No iterated morphism generates any Arshon sequence of odd order

James D. Currie*
Department of Mathematics and Statistics
University of Winnipeg
Winnipeg, Manitoba R3B 2E9, Canada
e-mail: j.currie@uwinnipeg.ca

March 6, 2002

Abstract

We show that no Arshon sequence of odd order can be generated by an iterated morphism. This solves a problem of Kitaev and generalizes results of Berstel and of Kitaev.
Keywords: Combinatorics on words, non-repetitive words, Arshon's sequence, DOL systems

1 Introduction

Let $n \in \mathbb{N}$, $n > 2$. Consider the alphabet $\Sigma_n = \{1, 2, \ldots, n\}$. Define two morphisms $\psi_o, \psi_e : \Sigma_n \to \Sigma_n^*$ by

$$\psi_o(i) = i(i+1)(i+2)\cdots(n-1)n12\cdots(i-2)(i-1)$$

$$\psi_e(i) = (i-1)(i-2)\cdots 21n(n-1)\cdots(i+2)(i+1)i.$$

The subscripts here are o and e, standing for *odd* and *even*. Morphism $\psi_o(i)$ consists of the n letters of Σ_n arranged in cyclically increasing order modulo n, starting with i. Word $\psi_e(i)$ is just the reverse of $\psi_o(i)$. We define an operator ψ on Σ_n^*. Let $p = a_1a_2a_3a_4\cdots a_m$, $a_i \in \Sigma_n$. Define

$$\psi(p) = \psi_o(a_1)\psi_e(a_2)\psi_o(a_3)\psi_e(a_4)\cdots.$$

Example: In the case where $n = 3$, $\psi(1231) = \psi_o(1)\psi_e(2)\psi_o(3)\psi_e(1) = 123132312321$.

*The author was supported by an NSERC operating grant.

Consider the words $\psi^m(1)$. One sees by induction that $\psi^m(1)$ is a prefix of $\psi^{m+1}(1)$, $m \in \mathbb{N}$. We can therefore define the limit $w_n = \lim_{m\to\infty} \psi^m(1)$.

The infinite words w_n were introduced by Arshon [1]. Arshon proved that each w_n is *non-repetitive*; that is, word w_n contains no (connected) subword qq with q a non-empty word. Words avoiding repetitions or other patterns have been extensively studied. The usual way of generating such words is via D0L systems, i.e. by iterating some morphism. (See [2, 3, 5, 6].)

It is natural to ask whether Arshon's word w_n can be generated by iterating a morphism. In the case that n is even, the answer is yes. In this case

$$w_n = \lim_{m\to\infty} f^m(1), \text{where} f(i) = \begin{cases} \psi_o(i), i \text{ odd} \\ \psi_e(i), i \text{ even} \end{cases}$$

Berstel showed in [4] that there is no morphism $f : T^* \to T^*$ such that

$$w_3 = \lim_{m\to\infty} f^m(1),$$

even if T is allowed to be a proper superset of Σ_3. Berstel's proof uses a deep result of Cobham [7]. In his thesis [8] Kitaev gave a simpler, self-contained proof that there is no morphism $f : \Sigma_3^* \to \Sigma_3^*$ such that

$$w_3 = \lim_{m\to\infty} f^m(1).$$

A question mentioned by Kitaev is whether there is a morphism $f : \Sigma_n^* \to \Sigma_n^*$ such that

$$w_n = \lim_{m\to\infty} f^m(1)$$

in the case where n is an odd number greater than 3. We complete the results of Berstel and Kitaev by proving the following theorem:

Theorem 1.1 *Let $n \geq 5$ be odd. There is no morphism $f : \Sigma_n^* \to \Sigma_n^*$ such that*

$$w_n = \lim_{m\to\infty} f^m(1).$$

2 Notations and Preliminaries

Fix an odd $n \geq 5$. When we consider n as a letter of Σ_n, it will be convenient to interpret $n+1$ as 1, i.e. to work modulo n.

Suppose that v is a subword of w_n, $v = v_1v_2v_3v_4\cdots v_m$. It is relevant to the action of ψ on v to note how v lies in w_n modulo 2. We say that $v_1^o v_2^e v_3^o v_4^e \cdots$ is a subword of w_n if w_n has some prefix pv where $|p|$ is even. If w_n has some prefix pv where $|p|$ is odd, we say that $v_1^e v_2^o v_3^e v_4^o \cdots$ is a subword of w_n. It may be that both $v_1^o v_2^e v_3^o v_4^e v_5^o \cdots$ and $v_1^e v_2^o v_3^e v_4^o v_5^e \cdots$ are subwords of w_n.

Remark 2.1 Since $\psi(12)$ is a prefix of w_n, w_n has the subword $1^o 2^e 3^o \cdots (n-1)^e n^o 1^e n^o (n-1)^e \cdots 4^e 3^o 2^e$ as a prefix. We therefore see that for each odd $i \in \Sigma_n$ the word $i^o(i+1)^e$ is a subword of w_n. Also, for each even i in Σ_n, $(i+1)^o i^e$ is a subword of w_n, as is $i^e(i+1)^o$.

The last n letters of $\psi(\psi(1)1) = \psi(123\cdots(n-1)n1)$ are $n^e(n-1)^o(n-2)^e \cdots 3^e 2^o 1^e$. We therefore see that w_n also contains $(i+1)^o$ for each odd $i \in \Sigma_n$, and $(i+1)^e$ for each even $i \in \Sigma_n$. In fact, we see that $(i+1)^e i^o$ is also a subword of w_n when i is even, hence whether i is even or odd.

Finally, for odd i, $(i-1)^o$ occurs in w_n. Then w_n has some prefix $p(i-1)$ with $|p|$ even. It follows that $\psi(p(i-1)= \psi(p)\psi_o(i-1)$ is a prefix of w_n. We see that $|\psi(p)| = n|p|$ is even, whence w_n contains subword $(i-1)^o i^e (i+1)^o$. Thus regardless of whether i is even or odd, w_n contains $i^e(i+1)^o$ as a subword.

It will also be important to note how v lies in w_n modulo n. We imagine w_n to be punctuated by 'bar lines', like music, one bar line every n letters:

$$w_n = 123\cdots n|1n\cdots 432|345\cdots 12|32\cdots$$

We write that $v_1|v_2|v_3|v_4|\cdots|v_{m-1}|v_m$ is a subword of w_m to say that w_n has a prefix $pv_1v_2v_3v_4\cdots v_{m-1}v_m$ where $|pv_1| \equiv 0 \pmod n$, $|v_i| = n$, $i = 2, 3, \ldots, m-1$.

If $v = |v_1|v_2|v_3|v_4|\cdots|v_{m-1}|v_m|$ is a subword of w_n, we say that v **is concatenated from ψ-blocks**. Whether the letters in ψ-block v_i increase or decrease modulo n depends on the preimage of v in w_n under ψ. If v occurs as the image of $q = q_1^o q_2^e q_3^o q_4^e q_5^o \cdots$ then we write $v = |\overrightarrow{v_1}|\overleftarrow{v_2}|\overrightarrow{v_3}|\overleftarrow{v_4}|\cdots$. If q lies in w_n with the opposite parity then the arrows on the v_i are reversed.
Example: When $n = 5$,

$$w_n = 1234515\underline{432345123215451}234\cdots$$

We see that $\overleftarrow{432} \mid \overrightarrow{34512} \mid \overleftarrow{32154} \mid \overrightarrow{5}$ is a subword of w_n.

3 Mordents and Ambiguity

A **mordent** is a word of the form iji where $i, j \in \Sigma_n$. Since every ψ-block is increasing or decreasing, mordents in w_n cross bar lines:

Lemma 3.1 *If $piji$ is a prefix of w_n with $i, j \in \Sigma_n$, then $|p| \equiv -1$ (mod n) or $|p| \equiv -2$ (mod n).*

Lemma 3.2 *Let $j \in \Sigma_n$. The mordent $i^o j^e i^o$ is a subword of w_n where $i \equiv j+1$ (mod n).*

Proof: By Remark 2.1, $j^e(j+1)^o$ is a subword of w_n. Thus $pj(j+1)$ is a prefix of w_n, some p with $|p|$ odd. This means that $|\psi(p)| = n|p|$ is odd.

It follows that $\psi(p)\psi_e(j)\psi_o(j+1)$ is a prefix of w_n, whence

$$\psi((j)^e(j{+}1)^o) = |(j{-}1)(j{-}2)(j{-}3)\cdots(j{+}2)(j{+}1)j|(j{+}1)(j{+}2)(j{+}3)\cdots(j{-}2)(j{-}1)j|$$

is a subword of w_n. Taking into account the parity of the length of $\psi(p)$, w_n contains

$$|(j{-}1)^e(j{-}2)^o(j{-}3)^e\cdots(j{+}2)^e(j{+}1)^o j^e|(j{+}1)^o(j{+}2)^e(j{+}3)^o\cdots(j{-}2)^o(j{-}1)^e j^o|$$

which contains the mordent $i^o j^e i^o$ where $i \equiv j+1 \pmod{n}$.□

Two consecutive mordents in w_n may occur in one of four different ways:

$$\begin{array}{ll} i|jitkl|k & \text{(near mordents)} \\ ij|itk|lk & \text{(far mordents)} \\ \left.\begin{array}{r} i|jitk|lk \\ ij|itkl|k \end{array}\right\} & \text{(neutral mordents)} \end{array} \qquad (1)$$

In other words, a subword v of w_n **contains a pair of near (far, neutral) mordents** if v has a subword $ijitklk$ where $i,j,k,l \in \Sigma_n$ and $|t| = n-4$ $(n-2, n)$. We speak of v containing **consecutive neutral mordents** if v has a subword $i_1j_1i_1t_1i_2j_2i_2t_2i_3j_3i_3$ where $|t_1| = |t_2| = n$.

We say that a subword v of w_n is **unambiguous with respect to barlines** if whenever p_1v and p_2v are prefixes of w_n we must have $|p_1| \equiv |p_2| \pmod{n}$. This means that all occurrences of v in w_n are cut by barlines in the same way.

Since mordents must cross barlines, if subword v of w_n contains a pair of near mordents or a pair of far mordents then v is unambiguous with respect to bar lines. In fact, one sees quickly that near or far mordents cross barlines as we have sketched them in (1).

Lemma 3.3 *A pair of neutral mordents in w_n can occur only as the image under ψ of some mordent.*

Proof: For the sake of definiteness, consider a pair of neutral mordents in w_n in the situation $\overrightarrow{i} \mid \overleftarrow{ji}tk \mid \overrightarrow{lk}$. The cases with shifted parity $ij|itkl|k$, or with reversed arrows, are similar.

We have $k \equiv l+1 \pmod{n}$ by the arrow over lk. Similarly, considering arrows, one deduces that $k \equiv l+1 \equiv j+1 \equiv i+2 \pmod{n}$. It follows that $j = l$, and $\overrightarrow{i} \mid \overleftarrow{ji}tk \mid \overrightarrow{lk} = \overrightarrow{i} \mid \overleftarrow{ji}tk \mid \overrightarrow{jk}$ is contained in the image of under ψ of the mordent $j^o k^e j^o$.□

Corollary 3.4 *No subword of w_n contains consecutive neutral mordents.*

Proof: By Lemma 3.3, consecutive mordents in w_n would have to be contained in the image under ψ of a word $ijij$ in w_n. However, this is impossible, since w_n contains no repetitions.□

The longest subword of w_n not containing 3 consecutive mordents will have the form $i|v_1|v_2|v_3|j$ where $i,j \in \Sigma_n$ and the v_i are ψ-blocks.

Corollary 3.5 *Any subword of w_n of length $3n+3$ or greater contains a pair of near mordents or a pair of far mordents.*

Corollary 3.6 *Any subword of w_n of length $3n+3$ or greater is unambiguous with respect to bar lines.*

For each $i \in \Sigma_n$, define word u_i by $u_i = \psi_e(i+1)\psi_o(i)$. From Remark 2.1 we see that for each i, $|u_i|$ will be a subword of w_n, since $(i+1)^e i^o$ is. In fact, w will contain u_i in a prefix $p_1 u_i$, some p_1 concatenated from ψ-blocks.

From Lemma 3.2, w_n also contains u_i as a subword in the context

$$\underline{i|(i-1)(i-2)\cdots(i+1)i|(i+1)(i+2)\cdots(i-1)}i|$$

which arises in $\psi((i+1)^o i^e (i+1)^o)$. Here, w_n contains u_i as a subword in a prefix $p_2 i u_i$, such that $p_2 i$ is concatenated from ψ-blocks.

4 Proof of Theorem 1.1

Suppose that $f : \Sigma_n^* \to \Sigma_n^*$ is a morphism such that

$$w_n = \lim_{m\to\infty} f^m(1).$$

In particular notice that $f(w_n) = w_n$.

Lemma 4.1 *We must have $|f(\psi(1))| > n$.*

Proof: This is clear, since otherwise $f^m(1)$ is a prefix of $f^m(\psi(1))$, which is a prefix of $f(\psi(1))$ for all m, and

$$w_n \neq \lim_{m\to\infty} f^m(1).\square$$

For the sake of getting a contradiction, choose f to be a morphism satisfying

1. $w_n = f(w_n)$
2. $|f(\psi(1)| > n$
3. Subject to conditions 1 and 2, $|f(1)|$ is as small as possible.

Lemma 4.2 *Morphism f is non-erasing, that is $|f(i)| \neq 0$ for all $i \in \Sigma_n$.*

Proof: Clearly $|f(1)| \neq 0$, or iterating f on 1 could not give w_n. Suppose that $|f(j)| = 0$, some $j \in \Sigma_n$. Let j be the greatest element of Σ_n such that $|f(j)| = 0$. Then $|f(j+1)| \neq 0$. (This is true even if $j = n$, interpreting $n+1$ as 1, i.e. working modulo n.) However, by Lemma 3.2, w_n contains $(j+1)j(j+1)$, hence $f(j+1)f(j)f(j+1) = f(j+1)f(j+1)$, a repetition. This is impossible.$\square$

If $i \in \Sigma_n$, denote by $|v|_i$ the number of i's appearing in word v. Because w_n is concatenated from ψ-blocks, if p is a prefix of w_n then $|p|/n \leq |p|_i < |p|/n+1$. Thus if $i, j \in \Sigma_n$ and p is a prefix of w_n we have $|f(p)|_j + 1 \geq |f(p)|_i$.

Lemma 4.3 *We have $|f(\psi(1))| \equiv 0$ (mod n).*

Proof: If $|f(\psi(1))| \not\equiv 0 \pmod n$, then choose $i, j \in \Sigma_n$ such that $|f(\psi(1))|_i \geq |f(\psi(1)|_j + 1$. We notice that if v is any ψ-block of w_n, then $|f(v)|_i = |f(\psi(1))|_i$, $|f(v)|_j = |f(\psi(1))|_j$. Since $\psi^2(1) = \psi(123\cdots n)$ is concatenated from n ψ-blocks, we find that

$$\begin{array}{rcl} |f(\psi^2(1))|_j + 1 & \geq & |f(\psi^2(1))|_i \\ & = & |f(\psi(123\cdots n))|_i \\ & = & n|f(\psi(1)|_i \\ & \geq & n(|f(\psi(1)|_j + 1) \\ & = & n|f(\psi(1)|_j + n \\ & = & |f(\psi^2(1))|_j + n \end{array}$$

This is a contradiction.□

Corollary 4.4 *Every ψ-block v obeys $|f(v)| \equiv 0$ (mod n)*

Lemma 4.5 *There is some $i \in \Sigma_n$ such that $|f(i)| \not\equiv 0$ (mod n).*

Proof: Otherwise $|f(i)| \equiv 0 \pmod n$ for all $i \in \Sigma_n$, so that for each $i \in \Sigma_n$, $f(i)$ is concatenated from ψ-blocks.

Define a morphism $g : \Sigma_n^* \to \Sigma_n^*$ by

$$g(i) = \psi^{-1}(f(i)), i \in \Sigma_n.$$

This definition makes sense since each $f(i)$ is concatenated from ψ-blocks.

Since ψ^{-1} and f both fix w_n, we have $w = g(w)$. However, since $|g(1)| = |f(1)|/n$, the minimality of f forces us to deduce that $|g(\psi(1))| \leq n$. By the manner of its definition g is non-erasing, so that we conclude $g|(\psi(1))| = n$. Because $\psi(1)$ is a prefix of w_n and g fixes w_n, morphism g is forced to be the identity on $\psi(1)$, and hence on Σ_n.

Since $1 = g(1) = \psi^{-1}(f(1))$, we find that $f(1)$ must be $\psi_o(1)$ or $\psi_e(1)$. In addition, $f(1)$ is a prefix of w_n, which forces $f(1) = \psi_o(1)$. Similarly, we find that

$$f(i) = \begin{cases} \psi_o(i), & i \text{ odd} \\ \psi_e(i), & i \text{ even.} \end{cases}$$

However, this means that w_n commences with

$$f(123\cdots(n-1)n1) = \psi_o(1)\psi_e(2)\psi_o(3)\cdots\psi_e(n-1)\psi_o(n)\psi_o(1).$$

This is a contradiction, as w_n actually commences with

$$f(123\cdots(n-1)n1) = \psi_o(1)\psi_e(2)\psi_o(3)\cdots\psi_e(n-1)\psi_o(n)\psi_e(1).\square$$

To finish our proof we prove the following lemma, giving a contradiction:

Lemma 4.6 *If $i \in \Sigma_n$, then $|f(i)| \equiv 0$ (mod n).*

Proof: As we remarked at the end of the last section, word $u_i = \psi_e(i+1)\psi_o(i)$ occurs in w_n in a prefix $p_1 u_i$ for some p_1 concatenated from ψ-blocks. This implies that

$$f(p_1)f(u_i) \text{ is a prefix of } w_n. \tag{2}$$

Also, w_n contains u_i as a subword in a prefix

$$p_2 i\underline{|(i-1)(i-2)\cdots(i+1)i|(i+1)(i+2)\cdots(i-1)i}|.$$

where $p_2 i$ is concatenated from ψ-blocks. This means that

$$f(p_2)f(u_i) \text{ is a prefix of } w_n. \tag{3}$$

By assumption, $|f(\psi(1))| > n$. Since $|f(\psi(1))| \equiv 0 \pmod n$, $|f(\psi(1))| \geq 2n$. Since u_i is concatenated from two ψ-blocks, we have $|f(u_i)| \geq 4n$. Thus by Corollary 3.6, $f(u_i)$ is unambiguous. From (2) and (3) then find that $|f(p_2)| \equiv |f(p_1)| \pmod n$. Since $|f(p_1)| \equiv |f(p_2 i)| = |f(p_2)| + |f(i)| \equiv 0 \pmod n$, this implies that $|f(i)| \equiv 0$.□

References

[1] S. Aršon, Démonstration de l'existence des suites asymétriques infinies, *Mat. Sb., (N.S.)* **2**, 769–779; Zbl. **18**, 115.

[2] Kirby A. Baker, George. F. McNulty & Walter Taylor, Growth problems for avoidable words, *Theoret. Comput. Sci.* **69** (1989), no. 3, 319–345; MR **91f**:68109.

[3] Dwight R. Bean, Andrzej Ehrenfeucht & George McNulty, Avoidable Patterns in Strings of Symbols, *Pacific J. Math.* **85** (1979), 261–294; MR **81i**:20075.

[4] Jean Berstel, Mots sans carré et morphisms itéré, *Disc. Math.* **29** (1979), 235-244.

[5] Julien Cassaigne, *Motifs évitables et régularités dans les mots*, thèse de doctorat, L.I.T.P., Université Paris 6 (1994).

[6] James D. Currie, Open problems in pattern avoidance, *Amer. Math. Monthly* **100** (1993), 790–793.

[7] Alan Cobham, Uniform Tag Sequences, *Systems Theory* **6** (1972), 164–192.

[8] Sergey Kitaev, *Symbolic Sequences, Crucial Words and Iterations of a Morphism*, Thesis for Licentiate of Philosophy, Göteburg (2000).