

When Can Patterns Characterise Fractional Powers? On a Question of Shallit

Alper Ferudun*

October 20, 2026

Abstract

Shallit asked (Problem 38 in Oberwolfach Reports 7 (2010)) whether, for a real number $\alpha > 1$ that is not an integer, there is a set P of patterns, finite or infinite, such that a word contains an α -power, that is, a factor which is an exact p/q -power with $p/q \geq \alpha$, if and only if some factor of the word matches a pattern of P . If one set P must serve words over every finite alphabet, as in the source's example of overlaps, the answer is negative for every non-integer $\alpha > 1$. For a fixed alphabet of s letters, where P may depend on the alphabet, we show: over two letters such a set exists if and only if $\alpha < 2$; over $s \geq 3$ letters it exists if and only if $\alpha \leq 1 + 1/s$, for every non-integer $\alpha > 1$ outside the window $2 - 2^{-16} < \alpha < 2$, which remains open; over one letter it always exists. Such a set exists if and only if the words of critical exponent at least α are closed under nonerasing morphisms, so a negative answer is certified by a word x and a nonerasing morphism σ with $\text{ce}(x) \geq \alpha > \text{ce}(\sigma(x))$. The negative answer in the universal reading is a short elementary argument, possibly folklore, in which the alphabet grows as α approaches an integer; the source reports that Currie and Cassaigne suggested ways to show impossibility for restricted ranges of α , and we found no write-up of their arguments. The fixed-alphabet results use the Thue–Morse word and its gap words. Three lemmas on explicit binary words and 29 explicit ternary certificates (covering $5/3 < \alpha \leq 2 - 2^{-16}$) are finite statements decided by exact integer computation, each reproduced by several independent programs; all infinite families are treated by proofs. This is an unrefereed note.

1 Introduction

1.1 The question

We use the definitions of Shallit's abstract [Sha10]. An alphabet is a non-empty finite set of symbols. A *pattern* is a non-empty word p over a pattern alphabet Δ , whose letters are called variables. A morphism h is *nonerasing* if $h(a)$ is non-empty for every letter a , and a word w over an alphabet Σ *matches* p if $w = h(p)$ for some nonerasing morphism $h: \Delta^* \rightarrow \Sigma^*$. For integers $p > q \geq 1$, a word x is an *exact p/q -power* if $x = y^e y'$ with $e \geq 1$, y' a prefix of y , $|x| = p$ and $|y| = q$; a word *has an occurrence of an α -power* if it has a factor that is an exact p/q -power with $p/q \geq \alpha$. For example [Sha10, p. 2231], a word has a factor of the form $axaxa$ (a a letter), an overlap, if and only if some factor matches xxx or $xyxyx$. Shallit's question reads as follows (in our paraphrase).

Problem 1.1 ([Sha10, Problem 38]). *Given a real number $\alpha > 1$ that is not an integer, does there exist a set P of patterns, finite or infinite, with the property that a word x has an occurrence of an α -power exactly when some factor of x matches a pattern from P ?*

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: <https://github.com/AlperTheKing>.
Unrefereed preprint. CC BY 4.0.

In a remark after the problem, Shallit reports that Currie and Cassaigne suggested ways to show that this is not possible, “at least for some restricted ranges of α ” [Sha10, p. 2231], and that further work is needed to cover all $\alpha > 1$. For integer $\alpha = m$ the single pattern x^m works, which is why integers are excluded.

The problem does not say whether P may depend on the alphabet of x . We consider two readings.

- (U) *Universal reading*: one set P must work for all words over every finite alphabet, as the patterns xxx , $xyxyx$ do for overlaps.
- (F_s) *Fixed alphabet*: an alphabet Σ with s letters is fixed, P may depend on Σ , and the equivalence is required only for words $x \in \Sigma^*$.

A positive answer in reading (U) gives one in every reading (F_s), and a negative answer in some reading (F_s) gives one in reading (U). Reading (U) matches the overlap example, and the hope expressed in the remark, to cover all $\alpha > 1$, can come true only in reading (U): over a fixed alphabet of s letters a set P exists for $1 < \alpha \leq 1 + \frac{1}{s}$ (Proposition 4.1(b)). On the other hand, the phrase “for some restricted ranges of α ” fits the fixed-alphabet reading better: in reading (U) the negative answer is a short argument that works for all α at once (Theorem 1.2). We state both so that the reader can judge which one the proposer meant.

1.2 Results

Write $\text{ce}(x)$ for the critical exponent of a finite word x (Section 2); a word has an occurrence of an α -power if and only if $\text{ce}(x) \geq \alpha$ (Lemma 2.1).

Theorem 1.2 (reading (U)). *For every non-integer $\alpha > 1$ there is no set of patterns as in Problem 1.1 that works for words over all finite alphabets.*

Theorem 1.3 (readings (F_s)). *Let $\alpha > 1$ be a non-integer real number and Σ an alphabet of s letters. A set P of patterns such that a word $x \in \Sigma^*$ has an occurrence of an α -power if and only if some factor of x matches a pattern of P*

- (a) *exists for every α if $s = 1$;*
- (b) *exists if and only if $\alpha < 2$, if $s = 2$; one can take $P = \{xx, xyx\}$ for $\alpha \leq \frac{3}{2}$ and $P = \{xx\}$ for $\frac{3}{2} < \alpha < 2$;*
- (c) *exists if and only if $\alpha \leq 1 + \frac{1}{s}$, if $s \geq 3$ and $\alpha \notin (2 - 2^{-16}, 2)$; one can take $P = \{xx, xyx\}$.*

For $s \geq 3$ and $2 - 2^{-16} < \alpha < 2$ the question remains open.

alphabet	a set P exists	no set P exists	open
one letter	every α	–	–
two letters	$1 < \alpha < 2$	every non-integer $\alpha > 2$	–
$s \geq 3$ letters	$1 < \alpha \leq 1 + \frac{1}{s}$	$1 + \frac{1}{s} < \alpha \leq 2 - 2^{-16}$, and every non-integer $\alpha > 2$	$2 - 2^{-16} < \alpha < 2$
all alphabets (U)	never	every non-integer $\alpha > 1$	–

Table 1: The answer to Problem 1.1 in the two readings (Theorems 1.2 and 1.3).

1.3 Method, credit and limits

By Lemma 2.2, a set P exists over Σ if and only if the set of words of critical exponent at least α is closed under nonerasing endomorphisms of Σ^* . A negative answer is therefore certified by one word x and one nonerasing morphism σ with $\text{ce}(x) \geq \alpha > \text{ce}(\sigma(x))$, and a certificate over s letters is one over every larger alphabet and in reading (U). Theorem 1.2 follows from a certificate (Proposition 3.1) in which a letter of a periodic word is inflated into many new letters; the number of letters grows as α approaches an integer from above, so it does not settle any fixed alphabet. This argument is elementary and may coincide with the ideas of Currie and Cassaigne mentioned in [Sha10], or be folklore; we claim no priority for it. The fixed-alphabet results use the Thue–Morse word, its desubstitution (Lemma 4.2, proved here) and the square-free ternary words obtained from it by counting the letters 1 between consecutive letters 0, a classical construction of Thue [Thu12] (see [Ber]); the source itself mentions this trick in connection with its Problem 39 [Sha10, p. 2232].

Three statements on explicit binary words (Lemmas 5.2 to 5.4) and 29 ternary certificates (Proposition 6.1(b)) are finite statements decided by exact computation (Section 7). The open window remains because these certificates form an infinite sequence of which only finitely many have been computed (Remark 6.2). No printed answer to Problem 1.1 was found in the sources searched (listed in Section 7); a search that finds nothing is not a proof of novelty.

2 Patterns and morphisms

Positions in a word $w = w_0w_1 \cdots w_{n-1}$ are numbered from 0. An integer $p \geq 1$ is a *period* of w if $w_i = w_{i+p}$ for $0 \leq i < n - p$ (so every $p \geq n$ is a period); $\text{per}(w)$ is the least period of a non-empty word, and the *critical exponent* of a finite non-empty word x is

$$\text{ce}(x) = \max\{|f|/\text{per}(f) : f \text{ a non-empty factor of } x\}.$$

For a period p we call $|f|/p$ the *ratio* of a factor f with period p ; thus $\text{ce}(x)$ is the largest ratio of a factor of x with some period. For the two-sided periodic word $M^{\mathbb{Z}}$ of a non-empty word M the same notions apply to its finite factors.

Lemma 2.1. *Let $\alpha > 1$. A word x has an occurrence of an α -power if and only if $\text{ce}(x) \geq \alpha$.*

Proof. A factor which is an exact p/q -power with $p/q \geq \alpha$ has length p and period q , so its least period is at most q and its exponent is at least $p/q \geq \alpha$. Conversely, if a factor f has $|f|/\text{per}(f) \geq \alpha > 1$, then $|f| > \text{per}(f)$ and f is an exact $|f|/\text{per}(f)$ -power. $\square$

Lemma 2.2 (closure criterion). *Let Σ be an alphabet, $\alpha > 1$, and $L_\alpha = \{x \in \Sigma^* : x \neq \varepsilon, \text{ce}(x) \geq \alpha\}$. There is a set P of patterns such that, for every $x \in \Sigma^*$, $x \in L_\alpha$ if and only if some factor of x matches a pattern of P , if and only if $\sigma(L_\alpha) \subseteq L_\alpha$ for every nonerasing morphism $\sigma : \Sigma^* \rightarrow \Sigma^*$.*

Proof. Suppose P exists, let $x \in L_\alpha$ and let σ be nonerasing. Some factor f of x equals $h(p_0)$ with $p_0 \in P$ and h nonerasing. Then $\sigma(f) = (\sigma \circ h)(p_0)$ is a factor of $\sigma(x)$ that matches p_0 , since $\sigma \circ h$ is nonerasing; so $\sigma(x) \in L_\alpha$. Conversely, suppose L_α is closed, and let P be L_α itself, each word read as a pattern over a copy of Σ . Every $x \in L_\alpha$ matches itself. If a factor f of x equals $h(p)$ with $p \in L_\alpha$ and h nonerasing, then h , read on Σ and extended by the identity to the letters not in p , is a nonerasing endomorphism of Σ^* , so $f \in L_\alpha$ by closure, and $x \in L_\alpha$ because $\text{ce}(x) \geq \text{ce}(f)$. $\square$

Corollary 2.3. *Let $\alpha > 1$, let x be a word over an alphabet Σ and σ a nonerasing endomorphism of Σ^* with $\text{ce}(x) \geq \alpha > \text{ce}(\sigma(x))$. Then no set P as in Problem 1.1 exists over any alphabet $\Sigma' \supseteq \Sigma$, nor in reading (U).*

Proof. Extend σ to Σ' by the identity; it remains nonerasing, and L_α over Σ' is not closed. Apply Lemma 2.2; a set P in reading (U) would in particular work over Σ' . $\square$

We call such a pair (x, σ) a *certificate* for the interval $(\text{ce}(\sigma(x)), \text{ce}(x)]$: it rules out every α in this interval. If $\alpha = m$ is an integer, L_m is closed, since $\sigma(y^m) = \sigma(y)^m$; this is the reason for excluding integers.

3 The universal reading

Throughout, $\alpha = k + \beta$ with $k = \lfloor \alpha \rfloor \geq 1$ and $0 < \beta < 1$.

Proposition 3.1. *Let $k \geq 1$ and $1 \leq r < q < N$ be integers, let $a_1, \dots, a_N$ be distinct letters, $x = (a_1 \cdots a_q)^k a_1 \cdots a_r$, and let σ be the morphism with $\sigma(a_q) = a_q a_{q+1} \cdots a_N$ and $\sigma(a_i) = a_i$ for $i \neq q$. Then $\text{ce}(x) = k + r/q$ and $\text{ce}(\sigma(x)) = k + r/N$.*

Proof. Since $r < q$, the letter a_q does not occur in the final block $a_1 \cdots a_r$, so $\sigma(x) = Y^k a_1 \cdots a_r$ with $Y = a_1 \cdots a_N$. Thus $\sigma(x)$ is a prefix of Y^ω , and in Y^ω two positions carry the same letter if and only if their distance is a multiple of N . A factor f of $\sigma(x)$ with a period $p < |f|$ satisfies $f_0 = f_p$, so $N \mid p$ and $|f|/p \leq |\sigma(x)|/N = k + r/N$; a factor without a period smaller than its length has ratio 1. The word $\sigma(x)$ itself has period N and ratio $k + r/N$. The same argument with q in place of N gives $\text{ce}(x) = k + r/q$. $\square$

Proof of Theorem 1.2. Let q be an integer with $q > 1/(1 - \beta)$ and $r = \lceil \beta q \rceil$. Then $1 \leq r < \beta q + 1 < q$ and $r/q \geq \beta$. Let N be an integer with $N > r/\beta$; then $r/N < \beta \leq r/q$, so $N > q$. By Proposition 3.1, $\text{ce}(x) = k + r/q \geq \alpha > k + r/N = \text{ce}(\sigma(x))$, and Corollary 2.3 applies. $\square$

This argument is elementary; it may be folklore or coincide with the ideas of Currie and Cassaigne mentioned in [Sha10], and we claim no priority for it.

The certificate uses $N > r/\beta \geq 1/\beta$ letters. Over a fixed alphabet of s letters it is available only when $N \leq s$, and it cannot exist for $\alpha \leq 1 + \frac{1}{s}$ (Proposition 4.1(b) below).

4 Fixed alphabets: general constructions

Proposition 4.1. *Let Σ have s letters.*

- (a) *If $s = 1$, the set $P = \{z^{\lceil \alpha \rceil}\}$ works for every $\alpha > 1$.*
- (b) *If $1 < \alpha \leq 1 + \frac{1}{s}$, the set $P = \{xx, xyx\}$ works.*
- (c) *If $s = 2$ and $\frac{3}{2} < \alpha < 2$, the set $P = \{xx\}$ works.*

Proof. (a) A non-empty unary word a^n has $\text{ce}(a^n) = n$, so $\text{ce}(a^n) \geq \alpha$ if and only if $n \geq \lceil \alpha \rceil$, if and only if a^n has a factor matching $z^{\lceil \alpha \rceil}$. (b) A factor $h(xx)$ or $h(xyx)$, with h nonerasing, contains the first letter of $h(x)$ twice, and conversely a factor awa (a a letter) is $h(xx)$ if w is empty and $h(xyx)$ otherwise. So a word has a factor matching P if and only if some letter occurs twice in it. If so, take two equal letters at minimal distance d ; the factor awa between them has d distinct letters $a, w_0, \dots, w_{d-2}$, so $d \leq s$, and it has period d and ratio $(d+1)/d \geq 1 + \frac{1}{s} \geq \alpha$. If no letter is repeated, every factor has ratio $1 < \alpha$. By Lemma 2.1 this is the claim. (c) The square-free binary words are the factors of 010 and 101; their critical exponent is at most $\frac{3}{2} < \alpha$. Every other binary word contains a square and has critical exponent at least $2 > \alpha$. $\square$

Let μ be the Thue–Morse morphism $0 \mapsto 01, 1 \mapsto 10$ (on other letters $a \mapsto ab, b \mapsto ba$). The following desubstitution property is standard (compare [AS03, Lot97]); its consequence that $\mu^n(0)$ is overlap-free is due to Thue [Thu06] (see [Ber]). We include the short proof.

Lemma 4.2. *Let w be a binary word and f a factor of $\mu(w)$ with a period p and $|f| > 2p$. Then p is even and w has a factor with period $p/2$ and ratio at least $|f|/p$. Consequently, for $n \geq 1$, every factor f of $\mu^n(w)$ with a period p and $|f| > 2p$ has $2^n \mid p$, and w has a factor with period $p/2^n$ and ratio at least $|f|/p$. In particular $\text{ce}(\mu^n(w)) \leq \max(2, \text{ce}(w))$, and $\mu^n(0)$ is overlap-free.*

Proof. $\mu(w)$ is a product of the blocks 01 and 10 starting at even positions, so two equal adjacent letters of $\mu(w)$ start at an odd position. Let f start at position t of $\mu(w)$, and $L = |f| \geq 2p + 1$.

If p were odd: for $0 \leq i \leq L - p - 2$, $f_i = f_{i+1}$ would imply $f_{i+p} = f_{i+p+1}$, two pairs of equal adjacent letters starting at the positions $t + i$ and $t + i + p$ of different parity. Hence the prefix of f of length $L - p \geq p + 1$ alternates, and $f_p \neq f_0$ because p is odd, against the period p .

So $p = 2p'$. If t is odd, the letters of $\mu(w)$ at the even positions $t - 1$ and $t - 1 + p$ are the complements of those at t and $t + p$, which are equal; so f extended by one letter to the left still has period p and a larger ratio. We may therefore assume $t = 2a$. If L is even, $f = \mu(g)$ with $g = w_a \cdots w_{a+L/2-1}$, and since the first letter of $\mu(c)$ is c , the relations $f_{2i} = f_{2i+p}$ give $g_i = g_{i+p'}$: g has period p' and ratio $(L/2)/p' = L/p$. If L is odd, f is $\mu(w_a \cdots w_{a+(L-3)/2})$ followed by the letter $w_{a+(L-1)/2}$, and the same argument shows that $g = w_a \cdots w_{a+(L-1)/2}$ has period p' and ratio $(L+1)/p > L/p$.

The ratio stays larger than 2, so the statement can be iterated n times. A factor of $\mu^n(w)$ with ratio larger than 2 thus produces a factor of w with at least the same ratio, which gives $\text{ce}(\mu^n(w)) \leq \max(2, \text{ce}(w))$; for $w = 0$ there is no factor of ratio larger than 1, so $\mu^n(0)$ has no factor f with $|f| > 2\text{per}(f)$. $\square$

Proposition 4.3. *Let $k \geq 2$ and $n \geq 1$. For $x = (01)^k 0$ and the morphism $\sigma_n(0) = 0, \sigma_n(1) = \mu^n(0)$ with its first letter deleted, $\text{ce}(x) = k + \frac{1}{2}$ and $\text{ce}(\sigma_n(x)) = k + 2^{-n}$. Hence there is no set P over two letters for $k < \alpha \leq k + \frac{1}{2}$.*

Proof. x is a prefix of $(01)^\omega$, in which equal letters are at even distance; so a factor with a period $p < |f|$ has p even and ratio at most $(2k+1)/2$, attained by x . Since $\sigma_n(01) = \mu^n(0)$ and $\mu^n(0)$ begins with 0, $u = \sigma_n(x) = \mu^n(0^k)0$ is a prefix of $\mu^n(0^{k+1})$. A factor of u with a period p and length larger than $2p$ has $2^n \mid p$ by Lemma 4.2, hence ratio at most $|u|/2^n = k + 2^{-n}$; all other factors have ratio at most $2 \leq k$; and u has period 2^n . So $\text{ce}(u) = k + 2^{-n}$. For $k < \alpha \leq k + \frac{1}{2}$ choose n with $2^{-n} < \alpha - k$ and apply Corollary 2.3. $\square$

Proposition 4.4. *Let $k \geq 2$ and $0 < \beta < 1$, and let $j \geq 0$ and an even $r \geq 2$ satisfy $2^j/(2^j + 1) \geq \beta > 2^j/(2^{j+r} + 1)$. Over the letters a, b, c let $T = \mu^j(a)$ and let U be $\mu^{j+r}(a)$ without its last 2^j letters, so that $\mu^{j+r}(a) = UT$. For $x = (Tc)^k T$ and the morphism $\sigma(a) = a, \sigma(b) = b, \sigma(c) = cU$, $\text{ce}(x) = k + 2^j/(2^j + 1)$ and $\text{ce}(\sigma(x)) = k + 2^j/(2^{j+r} + 1)$. Hence over $s \geq 3$ letters there is no set P for any non-integer $\alpha > 2$.*

Proof. Such j and r exist since $2^j/(2^j + 1) \rightarrow 1$ and $2^j/(2^{j+r} + 1) \rightarrow 0$. As $\mu^2(a) = abba$ ends with a , so does $\mu^r(a)$ for even r , and $\mu^{j+r}(a) = \mu^j(\mu^r(a))$ ends with $\mu^j(a) = T$. Now $\sigma(x) = (TcU)^k T$ is a prefix of V^ω with $V = TcU$, $|V| = 2^{j+r} + 1$, and c occurs in V^ω exactly at the positions congruent to $|T|$ modulo $|V|$. Let f be a factor of $\sigma(x)$ with period p . If f contains no c , it is a factor of T or of $UT = \mu^{j+r}(a)$, which are overlap-free by Lemma 4.2, so its ratio is at most 2. If f contains c at position i and $|V| \nmid p$, then neither $i + p$ nor $i - p$ is a position of f (it would carry a c at distance p from another c), so $|f| \leq i + p < 2p$. If $|V| \mid p$, the ratio is at most $|\sigma(x)|/|V| = k + 2^j/(2^{j+r} + 1)$, and $\sigma(x)$ attains it. Since $k \geq 2$, this is $\text{ce}(\sigma(x))$. The same three cases for x , a prefix of $(Tc)^\omega$, give $\text{ce}(x) = k + 2^j/(2^j + 1)$. For $\alpha = k + \beta$ apply Corollary 2.3. $\square$

Proposition 4.5. *Let $s \geq 3$. For $1 \leq r < q < N \leq s$, the certificate of Proposition 3.1 with $k = 1$ uses $N \leq s$ letters and rules out every $\alpha \in (1 + r/N, 1 + r/q]$. The union of these intervals is $(\frac{4}{3}, \frac{3}{2}]$ for $s = 3$ and $(1 + \frac{1}{s}, 1 + \frac{s-2}{s-1}]$ for $s \geq 4$.*

Proof. For $s = 3$ only $(r, q, N) = (1, 2, 3)$ occurs. Let $s \geq 4$. Each interval lies in $(1 + \frac{1}{s}, 1 + \frac{s-2}{s-1}]$, since $r/N \geq 1/s$ and $r/q \leq (q-1)/q \leq (s-2)/(s-1)$. The intervals $(1 + \frac{r}{s}, 1 + \frac{r}{r+1}]$, $r = 1, \dots, s-2$, cover this range: their right ends increase, and $\frac{r+1}{s} \leq \frac{r}{r+1}$, i.e. $(r+1)^2 \leq rs$, holds for $r \leq s-3$ because $r+2 + \frac{1}{r} \leq s$. $\square$

5 The binary alphabet above exponent two

Proposition 4.3 covers $(k, k + \frac{1}{2}]$. For the remaining ranges we use a family of words derived from the Thue–Morse word.

Gap words. For a binary word z beginning with 0, write $z = 01^{e_1}01^{e_2} \dots 01^{e_N}$ and call $\gamma(z) = e_1e_2 \dots e_N$ its *gap word*, a word over the digits. Let $\mathcal{F}$ consist of the one-letter word 2 ($N = 1$), of $G_n = \gamma(\mu^n(0))$ for $n \geq 2$ ($N = 2^{n-1}$) and of $H_n = \gamma(\mu^n(001))$ for $n \geq 1$ ($N = 3 \cdot 2^{n-1}$). For example $G_2 = 20$, $H_1 = 120$, $G_3 = 2101$, $H_2 = 202101$. For a word $E = e_1 \dots e_N$ over $\{0, 1, 2\}$ put

$$\tau(E) = 0^{e_1}10^{e_2}1 \dots 0^{e_N}1, \quad M(E) = A^{e_1}BA^{e_2}B \dots A^{e_N}B, \quad A = 001, \quad B = 010010011,$$

so that $M(E) = \sigma(\tau(E))$ for the morphism $\sigma(0) = A$, $\sigma(1) = B$. In the two-sided word $M(E)^{\mathbb{Z}}$ we index the blocks $A^{e_t}B$ by $t \in \mathbb{Z}$ (with $e_{t+N} = e_t$) and call the starting position d_t of the B that ends block t *designated*; thus $d_t - d_{t-1} = 9 + 3e_t \in [9, 15]$.

Lemma 5.1. *Let $E \in \mathcal{F}$ with $N = |E| \geq 2$.*

- (i) *Every letter of E is 0, 1 or 2, and $e_1 + \dots + e_N = N$; hence $|\tau(E)| = 2N$ and $|M(E)| = 12N$.*
- (ii) *$E^{\mathbb{Z}}$ contains no square uu with $N \nmid |u|$. In particular consecutive letters of $E^{\mathbb{Z}}$ are distinct, and every factor of $E^{\mathbb{Z}}$ of length less than $2N$ is square-free.*

Proof. Let $z = \mu^n(v)$ with $v = 0$ and $n \geq 2$, or $v = 001$ and $n \geq 1$, so that $E = \gamma(z)$ and $z^{\mathbb{Z}}$ has gap sequence $E^{\mathbb{Z}}$. The word $z^{\mathbb{Z}}$ is a product of blocks 01, 10, so it has no factor 000 or 111, and z has $|z|/2$ letters 0 and $|z|/2$ letters 1. This gives (i). A square uu with $|u| = q$ in $E^{\mathbb{Z}}$ gives a factor $XX0$ of $z^{\mathbb{Z}}$, where X begins with 0, contains exactly q letters 0, and has length ℓ . This factor has period ℓ and length $2\ell + 1$, and it is a factor of $\mu^n(v^m)$ for large m . By Lemma 4.2, $2^n \mid \ell$ and v^m has a factor with period $\ell/2^n$ and ratio larger than 2. If $v = 001$, then $3 \mid \ell/2^n$: in $(001)^m$ the letters 1 sit at the positions congruent to 2 modulo 3, so for a shift $p' \not\equiv 0 \pmod{3}$ the comparison of positions i and $i + p'$ fails at least once among any three consecutive i ; a factor with period $p' \geq 2$ thus has length at most $p' + 2 \leq 2p'$, and 000 does not occur. In both cases $|z|$ divides ℓ , so the window X of length ℓ of the $|z|$ -periodic word $z^{\mathbb{Z}}$ contains exactly $(\ell/|z|)N$ letters 0, that is, $N \mid q$. The last two assertions follow, since a square inside a factor of length less than $2N$ has $|u| < N$. $\square$

The following three lemmas are finite statements about explicit words, proved by exact computation (Section 7).

Lemma 5.2. *For all $(a, b) \in \{0, 1, 2\}^2$, the word B occurs in BA^aBA^bB only at the positions 0, $9 + 3a$ and $18 + 3a + 3b$. Consequently, for every two-sided sequence (e_t) over $\{0, 1, 2\}$, every occurrence of B in $\dots A^{e_t}BA^{e_{t+1}}B \dots$ starts at a designated position.*

Proof. The first statement is checked by computer for the nine pairs. For the second, let an occurrence start at i with $d_{t-1} \leq i < d_t$. It ends before $d_t + 9$, so it lies in the stretch $\mathbf{BA}^{e_t}\mathbf{B}$ that starts at d_{t-1} , a prefix of $\mathbf{BA}^{e_t}\mathbf{BA}^{e_{t+1}}\mathbf{B}$; hence $i = d_{t-1}$. $\square$

Lemma 5.3. *For every square-free word $w = w_1 \cdots w_{14}$ over $\{0, 1, 2\}$ and every $p < 42$, every factor of $\mathbf{BA}^{w_1}\mathbf{BA}^{w_2}\mathbf{B} \cdots \mathbf{A}^{w_{14}}\mathbf{B}$ with period p has length at most $8p/3$.*

There are 456 such words; the largest ratio is exactly $\frac{8}{3}$, attained for $p = 3$.

Lemma 5.4. *Let $E \in \mathcal{F}$ with $N \leq 6$, that is, $E \in \{2, G_2, H_1, G_3, H_2\}$, and $M = M(E)$. Then M is primitive, and every factor of $M^{\mathbb{Z}}$ with a period p that is not a multiple of $|M|$ has ratio at most $\frac{8}{3}$.*

Reduction to a finite computation. Let $m = |M|$. Whether $M_i^{\mathbb{Z}} = M_{i+p}^{\mathbb{Z}}$ depends only on i and p modulo m . If, for some p with $m \nmid p$, it held for all i , then M would equal a non-trivial conjugate of itself and hence be a proper power. For primitive M let r_p be the longest run of consecutive i with $M_i^{\mathbb{Z}} = M_{i+p}^{\mathbb{Z}}$, computed on the cyclic word; then the longest factor with period p has length $p + r_p$, and since r_p depends only on p modulo m , the largest ratio for $p \equiv p' \pmod{m}$ is $(p' + r_{p'})/p'$ with $1 \leq p' < m$. The computation of these $m - 1$ numbers gives the lemma (here $m \leq 72$). Alternatively, by the theorem of Fine and Wilf [FW65] a factor with period p , $m \nmid p$, and length at least $p + m$ would have period $\gcd(p, m) < m$ and contain a conjugate of M , which would be a proper power; so such factors are shorter than $p + m$, their ratio is less than $1 + m/p < 2$ if $p > m$, and for $p < m$ they lie in M^3 . Both computations give the maximum $\frac{8}{3}$, attained for $p = 3$. $\square$

Lemma 5.5. *Let $E \in \mathcal{F}$ with $N \geq 8$ and $M = M(E)$. Every factor f of $M^{\mathbb{Z}}$ with a period p that is not a multiple of $|M|$ satisfies $|f| \leq 8p/3$.*

Proof. Suppose $|f| > 8p/3$.

Case $p < 42$. The prefix f' of f of length $\lfloor 8p/3 \rfloor + 1 \leq 110$ has period p . If it starts at a position i with $d_{t-1} \leq i < d_t$, it ends before $d_t + 109 \leq d_{t+13} + 9$, since $d_{t+13} - d_t \geq 13 \cdot 9$. So f' is a factor of $\mathbf{BA}^{e_t}\mathbf{B} \cdots \mathbf{A}^{e_{t+13}}\mathbf{B}$, and $e_t \cdots e_{t+13}$ is a factor of $E^{\mathbb{Z}}$ of length $14 < 2N$, square-free by Lemma 5.1(ii). Lemma 5.3 contradicts $|f'| > 8p/3$.

Case $p \geq 42$. Write $f = M^{\mathbb{Z}}[s, s + L)$. Call t *admissible* if $s \leq d_t$ and $d_t + p + 9 \leq s + L$. For admissible t the copy of $\mathbf{B}$ at d_t lies in f together with its shift by p , so $\mathbf{B}$ occurs at $d_t + p$, which is designated by Lemma 5.2: $d_t + p = d_{\varphi(t)}$. The admissible t form an interval $[t_a, t_b]$ of integers.

Suppose first that some t is admissible. If t and $t + 1$ are admissible, the stretch $M^{\mathbb{Z}}[d_t, d_{t+1} + 9) = \mathbf{BA}^{e_{t+1}}\mathbf{B}$ is repeated at distance p . A designated position j strictly between $d_{\varphi(t)}$ and $d_{\varphi(t+1)}$ would carry a copy of $\mathbf{B}$ inside the repeated stretch, and shifting it back by p would give an occurrence of $\mathbf{B}$ strictly between d_t and d_{t+1} , against Lemma 5.2. So $\varphi(t + 1) = \varphi(t) + 1$ and $e_{\varphi(t)+1} = e_{t+1}$. Hence $\varphi(t) = t + q$ on $[t_a, t_b]$ for a fixed $q \geq 1$, $e_{u+q} = e_u$ for $t_a < u \leq t_b$, and $p = d_{t_a+q} - d_{t_a}$ is the sum of the q consecutive gaps $9 + 3e_u$, $t_a < u \leq t_a + q$. If $N \mid q$, this sum is $(q/N)|M|$, against the hypothesis; so $N \nmid q$. The factor $e_{t_a+1} \cdots e_{t_b+q}$ of $E^{\mathbb{Z}}$ has period q and length $t_b - t_a + q$; by Lemma 5.1(ii) it contains no square of period q , so $t_b - t_a \leq q - 1$. Then $d_{t_b} - d_{t_a}$ is a sum of at most $q - 1$ of the q gaps that make up p , and each gap is at least 9, so $d_{t_b} - d_{t_a} \leq p - 9$. Since $t_a - 1$ and $t_b + 1$ are not admissible and consecutive designated positions differ by at most 15, we have $d_{t_a} - s \leq 14$ and $(s + L - p - 9) - d_{t_b} \leq 14$. Adding, $L - p - 9 \leq p + 19$, that is, $L \leq 2p + 28$.

If no t is admissible, the integer interval $[s, s + L - p - 9]$ contains no designated position, so it has at most 14 elements: $L \leq p + 22$.

In both cases $L \leq 2p + 28 \leq 8p/3$ because $p \geq 42$, a contradiction. $\square$

Proposition 5.6. *Let $k \geq 2$. For every $\alpha \in (k, k + 1)$ there is no set P over two letters.*

Proof. Let $E \in \mathcal{F}$, $x = \tau(E)^k \tau(E)^-$, where w^- denotes w without its last letter, and $\sigma(0) = A$, $\sigma(1) = B$. The word x has period $|\tau(E)|$ and length $(k+1)|\tau(E)| - 1$, so $\text{ce}(x) \geq k+1 - 1/|\tau(E)|$. Since $\tau(E)$ ends with 1, $\sigma(x) = M^k M'$ with $M = M(E)$ and M' equal to M without its final B; this is a factor of $M^{\mathbb{Z}}$ of length $(k+1)|M| - 9$. A factor of $\sigma(x)$ whose period is not a multiple of $|M|$ has ratio at most $\frac{8}{3}$ (Lemma 5.4 for $N \leq 6$, Lemma 5.5 for $N \geq 8$); a factor whose period is a multiple of $|M|$ has ratio at most $|\sigma(x)|/|M| = k+1 - 9/|M|$. By Corollary 2.3 every α in

$$J_E = (\max(\frac{8}{3}, k+1 - 9/|M|), k+1 - 1/|\tau(E)|]$$

is ruled out. For $N \geq 2$, Lemma 5.1(i) gives $J_E = (\max(\frac{8}{3}, k+1 - \frac{3}{4N}), k+1 - \frac{1}{2N}]$, and for $E = 2$, $|\tau(E)| = 3$ and $|M| = 15$, so $J_2 = (\max(\frac{8}{3}, k + \frac{2}{5}), k + \frac{2}{3}]$. The sizes $N \geq 2$ in $\mathcal{F}$ are 2, 3, 4, 6, 8, 12, $\dots$, and consecutive ones satisfy $N' \leq \frac{3}{2}N$, hence $\frac{3}{4N'} \geq \frac{1}{2N}$: consecutive intervals $(k+1 - \frac{3}{4N'}, k+1 - \frac{1}{2N}]$ overlap or touch, and their union is $(k + \frac{5}{8}, k+1)$.

If $k \geq 3$, then $\frac{8}{3} < k + \frac{2}{5}$, the maxima disappear, and $(k, k + \frac{1}{2}]$ (Proposition 4.3), $J_2 = (k + \frac{2}{5}, k + \frac{2}{3}]$ and $(k + \frac{5}{8}, k+1)$ cover $(k, k+1)$. If $k = 2$, the intervals J_E with $N \geq 2$ cover $(\frac{8}{3}, 3)$, Proposition 4.3 covers $(2, \frac{5}{2}]$, and the certificate $x = 00100100$, $\sigma(0) = 0$, $\sigma(1) = 101101$, with $\sigma(x) = 001011010010110100$, $\text{ce}(x) = \frac{8}{3}$ and $\text{ce}(\sigma(x)) = \frac{9}{4}$, covers $(\frac{9}{4}, \frac{8}{3}]$. $\square$

For $N \leq 24$ and $k = 2, 3, 4$ the computation gives exactly $\text{ce}(x) = k+1 - 1/|\tau(E)|$ and $\text{ce}(\sigma(x)) = \max(\frac{8}{3}, k+1 - 9/|M|)$; for example $E = G_3$, $k = 2$ gives $\text{ce}(x) = \frac{23}{8}$ and $\text{ce}(\sigma(x)) = \frac{45}{16}$.

6 Three or more letters below exponent two

By Corollary 2.3, certificates over three letters serve every alphabet with $s \geq 3$ letters. Let $N_0 = 3$, and let $N_1 < N_2 < \dots < N_{29}$ be the numbers 2^j ($2 \leq j \leq 16$) and $3 \cdot 2^j$ ($1 \leq j \leq 14$), so that $N_1 = 4$, $N_2 = 6$, $N_3 = 8$, $N_4 = 12$, $\dots$, $N_{29} = 65536$. For $N = N_i$, $i \geq 1$, let E_N be the word G_n or H_n of Section 5 with N letters, written over $\{a, b, c\}$ by $0 \mapsto a$, $1 \mapsto b$, $2 \mapsto c$. Since $\mu^n(0)$ for $n \geq 3$ and $\mu^n(001)$ for $n \geq 2$ begin with 011, E_N begins with c ; write $E_N = c u_N$ and $x_N = u_N c u_N$, a word of length $2N - 1$. Let $\sigma_0(a) = a$, $\sigma_0(b) = bc$, $\sigma_0(c) = bac$.

Proposition 6.1. *Over three letters:*

(a) $x = aba$ and $\sigma(a) = a$, $\sigma(b) = bc$, $\sigma(c) = c$ give $\sigma(x) = abca$, $\text{ce}(x) = \frac{3}{2}$, $\text{ce}(\sigma(x)) = \frac{4}{3}$; $x = abacaba$ and $\sigma(a) = a$, $\sigma(b) = b$, $\sigma(c) = cbc$ give $\sigma(x) = abacbcaba$, $\text{ce}(x) = \frac{7}{4}$, $\text{ce}(\sigma(x)) = \frac{3}{2}$. So every $\alpha \in (\frac{4}{3}, \frac{7}{4}]$ is ruled out.

(b) For $1 \leq i \leq 29$, $\text{ce}(x_{N_i}) = 2 - 1/N_i$ and $\text{ce}(\sigma_0(x_{N_i})) = 2 - 1/N_{i-1}$. So every $\alpha \in (\frac{5}{3}, 2 - 2^{-16}]$ is ruled out.

Proof. (a) can be checked directly. (b) The word x_N is a factor of $E_N^{\mathbb{Z}}$ of length $2N - 1$, square-free by Lemma 5.1(ii); so a factor with period p has length at most $2p - 1$, and also at most $2N - 1$, and its ratio is at most $2 - 1/N$, attained by x_N itself. The values $\text{ce}(\sigma_0(x_{N_i}))$ were computed exactly (Section 7). The certificates rule out the intervals $(2 - 1/N_{i-1}, 2 - 1/N_i]$, whose union is $(\frac{5}{3}, 2 - 2^{-16}]$. $\square$

Proof of Theorem 1.3. (a) is Proposition 4.1(a). (b) For $1 < \alpha < 2$ use Proposition 4.1(b) ($\alpha \leq \frac{3}{2}$) and (c); for non-integer $\alpha > 2$, Proposition 5.6. (c) For $\alpha \leq 1 + \frac{1}{s}$ use Proposition 4.1(b). For non-integer $\alpha > 2$, Proposition 4.4 (or Proposition 5.6 with Corollary 2.3). For $1 + \frac{1}{s} < \alpha \leq 2 - 2^{-16}$: if $s = 3$, Proposition 4.5 and Proposition 6.1 cover $(\frac{4}{3}, \frac{3}{2}]$, $(\frac{4}{3}, \frac{7}{4}]$ and $(\frac{5}{3}, 2 - 2^{-16}]$; if $s \geq 4$, Proposition 4.5 covers $(1 + \frac{1}{s}, 1 + \frac{s-2}{s-1}]$, which contains $(1 + \frac{1}{s}, \frac{4}{3}]$, and Proposition 6.1 with Corollary 2.3 covers $(\frac{4}{3}, 2 - 2^{-16}]$. $\square$

Only the following values rest on computations: for $s = 2$, the ranges $(k + \frac{1}{2}, k + 1)$; for $s \geq 3$, the values in $(\frac{7}{4}, 2 - 2^{-16}]$ that Proposition 4.5 does not cover. For $s \geq 3$ the range $\alpha > 2$ is covered by the computer-free Proposition 4.4.

Remark 6.2 (the open window). Let $\theta(a) = b$, $\theta(b) = ca$, $\theta(c) = cba$ and let π be the renaming $a \mapsto c$, $b \mapsto a$, $c \mapsto b$; then $\sigma_0 = \pi \circ \theta$, so $\text{ce}(\sigma_0(w)) = \text{ce}(\theta(w))$. Since $\mu(01^e) = 01(10)^e$ has the gap word b , ca , cba for $e = 0, 1, 2$, we have $\gamma(\mu(z)) = \theta(\gamma(z))$ for binary z beginning with 0 (letters renamed), hence $E_{2N} = \theta(E_N)$, and $\theta(x_N)$ is $E_{2N}E_{2N}$ without its first three letters: a factor of $E_{2N}^{\mathbb{Z}}$ of length $4N - 3$, square-free by Lemma 5.1(ii). Let $N = N_i$ and $N' = N_{i-1}$, so $N \leq \frac{3}{2}N'$. Factors of $\theta(x_N)$ with a period $p \leq N'$ have ratio at most $2 - 1/p \leq 2 - 1/N'$, and those with $p \geq 2N$ have ratio at most $(4N - 3)/(2N) \leq 2 - 1/N'$. Hence the chain of Proposition 6.1(b) continues for all i , which would close the window, if and only if, for every i and every p with $N' < p < 2N$, each factor of $\theta(x_N)$ with period p has length at most $(2 - 1/N')p$; for $N' < p \leq 2N'$ this says that $\theta(x_N)$ has no factor udu with $|u| = p - 1$ and d a letter. The weaker statement that $\theta(x_N)$ has no factor udu with $N - 1 \leq |u| \leq 2N - 2$ is, by the same argument, equivalent to $\text{ce}(\sigma_0(x_N)) < \text{ce}(x_N)$, that is, to (x_N, σ_0) being a certificate for a non-empty interval; one of the independent verification runs checked it numerically for $N \leq 256$. It does not suffice for the chain. We have no proof of either statement for all N . Both are statements about factors of words derived from the Thue–Morse word, and statements of this kind can often be decided with the automatic theorem prover Walnut [Mou16, Sha22a]; a proof along these lines might close the window, but we have not tried.

Remark 6.3. The same computation for $N = 98304$ and $N = 131072$ (the next two sizes) gives $\text{ce}(x_N) = 2 - 1/N$ and $\text{ce}(\sigma_0(x_N)) = 2 - 1/N'$ again, with the preceding size N' , which would shrink the window to $(2 - 2^{-17}, 2)$. These two values were found by one of the verification runs and are reproduced by the program of this note and by the last verification run (Section 7); the statements above keep the bound $2 - 2^{-16}$, the range covered by all verification runs.

Remark 6.4. By Dejean’s theorem [Dej72], only finitely many ternary words have critical exponent less than $\frac{7}{4}$. The program with which the results were first obtained enumerates them (533 words up to renaming of the letters, the longest of length 38) together with all their factorisations into at most three distinct blocks (168,651), and finds that for three letters and $1 < \alpha \leq \frac{7}{4}$ the ruled-out values are exactly $(\frac{4}{3}, \frac{7}{4}]$, in agreement with Proposition 4.1(b) and Proposition 6.1(a). This exhaustive classification is a cross-check and is not used in any proof.

7 Computations, verification and scope

Computations. For a word w and a period p , the longest factor with period p has length $p + r_p$, where r_p is the longest run of consecutive indices i with $w_i = w_{i+p}$; so $\text{ce}(w) = \max(1, \max_p(p + r_p)/p)$, an exact rational number computed with integers only. The statements that depend on such computations are Lemmas 5.2 to 5.4, the certificate for $(\frac{9}{4}, \frac{8}{3}]$ in the proof of Proposition 5.6 (small enough to be checked directly), and the 29 values $\text{ce}(\sigma_0(x_{N_i}))$ of Proposition 6.1(b), the largest for a word of length 262,141. Each is a finite statement, reproduced by three independent implementations in the verification runs described below, besides the programs of the note. The infinite families (Lemmas 4.2, 5.1 and 5.5 and the propositions of Sections 3 and 4) are proved in the text; the programs only test them on instances. The program `check_note.py` of the reproducibility package (standard library only) repeats all finite checks of the note in exact arithmetic, tests its critical-exponent routine against a direct loop and against the definition of an exact p/q -power, and tests the propositions on many parameters; its default run takes a few seconds, the run with all 29 sizes about two minutes, and the run of Remark 6.3 about seven minutes. The package also contains the programs with which the results were first obtained and those of the three verification runs, with their recorded outputs.

Verification. The results were first obtained, with written proofs and programs, in a first written version. Two independent verification runs, both AI-assisted, examined this version; each read the source [Sha10] in the publisher’s open PDF, re-derived the proofs line by line and wrote its own programs, and neither read the other’s report. Both found no mathematical error. The first run checked with its own code Lemmas 5.2 and 5.3, Lemma 5.1 for $N \leq 3072$, the bound of Lemmas 5.4 and 5.5 for every word of $\mathcal{F}$ with $N \leq 192$, the formulas after Proposition 5.6, the interval unions, all 29 certificates of Proposition 6.1(b), and Proposition 3.1 and Section 4 on many parameters (up to 1002 letters). The second run wrote a C program for the exact critical exponent, compared it with the definition on 3,006 words, and a second one for the bound on $M^{\mathbb{Z}}$ by cyclic runs; it checked Lemmas 5.2 and 5.3, Lemma 5.1 and the cyclic bound for $N \leq 2048$, the exact critical exponents in Proposition 5.6 for $N \leq 512$, the 29 certificates and the two of Remark 6.3, and the propositions of Sections 3 and 4 on many parameters; it also ran its checks on deliberately wrong inputs, which were detected, and reproduced the recorded outputs of the first-version programs.

The two runs required corrections, all of which were made: both readings are stated wherever the status is described; Lemma 5.2 is stated for all nine pairs, so that its consequence does not depend on Lemma 5.1; the proof of Lemma 5.4 explains why the finite computation suffices; in Lemma 5.5 the case without admissible t gives $L \leq p + 22$; the page range of [Sha10] and the description of the literature search were corrected. In writing this note, Remark 6.2 was corrected: the condition stated in the first version gives only $\text{ce}(\sigma_0(x_N)) < \text{ce}(x_N)$, not the chain. The program `check_note.py` was written for this note and is a third implementation of the finite checks.

The text of this note was then examined by a further independent verification run, also AI-assisted. It downloaded and read the source again, re-derived every proof of the note line by line, and wrote new programs (standard library only, exact integer arithmetic, with a bit-parallel computation of critical exponents that differs from the routines above), which it compared with a computation over all factors on 1,517 words and with the definition of an exact p/q -power on 16,263 pairs of a word and an exponent. With them it recomputed Lemmas 5.2 and 5.3, Lemma 5.4 and the same cyclic bound for every word of $\mathcal{F}$ with $N \leq 3072$, Lemma 5.1 for $N \leq 6144$, the certificate for $(\frac{9}{4}, \frac{8}{3}]$, the formulas after Proposition 5.6, the certificates of Proposition 6.1(a), all 29 certificates of Proposition 6.1(b) and the two of Remark 6.3, the identities of Remark 6.2, the interval unions and the assembly of Theorem 1.3(c) for $3 \leq s \leq 40$, the propositions of Sections 3 and 4 on many parameters (up to 10,001 letters) and the counts of Remark 6.4; all 17 deliberately corrupted inputs were detected. It also re-ran every program of the package from the archive of the previous version; all 20 recorded outputs were reproduced up to timing. It found no error in a statement and required corrections of the text, which were made: in the proof of Proposition 4.1(b) a false intermediate sentence (a factor matching xx need not begin and end with the same letter) was replaced; the discussion of the two readings now notes that the hope to cover all $\alpha > 1$ can come true only in reading (U); the abstract no longer calls the ideas of Currie and Cassaigne unpublished; the caveat on Theorem 1.2 is repeated after its proof; and the descriptions of the verification and of the searches were completed.

Scope and priority. The question is Shallit’s [Sha10, Problem 38]. The note answers it negatively in reading (U) (Theorem 1.2) and, for a fixed alphabet, completely for one and two letters and for $s \geq 3$ letters outside the window $(2 - 2^{-16}, 2)$ (Theorem 1.3); the window stays open. The argument for Theorem 1.2 is elementary. The remark after the problem in [Sha10] credits Currie and Cassaigne with ideas showing impossibility for restricted ranges of α ; we found no write-up of their arguments, Theorem 1.2 may coincide with them or be folklore, and we claim no priority for it. The tools are classical: the overlap-freeness of the Thue–Morse word [Thu06] and Thue’s square-free ternary word obtained from it by counting letters between consecutive zeros [Thu12] (we follow the account in Berstel’s survey [Ber] for which result is in which paper; Thue’s papers were not read),

the desubstitution of Lemma 4.2, proved here and standard (compare [AS03, Lot97], which were not consulted for this note), the theorem of Fine and Wilf [FW65], and, for Remark 6.4 only, Dejean’s theorem [Dej72].

What was read: Shallit’s abstract [Sha10] (pp. 2230–2236 of the report) in full, in the open PDF of the publisher, and the parts of [Ber] that describe Thue’s results, in a text extraction; both were read again by the last verification run. The searches were made on 19 and 20 October 2026: queries to arXiv, zbMATH, Crossref, OpenAlex and Semantic Scholar (for patterns and fractional powers, for Shallit’s papers on patterns, for nonerasing morphisms and powers, for Currie, Cassaigne and patterns, for critical exponents and closure under morphisms), the lists of works citing [Sha10] in OpenAlex and Semantic Scholar, and eight web searches; the last verification run repeated part of them. Part of these requests were refused by arXiv, OpenAlex and Semantic Scholar with HTTP status 429 (rate limit) or 503 and were repeated at most once. Semantic Scholar lists two works citing [Sha10], [FLM⁺20] and [RSW22]; OpenAlex lists [FLM⁺20] and a 2026 preprint on an antimorphic Fine–Wilf theorem [Kri26]. Their titles, and the abstracts of [FLM⁺20] and [RSW22], show other topics. The other works found concern the avoidance of patterns and fractional powers, not their characterisation [CM21, RS12, PR18, Sha22b, LSW00]. No printed answer to Problem 1.1 was found in the sources searched; a search that finds nothing is not a proof of novelty, and no priority is claimed.

AI-assisted tools were used to search the problem corpus and the literature, to find the arguments, to write and run the verification programs, and to help prepare this manuscript. The author has reviewed the proofs, the computations and the sources and is responsible for all claims. The note has not been peer reviewed.

References

- [AS03] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, Cambridge, 2003.
doi:10.1017/CB09780511546563.
- [Ber] Jean Berstel. Wörter ohne Wiederholungen. Séminaire Lotharingien de Combinatoire, archive of early papers. Survey; Sections 3 and 4 describe the contents of [Thu06, Thu12]. URL: <https://www.mat.univie.ac.at/~slc/opapers/sc10berstel.pdf>.
- [CM21] James D. Currie and Lucas Mol. The undirected repetition threshold and undirected pattern avoidance. *Theoretical Computer Science*, 866:56–69, 2021.
doi:10.1016/j.tcs.2021.03.010.
- [Dej72] Françoise Dejean. Sur un théorème de Thue. *Journal of Combinatorial Theory, Series A*, 13(1):90–99, 1972. doi:10.1016/0097-3165(72)90011-8.
- [FLM⁺20] Pamela Fleischmann, Marie Lejeune, Florin Manea, Dirk Nowotka, and Michel Rigo. Reconstructing words from right-bounded-block words. In *Developments in Language Theory (DLT 2020)*, Lecture Notes in Computer Science, pages 96–109. Springer, 2020. arXiv:2001.11218, doi:10.1007/978-3-030-48516-0_8.
- [FW65] N. J. Fine and H. S. Wilf. Uniqueness theorems for periodic functions. *Proceedings of the American Mathematical Society*, 16(1):109–114, 1965.
doi:10.1090/S0002-9939-1965-0174934-9.
- [Kri26] Alec Kriebel. An alternating antimorphic Fine–Wilf theorem. Zenodo preprint, 2026.
doi:10.5281/zenodo.23130726.

- [Lot97] M. Lothaire. *Combinatorics on Words*. Cambridge University Press, Cambridge, 1997. doi:10.1017/CB09780511566097.
- [LSW00] John Loftus, Jeffrey Shallit, and Ming-wei Wang. New problems of pattern avoidance. In *Developments in Language Theory (DLT 1999)*, pages 185–199. World Scientific, 2000. doi:10.1142/9789812792464_0016.
- [Mou16] Hamoon Mousavi. Automatic theorem proving in Walnut, 2016. arXiv:1603.06017.
- [PR18] Lara Pudwell and Eric Rowland. Avoiding fractional powers over the natural numbers. *Electronic Journal of Combinatorics*, 25(2):P2.27, 2018. doi:10.37236/6678.
- [RS12] Eric Rowland and Jeffrey Shallit. Avoiding $3/2$ -powers over the natural numbers. *Discrete Mathematics*, 312(6):1282–1288, 2012. doi:10.1016/j.disc.2011.12.019.
- [RSW22] Michel Rigo, Manon Stipulanti, and Markus A. Whiteland. Binomial complexities and Parikh-collinear morphisms. In *Developments in Language Theory (DLT 2022)*, Lecture Notes in Computer Science, pages 251–262. Springer, 2022. Extended version: arXiv:2201.04603. doi:10.1007/978-3-031-05578-2_20.
- [Sha10] Jeffrey Shallit. On patterns and pattern avoidance. *Oberwolfach Reports*, 7(3):2230–2236, 2010. Extended abstract in Report No. 37/2010, *Mini-Workshop: Combinatorics on Words* (organisers V. Berthé, J. Karhumäki, D. Nowotka, J. Shallit), pp. 2195–2244; issue published 2011. Problem 38 and the Remark following it are on p. 2231. doi:10.4171/OWR/2010/37.
- [Sha22a] Jeffrey Shallit. *The Logical Approach to Automatic Sequences: Exploring Combinatorics on Words with Walnut*. Cambridge University Press, Cambridge, 2022. doi:10.1017/9781108775267.
- [Sha22b] Jeffrey Shallit. Using automata and a decision procedure to prove results in pattern matching (invited talk). In *33rd Annual Symposium on Combinatorial Pattern Matching (CPM 2022)*, volume 223 of *LIPIcs*, pages 2:1–2:3. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.CPM.2022.2.
- [Thu06] Axel Thue. Über unendliche Zeichenreihen. *Norske Vid. Selsk. Skr. I. Mat.-Nat. Kl. Christiania*, pages 1–22, 1906.
- [Thu12] Axel Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Norske Vid. Selsk. Skr. I. Mat.-Nat. Kl. Christiania*, pages 1–67, 1912.