The Repetition Threshold for Ternary Rich Words

James D. Currie[a]   Lucas Mol[b]   Jarkko Peltomäki[c]

Abstract

In 2017, Vesti proposed the problem of determining the repetition threshold for infinite rich words, i.e., for infinite words in which all factors of length n contain n distinct nonempty palindromic factors. In 2020, Currie, Mol, and Rampersad proved a conjecture of Baranwal and Shallit that the repetition threshold for binary rich words is $2 + \sqrt{2}/2$. In this paper, we prove a structure theorem for $16/7$-power-free ternary rich words. Using the structure theorem, we deduce that the repetition threshold for ternary rich words is $1 + 1/(3 - \mu) \approx 2.25876324$, where μ is the unique real root of the polynomial $x^3 - 2x^2 - 1$.

Mathematics Subject Classifications: 68R15

1 Introduction

The study of repetitions in words goes back to the works of Thue at the beginning of the twentieth century [39, 40], which have been translated to English by Berstel [9]. Many extensions and variations of Thue's results have been proven since then; see, for example, the surveys of Ochem, Rao, and Rosenfeld [33] and Rampersad and Shallit [35]. We use standard notation and terminology related to repetitions in words in the remainder of this section; the unfamiliar reader should refer to Section 2.

In this paper, we study repetitions in so-called rich words. A *palindrome* is a word that reads the same forwards and backwards. Droubay, Justin, and Pirillo [23] were first to observe that every word of length n contains at most $n + 1$ distinct palindromes as factors, including the empty word. A word of length n is called *rich* if it has $n + 1$ distinct palindromes as factors; an infinite word is rich if all of its finite factors are rich. Since their (implicit) introduction by Droubay, Justin, and Pirillo, rich words have been well-studied. It is known that the language of infinite rich words contains several highly

[a]Department of Mathematics and Statistics, The University of Winnipeg, Winnipeg, Canada (`j.currie@uwinnipeg.ca`).

[b]Department of Mathematics and Statistics, Thompson Rivers University, Kamloops, Canada (`lmol@tru.ca`).

[c]Information Technology, Åbo Akademi University, Turku, Finland (`r@turambar.org`).

structured classes of words, including Sturmian words, episturmian words, and complementary symmetric Rote words (see [10, 23]). Infinite rich words have been characterized in terms of a condition on complete returns to palindromes [28], in terms of a relation between factor and palindromic complexity [11], and in terms of a condition on bispecial factors [4]. For any fixed integer $k \geqslant 2$, the number of rich words of length n over k letters is known to grow superpolynomially [29] and subexponentially [38] in n.

Vesti [41] proposed the problem of determining the repetition threshold for the language of infinite rich words over k letters. This problem has been resolved in the binary case through the combined effort of several authors. Baranwal and Shallit [8] constructed an infinite binary rich word with critical exponent $2 + \sqrt{2}/2 \approx 2.707$ and conjectured that $2 + \sqrt{2}/2$ is in fact the repetition threshold for infinite binary rich words. This conjecture was confirmed by Currie, Mol, and Rampersad [16], who proved a structure theorem for 14/5-power-free infinite binary rich words. Roughly speaking, the structure theorem says that every 14/5-power-free infinite binary rich word contains all of the factors of one of two specific infinite binary words (one being the word of Baranwal and Shallit). Since both of these words turn out to be rich and have critical exponent $2 + \sqrt{2}/2$, Baranwal and Shallit's conjecture follows from the structure theorem.

A general lower bound on the repetition threshold for the language of infinite rich words over k letters follows from a result of Pelantová and Starosta [34], which says that every infinite rich word, over *any* finite alphabet, contains infinitely many factors of exponent 2. It follows that the repetition threshold (and also the *asymptotic* repetition threshold) for the language of infinite rich words over k letters is at least 2 for every $k \geqslant 2$. In fact, Dvořáková, Klouda, and Pelantová [24] have recently shown that the asymptotic repetition threshold for the language of infinite rich words over k letters is equal to 2 for all $k \geqslant 2$. But to date, Vesti's problem of determining the (ordinary) repetition threshold for the language of infinite rich words over k letters has only been resolved in the binary case.

In this paper, we resolve Vesti's problem in the ternary case by proving the following theorem.

Theorem 1.1. *The repetition threshold for the language of infinite ternary rich words is*

$$1 + \frac{1}{3 - \mu_1} \approx 2.25876324,$$

where μ_1 is the unique real root of the polynomial $x^3 - 2x^2 - 1$.

In order to prove Theorem 1.1, we first prove a structure theorem for infinite 16/7-power-free ternary rich words, similar to the one established by Currie, Mol, and Rampersad [16] in the binary case. Roughly speaking, we show that every infinite 16/7-power-free ternary rich word "looks like" one specific word. Structure theorems like this have existed in the combinatorics on words literature from the very beginning. Thue [40] established structure theorems for infinite square-free ternary words avoiding various sets of factors. Another well-known (and incredibly useful) structure theorem is the one for 7/3-power-free
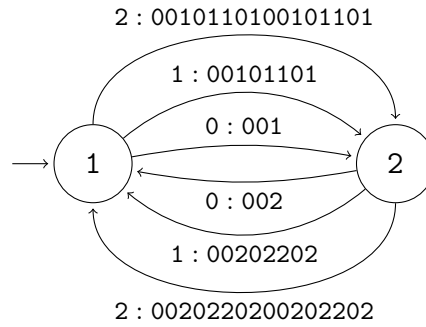
Figure 1: The transducer τ. We will also refer to the related transducer $\overline{\tau}$, which has the same states and transitions, but starts in state 2.

binary words due to Karhumäki and Shallit [30], which says that every infinite 7/3-power-free binary word contains all factors of the Thue-Morse word. In recent years, a surprising number of new structure theorems have been established; see [2, 7, 15, 17, 19, 25], for example.

For our structure theorem, we let $f, g \colon \Sigma_3^* \to \Sigma_3^*$ be the morphisms defined by

$$
\begin{aligned}
f(0) &= 01 & g(0) &= 20 \\
f(1) &= 022 & g(1) &= 21 \\
f(2) &= 02 & g(2) &= 2,
\end{aligned}
$$

and we let τ be the transducer drawn in Figure 1.

Theorem 1.2 (Structure Theorem)**.** *Suppose that $\mathbf{z} \in \Sigma_3^\omega$ is a 16/7-power-free rich word. Then for all $n \geqslant 0$, a suffix of $\mathbf{z}$ can be obtained from $\tau(g(f^n(\mathbf{x}_n)))$ by permuting the letters, where $\mathbf{x}_n \in \Sigma_3^\omega$. In particular, $\mathbf{z}$ contains all of the factors obtained from $\tau(g(f^\omega(0)))$ by permuting the letters.*

We then show that the word $\tau(g(f^\omega(0)))$ is rich and has critical exponent $1+1/(3-\mu_1)$. Theorem 1.1 then follows easily from these two facts and Theorem 1.2.

The remainder of the paper is laid out as follows. In Section 2, we describe the notation and terminology used in the paper, and provide a more comprehensive summary of results related to Theorem 1.1. In Section 3, we prove Theorem 1.2. In Section 4, we prove that $\tau(g(f^\omega(0)))$ is rich, and in Section 5, we prove that $\tau(g(f^\omega(0)))$ has critical exponent $1 + 1/(3 - \mu_1)$. Finally, in Section 6, we ask some questions related to Vesti's problem over larger alphabets.

2 Notation, Terminology, and Background

2.1 Words

An *alphabet* is a nonempty finite set of symbols, which we refer to as *letters*. Throughout, we let Σ_k denote the alphabet $\{0, 1, \ldots, k-1\}$. A *word* over an alphabet A is a finite or

infinite sequence of letters from A. The *length* of a finite word w, denoted by $|w|$, is the number of letters that make up w. We let ε denote the *empty word*, which is the unique word of length 0. For a letter a, we let $|w|_a$ denote the number of occurrences of a in w.

For a finite word x and a finite or infinite word y, the *concatenation* of x and y, denoted by xy, is the word consisting of all of the letters of x followed by all of the letters of y. Suppose that a finite or infinite word w can be written in the form $w = xyz$, where x, y, and z are possibly empty words. Then the word y is called a *factor* of w, the word x is called a *prefix* of w, and the word z is called a *suffix* of w. If x and z are nonempty, then y is called an *internal factor* of w, and we say that y *appears internally* in w. If yz is nonempty, then x is called a *proper prefix* of w, and if xy is nonempty, then z is called a *proper suffix* of w. For an infinite word $\mathbf{u}$, the *language* of $\mathbf{u}$, denoted by $\mathrm{Fact}(\mathbf{u})$, is the set of all finite factors of $\mathbf{u}$.

For a set of finite words A, we let A^* denote the set of finite words obtained by concatenating elements of A, and we let A^ω denote the set of infinite words obtained by concatenating elements of A. So in particular, Σ_k^* is the set of finite words over Σ_k, and Σ_k^ω is the set of infinite words over Σ_k.

Let $\mathbf{w}$ be an infinite word, and let w be a finite factor of $\mathbf{w}$. We say that w is *recurrent* in $\mathbf{w}$ if it occurs infinitely many times in $\mathbf{w}$ and that w is *uniformly recurrent* in $\mathbf{w}$ if there is an integer k such that every factor of $\mathbf{w}$ of length k contains w. We say that the infinite word $\mathbf{w}$ is *(uniformly) recurrent* if every finite factor of $\mathbf{w}$ is (uniformly) recurrent. A *complete return word* to w in $\mathbf{w}$ is a factor of $\mathbf{w}$ that contains w as a proper prefix and a proper suffix but not as an internal factor. If w is recurrent in $\mathbf{w}$, then every occurrence of w in $\mathbf{w}$ is followed by a successive occurrence, which gives rise to a complete return word to w. Evidently, if w is uniformly recurrent in $\mathbf{w}$, then there are only finitely many complete return words to w in $\mathbf{w}$. We say that a factor of w is *unioccurrent* if it occurs exactly once in w as a factor.

2.2 Morphisms and Transducers

For alphabets A and B, a *morphism* from A^* to B^* is a function $h\colon A^* \to B^*$ that satisfies $h(uv) = h(u)h(v)$ for all words $u, v \in A^*$. Let $h\colon A^* \to A^*$ be a morphism. For all words $x \in A^*$, we define $h^0(x) = x$, and $h^n(x) = h(h^{n-1}(x))$ for all integers $n \geqslant 1$. For a letter $a \in A$, we say that h is *prolongable* on a if $h(a) = ax$ for some non-empty word $x \in X^*$ and $h^n(x) \neq \varepsilon$ for all $n \geqslant 0$. If h is prolongable on a with $h(a) = ax$, then it is easy to show that for every integer $n \geqslant 1$, we have

$$h^n(a) = axh(x)h^2(x) \cdots h^{n-1}(x).$$

Thus, each $h^n(a)$ is a prefix of the infinite word

$$h^\omega(a) = axh(x)h^2(x)h^3(x) \cdots .$$

Note that $h^\omega(a)$ is a fixed point of h, i.e., $h(h^\omega(a)) = h^\omega(a)$, where the morphism h is extended to infinite words in the natural way.

For the formal definition of finite-state transducer, see [1, Section 4.3]. We only make use of the transducer τ drawn in Figure 1 (and the related transducer $\bar{\tau}$), whose behavior is very simple. These transducers simply alternate between the two states on every input letter. In fact, one could define τ and $\bar{\tau}$ in terms of the morphisms $t_1, t_2 \colon \Sigma_3^* \to \Sigma_3^*$ defined as follows:

$$
\begin{array}{ll}
t_1(0) = \texttt{001} & t_2(0) = \texttt{002} \\
t_1(1) = \texttt{00101101} & t_2(1) = \texttt{00202202} \\
t_1(2) = \texttt{0010110100101101} & t_2(2) = \texttt{0020220200202202}.
\end{array}
$$

For a finite or infinite word $w = w_0 w_1 w_2 w_3 \cdots$ over Σ_3, where the w_i are letters, we have

$$\tau(w) = t_1(w_0) t_2(w_1) t_1(w_2) t_2(w_3) \cdots$$

and

$$\bar{\tau}(w) = t_2(w_0) t_1(w_1) t_2(w_2) t_1(w_3) \cdots.$$

Essentially, the output of τ is obtained from the input word by applying t_1 to all letters of even index and t_2 to all letters of odd index, and the output of $\bar{\tau}$ is obtained by applying t_2 to all letters of even index and t_1 to all letters of odd index. We note that the well-known Arshon words can be defined in terms of similar maps (c.f. [32]).

2.3   Repetitions in Words

Let $w = w_0 w_1 \cdots w_{n-1}$ be a finite word, where the w_i are letters. For an integer p with $1 \leqslant p \leqslant n$, we say that p is a *period* of w if $w_{i+p} = w_i$ for all $0 \leqslant i < n - p$. In this case, we say that n/p is an *exponent* of w, and that w is an n/p-*power*. The smallest period of w is called *the* period of w, and it corresponds to the largest exponent of w, which is called *the* exponent of w.

Now let $\mathbf{w}$ be a finite or infinite word, and let $\alpha > 1$ be a real number. We say that $\mathbf{w}$ is α-*power-free* if it contains no factors of exponent greater than or equal to α, and that $\mathbf{w}$ is α^+-*power-free* if it contains no factors of exponent strictly greater than α. The *critical exponent* of $\mathbf{w}$, denoted by $\mathrm{ce}(\mathbf{w})$, is defined by

$$\mathrm{ce}(\mathbf{w}) = \sup\{r \in \mathbb{Q} : \mathbf{w} \text{ has a factor of exponent } r\},$$

or equivalently by

$$\mathrm{ce}(\mathbf{w}) = \inf\{\alpha \in \mathbb{R} : \mathbf{w} \text{ is } \alpha\text{-power-free}\}.$$

Roughly speaking, the critical exponent of $\mathbf{w}$ describes the largest exponents among all factors of $\mathbf{w}$.

Given a language L of infinite words, it is natural to try to find the infimum of the critical exponents among all words in L. The words in L with the smallest critical exponents are in some sense the "least repetitive" words in L. The *repetition threshold* of L, denoted by $\mathrm{RT}(L)$, is defined by

$$\mathrm{RT}(L) = \inf\{\mathrm{ce}(\mathbf{w}) : \mathbf{w} \in L\}.$$

One of the most celebrated results in the area of repetitions in words is Dejean's theorem, which describes the repetition threshold of the language of all infinite words over k letters:

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

The case $k = 2$ was proven by Thue [40], and Dejean [20] proved the case $k = 3$ and correctly conjectured the remaining cases, which have been proven through the work of many different authors. In particular, Carpi [12] proved all but finitely many cases, and the last cases were proven independently by Currie and Rampersad [18] and Rao [37].

Especially since the completion of the proof of Dejean's theorem, much work has been done on determining the repetition thresholds of various narrower languages. The repetition threshold for the language of all Sturmian words was determined by Carpi and de Luca [13], who showed that the Fibonacci word has the least critical exponent among all Sturmian words, namely $(5 + \sqrt{5})/2$. Sturmian words are episturmian, balanced, and rich, and research has been done on determining the repetition thresholds of k-ary words of these three types in the last decade. For all $k \geqslant 2$, we let E_k, B_k, and R_k denote the languages of k-ary episturmian, balanced, and rich words, respectively.

For episturmian words, Dvořáková and Pelantová [26] recently established that

$$\mathrm{RT}(E_k) = 2 + \frac{1}{t_k - 1}$$

for all $k \geqslant 2$, where t_k is the unique positive root of the polynomial $x^k - x^{k-1} - \cdots - x - 1$.

For balanced words, the following is known:

$$\mathrm{RT}(B_k) = \begin{cases} 2 + \frac{1+\sqrt{5}}{2}, & \text{if } k = 2 \text{ [13]}; \\ 2 + \frac{\sqrt{2}}{2}, & \text{if } k = 3 \text{ [36]}; \\ 1 + \frac{1+\sqrt{5}}{4}, & \text{if } k = 4 \text{ [36]}; \\ 1 + \frac{1}{k-3}, & \text{if } 5 \leqslant k \leqslant 10 \text{ [6, 21]}; \\ 1 + \frac{1}{k-2}, & \text{if } k = 11 \text{ or both } k \geqslant 12 \text{ and } k \text{ is even [27]}. \end{cases}$$

Dvořáková, Opočenská, Pelantová, and Shur [27] conjecture that $\mathrm{RT}(B_k) = 1 + 1/(k-2)$ for all $k \geqslant 13$ with k odd, but this conjecture remains open at the time of writing.

As described in the introduction, for rich words, it is known that $\mathrm{RT}(R_k) \geqslant 2$ for all $k \geqslant 2$, and that $\mathrm{RT}(R_2) = 2 + \sqrt{2}/2$.

It turns out that there are many infinite words $\mathbf{w}$ in which only "short" factors of $\mathbf{w}$ have exponent equal to (or even close to) the critical exponent of $\mathbf{w}$. If we ignore such short factors, and consider only arbitrarily long factors of $\mathbf{w}$, then we obtain what is called the *asymptotic critical exponent* of $\mathbf{w}$, first defined by Cassaigne [14]. It is denoted by $\mathrm{ce}^*(\mathbf{w})$ and defined by

$$\mathrm{ce}^*(\mathbf{w}) = \limsup_{n \to \infty}\{r \in \mathbb{Q} : \mathbf{w} \text{ has a factor of exponent } r \text{ and period } n\}.$$

The *asymptotic repetition threshold* of L, denoted by $\mathrm{RT}^*(L)$, is defined by

$$\mathrm{RT}^*(L) = \inf\{\mathrm{ce}^*(\mathbf{w}) \colon \ \mathbf{w} \in L\}.$$

For every language of infinite words L, we clearly have $\mathrm{RT}^*(L) \leqslant \mathrm{RT}(L)$. We have equality for some languages L, and strict inequality for others. For example, Cassaigne [14] observed that $\mathrm{RT}^*(S) = \mathrm{RT}(S)$, where S is the language of Sturmian words, but demonstrated that $\mathrm{RT}^*(\Sigma_k^\omega) = 1 < \mathrm{RT}(\Sigma_k^\omega)$ for all $k \geqslant 2$. We refer the reader to the recent work of Dvořáková, Klouda, and Pelantová [24] for a summary of results on the asymptotic repetition thresholds of episturmian, balanced, and rich words. In particular, as mentioned in the introduction, they show that for rich words, we have $\mathrm{RT}^*(R_k) = 2$ for all $k \geqslant 2$.

3   The Structure Theorem

In this section, we prove Theorem 1.2. Suppose that $\mathbf{z}$ is an infinite 16/7-power-free ternary rich word. In Section 3.1, we show that up to permutation of the letters, some suffix of $\mathbf{z}$ has the form $\tau(\mathbf{y})$ for some word $\mathbf{y} \in \Sigma_3^\omega$. (In fact, we show this under the slightly weaker assumption that $\mathbf{z}$ is an infinite 7/3-power-free ternary rich word.) In Section 3.2, we show that a suffix of $\mathbf{y}$ has the form $g(\mathbf{x})$ for some word $\mathbf{x} \in \Sigma_3^\omega$. In Section 3.3, we describe several families of factors that cannot appear in $\mathbf{x}$ if the word $\tau(g(f^n(\mathbf{x})))$ is 16/7-power-free and rich, where n is any nonnegative integer. Finally, in Section 3.4, we show that if $\tau(g(f^n(\mathbf{x})))$ is 16/7-power-free and rich for some $n \geqslant 0$ and $\mathbf{x} \in \Sigma_3^\omega$, then a suffix of $\mathbf{x}$ has the form $f(\mathbf{x}')$ for some word $\mathbf{x}' \in \Sigma_3^\omega$, which allows us to complete the proof of Theorem 1.2 by mathematical induction.

3.1   The First Layer

Throughout this subsection, let $\mathbf{z}$ be an infinite 7/3-power-free ternary rich word. We first observe that $\mathbf{z}$ must contain at least one of the factors 001002, 112110, and 220221, as a computer backtracking search shows that a longest 7/3-power-free ternary rich word containing none of these three factors has length 388.[1] Note that 001002, 112110, and 220221 can be obtained from one another by permuting the letters. Thus, by permuting the letters and removing a prefix of $\mathbf{z}$ if necessary, we assume henceforth that $\mathbf{z}$ has prefix 001002.

By performing further backtracking searches, we confirm that several short words, and all words obtained from them by permuting the letters, do not appear in $\mathbf{z}$. The results are summarized in Table 1. Our next lemma describes two more short words that do not appear in $\mathbf{z}$. To prove it, we make use of the following well-known characterization of finite rich words.

Lemma 3.1 ([23], c.f. [28])**.** *A finite word w is rich if and only if every prefix (resp. suffix) of w has a unioccurrent palindromic suffix (resp. prefix).*

---

[1]Code to verify all backtracking searches in this paper can be found at https://github.com/japeltom/ternary-rich-words-verification.

p	Length of a longest 7/3-power-free ternary rich word with prefix p
102	152
0011	498
00100200	502

Table 1: Some words that do not appear in $\mathbf{z}$.

Lemma 3.2. *The word* $\mathbf{z}$ *contains neither* 12 *nor* 21.

Proof. Suppose that $\mathbf{z}$ contains one of the factors 12 or 21. Let p be the shortest prefix of $\mathbf{z}$ that ends in 12 or 21, and let s be the suffix of p of length 2. By the minimality of p, the reversal of s does not appear in p, and it follows that the only nonempty palindromic suffix of p has length 1. Since p has prefix 001002, every word of length 1 occurs more than once in p. Hence p does not have a unioccurrent palindromic suffix, and we conclude by Lemma 3.1 that p is not rich, which is a contradiction. $\qquad\square$

Let

$$A_1 = 001,$$
$$B_1 = 00101101,$$
$$A_2 = 002, \text{ and}$$
$$B_2 = 00202202.$$

Note that A_1 can be obtained from A_2 (and vice versa) by swapping 1 and 2. The same can be said for B_1 and B_2. We will frequently need to deal with such pairs of words. For a word $x \in \Sigma_3^*$, the *sister* of x is the word obtained from x by swapping the letters 1 and 2. With this terminology, the words A_1 and A_2 are sisters, as are the words B_1 and B_2. Note also that for every word $w \in \Sigma_3^*$, the words $\tau(w)$ and $\overline{\tau}(w)$ are sisters.

Note that A_1, B_1, A_2, and B_2 have common prefix 00, and that 00 occurs only as a prefix of A_1, B_1, A_2, or B_2 in a word belonging to $\{A_1, B_1, A_2, B_2\}^*$.

Lemma 3.3. $\mathbf{z} \in \{A_1, B_1, A_2, B_2\}^\omega$.

Proof. By assumption, $\mathbf{z}$ begins with 00. In fact, a computer backtracking search shows that the longest 7/3-power-free ternary rich word with no 00 has length 57, so that the factor 00 is recurrent in $\mathbf{z}$. Thus, it suffices to show that the only complete returns to 00 in $\mathbf{z}$ are $A_1$00, $B_1$00, $A_2$00, or $B_2$00. Since the cube 000 does not appear in $\mathbf{z}$, every occurrence of 00 in $\mathbf{z}$ must be followed by 1 or 2. We show that the only complete returns to 00 starting with 001 in $\mathbf{z}$ are $A_1$00 and $B_1$00. The argument that the only complete returns to 00 starting with 002 in $\mathbf{z}$ are $A_2$00 and $B_2$00 is symmetric.

Consider the tree drawn in Figure 2, which shows all possible complete returns to 00 in $\mathbf{z}$ starting with 001. From Lemma 3.2, we know that 1 is never followed by 2 in $\mathbf{z}$, so these branches are suppressed in the tree. We explain below why the words corresponding to the red leaves in the tree cannot appear in $\mathbf{z}$.

Figure 2: The tree showing all possible complete returns to 00 in **z** starting with 001.

- 001010 has the $5/2$-power 01010 as a suffix.

- 00101100 has a permutation of the word 0011 from Table 1 as a suffix.

- 0010110101 has the $5/2$-power 10101 as a suffix.

- 0010110102 has the word 102 from Table 1 as a suffix.

- 001011011 has the $7/3$-power 1011011 as a suffix.

- 00101102 has the word 102 from Table 1 as a suffix.

- 0010111 has the cube 111 as a suffix.

- 00102 has the word 102 from Table 1 as a suffix.

- 0011 is in Table 1.

This means that the words corresponding to the green leaves in the tree are the only possible complete returns to 00 starting with 001 in **z**. Therefore, we conclude that $\mathbf{z} \in \{A_1, B_1, A_2, B_2\}^\omega$, as desired. $\square$

Let $q \colon \{\mathtt{a_1}, \mathtt{b_1}, \mathtt{a_2}, \mathtt{b_2}\}^* \to \{\mathtt{0}, \mathtt{1}, \mathtt{2}\}^*$ be the morphism that sends each lowercase letter to the word denoted by the corresponding uppercase letter. Then Lemma 3.3 says that $\mathbf{z} = q(\mathbf{v})$ for some word $\mathbf{v} \in \{\mathtt{a_1}, \mathtt{b_1}, \mathtt{a_2}, \mathtt{b_2}\}^\omega$.

Lemma 3.4. *The word* **v** *corresponds to an infinite walk on the directed graph G drawn in Figure 3, and contains neither* $\mathtt{b_1 b_1 b_1}$ *nor* $\mathtt{b_2 b_2 b_2}$.

Proof. First of all, if **v** contains $\mathtt{b_1 b_1 b_1}$ or $\mathtt{b_2 b_2 b_2}$, then **z** contains the cube $B_1 B_1 B_1$ or its sister, contradicting the assumption that **z** is $7/3$-power-free.

We know from Lemma 3.3 that $\mathbf{v} \in \{\mathtt{a_1}, \mathtt{b_1}, \mathtt{a_2}, \mathtt{b_2}\}^\omega$. So to show that **v** corresponds to an infinite walk on G, it suffices to show that the following factors do not appear in **v**:

- $\mathtt{a_1 a_1}$, $\mathtt{a_1 b_1}$, $\mathtt{a_1 a_2}$, and $\mathtt{b_1 a_1}$; and

- $\mathtt{a_2 a_2}$, $\mathtt{a_2 b_2}$, $\mathtt{a_2 a_1}$, and $\mathtt{b_2 a_2}$.

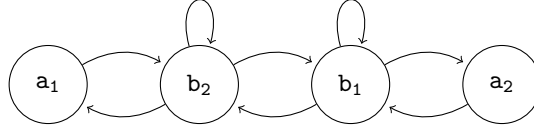By symmetry, we need only show that the first four factors do not appear in **v**.

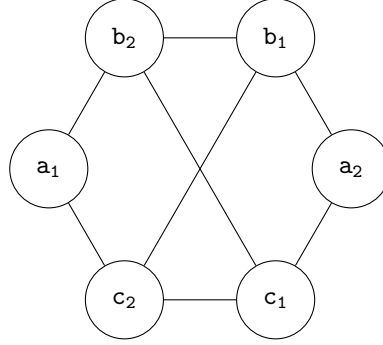Figure 3: The directed graph G showing possible transitions between letters in $\mathbf{v}$.

Figure 4: The graph G' showing possible transitions between letters in $\mathbf{v}'$.

- If $\mathbf{v}$ contains $\mathsf{a_1 a_1}$, then $\mathbf{z}$ contains the $8/3$-power $A_1 A_1 00 = 00100100$.

- If $\mathbf{v}$ contains $\mathsf{a_1 b_1}$, then $\mathbf{z}$ contains $A_1 B_1$, which has the $7/3$-power 0010010 as a prefix.

- If $\mathbf{v}$ contains $\mathsf{b_1 a_1}$, then $\mathbf{z}$ contains $B_1 A_1 00$, which has the $7/3$-power 0100100 as a suffix.

- If $\mathbf{v}$ contains $\mathsf{a_1 a_2}$, then $\mathbf{z}$ contains $A_1 A_2 00 = 00100200$, which is in Table 1.

This completes the proof. $\qquad\square$

Let $\mathbf{v}'$ be the word obtained from $\mathbf{v}$ by replacing each occurrence of the factor $\mathsf{b_1 b_1}$ with the new letter $\mathsf{c_1}$, and each occurrence of the factor $\mathsf{b_2 b_2}$ with the new letter $\mathsf{c_2}$. In other words, we have $\mathbf{v} = p(\mathbf{v}')$, where $p \colon \{\mathsf{a_1}, \mathsf{b_1}, \mathsf{c_1}, \mathsf{a_2}, \mathsf{b_2}, \mathsf{c_2}\}^* \to \{\mathsf{a_1}, \mathsf{b_1}, \mathsf{a_2}, \mathsf{b_2}\}^*$ is the morphism that sends $\mathsf{c_1}$ to $\mathsf{b_1 b_1}$ and $\mathsf{c_2}$ to $\mathsf{b_2 b_2}$, and all other letters to themselves. From Lemma 3.4, we see that $\mathbf{v}'$ corresponds to an infinite walk on the graph G' drawn in Figure 4. Note in particular that $\mathbf{v}'$ alternates between letters of index 1 and index 2. It follows that we can write $\mathbf{v}' = \sigma(\mathbf{v}'')$, where $\mathbf{v}''$ is an infinite word over $\{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ and $\sigma \colon \{\mathsf{a}, \mathsf{b}, \mathsf{c}\} \to \{\mathsf{a_1}, \mathsf{b_1}, \mathsf{c_1}, \mathsf{a_2}, \mathsf{b_2}, \mathsf{c_2}\}$ is the map that adds index 1 and index 2 alternately to each letter of the input word, starting with 1.

We have shown that

$$\mathbf{z} = q(\mathbf{v}) = q(p(\mathbf{v}')) = q(p(\sigma(\mathbf{v}''))),$$

where $\mathbf{v}''$ is an infinite word over $\{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$. In the sequel, we prefer to work with a word over the alphabet $\Sigma_3 = \{0, 1, 2\}$ instead of $\{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$, so we write $\mathbf{v}'' = r(\mathbf{y})$, where

p	Length of a longest word $y \in \Sigma_3^*$ with prefix p such that $\tau(y)$ is a 16/7-power-free rich word
201	141
210	144
211	101

Table 2: Some words that do not appear in **y**.

$\mathbf{y} \in \{0, 1, 2\}^\omega$ and $r \colon \Sigma_3^* \to \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ is defined by $r(0) = \mathsf{a}$, $r(1) = \mathsf{b}$, and $r(2) = \mathsf{c}$. Thus we have

$$\mathbf{z} = q(p(\sigma(r(\mathbf{y})))).$$

We can express the composition of the maps r, σ, p, and q as the single finite-state transducer τ, which gives the following result.

Proposition 3.5. *We can write* $\mathbf{z} = \tau(\mathbf{y})$, *where* $\mathbf{y} \in \Sigma_3^\omega$, *and* τ *is the transducer drawn in Figure 1.*

3.2 The Second Layer

Throughout this subsection, let **z** be an infinite 16/7-power-free ternary rich word. Since $16/7 < 7/3$, we may use all of the results from Section 3.1. In particular, we may assume without loss of generality that **z** starts with 001002, and we know from Proposition 3.5 that $\mathbf{z} = \tau(\mathbf{y})$ for some word $\mathbf{y} \in \Sigma_3^\omega$, where τ is the transducer drawn in Figure 1. In this subsection, we begin to describe the structure of the word **y**.

First of all, by backtracking, we confirm that several short words do not appear in **y**. The results are summarized in Table 2.

Proposition 3.6. *A suffix of* **y** *has the form* $g(\mathbf{x})$ *for some word* $\mathbf{x} \in \Sigma_3^\omega$.

Proof. A computer backtracking search shows that the longest word $y \in \{0, 1\}^*$ such that $\tau(y)$ is 16/7-power-free and rich has length 18. So the letter 2 appears infinitely many times in **y**. Thus, it suffices to show that the only complete return words to 2 in **y** are $g(0)2 = 202$, $g(1)2 = 212$, and $g(2)2 = 22$.

Consider the tree drawn in Figure 5, which shows all possible complete returns to 2 in **y**. We explain why the words corresponding to the red leaves in the tree cannot appear in **y**.

- 200 has suffix 00. If **y** contains 00, then **z** contains $\tau(00)00 = 00100200$ or its sister. Since 00100200 is in Table 1, this is impossible.

- 201 is in Table 2.
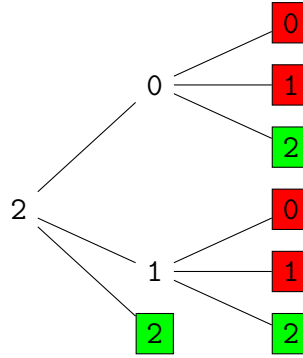
- 210 is in Table 2.

- 211 is in Table 2.

Figure 5: The tree showing all possible complete returns to 2 in **y**.

This means that the words corresponding to the green leaves in the tree are the only possible complete return words to 2 in **y**. Therefore, we conclude that a suffix of **y** is in $\{20, 21, 2\}^\omega$, as desired. $\qquad\square$

3.3 Forbidden Factors in the Inner Layers

In this subsection, we describe several families of factors that cannot appear in **x** if the word $\tau(g(f^n(\mathbf{x})))$ is 16/7-power-free and rich, where n is any nonnegative integer. Lemma 3.8 and Lemma 3.10 describe factors in **x** that lead to repetitions in $\tau(g(f^n(\mathbf{x})))$ with exponent at least 16/7, while Lemma 3.11 describes factors in **x** that lead to non-richness in $\tau(g(f^n(\mathbf{x})))$.

Note that when the transducer τ is applied to an r-power x, the alternation of τ between 1's and 2's may break the repetition in x. For example, if $x = 000$, then x is a cube, but $\tau(x) = 001002001$ is only a 3/2-power. Note, however, that if the period of x is even, then the alternation between 1's and 2's will line up after each period. For example, $x = 0101$ is a square with period 2, and we see that $\tau(x) = 00100202202001 00202202$ is also a square.

If $x \in \Sigma_3^*$ is an r-power of the form $x = p^r$, where p has an even number of 2's, then we say that x is an *even-r-power*. We say that a word is *even-r-power-free* if it contains no even-r-powers.

Lemma 3.7. *Let $x \in \Sigma_3^*$ be an even-r-power, and write $x = p^r$, where p has an even number of 2's. Then for all $n \geqslant 0$, the word $\tau(g(f^n(x)))$ has period $|\tau(g(f^n(p)))|$.*

Proof. We first claim that $f^n(p)$ has an even number of 2's for all $n \geqslant 0$. Note that $f(0)$ and $f(1)$ both have an even number of 2's, so the proof of the claim can be completed by a straightforward induction on n.

Now let $n \geqslant 0$. Since $f^n(p)$ has an even number of 2's, and $g(0)$ and $g(1)$ both have even length, it follows that $g(f^n(p))$ has even length. So by the observation preceding the lemma, we conclude that $\tau(g(f^n(x)))$ has period $|\tau(g(f^n(p)))|$. $\qquad\square$

Lemma 3.8. *Let $\mathbf{x} \in \Sigma_3^\omega$, and suppose for some $n \geqslant 0$ that $\tau(g(f^n(\mathbf{x})))$ is 16/7-power-free. Then $\mathbf{x}$ is 5-power-free and even-3-power-free.*

Proof. First, suppose towards a contradiction that $\mathbf{x}$ contains a 5-power, say $x = p^5$. Then we can write $x = (p^2)^{5/2}$, and p^2 has an even number of 2's. So by Lemma 3.7, the word $\tau(g(f^n(x)))$ has period $|\tau(g(f^n(p^2)))|$ and exponent $5/2$. This contradicts the assumption that $\tau(g(f^n(\mathbf{x})))$ is 16/7-power-free, so we see that $\mathbf{x}$ is 5-power-free.

Similarly, suppose that $\mathbf{x}$ contains an even-3-power, say $x = p^3$, where p has an even number of 2's. Then by Lemma 3.7, the word $\tau(g(f^n(x)))$ has period $|\tau(g(f^n(p)))|$ and exponent 3. This is a contradiction, and we conclude that $\mathbf{x}$ is even-3-power-free. $\qquad\square$

For finite words u and v over an alphabet Σ, we write $u \preceq v$ if $|u|_a \leqslant |v|_a$ for all $a \in \Sigma$. Note that non-erasing morphisms preserve the relation $\preceq$, and that $u \preceq v$ implies $|u| \leqslant |v|$. Finally, note that if $u, v \in \Sigma_3^*$ and $u \preceq v$, then $|\tau(u)| \leqslant |\tau(v)|$.

Lemma 3.9. *For all $n \geqslant 0$, we have*

$$|f^n(2)| \leqslant |f^n(0)| \leqslant |f^n(1)| \leqslant 2|f^n(2)|$$

and

$$|\tau(g(f^n(2)))| \leqslant |\tau(g(f^n(0)))| \leqslant |\tau(g(f^n(1)))| \leqslant 2|\tau(g(f^n(2)))|.$$

Proof. We show only that $|f^n(2)| \leqslant |f^n(0)|$ and $|\tau(g(f^n(2)))| \leqslant |\tau(g(f^n(0)))|$ for all $n \geqslant 0$. (The other inequalities can be proven in a similar manner.) Both inequalities are verified by inspection for $n \in \{0, 1\}$. When $n = 2$, we have

$$f^2(2) = 0102 \preceq 01022 = f^2(0),$$

so that $f^n(2) \preceq f^n(0)$ for all $n \geqslant 2$ by induction, and the desired inequalities follow immediately from the observations preceding the lemma. $\qquad\square$

Lemma 3.10. *Let $\mathbf{x} \in \Sigma_3^\omega$, and suppose for some $n \geqslant 0$ that $\tau(g(f^n(\mathbf{x})))$ is 16/7-power-free and rich. Then no factor from the set $F = F_1 \cup F_2 \cup F_3 \cup F_4$ appears internally in $\mathbf{x}$, where*

$$F_1 = \{00, 11, 212, 0101, 1010, 2222, 1222, 2221, 022022, 220220\},$$
$$F_2 = \{202202, 1022021, 1202201, 1(20102)^2 1, (021012)^{13/6}, (012021)^{13/6},$$
$$(21012010)^{21/8}, (21012210120)^{27/11}, (2101221012010)^{31/13}\},$$
$$F_3 = \{2(2101)^{17/4} 2, (2101210122101)^{31/13}\}, \text{ and}$$
$$F_4 = \{(0222)^{17/4} 1, (22010)^{12/5}, (022201)^{29/6}, (0222010222)^{12/5},$$
$$(0222022201)^{12/5}, (0222010222010222022201)^{5/2}\}.$$

Proof. Let $w \in F$, and suppose that w appears internally in $\mathbf{x}$. Then the word $\tau(g(f^n(\mathbf{x})))$ contains the word $\tau(g(f^n(awb)))$ or its sister for some $a, b \in \{0, 1, 2\}$. The general idea is to show that for all $a, b \in \{0, 1, 2\}$, the word $\tau(g(f^n(awb)))$ has a factor of exponent at least 16/7, which is a contradiction. We do this for each word in F_1 below. The proofs

for the words in F_2, F_3, and F_4 are similar, and are omitted.[2] Sometimes, we show the stronger statement that for all $b \in \{0, 1, 2\}$, the word $\tau(g(f^n(wb))$ has a factor of exponent at least $16/7$, which implies that w does not appear in $\mathbf{x}$ at all (i.e., neither internally nor as a prefix).

00: For $n = 0$, we check directly that for all $b \in \{0, 1, 2\}$, the word $\tau(g(00b))$ contains a factor of exponent at least $16/7$. So we may assume that $n \geqslant 1$. For all $b \in \{0, 1, 2\}$, we check that the word $f(00b)$ contains the even-$5/2$-power 01010. But this means that $\tau(g(f^n(00b)))$ contains the word $\tau(g(f^{n-1}(01010)))$ or its sister. By Lemma 3.7, the word $\tau(g(f^{n-1}(01010)))$ has period $|\tau(g(f^{n-1}(01)))|$. Further, by Lemma 3.9, we have $|\tau(g(f^{n-1}(01)))| \leqslant 3|\tau(g(f^{n-1}(0)))|$, so that $\tau(g(f^{n-1}(01010)))$ has exponent at least $7/3$.

11: For $n = 0$, we check directly that for all $b \in \{0, 1, 2\}$, the word $\tau(g(11b))$ contains a factor of exponent at least $16/7$. So we may assume that $n \geqslant 1$. For all $b \in \{0, 1, 2\}$, we check that the word $f(11b)$ contains the even-$7/3$-power 0220220. But this means that $\tau(g(f^n(11b)))$ contains the word $\tau(g(f^{n-1}(0220220)))$ or its sister, and by Lemma 3.7 and Lemma 3.9, the word $\tau(g(f^{n-1}(0220220)))$ has exponent at least $7/3$.

212: For $n = 0, 1, 2$, we check directly that for all $b \in \{0, 1, 2\}$, the word $\tau(g(f^n(212b)))$ has a factor of exponent at least $16/7$. So we may assume that $n \geqslant 3$. For all $b \in \{0, 1, 2\}$, we check that the word $f^3(212b)$ contains the even-$31/13$-power $p^{31/13}$, where $p = 2010201022010$. This means that $\tau(g(f^n(212b)))$ contains $\tau(g(f^{n-3}(p^{31/13})))$ or its sister. We write $p^{31/13} = (p_1 p_2 p_3)^2 p_1$, where $p_1 = 20102$, $p_2 = 0102$, and $p_3 = 2010$. Observe that $p_2, p_3 \preceq p_1$, so that $|\tau(g(f^{n-3}(p_1 p_2 p_3)))| \leqslant 3|\tau(g(f^{n-3}(p_1)))|$. From Lemma 3.7, it follows that $\tau(g(f^{n-3}(p^{31/13})))$ has exponent at least $7/3$.

2222: For $n = 0, 1$, we check that for all $b \in \{0, 1, 2\}$, the word $\tau(g(f^n(2222b)))$ has a factor of exponent at least $16/7$. So we may assume that $n \geqslant 2$. For all $b \in \{0, 1, 2\}$, we check that $f^2(2222b)$ contains the 5-power $(0102)^5$. Thus $\tau(g(f^n(2222a)))$ contains $\tau(g(f^{n-2}((0102)^5)))$ or its sister, and by Lemma 3.7, the word $\tau(g(f^{n-2}((0102)^5)))$ has exponent $5/2$.

1222: We have already seen that $\mathbf{x}$ does not contain the factor 2222, so it suffices to show that for all $b \in \{0, 1\}$, the word $\tau(g(f^n(1222b)))$ contains a factor of exponent at least $16/7$. We check this directly for $n \leqslant 3$, so we may assume that $n \geqslant 4$. For all $b \in \{0, 1\}$, we check that $f^4(1222b)$ contains the even-$19/8$-power $p^{19/8}$, where $p = (02010220102010220102)^2$. We write $p^{19/8} = (p_1 p_2 p_3)^2 p_1$, where $p_1 = 020102201020102$, $p_2 = 201020201022$, and $p_3 = 0102010220102$. Observe that $p_2, p_3 \preceq p_1$. So by Lemma 3.7, we see that $\tau(g(f^{n-4}(p^{19/8})))$ has exponent at least $7/3$.

2221: We have already seen that $\mathbf{x}$ does not contain the factor 2222, so it suffices to show that for all $a \in \{0, 1\}$ and $b \in \{0, 1, 2\}$, the word $\tau(g(f^n(a2221b)))$ has a factor of exponent at least $16/7$. We check this directly for $n \leqslant 3$, so we may assume that $n \geqslant 4$. For all $a \in \{0, 1\}$ and $b \in \{0, 1, 2\}$, we check that $f^4(a2221b)$ contains the even-$19/8$-power $p^{19/8}$, where $p = (20102010220102020102)^2$. But one can show that $\tau(g(f^{n-4}(p^{19/8})))$ has exponent at least $7/3$ as in the proof for 1222.

1010: We have already seen that $\mathbf{x}$ contains no 00, so it suffices to show that for all

---

[2]Please consult https://github.com/japeltom/ternary-rich-words-verification for the arguments in the omitted cases.

$b \in \{1, 2\}$, the word $\tau(g(f^n(1010b)))$ has a factor of exponent at least 16/7. We check this directly for $n = 0$, so we may assume that $n \geqslant 1$. For all $b \in \{1, 2\}$, we check that the word $f(1010b)$ contains the even-12/5-power $(02201)^{12/5}$. But by Lemma 3.7 and Lemma 3.9, the word $\tau(g(f^{n-1}((02201)^{12/5})))$ has exponent at least 7/3.

0101: We have already seen that $\mathbf{x}$ contains no 00, no 11, and no 1010, so it suffices to show that the word $\tau(g(f^n(201012)))$ has a factor of exponent at least 16/7. For $n = 0$, we check directly that $\tau(g(201012))$ contains a factor of exponent at least 16/7, so we may assume that $n \geqslant 1$. We observe that $f(201012)$ contains the even-12/5-power $(20102)^{12/5}$. But by Lemma 3.7 and Lemma 3.9, the word $\tau(g(f^{n-1}((20102)^{12/5})))$ has exponent at least 7/3.

022022 and 220220: Let $u \in \{022022, 220220\}$. For $n = 0, 1$, we check directly that for all $a, b \in \{0, 1, 2\}$, the word $\tau(g(f^n(aub)))$ contains a factor of exponent at least 16/7. So we may assume that $n \geqslant 2$. For all $a, b \in \{0, 1, 2\}$, we check that $f^2(aub)$ contains one of the following even-31/13-powers: $(2010220102010)^{31/13}$, $(2010201020102)^{31/13}$. By an argument similar to the one used for 212 above, we see that $\tau(g(f^n(aub)))$ contains a factor of exponent at least 7/3. $\qquad\square$

In Lemma 3.8 and Lemma 3.10, we demonstrated that certain factors in $\mathbf{x}$ lead to long repetitions in $\tau(g(f^n(\mathbf{x})))$. We now wish to demonstrate that certain factors in $\mathbf{x}$ lead to non-rich factors in $\tau(g(f^n(\mathbf{x})))$. The maps f and g belong to the well-studied class P_{ret} (see [5, 22]). In particular, it is known that morphisms in class P_{ret} preserve non-richness of infinite words. However, the alternation of τ complicates things for us here.

We say that a word $w \in \Sigma_3^*$ is *poor* if every palindromic prefix of w with an even number of 2's occurs at least once more in w with an even number of 2's before the occurrence. In other words, w is poor if for every $u \in \Sigma_3^*$, if u is a palindromic prefix of w with an even number of 2's, then there exist words $p, s \in \Sigma_3^*$ such that $w = pus$, $p \neq \varepsilon$, and p has an even number of 2's. The following are examples of poor words:

- 2012, whose only palindromic prefix with an even number of 2's is the empty word, which occurs again as a suffix with an even number of 2's before it;

- 01220, which has two palindromic prefixes with an even number of 2's, namely ε and 0, both of which occur as suffixes with an even number of 2's before them; and

- 0220102020220, which has three palindromic prefixes with an even number of 2's, namely ε, 0, and 0220, all of which occur as suffixes with an even number of 2's before them.

On the other hand, the word 0120 is neither rich nor poor—the palindromic prefix 0 occurs just once more as a suffix, and there are an *odd* number of 2's before this second occurrence. We will prove the following.

Lemma 3.11. *Let $\mathbf{x} \in \Sigma_3^\omega$, and suppose for some $n \geqslant 0$ that $\tau(g(f^n(\mathbf{x})))$ is rich. Then $\mathbf{x}$ contains no poor factor.*

We first prove several intermediate lemmas, for which we make use of one more term. We say that a word $w \in \Sigma_3^*$ is *middle-class* if it begins in 2 and every odd-length palindromic prefix of w occurs at least once more in w starting at an even index. In other words, w is middle-class if it begins in 2 and for every word $u \in \Sigma_3^*$, if u is a palindromic prefix of w of odd length, then there exist words $p, s \in \Sigma^*$ such that $w = pus$, $p \neq \varepsilon$, and p has even length. The following are examples of middle-class words:

- 2202122, whose only palindromic prefix of odd length is 2, which occurs again starting at even index 6; and

- 202122202, whose only palindromic prefixes of odd length are 2 and 202, which both occur again starting at even index 6.

On the other hand, the word 2012 is not middle-class. It has palindromic prefix 2 of odd length, which occurs again only at odd index 3.

Roughly speaking, we will show the following:

- f sends poor words to poor words;

- g sends poor words to middle-class words; and

- τ sends middle-class words to non-rich words.

Throughout, we use the following result.

Lemma 3.12. *Let $u, w \in \Sigma_3^*$.*

(i) *If $f(u)0$ is a palindromic prefix of $f(w)0$, then u is a palindromic prefix of w.*

(ii) *If $g(u)2$ is a palindromic prefix of $g(w)2$, then u is a palindromic prefix of w.*

(iii) *If u starts with 2 and $\tau(u)00$ is a palindromic prefix of $\tau(w)00$, then u is a palindromic prefix of w of odd length.*

Proof. Parts (i) and (ii) can be proven in a manner similar to [16, Lemma 4]. The argument for part (iii) is similar; the alternation of τ forces u to have odd length. $\square$

Lemma 3.13. *Let $w \in \Sigma_3^*$. If w is poor, then $f(w)0$ is poor.*

Proof. Let p be a palindromic prefix of $f(w)0$ with an even number of 2's. Note that $f(w)0$ starts with 0. If $p = \varepsilon$, then p occurs again in $f(w)0$ after the prefix 0, which has an even number of 2's. So we may assume that $p \neq \varepsilon$. Since p is a palindromic prefix of $f(w)0$, and $f(w)0$ starts with 0, we see that p must begin and end in 0. Hence, we can write $p = f(u)0$ for some word $u \in \Sigma_3^*$, and by Lemma 3.12, we see that u is a palindromic prefix of w. Since p has an even number of 2's, and since 2 is the only letter whose f-image has an odd number of 2's, we conclude that the word u must also contain an even number of 2's. Since w is poor, the word u occurs at least once more in w with an even number of 2's before the occurrence. It follows that the word $p = f(u)0$ occurs at least once more in $f(w)0$ with an even number of 2's before the occurrence, i.e., that $f(w)0$ is poor. $\square$

Lemma 3.14. *Let $w \in \Sigma_3^*$. If w is poor, then $g(w)2$ is middle-class.*

Proof. First note that $g(w)2$ begins in 2. Let p be an odd-length palindromic prefix of $g(w)2$. Since p is palindromic, it must begin and end with 2, hence we can write $p = g(u)2$ for some word $u \in \Sigma_3^*$, and by Lemma 3.12, the word u is a palindromic prefix of w. Since p has odd length, we see that $g(u)$ has even length, and since 2 is the only letter whose g-image has odd length, it follows that u has an even number of 2's. Since w is poor, the word u occurs at least once more in w with an even number of 2's before the occurrence. It follows that $p = g(u)2$ occurs at least once more in $g(w)2$ starting at an even index. Therefore, we conclude that the word $g(w)2$ is middle-class. $\qquad\square$

Lemma 3.15. *Let $w \in \Sigma_3^*$. If w is middle-class, then $\tau(w)00$ is not rich.*

Proof. Let p be the longest palindromic prefix of $\tau(w)00$. Since w is middle-class, it begins in 2, and we see that p must have prefix $\tau(2)00$. Since p is palindromic, it must also end in $\tau(2)00$, and not its sister. Note that every occurrence of $\tau(2)00$ in $\tau(w)00$ corresponds in an obvious manner to an occurrence of 2 in w at an even index. So we can write $p = \tau(u)00$ for some word $u \in \Sigma_3^*$. Further, by Lemma 3.12, the word u is a palindromic prefix of w of odd length. Since w is middle-class, the word u occurs at least once more in w starting at an even index. It follows that $p = \tau(u)00$ occurs at least twice in $\tau(w)00$. Since p is the longest palindromic prefix of w, it follows that every palindromic prefix of w occurs at least twice in w. By Lemma 3.1, we conclude that w is not rich. $\quad\square$

We are now ready to prove Lemma 3.11.

Proof of Lemma 3.11. Suppose towards a contradiction that $\mathbf{x}$ contains a poor factor w. By a straightforward induction using Lemma 3.13, we see that the word

$$W_1 = f(f(\cdots(f(w)0)\cdots)0)0,$$

where f is applied n times, is poor. Hence, by Lemma 3.14, the word $W_2 = g(W_1)2$ is middle-class, and by Lemma 3.15, the word $W_3 = \tau(W_2)00$ is not rich. But W_3 or its sister is a factor of $\tau(g(f^n(\mathbf{x})))$, which contradicts the assumption that $\tau(g(f^n(\mathbf{x})))$ is rich. $\qquad\square$

3.4 The Inner Layers

Proposition 3.16. *Let $\mathbf{x} \in \Sigma_3^\omega$, and suppose for some $n \geqslant 0$ that $\tau(g(f^n(\mathbf{x})))$ is $16/7$-power-free and rich. Then a suffix of $\mathbf{x}$ has the form $f(\mathbf{x}')$ for some word $\mathbf{x}' \in \Sigma_3^\omega$.*

Proof. First observe the following.

- By Lemma 3.8, the word $\mathbf{x}$ is 5-power-free and even-3-power-free.

- By Lemma 3.10, no factor from $F = F_1 \cup F_2 \cup F_3 \cup F_4$ appears internally in $\mathbf{x}$.

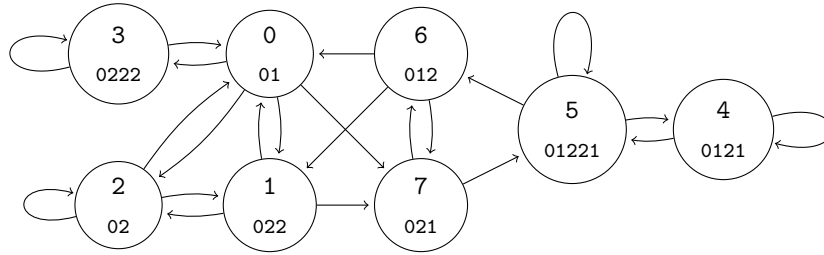- By Lemma 3.11, the word $\mathbf{x}$ contains no poor factor.

Figure 6: The directed graph H showing possible transitions between letters in $\mathbf{x}'$.

So by taking a suffix if necessary, we may assume that $\mathbf{x}$ is 5-power-free, even-3-power-free, and contains neither poor words nor words from F as factors. We use these properties frequently throughout the remainder of the proof without further reference.

For ease of writing, we consider an extension $\hat{f}$ of f to Σ_8^*, defined by

$$
\begin{aligned}
&\hat{f}(0) = 01, &\quad &\hat{f}(4) = 0121, \\
&\hat{f}(1) = 022, &\quad &\hat{f}(5) = 01221, \\
&\hat{f}(2) = 02, &\quad &\hat{f}(6) = 012, \\
&\hat{f}(3) = 0222, &\quad &\hat{f}(7) = 021.
\end{aligned}
$$

Observe that for all $a \in \Sigma_3$, we have $\hat{f}(a) = f(a)$, so $\hat{f}$ is indeed an extension of f.

Claim 3.17. *A suffix of $\mathbf{x}$ has the form $\hat{f}(\mathbf{x}')$ for some word $\mathbf{x}' \in \Sigma_8^\omega$. Further, the word $\mathbf{x}'$ corresponds to an infinite walk on the graph H drawn in Figure 6.*

Proof of Claim 3.17. The longest word over $\{1, 2\}$ that contains no factor from F_1 has length 4, so the letter 0 must occur infinitely many times in $\mathbf{x}$, and by taking a suffix if necessary, we may assume that $\mathbf{x}$ starts with 0. Thus, it suffices to show that the only possible complete returns to 0 in $\mathbf{x}$ have the form $\hat{f}(a)0$ for some $a \in \Sigma_8$. Consider the tree drawn in Figure 7, which shows all possible complete returns to 0 in $\mathbf{x}$. The words corresponding to red leaves in the tree have a suffix in F_1, while the words corresponding to yellow leaves in the tree are poor. So the words corresponding to green leaves in the tree are the only possible complete returns to 0 in $\mathbf{x}$. Thus, we can write $\mathbf{x} = \hat{f}(\mathbf{x}')$ for some $\mathbf{x}' \in \Sigma_8^\omega$.

It remains to show that $\mathbf{x}'$ corresponds to an infinite walk on the graph H drawn in Figure 6. First note that the image of every letter under $\hat{f}$ has prefix 0. So if u is a factor of $\mathbf{x}'$, then $\hat{f}(u)0$ is a factor of $\mathbf{x}$. We use this fact to show that some letters cannot appear immediately after others in $\mathbf{x}'$. Consider the letter 4, for example. It cannot be followed by the letter 0, because $\hat{f}(40)0 = 0121010$ contains the factor $1010 \in F_1$. It cannot be followed by a letter from $\{1, 2, 3, 7\}$, because for all $a \in \{1, 2, 3, 7\}$, the word $\hat{f}(4a)$ contains the poor word 2102. Finally, it cannot be followed by the letter 6, because $\hat{f}(46)0 = 01210120$ is poor. So the letter 4 can only be followed by 4 or 5 in $\mathbf{x}'$. By performing a similar analysis on all letters of Σ_8, ruling out the transition from a to b

Figure 7: The tree showing all possible complete returns to `0` in **x**.

if the word $\hat{f}(ab)0$ contains a poor word or a word in F_1, we obtain the directed graph H. ∎

So $\mathbf{x}$ has the form $\hat{f}(\mathbf{x}')$ for some word $\mathbf{x}' \in \Sigma_8^\omega$. Our goal now is to show that in fact we have $\mathbf{x}' \in \Sigma_3^\omega$.

Claim 3.18. *The word $\mathbf{x}'$ contains neither 6 nor 7, i.e., we have $\mathbf{x}' \in \Sigma_6^\omega$.*

Proof of Claim 3.18. Consider an occurrence of 6 in $\mathbf{x}'$. From the digraph H, we see that 6 is followed by a factor of the form ua, where $u \in \{0, 1\}^*$ and $a \in \{2, 3, 7\}$. As 00, 11, 0101, and 1010 are in F_1, we see that $u \in \{\varepsilon, 0, 1, 01, 10, 010, 101\}$. If $a \neq 7$, then the only palindromic prefixes of $\hat{f}(6ua)0$ are ε and 0, because $\hat{f}(6ua)0$ has prefix 012 and contains no 21. Further, both palindromic prefixes occur as a suffix in $\hat{f}(6ua)0$ after an even number of 2's, meaning that $\hat{f}(6ua)0$ is poor. Thus every time the letter 6 appears in $\mathbf{x}'$, it appears as a prefix of $6u7$ for some $u \in \{\varepsilon, 0, 1, 01, 10, 010, 101\}$. But for each $u \in \{1, 01, 10, 010, 101\}$, we check that $\hat{f}(6u7)$ contains a factor in F_2. So every 6 in $\mathbf{x}'$ occurs as a prefix of either 67 or 607.

Now consider an occurrence of 7 in $\mathbf{x}'$. From the digraph H, we see that 7 is followed by a factor of the form $v6$, where $v \in \{4, 5\}^*$. First note that 55 is not a factor of v, since in that case $\hat{f}(7v6)$ contains the factor $21\hat{f}(55)012 = (21012)^3$, which is an even-3-power. Next, note that 4 is not a factor of v, since in that case v must have suffix 45, and in turn $\hat{f}(7v6)0$ has suffix $\hat{f}(456)0$, which is poor. So we must have $v \in \{\varepsilon, 5\}$, i.e., every 7 in $\mathbf{x}'$ occurs as a prefix of either 76 or 756.

So suppose that the letter 7 appears in $\mathbf{x}'$. By taking a suffix of $\mathbf{x}$ if necessary, we may assume that $\mathbf{x}'$ starts with 7. Then we can write $\mathbf{x}' = \phi(\mathbf{x}'')$ for some $\mathbf{x}'' \in \Sigma_4^\omega$, where

$$
\begin{aligned}
\phi(0) &= 76, \\
\phi(1) &= 760, \\
\phi(2) &= 756, \\
\phi(3) &= 7560.
\end{aligned}
$$

In turn, we have $\mathbf{x} = \hat{f}(\phi(\mathbf{x}''))$. Observe that for every letter $a \in \Sigma_4$, the word $\hat{f}(\phi(a))$ has an even number of 2's. Since $\mathbf{x}$ is even-3-power-free, we see that $\mathbf{x}''$ must be 3-power-free. We also claim that $\mathbf{x}''$ contains no factor from the set

$$F_\phi = \{00, 11, 22, 33, 01, 20, 31\}.$$

Let $w \in F_\phi$, and suppose that w appears in $\mathbf{x}''$. Note that for every letter $a \in \Sigma_4$, the word $\hat{f}(\phi(a))$ has prefix $\hat{f}(\phi(0)) = 021012$. It follows that the word $\hat{f}(\phi(w0))$ appears in $\mathbf{x}$. However, we check that $\hat{f}(\phi(w0))$ contains a factor from F_2, hence we conclude that $\mathbf{x}''$ contains no factor from F_ϕ.

Now we run a backtracking algorithm that searches through the tree of all words $u \in \Sigma_4^*$ such that

(i) u is 3-power-free;

(ii) u contains no factor from F_ϕ; and

(iii) $\hat{f}(\phi(u))$ contains no poor factor.

We find that the longest such word has length 8. Thus, we conclude that the letter 7 does not appear in $\mathbf{x}'$. Since every 6 in $\mathbf{x}'$ appears as a prefix of either 67 or 607, it follows that the letter 6 does not appear in $\mathbf{x}'$ either. ∎

Claim 3.19. *The word $\mathbf{x}'$ contains neither 4 nor 5, i.e., we have $\mathbf{x}' \in \Sigma_4^\omega$.*

Proof of Claim 3.19. Suppose that 4 or 5 appears in $\mathbf{x}'$. By Claim 3.18, we have $\mathbf{x}' \in \Sigma_6^\omega$. So by considering the digraph H, we see that $\mathbf{x}' \in \{4,5\}^\omega$. Since $\mathbf{x}$ contains no 5-power, we see that $\mathbf{x}'$ contains no 5-power. Further, we claim that some suffix of $\mathbf{x}'$ contains no factor from the set $F' = \{55, 444, 5445\}$. Observe that for all $a \in \{4,5\}$, the word $\hat{f}(a)$ has prefix 012 and suffix 21. So if some word u appears internally in $\mathbf{x}'$, then $\mathbf{x}$ contains the word $21\hat{f}(u)012$. We handle each word in F' separately.

- 55: Suppose that 55 appears internally in $\mathbf{x}'$. It follows that $\mathbf{x}$ contains the word $21\hat{f}(55)012 = (21012)^3$. But this contradicts the fact that $\mathbf{x}$ is even-3-power-free.

- 444: We first claim that 4444 does not appear internally in $\mathbf{x}$. For if 4444 appears internally in $\mathbf{x}'$, then $\mathbf{x}$ contains the 5-power $21\hat{f}(4444)01 = (2101)^5$; a contradiction. So by taking a suffix if necessary, we may assume that $\mathbf{x}'$ contains no 4444. Now suppose that 444 appears internally in $\mathbf{x}'$. Then $\mathbf{x}'$ contains the factor 54445. But this is impossible, since $\hat{f}(54445)$ contains the factor $2(2101)^{17/4}2 \in F_3$.

- 5445: Suppose that 5445 appears in $\mathbf{x}'$ with at least two letters before it. Then since 55 does not appear internally in $\mathbf{x}'$, we see that $\mathbf{x}'$ must contain an internal occurrence of 454454. But then $\mathbf{x}$ contains the word $21\hat{f}(454454)012 = (2101210122101)^{31/13} \in F_3$, a contradiction.

Now we run a backtracking algorithm that searches through the tree of all words $u \in \{4,5\}^*$ such that

(i) u is 5-power-free;

(ii) u contains no factor from F'; and

(iii) $\hat{f}(u)$ contains no poor factor.

The longest such word has length 11. Thus, we conclude that neither 4 nor 5 appears in $\mathbf{x}'$. ∎

Claim 3.20. *The word $\mathbf{x}'$ contains no 3, i.e., we have $\mathbf{x}' \in \Sigma_3^\omega$.*

Proof of Claim 3.20. Suppose that the letter 3 appears in $\mathbf{x}'$. From Claim 3.19, we have $\mathbf{x}' \in \Sigma_4^\omega$. We claim that the letter 2 does not appear in $\mathbf{x}'$. For if it did, then $\mathbf{x}'$ would have a factor of the form $3u2$ or $2u3$, where $u \in \{0,1\}^*$. But for all $u \in \{0,1\}^*$, the words $\hat{f}(3u2)0$ and $\hat{f}(2u3)0$ are poor; their only palindromic prefixes with an even number of 2's are ε and 0, and they occur again (as suffixes) after an even number of 2's.

So we have $\mathbf{x}' \in \{0,1,3\}^\omega$. Observe that $\mathbf{x}'$ does not contain the factor 3333. For if it did, then we see from the digraph H that $\mathbf{x}'$ would contain either 33333 or 33330. But 33333 is a 5-power, and $\hat{f}(33330) = (0222)^{17/4}1$ is in F_4. By taking a suffix if necessary, assume that $\mathbf{x}'$ starts with 0. Then by considering the digraph H, and remembering that $\mathbf{x}'$ does not contain 3333, we see that we must have $\mathbf{x}' = \psi(\mathbf{x}'')$ for some word $\mathbf{x}'' \in \{0,1,2,3\}^\omega$, where

$$\psi(0) = 03,$$
$$\psi(1) = 033,$$
$$\psi(2) = 0333,$$
$$\psi(3) = 01.$$

In turn, we have $\mathbf{x} = \hat{f}(\psi(\mathbf{x}''))$. Since $\mathbf{x}$ contains no 5-power, we see that $\mathbf{x}''$ contains no 5-power. Note that for all $a \in \Sigma_4$, the word $\hat{f}(\psi(a))$ has prefix 01022 and suffix 22. Further, for all $a \in \Sigma_3$, the word $\hat{f}(\psi(a))$ has suffix 0222, and for all $v \in \Sigma_3^*$ with $|v| \geqslant 2$, the word $\hat{f}(\psi(v))$ has prefix 0102220.

Now we claim that some suffix of $\mathbf{x}''$ belongs to Σ_2^ω. Suppose first that the letter 3 appears internally in $\mathbf{x}''$. Then $\mathbf{x}$ contains the word $22\hat{f}(\psi(3))01022 = (22010)^{12/5} \in F_4$, a contradiction. So by taking a suffix if necessary, we may assume that $\mathbf{x}'' \in \Sigma_3^\omega$. Suppose now that the letter 2 appears in $\mathbf{x}''$. If 20 appears in $\mathbf{x}''$, then $\mathbf{x}$ contains the poor factor $\hat{f}(\psi(20))010$, a contradiction. If 21 appears in $\mathbf{x}''$, then $\mathbf{x}$ contains the word $\hat{f}(\psi(21))010222$, which has suffix $(0222022201)^{12/5} \in F_4$, a contradiction. So every occurrence of 2 in $\mathbf{x}''$ must be followed by another 2. But then $\mathbf{x}''$ contains the 5-power 22222, a contradiction.

So we may assume that $\mathbf{x}'' \in \Sigma_2^\omega$. We now claim that some suffix of $\mathbf{x}''$ contains no factor from the set

$$F_\psi = \{11, 000, 1001\}.$$

We handle each word in F_ψ separately.

- 11: Suppose that 11 appears internally in $\mathbf{x}''$. Then $\mathbf{x}$ contains the word

$$0222\hat{f}(\psi(11)) = (0222010222)^{12/5} \in F_4,$$

  a contradiction.

- 000: Suppose that 000 appears internally in $\mathbf{x}''$. Then $\mathbf{x}$ contains the word

$$0222\hat{f}(\psi(000))0102220 = (022201)^{29/6} \in F_4,$$

  a contradiction.

- 1001: Suppose that 1001 appears with at least two letters before it in $\mathbf{x}''$. Since 11 does not appear internally in $\mathbf{x}''$, we see that the word 010010 appears internally in $\mathbf{x}''$. But then $\mathbf{x}$ contains the word

$$0222\hat{f}(\psi(010010))0102220 = (0222010222010222022201)^{5/2} \in F_4,$$

a contradiction.

Now we run a backtracking algorithm that searches through the tree of all words $u \in \Sigma_2^*$ such that

(i) u is 5-power-free;

(ii) u contains no factor from F_ψ; and

(iii) $\hat{f}(\psi(u))$ contains no poor factor.

The longest such word has length 11. Thus, we conclude that 3 does not appear in $\mathbf{x}'$. ■

Since $\mathbf{x}' \in \Sigma_3^\omega$, we have $\mathbf{x} = \hat{f}(\mathbf{x}') = f(\mathbf{x}')$, and this completes the proof of the proposition. □

Proof of Theorem 1.2. We proceed by induction on n. As argued at the beginning of Section 3.1, by permuting the letters and taking a suffix if necessary, we may assume that $\mathbf{z}$ has prefix 001002. Then by Proposition 3.5, we have $\mathbf{z} = \tau(\mathbf{y})$ for some word $\mathbf{y} \in \Sigma_3^\omega$, and by Proposition 3.6, some suffix of $\mathbf{y}$ has the form $g(\mathbf{x}_0)$ for some word $\mathbf{x}_0 \in \Sigma_3^\omega$. This completes the base case.

Now suppose for some integer $k \geqslant 0$ that some suffix of $\mathbf{z}$ has the form $\tau(g(f^k(\mathbf{x}_k)))$ for some word $\mathbf{x}_k \in \Sigma_3^\omega$. Then by Proposition 3.16, some suffix of $\mathbf{x}_k$ has the form $f(\mathbf{x}_{k+1})$ for some word $\mathbf{x}_{k+1} \in \Sigma_3^\omega$. It follows that some suffix of $\mathbf{z}$ has the form $\tau(g(f^{k+1}(\mathbf{x}_{k+1})))$. Therefore, we conclude that the theorem statement holds by mathematical induction. □

4 Richness

Throughout this section, let $\mathbf{x} = f^\omega(0)$, $\mathbf{y} = g(\mathbf{x})$, and $\mathbf{z} = \tau(\mathbf{y})$. For an infinite word $\mathbf{u}$, we let $C_\mathbf{u} \colon \mathbb{N} \to \mathbb{N}$ denote the *factor complexity* of $\mathbf{u}$, and we let $P_\mathbf{u} \colon \mathbb{N} \to \mathbb{N}$ denote the *palindromic complexity* of $\mathbf{u}$. That is, $C_\mathbf{u}(n)$ is the number of distinct factors of $\mathbf{u}$ of length n for all $n \geqslant 0$, and $P_\mathbf{u}(n)$ is the number of distinct palindromic factors of $\mathbf{u}$ of length n for all $n \geqslant 0$.

The goal of this section is to prove the following result.

Proposition 4.1. *The word $\mathbf{z}$ is rich.*

Along the way, we will also show that $\mathbf{x}$ and $\mathbf{y}$ are rich. In Section 4.1 and Section 4.2, we determine the factor complexity and the palindromic complexity, respectively, of the word $\mathbf{z}$. Proposition 4.1 follows immediately from these results and the following result that characterizes infinite rich words in terms of their factor complexity and palindromic complexity functions. Notice that the language of $\mathbf{z}$ is closed under reversal as $\mathbf{z}$ is uniformly recurrent and contains arbitrarily long palindromes (see the results of Section 4.2).

Proposition 4.2. *[11, Thm. 1.1] Let $\mathbf{u}$ be an infinite word whose language is closed under reversal. Then $\mathbf{u}$ is rich if and only if*

$$P_{\boldsymbol{u}}(n) + P_{\boldsymbol{u}}(n+1) = C_{\boldsymbol{u}}(n+1) - C_{\boldsymbol{u}}(n) + 2 \tag{1}$$

for all $n \geqslant 0$.

Before we proceed with determining the factor complexity and palindromic complexity functions of $\mathbf{z}$, we introduce some terminology and prove a preliminary result. Let $\mathbf{u}$ be an infinite word, and let $w \in \text{Fact}(\mathbf{u})$. A *left extension* (resp. *right extension*) of w in $\mathbf{u}$ is a word of the form $aw \in \text{Fact}(\mathbf{u})$ (resp. $wa \in \text{Fact}(\mathbf{u})$), where a is a letter. A *bi-extension* of w in $\mathbf{u}$ is a word of the form $awb \in Fact(\mathbf{u})$, where a and b are letters. If w is a palindrome, then a *palindromic extension* of w in $\mathbf{u}$ is a bi-extension of w that is a palindrome. We say that w is *left-special* (resp. *right-special*) in $\mathbf{u}$ if it has at least two distinct left (resp. right) extensions, and we say that w is *bispecial* in $\mathbf{u}$ if it is both left-special and right-special.

Lemma 4.3. *If z is a factor of $\mathbf{z}$, then its sister is a factor of $\mathbf{z}$.*

Proof. We first claim that for each $x \in \text{Fact}(\mathbf{x})$, there exist u and u' such that both ux and $u'x$ are prefixes of $\mathbf{x}$, and $|u|_2$ and $|u'|_2$ have different parity. Let $x \in \text{Fact}(\mathbf{x})$. Then x occurs in $f^n(0)$ for some $n \geqslant 0$, i.e., there is some $n \geqslant 0$ such that $f^n(0)$ has prefix ux for some word u. Set $v_n = f^{n+2}(0)f^{n+1}(02)f^n(0)$. It is not difficult to see that v_n is a prefix of $f^{n+3}(0)$, and hence of $\mathbf{x}$. Letting $u' = f^{n+2}(0)f^{n+1}(02)u$, we see that v_n (and hence $\mathbf{x}$) has both ux and $u'x$ as prefixes. Note that $|f^k(0)|_2$ is even for all $k \geqslant 0$ and $|f^k(2)|_2$ is odd for all $k \geqslant 0$. So $|f^{n+2}(0)f^{n+1}(02)|_2$ is odd, and we conclude that $|u|_2$ and $|u'|_2$ have different parity.

Now let $y \in \text{Fact}(\mathbf{y})$. Then there exists a factor $x \in \text{Fact}(\mathbf{x})$ such that $g(x) = \alpha y \beta$ for some words α and β. Let u and u' be as above, so that $g(u)\alpha y$ and $g(u')\alpha y$ are prefixes of $\mathbf{y}$. Since $|g(v)|$ is even if and only if $|v|_2$ is even, it follows that the lengths of $g(u)\alpha$ and $g(u')\alpha$ have different parities. So y occurs starting at both an even position and an odd position in $\mathbf{y}$.

Finally, let $z \in \text{Fact}(\mathbf{z})$. Then z occurs in $\tau(y)$ for some prefix y of $\mathbf{y}$. From above, we see that y must also occur starting at an odd position in $\mathbf{y}$, hence $\overline{\tau}(y)$ also occurs in $\mathbf{z}$, and $\overline{\tau}(y)$ contains the sister of z. $\square$

4.1 Factor complexity

Proposition 4.4. *For all $n \geqslant 1$, the word $\mathbf{x} = f^\omega(0)$ has exactly two right-special factors u and v of length n. The word u ends with 0 and $u1, u2 \in \text{Fact}(\mathbf{x})$, and the word v ends with 2 and $v0, v2 \in \text{Fact}(\mathbf{x})$.*

Proof. Since $00, 11, 12, 21 \notin \text{Fact}(\mathbf{x})$, every non-empty factor of $\mathbf{x}$ has at most two right extensions in $\text{Fact}(\mathbf{x})$. Moreover, there is no right-special factor ending in 1 since neither 11 nor 12 belongs to $\text{Fact}(\mathbf{x})$. Both 0 and 2 are right-special as $01, 02, 20$, and 22 all belong to $\text{Fact}(\mathbf{x})$.

We claim that **x** contains arbitrarily long right-special factors ending with 0 and arbitrarily long right-special factors ending with 2. Let u and v be right-special factors of length n ending respectively with 0 and 2. Such factors exist when $n = 1$. The word u has right extensions $u1$ and $u2$, hence $f(u)f(1)$ and $f(u)f(2)0$ are also factors of **x**. As $f(u)f(1) = f(u)f(2)2$, we see that $f(u)f(2)$ is right-special, ends in 2, and has length greater than n. Similarly, v has right extensions $v0$ and $v2$, and we find that $f(v)0$ is right-special, ends in 0, and has length greater than n. This proves the claim.

It suffices to show that **x** never has three right-special factors of the same length. Suppose, towards a contradiction, that **x** has three right-special factors of the same length. Let n be the least integer such that **x** has three right-special factors of length n. By enumerating all right-special factors of **x** of length at most 3, we see that $n \geqslant 4$. Now each right-special factor of length n contains a right-special factor of length $n - 1$ as a suffix. By the minimality of n, there are only two right-special factors of length $n - 1$. Thus, two distinct right-special factors of length n have a common suffix u of length $n - 1$, i.e., they have the form au and bu, where $a, b \in \Sigma_3$ and $|u| = n - 1$. So u is bispecial, meaning that u does not begin (or end) in 1 or 22. Hence, by deleting at most one letter at the beginning of u (if 20 is a prefix of u) and at most two letters at the end of u (if 02 is a suffix of u), we obtain a factor $u' = f(v)$ of u for some v. Since $|u| \geqslant 4$, we have $1 \leqslant |v| < |u|$. Further, we see that v has four distinct bi-extensions in $\mathrm{Fact}(\mathbf{x})$. Let cv and dv be the left extensions of v in $\mathrm{Fact}(\mathbf{x})$. Then cv and dv are right-special and end in the same letter. Thus **x** has at least three distinct right-special factors of length $|cv| < |au| = n$, and this contradicts the minimality of n. $\square$

Proposition 4.5. *For all $n \geqslant 2$, the word $\mathbf{y} = g(f^\omega(0))$ has exactly two right-special factors u and v of length n. The word u ends with 02 and $u1, u2 \in \mathrm{Fact}(\mathbf{y})$, and the word v ends with 22 and $v0, v2 \in \mathrm{Fact}(\mathbf{y})$.*

Proof. By considering the images under g of the right extensions of the right-special factors of **x** described in Proposition 4.4, we see that the claimed words u and v exist for all $n \geqslant 2$. Thus it suffices to show that **y** has at most two right-special factors of length n for all $n \geqslant 2$. The claim is easily checked for $n = 2$.

Assume for a contradiction that there exists a least integer $n \geqslant 3$ such that there are three right-special factors of length n. As in the proof of Proposition 4.4, there exists a right-special factor u of length $n - 1$ such that au and bu are right-special for distinct letters a and b. Since 2 is the unique left-special letter and the unique right-special letter, we see that u begins and ends in 2. Therefore $u = g(v)2$ for a factor v of **x**. Since au and bu both have two right extensions in **y**, we deduce that **x** has two right-special factors of length $|v| + 1$ that end in the same letter. This contradicts Proposition 4.4. $\square$

Proposition 4.6. *For all $n \geqslant 4$, the word $\mathbf{z} = \tau(g(f^\omega(0)))$ has 4 right-special factors of length n, each with exactly two right extensions.*

Proof. This proof is essentially the same as for **x** and **y**, but the presence of τ complicates the analysis slightly. It is straightforward to verify the statement for $n \leqslant 9$.

Let s be a right-special factor of $\mathbf{z}$ such that $|s| \geqslant 10$. First note that s contains at least one of the letters 1 and 2. We assume that the letter 1 occurs closer to the end of the word s than the letter 2, i.e., that s has suffix 1, 10, or 100. (A symmetric argument applies if s has suffix 2, 20, or 200.) We first argue that s has suffix 10 or 100. Suppose otherwise that s has suffix 1. Then we must have $s0, s1 \in \text{Fact}(\mathbf{z})$, since $12 \notin \text{Fact}(\mathbf{z})$. The suffix 11 of $s1$ forces s to have suffix 00101, but then $s0$ has suffix 001010. Since $001010 \notin \text{Fact}(\mathbf{z})$, this is a contradiction. So s has suffix 10 or 100. If s has suffix 10, then we have $s0, s1 \in \text{Fact}(\mathbf{z})$ and $s2 \notin \text{Fact}(\mathbf{z})$, since $102 \notin \text{Fact}(\mathbf{z})$. If s has suffix 100, then we have $s1, s2 \in \text{Fact}(\mathbf{z})$ and $s0 \notin \text{Fact}(\mathbf{z})$, since $000 \notin \text{Fact}(\mathbf{z})$. In particular, we see that every right-special factor of $\mathbf{z}$ of length at least 10 has exactly two right extensions.

Now it suffices to show that $\mathbf{z}$ contains a unique right-special factor of length n with suffix 10 and a unique right-special factor of length n with suffix 100 for each $n \geqslant 10$, as Lemma 4.3 guarantees that their sisters will also be right-special factors of $\mathbf{z}$. To see that at least one right-special of length n of each type exists, we consider the images under τ (or $\overline{\tau}$) of the right extensions of the right-special factors of $\mathbf{y}$ described in Proposition 4.5.

Suppose for a contradiction that there is a least integer $n \geqslant 10$ such that there exist two right-special factors of length n with suffix 10. By the minimality of n, they have the form au and bu for distinct letters a, b and a right-special factor u of length $n-1$, and their right extensions are $au0$, $au1$, $bu0$, and $bu1$. Since $1100, 10101 \notin \text{Fact}(\mathbf{z})$, it must be that u has suffix 0010. Further, u cannot have suffix 10010 as this factor is not right-special. Therefore u has suffix 20010. Since $\text{Fact}(\mathbf{z})$ is closed under reversal and u is left-special, we see from our work above that u has prefix 0.

First assume that $a = 0$. Then $b \in \{1, 2\}$, and we let c be the other letter in $\{1, 2\}$. Since $a00, b0c \notin \text{Fact}(\mathbf{z})$, it must be that u has prefix $0b$. Further, as $00bb, b0b0b \notin \text{Fact}(\mathbf{z})$, we see that u must begin with $0b00c$. It follows that $u = 0b00cu'20010$ for some word u'. Now $00cu'2$ must equal $\tau(v)$ (if $c = 1$) or $\overline{\tau}(v)$ (if $c = 2$) for a nonempty factor v of $\mathbf{y}$. Further, since $au0 \in \text{Fact}(\mathbf{z})$, we must have $0v0 \in \text{Fact}(\mathbf{y})$. Since $au1$ is also in $\text{Fact}(\mathbf{z})$, we must have $0v1$ or $0v2$ in $\text{Fact}(\mathbf{y})$ as well. So $0v$ is right-special in $\mathbf{y}$, and from Proposition 4.5, it must be $0v2 \in \text{Fact}(\mathbf{y})$. By considering the right extensions $bu0$ and $bu1$ of bu, and using the fact that $\text{Fact}(\mathbf{z})$ is closed under reversal, we see that $2v0$ and $2v2$ must also be factors of $\mathbf{y}$. So $0v$ and $2v$ are distinct right-special factors of $\mathbf{y}$, both with right extensions by 0 and 2 belonging to $\text{Fact}(\mathbf{y})$, and this contradicts Proposition 4.5.

We may now assume that $a, b \neq 0$. In this case, we see that $u = u'0010$ with $u' = \tau(1v)$ or $u' = \overline{\tau}(1v)$ for a factor $1v$ of $\mathbf{y}$. By an argument similar to the one in the previous paragraph, one can show that $1v$ and $2v$ are right-special in $\mathbf{y}$, and derive a contradiction with Proposition 4.5. Thus we have shown that $\mathbf{z}$ contains a unique right-special factor of length n with suffix 10 for all $n \geqslant 10$.

The proof that $\mathbf{z}$ contains a unique right-special factor of length n with suffix 100 for all $n \geqslant 10$ is similar, and is omitted. $\square$

Note that it now follows easily from Proposition 4.4, Proposition 4.5, and Proposition 4.6, along with a determination of initial values by inspection, that $C_{\mathbf{x}}(n) = C_{\mathbf{y}}(n) = 2n + 1$ for all $n \geqslant 0$, and that $C_{\mathbf{z}}(n) = 4n + 2$ for all $n \geqslant 4$, with $C_{\mathbf{z}}(0) = 1$, $C_{\mathbf{z}}(1) = 3$, $C_{\mathbf{z}}(2) = 7$, and $C_{\mathbf{z}}(3) = 12$.

We remark that [3, Corollary 1.4] states that an infinite word with factor complexity $2n+1$ for all n is rich provided that its language is closed under reversal. Since $\mathbf{x}$ and $\mathbf{y}$ are uniformly recurrent and contain arbitrarily long palindromes (see the results of Section 4.2), their languages are closed under reversal, and we immediately obtain the following.

Corollary 4.7. *The words* $\mathbf{x}$ *and* $\mathbf{y}$ *are rich.*

We note, however, that we cannot deduce that $\mathbf{z}$ is rich directly from its factor complexity. So we proceed with the determination of the palindromic complexity of $\mathbf{z}$ in the next subsection.

4.2 Palindromic Complexity

In order to determine the palindromic complexity of $\mathbf{z}$, we first show that all palindromic factors in $\mathbf{x}$ and $\mathbf{y}$ have unique palindromic extensions. Note that this property does not hold for all words of factor complexity $2n+1$ (c.f. [3, Section 4]).

Proposition 4.8. *Every palindromic factor of* $\mathbf{x}$ *has a unique palindromic extension in* $\mathrm{Fact}(\mathbf{x})$.

Proof. We proceed by induction on the length of the palindromic factor. First, we check that every palindromic factor of $\mathbf{x}$ of length at most 2 has a unique palindromic extension.

Now suppose for some $n \geqslant 3$ that every palindromic factor of $\mathbf{x}$ of length less than n has a unique palindromic extension, and let w be a palindromic factor of $\mathbf{x}$ of length n. Since w is a palindrome of length at least 3, we have $w = aw'a$ for some letter $a \in \Sigma_3$ and word $w' \in \Sigma_3^*$ with $|w'| \geqslant 1$.

Case 1: Suppose that $w = 0w'0$. Then $w = f(v)0$ for some word $v \in \mathrm{Fact}(\mathbf{x})$ of length less than w. By Lemma 3.12, the word v is a palindrome, and since $|v| < |w|$, the word v has a unique palindromic extension in $\mathrm{Fact}(\mathbf{x})$, say ava, where $a \in \Sigma_3$. Then $\mathbf{x}$ contains the factor $f(ava)$, which contains a palindromic extension of w. Now suppose (towards a contradiction) that there are two distinct palindromic extensions of w in $\mathrm{Fact}(\mathbf{x})$. Since $00 \notin \mathrm{Fact}(\mathbf{x})$, we must have $1w1 = 10w'01$ and $2w2 = 20w'02$ in $\mathrm{Fact}(\mathbf{x})$. Note that the only possible preimage of $1w1$ is $0v0$, while the possible preimages of $2w2$ are $1v1$, $1v2$, $2v1$, and $2v2$. So $\mathrm{Fact}(\mathbf{x})$ must contain $0v0$ and at least one word from the set $\{1v1, 1v2, 2v1, 2v2\}$. Since v has a unique palindromic extension in $\mathrm{Fact}(\mathbf{x})$ by the inductive hypothesis, we see that $\mathrm{Fact}(\mathbf{x})$ must contain $0v0$ and either $1v2$ or $2v1$. Since $0v0 \in \mathrm{Fact}(\mathbf{x})$ and $00 \notin \mathrm{Fact}(\mathbf{x})$, we see that the first and last letter of v is either a 1 or a 2. But this is impossible, since $1v2$ or $2v1$ is also in $\mathrm{Fact}(\mathbf{x})$, and no word from $\{11, 12, 21\}$ belongs to $\mathrm{Fact}(\mathbf{x})$.

Case 2: Suppose that $w = 1w'1$. Since every occurrence of 1 in $\mathbf{x}$ is preceded and followed by 0, we see that $0w0$ is the unique palindromic extension of w in $\mathrm{Fact}(\mathbf{x})$.

Case 3: Suppose that $w = 2w'2$. Since $21 \notin \mathrm{Fact}(\mathbf{x})$, we see that w' must begin in 0 or 2. If w' begins in 2, then we must have $|w'| \geqslant 2$, since $222 \notin \mathrm{Fact}(\mathbf{x})$. So we can write $w = 22w''22$ for some $w'' \in \Sigma_3^*$, and since every occurrence of 22 in $\mathbf{x}$ is preceded

and followed by 0, we see that the unique palindromic extension of w in $\text{Fact}(\mathbf{x})$ is $0w0$. So we may assume that w' begins in 0. Hence we can write $w' = f(v)0$ for some word $v \in \Sigma_3^*$, and by Lemma 3.12, the word v is a palindrome. Since $|v| < |w'| < |w|$, the words v and w' have a unique palindromic extension in $\text{Fact}(\mathbf{x})$. So $w = 2w'2$ must be the unique palindromic extension of w' in $\text{Fact}(\mathbf{x})$. It follows that $0v0 \notin \text{Fact}(\mathbf{x})$, for $f(0v0)$ contains the palindromic extension $1w'1$ of w'. Now, if the unique palindromic extension of v in $\text{Fact}(\mathbf{x})$ is $1v1$, then $\mathbf{x}$ contains the factor $f(1v1)$, which contains the palindromic extension $2w2$. If the unique palindromic extension of v in $\text{Fact}(\mathbf{x})$ is $2v2$, then $\mathbf{x}$ contains the palindromic extension $f(2v2)0 = 0w0$ of w. Finally, if both $0w0$ and $2w2$ are in $\text{Fact}(\mathbf{x})$, then both $2v2$ and $1v1$ would be in $\text{Fact}(\mathbf{x})$, a contradiction.

By mathematical induction, we conclude that every palindromic factor of $\mathbf{x}$ has a unique palindromic extension. $\qquad\square$

Proposition 4.9. *Every palindromic factor of* $\mathbf{y}$ *has a unique palindromic extension in* $\text{Fact}(\mathbf{y})$.

Proof. Let w be a palindromic factor of $\mathbf{y}$. First note that if $w = \varepsilon$, then the unique palindromic extension of w is 22. Next note that every occurrence of 0 or 1 in $\mathbf{y}$ is preceded and followed by 2. So if w begins in 0 or 1, then the unique palindromic extension of w in $\text{Fact}(\mathbf{y})$ is $2w2$.

So we may assume that w begins in 2. By inspecting all factors of $\mathbf{y}$ of length 3, we see that the unique palindromic extension of 2 is 222. So we may assume that $|w| \geqslant 2$, which means that we can write $w = 2w'2$ for some word $w' \in \Sigma_3^*$. But then we have $w = g(v)2$ for some nonempty word $v \in \text{Fact}(\mathbf{x})$, and by Lemma 3.12, the word v is a palindrome. By Proposition 4.8, the word v has a unique palindromic extension in $\text{Fact}(\mathbf{x})$. Observe that for all $a \in \Sigma_3^*$, we have $ava \in \text{Fact}(\mathbf{x})$ if and only if $g(ava)2 \in \text{Fact}(\mathbf{y})$, and that $g(ava)2$ contains the palindromic extension awa of w. So if ava is the unique palindromic extension of v in $\text{Fact}(\mathbf{x})$, then awa is the unique palindromic extension of w in $\text{Fact}(\mathbf{y})$. $\qquad\square$

Lemma 4.10. *Let* w *be a palindromic factor of* $\mathbf{y}$. *If* $0w0 \in \text{Fact}(\mathbf{y})$, *then no word from the set* $\{1w1, 1w2, 2w1, 2w2\}$ *is in* $\text{Fact}(\mathbf{y})$.

Proof. Suppose that $0w0 \in \text{Fact}(\mathbf{y})$. By Proposition 4.9, the word w has a unique palindromic extension in $\text{Fact}(\mathbf{y})$, so we see that neither $1w1$ nor $2w2$ belongs to $\text{Fact}(\mathbf{y})$. It remains to show that $1w2$ and $2w1$ do not belong to $\text{Fact}(\mathbf{y})$.

First observe that since $0w0 \in \text{Fact}(y)$, we have $20w02 \in \text{Fact}(\mathbf{y})$, and we can write $20w02 = g(v)2$ for some word $v \in \text{Fact}(\mathbf{x})$ of the form $0v'0$. (Note in particular that $w = g(v')2$.) By Lemma 3.12, we see that v and v' are palindromes. Since $00 \notin \text{Fact}(\mathbf{x})$, the word v' begins with 1 or 2.

Now suppose, towards a contradiction, that $1w2 \in \text{Fact}(\mathbf{y})$. Then we have $21w2 \in \text{Fact}(\mathbf{y})$, which means that $g^{-1}(21w) = 1v' \in \text{Fact}(\mathbf{x})$. Since neither 11 nor 12 belongs to $\text{Fact}(\mathbf{x})$, we see that v' begins with 0. But this contradicts the fact that v' begins with 1 or 2. So we conclude that $1w2 \notin \text{Fact}(\mathbf{y})$.

Finally, suppose that $2w1 \in \text{Fact}(\mathbf{y})$. Then we have $g^{-1}(w1) = v'1 \in \text{Fact}(\mathbf{y})$. Since neither 11 nor 21 belongs to $\text{Fact}(\mathbf{x})$, we see that v' ends with 0. But this contradicts

the fact that v' is a palindrome that begins with 1 or 2. So we conclude that $2w1 \notin \mathrm{Fact}(\mathbf{y})$. $\qquad\square$

Proposition 4.11. $P_{\mathbf{z}}(n) = \begin{cases} 1, & \text{if } n = 0; \\ 3, & \text{if } n = 1 \text{ or } n = 2; \\ 4, & \text{if } n = 3, \text{ or } n \geqslant 4 \text{ and } n \text{ is even}; \\ 2, & \text{if } n \geqslant 5 \text{ and } n \text{ is odd}. \end{cases}$

Proof. By generating all factors of $\mathbf{z}$ of length at most 20, we verify the formula given in the theorem statement for all $n \leqslant 20$. Now it suffices to show that every palindromic factor of $\mathbf{z}$ of length at least 19 has a unique palindromic extension in $\mathrm{Fact}(\mathbf{z})$.

Let w be a palindromic factor of $\mathbf{z}$ of length at least 19, so that w contains both 1 and 2. We assume that 1 appears before 2 in w; a symmetric argument applies if 2 appears before 1. Let $u00$ be the longest prefix of w that does not contain the letter 2. Then u must be a nonempty suffix of $\tau(0)$, $\tau(1)$, or $\tau(2)$, and we can write

$$w = u00w'00\tilde{u},$$

where w' begins and ends in 2 and $\tilde{u}$ is the reversal of u. So we see that

$$w = u\overline{\tau}(y)00\tilde{u}$$

for some word $y \in \mathrm{Fact}(\mathbf{y})$. It is not hard to see that y must be a palindrome of odd length.

Case 1: Suppose that $u = 01$. First note that $2w2 \notin \mathrm{Fact}(\mathbf{z})$, since $201 \notin \mathrm{Fact}(\mathbf{z})$. Note further that $0w0 \in \mathrm{Fact}(\mathbf{z})$ if and only if $0y0 \in \mathrm{Fact}(\mathbf{y})$, and that $1w1 \in \mathrm{Fact}(\mathbf{z})$ if and only if $ayb \in \mathrm{Fact}(\mathbf{y})$ for some $a, b \in \{1, 2\}$.

Suppose first that $0y0 \in \mathrm{Fact}(\mathbf{y})$. Then we see from Lemma 4.10 that the words $1y1$, $1y2$, $2y1$, and $2y2$ do not belong to $\mathrm{Fact}(\mathbf{y})$. Hence, from the biconditional statements above, we conclude that $0w0 \in \mathrm{Fact}(\mathbf{z})$ and $1w1 \notin \mathrm{Fact}(\mathbf{z})$, i.e., that $0w0$ is the unique palindromic extension of w in $\mathrm{Fact}(\mathbf{z})$.

Suppose otherwise that $0y0 \notin \mathrm{Fact}(\mathbf{y})$. Then either $1y1$ or $2y2$ belongs to $\mathrm{Fact}(\mathbf{y})$, since y has a unique palindromic extension in $\mathrm{Fact}(\mathbf{y})$ by Proposition 4.9. Hence, from the biconditional statements above, we conclude that $1w1$ is the unique palindromic extension of w in $\mathrm{Fact}(\mathbf{z})$.

Case 2: Suppose that $u = 00101101$. First note that $0w0 \notin \mathrm{Fact}(\mathbf{z})$, since $000 \notin \mathrm{Fact}(\mathbf{z})$. Note further that $1w1 \in \mathrm{Fact}(\mathbf{z})$ if and only if $2y2 \in \mathrm{Fact}(\mathbf{y})$, and that $2w2 \in \mathrm{Fact}(\mathbf{z})$ if and only if $1y1 \in \mathrm{Fact}(\mathbf{y})$.

From the structure of w, we see that $\mathrm{Fact}(\mathbf{y})$ contains a factor of the form ayb, where $a, b \in \{1, 2\}$. So by Lemma 4.10, we have $0y0 \notin \mathrm{Fact}(\mathbf{y})$. But y has a unique palindromic extension in $\mathrm{Fact}(\mathbf{y})$ by Proposition 4.9, so exactly one of the words $1y1$ and $2y2$ belongs to $\mathrm{Fact}(\mathbf{y})$. By the biconditionals above, we conclude that w has a unique palindromic extension in $\mathrm{Fact}(\mathbf{z})$.

Case 3: Say $u \notin \{\texttt{01}, \texttt{00101101}\}$. It is not hard to show by inspection of $\tau(0)$, $\tau(1)$, and $\tau(2)$ that w has a unique palindromic extension. For example, if $u = \texttt{1}$, then the unique palindromic extension of w in Fact($\mathbf{z}$) is $\texttt{0}w\texttt{0}$, since $\tau(0)$, $\tau(1)$, and $\tau(2)$ have common suffix $\texttt{01}$. If $u = \texttt{001}$, then $w = \tau(\texttt{0}y\texttt{0})\texttt{00}$, and the unique palindromic extension of w in Fact($\mathbf{z}$) is $\texttt{2}w\texttt{2}$. $\qquad\square$

We now have the means to prove Proposition 4.1.

Proof of Proposition 4.1. First of all, when $n \leqslant 3$, it is straightforward to verify that the equation (1) holds. Let $n \geqslant 4$. Since each right-special factor of $\mathbf{z}$ of length n has exactly two right extensions by Proposition 4.6, the quantity $C_{\mathbf{z}}(n+1) - C_{\mathbf{z}}(n)$ equals the number of right-special factors of $\mathbf{z}$ of length n. Hence the right side of (1) equals 6. Proposition 4.11 implies that the left side of the equation also equals 6. Therefore $\mathbf{z}$ is rich by Proposition 4.2. $\qquad\square$

5 The Critical Exponent

Throughout this section, let $\mathbf{x} = f^{\omega}(0)$, $\mathbf{y} = g(\mathbf{x})$, and $\mathbf{z} = \tau(\mathbf{y})$. We will write $\mathbf{x} = x_0 x_1 x_2 \cdots$, $\mathbf{y} = y_0 y_1 y_2 \cdots$, and $\mathbf{z} = z_0 z_1 z_2 \cdots$, where the x_i's, y_i's, and z_i's are letters.

Theorem 5.1. *The critical exponent of $\mathbf{z}$ is*

$$1 + \frac{1}{3 - \mu_1} \approx 2.25876324,$$

where $\mu_1 \approx 2.20557$ is the unique real root of the polynomial $x^3 - 2x^2 - 1$.

We prove Theorem 5.1 by adapting the method of Krieger [31] for finding the critical exponent of a fixed point of a non-erasing morphism. The basic idea is that every factor of exponent greater than $9/4$ in $\mathbf{z}$ belongs to one of only finitely many sequences of "unstretchable" repetitions, which cannot be extended periodically (i.e., "stretched") in $\mathbf{z}$. Each of these sequences is obtained as follows:

- We start from a short unstretchable repetition w_0 in $\mathbf{x}$ and repeatedly apply f and stretch to the left and right as far as possible. This gives us a sequence $w_0, w_1, w_2, \ldots$ of unstretchable repetitions in $\mathbf{x}$, which we call an "inner stretch sequence".

- We apply g and τ to each word in the sequence $w_0, w_1, w_2 \ldots$, and stretch to the left and right as far as possible. This gives us a sequence $W_0, W_1, W_2, \ldots$ of unstretchable repetitions in $\mathbf{z}$, which we call an "outer stretch sequence".

Once we identify the outer stretch sequences that contain repetitions of exponent greater than $9/4$ in $\mathbf{z}$, we analyze the exponents of the words in these outer stretch sequences, and show that their supremum is $1 + 1/(3 - \mu_1)$.

In Section 5.1, we formally define unstretchable repetitions and stretch sequences. In Section 5.2, we show that every unstretchable repetition of exponent greater than $9/4$ in $\mathbf{z}$ belongs to one of only finitely many outer stretch sequences. Finally, in Section 5.3, we show that the supremum of the exponents of the words in these outer stretch sequences is $1 + 1/(3 - \mu_1)$.

5.1 Unstretchable Repetitions and Stretch Sequences

In order to define an unstretchable repetition, we need to consider a specific occurrence of the repetition. An *occurrence* of a factor w in an infinite word $\mathbf{u} = u_0 u_1 u_2 \cdots$ is a triple (w, i, j) such that $w = u_i \cdots u_j$. We usually omit the reference to the triple, and simply refer to an occurrence (w, i, j) as $w = u_i \cdots u_j$. The set of all occurrences of all factors of $\mathbf{u}$ is denoted by $\mathrm{Occ}(\mathbf{u})$. Note that every occurrence of a word w in $\mathbf{x}$ corresponds in a natural way to an occurrence of $f(w)$ in $\mathbf{x}$. More precisely, if $w = x_i \cdots x_j$, then $f(w) = x_k \cdots x_\ell$, where $k = |f(x_0 \cdots x_{i-1})|$ and $\ell = |f(x_0 \cdots x_j)| - 1$. Similarly, every occurrence of a word w in $\mathbf{x}$ corresponds to an occurrence of $g(w)$ in $\mathbf{y}$, and every occurrence of a word w in $\mathbf{y}$ corresponds to an occurrence of either $\tau(w)$ or $\overline{\tau}(w)$ in $\mathbf{z}$. If $w = y_i \cdots y_j \in \mathrm{Occ}(\mathbf{y})$, then we write

$$\tilde{\tau}(w) = z_k \cdots z_\ell,$$

where $k = |\tau(y_0 \cdots y_{i-1})|$ and $\ell = |\tau(y_0 \cdots y_j)| - 1$, i.e.,

$$\tilde{\tau}(w) = \begin{cases} \tau(w), & \text{if } i \text{ is even;} \\ \overline{\tau}(w), & \text{if } i \text{ is odd.} \end{cases}$$

Suppose that $w = u_i \cdots u_j \in \mathrm{Occ}(\mathbf{u})$ has period q. We say that the pair (w, q) is *left-stretchable* if $u_{i-1} \cdots u_j$ also has period q, i.e., if $u_{i-1} = u_{i+q-1}$. Similarly, we say that (w, q) is *right-stretchable* if $u_i \cdots u_{j+1}$ has period q. The *left stretch* of (w, q) is the longest word λ such that λw has period q and $\lambda w = u_{i-|\lambda|} \cdots u_j$. The *right stretch* of (w, q) is the longest word ρ such that $w\rho$ has period q and $w\rho = u_i \cdots u_{j+|\rho|}$. (Such a word ρ exists as long as $\mathbf{u}$ has finite critical exponent.) We say that (w, q) is *unstretchable* if it is neither left-stretchable nor right-stretchable, i.e., if $\lambda = \rho = \varepsilon$. Sometimes we say that "w is unstretchable with respect to q" instead of "(w, q) is unstretchable".

Suppose that $w = u^r = x_i \cdots x_j \in \mathrm{Occ}(\mathbf{x})$ is unstretchable with respect to $q = |u|$. Note that the word $f(w)$ has period $|f(u)|$, the word $f^2(w)$ has period $|f^2(u)|$, and so on. Further, each time we apply f, we can stretch the resulting repetition as far as possible to the left and the right in $\mathbf{x}$, to obtain a sequence of unstretchable repetitions $w_0, w_1, w_2, \ldots$ with periods $|u|, |f(u)|, |f^2(u)|, \ldots$, respectively. Formally, the *inner stretch sequence* of (w, q) is the sequence

$$(w_0, q_0), (w_1, q_1), (w_2, q_2), \ldots$$

where $q_n = |f^n(u)|$ for all $n \geqslant 0$, and w_n is defined recursively by $(w_0, q_0) = (w, q)$ and

$$w_n = \lambda_n f(w_{n-1}) \rho_n$$

for all $n \geqslant 1$, where λ_n (resp. ρ_n) is the left (resp. right) stretch of $(f(w_{n-1}), q_n)$ in $\mathbf{x}$. Thus we have

$$\begin{aligned}
w_0 &= w, \\
w_1 &= \lambda_1 f(w) \rho_1, \\
w_2 &= \lambda_2 f(\lambda_1 f(w) \rho_1) \rho_2, \\
w_3 &= \lambda_3 f(\lambda_2 f(\lambda_1 f(w) \rho_1) \rho_2) \rho_3,
\end{aligned}$$

and so on. Since w_n has period q_n, we can write $w_n = v_n^{r_n}$, where v_n is the prefix of w_n of length q_n and

$$r_n = \frac{|w_n|}{q_n} = \frac{|f^n(w)| + \sum_{k=1}^{n} |f^{n-k}(\lambda_k \rho_k)|}{|f^n(u)|}$$

is an exponent of w_n. We call $r_0, r_1, r_2, \ldots$ the *inner power sequence* of (w, q).

Now suppose that u has an even number of 2's. Then it is not hard to show that for all $n \geqslant 0$, the word v_n has an even number of 2's. It follows that $g(w_n)$ has even period $|g(v_n)| = |g(f^n(u))|$ and in turn that $\tau(g(w_n))$ has period $|\tau(g(v_n))| = |\tau(g(f^n(u)))|$. For all $n \geqslant 0$, define $Q_n = |\tau(g(f^n(u)))|$ and

$$W_n = \lambda_n^* \tilde{\tau}(g(w_n)) \rho_n^*$$

where λ_n^* (resp. ρ_n^*) is the left (resp. right) stretch of $(\tilde{\tau}(g(w_n)), Q_n)$ in $\mathbf{z}$. The sequence

$$(W_0, Q_0), (W_1, Q_1), (W_2, Q_2), \ldots$$

is called the *outer stretch sequence* of (w, q). Since W_n has period Q_n, we can write $W_n = V_n^{R_n}$, where V_n is the prefix of W_n of length Q_n and

$$R_n = \frac{|W_n|}{Q_n} = \frac{|\tau(g(f^n(w)))| + \sum_{k=1}^{n} |\tau(g(f^{n-k}(\lambda_k \rho_k)))| + |\lambda_n^* \rho_n^*|}{|\tau(g(f^n(u)))|}$$

is an exponent of W_n. We call $R_0, R_1, R_2, \ldots$ the *outer power sequence* of (w, q).

Example 5.2. Consider the occurrence $w = x_2 = 0$ in $\mathbf{x}$. Note that $(w, 1)$ is unstretchable, since $x_1, x_3 \neq 0$. Let us find the first few terms of the inner stretch sequence of $(w, 1)$. By definition, we have $w_0 = w$ and $q_0 = 1$. Next, we have $q_1 = |f(0)| = 2$ and $f(w_0) = x_5 x_6 = 01$. Since $x_4 = 2$ and $x_7 x_8 = 02$, we find $\lambda_1 = \varepsilon$ and $\rho_1 = x_7 = 0$, so that

$$w_1 = \lambda_1 \cdot f(w_0) \cdot \rho_1 = x_5 x_6 x_7 = 010.$$

Continuing in this manner, we find $q_2 = |f^2(0)| = 5$ and

$$w_2 = \lambda_2 \cdot f(w_1) \cdot \rho_2 = 2 \cdot 0102201 \cdot 02 = (20102)^2,$$

and $q_3 = |f^3(0)| = 11$ and

$$w_3 = \lambda_3 \cdot f(w_2) \cdot \rho_3 = \varepsilon \cdot (02010220102)^2 \cdot 0 = (02010220102)^{23/11}.$$

Thus, the inner stretch sequence of $(w, 1)$ begins with

$$\begin{aligned}
(w_0, q_0) &= (0, 1), \\
(w_1, q_1) &= \left((01)^{3/2}, 2\right), \\
(w_2, q_2) &= \left((20102)^2, 5\right), \text{ and} \\
(w_3, q_3) &= \left((02010220102)^{23/11}, 11\right),
\end{aligned}$$

and the inner power sequence of $(w, 1)$ begins with

$$1, \tfrac{3}{2}, 2, \tfrac{23}{11}.$$

Now let us find the initial term of the outer stretch sequence of $(w, 1)$. We have $Q_0 = |\tau(g(0))| = 19$ and

$$\begin{aligned}
W_0 &= \lambda_0^* \cdot \tau(g(0)) \cdot \rho_0^* \\
&= 02 \cdot \tau(20) \cdot \tau(2)0010 \\
&= (0200101101001011010)^{41/19}.
\end{aligned}$$

Thus, the initial term in the outer stretch sequence of $(w, 1)$ is

$$(W_0, Q_0) = \left((0200101101001011010)^{41/19}, 19\right).$$

Continuing in this manner, one can show that the outer power sequence of $(w, 1)$ begins with

$$\tfrac{41}{19}, \tfrac{94}{43}, \tfrac{210}{94}, \tfrac{467}{207}, \cdots$$

In fact, we will see that this is the sequence that is responsible for the critical exponent of $\mathbf{z}$. Note that $467/207 \approx 2.25604$ is just barely larger than $9/4$.

5.2 Repetitions of Exponent Greater than 9/4 in z

In this subsection, we show that every unstretchable repetition of exponent greater than $9/4$ in $\mathbf{z}$ belongs to one of only a narrow family of outer stretch sequences. More precisely, we prove the following proposition.

Proposition 5.3. *Suppose that $W = U^R = z_k \cdots z_\ell \in \mathrm{Occ}(\mathbf{z})$ is unstretchable with respect to $Q = |U|$, where $R > 9/4$. Then (W, Q) is in the outer stretch sequence of $(w, |w|)$ for some occurrence $w = x_i \cdots x_j \in \mathrm{Occ}(\mathbf{x})$, where the word w belongs to the set*

$$S = \{0, 1, 22, 202, 1022, 0220, 2201, 10202, 02020, 20201\}.$$

First, we show that unstretchable repetitions of exponent greater than $9/4$ in $\mathbf{z}$ come from unstretchable repetitions of exponent greater than 2 in $\mathbf{y}$ with periods of even length by applying τ (or $\overline{\tau}$) and stretching to the left and right.

Lemma 5.4. *Suppose that $W = U^R = z_k \cdots z_\ell \in \mathrm{Occ}(\mathbf{z})$ is unstretchable with respect to $Q = |U|$, where $R > 9/4$. Then we have*

$$W = \lambda \tau(w) \rho \quad \text{or} \quad W = \lambda \overline{\tau}(w) \rho$$

for some words $\lambda, \rho \in \mathrm{Occ}(\mathbf{z})$ and a word $w = y_i \cdots y_j \in \mathrm{Occ}(\mathbf{y})$ such that

- $w = u^r$ *has period $q = |u|$, where u is a prefix of w satisfying $|\tau(u)| = |U|$;*

- (w, q) *is unstretchable;*

- *u has even length; and*

- $r > 2$.

Proof. First of all, we check that $\mathbf{z}$ has no factor of the form U^R where $R > 9/4$ and $|U| \leqslant 136$. So we may assume that $|U| > 136$. Then U contains both 1 and 2. By Lemma 4.3, we may assume that 2 appears before 1 in U. Let $\lambda 00$ be the longest prefix of W that does not contain the letter 1, and let ρ be the longest suffix of W that does not contain both 1 and 2. Notice that our assumptions imply that λ exists and ends with 2. Then λ must be a suffix of $\overline{\tau}(a)$ for some $a \in \{0, 1, 2\}$, and ρ must be a prefix of $\tau(b)00$ (or its sister) for some $b \in \{0, 1, 2\}$, and we can write

$$W = \lambda \tau(w) \rho$$

for some word $w = y_i \cdots y_j \in \mathrm{Occ}(\mathbf{y})$. Since $W = U^R$ where $R > 9/4$ and $|U| \geqslant 136$, we have

$$|W| > \tfrac{9}{4}|U| = 2|U| + \tfrac{1}{4}|U| \geqslant 2|U| + 34.$$

Further, since $W = \lambda \tau(w) \rho$ and $|\lambda| \leqslant 16$ and $|\rho| \leqslant 18$, we have

$$|\tau(w)| = |W| - |\lambda| - |\rho| > 2|U|.$$

Write $\tau(w) = U'V'$, where U' is the prefix of $\tau(w)$ of length $|U|$. Since $|\tau(w)| > 2|U|$, we have $|V'| > |U'|$. Since W has period $|U|$ and λ ends in 2, we see that U' ends in 2. Similarly, since U' starts in 001, we see that V' starts in 001. It follows that we can write $U' = \tau(u)$ and $V' = \tau(v)$ for some words $u, v \in \mathrm{Fact}(\mathbf{y})$. Since τ is injective, we must have $w = uv$, and we see that w has period $q = |u|$. Since $|V'| > |U'|$, we must have $|v| > |u|$, which means that $w = u^r$ for some rational number $r > 2$. Since $U' = \tau(u)$ starts in 001 and ends in 2, we see that u must have even length.

It remains to show that (w, q) is unstretchable. Suppose that (w, q) is left-stretchable. (The proof is similar if (w, q) is right-stretchable.) Then $y_{i-1} \cdots y_j$ has period q, and it follows that $1\overline{\tau}(y_{i-1} \cdots y_j)\rho$ has period Q and properly contains $W = z_k \cdots z_\ell$, which contradicts the assumption that (W, Q) is unstretchable. $\square$

Next, we show that unstretchable repetitions in $\mathbf{y}$ with exponent greater than 2 and period of even length come from unstretchable repetitions in $\mathbf{x}$ with exponent at least 2 with an even number of 2's in the repeated prefix.

Lemma 5.5. *Suppose that $W = U^R = y_k \cdots y_\ell \in \mathrm{Occ}(\mathbf{y})$ is unstretchable with respect to $Q = |U|$, where $R > 2$ and U has even length. Then we have*

$$W = g(w)2$$

where $w = x_i \cdots x_j \in \mathrm{Occ}(\mathbf{x})$ satisfies the following:

- *$w = u^r$ has period $q = |u|$, where u is a prefix of w satisfying $|g(u)| = |U|$;*

- *(w, q) is unstretchable;*

- *u has an even number of 2's; and*

- $r \geqslant 2$.

Proof. First note that since (W, Q) is unstretchable, it must begin and end in 2. For if $y_k \in \{0, 1\}$, then $y_k = y_{k+Q}$, and since every occurrence of 0 or 1 in $\mathbf{y}$ is preceded by 2, we have $y_{k-1} = y_{k+Q-1}$, which contradicts the assumption that W is unstretchable. Similarly $y_\ell = 2$. Since $R > 2$, we may write $W = UV2$, where V is a nonempty prefix of W satisfying $|V| \geqslant |U|$. Since W begins in 2, we see that V begins in 2. So we can write $U = g(u)$ and $V = g(v)$ for some words $u, v \in \text{Occ}(\mathbf{x})$. Since U has even length, we see that u must have an even number of 2's. Let $w = uv$, so that $W = UV2 = g(u)g(v)2 = g(w)2$. Now since $g(v)2$ is a prefix of $g(w)$, we see that v is a prefix of w. Further, since $|V| \geqslant |U|$, we must have $|v| \geqslant |u|$. Thus, we have $w = u^r$ for some rational number $r \geqslant 2$.

It remains to show that (w, q) is unstretchable, where $q = |u|$. Write $w = x_i \cdots x_j$, and suppose towards a contradiction that (w, q) is left-stretchable. (The proof is similar if (w, q) is right-stretchable.) Then $x_{i-1} \cdots x_j$ has period q, and it follows that $g(x_{i-1} \cdots x_j)2 = g(x_{i-1})W$ has period Q, which contradicts the assumption that (W, Q) is unstretchable. $\square$

In order to prove Proposition 5.3, it remains to show that every unstretchable repetition in $\mathbf{x}$ with exponent at least 2 and an even number of 2's in the repeated prefix belongs to the inner stretch sequence of $(w, |w|)$ for some word $w \in S$. We start by showing that every sufficiently long unstretchable repetition in $\mathbf{x}$ comes from a shorter unstretchable repetition in $\mathbf{x}$.

Lemma 5.6. *Suppose that $W = x_k \cdots x_\ell \in \text{Occ}(\mathbf{x})$ has period Q and that (W, Q) is unstretchable. Write $W = UV$, where $|U| = Q$, and suppose that $|U| \geqslant 3$ and $|V| \geqslant 3$. Then $W = \lambda f(w)\rho$ for some words $\lambda, w, \rho \in \text{Occ}(\mathbf{x})$ such that*

- $\lambda \in \{\varepsilon, 2\}$ *and* $\rho \in \{0, 02\}$;

- $w = x_i \cdots x_j$ *has period* $q = |u|$, *where u is a prefix of w satisfying $|f(u)| = |U|$; and*

- (w, q) *is unstretchable.*

Proof. We begin by proving a simple claim.

Claim 5.7. *W has prefix 0 or 201 and suffix 0 or 102.*

Proof. We prove only that W has prefix 0 or 201. The proof that W has suffix 0 or 102 uses symmetric arguments.

We begin by showing that 1 is not a prefix of W. Suppose towards a contradiction that W has prefix 1. Then both U and V begin in 1. But the unique left extension of 1 in $\text{Fact}(\mathbf{x})$ is 01, which means that $x_{k-1} = x_{k+q-1} = 0$. This contradicts the assumption that W is unstretchable. So W has prefix 0 or 2.

If W has prefix 0, then we are done, so suppose that W has prefix 2. Then W has prefix 201, 202, or 220, and V has the same prefix of length 3 as W, since $|V| \geqslant 3$.

The unique left extension of 202 in Fact($\mathbf{x}$) is 0202, and the unique left extension of 220 in Fact($\mathbf{x}$) is 0220. So if W has prefix 202 or 220, then $x_{k-1} = x_{k+q-1} = 0$, which contradicts the assumption that W is unstretchable. Thus, we conclude that W must have prefix 201. ■

We now consider two cases, based on the prefix of W.

Case 1: Suppose that W has prefix 0. Then both U and V have prefix 0. Further, since W has suffix 0 or 102, we can write $V = V'\rho$ where $\rho \in \{0, 02\}$. Let $w = x_i \cdots x_j$ be the unique preimage of UV' in $\mathbf{x}$. Write $w = uv$, where $f(u) = U$ and $f(v) = V'$. Let $\lambda = \varepsilon$, so that $W = \lambda f(w)\rho$. Since $f(w)\rho = f(u)f(v)\rho$ has period $|U| = |f(u)|$, we see that $w = uv$ must have period $|u| = q$.

It remains to show that (w, q) is unstretchable. Suppose first that (w, q) is right stretchable. But then $x_i \cdots x_{j+1}$ has period q, and it follows that $f(x_i \cdots x_{j+1})0 = x_k \cdots x_{\ell'}$ has period Q and contains the occurrence $W = x_k \cdots x_\ell$ as a proper prefix. But this contradicts the assumption that (W, Q) is unstretchable. Suppose now that w is left stretchable, i.e., that $x_{i-1} \cdots x_j$ has period q. But then $f(x_{i-1} \cdots x_j)\rho = f(x_{i-1})W = x_{k'} \cdots x_\ell$ has period Q, and contains the occurrence $W = x_k \cdots x_\ell$ as a proper suffix. But this contradicts the assumption that (W, Q) is unstretchable.

Case 2: Suppose that W has prefix 201. Then both U and V have prefix 201, and since W has suffix 0 or 102, we can write $U = 2U'$ and $V = 2V'\rho$, where $\rho \in \{0, 02\}$. So $W = 2U'2V'\rho$, and both U' and V' have prefix 01. Let $w = x_i \cdots x_j$ be the preimage of $U'2V'$ in f. Write $w = uv$, where $f(u) = U'2$ and $f(v) = V'$. Let $\lambda = 2$ so that we have $W = \lambda f(w)\rho$. Since $f(w)\rho = f(u)f(v)\rho$ has period $|U| = |U'2| = |f(u)|$, we see that $w = uv$ must have period $|u| = q$.

It remains to show that (w, q) is unstretchable. Suppose first that (w, q) is right stretchable, i.e., that $x_i \cdots x_{j+1}$ has period q. Then $2f(x_i \cdots x_{j+1})0 = x_k \cdots x_{\ell'}$ has period Q and contains the occurrence $W = x_k \cdots x_\ell$ as a proper prefix. But this contradicts the assumption that (W, Q) is unstretchable. Suppose now that w is left stretchable, i.e., that $x_{i-1} \cdots x_j$ has period q. Then $f(x_{i-1} \cdots x_j)\rho = x_{k'} \cdots x_\ell$ has period Q and contains the occurrence $W = x_k \cdots x_\ell$ as a proper suffix. But this contradicts the assumption that (W, Q) is unstretchable. □

The next lemma describes the left and right stretch of an occurrence $w = x_i \cdots x_j \in$ Occ($\mathbf{x}$) based on the context in which it appears, i.e., based on $x_{i-1} \cdots x_{j+1}$.

Lemma 5.8. *Suppose that $w = x_i \cdots x_j \in$ Occ($\mathbf{x}$) has period q and that (w, q) is unstretchable, where $i \geqslant 1$. Let $Q = |f(x_i \cdots x_{i+q-1})|$ be the corresponding period of $f(w) = x_k \cdots x_\ell$, and let λ (resp. ρ) be the left (resp. right) stretch of $(f(w), Q)$. Then*

$$\lambda = \begin{cases} 2, & \text{if } x_{i-1}, x_{i+q-1} \in \{1, 2\}; \\ \varepsilon, & \text{otherwise;} \end{cases}$$

and

$$\rho = \begin{cases} 02, & \text{if } x_{j+1}, x_{j-q+1} \in \{1, 2\}; \\ 0, & \text{otherwise.} \end{cases}$$

x_{i-1}	w							x_{j+1}
x_{i-1}	x_i	$\cdots$	x_{i+q-1}	$\cdots$	x_{j-q+1}	$\cdots$	x_j	x_{j+1}

$$\underbrace{}_{\text{length } q} \qquad \underbrace{}_{\text{length } q}$$

Figure 8: An occurrence $w = x_i \cdots x_j$ such that (w, q) is unstretchable. (We have illustrated the case that $|w| > 2q$.)

Proof. First note that since (w, q) is unstretchable, we must have $x_{i-1} \neq x_{i+q-1}$ and $x_{j+1} \neq x_{j-q+1}$ (see Figure 8). If $x_{i-1}, x_{i+q-1} \in \{1, 2\}$, then $\{x_{i-1}, x_{i+q-1}\} = \{1, 2\}$, and it follows that λ is the longest common suffix of $f(1)$ and $f(2)$, which is 2. Otherwise, we have $x_{i-1} = 0$ (and $x_{i+q-1} \in \{1, 2\}$) or $x_{i+q-1} = 0$ (and $x_{i-1} \in \{1, 2\}$). In either case, the longest common suffix of $f(x_{i-1})$ and $f(x_{i+q-1})$ is ε, hence $\lambda = \varepsilon$.

The argument for ρ is similar. If $x_{j+1}, x_{j-q+1} \in \{1, 2\}$, then $\{x_{j+1}, x_{j-q+1}\} = \{1, 2\}$, and it follows that $\rho = 02$. Otherwise, we have $x_{j+1} = 0$ (and $x_{j-q+1} \neq 0$) or $x_{j-q+1} = 0$ (and $x_{j+1} \neq 0$), and it follows that $\rho = 0$. $\qquad\square$

We are finally ready to prove that every unstretchable repetition in $\mathbf{x}$ with exponent at least 2 and an even number of 2's in the repeated prefix belongs to the inner stretch sequence of $(w, |w|)$ for some word $w \in S$.

Lemma 5.9. *Suppose that $W = U^R = x_k \cdots x_\ell \in \mathrm{Occ}(\mathbf{x})$ is unstretchable with respect to $Q = |U|$, where $R \geqslant 2$ and U contains an even number of 2's. Then (W, Q) is in the inner stretch sequence of $(w, |w|)$ for some word $w = x_i \cdots x_j \in \mathrm{Occ}(\mathbf{x})$, where w belongs to the set*

$$S = \{0, 1, 22, 202, 1022, 0220, 2201, 10202, 02020, 20201\}.$$

Proof. Let $w = u^r = uv = x_i \cdots x_j \in \mathrm{Occ}(\mathbf{x})$ be a shortest word such that w is unstretchable with respect to $q = |u|$ and (W, Q) is in the inner stretch sequence of (w, q). Note that (W, Q) is in the inner stretch sequence of itself, so there is such a pair (w, q). We claim that $|u| < 3$ or $|v| < 3$. Suppose otherwise that $|u| \geqslant 3$ and $|v| \geqslant 3$. Then by Lemma 5.6, we can write $w = \lambda f(w')\rho$ for some words $\lambda, w', \rho \in \mathrm{Occ}(\mathbf{x})$ such that $\lambda \in \{\varepsilon, 2\}$, $\rho \in \{0, 02\}$, $w' = (u')^r$ where u' is the prefix of w' satisfying $|f(u')| = |u|$ and w' is unstretchable with respect to $q' = |u'|$. But as $w = \lambda f(w')\rho$ has period q, we see that λ is the left stretch of $(f(w'), q)$, and ρ is the right stretch of $(f(w'), q)$, so that (w, q), and (W, Q) in turn, are in the inner stretch sequence of (w', q'). Since $|w'| < |w|$, this contradicts the minimality of w.

So we have $|u| < 3$ or $|v| < 3$. We now claim that $|u| \leqslant 5$. Suppose otherwise that $|u| > 5$, whence $|v| \leqslant 2$. Then the nth term of the inner power sequence of (w, q) is given

by

$$r_n = \frac{|f^n(w)| + \sum_{k=1}^n |f^{n-k}(\lambda_k \rho_k)|}{|f^n(u)|}$$

$$= \frac{|f^n(u)| + |f^n(v)| + \sum_{k=1}^n |f^{n-k}(\lambda_k \rho_k)|}{|f^n(u)|}$$

$$= 1 + \frac{|f^n(v)|}{|f^n(u)|} + \frac{\sum_{k=1}^n |f^{n-k}(\lambda_k \rho_k)|}{|f^n(u)|} \tag{2}$$

Since $|v| \leqslant 2$ and $\mathtt{11} \notin \mathrm{Fact}(\mathbf{x})$, it follows from Lemma 3.9 that

$$|f^n(v)| \leqslant |f^n(\mathtt{01})|. \tag{3}$$

By inspecting all factors of $\mathbf{x}$ of length 6 and using Lemma 3.9, we find that

$$|f^n(u)| \geqslant 2|f^n(\mathtt{01})|. \tag{4}$$

Since $\lambda_k \in \{\varepsilon, \mathtt{2}\}$ and $\rho_k \in \{\mathtt{0}, \mathtt{02}\}$, we have

$$|f^{n-k}(\lambda_k \rho_k)| \leqslant |f^{n-k}(\mathtt{202})| = |f^{n-k+1}(\mathtt{1})|$$

for all $k \in \{1, 2, \ldots n\}$. Hence we have

$$\sum_{k=1}^n |f^{n-k}(\lambda_k \rho_k)| \leqslant \sum_{k=1}^n |f^{n-k+1}(\mathtt{1})| = \sum_{\ell=1}^n |f^\ell(\mathtt{1})| < |f^n(\mathtt{01})|, \tag{5}$$

where the last inequality can be proven by a straightforward induction. Substituting inequalities (3), (4), and (5) into (2), we find

$$r_n < 1 + \frac{|f^n(\mathtt{01})|}{2|f^n(\mathtt{01})|} + \frac{|f^n(\mathtt{01})|}{2|f^n(\mathtt{01})|} = 2.$$

But $W = U^R$ is in the inner stretch sequence of (w, q), and we assumed that $R \geqslant 2$, so this is a contradiction.

Therefore, we have $|u| \leqslant 5$, which means that there are only finitely many possibilities for u. Since U has an even number of 2's and U is a conjugate of $f^n(u)$ for some $n \geqslant 0$, we see that u must also have an even number of 2's. For each factor u of $\mathbf{x}$ of length at most 5 with an even number of 2's, we find the longest possible periodic extension $w = \lambda u \rho$ of u in $\mathrm{Fact}(\mathbf{x})$, so that λ is the left stretch of $(u, |u|)$ and ρ is the right stretch of $(u, |u|)$ for some particular occurrence of u in $\mathbf{x}$, and (w, q) is unstretchable. By inspection, the only such pairs (w, q) that are not contained in the inner stretch sequence of (w', q') for some word w' with $|w'| < |w|$ belong to S and satisfy $w = u$ (i.e., $v = \varepsilon$). $\qquad \square$

Proposition 5.3 now follows easily from Lemma 5.4, Lemma 5.5, and Lemma 5.9.

5.3 The Critical Exponent

In this subsection, we complete the proof of Theorem 5.1. Essentially all that remains is to analyze the narrow family of outer power sequences described in Proposition 5.3.

For $w \in S$, we let $R_0(w), R_1(w), R_2(w), \ldots$ be the outer power sequence of $(w, |w|)$ for some occurrence $w = x_i \cdots x_j$ with $i \geqslant 1$. We will see shortly that $R_n(w)$ is independent of i. Note that we ignore the case $i = 0$, where all of the left stretches involved in the stretch sequence of w will be empty, but taking $i \geqslant 1$ gives us "enough space" to stretch as far as we need to on the left. Since $\mathbf{x}$ is recurrent, we see that each term in the outer power sequence of an occurrence of w starting at $i = 0$ will be no greater than the corresponding term in the outer power sequence of some later occurrence of w. Thus, since we are interested only in the largest exponents in $\mathbf{z}$, it is safe to ignore the case $i = 0$.

In the next lemma, we derive an expression for $R_n(w)$ in terms of the lengths of the words $\tau(g(f^n(w)))$ and $\tau(g(f^k(0)))$ for $k \leqslant n$. An important consequence is that $R_n(w) \leqslant R_n(0)$ for all $w \in S$ and $n \geqslant 0$, which means that the sequence $(R_n(0))_n$ is responsible for the critical exponent of $\mathbf{z}$.

Lemma 5.10. *For all $w \in S$ and $n \geqslant 0$, we have*

$$R_{2n}(w) = 1 + \frac{\sum_{k=0}^{n} |\tau(g(f^{2k}(0)))| + 3}{|\tau(g(f^{2n}(w)))|} \tag{6}$$

and

$$R_{2n+1}(w) = 1 + \frac{\sum_{k=0}^{n} |\tau(g(f^{2k+1}(0)))| + 10}{|\tau(g(f^{2n+1}(w)))|}. \tag{7}$$

It follows that $R_n(w) \leqslant R_n(0)$ for all $w \in S$ and $n \geqslant 0$.

Proof. Let $w = x_i \cdots x_j$ be an occurrence of w in $\mathbf{x}$ with $i \geqslant 1$. Throughout this proof, we use the notation of Section 5.1 to describe the inner and outer stretch sequences of $(w, |w|)$. In particular, we let $(w_0, q_0), (w_1, q_1), \ldots$ and $(W_0, Q_0), (W_1, Q_1), \ldots$ denote the inner and outer stretch sequences of $(w, |w|)$, and we let $\lambda_k, \lambda_k^*, \rho_k$, and ρ_k^* denote the left and right stretches applied along the way. For all $k \geqslant 0$, we write

$$w_k = x_{i_k} \cdots x_{j_k}.$$

We start by proving the following claim.

Claim 5.11. (a) *For all $k \geqslant 1$, we have* $\lambda_k = \begin{cases} \varepsilon, & \text{if } k \text{ is odd;} \\ 2, & \text{if } k \text{ is even.} \end{cases}$

(b) *For all $k \geqslant 1$, we have* $\rho_k = \begin{cases} 0, & \text{if } k \text{ is odd;} \\ 02, & \text{if } k \text{ is even.} \end{cases}$

(c) *For all $k \geqslant 0$, we have* $|\lambda_k^* \rho_k^*| = \begin{cases} 34, & \text{if } k \text{ is odd;} \\ 22, & \text{if } k \text{ is even.} \end{cases}$

Proof. For parts (a) and (b), we proceed by induction on k. For the base case, we inspect each word w in S individually and examine all possible contexts $x_{i-1} \cdots x_{j+1}$ in which the occurrence $w = x_i \cdots x_j$ might appear. In every case, we see by Lemma 5.8 that $\lambda_1 = \varepsilon$ and $\rho_1 = 0$.

Now suppose that (a) and (b) hold for some integer $k \geqslant 1$. If k is odd, then we have $\lambda_k = \varepsilon$ and $\rho_k = 0$. Thus we can write

$$w_k = x_{i_k} \cdots x_{j_k} = f(w_{k-1})0,$$

So we see that $x_{i_k} = x_{i_k + q_k} = 0$. Since $00 \notin \mathrm{Fact}(\mathbf{x})$ and (w_k, q_k) is unstretchable, we must have $\{x_{i_k - 1}, x_{i_k + q_k - 1}\} = \{1, 2\}$. Therefore, by Lemma 5.8, we have $\lambda_{k+1} = 2$. Similarly, we have $x_{j_k} = x_{j_k - q_k} = 0$, hence $\{x_{j_k + 1}, x_{j_k - q_k + 1}\} = \{1, 2\}$ and $\rho_{k+1} = 02$. On the other hand, if k is even, then we have $\lambda_k = 2$ and $\rho_k = 02$. Thus we can write

$$w_k = x_{i_k} \cdots x_{j_k} = 2f(w_{k-1})02.$$

So we see that $x_{i_k} = x_{i_k + q_k} = 2$. Since $12 \notin \mathrm{Fact}(\mathbf{x})$ and (w_k, q_k) is unstretchable, we must have $\{x_{i_k - 1}, x_{i_k + q_k - 1}\} = \{0, 2\}$. Therefore, by Lemma 5.8, we have $\lambda_{k+1} = \varepsilon$. Similarly, we have $x_{j_k} = x_{j_k - q_k} = 2$, hence $\{x_{j_k + 1}, x_{j_k - q_k + 1}\} = \{0, 2\}$ and $\rho_{k+1} = 0$.

Finally, we prove (c). First suppose that k is odd. Using (a) and (b), we see, as in the second paragraph of the proof, that $x_{i_k} = x_{i_k + q_k} = 0$ and $\{x_{i_k - 1}, x_{i_k + q_k - 1}\} = \{1, 2\}$. Since $w \in S$, we see that $|g(w_k)|$ is even. Therefore λ_k^* is the longest common suffix of $\tau(g(1))$ and $\overline{\tau}(g(2))$ or their sisters (observe that $|g(1)|$ is even and $|g(2)|$ is odd). This suffix is 00202202 or its sister, i.e., $|\lambda_k^*| = 8$. Similarly, from (a) and (b), we see that $x_{j_k} = x_{j_k - q_k} = 0$, so that $\{x_{j_k + 1}, x_{j_k - q_k + 1}\} = \{1, 2\}$, and it follows that ρ_k^* is the longest common prefix of $\tau(g(1)2)$ and $\tau(g(2)2)$ (or their sisters), which has length 26. So we conclude that $|\lambda_k^* \rho_k^*| = 8 + 26 = 34$.

Now suppose that k is even. First suppose that $k \geqslant 2$. From (a) and (b), we see that $x_{i_k} = x_{i_k + q_k} = 2$ and $x_{j_k} = x_{j_k - q_k} = 2$. So $\{x_{i_k - 1}, x_{i_k + q_k - 1}\} = \{0, 2\}$, and we see that λ_k^* is the longest common suffix of $\tau(g(0))$ and $\overline{\tau}(g(2))$ (or their sisters), which has length 2. Similarly, we have $\{x_{j_k + 1}, x_{j_k - q_k + 1}\} = \{0, 2\}$, and it follows that ρ_k^* is the longest common prefix of $\tau(g(0)2)$ and $\tau(g(2)2)$ (or their sisters), which has length 20. Finally, in the case that $k = 0$, we examine the context $x_{i-1} \cdots x_{j+1}$ in which w appears, and confirm by inspection that λ_0^* and ρ_0^* are the same as λ_k^* and ρ_k^* for $k \geqslant 2$. Thus, we conclude that $|\lambda_k^* \rho_k^*| = 2 + 20 = 22$. ∎

We now proceed with the proof of (6). Let $n \geqslant 0$. By definition, we have

$$R_{2n}(w) = \frac{|\tau(g(f^{2n}(w)))| + \sum_{k=1}^{2n} |\tau(g(f^{2n-k}(\lambda_k \rho_k)))| + |\lambda_{2n}^* \rho_{2n}^*|}{|\tau(g(f^{2n}(w)))|} \tag{8}$$

We first show that

$$\sum_{k=1}^{2n} |\tau(g(f^{2n-k}(\lambda_k \rho_k)))| = \sum_{k=1}^{n} |\tau(g(f^{2k}(0)))|. \tag{9}$$

Starting from the left side, we group the terms in pairs, and then use Claim 5.11(a) and (b), and the facts that $f(\texttt{0}) = \texttt{01}$ and $f(\texttt{1}) = \texttt{022}$, as follows.

$$
\sum_{k=1}^{2n} |\tau(g(f^{2n-k}(\lambda_k\rho_k)))| = \sum_{\ell=1}^{n} \left(|\tau(g(f^{2n-2\ell}(\lambda_{2\ell}\rho_{2\ell})))| + |\tau(g(f^{2n-2\ell+1}(\lambda_{2\ell-1}\rho_{2\ell-1})))| \right)
$$

$$
= \sum_{\ell=1}^{n} \left(|\tau(g(f^{2n-2\ell}(\texttt{202})))| + |\tau(g(f^{2n-2\ell+1}(\texttt{0})))| \right)
$$

$$
= \sum_{\ell=1}^{n} \left(|\tau(g(f^{2n-2\ell+1}(\texttt{1})))| + |\tau(g(f^{2n-2\ell+1}(\texttt{0})))| \right)
$$

$$
= \sum_{\ell=1}^{n} |\tau(g(f^{2n-2\ell+2}(\texttt{0})))|
$$

$$
= \sum_{k=1}^{n} |\tau(g(f^{2k}(\texttt{0})))|
$$

Substituting (9) into (8) and using Claim 5.11(c) and the fact that $|\tau(g(\texttt{0}))| = 19$, we obtain

$$
R_{2n}(w) = 1 + \frac{\sum_{k=1}^{n} |\tau(g(f^{2k}(\texttt{0})))| + 22}{|\tau(g(f^{2n}(w)))|} = 1 + \frac{\sum_{k=0}^{n} |\tau(g(f^{2k}(\texttt{0})))| + 3}{|\tau(g(f^{2n}(w)))|}.
$$

The analogous expression (7) for R_{2n+1} can be established in a similar manner; we omit the details.

Finally, by Lemma 3.9, we have $|\tau(g(f^n(w)))| \geqslant |\tau(g(f^n(\texttt{0})))|$ for all $w \in S$. Thus we see that $R_{2n}(w) \leqslant R_{2n}(\texttt{0})$ for all $n \geqslant 0$. $\qquad\square$

With Lemma 5.10 in hand, we now wish to find a simple expression for $|\tau(g(f^n(\texttt{0})))|$. We do so by finding a linear recurrence for $|\tau(g(f^n(\texttt{0})))|$.

Lemma 5.12. *Write* $a_n = |\tau(g(f^n(\texttt{0})))|$ *for* $n \geqslant 0$. *The sequence* (a_n) *satisfies the recurrence*

$$
a_n = 2a_{n-1} + a_{n-3} \tag{10}
$$

with initial values 19, 43, 94.

Proof. Recall that the Parikh vector $P(w)$ of a ternary word $w \in \{\texttt{0}, \texttt{1}, \texttt{2}\}^*$ is the vector $(|w|_{\texttt{0}}, |w|_{\texttt{1}}, |w|_{\texttt{2}})^T$. Let

$$
M_f = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 2 & 1 \end{pmatrix} \quad \text{and} \quad M_g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}
$$

be the incidence matrices of the morphisms f and g, so that $P(f(w)) = M_f P(w)$ and $P(g(w)) = M_g P(w)$ for all $w \in \{\texttt{0}, \texttt{1}, \texttt{2}\}^*$. From the form of the transducer τ, it is straightforward to see that $a_n = (3, 8, 16)P(g(f^n(\texttt{0}))) = (3, 8, 16)M_g M_f^n (1, 0, 0)^T$.

The characteristic polynomial of the matrix M_f equals $x^3 - 2x - 1$ so, by the Cayley-Hamilton Theorem, we have $M_f^3 - 2M_f^2 - I = 0$. Thus $0 = (3, 8, 16)M_g(M_f^3 - 2M_f^2 - I)M_f^n(1, 0, 0)^T = a_{n+3} - 2a_{n+2} - a_n$ for all $n \geqslant 0$. This proves the claim (after checking that the initial values are 19, 43, 94). $\square$

Lemma 5.12 gives us a linear recurrence for $a_n = |\tau(g(f^n(0)))|$. Using a computer algebra system to solve the recurrence, we obtain the following.

Corollary 5.13. *For all $n \geqslant 0$, we have*

$$|\tau(g(f^n(0)))| = \kappa_1 \mu_1^n + \kappa_2 \mu_2^n + \kappa_3 \mu_3^n = \kappa_1 \mu_1^n + 2\operatorname{Re}\left(\kappa_2 \mu_2^n\right),$$

where

$$\begin{aligned}
\mu_1 &\approx 2.20557, \\
\mu_2 &\approx -0.10278 + 0.66546i, \text{ and} \\
\mu_3 &= \overline{\mu_2}
\end{aligned}$$

are the roots of the polynomial $x^3 - 2x^2 - 1$, and

$$\begin{aligned}
\kappa_1 &\approx 19.31167, \\
\kappa_2 &\approx -0.15583 - 0.28157i, \text{ and} \\
\kappa_3 &= \overline{\kappa_2}.
\end{aligned}$$

We are now ready to prove that $\operatorname{ce}(\mathbf{z}) = 1 + 1/(3 - \mu_1)$.

Proof of Theorem 5.1. The critical exponent of $\mathbf{z}$ is the supremum of the exponents of all unstretchable repetitions in $\mathbf{z}$. We have already seen in Example 5.2 that $\mathbf{z}$ has factors of exponent greater than $\frac{9}{4}$, and by Proposition 5.3, every unstretchable repetition with exponent greater than $\frac{9}{4}$ in $\mathbf{z}$ belongs to the outer stretch sequence of $(w, |w|)$ for some $w = x_i \cdots x_j \in \operatorname{Occ}(\mathbf{x})$ with $w \in S$. It follows that

$$\operatorname{ce}(\mathbf{z}) = \sup\{R_n(w) \colon w \in S, n \geqslant 0\}.$$

By Lemma 5.10, we have $R_n(w) \leqslant R_n(0)$ for all $w \in S$ and $n \geqslant 0$, hence

$$\operatorname{ce}(\mathbf{z}) = \sup\{R_n(0) \colon n \geqslant 0\}.$$

So we need to show that $\sup\{R_n(0) \colon n \geqslant 0\} = 1 + 1/(3 - \mu_1)$. It suffices to show that

$$\lim_{n \to \infty} R_{2n} = 1 + \frac{1}{3 - \mu_1} \tag{11}$$

and that

$$R_n \leqslant 1 + \frac{1}{3 - \mu_1} \text{ for all } n \geqslant 0. \tag{12}$$

We first establish (11). By Lemma 5.10, we have

$$R_{2n} = 1 + \frac{\sum_{k=0}^{n} |\tau(g(f^{2k}(0)))| + 3}{|\tau(g(f^{2n}(0)))|}.$$

Evidently, we have

$$\lim_{n\to\infty} R_{2n} = 1 + \lim_{n\to\infty} S_{2n},$$

where

$$S_{2n} = \frac{\sum_{k=0}^{n} |\tau(g(f^{2k}(0)))|}{|\tau(g(f^{2n}(0)))|},$$

as long as $\lim_{n\to\infty} S_{2n}$ exists. By Corollary 5.13, we have

$$\lim_{n\to\infty} S_{2n} = \lim_{n\to\infty} \frac{\sum_{k=0}^{n} (\kappa_1 \mu_1^{2k} + \kappa_2 \mu_2^{2k} + \kappa_3 \mu_3^{2k})}{\kappa_1 \mu_1^{2n} + \kappa_2 \mu_2^{2n} + \kappa_3 \mu_3^{2n}}$$

Breaking the sum in the numerator into three geometric sums, we obtain

$$\lim_{n\to\infty} S_{2n} = \lim_{n\to\infty} \frac{\kappa_1 \cdot \frac{\mu_1^{2n+2}-1}{\mu_1^2-1} + \kappa_2 \cdot \frac{\mu_2^{2n+2}-1}{\mu_2^2-1} + \kappa_3 \cdot \frac{\mu_3^{2n+2}-1}{\mu_3^2-1}}{\kappa_1 \mu_1^{2n} + \kappa_2 \mu_2^{2n} + \kappa_3 \mu_3^{2n}}$$

Finally, dividing through by $\kappa_1 \mu_1^{2n}$ in the numerator and denominator, and using the fact that $|\mu_1| > 1$ and $|\mu_2|, |\mu_3| < 1$, we find

$$\lim_{n\to\infty} S_{2n} = \lim_{n\to\infty} \frac{\frac{1}{\mu_1^2-1} \cdot \frac{\mu_1^{2n+2}-1}{\mu_1^{2n}} + \frac{\kappa_2}{\kappa_1} \cdot \frac{1}{\mu_2^2-1} \cdot \frac{\mu_2^{2n+2}-1}{\mu_1^{2n}} + \frac{\kappa_3}{\kappa_1} \cdot \frac{1}{\mu_3^2-1} \cdot \frac{\mu_3^{2n+2}-1}{\mu_1^{2n}}}{1 + \frac{\kappa_2}{\kappa_1} \cdot \frac{\mu_2^{2n}}{\mu_1^{2n}} + \frac{\kappa_3}{\kappa_1} \cdot \frac{\mu_3^{2n}}{\mu_1^{2n}}}$$

$$= \frac{\mu_1^2}{\mu_1^2 - 1} = \frac{1}{3 - \mu_1}.$$

Now we establish (12). Let $n \geqslant 0$. We show only that

$$R_{2n} \leqslant 1 + \frac{\mu_1^2}{\mu_1^2 - 1}. \tag{13}$$

The proof for R_{2n+1} is similar. By Lemma 5.10, we see that (13) is equivalent to

$$\frac{3 + \sum_{k=0}^{n} |\tau(g(f^{2k}(0)))|}{|\tau(g(f^{2n}(0)))|} \leqslant \frac{\mu_1^2}{\mu_1^2 - 1}, \tag{14}$$

and by Corollary 5.13, this is equivalent to

$$\frac{3 + \kappa_1 \sum_{k=0}^{n} \mu_1^{2k} + 2\operatorname{Re}\left(\kappa_2 \sum_{k=0}^{n} \mu_2^{2k}\right)}{\kappa_1 \mu_1^{2n} + 2\operatorname{Re}\left(\kappa_2 \mu_2^{2n}\right)} \leqslant \frac{\mu_1^2}{\mu_1^2 - 1}. \tag{15}$$

By evaluating the geometric sums and doing some basic algebra, we find that (15) is equivalent to

$$3 + 2\operatorname{Re}\left(\kappa_2 \cdot \frac{\mu_2^{2n+2} - 1}{\mu_2^2 - 1}\right) \leqslant \frac{1}{\mu_1^2 - 1}\left(\kappa_1 + 2\mu_1^2\operatorname{Re}\left(\kappa_2\mu_2^{2n}\right)\right). \qquad (16)$$

So it suffices to prove (16). Starting from the left side and using the fact that $|\mu_2| < 1$ and $n \geqslant 0$, we have

$$3 + 2\operatorname{Re}\left(\kappa_2 \cdot \frac{\mu_2^{2n+2} - 1}{\mu_2^2 - 1}\right) \leqslant 3 + 2 \cdot |\kappa_2| \cdot \frac{|\mu_2|^{2n+2} + 1}{|\mu_2^2 - 1|}$$
$$\leqslant 3 + 2 \cdot |\kappa_2| \cdot \frac{|\mu_2|^2 + 1}{|\mu_2^2 - 1|}.$$

Starting from the right side, we have

$$\frac{1}{\mu_1^2 - 1}\left(\kappa_1 + 2\mu_1^2\operatorname{Re}\left(\kappa_2\mu_2^{2n}\right)\right) \geqslant \frac{1}{|\mu_1^2 - 1|}\left(|\kappa_1| - 2 \cdot |\mu_1|^2 \cdot |\kappa_2| \cdot |\mu_2|^{2n}\right)$$
$$\geqslant \frac{1}{|\mu_1^2 - 1|}\left(|\kappa_1| - 2 \cdot |\mu_1|^2 \cdot |\kappa_2|\right).$$

Finally, we use the values of κ_1, κ_2, μ_1, and μ_2 obtained in proving Corollary 5.13 to verify computationally that

$$3 + 2 \cdot |\kappa_2| \cdot \frac{|\mu_2|^2 + 1}{|\mu_2^2 - 1|} \leqslant \frac{1}{|\mu_1^2 - 1|}\left(|\kappa_1| - 2 \cdot |\mu_1|^2 \cdot |\kappa_2|\right).$$

So we conclude that (16) holds, as desired. $\qquad\square$

6 Alphabets of Size 4 and More

In this section, we briefly discuss the repetition threshold for rich words over alphabets of size 4 and more. By [16, Thm. 2] and Theorem 5.1, we have $\mathrm{RT}(R_2) = 2 + \sqrt{2}/2 \approx 2.7071$ and $\mathrm{RT}(R_3) = 1 + 1/(3 - \mu_1) \approx 2.2588$. Moreover, $\mathrm{RT}^*(R_k) = 2$ for all $k \geqslant 2$ [24]. It is natural to ask the following question.

Question 6.1. What is the number $\mathrm{RT}(R_4)$?

By performing an exhaustive computational search, we have shown that $\mathrm{RT}(R_4) > 2.117$ (a longest 2.117-power-free rich word over 4 letters has length 46628). Since we have found a 2.12-power-free rich word of length one million over 4 letters, we believe this to be close to the true value of $\mathrm{RT}(R_4)$, but we have not attempted to work on this.[3]

Based on the known results and the computational evidence we have gathered, we propose the following conjecture.

Conjecture 6.2. $\lim_{k \to \infty} \mathrm{RT}(R_k) = 2$.

[3]The code used to verify these claims, and a text file containing a 2.12-power-free rich word of length one million over 4 letters, can be found at https://github.com/japeltom/rich-repetition-threshold.

References

[1] J.-P. Allouche and J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, 2003.

[2] G. Badkobeh and P. Ochem, More characterizations of morphic words, preprint, 2023. Available at https://arxiv.org/abs/2312.10757.

[3] Ľ. Balková, E. Pelantová, and Š. Starosta, Palindromes in infinite ternary words. *RAIRO Theor. Inform. Appl.* **43** (2009), 687–702.

[4] Ľ. Balková, E. Pelantová, and Š. Starosta, Sturmian jungle (or garden?) on multiliteral alphabets, *RAIRO Theor. Inform. Appl.* **44** (2010), 443–470.

[5] Ľ. Balková, E. Pelantová, and Š. Starosta, Infinite words with finite defect, *Adv. Appl. Math.* **47**(3) (2011), 562–574.

[6] A. R. Baranwal, *Decision Algorithms for Ostrowski-Automatic Sequences*, Master's Thesis, University of Waterloo, 2020.

[7] A. R. Baranwal, J. D. Currie, L. Mol, P. Ochem, N. Rampersad, and J. Shallit, Antisquares and critical exponents, *Discrete Math. Theor. Comput. Sci.* **25**(2) (2023), article no. 11.

[8] A. R. Baranwal and J. Shallit, Repetitions in infinite palindrome-rich words. In: R. Mercaş and D. Reidenbach (eds.) *Proceedings of WORDS 2019*, Lecture Notes in Computer Science, vol. 11682, pp. 93–105, Springer, 2019.

[9] J. Berstel, *Axel Thue's papers on repetitions in words: A translation*, Publications du LaCIM (Université du Québec à Montréal), vol. 20, 1995.

[10] A. Blondin Massé, S. Brlek, S. Labbé, and L. Vuillon, Palindromic complexity of codings of rotations, *Theoret. Comput. Sci.* **412**(46) (2011), 6455–6463.

[11] M. Bucci, A. De Luca, A. Glen, and L. Q. Zamboni, A connection between palindromic and factor complexity using return words, *Adv. in Appl. Math.* **42** (2009), 69–74.

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

[13] A. Carpi and A. de Luca, Special factors, periodicity, and an application to Sturmian words, *Acta. Inform.* **36** (2000), 983–1006.

[14] J. Cassaigne, On extremal properties of the Fibonacci word, *RAIRO Theor. Inform. Appl.* **42**(4) (2008), 701–715.

[15] J. D. Currie, The analog of overlap-freeness for the period-doubling sequence, *J. Integer Seq.* **26** (2023), article no. 23.8.2.

[16] J. D. Currie, L. Mol, and N. Rampersad, The repetition threshold for binary rich words, *Discrete Math. Theor. Comput. Sci.* **22**(1) (2020), article no. 6.

[17] J. D. Currie, P. Ochem, N. Rampersad, and J. Shallit, Properties of a ternary infinite word, *RAIRO Theor. Inform. Appl.* **57** (2023), article no. 1.

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

[19] J. D. Currie and N. Rampersad, The analogue of overlap-freeness for the Fibonacci morphism, preprint, 2023. Available at https://arxiv.org/abs/2311.12962.

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

[21] F. Dolce, Ľ. Dvořáková, and E. Pelantová, On balanced sequences and their critical exponent, *Theoret. Comput. Sci.* **939** (2023), 18–47.

[22] F. Dolce and E. Pelantová, On morphisms preserving palindromic richness, *Fund. Inform.* **185**(1) (2022), 1–25.

[23] X. Droubay, J. Justin, and G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, *Theoret. Comput. Sci.* **255**(1–2) (2001), 539–553.

[24] Ľ. Dvořáková, K. Klouda, and E. Pelantová, The asymptotic repetition threshold of sequences rich in palindromes, *European J. Combin.* **126** (2025), article no. 104124.

[25] Ľ. Dvořáková, P. Ochem, and D. Opočenská, Critical exponents of binary words with few distinct palindromes, *Electron. J. Combin.* **31**(2) (2024), article no. P2.29.

[26] Ľ. Dvořáková and E. Pelantová, The repetition threshold of episturmian sequences, *European J. Combin.* **120** (2024), article no. 104001.

[27] Ľ. Dvořáková, D. Opočenská, E. Pelantová, and A. M. Shur, On minimal critical exponent of balanced sequences, *Theoret. Comput. Sci.* **922** (2022), 158–169.

[28] A. Glen, J. Justin, S. Widmer, and L. Q. Zamboni, Palindromic richness, *European J. Combin.* **30**(2) (2009), 510–531.

[29] C. Guo, J. Shallit, and A. M. Shur, Palindromic rich words and run-length encodings, *Inform. Process. Lett.* **116**(12) (2016), 735–738.

[30] J. Karhumäki and J. Shallit, Polynomial versus exponential growth in repetition-free binary words, *J. Combin. Theory Ser. A* **105**(2) (2004), 335–347.

[31] D. Krieger, On critical exponents in fixed points of non-erasing morphisms, *Theoret. Comput. Sci.* **376**(1–2) (2007), 70–88.

[32] D. Krieger, The critical exponent of the Arshon words, *RAIRO Theor. Inform. Appl.* **44** (2010), 139–150.

[33] P. Ochem, M. Rao, and M. Rosenfeld, Avoiding or limiting regularities in words. In: V. Berthé and M. Rigo (eds) *Sequences, Groups, and Number Theory*, Trends in Mathematics, pp. 177–212, Birkhäuser, 2018.

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

[35] N. Rampersad and J. Shallit, Repetitions in words. In: V. Berthé and M. Rigo (eds) *Combinatorics, Words and Symbolic Dynamics*, Encyclopedia of Mathematics and its Applications, vol. 159, pp. 101–150, Cambridge University Press, 2016.

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

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

[38] J. Rukavička, On the number of rich words. In: É. Charlier, J. Leroy, and M. Rigo (eds) *Developments in Language Theory, DLT 2017*, Lecture Notes in Computer Science, vol. 10396, pp. 345–352, Springer, 2017.

[39] A. Thue, Über unendliche Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **7** (1906) 1-22.

[40] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **1** (1912), 1–67.

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