Rich square-free words

Jetro Vesti

University of Turku, Department of Mathematics and Statistics, 20014 Turku, Finland

Abstract

A word w is *rich* if it has $|w|+1$ many distinct palindromic factors, including the empty word. A word is *square-free* if it does not have a factor uu, where u is a non-empty word.

Pelantová and Starosta (Discrete Math. 313 (2013)) proved that every infinite rich word contains a square. We will give another proof for that result. Pelantová and Starosta denoted by $r(n)$ the length of a longest rich square-free word on an alphabet of size n. The exact value of $r(n)$ was left as an open question. We will give an upper and a lower bound for $r(n)$, and make a conjecture that our lower bound is exact.

We will also generalize the notion of repetition threshold for a limited class of infinite words. The repetition thresholds for episturmian and rich words are left as an open question.

Keywords: Combinatorics on words, Palindromes, Rich words, Square-free words, Repetition threshold.
2000 MSC: 68R15

1. Introduction

In recent years, rich words and palindromes have been studied extensively in combinatorics on words. A word is a *palindrome* if it is equal to its reversal. In [DJP], the authors proved that every word w has at most $|w|+1$ many distinct palindromic factors, including the empty word. The class of words which achieve this limit was introduced in [BHNR] with the term *full* words. When the authors of [GJWZ] studied these words thoroughly they called them *rich* (in palindromes). Rich words have been studied in various papers, for example in [AFMP], [BDGZ1], [BDGZ2], [DGZ], [RR] and [V].

The *defect* of a finite word w, denoted $D(w)$, is defined as $D(w) = |w| + 1 - |\mathrm{Pal}(w)|$, where $\mathrm{Pal}(w)$ is the set of palindromic factors in w. The *defect* of an infinite word w is

Email address: `jejove@utu.fi` (Jetro Vesti)

defined as $D(w) = \sup\{D(u) |\ u \text{ is a factor of } w\}$. In other words, the defect tells how many palindromes the word is lacking. Rich words are exactly those whose defect is equal to 0.

The authors of [PS] proved, in Theorem 4 of the article, that every recurrent word with finite Θ-defect contains infinitely many overlapping factors. An *overlapping* word is a word of form uuv, where v is a non-empty prefix of u. A word is a Θ-*palindrome* if it is a fixed point of an involutive antimorphism Θ. The reversal mapping R is an involutive antimorphism, which means that if $\Theta = R$ then Θ-defect is equal to the defect. This means Theorem 4 in [PS] holds also for normal defect and normal palindromes. In this article we will restrict ourselves to the case where Θ is the reversal mapping.

Since every rich word has a finite defect and every overlapping factor uuv has a square uu, a corollary of Theorem 4 in [PS] is that every recurrent rich word contains a square. This was noted in [PS] as Remark 6, where the word *recurrent* was replaced with *infinite*. This can be done, since every infinite rich word x has a recurrent point y in the shift orbit closure of x (see e.g. Section 4 of [Q]). We know y has a square, which means x has a square. In Corollary 2.8 of our article, we will give another proof of the result in Remark 6 of [PS].

In Remark 6 of [PS] there was also noted that since every rich square-free word is finite, we can look for a longest one. The length of a longest such word, on an alphabet of size n, was denoted by $r(n)$. An explicit formula for $r(n)$ was left as an open question.

In Section 2.1 we will construct recursively a sequence of rich square-free words, the lengths of which give us a lower bound for $r(n)$. We will also make a conjecture that $r(n)$ can be achieved using these words. In Section 2.2 we will prove an upper bound for $r(n)$.

1.1. Repetition threshold

Square-free words are a special case of unavoidable repetitions of words, which has been a central topic in combinatorics on words since Thue (see [T1] and [T2]). The *repetition threshold*, on an alphabet of size n, is the smallest number r such that there exists an infinite word which avoids greater than r-powers. This number is denoted by $RT(n)$ and it was first studied in [D], where Dejean gave her famous conjecture. This conjecture has now been proven, in many parts and by several authors (see [R] and [CR]).

The repetition threshold can be studied also for a limited class of infinite words. In [MP], it was proven that the infinite Fibonacci word does not contain a power with exponent greater than $2+\varphi$, where φ is the golden ratio $\frac{\sqrt{5}+1}{2}$, but every smaller fractional power is contained. In [CD], the authors proved that among *Sturmian* words, the Fibonacci word is optimal with respect to this property. Sturmian words are equal to *episturmian* words when $n = 2$ (see [DJP]). This means the *episturmian repetition threshold* for $n = 2$ is $2 + \varphi$, denoted $ERT(2) = 2 + \varphi$. From [GJ], we get that the n-bonacci word is episturmian and it has critical exponent $2 + 1/(\varphi_n - 1)$, where φ_n is the generalized golden ratio. This means $ERT(n) \leq 2 + 1/(\varphi_n - 1)$. Notice, from [HPS] we get that φ_n converges to 2.

In the same way, we define the *rich repetition threshold* $RRT(n)$. From [PS] we get that $RRT(n) \geq 2$. Since episturmian words are rich (see [DJP]), we also know $RRT(n) \leq 2 + 1/(\varphi_n - 1)$ and $ERT(n) \geq 2$. This means $2 \leq RRT(n), ERT(n) \leq 2 + 1/(\varphi_n - 1)$. The exact values of $ERT(n)$ and $RRT(n)$ are left as an open problem.

Open problem 1.1. *Determine the repetition threshold for episturmian words and for rich words, on an alphabet of size n.*

1.2. Preliminaries

An *alphabet* A is a non-empty finite set of symbols, called *letters*. A *word* is a finite sequence of letters from A. The *empty* word ϵ is the empty sequence. The set A^* of all finite words over A is a *free monoid* under the operation of concatenation. The set $\mathrm{Alph}(w)$ is the set of all letters that occur in w. If $|\mathrm{Alph}(w)| = n$ then we say that w is *n-ary*.

An *infinite word* is a sequence indexed by $\mathbb{N}$ with values in A. We denote the set of all infinite words over A by A^ω and define $A^\infty = A^* \cup A^\omega$.

The *length* of a word $w = a_1 a_2 \ldots a_n$, with each $a_i \in A$, is denoted by $|w| = n$. The empty word ϵ is the unique word of length 0. By $|w|_a$ we denote the number of occurrences of a letter a in w.

A word x is a *factor* of a word $w \in A^\infty$, denoted $x \in w$, if $w = uxv$ for some $u \in A^*$,$v \in A^\infty$. If x is not a factor of w, we denote $x \notin w$. If $u = \epsilon$ (resp. $v = \epsilon$) then we say that x is a *prefix* (resp. *suffix*) of w. If $w = uv \in A^*$ is a word, we use the notation $u^{-1}w = v$ or $wv^{-1} = u$ to mean the removal of a prefix or a suffix of w. We say that a prefix or a suffix of w is *proper* if it is not the whole w.

A factor x of a word w is said to be *unioccurrent* in w if x has exactly one occurrence in w. Two occurrences of factor x are said to be *consecutive* if there is no occurrence of x between them. A factor of w having exactly two occurrences of a non-empty factor u, one as a prefix and the other as a suffix, is called a *complete return* to u in w.

The *reversal* of $w = a_1 a_2 \ldots a_n$ is defined as $\widetilde{w} = a_n \ldots a_2 a_1$. A word w is called a *palindrome* if $w = \widetilde{w}$. The empty word ϵ is assumed to be a palindrome.

Other basic definitions and notation in combinatorics on words can be found from Lothaire's books [Lot1] and [Lot2].

Proposition 1.2. *([DJP], Prop. 2) A word w has at most $|w| + 1$ distinct palindromic factors, including the empty word.*

Definition 1.3. *A word w is* rich *if it has exactly $|w| + 1$ distinct palindromic factors, including the empty word. An infinite word is* rich *if all of its factors are rich.*

Proposition 1.4. *([GJWZ], Thm. 2.14) A finite or infinite word w is rich if and only if all complete returns to any palindromic factor in w are themselves palindromes.*

Let $w = vu$ be a word and u its longest palindromic suffix. The *palindromic closure* of w is defined as $w^{(+)} = vu\widetilde{v}$. If u is the longest *proper* palindromic suffix of w, called lpps, we define the *proper palindromic closure* of w the same way as $w^{(++)} = vu\tilde{v}$. We refer to the longest proper palindromic prefix of w as lppp and define the *proper palindromic prefix closure* of w as ${}^{(++)}w = \widetilde{\widetilde{w}^{(++)}}$.

Proposition 1.5. *([GJWZ], Prop. 2.6) Palindromic closure preserves richness.*

Proposition 1.6. *([GJWZ], Prop. 2.8) Proper palindromic (prefix) closure preserves richness.*

2. The length of a longest rich square-free word

A word of form uu, where $u \neq \epsilon$, is called a *square* and a word w which does not have a square as a factor is called *square-free*. For example 1212 is a square and 01210 is square-free.

In [PS], Theorem 4 and Remark 6, it was proved that every infinite rich word contains a square. This means that every rich square-free word is of finite length. The length of a longest such word, on an alphabet of size n, is denoted with $r(n)$. The explicit formula for $r(n)$ was left as an open problem in [PS].

The first seven exact values of $r(n)$ are $r(1) = 1, r(2) = 3, r(3) = 7, r(4) = 15, r(5) = 33, r(6) = 67$ and $r(7) = 145$. These can be found from https://oeis.org/A269560. The longest rich square-free word on a given alphabet is not unique. Here are all the longest non-isomorphic ones, up to permutating the letters and taking the reversal, for $n = 1, \ldots, 7$:

$w_{1,1} = 1$
$w_{2,1} = 121$
$w_{3,1} = 2131213$
$w_{3,2} = 1213121$
$w_{4,1} = 131214121312141$
$w_{4,2} = 123121412131214$
$w_{4,3} = 213121343121312$
$w_{4,4} = 121312141213121$
$w_{5,1} = 421242131213531213124213121353135$
$w_{5,2} = 131242131213531213124213121353135$
$w_{6,1} = 1513121315131214121312141614121312141213151312141213121416141214161$
$w_{6,2} = 1214121315131214121312141614121312141213151312141213121416141214161$
$w_{6,3} = 4212421312135312131242131213531356531353121312421312135312131242124$
$w_{6,4} = 1312421312135312131242131213531356531353121312421312135312131242124$
$w_{6,5} = 5313531213124213121353121312421316131242131213531213124213121353135$

$w_{6,6} = 131242131213531213124213121353135653135312131242131213531213124213 1$
$w_{7,1} = 2421312135312131242131613124213121353121312421312135313575313531213 12$
$421312135312131242131613124213121353121312421312135313575313531213531357535 7$
$w_{7,2} = 2421312135312131242131613124213121353121312421312135313575313531213 12$
$421312135312131242131613124213121353121312421312135313575313531213531357531 3$
$w_{7,3} = 2421312135312131242124642124213121353121312421312135313575313531213 12$
$421312135312131242124642124213121353121312421312135313575313531213531357535 7$
$w_{7,4} = 2421312135312131242124642124213121353121312421312135313575313531213 12$
$421312135312131242124642124213121353121312421312135313575313531213531357531 3$

We can see that

$$w_{2,1} = w_{1,1}2w_{1,1},\ w_{3,2} = w_{2,1}3w_{2,1},\ w_{4,3} = w_{3,1}4\widetilde{w_{3,1}},\ w_{4,4} = w_{3,2}4w_{3,2},$$

$$w_{6,3} = w_{5,1}6\widetilde{w_{5,1}},\ w_{6,4} = w_{5,2}6\widetilde{w_{5,1}},\ w_{6,5} = \widetilde{w_{5,2}}6w_{5,2} \text{ and } w_{6,6} = w_{5,2}6\widetilde{w_{5,2}}.$$

Generally, we can construct rich square-free words by using a basic recursion

$$b_n = ba\widetilde{b},$$

where b is a longest rich square-free word over an $(n-1)$-ary alphabet A and $a \notin A$ is a new letter. It is very easy to see that b_n is rich and square-free. This gives us a recursive lower bound for $r(n)$: $r(n) \geq 2r(n-1)+1$, for all $n \geq 2$. We will use this inequality excessively later in Section 2.2, when we prove an upper bound for $r(n)$. The closed-form solution for the recursion $r(1) = 1, r(n) \geq 2r(n-1)+1$ is $r(n) \geq 2^n - 1$.

The case $n = 5$ reveals that the basic recursion $b_n = ba\widetilde{b}$ is not always optimal, since neither $w_{5,1}$ nor $w_{5,2}$ is of that form: $|w_{5,1}| = r(5) = 33 > 31 = 2 \cdot r(4) + 1$.

We can also see that

$$w_{3,1} = 2w_{1,1}3w_{1,1}2w_{1,1}3,\ w_{4,1} = 13w_{2,1}4w_{2,1}3w_{2,1}41,\ w_{4,2} = 213w_{2,1}4w_{2,1}3w_{2,1}4,$$

$$w_{5,1} = 42124w_{3,1}5\widetilde{w_{3,1}}4w_{3,1}53135,\ w_{5,2} = 13124w_{3,1}5\widetilde{w_{3,1}}4w_{3,1}53135,$$

$$w_{6,1} = 1513121315w_{4,1}6\widetilde{w_{4,1}}5w_{4,1}6141214161,\ w_{6,2} = 1214121315w_{4,1}6\widetilde{w_{4,1}}5w_{4,1}6141214161,$$

$$w_{7,1} = u_{1,2}6w_{5,2}7\widetilde{w_{5,2}}6w_{5,2}7v_{1,3},\ w_{7,2} = u_{1,2}6w_{5,2}7\widetilde{w_{5,2}}6w_{5,2}7v_{2,4},$$

$$w_{7,3} = u_{3,4}6w_{5,1}7\widetilde{w_{5,1}}6w_{5,1}7v_{1,3},\ w_{7,4} = u_{3,4}6w_{5,1}7\widetilde{w_{5,1}}6w_{5,1}7v_{2,4},$$

$$\text{where } u_{1,2} = 242131213531213124213 1, u_{3,4} = 2421312135312131242124,$$

$$v_{1,3} = 53135312135313575357 \text{ and } v_{2,4} = 53135312135313575313.$$

This gives us a hint how to get, in some cases, a better recursion than the basic recursion. We will define this recursion explicitly in the next subsection.

2.1. A lower bound

In this subsection, we will prove another lower bound for $r(n)$. We will use an alphabet $\{A_0, A_1, A_2, A_3, B_3, A_4, B_4, A_5, B_5, \ldots\}$. The following construction of rich square-free words w_n is recursive. The first six words are

$$w_1 = A_1, w_2 = A_0A_2A_0, w_3 = v_3A_3w_1B_3w_1A_3w_1B_3u_3, w_4 = v_4A_4w_2B_4w_2A_4w_2B_4u_4,$$

$$w_5 = v_5A_5w_3B_5w_3A_5w_3B_5u_5, w_6 = v_6A_6w_4B_6w_4A_6w_4B_6u_6,$$

where $v_3, u_3 = \epsilon$, $v_4, u_4 = A_0$, $v_5 = A_5A_3A_1A_3$, $u_5 = B_3A_1A_3A_1$, $v_6 = A_0A_6A_0A_4A_0A_2A_0A_4A_0$ and $u_6 = A_0B_4A_0A_2A_0A_4A_0A_2A_0$. Notice that w_6 is isomorphic ($\cong$) to $w_{6,2}$, $w_5 \cong w_{5,2}$, $w_4 \cong w_{4,1}$ and $w_3 \cong w_{3,1}$. For $n \geq 7$, we define

$$w_n = v_nA_nw_{n-2}B_n\widetilde{w_{n-2}}A_nw_{n-2}B_nu_n,$$

where $v_n = (P_nc_n)^{-1}\widetilde{v_{n-4}}A_{n-2}\widetilde{v_{n-2}}A_nv_{n-2}A_{n-2}v_{n-4}A_{n-4}w_{n-6}B_{n-4}\widetilde{w_{n-6}}A_{n-4}\widetilde{v_{n-4}}A_{n-2}\widetilde{v_{n-2}}$ and $u_n = \widetilde{u_{n-2}}B_{n-2}\widetilde{w_{n-4}}A_{n-2}w_{n-4}B_{n-2}\widetilde{u_{n-4}}B_{n-4}\widetilde{w_{n-6}}(d_n\widetilde{P_n})^{-1}$, where P_n is the largest common prefix of w_{n-6} and $\widetilde{v_{n-4}}$, c_n is the first letter of $(P_n)^{-1}\widetilde{v_{n-4}}A_{n-2}$ and d_n is the first letter of $(P_n)^{-1}w_{n-6}B_{n-4}$.

We can see that $\mathrm{Alph}(w_{2k}) = \{A_0, A_2, A_4, B_4, A_6, B_6, \ldots, A_{2k}, B_{2k}\}$ and $\mathrm{Alph}(w_{2k+1}) = \{A_1, A_3, B_3, A_5, B_5, \ldots, A_{2k+1}, B_{2k+1}\}$. This means we really have $\mathrm{Alph}(w_n) = n$. We also have $c_n \neq d_n$, since $A_{n-2} \notin w_{n-6}$ and $B_{n-4} \notin \widetilde{v_{n-4}}$.

Before we prove that w_n is rich and square-free, we will make some notation in order to make the proof look simpler. We mark that $E_n = A_nw_{n-2}B_n\widetilde{w_{n-2}}A_nw_{n-2}B_n$, $F_n = (P_nc_n)^{-1}\widetilde{v_{n-4}}A_{n-2}\widetilde{v_{n-2}}$, $G_n = \widetilde{w_{n-6}}A_{n-4}\widetilde{v_{n-4}}A_{n-2}\widetilde{v_{n-2}}$ and $H_n = \widetilde{P_n}A_{n-4}w_{n-6}B_{n-4}G_n$. Now $w_n = v_nE_nu_n$, $v_n = F_nA_n\widetilde{G_n}B_{n-4}G_n$ and $w_{n-2} = \widetilde{H_n}d_n\widetilde{u_n}$. We can also see that H_n is a suffix of v_n and F_n is a suffix of G_n.

Proposition 2.1. *The word w_n is square-free for all $n \geq 1$.*

Proof. We prove the claim by induction. It is easy to check that w_n is square-free when $1 \leq n \leq 6$. Suppose w_n is square-free for all $n < k$, where $k \geq 7$. Now we need to prove that w_k is square-free.

The word $A_kw_{k-2}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_ku_k$ is square-free because w_{k-2} is square-free, $A_k, B_k \notin w_{k-2}$ and $\widetilde{u_k}$ is a proper suffix of w_{k-2}. The words G_k and F_k are suffixes of $\widetilde{w_{k-2}}$ and $A_k, B_{k-4} \notin G_k, F_k$, which means that $v_k = F_kA_k\widetilde{G_k}B_{k-4}G_k$ is square-free.

Now, the only way $w_k = F_kA_k\widetilde{G_k}B_{k-4}G_kA_kw_{k-2}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_ku_k$ can have a square is either 1) $xA_kw_{k-2}B_kyxA_kw_{k-2}B_ky$, where x is a suffix of both v_k and $\widetilde{w_{k-2}}$, and y is a prefix of both u_k and $\widetilde{w_{k-2}}$, or 2) xA_kyxA_ky, where x is a suffix of both F_k and $\widetilde{G_k}B_{k-4}G_k$, and y is a prefix of both $\widetilde{w_{k-2}}$ and $\widetilde{G_k}B_{k-4}G_k$.

1) Case $xA_kw_{k-2}B_kyxA_kw_{k-2}B_ky$. Now $yx = \widetilde{w_{k-2}} = u_kd_kH_k$. Because y is a prefix of u_k and x is suffix of v_k, we have that d_kH_k is a suffix of v_k. We also know that c_kH_k is always a suffix of v_k. This is a contradiction since $c_k \neq d_k$.

2) Case xA_kyxA_ky. Now y is a prefix of $\widetilde{w_{k-2}}$, which means that x has to have a suffix $P_k^{-1}\widetilde{v_{k-4}}A_{k-2}\widetilde{v_{k-2}}$. This is a contradiction, since x is also a suffix of $(P_kc_k)^{-1}\widetilde{v_{k-4}}A_{k-2}\widetilde{v_{k-2}}$. □

Proposition 2.2. *The word w_n is rich for all $n \geq 1$.*

Proof. We prove the claim by induction. It is easy to check that w_n is rich when $1 \leq n \leq 6$. Suppose w_n is rich for all $n < k$, where $k \geq 7$. Now we need to prove that w_k is rich.

Since w_{k-2} is rich and $A_k, B_k \notin w_{k-2}$, we get that $A_kw_{k-2}B_k$ is rich. Proposition 1.5 gives now that $A_kw_{k-2}B_k\widetilde{w_{k-2}}A_k$ is rich. The lpps of $A_kw_{k-2}B_k\widetilde{w_{k-2}}A_k$ is A_k, which means $A_kw_{k-2}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_k\widetilde{w_{k-2}}A_k$ is rich by Proposition 1.6. The word u_k is a prefix of $\widetilde{w_{k-2}}$, so the factor $A_kw_{k-2}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_ku_k$ is also rich.

The lppp of $A_kw_{k-2}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_ku_k$ is $A_kw_{k-2}B_k\widetilde{w_{k-2}}A_k$, which means that also the proper palindromic prefix closure $\widetilde{u_k}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_ku_k$ is rich. The word H_k is a suffix of $\widetilde{w_{k-2}}$, which means $H_kA_kw_{k-2}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_ku_k = H_kE_ku_k$ is also rich.

The word $c_kH_kE_ku_k$ has a palindromic prefix $PP = c_k\widetilde{P_k}A_{k-4}w_{k-6}B_{k-4}\widetilde{w_{k-6}}A_{k-4}P_kc_k$. The following paragraph proves that it is unioccurrent in $c_kH_kE_ku_k$.

The letter B_{k-4} occurs only once in c_kH_k, in the middle of our palindromic prefix PP. This occurrence of B_{k-4} is preceded by $c_k\widetilde{P_k}A_{k-4}w_{k-6}$ and succeeded by $\widetilde{w_{k-6}}A_{k-4}P_kc_k$. The last occurrence of B_{k-4} in E_ku_k is succeeded by $\widetilde{w_{k-6}}(d_k\widetilde{P_k})^{-1}$ and nothing more. Since the word $\widetilde{w_{k-6}}(d_k\widetilde{P_k})^{-1}$ is clearly a proper prefix of $\widetilde{w_{k-6}}A_{k-4}P_kc_k$, this last occurrence of B_{k-4} in E_ku_k cannot occur in a factor PP. All other occurrences of B_{k-4} in E_ku_k are preceded by $B_{k-4}\widetilde{w_{k-6}}A_{k-4}w_{k-6}$ or succeeded by $\widetilde{w_{k-6}}A_{k-4}w_{k-6}B_{k-4}$. The word $B_{k-4}\widetilde{w_{k-6}}A_{k-4}w_{k-6}$ has a suffix $d_k\widetilde{P_k}A_{k-4}w_{k-6}$, which means that it cannot have a suffix $c_k\widetilde{P_k}A_{k-4}w_{k-6}$ because $c_k \neq d_k$. These mean that no B_{k-4} in $c_kH_kE_ku_k$ can occur in a factor PP, except the first one.

Since PP is unioccurrent palindromic prefix in $c_kH_kE_ku_k$, we get that $c_kH_kE_ku_k$ is rich and PP is the lppp of $c_kH_kE_ku_k$. Now, all we need to do is to take the proper palindromic prefix closure of $c_kH_kE_ku_k$, which is rich by Proposition 1.6. It has a suffix w_k, which concludes the proof:

$$^{(++)}(c_kH_kE_ku_k) = \widetilde{u_k}B_k\widetilde{w_{k-2}}A_kw_{k-2}B_k\widetilde{w_{k-2}}A_k\widetilde{G_k}B_{k-4}G_kE_ku_k$$

$$\overset{*}{=} XF_kA_k\widetilde{G_k}B_{k-4}G_kE_ku_k = Xv_kE_ku_k = Xw_k \text{ } (^*F_k \text{ is a suffix of } \widetilde{w_{k-2}}).$$

□

Now we know that w_n is rich and square-free, which means $r(n) \geq |w_n|$ for all $n \geq 1$. We can compute that $|w_7| = 145$, $|w_8| = 291$, $|w_9| = 629$ and $|w_{10}| = 1255$. Notice that $w_7 = w_{7,4}$, which means our lower bound is exact when $n = 7$. The cases $r(8)$ and $r(9)$ are too large to compute the exact value. However, by creating a partial tree of rich square-free words for $n = 8$ and 9, by leaving some branches out of it, the longest words we could find were of length 291 and 629, respectively. These are exactly the lengths of $|w_8|$ and $|w_9|$. Notice that $|w_8| = 291 = 2 \cdot 145 + 1 = 2|w_7| + 1$, which means the basic recursion b_n is as good as our recursion w_n when $n = 8$. Notice also that $|w_9| = 629 > 583 = 2 \cdot 291 + 1 = 2|w_8| + 1$ and $|w_{10}| = 1255 < 1259 = 2 \cdot 629 + 1 = 2|w_9| + 1$, which mean w_n is better than b_n when $n = 9$ and b_n is better than w_n when $n = 10$.

The previous paragraph suggests that it is reasonable to make the following conjecture.

Conjecture 2.3. *$r(n) = max\{|w_n|, 2 \cdot |w_{n-1}| + 1\}$ for all $n \geq 1$.*

The recursion for the length of w_n might be too complex to be solved in a closed-form, but we want to get at least an estimate for it. Let us first estimate the length of v_n, which will be used in Proposition 2.5.

Lemma 2.4. *$|v_n| \geq 3|v_{n-2}| + 2|w_{n-6}| + 2|v_{n-4}| + 6$, for $n \geq 7$.*

Proof.

$$|v_n| = |(P_n c_n)^{-1} \widetilde{v_{n-4}} A_{n-2} \widetilde{v_{n-2}} A_n v_{n-2} A_{n-2} v_{n-4} A_{n-4} w_{n-6} B_{n-4} \widetilde{w_{n-6}} A_{n-4} \widetilde{v_{n-4}} A_{n-2} \widetilde{v_{n-2}}|$$

$$\geq |\widetilde{v_{n-2}} A_n v_{n-2} A_{n-2} v_{n-4} A_{n-4} w_{n-6} B_{n-4} \widetilde{w_{n-6}} A_{n-4} \widetilde{v_{n-4}} A_{n-2} \widetilde{v_{n-2}}|$$

$$\geq 3|v_{n-2}| + 2|w_{n-6}| + 2|v_{n-4}| + 6,$$

where $|(P_n c_n)^{-1} \widetilde{v_{n-4}} A_{n-2}| \geq 0$, since c_n is a letter and P_n is a prefix of $\widetilde{v_{n-4}}$. □

Proposition 2.5. *$r(n) \geq |w_n| > 2,008^n$ for $n \geq 5$.*

Proof. From our recursion of w_n, we get that for $n \geq 11$:

$$|w_n| = |v_n A_n w_{n-2} B_n \widetilde{w_{n-2}} A_n w_{n-2} B_n u_n| = 3|w_{n-2}| + |v_n| + |u_n| + 4$$

$$= 3|w_{n-2}| + |(P_n c_n)^{-1} \widetilde{v_{n-4}} A_{n-2} \widetilde{v_{n-2}} A_n v_{n-2} A_{n-2} v_{n-4} A_{n-4} w_{n-6} B_{n-4} \widetilde{w_{n-6}} A_{n-4} \widetilde{v_{n-4}} A_{n-2} \widetilde{v_{n-2}}|$$

$$+ |\widetilde{u_{n-2}} B_{n-2} \widetilde{w_{n-4}} A_{n-2} w_{n-4} B_{n-2} \widetilde{u_{n-4}} B_{n-4} \widetilde{w_{n-6}} (d_n \widetilde{P_n})^{-1}| + 4$$

$$= 3|w_{n-2}| + |(P_n c_n)^{-1} \widetilde{v_{n-4}} A_{n-2} \widetilde{v_{n-2}} A_n v_{n-2} A_{n-2} v_{n-4}| - |d_n \widetilde{P_n}| + 4$$

$$+ |\widetilde{u_{n-2}} B_{n-2} \widetilde{w_{n-4}} A_{n-2} w_{n-4} B_{n-2} \widetilde{u_{n-4}} B_{n-4} \widetilde{w_{n-6}}| + |A_{n-4} w_{n-6} B_{n-4} \widetilde{w_{n-6}} A_{n-4} \widetilde{v_{n-4}} A_{n-2} \widetilde{v_{n-2}}|$$

$$= 4|w_{n-2}| + |(P_n c_n)^{-1}\widetilde{v_{n-4}}A_{n-2}\widetilde{v_{n-2}}A_n v_{n-2}A_{n-2}v_{n-4}| - |d_n\widetilde{P_n}| + 4$$
$$= 4|w_{n-2}| + 2(|\widetilde{v_{n-4}}| - |P_n|) + 2|v_{n-2}| + |A_{n-2}A_nA_{n-2}| - |d_n| - |c_n| + 4$$
$$\geq 4|w_{n-2}| + 2|v_{n-2}| + 5 \geq 4|w_{n-2}| + 2(3|v_{n-4}| + 2|w_{n-8}| + 2|v_{n-6}| + 6) + 5$$
$$\geq 4|w_{n-2}| + 2(3(3|v_{n-6}| + 2|w_{n-10}| + 2|v_{n-8}| + 6) + 2|w_{n-8}| + 2|v_{n-6}| + 6) + 5$$
$$> 4|w_{n-2}| + 4|w_{n-8}| + 12|w_{n-10}|.$$

From our recursion of w_n we also know that $|w_{10}| = 1255 > 1164 = 4|w_8|$, $|w_9| = 629 > 580 = 4|w_7|$, $|w_8| = 291 > 268 = 4|w_6|$ and $|w_7| = 145 > 132 = 4|w_5|$.

Now, for $n \geq 15$ we have

$$|w_n| > 4|w_{n-2}| + 4|w_{n-8}| + 12|w_{n-10}| > 4(4|w_{n-4}| + 4|w_{n-10}|) + 4|w_{n-8}| + 12|w_{n-10}|$$
$$= 16|w_{n-4}| + 4|w_{n-8}| + 28|w_{n-10}| > 16 \cdot 4 \cdot 4 \cdot 4|w_{n-10}| + 4 \cdot 4|w_{n-10}| + 28|w_{n-10}|$$
$$= 1068|w_{n-10}| > 2,008^{10}|w_{n-10}|.$$

We can also easily check that $|w_n| > 2,008^n$ for all $5 \leq n \leq 14$. This means we have our result

$$|w_n| > 2,008^n \text{ for } n \geq 5.$$

□

From the basic recursion b_n alone, we get $r(n) \geq 2^n - 1$. Our new recursion gives a slightly better bound $r(n) > 2,008^n$, which can be improved easily if we do not estimate the length of w_n in Proposition 2.5 that roughly. We only mention that it can be improved at least to $2,0178^n$, but we will not do it here.

2.2. An upper bound

In this subsection, we will prove an upper bound for $r(n)$. First, we will prove two useful lemmas. For that, let us mention that every square-free palindrome has to be of odd length, because palindromes of even length create a square of two letters to the middle, for example 12011021 has a square 11 in the middle.

Lemma 2.6. *The middle letter of a rich square-free palindrome is unioccurrent.*

Proof. Since all square-free palindromes are of odd length, there always exists the middle letter. Then, suppose the contrary: $zb\widetilde{z}$ is rich and square-free and the letter b has another occurrence inside z. We can take the other occurrence of b to be consecutive to the b in the middle and get that $zb\widetilde{z} = z_1bz_2bz_2b\widetilde{z}_1$, where z_2 is a palindrome because of Proposition 1.4. We get a contradiction because bz_2bz_2 is a square. □

Lemma 2.7. *Suppose $w = u_1a_1u_2a_1 \cdots a_1u_{k-1}a_1u_k \in \{a_1, a_2, \dots, a_n\}^*$ is rich and square-free, where $n, k \geq 3$ (possibly $u_k = \epsilon$), Alph$(u_1) = \{a_2, \dots, a_n\}$ and $\forall i : a_1 \notin$ Alph(u_i).*

$$\textit{For } 2 \leq i \leq k-1: \ \textit{Alph}(u_{i+1}) \subseteq \textit{Alph}(u_i) \setminus \{a_i\}, \ \textit{where } u_i = v_ia_i\widetilde{v_i}.$$

Proof. Since $\forall i : a_1 \notin \text{Alph}(u_i)$, we get from Proposition 1.4 that $u_2, \dots, u_{k-1}$ are palindromes, and because w is square-free, they are of odd length and non-empty. By permutating the letters, we can suppose for $2 \leq i \leq k-1$: a_i is the middle letter of $u_i = v_ia_i\widetilde{v_i}$, where $a_i \notin \text{Alph}(v_i)$ by Lemma 2.6.

We will prove the claim by induction on i.

1) The base case $i = 2$. Since $a_2 \in \text{Alph}(u_1) = \{a_2, \dots, a_n\}$, we get from Proposition 1.4 that $u_1 = v_1a_2\widetilde{v_2}$. If $a_2 \in \text{Alph}(u_3)$ then, by Proposition 1.4, we have $u_3 = v_2a_2v_3'$, which creates a square $(a_2\widetilde{v_2}a_1v_2)^2$ in $u_1a_1u_2a_1u_3 = v_1a_2\widetilde{v_2}a_1v_2a_2\widetilde{v_2}a_1v_2a_2v_3'$. This means $a_2 \notin \text{Alph}(u_3)$.

Suppose then that $b \in \text{Alph}(u_3) \setminus \text{Alph}(u_2)$, which implies $b \in \text{Alph}(v_1)$. The word between the first occurrence of b in u_3 and the last occurrence of b in v_1 is a palindrome by Proposition 1.4: $u_1a_1u_2a_1u_3 = t_1bt_2a_2\widetilde{v_2}a_1v_2a_2\widetilde{v_2}a_1v_2a_2\widetilde{t_2}bt_3$, where $v_1 = t_1bt_2$ and $u_3 = v_2a_2\widetilde{t_2}bt_3$. We get a contradiction since we have a square $(a_2\widetilde{v_2}a_1v_2)^2$. This means $\text{Alph}(u_3) \subseteq \text{Alph}(u_2) \setminus \{a_2\}$.

2) The induction hypothesis. We can now suppose $k \geq 4$, since the base case proves our claim if $k = 3$. Suppose then that for every j, where $2 \leq j \leq i < k-1$, we have: $\text{Alph}(u_{j+1}) \subseteq \text{Alph}(u_j) \setminus \{a_j\}$.

3) The induction step. Now we need to prove that $\text{Alph}(u_{i+2}) \subseteq \text{Alph}(u_{i+1}) \setminus \{a_{i+1}\}$. From the induction hypothesis we get that $a_{i+1} \in \text{Alph}(u_{i+1}) \subseteq \text{Alph}(u_i) \setminus \{a_i\}$, which means $u_i = v_{i+1}a_{i+1}xa_i\widetilde{x}a_{i+1}\widetilde{v_{i+1}}$ by Proposition 1.4. If $a_{i+1} \in \text{Alph}(u_{i+2})$ then, by Proposition 1.4, we have $u_{i+2} = v_{i+1}a_{i+1}y$, which creates a square $(a_{i+1}\widetilde{v_{i+1}}a_1v_{i+1})^2$ inside $u_ia_1u_{i+1}a_1u_{i+2} = v_{i+1}a_{i+1}xa_i\widetilde{x}a_{i+1}\widetilde{v_{i+1}}a_1v_{i+1}a_{i+1}\widetilde{v_{i+1}}a_1v_{i+1}a_{i+1}y$. This means $a_{i+1} \notin \text{Alph}(u_{i+2})$

Suppose then that $c \in \text{Alph}(u_{i+2}) \setminus \text{Alph}(u_{i+1})$, which implies $c \in \text{Alph}(u_1a_1 \dots a_1u_i)$. Without loss of generality, we can assume that c is the letter from $\text{Alph}(u_{i+2}) \setminus \text{Alph}(u_{i+1})$ that has the rightmost occurrence in $u_1a_1 \dots a_1u_i$. The word between the leftmost occurrence of c in $u_{i+2} = zcz'$ and the rightmost occurrence of c in $u_1a_1 \dots a_1u_i$ has to be a palindrome by Proposition 1.4. We divide this into two cases.

- Suppose $c \notin \text{Alph}(u_i)$. Now $c\widetilde{z}a_1u_{i+1}Pu_{i+1}a_1zc$ is a palindrome, where $\text{Alph}(P) \subseteq \text{Alph}(a_1u_{i+1})$ because of the way we chose c. Now the middle letter of the palindrome $a_1u_{i+1}Pu_{i+1}a_1$ belongs to P and therefore has other occurrences inside it, in a_1u_{i+1} and in $u_{i+1}a_1$. This is a contradiction by Lemma 2.6.

- Suppose $c \in \text{Alph}(u_i)$. Now $c\widetilde{z}a_1v_{i+1}a_{i+1}\widetilde{v_{i+1}}a_1zc$ is a palindrome, where a_{i+1} is its middle letter and $c\widetilde{z}$ is a suffix of u_i. If $a_{i+1} \in \text{Alph}(z)$ then it is not unioccurrent in the palindrome $\widetilde{z}a_1v_{i+1}a_{i+1}\widetilde{v_{i+1}}a_1z$ and we get a contradiction by Lemma 2.6. Since $a_{i+1} \in \text{Alph}(u_i)$ by the induction hypothesis, we can take the rightmost occurrence of it in u_i and get that $a_{i+1}v_i'c\widetilde{z}a_1v_{i+1}a_{i+1}$ is a palindrome, where $v_i'c\widetilde{z} = \widetilde{v_{i+1}}$. We get a contradiction since this would mean $c \in \text{Alph}(v_{i+1}) \subset \text{Alph}(u_{i+1})$.

Both cases yield a contradiction, which means $\text{Alph}(u_{i+2}) \subseteq \text{Alph}(u_{i+1}) \setminus \{a_{i+1}\}$. □

Corollary 2.8. *All rich square-free words are finite.*

Proof. We prove this by induction. Suppose w is rich and square-free word for which $|\text{Alph}(w)| = n \geq 4$. Suppose that all rich square-free words on an alphabet of size $n-1$ or smaller are finite. Cases $n = 1, 2, 3$ are trivial.

Suppose that a_1 is the letter of w for which $w = u_1a_1w'$, where $\text{Alph}(u_1) = \text{Alph}(w)\setminus\{a_1\}$. We partition w such that $w = u_1a_1u_2a_1u_3a_1u_4a_1\ldots$, where $a_1 \notin \text{Alph}(u_i)$ for all i. From Lemma 2.7 we now get that $|\text{Alph}(u_i)| > |\text{Alph}(u_{i+1})|$ for all $i \geq 2$. This means there are finitely many words u_i, at most n, and they are all over an alphabet of size $n-1$ or smaller, which concludes the proof. □

The above corollary gives another proof for the result mentioned in Remark 6 of [PS]. The proof of the above corollary also gives us a way to get an upper bound for $r(n)$: $r(n) \leq r(n-1) + 1 + \sum_{i=1}^{n-1}(r(n-i)+1)$. This bound can be easily improved if we examine the word also from the right side, i.e. we suppose that a_1 is the letter of w for which $w = w'a_1u_1$, where $\text{Alph}(u_1) = \text{Alph}(w) \setminus \{a_1\}$. This notice makes it reasonable to make the following definition.

Definition 2.9. *Let $w = uav$ be a word, where a is a letter. If $Alph(u) = Alph(w) \setminus \{a\}$ then the leftmost occurrence of the letter a in w is called the* left special letter *of w. If $Alph(v) = Alph(w) \setminus \{a\}$ then the rightmost occurrence of the letter a in w is called the* right special letter *of w.*

In Subsection 2.1, where we constructed the words w_n for our lower bound, the rightmost occurrence of A_n is always the right special letter of w_n and the leftmost occurrence of B_n is always the left special letter of w_n, for $n \geq 3$. In Lemma 2.7 and Corollary 2.8, the first occurrence of letter a_1 is the left special letter of w.

Before we go to our upper bound for $r(n)$, we will state a helpful lemma.

Lemma 2.10. *Suppose $w_n = xB_nyA_nz$ is a rich square-free n-ary word, where $n \geq 3$ and the letters A_n and B_n are the right and left special letters of w_n, respectively. Now $Alph(y) = Alph(w_n) \setminus \{A_n, B_n\}$ and $A_n \neq B_n$.*

Proof. First we prove that $A_n, B_n \notin y$. Suppose to the contrary that $B_n \in y$ (case $A_n \in y$ is symmetric). We can take the leftmost occurrence of B_n in y and get that $w_n = xB_n y_1 c\widetilde{y_1} B_n y_2 A_n z$, where $B_n \notin y_1 c\widetilde{y_1}$ and c is a letter. Since A_n is the right special letter of w_n, we have that $c \in z$. Since B_n is the left special letter of w_n, we get from Lemma 2.7 that $c \notin y_2 A_n z$, i.e. $c \notin z$. This is a contradiction.

Then we prove that $A_n \neq B_n$. Suppose to the contrary that $A_n = B_n$. Now, since $A_n, B_n \notin y$, we get from Proposition 1.4 that y is a palindrome. From Lemma 2.7 we get that the middle letter of y cannot be in x nor in z. This is a contradiction, since x and z has to contain all the letters except A_n.

Then we prove that if $a \in \mathrm{Alph}(w_n) \setminus \{A_n, B_n\}$ then $a \in y$. Suppose to the contrary that $a \in \mathrm{Alph}(w_n) \setminus (\{A_n, B_n\} \cup \mathrm{Alph}(y))$. Since A_n and B_n are the right and left special letters, we have that $a \in x, z$. If we take the leftmost occurrence of a in z and the rightmost occurrence of a in x, then we get from Proposition 1.4 that $w = x'auB_n y A_n vaz'$, where $auB_n y A_n va$ is a palindrome, $x = x'au$ and $z = vaz'$. The middle letter of the palindrome $auB_n y A_n va$ cannot be inside u nor v, since it would mean $B_n \in u$ or $A_n \in v$, which is impossible since A_n and B_n are special letters. The middle letter of $auB_n y A_n va$ cannot be inside y neither, since that would mean $B_n \in yA_n$ or $A_n \in B_n y$, which we proved above to be impossible. The only possibility is that the middle letter of $auB_n y A_n va$ is either A_n or B_n. Since these cases are symmetric, we can suppose B_n is the middle letter. This means $w = x'a\widetilde{v}A_n\widetilde{y}B_n y A_n vaz'$. Since B_n is the left special letter of w, we have $B_n \in z = vaz'$ and $B_n \notin v$. This means $B_n \in z'$. If we take the leftmost occurrence of B_n in z', we get $w = x'a\widetilde{v}A_n\widetilde{y}B_n y A_n vav'B_n z''$, where $B_n y A_n vav' B_n$ is a palindrome which has A_n as the middle letter. This means $\widetilde{y} = vav'$ and hence $a \in y$, which is a contradiction. ☐

There are only three cases how the right and left special letters can appear inside a word, with respect to each other. If w_n is a rich square-free n-ary word which has A_n and B_n as the right and left special letters, respectively, then one the following cases must hold (the visible occurrences of A_n and B_n in w_n are the special letters):

1) $w_n = xB_n y A_n z$. Now $A_n \neq B_n$ by Lemma 2.10.

2) $w_n = xA_n y B_n z$. Now $A_n \neq B_n$ by the definition of special letters.

3) $w_n = xA_n z = xB_n z$. Now $A_n = B_n$.

Proposition 2.11. *Suppose w_n is a rich square-free n-ary word, where $n \geq 3$.*

1) If $w_n = xB_n y A_n z$, where the letters A_n and B_n are the right and left special letters of w_n, respectively, then $|w_n| \leq 2r(n-1) + r(n-2) + 2$.

2) If $w_n = xA_n y B_n z$, where the letters A_n and B_n are the right and left special letters of w_n, respectively, then $|w_n| \leq r(n-1) + r(n-2) + r(n-3) + 2 \leq 2r(n-1)$ and $|x|, |z| \leq r(n-2) + r(n-3) + 1$, where $r(n-3) = 0$ if $n = 3$.

3) If $w_n = xA_nz = xB_nz$, where the letter $A_n = B_n$ is both the right and left special letter of w_n, then $|w_n| \leq 2r(n-1)+1$.

Proof. Let us denote $A = \text{Alph}(w_n)$.

1) By the definition of special letters, we have that $\text{Alph}(x) = A\setminus\{B_n\}$ and $\text{Alph}(z) = A\setminus\{A_n\}$. These mean $|x|, |z| \leq r(n-1)$. From Lemma 2.10 we get that $\text{Alph}(y) = A\setminus\{A_n, B_n\}$, which means $|y| \leq r(n-2)$, since $A_n \neq B_n$. Now

$$|w_n| = |x|+|B_n|+|y|+|A_n|+|z| \leq r(n-1)+1+r(n-2)+1+r(n-1) = 2r(n-1)+r(n-2)+2.$$

2) If $A_n \notin x$, then $|x| \leq r(n-2)$. If $A_n \in x$ then we can take the rightmost occurrence of it in x and get that $xA_n = x_2A_nx_1c\widetilde{x_1}A_n$, where $A_n \notin x_1c\widetilde{x_1}$ and by Lemma 2.7 $c \notin x_2A_nx_1$. Now $\text{Alph}(x_2A_nx_1) = A \setminus \{c, B_n\}$ and $\text{Alph}(\widetilde{x_1}) = A \setminus \{c, A_n, B_n\}$, where $c \neq B_n$ since B_n is the left special letter of w_n. This means $|x| = |x_2A_nx_1| + |c| + |\widetilde{x_1}| \leq r(n-2) + r(n-3) + 1$, where $r(n-3) = 0$ if $n = 3$. The same holds for z.

We have $\text{Alph}(yB_nz) = A \setminus \{A_n\}$, which means $|yB_nz| \leq r(n-1)$. Now

$$|w_n| = |x|+|A_n|+|yB_nz| \leq [r(n-2)+r(n-3)+1]+1+r(n-1) = r(n-1)+r(n-2)+r(n-3)+2.$$

From the basic recursion we know that $r(n) \geq 2r(n-1)+1$. This means that $r(n-1)+r(n-2)+r(n-3)+2 \leq r(n-1)+r(n-2)+2r(n-3)+2 \leq r(n-1)+2r(n-2)+1 \leq 2r(n-1)$, which we needed to prove.

3) By the definition of special letters, we have that $\text{Alph}(x) = \text{Alph}(z) = A\setminus\{A_n\}$, which means $|x|, |z| \leq r(n-1)$. Now

$$|w_n| = |x| + |A_n| + |z| \leq r(n-1) + 1 + r(n-1) = 2r(n-1) + 1.$$

□

Corollary 2.12. *$r(n) \leq 2r(n-1) + r(n-2) + 2$, for $n \geq 3$.*

Proof. We get our claim from Proposition 2.11, since the proposition covered all the three different possible cases for w_n. □

We do not solve the recursion $r(n) \leq 2r(n-1) + r(n-2) + 2, r(2) = 3, r(1) = 1$, in a closed-form, but we will estimate it. We use the inequality $r(n) \geq 2r(n-1) + 1$ from the basic recursion, and the fact that $r(4) = 15 > 13$. For $n \geq 8$ we have

$$r(n) \leq 2r(n-1)+r(n-2)+2 \leq 2(2r(n-2)+r(n-3)+2)+r(n-2)+2 = 5r(n-2)+2r(n-3)+6$$

$$\leq 5(2r(n-3) + r(n-4) + 2) + 2r(n-3) + 6 = 12r(n-3) + 5r(n-4) + 16$$

$$< 12r(n-3)+5r(n-4)+16+(r(n-4)-13) = 12r(n-3)+6r(n-4)+3 \le 15r(n-3)$$
$$< 2,47^3 r(n-3) < 2,47^n,$$

where the last inequality comes from the fact that $r(n) < 2,47^n$ for $1 \le n \le 7$. Together with the lower bound, we now have $2,008^n < r(n) < 2,47^n$ for $n \ge 5$.

This upper bound can still be improved. The cases 2 and 3 from Proposition 2.11 already give better or equal upper bounds than the basic recursion, i.e. $r(n) \le 2r(n-1)+1$. This means we need to look closer only for the case 1.

Proposition 2.13. *$r(n) \le 5r(n-2)+4$, for $n \ge 7$.*

Proof. Suppose $w_n = xB_nyA_nz$ is a rich square-free n-ary word, where $n \ge 7$ and the letters A_n and B_n are the right and left special letters of w_n, respectively. This means $A_n \ne B_n$. If w_n is not of this form, then we already know from Proposition 2.11 that $|w_n| \le 2r(n-1)+1$, which means we can use the upper bound of Corollary 2.12 and get that

$$|w_n| \le 2(2r(n-2)+r(n-3)+2)+1 = 4r(n-2)+2r(n-3)+5 \le 5r(n-2)+4,$$

where the last inequality comes from the basic recursion $r(n) \ge 2r(n-1)+1$. From now on, we will use the basic recursion without mentioning it.

By the definition of special letters, we have that $A_n \in x$ and $B_n \in z$. From Lemma 2.10 we know that $A_n, B_n \notin y$. Since $A_n \ne B_n$, we can take the rightmost occurrence of A_n in x and the leftmost occurrence of B_n in z and get, by Proposition 1.4, that $w_n = x_1A_n\widetilde{y}B_nyA_n\widetilde{y}B_nz_1$.

We divide this proof into three different cases depending whether $A_n \in x_1$ or $A_n \notin x_1$ and whether $B_n \in z_1$ or $B_n \notin z_1$.

Case 1) $A_n \notin x_1, B_n \notin z_1$.

Now we have $A_n, B_n \notin x_1, z_1, y$. This means $|x_1|, |z_1|, |y| \le r(n-2)$. Together we get

$$|w_n| = |x_1A_n\widetilde{y}B_nyA_n\widetilde{y}B_nz_1| \le 5r(n-2)+4.$$

Case 2) $A_n \in x_1, B_n \notin z_1$ (the case $A_n \notin x_1, B_n \in z_1$ is symmetric).

If we take the rightmost occurrence of A_n in x_1 we get, by Proposition 1.4, Lemma 2.6 and Lemma 2.7, that $w_n = x_2A_n\widetilde{x_B}Bx_BA_n\widetilde{y}B_nyA_n\widetilde{y}B_nz_1$, where B ($\ne A_n, B_n$) is a letter, $A_n, B \notin x_B$, $B \notin x_2$ and $x_1 = x_2A_n\widetilde{x_B}Bx_B$. Since B_n is a left special letter of w_n, we have that $B_n \notin x_2A_n\widetilde{x_B}$ and $B_n \notin x_B$. We also have $A_n, B_n \notin y, z_1$. Together we have $|y|, |z_1|, |x_2A_n\widetilde{x_B}| \le r(n-2)$ and $|x_B| \le r(n-3)$.

Let us mark the left special letter of $\widetilde{y}$ with B_{n-2}. Now we divide this into two cases whether $B \ne B_{n-2}$ or $B = B_{n-2}$.

Case 2.1) $B \ne B_{n-2}$.

Since B_{n-2} is the left special letter of $\widetilde{y}$, we must have $B_{n-2} \notin x_B$. Otherwise we would have, by Proposition 1.4, that $B \in x_B$, which is impossible by Lemma 2.6. From Lemma 2.7 we now get that $B_{n-2} \notin x_2$. Earlier, we already noted that $B_n, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B \notin x_B$. Together we now get $|x_2 A_n \widetilde{x_B}| \leq r(n-3)$ and $|x_B| \leq r(n-4)$, and therefore

$$|w_n| = |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_1|$$

$$\leq r(n-3) + 1 + r(n-4) + [3r(n-2)+4] + r(n-2) = 4r(n-2) + r(n-3) + r(n-4) + 5$$

$$< 4r(n-2) + r(n-3) + r(n-4) + 5 + r(n-4) \leq 5r(n-2) + 3,$$

where we added the extra $r(n-4)$ after the second inequality only to make the use of the basic recursion simpler.

Case 2.2) $B = B_{n-2}$.

If we can prove that $|z_1| \leq r(n-3)$, then we get

$$|w_n| = |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_1|$$

$$\leq r(n-2) + 1 + r(n-3) + [3r(n-2)+4] + r(n-3) = 4r(n-2) + 2r(n-3) + 5 \leq 5r(n-2) + 4.$$

So we need to prove there exists some letter, different from A_n and B_n, such that it does not belong to z_1. We divide this into three cases depending of which form $\widetilde{y}$ is.

Case 2.2.1) $\widetilde{y} = y_1 A_{n-2} y_3 B_{n-2} y_2$, where the letters A_{n-2} and B_{n-2} are the right and left special letters of $\widetilde{y}$, respectively.

Because $B = B_{n-2}$, we have $\widetilde{x_B} = y_1 A_{n-2} y_3$, by Proposition 1.4 and Lemma 2.6. Now $A_{n-2} \notin z_1$, since otherwise we could take the leftmost occurrence of A_{n-2} in z_1 and get a square in w_n:

$$\widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2},$$

where the rightmost $\widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2}$ is a prefix of z_1 and the leftmost $\widetilde{y_1}$ is a suffix x_1.

Case 2.2.2) $\widetilde{y} = y_1 B_{n-2} y_2$, where B_{n-2} is also the right special letter of $\widetilde{y}$.

Because $B = B_{n-2}$, we have $\widetilde{x_B} = y_1$. Now $B_{n-2} \notin z_1$, since otherwise, similar to Case 2.2.1, we could take the leftmost occurrence of B_{n-2} in z_1 and get a square in w_n:

$$\widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2}.$$

Case 2.2.3) $\widetilde{y} = y_1 A_{n-2} y_3 B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} y_2$, where the rightmost A_{n-2} and the leftmost B_{n-2} are the right and left special letters of $\widetilde{y}$, respectively.

Because $B = B_{n-2}$, we have $\widetilde{x_B} = y_1 A_{n-2} y_3$. Again $A_{n-2} \notin z_1$, since otherwise, similar to Case 2.2.1, we could take the leftmost occurrence of A_{n-2} in z_1 and get a square in w_n:

$$y_3 B_{n-2} \widetilde{y_3} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{y_3} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2}.$$

Case 3) $A_n \in x_1, B_n \in z_1$.

If we take the rightmost occurrence of A_n in x_1 and the leftmost occurrence of B_n in z_1, we get that $w_n = x_2 A_n \widetilde{x_B} B x_B A_n \widetilde{y} B_n y A_n \widetilde{y} B_n z_A A \widetilde{z_A} B_n z_2$, where A, B ($\neq A_n, B_n$) are letters and $x_1 = x_2 A_n \widetilde{x_B} B x_B$, $z_1 = z_A A \widetilde{z_A} B_n z_2$. Similar to Case 2, we have $|y|, |x_1|, |z_1|, |x_2 A_n \widetilde{x_B}|$, $|\widetilde{z_A} B_n z_2| \leq r(n-2)$ and $|x_B|, |z_A| \leq r(n-3)$.

We divide this case now into three cases depending of which form $\widetilde{y}$ is.

Case 3.1) $\widetilde{y} = y_1 B_{n-2} y_2$, where B_{n-2} is both the right and left special letter of $\widetilde{y}$.

If $A = B = B_{n-2}$ then $x_B = \widetilde{y_1}$ and $z_A = \widetilde{y_2}$. This would create a square in w_n:

$$B_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2}.$$

Now we divide this into two possible cases: $A, B \neq B_{n-2}$ and $A = B_{n-2}$,$B \neq B_{n-2}$.

Case 3.1.1) $A, B \neq B_{n-2}$.

Similar to Case 2.1, we get $B_n, B_{n-2}, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B \notin x_B$. In the same way, we get $A_n, A, B_{n-2} \notin \widetilde{z_A} B_n z_2$ and $A_n, A, B_n, B_{n-2} \notin z_A$. Together we have

$$|w_n| = |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_A| + |A| + |\widetilde{z_A} B_n z_2|$$

$$\leq r(n-3)+1+r(n-4)+[3r(n-2)+4]+r(n-4)+1+r(n-3) = 3r(n-2)+2r(n-3)+2r(n-4)+6$$

$$< 3r(n-2) + 2r(n-3) + 2r(n-4) + 6 + 2r(n-4) \leq 5r(n-2) + 2.$$

Case 3.1.2) $A = B_{n-2}$ and $B \neq B_{n-2}$ (the case $A \neq B_{n-2}$ and $B = B_{n-2}$ is symmetric).

Now $z_A = \widetilde{y_2}$. Let us mark $y_1 = u_1 B_{n-4} u_2$ and $y_2 = v_1 A_{n-4} v_2$, where B_{n-4} and A_{n-4} are the left special letters of y_1 and y_2, respectively.

We prove $B \neq B_{n-4}$. Suppose to the contrary that $B = B_{n-4}$. Since B_{n-2} is the right and left special letter of $\widetilde{y}$, we have that $A_{n-4} \in y_1$. If we take the rightmost occurrence of A_{n-4} in y_1 then we get from Proposition 1.4 that $A_{n-4} \widetilde{v_1}$ is a suffix of y_1 and hence A_{n-4} is the right special letter of y_1. There are now three different cases how A_{n-4} and B_{n-4} can appear inside y_1 with respect to each other. These all yield a square and hence a contradiction:

- If $y_1 = u_1' A_{n-4} u_3 B_{n-4} \widetilde{u_3} A_{n-4} u_3 B_{n-4} u_2'$, where $u_1 = u_1' A_{n-4} u_3$, $u_2 = \widetilde{u_3} A_{n-4} u_3 B_{n-4} u_2'$ and $v_1 = \widetilde{u_2'} B_{n-4} \widetilde{u_3}$, then we have a square in w_n:

$$A_{n-4} u_3 B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1'} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{u_2'} B_{n-4} \widetilde{u_3} A_{n-4} u_3 B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1'} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{u_2'} B_{n-4} \widetilde{u_3}.$$

- If $y_1 = u_1 B_{n-4} u_2 = u_1 A_{n-4} \widetilde{v_1}$ (i.e. $A_{n-4} = B_{n-4}$), then $u_2 = \widetilde{v_1}$ and we have a square in w_n:

$$B_{n-4} \widetilde{u_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{u_2}.$$

- If $y_1 = u_1' A_{n-4} u_3 B_{n-4} u_2$, where $u_1 = u_1' A_{n-4} u_3$ and $v_1 = \widetilde{u_2} B_{n-4} \widetilde{u_3}$, then we have a square in w_n:

$$B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1'} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1'} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{u_2}.$$

This means $B \neq B_{n-4}$. Similar to Case 2.1 we now get that $B_n, B_{n-2}, B_{n-4}, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B_{n-4}, B \notin x_B$. Together we have

$$|w_n| = |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_A| + |B_{n-2}| + |\widetilde{z_A} B_n z_2|$$

$$\leq r(n-4) + 1 + r(n-5) + [3r(n-2) + 4] + r(n-3) + 1 + r(n-2)$$

$$= 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 6$$

$$< 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 6 + r(n-5) \leq 5r(n-2) + 3.$$

Case 3.2) $\widetilde{y} = y_1 A_{n-2} y_3 B_{n-2} y_2$, where the letters A_{n-2} and B_{n-2} are the right special letter and the left special letter of $\widetilde{y}$, respectively.

If $A = A_{n-2}$, $B = B_{n-2}$ then we would have a square in w_n:

$$A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3}.$$

This means we can divide this case, similar to Case 3.1, into two different cases: $A = A_{n-2}$, $B \neq B_{n-2}$ and $A \neq A_{n-2}$, $B \neq B_{n-2}$.

Case 3.2.1) $A \neq A_{n-2}$, $B \neq B_{n-2}$.

Similar to Case 2.1, we get $B_n, B_{n-2}, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B \notin x_B$. In the same way, we get $A_n, A_{n-2}, A \notin \widetilde{z_A} B_n z_2$ and $A_n, A_{n-2}, A, B_n \notin z_A$. Together we have

$$|w_n| = |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_A| + |A| + |\widetilde{z_A} B_n z_2|$$

$$\leq r(n-3) + 1 + r(n-4) + [3r(n-2) + 4] + r(n-4) + 1 + r(n-3) = 3r(n-2) + 2r(n-3) + 2r(n-4) + 6$$

$$< 3r(n-2) + 2r(n-3) + 2r(n-4) + 6 + 2r(n-4) \leq 5r(n-2) + 2.$$

Case 3.2.2) $A = A_{n-2}$, $B \neq B_{n-2}$ (the case $A \neq A_{n-2}$,$B = B_{n-2}$ is symmetric).

Now $z_A = \widetilde{y_2} B_{n-2} \widetilde{y_3}$. We divide this case into two cases: $A_{n-2} \notin y_1$ and $A_{n-2} \in y_1$.

Case 3.2.2.1) $A_{n-2} \notin y_1$.

We must have $A_{n-2} \notin x_1$. Otherwise we could take the rightmost occurrence of A_{n-2} in x_1 and get a square in w_n:

$$A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3}.$$

Similar to Case 2.1, we have $B_n, B_{n-2} \notin x_1$. Since B_n and B_{n-2} are the left special letters of w_n and $\widetilde{y}$, respectively, we have $B_n, B_{n-2} \notin y_1$. Together with the previous paragraph we get that $A_{n-2}, B_n, B_{n-2} \notin x_1 A_n y_1$. Since A_{n-2} is the right special letter of $\widetilde{y}$ we have $A_n, B_n, A_{n-2} \notin y_3 B_{n-2} y_2$. These mean $|x_1 A_n y_1| \leq r(n-3)$ and $|y_3 B_{n-2} y_2| \leq r(n-3)$. Together we have

$$|w_n| = |x_1 A_n y_1| + |A_{n-2}| + |y_3 B_{n-2} y_2| + |B_n y A_n \widetilde{y} B_n| + |z_A| + |A_{n-2}| + |\widetilde{z_A} B_n z_2|$$

$$\le r(n-3)+1+r(n-3)+[2r(n-2)+3]+r(n-3)+1+r(n-2)$$

$$= 3r(n-2)+3r(n-3)+5 < 3r(n-2)+3r(n-3)+5+r(n-3) \le 5r(n-2)+3.$$

Case 3.2.2.2) $A_{n-2} \in y_1$.

If we take the rightmost occurrence of A_{n-2} in y_1, we get $\widetilde{y} = y_1' A_{n-2} y_4 B_y \widetilde{y_4} A_{n-2} y_3 B_{n-2} y_2$, where B_y is a letter, $y_1 = y_1' A_{n-2} y_4 B_y \widetilde{y_4}$ and $A_{n-2} \notin y_4 B_y \widetilde{y_4}$. Let us mark $y_3 = u_1 B_{n-4} u_2$, where B_{n-4} is the left special letter of y_3. We will prove $B_{n-4} \notin x_1$.

Suppose $B_{n-4} \notin y_1$. Now $B_{n-4} \notin x_1$, since otherwise we could take the rightmost occurrence of B_{n-4} in x_1 and get a square in w_n:

$$A_{n-2}\widetilde{y_1}A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}A_{n-2}\widetilde{y_1}A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}.$$

Suppose $B_{n-4} \in y_1$. Because of Lemma 2.6, we must have that $B_y = B_{n-4}$ and $y_4 = \widetilde{u}_1$. Also now $B_{n-4} \notin x_1$, since otherwise we would have a square in w_n:

$$B_{n-4}\widetilde{u}_1A_{n-2}\widetilde{y'}_1A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}A_{n-2}\widetilde{u}_1B_{n-4}\widetilde{u}_1A_{n-2}\widetilde{y'}_1A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}A_{n-2}\widetilde{u}_1.$$

This means we have $B_{n-4} \notin x_1$.

If $B_y = B_{n-4}$ then we get from Lemma 2.7 that $B_{n-4} \notin y_1' A_{n-2} y_4$. If $B_y \ne B_{n-4}$ then, since B_{n-4} is the left special letter of y_3, we also get from Lemma 2.6 and 2.7 that $B_{n-4} \notin y_1' A_{n-2} y_4$. These mean $B_{n-4} \notin x_1 A_n y_1' A_{n-2} y_4$.

From Lemma 2.6 we get that $B_y \notin \widetilde{y_4}$, which means $A_n, A_{n-2}, B_n, B_{n-2}, B_y \notin \widetilde{y_4}$. Since A_{n-2} is the right special letter of $\widetilde{y}$, we have that $A_n, A_{n-2}, B_n \notin y_3 B_{n-2} y_2$. Together we have

$$|w_n| = |x_1A_ny_1'A_{n-2}y_4| + |B_y| + |\widetilde{y_4}| + |A_{n-2}| + |y_3B_{n-2}y_2| + |B_nyA_n\widetilde{y}B_n| + |z_A| + |A_{n-2}| + |\widetilde{z_A}B_nz_2|$$

$$\le r(n-3)+1+r(n-5)+1+r(n-3)+[2r(n-2)+3]+r(n-3)+1+r(n-2)$$

$$= 3r(n-2)+3r(n-3)+r(n-5)+6 < 3r(n-2)+3r(n-3)+r(n-5)+6+3r(n-5) \le 5r(n-2)+1.$$

Case 3.3) $\widetilde{y} = y_1 A_{n-2} y_3 B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} y_2$, where the rightmost A_{n-2} and the leftmost B_{n-2} are the right and left special letters of $\widetilde{y}$, respectively.

If $A = A_{n-2}$, $B = B_{n-2}$ then we would have a square in w_n:

$$B_{n-2}\widetilde{y_3}A_{n-2}\widetilde{y_1}A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}A_{n-2}y_3B_{n-2}\widetilde{y_3}A_{n-2}\widetilde{y_1}A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}A_{n-2}y_3.$$

This means we can divide this case, similar to Case 3.1, into two different cases: $A = A_{n-2}$, $B \ne B_{n-2}$ and $A \ne A_{n-2}$, $B \ne B_{n-2}$.

Case 3.3.1) $A \ne A_{n-2}$ and $B \ne B_{n-2}$.

Similar to Case 2.1, we get $B_n, B_{n-2}, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B \notin x_B$. In the same way, we get $A_n, A_{n-2}, A \notin \widetilde{z_A} B_n z_2$ and $A_n, A_{n-2}, A, B_n \notin z_A$. Together we have

$$|w_n| = |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_A| + |A| + |\widetilde{z_A} B_n z_2|$$

$$\leq r(n-3)+1+r(n-4)+[3r(n-2)+4]+r(n-4)+1+r(n-3) = 3r(n-2)+2r(n-3)+2r(n-4)+6$$

$$< 3r(n-2) + 2r(n-3) + 2r(n-4) + 6 + 2r(n-4) \leq 5r(n-2) + 2.$$

Case 3.3.2) $A = A_{n-2}$, $B \neq B_{n-2}$ (the case $A \neq A_{n-2}$, $B = B_{n-2}$ is symmetric).

Let A_{n-4} be the right special letter of y_3. We will divide this into two cases: $A_{n-4} \notin y_2$ and $A_{n-4} \in y_2$.

Case 3.3.2.1) $A_{n-4} \notin y_2$.

If $A_{n-4} \in z_2$ then we could take the leftmost occurrence of it in z_2, which would create a square in w_n:

$$\widetilde{y_3} A_{n-2} y_3 B_{n-2} y_2 B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} y_2 B_n \widetilde{y_2} B_{n-2}.$$

This means $A_{n-4} \notin z_2$. Let us now mark $y_3 = u_1 A_{n-4} u_2$, where the letter A_{n-4} is the right special letter. We get that $A_{n-4} \notin u_2 B_{n-2} y_2 B_n z_2$. Similar to Case 2.1, we also have $A_n, A_{n-2} \notin u_2 B_{n-2} y_2 B_n z_2$. From Proposition 2.11 we get that $|u_1| \leq r(n-5) + r(n-6) + 1$. Similar to Case 2.1, we get $B_n, B_{n-2}, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B \notin x_B$. Together we have

$$|w_n| = |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |\widetilde{y_2} B_{n-2} \widetilde{u_2}| + |A_{n-4} \widetilde{u_1} A_{n-2} u_1 A_{n-4}| + |u_2 B_{n-2} y_2 B_n z_2|$$

$$\leq r(n-3) + 1 + r(n-4) + [3r(n-2) + 4] + r(n-4) + [2(r(n-5) + r(n-6) + 1) + 3] + r(n-3)$$

$$= 3r(n-2) + 2r(n-3) + 2r(n-4) + 2r(n-5) + 2r(n-6) + 10$$

$$< 3r(n-2) + 2r(n-3) + 2r(n-4) + 2r(n-5) + 2r(n-6) + 10 + 2r(n-6) \leq 5r(n-2) + 2.$$

Case 3.3.2.2) $A_{n-4} \in y_2$.

We will divide this case into three cases depending of which form y_3 is.

Case 3.3.2.2.1) $y_3 = u_1 A_{n-4} u_3 B_{n-4} u_2$, where A_{n-4} and B_{n-4} are the right and left special letters of y_3, respectively.

Since $A_{n-4} \in y_2$, we have that $y_2 = \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} y_2'$, where the A_{n-4} is the leftmost occurrence of A_{n-4} in y_2. If $B_{n-4} \in y_1$ then $y_1 = y_1' B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1}$, where the B_{n-4} is the rightmost occurrence of B_{n-4} in y_1. This would create a square in $\widetilde{y}$:

$$B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} y_3 B_{n-2} \widetilde{u_2}.$$

So $B_{n-4} \notin y_1$. Now, if $B = B_{n-4}$ then $x_B = \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1}$ by Lemma 2.6, since $B_{n-4} \notin y_1$. This would create a square in w_n:

$$B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2}$$

$$B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2}.$$

So $B \neq B_{n-4}$. This means that, in similar way as in Case 2.1, we get $B_n, B_{n-2}, B_{n-4}, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B_{n-4}, B \notin x_B$. Together we have

$$\begin{aligned} |w_n| &= |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_A| + |A_{n-2}| + |\widetilde{z_A} B_n z_2| \\ &\leq r(n-4) + 1 + r(n-5) + [3r(n-2) + 4] + r(n-3) + 1 + r(n-2) \\ &= 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5 \\ &< 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5 + r(n-5) \leq 5r(n-2) + 2. \end{aligned}$$

Case 3.3.2.2.2) $y_3 = u_1 B_{n-4} u_2$, where B_{n-4} is both the right and left special letter.

This case is very similar to the previous, Case 3.3.2.2.1.

Now B_{n-4} is both the right and left special letter, which means $A_{n-4} = B_{n-4}$. Since this case is a subcase of Case 3.3.2.2, we have that $A_{n-4} = B_{n-4} \in y_2$, which means $y_2 = \widetilde{u_2} B_{n-4} y_2'$. If $B_{n-4} \in y_1$ then $y_1 = y_1' B_{n-4} \widetilde{u_1}$ and we would have a square in $\widetilde{y}$:

$$B_{n-4} \widetilde{u_1} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_1} A_{n-2} y_3 B_{n-2} \widetilde{u_2}.$$

So $B_{n-4} \notin y_1$. If $B = B_{n-4}$ then $x_B = \widetilde{u_1} A_{n-2} \widetilde{y_1}$. This would create a square in w_n:

$$B_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2}$$

$$B_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2}.$$

So $B \neq B_{n-4}$. This means that, in similar way as in Case 2.1, we get $B_n, B_{n-2}, B_{n-4}, B \notin x_2 A_n \widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B_{n-4}, B \notin x_B$. Again, we have

$$\begin{aligned} |w_n| &= |x_2 A_n \widetilde{x_B}| + |B| + |x_B| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n| + |z_A| + |A_{n-2}| + |\widetilde{z_A} B_n z_2| \\ &\leq r(n-4) + 1 + r(n-5) + [3r(n-2) + 4] + r(n-3) + 1 + r(n-2) \\ &= 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5 \\ &< 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5 + r(n-5) \leq 5r(n-2) + 2. \end{aligned}$$

Case 3.3.2.2.3) $y_3 = u_1 A_{n-4} u_3 B_{n-4} \widetilde{u_3} A_{n-4} u_3 B_{n-4} u_2$, where the rightmost A_{n-4} and the leftmost B_{n-4} are the right and left special letters of y_3, respectively.

We divide this case into two subcases: $B_{n-4} \notin y_1$ and $B_{n-4} \in y_1$.

Case 3.3.2.2.3.1) $B_{n-4} \notin y_1$.

Now $B \neq B_{n-4}$, since otherwise we would have a square in w_n:

$$A_{n-4}u_3B_{n-4}\widetilde{u_3}A_{n-4}\widetilde{u_1}A_{n-2}\widetilde{y_1}A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}A_{n-2}y_3B_{n-2}\widetilde{u_2}B_{n-4}\widetilde{u_3}$$

$$A_{n-4}u_3B_{n-4}\widetilde{u_3}A_{n-4}\widetilde{u_1}A_{n-2}\widetilde{y_1}A_n\widetilde{y}B_n\widetilde{y_2}B_{n-2}\widetilde{y_3}A_{n-2}y_3B_{n-2}\widetilde{u_2}B_{n-4}\widetilde{u_3}.$$

Similar to Case 2.1, we get $B_n, B_{n-2}, B_{n-4}, B \notin x_2A_n\widetilde{x_B}$ and $A_n, B_n, B_{n-2}, B_{n-4}, B \notin x_B$. Again, we have

$$\begin{aligned}
|w_n| = |x_2A_n\widetilde{x_B}| + |B| + |x_B| + |A_n\widetilde{y}B_nyA_n\widetilde{y}B_n| + |z_A| + |A_{n-2}| + |\widetilde{z_A}B_nz_2|\\
\le r(n-4) + 1 + r(n-5) + [3r(n-2)+4] + r(n-3) + 1 + r(n-2)\\
= 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5\\
< 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5 + r(n-5) \le 5r(n-2) + 2.
\end{aligned}$$

Case 3.3.2.2.3.2) $B_{n-4} \in y_1$.

Now $y_1 = y_1'B_{n-4}\widetilde{u_3}A_{n-4}\widetilde{u_1}$, where the B_{n-4} is the rightmost occurrence of B_{n-4} in y_1, and $y_2 = \widetilde{u_2}B_{n-4}\widetilde{u_3}A_{n-4}y_2'$, where the A_{n-4} is the leftmost occurrence of A_{n-4} in y_2. Remember that we really have $A_{n-4} \in y_2$, since this is a subcase of Case 3.3.2.2.

If $A_{n-2} \in y_1$ then we can take the rightmost occurrence of A_{n-2} in y_1' and get that $y_1 = y_1''A_{n-2}u_1A_{n-4}u_3B_{n-4}\widetilde{u_3}A_{n-4}\widetilde{u_1}$, which creates a square in $\widetilde{y}$:

$$A_{n-4}u_3B_{n-4}\widetilde{u_3}A_{n-4}\widetilde{u_1}A_{n-2}y_3B_{n-2}\widetilde{u_2}B_{n-4}\widetilde{u_3}A_{n-4}u_3B_{n-4}\widetilde{u_3}A_{n-4}\widetilde{u_1}A_{n-2}y_3B_{n-2}\widetilde{u_2}B_{n-4}\widetilde{u_3}.$$

This means $A_{n-2} \notin y_1$.

Now we divide this case into two subcases: $B \neq A_{n-2}$ and $B = A_{n-2}$.

Case 3.3.2.2.3.2.1) $B \neq A_{n-2}$.

Now, in similar way as in Case 2.1, we get that $A_{n-2}, B_n, B_{n-2}, B \notin x_2A_n\widetilde{x_B}$ and $A_n, A_{n-2}, B_n, B_{n-2}, B \notin x_B$. Again, we have

$$\begin{aligned}
|w_n| = |x_2A_n\widetilde{x_B}| + |B| + |x_B| + |A_n\widetilde{y}B_nyA_n\widetilde{y}B_n| + |z_A| + |A_{n-2}| + |\widetilde{z_A}B_nz_2|\\
\le r(n-4) + 1 + r(n-5) + [3r(n-2)+4] + r(n-3) + 1 + r(n-2)\\
= 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5\\
< 4r(n-2) + r(n-3) + r(n-4) + r(n-5) + 5 + r(n-5) \le 5r(n-2) + 2.
\end{aligned}$$

Case 3.3.2.2.3.2.2) $B = A_{n-2}$.

Now we have that $x_B = \widetilde{y_1} = u_1 A_{n-4} u_3 B_{n-4} \widetilde{y_1'}$. We will first show that $A_{n-4} \notin y_1', x_2$ and $B_{n-4} \notin y_2', z_2$.

If $A_{n-4} \in y_1'$ then we have $y_1 = y_1'' A_{n-4} u_3 B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1}$. This creates a square in w_n:

$$u_3 B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4}$$

$$u_3 B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4}.$$

So $A_{n-4} \notin y_1'$. If $B_{n-4} \in y_2'$ then we have that $y_2 = \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} u_3 B_{n-4} y_2''$. Also this creates a square in w_n:

$$B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} u_3$$

$$B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} \widetilde{y_1} A_n \widetilde{y} B_n \widetilde{y_2} B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} u_3.$$

So $B_{n-4} \notin y_2'$. If $A_{n-4} \in x_2$ then we could take the rightmost occurrence of A_{n-4} in x_2 and get a square in w_n:

$$A_{n-4} u_3 B_{n-4} \widetilde{y_1'} A_n y_1' B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} u_1 A_{n-4} u_3 B_{n-4} \widetilde{y_1'} A_n y_1' B_{n-4} \widetilde{u_3} A_{n-4} \widetilde{u_1} A_{n-2} u_1.$$

So $A_{n-4} \notin x_2$. If $B_{n-4} \in z_2$ then we could take the leftmost occurrence of B_{n-4} in z_2 and get a square in w_n:

$$u_2 B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} y_2' B_n \widetilde{y_2'} A_{n-4} u_3 B_{n-4}$$

$$u_2 B_{n-2} \widetilde{y_3} A_{n-2} y_3 B_{n-2} \widetilde{u_2} B_{n-4} \widetilde{u_3} A_{n-4} y_2' B_n \widetilde{y_2'} A_{n-4} u_3 B_{n-4}.$$

So $B_{n-4} \notin z_2$. Now we know that $A_{n-4} \notin x_2 A_n y_1' B_{n-4} \widetilde{u_3}$ and $B_{n-4} \notin \widetilde{u_3} A_{n-4} y_2' B_n z_2$.

Similar to Case 2.1, we get $A_{n-2}, B_n, B_{n-2} \notin x_2 A_n y_1' B_{n-4} \widetilde{u_3}$ and $A_n, A_{n-2}, B_n, B_{n-2} \notin u_3 B_{n-4} \widetilde{y_1'}$ and $A_n, A_{n-2} \notin \widetilde{u_3} A_{n-4} y_2' B_n z_2$. From Proposition 2.11 we get that $|u_1|, |u_2| \leq r(n-6) + r(n-7) + 1$, where $r(n-7) = 0$ if $n = 7$. Since $A_n, B_n \notin y$ and A_{n-2} is the right special letter of $\widetilde{y}$, we trivially have $A_n, A_{n-2}, B_n \notin \widetilde{y_2} B_{n-2} \widetilde{y_3}$. From Lemma 2.10 we also get easily that $A_n, A_{n-2}, B_n, B_{n-2} \notin y_3$. Together we finally have

$$|w_n| = |x_2 A_n y_1' B_{n-4} \widetilde{u_3}| + |A_{n-4} \widetilde{u_1} A_{n-2} u_1 A_{n-4}| + |u_3 B_{n-4} \widetilde{y_1'}| + |A_n \widetilde{y} B_n y A_n \widetilde{y} B_n|$$

$$+ |\widetilde{y_2} B_{n-2} \widetilde{y_3}| + |A_{n-2}| + |y_3| + |B_{n-2}| + |\widetilde{u_2}| + |B_{n-4}| + |\widetilde{u_3} A_{n-4} y_2' B_n z_2|$$

$$\leq r(n-4) + [2r(n-6) + 2r(n-7) + 5] + r(n-5) + [3r(n-2) + 4] + r(n-3) + 1 + r(n-4) + 1 +$$

$$[r(n-6) + r(n-7) + 1] + 1 + r(n-3) = 3r(n-2) + 2r(n-3) + 2r(n-4) + r(n-5) + 3r(n-6) + 3r(n-7) + 13$$

$$< 3r(n-2) + 2r(n-3) + 2r(n-4) + r(n-5) + 3r(n-6) + 3r(n-7) + 13 + r(n-6) + r(n-7) \leq 5r(n-2) + 2.$$

□

As we can see, improving our upper bound was very exhausting. If we would like to achieve Conjecture 2.3, we would need to use a slightly different approach.

Let us still estimate our upper bound in a closed form. Suppose first $n \geq 7$ is even:

$$r(n) \leq 5r(n-2)+4 \leq 5(5r(n-4)+4)+4 \leq \ldots \leq 5^{(n-6)/2} r(6) + 4(5^{(n-8)/2} + \ldots + 5 + 1)$$

$$< 5^{(n-6)/2} \cdot (5^3 - 58) + (5^{(n-8)/2+1} + \ldots + 5) = 5^{n/2} - 58 \cdot 5^{(n-6)/2} + (5^{(n-8)/2+1} + \ldots + 5) < 5^{n/2} < 2,237^n.$$

Suppose then that $n \geq 7$ is odd:

$$r(n) \leq 5r(n-2)+4 \leq 5(5r(n-4)+4)+4 \leq \ldots \leq 5^{(n-5)/2} r(5) + 4(5^{(n-7)/2} + \ldots + 5 + 1)$$

$$< 5^{(n-5)/2} \cdot (5^{2,5} - 22) + (5^{(n-7)/2+1} + \ldots + 5) = 5^{n/2} - 22 \cdot 5^{(n-5)/2} + (5^{(n-7)/2+1} + \ldots + 5) < 5^{n/2} < 2,237^n.$$

Together with the lower bound, we finally get that $2,008^n < r(n) < 2,237^n$, for $n \geq 5$.

Acknowledgements

I want to thank Š. Starosta for making me aware of this problem, L. Zamboni and T. Harju for insightful comments, and J. Peltomäki for calculating the exact value of $r(7)$ and adding all the values of $r(n)$ up to $n = 7$ to the OEIS database https://oeis.org/A269560.

References

References

[AFMP] P. Ambrož, C. Frougny, Z. Masáková, E. Pelantová: *Palindromic complexity of infinite words associated with simple Parry numbers*, Ann. Inst. Fourier (Grenoble) 56 (2006), pp. 2131-2160.

[BDGZ1] M. Bucci, A. De Luca, A. Glen, L. Q. Zamboni: *A connection between palindromic and factor complexity using return words*, Advances in Applied Mathematics 42 (2009), pp. 60-74.

[BDGZ2] M. Bucci, A. De Luca, A. Glen, L. Q. Zamboni: *A new characteristic property of rich words*, Theoret. Comput. Sci. 410 (2009), pp. 2860-2863.

[BHNR] S. Brlek, S. Hamel, M. Nivat, C. Reutenauer: *On the palindromic complexity of infinite words*, Internat. J. Found. Comput. 15 (2004), pp. 293-306.

[CD] A. Carpi, A. de Luca: *Special factors, periodicity, and an application to Sturmian words*, Acta Informatica 36 (2000), pp. 983-1006.

[CR] J. D. Currie, N. Rampersad: *A proof of Dejeans conjecture*, Math. Comput. 80(274) (2011), pp. 1063-1070.

[D] F. Dejean: *Sur un théorème de Thue*, J. Combin. Theory Ser. A 13 (1972), pp. 90-99.

[DGZ] A. de Luca, A. Glen, L. Q. Zamboni: *Rich, Sturmian, and trapezoidal words*, Theoret. Comput. Sci. 407 (2008), pp. 569-573.

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

[GJ] A. Glen, J. Justin: *Episturmian words: A survey*, Theor. Inform. Appl. 43(3) (2009), pp. 403-442.

[GJWZ] A. Glen, J. Justin, S. Widmer, L. Q. Zamboni: *Palindromic richness*, European Journal of Combinatorics 30 (2009), pp. 510-531.

[HPS] K. Hare, H. Prodinger, J. Shallit: *Three Series for the Generalized Golden Mean*, Fibonacci Quart. 52 (2014), pp. 307-313.

[Lot1] M. Lothaire: *Combinatorics on Words,* in: Encyclopedia of Mathematics and its Applications, vol 17, Addison-Wesley, Reading, MA, 1983.

[Lot2] M. Lothaire: *Algebraic Combinatorics on Words,* in: Encyclopedia of Mathematics and its Applications, vol 90, Cambridge University Press, UK, 2002.

[MP] F. Mignosi, G. Pirillo: *Repetitions in the Fibonacci infinite word*, Informatique théorique et applications, tome 26, n. 3 (1992), pp. 199-204.

[PS] E. Pelantová, Š. Starosta: *Languages invariant under more symmetries: overlapping factors versus palindromic richness*, Discrete Math. 313 (2013), pp. 2432-2445.

[Q] M. Queffélec: *Substitution Dynamical Systems Spectral Analysis,* Lecture Notes in Mathematics. 2nd ed. 2010. Springer-Verlag, Berlin.

[R] M. Rao: *Last cases of Dejeans conjecture*, Theoret. Comput. Sci. 412(27) (2011), pp. 3010-3018.

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

[T1] A. Thue: *Uber unendliche Zeichenvihen*, Kra. Vidensk. Selsk. Skrifter. I Mat.-Nat.Kl., Christiana, Nr. 7, 1906.

[T2] A. Thue: *Uber die gegenseitige Lage gleicher Teile gewisser Zeichenreihen*, Kra. Vidensk. Selsk. Skrifter. I Mat.-Nat.Kl., Christiana, Nr. 12, 1912.

[V] J. Vesti: *Extensions of rich words*, Theoret. Comput. Sci. 548 (2014), pp. 14-24.