Discrete Mathematics and Theoretical Computer Science DMTCS vol. **22**:1, 2020, #6

The repetition threshold for binary rich words

James D. Currie* Lucas Mol Narad Rampersad†

Department of Mathematics and Statistics, University of Winnipeg
515 Portage Avenue, Winnipeg, Manitoba R3B 2E9 (Canada)

received 27th Sep. 2019, revised 10th Jan. 2020, accepted 13th Jan. 2020.

A word of length n is rich if it contains n nonempty palindromic factors. An infinite word is rich if all of its finite factors are rich. Baranwal and Shallit produced an infinite binary rich word with critical exponent $2+\sqrt{2}/2$ (≈ 2.707) and conjectured that this was the least possible critical exponent for infinite binary rich words (i.e., that the repetition threshold for binary rich words is $2+\sqrt{2}/2$). In this article, we give a structure theorem for infinite binary rich words that avoid $14/5$-powers (i.e., repetitions with exponent at least 2.8). As a consequence, we deduce that the repetition threshold for binary rich words is $2+\sqrt{2}/2$, as conjectured by Baranwal and Shallit. This resolves an open problem of Vesti for the binary alphabet; the problem remains open for larger alphabets.

Keywords: rich word, repetition threshold, critical exponent, palindrome

1 Introduction

A *palindrome* is a word that is equal to its reversal, i.e., it reads the same forwards and backwards. It is well-known that a word of length n contains at most n distinct nonempty palindromes [13]. Words of length n that contain n distinct nonempty palindromes are called *palindrome-rich*, or simply *rich*. An infinite word is *rich* if all of its factors are rich. Rich words were introduced in [6] (where they were called *full* words), were first studied systematically in [15], and have since been studied by many authors [11, 24, 30, 33, 34].

Let u be a finite nonempty word, and let $u = u_1 \dots u_n$, where the u_i are letters. A positive integer p is a *period* of u if $u_i = u_{i+p}$ for all $1 \leq i \leq n-p$. Let $e = |u|/p$ and let z be the prefix of u of length p. Then we say that e is an *exponent* of p, and write $u = z^e$. We say that u is *primitive* if the only integer exponent of u is 1.

For a real number $\alpha \geq 1$, a finite or infinite word w is called *α-free* if it contains no nonempty factor of exponent greater than or equal to α. Otherwise, we say that w *contains* an α-power. The *critical exponent* of w is the supremum of the set of all rational numbers α such that w contains an α-power. The *repetition threshold* for a language L is the infimum of the set of all real numbers $\alpha > 1$ such that there is an infinite

*The work of James D. Currie is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number 2017-03901].

†The work of Narad Rampersad is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number 2019-04111].

ISSN 1365–8050

α-free word in L. In other words, the repetition threshold for L is the smallest possible critical exponent among all infinite words in L.

The repetition threshold for the language of all words on a fixed alphabet of size k, denoted $\mathrm{RT}(k)$, was introduced by Dejean [12], who conjectured that

$$\mathrm{RT}(k) = \begin{cases} 2, & \text{if } k = 2; \\ 7/4, & \text{if } k = 3; \\ 7/5, & \text{if } k = 4; \\ k/(k-1), & \text{if } k \geq 5. \end{cases}$$

This conjecture was eventually proven through the work of many authors [7–9, 12, 21–23, 29]. Rampersad et al. [28] recently proposed the problem of determining the repetition threshold for the language of *balanced* words over a fixed alphabet of size k. Both Rampersad et al. [28] and Baranwal and Shallit [4] have made some progress on this problem.

We are concerned with repetitions in rich words. Vesti [34] proposed the problem of determining the repetition threshold for the language of rich words over k letters, denoted $\mathrm{RRT}(k)$. Vesti noted that $2 \leq \mathrm{RRT}(k) \leq 2 + 1/(\varphi_k - 1)$ for all $k \geq 2$, where φ_k is the generalized golden ratio. The lower bound follows from the fact that every infinite rich word contains a square [24]. The upper bound follows from the fact that the k-bonacci word is rich and has critical exponent $2 + 1/(\varphi_k - 1)$ [14]. Baranwal and Shallit [3] demonstrated that there is an infinite binary rich word with critical exponent $2 + \sqrt{2}/2$, and conjectured that this is the smallest possible critical exponent among all infinite binary rich words, i.e., that $\mathrm{RRT}(2) = 2 + \sqrt{2}/2$. In this article, we prove a structure theorem for infinite $14/5$-free binary rich words. We use this theorem to confirm Baranwal and Shallit's conjecture.

We use the following notation throughout the paper. Let $\Sigma_k = \{\mathtt{0}, \mathtt{1}, \ldots, \mathtt{k-1}\}$. Define $f : \Sigma_3^* \to \Sigma_2^*$ and $g, h : \Sigma_3^* \to \Sigma_3^*$ by

$$\begin{aligned} f(\mathtt{0}) &= \mathtt{0} \\ f(\mathtt{1}) &= \mathtt{01} \\ f(\mathtt{2}) &= \mathtt{011} \end{aligned}$$

$$\begin{aligned} g(\mathtt{0}) &= \mathtt{011} \\ g(\mathtt{1}) &= \mathtt{0121} \\ g(\mathtt{2}) &= \mathtt{012121} \end{aligned}$$

$$\begin{aligned} h(\mathtt{0}) &= \mathtt{01} \\ h(\mathtt{1}) &= \mathtt{02} \\ h(\mathtt{2}) &= \mathtt{022} \end{aligned}$$

Note that $f(h^\omega(\mathtt{0}))$ is the infinite binary rich word with critical exponent $2 + \sqrt{2}/2$ constructed by Baranwal and Shallit [3]. Also, note that f, g, and h are injective. Furthermore, these three morphisms all belong to the well-studied family of *class P* morphisms [16], which are connected to the study of rich words [2].

We prove the following structure theorem for infinite $14/5$-free binary rich words.[(i)]

Theorem 1. *Let $w \in \Sigma_2^\omega$ be a $14/5$-free rich word. For every $n \geq 1$, a suffix of w has the form $f(h^n(w_n))$ or $f(g(h^n(w_n)))$ for some word $w_n \in \Sigma_3^\omega$.*

We then demonstrate that, like $f(h^\omega(\mathtt{0}))$, the word $f(g(h^\omega(\mathtt{0})))$ has critical exponent $2 + \sqrt{2}/2$. This gives the following.

Theorem 2. *The repetition threshold for binary rich words is $2 + \sqrt{2}/2$.*

Our structure theorem is somewhat reminiscent of the well-known structure theorem for overlap-free binary words due to Restivo and Salemi [31, 32], and its extension to $7/3$-free binary words by Karhumäki and Shallit [18]. However, we deal only with infinite words.

2 A structure theorem

In this section, we prove Theorem 1. Throughout, we say that a word $w \in \Sigma_2^\omega$ is *good* if it is both rich and $14/5$-free. In particular, a good word is cube-free.

We begin by proving several properties of the morphisms f, g, and h. For every $\phi \in \{f, g, h\}$, one verifies by computer using a straightforward backtracking algorithm that the longest word $u \in \{\mathtt{1}, \mathtt{2}\}^*$ such that $\phi(u)$ is cube-free has length 6. This gives the following.

Observation 3. *Let $\phi \in \{f, g, h\}$ and $u \in \Sigma_3^\omega$. If $\phi(u)$ is cube-free, then u contains a $\mathtt{0}$.*

We now show that the morphisms f, g, and h preserve non-richness of ω-words. We require two short lemmas. The first can be derived from [2, Lemma 5.2], but we give a proof here for completeness.

Lemma 4. *Let $\phi \in \{f, g, h\}$ and let $u, v \in \Sigma_3^*$. Suppose $\phi(u)\mathtt{0}$ is a palindromic suffix of $\phi(v)\mathtt{0}$. Then u is a palindromic suffix of v.*

Proof: Since $\phi(u)\mathtt{0}$ is a suffix of $\phi(v)\mathtt{0}$ and ϕ is injective, we have that u is a suffix of v. For any $u \in \Sigma_3^*$, we have $\mathtt{0}(\phi(u))^R = \phi(u^R)\mathtt{0}$. Since $\phi(u)\mathtt{0}$ is a palindrome, $(\phi(u)\mathtt{0})^R = \mathtt{0}(\phi(u))^R = \phi(u^R)\mathtt{0}$. Since ϕ is injective, we have $u = u^R$. Thus u is a palindromic suffix of v. □

In order to prove the next lemma, we use the fact that a word w is rich if and only if every nonempty prefix p of w has a nonempty palindromic suffix that appears only once in p [15].

Lemma 5. *Let $\phi \in \{f, g, h\}$. Suppose that $w \in \Sigma_3^*$ is non-rich. Then $\phi(w)\mathtt{0}$ is non-rich.*

Proof: Let w' be a prefix of w such that every palindromic suffix of w' occurs at least twice in w'. We claim that $\phi(w')\mathtt{0}$ is a prefix of $\phi(w)\mathtt{0}$ such that every palindromic suffix of $\phi(w')\mathtt{0}$ occurs at least twice in $\phi(w')\mathtt{0}$. Any palindromic suffix of $\phi(w')\mathtt{0}$ has the form $\phi(u)\mathtt{0}$ for some u. Then by Lemma 4, we know that u is a palindromic suffix of w'. However, by hypothesis, w' contains two occurrences of u. Consequently, $\phi(w')\mathtt{0}$ contains two occurrences of the palindrome $\phi(u)\mathtt{0}$. We conclude that $\phi(w)\mathtt{0}$ is non-rich, as required. □

The fact that the morphisms f, g, and h preserve non-richness of ω-words now follows as an easy corollary.

[(i)] Note that $g = \tilde{g} \circ h$, where $\tilde{g} : \Sigma_3^* \to \Sigma_2^*$ is defined by $\tilde{g}(\mathtt{0}) = \mathtt{01}$, $\tilde{g}(\mathtt{1}) = \mathtt{1}$, and $\tilde{g}(\mathtt{2}) = \mathtt{21}$. Thus, in the statement of Theorem 1, one could replace g with $\tilde{g}$. For convenience, we have elected to work with the morphism g throughout.

Table row	v	$\|vs\|$
1	`00`	2
2	`0121012`	49
3	`021`	22
4	`0221`	19
5	`11010`	24
6	`11011`	29
7	`1102`	30
8	`112`	$*$
9	`120`	22
10	`122`	17
11	`21010`	6
12	`2101210`	48
13	`211`	3

Tab. 1: Forbidden factors in every ω-word u such that $f(u)$ is good.

Corollary 6. *Let $\phi \in \{f, g, h\}$ and $u \in \Sigma_3^\omega$. If $\phi(u)$ is rich, then u is rich.*

By straightforward induction arguments using Observation 3 and Corollary 6, we obtain the following.

Lemma 7. *Let ϕ be a morphism of the form $f \circ h^n$ or $f \circ g \circ h^n$ for some $n \geq 0$. If $\phi(u)$ is good for some $u \in \Sigma_3^\omega$, then the word u is cube-free, rich, and contains a `0`.*

We use Lemma 7 frequently throughout this section, sometimes without reference.

If w is good, then w avoids the cube `111`, so the following observation is immediate.

Observation 8. *If $w \in \Sigma_2^\omega$ is good, then a suffix of w has the form $f(u)$ for some word $u \in \Sigma_3^\omega$.*

So we may now restrict our attention to good words of the form $f(u)$, where $u \in \Sigma_3^\omega$. By Lemma 7, if $u \in \Sigma_3^\omega$ is a word such that $f(u)$ is good, then every factor of u is rich, i.e., no non-rich word is a factor of u. There are a variety of other short factors that cannot appear in such a word u. One checks by backtracking that for each word v in Table 1, there is a longest right-extension $vs \in \Sigma_3^*$ of v such that $f(vs)$ is not good. Table 1 indicates in each case the length of such a longest extension vs. (The notation $*$ indicates that $f(v)$ already fails to be good.) Hence, none of the factors in Table 1 can appear in $u \in \Sigma_3^\omega$ if $f(u)$ is good. We use this fact frequently throughout this section. We also remark that the choice of the constant $14/5$ in the definition of "good" becomes relevant at this backtracking step. If we replace $14/5$ with 3 in the definition of "good", then for certain v the backtracking search runs for a very long time without finding a longest right-extension vs such that $f(vs)$ is not good.

We will prove that if $f(u)$ is good for some ω-word u, then u either has a suffix of the form $g(W)$, or a suffix of the form $h(W)$. It turns out that if u contains the factor `0110`, then we are forced into the former structure. Otherwise, if h does not contain `0110`, then we are forced into the latter structure. We handle the case that u contains the factor `0110` first. In fact, we show that in this case, a suffix of u must have the form $h(g(U))$.

Lemma 9. *Suppose $f(u)$ is good, where $u \in \Sigma_3^\omega$, and u contains the factor `0110`. Then*

1. *The word u has a suffix of the form $g(W)$ for some word $W \in \Sigma_3^\omega$.*

2. *A suffix of W has the form $h(U)$ for some word $U \in \Sigma_3^\omega$.*

Proof: (1) Replacing u by a suffix if necessary, write $u = u_1u_2u_3u_4\cdots$, where $u_1 = \texttt{011}$ and each u_i starts with `0` and contains no other `0`. To show that $u = g(W)$ for some $W \in \Sigma_3^\omega$, it will suffice to show that every u_i is one of `011`, `0121` or `012121`. The proof is by induction on i. The base case is immediate since $u_1 = \texttt{011}$.

Now suppose for some $i \geq 1$ that $u_i \in \{\texttt{011}, \texttt{0121}, \texttt{012121}\}$. Consider the tree in Figure 1, which shows all candidates for $u_{i+1}\texttt{0}$. We explain why the word ending at every unboxed leaf of the tree cannot be a prefix of $u_{i+1}\texttt{0}$, from which we conclude that $u_{i+1} \in \{\texttt{011}, \texttt{0121}, \texttt{012121}\}$. We use the following facts:

- By Lemma 7, the word u is cube-free and rich.
- No word in Table 1 is a factor of u.
- The word u_i must have suffix `11` or `21` by the induction hypothesis; so if neither $\texttt{11}x$ nor $\texttt{21}x$ appears in u, then x cannot be a prefix of $u_{i+1}\texttt{0}$.

We discuss each unboxed leaf of the tree in lexicographic order.

- `00`: The word `00` is in Table 1.
- `010`: The words `11010` and `21010` are in Table 1.
- `0111`: The word `111` is a cube.
- `0112`: The word `112` is in Table 1.
- `0120`: The word `0120` is not rich.
- `01211`: The word `211` is in Table 1.
- `012120`: The word `012120` is not rich.
- `0121211`: The word `211` is in Table 1.
- `0121212`: The word `121212` is a cube.
- `012122`: The word `122` is in Table 1.
- `0122`: The word `122` is in Table 1.
- `02`: The word `1102` is in Table 1, and the word `2102` is not rich.

(2) To begin with, we show that `00`, `11`, `12`, and `21` are not factors of W. If W contains `00`, then u contains $g(\texttt{00}) = \texttt{011011}$, but this is impossible since `11011` is in Table 1. If W contains `11` or `12`, then u contains $g(\texttt{11}) = \texttt{01210121}$ or $g(\texttt{12}) = \texttt{0121012121}$; but this is impossible since `0121012` is in Table 1. Finally, if W contains `21`, then u contains $g(\texttt{21})\texttt{0}$=`01212101210`, but this is impossible since `2101210` is in Table 1.

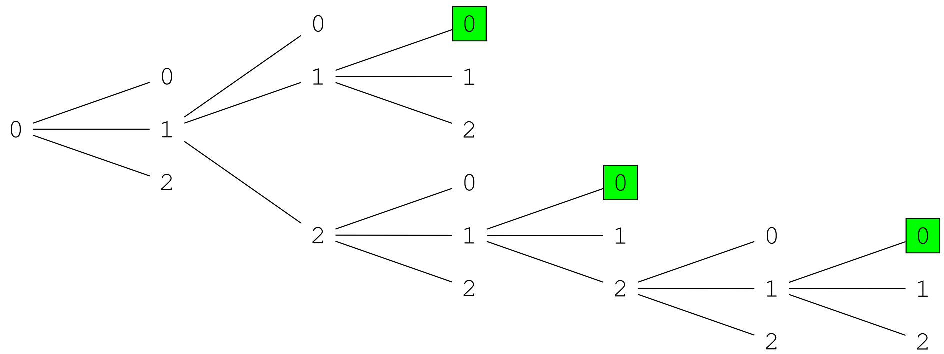

Fig. 1: The tree showing all possible prefixes of $u_{i+1}\mathtt{0}$.

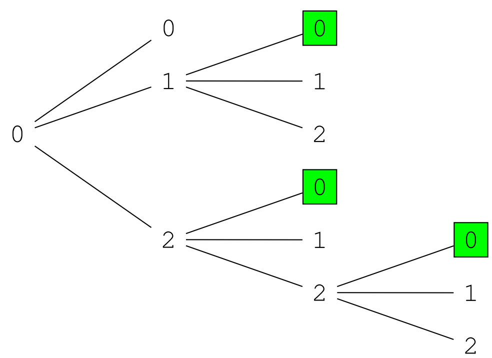

Fig. 2: The tree showing all possible prefixes of $W_i\mathtt{0}$.

By Lemma 7, the word W contains a $\mathtt{0}$. Replacing W by a suffix if necessary, write $W = W_1W_2W_3W_4\cdots$, where each W_i starts with $\mathtt{0}$ and contains no other $\mathtt{0}$. Let $i \geq 1$. As above, we enumerate the possible prefixes of W_i0 in the tree of Figure 2. It is easy to verify that the word ending at every unboxed leaf of the tree ends in one of the factors $\mathtt{00}$, $\mathtt{11}$, $\mathtt{12}$, $\mathtt{21}$, or the cube $\mathtt{222}$, so we conclude that $W_i \in \{\mathtt{01}, \mathtt{02}, \mathtt{022}\}$ as desired. □

We now show that there are several factors that do not appear in relevant preimages of good words. Define $F = \{\mathtt{1221}, \mathtt{00}, \mathtt{10101}, \mathtt{212}, \mathtt{11}\}$.

Lemma 10. *Let $u \in \Sigma_3^\omega$. Suppose that for some positive integer n, one of $f(g(h^n(u)))$ and $f(h^n(u))$ is good. Then a suffix of u does not contain any of the factors in the set F.*

Proof: By Lemma 7, the word u must be cube-free and rich, and we may assume, by taking a suffix if necessary, that u begins in $\mathtt{0}$.

$\mathtt{1221}$: Since $h(\mathtt{1221})$ contains a cube, $\mathtt{1221}$ cannot be a factor of u.

$\mathtt{00}$: For any letter $x \in \{0, 1, 2\}$, all of $f(h(\mathtt{00}x))$, $g(h(\mathtt{00}x))$, and $h^2(\mathtt{00}x)$ contain a cube. Suppose towards a contradiction that $\mathtt{00}$ is a factor of u. If $n = 1$, then $f(h(u))$ and $f(g(h(u)))$ contain factors of the form $f(h(\texttt{00}x))$ and $f(g(h(\texttt{00}x)))$, respectively, giving a cube. Otherwise, if $n \geq 2$, then $f(h^n(u))$ and $f(g(h^n(u)))$ contain factors of the form $f(h^{n-2}(h^2(\texttt{00}x)))$ and $f(g(h^{n-2}(h^2(\texttt{00}x))))$, respectively, giving a cube. Since u is cube-free, this is impossible, and we conclude that $\texttt{00}$ is a not a factor of u.

$\texttt{10101}$: All of $f(h(\texttt{10101}x))$, $g(h(\texttt{10101}x))$, and $h^2(\texttt{10101}x)$ contain cubes. By an argument similar to the one used for $\texttt{00}$, we see that the factor $\texttt{10101}$ cannot be a factor of u.

$\texttt{212}$: First note that $f(\texttt{0})$ is a prefix of $f(\texttt{1})$, which is a prefix of $f(\texttt{2})$. It follows that if $v \in \Sigma_3^*$, then $f(v\texttt{0}v\texttt{0}v\texttt{1})$ and $f(v\texttt{1}v\texttt{1}v\texttt{2})$ contain cubes. Next, note that $g(v\texttt{0}v\texttt{0}v\texttt{2}) = (V\texttt{1}V\texttt{1}V\texttt{2})\texttt{121}$, where $V = g(v)\texttt{01}$. Since $g(\texttt{1})$ is a prefix of $g(\texttt{2})$, we see that $g(v\texttt{1}v\texttt{1}v\texttt{2})$ contains a cube. Similarly, note that $h(v\texttt{0}v\texttt{0}v\texttt{2}) = (V\texttt{1}V\texttt{1}V\texttt{2})\texttt{2}$, where $V = h(v)\texttt{0}$. Further, since $h(\texttt{1})$ is a prefix of $h(\texttt{2})$, we see that $h(v\texttt{1}v\texttt{1}v\texttt{2})$ contains a cube. Finally, note that $h(\texttt{212}) = \texttt{02202022}$ ends in a factor of the form $v\texttt{0}v\texttt{0}v\texttt{2}$, where $v = \texttt{2}$.

Suppose that $\texttt{212}$ is a factor of u. It follows by induction that $h^n(u)$ contains either a cube, a factor of the form $v\texttt{0}v\texttt{0}v\texttt{2}$ (in the case $n = 1$), or a factor of the form $v\texttt{1}v\texttt{1}v\texttt{2}$. It follows that $g(h^n(u))$ contains a factor of the form $V\texttt{1}V\texttt{1}V\texttt{2}$, or a cube, so that $f(h^n(u))$ and $f(g(h^n(u)))$ both contain cubes. This is impossible.

$\texttt{11}$: Suppose that $\texttt{11}$ is a factor of u. The words $\texttt{111}$, $h(\texttt{112})$, and $h(\texttt{211})\texttt{0}$ all contain a cube, hence $\texttt{11}$ is preceded and followed by $\texttt{0}$. Thus, $\texttt{0110}$ is a factor of u. However, all of $f(h(\texttt{0110}))$, $g(h(\texttt{0110}))$, and $h^2(\texttt{0110})$ contain a cube. By an argument similar to the one used for $\texttt{00}$, we conclude that $\texttt{11}$ is not a factor of u. $\square$

We now prove that any cube-free rich word $u \in \Sigma_3^\omega$ that avoids the finite list of factors from Lemma 10 must have a suffix of the form $h(W)$. Together, Lemma 10 and the following lemma will form the inductive step of our structure theorem.

Lemma 11. *Suppose that $u \in \Sigma_3^\omega$ is cube-free and rich. If u does not contain any of the factors in the set F, then u has a suffix of the form $h(W)$ for some word $W \in \Sigma_3^\omega$.*

Proof: Taking a suffix of u if necessary, write $u = u_1u_2u_3u_4\cdots$, where each u_i starts with $\texttt{0}$ and contains no other $\texttt{0}$. It will suffice to show that every u_i is one of $\texttt{01}$, $\texttt{02}$ or $\texttt{022}$. For an arbitrary $i \geq 1$, as in the proof of Lemma 9, we consider the tree of possible prefixes of $u_i\texttt{0}$, drawn in Figure 3. We explain why the word ending at every unboxed leaf of the tree cannot be a factor of u.

- $\texttt{00}$: The word $\texttt{00}$ is in F.
- $\texttt{011}$: The word $\texttt{11}$ is in F.
- $\texttt{0120}$: The word $\texttt{0120}$ is not rich.
- $\texttt{01211}$: The word $\texttt{11}$ is in F.
- $\texttt{01212}$: The word $\texttt{212}$ is in F.
- $\texttt{01220}$: The word $\texttt{01220}$ is not rich.
- $\texttt{01221}$: The word $\texttt{1221}$ is in F.
- $\texttt{01222}$: The word $\texttt{222}$ is a cube.

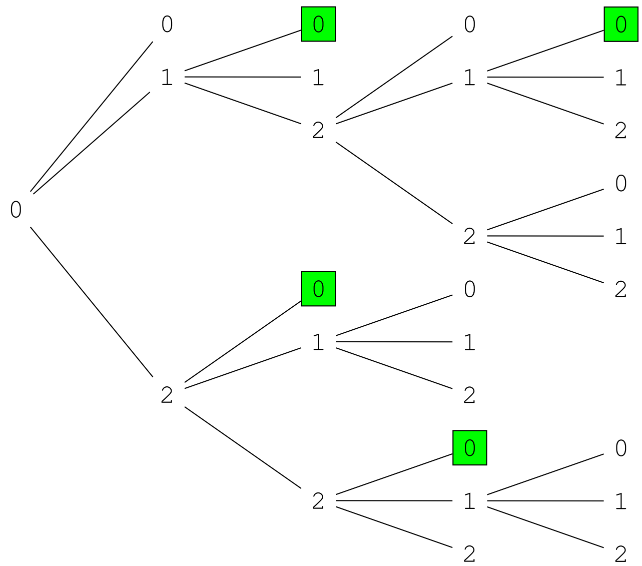

Fig. 3: The tree showing all possible prefixes of $u_i\mathtt{0}$.

- 0210: The word 0210 is not rich.
- 0211: The word 11 is in F.
- 0212: The word 212 is in F.
- 02210: The word 02210 is not rich.
- 02211: The word 11 is in F.
- 02212: The word 212 is in F.
- 0222: The word 222 is a cube.

Thus, we conclude from Figure 3 that $u_i \in \{\mathtt{01}, \mathtt{0121}, \mathtt{02}, \mathtt{022}\}$ for all $i \geq 1$. Suppose towards a contradiction that for some $i \geq 1$, we have $u_i = \mathtt{0121}$. Because u does not have the non-rich word 2102 as a factor, we see that $u_{i+1} \neq \mathtt{02}, \mathtt{022}$. Suppose that $u_{i+1} = \mathtt{01}$. Then $u_{i+2} \in \{\mathtt{01}, \mathtt{0121}, \mathtt{02}, \mathtt{022}\}$, forcing u to contain one of 10101, or 210102. However, this is impossible since 10101 is in F, and 210102 is not rich. We conclude that $u_{i+1} = \mathtt{0121}$. By the same argument, $u_{i+2} = \mathtt{0121}$, and u contains the cube $(\mathtt{0121})^3$. This is impossible. It follows that we cannot have $u_i = \mathtt{0121}$, so that $u_i \in \{\mathtt{01}, \mathtt{02}, \mathtt{022}\}$, as desired. □

Finally, we still need to handle the case that $f(u)$ is good, but $u \in \Sigma_3^*$ does not contain the factor 0110.

Lemma 12. *Suppose $f(u)$ is good for some word $u \in \Sigma_3^\omega$ that does not contain the factor* 0110. *Then u has a suffix of the form $h(W)$.*

Proof: By Lemma 7, we know that u is cube-free and rich, and by taking a suffix if necessary, we may assume that u begins in 0. By Lemma 11, it suffices to show that u does not contain any of the words in F.

1221, 00: The words 122 and 00 are in Table 1.

10101: Since $f(0)$ is a prefix of $f(1)$ and $f(2)$, the word $f(10101x)$ begins with a cube. Since $f(u)$ is cube-free, we conclude that u cannot contain the factor 10101.

212: Backtracking by computer as we did to create Table 1, with the additional restriction that 0110 is not allowed, one finds that the longest right extension of 212 has length 21. Hence 212 is not a factor of u.

11: The word 11 cannot be preceded or followed by 1 in u, since u is cube-free. Further, the word 11 cannot be preceded or followed by 2 in u, since 112 and 211 are in Table 1. However, then if 11 is a factor of u, so is 0110, contrary to assumption. □

We are now ready to prove our structure theorem.

Proof Proof of Theorem 1: The proof is by induction on n. We first establish the base case $n = 1$. By Observation 8, a suffix of w has the form $f(w_0)$ for some word $w_0 \in \Sigma_3^\omega$. If w_0 contains the factor 0110, then by Lemma 9, there is a suffix of w_0 that has the form $g(h(w_1))$. Otherwise, if w_0 does not contain the factor 0110, then by Lemma 12, there is a suffix of w_0 that has the form $h(w_1)$. Therefore, a suffix of w has the form $f(h(w_1))$ or $f(g(h(w_1)))$ for some $w_1 \in \Sigma_3^*$, establishing the base case.

Suppose now that for some $n \geq 1$, a suffix of w has the form $f(h^n(w_n))$ or $f(g(h^n(w_n)))$ for some $w_n \in \Sigma_3^\omega$. By Lemma 10, there is a suffix of w_n that does not contain any of the factors in $F = \{1221, 00, 10101, 212, 11\}$. By Lemma 7, we know that w_n is cube-free and rich. Therefore, by Lemma 11, a suffix of w_n has the form $h(w_{n+1})$ for some $w_{n+1} \in \Sigma_3^\omega$. We conclude that a suffix of w has the form $f(h^{n+1}(w_{n+1}))$ or $f(g(h^{n+1}(w_{n+1})))$. □

3 The repetition threshold

Baranwal and Shallit [3] showed that the word $f(h^\omega(0))$ is rich and has critical exponent $2 + \sqrt{2}/2$. They showed both properties using the Walnut theorem prover. We show that the word $f(g(h^\omega(0)))$ has the same properties using a different method, which relies heavily on a connection to Sturmian words; it turns out that both $f(h^\omega(0))$ and $f(g(h^\omega(0)))$ are *complementary symmetric Rote words*[(ii)].

A word $w \in \Sigma_2^\omega$ is a *complementary symmetric Rote word* if its factorial language is closed under complementation and it has factor complexity $\mathcal{C}(n) = 2n$ for all $n \geq 1$. For any infinite binary word $w = (w_n)_{n\geq 0}$, let $\Delta(w) = ((w_n + w_{n+1}) \bmod 2)_{n\geq 0}$, i.e., $\Delta(w)$ is the sequence of first differences of w modulo 2. We use the fact that a word $w \in \Sigma_2^\omega$ is a complementary symmetric Rote word if and only if $\Delta(w)$ is a Sturmian word [27, Theorem 3].

Let $u = f(g(h^\omega(0)))$. We begin by showing that $\Delta(u)$ is a certain Sturmian word v, from which we conclude that u is a complementary symmetric Rote word. In particular, this implies that u is rich [5]. We then relate the repetitions in v to those in u, and use the theory of repetitions in Sturmian words to establish that the critical exponent of u is $2 + \sqrt{2}/2$. We note that a similar calculation would provide an alternate proof of Baranwal and Shallit's result that the critical exponent of $f(h^\omega(0))$ is $2 + \sqrt{2}/2$.

(ii) This very useful observation was communicated to us by Edita Pelantová.

Define $\lambda, \mu : \Sigma_3^* \to \Sigma_2^*$ by

$$\begin{aligned}\lambda(\mathtt{0}) &= \mathtt{0}\\ \lambda(\mathtt{1}) &= \mathtt{11}\\ \lambda(\mathtt{2}) &= \mathtt{101}\end{aligned}$$

$$\begin{aligned}\mu(\mathtt{0}) &= \mathtt{01111}\\ \mu(\mathtt{1}) &= \mathtt{01110111}\\ \mu(\mathtt{2}) &= \mathtt{0111011110111}\end{aligned}$$

We extend the map Δ to finite binary words in the obvious manner in order to prove the following straightforward lemma.

Lemma 13. *Let $w \in \Sigma_3^*$. Then*

1. $\Delta(f(w)\mathtt{0}) = \lambda(w)$*, and*
2. $\Delta(f(g(w))\mathtt{0}) = \mu(w)$.

Proof: One checks that $\Delta(f(a)\mathtt{0}) = \lambda(a)$ and $\Delta(f(g(a))\mathtt{0}) = \mu(a)$ for all $a \in \Sigma_3$.

For (1), we proceed by induction on the length n of w. When $n = 0$, we have $\Delta(f(\varepsilon)\mathtt{0}) = \varepsilon = \lambda(\varepsilon)$, so the statement holds. Suppose for some $n \geq 0$ that the statement holds for all words w of length n. Let x be a word of length $n+1$. Then $x = ya$ for some $y \in \Sigma_3^n$ and $a \in \Sigma_3$. Then $\Delta(f(x)\mathtt{0}) = \Delta(f(y)\mathtt{0})\Delta(f(a)\mathtt{0}) = \lambda(y)\lambda(a) = \lambda(x)$.

The proof of (2) is similar. □

Define morphisms $\xi, \eta : \Sigma_2^* \to \Sigma_2^*$ by

$$\begin{aligned}\xi(\mathtt{0}) &= \mathtt{011}\\ \xi(\mathtt{1}) &= \mathtt{01}\end{aligned}$$

$$\begin{aligned}\eta(\mathtt{0}) &= \mathtt{011}\\ \eta(\mathtt{1}) &= \mathtt{1}.\end{aligned}$$

Note that both ξ and η are Sturmian morphisms (see [19, Section 2.3]). By checking the images of all letters in Σ_3, one verifies that $\lambda \circ h = \xi \circ \lambda$ and $\mu = \eta \circ \xi \circ \lambda$.

Lemma 14.

1. $\Delta(f(h^\omega(\mathtt{0}))) = \xi^\omega(\mathtt{0})$
2. $\Delta(f(g(h^\omega(\mathtt{0})))) = \eta(\xi^\omega(\mathtt{0}))$

Proof: For (1), we show that $\Delta(f(h^n(\mathtt{0}))\mathtt{0}) = \xi^n(\mathtt{0})$ for every $n \geq 0$. First of all, we have $\Delta(f(h^n(\mathtt{0}))\mathtt{0}) = \lambda(h^n(\mathtt{0}))$ by Lemma 13, so it suffices to show that $\lambda(h^n(\mathtt{0})) = \xi^n(\mathtt{0})$. We proceed by induction on n. The statement is easily verified when $n = 0$. Suppose for some $n \geq 0$ that $\lambda(h^n(\mathtt{0})) = \xi^n(\mathtt{0})$. Using the fact that $\lambda \circ h = \xi \circ \lambda$, we obtain

$$\lambda(h^{n+1}(\mathtt{0})) = \xi(\lambda(h^n(\mathtt{0}))) = \xi(\xi^n(\mathtt{0})) = \xi^{n+1}(\mathtt{0}),$$

which completes the proof of (1).

For (2), we show that $\Delta(f(g(h^n(0)))0) = \eta(\xi^{n+1}(0))$ for every $n \geq 0$. By Lemma 13, we have $\Delta(f(g(h^n(0)))0) = \mu(h^n(0))$, so it suffices to show that $\mu(h^n(0)) = \eta(\xi^{n+1}(0))$. Using the facts that $\mu = \eta \circ \xi \circ \lambda$ and $\lambda(h^n(0)) = \xi^n(0)$, we obtain

$$\mu(h^n(0)) = \eta(\xi(\lambda(h^n(0)))) = \eta(\xi(\xi^n(0))) = \eta(\xi^{n+1}(0)),$$

which completes the proof of (2). □

Since $\xi^\omega(0)$ and $\eta(\xi^\omega(0))$ are Sturmian words, we have proved that $f(h^\omega(0))$ and $f(g(h^\omega(0)))$ are complementary symmetric Rote words. Since all complementary symmetric Rote words are rich [5, Theorem 25], the following is immediate.

Theorem 15. *The words $f(h^\omega(0))$ and $f(g(h^\omega(0)))$ are rich.*

Now we analyze the repetitions in $u = f(g(h^\omega(0)))$. Let $v = \Delta(u) = \eta(\xi^\omega(0))$ (by Lemma 14). The relation between the repetitions in u and those in v is given by the following lemma.

Lemma 16. *For any infinite binary word $x = (x_n)_{n\geq 0}$, let $y = (y_n)_{n\geq 0} = \Delta(x)$. If x contains a repetition*

$$(x_i x_{i+1} \cdots x_{i+\ell-1})^e x_i x_{i+1} \cdots x_{i+t-1}$$

for some positive integers $e \geq 2$, $\ell \geq 1$, and $t \leq \ell$, then y contains a repetition

$$(y_i y_{i+1} \cdots y_{i+\ell-1})^e y_i y_{i+1} \cdots y_{i+t-2}$$

where the number of 1's in $y_i y_{i+1} \cdots y_{i+\ell-1}$ is even.

Proof: The fact that y contains such a repetition is immediate. To see that the number of 1's in $y_i y_{i+1} \cdots y_{i+\ell-1}$ is even, note first that

$$\sum_{j=0}^{r} y_{i+j} \bmod 2 = (x_i + x_{i+r+1}) \bmod 2.$$

Hence if $x_i = x_{i+\ell}$, we have

$$\sum_{j=0}^{\ell-1} y_{i+j} \bmod 2 = (x_i + x_{i+\ell}) \bmod 2 = 0.$$

It follows that the number of 1's in $y_i y_{i+1} \cdots y_{i+\ell-1}$ is even, as required. □

We now analyze the repetitions in v. We first need to review some basic definitions from the theory of Sturmian words and the theory of continued fractions. Consider a real number α with continued fraction expansion $\alpha = [d_0; d_1, d_2, d_3, \ldots]$, where $d_0 = 0$ and d_i is a positive integer for all $i > 0$.

The *characteristic Sturmian word with slope α* (see [1, Chapter 9]) is the infinite word c_α obtained as the limit of the sequence of *standard words* s_n defined by

$$s_0 = 0, \quad s_1 = 0^{d_1 - 1}1, \quad s_n = s_{n-1}^{d_n} s_{n-2}, \quad n \geq 2.$$

For $n \geq 2$, we also define the *semi-standard words*

$$s_{n,t} = s_{n-1}^t s_{n-2},$$

for every $1 \leq t < d_n$. The slope α is the frequency of 1's in c_α. It is known that any Sturmian word with the same frequency of 1's has the same set of factors as c_α.

We also make use of the *convergents* of α, namely

$$\frac{p_n}{q_n} = [0; d_1, d_2, d_3, \ldots, d_n],$$

where

$$\begin{aligned} p_{-2} &= 0, \quad p_{-1} = 1, \quad p_n = d_n p_{n-1} + p_{n-2} \text{ for } n \geq 0; \\ q_{-2} &= 1, \quad q_{-1} = 0, \quad q_n = d_n q_{n-1} + q_{n-2} \text{ for } n \geq 0. \end{aligned}$$

Note that $|s_n| = q_n$ for $n \geq 0$. We use the well-known fact that $q_{n-1}/q_n = [0; d_n, d_{n-1}, \ldots, d_1]$.

We can now prove the main theorem concerning the critical exponent of u.

Theorem 17. *The critical exponent of u is $2 + \sqrt{2}/2$.*

Proof: Let $\bar{\xi} : \Sigma_2^* \to \Sigma_2^*$ be the Sturmian morphism defined by $\mathtt{0} \to \mathtt{01}$, and $\mathtt{1} \to \mathtt{001}$. Let $\bar{\eta} : \Sigma_2^* \to \Sigma_2^*$ be the Sturmian morphism defined by $\mathtt{0} \to \mathtt{0}$ and $\mathtt{1} \to \mathtt{001}$. Let $\bar{v} = \bar{\eta}(\bar{\xi}^\omega(\mathtt{0}))$. The morphisms $\bar{\xi}$ and $\bar{\eta}$ are obtained by conjugating and complementing ξ and η, so the factors of $\bar{v}$ are exactly the complements of the factors of v. Clearly, the periods and exponents of the repetitions in v and $\bar{v}$ are identical, so we analyze the repetitions in $\bar{v}$ instead. To analyze the repetitions in $\bar{v}$ it suffices to consider the repetitions in the characteristic word with the same slope as $\bar{v}$.

The matrix of $\bar{\xi}$ is $M_{\bar{\xi}} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ and the matrix of $\bar{\eta}$ is $M_{\bar{\eta}} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. The frequency vector of $\mathtt{0}$'s and $\mathtt{1}$'s in $\xi^\omega(\mathtt{0})$ is the normalized eigenvector $\mathbf{v}$ of $M_{\bar{\xi}}$ corresponding to the dominant eigenvalue $1+\sqrt{2}$. We have $\mathbf{v} = (2-\sqrt{2}, \sqrt{2}-1)^T$. We then compute $M_{\bar{\eta}}\mathbf{v}$ and normalize to find that the frequency of 1's in $\bar{v}$ is $\alpha = (3-\sqrt{2})/7$. We therefore consider the characteristic word c_α with slope α in place of $\bar{v}$.

Since $\alpha = [0; 4, \overline{2}]$, we see that c_α is the infinite word obtained as the limit of the sequence of standard words s_k defined by

$$s_0 = \mathtt{0}, \quad s_1 = s_0^{4-1}\mathtt{1}, \quad s_k = s_{k-1}^2 s_{k-2}, \quad k \geq 2$$

We have $s_1 = \mathtt{0001}$, $s_2 = \mathtt{000100010}$, $s_3 = \mathtt{0001000100001000100001}$, etc. We will also need the semi-standard words

$$s_{k,1} = s_{k-1}s_{k-2}, \quad k \geq 2.$$

Note that the number of $\mathtt{0}$'s in s_k is always odd and the number of $\mathtt{0}$'s in $s_{k,1}$ is always even. Also note that by [26, Proposition 4.6.12], the critical exponent of c_α is $3+\sqrt{2}$. Write $c_\alpha = (c_n)_{n \geq 0}$.

Now suppose that u contains a repetition

$$y^e y' = (u_i u_{i+1} \cdots u_{i+\ell-1})^e u_i u_{i+1} \cdots u_{i+t-1}$$

for some positive integers $e \geq 2$, $\ell \geq 1$, and $t \leq \ell$. By Lemma 16, we see that v contains a repetition

$$(v_i v_{i+1} \cdots v_{i+\ell-1})^e v_i v_{i+1} \cdots v_{i+t-2},$$

where the number of 1's in $v_{i+1}\cdots v_{i+\ell-1}v_\ell$ is even. It follows that $\bar{v}$, and hence c_α, contains a repetition

$$z^e z' = (c_j c_{j+1} \cdots c_{j+\ell-1})^e c_j c_{j+1} \cdots c_{j+t-2},$$

where the number of 0's in z is even. The remainder of the argument is very similar to that of [28, Proposition 6].

Suppose that z is not primitive. Since the critical exponent of c_α is $3+\sqrt{2}$, the exponent of z cannot be greater than 2. Thus z is a square, and we get that the exponent of $z^e z'$ is at most

$$\frac{3+\sqrt{2}}{2} < 2 + \frac{\sqrt{2}}{2}.$$

So we may assume that z is primitive. By [25, Corollary 4.6] (originally due to Damanik and Lenz [10]), the word z is either a conjugate of one of the standard words s_k, or a conjugate of one of the semi-standard words $s_{k,1}$. However, s_k has an odd number of $\mathtt{0}$'s, so this case is ruled out.

Thus we may assume that z is a conjugate of $s_{k,1}$ for some $k \geq 2$. Hence $|z| = q_{k-2} + q_{k-1}$ for some $k \geq 2$. From [17, Theorem 4(i)], one finds that the longest factor of c_α with period $q_{k-2} + q_{k-1}$ has length $2(q_{k-2} + q_{k-1}) + q_{k-1} - 2$. It follows that $z^e z'$ has exponent at most

$$\frac{2(q_{k-2} + q_{k-1}) + q_{k-1} - 2}{q_{k-2} + q_{k-1}}$$

for some $k \geq 2$. In turn, it must be the case that $y^e y'$ has exponent

$$\begin{aligned} E_k &= \frac{2(q_{k-2} + q_{k-1}) + q_{k-1} - 1}{q_{k-2} + q_{k-1}} \\ &= 2 + \frac{q_{k-1} - 1}{q_{k-2} + q_{k-1}} \qquad (1) \\ &= 2 + \frac{1 - 1/q_{k-1}}{1 + q_{k-2}/q_{k-1}} \qquad (2) \end{aligned}$$

for some $k \geq 2$.

We claim that $\lim_{k\to\infty} E_k = 2 + \sqrt{2}/2$, and that the sequence $(E_k)_{k\geq 2}$ is increasing. It follows that the exponent of $y^e y'$ is at most $2 + \sqrt{2}/2$. Moreover, by the discussion above, the word u has a factor of exponent E_k for every $k \geq 2$. Thus, we conclude from the claim that u has critical exponent $2 + \sqrt{2}/2$. We now complete the proof of the claim.

First we show that $\lim_{k\to\infty} E_k = 2+\sqrt{2}/2$. Since $q_{k-2}/q_{k-1} = [0; \underbrace{2, 2, \ldots, 2}_{k-2}, 4]$, we see immediately that $\lim_{k\to\infty} q_{k-2}/q_{k-1} = [0; \overline{2}] = \sqrt{2} - 1$. From (2), we obtain

$$\lim_{k\to\infty} E_k = 2 + \sqrt{2}/2.$$

Finally, we show that the sequence $(E_k)_{k\geq 2}$ is increasing. Let $k \geq 2$. Starting from (1), using algebra and the recursion $q_k = 2q_{k-1} + q_{k-2}$, one finds that $E_{k+1} > E_k$ if and only if

$$2q_{k-1} > q_{k-1}^2 - q_k q_{k-2}. \qquad (3)$$

When $k = 2$, we have $q_{k-1}^2 - q_k q_{k-2} = 4^2 - 9 \cdot 1 = 7$. Suppose for some $k \geq 2$ that $q_{k-1}^2 - q_k q_{k-2} = 7(-1)^k$. Then

$$q_k^2 - q_{k+1}q_{k-1} = (2q_{k-1} + q_{k-2})q_k - (2q_k + q_{k-1})q_{k-1} = q_{k-2}q_k - q_{k-1}^2 = 7(-1)^{k+1}.$$

Thus, by mathematical induction, we have $q_{k-1}^2 - q_k q_{k-2} = 7(-1)^k$ for all $k \geq 2$. In particular, the right-hand side of (3) is at most 7 for all $k \geq 2$. Since $q_{k-1} \geq q_1 = 4$ for all $k \geq 2$, we conclude that (3) is satisfied for all $k \geq 2$. Therefore, we have $E_{k+1} > E_k$ for all $k \geq 2$.

This completes the proof of the claim, and hence the theorem. □

Since $f(h^\omega(\texttt{0}))$ and $f(g(h^\omega(\texttt{0})))$ both have critical exponent $2 + \sqrt{2}/2$, Theorem 2 now follows immediately from Theorem 1.

4 Future Prospects

For $k \geq 3$, it remains an open problem to determine the repetition threshold $\mathrm{RRT}(k)$ for the language of rich words on k letters. In fact, we even lack a conjecture for the value of $\mathrm{RRT}(k)$ in these cases. Baranwal and Shallit [3] have established that $\mathrm{RRT}(3) \geq 9/4$, but did not explicitly conjecture that $\mathrm{RRT}(3) = 9/4$.

Acknowledgements

After reading an early draft of this work, Edita Pelantová pointed out to us that $f(h^\omega(\texttt{0}))$ and $f(g(h^\omega(\texttt{0})))$ are complementary symmetric Rote words. We are very grateful to her for this observation, as it allowed us to prove some stronger results than were in our original draft. We take this opportunity to refer the reader to the recent manuscript by Medková, Pelantová, and Vuillon [20] for more results on complementary symmetric Rote words. We thank Jeffrey Shallit for his comments on our earlier draft as well. We also thank the referee, who provided some helpful comments.

References

[1] J.-P. Allouche and J. Shallit, *Automatic Sequences*, Cambridge, 2003.

[2] L. Balková, E. Pelantová, Š. Starosta, Infinite words with finite defect, Adv. Appl. Math. 47 (2011), 562–574.

[3] A. Baranwal and J. Shallit, Repetitions in infinite palindrome-rich words, in: R. Mercas and D. Reidenbach (Eds.), *Proc. WORDS 2019*, Lecture Notes in Computer Science, Vol. 11682, Springer, (2019), 93–105.

[4] A. Baranwal and J. Shallit, Critical exponent of infinite balanced words via the Pell number system, in: R. Mercas and D. Reidenbach (Eds.), *Proc. WORDS 2019*, Lecture Notes in Computer Science, Vol. 11682, Springer, (2019), 80–92.

[5] A. Blondin Massé, S. Brlek, S. Labbé, L. Vuillon, Palindromic complexity of codings of rotations, Theoret. Comput. Sci. 412 (2011), 6455–6463.

[6] S. Brlek, S. Hamel, M. Nivat, C. Reutenauer, The palindromic complexity of infinite words, Internat. J. Found. Comput. Sci. 15 (2004), 293–306.

[7] A. Carpi, On Dejean's conjecture over large alphabets, Theoret. Comput. Sci. 385 (2007), 137–151.

[8] J. D. Currie and N. Rampersad, Dejean's conjecture holds for $n \geq 27$, RAIRO - Theor. Inform. Appl. 43 (2009), 775–778.

[9] J. D. Currie and N. Rampersad, A proof of Dejean's conjecture, Math. Comp. 80 (2011), 1063–1070.

[10] D. Damanik and D. Lenz, The index of Sturmian sequences, European J. Combinatorics 23 (2002), 23–29.

[11] A. de Luca, A. Glen, L. Q. Zamboni, Rich, Sturmian, and trapezoidal words, Theoret. Comput. Sci. 407 (2008), 569–573.

[12] F. Dejean, Sur un théorème de Thue, J. Combin. Theory Ser. A 13 (1972), 90–99.

[13] X. Droubay, J. Justin, G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, Theoret. Comput. Sci. 255 (2001), 539–553.

[14] A. Glen and J. Justin, Episturmian words: A survey, RAIRO – Theoret. Inform. Appl. 43 (2009), 403–442.

[15] A. Glen, J. Justin, S. Widmer, L. Q. Zamboni, Palindromic richness, European J. Combinatorics 30 (2009), 510-531.

[16] A. Hof, O. Knill, B. Simon, Singular continuous spectrum for palindromic Schrödinger operators, Comm. Math. Phys. 174 (1995), 149–159.

[17] J. Justin and G. Pirillo, Fractional powers in Sturmian words, Theoret. Comput. Sci. 255 (2001), 363–376.

[18] J. Karhumäki and J. Shallit, Polynomial versus exponential growth in repetition-free binary words, J. Combin. Theory Ser. A 105 (2004), 335–347.

[19] M. Lothaire, Algebraic Combinatorics on Words, Cambridge, 2002.

[20] K. Medková, E. Pelantová, L. Vuillon, Derivated sequences of complementary symmetric Rote words. Preprint available at `https://arxiv.org/abs/1812.03748`

[21] M. Mohammad-Noori and J. D. Currie, Dejean's conjecture and Sturmian words, European J. Combin. 28 (2007), 876–890.

[22] J. Moulin-Ollagnier, Proof of Dejean's conjecture for alphabets with 5, 6, 7, 8, 9, 10, and 11 letters, Theoret. Comput. Sci. 95 (1992), 187–205.

[23] J. J. Pansiot, A propos d'une conjecture de F. Dejean sur les répétitions dans les mots, Discrete Appl. Math. 7 (1984), 297–311.

[24] E. Pelantová and Štěpán Starosta, Languages invariant under more symmetries: Overlapping factors versus palindromic richness, Discrete Math. 313 (2013), 2432–2445.

[25] J. Peltomäki, Characterization of repetitions in Sturmian words: A new proof. Inform. Process. Lett. 115 (2015), 886–891.

[26] J. Peltomäki, *Privileged Words and Sturmian Words*, PhD thesis, TUCS Dissertations No. 214, August 2016. Available at `http://www.utupub.fi/handle/10024/124473`

[27] G. Rote, Sequences with subword complexity $2n$, J. Number Theory 46 (1994), 196–213.

[28] N. Rampersad, J. Shallit, É. Vandomme, Critical exponents of infinite balanced words, Theoret. Comput. Sci 777 (2019), 454–463.

[29] M. Rao, Last cases of Dejean's conjecture, Theoret. Comput. Sci. 412 (2011), 3010–3018.

[30] A. Restivo and G. Rosone, Burrows-Wheeler transform and palindromic richness, Theoret. Comput. Sci. 410 (2009), 3018–3026.

[31] A. Restivo and S. Salemi, On weakly square free words, Bull. European Assoc. Theoret. Comput. Sci. 21 (1983), 49–56.

[32] A. Restivo and S. Salemi, Overlap free words on two symbols, in: M. Nivat and D. Perrin (Eds.), *Automata on Infinite Words*, Lecture Notes in Computer Science, Vol. 192, Springer, Berlin (1985), 198–206.

[33] J. Vesti, Extensions of rich words, Theoret. Comput. Sci. 548 (2014), 14–24.

[34] J. Vesti, Rich square-free words, Theoret. Comput. Sci. 687 (2017), 48–61.