RAIRO-Theor. Inf. Appl. 44 (2010) 139–150 Available online at:
DOI: 10.1051/ita/2010009 www.rairo-ita.org

THE CRITICAL EXPONENT OF THE ARSHON WORDS

Dalia Krieger¹

Abstract. Generalizing the results of Thue (for n=2) [Norske Vid. Selsk. Skr. Mat. Nat. Kl. 1 (1912) 1–67] and of Klepinin and Sukhanov (for n=3) [Discrete Appl. Math. 114 (2001) 155–169], we prove that for all $n \geq 2$, the critical exponent of the Arshon word of order n is given by (3n-2)/(2n-2), and this exponent is attained at position 1.

Mathematics Subject Classification. 68R15.

1. Introduction

In 1935, the Russian mathematician Solomon Efimovich Arshon* [1,2] gave an algorithm to construct an infinite cube-free word over 2 letters, and an algorithm to construct an infinite square-free word over n letters for each $n \geq 3$. The binary word he constructed turns out to be exactly the celebrated *Thue-Morse word*, $\mathbf{t} = 01101001\dots[4,11]$; the square-free words are now known as the *Arshon words*. For $n \geq 2$, the Arshon word of order n is denoted by $\mathbf{a}_n = a_0a_1a_2\dots$

A square or a 2-power is a word of the form xx, where x is a nonempty word. Similarly, an n-power is a word of the form $x^n = xx \dots x$ (n times). The notion of integral powers can be generalized to fractional powers. A non-empty finite word z over a finite alphabet Σ is a fractional power if it has the form $z = x^n y$, where x is a non-empty word, x is a positive integer, and x is a prefix of x, possibly empty. If |z| = p and |x| = q, we say that x is a (p/q)-power, or x is a (p/q)-power, or x is a (p/q)-power, or x is a (p/q)-power.

 $Keywords\ and\ phrases.$ Arshon words, critical exponent.

¹ Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, POB 26, Rehovot 76100, Israel; dalia.krieger@gmail.com

^{*}Vilenkin, in his 1991 article "Formulas on cardboard" [12], says that Arshon was arrested by the Soviet regime and died in prison, most likely in the late 1930's or early 1940's.

140 D. Krieger

Let $\alpha > 1$ be a real number. A right-infinite word \mathbf{w} over Σ is said to be α -power-free (resp. α^+ -power-free), or to avoid α -powers (resp. α^+ -powers), if no subword of it is an r-power for any rational $r \geq \alpha$ (resp. $r > \alpha$). Otherwise, \mathbf{w} contains an α -power. The critical exponent of \mathbf{w} , denoted by $E(\mathbf{w})$, is the supremum of the set of exponents $r \in \mathbb{Q}_{\geq 1}$, such that \mathbf{w} contains an r-power; it may or may not be attained.

Arshon constructed the words \mathbf{a}_n , $n \geq 3$, as square-free infinite words. But actually, these words avoid smaller powers. In 2001, Klepinin and Sukhanov [8] proved that $E(\mathbf{a}_3) = 7/4$, and the bound is attained; that is, \mathbf{a}_3 avoids $(7/4)^+$ -powers. In this paper we generalize the result of Klepinin and Sukhanov, and prove the following theorem:

Theorem 1.1. Let $n \geq 2$, and let $\mathbf{a}_n = a_0 a_1 a_2 \dots$ be the Arshon word of order n. Then the critical exponent of \mathbf{a}_n is given by $E(\mathbf{a}_n) = (3n-2)/(2n-2)$, and $E(\mathbf{a}_n)$ is attained by a subword beginning at position 1.

2. Definitions and notation

Let $\Sigma_n = \{0, 1, \dots, n-1\}$ be an alphabet of size $n, n \geq 3$. In what follows, we use the notation $a \pm 1$, where $a \in \Sigma_n$, to denote the next or previous letter in lexicographic order, and similarly we use the notation a+b, a-b, where $a, b \in \Sigma_n$; all sums of letters are taken modulo n.

Define two morphisms over Σ_n as follows:

$$\varphi_{e,n}(a) = a(a+1)\dots(n-1)01\dots(a-2)(a-1), \quad a = 0,1,\dots,n-1;$$

 $\varphi_{o,n}(a) = (a-1)(a-2)\dots10(n-1)\dots(a+1)a, \quad a = 0,1,\dots,n-1.$

The letters 'e' and 'o' stand for "even" and "odd", respectively. Both $\varphi_{e,n}$ and $\varphi_{o,n}$ are n-uniform (that is, $|\varphi_{e,n}(a)| = n$ for all $a \in \Sigma_n$, and similarly for $\varphi_{o,n}$) and marked (that is, $\varphi_{e,n}(a)$ and $\varphi_{e,n}(b)$ have no common prefix or suffix for all $a \neq b \in \Sigma_n$, and similarly for $\varphi_{o,n}$). The Arshon word of order n can be generated by alternately iterating $\varphi_{e,n}$ and $\varphi_{o,n}$: define an operator $\varphi_n : \Sigma_n^* \to \Sigma_n^*$ by

$$\varphi_n(a_i) = \begin{cases} \varphi_{e,n}(a_i), & \text{if } i \text{ is even;} \\ \varphi_{o,n}(a_i), & \text{if } i \text{ is odd.} \end{cases}$$
 (2.1)

That is, if $u = a_0 a_1 \dots a_m \in \Sigma_n^*$, then $\varphi_n(u) = \varphi_{e,n}(a_0) \varphi_{o,n}(a_1) \varphi_{e,n}(a_2) \varphi_{o,n}(a_3) \dots$ The Arshon word of order n is given by

$$\mathbf{a}_n = \lim_{k \to \infty} \varphi_n^k(0). \tag{2.2}$$

Note that $\varphi_n^k(0)$ is a prefix of $\varphi_n^{k+1}(0)$ for all $k \geq 0$, and the limit is well defined.

Example 2.1. For n = 3, the even and odd Arshon morphisms are given by

$$\varphi_{e,3}: \left\{ \begin{array}{ccc} 0 & \to & 012 \\ 1 & \to & 120, & \varphi_{o,3}: \\ 2 & \to & 201 \end{array} \right. : \left\{ \begin{array}{ccc} 0 & \to & 210 \\ 1 & \to & 021, \\ 2 & \to & 102 \end{array} \right.$$

and the Arshon word of order 3 is given by

$$\mathbf{a}_3 = \lim_{k \to \infty} \varphi_3^k(0) = \underbrace{012}_{\varphi_{e,3}(0)} \underbrace{021}_{\varphi_{o,3}(1)} \underbrace{201}_{\varphi_{e,3}(2)} \underbrace{210}_{\varphi_{o,3}(0)} \dots$$

It is not difficult to see that when n is even, the i'th letter of \mathbf{a}_n is even if and only if i is an even position (for a formal proof, see Séébold [9,10]). Therefore, when n is even, the map φ_n becomes a morphism, denoted by α_n :

$$\alpha_n(a) = \begin{cases} \varphi_{e,n}(a), & \text{if } a \text{ is even;} \\ \varphi_{o,n}(a), & \text{if } a \text{ is odd.} \end{cases}$$
 (2.3)

When n is odd no such partition exists, and indeed, \mathbf{a}_n cannot be generated by iterating a morphism. This fact was proved for \mathbf{a}_3 by Berstel [3] and Kitaev [6,7], and for any odd n by Currie [5].

An occurrence of a subword within \mathbf{a}_n is a triple (z,i,j), where z is a subword of \mathbf{a}_n , $0 \le i \le j$, and $a_i \dots a_j = z$. In other words, z occurs in \mathbf{a}_n at positions i, \dots, j . We usually refer to an occurrence (z,i,j) as $z = a_i \dots a_j$. The set of all subwords of \mathbf{a}_n is denoted by $\mathrm{Sub}(\mathbf{a}_n)$. The set of all occurrences of subwords within \mathbf{a}_n is denoted by $\mathrm{Occ}(\mathbf{a}_n)$. An occurrence (z,i,j) contains an occurrence (z',i',j') if $i \le i'$ and $j \ge j'$.

A subword v of \mathbf{a}_n admits an interpretation by φ_n if there exists a subword $v' = v_0 v_1 \dots v_k v_{k+1}$ of \mathbf{a}_n , $v_i \in \Sigma_n$, such that $v = y_0 \varphi_n(v_1 \dots v_k) x_{k+1}$, where y_0 is a suffix of $\varphi_n(v_0)$ and x_{k+1} is a prefix of $\varphi_n(v_{k+1})$. The word v' is called an ancestor of v.

For an occurrence $z \in \text{Occ}(\mathbf{a}_n)$, we denote by inv(z) the inverse image of z under φ_n . That is, inv(z) is the shortest occurrence $z' \in \text{Occ}(\mathbf{a}_n)$ such that $\varphi_n(z')$ contains z. Note that the word (rather than occurrence) inv(z) is an ancestor of the word z, but not necessarily a unique one.

Following Currie [5], we refer to the decomposition of \mathbf{a}_n into images under φ_n as the φ -decomposition, and to the images of the letters as φ -blocks. We denote the borderline between two consecutive φ -blocks by '|'; e.g., i|j means that i is the last letter of a block and j is the first letter of the following block. If $z = a_i \dots a_j \in \operatorname{Occ}(\mathbf{a}_n)$ begins at an even position we write $z = a_i^{(e)} a_{i+1}^{(o)} a_{i+2}^{(e)} \dots$, and similarly for an occurrence that begins at an odd position.

142 d. krieger

3. General properties of the Arshon words

Lemma 3.1. For all $n \geq 2$, \mathbf{a}_n contains a (3n-2)/(2n-2)-power beginning at position 1.

Proof. For n=2, $\mathbf{a}_2=\mathbf{t}=0110\ldots$, which contains the 2-power 11 at position 1. For $n\geq 3$, \mathbf{a}_n begins with

$$\varphi_{e,n}(0)\varphi_{o,n}(1)\varphi_{e,n}(2) = 012\dots(n-1)|0(n-1)\dots21|2\dots(n-1)01$$
$$= 0 (12\dots(n-1)0(n-1)\dots2)^{(3n-2)/(2n-2)} 1. \quad \Box$$

Example 3.2.

Corollary 3.3. The critical exponent of \mathbf{a}_n satisfies $(3n-2)/(2n-2) \leq E(\mathbf{a}_n) \leq 2$ for all $n \geq 2$.

Proof. For n=2, it is well known that $E(\mathbf{a}_n)=E(\mathbf{t})=2$ [4,11]. For $n\geq 3$, we know by Arshon [1,2] that \mathbf{a}_n is square-free, and so $E(\mathbf{a}_n)\leq 2$. The lower bound follows from Lemma 3.1.

Lemma 3.4. Let $n \geq 3$, and let $i, j \in \Sigma_n$.

- (1) If $ij \in Occ(\mathbf{a}_n)$, then $j = i \pm 1$.
- (2) The borderline between two consecutive φ -blocks has the form i|ji or ij|i. Moreover, a word of the form iji can occur only at a borderline.

Proof. If ij occurs within a φ -block, then $j=i\pm 1$ by definition of φ_n . Suppose i is the last letter of a φ -block and j is the first letter of the next φ -block, and let $kl=\operatorname{inv}(ij)$. Assume $j\neq i\pm 1$, and suppose further that ij is the first pair that satisfies this inequality. Then $l=k\pm 1$, and so there are four cases:

$$\begin{array}{lclcl} \varphi_n(kl) & = & \varphi_{e,n}(k)\varphi_{o,n}(k+1), & \varphi_n(kl) & = & \varphi_{o,n}(k)\varphi_{e,n}(k+1), \\ \\ \varphi_n(kl) & = & \varphi_{e,n}(k)\varphi_{o,n}(k-1), & \varphi_n(kl) & = & \varphi_{o,n}(k)\varphi_{e,n}(k-1). \end{array}$$

But it is easy to check that for all the cases above, $j = i \pm 1$, a contradiction.

For the second assertion, observe that by definition of φ_n , a φ -block is either strictly increasing or strictly decreasing, and two consecutive blocks have alternating directions. By the above, a change of direction can have only the form i|ji or ij|i.

Definition 3.5 (Currie [5]). A mordent is a word of the form iji, where $i, j \in \Sigma_n$ and $j = i \pm 1$. Two consecutive mordents occurring in \mathbf{a}_n are either near mordents,

far mordents, or neutral mordents, according to the position of the borderlines:

```
\begin{split} i|ji \ u \ kl|k &= \text{near mordents, } |u| = n-4; \\ ij|i \ u \ k|lk &= \text{far mordents, } |u| = n-2; \\ i|ji \ u \ k|lk &= \text{neutral mordents, } |u| = n-3; \\ ij|i \ u \ kl|k &= \text{neutral mordents, } |u| = n-3. \end{split}
```

Note that for n = 3, near mordents are overlapping: $\mathbf{a}_3 = 012|021|201|\dots$

Since \mathbf{a}_n is square-free, a p/q-power occurring in \mathbf{a}_n has the form xyx, where q = |xy|, p = |xyx|, and both x, y are nonempty.

Definition 3.6. Let $z = a_i \dots a_j \in \text{Occ}(\mathbf{a}_n)$ be a p/q-power. We say that z is left-stretchable (resp. right-stretchable) if the q-period of z can be stretched left (resp. right), i.e., if $a_{i-1} = a_{i+q-1}$ (resp. $a_{j+1} = a_{j-q+1}$). If the q-period of z can be stretched neither left nor right, we say that z is an unstretchable p/q-power.

Since the critical exponent is a supremum, it is enough to consider unstretchable powers when computing it.

Lemma 3.7. Let $n \ge 3$. Let $z = xyx = (xy)^{p/q} \in Occ(\mathbf{a}_n)$ be an unstretchable power such that $|x| \le n$ and x contains no mordents. Then $p/q \le (3n-2)/(2n-2)$.

Proof. Since $|x| \le n$, it is enough to consider y such that $|y| \le n-2$, for otherwise we would get that p/q < (3n-2)/(2n-2). Therefore, $|xy| = q \le 2n-2$ and $|z| \le 3n-2$. We get that xy is contained in at most 3 consecutive φ -blocks and z is contained in at most 4 consecutive φ -blocks. Suppose z is not contained in 3 consecutive φ -blocks. Let $B_0B_1B_2B_3$ be the blocks containing z, and assume that B_0 is even (the other case is similar). Since $|x| \le n$, necessarily xy begins in B_0 and ends in B_2 . Since x contains no mordents, x has to start at the last letter of B_0 : otherwise, we would get that x cannot extend beyond the first letter of B_1 , and since $|y| \le n-2$, we would get that z is contained in 3 φ -blocks. Therefore, the letters of x are decreasing. Now, since $|xy| \le 2n-2$, the second occurrence of x begins at least 3 letters from the end of B_2 . Since B_2 is an even block, we get a contradiction if |x| > 1. But if |x| = 1 then z is contained in $B_0B_1B_2$. We can assume therefore that z is contained in 3 consecutive φ -blocks, $B_0B_1B_2$. We assume that B_0 is even (the other case is symmetric).

If xy is contained in one block then, because B_0, B_2 are even and B_1 is odd, necessarily |x| = 1, and so $p/q \le 3/2 < (3n-2)/(2n-2)$. If xy begins in B_0 and ends in B_2 , then, since $|y| \le n-2$, the first x occurrence has to end at the third letter of B_1 or later. Since x contains no mordents, this implies that xy begins at the last letter of B_0 and the letters of x are decreasing. Since x is even, again |x| = 1.

Assume xy begins in B_0 and ends in B_1 . Again, because B_0 is even and B_1 is odd, in order for x to contain more than one letter the second x occurrence has to start either at the last letter of B_1 , or at the first letter of B_2 .

144 D. KRIEGER

Let $B_0 = \varphi_{e,n}(i)$. Then there are four cases for B_1, B_2 :

- (1) $B_1 = \varphi_{o,n}(i+1), B_2 = \varphi_{e,n}(i);$
- (2) $B_1 = \varphi_{o,n}(i-1), B_2 = \varphi_{e,n}(i);$
- (3) $B_1 = \varphi_{o,n}(i+1), B_2 = \varphi_{e,n}(i+2);$
- (4) $B_1 = \varphi_{o,n}(i-1), B_2 = \varphi_{e,n}(i-2).$

We now check what the maximal possible exponent is in each of these cases. Without loss of generality, we can assume i = 0. We use the notation z = xyx', where x' is the second occurrence of x in z.

Case 1: $B_0B_1B_2 = |01...(n-1)|0(n-1)...1|01...(n-1)|$.

If x' starts at the last letter of B_1 then |x| = 1, since 10 does not occur anywhere before. If x' starts at the first letter of B_2 , the only possible power is the 3n/2n-power $B_0B_1B_0$, which contradicts the hypothesis $|y| \le n-2$.

Case 2: $B_0B_1B_2 = |01...(n-1)|(n-2)(n-3)...0(n-1)|01...(n-1)|$. By the same argument, either |x| = 1 or z is a 3n/2n-power.

Case 3: $B_0B_1B_2 = |01...(n-1)|0(n-1)...1|23...(n-1)01|$.

If x' starts at the last letter of B_1 , we get the (3n-2)/(2n-2)-power described in Lemma 3.1. If x' starts at the first letter of B_2 , then x has to start at the 2 in B_0 . But then the power is left-stretchable, to the (3n-2)/(2n-2)-power described above.

Case 4: $B_0B_1B_2 = |01...(n-1)|(n-2)(n-3)...0(n-1)|(n-2)(n-1)0...(n-3)|$.

If x' starts at the last letter of B_1 , then x has to start at the last letter of B_0 . But then |x| = 2, since $(n-1) \neq (n-3)$. We get that z is an (n+2)/n-power, and (n+2)/n < (3n-2)/(2n-2) for all $n \geq 3$. If x' starts at the first letter of B_2 , then x has to start at the second last letter of B_0 . Again, |x| = 2, and z is an (n+4)/(n+2)-power, where (n+4)/(n+2) < (3n-2)/(2n-2) for all $n \geq 2$.

In what follows, we will show that in order to compute $E(\mathbf{a}_n)$, it is enough to consider powers xyx such that $|x| \leq n$ and x contains no mordents.

Definition 3.8. Let z be a subword of \mathbf{a}_n . We say that (z_1, z_2) is a synchronization point of z under φ_n if $z = z_1 z_2$, and whenever $\varphi_n(u) = v_1 z v_2$ for some $u, v_1, v_2 \in \operatorname{Sub}(\mathbf{a}_n)$, we have $u = u_1 u_2$, $\varphi_n(u_1) = v_1 z_1$, and $\varphi_n(u_2) = z_2 v_2$. That is, $z_1 | z_2$ is always a borderline in the φ -decomposition of z, regardless of the position in \mathbf{a}_n where z occurs. We say that a subword $z \in \operatorname{Sub}(\mathbf{a}_n)$ is synchronized if it can be decomposed unambiguously under φ_n , in which case it has a unique ancestor.

Lemma 3.9. If $z \in Sub(\mathbf{a}_n)$ has a synchronization point then z is synchronized.

Proof. Suppose z has a synchronization point, z = u|v. If |u| = |v| = 1 then z cannot have a synchronization point at u|v, since uv occurs either in $\varphi_{e,n}(u)$ or in $\varphi_{o,n}(u+1)$. Therefore, at least one of u, v has length > 1. Suppose |u| > 1. If the last two characters of u are increasing, we know that an even φ -block ends at u and an odd φ -block starts at v, and vice versa if the last two characters of u are

decreasing. Since both $\varphi_{e,n}$ and $\varphi_{o,n}$ are uniform marked morphisms, and since we know φ -blocks alternate between even and odd, we can infer inv(z) unambiguously from u|v.

Lemma 3.10. Let $n \geq 3$, and let $z = xyx = (xy)^{p/q} \in Occ(\mathbf{a}_n)$ be an unstretchable p/q-power, such that x has a synchronization point. Then there exists an r/s-power $z' \in Occ(\mathbf{a}_n)$, such that p = nr, q = ns, and $z = \varphi_n(z')$.

Proof. Since x has a synchronization point, it has a unique decomposition under φ_n . Suppose x does not begin at a borderline of φ -blocks. Then x=t|w, where t is a nonempty suffix of a φ -block, and z=t|wyt|w. But since the interpretation is unique, both occurrences of t must be preceded by a word s, such that st is a φ -block. Thus z can be stretched by s to the left, a contradiction. Therefore, x begins at a borderline, and so y ends at a borderline. For the same reason, x must end at a borderline, and so y must begin at a borderline. We get that both x and y have an exact decomposition into φ -blocks, and this decomposition is unique. In particular, both occurrences of x have the same inverse image under φ_n . Let k,l be the number of φ -blocks composing x,y, respectively. Then p=n(2k+l), q=n(k+l), and $\varphi_n^{-1}(z)=\varphi_n^{-1}(x)\varphi_n^{-1}(y)\varphi_n^{-1}(x)$ is a (2k+l)/(k+l)-power. \square

Corollary 3.11. To compute $E(\mathbf{a}_n)$, it is enough to consider powers z = xyx such that x has no synchronization points.

4. Arshon words of even order

To illustrate the power structure in Arshon words of even order, consider a_4 :

		$arphi_{e,4}$	$\varphi_{o,4}$	α_4
0	\rightarrow	0123	3210	0123
1	\longrightarrow	1230	0321	0321
2	\longrightarrow	2301	1032	2301
3	\longrightarrow	3012	2103	2103

 $\mathbf{a}_4 = 0123|0321|2301|2103|0123|2103|2301|0321|2301|2103|0123|0321|2301|0321|\dots$

Lemma 4.1. Let $n \geq 4$ be even, and let $x \in Sub(\mathbf{a}_n)$ be a subword that has no synchronization point. Then $|x| \leq n$ and x contains no mordents.

Proof. In general, a mordent iji can admit two possible borderlines: ij|i or i|ji. However, if n is even, all images under α_n begin with an even letter and end with an odd letter; images of odd letters under $\varphi_{e,n}$ and images of even letters under $\varphi_{o,n}$ are never manifested. Therefore, every mordent admits exactly one interpretation: if i is even and j is odd the interpretation has to be ij|i, and vice versa for odd i. Thus, if x has no synchronization point it contains no mordents.

Suppose x contains no mordents. Then $|x| \le n+2$, and the letters of x are either increasing or decreasing. Assume they are increasing. If |x| = n+2 then x has exactly one interpretation, $x = i|(i+1)\dots(i-1)i|(i+1)$, or else we would get

146 d. krieger

that \mathbf{a}_n contains two consecutive even blocks. If |x| = n + 1 then a priori x has two possible interpretations: $x = i|(i+1)\dots(i-1)i|$ or $x = |i(i+1)\dots(i-1)|i|$. However, the first case is possible if and only if i is odd, since for an even n no φ -block ends with an even letter. Similarly, the second case is possible if and only if i is even.

Lemma 4.1, together with Corollary 3.11 and Lemma 3.7, completes the proof of Theorem 1.1 for all even $n \ge 4$.

5. Arshon words of odd order

To illustrate the power structure in Arshon words of odd order, consider a_5 :

		$arphi_{e,5}$	$\varphi_{o,5}$
0	\rightarrow	01234	43210
1	\rightarrow	12340	04321
2	\rightarrow	23401	10432
3	\longrightarrow	34012	21043
4	\longrightarrow	40123	32104

 $\mathbf{a}_5 =$

 $01234|04321|23401|21043|40123|43210|40123|21043|23401|04321|23401|21043|\dots$

Lemma 5.1. Let $n \geq 3$ be odd. Then every subword $z \in Sub(\mathbf{a}_n)$ with $|z| \geq 3n$ has a unique interpretation under φ_n .

Proof. Consider a subword that contains a pair of consecutive mordents, $z=iji\ u\ klk$. If |u|=n-4 (that is, these are near mordents), then z contains two synchronization points, $z=i|ji\ u\ kl|k$: otherwise, we get a φ -block that contains a repeated letter, a contradiction. Similarly, if |u|=n-2 (a pair of far mordents), z contains the synchronization points $z=ij|i\ u\ k|lk$. To illustrate, consider a_5 : let $z=a_4\ldots a_{10}=404$ 3 212. A borderline 40|4 implies that 43212 is a φ -block, a contradiction; a borderline 2|12 implies that 40432 is a φ -block, again a contradiction. Now let $z=a_8\ldots a_{16}=212$ 340 121. A borderline 2|12 implies that 121 is a prefix of a φ -block, while a borderline 12|1 implies that 212 is a suffix of a φ -block. Again, we get a contradiction.

If |u|=n-3 (neutral mordents), then z has two possible interpretations, either z=i|ji|u|k|lk or z=ij|i|u|k|lk. However, by Currie [5], \mathbf{a}_n does not contain two consecutive pairs of neutral mordents: out of three consecutive mordents, at least one of the pairs is either near or far. (It is also easy to see that this is the case by a simple inverse image analysis: an occurrence of the form ij|i|u|k|l|k|v|rs|r or i|ji|u|k|lk|v|rs|r implies that \mathbf{a}_n contains a square of the form $abab, a, b \in \Sigma_n$, a contradiction: by Arshon, \mathbf{a}_n is square-free.)

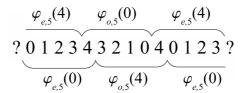


FIGURE 1. Two interpretations under φ_5 .

Let $z \in \text{Occ}(\mathbf{a}_n)$ satisfy |z| = 3n. If z contains a pair of near or far mordents, then z has a unique ancestor. Otherwise, z contains a pair of neutral mordents, $iji \ u \ klk$, where |u| = n - 3, and there are two possible interpretations: $i|ji \ u \ k|lk$ or $ij|i \ u \ kl|k$. Let i'j'i' be the mordent on the left of iji, and let k'l'k' be the mordent on the right of klk. Since no two consecutive neutral mordents occur, i'j'i' and k'l'k' must form near or far mordents with iji and klk.

If the interpretation is $i|ji \ u \ k|lk$, then k'l'k' forms a near pair with klk, while i'j'i' forms a far pair with iji. Therefore, k'l'k' is n-4 letters away from klk, while i'j'i' is n-2 letters away from iji. By assumption, z does not contain a near pair or a far pair, therefore z can contain at most n-2 letters to the right of klk, and at most n letters to the left of iji. Since |z|=3n, this means that either $z=j'|i' \ x \ i|ji \ u \ k|lk \ v \ k'$ or $z=|i' \ x \ i|ji \ u \ k|lk \ v \ k'l'|$, where |x|=n-2 and |v|=n-4. Similarly if the interpretation is $ij|i \ u \ kl|k$, then either $z=|j'i' \ v \ ij|i \ u \ kl|k \ x \ k'|i'$ or $z=i' \ v \ ij|i \ u \ kl|k \ x \ k'|l'$, where |x|=n-2 and |v|=n-4. In any case, z contains enough letters to determine if the far mordent is on the left or on the right, and the interpretation is unique.

Example 5.2. For n=5, the occurrence $z=a_{21}\dots a_{34}=01234321040123$, of length 3n-1=14, has two possible interpretations under φ_5 , as illustrated in Figure 1. However, if either of the left or right question marks is known, the ambiguity is solved: the top interpretation is valid if and only if the left question mark equals 4 (so as to complete the φ -block) and the right question mark equals 2 (so as to complete the near mordent). The bottom interpretation is valid if and only if the left question mark equals 1 (so as to complete the near mordent) and the right question mark equals 4 (so as to complete the φ -block).

Note. Lemma 5.1 is an improvement of a similar lemma of Currie [5], who proved that every occurrence of length 3n + 3 or more has a unique interpretation.

Corollary 5.3. The critical exponent of an odd Arshon word is the largest exponent of powers of the form z = xyx, such that |x| < 3n.

To compute $E(\mathbf{a}_n)$ we need to consider subwords of the form xyx, with x unsynchronized. Moreover, the two occurrences of x should have different interpretations, or else we could take an inverse image under φ_n . For a fixed n, it would suffice to run a computer check on a finite number of subwords of \mathbf{a}_n ; this is exactly the technique Klepinin and Sukhanov employed in [8]. For a general n, we need a more careful analysis.

148 D. KRIEGER

Lemma 5.4. Let $n \geq 3$, n odd. For all mordents in \mathbf{a}_n ,

- (1) $\operatorname{inv}(i(i+1)i) = (i+2)^{(e)}(i+1)^{(o)} \text{ or } \operatorname{inv}(i(i+1)i) = (i+1)^{(e)}(i+2)^{(o)};$ (2) $\operatorname{inv}(i(i-1)i) = (i-1)^{(o)}(i)^{(e)} \text{ or } \operatorname{inv}(i(i-1)i) = (i)^{(o)}(i-1)^{(e)}.$

Proof. A mordent iji can admit two possible borderlines: ij|i or i|ji. Consider the mordent i(i+1)i. If the borderline is i(i+1)|i, then i(i+1) is a suffix of an increasing φ -block, and so the block must be an image under $\varphi_{e,n}$. By definition of $\varphi_{e,n}$, i(i+1) is the suffix of $\varphi_{e,n}(i+2)$. Since even and odd blocks alternate, the next block must be an image under $\varphi_{o,n}$, and by definition, i is the prefix of

If the borderline is i|(i+1)i, then (i+1)i is the prefix of a decreasing φ -block, and by similar considerations this block is $\varphi_{o,n}(i+2)$, while the previous block is $\varphi_{e,n}(i+1)$. The assertion for i(i-1)i is proved similarly.

Lemma 5.5. Let $n \geq 3$, n odd, and let $z \in Occ(\mathbf{a}_n)$.

- (1) If $z = i^{(e)}ui^{(o)}$ or $z = i^{(o)}ui^{(e)}$ for some $i \in \Sigma_n$, then $|u| \ge n 1$; (2) If $z = i^{(e)}u(i\pm 1)^{(e)}$ or $z = i^{(o)}u(i\pm 1)^{(o)}$ for some $i \in \Sigma_n$, then $|u| \ge n 2$.

Proof. Let $z = i^{(e)}ui^{(o)}$, and suppose $i^{(o)}$ does not occur in u (otherwise, if u = $u'i^{(o)}u''$, take $z=i^{(e)}u'i^{(o)}$). If |u|< n-1 then z must contain a mordent in order for i to be repeated. But then the two occurrences of i have the same parity, a contradiction. The rest of the cases are proved similarly.

Lemma 5.6. Let $n \geq 3$ be odd, and let $z = xyx = (xy)^{p/q} \in Occ(\mathbf{a}_n)$ be an unstretchable power, such that x is unsynchronized and contains a mordent. Then $p/q < E(\mathbf{a}_n)$.

Proof. Suppose x contains the mordent i(i+1)i (the case of i(i-1)i is symmetric). Then the two occurrences of the mordent have different interpretations, else we could take an inverse image under φ_n and get a power with the same exponent. By Lemma 5.4, there are two different cases, according to which interpretation comes first:

$$\underbrace{ \begin{array}{c} \varphi_{e,n}(i+2) \\ \dots i(i+1) \end{array} | \underbrace{ \begin{array}{c} \varphi_{o,n}(i+1) \\ \dots i(i+1) \end{array} | \begin{array}{c} \varphi_{o,n}(i+1) \\ \dots \dots \end{array}}_{ \begin{array}{c} p_{e,n}(i+1) \\ \dots \\ \end{array} | \underbrace{ \begin{array}{c} \varphi_{e,n}(i+1) \\ (i+1) \dots (i-1)i \end{array} | \underbrace{ \begin{array}{c} \varphi_{o,n}(i+2) \\ (i+1)i \dots \end{array}}_{ \begin{array}{c} \varphi_{o,n}(i+1) \\ \dots \\ \end{array}}_{ \begin{array}{c} p_{e,n}(i+1) \\ \dots \\ \end{array} | \underbrace{ \begin{array}{c} \varphi_{e,n}(i+2) \\ (i+1)i \dots (i+1) \end{array} | \underbrace{ \begin{array}{c} \varphi_{o,n}(i+1) \\ (i+1)i \dots \end{array}}_{ \begin{array}{c} p_{o,n}(i+1) \\ \dots \\ \end{array}}_{ \begin{array}{c} p_{o,n}(i+1)$$

By Lemma 5.5, in both cases there must be at least n-1 additional φ -blocks between the blocks containing the two i(i+1)i occurrences. Thus, in both cases $q \ge n^2 + n - 1$ (note that q is the length of the period, and can be measured from the beginning of i(i+1)i in the first x to just before i(i+1)i in the second x). Now, x is unsynchronized, and so by Lemma 5.1 |x| < 3n. Thus, $|x|/q \le$ $(3n-1)/(n^2+n-1) < n/(2n-2)$ for all $n \ge 3$, and so p/q = (|x|+q)/q < 1 $(3n-2)/(2n-2) \le E(\mathbf{a}_n).$

By Lemma 5.6, in order to compute $E(\mathbf{a}_n)$ it is enough to consider powers xyx such that x is unsynchronized and contains no mordents. The longest subword that contains no mordents is of length n+2, but such subword implies a far pair, and has a unique ancestor. Therefore, we can assume $|x| \leq n+1$.

Lemma 5.7. Let $n \geq 3$ be odd, and let $z = xyx = (xy)^{p/q} \in Occ(\mathbf{a}_n)$ be an unstretchable power, such that x is unsynchronized, x contains no mordents, and |x| = n + 1. Then $p/q < E(\mathbf{a}_n)$.

Proof. Since |x| = n + 1 and x contains no mordents, necessarily x = ivi, where $i \in \Sigma_n$ and either

$$v = (i+1)\dots(n-1)01\dots(i-2)(i-1),$$

or

$$v = (i-1) \dots 10(n-1) \dots (i+2)(i+1).$$

Suppose the letters of v are increasing, and assume without loss of generality that i = 0. Then x admits two possible interpretations: x = 01...(n-1)|0 or x = 0|1...(n-1)0. The ancestors of the first and second case are $\operatorname{inv}(x) = 0^{(e)}1^{(o)}$ and $\operatorname{inv}(x) = 0^{(o)}1^{(e)}$, respectively. Any other interpretation is impossible, since it implies \mathbf{a}_n contains two consecutive even φ -blocks.

As in the previous lemma, we can assume that the two x occurrences of z have different inverse images. There are two possible cases:

$$\underbrace{01\dots(n-1)}^{\varphi_{e,n}(0)} |\underbrace{0\dots}_{0\dots}| \underbrace{0\dots -2}^{n-2} \underbrace{\varphi_{-\text{blocks}}}_{0\dots\dots} |\underbrace{\dots}_{0\dots}_{1\dots(n-1)0},$$

$$\underbrace{\cdots 0}_{\varphi_{o,n}(0)} \underbrace{|\overbrace{1\dots(n-1)0}|}_{\varphi_{e,n}(1)} |\underbrace{-\frac{\varphi_{e,n}(0)}{n-2}}_{n-2} \underbrace{\varphi_{e,n}(0)}_{\varphi_{o,n}(1)} |\underbrace{\varphi_{o,n}(1)}_{0\dots}.$$

By Lemma 5.5, in both cases y contains at least n-2 additional φ -blocks. Therefore, $q \geq n^2-n+1$, and so $|x|/q \leq (n+1)/(n^2-n+1) < n/(2n-2)$ for all $n \geq 3$. Again, $p/q < (3n-2)/(2n-2) \leq E(\mathbf{a}_n)$.

By Lemma 5.7, to compute $E(\mathbf{a}_n)$ for an odd $n \geq 3$ it is enough to consider powers of the form z = xyx such that $|x| \leq n$ and x contains no mordent. By Lemma 3.7, such powers have exponent at most (3n-2)/(2n-2). This completes the proof of Theorem 1.1.

Acknowledgements. I would like to thank Yuri Pritykin, for providing many interesting details about Arshon's life.

D. KRIEGER

References

- [1] S.E. Arshon, A proof of the existence of infinite asymmetric sequences on n symbols. Matematicheskoe Prosveshchenie (Mathematical Education) 2 (1935) 24–33 (in Russian). Available electronically at http://ilib.mccme.ru/djvu/mp1/mp1-2.htm.
- [2] S.E. Arshon, A proof of the existence of infinite asymmetric sequences on n symbols. *Mat. Sb.* **2** (1937) 769–779 (in Russian, with French abstract).
- [3] J. Berstel, Mots sans carré et morphismes itérés. Discrete Math. 29 (1979) 235–244.
- [4] J. Berstel, Axel Thue's papers on repetitions in words: a translation. Publications du Laboratoire de Combinatoire et d'Informatique Mathématique **20**, Université du Québec à Montréal (1995)
- [5] J.D. Currie, No iterated morphism generates any Arshon sequence of odd order. Discrete Math. 259 (2002) 277–283.
- [6] S. Kitaev, Symbolic sequences, crucial words and iterations of a morphism. Ph.D. thesis, Göteborg, Sweden (2000).
- [7] S. Kitaev, There are no iterative morphisms that define the Arshon sequence and the σ -sequence. J. Autom. Lang. Comb. 8 (2003) 43–50.
- [8] A.V. Klepinin and E.V. Sukhanov, On combinatorial properties of the Arshon sequence. Discrete Appl. Math. 114 (2001) 155–169.
- [9] P. Séébold, About some overlap-free morphisms on a n-letter alphabet. J. Autom. Lang. Comb. 7 (2002) 579–597.
- [10] P. Séébold, On some generalizations of the Thue–Morse morphism. Theoret. Comput. Sci. 292 (2003) 283–298.
- [11] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. Norske Vid. Selsk. Skr. Mat. Nat. Kl. 1 (1912) 1–67.
- [12] N.Ya. Vilenkin, Formulas on cardboard. Priroda 6 (1991) 95-104 (in Russian). English summary available at http://www.ams.org/mathscinet/index.html, review no. MR1143732.