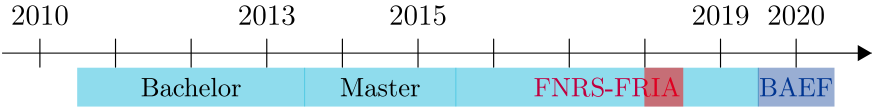
Avoiding fractional powers on the alphabet $\mathbb{N}$

Joint work with Eric Rowland (Hofstra University)

Manon Stipulanti (Hofstra University)   BAEF Fellow

Virtual New York Combinatorics Seminar
CUNY Graduate Center (NY, USA)
March 27, 2020

A little about me



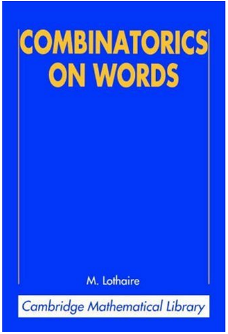
Scientific interests:

- combinatorics on words
- numeration systems
- formal language theory
- automata theory

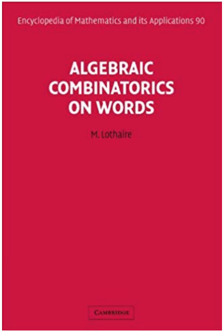
Combinatorics on words (CoW)

- Defined as the study of sequences of symbols

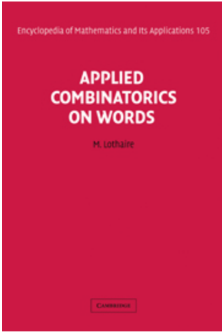
  words letters
- Lothaire's books (collective work):



1983



2002



2005

First two: overview of CoW (with two different sets of authors)
Third one: applications of CoW

Combinatorics on words is a relatively new area of discrete mathematics (± 1900).

It was initiated by the work of the Norwegian mathematician Axel Thue in 1906.

Axel Thue (1863–1922)

Topics of interest in CoW:

- regularities and patterns in words
- important types/families of words
  (e.g. automatic, regular, de Bruijn, Lyndon, Sturmian)
- equations on words
- etc.

Some definitions

- A word is a (finite or infinite) sequence of letters belonging to a (finite or infinite) set called the alphabet.

  Example: $A = \{\texttt{a}, \texttt{b}, \texttt{c}, \texttt{d}, \dots, \texttt{z}\}$
  $\texttt{repetition}, \texttt{aaaaaa}\cdots$

- A^* is the set of all finite words on A.
- The length $|w|$ of a word $w \in A^*$ is the number of letters of w.

  Example: $|\texttt{repetition}| = 10$

- ε is the empty word of length 0.
- A factor of a word $w \in A^*$ is a block of consecutive letters of w, i.e., t is a factor of w if $\exists u, v \in A^*$ s.t. $w = utv$.

  Example: `pet` is a factor of `repetition`

- A prefix (resp. suffix) of a word $w \in A^*$ is a factor starting (resp. ending) w.

  Example: `repetition`

Infinite square-free words

In his 1906 and 1912 papers, Thue studied square-free words.

- A square is a non-empty word of the form ww.

  Example: `couscous`, `murmur` are squares in English

- A word is square-free if is does not contain any squares.

  Example: `word` is square-free, `repetition` is not

More precisely, Thue was wondering about infinite square-free words.

Of course, there are no infinite square-free words on one letter $\{0\}$.

Question: Are there infinite square-free words on two letters $\{0, 1\}$?

0 1 0 ✗

So the answer is no.

We say that squares are not avoidable on two letters, or that avoiding squares cannot be done on two letters.

Thue's work on square-free words

Question: What about larger alphabets?

In 1906, Thue provided an infinite square-free word on four letters

$$03121\,01213\,01321\,01231\,01321\cdots,$$

and then in 1912, an infinite square-free word on three letters

$$012021012102\cdots.$$

Avoiding squares (Thue, 1906&1912)

There exist infinite square-free words on any alphabet with three or more letters.

Let's see how to build those two words.

More definitions

- A morphism is a map $\varphi\colon A^* \to B^*$ satisfying $\varphi(uv) = \varphi(u)\varphi(v)$ for all $u, v \in A^*$. It suffices to define φ on letters of A.
  It extends naturally to infinite words by
  $$\varphi(a_0a_1a_2\cdots) = \varphi(a_0)\varphi(a_1)\varphi(a_2)\cdots, \quad a_i \in A \;\forall\, i \geq 0.$$
  Example: $\mu\colon \{0,1\}^* \to \{0,1\}^*, 0 \mapsto 01, 1 \mapsto 10$
- Let $\varphi\colon A^* \to A^*$ s.t. $\varphi(a) = ax$, $a \in A$, $x \in A^*$, $\varphi^n(x) \neq \varepsilon \;\forall\, n \geq 0$.
  Iterating φ on a
  $$\begin{aligned}\varphi(a) &= ax \\ \varphi^2(a) &= ax\varphi(x) \\ \varphi^3(a) &= ax\varphi(x)\varphi^2(x) \\ &\vdots \\ \varphi^\omega(a) &= ax\varphi(x)\varphi^2(x)\cdots.\end{aligned}$$
  Example: The Thue–Morse word is the infinite word
  $$\mu^\omega(0) = 0\;1\;\underbrace{10}_{\mu(1)}\;\underbrace{1001}_{\mu^2(1)}\;\underbrace{10010110}_{\mu^3(1)}\;\underbrace{1001011001101001}_{\mu^4(1)}\cdots$$

Avoiding squares on four letters (Thue, 1906)

The infinite square-free word on four letters is the infinite word

$$\varphi^\omega(0) = 03121\,01213\,01321\,01231\,01321\cdots,$$

where the morphism $\varphi\colon \{0,1,2,3\}^* \to \{0,1,2,3\}^*$ is defined by $\varphi(0) = 03121$, $\varphi(1) = 01321$, $\varphi(2) = 01231$, $\varphi(3) = 01213$.

Avoiding squares on three letters (Thue, 1912)

The infinite square-free word on three letters

$$012021012102\cdots$$

is the pre-image of the Thue–Morse word $\mu^\omega(0)$ under the morphism $0 \mapsto 011$, $1 \mapsto 01$, $2 \mapsto 0$.

$$\begin{array}{ccccccccccccccc} \mu^\omega(0) = & 011 & 01 & 0 & 011 & 0 & 01 & 011 & 01 & 0 & 01 & 011 & 0 & \cdots \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ & 0 & 1 & 2 & 0 & 2 & 1 & 0 & 1 & 2 & 1 & 0 & 2 & \cdots \end{array}$$

Lexicographic order

We could add a condition on the words: lexicographic leastness.

More definitions (again):

- The alphabet A is endowed with an order $<$ on the letters.

  Example: $A = \{0, 1\}$ with $0 < 1$

- The lexicographic order $<_{\text{lex}}$ is a total order on A^* induced by the order $<$ on A:

$$x <_{\text{lex}} y \quad \Leftrightarrow \quad y = xw \text{ for some } w \in A^*$$

$$\text{or } \exists p \in A^* \text{ and } a, b \in A \text{ s.t. } \begin{cases} a < b \\ x = pau \\ y = pbv \\ \text{for some } u, v \in A^* \end{cases}$$

We write $x \leq_{\text{lex}} y$ if $x <_{\text{lex}} y$ or $x = y$.

Example: $0 <_{\text{lex}} 00 <_{\text{lex}} 011$

We study lexicographically least words on the alphabet $\mathbb{N}$ of non-negative integers avoiding some patterns.

Question: What is the lex. least word on $\mathbb{N}$ avoiding squares?

Let's find its first few letters together!

$$01020103\cdots$$

Question: Can we describe it as Thue did, with a morphism?

Let $\varphi\colon \mathbb{N}^* \to \mathbb{N}^*, n \mapsto 0(n+1)$.

$$\begin{aligned}\varphi(0) &= 01\\ \varphi^2(0) &= 0102\\ \varphi^3(0) &= 01020103\\ &\vdots\\ \varphi^\omega(0) &= 010201030102010401020103010201050\cdots\end{aligned}$$

Theorem (Guay-Paquet and Shallit, 2009)

Let $\varphi\colon \mathbb{N}^* \to \mathbb{N}^*, n \mapsto 0(n+1)$.
The lexicographically least square-free word on $\mathbb{N}$ is $\varphi^\omega(0)$.

Their proof relies on two main ingredients:

1. Show that $\varphi^\omega(0)$ is square-free.
2. Show that $\varphi^\omega(0)$ is lexicographically least (decreasing any letter in $\varphi^\omega(0)$ to any smaller number introduces a square ending in that (same) position).

Lexicographically least word on $\mathbb{N}$ avoiding a-powers

Definitions: Let $a \geq 2$ be an integer.

- An a-power is a non-empty word of the form $w^a = \underbrace{ww\cdots w}_{a \text{ times}}$.

  Example: $a = 2$ squares
  $a = 3$ cubes

- A word is a-power-free if is does not contain any a-powers.
- Let $\varphi_a \colon \mathbb{N}^* \to \mathbb{N}^*, n \mapsto 0^{a-1}(n+1)$.

  Example: $\varphi_3(0) = 001$
  $\varphi_3^2(0) = 001001002$
  $\varphi_3^3(0) = 001001002001001002001001003$

Theorem (Guay-Paquet and Shallit, 2009)

Let $a \geq 2$ be an integer.
The lexicographically least a-power-free word on $\mathbb{N}$ is $\varphi_a^\omega(0)$.

The proof is similar.

Fractional powers

Example: $(0111)^{3/2} = \underbrace{0111}_{0111}\ \underbrace{01}_{\text{half of } 0111}$ is the $\frac{3}{2}$-power of 0111

Definition

Let $a, b \geq 1$ be integers with $\gcd(a, b) = 1$.
If $v = v_0 v_1 \cdots v_{\ell-1}$ is a non-empty word whose length ℓ is divisible by b, the $\frac{a}{b}$-power of v is the word

$$v^{a/b} = v^{\lfloor a/b \rfloor} v_0 v_1 \cdots v_{\ell \cdot \{a/b\}-1}.$$

Note that $|v^{a/b}| = \lfloor a/b \rfloor \cdot |v| + \{a/b\} \cdot |v| = (a/b)|v|$.

Example: $(0111)^{5/4} = 0111\ 0$ is the $\frac{5}{4}$-power of 0111

- $\frac{3}{2}$-powers look like $xyx = (xy)^{3/2}$ where $|y| = |x|$ since $|xyx| = \frac{3}{2} \cdot |xy|$
- $\frac{5}{4}$-powers look like $xyx = (xy)^{5/4}$ where $|y| = 3|x|$ since $|xyx| = \frac{5}{4} \cdot |xy|$

Lexicographically least word on $\mathbb{N}$ avoiding a/b-powers

A word is $\frac{a}{b}$-power-free if it contains no $\frac{a}{b}$-powers.

Notation

For $\frac{a}{b} > 1$, $\mathbf{w}_{a/b}$ is the lexicographically least $\frac{a}{b}$-power-free word on $\mathbb{N}$.

Theorem (Guay-Paquet and Shallit, 2009)

Let $a \geq 2$ be an integer and let $\varphi_a \colon \mathbb{N}^* \to \mathbb{N}^*, n \mapsto 0^{a-1}(n+1)$.
Then $\mathbf{w}_{a/1} = \mathbf{w}_a = \varphi_a^{\omega}(0)$.

Goal: Study those words $\mathbf{w}_{a/b}$ for $\frac{a}{b} > 1$.

Outside the interval $(1, 2)$

There is no need to look at fractions $\frac{a}{b} \geq 2$.

Theorem (Pudwell and Rowland, 2018)

Let a, b be positive integers with $\gcd(a, b) = 1$ and $\frac{a}{b} \geq 2$.
Then $\mathbf{w}_{a/b} = \mathbf{w}_a$.

Idea of the proof:

1. $\mathbf{w}_a$ is $\frac{a}{b}$-power-free, so $\mathbf{w}_{a/b} \leq_{\text{lex}} \mathbf{w}_a$.
2. $\mathbf{w}_{a/b}$ is a-power-free, so $\mathbf{w}_a \leq_{\text{lex}} \mathbf{w}_{a/b}$.

Consequence: We may assume that $\gcd(a, b) = 1$ and $1 < \frac{a}{b} < 2$.

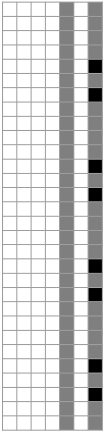
Lexicographically least word on $\mathbb{N}$ avoiding $\frac{5}{3}$-powers

$\frac{5}{3}$-powers look like $xyx = (xy)^{5/3}$ with $2|y| = |x|$ ($|xyx| = \frac{5}{3} \cdot |xy|$)

$$\mathbf{w}_{5/3} = 0000101\,0000101\,0000101\,0000101\,0000102\,0000101 \cdots$$

$00000 = (000)^{5/3}$, $00100 = (001)^{5/3}$

$$\mathbf{w}_{5/3} = \begin{array}{l} 0000101 \\ 0000101 \\ 0000101 \\ 0000101 \\ 0000102 \\ 0000101 \\ 0000102 \\ 0000101 \\ \vdots \end{array}$$



6 columns are constant: $000 \cdots$ or $111 \cdots$

7th column is self-similar: $w(7i + 6) = w(i) + 1$

Morphic structure of $\mathbf{w}_{5/3}$

$$\mathbf{w}_{5/3} = 0000101\,0000101\,0000101\,0000101\,0000102\,0000101\cdots$$

Theorem (Pudwell and Rowland, 2018)

Let $\varphi\colon \mathbb{N}^* \to \mathbb{N}^*, n \mapsto 000010(n+1)$ be a 7-uniform morphism.
Then $\mathbf{w}_{5/3} = \varphi^\omega(0)$.

As before, the proof relies on two main ingredients:

1. Show that $\varphi^\omega(0)$ is $\frac{5}{3}$-power-free.
2. Show that $\varphi^\omega(0)$ is lexicographically least.

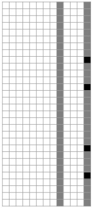
Observation: 7 columns $\rightsquigarrow$ 7 letters in $\varphi(n)$ $\rightsquigarrow$ 7-uniform morphism

Definition: Let $k \geq 0$ be an integer.
A morphism $\varphi\colon A^* \to B^*$ is k-uniform if $|\varphi(a)| = k$ for all $a \in A$.

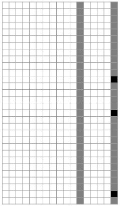
Example: $|\varphi(n)| = |000010(n+1)| = 7$, so φ is 7-uniform

$\mathbf{w}_{5/3}$ is not alone

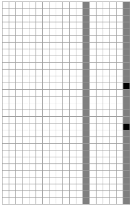
$$\frac{5}{3} < \frac{12}{7} < \frac{19}{11} < \frac{16}{9} < \frac{9}{5} < \frac{13}{7} < \frac{17}{9} < 2$$



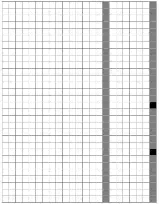
$\mathbf{w}_{9/5}$



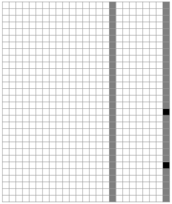
$\mathbf{w}_{12/7}$



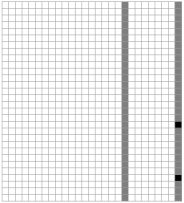
$\mathbf{w}_{13/7}$



$\mathbf{w}_{16/9}$



$\mathbf{w}_{17/9}$



$\mathbf{w}_{19/11}$

Theorem (Pudwell and Rowland, 2018)

Let $\frac{5}{3} \leq \frac{a}{b} < 2$ with b odd. Then $\mathbf{w}_{a/b} = \varphi^\omega(0)$, where $\varphi\colon \mathbb{N}^* \to \mathbb{N}^*, n \mapsto 0^{a-1}\,1\,0^{a-b-1}\,(n+1)$ is a $(2a-b)$-uniform morphism.

Example: $\frac{a}{b} = \frac{5}{3}$, $2a - b = 7$ $\quad$ $\frac{a}{b} = \frac{9}{5}$, $2a - b = 13$

$\varphi(n) = 000010(n+1)$ $\quad$ $\varphi(n) = 000000001000(n+1)$

Pudwell and Rowland (2018) identified other families of words $\mathbf{w}_{a/b}$ and the structure of some words $\mathbf{w}_{a/b}$ when $\mathbf{w}_{a/b} \notin$ these families.

Theorem (Pudwell and Rowland, 2018)

For each $\frac{a}{b}$ in the table below, $\mathbf{w}_{a/b} = \varphi^\omega(0)$ for some k-uniform morphism $\varphi\colon \mathbb{N}^* \to \mathbb{N}^*, n \mapsto u(n+d)$, where u is a word of length $k-1$.

$\frac{a}{b}$	d	k	$\frac{a}{b}$	d	k
$\frac{7}{4} \approx 1.75$	2	50847	$\frac{37}{26} \approx 1.42308$	1	2359
$\frac{8}{5} \approx 1.6$	2	733	$\frac{37}{28} \approx 1.32143$	1	5349
$\frac{13}{9} \approx 1.44444$	1	45430	$\frac{41}{28} \approx 1.46429$	1	2103
$\frac{17}{10} \approx 1.7$	2	55657	$\frac{49}{34} \approx 1.44118$	1	4171
$\frac{15}{11} \approx 1.36364$	1	6168	$\frac{55}{38} \approx 1.44737$	1	5269
$\frac{16}{13} \approx 1.23077$	1	12945	$\frac{53}{40} \approx 1.325$	1	9933
$\frac{18}{13} \approx 1.38462$	1	4188	$\frac{59}{42} \approx 1.40476$	1	5861
$\frac{19}{13} \approx 1.46154$	1	7698	$\frac{65}{46} \approx 1.41304$	1	7151
$\frac{21}{16} \approx 1.3125$	2	25441	$\frac{67}{46} \approx 1.45652$	1	7849
$\frac{25}{17} \approx 1.47059$	1	11705	$\frac{71}{50} \approx 1.42$	1	8569
$\frac{31}{22} \approx 1.40909$	1	1645	$\frac{73}{50} \approx 1.46$	1	9331
$\frac{33}{23} \approx 1.43478$	1	24995	$\frac{77}{54} \approx 1.42593$	1	10115

The values of k and d depend on $\frac{a}{b}$.

What's left then?

Well,

$$\mathbf{w}_{3/2} \text{ is slightly different}$$

and

$$\mathbf{w}_{5/4} \text{ was open (at the time).}$$

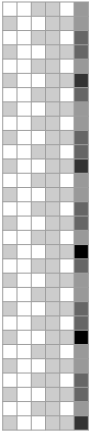
Let's look at both!

Lexicographically least word on $\mathbb{N}$ avoiding $\frac{3}{2}$-powers

$\frac{3}{2}$-powers look like $xyx = (xy)^{3/2}$ where $|y| = |x|$

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

$$\begin{aligned}\mathbf{w}_{3/2} = {} & 001102 \\ & 100112 \\ & 001103 \\ & 100113 \\ & 001102 \\ & 100114 \\ & 001103 \\ & 100112 \\ & \vdots\end{aligned}$$



5 columns are periodic:
$000\cdots$, $111\cdots$,
$0101\cdots$ or $1010\cdots$

6th column is self-similar:
$w(6i+5) = w(i) + 2$

Morphic description of $\mathbf{w}_{3/2}$

The trick is to use two types of letters.

Definitions:

- The alphabet is $\Sigma_2 = \{n_j : n \in \mathbb{Z}, j \in \{0,1\}\}$.

  Example: $0_0, 0_1, 1_0, 1_1, 2_0, 2_1 \in \Sigma_2$

- A coding is a 1-uniform morphism.

  Example: $\tau\colon \Sigma_2^* \to \mathbb{Z}^*, n_j \mapsto n$
  $\tau(0_0) = 0 = \tau(0_1)$

- We use the 6-uniform morphism $\varphi\colon \Sigma_2^* \to \Sigma_2^*$ defined by

$$\varphi(n_0) = 0_0 0_1 1_0 1_1 0_0 (n+2)_1$$
$$\varphi(n_1) = 1_0 0_1 0_0 1_1 1_0 (n+2)_1$$

  Example: $\varphi(0_0) = 0_0 0_1 1_0 1_1 0_0 2_1$

$$\varphi(n_0) = 0_0 0_1 1_0 1_1 0_0 (n+2)_1 \qquad \varphi(n_1) = 1_0 0_1 0_0 1_1 1_0 (n+2)_1$$

Iterate φ on 0_0:

$$\begin{aligned}
\varphi(0_0) &= 0_0 0_1 1_0 1_1 0_0 2_1 \\
\varphi^2(0_0) &= 0_0 0_1 1_0 1_1 0_0 2_1\ 1_0 0_1 0_0 1_1 1_0 2_1\ 0_0 0_1 1_0 1_1 0_0 3_1 \\
&\quad 1_0 0_1 0_0 1_1 1_0 3_1\ 0_0 0_1 1_0 1_1 0_0 2_1\ 1_0 0_1 0_0 1_1 1_0 4_1 \\
\varphi^3(0_0) &= 0_0 0_1 1_0 1_1 0_0 2_1\ 1_0 0_1 0_0 1_1 1_0 2_1\ 0_0 0_1 1_0 1_1 0_0 3_1 \\
&\quad 1_0 0_1 0_0 1_1 1_0 3_1\ 0_0 0_1 1_0 1_1 0_0 2_1\ 1_0 0_1 0_0 1_1 1_0 4_1 \\
&\quad 0_0 0_1 1_0 1_1 0_0 3_1\ 1_0 0_1 0_0 1_1 1_0 2_1\ 0_0 0_1 1_0 1_1 0_0 2_1 \cdots
\end{aligned}$$

$$\mathbf{w}_{3/2} = 001102\ 100112\ 001103\ 100113\ 001102\ 100114\ 001103 \cdots$$

Theorem (Rowland and Shallit, 2012)

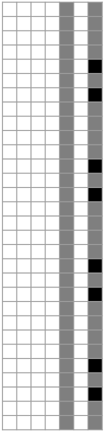
$\mathbf{w}_{3/2} = \tau(\varphi^\omega(0_0))$.

Idea of proof:

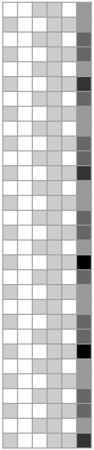
1. $\tau(\varphi^\omega(0_0))$ is $\frac{3}{2}$-power-free.
2. $\tau(\varphi^\omega(0_0))$ is lexicographically least.

Why is this case different?

$$\mathbf{w}_{5/3} = \begin{array}{l} 0000101 \\ 0000101 \\ 0000101 \\ 0000101 \\ 0000102 \\ 0000101 \\ 0000102 \\ 0000101 \\ \vdots \end{array}$$



$$\mathbf{w}_{3/2} = \begin{array}{l} 001102 \\ 100112 \\ 001103 \\ 100113 \\ 001102 \\ 100114 \\ 001103 \\ 100112 \\ \vdots \end{array}$$



Let's look at the background (=periodic columns).

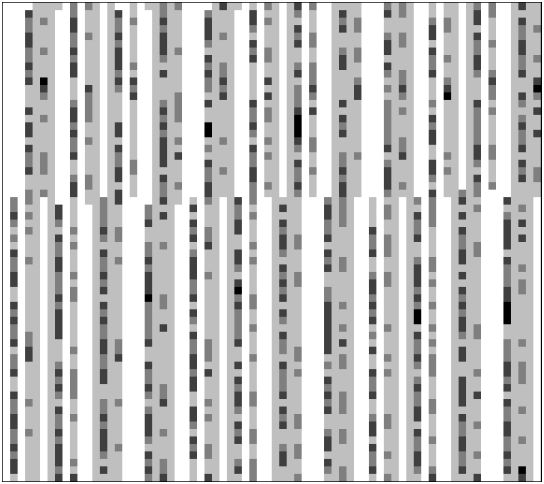
- $\mathbf{w}_{5/3}$: one constant background 00010
- $\mathbf{w}_{3/2}$: two alternating backgrounds 00110 and 10011

This is why we need two types of letters.

Lexicographically least word on $\mathbb{N}$ avoiding $\frac{5}{4}$-powers

$\frac{5}{4}$-powers look like $xyx = (xy)^{5/4}$ where $|y| = 3|x|$

$$\mathbf{w}_{5/4} = 000011110202101001011121200001311010210130\cdots$$



$\mathbf{w}_{5/4}$ partitioned into 72 columns

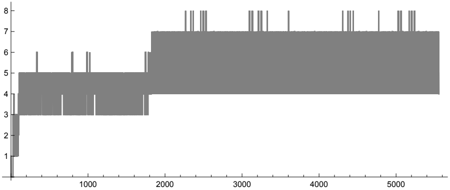
How did we discover the structure of $\mathbf{w}_{5/4}$?

The strategy in previous papers is to find a single column that is not eventually periodic.

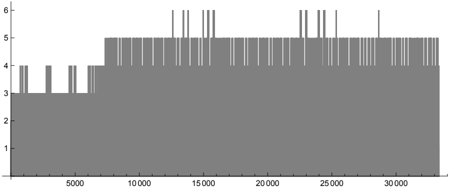
Here, 72 seems to be the largest number of columns (width) that gives exactly one column that is not eventually periodic.
That column is the 32nd column.

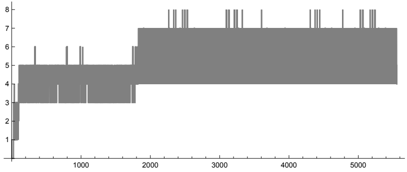
Write $\mathbf{w}_{5/4} = w(0)w(1)\cdots$.

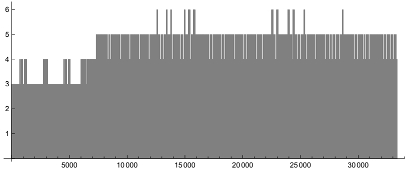
$w(72i+31)_{i\geq 0}$



$w(i)_{i\geq 0}$







Peaks in these plots:

- Occurrences of 8 in $w(72i+31)_{i\geq 0}$ $\leftrightsquigarrow$ occurrences of 6 in $w(i)_{i\geq 0}$
- Positions of 8 in $w(i)_{i\geq 0}$

$$163183, 168151, 170311, 177367, 179527, 179959, \ldots$$

Positions of 6 in $w(i)_{i\geq 0}$

$$12607, 13435, 13795, 14971, 15331, 15403, \ldots$$

Lining up the peaks:

$$8 = w(168151) = w(72 \cdot 69 + 163183)$$
$$6 = w(13435) = w(12 \cdot 69 + 12607)$$

Conjecture: $w(72i + 163183) = w(12i + 12607) + 2 \quad \forall i \geq 0$

Now we would like to get rid of the coefficient 12 of i in the relationship:

$$(w(6i+163183)-w(i+12607))_{i\geq 0} = 2,3,2,3,2,1,2,1,2,3,2,3,2,1,2,1\ldots$$

This gives a periodic sequence with period $2,3,2,3,2,1,2,1$.
Actually, this periodic difference begins 6687 terms earlier:

Conjecture

Let $w(i)$ be the ith letter of the word $\mathbf{w}_{5/4}$.
For all $i \geq 0$,

$$w(6i+123061) = w(i+5920) + \begin{cases} 3 & \text{if } i \equiv 0,2 \mod 8 \\ 1 & \text{if } i \equiv 4,6 \mod 8 \\ 2 & \text{if } i \equiv 1 \mod 2 \end{cases}$$

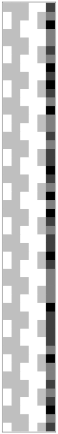
$(163183 - 6 \cdot 6687 = 123061$ and $12607 - 6687 = 5920)$

The conjecture suggests we should look at 6 columns instead of 72.

Partition into 6 columns instead of 72

$$\mathbf{w}_{5/4} = 0000111102021010010112120000131101021013 02\cdots$$

$$\mathbf{w}_{5/4} = \mathbf{p}\begin{array}{l}111003\\011012\\010014\\110002\\111002\\011013\\010012\\110004\\\vdots\end{array}$$



5 columns are periodic:
$1001\cdots$, $111\cdots$, $1100\cdots$
$000\cdots$ or $0110\cdots$

6th column is self-similar
(satisfies our conjecture)

Morphic description of $\mathbf{w}_{5/4}$

Definitions:

- Alphabet $\Sigma_8 = \{n_j : n \in \mathbb{Z}, j \in \{0, 1, \ldots, 7\}\}$
- 6-uniform morphism $\varphi \colon \Sigma_8^* \to \Sigma_8^*$ defined by

$$\begin{aligned} \varphi(n_0) &= 0_0 1_1 0_2 0_3 1_4 (n+3)_5 & \varphi(n_4) &= 0_0 1_1 0_2 0_3 1_4 (n+1)_5 \\ \varphi(n_1) &= 1_6 1_7 0_0 0_1 0_2 (n+2)_3 & \varphi(n_5) &= 1_6 1_7 0_0 0_1 0_2 (n+2)_3 \\ \varphi(n_2) &= 1_4 1_5 1_6 0_7 0_0 (n+3)_1 & \varphi(n_6) &= 1_4 1_5 1_6 0_7 0_0 (n+1)_1 \\ \varphi(n_3) &= 0_2 1_3 1_4 0_5 1_6 (n+2)_7 & \varphi(n_7) &= 0_2 1_3 1_4 0_5 1_6 (n+2)_7 \end{aligned}$$

- Coding $\tau \colon \Sigma_8^* \to \mathbb{Z}^*, n_j \mapsto n$

Theorem (Rowland and S., 2020)

There exist a length-6764 word $\mathbf{p}$ on $\mathbb{N} = \{0, 1, \ldots\}$ and a length-20226 word $\mathbf{z}$ on Σ_8 such that $\mathbf{w}_{5/4} = \mathbf{p}\, \tau(\varphi(\mathbf{z})\varphi^2(\mathbf{z}) \cdots)$.

$\mathbf{p}$ = length-6764 prefix of $\mathbf{w}_{5/4} = 000011110202101001011121200\cdots$

$\mathbf{z} = 0_2 0_3 3_4 0_5 1_6 1_7 (-1_0) 2_1 0_2 2_3 2_4 0_5 3_6 0_7 (-1_0) 1_1 (-1_2) 1_3 2_4 2_5 \cdots$

Remark: $\tau(\mathbf{z})$ cannot be a factor of $\mathbf{w}_{5/4}$.

Idea of the proof

As before, two steps:

1. Show that $\mathbf{p}\,\tau(\varphi(\mathbf{z})\varphi^2(\mathbf{z})\cdots)$ avoids $\frac{5}{4}$-powers.
2. Show that $\mathbf{p}\,\tau(\varphi(\mathbf{z})\varphi^2(\mathbf{z})\cdots)$ is lexicographically least (decreasing any letter in $\mathbf{p}\,\tau(\varphi(\mathbf{z})\varphi^2(\mathbf{z})\cdots)$ to any smaller number introduces a $\frac{5}{4}$-power ending in that (same) position).

In Step 1 for previously studied $\mathbf{w}_{a/b}$, we show that the corresponding φ is $\frac{a}{b}$-power-free. A morphism φ is $\frac{a}{b}$-power-free if

$$w \text{ is } \tfrac{a}{b}\text{-power-free} \Rightarrow \varphi(w) \text{ is } \tfrac{a}{b}\text{-power-free.}$$

However, the morphism for $\mathbf{w}_{5/4}$ is not $\frac{5}{4}$-power-free.

Example: For $n, m \in \mathbb{Z}$, the word $0_4 n_5 m_6$ is $\frac{5}{4}$-power-free, but its image under φ is not

$$\varphi(0_4 n_5 m_6) = 0_0 1_1 0_2 0_3 1_4 1_5 \;\; 1_6 1_7 0_0 0_1 0_2 (n+2)_3 \;\; 1_4 1_5 1_6 0_7 0_0 (m+1)_1$$

We need a new notion!

Pre-$\frac{5}{4}$-power-freeness

A word is pre-$\frac{5}{4}$-power-free if every factor xyx' with $|x| = \frac{1}{3}|y| = |x'|$ satisfies $\varphi(x) \neq \varphi(x')$.

Example: $0_0 n_1 n_2 n_3 2_4$ is not pre-$\frac{5}{4}$-power-free because

$$\varphi(0_0) = 0_0 1_1 0_2 0_3 1_4 3_5 = \varphi(2_4)$$

Stronger condition:

Proposition (Rowland and S., 2020)

If w is pre-$\frac{5}{4}$-power-free, then w is $\frac{5}{4}$-power-free.

φ preserves pre-$\frac{5}{4}$-power-freeness on some sub-alphabet of Σ_8:

Proposition (Rowland and S., 2020)

Let Γ be the set

$$\{-3_0, -3_2, -2_0, -2_1, -2_2, -2_3, -2_5, -2_7, -1_1, -1_3, -1_4, -1_5, -1_6, -1_7, 0_4, 0_6\}.$$

If $w \in (\Sigma_8 \setminus \Gamma)^*$ is pre-$\frac{5}{4}$-power-free, then $\varphi(w)$ is pre-$\frac{5}{4}$-power-free.

A better idea of the proof

1. To prove $\frac{5}{4}$-power-freeness, we need a sequence of results:
   - $\mathbf{z}\varphi(\mathbf{z})\varphi^2(\mathbf{z})\cdots$ is pre-$\frac{5}{4}$-power-free.
   - $\varphi(\mathbf{z}\varphi(\mathbf{z})\varphi^2(\mathbf{z})\cdots)$ is $\frac{5}{4}$-power-free (two previous propositions).
   - $\tau(\varphi(\mathbf{z})\varphi^2(\mathbf{z})\cdots)$ is $\frac{5}{4}$-power-free.
   - $\mathbf{p}\tau(\varphi(\mathbf{z})\varphi^2(\mathbf{z})\cdots)$ is $\frac{5}{4}$-power-free.
2. To prove lexicographic leastness, we use a case analysis and a complicated induction.

Both steps involve large finite checks carried out programmatically.

Corollary (Rowland and S., 2020)

Let $w(i)$ be the ith letter of the word $\mathbf{w}_{5/4}$.
For all $i \geq 0$,

$$w(6i + 123061) = w(i + 5920) + \begin{cases} 3 & \text{if } i \equiv 0, 2 \mod 8 \\ 1 & \text{if } i \equiv 4, 6 \mod 8 \\ 2 & \text{if } i \equiv 1 \mod 2 \end{cases}$$

Our conjecture was true.

Regularity in the sense of Allouche and Shallit (1992)

Definition: Let $k \geq 2$ be an integer. Let $s(i)_{i\geq 0}$ be a sequence.

- The *k-kernel* of $s(i)_{i\geq 0}$ is the set of subsequences

$$\ker_k(s) = \{s(k^e i + j)_{i\geq 0} \colon e \geq 0 \text{ and } 0 \leq j \leq k^e - 1\}.$$

- $s(i)_{i\geq 0}$ is *k-regular* if there exists a finite number of sequences $s_1(i)_{i\geq 0}, \ldots, s_r(i)_{i\geq 0}$ such that each sequence in $\ker_k(s)$ is a $\mathbb{Z}$-linear combination of $s_1(i)_{i\geq 0}, \ldots, s_r(i)_{i\geq 0}$.

Sum-of-digits function S_2 in base 2

$i = \sum_t d_t 2^t$ with $d_t \in \{0,1\}$ (i as a sum of powers of 2)

$S_2(i) = \sum_t d_t = \#$ of 1's in the binary expansion of i

i	0	1	2	3	4	5	6	7	8	9
bin. exp. of i	ε	1	10	11	100	101	110	111	1000	1001
$S_2(i)$	0	1	1	2	1	2	2	3	1	2

$\ker_2(S_2)$ is the set

$$\{\underbrace{S_2(i)_{i\geq 0}}_{e=0,\, j=0}, \underbrace{S_2(2i)_{i\geq 0}, S_2(2i+1)_{i\geq 0}}_{e=1,\, 0\leq j\leq 2^1-1}, \underbrace{S_2(4i)_{i\geq 0}, \ldots, S_2(4i+3)_{i\geq 0}}_{e=2,\, 0\leq j\leq 2^2-1}, \ldots\}$$

$$S_2(2i) = S_2(i)$$

$$S_2(2i+1) = S_2(i) + 1$$

Any sequence in $\ker_2(S_2)$ is a $\mathbb{Z}$-linear combination of $S_2(i)_{i\geq 0}$ and the constant sequence $1, 1, 1, \ldots$.

Example: $S_2(4i+3) = 1 \cdot S_2(i) + 2 \cdot 1$.

So S_2 is 2-regular.

The sequences of letters in the words $\mathbf{w}_{a/b}$ I have talked about form k-regular sequences for some value of $k \geq 2$:

- $\mathbf{w}_{a/1} = \mathbf{w}_a$ with $a \geq 2 \rightsquigarrow a$-regular
- $\mathbf{w}_{a/b}$ with $\frac{5}{3} \leq \frac{a}{b} < 2$ and b odd $\rightsquigarrow (2a - b)$-regular
- $\mathbf{w}_{3/2} \rightsquigarrow$ 6-regular

The value of k is not unique.

Theorem (Allouche and Shallit, 1992)

For $e \geq 1$, a sequence is k-regular $\Leftrightarrow$ it is k^e-regular.

Two integers $k, \ell \geq 2$ are multiplicative dependent if there exist integers $e, e' > 0$ such that $k^e = \ell^{e'}$.

Corollary (Pudwell and Rowland, 2018)

Let a, b be relatively prime positive integers such that $\frac{a}{b} > 1$. The values of k for which $\mathbf{w}_{a/b}$ is k-regular are multiplicative dependent.

Pudwell and Rowland (2018) identified a family of regular sequences.

Theorem (Pudwell and Rowland, 2018)

Let $k \geq 2$ and $d \geq 0$. Let u be a word on $\mathbb{N}$ with $|u| = k - 1$. Let v be a non-empty finite word on $\mathbb{N} \cup \{0'\}$ whose first letter is $0'$ and whose remaining letters are in $\mathbb{N}$. Let

$$\varphi\colon (\mathbb{N} \cup \{0'\})^* \to (\mathbb{N} \cup \{0'\})^*, n \mapsto \begin{cases} v\varphi(0) & \text{if } n = 0' \\ u(n+d) & \text{if } n \in \mathbb{N}. \end{cases}$$

Then the sequence of letters in $\tau(\varphi^\omega(0'))$ is a k-regular sequence.

This result can be applied to show that, for many fractions $\frac{a}{b} > 1$, the sequence of letters in $\mathbf{w}_{a/b}$ is k-regular for some $k \geq 2$.
But it does not apply to $\mathbf{w}_{5/4}$.

A generalization of the previous result

Theorem (Rowland and S., 2020)

Let $k \geq 2$, $\ell \geq 1$ and $d_0, d_1, \ldots, d_{\ell-1} \in \mathbb{Z}$. Let u be a word on $\mathbb{Z}$ of length $k\ell$. Let r, s be nonnegative integers such that $r - s + k - 1 \geq 0$. Let w be an infinite word on $\mathbb{Z}$ such that, for all $0 \leq m \leq k-1$ and all $i \geq 0$,

$$w(ki + r + m) = \begin{cases} u((ki+m) \bmod k\ell) & \text{if } 0 \leq m \leq k-2 \\ w(i+s) + d_{i \bmod \ell} & \text{if } m = k-1. \end{cases}$$

Then $w(i)_{i\geq 0}$ is k-regular.

Corollary (Rowland and S., 2020)

The sequence of letters in $\mathbf{w}_{5/4}$ is a 6-regular sequence.

References

- J.-P. Allouche, J. Shallit, The ring of k-regular sequences, *Theoret. Comput. Sci.* **98** (1992), 163–197.
- J. Berstel, Axel Thue's Papers on Repetitions in Words: A Translation, *Publications du LaCIM* **20**, Université du Québec à Montréal, 1995.
- J. Berstel, D. Perrin, The origins of combinatorics on words, *European J. Combin.* **28** (2007), 996–1022.
- M. Guay-Paquet, J. Shallit, Avoiding squares and overlaps over the natural numbers, *Discrete Math.* **309** (2009), 6245–6254.
- M. Lothaire, *Combinatorics On Words*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1997. Corrected reprint of the 1983 original.
- M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002.
- M. Lothaire, *Applied Combinatorics on Words*, Cambridge University Press, 2005.

References

L. Pudwell, E. Rowland, Avoiding fractional powers over the natural numbers, *Electron. J. Combin.* **25** (2018), Paper 2.27, 46 pp.

E. Rowland, J. Shallit, Avoiding 3/2-powers over the natural numbers, *Discrete Math.* **312** (2012), 1282–1288.

E. Rowland, M. Stipulanti, Avoiding 5/4-powers on the alphabet of nonnegative integers, in preparation.

N. Sloane et al., The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.

A. Thue, Über unendliche Zeichenreihen, *Norske Vid. Selsk. Skr. I Math-Nat. Kl.* **7** (1906), 1–22.

A. Thue, Über die gegenseitige Loge gleicher Teile gewisser Zeichenreihen, *Norske Vid. Selsk. Skr. I Math-Nat. Kl. Chris.* **1** (1912), 1–67.