Words whose factors are not shuffled squares

Laurent Bulteau, Vincent Jugé & Stéphane Vialette

LIGM – Université Gustave Eiffel & CNRS

Séminaire Flajolet
02/02/2023

Squares, powers and square-free words

A word u is a **square** if there exists a word v such that $u = v \cdot v$.

Examples: $01 \cdot 01$ and $010101 \cdot 010101$

Counter-examples: 0110, 010101 and 01201

Squares, powers and square-free words

A word u is a **square** if there exists a word v such that $u = v \cdot v$;
k^{th} **power** if there exists a word v such that $u = v^k$;
power if u is a k^{th} power for some integer $k \geqslant 2$.

Examples: $01 \cdot 01$ and $010101 \cdot 010101$ are **squares**;
$01 \cdot 01 \cdot 01$ and $0101 \cdot 0101 \cdot 0101$ are **cubes**;
$01 \cdot 01 \cdot 01 \cdot 01 \cdot 01$ is a **power**.

Counter-examples: 0110, 010101 and 01201 are not **squares**;
0110 and 0101, 01201 are not **cubes**;
0110 and 01201 are not **powers**.

Squares, powers and square-free words

A word u is a **square** if there exists a word v such that $u = v \cdot v$;
k^{th} **power** if there exists a word v such that $u = v^k$;
power if u is a k^{th} power for some integer $k \geqslant 2$.

Examples: $01 \cdot 01$ and $010101 \cdot 010101$ are **squares**;
$01 \cdot 01 \cdot 01$ and $0101 \cdot 0101 \cdot 0101$ are **cubes**;
$01 \cdot 01 \cdot 01 \cdot 01 \cdot 01$ is a **power**.

Counter-examples: 0110, 010101 and 01201 are not **squares**;
0110 and 0101, 01201 are not **cubes**;
0110 and 01201 are not **powers**.

A word u is **square-free** if none of its non-empty (contiguous) factors is a square (or a power).

Examples: 010, 0102010 and 01020120210 are **square-free**.
Counter-examples: 0101, 0110 and binary words u of length $|u| \geqslant 4$ are not **square-free**.

Why studying squares, powers and square-free words?

> “*A problem that is easy to state, yet difficult to solve, should be studied primarily for its own sake, even if it does not yet have concrete applications.*” Axel Thue, 1912

Why studying squares, powers and square-free words?

> *"A problem that is easy to state, yet difficult to solve, should be studied primarily for its own sake, even if it does not yet have concrete applications."* Axel Thue, 1912

Squares, powers and square-free words arise in many areas:

- group theory;
- bio-informatics;
- compression algorithms;
- automata theory;
- ...

Why studying squares, powers and square-free words?

> *“A problem that is easy to state, yet difficult to solve, should be studied primarily for its own sake, even if it does not yet have concrete applications.”* Axel Thue, 1912

Squares, powers and square-free words arise in many areas:

- group theory;
- bio-informatics;
- compression algorithms;
- automata theory;
- …

Some questions of interest:

Given a finite alphabet $A_n = \{0, 1, \ldots, n-1\}$,

- how difficult is it to check whether a word $u \in A_n^*$ is a square? a k^{th} power? a power?
- how difficult is it to check whether a word $u \in A_n^*$ is square-free?

Why studying squares, powers and square-free words?

> *“A problem that is easy to state, yet difficult to solve, should be studied primarily for its own sake, even if it does not yet have concrete applications.”* Axel Thue, 1912

Squares, powers and square-free words arise in many areas:

- group theory;
- bio-informatics;
- compression algorithms;
- automata theory;
- ...

Some questions of interest:

Given a finite alphabet $A_n = \{0, 1, \ldots, n-1\}$,

- how difficult is it to check whether a word $u \in A_n^*$ is a square? a k^{th} power? a power?
- how difficult is it to check whether a word $u \in A_n^*$ is square-free?
- are there arbitrarily long square-free words in A_n^*?

Why studying squares, powers and square-free words?

> “*A problem that is easy to state, yet difficult to solve, should be studied primarily for its own sake, even if it does not yet have concrete applications.*” Axel Thue, 1912

Squares, powers and square-free words arise in many areas:

- group theory;
- bio-informatics;
- compression algorithms;
- automata theory;
- ...

Some questions of interest:

Given a finite alphabet $A_n = \{0, 1, \ldots, n-1\}$,

- how difficult is it to check whether a word $u \in A_n^*$ is a square? a k^{th} power? a power?
- how difficult is it to check whether a word $u \in A_n^*$ is square-free?
- are there arbitrarily long square-free words in A_n^*? infinite square-free words in A_n^ω?

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore)**:** Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906)**:** The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** $\mathbf{TM}$ defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

0

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1$$

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore)**:** Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906)**:** The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10$$

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10 \cdot 1001$$

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10 \cdot 1001 \cdot 10010110$$

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10 \cdot 1001 \cdot 10010110 \cdot 1001011001101001 \cdots$$

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10 \cdot 1001 \cdot 10010110 \cdot 1001011001101001 \cdots$$

Proof: Assume that there is a shortest cube-free word w for which $\mathbf{TM}(w)$ has a factor x^3:

- if $\mathbf{TM}(w) = x^3$, $|x|$ is even, and there exists a word u such that $\mathbf{TM}(u) = x$ and $w = u^3$;

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10 \cdot 1001 \cdot 10010110 \cdot 1001011001101001 \cdots$$

Proof: Assume that there is a shortest cube-free word w for which $\mathbf{TM}(w)$ has a factor x^3:

- if $\mathbf{TM}(w) = x^3$, $|x|$ is even, and there exists a word u such that $\mathbf{TM}(u) = x$ and $w = u^3$;
- if $\mathbf{TM}(w) = \lambda \cdot x^3$ or $\mathbf{TM}(w) = x^3 \cdot \lambda$ for some letter $\lambda \in A_2$,

$$\mathbf{1}_{\lambda=0} \equiv 3|x|_0 + \mathbf{1}_{\lambda=0} \equiv |\mathbf{TM}(w)|_0 \equiv |\mathbf{TM}(w)|_1 \equiv 3|x|_1 + \mathbf{1}_{\lambda=1} \equiv \mathbf{1}_{\lambda=1} \pmod 3;$$

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore)**:** Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906)**:** The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10 \cdot 1001 \cdot 10010110 \cdot 1001011001101001 \cdots$$

Proof: Assume that there is a shortest cube-free word w for which $\mathbf{TM}(w)$ has a factor x^3:

- if $\mathbf{TM}(w) = x^3$, $|x|$ is even, and there exists a word u such that $\mathbf{TM}(u) = x$ and $w = u^3$;
- if $\mathbf{TM}(w) = \lambda \cdot x^3$ or $\mathbf{TM}(w) = x^3 \cdot \lambda$ for some letter $\lambda \in A_2$,

$$\mathbf{1}_{\lambda=0} \equiv 3|x|_0 + \mathbf{1}_{\lambda=0} \equiv |\mathbf{TM}(w)|_0 \equiv |\mathbf{TM}(w)|_1 \equiv 3|x|_1 + \mathbf{1}_{\lambda=1} \equiv \mathbf{1}_{\lambda=1} \pmod 3;$$

- if $\mathbf{TM}(w) = \lambda \cdot x^3 \cdot \mu$ for some letters $\lambda \in A_2$ and $\mu \in A_2$, $|x|$ is even, and we can factor x as $x = a \cdot y \cdot b$, where $|y|$ is even; then, $\mathbf{TM}(w) = \lambda a \cdot y \cdot ba \cdot y \cdot ba \cdot y \cdot b\mu$: there exists a word u such that $\mathbf{TM}(u) = y$, and $\mathbf{TM}(\lambda) = \lambda a = \mathbf{TM}(b) = b\mu$, so that $w = \lambda \cdot u \cdot \lambda \cdot u \cdot \lambda \cdot u \cdot \lambda$.

Are there infinite square-free or k^{th}-power-free words? (1/2)

Lemma (Folklore): Square-free words over the alphabet $A_2 = \{0, 1\}$ are ε, 0, 01, 010, 1, 10, 101.

Theorem (Thue, 1906): The **Thue-Morse** word is cube-free. This is the **infinite fixed-point** of the **morphism of monoids** **TM** defined by $\mathbf{TM}(0) = 01$ and $\mathbf{TM}(1) = 10$, and starting with 0:

$$0 \cdot 1 \cdot 10 \cdot 1001 \cdot 10010110 \cdot 1001011001101001 \cdots$$

Proof: Assume that there is a shortest cube-free word w for which $\mathbf{TM}(w)$ has a factor x^3:

- if $\mathbf{TM}(w) = x^3$, $|x|$ is even, and there exists a word u such that $\mathbf{TM}(u) = x$ and $w = u^3$;
- if $\mathbf{TM}(w) = \lambda \cdot x^3$ or $\mathbf{TM}(w) = x^3 \cdot \lambda$ for some letter $\lambda \in A_2$,

$$\mathbf{1}_{\lambda=0} \equiv 3|x|_0 + \mathbf{1}_{\lambda=0} \equiv |\mathbf{TM}(w)|_0 \equiv |\mathbf{TM}(w)|_1 \equiv 3|x|_1 + \mathbf{1}_{\lambda=1} \equiv \mathbf{1}_{\lambda=1} \pmod 3;$$

- if $\mathbf{TM}(w) = \lambda \cdot x^3 \cdot \mu$ for some letters $\lambda \in A_2$ and $\mu \in A_2$, $|x|$ is even, and we can factor x as $x = a \cdot y \cdot b$, where $|y|$ is even; then, $\mathbf{TM}(w) = \lambda a \cdot y \cdot ba \cdot y \cdot ba \cdot y \cdot b\mu$: there exists a word u such that $\mathbf{TM}(u) = y$, and $\mathbf{TM}(\lambda) = \lambda a = \mathbf{TM}(b) = b\mu$, so that $w = \lambda \cdot u \cdot \lambda \cdot u \cdot \lambda \cdot u \cdot \lambda$.

In conclusion, the Thue-Morse morphism **TM** is **cube-free**: if w is cube-free, so is $\mathbf{TM}(w)$!

Are there infinite square-free or k^{th}-power-free words? (2/2)

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Are there infinite square-free or k^{th}-power-free words? (2/2)

Theorem (Thue, 1912; Crochemore, 1982)**:** An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Open question: Can we decide whether a given ℓ-uniform morphism is cube-free?

Are there infinite square-free or k^{th}-power-free words? (2/2)

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Open question: Can we decide whether a given ℓ-uniform morphism is cube-free?

Corollary (Thue, 1906; Leech, 1957): The **Leech morphism** $\mathbf{L}$ defined by

$$\mathbf{L}(0) = 0121021201210 \qquad \mathbf{L}(1) = 1202102012021 \qquad \mathbf{L}(2) = 2010210120102$$

is **square-free**. Thus, the infinite fixed-point of $\mathbf{L}$ starting with 0 is square-free.

Are there *many* square-free or k^{th}-power-free words? (1/3)

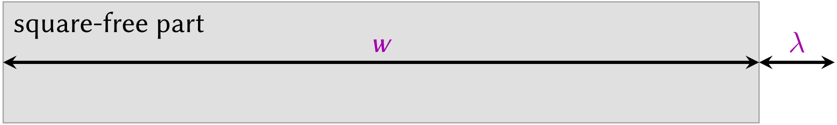
How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Are there *many* square-free or k^{th}-power-free words? (1/3)

How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Analytic approach when $\ell \geqslant 4$:

1. For all square-free words $w \in A_\ell^n$ and letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ has a factorisation $w \cdot \lambda = p \cdot s \cdot s$, where $p \cdot s$ is square-free: knowing $p \cdot s$ determines w and λ.



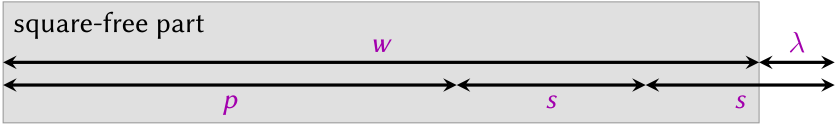
Are there *many* square-free or k^{th}-power-free words? (1/3)

How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Analytic approach when $\ell \geqslant 4$:

1. For all square-free words $w \in A_\ell^n$ and letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ has a factorisation $w \cdot \lambda = p \cdot s \cdot s$, where $p \cdot s$ is square-free: knowing $p \cdot s$ determines w and λ.

if $s \neq \varepsilon$:



Are there *many* square-free or k^{th}-power-free words? (1/3)

How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Analytic approach when $\ell \geqslant 4$:

1. For all square-free words $w \in A_\ell^n$ and letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ has a factorisation $w \cdot \lambda = p \cdot s \cdot s$, where $p \cdot s$ is square-free: knowing $p \cdot s$ determines w and λ.
2. Let W_ℓ^n be the set of **square-free** words in A_ℓ^n: $|W_\ell^n|\,|A_\ell| \leqslant |W_\ell^0| + |W_\ell^1| + \cdots + |W_\ell^{n+1}|$.

Are there *many* square-free or k^{th}-power-free words? (1/3)

How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Analytic approach when $\ell \geqslant 4$:

1. For all square-free words $w \in A_\ell^n$ and letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ has a factorisation $w \cdot \lambda = p \cdot s \cdot s$, where $p \cdot s$ is square-free: knowing $p \cdot s$ determines w and λ.
2. Let W_ℓ^n be the set of **square-free** words in A_ℓ^n: $|W_\ell^n|\,|A_\ell| \leqslant |W_\ell^0| + |W_\ell^1| + \cdots + |W_\ell^{n+1}|$.
3. We prove by induction that $|W_\ell^{n+1}| \geqslant (\ell - 2)\,|W_\ell^n|$, so that $|W_\ell^n| \geqslant (\ell - 2)^n$:

$$\frac{|W_\ell^{n+1}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{|W_\ell^{n-k}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{1}{(\ell - 2)^k} = (\ell - 2) + \frac{\ell - 4}{\ell - 3} \geqslant \ell - 2.$$

Are there *many* square-free or k^{th}-power-free words? (1/3)

How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Analytic approach when $\ell \geqslant 4$:

1. For all square-free words $w \in A_\ell^n$ and letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ has a factorisation $w \cdot \lambda = p \cdot s \cdot s$, where $p \cdot s$ is square-free: knowing $p \cdot s$ determines w and λ.
2. Let W_ℓ^n be the set of **square-free** words in A_ℓ^n: $|W_\ell^n|\,|A_\ell| \leqslant |W_\ell^0| + |W_\ell^1| + \cdots + |W_\ell^{n+1}|$.
3. We prove by induction that $|W_\ell^{n+1}| \geqslant (\ell - 2)\,|W_\ell^n|$, so that $|W_\ell^n| \geqslant (\ell - 2)^n$:

$$\frac{|W_\ell^{n+1}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{|W_\ell^{n-k}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{1}{(\ell - 2)^k} = (\ell - 2) + \frac{\ell - 4}{\ell - 3} \geqslant \ell - 2.$$

Corollaries (Folklore): $|W_\ell^n| \geqslant (\ell - 2)^n$ when $\ell \geqslant 4$.

There are infinite square-free words in A_ℓ^ω when $\ell \geqslant 4$.

Are there *many* square-free or k^{th}-power-free words? (1/3)

How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Analytic approach when $\ell \geqslant 4$:

1. For all square-free words $w \in A_\ell^n$ and letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ has a factorisation $w \cdot \lambda = p \cdot s \cdot s$, where $p \cdot s$ is square-free: knowing $p \cdot s$ determines w and λ.
2. Let W_ℓ^n be the set of **square-free** words in A_ℓ^n: $|W_\ell^n|\,|A_\ell| \leqslant |W_\ell^0| + |W_\ell^1| + \cdots + |W_\ell^{n+1}|$.
3. We prove by induction that $|W_\ell^{n+1}| \geqslant (\ell - 2)\,|W_\ell^n|$, so that $|W_\ell^n| \geqslant (\ell - 2)^n$:

$$\frac{|W_\ell^{n+1}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{|W_\ell^{n-k}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{1}{(\ell - 2)^k} = (\ell - 2) + \frac{\ell - 4}{\ell - 3} \geqslant \ell - 2.$$

Corollaries (Folklore): $|W_\ell^n| \geqslant (\ell - 2)^n$ when $\ell \geqslant 4$.

There are infinite square-free words in A_ℓ^ω when $\ell \geqslant 4$;
cube-free words in A_ℓ^ω when $\ell \geqslant 3$;
5^{th}-power-free words in A_ℓ^ω when $\ell \geqslant 2$.

Are there *many* square-free or k^{th}-power-free words? (1/3)

How many square-free words of length n over the alphabet $A_\ell = \{0, 1, \ldots, \ell - 1\}$ are there?

Analytic approach when $\ell \geqslant 4$:

1. For all square-free words $w \in A_\ell^n$ and letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ has a factorisation $w \cdot \lambda = p \cdot s \cdot s$, where $p \cdot s$ is square-free: knowing $p \cdot s$ determines w and λ.
2. Let W_ℓ^n be the set of **square-free** words in A_ℓ^n: $|W_\ell^n|\,|A_\ell| \leqslant |W_\ell^0| + |W_\ell^1| + \cdots + |W_\ell^{n+1}|$.
3. We prove by induction that $|W_\ell^{n+1}| \geqslant (\ell - 2)\,|W_\ell^n|$, so that $|W_\ell^n| \geqslant (\ell - 2)^n$:

$$\frac{|W_\ell^{n+1}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{|W_\ell^{n-k}|}{|W_\ell^n|} \geqslant (\ell - 1) - \sum_{k \geqslant 1} \frac{1}{(\ell - 2)^k} = (\ell - 2) + \frac{\ell - 4}{\ell - 3} \geqslant \ell - 2.$$

Corollaries (Folklore): $|W_\ell^n| \geqslant (\ell - 2)^n$ when $\ell \geqslant 4$.

There are infinite square-free words in A_ℓ^ω when $\ell \geqslant 4$;

⚠ Non-constructive existence results!

cube-free words in A_ℓ^ω when $\ell \geqslant 3$;

5^{th}-power-free words in A_ℓ^ω when $\ell \geqslant 2$.

Are there *many* square-free words? (2/3)

Theorem (Brandenburg, 1983; Brinkhuis, 1983; Kolpakov, 2006; Shur, 2009): $|W_3^n| \geqslant (5/4)^n$.

Proof that $|W_3^n| \geqslant 2^{n/72} \approx 1.01^n$: If $w = 1201021$, the **A_4-to-A_3 morphism** φ defined by

$$\varphi(0) = \mathbf{L}(01020) \cdot w \qquad \varphi(2) = \mathbf{L}(01210) \cdot w$$
$$\varphi(1) = \mathbf{L}(02010) \cdot w \qquad \varphi(3) = \mathbf{L}(02120) \cdot w$$

is square-free.

Are there *many* square-free words? (2/3)

Theorem (Brandenburg, 1983; Brinkhuis, 1983; Kolpakov, 2006; Shur, 2009): $|W_3^n| \geqslant (5/4)^n$.

Proof that $|W_3^n| \geqslant 2^{n/72} \approx 1.01^n$: If $w = 1201021$, the **A_4-to-A_3 morphism** φ defined by

$$\varphi(0) = \mathbf{L}(01020) \cdot w \qquad \varphi(2) = \mathbf{L}(01210) \cdot w$$
$$\varphi(1) = \mathbf{L}(02010) \cdot w \qquad \varphi(3) = \mathbf{L}(02120) \cdot w$$

is square-free. Thus, there are at least 2^n square-free words $x \in W_4^n \subseteq A_4^n$.

Are there *many* square-free words? (2/3)

Theorem (Brandenburg, 1983; Brinkhuis, 1983; Kolpakov, 2006; Shur, 2009): $|W_3^n| \geqslant (5/4)^n$.

Proof that $|W_3^n| \geqslant 2^{n/72} \approx 1.01^n$: If $w = 1201021$, the **A_4-to-A_3 morphism** φ defined by

$$\varphi(0) = \mathbf{L}(01020) \cdot w \qquad \varphi(2) = \mathbf{L}(01210) \cdot w$$
$$\varphi(1) = \mathbf{L}(02010) \cdot w \qquad \varphi(3) = \mathbf{L}(02120) \cdot w$$

is square-free. Thus, there are at least 2^n square-free words $x \in W_4^n \subseteq A_4^n$,
at least 2^n square-free words $\varphi(x) \in W_3^{72n}$, and

$$|W_3^n|^{72} \geqslant |W_3^{72n}| \geqslant 2^n.$$

Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n - 1$.

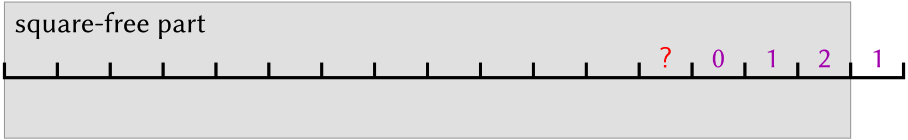
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(?012)|$



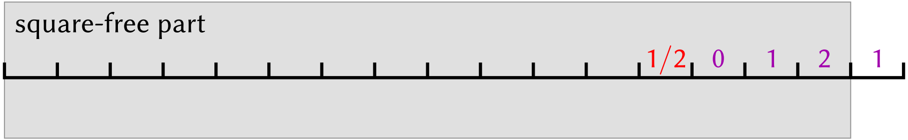
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4}$



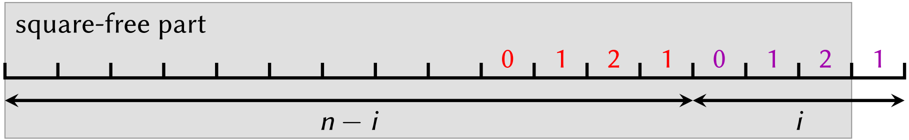
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$



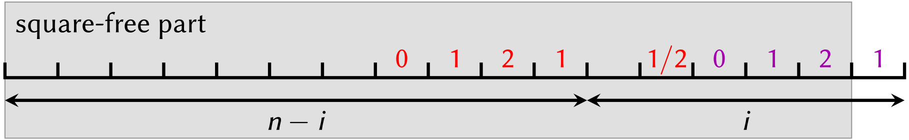
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$



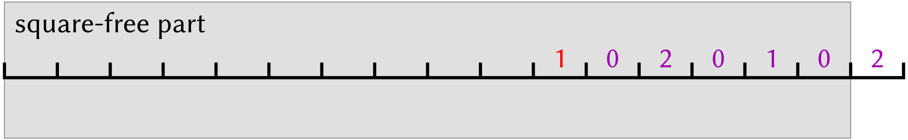
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$;
- $|W_3^n(020102)| \geqslant |W_3^{n-1}(102010)| + \mathbf{1}_{n=6}$.



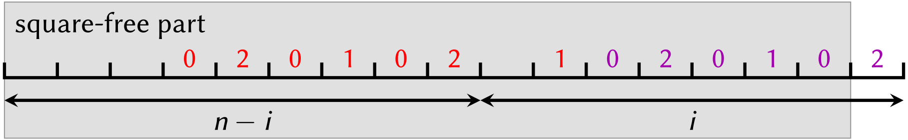
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$;
- $|W_3^n(020102)| \geqslant |W_3^{n-1}(102010)| + \mathbf{1}_{n=6} - \sum_{i \geqslant 6} |W_3^{n-i}(020102)|$.



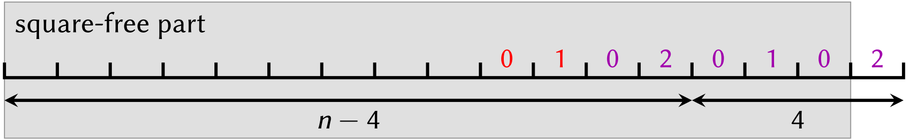
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$;
- $|W_3^n(020102)| \geqslant |W_3^{n-1}(102010)| + \mathbf{1}_{n=6} - \sum_{i \geqslant 6} |W_3^{n-i}(020102)| - |W_3^{n-4}(??0102)|$.



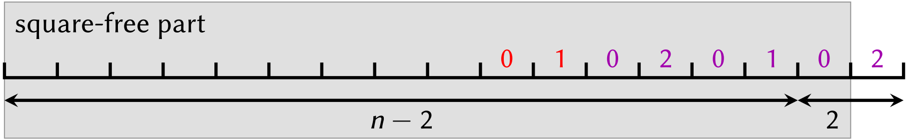
Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$;
- $|W_3^n(020102)| \geqslant |W_3^{n-1}(102010)| + \mathbf{1}_{n=6} - \sum_{i \geqslant 6} |W_3^{n-i}(020102)| - |W_3^{n-2}(010201)|$.



Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$;
- $|W_3^n(020102)| \geqslant |W_3^{n-1}(102010)| + \mathbf{1}_{n=6} - \sum_{i \geqslant 6} |W_3^{n-i}(020102)| - |W_3^{n-2}(010201)|$.

2. Group the terms $|W_3^n(w)|$ in a vector $\mathbf{W}_3^n$, and then in a power series $\mathbf{W}_3(z) = \sum_{n \geqslant 0} \mathbf{W}_3^n z^n$:
Inequalities about $|W_3^n(w)|$ rewrite as $\mathbf{M}(z)\,\mathbf{W}_3(z) \geqslant (1-z)z^k \mathbf{1}$, where $\mathbf{M}(z)$ is polynomial.

Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k): Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$;
- $|W_3^n(020102)| \geqslant |W_3^{n-1}(102010)| + \mathbf{1}_{n=6} - \sum_{i \geqslant 6} |W_3^{n-i}(020102)| - |W_3^{n-2}(010201)|$.

2. Group the terms $|W_3^n(w)|$ in a vector $\mathbf{W}_3^n$, and then in a power series $\mathbf{W}_3(z) = \sum_{n \geqslant 0} \mathbf{W}_3^n z^n$: Inequalities about $|W_3^n(w)|$ rewrite as $\mathbf{M}(z)\,\mathbf{W}_3(z) \geqslant (1-z)z^k\mathbf{1}$, where $\mathbf{M}(z)$ is polynomial.
3. For $k = 23$ and $z = 4/5$, there exists a non-zero vector $\mathbf{v} \geqslant 0$ such that $\mathbf{v}\mathbf{M}(z) \leqslant 0$, hence $\mathbf{v}\mathbf{M}(z)\mathbf{W}_3(z) \leqslant 0 < (1-z)z^k\mathbf{v}\mathbf{1}$, hence $\mathbf{W}_3(z)$ is **divergent** and $\limsup |W_3^n|^{1/n} \geqslant 5/4$.

Are there *many* square-free words? (3/3)

Finer analytic approach towards proving that $|W_3^n| \geqslant (5/4)^n$:

1. Let $W_3^n(w)$ be the set of words $x \in W_3^n$ that end with w, for all $w \in W_3^k$ (for a given k):
   Find linear inequalities that connect $|W_3^n(w)|$ to terms $|W_3^m(w')|$ when $m \leqslant n-1$.

Examples:

- $|W_3^n(0121)| \geqslant |W_3^{n-1}(1012)| + |W_3^{n-1}(2012)| + \mathbf{1}_{n=4} - \sum_{i \geqslant 4} |W_3^{n-i}(0102)|$;
- $|W_3^n(020102)| \geqslant |W_3^{n-1}(102010)| + \mathbf{1}_{n=6} - \sum_{i \geqslant 6} |W_3^{n-i}(020102)| - |W_3^{n-2}(010201)|$.

2. Group the terms $|W_3^n(w)|$ in a vector $\mathbf{W}_3^n$, and then in a power series $\mathbf{W}_3(z) = \sum_{n \geqslant 0} \mathbf{W}_3^n z^n$:
   Inequalities about $|W_3^n(w)|$ rewrite as $\mathbf{M}(z)\,\mathbf{W}_3(z) \geqslant (1-z)z^k\mathbf{1}$, where $\mathbf{M}(z)$ is polynomial.
3. For $k = 23$ and $z = 4/5$, there exists a non-zero vector $\mathbf{v} \geqslant 0$ such that $\mathbf{v}\mathbf{M}(z) \leqslant 0$, hence
   $\mathbf{v}\mathbf{M}(z)\mathbf{W}_3(z) \leqslant 0 < (1-z)z^k\mathbf{v}\mathbf{1}$, hence $\mathbf{W}_3(z)$ is **divergent** and $\limsup |W_3^n|^{1/n} \geqslant 5/4$.
4. The sequence $(\log |W_3^n|)_{n \geqslant 0}$ is sub-additive, hence $|W_3^n| \geqslant (5/4)^n$ for all $n \geqslant 0$.

Shuffles

A **shuffle** of two words u and v is a word w obtained by merging u and v from left to right, choosing the next symbol arbitrarily from u or from v: we write $w \in u \sqcup\!\sqcup v$.

Examples: 0123, 0213, 0231, 2013, 2031 and 2301 are the **shuffles** of 01 and 23;
0101 = 0101 and 0011 = 0011 = 0011 = 0011 are the **shuffles** of 01 and 01.

Counter-examples: 1023 and 0122 are not **shuffles** of 01 and 23.

Shuffles

A **shuffle** of two words u and v is a word w obtained by merging u and v from left to right, choosing the next symbol arbitrarily from u or from v: we write $w \in u \sqcup\!\sqcup v$.

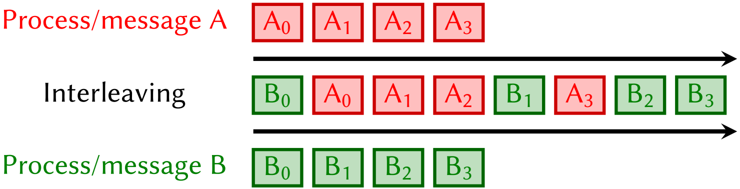
Examples: 0123, 0213, 0231, 2013, 2031 and 2301 are the **shuffles** of 01 and 23;
0101 = 0101 and 0011 = 0011 = 0011 = 0011 are the **shuffles** of 01 and 01.

Counter-examples: 1023 and 0122 are not **shuffles** of 01 and 23.

Playing cards

Single-thread concurrent execution

Message synchronisation



Shuffled squares and powers

A word u is a **shuffled square** if there exists a word v such that $u \in v \sqcup\!\sqcup v$.

Examples: 0101 and 0011

Counter-examples: 0110, 010101 and 000001111101

Shuffled squares and powers

A word u is a **shuffled square** if there exists a word v such that $u \in v \sqcup\!\sqcup v$;
shuffled k^{th} power if there exists a word v such that $u \in v \sqcup\!\sqcup \cdots \sqcup\!\sqcup v$ (with k vs);
shuffled power if u is a shuffled k^{th} power for some integer $k \geqslant 2$.

Examples: 0101 and 0011 are **shuffled squares**;
010101, 010011, 001101, 001011 and 000111 are **shuffled cubes**;
000001111101 is a **shuffled power**.

Counter-examples: 0110, 010101 and 000001111101 are not **shuffled squares**;
0101 and 010110, 000001111101 are not **shuffled cubes**;
0110, 010110 and 000001111110 are not **shuffled powers**.

Shuffled squares and powers

A word u is a **shuffled square** if there exists a word v such that $u \in v \sqcup\!\sqcup v$;
shuffled k^{th} power if there exists a word v such that $u \in v \sqcup\!\sqcup \cdots \sqcup\!\sqcup v$ (with k vs);
shuffled power if u is a shuffled k^{th} power for some integer $k \geqslant 2$.

Examples: 0101 and 0011 are **shuffled squares**;
010101, 010011, 001101, 001011 and 000111 are **shuffled cubes**;
000001111101 is a **shuffled power**.

Counter-examples: 0110, 010101 and 000001111101 are not **shuffled squares**;
0101 and 010110, 000001111101 are not **shuffled cubes**;
0110, 010110 and 000001111110 are not **shuffled powers**.

A word u is **shuffled-square-free** if none of its non-empty factors is a **shuffled square**.
shuffled-power-free if none of its non-empty factors is a **shuffled power**.

Examples: 010, 0102010 and 01202102012 are **shuffled-square-free**.
Counter-examples: 010212 and 2010201202 are not **shuffled-square-free**;
Ternary words u of length $|u| \geqslant 12$ are not **shuffled-square-free**.

Being $\square$-free vs being $⧢^2$-free

Lemma (Folklore)**:** One can check whether u is a $\square$ in time $\mathcal{O}(|u|)$.

Theorem (Buss & Soltys, 2013; Rizzi & Vialette, 2013; Bulteau & Vialette, 2020)**:**
Checking whether u is a $⧢^{\mathbf{2}}$ is NP-hard, even if $|A| = 2$.

Being $\square$-free vs being $⧢^2$-free

Lemma (Folklore): One can check whether u is a $\square$ in time $\mathcal{O}(|u|)$.

Theorem (Buss & Soltys, 2013; Rizzi & Vialette, 2013; Bulteau & Vialette, 2020):
Checking whether u is a $⧢^2$ is NP-hard, even if $|A| = 2$.

Theorem (Crochemore, 1983): If A is fixed, one can check whether u is **$\square$-free** in time $\mathcal{O}(|u|)$.

Theorem (Bulteau, Jugé & Vialette, 2023): Checking whether u is **$⧢^2$-free** is NP-hard.

Being $\square$-free vs being $\sqcup\!\sqcup^2$-free

Lemma (Folklore): One can check whether u is a $\square$ in time $\mathcal{O}(|u|)$.

Theorem (Buss & Soltys, 2013; Rizzi & Vialette, 2013; Bulteau & Vialette, 2020):
Checking whether u is a $\sqcup\!\sqcup^2$ is NP-hard, even if $|A| = 2$.

Theorem (Crochemore, 1983): If A is fixed, one can check whether u is **$\square$-free** in time $\mathcal{O}(|u|)$.

Theorem (Bulteau, Jugé & Vialette, 2023): Checking whether u is **$\sqcup\!\sqcup^2$-free** is NP-hard.

Lemma (Folklore): When k divides ℓ, every ℓ^{th} **power** is a k^{th} **power**.

Lemma (Folklore): 0102013231023123 is a **shuffled** 4^{th} **power** of 0123 but is **$\sqcup\!\sqcup^2$-free**.

Being $\square$-free vs being $\sqcup\!\sqcup^2$-free

Lemma (Folklore)**:** One can check whether u is a $\square$ in time $\mathcal{O}(|u|)$.

Theorem (Buss & Soltys, 2013; Rizzi & Vialette, 2013; Bulteau & Vialette, 2020)**:**
Checking whether u is a $\sqcup\!\sqcup^{\mathbf{2}}$ is NP-hard, even if $|A| = 2$.

Theorem (Crochemore, 1983)**:** If A is fixed, one can check whether u is **$\square$-free** in time $\mathcal{O}(|u|)$.

Theorem (Bulteau, Jugé & Vialette, 2023)**:** Checking whether u is **$\sqcup\!\sqcup^{\mathbf{2}}$-free** is NP-hard.

Lemma (Folklore)**:** When k divides ℓ, every ℓ^{th} **power** is a k^{th} **power**.

Lemma (Folklore)**:** 0102013231023123 is a **shuffled** 4^{th} **power** of 0123 but is **$\sqcup\!\sqcup^{\mathbf{2}}$-free**.

Theorem: There are infinite **$\square$-free** words in A^ω if and only if $|A| \geqslant 3$.

Open question: For which alphabets A are there infinite **$\sqcup\!\sqcup^{\mathbf{2}}$-free** words in A^ω?

When does A_ℓ^ω contain infinite $\sqcup\!\sqcup^2$-free words? (1/3)

Lemma: When $\ell \leqslant 3$, it does not.

Conjecture: When $\ell \geqslant 4$, it does.

When does A_ℓ^ω contain infinite $⧢^2$-free words? (1/3)

Lemma: When $\ell \leqslant 3$, it does not.

Conjecture: When $\ell \geqslant 4$, it does.

Theorem (Currie, 2014; Müller, 2015; Guégan & Ochem, 2016; Bulteau, Jugé & Vialette, 2023)**:**
When $\ell \geqslant 6$, it does.

Proof relying on an **analytic approach** when ℓ is large enough**:**

When does A_ℓ^ω contain infinite $⧢^2$-free words? (1/3)

Lemma: When $\ell \leqslant 3$, it does not.

Conjecture: When $\ell \geqslant 4$, it does.

Theorem (Currie, 2014; Müller, 2015; Guégan & Ochem, 2016; Bulteau, Jugé & Vialette, 2023)**:**
When $\ell \geqslant 6$, it does.

Proof relying on an **analytic approach** when ℓ is large enough**:**

1. For all $⧢^2$-free words $w \in A_\ell^n$ and all letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ is $⧢^2$-free or has a factorisation $w \cdot \lambda = x \cdot y$, where x is $⧢^2$-free and $y \in u ⧢ u$ for some word u.

When does A_ℓ^ω contain infinite $⧢^2$-free words? (1/3)

Lemma: When $\ell \leqslant 3$, it does not.

Conjecture: When $\ell \geqslant 4$, it does.

Theorem (Currie, 2014; Müller, 2015; Guégan & Ochem, 2016; Bulteau, Jugé & Vialette, 2023)**:**
When $\ell \geqslant 6$, it does.

Proof relying on an **analytic approach** when ℓ is large enough**:**

1. For all $⧢^2$-free words $w \in A_\ell^n$ and all letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ is $⧢^2$-free or has a factorisation $w \cdot \lambda = x \cdot y$, where x is $⧢^2$-free and $y \in u ⧢ u$ for some word u.
2. Let W_ℓ^n be the set of $⧢^2$-free words in A_ℓ^n. Once the length $k = |u|$ is fixed, there are $|W_\ell^{n+1-2k}|$ choices for x, ℓ^k choices for u and up to 2^{2k} choices for y:

$$|W_\ell^n|\,\ell \leqslant |W_\ell^{n+1}| + |W_\ell^{n-1}|\,(4\ell) + \cdots + |W_\ell^{n+1-2k}|\,(4\ell)^k + \cdots$$

When does A_ℓ^ω contain infinite $\sqcup\!\sqcup^2$-free words? (1/3)

Lemma: When $\ell \leqslant 3$, it does not.

Conjecture: When $\ell \geqslant 4$, it does.

Theorem (Currie, 2014; Müller, 2015; Guégan & Ochem, 2016; Bulteau, Jugé & Vialette, 2023)**:**
When $\ell \geqslant 6$, it does.

Proof relying on an **analytic approach** when ℓ is large enough**:**

1. For all $\sqcup\!\sqcup^2$-free words $w \in A_\ell^n$ and all letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ is $\sqcup\!\sqcup^2$-free or has a factorisation $w \cdot \lambda = x \cdot y$, where x is $\sqcup\!\sqcup^2$-free and $y \in u \sqcup\!\sqcup u$ for some word u.
2. Let W_ℓ^n be the set of $\sqcup\!\sqcup^2$-free words in A_ℓ^n. Once the length $k = |u|$ is fixed, there are $|W_\ell^{n+1-2k}|$ choices for x, ℓ^k choices for u and up to 2^{2k} choices for y:
$$|W_\ell^n|\,\ell \leqslant |W_\ell^{n+1}| + |W_\ell^{n-1}|\,(4\ell) + \cdots + |W_\ell^{n+1-2k}|\,(4\ell)^k + \cdots$$
3. We prove by induction that $|W_\ell^{n+1}| \geqslant 2\ell|W_\ell^n|/3$ when $\ell \geqslant 27$, so that $W_\ell^n \neq \emptyset$:
$$\frac{|W_\ell^{n+1}|}{|W_\ell^n|} \geqslant \ell - \sum_{k \geqslant 1} \frac{|W_\ell^{n+1-2k}|}{|W_\ell^n|}(4\ell)^k \geqslant \ell - \sum_{k \geqslant 1} \frac{3^{2k-1}}{(2\ell)^{2k-1}}(4\ell)^k = \frac{2\ell}{3} + \frac{\ell(\ell-27)}{3\ell-27} \geqslant \frac{2\ell}{3}.$$

When does A_ℓ^ω contain infinite $⧢^2$-free words? (2/3)

Theorem: When $\ell \geqslant 6$, it does.

Proof relying on a **finer analytic approach** when ℓ is not so large:

1. For all $⧢^2$-free words $w \in A_\ell^n$ and all letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ is $⧢^2$-free or has a factorisation $w \cdot \lambda = x \cdot y$, where x is $⧢^2$-free and $y \in u ⧢ u$ for some word u. Moreover,

When does A_ℓ^ω contain infinite $⧢^2$-free words? (2/3)

Theorem: When $\ell \geqslant 6$, it does.

Proof relying on a **finer analytic approach** when ℓ is not so large:

1. For all $⧢^2$-free words $w \in A_\ell^n$ and all letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ is $⧢^2$-free or has a factorisation $w \cdot \lambda = x \cdot y$, where x is $⧢^2$-free and $y \in u ⧢ u$ for some word u. Moreover,
   - either $u = \lambda$ or $|u| \geqslant 2$ and $x \cdot y_0 \cdot y_1$ is $⧢^2$-free;
   - the suffix y can be chosen **minimal**: no strict factor of y is a $⧢^2$.

When does A_ℓ^ω contain infinite $\sqcup\!\sqcup^2$-free words? (2/3)

Theorem: When $\ell \geqslant 6$, it does.

Proof relying on a **finer analytic approach** when ℓ is not so large:

1. For all $\sqcup\!\sqcup^2$-free words $w \in A_\ell^n$ and all letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ is $\sqcup\!\sqcup^2$-free or has a factorisation $w \cdot \lambda = x \cdot y$, where x is $\sqcup\!\sqcup^2$-free and $y \in u \sqcup\!\sqcup u$ for some word u. Moreover,
   - either $u = \lambda$ or $|u| \geqslant 2$ and $x \cdot y_0 \cdot y_1$ is $\sqcup\!\sqcup^2$-free;
   - the suffix y can be chosen **minimal**: no strict factor of y is a $\sqcup\!\sqcup^2$.
2. Let S_ℓ^k be the set of minimal $\sqcup\!\sqcup^2$ of length $2k$: once the length $k = |u|$ and the word $y \in S_\ell^k$ are fixed, there are $|W_\ell^{n+3-2k}|/\ell(\ell-1)$ choices for $x \cdot y_0 \cdot y_1$:

$$|W_\ell^n|\,\ell \leqslant |W_\ell^{n+1}| + |W_\ell^n| + \frac{|W_\ell^{n-1}|\,|S_\ell^2| + |W_\ell^{n-3}|\,|S_\ell^3| + \cdots + |W_\ell^{n+3-2k}|\,|S_\ell^k| + \cdots}{\ell(\ell-1)}$$

When does A_ℓ^ω contain infinite $⧢^2$-free words? (2/3)

Theorem: When $\ell \geqslant 6$, it does.

Proof relying on a **finer analytic approach** when ℓ is not so large:

1. For all $⧢^2$-free words $w \in A_\ell^n$ and all letters $\lambda \in A_\ell$, the word $w \cdot \lambda$ is $⧢^2$-free or has a factorisation $w \cdot \lambda = x \cdot y$, where x is $⧢^2$-free and $y \in u ⧢ u$ for some word u. Moreover,
   - either $u = \lambda$ or $|u| \geqslant 2$ and $x \cdot y_0 \cdot y_1$ is $⧢^2$-free;
   - the suffix y can be chosen **minimal**: no strict factor of y is a $⧢^2$.
2. Let S_ℓ^k be the set of minimal $⧢^2$ of length $2k$: once the length $k = |u|$ and the word $y \in S_\ell^k$ are fixed, there are $|W_\ell^{n+3-2k}|/\ell(\ell-1)$ choices for $x \cdot y_0 \cdot y_1$:

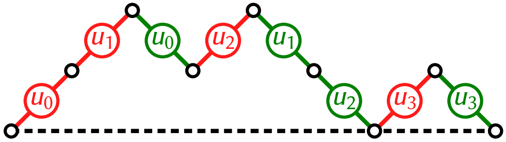
$$|W_\ell^n|\,\ell \leqslant |W_\ell^{n+1}| + |W_\ell^n| + \frac{|W_\ell^{n-1}|\,|S_\ell^2| + |W_\ell^{n-3}|\,|S_\ell^3| + \cdots + |W_\ell^{n+3-2k}|\,|S_\ell^k| + \cdots}{\ell(\ell-1)}$$

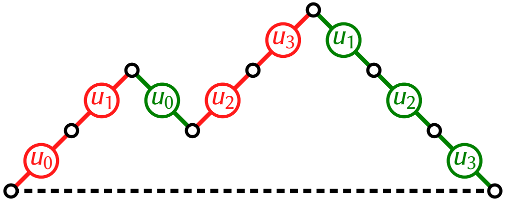
3. We can prove that $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.
4. We can prove by induction that $|W_\ell^{n+1}| \geqslant 37\sqrt{\ell-1}\,|W_\ell^n|/16$ when $\ell \geqslant 7$, so that $W_\ell^n \neq \emptyset$.

Intermezzo: counting minimal $⧢^2$ words and $⧢^2$-free words

When choosing a word $y \in u ⧢ u$ whose strict factors are $⧢^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;

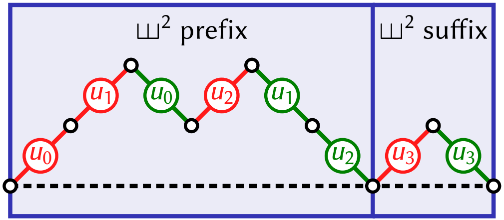


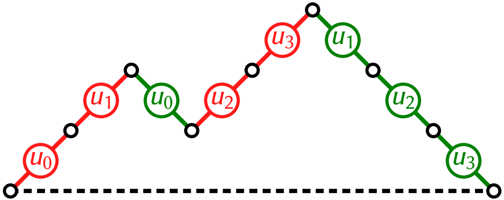


Intermezzo: counting minimal $\sqcup\!\sqcup^2$ words and $\sqcup\!\sqcup^2$-free words

When choosing a word $y \in u \sqcup\!\sqcup u$ whose strict factors are $\sqcup\!\sqcup^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;

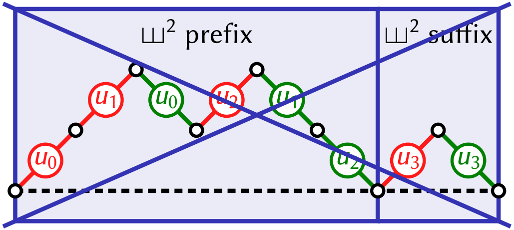


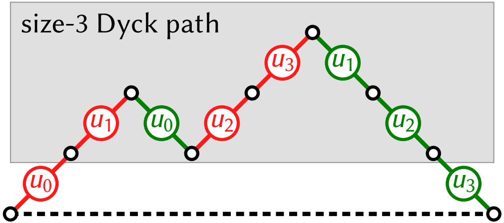


Intermezzo: counting minimal $\sqcup\!\sqcup^2$ words and $\sqcup\!\sqcup^2$-free words

When choosing a word $y \in u \sqcup\!\sqcup u$ whose strict factors are $\sqcup\!\sqcup^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;

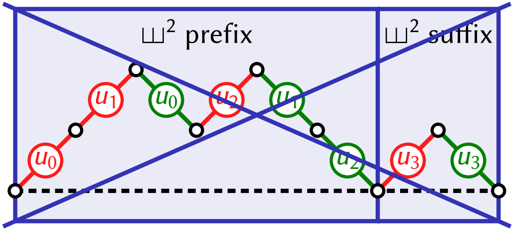


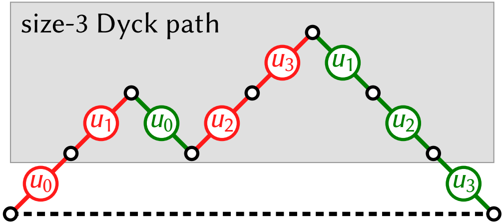


Intermezzo: counting minimal $\shuffle^2$ words and $\shuffle^2$-free words

When choosing a word $y \in u \shuffle u$ whose strict factors are $\shuffle^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

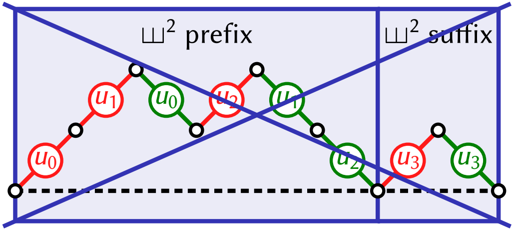


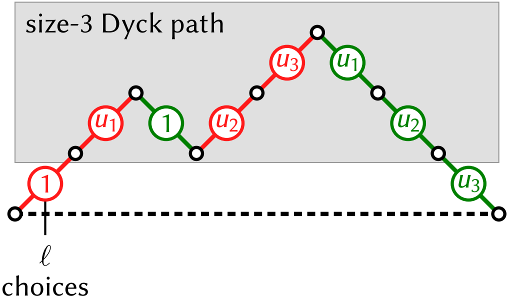


Intermezzo: counting minimal $ш^2$ words and $ш^2$-free words

When choosing a word $y \in u \sqcup\!\sqcup u$ whose strict factors are $ш^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S^k_\ell| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

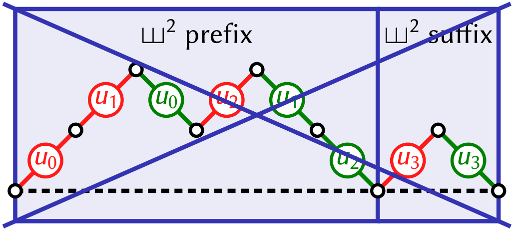


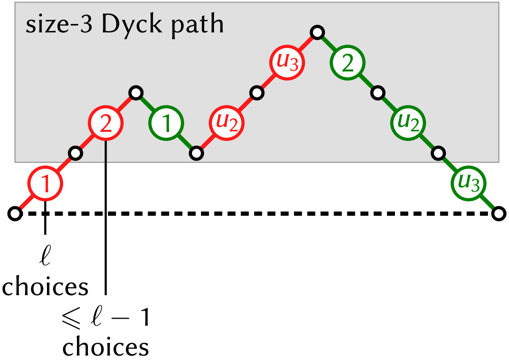


Intermezzo: counting minimal $⧢^2$ words and $⧢^2$-free words

When choosing a word $y \in u ⧢ u$ whose strict factors are $⧢^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

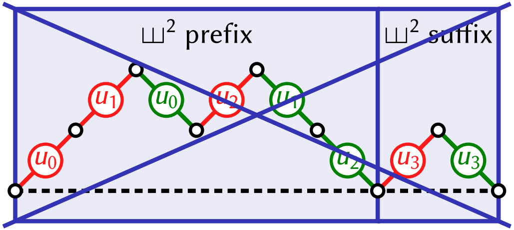


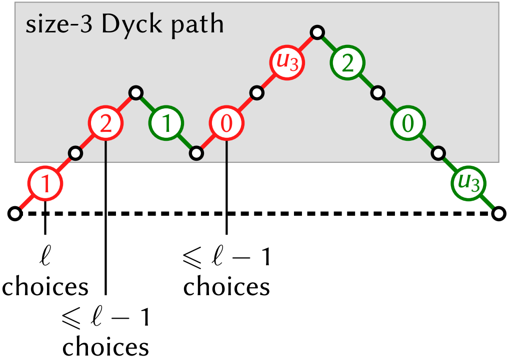


Intermezzo: counting minimal $⧢^2$ words and $⧢^2$-free words

When choosing a word $y \in u ⧢ u$ whose strict factors are $⧢^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

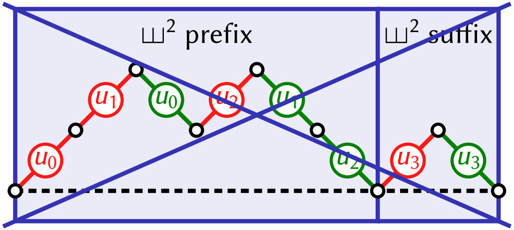


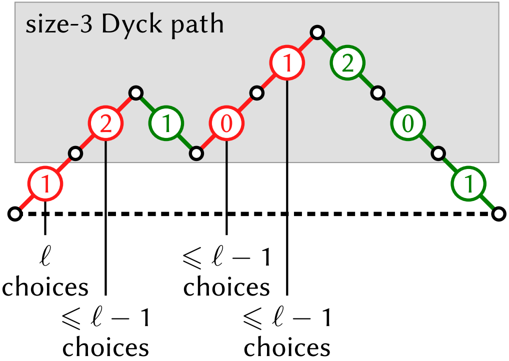


Intermezzo: counting minimal $\sqcup\!\sqcup^2$ words and $\sqcup\!\sqcup^2$-free words

When choosing a word $y \in u \sqcup\!\sqcup u$ whose strict factors are $\sqcup\!\sqcup^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.



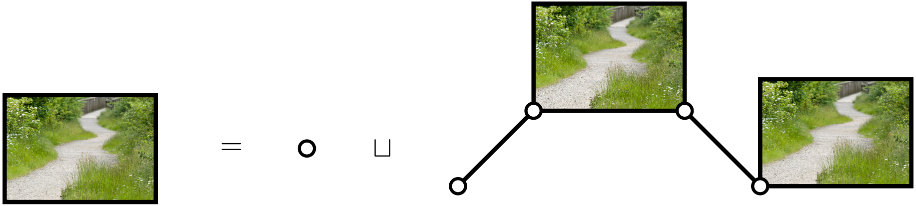


Intermezzo: counting minimal $\sqcup\!\sqcup^2$ words and $\sqcup\!\sqcup^2$-free words

When choosing a word $y \in u \sqcup\!\sqcup u$ whose strict factors are $\sqcup\!\sqcup^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

The generating series $\mathcal{C}(z) = \sum_{k \geqslant 0} \mathrm{Cat}(k) z^k$ satisfies the equation $\mathcal{C}(z) = 1 + z\mathcal{C}(z)^2$.

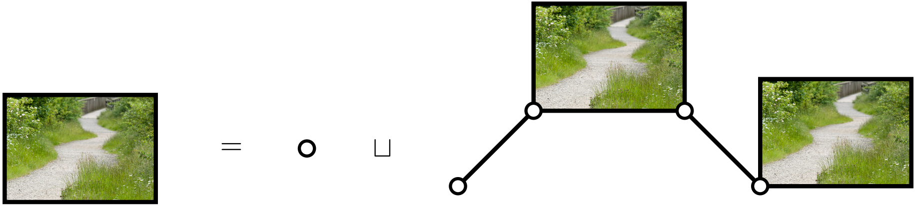


Intermezzo: counting minimal $\sqcup\!\sqcup^2$ words and $\sqcup\!\sqcup^2$-free words

When choosing a word $y \in u \sqcup\!\sqcup u$ whose strict factors are $\sqcup\!\sqcup^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

The generating series $\mathcal{C}(z) = \sum_{k \geqslant 0} \mathrm{Cat}(k) z^k$ satisfies the equation $\mathcal{C}(z) = 1 + z\mathcal{C}(z)^2$.



$$\mathcal{C}(z) \quad = \quad 1 \quad + \quad z \times \mathcal{C}(z) \times 1 \times \mathcal{C}(z)$$

Intermezzo: counting minimal $⧢^2$ words and $⧢^2$-free words

When choosing a word $y \in u ⧢ u$ whose strict factors are $⧢^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

The generating series $\mathcal{C}(z) = \sum_{k \geqslant 0} \mathrm{Cat}(k) z^k$ satisfies the equation $\mathcal{C}(z) = 1 + z\mathcal{C}(z)^2$, hence

$$\mathcal{C}(z) = \frac{2}{1+\sqrt{1-4z}} \text{ when } |z| < \frac{1}{4}.$$

Intermezzo: counting minimal $⧢^2$ words and $⧢^2$-free words

When choosing a word $y \in u ⧢ u$ whose strict factors are $⧢^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

The generating series $\mathcal{C}(z) = \sum_{k \geqslant 0} \mathrm{Cat}(k) z^k$ satisfies the equation $\mathcal{C}(z) = 1 + z\mathcal{C}(z)^2$, hence

$$\mathcal{C}(z) = \frac{2}{1+\sqrt{1-4z}} \text{ when } |z| < \frac{1}{4}.$$

We prove now by induction that $|W_\ell^{n+1}| \geqslant \lambda |W_\ell^n|$ when $\ell \geqslant 7$, where $\lambda = 37\sqrt{\ell-1}/16 \geqslant 5.65$:

$$\frac{|W_\ell^{n+1}|}{|W_\ell^n|} \geqslant (\ell-1) - \frac{1}{\ell(\ell-1)} \sum_{k \geqslant 2} \lambda^{3-2k} |S_\ell^k| \geqslant (\ell-1) - \frac{\lambda}{\ell-1}\left(\mathcal{C}\left(\frac{\ell-1}{\lambda^2}\right) - 1\right)$$

Intermezzo: counting minimal $\sqcup\!\sqcup^2$ words and $\sqcup\!\sqcup^2$-free words

When choosing a word $y \in u \sqcup\!\sqcup u$ whose strict factors are $\sqcup\!\sqcup^2$-free, with $|u| = k$:

- each letter u_{i+1} precedes the letter u_i: $\mathrm{Cat}(k-1) = \frac{1}{k}\binom{2k-2}{k-1}$ interleavings are possible;
- each letter u_{i+1} must be distinct from its predecessor in y: given an interleaving, there are up to $\ell(\ell-1)^{k-1}$ choices for u. Hence, $|S_\ell^k| \leqslant \ell(\ell-1)^{k-1}\,\mathrm{Cat}(k-1)$.

The generating series $\mathcal{C}(z) = \sum_{k \geqslant 0} \mathrm{Cat}(k) z^k$ satisfies the equation $\mathcal{C}(z) = 1 + z\mathcal{C}(z)^2$, hence

$$\mathcal{C}(z) = \frac{2}{1+\sqrt{1-4z}} \text{ when } |z| < \frac{1}{4}.$$

We prove now by induction that $|W_\ell^{n+1}| \geqslant \lambda |W_\ell^n|$ when $\ell \geqslant 7$, where $\lambda = 37\sqrt{\ell-1}/16 \geqslant 5.65$:

$$\frac{|W_\ell^{n+1}|}{|W_\ell^n|} \geqslant (\ell-1) - \frac{1}{\ell(\ell-1)} \sum_{k \geqslant 2} \lambda^{3-2k} |S_\ell^k| \geqslant (\ell-1) - \frac{\lambda}{\ell-1}\left(\mathcal{C}\left(\frac{\ell-1}{\lambda^2}\right) - 1\right) = \lambda + \frac{P(\lambda)}{\lambda},$$

where P is an explicit polynomial such that $P(x) > 0$ whenever $x \geqslant 5.65$.

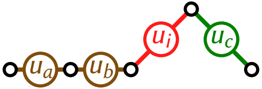
When does A_ℓ^ω contain infinite $\sqcup\!\sqcup^2$-free words? (3/3)

Theorem: When $\ell \geqslant 6$, it does.

Proof relying on **finer upper bounds** on $|S_\ell^k|$ when $\ell = 6$:

When a letter u_i is placed just before a peak, it can take at most $\ell - 2$ values (unless $i = 2$):

- if $u_b = u_c$, then $u_i \neq u_a$ and $u_i \neq u_b$;
- if $u_b \neq u_c$, then $u_i \neq u_b$ and $u_i \neq u_c$.



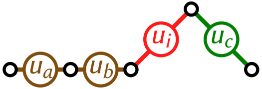
$a, b, c < i$

When does A_ℓ^ω contain infinite $\sqcup\!\sqcup^2$-free words? (3/3)

Theorem: When $\ell \geqslant 6$, it does.
Proof relying on **finer upper bounds** on $|S_\ell^k|$ when $\ell = 6$:

When a letter u_i is placed just before a peak, it can take at most $\ell - 2$ values (unless $i = 2$):

- if $u_b = u_c$, then $u_i \neq u_a$ and $u_i \neq u_b$;
- if $u_b \neq u_c$, then $u_i \neq u_b$ and $u_i \neq u_c$.



$a, b, c < i$

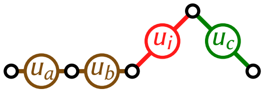
Remark (Narayana, 1959): There are $\mathrm{Cat}(k,p) = \frac{1}{k}\binom{k}{p}\binom{k}{p-1}$ size-k Dyck trees with p peaks.

When does A_ℓ^ω contain infinite $\sqcup^2$-free words? (3/3)

Theorem: When $\ell \geqslant 6$, it does.
Proof relying on **finer upper bounds** on $|S_\ell^k|$ when $\ell = 6$:

When a letter u_i is placed just before a peak, it can take at most $\ell - 2$ values (unless $i = 2$):

- if $u_b = u_c$, then $u_i \neq u_a$ and $u_i \neq u_b$;
- if $u_b \neq u_c$, then $u_i \neq u_b$ and $u_i \neq u_c$.



$a, b, c < i$

Remark (Narayana, 1959): There are $\mathrm{Cat}(k,p) = \frac{1}{k}\binom{k}{p}\binom{k}{p-1}$ size-k Dyck trees with p peaks.

Lemma: The generating series $\mathcal{C}(x,y) = \sum_{k,p \geqslant 0} \mathrm{Cat}(k,p) x^k y^p$ satisfies the relation

$$\mathcal{C}(x,y) = 1 + xy\mathcal{C}(x,y) + x(\mathcal{C}(x,y) - 1)\mathcal{C}(x,y).$$

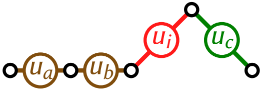
When does A_ℓ^ω contain infinite $\sqcup\!\sqcup^2$-free words? (3/3)

Theorem: When $\ell \geqslant 6$, it does.

Proof relying on **finer upper bounds** on $|S_\ell^k|$ when $\ell = 6$:

When a letter u_i is placed just before a peak, it can take at most $\ell - 2$ values (unless $i = 2$):

- if $u_b = u_c$, then $u_i \neq u_a$ and $u_i \neq u_b$;
- if $u_b \neq u_c$, then $u_i \neq u_b$ and $u_i \neq u_c$.



$a, b, c < i$

Remark (Narayana, 1959): There are $\mathrm{Cat}(k,p) = \frac{1}{k}\binom{k}{p}\binom{k}{p-1}$ size-k Dyck trees with p peaks.

Lemma: The generating series $\mathcal{C}(x,y) = \sum_{k,p\geqslant 0} \mathrm{Cat}(k,p) x^k y^p$ satisfies the relation

$$\mathcal{C}(x,y) = 1 + xy\mathcal{C}(x,y) + x(\mathcal{C}(x,y) - 1)\mathcal{C}(x,y).$$

Corollary: We can prove that $|W_\ell^{n+1}| \geqslant \lambda |W_\ell^n|$, where $\lambda \approx 4.56594$ is the largest root of

$$P(X) = X^5 - 5X^4 + 18X^3 - 210X^2 + 626X - 5.$$

Open questions

1. When A is fixed, how hard is recognising $\sqcup\!\sqcup^2$-**free** words?

Open questions

1. When A is fixed, how hard is recognising **$\sqcup\!\sqcup^2$-free** words?
2. Can we extend our analytic proof to the cases $|A| = 4$? $|A| = 5$?

Open questions

1. When A is fixed, how hard is recognising **$\shuffle^2$-free** words?
2. Can we extend our analytic proof to the cases $|A| = 4$? $|A| = 5$?
3. Can we decide whether a given word is a factor of an infinite **$\shuffle^2$-free** word?

Open questions

1. When A is fixed, how hard is recognising $\sqcup\!\sqcup^2$**-free** words?
2. Can we extend our analytic proof to the cases $|A| = 4$? $|A| = 5$?
3. Can we decide whether a given word is a factor of an infinite $\sqcup\!\sqcup^2$**-free** word?
4. A morphism of monoids $\varphi\colon A^* \to B^*$ is $\square$**-free** if $\varphi(x)$ is $\square$**-free** whenever x is $\square$**-free**.
   Can we meaningfully generalise this notion to $\sqcup\!\sqcup^2$**-free** morphisms?

Open questions

1. When A is fixed, how hard is recognising **$\sqcup\!\sqcup^2$-free** words?
2. Can we extend our analytic proof to the cases $|A| = 4$? $|A| = 5$?
3. Can we decide whether a given word is a factor of an infinite **$\sqcup\!\sqcup^2$-free** word?
4. A morphism of monoids $\varphi\colon A^* \to B^*$ is **$\square$-free** if $\varphi(x)$ is **$\square$-free** whenever x is **$\square$-free**.
   Can we meaningfully generalise this notion to **$\sqcup\!\sqcup^2$-free** morphisms?
5. What about **shuffled-cube-free** or **shuffled-power-free** words?
   Does there exist an infinite **shuffled-power-free** word on a finite alphabet? in A_4^ω?

Open questions

1. When A is fixed, how hard is recognising **$\sqcup\!\sqcup^2$-free** words?
2. Can we extend our analytic proof to the cases $|A| = 4$? $|A| = 5$?
3. Can we decide whether a given word is a factor of an infinite **$\sqcup\!\sqcup^2$-free** word?
4. A morphism of monoids $\varphi \colon A^* \to B^*$ is **$\square$-free** if $\varphi(x)$ is **$\square$-free** whenever x is **$\square$-free**.
   Can we meaningfully generalise this notion to **$\sqcup\!\sqcup^2$-free** morphisms?
5. What about **shuffled-cube-free** or **shuffled-power-free** words?
   Does there exist an infinite **shuffled-power-free** word on a finite alphabet? in A_4^ω?

QTUHEASNTKIYOONUS!?

Bibliography

Axel Thue, *Über unendliche Zeichenreihen* 1906

Axel Thue, *Über die gegenseitige Lage gleicher teile gewisser Zeichenreihen* 1912

John Leech, *A problem on strings of beads* 1957

Tadepalli Venkata Narayana, *A partial order and its applications to probability theory* 1959

Maxime Crochemore, *Sharp characterizations of square-free morphisms* 1982

Franz-Josef Brandenburg, *Uniformly growing k^{th} power-free homomorphisms* 1983

Jan Brinkhuis, *Non-repetitive sequences on three symbols* 1983

Maxime Crochemore, *Recherche linéaire d'un carré dans un mot* 1983

Roman Kolpakov, *Efficient lower bounds on the number of repetition-free words* 2007

Arseny Shur, *Two-sided bounds for the growth rates of power-free languages* 2009

Sam Buss & Michael Soltys, *Unshuffling a square is NP-hard* 2013

Romeo Rizzi & Stéphane Vialette, *On recognizing words that are squares for the shuffle product* 2013

James Currie, *Shuffle squares are avoidable* (unpublished manuscript) 2014

Mike Müller, *Avoiding and enforcing repetitive structures in words* 2015

Guillaume Guégan & Pascal Ochem, *A short proof that shuffle squares are 7-avoidable* 2016

Laurent Bulteau & Stéphane Vialette, *Recognizing binary shuffle squares is NP-hard* 2020

Laurent Bulteau, Vincent Jugé & Stéphane Vialette, *On shuffled-square-free words* 2023

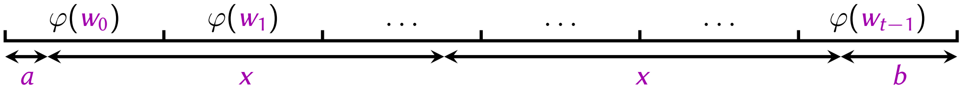
Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Is this morphism square-free?

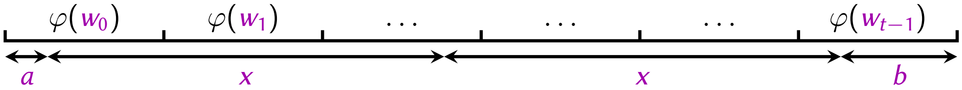
Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.



Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.
Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.
Claim #0: φ is injective.



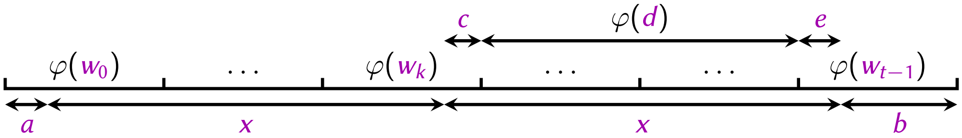
Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.



Is this morphism square-free?

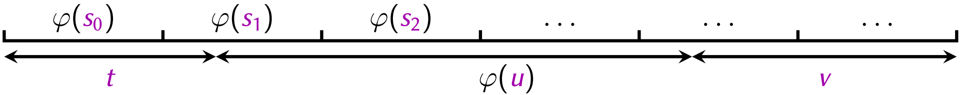
Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.



Is this morphism square-free?

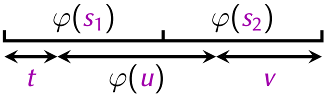
Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.



Is this morphism square-free?

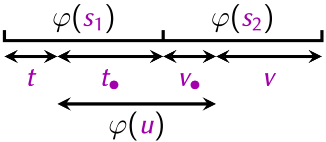
Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.



Is this morphism square-free?

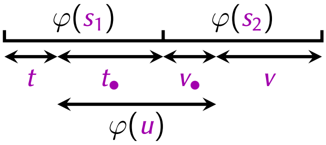
Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

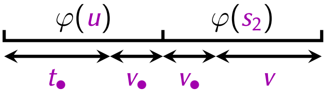
Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.



$u = s_2$, because



Is this morphism square-free?

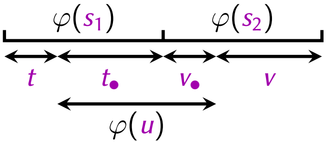
Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

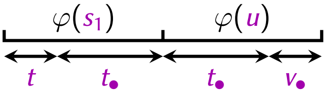
Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.



$u = s_1$, because



Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

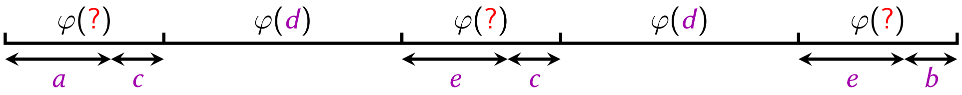
Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.

Conclusion: w is not minimal!



Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

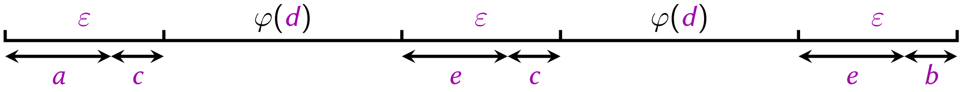
Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\bar{t}$ and $\bar{v}$ such that $t = \varphi(\bar{t})$ and $v = \varphi(\bar{v})$.

Conclusion: w is not minimal!



Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

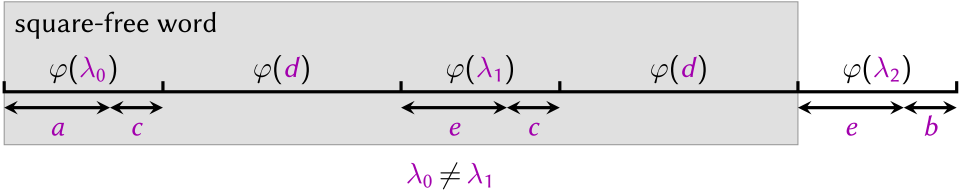
Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.

Conclusion: w is not minimal!



Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

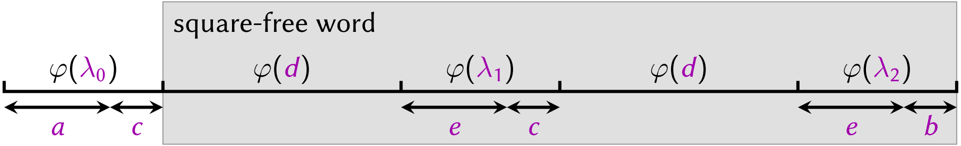
Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\bar{t}$ and $\bar{v}$ such that $t = \varphi(\bar{t})$ and $v = \varphi(\bar{v})$.

Conclusion: w is not minimal!



$\lambda_0 \neq \lambda_1 \neq \lambda_2$

Is this morphism square-free?

Theorem (Thue, 1912; Crochemore, 1982): An **ℓ-uniform** morphism φ is **square-free** if $\varphi(w)$ is square-free whenever $|w| \leqslant 3$ and w is square-free.

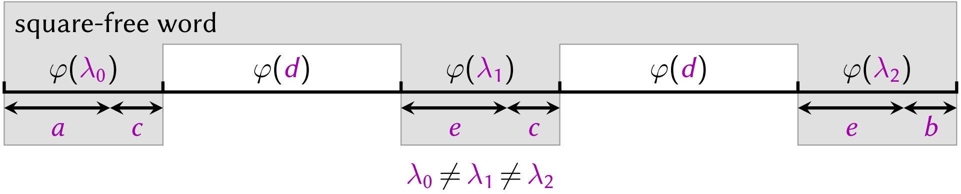
Proof: Assume that there is a shortest square-free word $w = w_0 w_1 \cdots w_{t-1}$, of length $|w| \geqslant 4$, for which $\varphi(w)$ has a factorisation $\varphi(w) = a \cdot x^2 \cdot b$ where $0 \leqslant |a| < \ell$, $x \neq \varepsilon$ and $0 \leqslant |b| < \ell$.

Claim #0: φ is injective.

Claim #1: x has a factorisation $x = c \cdot \varphi(d) \cdot e$, where $0 \leqslant |c| < \ell$, $d \neq \varepsilon$ and $0 \leqslant |e| < \ell$.

Claim #2: in each factorisation of the form $\varphi(s) = t \cdot \varphi(u) \cdot v$ where s is square-free and $u \neq \varepsilon$, there exist words $\overline{t}$ and $\overline{v}$ such that $t = \varphi(\overline{t})$ and $v = \varphi(\overline{v})$.

Conclusion: w is not minimal!



The vector **v**

By identifing isomorphic words w to let them start with 01 and ordering coordinates w of $\mathbf{W}_3^n$ in lexicographic order, we can choose **v** as the inverse of

770	980	732	826	744	640	662	790	580	1380	892	719	1164	1226	688	1198	962	926	1154	600	1148	530	694	1346	598	856	920	1000	1120	678	∞	∞	614
1208	666	702	1116	1258	814	946	602	780	1236	686	1564	932	1208	956	740	820	692	572	988	1264	798	∞	∞	1122	936	802	1006	1494	632	580	784	934
986	564	732	580	1042	954	1578	∞	826	1034	1188	520	946	1206	676	846	1174	1110	∞	952	752	934	706	702	692	1790	1070	744	1140	616	700	1184	716
0	∞	912	720	526	636	794	982	540	640	1146	680	754	538	1042	918	860	672	896	∞	∞	∞	694	1004	1126	698	712	808	712	1012	600	1068	1120
748	986	696	1188	1074	668	904	1198	554	974	1170	956	1140	1736	740	1244	684	589	1098	896	768	1028	766	814	682	873	624	1996	880	1626	870	828	∞
536	810	930	569	720	∞	607	738	892	∞	758	658	1228	922	722	904	718	660	810	658	756	772	808	980	878	568	1110	776	2822	940	734	576	934
854	1086	902	∞	816	585	758	834	∞	546	938	716	∞	644	896	764	572	764	1072	604	666	∞	872	1572	1022	726	1552	1094	874	1230	680	684	726
688	∞	1072	642	742	946	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	∞	553	1550	778	∞	872	750	∞	538	790
1030	598	866	774	784	856	1550	∞	626	938	740	710	928	∞	744	1020	846	∞	870	994	1072	540	764	752	572	683	874	712	1382	590	814	1158	1492
822	950	758	1670	1012	626	574	498	∞	944	562	728	2628	860	922	1354	620	720	3068	1164	1066	706	564	714	950	∞	956	1092	714	724	1760	894	866
1040	728	938	716	∞	660	564	980	878	604	1144	770	1542	1160	730	696	838	896	2144	840	698	725	1160	1544	734	1144	619	846	980	576	648	552	738
938	728	1118	856	894	1762	694	∞	1096	956	∞	952	714	560	702	1068	1164	1776	720	624	1542	912	860	2182	728	566	942	1548	502	586	626	1086	1642
760	952	822	∞	∞	∞	∞	∞	∞	764	540	1088	978	888	∞	832	1020	742	700	928	732	740	906	626	∞	1554	858	784	802	868	580	1056	782
538	∞	764	872	∞	804	1554	553	574	878	1310	964	530	694	1344	582	2008	922	966	1130	678	696	838	∞	586	798	946	742	642	∞	∞	628	722
683	676	1236	872	1092	1546	752	1020	1564	870	∞	678	588	1128	754	572	∞	990	674	862	920	1042	538	748	680	1044	646	540	982	820	618	526	730
876	∞	1552	714	1182	710	846	676	1196	950	524	1092	1030	828	∞	∞	954	1042	576	732	560	1016	934	828	584	632	1492	1004	836	936	1146	676	1212
790	1316	984	582	676	∞	740	956	1200	982	1562	686	1148	800	606	944	862	1256	1154	672	665	1200	608	1094	874	678	1138	970	920	858	584	1346	694
538	1070	616	1154	898	1018	1198	702	1224	1082	720	892	1386	576	790	674	646	758	830	732	978	593	738	554	740	∞	848	550	2086	626	1574	864	1374
958	796	902	828	544	1146	729	1064	896	874	692	1198	730	589	786	1270	1490	1116	738	565	1688	762	1022	668	1050	708	806	664	1246	950	∞	1362	694
564	726	1276	982	1048	716	1142	1132	886	1194	900	563	1064	1098	650	696	∞	1074	738	502	880	2572	760	998	886	1056	538	738	∞	914	1594	752	582
1564	860	798	1006	776	762	946	564	874	692	572	792	590	683	1706	1490	1236	774	1388	540	852	824	∞	1064	960	698	926	756	1018	688	886	798	794
584	864	892	1088	982	1228	774	1204	910	866	1156	572	978	570	694	1344	826	∞	696	566	842	567	∞	926	928	1224	678	872	1092	1568	862	1220	670
642	1066	2038	756	944	618	732	1046	714	∞	∞	1012	748	572	658	575	778	622	∞	870	726	1156	1080	896	954	852	609	524	938	764	976	613	714
1110	724	4094	938	568	822	978	1220	641	808	740	924	718	732	846	798	928	∞	656	920	1224	986	696	678	1108	1122	554	∞	1200	956	1254	914	840
1118	1784	942	760	968	738	497	808	1010	710	1118	1000	808	580	712	668	694	∞	∞	1128	968	548	∞	724	∞	860	824	738	580	1072	940	536	830
1384	884	888	1594	2010	626	950	∞	840	978	1174	928	746	612	738	1134	558	∞	1070	690	578	640	497	1050	1004	758	881	1604	890	954	918	946	696
0	894	828	966	716	538	1042	678	1358	1546	1166	714	750	868	526	732	632	1042	1044	∞	864	974	960	908	606	1060	850	516	932	678	848	1154	984
890	1210	722	930	1076	1024	754	1078	742	575	636	∞																					

Missing computational details

1. For all $k \geqslant 2$, $|S_k| \geqslant \Sigma_k$, where

$$\Sigma_k = \sum_{p \geqslant 1} \ell(\ell-1)^{k-p-1}(\ell-2)^p \mathrm{Cat}(k-1,p) + \ell(\ell-1)^{k-p-1}(\ell-2)^{p-1}\mathrm{Cat}(k-2,p-1).$$

2. The generating series $\mathcal{C}(x,y) = \sum_{k,p \geqslant 0} \mathrm{Cat}(k,p) x^k y^p$ coincides with

$$\mathcal{C}(x,y) = \frac{1+x-xy-\sqrt{(1+x-xy)^2-4x}}{2x}.$$

3. We prove by induction that $|W_{k+1}| \geqslant \lambda |W_k|$:

$$\frac{|W_{n+1}|}{|W_n|} \geqslant (\ell-1) - \frac{1}{\ell(\ell-1)} \sum_{k \geqslant 2} \lambda^{3-2k}\Sigma_k = (\ell-1) + \frac{\lambda}{\ell-1} - \frac{\lambda^2+1}{\lambda(\ell-1)}\mathcal{C}\left(\frac{\ell-1}{\lambda^2}, \frac{\ell-2}{\ell-1}\right) = \lambda.$$

Going to $|A| = 4$ or $|A| = 5$: a (too) optimistic view?

If there are $|W_\ell^n| \approx \alpha^n$ $\sqcup\!\sqcup^2$-free words in A_ℓ^n, we roughly estimate $|S_\ell^k|$ as follows: there are

1. some words $y \in S_\ell^k$ with multiple decompositions: 010201320232 = 010201320232;

Going to $|A| = 4$ or $|A| = 5$: a (too) optimistic view?

If there are $|W_\ell^n| \approx \alpha^n$ $\sqcup\!\sqcup^2$-free words in A_ℓ^n, we roughly estimate $|S_\ell^k|$ as follows: there are

1. some words $y \in S_\ell^k$ with multiple decompositions: 010201320232 = 010201320232;
2. $\mathrm{Cat}(k-1) \approx 2^{2k}/k^{3/2}$ possible interleavings;
3. ℓ^k ways of choosing u;
4. a probability $(\alpha/\ell)^{2k}$ not to get $\sqcup\!\sqcup^2$ factors, because u_i and u_i are $\approx \sqrt{k}$ steps apart;

Going to $|A| = 4$ or $|A| = 5$: a (too) optimistic view?

If there are $|W_\ell^n| \approx \alpha^n$ $⧢^2$-free words in A_ℓ^n, we roughly estimate $|S_\ell^k|$ as follows: there are

1. some words $y \in S_\ell^k$ with multiple decompositions: 010201320232 = 010201320232;
2. $\mathrm{Cat}(k-1) \approx 2^{2k}/k^{3/2}$ possible interleavings;
3. ℓ^k ways of choosing u;
4. a probability $(\alpha/\ell)^{2k}$ not to get $⧢^2$ factors, because u_i and u_i are $\approx \sqrt{k}$ steps apart; thus,

$$|W_\ell^{n+3-2k}|\,|S_\ell^k| \approx \alpha^n (4/\ell)^k k^{-3/2},$$

and we obtain a converging series for $\ell = 4$ and $\ell = 5$.

Going to $|A| = 4$ or $|A| = 5$: a (too) optimistic view?

If there are $|W_\ell^n| \approx \alpha^n$ $\sqcup\!\sqcup^2$-free words in A_ℓ^n, we roughly estimate $|S_\ell^k|$ as follows: there are

1. some words $y \in S_\ell^k$ with multiple decompositions: 010201320232 = 010201320232;
2. $\mathrm{Cat}(k-1) \approx 2^{2k}/k^{3/2}$ possible interleavings;
3. ℓ^k ways of choosing u;
4. a probability $(\alpha/\ell)^{2k}$ not to get $\sqcup\!\sqcup^2$ factors, because u_i and u_i are $\approx \sqrt{k}$ steps apart; thus,

$$|W_\ell^{n+3-2k}|\,|S_\ell^k| \approx \alpha^n (4/\ell)^k k^{-3/2},$$

and we obtain a converging series for $\ell = 4$ and $\ell = 5$. Next, we control its sum by working on sets $W_\ell^n(w)$, i.e., considering separately suffixes of x, w and y and prefixes of y.

Going to $|A| = 4$ or $|A| = 5$: a (too) optimistic view?

If there are $|W_\ell^n| \approx \alpha^n$ $\sqcup\!\sqcup^2$-free words in A_ℓ^n, we roughly estimate $|S_\ell^k|$ as follows: there are

1. some words $y \in S_\ell^k$ with multiple decompositions: 010201320232 = 010201320232;
2. $\mathrm{Cat}(k-1) \approx 2^{2k}/k^{3/2}$ possible interleavings;
3. ℓ^k ways of choosing u;
4. a probability $(\alpha/\ell)^{2k}$ not to get $\sqcup\!\sqcup^2$ factors, because u_i and u_i are $\approx \sqrt{k}$ steps apart; thus,

$$|W_\ell^{n+3-2k}|\,|S_\ell^k| \approx \alpha^n (4/\ell)^k k^{-3/2},$$

and we obtain a converging series for $\ell = 4$ and $\ell = 5$. Next, we control its sum by working on sets $W_\ell^n(w)$, i.e., considering separately suffixes of x, w and y and prefixes of y.

In practice, experiments suggest that $|S_\ell^k| \approx \beta^k$ for some $\beta < \alpha^2$…