Extremal words avoiding partial repetitions

Eric Rowland



Joint work with Jeff Shallit and Lara Pudwell.

2016 January 14

Squares on a binary alphabet



Axel Thue (1863-1922)

A square is a nonempty word of the form *ww*.

Are squares are avoidable on a binary alphabet? Are there arbitrarily long square-free words on $\{0, 1\}$?

010X

Are squares avoidable on $\{0, 1, 2\}$?

01020120210120102012021020102101201020120210...

Theorem (Thue 1906)

There exist arbitrarily long square-free words on 3 letters.

The backtracking algorithm builds the lexicographically least word.

Open problem (Allouche–Shallit, Automatic Sequences §1.10)

Characterize the lex. least square-free word over $\{0, 1, 2\}$.

An overlap is a word of the form *cxcxc*, where *c* is a letter; equivalently, *wwc*, where *c* is the first letter of *w*.

Theorem (Thue 1912)

There exist arbitrarily long overlap-free words on 2 letters.

Let $\varphi(0) = 01$ and $\varphi(1) = 10$. The Thue–Morse word is overlap-free:

 $\varphi^{\infty}(0) = 01101001100101101001011001011001001\cdots$

It follows that cubes are avoidable on a binary alphabet.

Thue–Morse word

• Curious products (Robbins 1958):

$$\prod_{n\geq 0} (n+1)^{(-1)^{t_n}} = \frac{1\cdot 4\cdot 6\cdot 7\cdot 10\cdot 11\cdot 13\cdot 16\cdots}{2\cdot 3\cdot 5\cdot 8\cdot 9\cdot 12\cdot 14\cdot 15\cdots} = \frac{1}{\sqrt{2}}$$

• Multigrades (Prouhet 1851):

$$\begin{array}{l} 0^{0}+3^{0}+5^{0}+6^{0}=1^{0}+2^{0}+4^{0}+7^{0}=4\\ 0^{1}+3^{1}+5^{1}+6^{1}=1^{1}+2^{1}+4^{1}+7^{1}=14\\ 0^{2}+3^{2}+5^{2}+6^{2}=1^{2}+2^{2}+4^{2}+7^{2}=70 \end{array}$$

Theorem (Allouche–Currie–Shallit 1998)

The lex. least overlap-free word on $\{0,1\}$ is $001001\varphi^{\infty}(1)$.

Eric Rowland

Infinite alphabet

On an infinite alphabet, the backtracking algorithm doesn't backtrack.

Are squares avoidable on $\mathbb{Z}_{>0} = \{0, 1, 2, ...\}$? Yes.

01020103010201040102010301020105 ····

Theorem (Guay-Paquet–Shallit 2009)

Let $\varphi(n) = 0(n+1)$. The lex. least square-free word on $\mathbb{Z}_{\geq 0}$ is $\varphi^{\infty}(0)$.

The *n*th letter of each word is $\nu_2(n + 1)$. The *k*-adic valuation $\nu_k(n)$ is the exponent of the highest power of *k* dividing *n*.

Integer powers

More generally, let $a \ge 2$. Let $\varphi(n) = 0^{a-1}(n+1)$. The lex. least *a*-power-free word on $\mathbb{Z}_{\ge 0}$ is $\varphi^{\infty}(0) = (\nu_a(n+1))_{n\ge 0}$.

 $\bm{w}_5 = 000010000100001000020000100001\cdots$



The letters of w_5 satisfy a recurrence with constant coefficients.

k-regular sequences

An integer sequence $s(n)_{n\geq 0}$ is *k*-regular if the set

$$\{s(k^e n + r) : e \geq 0 \text{ and } 0 \leq r \leq k^e - 1\}$$

is contained in a finite-dimensional \mathbb{Q} -vector space.

For example, let $s(n) = |\{0 \le m \le n : \binom{n}{m} \not\equiv 0 \mod 8\}|$.



$$s(2n+1) = 2s(n)$$

$$s(4n+0) = s(2n)$$

$$s(8n+2) = -2s(n) + 2s(2n) + s(4n+2)$$

$$s(8n+6) = 2s(4n+2)$$

A 2-regular sequence is also 4-regular, and vice versa.

k and *l* are multiplicatively dependent if there exist positive integers *s* and *t* such that $k^s = l^t$.

For example, 4 and 8 (since $4^3 = 8^2$).

Theorem (Bell 2006)

If k and I are multiplicatively independent and $s(n)_{n\geq 0}$ is both k-regular and I-regular, then $\sum_{n\geq 0} s(n)x^n$ is the power series of a rational function whose poles are roots of unity.

Fractional powers

011101 = $(0111)^{3/2}$ is a $\frac{3}{2}$ -power. If |x| = |y| = |z|, then $xyzxyzx = (xyz)^{7/3}$ is a $\frac{7}{3}$ -power.

Definition

A word w is an $\frac{a}{b}$ -power if

$$w = v^e x$$

where $e \ge 0$ is an integer, x is a prefix of v, and $\frac{|w|}{|v|} = \frac{a}{b}$.

Notation

For $\frac{a}{b} > 1$, let $\mathbf{w}_{a/b}$ be the lex. least $\frac{a}{b}$ -power-free word on $\mathbb{Z}_{\geq 0}$.

We assume gcd(a, b) = 1 from now on.

Eric Rowland

Avoiding 3/2-powers

 ${\bm w}_{3/2} = 001102100112001103100113001102100114001103\cdots$

$$\begin{split} \textbf{w}_{3/2} &= 001102 \\ 100112 \\ 001103 \\ 100113 \\ 001102 \\ 100114 \\ 001103 \\ 100112 \\ . \end{split}$$

Theorem (Rowland–Shallit 2012)

The sequence of letters in $\mathbf{w}_{3/2}$ is 6-regular.

.

Why 6?

Eric Rowland

 $\textbf{w}_{5/2} = 000010000100001000020000100001 \cdots = \textbf{w}_{5}$

Theorem

If $\frac{a}{b} \geq 2$, then $\mathbf{w}_{a/b} = \mathbf{w}_a$.

Proof (one direction).

The *a*-power $v^a = (v^b)^{a/b}$ is also an $\frac{a}{b}$ -power. So $\mathbf{w}_{a/b}$ is *a*-power-free. Thus $\mathbf{w}_a \leq \mathbf{w}_{a/b}$ lexicographically.

Therefore it suffices to consider $1 < \frac{a}{b} < 2$.

w_{5/3} wrapped into 100 columns

 ${\bm w}_{5/3} = 00001010000101000010100001020000101\cdots$



w_{5/3} wrapped into 7 columns

 $\mathbf{w}_{5/3} = 000010100001010000101000010100001020000101\cdots$



Theorem

 $\mathbf{w}_{5/3} = \varphi^{\infty}(0)$, where $\varphi(n) = 000010(n+1)$ is a 7-uniform morphism.

Eric Rowland

Extremal words avoiding partial repetitions

$\mathbf{w}_{7/4}$ wrapped into 100 columns



w7/4 wrapped into 50847 columns



Theorem

 $\mathbf{w}_{7/4} = \varphi^{\infty}(\mathbf{0})$ for some 50847-uniform morphism $\varphi(n) = u(n+2)$.

w_{8/5} wrapped into 733 columns

$\bm{w}_{8/5} = 0000001001000001001000000100110000000100\cdots$

Theorem

$\mathbf{w}_{8/5} = \varphi^{\infty}(\mathbf{0})$ for the 733-uniform morphism

w_{6/5} wrapped into 1001 columns

${\bm w}_{6/5} = 000001111102020201011101000202120210110010\cdots$



There is a transient region.

Introduce a new letter 0', and let $\tau(0') = 0$ and $\tau(n) = n$ for $n \in \mathbb{Z}_{\geq 0}$.

Theorem

There exist words u, v of lengths |u| = 1001 - 1 and |v| = 29949 such that $\mathbf{w}_{6/5} = \tau(\varphi^{\infty}(0'))$, where

$$\varphi(n) = \begin{cases} v \,\varphi(0) & \text{if } n = 0' \\ u \,(n+3) & \text{if } n \ge 0. \end{cases}$$

$w_{5/4}$ wrapped into 144 columns

$\bm{w}_{5/4} = 000011110202101001011212000013110102101302\cdots$



We don't know the structure of $\mathbf{w}_{5/4}$.

Catalogue

For many words $\mathbf{w}_{a/b}$, there is a related *k*-uniform morphism. A *k*-uniform morphism generates a *k*-regular sequence.

$$\frac{a}{b} = \frac{3}{2} \rightarrow k = 6$$

$$\frac{a}{b} = \frac{5}{3} \rightarrow k = 7$$

$$\frac{a}{b} = \frac{7}{4} \rightarrow k = 50847$$

$$\frac{a}{b} = \frac{8}{5} \rightarrow k = 733$$

$$\frac{a}{b} = \frac{6}{5} \rightarrow k = 1001$$

$$\frac{a}{b} = \frac{5}{4} \rightarrow k = ?$$

Question

Is every $\mathbf{w}_{a/b}$ k-regular for some k? How is k related to $\frac{a}{b}$?

Eric Rowland

Extremal words avoiding partial repetitions

A family related to $\mathbf{w}_{5/3}$



Theorem

Let $\frac{5}{3} \leq \frac{a}{b} < 2$ and b odd. Let φ be the (2a - b)-uniform morphism

$$\varphi(n) = 0^{a-1} 1 0^{a-b-1} (n+1)$$

for all $n \in \mathbb{Z}_{\geq 0}$. Then $\mathbf{w}_{a/b} = \varphi^{\infty}(0)$.

- Show that φ preserves $\frac{a}{b}$ -power-freeness. That is, if w is $\frac{a}{b}$ -power-free then $\varphi(w)$ is $\frac{a}{b}$ -power-free. Since 0 is $\frac{a}{b}$ -power-free, it follows that $\varphi^{\infty}(0)$ is $\frac{a}{b}$ -power-free.
- 2 Show that decrementing any letter in $\mathbf{w}_{a/b}$ introduces an $\frac{a}{b}$ -power.

Other intervals

We have 30 symbolic $\frac{a}{b}$ -power-free morphisms, found experimentally.

Theorem

Let $\frac{3}{2} < \frac{a}{b} < \frac{5}{3}$ and gcd(*b*, 5) = 1. The (5*a* - 4*b*)-uniform morphism $\varphi(n) = 0^{a-1} 1 0^{a-b-1} 1 0^{2a-2b-1} 1 0^{a-b-1} (n+1)$

is $\frac{a}{b}$ -power-free.

Theorem

Let $\frac{6}{5} < \frac{a}{b} < \frac{5}{4}$ and $\frac{a}{b} \notin \{\frac{11}{9}, \frac{17}{14}\}$. The a-uniform morphism $\varphi(n) = 0^{6a-7b-1} 1 0^{-3a+4b-1} 1 0^{-8a+10b-1} 1 0^{6a-7b-1} (n+1)$

is $\frac{a}{b}$ -power-free.

Theorem 50. Let a, b be relatively prime positive integers such that $\frac{10}{9} < \frac{a}{b} < \frac{29}{26}$ and $\frac{a}{b} \neq \frac{39}{35}$ and gcd(b, 67) = 1. Then the (67a - 30b)-uniform morphism

 $\varphi(n) = 0^{-7a+8b-1} 1 0^{10a-11b-1} 1 0^{10a-11b-1} 1 0^{a-b-1} 1 0^{a-b-1} 1 0^{-26a+29b-1} 1 0^{28a-31b-1} 1 0^{2a-2b-1} 1$ $0^{a-b-1} 1 0^{-25a+28b-1} 1 0^{10a-11b-1} 1 0^{2a-2b-1} 1 0^{a-b-1} 1 0^{10a-11b-1} 1 0^{3a-3b-1} 1 0^{3a-3b-1}$ $0^{-25a+28b-1}$ $10^{10a-11b-1}$ $10^{3a-3b-1}$ $10^{10a-11b-1}$ $10^{-8a+9b-1}$ 10^{a-b-1} $10^{10a-11b-1}$ $10^{10a-11b-1}$ $0^{-25a+28b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{-8a+9b-1}$ 1 0^{a-b-1} 1 $0^{10a-11b-1}$ 1 $0^{2a-2b-1}$ 2 0^{a-b-1} 1 $0^{10a-11b-1}$ 10^{-25a+28b-1} 10^{2a-2b-1} 20^{a-b-1} 10^{10a-11b-1} 10^{3a-3b-1} 10^{10a-11b-1} 1 $0^{-25a+28b-1}$ 1 $0^{3a-3b-1}$ 1 $0^{10a-11b-1}$ 1 0^{a-b-1} 1 0^{a-b-1} 2 0^{a-b-1} 1 $0^{10a-11b-1}$ 1 $a^{-25a+28b-1}1a^{a-b-1}1a^{a-b-1}2a^{a-b-1}1a^{2a-2b-1}1a^{11a-12b-1}1a^{10a-11b-1}1$ $0^{2a-2b-1}$ 1 $0^{-24a+27b-1}$ 1 $0^{2a-2b-1}$ 1 0^{a-b-1} 1 $0^{10a-11b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{2a-2b-1}$ 1 0^{a-b-1} 1 $0^{-25a+28b-1}$ 1 $0^{27a-30b-1}$ 1 $0^{-24a+27b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{-8a+9b-1}$ 1 $a^{11a-12b-1}$, $a^{2a-2b-1}$, a^{a-b-1} , $a^{-25a+28b-1}$, $a^{10a-11b-1}$, $a^{2a-2b-1}$, a^{a-b-1} , $a^{2a-2b-1}$, a $0^{10a-11b-1}$ $10^{-25a+28b-1}$ $10^{28a-31b-1}$ $10^{-25a+28b-1}$ $10^{10a-11b-1}$ $10^{10a-11b-1}$ $10^{-7a+8b-1}$ $10^{-7a+8b-1}$ $0^{10a-11b-1}$ $10^{-8a+9b-1}$ 10^{a-b-1} $10^{10a-11b-1}$ $10^{-25a+28b-1}$ $10^{10a-11b-1}$ $10^{-8a+9b-1}$ 10⁴-b-110¹⁰s-11b-110³a-3b-110¹⁰a-11b-110⁻²⁵s+28b-110³a-3b-110¹⁰a-11b-11 $0^{a-b-1} + 0^{9a-10b-1} + 0^{-7a+8b-1} + 0^{10a-11b-1} + 0^{-25a+28b-1} + 0^{a-b-1} + 0^{9a-10b-1} + 0^{10a-10b-1} + 0^{10a$ $0^{-7a+8b-1}$ 1 $0^{2a-2b-1}$ 1 0^{a-b-1} 1 $0^{10a-11b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{2a-2b-1}$ 1 0^{a-b-1} 1 $0^{-25a+28b-1}$ 1 $0^{3a-3b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{3a-3b-1}$ 1 $0^{-25a+28b-1}$ 1 $0^{27a-30b-1}$ 1 $0^{a-b-1} + 0^{-25a+28b-1} + 0^{10a-11b-1} + 0^{10a-11b-1} + 0^{-8a+9b-1} + 0^{a-b-1} + 0^{10a-11b-1} + 0^{1$ $0^{2a-2b-1} 2 0^{a-b-1} 1 0^{-25a+28b-1} 1 0^{10a-11b-1} 1 0^{2a-2b-1} 2 0^{a-b-1} 1 0^{10a-11b-1} 1$ $0^{3a-3b-1}$, $0^{-25a+28b-1}$, $0^{10a-11b-1}$, $0^{3a-3b-1}$, $0^{10a-11b-1}$, 0^{a-b-1} , 0^{a-b-1} , 0^{a-b-1} 1 $0^{-25a+28b-1}$ 1 $0^{10a-11b-1}$ 1 0^{a-b-1} 1 0^{a-b-1} 2 0^{a-b-1} 1 $0^{2a-2b-1}$ 1 $0^{11a-12b-1} 10^{-25a+28b-1} 10^{2a-2b-1} 10^{11a-12b-1} 10^{2a-2b-1} 10^{a-b-1} 10^{10a-11b-1} 1$ $0^{-25a+28b-1}$ 10^{2a-2b-1} 10^{a-b-1} 10^{10a-11b-1} 10^{-8a+9b-1} 10^{11a-12b-1} 10^{10a-11b-1} 1 $0^{-25a+28b-1} + 0^{27a-30b-1} + 0^{-24a+27b-1} + 0^{2a-2b-1} + 0^{a-b-1} + 0^{10a-11b-1} +$ $0^{2a-2b-1}$ 10^{a-b-1} $10^{-25a+28b-1}$ $10^{10a-11b-1}$ $10^{-7a+8b-1}$ $10^{10a-11b-1}$ $10^{10a-11b-1}$ $10^{10a-11b-1}$ $10^{10a-11b-1}$ $0^{-25a+28b-1} \, {}_{1} \, 0^{28a-31b-1} \, {}_{1} \, 0^{-25a+28b-1} \, {}_{1} \, 0^{27a-30b-1} \, {}_{1} \, 0^{a-b-1} \, {}_{1} \, 0^{-25a+28b-1} \, {}_{1} \, 0^{10a-11b-1} \, {}_{$ $0^{10a-11b-1}$ 10^{-8a+9b-1} 10^{a-b-1} 10^{10a-11b-1} 10^{3a-3b-1} 10^{-25a+28b-1} 10^{10a-11b-1} 1 $0^{3a-3b-1}$ $10^{10a-11b-1}$ 10^{a-b-1} $10^{-26a+29b-1}$ $10^{26a-31b-1}$ $10^{-25a+28b-1}$ $10^{10a-11b-1}$ 1 $0^{a-b-1} 1 0^{9a-10b-1} 1 0^{-7a+8b-1} 1 0^{2a-2b-1} 1 0^{a-b-1} 1 0^{10a-11b-1} 1 0^{-25a+28b-1} 1$ $0^{2a-2b-1}$ 1 0^{a-b-1} 1 $0^{10a-11b-1}$ 1 $0^{3a-3b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{-25a+28b-1}$ 1 $0^{3a-3b-1}$ 1 $0^{10a-11b-1}$ $10^{-8a+9b-1}$ 10^{a-b-1} $10^{10a-11b-1}$ $10^{10a-11b-1}$ $10^{-25a+28b-1}$ $10^{27a-30b-1}$ $10^{10a-11b-1}$ $0^{a-b-1} 10^{-25a+28b-1} 10^{2a-2b-1} 20^{a-b-1} 10^{10a-11b-1} 10^{10a-11b-1} 10^{2a-2b-1} 2$ 0^{a-b-1} 1 $0^{-25a+28b-1}$ 1 $0^{3a-3b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{3a-3b-1}$ 1 $0^{-25a+28b-1}$ 1 $0^{a-b-1} 10^{a-b-1} 20^{a-b-1} 10^{10a-11b-1} 10^{10a-11b-1} 10^{a-b-1} 10^{a-b-1} 20^{a-b-1} 10^{a-b-1} 20^{a-b-1} 10^{a-b-1} 20^{a-b-1} 10^{a-b-1} 10$ $0^{a-b-1} \, {}_1 \, 0^{2a-2b-1} \, {}_1 \, 0^{-24a+27b-1} \, {}_1 \, 0^{10a-11b-1} \, {}_1 \, 0^{2a-2b-1} \, {}_1 \, 0^{11a-12b-1} \, {}_1 \, 0^{2a-2b-1} \, {}_1$ $0^{a-b-1} 10^{-25a+28b-1} 10^{10a-11b-1} 10^{2a-2b-1} 10^{a-b-1} 10^{10a-11b-1} 10^{-8a+9b-1} 10^{10a-11b-1} 10^{-8a+9b-1} 10^{10a-11b-1} 10^{-8a+9b-1} 10^{10a-11b-1} 10$ $0^{11a-12b-1}$ $10^{-25a+28b-1}$ $10^{10a-11b-1}$ $10^{-8a+9b-1}$ $10^{11a-12b-1}$ $10^{2a-2b-1}$ 10^{a-b-1} $10^{2a-2b-1}$ $0^{10a-11b-1}$ $10^{-25a+28b-1}$ $10^{2a-2b-1}$ 10^{a-b-1} $10^{10a-11b-1}$ $10^{10a-11b-1}$ $10^{-7a+8b-1}$ $10^{-7a+8b-1}$ $0^{10a-11b-1}$ 1 $0^{-25a+28b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{-7a+8b-1}$ 1 $0^{10a-11b-1}$ 1 $0^{-8a+9b-1}$ 1 0^{a-b-1} 1 $0^{10a-11b-1}$ $10^{10a-11b-1}$ $10^{-25a+28b-1}$ $10^{27a-30b-1}$ 10^{a-b-1} $10^{-25a+28b-1}$ $10^{3a-3b-1}$ $10^{10a-11b-1}$ $0^{10a-11b-1}$ $10^{10a-11b-1}$ $10^{2a-3b-1}$ $10^{-25a+28b-1}$ 10^{a-b-1} $10^{9a-10b-1}$ (n+1)

with 279 nonzero letters, locates words of length 5a - 4b and is a power-free.

Extremal words avoiding partial repetitions

Coverage of $\frac{a}{b}$ -power-free morphisms



A family with a transient



The interval $\frac{9}{7} < \frac{a}{b} < \frac{4}{3}$

Theorem

Let $\frac{9}{7} < \frac{a}{b} < \frac{4}{3}$ and gcd(b, 6) = 1. Let

$$\varphi(0') = 0'0^{a-2} \, 1 \, 0^{a-b-1} \, 1 \, 0^{a-b-1} \, 1 \varphi(0)$$

and

$$\begin{split} \varphi(n) &= 0^{a-b-1} 1 0^{2a-2b-1} 1 0^{-a+2b-1} 1 0^{2a-2b-1} 1 0^{a-b-1} 1 0^{-2a+3b-1} 1 0^{4a-5b-1} 1 \\ & 0^{-a+2b-1} 1 0^{2a-2b-1} 1 0^{a-b-1} 1 0^{-2a+3b-1} 1 0^{-2a+3b-1} 1 0^{5a-6b-1} 1 \\ & 0^{-2a+3b-1} 1 0^{4a-5b-1} 1 0^{a-b-1} 1 0^{-2a+3b-1} 1 0^{3a-3b-1} 1 0^{-2a+3b-1} 1 \\ & 0^{a-b-1} 1 0^{-3a+4b-1} 1 0^{5a-6b-1} 1 0^{2a-2b-1} 1 0^{a-b-1} 1 0^{-2a+3b-1} 1 \\ & 0^{3a-3b-1} 1 0^{-2a+3b-1} 1 0^{4a-5b-1} 1 0^{a-b-1} 1 0^{-2a+3b-1} 1 0^{2a-2b-1} 2 \\ & 0^{a-b-1} 1 0^{-2a+3b-1} 1 0^{3a-3b-1} 1 0^{-2a+3b-1} 1 0^{a-b-1} 1 0^{a-b-1} (n+2), \end{split}$$

for $n \in \mathbb{Z}_{\geq 0}$. Then $\mathbf{w}_{a/b} = \tau(\varphi^{\infty}(0'))$.

The same proof technique applies to symbolic and explicit rationals...

The letters of $\boldsymbol{w}_{7/4}$ form a 50847-regular sequence. The letters of $\boldsymbol{w}_{8/5}$ form a 733-regular sequence. The letters of $\boldsymbol{w}_{13/9}$ form a 45430-regular sequence. The letters of $\boldsymbol{w}_{17/10}$ form a 55657-regular sequence. etc.

Is there some way to make sense of them?



There exist words u, v on $\{0, 1, 2\}$ of lengths |u| = 353 - 1 and |v| = 75019 such that $\mathbf{w}_{27/23} = \tau(\varphi^{\omega}(0'))$, where

$$arphi(n) = egin{cases} v\,arphi(0) & ext{if } n=0' \ u\,(n+0) & ext{if } n\geq 0. \end{cases}$$

 $\bm{w}_{27/23}$ is also the lex least $\frac{27}{23}\text{-power-free word on }\{0,1,2\}.$