Is *not* prefix palindromic length of a *k*-automatic word *k*-regular?

A. Frid (with J. Peltomäki and E. Laborde)

January 4, 2021

A. Frid (with J. Peltomäki and E. Laborde)

Prefix palindromic length

January 4, 2021 1/31

Prefix palindromic length

The palindromic length PL(u) is the minimal number of concatenated palindromes needed to express the word u.

$$abbaba = (abba) (b) (a) = (a) (bb) (aba)$$

 $PL(abbaba) = 3$

The prefix palindromic length $PPL_{\mathbf{u}}(n)$ of an infinite word \mathbf{u} is the palindromic length of the prefix of length *n* of \mathbf{u} .

< ロ > < 同 > < 回 > < 回 > < 回 >

Prefix palindromic length

$\mathbf{t} = abbabaabba \cdots$



A. Frid (with J. Peltomäki and E. Laborde)

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

What is known on the prefix palindromic length?

Very little. If **u** is periodic, it can be bounded

 $\mathbf{u} = abababab \cdots$

or not

 $\mathbf{u} = abcabcabcabc \cdots$.

Conjecture

Conjecture (F., Puzynina, Zamboni, 2013)

For every non ultimately periodic word **u**, the function $PPL_{\mathbf{u}}(n)$ is unbounded.

Proven for *u* avoiding some power *k*, that is, words $v^k = \underbrace{v \cdots v}_k$ (2013).

The proof is not constructive. The number of palindromes starting at the same point is logarithmical.



Conjecture

Conjecture (F., Puzynina, Zamboni, 2013)

For every non ultimately periodic word **u**, the function $PPL_{\mathbf{u}}(n)$ is unbounded.

The conjecture is also proven

- For all morphic words ([FPZ 2013] + [Klouda, Starosta, 2015])
- In the greedy versions [Bucci, Richomme 2018]
- For all Sturmian words [F. 2018]

Conjecture

Conjecture (F., Puzynina, Zamboni, 2013)

For every non ultimately periodic word **u**, the function $PPL_{\mathbf{u}}(n)$ is unbounded.

So, the conjecture holds for *almost all* infinite words, and here are the remaining cases.



What about known words?

It is easier to find a decomposition to few palindromes than to prove it does not exist.

Upper bounds : F. 2017 - Fibonacci word; Ambrož, Kadlec, Masáková, Pelantová, 2019: morphic words from class *P*

Lower bounds: F. 2019; Shuo LI 2020 - for some Toeplitz words including the period doubling word.

< ロ > < 同 > < 三 > < 三 >

Thue-Morse word

$$\sigma: \begin{cases} a \to abba, \\ b \to baab. \end{cases}$$

 $\mathbf{t} = \sigma(\mathbf{t}) = abba \ baab \ baab \ abba \ baab \ abba \ baab$



Thue-Morse word

Theorem (F., 2019)

The following identities hold for all $n \ge 0$ *:*

$$PPL_{t}(4n) = PPL_{t}(n),$$

$$PPL_{t}(4n+1) = PPL_{t}t(n) + 1,$$

$$PPL_{t}(4n+2) = \min(PPL_{t}(n), PPL_{t}(n+1)) + 2,$$

$$PPL_{t}(4n+3) = PPL_{t}(n+1) + 1.$$

Formula rediscovered by Shuo LI later the same year.

A. Frid (with J. Peltomäki and E. Laborde)

Morphism for first differences

$$d_{\mathbf{t}}(n) = PPL_{\mathbf{t}}(n+1) - PPL_{\mathbf{t}}(n) \in \{\mathbf{n}, \mathbf{e}, \mathbf{e$$

Proposition

The sequence $(d_t(n))$ is the fixed point of the morphism



Growth



$$\limsup \frac{PPL_{\mathbf{t}}(n)}{\ln n} = \frac{3}{4\ln 2}.$$

A. Frid (with J. Peltomäki and E. Laborde)

January 4, 2021 12/31

2

<ロト < 四ト < 三ト < 三ト

k-regularity?

The Thue-Morse word **t** is 2-automatic (and 4-automatic).

Its sequence $PPL_t(n)$ is 4-regular (and 2-regular).

The first differences $d_t(n)$ are 2- and 4-automatic.

Question Is $PPL_{\mathbf{u}}(n)$ a *k*-regular sequence whenever **u** is *k*-automatic?

• • • • • • • • • • • • •

k-automatic words

Definition

A word $\mathbf{u} = \mathbf{u}[0] \cdots \mathbf{u}[n] \cdots$ is *k*-automatic if there exists a DFAO A such that every symbol $\mathbf{u}[n]$ of \mathbf{u} can be obtained as the output of A with the base-k representation of n as the input.

Morphic definition

A word **u** is *k*-automatic if and only if $\mathbf{u} = \psi(\varphi^{\omega}(a))$ for some *k*-uniform morphism φ and a uniform morphism ψ .

Kernel definition

A word **u** is *k*-automatic if and only if its *k*-kernel

$$\ker_k(\mathbf{u}) = \{ (\mathbf{u}[k^e n + b])_{n \ge 0} : e \ge 0, 0 \le b < k^e \}.$$

is finite.

Equivalence

The three definitions above are equivalent.

イロン イロン イヨン イヨン

k-regularity

Definition

A \mathbb{Z} -valued sequence is *k*-regular if the \mathbb{Z} -module generated by its *k*-kernel is finitely generated.

Charlier, Rampersad, Shallit, 2012

Many functions of k-automatic words are k-regular, including

- factor complexity,
- number of different palindrome factors of length *n*.

Walnut software

What about our case?

- That method does not work for the PPL function.
- The situation looks complicated even for simplest examples like the period-doubling word.
- On the good side: $PPL_{\mathbf{u}}$ is k-regular if and only if $d_{\mathbf{u}}$ is k-automatic.



First differences

Proposition (Particular case of a lemma by Saarela, 2017)

For every infinite word **u**, the first differences of the prefix palindromic length is a sequence over the alphabet $\{-1, 0, 1\}$ which we prefer to denote as $\{-, 0, +\}$.

So, it is a word over a finite alphabet which is *k*-automatic if and only if $PPL_{\mathbf{u}}$ is *k*-regular.

The best result we get

Theorem

If a *k*-automatic word **u** contains a finite number of distinct palindromes, then the PPL-difference sequence $d_{\mathbf{u}}$ is *k*-automatic.

IDEA OF THE PROOF: If p is the length of a longest palindrome in \mathbf{u} , we have

 $PPL(n) = \min\{PPL(n-k)+1 : k = 1, ..., p, \mathbf{u}[n-k+1..n] \text{ is a palindrome}\}.$

Just a finite number of cases to check and a transducer from *another* theorem by Cobham (1972).

< ロ > < 同 > < 回 > < 回 > < 回 >

Paperfolding word

$$\mathbf{u}_{pf} = \psi(\varphi_{pf}^{\omega}(a)) = 0010011000110110\cdots,$$

where

$$arphi_{pf}$$
: $\begin{cases} a \mapsto ab, \\ b \mapsto cb, \\ c \mapsto ad, \\ d \mapsto cd, \end{cases}$

and the coding ψ is defined as $\psi(a) = \psi(b) = 0$, $\psi(c) = \psi(d) = 1$.

A. Frid (with J. Peltomäki and E. Laborde)

イロト イポト イヨト イヨト

PPL of the paperfolding word

The longest palindromes in the paperfolding word are of length 13.

Theorem

The sequence d_{pf} over the alphabet $\{-, 0, +\}$ is equal to $d_{pf} = \gamma_{pf}(\mu_{pf}^{\omega}(a_0))$, where

$$\mu_{pf} \colon \begin{cases} a_0 \mapsto a_0 b_a, \\ a_b \mapsto a_b b_a, \quad b_a \mapsto c_b b_c, \quad c_b \mapsto a_b d_a, \quad d_a \mapsto c_b d_c, \\ a_d \mapsto a_d b_a, \quad b_c \mapsto c_d b_c, \quad c_d \mapsto a_d d_a, \quad d_c \mapsto c_d d_c \end{cases}$$

and

A. Frid (with J. Peltomäki and E. Laborde)

< □ > < □ > < □ > < □ >

PPL of the paperfolding word

with P = 0 + 00 + 00 - 0 + -0 + 0.

э

イロト イポト イヨト イ

Rudin-Shapiro word

The Rudin-Shapiro word \mathbf{u}_{rs} is the 2-automatic word

$$\mathbf{u}_{rs} = \psi(\varphi_{rs}^{\omega}(a)) = 00010010000111010\cdots$$

where

$$arphi_{rs} \colon egin{cases} a o ab, \ b o ac, \ c o db, \ d o dc, \end{cases}$$

and the coding ψ is defined by $\psi(a) = \psi(b) = 0$, $\psi(c) = \psi(d) = 1$.

The longest palindromes in the Rudin-Shapiro word are of length 14.

PPL of the Rudin-Shapiro word

Theorem

The sequence d_{rs} over the alphabet $\{-, 0, +\}$ is equal to

$$d_{rs} = \gamma_{rs}(\mu_{rs}^{\omega}(A)),$$

where

$$\mu_{rs}: \begin{cases} A \to AB, \\ B \to CD, \\ C \to EB, \\ D \to ED, \\ E \to CB \end{cases}$$

and

PPL of the Rudin-Shapiro word

with P = 0 - + 00 + 00 + 00 +.

э

Remarks

- The results on the paperfolding word and on the Rudin-Shapiro word are not direct corollaries of the general theorem.
- The Thue-Morse word contains an infinite number of palindromes but its d_t is also 2-automatic.
- We also have a conjecture on the Sierpiński word. $(0 \rightarrow 010, 1 \rightarrow 111)$

Period-doubling word

$$\mathbf{u}_{pd} = \varphi_{pd}^{\omega}(a) = abaaabababaaabaa \cdots,$$

where

$$\varphi_{pd} \colon \begin{cases} a \to ab, \\ b \to aa. \end{cases}$$

The growth of sup PPL is logarithmic (Ambrož et al., Li)

A. Frid (with J. Peltomäki and E. Laborde)

• • • • • • • • • • • •

2-kernel?

 $\ker_2(\mathbf{u}) = \{\mathbf{u}[b]\mathbf{u}[2^e + b]\mathbf{u}[2 \cdot 2^e + b] \cdots \mathbf{u}[2^e n + b] \cdots : e \ge 0, 0 \le b < 2^e\}.$

u is 2-automatic if and only if its 2-kernel is finite.

Let k_n be the number of elements of the 2-kernel observed in the prefix of $d_{pd}(n)$ pf length 4^n . Then

4 ^{<i>m</i>}	44	45	46	47	48	49	4 ¹⁰	$4^{11} = 4194304$
k _m	66	145	297	584	1046	1816	3047	5051
k_m/k_{m-1}	3.0	2.197	2.048	1.966	1.791	1.736	1.678	1.658

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

Fibonacci word

 $\mathbf{u}_f = abaababaabaab \cdots$ is the fixed point of

$$\varphi_f \colon \begin{cases} a \to ab, \\ b \to a. \end{cases}$$

It is *Fibonacci-automatic*, and all the regularity results can be generalized to it if we use the Fibonacci numeration system (Du, Mousavi, Schaeffer, Shallit, 2016).

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

Fibonacci-kernel of $d_{\mathbf{f}}(n)$?

$ \varphi_f^{3m}(a) $	5	21	89	377	1597	6765	28657	121393
k _m	3	11	31	88	207	504	1139	2377
k_m/k_{m-1}		3.67	2.82	2.85	2.35	2.43	2.26	2.09

イロン イロン イヨン イヨン

Main conjecture(s)

Conjecture

The PPL-difference sequence of the period-doubling word $(a \rightarrow ab, b \rightarrow aa)$ is **NOT** 2-automatic.

Conjecture

The PPL-difference sequence of the Fibonacci word $(a \rightarrow ab, b \rightarrow a)$ is **NOT** Fibonacci-automatic.

Conjecture

The PPL-difference sequence of a *k*-automatic word containing an infinite number of palindromes is **NOT** obliged to be *k*-automatic.