Journal of Integer Sequences, Vol. 26 (2023),
Article 23.8.2

The Analog of Overlap-Freeness for the Period-Doubling Sequence

James D. Currie
Department of Mathematics and Statistics
The University of Winnipeg
Winnipeg, MB R3B 2E9
Canada
j.currie@uwinnipeg.ca

Abstract

Good words are defined to be binary words avoiding the factors 11 and 1001, and the patterns 0000 and 00010100. We show that good words bear the same relationship to the period-doubling sequence that overlap-free words bear to the Thue-Morse sequence. We prove an analog of Fife's theorem for good words, exhibit the lexicographically least and greatest infinite good words, and determine the patterns avoided by the period doubling word.

1 Introduction

Let A be a finite set. A *word* over *alphabet* A is a finite or infinite sequence over A. We use lower case letters to denote finite words, writing, for example, word $w = w_1 w_2 \cdots w_n$, where each $w_i \in A$. The *length* of w is denoted by $|w| = n$. The *empty word*, of length 0, is denoted by ϵ. A *non-empty word* is a word of positive length. The *concatenation* of two words $u = u_1 u_2 \cdots u_n$ and $v = v_1 v_2 \cdots v_m$ is given by $uv = u_1 u_2 \cdots u_n v_1 v_2 \cdots v_m$. If u, v, w, z are words and $w = uzv$, we say that word u is a *prefix* of w, word v is a *suffix* of w, and word z is a *factor* of w. We say that $\langle u, z, v\rangle$ is an *occurrence of z in w*, and that z *occurs* in w with *index* $|u|$. The set of finite words over A is denoted by A^*.

Let A and B be alphabets. A *morphism* from A^* to B^* is a function respecting concatenation; i.e., $f(xy) = f(x)f(y)$ for all $x, y \in A^*$. Thus f is generated by its values on the elements of A. If $f(x)$ is non-empty whenever x is non-empty, we call f *non-erasing*. Let $p \in A^*$ and $w \in B^*$. We say that w *encounters the pattern* p if we can write $w = uf(p)v$ for some $u, v \in B^*$, and some non-erasing morphism f. Otherwise, we say that w *avoids* p. An *overlap* is a word v such that we can write $v = xyxyx$ for words x and y where x is non-empty. If no factor of w is an overlap, then w is *overlap-free*. A word is overlap-free if and only if it avoids patterns xxx and $xyxyx$. If x is non-empty and k is a positive integer, we denote by x^k the word consisting of x repeated k times in a row. A *fourth power* is a word x^4 where $x \neq \epsilon$. The *reversal* of finite word $w = w_1 w_2 \cdots w_n$ is the word $w_n w_{n-1} \cdots w_1$. A word over $\{0, 1\}$ is called a *binary word*. The complement of a binary word w is obtained by replacing 0's with 1's and vice versa.

Example 1. Consider the word $u = \alpha\beta\beta\gamma\beta\beta\gamma\beta\alpha\alpha \in \{\alpha, \beta, \gamma\}^*$. Then u has the factor $\beta\beta\gamma$. There are two occurrences of $\beta\beta\gamma$ in u, namely $\langle \alpha, \beta\beta\gamma, \beta\beta\gamma\beta\alpha\alpha \rangle$ and $\langle \alpha\beta\beta\gamma, \beta\beta\gamma, \beta\alpha\alpha \rangle$, so that $\beta\beta\gamma$ occurs in u with index 1, and with index 4. Factor $\beta\beta\gamma\beta\beta\gamma\beta$ of u is an overlap, with $x = \beta$, $y = \beta\gamma$. Thus u is not overlap-free.

The word $v = 0101010101 = (01)^5$ contains overlaps 01010 and 10101, and fourth powers $(01)^4$ and $(10)^4$. The reversal of v is 1010101010, which is also the complement of v.

The word *bananas* can be written $bananas = uf(pear)v$, where $u = ba$, $f(p) = f(a) = n$, $f(e) = a$, $f(r) = as$, $v = \epsilon$. Thus *bananas* encounters the pattern *pear*. It also encounters *apple*, letting $u = v = \epsilon$, $f(a) = b$, $f(p) = an$, $f(l) = a$, $f(e) = s$. However, *apple* avoids *bananas* (since the image of *bananas* under any non-erasing morphism has length 7 or more, whereas $|apple| = 5$).

We use bold-face letters for infinite words, writing $\boldsymbol{w} = w_1 w_2 w_3 \cdots$, where each $w_i \in A$. Thus we are concerned with infinite words where the domain is the positive integers, and we refer to them as *one-way infinite words* (in contrast to *two-way infinite words* where the domain would be the set of all integers). The set of finite words over A is denoted by A^*, and the set of one-way infinite words is denoted by A^ω.

Iteration of a morphism f is denoted by exponentiation:

$$f^i(x) = \begin{cases} x, & \text{if } i = 0; \\ f(f^{i-1}(x)), & \text{if } i > 0. \end{cases}$$

If $f : A^* \to A^*$ is a morphism such that for some $a \in A$, $|f(a)| > 1$ and the first letter of $f(a)$ is a, then $f^{n-1}(a)$ is a prefix of $f^n(a)$ for every positive integer n. We can then define $\boldsymbol{w} = \lim_{n\to\infty} f^n(a)$ to be the unique one-way infinite word such that for each n, word $f^n(a)$ is a prefix of $\boldsymbol{w}$; thus $\boldsymbol{w}$ is a fixed point of f.

The famous Thue-Morse sequence $\mathbf{t}$ is a fixed point of the binary morphism μ given by $\mu(0) = 01$, $\mu(1) = 10$, namely

$$\mathbf{t} = \lim_{n\to\infty} \mu^n(0).$$

It is sequence A010060 in the On-Line Encyclopedia of Integer Sequences (OEIS) [16].

Thue [18] introduced $\mathbf{t}$ and proved that it is overlap-free. The sequence is also nascent in an earlier paper of Thue [17], where it could be obtained by applying Satz 6 to Satz 3. It is also implicit in an early memoir of Prouhet [12] on multigrades.

Theorem 2. *Let w be an overlap-free binary word. Then $\mu(w)$ is overlap-free.*

He also showed that, in the case of two-sided infinite words and circular words, every overlap-free binary word is the image under μ of an overlap-free word [18]. The analysis of words with 'ends' is more complicated, but finite overlap-free binary words also arise via iterating μ. (See Restivo and Salemi [14] for example.)

Theorem 3. *Let $w \in \{0,1\}^*$ be a finite overlap-free word. Then we can write $w = a\mu(u)b$, where $a, b \in \{\epsilon, 0, 00, 1, 11\}$, and u is overlap-free. If $|w| \geq 7$ this factorization is unique. If $\boldsymbol{w}$ is a one-sided infinite overlap-free word, then we can write $\boldsymbol{w} = a\mu(\boldsymbol{u})$, for some one-sided infinite overlap-free word $\boldsymbol{u}$ where $a \in \{\epsilon, 0, 00, 1, 11\}$.*

Restivo and Salemi used their version of Theorem 3 to give a rough enumeration of binary overlap-free words. Kobayashi [10] gave a better enumeration, obtaining a good lower bound by counting finite words which extend to infinite overlap-free words. For this, he used the deep theorem of Fife [7] characterizing the infinite overlap-free words. The problem of enumerating binary overlap-free words was finally completely solved by Jungers, Protasov, and Blondel [9], and by Guglielmi and Protasov [8].

Because of Theorem 3, the word $\mathbf{t}$ turns up frequently in the study of overlap-free binary words. An example is the following result of Berstel [1] (later greatly generalized by Allouche et al. [2]):

Theorem 4. *The lexicographically greatest one-sided infinite overlap-free binary word starting with 0 is $\mathbf{t}$.*

The relationship between $\mathbf{t}$ and μ was also key to establishing the following theorem of Shur [15]:

Theorem 5. *Suppose $\mathbf{t}$ encounters a pattern $p \in \{0,1\}^*$. Then either p is a factor of $\mathbf{t}$, or p is one of 00100 and 11011.*

Another famous binary sequence is the period-doubling sequence, which is the fixed point

$$\boldsymbol{d} = 0100010101000100010001010100\,0101\cdots$$

of the binary morphism δ where $\delta(0) = 01$, $\delta(1) = 00$. This sequence has been much studied in the context of quasi-crystal spectral theory. (See Damanik [6], for example.) It is sequence A096268 in the On-Line Encyclopedia of Integer Sequences (OEIS) [16].

Call a binary word w *good* if it does not contain the factor 11 or 1001, and does not encounter either of the patterns 0000 or 00010100. We will show that $\boldsymbol{d}$ is good. In fact, we show that the period doubling morphism δ has the same relationship to good words that μ has to overlap-free words, namely

Theorem 6. *Suppose w is good. Then $\delta(w)$ is good.*

Theorem 7. *Let w be a finite good word. Then we can write $w = a\delta(u)b$, where $a \in \{\epsilon, 0, 1\}$, $b \in \{\epsilon, 0\}$, and u is good. If $|w| \geq 4$ this factorization is unique. If $\boldsymbol{w}$ is a one-sided infinite good word, then we can write $\boldsymbol{w} = a\delta(\boldsymbol{u})$, for a one-sided infinite good word $\boldsymbol{u}$ where $a \in \{\epsilon, 0, 1\}$.*

We build on these theorems to

- Give a version of Fife's theorem for good words, characterizing infinite good words;
- Exhibit lexicographically extremal one-sided infinite good words;
- Characterize the binary patterns avoided by $\boldsymbol{d}$.

2 Good words

Unless otherwise specified, our words and morphisms are over the binary alphabet $\{0, 1\}$. We record morphisms inline, i.e., $g = [g(0), g(1)]$.

Lemma 8. *Let u be a finite binary word. Suppose $\delta(u)$ is good. Then u is good.*

Proof. If u encounters pattern 0000 or 00010100, so does $\delta(u)$. If u has factor 11 or 1001, then $\delta(u)$ has factor $\delta(11) = 0000$ or $\delta(1001) = 00010100$, and thus encounters pattern 0000 or 00010100. □

Remark 9. Let $w = w_1 w_2 w_3 \cdots w_{2n}$ with $w_i \in \{0, 1\}$. Word w can be written as $w = \delta(u)$ for some u if and only if $w_i = 0$ for each odd index i.

Lemma 10. *Let w be a finite binary word with no factor 11, 1001, or 0000. Then we can write $w = a\delta(u)b$ where $a \in \{\epsilon, 0, 1\}$ and $b \in \{\epsilon, 0\}$. If $|w| \geq 4$ this factorization is unique. If $\boldsymbol{w}$ is a one-sided infinite word with no factor 11, 1001, or 0000, then we can write $\boldsymbol{w} = a\delta(\boldsymbol{u})$, some one-sided infinite word $\boldsymbol{u}$ where $a \in \{\epsilon, 0, 1\}$.*

Proof. First we demonstrate the existence of the factorization for finite words: If $|w|_1 \leq 1$ then w is a factor of 0001000, and the result is established by a finite check. Suppose $|w|_1 \geq 2$. If 10^k1 is a factor of w, the conditions on w force $k = 1$ or $k = 3$. By induction, if $1u1$ is a factor of w then $|1u|$ is even; therefore, all the 1's in w have an index in w of the same parity. If the parity of the indices of 1's is even, let $|a| = 0$; let $|b|$ be 0 (resp., 1) if $|w|$ is even (resp., odd). If the parity of the indices of 1's is odd, let $|a| = 1$; let $|b|$ be 1 (resp., 0) if $|w|$ is even (resp., odd). By Remark 9 we can write $a^{-1}wb^{-1} = \delta(u)$ for some u.

We have shown that we can write $w = a\delta(u)b$ and $a, b \in \{\epsilon, 0, 1\}$. Suppose that $b = 1$. Recall that $|w|_1 \geq 2$. The parity of the indices of a and b is different, so we cannot have $a = 1$. It follows that $|\delta(u)|_1 \geq 1$. Then $\delta(u)1$ must have suffix 011, 01001, or 00001, none of which are good. This is a contradiction so in fact $b \in \{\epsilon, 0\}$.

Now we show the uniqueness of the factorization for finite words: Suppose $|w| \geq 4$ and w has two factorizations $w = a_1\delta(u_1)b_1 = a_2\delta(u_2)b_2$. If $|a_1| = |a_2|$ we are forced to choose $a_1 = a_2$, $u_1 = u_2$, and $b_1 = b_2$, so the factorizations are identical. Suppose without loss of generality then that $|a_1| = 0$, $|a_2| = 1$. Then by Remark 9, every 1 in $a_1\delta(u_1)b_1$ has even index, but every 1 in $a_2\delta(u_2)b_2$ has even index. Since $a_1\delta(u_1)b_1 = w = a_2\delta(u_2)b_2$, we conclude that $|w|_1 = 0$, so that w has 0000 as a prefix, which is impossible.

Now suppose that $\boldsymbol{w}$ is a one-sided infinite good word. For each non-negative n, let the length n prefix of $\boldsymbol{w}$ be p_n. We have proved that finite words can be factored, so write each $p_n = a_n\delta(u_n)b_n$ where $a_n, b_n \in \{\epsilon, 0, 1\}$. Word p_4 cannot be 0000, so that $|w_4|_1 > 0$. By Remark 9, the index of the first 1 in w_4 determines a_4 and all subsequent a_n, so that $a_n = a_4$ for $n \geq 4$. This implies that $\delta(u_n)$ is a prefix of $\delta(u_{n+1})$ for $n \geq 4$, so that u_{n+4} is a prefix of u_{n+5} for all n. Let $\boldsymbol{u} = \lim_{n\to\infty} u_{n+4}$. Then

$$\begin{aligned}
\boldsymbol{w} &= \lim_{n\to\infty} w_n \\
&= \lim_{n\to\infty} w_{n+4} \\
&= \lim_{n\to\infty} w_{n+4}b_{n+4}^{-1} \\
&= \lim_{n\to\infty} a_4\delta(u_{n+4}) \\
&= a_4 \lim_{n\to\infty} \delta(u_{n+4}) \\
&= a_4\delta(\lim_{n\to\infty} u_{n+4}) \\
&= a_4\delta(\boldsymbol{u}).
\end{aligned}$$

□

Proof of Theorem 7. This is immediate from Lemmas 8 and 10. □

Call a non-erasing morphism g *even* if, for every letter u, $|g(u)|$ is even.

Lemma 11. *Let p be a pattern and let w be a binary word. Suppose that $g(p)$ is a factor of $\delta(w)$ where g is an even morphism. Then w encounters pattern p.*

Proof. Write $p = u_1 \cdots u_n$ with the u_i letters. Write $\delta(w) = aU_1 \cdots U_nb$, where $U_i = g(u_i)$ for each i. If $|a|$ is even, $w = \delta^{-1}(a)\delta^{-1}(U_1 \cdots U_n)\delta^{-1}(b)$, and w contains the instance $\delta^{-1}(g(p))$ of p.

If $|a|$ is odd, then 0 is the last letter of a and of each U_i. Thus

$$\delta(w) = a0^{-1}0U_10^{-1}0U_20^{-1}0 \cdots U_n0^{-1}0b,$$

and w contains the instance $\delta^{-1}(h(p))$ of p, where h is the morphism defined on the letters of p by $h(u_i) = 0g(u_i)0^{-1}$. □

Lemma 12. *Let w, u, v be binary words and suppose that uvu is a factor of $\delta(w)$. If $|u|_1 > 0$ then $|uv|$ is even.*

Proof. Since uvu has period $|uv|$, $\delta(w)$ has a factor $1z1$ where $|1z| = |uv|$. The Lemma follows by Remark 9. $\square$

Proof of Theorem 6. To begin with, we show that $\delta(w)$ does not contain 11, 0000, 1001, or 00010100 as a factor. By Remark 9, $\delta(w)$ does not have a factor 11 or 1001. If $\delta(w)$ has factor 0000, write $\delta(w) = a0000b$ for words a and b. If $|a|$ is even, then w has prefix $\delta^{-1}(a0000)$, which ends in 11. This is impossible; if $|a|$ is odd, then the last letter of a is 0, so that w has prefix $\delta^{-1}((a0^{-1})0000)$, which again ends in 11. Finally, if 00010100 is a factor of $\delta(w)$, then w contains factor $\delta^{-1}(00010100) = 1001$, which is impossible. (The index of 00010100 in $\delta(w)$ must be odd by Remark 9)

Suppose now that $\delta(w)$ encounters pattern $p = 0000$, so that $XXXX$ is a factor of $\delta(w)$ for some non-empty X. Since $\delta(w)$ does not have 0000 as a factor, we must have $|X|_1 > 0$. Using $u = X$ and $v = \epsilon$ in Lemma 12, we conclude that $|X|$ is even. Then Lemma 11 implies that w encounters 0000, which is a contradiction.

Suppose that $\delta(w)$ encounters pattern $p = 00010100$, so that X^3YXYXX is a factor of $\delta(w)$ for some non-empty X and Y. Suppose that $|X|_1 > 0$. Since XX is a factor of $\delta(w)$, letting $u = X$ and $v = \epsilon$ in Lemma 12 implies that $|X|$ is even. Again, since XYX is a factor of $\delta(w)$, letting $u = X$ and $v = Y$ in Lemma 12 implies that $|XY|$ is even. It follows that $|Y|$ is even. Then Lemma 11 implies that w encounters 00010100, which is a contradiction. We therefore conclude that $|X|_1 = 0$, so that $X = 0^n$ for some $n \geq 1$. Since $XX = 0^{2n}$ is a factor of $\delta(w)$, but 0000 is not, we conclude that $n = 1$ and $X = 0$. Thus $\delta(w)$ contains the factor $XXXYXYXX = 000Y0Y00$.

Since 0000 is not a factor of $\delta(w)$, the first letter of Y is 1. If $Y = 1$, then $\delta(w)$ has factor 00010100, which is impossible. Therefore $|Y| > 1$.

Suppose Y ends in 0. If Y ends in 10, then $\delta(w)$ contains factor 1001 (inside $Y0Y$), which is impossible. If $|Y|$ ends in 00, then $000Y0Y00$ ends in $Y00$, hence 0000, which is again impossible. Thus Y ends in 1, hence in 01. Write $Y = 1Z1$ for some non-empty word Z. Word $\delta(w)$ has the factor $0001Z101Z100$.

If $|1Z1| = 3$, then $\delta(w)$ contains $01Z101Z1 = 01010101$, an instance of 0000, already proved impossible. It follows that $|1Z1| \geq 4$. If Z begins 01, then $Z10$ begins either 0101 or 0100. However, if $Z10$ begins 0101 then, since Z ends in 0, $Z101Z10$ again contains 01010101; if $Z10$ begins 0100 then $0001Z10$ begins 00010100, which is not a factor of $\delta(w)$. We conclude that Z does not begin 01 and therefore begins 00.

If $Z10$ begins 001 then $1Z10$ has the impossible factor 1001. Thus $Z10$ begins 0001. We now consider suffixes of Z. If Z ends 10, then $01Z$ ends either 1010 or 0010. However, if $01Z$ ends 1010 then, since Z begins with 0, $01Z101Z$ contains 10101010, an instance of 0000, which is impossible; if $01Z$ ends in 0010 then $101Z100$ ends 00010100, which is not a factor of $\delta(w)$. We conclude that Z does not end 10 and therefore ends 00. If $01Z$ ends 0100 then $\delta(w)$ contains $01Z1$, hence the impossible factor 1001. Thus $01Z$ ends in 1000. Since Z begins 00, this forces $01Z101Z$ to contain 00010100, which is impossible. $\square$

Corollary 13. *The period-doubling word $\boldsymbol{d}$ is good.*

Theorem 14. *Let* $\mathbf{w}$ *be a one-sided infinite binary word. Then* $\boldsymbol{w}$ *is good if and only if* $\delta(\boldsymbol{w})$ *is good.*

Proof. If w_n is the length n prefix of $\mathbf{w}$ then $\mathbf{w} = \lim_{n\to\infty} w_n$ and $\delta(\mathbf{w}) = \lim_{n\to\infty} \delta(w_n)$. By Theorem 6 and Lemma 8, w_n is good if and only if $\delta(w_n)$ is good.

$$\begin{aligned} \mathbf{w} \text{ is good } &\iff \text{each } w_n \text{ is good} \\ &\iff \text{each } \delta(w_n) \text{ is good} \\ &\iff \lim_{n\to\infty} \delta(w_n) = \delta(\lim_{n\to\infty} w_n) = \delta(\mathbf{w}) \text{ is good.} \end{aligned}$$

□

Remark 15. While the set of binary overlap-free words is closed under complementation and reversal, the same is not true of good words. For example, 00101000 is good, but neither its complement nor its reversal is good.

3 An analog of Fife's theorem

We characterize the one-sided infinite good words, developing a theory analogous to Fife's [7] theory for overlap-free binary words; however, we follow Rampersad's [13] exposition of Fife, with appropriate modifications, rather than Fife's original paper. Let G be the set of one-sided infinite good words.

Lemma 16. *Suppose* $\boldsymbol{u}$ *is a one-sided infinite good word. Then* $1\boldsymbol{u}$ *is good if and only if* $01\boldsymbol{u}$ *is good.*

Proof. Clearly if $01\boldsymbol{u}$ is good then $1\boldsymbol{u}$ is good. Suppose that $1\boldsymbol{u}$ is good but $01\boldsymbol{u}$ is not. By Theorem 7, write $1\boldsymbol{u} = 1\delta(\boldsymbol{v})$ for some $\boldsymbol{v}$. Since $1\boldsymbol{u}$ is good, $01\boldsymbol{u}$ must have a prefix which is either factor 11 or 1001, or a pattern instance $g(0000)$ or $g(00010100)$ where $g = [X, Y]$ is some non-erasing morphism. Clearly neither of 11 and 1001 can be a prefix of $01\boldsymbol{u}$, so $01\boldsymbol{u}$ has a prefix $XXXX$ or $XXXYXYXX$ where X and Y are non-empty.

Suppose $01\boldsymbol{u}$ has prefix $XXXX$. Write $X = 0X'$. Since $|XXXX|$ is even, word $01\boldsymbol{u} = \delta(0\boldsymbol{v})$ has $XXXX0 = 0X'0X'0X'0X'0$ as a prefix. But now the good word $1\boldsymbol{u}$ contains the fourth power $X'0X'0X'0X'0$, which is a contradiction.

Now suppose that $01\boldsymbol{u}$ has prefix $XXXYXYXX$. Write $X = 0X'$. If $X' = \epsilon$, then $X = 0$. However then the length 2 prefix of $01\boldsymbol{u}$ is $XX = 00$, which is impossible. Thus $X' \neq \epsilon$, forcing 01 to be a prefix of X, so that $|X|_1 > 0$. Since $01\boldsymbol{u} = \delta(0\boldsymbol{v})$, considering the second XX in $XXXYXYXX$, by Lemma 12 with $u = X$, $v = \epsilon$ we find that $|X|$ is even. Again, XYX is a factor of $\delta(0\boldsymbol{v})$, so applying Lemma 12 with $u = X$, $v = Y$ shows that $|XY|$ is also even. Since both $|X|$ and $|XY|$ are even, $|Y|$ is even. Since $|XXX|$ is even, $XXX0$ is a prefix of $01\boldsymbol{u} = \delta(0\boldsymbol{v})$, so that we can write $Y = 0Y'$. Since $|XXXYXYXX|$ is even, word $01\boldsymbol{u} = \delta(0\boldsymbol{v})$ has $XXXYXYXX0 = 0X'0X'0X'0Y'0X'0Y'0X'0X'0$ as a prefix. But now the good word $1\boldsymbol{u}$ contains $X'0X'0X'0Y'0X'0Y'0X'0X'0 = g'(00010100)$ where $g' = [X'0, Y'0]$. This is a contradiction. □

Remark 17. This result uses the fact that $\boldsymbol{u}$ is one-sided infinite. If $u = 010101$, then $1u$ is good, but $01u = (01)^4$ is not.

Let G be the set of one-sided infinite good words. For $w \in \{0,1\}^*$, let $G_w = G \cap w\{0,1\}^\omega$.

Lemma 18. *Let* $\boldsymbol{w}$ *be a one-sided infinite binary word.*

(a) $\delta(\boldsymbol{w}) \in G \iff \boldsymbol{w} \in G$*;*

(b) $1\delta(\boldsymbol{w}) \in G \iff 0\boldsymbol{w} \in G$*;*

(c) $0\delta(\boldsymbol{w}) \in G \iff (1\boldsymbol{w} \in G)$ *or* $(\boldsymbol{w} \in G_{001})$*.*

Remark 19. The cases in (c) are disjoint, since if 001 is a prefix of $\mathbf{w}$, then $1\mathbf{w}$ has prefix 1001 and is not good.

Proof of (a). This is just Theorem 14. □

Proof of (b). If $0\boldsymbol{w} \in G$, then by Theorem 14, $\delta(0\boldsymbol{w}) = 01\delta(\boldsymbol{w}) \in G$, so in particular $1\delta(\boldsymbol{w}) \in G$.

In the other direction, suppose $1\delta(\boldsymbol{w}) \in G$. Applying Lemma 16 to prefixes gives $01\delta(\boldsymbol{w}) = \delta(0\boldsymbol{w}) \in G$. By Theorem 14, $0\boldsymbol{w} \in G$. □

Proof of (c). If $1\boldsymbol{w} \in G$, then by Theorem 14 $\delta(1\boldsymbol{w}) = 00\delta(\boldsymbol{w}) \in G$, so in particular $0\delta(\boldsymbol{w}) \in G$.

Suppose $\boldsymbol{w} \in G_{001}$. By Theorem 14, $\delta(\boldsymbol{w}) \in G$. Suppose $0\delta(\boldsymbol{w}) \notin G$. It, therefore, has a prefix $XXXX$ or $XXXYXYXX$ where X is non-empty. Since 001 is a prefix of $\mathbf{w}$, $p = 0010100$ is a prefix of $0\delta(\mathbf{w})$. The prefix XXX of $0\delta(\mathbf{w})$ has period $|X|$. If $|X| < |p|$, then $|X|$ is a period of p, which has least period 5. We conclude that $|X| \geq 5$. This implies that, $|X|_1 > 0$. The second XX in XXX is a factor of $\delta(\mathbf{w})$. By Lemma 12 with $u = X$, and $v = \epsilon$, we conclude that $|X|$ is even. Therefore $|X| \geq 6$. Since $0^{-1}X$ is an odd-length prefix of $\delta(\boldsymbol{w})$, 0 is a suffix of X.

If $|X| = 6$ then the prefix $0^{-1}XXX$ of $\delta(\mathbf{w})$ ends in $0XX = 0\ 001010\ 001010$, which has prefix 00010100. This is impossible, since $\delta(\mathbf{w}) \in G$. If $|X| > 6$, then $p = 0010100$ is a prefix of X, and 0 is a suffix of X, so factor XX of $\delta(\mathbf{w})$ contains 0 0010100, again an impossibility.

In the other direction, suppose that $0\delta(\mathbf{w}) \in G$. Then $\boldsymbol{w} \in G$ by Theorem 14, and we show that either $1\boldsymbol{w} \in G$, or $\boldsymbol{w} \in 001\{0,1\}^*$. Suppose that $1\boldsymbol{w} \notin G$ and 001 is not a prefix of $\mathbf{w}$. It follows that a prefix of $1\boldsymbol{w}$ has the form 11, 1001, $XXXX$, or $XXXYXYXX$ with non-empty X and/or Y.

By Theorem 7, write $\boldsymbol{w} = a\delta(\boldsymbol{u})$ where $a \in \{\epsilon, 0, 1\}$.

If 11 is a prefix of $1\boldsymbol{w}$, then $a = 1$, and 10 is a prefix of $\boldsymbol{w}$. However, then $0\delta(\boldsymbol{w})$ has prefix $0\delta(10) = 00001$, and is not good.

If 1001 is a prefix of $1\boldsymbol{w}$, then 001 is a prefix of $\mathbf{w}$, which is a contradiction.

Suppose $XXXX$ is a non-empty prefix of $1\boldsymbol{w}$. Then the final XX of $XXXX$ is a factor of $\delta(\boldsymbol{u})$ and $|X|_1 > 0$, so by Lemma 12, $|X|$ is even. Since X is a prefix of $1\boldsymbol{w}$, the first letter of X is 1. Since XX is a factor of $\boldsymbol{w}$, and $\boldsymbol{w}$ has no factor 11, the last letter of X is 0. Write $X = 1X'0$.

Word $\boldsymbol{w}$ has prefix $X'01X'01X'01X'0c$ for some $c \in \{0,1\}$. This implies that $0\delta(\boldsymbol{w})$ has prefix

$$\begin{aligned}0\delta(X'01X'01X'01X'0c) &= 0\delta(X'0)\delta(1)\delta(X'0)\delta(1)\delta(X'0)\delta(1)\delta(X'0)\delta(c)\\ &= 0\delta(X'0)00\delta(X'0)00\delta(X'0)00\delta(X'0)0\overline{c}\\ &= (0\delta(X'0)0)^4\overline{c},\end{aligned}$$

and $0\delta(\boldsymbol{w})$ contains a fourth power. This is a contradiction, since $0\delta(\boldsymbol{w}) \in G$.

Suppose $XXXYXYXX$ is a prefix of $1\boldsymbol{w}$. This means that XX and XYX are factors of $\delta(\boldsymbol{u})$, and $|X|_1 > 0$. We conclude by Lemma 12 that $|X|$ and $|Y|$ are even. The first letter of X is 1. Since XX and YX are factors of $\boldsymbol{w}$, and $\boldsymbol{w}$ has no factor 11, the last letter of each of X and Y is 0. Write $X = 1X'0$ and $Y = cY'0$ where $c \in \{0,1\}$. Since $X'01$ is an even length prefix of $a\delta(\boldsymbol{u})$, Remark 9 forces $a = \epsilon$.

Then $\boldsymbol{w}$ has prefix $X'01X'01X'01Y'01X'01Y'01X'01X'0d$ for some $d \in \{0,1\}$. Therefore $0\delta(\boldsymbol{w})$ has prefix

$$\begin{aligned}&0\delta(X'01X'01X'0cY'01X'0cY'01X'01X'0d)\\ &= 0X''0X''0X''\overline{c}Y''0X''\overline{c}Y''0X''0X''\overline{d}\\ &= h(00010100)\overline{d},\end{aligned}$$

where $X'' = \delta(X')010$, $Y'' = \delta(Y')010$, and $h = [0\delta(X')010, \overline{c}\delta(Y')010]$. This is a contradiction, since $0\delta(\boldsymbol{w}) \in G$. □

Suppose $w \in \{0,1\}^*$ has a suffix $\delta^n(01)$, $n \geq 0$, and let n be as large as possible. Write $w = y\delta^n(01)$. Define mappings α, β and γ on w by

$$\begin{aligned}\alpha(w) &= w\delta^n(00) = y\delta^{n+1}(01)\\ \beta(w) &= w\delta^n(0100) = y\delta^{n+1}(001)\\ \gamma(w) &= w\delta^n(010100) = y\delta^{n+1}(0001).\end{aligned}$$

For example, if $w = 0001000101$, then $y = 00$, $n = 2$, $\delta^n(0) = 0100$, $\delta^n(1) = 0101$, so that

$$\begin{aligned}\alpha(w) &= 00\ 0100\ 0101\ 0100\ 0100\\ \beta(w) &= 00\ 0100\ 0101\ 0100\ 0101\ 0100\ 0100\\ \gamma(w) &= 00\ 0100\ 0101\ 0100\ 0101\ 0100\ 0101\ 0100\ 0100.\end{aligned}$$

Let $\boldsymbol{f} = f_1 f_2 f_3 \cdots$, where each $f_i \in B = \{\alpha, \beta, \gamma\}$. Suppose w has some suffix $\delta^n(01)$. Then w is a proper prefix of $f_1(w)$, which is a proper prefix of $f_2(f_1(w))$, which is a proper prefix of $f_3(f_2(f_1(w)))$, etc. We define the infinite composition $\boldsymbol{x} = (\cdots \circ f_3 \circ f_2 \circ f_1)(01)$ to be the one-sided infinite word

$$\boldsymbol{x} = \lim_{n\to\infty} f_n(f_{n-1}(\cdots f_2(f_1(w))\cdots)),$$

which has each $f_n(f_{n-1}(\cdots f_2(f_1(w))\cdots)$ as a prefix. Following Rampersad, we use the notation $\boldsymbol{x} = w \bullet \boldsymbol{f}$.

Let $u \in B^k$. Then

$$01 \bullet u\boldsymbol{f} = (01 \bullet u)\delta^k(01)^{-1}\delta^k(\boldsymbol{x}). \tag{1}$$

Define the sets I and F by

$$\begin{aligned} I &= (\beta+\gamma)(\alpha\alpha)^*\alpha(\gamma+\beta\beta) \cup \gamma(\alpha\beta)^*\alpha\gamma, \\ F &= B^\omega - B^* I B^\omega. \end{aligned}$$

Theorem 20. *Let $\boldsymbol{x} \in \{0,1\}^\omega$. If $\boldsymbol{x}$ begins with 01, then $\boldsymbol{x}$ is good if and only if $\boldsymbol{x} = 01 \bullet \boldsymbol{f}$ for some $\boldsymbol{f} \in F$.*

We follow the notation of Berstel, also used by Rampersad. Here I stands for 'ideal', and $B^* I B^\omega$ is the ideal generated by I, consisting of the forbidden factors for F.

Let $W = \{\boldsymbol{f} \in B^\omega : 01 \bullet \boldsymbol{f} \in G\}$. To prove Theorem 20 it is enough to prove that $W = F$. Let $L \subseteq \Sigma^\omega$ and let $x \in \Sigma^*$. We define the (left) quotient $x^{-1}L$ by $x^{-1}L = \{\boldsymbol{y} \in \Sigma^\omega : x\boldsymbol{y} \in L\}$. The next lemma establishes several identities concerning quotients of the set W. They are proved using (1) and Lemma 18. The identities demonstrate that W is precisely the set of infinite labeled paths through the automaton A_{01} given in Figure 1. These are just the labeled paths omitting factors in I, so that $W = F$. Thus, proving Lemma 21 establishes Theorem 20.

Lemma 21. *The following identities hold:*

(a) $W = \alpha^{-1}W$;

(b) $(\beta\alpha\alpha)^{-1}W = (\beta\beta)^{-1}W = \beta^{-1}W$;

(c) $(\beta\alpha\beta)^{-1}W = \gamma^{-1}W = (\beta\gamma)^{-1}W = (\gamma\gamma)^{-1}W$;

(d) $(\beta\alpha)^{-1}W = (\gamma\alpha)^{-1}W$;

(e) $(\beta\alpha\gamma)^{-1}W = (\gamma\beta)^{-1}W = \emptyset$.

Each set of identities corresponds to the state of A_{01} with the same label as the identities. The non-accepting sink (e) is not shown in the figure.

Proof. Suppose $\boldsymbol{f} \in B^*$. Let $01 \bullet \boldsymbol{f} = \boldsymbol{x}$. Thus 01 is a prefix of $\boldsymbol{x}$.

(a) We have

$$\begin{aligned} \alpha\boldsymbol{f} \in W &\iff 01 \bullet \alpha\boldsymbol{f} \in G \\ &\iff (01 \bullet \alpha)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \\ &\iff \cancel{0100(0100)}^{-1}\delta(\boldsymbol{x}) \in G \\ &\iff \boldsymbol{x} \in G \\ &\iff 01 \bullet \boldsymbol{f} \in G \\ &\iff \boldsymbol{f} \in W, \end{aligned}$$

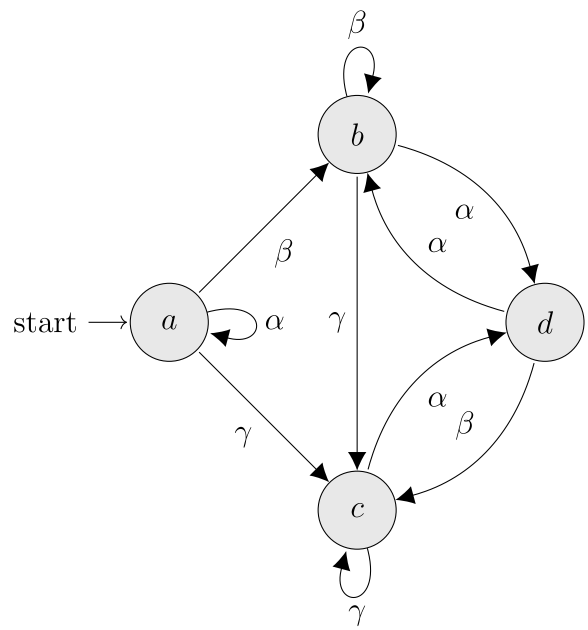

Figure 1: 'Fife' automaton A_{01} for G_{01}.

so that $\alpha^{-1}W = W$.

(b) Here

$$\begin{aligned}
\beta\boldsymbol{f} \in W &\iff 01 \bullet \beta\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \\
&\iff 01\ \cancel{0100(0100)^{-1}}\delta(\boldsymbol{x}) \in G \\
&\iff \delta(0\boldsymbol{x}) \in G \iff 0\boldsymbol{x} \in G.
\end{aligned}$$

Similarly we find that

$$\begin{aligned}
\beta\beta\boldsymbol{f} \in W &\iff 01 \bullet \beta\beta\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\beta)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 010100\ \cancel{01000101(01000101)^{-1}}\delta^2(\boldsymbol{x}) \in G \\
&\iff \delta(0\delta(0\boldsymbol{x})) \in G \\
&\iff 0\delta(0\boldsymbol{x}) \in G \\
&\iff 10\boldsymbol{x} \in G \text{ or } 0\boldsymbol{x} \in G_{001} \\
&\iff 0\boldsymbol{x} \in G.
\end{aligned}$$

Here we use the fact that 01 is a prefix of $\boldsymbol{x}$, so that $10\boldsymbol{x} \notin G$ and $0\boldsymbol{x} \in 001\{0,1\}^{\omega}$.

Finally we get

$$\begin{aligned}
\beta\alpha\alpha\boldsymbol{f} \in W &\iff 01 \bullet \beta\alpha\alpha\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\alpha\alpha)\delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \\
&\iff 01\delta^3(\boldsymbol{x}) \in G \\
&\iff \delta(0\delta^2(\boldsymbol{x})) \in G \\
&\iff 0\delta^2(\boldsymbol{x}) \in G \\
&\iff 1\delta(\boldsymbol{x}) \in G \text{ or } \delta(\boldsymbol{x}) \in G_{001} \\
&\iff 1\delta(\boldsymbol{x}) \in G \\
&\iff 0\boldsymbol{x} \in G.
\end{aligned}$$

Here again note that 01 is a prefix of $\delta(\boldsymbol{x})$. We have shown that $(\beta\alpha\alpha)^{-1}W = (\beta\beta)^{-1}W = \beta^{-1}W$, as desired.

(c) We have

$$\begin{aligned}
\gamma\boldsymbol{f} \in W &\iff 01 \bullet \gamma\boldsymbol{f} \in G \\
&\iff (01 \bullet \gamma)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \\
&\iff 0101\ \cancel{0100(0100)}^{-1}\delta(\boldsymbol{x}) \in G \\
&\iff \delta(00\boldsymbol{x}) \in G \\
&\iff 00\boldsymbol{x} \in G.
\end{aligned}$$

We also find

$$\begin{aligned}
\beta\alpha\beta\boldsymbol{f} \in W &\iff 01 \bullet \beta\alpha\beta\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\alpha\beta)\delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \\
&\iff 0101000101\delta^3(\boldsymbol{x}) \in G \\
&\iff \delta(0\delta^2(0\boldsymbol{x})) \in G \\
&\iff 0\delta^2(0\boldsymbol{x}) \in G \\
&\iff 1\delta(0\boldsymbol{x}) \in G \text{ or } \delta(0\boldsymbol{x}) \in G_{001} \\
&\iff 1\delta(0\boldsymbol{x}) \in G \\
&\iff 00\boldsymbol{x} \in G.
\end{aligned}$$

Similarly, we get

$$\begin{aligned}
\beta\gamma\boldsymbol{f} \in W &\iff 01 \bullet \beta\gamma\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\gamma)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 0101000100\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff \delta(0\delta(00\boldsymbol{x})) \in G \\
&\iff 0\delta(00\boldsymbol{x}) \in G \\
&\iff 100\boldsymbol{x} \in G \text{ or } 00\boldsymbol{x} \in G_{001} \\
&\iff 00\boldsymbol{x} \in G.
\end{aligned}$$

We also have

$$\begin{aligned}
\gamma\gamma\boldsymbol{f} \in W &\iff 01 \bullet \gamma\gamma\boldsymbol{f} \in G \\
&\iff (01 \bullet \gamma\gamma)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 010101000100\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff \delta^2(100\boldsymbol{x}) \in G \\
&\iff 100\boldsymbol{x} \in G.
\end{aligned}$$

Finally, suppose $00\boldsymbol{x} \in G$. Then 0001 is a prefix of $00\boldsymbol{x}$, and by Theorem 7 we can write $00\boldsymbol{x} = \delta(10\boldsymbol{y})$, some $10\boldsymbol{y} \in G$. Then by Lemma 16, $010\boldsymbol{y} \in G$. Thus $\delta(010\boldsymbol{y}) = 0100\boldsymbol{x} \in G$, and in particular, $100\boldsymbol{x} \in G$. We conclude that $100\boldsymbol{x} \in G \iff 00\boldsymbol{x} \in G$, which gives the desired result.

(d) In this case we see that

$$\begin{aligned}
\beta\alpha\boldsymbol{f} \in W &\iff 01 \bullet \beta\alpha\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\alpha)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 01\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff \delta(0\delta(\boldsymbol{x})) \in G \\
&\iff 0\delta(\boldsymbol{x}) \in G \\
&\iff 1\boldsymbol{x} \in G \text{ or } \boldsymbol{x} \in G_{001} \\
&\iff 1\boldsymbol{x} \in G.
\end{aligned}$$

Also we have

$$\begin{aligned}
\gamma\alpha\boldsymbol{f} \in W &\iff 01 \bullet \gamma\alpha\boldsymbol{f} \in G \\
&\iff (01 \bullet \gamma\alpha)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 0101\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff \delta(00\delta(\boldsymbol{x})) \in G \\
&\iff \delta^2(1\boldsymbol{x}) \in G \iff 1\boldsymbol{x} \in G.
\end{aligned}$$

Thus $(\beta\alpha)^{-1}W = (\gamma\alpha)^{-1}W$.

(e) Here

$$\begin{aligned}
\beta\alpha\gamma\boldsymbol{f} \in W &\iff 01 \bullet \beta\alpha\gamma\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\alpha\gamma)\delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \\
&\iff 010100010101000101\delta^3(\boldsymbol{x}) \in G \\
&\iff \delta(0\delta^2(00\boldsymbol{x})) \in G \\
&\iff 0\delta^2(00\boldsymbol{x}) \in G \\
&\iff 1\delta(00\boldsymbol{x}) \in G \text{ or } \delta(00\boldsymbol{x}) \in G_{001} \\
&\iff 1\delta(00\boldsymbol{x}) \in G \\
&\iff 000\boldsymbol{x} \in G.
\end{aligned}$$

But $000\boldsymbol{x}$ has prefix 0000. Thus $(\beta\alpha\gamma)^{-1}W = \emptyset$. We also find

$$\begin{aligned}
\gamma\beta\boldsymbol{f} \in W &\iff 01 \bullet \gamma\beta\boldsymbol{f} \in G \\
&\iff (01 \bullet \gamma\beta)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 01010100\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff \delta^2(10\boldsymbol{x}) \in G \\
&\iff 10\boldsymbol{x} \in G.
\end{aligned}$$

But $10\boldsymbol{x}$ has prefix 1001, so $(\gamma\beta)^{-1}W = \emptyset$.

$\square$

We have

$$G = G_{01} \cup G_{001} \cup G_{0001} \cup G_1.$$

We can write Theorem 20 as

$$G_{01} = 01 \bullet W.$$

By Lemma 16, $\boldsymbol{x} \in G_1$ if and only if $0\boldsymbol{x} \in G_{01}$, so that

$$G_1 = 0^{-1}G_{01}.$$

Also, by Theorem 7 and Theorem 6,

$$G_{0001} = \delta(G_1),$$

since $G_{10} = G_1$. If $\boldsymbol{x} \in G_{001}$, we can write $\boldsymbol{x} = 0\delta(0\boldsymbol{y})$ for some $\boldsymbol{y}$. However, by Lemma 18 we get

$$\begin{aligned}
0\delta(0\boldsymbol{y}) \in G &\iff 10\boldsymbol{y} \in G \text{ or } 0\boldsymbol{y} \in G_{001} \\
&\iff 010\boldsymbol{y} \in G \text{ or } 0\boldsymbol{y} \in G_{001} \text{ by Lemma 16} \\
&\iff 010\boldsymbol{y} \in G_{01} \text{ or } 0\boldsymbol{y} \in G_{001} \\
&\iff 0\boldsymbol{y} \in (01)^{-1}G_{01} \text{ or } 0\boldsymbol{y} \in G_{001} \\
&\iff 0\delta(0\boldsymbol{y}) \in 0\delta((01)^{-1}G_{01} \cup G_{001}), \text{ so that}
\end{aligned}$$

$$G_{001} = 0\delta((01)^{-1}G_{01} \cup G_{001}).$$

We summarize these results in a theorem.

Theorem 22. *The following identities hold:*

$$\begin{aligned} G_{01} &= 01 \bullet W; \\ G_{001} &= 0\delta((01)^{-1}G_{01} \cup G_{001}); \\ G_{0001} &= \delta(G_1); \\ G_1 &= 0^{-1}G_{01}. \end{aligned}$$

Since I is a regular language one could give an enumeration of finite prefixes of G_{01} (and hence G_{0001}, G_1) following the approach of Kobayashi [10]. This would give a lower bound on good words of length n. Obtaining an enumeration of all prefixes of G would involve dealing with the recursion in the equation for G_{001}, and would give a better lower bound on the number of length n good words.

It would be nice to remove the recursion in the equation for G_{001}. With $\mathcal{I}$, $\mathcal{F}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, $\hat{\bullet}$, etc., corresponding to I, F, α, β, γ, $\bullet$, etc., but for overlaps, Rampersad's formulation of Fife's theorem has the following (nonrecursive) form:

Theorem 23. *Let* $\boldsymbol{x} \in \{0,1\}^\omega$.

1. *If* $\boldsymbol{x}$ *begins with* 01, *then* $\boldsymbol{x}$ *is overlap-free if and only if* $\boldsymbol{x} = 01\hat{\bullet}\boldsymbol{f}$ *for some* $\boldsymbol{f} \in \mathcal{F}$.
2. *If* $\boldsymbol{x}$ *begins with* 001, *then* $\boldsymbol{x}$ *is overlap-free if and only if* $\boldsymbol{x} = 01\hat{\bullet}\boldsymbol{f}$ *for some* $\beta\boldsymbol{f} \in \mathcal{F}$.

Remark 24. In the second part of this theorem the condition can be rewritten as saying that $\boldsymbol{x} \in 001\{0,1\}^\omega$ is not overlap-free, if and only if $\hat{\beta}\boldsymbol{f}$ has a factor in $\mathcal{I}$, where $\boldsymbol{f}$ describes the canonical decomposition of $\boldsymbol{x}$. In particular, if $\boldsymbol{f}$ has a factor in $\mathcal{I}$, then $\boldsymbol{x}$ is not overlap-free. No result analogous to this seems possible for good words; we shall see that there are words $\boldsymbol{x} \in G_{001}$ with description $\boldsymbol{x} = 001 \bullet \boldsymbol{f}$ such that $\boldsymbol{f}$ has a factor $\beta\alpha\beta\beta \in I$. On the other hand, $\alpha\alpha\alpha\gamma \notin I$, but cannot be a factor of an infinite word $\boldsymbol{f}$ such that $001 \bullet \boldsymbol{f} \in G$.

Let $\boldsymbol{f} \in B^\omega$ such that $001 \bullet \boldsymbol{f} = \boldsymbol{x}$. Let $w \in B^k$. Then

$$001 \bullet w\boldsymbol{f} = (001 \bullet w)\delta^k(01)^{-1}\delta^k(\boldsymbol{x}). \tag{2}$$

Let F_{001} be the set of infinite words walkable on the automaton A_{001} of Figure 2.

Theorem 25. *Let* $\boldsymbol{x} \in \{0,1\}^\omega$. *If* $\boldsymbol{x}$ *begins with* 001, *then* $\boldsymbol{x}$ *is good if and only if* $\boldsymbol{x} = 001 \bullet \boldsymbol{f}$ *for some* $\boldsymbol{f} \in F_{001}$.

Let $W_{001} = \{\boldsymbol{f} \in B^\omega : 001 \bullet \boldsymbol{f} \in G\}$. We prove that $W_{001} = F_{001}$.

The identities in the following lemma correspond to the states of A_{001}, except for state $\beta\gamma\alpha$, which is labeled for its shortest path from the state a. The non-accepting sink (d) is omitted from the figure. Proving the identities thus proves Theorem 25. One notes that $\beta\alpha\beta\beta \in I \cap F_{001}$. However, $\alpha\alpha\alpha\gamma$ is a prefix of words in F, but cannot be a factor of a word $\boldsymbol{f} \in F_{001}$.

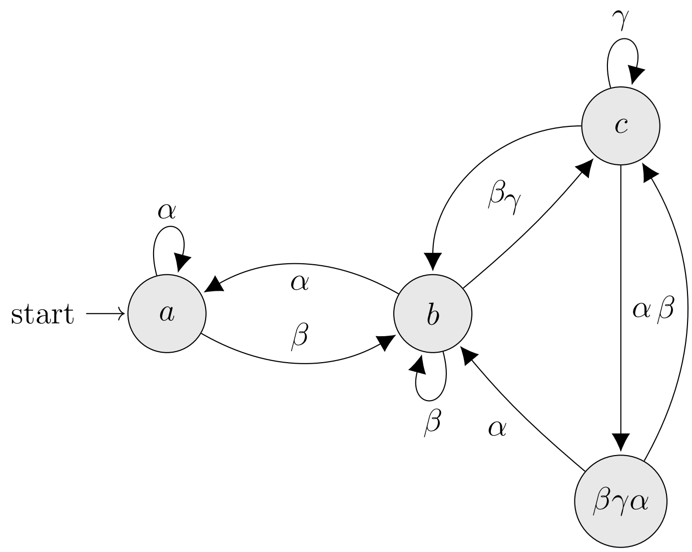

Figure 2: ‘Fife’ automaton A_{001} for G_{001}.

Lemma 26. *The following identities hold:*

(a) $(\beta\alpha)^{-1}W_{001} = W_{001} = \alpha^{-1}W_{001}$*;*

(b) $(\beta\gamma\alpha\alpha)^{-1}W_{001} = (\beta\gamma\beta)^{-1}W_{001} = (\beta\beta)^{-1}W_{001} = \beta^{-1}W_{001}$*;*

(c) $(\beta\gamma\alpha\beta)^{-1}W_{001} = (\beta\gamma\gamma)^{-1}W_{001} = (\beta\gamma)^{-1}W_{001}$*;*

(d) $(\beta\gamma\alpha\gamma)^{-1}W_{001} = \gamma^{-1} = \emptyset$*.*

Proof. Suppose $\boldsymbol{f} \in B^*$. Let $001 \bullet \boldsymbol{f} = \boldsymbol{x}$. Thus 001 is a prefix of $\boldsymbol{x}$.

(a) We have

$$\begin{aligned} \alpha\boldsymbol{f} \in W_{001} &\iff 001 \bullet \alpha\boldsymbol{f} \in G \\ &\iff (001 \bullet \alpha)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \\ &\iff 0\ \cancel{0100(0100)}^{-1}\delta(\boldsymbol{x}) \in G \\ &\iff 1\boldsymbol{x} \in G \text{ or } \boldsymbol{x} \in G_{001} \\ &\iff \boldsymbol{x} \in G \\ &\iff \boldsymbol{f} \in W_{001}. \end{aligned}$$

We also find that

$$
\begin{aligned}
\beta\alpha\boldsymbol{f} \in W_{001} &\iff 001 \bullet \beta\alpha\boldsymbol{f} \in G \\
&\iff (001 \bullet \beta\alpha)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 001\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 0\delta(0\delta(\boldsymbol{x})) \in G \\
&\iff 10\delta(\boldsymbol{x}) \in G \text{ or } 0\delta(\boldsymbol{x}) \in G_{001} \\
&\iff 0\delta(\boldsymbol{x}) \in G \\
&\iff 1\boldsymbol{x} \in G \text{ or } \boldsymbol{x} \in G_{001} \\
&\iff \boldsymbol{x} \in G \\
&\iff \boldsymbol{f} \in W_{001},
\end{aligned}
$$

so that $(\beta\alpha)^{-1}W_{001} = \alpha^{-1}W_{001} = W_{001}$.

(b) Here we have

$$
\begin{aligned}
\beta\boldsymbol{f} \in W_{001} &\iff 001 \bullet \beta\boldsymbol{f} \in G \\
&\iff (001 \bullet \beta)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \\
&\iff 001\ \cancel{0100(0100)}^{-1}\delta(\boldsymbol{x}) \in G \\
&\iff 0\delta(0\boldsymbol{x}) \in G \\
&\iff 10\boldsymbol{x} \in G \text{ or } 0\boldsymbol{x} \in G_{001} \\
&\iff 10\boldsymbol{x} \in G.
\end{aligned}
$$

In the same way we get

$$
\begin{aligned}
\beta\beta\boldsymbol{f} \in W_{001} &\iff 001 \bullet \beta\beta\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\beta)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 010100\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff \delta(0\delta(0\boldsymbol{x})) \in G \\
&\iff 0\delta(0\boldsymbol{x}) \in G \\
&\iff 10\boldsymbol{x} \in G \text{ or } 0\boldsymbol{x} \in G_{001} \\
&\iff 10\boldsymbol{x} \in G.
\end{aligned}
$$

Finally we find that

$$\begin{aligned}
\beta\gamma\alpha\alpha\boldsymbol{f} \in W_{001} &\iff 001 \bullet \beta\gamma\alpha\alpha\boldsymbol{f} \in G \\
&\iff (01 \bullet \beta\gamma\alpha\alpha)\delta^4(01)^{-1}\delta^4(\boldsymbol{x}) \in G \\
&\iff 00101000100\delta^4(\boldsymbol{x}) \in G \\
&\iff 0\delta(0\delta^2(1\delta(\boldsymbol{x}))) \in G \\
&\iff 10\delta^2(1\delta(\boldsymbol{x})) \in G \text{ or } 0\delta^2(1\delta(\boldsymbol{x})) \in G_{001} \\
&\iff 0\delta^2(1\delta(\boldsymbol{x})) \in G \\
&\iff 1\delta(1\delta(\boldsymbol{x})) \in G \text{ or } \delta(1\delta(\boldsymbol{x})) \in G_{001} \\
&\iff 1\delta(1\delta(\boldsymbol{x})) \in G \\
&\iff 01\delta(\boldsymbol{x}) \in G \\
&\iff \delta(0\boldsymbol{x}) \in G \\
&\iff 0\boldsymbol{x} \in G \\
&\iff 10\boldsymbol{x} \in G.
\end{aligned}$$

Thus $(\beta\gamma\alpha\alpha)^{-1}W_{001} = (\beta\gamma\beta)^{-1}W_{001} = (\beta\beta)^{-1}W_{001} = (\beta\alpha)^{-1}W_{001} = \beta^{-1}W_{001}$. We have $0\boldsymbol{x} \in G \iff 10\boldsymbol{x} \in G$ by the same argument as at the end of case (c) of Lemma 21.

(c) Here we have

$$\begin{aligned}
\beta\gamma\boldsymbol{f} \in W_{001} &\iff 001 \bullet \beta\gamma\boldsymbol{f} \in G \\
&\iff (001 \bullet \beta\gamma)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 00101000100\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \\
&\iff 0\delta(0\delta(00\boldsymbol{x})) \in G \\
&\iff 10\delta(00\boldsymbol{x}) \in G \text{ or } 0\delta(00\boldsymbol{x}) \in G_{001} \\
&\iff 0\delta(00\boldsymbol{x}) \in G \\
&\iff 100\boldsymbol{x} \in G \text{ or } 00\boldsymbol{x} \in G_{001} \\
&\iff 100\boldsymbol{x} \in G.
\end{aligned}$$

We also have that

$$
\begin{aligned}
\beta\gamma\gamma\boldsymbol{f} \in W_{001} &\iff 001 \bullet \beta\gamma\gamma\boldsymbol{f} \in G \\
&\iff (001 \bullet \beta\gamma\gamma)\delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \\
&\iff 0010100010001000101010001 01\delta^3(\boldsymbol{x}) \in G \\
&\iff 0\delta(0\delta^2(100\boldsymbol{x})) \in G \\
&\iff 0\delta^2(100\boldsymbol{x}) \in G \text{ or } \delta^2(100\boldsymbol{x}) \in G_{001} \\
&\iff 0\delta^2(100\boldsymbol{x}) \in G \\
&\iff 1\delta(100\boldsymbol{x}) \in G \text{ or } \delta(100\boldsymbol{x}) \in G_{001} \\
&\iff 1\delta(100\boldsymbol{x}) \in G \\
&\iff 0100\boldsymbol{x} \in G \\
&\iff 100\boldsymbol{x} \in G.
\end{aligned}
$$

Finally we compute that

$$
\begin{aligned}
\beta\gamma\alpha\beta\boldsymbol{f} \in W_{001} &\iff 001 \bullet \beta\gamma\alpha\beta\boldsymbol{f} \in G \\
&\iff (001 \bullet \beta\gamma\alpha\beta)\delta^4(01)^{-1}\delta^4(\boldsymbol{x}) \in G \\
&\iff 001010001000100010101000100\delta^4(\boldsymbol{x}) \in G \\
&\iff 0\delta(0\delta^2(1\delta(0\boldsymbol{x}))) \in G \\
&\iff 10\delta^2(1\delta(0\boldsymbol{x})) \in G \text{ or } 0\delta^2(1\delta(0\boldsymbol{x})) \in G_{001} \\
&\iff 0\delta^2(1\delta(0\boldsymbol{x})) \in G \\
&\iff 1\delta(1\delta(0\boldsymbol{x})) \in G \text{ or } \delta(1\delta(0\boldsymbol{x})) \in G_{001} \\
&\iff 1\delta(1\delta(0\boldsymbol{x})) \in G \\
&\iff 01\delta(0\boldsymbol{x}) \in G \\
&\iff \delta(00\boldsymbol{x}) \in G \\
&\iff 00\boldsymbol{x} \in G \\
&\iff 100\boldsymbol{x} \in G.
\end{aligned}
$$

(d) In this instance we see that

$$
\begin{aligned}
\gamma\boldsymbol{f} \in W_{001} &\iff 001 \bullet \gamma\boldsymbol{f} \in G \\
&\iff (001 \bullet \gamma)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \\
&\iff 00101\ \cancel{0100(0100)}^{-1}\delta(\boldsymbol{x}) \in G \\
&\iff 0\delta(00\boldsymbol{x})) \in G.
\end{aligned}
$$

But $00\boldsymbol{x}$ has prefix 0000. Thus $\gamma^{-1}W_{001} = \emptyset$.

In the same way we find that

$$
\begin{aligned}
&\beta\gamma\alpha\gamma\boldsymbol{f} \in W_{001} \\
\iff &001 \bullet \beta\gamma\alpha\gamma\boldsymbol{f} \in G \\
\iff &(001 \bullet \beta\gamma\alpha\gamma)\delta^4(01)^{-1}\delta^4(\boldsymbol{x}) \in G \\
\iff &0010100010001000101010001000100010101000100\delta^4(\boldsymbol{x}) \in G \\
\iff &0\delta(0\delta(^2(1\delta(00\boldsymbol{x}))) \in G.
\end{aligned}
$$

We see that $001 \bullet \beta\gamma\alpha\gamma\boldsymbol{f}$ contains the prefix 0000 of $00\boldsymbol{x}$, so that $(\beta\gamma\alpha\gamma)^{-1}W_{001} = \emptyset$.

□

4 Lexicographically extremal good words

If u is a word of positive length we let u^- denote the word obtained from u by deleting its last letter. The lexicographic order on finite binary words is given recursively by

$$
u < v \iff \begin{cases} v \neq \epsilon, & \text{if } u = \epsilon; \\ (u^- < v^-) \text{ or } ((u = u^-0) \text{ and } (v = u^-1)), & \text{otherwise.} \end{cases}
$$

This extends to an ordering of one-sided infinite binary words; for infinite binary words $\boldsymbol{u}$ and $\boldsymbol{v}$, we say that $\boldsymbol{u} < \boldsymbol{v}$ exactly when there is a prefix u of $\boldsymbol{u}$ and a prefix v of $\boldsymbol{v}$ with $|u| = |v|$ and $u < v$.

The morphism δ is order-reversing on infinite words: Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be infinite binary words such that $\boldsymbol{u} < \boldsymbol{v}$. Write $\boldsymbol{u} = u'0\boldsymbol{u}'$, $\boldsymbol{v} = u'1\boldsymbol{v}'$ where u' is the longest common prefix of $\boldsymbol{u}$ and $\boldsymbol{v}$. Then $\delta(u')01$ is a prefix of $\delta(\boldsymbol{u})$, while $\delta(u')00$ is a prefix of $\delta(\boldsymbol{v})$, so that $\delta(\boldsymbol{u}) > \delta(\boldsymbol{v})$.

For each non-negative integer n, let ℓ_n (resp., m_n) be the lexicographically least (resp., greatest) word of length n such that ℓ_n (resp., m_n) is the prefix of a one-sided infinite good word.

Lemma 27. *Let n be a non-negative integer. Word ℓ_n is a prefix of ℓ_{n+1}. Word m_n is a prefix of m_{n+1}.*

Proof. We prove the result for the ℓ_n; the proof for the m_n is similar. Let $\ell_n\mathbf{r}$ be a one-sided infinite good word. Let p be the length $n+1$ prefix of $\ell_n\mathbf{r}$, and let q be the length n prefix of ℓ_{n+1}. We need to show that $q = \ell_n$. Both p and q are prefixes of one-sided infinite good words. By definition we have $\ell_{n+1} \leq p$ and $\ell_n \leq q$. If $\ell_n < q$, then $p^- = \ell_n < q = \ell_{n+1}^-$, so that $p < \ell_{n+1}$. This is a contradiction. Therefore $\ell_n = q$, as desired. □

Let $\boldsymbol{\ell} = \lim_{n\to\infty} \ell_n$, $\mathbf{m} = \lim_{n\to\infty} m_n$.

Lemma 28. *Word $\boldsymbol{\ell}$ is the lexicographically least one-sided infinite good word. Word $\mathbf{m}$ is the lexicographically greatest one-sided infinite good word.*

Proof. We show that $\boldsymbol{\ell}$ is lexicographically least. The proof that $\mathbf{m}$ is lexicographically greatest is similar. Let $\mathbf{w}$ be a one-sided infinite good word. For each n let w_n be the length n prefix of $\mathbf{w}$, so that $\mathbf{w} = \lim_{n\to\infty} w_n$.

If for some n we have $w_n > \ell_n$, then $\mathbf{w} > \boldsymbol{\ell}$.

Otherwise $w_n \leq \ell_n$ for all n. By the definition of the ℓ_n we have $w_n \geq \ell_n$, so that $w_n = \ell_n$ for all n. Thus $\mathbf{w} = \lim_{n\to\infty} w_n = \lim_{n\to\infty} \ell_n = \boldsymbol{\ell}$.

In all cases we find $\mathbf{w} \geq \boldsymbol{\ell}$. □

Lemma 29. *We have* $\boldsymbol{\ell} = \delta(\mathbf{m})$.

Proof. Word 0001 is the least good word of length 4. Also, 0001 is a factor of $\boldsymbol{d}$, so there are one-sided infinite good words (which are suffixes of $\boldsymbol{d}$) with prefix 0001. Therefore, $\ell_4 = 0001$. Since 0001 is a prefix of $\boldsymbol{\ell}$, and $\boldsymbol{\ell}$ is good, we can write $\boldsymbol{\ell} = \delta(\mathbf{m}')$ for some $\mathbf{m}'$ by Theorem 7. Since $\boldsymbol{\ell}$ is good, $\mathbf{m}'$ is good by Lemma 8. It follows that $\mathbf{m}' \leq \mathbf{m}$. However if $\mathbf{m}' < \mathbf{m}$ then $\delta(\mathbf{m}) < \delta(\mathbf{m}') = \boldsymbol{\ell}$ since δ is order-reversing. This is impossible, since $\boldsymbol{\ell}$ is least. Therefore $\mathbf{m}' = \mathbf{m}$, and $\boldsymbol{\ell} = \delta(\mathbf{m})$. □

Lemma 30. *We have* $0\mathbf{m} = \delta(\boldsymbol{\ell})$.

Proof. Since 1 is the greatest good word of length 1, and $\boldsymbol{d}$ has suffixes that begin with 1, we have $m_1 = 1$. By Lemma 16, $0\mathbf{m}$ is good. Since 01 is a prefix of $0\mathbf{m}$, by Theorem 7 we can write $0\mathbf{m} = \delta(\boldsymbol{\ell}')$ for some $\boldsymbol{\ell}'$. Since $0\mathbf{m}$ is good, $\boldsymbol{\ell}'$ is good, so that $\boldsymbol{\ell}' \geq \boldsymbol{\ell}$. However, δ is order-reversing, so that if $\boldsymbol{\ell}' > \boldsymbol{\ell}$, then $\delta(\boldsymbol{\ell}) > \delta(\boldsymbol{\ell}') = \boldsymbol{m}$, contradicting the maximality of $\boldsymbol{m}$. Thus $\boldsymbol{\ell}' = \boldsymbol{\ell}$ and $0\mathbf{m} = \delta(\boldsymbol{\ell})$. □

Theorem 31. *Word $\boldsymbol{m}$ is the fixed point*

$$\boldsymbol{m} = h_1(\boldsymbol{m}), \tag{3}$$

where $h_1 = [1000, 1010]$. *Word $\boldsymbol{\ell}$ is the fixed point*

$$\boldsymbol{\ell} = h_2(\boldsymbol{\ell}), \tag{4}$$

where $h_2 = [0001, 0101]$.

Proof. From Lemma 29 and Lemma 30 we find $\boldsymbol{m} = 0^{-1}\delta^2(\boldsymbol{m})$, so that for each n we have $m_{4n-2} = 0^{-1}\delta^2(m_n)$. However, $\delta^2(a) = 0h_1(a)0^{-1}$ for $a \in \{0,1\}$, so that $\delta^2(w) = 0h_1(w)0^{-1}$ for $w \in \{0,1\}^*$. Thus $m_{4n-2} = h_1(m_n)0^{-1}$. Then

$$\begin{aligned}
\boldsymbol{m} &= \lim_{n\to\infty} m_n \\
&= \lim_{n\to\infty} m_{4n-2} \\
&= \lim_{n\to\infty} h_1(m_n)0^{-1} \\
&= \lim_{n\to\infty} h_1(m_n) \\
&= h_1(\lim_{n\to\infty} m_n) \\
&= h_1(\boldsymbol{m}).
\end{aligned}$$

The proof for $\boldsymbol{\ell}$ is similar, using the fact that $h_2 = (01)^{-1}\delta^2 01$. □

Corollary 32. *Every finite factor of $\boldsymbol{\ell}$ or $\boldsymbol{m}$ is a factor of $\boldsymbol{d}$, and vice versa. However, words $\boldsymbol{\ell}$, $\boldsymbol{m}$, and $\boldsymbol{d}$ have no common suffix.*

Proof. Because δ^2, h_1, and h_2 are conjugates of each other, their fixed points have the same factors. For example, let u be a factor of $\boldsymbol{d} = \lim_{n\to\infty} \delta^n(0)$. Then for some n, u is a factor of $\delta^{2n}(0) = 0h_1^{2n}(0)0^{-1}$, which is a factor of $h_1^{2n}(1010) = h_1^{2n+1}(1)$, which is a prefix of $\boldsymbol{m} = \lim_{n\to\infty} h_1^n(1)$. Thus u is a factor of $\boldsymbol{m}$.

Now suppose $\boldsymbol{d}$ and $\boldsymbol{\ell}$ have a common suffix $\boldsymbol{s}$. (The proofs for the other pairs of fixed points are similar.) Write $\boldsymbol{d} = d_i\boldsymbol{s}$ and $\boldsymbol{\ell} = \ell_j\boldsymbol{s}$ for some i and j, where d_n is the prefix of $\boldsymbol{d}$ of length n. Then

$$\begin{aligned} d_i\boldsymbol{s} &= \boldsymbol{d} \\ &= \delta^2(\boldsymbol{d}) \\ &= \delta^2(d_i)\delta^2(\boldsymbol{s}), \end{aligned}$$

so that $\boldsymbol{s} = s_{3i}\delta^2(\boldsymbol{s})$ where is s_n is the prefix of $\boldsymbol{s}$ of length n. Similarly,

$$\begin{aligned} \ell_j\boldsymbol{s} &= \boldsymbol{\ell} \\ &= h_2(\boldsymbol{\ell}) \\ &= h_2(\ell_j)h_2(\boldsymbol{s}), \end{aligned}$$

so that $\boldsymbol{s} = s_{3j}h_2(\boldsymbol{s})$. Thus

$$\begin{aligned} s_{3j}h_2(\boldsymbol{s}) &= \boldsymbol{s} \\ &= s_{3i}\delta^2(\boldsymbol{s}) \\ &= s_{3i}01h_2(\boldsymbol{s}) \\ &= s_{3i+2}h_2(\boldsymbol{s}). \end{aligned}$$

It follows that $h_2(\boldsymbol{s})$ has period $|3(i-j)+2|$. However, $h_2(\boldsymbol{s})$ is good and does not contain fourth powers. Therefore, $3(i-j)+2 = 0$. This is impossible since i and j are integers. $\square$

5 Binary patterns in $\boldsymbol{d}$

Every word encounters its factors as patterns. If a word encounters some binary pattern p, it necessarily encounters the complement $\overline{p}$. Shur [15] has shown that, up to complementation, the only binary patterns encountered by $\mathbf{t}$ are its factors and 00100. The situation with $\boldsymbol{d}$ is more complicated.

Lemma 33. *Any factor $0u$ of $\boldsymbol{d}$ can be written as $h_1(p)$ for some word p where $h_1 = [0, 01]$. Any factor $u0$ of $\boldsymbol{d}$ can be written as $h_2(p)$ for some word p where $h_2 = [0, 10]$. Word $\boldsymbol{d}$ thus has an inverse image under each of h_1 and h_2.*

Proof. Every occurrence of 1 in $\boldsymbol{d}$ is preceded and followed by 0. $\square$

Corollary 34. *Let $0u$ be a factor of $\boldsymbol{d}$ such that $|0u| \geq 13$. Then $0u$ can be written as $h_1(p)$ where p is neither a factor of $\boldsymbol{d}$ nor the complement of a factor of $\boldsymbol{d}$. Thus $\boldsymbol{d}$ encounters infinitely many patterns p such that neither of p and $\overline{p}$ is a factor of $\boldsymbol{d}$.*

Proof. The longest factor of $\boldsymbol{d}$ not containing 010001 is 100010101000, which has length 12. Therefore, $0u = h_1(p)$ has the factor $010001 = h_1(1001)$, and p has the factor 1001. Neither of 1001 and 0110 is a factor of $\boldsymbol{d}$ so neither p nor $\overline{p}$ is a factor of $\mathbf{d}$. $\square$

For every particular pattern p, the automatic proving system Walnut [11] can in theory, given enough computing power and time, determine whether $\boldsymbol{d}$ encounters p. However, in the next theorem, we effectively characterize all binary patterns p encountered by $\boldsymbol{d}$. The remainder of this section is devoted to its proof.

Theorem 35. *Word p is a binary pattern encountered by $\boldsymbol{d}$ if and only if one of the following holds:*

1. *One of p and $\overline{p}$ is a factor of $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$.*
2. *One of p and $\overline{p}$ is among*

   0010100, 01001001000, 00100100100, 001001001000, 00010010010,

   000100100100, 0010001000100, 00100010001000, 00010001000100, *and* 000100010001000.

The two possibilities are distinct.

The following analog of Lemma 12 will be useful for analyzing the patterns appearing in $\mathbf{w}$:

Lemma 36. *Let u, v be binary words such that uvu is a factor of $\boldsymbol{d}$, $|u| \geq 3$, and either*

- *we have $|u|_{00} \geq 1$, or*
- *we have $|u|_{10101} \geq 1$.*

Then $|uv| \equiv 0 \pmod 4$.

Proof. Suppose that 00 is a factor of u, and $|u| \geq 3$. Then word u contains one of the factors 000, 001, and 100. These can only arise in $\boldsymbol{d}$ as suffixes of some prefix of $\boldsymbol{d}$ of the form $\delta^2(p0)0$, $\delta^2(p0)01$, and $\delta^2(p0)$, respectively. The index of every occurrence of factor 000, 001, or 100, and thus the index of every occurrence of u, in $\boldsymbol{d}$ is therefore fixed modulo 4, and the result follows.

Suppose $|u|_{10101} \geq 1$. The factor 10101 only occurs in $\boldsymbol{d}$ as a suffix of some prefix of $\boldsymbol{d}$ of the form $\delta^2(p10)(00)^{-1}$ and the index of every occurrence of u in $\boldsymbol{d}$ is again fixed, modulo 4. $\square$

Theorem 37. *Suppose that $\boldsymbol{d}$ encounters binary pattern p. Then one of p and $\overline{p}$ either*

1. *is a factor of* $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, *or of* $h_2^{-1}(\boldsymbol{d})$, *or*

2. *has the property that all its factors of the form* 10^k1 *have the same length.*

Proof. Without loss of generality, replacing p by $\overline{p}$ if necessary, assume that 0 is the first letter of p. Assume that neither p nor $\overline{p}$ is a factor of $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$. Since 000 is a factor of $\boldsymbol{d}$ and $\boldsymbol{d}$ does not encounter 0000, we must have $|p|_1 > 0$. Let $g(p)$ be a factor of $\delta^n(0)$ where $g = [X, Y]$ is a non-erasing morphism, and n is as small as possible. By Lemma 11, $|X|$ and $|Y|$ are not both even. We consider cases based on $|X|_1$ and $|Y|_1$.

Case 1 We have $|X|_1, |Y|_1 > 0$: If XX (resp., YY, XYX, YXY) is a factor of $\delta^n(0)$, then by Lemma 12, $|X|$ (resp., $|Y|$, $|XY|$, $|YXY|$) is even. It follows that not both XX and YY are factors of $g(p)$. Also, not both XX and XYX are factors of $g(p)$, or else $|X|$ and $|XY|$ are even, forcing $|Y|$ to be even. Similarly, not both YY and YXY are factors of $g(p)$.

If neither of XX and YY is a factor of $g(p)$, then $g(p)$ is an alternating string of X's and Y's, and thus a prefix of $XYXYXYX$. (Since $\boldsymbol{d}$ is good, the fourth power $(XY)^4$ is not a factor of $\boldsymbol{d}$.) But then p is a prefix of 0101010, which is a factor of $\boldsymbol{d}$. This is a contradiction.

Suppose then, that XX is a factor of $g(p)$. Then neither of YY and XYX is a factor of $g(p)$. If $|p|_1 \geq 2$, then $g(p)$ would have a prefix of the form X^rYX^sY, $r, s \geq 1$. This contains a factor XYX, which is impossible. It follows that $|p|_1 \leq 1$. Because $\boldsymbol{d}$ is good, $g(p)$ cannot have a factor $XXXX$, and we conclude that p is a factor of 0001000. But 0001000 is seen to be a factor of $\boldsymbol{d}$.

Finally, suppose that YY is a factor of $g(p)$. Then XX and YXY are not factors of $g(p)$. Therefore, p begins with 0, and has a factor 11, but not 00, 101, or 1111. (A factor 1111 in p would give a fourth power in the good word $\boldsymbol{d}$.) It follows that p is one of 011, 0110, 0111, and 01110. However $h_1(\overline{011}) = 0100$, $h_1(\overline{0110}) = 010001$, $\overline{0111} = 1000$, and $\overline{0111} = 1000$ are all factors of $\boldsymbol{d}$.

Case 2 We have $|X|_1 = |Y|_1 = 0$. This forces $g(p)$ to be a factor of 000, so that p is a binary word of length 3 or less. Each such word, or its complement, is a factor of $\boldsymbol{d}$.

Case 3 We have $|X|_1 > 0$ but $|Y|_1 = 0$. Write $Y = 0^n$ where $1 \leq n \leq 3$.

Case 3(a) We have $|X| \geq 3$, and either $|X|_{00} \geq 1$ or $|X|_{10101} \geq 1$. If p contains only a single 0 (its first letter) then, since fourth powers do not appear in $\boldsymbol{d}$, p must be a prefix of 0111, which is the complement of a factor of $\boldsymbol{d}$. Otherwise, p has a factor of the form 01^k0 for some k, $0 \leq k \leq 3$. Thus $XY^kX = X0^{nk}X$ is a factor of $\boldsymbol{d}$. By Lemma 36 with $u = X$, $v = 0^{nk}$, we have $|X| \equiv -nk \pmod 4$.

If $n = 2$, then since 0^4 is not a factor of $\boldsymbol{d}$, we have $0 \leq k \leq 1$. Also, $|X| \equiv -nk$ (mod 4), so $|X|$ is even, giving $k \equiv -|X|/2$ (mod 2) and k is determined; each pair of 0's in p is separated by the same number of 1's. Then $\overline{p}$ has the property that all its factors of the form 10^k1 have the same length. If $n = 1$ or $n = 3$, then since $0 \leq k \leq 3$, k is determined by the congruence $|X| \equiv -nk \pmod 4$, and again $\overline{p}$ has the property that all its factors of the form 10^k1 have the same length.

Case 3(b) Either $|X| < 3$, or $|X|_{00} = |X|_{10101} = 0$. If $|X|_{00} = 0$, then X does not contain a factor 00 or 11, and is, therefore, an alternating string of 0's and 1's. Thus the given conditions imply that $X = 00$, or X is a factor of 01010.

If $X = 00$, then $g(p)$ must be a factor of 000, which is a factor of $\boldsymbol{d}$.

Suppose then that X is a factor of 01010. If p contained at most a single 0 (its first letter), then p would be a prefix of 0111, the complement of a factor of $\boldsymbol{d}$. This is impossible. Therefore, assume that p has a factor of the form 01^k0. If $|X| \geq 3$, then X is one of 010, 101, 0101, 1010, or 01010. One checks that for each of these possibilities for X there is exactly one value of k such that $X0^kX$ is a factor of $\boldsymbol{d}$:

- If X is 010 or 01010, $k = 1$.
- If $X = 101$, then a factor XX would imply $\boldsymbol{d}$ has factor 11; a factor $X0X$ in $\boldsymbol{d}$ would extend on the right to a factor $X0X0 = (10)^4$ in $\boldsymbol{d}$; a factor $X00X$ would imply $\boldsymbol{d}$ has factor 1001; thus $k = 3$.
- If X is 0101 or 1010, then XX is a fourth power, while $X0X$ has the factor 1001; $X000X$ contains 0000; thus $k = 2$.

Thus if $|X| \geq 3$ then $\overline{p}$ has the property that all its factors of the form 10^k1 have the same length. We are left with the cases where $|X| \leq 2$ and $X \neq 00$, i.e., word X is among 1, 01, and 10.

$X = 1$. Since $\boldsymbol{d}$ has a factor XY^kX, we cannot have $Y = 00$ or else $\boldsymbol{d}$ contains 0000 or 1001. Thus $Y = 0$. Then $g = [1, 0]$, and p is the complement of a factor of $\boldsymbol{d}$.

$X = 01$. Since $|X|$ is even, we must have $|Y|$ odd, so that $n = 1$ or $n = 3$. If $n = 3$, then $YX = 0^41$ and $YY = 0^6$ are not factors of $\boldsymbol{d}$, forcing $|p|_1 = 1$, and p is a factor of 00010000, which is a factor of $\boldsymbol{d}$. Suppose then that $n = 1$. Now, however $g = [01, 0]$, and $h_1(\overline{p}) = g(p)$ is a factor of $\boldsymbol{d}$, so that $\overline{p}$ is a factor of $h_1^{-1}(\boldsymbol{d})$.

$X = 10$. Again $|X|$ is even, so we find $n = 3$ or $n = 1$. In the first case, XY contains 0000, which is impossible. Thus p is a factor of 000, a factor of $\boldsymbol{d}$. In the second case, $g = [10, 0]$, and $\overline{p}$ is a factor of $h_2^{-1}(\boldsymbol{d})$.

Case 4 We have $|X|_1 = 0$ but $|Y|_1 > 0$. Write $X = 0^n$ where $1 \leq n \leq 3$.

Case 4(a) We have $|Y| \geq 3$, and either $|Y|_{00} \geq 1$ or $|Y|_{10101} \geq 1$. If p contains at most a single 1 then, since fourth powers do not appear in $\boldsymbol{d}$, p must be a factor of 0001000, which is a factor of $\boldsymbol{d}$. Otherwise, p has some factor of the form 10^k1 for some k, $0 \le k \le 3$. Thus $YX^kY = Y0^{nk}Y$ is a factor of $\boldsymbol{d}$. By Lemma 36 with $u = Y$, $v = 0^{nk}$, we have $|Y| \equiv -nk \pmod 4$.

If $n = 2$, then since 0^4 is not a factor of $\boldsymbol{d}$, we have $0 \le k \le 1$. Also, $|Y| \equiv -nk$ so $|Y|$ is even, giving $k \equiv -|Y|/2 \pmod 2$ and k is determined. Therefore, p has the property that all its factors of the form 10^k1 have the same length.

If $n = 1$ or $n = 3$, then since $0 \le k \le 3$, k is determined by the congruence $|Y| \equiv -nk \pmod 4$, and again p has the property that all its factors of the form 10^k1 have the same length.

Case 4(b) Either $|Y| \le 3$, or $|Y|_{00} = |Y|_{10101} = 0$. If $|Y|_{00} = 0$, then Y does not contain a factor 00 or 11, and is an alternating string of 0's and 1's. Thus the given conditions imply that $Y = 00$, or Y is a factor of 01010.

If $Y = 00$, then $g(p)$ must be a factor of 000, which is a factor of $\boldsymbol{d}$.

Suppose then that Y is a factor of 01010. If p contains at most a single 1, then p must be a factor of 0001000, a factor of $\boldsymbol{d}$. Therefore, assume that p has a factor of the form 10^k1. If $|Y| > 2$, an analysis as in Case 3(a), shows that p has the property that all its factors of the form 10^k1 have the same length. The cases where $|Y| \le 2$ are impossible, with an analysis analogous to that of Case 3(a).

□

Lemma 38. *The set of factors of* $\boldsymbol{d}$ *is closed under reversal.*

Proof. To begin, we notice that if u is a binary word, then

$$(\delta(u))^R = 0^{-1}\delta(u^R)0.$$

Suppose v is a factor of $\boldsymbol{d}$. The set of length 2 factors of $\boldsymbol{d}$ is seen to be $\{00, 01, 10\}$, which is closed under reversal. Therefore, if $|v| \le 2$, then v^R is a factor of $\boldsymbol{d}$. Suppose that $|v| > 2$, and for every shorter factor u of $\boldsymbol{d}$, the reversal u^R is a factor of $\boldsymbol{d}$. However, $|v| > 2$ implies that v is a factor of $\delta(u)$, for some factor u of $\boldsymbol{d}$ which is shorter than v. Thus u and u^R are factors of $\boldsymbol{d}$. Then $(\delta(u))^R = 0^{-1}\delta(u^R)0$ is also a factor of $\boldsymbol{d}$, so that v^R is also. The result follows by induction. □

Corollary 39. *If* $\boldsymbol{d}$ *encounters pattern* p*, then* $\boldsymbol{d}$ *encounters pattern* p^R *also.*

Theorem 40. *The following are equivalent:*

1. *Word* p *is a binary pattern encountered by* $\boldsymbol{d}$*, but neither* p *nor* $\overline{p}$ *is a factor of* $\boldsymbol{d}$*,* $h_1^{-1}(\boldsymbol{d})$*, or* $h_2^{-1}(\boldsymbol{d})$*.*

2. *One of* p *and* $\overline{p}$ *is among*

   0010100*,* 01001001000*,* 00100100100*,* 001001001000*,* 00010010010*,* 000100100100*,* 0010001000100*,* 00100010001000*,* 00010001000100*, and* 000100010001000*.*

Proof. ($1 \Rightarrow 2$): Suppose that p is a binary pattern encountered by $\boldsymbol{d}$, but neither p nor $\overline{p}$ is a factor of $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$. By Theorem 37, replacing p by $\overline{p}$ if necessary, there is a number k such that there are exactly k 0's between subsequent 1's in p. Thus $p = 0^r(10^k)^s10^t$ for some k, r, s, t. Since $\boldsymbol{d}$ does not contain fourth powers, $k, r, s, t \leq 3$. For the same reason, if $s = 3$, then $r + s < k$. If $s = 0$, then p is a factor of 0001000, which is a factor of $\boldsymbol{d}$. We, therefore, have $s \geq 1$. We consider cases based on the value of k:

$k = 0$: In this case, to avoid fourth powers, we must also have $s \leq 2$, and p is a factor of 00011000 or 000111000. We note that $h_1(\overline{00011000}) = 01010100010101$ is a factor of $\boldsymbol{d}$. Thus p cannot be a factor of 00011000, which would make $\overline{p}$ a factor of $h_1^{-1}(\boldsymbol{d})$. We conclude that p is a factor of 000111000. Note that, $h_2(00011100) = h_1(00111000) = 00010101000$ is a factor of $\boldsymbol{d}$, so that p cannot be a proper factor of 000111000. It remains that we must have $p = 000111000$. To obtain a contradiction, we show that $\boldsymbol{d}$ does not encounter p.

Suppose $g(p)$ is a factor of $\boldsymbol{d}$, $g = [X, Y]$ a non-erasing morphism. By Lemma 11 assume that one of $|X|$ and $|Y|$ is odd. Both XX and YY are factors of $g(p) = XXXYYYXXX$, so by Lemma 12 we must have $|X|_1 = 0$ or $|Y|_1 = 0$.

Suppose $|X|_1 = 0$. Thus XXX ends in 000. Since $XXXY$ is a factor of $g(p)$, which is a factor of $\boldsymbol{d}$, the first letter of Y is 1. Since XXX begins with 000 and $YXXX$ is a factor of $g(p)$, the last letter of Y is 1. But then YY contains the factor 11, so that 11 is a factor of the good word $\boldsymbol{d}$, which is impossible.

Suppose $|Y|_1 = 0$. Switching X and Y in the argument of the previous paragraph shows that this is impossible also.

$k = 1$: We consider subcases based on whether s is 1, 2, or 3:

$s = 1$: If $r + t \geq 5$, then p contains 00010100 or its reverse. However $\boldsymbol{d}$ is good, and therefore does not encounter 00010100. By Corollary 39, $\boldsymbol{d}$ does not encounter the reverse of $\boldsymbol{d}$. We conclude that $r + t \leq 4$, so that p is a factor of 0101000, 0010100, or 0001010. However, 0101000 and 0001010 are factors of $\boldsymbol{d}$, so p cannot be a factor of one of these. Therefore p must be a factor of 0010100. Both 010100 and 010100 are factors of $\boldsymbol{d}$, so p cannot be a proper factor of 0010100, forcing $p = 0010100$.

$s = 2$: In this case p is a factor of 00010101000, which is a factor of $\boldsymbol{d}$.

$s = 3$: If $r = t = 0$, then $p = 1010101$ is the complement of a factor of $\boldsymbol{d}$. However, if $r > 0$ or $t > 0$, then p contains one of the fourth powers 01010101 and 10101010, so that $\boldsymbol{d}$ does not encounter p, which is a contradiction.

$k = 2$: We consider subcases based on whether s is 1, 2, or 3:

$s = 1$: If $r+t \leq 5$ then p is factor of 000100100 or of 001001000. However $h_2(000100100) = 00010001000 = h_1(001001000)$ is a factor of $\boldsymbol{d}$, so p is a factor of $h_1^{-1}(\boldsymbol{d})$ or of $h_2^{-1}(\boldsymbol{d})$, a contradiction. If $r = t = 3$, then $p = 0001001000 = h_2(00010100)$, and $\boldsymbol{d}$ encounters 00010100. Since $\boldsymbol{d}$ is good, this is impossible.

$s = 2$: If $r + t \le 3$, then p is a factor of 1001001000, 0100100100, 0010010010, or 0001001001; however each of

$$\begin{aligned} h_1(1001001000) &= 0100010001000, \\ h_2(0100100100) &= 0100010001000, \\ h_1(0010010010) &= 0001000100010, \text{ and} \\ h_2(0001001001) &= 0001000100010 \end{aligned}$$

is a factor of $\boldsymbol{d}$, so that p is a factor of $h_1^{-1}(\boldsymbol{d})$ or of $h_2^{-1}(\boldsymbol{d})$.

If $r = t = 3$, then $p = 0001001001000$. To get a contradiction, we show that $\boldsymbol{d}$ does not encounter p. Otherwise, suppose without loss of generality that $\boldsymbol{d}$ has a factor $g(p)$ where $g = [X, Y]$ is a non-erasing morphism and $|X|$, $|Y|$ are not both odd. Word $\boldsymbol{d}$ has the factor $XXXYXXYXXYXXX$. We cannot have $|X|_1 > 0$, or since XX and XYX are factors of $\boldsymbol{d}$, Lemma 12 would force both $|X|$ and $|Y|$ to be even. Write $X = 0^n$. Since XX is a factor of $\boldsymbol{d}$, we must have $n = 1$ and $X = 0$. Since $XXXY$ and $YXXX$ are factors of $\boldsymbol{d}$, the first and last letters of Y are 1. Then $YXXY$ contains the factor 1001, which cannot be a factor of $\boldsymbol{d}$. This is a contradiction.

We conclude that $4 \le r + t \le 5$, so that p is one of 01001001001000, 00100100100100, 00010010010010, 001001001001000, and 000100100100100.

$s = 3$: If $r + t \le 1$, then p is a factor of 01001001001 or 10010010010. However $h_2(01001001001) = 010001000100010$ and $h_1(010001000100010)$ are factors of $\boldsymbol{d}$, so p is a factor of $h_1^{-1}(\boldsymbol{d})$ or of $h_2^{-1}(\boldsymbol{d})$, which is a contradiction. On the other hand, if $r + t \ge 2$, then one of the fourth powers $(001)^4)$, $(010)^4$ and $(100)^4$ is a factor of p, and $\boldsymbol{d}$ contains a fourth power, which is impossible.

$k = 3$: We consider subcases based on whether s is 1, 2, or 3:

$s = 1$: In this case p is a factor of 00010001000, which is a factor of $\boldsymbol{d}$.

$s = 2$: If $r \le 1$ then p is a factor of 0100010001000 which is a factor of $\boldsymbol{d}$. This is impossible. Similarly, $t \le 1$ is impossible. Thus p must be one of 0010001000100, 00010001000100, 00100010001000, and 000100010001000.

$s = 3$: If $r, t \le 1$ then p is a factor of 010001000100010, which is a factor of $\boldsymbol{d}$, a contradiction. Suppose $r \ge 2$ or $t \ge 2$. Then p has $q = 001000100010001$ or its reverse as a factor. By Corollary 39, it suffices to show that $\boldsymbol{d}$ does not encounter q.

To get a contradiction, suppose that $g(q)$ is a factor of $\boldsymbol{d}$ for some non-erasing morphism $g = [X, Y]$ where one of $|X|$ and Y is odd. Thus word $\boldsymbol{d}$ has the factor

$$XXYXXXYXXXYXXXY.$$

Since XX and XYX are factors of $\boldsymbol{d}$, we must have $X = 0^n$ for some n; otherwise, Lemma 12 implies that $|X|$ and $|Y|$ are even, a contradiction. Since XX is a factor of $\boldsymbol{d}$, we have

$n = 1$ and $X = 0$. Since $YXXX = Y000$ is a factor of $\boldsymbol{d}$, the last letter of Y is 1. Therefore, Y is always followed by 0 in $\boldsymbol{d}$, so $XXYXXXYXXXYXXXY0$ is a factor of $\boldsymbol{d}$. However $XXYXXXYXXXYXXXY0 = (00Y0)^4$, and $\boldsymbol{d}$ contains a fourth power. This is impossible.

($1 \Leftarrow 2$): We show that $\boldsymbol{d}$ encounters each of the listed patterns, but none of these patterns or their complements is a factor of $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$.

First we show that $\boldsymbol{d}$ encounters the listed patterns. Note that each of these patterns is a factor of 0010100, 001001001000, 000100100100, or 000100010001000. Since 001001001000 is the reverse of 000100100100, by Lemma 39 it suffices to show that $\boldsymbol{d}$ encounters 0010100, 001001001000, or 000100010001000.

- $p = 0010100$: Word $\boldsymbol{d}$ has the factor

  $$00010001000 = g(0010100),$$

  where $g = [0, 010]$.

- $p = 001001001000$: Word $\boldsymbol{d}$ has the factor

  $$0001010100010101000010101000 = g(001001001000),$$

  where $g = [0, 010101]$.

- $p = 000100010001000$: Word $\boldsymbol{d}$ has the factor

  $$0001010100010101000010101000 = g(000100010001000),$$

  where $g = [0, 10101]$.

Next, we show that none of these patterns or their complements is a factor of $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$. Notice that $h_1(101) = 01001$ and $h_2(101) = 10010$, and 1001 is not a factor of $\boldsymbol{d}$, since $\boldsymbol{d}$ is good. The complements of all the listed patterns contain 101 as a factor, so none of the complements is a factor of $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$. Further, all the complements of the listed factors contain 11 as a factor, so none of the complements is a factor of $\boldsymbol{d}$. It therefore suffices to show that none of the listed patterns is a factor of $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$.

Each of the listed patterns has one of 0010100, 01001001000, 00100100100, 00010010010, and 0010001000100 as a factor. It, therefore, suffices to show that none of these five patterns is a factor of $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$.

Case 1: 0010100. If this was a factor of $\boldsymbol{d}$, the 1's would force it to appear in the context 00010100, which is impossible since $\boldsymbol{d}$ is good. Since 101 is a factor of 0010100, it is not a factor of $h_1^{-1}(\boldsymbol{d})$, or $h_2^{-1}(\boldsymbol{d})$ either.

Case 2: 01001001000. This has 1001 as a factor, and therefore is not a factor of $\boldsymbol{d}$.

Word $00100010001000 = h_1(01001001000)$ cannot be a factor of $\boldsymbol{d}$; if it preceded 0 then 0000 would be a factor of $\boldsymbol{d}$; if it preceded 10 then $(0100)^4$ is a factor of $\boldsymbol{d}$, which is impossible.

Word $h_2(01001001000)$ ends in the fourth power 0000 and cannot be a factor of $\boldsymbol{d}$.

Case 3: 00100100100. This has 1001 as a factor, and therefore is not a factor of $\boldsymbol{d}$.

Word $h_1(00100100100) = 00010001000100$ cannot be a factor of $\boldsymbol{d}$; if it preceded 1 then 1001 would be a factor of $\boldsymbol{d}$; if it preceded 00 then 0000 is a factor of $\boldsymbol{d}$; if it preceded 01 then $(0001)^4$ is a factor of $\boldsymbol{d}$, which is impossible.

Word $00100010001000 = h_2(00100100100)$ cannot be a factor of $\boldsymbol{d}$ as argued in Case 2.

Case 4: 00010010010. This has 1001 as a factor, and therefore is not a factor of $\boldsymbol{d}$.

Word $h_1(00010010010)$ begins with 0000 and cannot be a factor of $\boldsymbol{d}$.

Word $00010001000100 = h_2(00010010010)$ cannot be a factor of $\boldsymbol{d}$ as argued in Case 3.

Case 5: 0010001000100. This cannot be a factor of $\boldsymbol{d}$. Recall that $\boldsymbol{d}$ is recurrent and does not contain factors 1001 or 0000. However, a second occurrence of this factor in $\boldsymbol{d}$ would be in the context $0001000100010001 = (0001)^4$, which is impossible.

Word $h_1(0010001000100)$ begins 00010000, containing 0000, and cannot be a factor of $\boldsymbol{d}$.

Word $h_2(0010001000100)$ ends 00001000, containing 0000, and cannot be a factor of $\boldsymbol{d}$. □

6 Acknowledgment

The author was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number 2017-03901.

References

[1] J. Berstel, A rewriting of Fife's theorem about overlap-free words, in J. Karhumäki, H. Maurer, and G. Rozenberg, eds., *Results and Trends in Theoretical Computer Science*, Lect. Notes in Comp. Sci., Vol. 812, Springer-Verlag, 1994, pp. 19–29.

[2] J.-P. Allouche, J. Currie, and J. Shallit, Extremal infinite overlap-free binary words, *Electron. J. Combin.* **5** (1998), #R27.

[3] J.-P. Allouche and J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, in C. Ding, T. Helleseth, H. Niederreiter, eds., *Sequences and their Applications, Proceedings of SETA '98*, Springer-Verlag, 1998, pp. 1–16.

[4] A. Carpi, Overlap-free words and finite automata, *Theoret. Comput. Sci.* **115** (1993), 243–260.

[5] J. Cassaigne, Counting overlap-free binary words, in P. Enjalbert, A. Finkel, and K. Wagner, eds., *STACS 93, Proc. 10th Symp. Theoretical Aspects of Comp. Sci.*, Lect. Notes in Comp. Sci., Vol. 665, Springer-Verlag, 1993, pp. 216–225.

[6] D. Damanik, Local symmetries in the period-doubling sequence, *Discrete Appl. Math.* **100** (2000), 115–121.

[7] E. Fife, Binary sequences which contain no *BBb*, *Trans. Amer. Math. Soc.* **261** (1980), 115–136.

[8] N. Guglielmi and V. Protasov, Exact computation of joint spectral characteristics of linear operators. *Found. Comput. Math.* **13** (2013), 37–97.

[9] R. M. Jungers, V. Y. Protasov, and V. D. Blondel, Overlap-free words and spectra of matrices. *Theoret. Comput. Sci.* **410** (38-40), 3670–3684.

[10] Y. Kobayashi, Enumeration of irreducible binary words. *Discrete Appl. Math.* **20** (1988), 221–232.

[11] H. Mousavi, Automatic theorem proving in Walnut. Preprint, available at https://arxiv.org/abs/1603.06017 .

[12] E. Prouhet, Mémoire sur quelques relations entre les puissances des nombres, *C. R. Acad. Sci. Paris* **33** (1851), 225.

[13] N. Rampersad, *Overlap-Free Words and Generalizations*, Ph.D. thesis, University of Waterloo, 2007.

[14] A. Restivo and S. Salemi, Overlap-free words on two symbols, in M. Nivat and D. Perrin, eds., *Automata on Infinite Words*, Vol. 192 of *Lect. Notes in Comp. Sci.*, Springer-Verlag, 1984, pp. 198–206.

[15] A. M. Shur, Binary words avoided by the Thue-Morse sequence, *Semigroup Forum* **53** (1996), pp. 212–219.

[16] N. J. A. Sloane et al., The On-Line Encyclopedia of Integer Sequences, 2023. Available at https://oeis.org.

[17] A. Thue, Über unendliche Zeichenreihen. *Kra. Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl.* **7** (1906), pp. 1–22. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell et al., editors, Universitetsforlaget, Oslo, 1977, pp. 139–158.

[18] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Kra. Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl.* **1** (1912), pp. 1–67. Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell et al., editors, Universitetsforlaget, Oslo, 1977, pp. 413–477.

2020 *Mathematics Subject Classification*: Primary 68R15.
Keywords: combinatorics on words, word avoiding a pattern, period-doubling word, Thue-Morse word, Fife's theorem.

(Concerned with sequences A010060 and A096268.)

Received June 10 2023; revised version received September 18 2023; September 26 2023. Published in *Journal of Integer Sequences*, September 27 2023.

Return to Journal of Integer Sequences home page.