

Avoiding $3/2$ -Powers over the Natural Numbers

Eric Rowland

School of Computer Science

University of Waterloo

Waterloo, ON N2L 3G1

Canada

erowland@uwaterloo.ca

Jeffrey Shallit

School of Computer Science

University of Waterloo

Waterloo, ON N2L 3G1

Canada

shallit@cs.uwaterloo.ca

December 3, 2011

Abstract

In this paper we answer the following question: what is the lexicographically least sequence over the natural numbers that avoids $\frac{3}{2}$ -powers?

1 Introduction

Ever since the work of Thue more than a hundred years ago, mathematicians have been interested in avoiding patterns in words. Thue showed that it is possible to create an infinite sequence over a three-letter alphabet containing no nonempty squares (that is, no factors of the form xx), and over a two-letter alphabet containing no overlaps (that is, no factors of the form $axaxa$, where a is a single letter and x is a possibly empty word).

Dejean [3] instituted the study of fractional powers in 1972. We say that a word x is a p/q -power, for integers $p > q \geq 1$, if x can be written in the form $y^e y'$ for some integer $e \geq 1$, where y' is a prefix of y and $|x|/|y| = p/q$. For example, the German word *schematische* is a $3/2$ -power.

Recently there has been some interest in avoiding patterns over an infinite alphabet, say \mathbb{N} , the natural numbers. More precisely, we are interested in finding the *lexicographically least* infinite sequence over \mathbb{N} avoiding p/q -powers, by which we mean a sequence not containing any factor that is an α -power with $\alpha \geq p/q$. As shown in [4], this corresponds to running a backtracking algorithm without actually doing any backtracking; whenever a choice for the next symbol fails, we increment the choice by 1 until a valid choice is found.

Guay-Paquet and the second author [4] recently described the lexicographically least words avoiding squares and overlaps over \mathbb{N} . The lexicographically least square-free word over \mathbb{N} is

01020103010201040102010301020105

Indexing from 0, the n th symbol in this sequence is the exponent of the highest power of 2 dividing $n + 1$. This sequence is 2-regular in the sense of Allouche and Shallit [1, 2]. The lexicographically least overlap-free word can also be described compactly, but it is more complicated.

In this paper we are interested in $\mathbf{w}_{3/2}$, the lexicographically least word over \mathbb{N} avoiding $\frac{3}{2}$ -powers. As we will see, this word has a short description, and it is 6-regular.

Here are the first 100 terms of $\mathbf{w}_{3/2}$:

$\mathbf{w}_{3/2}[10i + j]$	$j = 0$	1	2	3	4	5	6	7	8	9
$i = 0$	0	1	2	0	3	1	0	2	1	3
1	0	1	2	0	4	1	0	2	1	4
2	0	1	2	0	3	1	0	2	1	5
3	0	1	2	0	4	1	0	2	1	3
4	0	1	2	0	3	1	0	2	1	4
5	0	1	2	0	4	1	0	2	1	5
6	0	1	2	0	3	1	0	2	1	3
7	0	1	2	0	4	1	0	2	1	4
8	0	1	2	0	3	1	0	2	1	6
9	0	1	2	0	4	1	0	2	1	3

2 The word $\mathbf{w}_{3/2}$

We define an infinite sequence $\mathbf{v} = a(0)a(1)a(2)\cdots$ as follows:

$$\begin{aligned}
 a(10n) &= a(10n + 3) = a(10n + 6) = 0, & n \geq 0; \\
 a(10n + 1) &= a(10n + 5) = a(10n + 8) = 1, & n \geq 0; \\
 a(10n + 2) &= a(10n + 7) = 2, & n \geq 0; \\
 a(10n + 4) &= \begin{cases} 3, & \text{if } n \equiv 0 \pmod{2}; \\ 4, & \text{if } n \equiv 1 \pmod{2}; \end{cases} & (1) \\
 a(10n + 9) &= \begin{cases} 3, & \text{if } n \equiv 0 \pmod{3}; \\ 4, & \text{if } n \equiv 1 \pmod{3}; \\ a(5(n - 2)/3 + 4) + 2, & \text{if } n \equiv 2 \pmod{3}. \end{cases}
 \end{aligned}$$

Thus we have

$$\mathbf{v} = 01203102130120410214012031021501204102130120310214\cdots.$$

Note that \mathbf{v} has both periodic aspects (in that some linearly-indexed subsequences are constant) and self-similar aspects (in that $a(60n + 59) = a(10n + 9) + 2$).

Our goal is to show that $\mathbf{v} = \mathbf{w}_{3/2}$. To do so we will prove that (a) \mathbf{v} is $\frac{3}{2}$ -power-free and (b) \mathbf{v} is lexicographically least among all words avoiding $\frac{3}{2}$ -powers.

3 \mathbf{v} is $\frac{3}{2}$ -power-free

If w is a $\frac{p}{q}$ -power with $\frac{p}{q} \geq \frac{3}{2}$, then w contains a factor of the form xyx , where either $|y| = |x|$ or $|y| = |x| - 1$. We show that neither of these factors appears in \mathbf{v} .

First, we observe that from the definition, \mathbf{v} is “pseudoperiodic” with period 10. More precisely,

$$\mathbf{v} \in (0120\{3, 4\}1021\{3, 4, \dots\})^\omega,$$

where by x^ω we mean the infinite sequence $xxx\dots$. From this we immediately see that if $a(n)a(n+1) = a(n+i)a(n+i+1)$ for some $n \geq 0$ and $i \geq 0$ then $i \equiv 0 \pmod{10}$. It now follows that if $|y| = |x| - 1 \geq 1$ then, since $|xy|$ is odd, xyx is not a factor of \mathbf{v} . Similarly, if $|x| = 1$ then xx is not a factor of \mathbf{v} .

We now show that no word xyx , with $|x| = |y| = k \geq 1$, occurs as a factor of \mathbf{v} . For each factor $a(n)a(n+1)\dots a(n+3k-1)$ of length $3k$ we show that this factor is not a $\frac{3}{2}$ -power. If $k \not\equiv 0 \pmod{5}$, then the pseudoperiodicity of \mathbf{v} implies that $a(n)a(n+1) \neq a(n+2k)a(n+2k+1)$, so this factor is not a $\frac{3}{2}$ -power. The remainder of the proof is divided up into two cases. In each case we exhibit an index i , $0 \leq i < k$, such that $a(n+i) \neq a(n+2k+i)$, implying that $a(n)a(n+1)\dots a(n+3k-1)$ is not a $\frac{3}{2}$ -power.

Case 1: $k \equiv 5 \pmod{10}$. Then $k \geq 5$.

If $k = 5$ then we can choose i such that either $n+i \equiv 4 \pmod{10}$ or $n+i \equiv 9 \pmod{10}$. In the former case we have $a(n+i) = 3$, $a(n+2k+i) = 4$ or vice versa. In the latter case we have $a(n+i) \neq a(n+2k+i)$ by (1).

Otherwise $k \geq 15$. Then we can choose i such that $n+i \equiv 4 \pmod{10}$. We have $2k \equiv 10 \pmod{20}$, and hence $a(n+i) \neq a(n+2k+i)$ from (1).

Case 2: $k \equiv 0 \pmod{10}$. Define $b(n) = a(10n+9)$ for $n \geq 0$. Then from (1) we have, for $n \geq 0$, that

$$\begin{aligned} b(6n) &= b(6n+3) = 3; \\ b(6n+1) &= b(6n+4) = 4; \\ b(6n+2) &= \begin{cases} 5, & \text{if } n \equiv 0 \pmod{2}; \\ 6, & \text{if } n \equiv 1 \pmod{2}; \end{cases} \\ b(6n+5) &= b(n) + 2. \end{aligned} \tag{2}$$

From this, an easy induction gives

$$b(n) = \begin{cases} 2t+3, & \text{if the base-6 representation of } n \text{ ends with } 05^t \text{ or } 35^t; \\ 2t+4, & \text{if the base-6 representation of } n \text{ ends with } 15^t \text{ or } 45^t; \\ 2t+5, & \text{if the base-6 representation of } n \text{ ends with } 025^t, 225^t \text{ or } 425^t; \\ 2t+6, & \text{if the base-6 representation of } n \text{ ends with } 125^t, 325^t \text{ or } 525^t. \end{cases} \tag{3}$$

We need to find an index i , $0 \leq i < k$, such that $a(n+i) \neq a(n+2k+i)$. We choose i such that $n+i \equiv 9 \pmod{10}$. Since $k = 10r$, there are r possible choices for i . It follows that we need to show, for each $n \geq 0$ and $r \geq 1$, that there exists j , $0 \leq j < r$ such that

$$b(n+j) \neq b(n+2r+j). \quad (4)$$

If $r \equiv 1, 2 \pmod{3}$, then in fact we can choose $j = 0$ and use (2).

Otherwise $r \equiv 0 \pmod{3}$. This is the most difficult case. To solve it, we first define two auxiliary sequences, as follows:

$$\begin{aligned} c(0) &= 0; \\ c(6n+1) &= c(6n+3) = c(6n+5) = 2; \\ c(6n+2) &= c(6n+4) = 5; \\ c(6n) &= 6c(n) + 5; \end{aligned}$$

$$\begin{aligned} d(0) &= 0; \\ d(6n+1) &= d(6n+5) = 3; \\ d(6n+2) &= d(6n+3) = d(6n+4) = 6; \\ d(6n) &= 6d(n). \end{aligned}$$

Write $r = 3s$, and let the base-6 representation of s end in exactly t zeroes. An easy induction gives, for $s \geq 1$, that

$$c(s) = \begin{cases} 6^{t+1} - 1, & \text{if the last nonzero digit of } s \text{ is even;} \\ 3 \cdot 6^t - 1, & \text{if the last nonzero digit of } s \text{ is odd;} \end{cases}$$

and

$$d(s) = \begin{cases} 6^{t+1}, & \text{if the last nonzero digit of } s \text{ is 2, 3, or 4;} \\ 3 \cdot 6^t, & \text{if the last nonzero digit of } s \text{ is 1 or 5.} \end{cases}$$

We now claim that for all $s \geq 1$ and $j \geq 0$, we have

$$b(d(s)j + c(s)) \neq b(d(s)j + c(s) + 6s). \quad (5)$$

Since $c(s) \leq d(s) \leq 3s = r$, (5) also provides a solution to (4).

To verify (5), we assume that s ends in exactly t 0's. Then there are 30 cases to consider, based on the last nonzero digit s' of s and the last digit j' of j (digit in the base-6 representation, of course). The claim can now be verified by a rather tedious examination of the 30 cases. We provide the details for three typical cases:

$s' = 1, j' = 0$: The base-6 expansion of $d(s)j + c(s)$ ends with $l25^t$ for some $l \in \{0, 3\}$, while the base-6 expansion of $d(s)j + c(s) + 6s$ ends with $(l+1)25^t$, so by (3) we see that $b(d(s)j + c(s))$ and $b(d(s)j + c(s) + 6s)$ are of different parity.

$s' = 1, j' = 3$: The base-6 expansion of $d(s)j + c(s)$ ends with 15^{t+1} or 45^{t+1} , while the base-6 expansion of $d(s)j + c(s) + 6s$ ends with 25^{t+1} or 5^{t+2} . From (3) we get $b(d(s)j + c(s)) = 2t + 6$, while $b(d(s)j + c(s) + 6s) \geq 2t + 7$.

$s' = 1, j' = 5$: The base-6 expansion of $d(s)j + c(s)$ ends with $l5^{t+1}$ for some $l \in \{2, 5\}$, while the base-6 expansion of $d(s)j + c(s) + 6s$ ends with $((l + 1) \bmod 6)5^{t+1}$. In either case, from (3) we have $b(d(s)j + c(s)) \geq 2t + 7$, while $b(d(s)j + c(s) + 6s) = 2t + 5$.

Below we summarize Equation (5) for each of the 30 cases. In 22 cases, both sides of (5) can be evaluated symbolically as piecewise expressions in $2t + k$. For the remaining 8 cases we bound one side of the inequality above the other.

$(s', j') \in \{(1, 1), (3, 0), (3, 2), (3, 4), (4, 0), (4, 3)\}$:

$$2t + 5 \neq 2t + 6$$

$(s', j') \in \{(2, 1), (2, 4), (3, 1), (3, 3), (3, 5), (5, 3)\}$:

$$2t + 6 \neq 2t + 5$$

$(s', j') \in \{(1, 0), (5, 0)\}$:

$$\begin{cases} 2t + 5 & \text{if } j \equiv 0 \pmod{12} \\ 2t + 6 & \text{if } j \equiv 6 \pmod{12} \end{cases} \neq \begin{cases} 2t + 6 & \text{if } j \equiv 0 \pmod{12} \\ 2t + 5 & \text{if } j \equiv 6 \pmod{12} \end{cases}$$

$(s', j') \in \{(1, 2), (5, 2)\}$:

$$\begin{cases} 2t + 6 & \text{if } j \equiv 2 \pmod{12} \\ 2t + 5 & \text{if } j \equiv 8 \pmod{12} \end{cases} \neq \begin{cases} 2t + 5 & \text{if } j \equiv 2 \pmod{12} \\ 2t + 6 & \text{if } j \equiv 8 \pmod{12} \end{cases}$$

$(s', j') \in \{(1, 4), (5, 4)\}$:

$$\begin{cases} 2t + 5 & \text{if } j \equiv 4 \pmod{12} \\ 2t + 6 & \text{if } j \equiv 10 \pmod{12} \end{cases} \neq \begin{cases} 2t + 6 & \text{if } j \equiv 4 \pmod{12} \\ 2t + 5 & \text{if } j \equiv 10 \pmod{12} \end{cases}$$

$(s', j') = (2, 2)$:

$$\begin{cases} 2t + 7 & \text{if } j \equiv 2 \pmod{12} \\ 2t + 8 & \text{if } j \equiv 8 \pmod{12} \end{cases} \neq 2t + 6$$

$(s', j') = (4, 2)$:

$$\begin{cases} 2t + 7 & \text{if } j \equiv 2 \pmod{12} \\ 2t + 8 & \text{if } j \equiv 8 \pmod{12} \end{cases} \neq 2t + 5$$

$(s', j') = (2, 0)$:

$$2t + 5 \neq \begin{cases} 2t + 7 & \text{if } (j \equiv 0 \pmod{12} \text{ and } s/6^t \equiv 2 \pmod{12}) \\ & \text{or } (j \equiv 6 \pmod{12} \text{ and } s/6^t \equiv 8 \pmod{12}) \\ 2t + 8 & \text{if } (j \equiv 0 \pmod{12} \text{ and } s/6^t \equiv 8 \pmod{12}) \\ & \text{or } (j \equiv 6 \pmod{12} \text{ and } s/6^t \equiv 2 \pmod{12}) \end{cases}$$

$(s', j') = (4, 4)$:

$$2t + 6 \neq \begin{cases} 2t + 8 & \text{if } (j \equiv 4 \pmod{12} \text{ and } s/6^t \equiv 4 \pmod{12}) \\ & \text{or } (j \equiv 10 \pmod{12} \text{ and } s/6^t \equiv 10 \pmod{12}) \\ 2t + 7 & \text{if } (j \equiv 4 \pmod{12} \text{ and } s/6^t \equiv 10 \pmod{12}) \\ & \text{or } (j \equiv 10 \pmod{12} \text{ and } s/6^t \equiv 4 \pmod{12}) \end{cases}$$

$(s', j') \in \{(1, 3), (4, 1)\}$:

$$b(d(s)j + c(s)) = 2t + 6 < 2t + 7 \leq b(d(s)j + c(s) + 6s)$$

$(s', j') \in \{(2, 3), (5, 1)\}$:

$$b(d(s)j + c(s)) = 2t + 5 < 2t + 7 \leq b(d(s)j + c(s) + 6s)$$

$(s', j') \in \{(1, 5), (4, 5)\}$:

$$b(d(s)j + c(s)) \geq 2t + 7 > 2t + 5 = b(d(s)j + c(s) + 6s)$$

$(s', j') \in \{(2, 5), (5, 5)\}$:

$$b(d(s)j + c(s)) \geq 2t + 7 > 2t + 6 = b(d(s)j + c(s) + 6s)$$

The proof that \mathbf{v} is $\frac{3}{2}$ -power free is now complete.

4 \mathbf{v} is the lexicographically least $\frac{3}{2}$ -power-free sequence

In this section we complete our characterization of $\mathbf{w}_{3/2}$ by showing that a $\frac{3}{2}$ -power is formed whenever any symbol of \mathbf{v} , other than 0, is decremented.

The symbol 0 appears in positions congruent to 0, 3, and 6 (mod 10), so it suffices to examine other positions.

If the symbol is in a position congruent to 1, 5, or 8 (mod 10), then it is a 1, and it follows 0, 02, 03, or 04. Decrementing this to 0 then produces the square 00 or one of the $\frac{3}{2}$ -powers 020, 030, or 040.

If the symbol is in a position congruent to 2 or 7 (mod 10), then it is a 2, and it follows 01 or 10. Decrementing to 0 produces 010 or 00, while decrementing to 1 produces 11 or 101.

If the symbol is in a position congruent to 4 (mod 10), then it is either 3 or 4, and it follows 0120. Decrementing to 0 produces 00; decrementing to 1 produces the $5/3$ -power 01201; and decrementing to 2 produces 202. If the symbol is 4, there is also the possibility of decrementing to 3, and it occurs at a position ≥ 14 . If the 4 is decremented to 3, then the immediately preceding 15 symbols (including the 3) form a $\frac{3}{2}$ -power.

The last case is that the symbol is in a position congruent to 9 (mod 10). Then this position is $10n+9$ for some n , and the symbol is $a(10n+9) = b(n)$. It therefore follows 1021. If it is decremented to 0, this produces 10210. If it is decremented to 1, it produces 11. If it is decremented to 2, it produces 212. We now have to handle the possibility of decrementing to m , for some m with $3 \leq m < b(n)$.

If $n \equiv 0 \pmod{3}$, then $b(n) = 3$, so there are no other possibilities to consider. If $n \equiv 1 \pmod{3}$, then $b(n) = 4$, so we also have to consider the possibility of decrementing to 3. In this case, the immediately preceding 15 symbols (including the 3) form a $\frac{3}{2}$ -power.

It remains to consider the case when $n \equiv 2 \pmod{3}$. Here $b(n)$ is at least 5. If we decrement $b(n)$ to 3, then the immediately preceding 30 symbols form a $\frac{3}{2}$ -power, while if we decrement to 4, then the immediately preceding 15 symbols form a $\frac{3}{2}$ -power.

For all other cases, we replace $b(n)$ by $m \geq 5$. We claim this gives the $\frac{3}{2}$ -power xyx in \mathbf{v} , where x is of length ℓ_m , in the preceding $3\ell_m$ symbols, where

$$\ell_m = \begin{cases} 30 \cdot 6^{\frac{m}{2}-3}, & \text{if } b(n) \text{ is odd and } m \text{ is even;} \\ 60 \cdot 6^{\frac{m+1}{2}-3}, & \text{if } b(n) \text{ is odd and } m \text{ is odd;} \\ 30 \cdot 6^{\frac{m}{2}-3}, & \text{if } b(n) \text{ is even and } m \text{ is even;} \\ 60 \cdot 6^{\frac{m+1}{2}-3}, & \text{if } b(n) \text{ is even, } m \text{ is odd, and } m \neq b(n) - 1; \\ 30 \cdot 6^{\frac{m+1}{2}-3}, & \text{if } b(n) \text{ is even and } m = b(n) - 1. \end{cases}$$

To see this, note that from (1) it is enough to show that

$$b(n+1-3\ell_m/10) \cdots b(n-\ell_m/5-1) = b(n+1-\ell_m/10) \cdots b(n-1) \quad (6)$$

and $b(n-\ell_m/5) = m$. This can be done by a tedious examination of each case in (3). We give here one representative case, leaving the rest to the reader.

Suppose the base-6 representation of n ends with 35^t for $t \geq 1$ and so $b(n) = 2t + 3$.

Case 1: $m > 6$ is even. We have $\ell_m/10 = 3 \cdot 6^{\frac{m}{2}-3}$. The base-6 expansion of $n+1-\ell_m/10$ is $35^{t-i-1}30^i$, while the base-6 expansion of $n-1$ is $35^{t-1}4$, where $1 \leq i \leq t-2$ and $m = 2i+6$.

On the other hand, the base-6 expansion of $n+1-3\ell_m/10$ is $35^{t-i-1}430^i$, while the base-6 expansion of $n-\ell_m/5-1$ is $35^{t-i-1}45^i4$. Using (3) we see that b takes the same values on these intervals. On the other hand, b takes the value $m = 2i+6$ at $35^{t-i-1}45^i5$. Thus changing $b(n)$ to m forms a $\frac{3}{2}$ -power.

Case 2: $m = 6$. Then $\ell_m/10 = 3$. It is now easy to see that $b(n-5)b(n-4) = b(n-2)b(n-1) = 34$, while $b(n-3) = 6$. Thus changing $b(n)$ to m forms a $\frac{3}{2}$ -power.

Case 3: m is odd. We have $\ell_m/10 = 6^{\frac{m+1}{2}-3}$. The base-6 expansion of $n + 1 - \ell_m/10$ is $35^{t-i-1}0^{i+1}$, while the base-6 expansion of $n - 1$ is $35^{t-1}4$, where $1 \leq i \leq t-2$ and $m = 2i+5$.

On the other hand, the base-6 expansion of $n + 1 - 3\ell_m/10$ is $35^{t-i-2}30^{i+1}$, while the base-6 expansion of $n - \ell_m/5 - 1$ is $35^{t-i-2}35^i4$. Using (3) we see that b takes the same values on these intervals. On the other hand, b takes the value $m = 2i + 5$ at $35^{t-i-2}35^{i+1}$. Thus changing $b(n)$ to m forms a $\frac{3}{2}$ -power.

5 Morphism description

The sequence $\mathbf{w}_{3/2}$ can also be generated as follows. Consider the morphisms φ and τ on the alphabet $\{0, \bar{0}, 1, \bar{1}, \dots\}$ defined by

$$\begin{aligned}\varphi(n) &= 3\bar{3}4\bar{4}3(\overline{n+2}) \\ \varphi(\bar{n}) &= 4\bar{3}3\bar{4}4(\overline{n+2}) \\ \tau(n) &= 0120n \\ \tau(\bar{n}) &= 1021n.\end{aligned}$$

Equation (2) is equivalent to the statement that the sequence $3\bar{b}(0)4\bar{b}(1)3\bar{b}(2)4\bar{b}(3)\dots$ is a fixed point of φ ; hence this sequence is $\varphi^\omega(3)$. It follows that $\mathbf{w}_{3/2} = \tau(\varphi^\omega(3))$.

6 Avoiding only $\frac{3}{2}$ -powers

In this section we consider a variant of $\mathbf{w}_{3/2}$. Whereas $\mathbf{w}_{3/2}$ is the lexicographically least word over \mathbb{N} not containing any α -power for $\alpha \geq 3/2$, we let $\mathbf{x}_{3/2}$ be the lexicographically least word over \mathbb{N} not containing any (exact) $\frac{3}{2}$ -power. Here are the first 144 terms of $\mathbf{x}_{3/2}$:

$\mathbf{x}_{3/2}[12i+j]$	$j=0$	1	2	3	4	5	6	7	8	9	10	11
$i=0$	0	0	1	1	0	2	1	0	0	1	1	2
1	0	0	1	1	0	3	1	0	0	1	1	3
2	0	0	1	1	0	2	1	0	0	1	1	4
3	0	0	1	1	0	3	1	0	0	1	1	2
4	0	0	1	1	0	2	1	0	0	1	1	3
5	0	0	1	1	0	3	1	0	0	1	1	4
6	0	0	1	1	0	2	1	0	0	1	1	2
7	0	0	1	1	0	3	1	0	0	1	1	3
8	0	0	1	1	0	2	1	0	0	1	1	5
9	0	0	1	1	0	3	1	0	0	1	1	2
10	0	0	1	1	0	2	1	0	0	1	1	3
11	0	0	1	1	0	3	1	0	0	1	1	5

Unlike $\mathbf{w}_{3/2}$, the word $\mathbf{x}_{3/2}$ contains squares, for example. However, the underlying structures of $\mathbf{w}_{3/2}$ and $\mathbf{x}_{3/2}$ are the same. Namely, let $\mathbf{y} = f(0)f(1)f(2)\cdots$, where

$$\begin{aligned} f(12n) &= f(12n+1) = f(12n+4) = f(12n+7) = f(12n+8) = 0, \quad n \geq 0; \\ f(12n+2) &= f(10n+3) = f(12n+6) = f(12n+9) = f(12n+10) = 1, \quad n \geq 0; \\ f(12n+5) &= \begin{cases} 2, & \text{if } n \equiv 0 \pmod{2}; \\ 3, & \text{if } n \equiv 1 \pmod{2}; \end{cases} \\ f(12n+11) &= \begin{cases} 2, & \text{if } n \equiv 0 \pmod{3}; \\ 3, & \text{if } n \equiv 1 \pmod{3}; \\ f(2n+1)+2, & \text{if } n \equiv 2 \pmod{3}. \end{cases} \end{aligned}$$

Note that $f(12n+11)+1 = b(n) = a(10n+9)$ for $n \geq 0$. We show that $\mathbf{y} = \mathbf{x}_{3/2}$. In particular, it follows that $\mathbf{x}_{3/2}$ is 6-regular.

First we show that \mathbf{y} is $\frac{3}{2}$ -power-free. As in Section 3, for each $n \geq 0$ and $k \geq 0$ we exhibit an index i , $0 \leq i < k$, such that $f(n+i) \neq f(n+2k+i)$.

Case 1: $k \equiv 1, 5, 7, 11 \pmod{12}$. We have $f(n) \neq f(n+2k)$ for all n , so let $i = 0$.

Case 2: $k \equiv 2, 4, 8, 10 \pmod{12}$. If n is even, let $i = 1$; if n is odd, let $i = 0$.

Case 3: $k \equiv 3, 9 \pmod{12}$. If n is even, let $i = 0$; if n is odd, let $i = 1$.

Case 4: $k \equiv 6 \pmod{12}$. If $k = 6$ then we can choose i such that either $n+i \equiv 5 \pmod{12}$ or $n+i \equiv 11 \pmod{12}$; in either case we have $f(n+i) \neq f(n+2k+i)$. Otherwise $k \geq 18$, and we can choose i such that $n+i \equiv 5 \pmod{12}$, and hence $f(n+i) \neq f(n+2k+i)$.

Case 5: $k \equiv 0 \pmod{12}$. Since $f(12n+11)+1 = b(n)$, this case follows immediately from Case 2 of Section 3.

Therefore \mathbf{y} is $\frac{3}{2}$ -power-free. Now we show that decrementing any nonzero symbol of \mathbf{y} introduces a $\frac{3}{2}$ -power.

The symbol in each position congruent to 0, 1, 4, 7, or 8 modulo 12 is 0, which cannot be decremented.

The symbol in each position congruent to 2, 3, 6, 9, or 10 modulo 12 is 1, and it follows 00, 01, 02, or 03. In each case, decrementing the 1 to 0 produces a $\frac{3}{2}$ -power.

The symbol in each position congruent to 5 modulo 12 is either 2 or 3, and it follows 00110. Decrementing to 0 produces 001100, and decrementing to 1 produces 101. If the symbol is 3, then decrementing to 2 produces a $\frac{3}{2}$ -power in the preceding 18 symbols (including the 2).

The symbol in each position congruent to 11 modulo 12 is at least 2. Decrementing to 0 produces 100110. Decrementing to 1 produces 111. If $n \equiv 0 \pmod{3}$, then $f(12n+11) = 2$, so there are no other possibilities to consider. If $n \equiv 1 \pmod{3}$, then $f(12n+11) = 3$, and decrementing to 2 produces a $\frac{3}{2}$ -power in the preceding 18 symbols.

If $n \equiv 2 \pmod{3}$, then $f(12n + 11) \geq 4$. Decrementing to 2 or 3 produces a $\frac{3}{2}$ -power in the preceding 36 or 18 symbols, respectively. It follows from (6) that if $f(12n + 11) > m \geq 4$ then decrementing $f(12n + 11) = b(n) - 1$ to m produces a $\frac{3}{2}$ -power in the preceding $\frac{18}{5}\ell_{m+1}$ symbols.

We have shown that \mathbf{y} is the lexicographically least $\frac{3}{2}$ -power-free word over \mathbb{N} ; hence $\mathbf{x}_{3/2} = \mathbf{y}$. The word $\mathbf{x}_{3/2}$ is generated by the morphism φ of the previous section as $\mathbf{x}_{3/2} = v(\varphi^\omega(3))$, where v is the morphism defined by

$$\begin{aligned} v(n) &= 00110(n-1) \\ v(\bar{n}) &= 10011(n-1). \end{aligned}$$

As already mentioned, $\mathbf{x}_{3/2}$ contains squares. However, the only squares in $\mathbf{x}_{3/2}$ are 00 and 11, as we now show. Suppose xx is a factor of $\mathbf{x}_{3/2}$ for some nonempty word x . The length of x cannot be even, because otherwise xx contains a $\frac{3}{2}$ -power. One checks from the definition of $f(n)$ that the length of x cannot be 3. Similarly, $f(n) \cdots f(n+4) \neq f(n+k) \cdots f(n+4+k)$ for all $n \geq 0$ and odd k , which precludes a factor xx where the length of x is odd and at least 5.

Hence the only square factors of $\mathbf{x}_{3/2}$ are 00 and 11. Since 000 and 111 are not factors of $\mathbf{x}_{3/2}$, it follows that $\mathbf{x}_{3/2}$ is overlap-free.

7 Remarks

A previous paper discussed the structure of the lexicographically least word over \mathbb{N} avoiding n 'th powers or overlaps [4]. In this paper, we discussed the structure of such a word avoiding $3/2$ -powers. It remains to gain a deeper understanding of the lexicographically least word avoiding α -powers for arbitrary α . Computer experiments strongly suggest that a similar structure exists for $4/3$ -powers. But it is still not known whether the lexicographically least word avoiding $5/2$ -powers uses only finitely many distinct letters.

References

- [1] J.-P. Allouche and J. Shallit. The ring of k -regular sequences. *Theoret. Comput. Sci.* **98** (1992), 163–197.
- [2] J.-P. Allouche and J. Shallit. The ring of k -regular sequences, II. *Theoret. Comput. Sci.* **307** (2003), 3–29.
- [3] F. Dejean. Sur un théorème de Thue. *J. Combin. Theory Ser. A* **13** (1972), 90–99.
- [4] M. Guay-Paquet and J. Shallit. Avoiding squares and overlaps over the natural numbers. *Discrete Math.* **309** (2009), 6245–6254.