2k} \cdots x_{I+k+p} \cdots x_{I+3k-1}}^{A_3} \; \overbrace{x_{I+3k} \cdots x_{2p}}^{A_4}. \tag{7}$$

Note that A_1, A_2, and A_3 are all k-aligned subwords.

Otherwise, if $2p-3k+1 < I \le p-k$, define

$$X = \overbrace{x_0 \cdots x_{I-1}}^{A_0} \; \overbrace{x_I \cdots x_{I+k-1}}^{A_1} \; \overbrace{x_{I+k} \cdots x_{I+p} \cdots x_{I+2k-1}}^{A_2} \; \overbrace{x_{I+2k} \cdots x_{I+k+p} \cdots x_{2p}}^{A_3}. \tag{8}$$

Note that A_1 and A_2 are both k-aligned subwords, and A_3 is a prefix of a k-aligned subword.

Suppose $x_I \equiv J \pmod{m}$. Then, from the fact that A_1 is a k-aligned subword, $x_{I+k-1} \equiv J + k - 1 \pmod{m}$. Then from (6) we get $x_{I+k+p-1} \equiv J + k - 1 \pmod{m}$. Since A_3 is a k-aligned subword or prefix of one, we get

$$x_{I+k+p} \equiv J + k \pmod{m}. \tag{9}$$

From $x_I \equiv J \pmod{m}$ and (6) we get $x_{I+p} \equiv J \pmod{m}$. Since A_2 is a k-aligned subword, we get $x_{I+k} \equiv J + k - p \pmod{m}$. From (6) we get

$$x_{I+k+p} \equiv J + k - p \pmod{m}. \tag{10}$$

Now combining Eqs. (9) and (10) gives $J + k \equiv J + k - p \pmod{m}$ or $p \equiv 0 \pmod{m}$, so $m \mid p$. If $p \geq 2m$, then since $m \geq k$ we get $p \geq 2k$, a contradiction. Hence $p = m$.

In this case by examining A_1 we get

$$x_i \equiv i + J - I \pmod{m} \text{ for } I \leq i < I + k. \tag{11}$$

Now $x_I = x_{I+p}$, so by examining A_2 we get $x_i \equiv i + J - I - p \pmod{m}$ for $I + k \leq i < I + 2k$. But $p = m$, so

$$x_i \equiv i + J - I \pmod{m} \text{ for } I + k \leq i < I + 2k. \tag{12}$$

Now $2k > p$, so (11) and (12) together cover all residue classes mod p, and so we get

$$x_i \equiv i + J - I \pmod{m} \text{ for } 0 \leq i \leq 2p. \tag{13}$$

Now consider $Y = \Delta X$. By Lemma 6, Y must be a square. From (13) we get $Y = \overbrace{1 \cdots 1}^{2p}$. But by Lemma 5

$$Y = (1 - (k-1)\nu_k(r+1)) \, \cdots \, (1 - (k-1)\nu_k(r+2p)),$$

where both sides are considered modulo m. Now $p > k$, so by Lemma 4, there exists an index g, $r + 1 \leq g \leq r + 2p$ such that $\nu_k(g) = 1$. Then $(1 - (k-1)\nu_k(g)) \bmod m = 1$, so $2 - k \equiv 1 \pmod{m}$. Hence $1 - k \equiv 0 \pmod{m}$ and hence $m \mid k - 1$. Then since $k \geq 2$, we have $k - 1 \geq m$ and so $k > m$. But $m = p$ and hence $k > p$. This contradicts the assumption of case 2(b) that $k < p$, and hence this case cannot occur.

Case 2(b)(ii): $I > p - k$. If further $I \leq 2p - 2k + 1$, define

$$X = \overbrace{x_0 \cdots x_{I+k-p} \cdots x_{I-1}}^{A_0} \; \overbrace{x_I \cdots x_p \cdots x_{I+k-1}}^{A_1} \; \overbrace{x_{I+k} \cdots x_{I+p} \cdots x_{I+2k-1}}^{A_2} \; \overbrace{x_{I+2k} \cdots x_{2p}}^{A_3} \tag{14}$$

Note that A_1 is a k-aligned subword, and from the inequality $I \leq 2p - 2k + 1$, we get $I + 2k - 1 \leq 2p$, so A_2 is also a k-aligned subword. Note that A_3 may be empty.

Otherwise, if $I > 2p - 2k + 1$, define

$$X = \overbrace{x_0 \cdots x_{I+k-p} \cdots x_{I-1}}^{A_0} \; \overbrace{x_I \cdots x_p \cdots x_{I+k-1}}^{A_1} \; \overbrace{x_{I+k} \cdots x_{I+p} \cdots x_{2p}}^{A_2} \tag{15}$$

In this case A_1 is a k-aligned subword, and A_2 is a prefix of a k-aligned subword.

Suppose $x_I \equiv J \pmod{m}$. Then since A_1 is a k-aligned subword, we have

$$x_p \equiv J + p - I \pmod{m}. \tag{16}$$

Now $x_I = x_{I+p}$, and A_2 is a k-aligned subword or prefix of one, so $x_{I+k} \equiv J + k - p \pmod{m}$. Now $I + k - p \geq 0$, so x_{I+k-p} lies in A_0, and

$$x_{I+k} = x_{I+k-p} \equiv J + k - p \pmod{m}. \tag{17}$$

Then A_0 is the suffix of a k-aligned subword, so from Eq. (17) we get $x_0 \equiv J - I \pmod{m}$. Then $x_p = x_0$, so

$$x_p \equiv J - I \pmod{m}. \tag{18}$$

Combining the congruences (16) and (18), we get $J + p - I \equiv J - I \pmod{m}$. Hence $p \equiv 0 \pmod{m}$, and so $m \mid p$. As before, if $p \geq 2m$, then since $m \geq k$ we get $p \geq 2k$, a contradiction. Hence $p = m$.

Now, by examining A_1 we get

$$x_i \equiv i + J - I \pmod{m} \text{ for } I \leq i < I + k. \tag{19}$$

Similarly, by examining A_0 we get

$$x_i \equiv i + J - I \pmod{m} \text{ for } 0 \leq i < I. \tag{20}$$

Combining (19) and (20), we get

$$x_i \equiv i + J - I \pmod{m} \text{ for } 0 \leq i < I + k. \tag{21}$$

By assumption for this case, we have $p < I + k$, so all residue classes mod p, are covered, and we have

$$x_i \equiv i + J - I \pmod{m} \text{ for } 0 \leq i \leq 2p. \tag{22}$$

The rest of the proof proceeds as in Case 2(b)(i). The argument there shows this case cannot occur.

Case 2(c): $p > 2k$. Consider the word $\Delta\mathbf{t}$. Then from Lemma 4 there must be an i, $r \leq i < r + p$ such that $\nu_k(i) = 1$. Then by Lemma 6 we know $\Delta\mathbf{t}$ contains a square. Then by Lemma 5 we have

$$1 - (k-1)\nu_k(i) \equiv 1 - (k-1)\nu_k(i+p) \pmod{m}.$$

Hence

$$k - 1 \equiv (k-1)\nu_k(i+p) \pmod{m}.$$

It follows that $\nu_k(i+p) \geq 1$, for if $\nu_k(i+p) = 0$ we would have $k - 1 \equiv 0 \pmod{m}$, and so $k - 1 \geq m$ and $k > m$, a contradiction.

Now $\nu_k(i) = 1$ and $\nu_k(i+p) \geq 1$. It follows from Lemma 3 that $k \mid p$. But in Case 2 we assumed $k \not\mid p$, a contradiction.

The proof of Theorem 1 is now complete. ■

4 Squares in the sequence $\mathbf{t}_{k,m}$

It is easy to show the following theorem about the existence of arbitrarily long squares in the sequence $\mathbf{t}_{k,m}$.

Theorem 7 *The sequence $\mathbf{t}_{k,m}$ contains arbitrarily long squares. More precisely we have*

(a) The sequence $\mathbf{t}_{k,m}$ contains the square of a single letter if and only if $\gcd(k-1, m) = 1$.

(b) For all integers $k \geq 2$, $m \geq 1$, the sequence $\mathbf{t}_{k,m}$ contains arbitrarily long squares.

Proof.
(a) By Lemma 6, there exists a square aa with $a \in \Sigma_m$ in the sequence $\mathbf{t}_{k,m}$ if and only if there exists an integer $n \geq 1$ such that $(k-1)\nu_k(n) \equiv 1 \pmod m$. Since $\nu_k(n)$ can take any integer value, this is equivalent to $\gcd(k-1, m) = 1$.

(b) If $m < k$, then in Theorem 1 we proved the existence of overlaps, hence squares. Now the image of a square by $\varphi_{k,m}$ is a longer square. Iterating $\varphi_{k,m}$ and using the fact that $\mathbf{t}_{k,m}$ is a fixed point of $\varphi_{k,m}$ gives arbitrarily long squares.

Suppose now that $m \geq k$. Then the first $2k-1$ terms of the sequence $\mathbf{t}_{k,m}$ are

$$0\,1\,2\,3\ \cdots\ (k-1)\,1\,2\,3\cdots\ (k-1)$$

which contains a square of length $2k-2$. The images of this square under iterates of $\varphi_{k,m}$ are arbitrarily large squares. ■

Remark. It would be interesting to determine the largest (fractional) power that occurs in the sequence $\mathbf{t}_{k,m}$. For $m \geq k$, we already know that 2 is sharp.

5 Palindromes in $\mathbf{t}_{k,m}$

In this section we examine the occurrence of palindromes in $\mathbf{t}_{k,m}$.

Theorem 8 *The sequence $\mathbf{t}_{k,m}$ with $k \geq 2$, $m \geq 1$ contains arbitrarily long palindromes if and only if $m \leq 2$.*

Proof. $\Longrightarrow$: Suppose that the sequence $\mathbf{t}_{k,m}$ contains some palindrome of even length larger than or equal to 4. Then it must contain the word $baab$ for some $a, b \in \Sigma_m$. If aa is contained in the image by $\varphi_{k,m}$ of some letter in Σ_m, then $m = 1$. Otherwise the first a must be the last letter of the image by $\varphi_{k,m}$ of some letter, and the second a must be the first letter of the image by $\varphi_{k,m}$ of some letter. It follows that $b \equiv a - 1 \pmod m$ and $b \equiv a + 1 \pmod m$. Hence $2 \equiv 0 \pmod m$ and this gives $m \leq 2$.

Now suppose that the sequence $\mathbf{t}_{k,m}$ contains a palindrome of odd length larger than or equal to 5, say $cbabc$, with $a, b, c \in \Sigma_m$. If bab is a subword of the image by $\varphi_{k,m}$ of some letter in Σ_m, then $a \equiv b + 1 \pmod m$ and $b \equiv a + 1 \pmod m$, hence $2 \equiv 0 \pmod m$. Hence again $m \leq 2$. If bab is not a subword of the image by $\varphi_{k,m}$ of some letter in Σ_m, then we have two possibilities according to whether $k \geq 3$ or $k = 2$.

If $k \geq 3$, then either cba is a suffix of the image by $\varphi_{k,m}$ of some letter, and bc a prefix of the image by $\varphi_{k,m}$ of some letter, or cb is a suffix of the image by $\varphi_{k,m}$ of some letter, and abc a prefix of the image by $\varphi_{k,m}$ of some letter. In the first case we have $c \equiv a - 2 \pmod{m}$, $b \equiv a - 1 \pmod{m}$, and $c \equiv b + 1 \pmod{m}$, hence $2 \equiv 0 \pmod{m}$. In the second case $c \equiv b - 1 \pmod{m}$, $b \equiv a + 1 \pmod{m}$, and $c \equiv a + 2 \pmod{m}$, hence $2 \equiv 0 \pmod{m}$. This gives $m \leq 2$ in both cases.

If $k = 2$, then either the first c is the last letter of the image by $\varphi_{k,m}$ of some letter, ba is the image by $\varphi_{k,m}$ of some letter, and bc is the image by $\varphi_{k,m}$ of some letter, or cb is the image by $\varphi_{k,m}$ of some letter, ab is the image by $\varphi_{k,m}$ of some letter, and the second c is the first letter of the image by $\varphi_{k,m}$ of some letter. In the first case, we must have $a \equiv b + 1 \pmod{m}$ and $c \equiv b + 1 \pmod{m}$, hence $cbabc = (b + 1)\ b\ (b + 1)\ b\ (b + 1)$ which is an overlap, hence $m < k = 2$ from Theorem 1 and this is impossible. In the second case, we must have $b \equiv c + 1 \pmod{m}$ and $b \equiv a + 1 \pmod{m}$, hence $a \equiv c \pmod{m}$. This gives $cbabc = c\ (c + 1)\ c\ (c + 1)\ c$ which is again an overlap, and we conclude as just above.

$\Longleftarrow$: Now let us suppose that $m \leq 2$. If $m = 1$, then $\mathbf{t}_{k,m} = 000\cdots$ and hence trivially contains arbitrarily large palindromes.

Now assume $m = 2$. If k is odd, then $\mathbf{t}_{k,m} = 01010101\cdots$ and hence trivially contains arbitrarily long palindromes.

If k is even, the sequence $\mathbf{t}_{k,m}$ is a fixed point of the morphism defined on $\{0, 1\}$ by

$$\begin{aligned} \varphi_{k,m}(0) &= (01)^{k/2} \\ \varphi_{k,m}(1) &= (10)^{k/2} \end{aligned}$$

and an easy induction shows that $\varphi_{k,m}^{2j}(0)$ is a palindrome of length k^{2j}. ■

Acknowledgements

We thank the referees for a careful reading of this paper.

References

[1] J.-P. Allouche. Schrödinger operators with Rudin-Shapiro potentials are not palindromic. *J. Math. Phys.* **38** (1997), 1843–1848.

[2] J.-P. Allouche, L. Davison, and M. Queffélec. Transcendence of Sturmian and morphic continued fractions. Preprint, 1999.

[3] J.-P. Allouche and J. Shallit. The ubiquitous Prouhet-Thue-Morse sequence. In C. Ding, T. Helleseth, and H. Niederreiter, editors, *Sequences and Their Applications, Proceedings of SETA '98*, pp. 1–16. Springer-Verlag, 1999.

[4] J.-P. Allouche and L. Q. Zamboni. Algebraic irrational binary numbers cannot be fixed points of non-trivial constant-length or primitive morphisms. *J. Number Theory* **69** (1998), 119–124.

[5] M. Baake. A note on palindromicity. Preprint, 1999.

[6] J. Berstel. Axel Thue's work on repetitions in words. In P. Leroux and C. Reutenauer, editors, *Séries Formelles et Combinatoire Algébrique*, number 11 in Publications du LACIM, pp. 65–80. Université du Québec à Montréal, June 1992.

[7] J. Berstel. A rewriting of Fife's theorem about overlap-free words. In J. Karhumäki, H. Maurer, and G. Rozenberg, editors, *Results and Trends in Theoretical Computer Science*, Vol. 812 of *Lecture Notes in Computer Science*, pp. 19–29. Springer-Verlag, 1994.

[8] J. Berstel. *Axel Thue's Papers on Repetitions in Words: a Translation*. Number 20 in Publications du LACIM. Université du Québec à Montréal, February 1995.

[9] A. Bovier and J.-M. Ghez. Spectral properties of one-dimensional Schrödinger operators with potentials generated by substitutions. *Comm. Math. Phys.* **158** (1993), 45–66. Erratum, **166** (1994), 431–432.

[10] M. Euwe. Mengentheoretische Betrachtungen über das Schachspiel. *Proc. Konin. Akad. Wetenschappen, Amsterdam* **32** (1929), 633–642.

[11] S. Ferenczi and C. Mauduit. Transcendence of numbers with a low complexity expansion. *J. Number Theory* **67** (1997), 146–161.

[12] E. D. Fife. Binary sequences which contain no BBb. *Trans. Amer. Math. Soc.* **261** (1980), 115–136.

[13] A. Hof, O. Knill, and B. Simon. Singular continuous spectrum for palindromic Schrödinger operators. *Comm. Math. Phys.* **174** (1995), 149–159.

[14] M. Morse. Recurrent geodesics on a surface of negative curvature. *Trans. Amer. Math. Soc.* **22** (1921), 84–100.

[15] P. Morton and W. J. Mourant. Digit patterns and transcendental numbers. *J. Austral. Math. Soc. Ser. A* **51** (1991), 216–236.

[16] M. Queffélec. Transcendance des fractions continue de Thue-Morse. *J. Number Theory* **73** (1998), 201–211.

[17] P. Séébold. Sequences generated by infinitely iterated morphisms. *Disc. Appl. Math.* **11** (1985), 255–264.

[18] A. Thue. Über unendliche Zeichenreihen. *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **7** (1906), 1–22. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo, 1977, pp. 139–158.

[19] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **1** (1912), 1–67. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo, 1977, pp. 413–478.