Journal of Integer Sequences, Vol. 16 (2013),
Article 13.9.4

Concerning Kurosaki's Squarefree Word

Serina Camungol and Narad Rampersad
Department of Mathematics and Statistics
University of Winnipeg
515 Portage Ave.
Winnipeg, Manitoba R3B 2E9
Canada
serina.camungol@gmail.com
narad.rampersad@gmail.com

Abstract

In 2008, Kurosaki gave a new construction of a (bi-)infinite squarefree word over three letters. We show that in fact Kurosaki's word avoids $(7/4)^+$-powers, which, as shown by Dejean, is optimal over a 3-letter alphabet.

1 Introduction

In 1906, Thue [8] constructed an infinite word over a 3-letter alphabet that avoided *squares*, i.e., repetitions of the form xx. In 1935, Arshon [1, 2] independently rediscovered this and gave another construction of an infinite squarefree word. In 1972, Dejean [3] generalized Thue's notion of repetition to *fractional powers*. She constructed an infinite ternary word avoiding $(7/4)^+$-powers, i.e., repetitions of the form x^α, where $\alpha > 7/4$. Furthermore, she showed that this was optimal over a ternary alphabet. In 2001, Klepinin and Sukhanov [5] examined Arshon's word in more depth and showed that it also avoided $(7/4)^+$-powers (Krieger [6] extended this result by analyzing Arshon words over larger alphabets.) In 2008, Kurosaki [7] gave a new construction of a squarefree word. In this short note, we show that Kurosaki's word also avoids $(7/4)^+$-powers.

Let us now recall some basic definitions. Let Σ be a finite alphabet and let x be a word over Σ. We denote the length of x by $|x|$. Write $x = x_1x_2\cdots x_n$, where each $x_i \in \Sigma$. The word x has a *period* p if $x_i = x_{i+p}$ for all i. The *exponent* of x is the quantity $|x|/p$, where

p is the least period of x. If x has exponent α, we say that x is an *α-power*. A 2-power is also called a *square*. A word y (finite or infinite) *avoids α-powers* (resp. *avoids α^+-powers*) if every factor of y has exponent less than (resp. at most) α.

2 Kurosaki's word avoids $(7/4)^+$-powers

We now give the definition of Kurosaki's word. We will use the notation from his paper. We define functions $\sigma, \rho : \{1,2,3\}^* \to \{1,2,3\}^*$ as follows: If a is a word over $\{1,2,3\}$, then

- $\sigma(a)$ is the word obtained by exchanging 1's and 2's in a;
- $\rho(a)$ is the word obtained by exchanging 2's and 3's in a.

We define the function $\varphi : \{1,2,3\}^* \to \{1,2,3\}^*$ as follows:

$$\varphi(a) = \sigma(a)a\rho(a)$$

Consider the finite words $\varphi^n(2)$, for $n = 1, 2, \ldots$. These words do not converge to a one-sided infinite word as n tends to infinity; however, since $\varphi^n(2)$ appears as the middle third of $\varphi^{n+1}(2)$ for all n, we see that this sequence of words defines, in the limit, a two-sided, or bi-infinite, word. It is this bi-infinite word

$$\cdots 123213231\ 213123132\ 312132123 \cdots$$

that we refer to as *Kurosaki's word* (see A217522 in Sloane's *Encyclopedia*).

Following Kurosaki, we write $\varphi^n(2) = \{a_i^n\}_{i=0}^{3^n-1}$. Proofs of Theorem 1, Lemma 2 and Lemma 3 can be found in [7].

Theorem 1. *$\{a_i^n\}_{i=0}^{3^n-1}$ is squarefree for all $n \geq 1$.*

Lemma 2. *Each triple of the form $a_{3k}^n a_{3k+1}^n a_{3k+2}^n$, $k = 0, 1, \ldots, 3^{n-1}-1$, is some permutation of 123.*

From now on, when we use the term *triple* we mean a permutation of 123 occurring at a position congruent to 0 modulo 3 in $\varphi^n(2)$. Next we define the operator f, which extracts the middle term of each triple: $f(\varphi^n(2)) = \{a_{3i+1}^n\}_{i=0}^{3^{n-1}-1}$.

Lemma 3. *The sequence $f(\varphi^n(2)) = \{a_{3i+1}^n\}_{i=0}^{3^{n-1}-1}$ is equal to $\varphi^{n-1}(2) = \{a_i^{n-1}\}_{i=0}^{3^{n-1}-1}$.*

Lemma 4. *$\{a_i^n\}_{i=0}^{3^n-1}$ has no factors of the form $a_{3q}a_{3q+1}\cdots a_{3q+11}$ such that $a_{3q} = a_{3q+3} = a_{3q+6} = a_{3q+9}$.*

Proof. For convenience we omit the superscript n from the terms a_i^n. Let $a_{3q}a_{3q+1}\cdots a_{3q+11}$ be a factor of $\{a_i^n\}_{i=0}^{3^n-1}$. As $a_{3q}a_{3q+1}a_{3q+2}$, $a_{3q+3}a_{3q+4}a_{3q+5}$, $a_{3q+6}a_{3q+7}a_{3q+8}$ and $a_{3q+9}a_{3q+10}a_{3q+11}$ are permutations of 123 (Lemma 2), it is not difficult to see that having $a_{3q} = a_{3q+3} = a_{3q+6} = a_{3q+9}$ creates a square of length 6 or 12, in contradiction to Theorem 1. $\square$

We now establish the main result.

Theorem 5. *$\{a_i^n\}_{i=0}^{3^n-1}$ is $(7/4)^+$-power-free for all $n \geq 1$.*

Proof. We will prove that $\varphi^n(2) = \{a_i^n\}_{i=0}^{3^n-1}$ does not contain a $(7/4)^+$-power of the form xyx, where $|x| + |y| = s$, $|x| = r$ and $r/s > 3/4$. We prove this by induction on n, considering various cases for the possible lengths of s. We first check the base cases $n = 1, 2, 3$. Clearly $\varphi^1(2) = 123$ and $\varphi^2(2) = 213123132$ avoid $(7/4)^+$-powers, and by observation $\varphi^3(2) = 123213231213123132312132123$ avoids $(7/4)^+$-powers. So we assume $\varphi^n(2)$ avoids $(7/4)^+$-powers and we show that $\varphi^{n+1}(2)$ is $(7/4)^+$-power-free.

Case 1. $s \equiv 0 \pmod 3$

Suppose towards a contradiction that $\varphi^{n+1}(2)$ contains a $(7/4)^+$-power of the form xyx, where $|x| + |y| = s$, $|x| = r$, $r/s > 3/4$ and $s \equiv 0 \pmod 3$. Let us write

$$xyx = a_p a_{p+1} \cdots a_{p+s-1} a_{p+s} a_{p+s+1} \cdots a_{p+s+r-1},$$

where $a_{p+i} = a_{p+s+i}$ for $0 \leq i \leq r-1$. Again, for convenience we omit the superscript n from the terms a_i^n. If $p \equiv r \equiv 0 \pmod 3$, then $f(\varphi^{n+1}(2))$ contains a $(7/4)^+$-power and by Lemma 3 the word $\varphi^n(2)$ contains a $(7/4)^+$-power, a contradiction to our assumption.

We will show that the repetition xyx can always be extended to the left to give a $(7/4)^+$-power that has the same period s, but now starts at a position congruent to 0 modulo 3. A similar argument shows that the repetition can also be extended to the right so that its ending position is congruent to 2 modulo 3, and therefore that the argument of the previous paragraph can be applied to obtain a contradiction.

If $p \equiv 1 \pmod 3$, let $p = 3q + 1$. The first two terms of x form a suffix of the triple $a_{3q}a_{3q+1}a_{3q+2}$. Lemma 2 allows us to uniquely determine a_{3q} from a_{3q+1} and a_{3q+2}. Similarly, the values of a_{3q+s+1} and a_{3q+s+2} uniquely determine the value of a_{3q+s}. Since $a_{3q+1}a_{3q+2} = a_{3q+s+1}a_{3q+s+2}$, we have $a_{3q} = a_{3q+s}$. Thus $a_{p-1}a_p \cdots a_{p+s+r-1}$ is a $(7/4)^+$-power with period s starting at position $3q$.

If $p \equiv 2 \pmod 3$, let $p = 3q + 2$. In order to determine the two elements preceding a_{3q+2}, we look at the triple succeeding a_{3q+2}, namely $a_{3q+3}a_{3q+4}a_{3q+5}$. By Lemma 2, $a_{3q}a_{3q+1}a_{3q+2}$ is some permutation of $a_{3q+3}a_{3q+4}a_{3q+5}$. We must have either $a_{3q+2} = a_{3q+4}$ or $a_{3q+2} = a_{3q+5}$ (for otherwise we would have a square, contrary to Theorem 1). Without loss of generality, suppose $a_{3q+2} = a_{3q+4}$; we will determine the values of a_{3q} and a_{3q+1}. If $a_{3q+1} = a_{3q+3}$, then $a_{3q+1}a_{3q+2}a_{3q+3}a_{3q+4}$ is a square, contrary to Theorem 1, so it must be that $a_{3q+1} = a_{3q+5}$, which implies $a_{3q} = a_{3q+3}$. In other words, $a_{3q}a_{3q+1}$ is uniquely determined by $a_{3q+2}a_{3q+3}a_{3q+4}a_{3q+5}$. Similarly, $a_{3q+s}a_{3q+s+1}$ is uniquely determined by $a_{3q+s+2}a_{3q+s+3}a_{3q+s+4}a_{3q+s+5}$. We deduce that $a_{p-2}a_p \cdots a_{p+s+r-1}$ is a $(7/4)^+$-power with period s starting at position $3q$.

Case 2. $s < 14$ and $s \not\equiv 0 \pmod 3$

As $\varphi^n(2)$ is $(7/4)^+$-power-free by the induction hypothesis, it is not hard to see that $\sigma(\varphi^n(2))$ and $\rho(\varphi^n(2))$ are also $(7/4)^+$-power-free. Therefore, if $\varphi^{n+1}(2)$ contains a $(7/4)^+$-power, it must be a factor of $\sigma(\varphi^n(2))\varphi^n(2)$ or $\varphi^n(2)\rho(\varphi^n(2))$. Note that here we only consider $n \geq 3$, as the smaller values of n are dealt with in our base cases, and $|\varphi^n(2)| \geq 27$, so if $\varphi^{n+1}(2)$ contains a $(7/4)^+$-power it cannot extend from $\sigma(\varphi^n(2))$ over into $\rho(\varphi^n(2))$. Furthermore, if $s < 14$, then we may assume that the entire repetition xyx has length at most $\lceil 13(7/4) \rceil = 23$.

Observe that for even n, $\varphi^n(2)$ has the form

$$213123132123213231321231 \cdots 312321312132123213123132,$$

so

$$\sigma(\varphi^n(2))\varphi^n(2) = \cdots 321312321231213123213231 \cdot 213123132123213231321231 \cdots$$

and

$$\varphi^n(2)\rho(\varphi^n(2)) = \cdots 312321312132123213123132 \cdot 312132123132312321231321 \cdots ,$$

where the single dots represent the boundaries between $\sigma(\varphi^n(2))$, $\varphi^n(2)$ and $\rho(\varphi^n(2))$.

Similarly, for odd n, $\varphi^n(2)$ has the form

$$123213231213123132312132 \cdots 213231213123132312132123,$$

so

$$\sigma(\varphi^n(2))\varphi^n(2) = \cdots 123132123213231321231213 \cdot 123213231213123132312132 \cdots$$

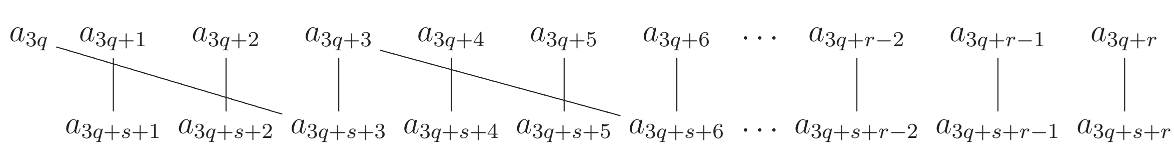
and

$$\varphi^n(2)\rho(\varphi^n(2)) = \cdots 213231213123132312132123 \cdot 132312321312132123213123 \cdots .$$

It suffices to check the factors of length at most 23 crossing the indicated boundaries; by inspection we are able to conclude that $\varphi^{n+1}(2)$ is $(7/4)^+$-power-free for $s < 14$.

Case 3. $s \geq 14$ and $s \not\equiv 0 \pmod 3$

This case is further divided into six subcases. We will prove this case for $p \equiv 1 \pmod 3$, where p is the starting index of our $(7/4)^+$-power and $s \equiv 2 \pmod 3$. The proofs for the other five cases are similar, all involving use of Lemma 4. Suppose towards a contradiction that $\varphi^{n+1}(2)$ contains a $(7/4)^+$-power of the form xyx, where $|x| + |y| = s$, $|x| = r$ and $s \equiv 2 \pmod 3$. We can write this as $a_{3q+1}a_{3q+2} \cdots a_{3q+s}a_{3q+s+1}a_{3q+s+2} \cdots a_{3q+s+r}$, where $a_{3q+i} = a_{3q+s+i}$ for $1 \leq i \leq r$. Now consider the triples $a_{3q}a_{3q+1}a_{3q+2}$ and $a_{3q+s+1}a_{3q+s+2}a_{3q+s+3}$. As $a_{3q+1} = a_{3q+s+1}$ and $a_{3q+2} = a_{3q+s+2}$ and both $a_{3q}a_{3q+1}a_{3q+2}$ and $a_{3q+s+1}a_{3q+s+2}a_{3q+s+3}$ are permutations of 123, it must be that $a_{3q} = a_{3q+s+3}$. Continuing in this manner we see that $a_{3q} = a_{3(q+1)} = \cdots = a_{3(q+j)}$, where $j \leq r$ is maximal and $j \equiv 0 \pmod 3$. This is easy to see using the diagram below, where the lines signify that the terms are equivalent.



However, as $s \geq 14$ it follows that $r \geq 11$ and that $j \geq 3$, so we always have $a_{3q} = a_{3(q+1)} = a_{3(q+2)} = a_{3(q+3)}$, in contradiction to Lemma 4.

□

3 Conclusion

Kurosaki also gives an alternate definition of his sequence. In this other formulation, the n-th term of the sequence is computed by a finite automaton that reads the *balanced ternary* representation of n. This representation is a base-3 representation with digit set $\{-1, 0, 1\}$. Kurosaki says that his sequence can be viewed as an analogue of the classical Thue–Morse sequence for the balanced ternary numeration system. His sequence certainly seems to have many interesting properties and perhaps should be better known.

We would also like to mention that Daniel Goč has verified the main result of this paper by a computer calculation using the techniques described by Goč, Henshall, and Shallit [4].

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).

[3] F. Dejean, Sur un théorème de Thue, *J. Combin. Theory Ser. A* **13** (1972), 90–99.

[4] D. Goč, D. Henshall, and J. Shallit, Automatic theorem-proving in combinatorics on words, in N. Moreira and R. Reis eds., *Conference on Implementation and Application of Automata: Proc. CIAA 2012*, Lect. Notes in Comp. Sci., Vol. 7381, Springer, 2012, pp. 180–191.

[5] A. V. Klepinin and E. V. Sukhanov, On combinatorial properties of the Arshon sequence, *Disc. Appl. Math.* **114** (2001), 155–169.

[6] D. Krieger, The critical exponent of the Arshon words, *RAIRO Theor. Inform. Appl.* **44** (2010), 139–150.

[7] T. Kurosaki, Direct definition of a ternary infinite square-free sequence, *Inform. Process. Lett.* **106** (2008), 175–179.

[8] A. Thue, Über unendliche Zeichenreihen, *Kra. Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl.* **7** (1906), 1–22.

2010 *Mathematics Subject Classification*: Primary 68R15.
Keywords: repetition, squarefree word, $(7/4)^+$-power.

(Concerned with sequence A217522.)

Received August 9 2013; revised version received October 3 2013. Published in *Journal of Integer Sequences*, October 12 2013.

Return to Journal of Integer Sequences home page.