Avoidable Binary Patterns in Partial Words*

F. Blanchet-Sadri[1] Robert Mercaş[2†] Sean Simmons[3]
Eric Weissenstein[4]

October 14, 2010

Abstract

The problem of classifying all the avoidable binary patterns in (full) words has been completely solved (see Chapter 3 of M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, 2002). In this paper, we classify all the avoidable binary patterns in partial words, or sequences that may have some undefined positions called holes. In particular we show that, if we do not substitute any variable of the pattern by a partial word consisting of only one hole, the avoidability index of the pattern remains the same as in the full word case.

Keywords: Combinatorics on words; Partial words; Binary patterns; Avoidable patterns; Avoidability index.

---

*This material is based upon work supported by the National Science Foundation under Grant No. DMS–0754154. The Department of Defense is also gratefully acknowledged. Part of this paper was presented at LATA 2010 [4]. We thank the referees of some preliminary versions of this paper for their very valuable comments and/or suggestions. A World Wide Web server interface has been established at `www.uncg.edu/cmp/research/unavoidablesets4` for automated use of the program.

†Author's work was partially supported by Research Grant No. 1323 U07 E30 N-2008/InvAct/Bel, G./BJ01 of the University Rovira i Virgili.

[1]Department of Computer Science, University of North Carolina, P.O. Box 26170, Greensboro, NC 27402–6170, USA, `blanchet@uncg.edu`

[2]GRLMC, Departament de Filologies Romàniques, Universitat Rovira i Virgili, Av. Catalunya 35, Tarragona, 43002, Spain, `robertmercas@gmail.com`

[3]Department of Mathematics, The University of Texas at Austin, 2515 Speedway Rm 8, Austin, TX 78712–0233, USA

[4]Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Amos Eaton 301, 110 8th Street, Troy, NY 12180, USA

1 Introduction

A *pattern* p is a word over an alphabet E of *variables*, denoted by $\alpha, \beta, \gamma, \ldots$, and the associated set, over a finite alphabet A, is built by replacing p's variables with non-empty words over A, so that the occurrences of the same variable be replaced with the same word.

The concept of *unavoidable pattern*, see Section 2, was introduced, in the context of (full) words, by Bean, Ehrenfeucht and McNulty [1] (and by Zimin who used the terminology "blocking sets of terms" [15]). Although they characterized such patterns (in fact, avoidability can be decided using the Zimin algorithm by reduction of patterns), there is no known characterization of the patterns unavoidable over a k-letter alphabet (also called k-unavoidable). An alternative is to find all unavoidable patterns for a fixed alphabet size. The unary patterns, or powers of a single variable α, were investigated by Thue [13, 14]: α is unavoidable, $\alpha\alpha$ is 2-unavoidable but 3-avoidable, and α^m with $m \geq 3$ is 2-avoidable. Schmidt proved that there are only finitely many binary patterns, or patterns over $E = \{\alpha, \beta\}$, that are 2-unavoidable [11, 12]. Later on, Roth showed that there are no binary patterns of length six or more that are 2-unavoidable [10]. The classification of unavoidable binary patterns was completed by Cassaigne [5] who showed that $\alpha\alpha\beta\beta\alpha$ is 2-avoidable.

In this paper, our goal is to classify all binary patterns with respect to partial word avoidability. A partial word is a sequence of symbols from a finite alphabet that may have some undefined positions, called holes, and denoted by $\diamond$'s. Here $\diamond$ is *compatible* with, or matches, every letter of the alphabet. In this context, in order for a pattern p to occur in a partial word, it must be the case that for each variable α of p, all its substituted partial words be pairwise compatible.

The contents of our paper is as follows: In Section 2, we recall some basic definitions regarding partial words, and define our terminology on unavoidable patterns in this context. In Section 3, we start our investigation of avoidability of binary patterns in partial words. There, using iterated morphisms, we construct binary partial words with infinitely many holes that avoid the patterns $\alpha\beta\alpha\beta\beta\alpha$ and $\alpha\alpha\beta\alpha\beta\beta$. In Section 4, using non-iterated morphisms, we construct such words that avoid the patterns $\alpha\alpha\beta\beta\alpha$ and $\alpha\beta\alpha\alpha\beta$. In Section 5, by proving that over a binary alphabet there exist partial words with infinitely many holes that non-trivially avoid the pattern $\alpha\beta\alpha\beta\alpha$, we reach one of our goals of proving that all binary patterns 2-avoidable in full words are also non-trivially 2-avoidable in partial words (an occurrence of a pattern is *non-trivial* if none of its variables is substituted by a partial word consisting of only one hole). In fact, our results show that the so-called avoidability index of a binary pattern remains unchanged in the context of non-trivial avoidability in partial words. Finally in Section 6, we conclude by characterizing almost all binary patterns in terms of (not restricted to non-trivial) avoidability in partial words.

2 Preliminaries

For more information regarding concepts on partial words, the reader is referred to [2]. Cassaigne's Chapter 3 of [8] provides background on unavoidable patterns in the context of full words.

Let A be a non-empty finite set of symbols called an *alphabet*. Each element $a \in A$ is called a *letter*. A *(full) word* over A is a sequence of letters from A. A *partial word* over A is a sequence of symbols from $A_\diamond = A \cup \{\diamond\}$, the alphabet A being augmented with the "hole" symbol $\diamond$ (a full word is a partial word without holes). We denote by $u(i)$ the symbol at position i of a partial word u. The *length* of u is denoted by $|u|$ and represents the number of symbols in u. The *empty word* is the sequence of length zero and is denoted by ε. The set containing all full words (respectively, non-empty full words) over A is denoted by A^* (respectively, A^+), while the set of all partial words (respectively, non-empty partial words) over A is denoted by $A_\diamond^*$ (respectively, $A_\diamond^+$).

For a partial word u, the powers of u are defined recursively by $u^0 = \varepsilon$ and for $n \geq 1$, $u^n = uu^{n-1}$. Furthermore, $\lim_{n\to\infty} u^n$ is denoted by u^ω.

If u and v are two partial words of equal length, then u is said to be *contained in* v, denoted $u \subset v$, if $u(i) = v(i)$ for all i such that $u(i) \in A$. Partial words u and v are *compatible*, denoted $u \uparrow v$, if there exists a partial word w such that $u \subset w$ and $v \subset w$. If u, v are non-empty compatible partial words, then uv is called a *square*.

A partial word u is a *factor* of a partial word v if there exist x, y such that $v = xuy$ (the factor u is *proper* if $u \neq \varepsilon$ and $u \neq v$). We say that u is a *prefix* of v if $x = \varepsilon$ and a *suffix* of v if $y = \varepsilon$.

Let E be a non-empty finite set of symbols, distinct from A, whose elements are denoted by α, β, γ, etc. Symbols in E are called *variables*, and words in E^* are called *patterns*. The *pattern language*, over A, associated with a pattern $p \in E^*$, denoted by $p(A_\diamond^+)$, is the subset of $A_\diamond^*$ containing all partial words compatible with $\varphi(p)$, where φ is any non-erasing morphism from E^* to A^* that maps every variable in E to an arbitrary non-empty word. A partial word $w \in A_\diamond^*$ *meets* the pattern p (or p *occurs* in w) if for some factorization $w = xuy$, we have $u \in p(A_\diamond^+)$. Otherwise, w *avoids* p.

To be more precise, let $p = \alpha_0 \cdots \alpha_m$, where $\alpha_i \in E$ for $i = 0, \ldots, m$. Define an *occurrence* of p in a partial word w as a factor $u_0 \cdots u_m$ of w, where for all $i, j \in \{0, \ldots, m\}$, if $\alpha_i = \alpha_j$, then $u_i \uparrow u_j$. Stated differently, for all $i \in \{0, \ldots, m\}$, $u_i \subset \varphi(\alpha_i)$, where φ is any non-erasing morphism from E^* to A^* as described earlier. These definitions also apply to (one-sided) infinite partial words w over A which are functions from $\mathbb{N}$ to $A_\diamond$.

Considering the pattern $p = \alpha\beta\beta\alpha$, the language associated with p over the alphabet $\{a, b\}$ is $p(\{a, b, \diamond\}^+) = \{u_1v_1v_2u_2 \mid u_1, u_2, v_1, v_2 \in \{a, b, \diamond\}^+$ such that $u_1 \uparrow u_2$ and $v_1 \uparrow v_2\}$. The partial word $ab\diamond ba\diamond bba$ meets p (take $\varphi(\alpha) = bb$ and $\varphi(\beta) = a$), while the word $\diamond babbbaaab\diamond$ avoids p.

Let p and p' be two patterns. If p' meets p, then p *divides* p', which we denote by $p \mid p'$. For example, $\alpha\alpha \nmid \alpha\beta\alpha$ but $\alpha\alpha \mid \alpha\beta\alpha\beta$. When both $p \mid p'$ and $p' \mid p$ hold, the patterns p and p' are *equivalent*, and this happens when and only when they differ by a permutation of E. For instance, $\alpha\alpha$ and $\beta\beta$ are equivalent.

Remark 1. *Let $p, p' \in E$ be such that p meets p'. If an infinite partial word avoids p', then it also avoids p.*

A pattern $p \in E^*$ is *k-avoidable* if there are infinitely many partial words in $A_\diamond^*$ with h holes, for any integer $h > 0$, that avoid p, where A is any alphabet of size k. Note that if there is a partial word over A with infinitely many holes that avoids p, then p is obviously k-avoidable. On the other hand, if, for some integer $h \geq 0$, every long enough partial word in $A_\diamond^*$ with h holes meets p, then p is *k-unavoidable* (it is also called unavoidable over A). Finally, a pattern $p \in E^*$ which is k-avoidable for some k is simply called *avoidable*, and a pattern which is k-unavoidable for every k is called *unavoidable*. The *avoidability index* of p is the smallest integer k such that p is k-avoidable, or is ∞ if p is unavoidable.

For the remaining of this paper, we only consider binary patterns, hence we can fix $E = \{\alpha, \beta\}$. We define $\overline{\alpha} = \beta$ and $\overline{\beta} = \alpha$ as being the *complements* of α and β respectively. The *complement* of a pattern p over E is then the pattern $\overline{p}$ built by complementing each variable of p. Similarly, if $A = \{a, b\}$, we can define $\overline{a} = b$ and $\overline{b} = a$, as well as the complement of a word over A.

In the context of full words all binary patterns' avoidability indices have been characterized.

Theorem 1 ([8])**.** *For full words, binary patterns fall into three categories:*

1. *The binary patterns $\varepsilon, \alpha, \alpha\beta, \alpha\beta\alpha$, and their complements, are unavoidable (or have avoidability index ∞).*

2. *The binary patterns $\alpha\alpha$, $\alpha\alpha\beta$, $\alpha\alpha\beta\alpha$, $\alpha\alpha\beta\beta$, $\alpha\beta\alpha\beta$, $\alpha\beta\beta\alpha$, $\alpha\alpha\beta\alpha\alpha$, $\alpha\alpha\beta\alpha\beta$, their reverses, and complements, have avoidability index 3.*

3. *All other binary patterns, and in particular all binary patterns of length six or more, have avoidability index 2.*

Theorem 1 gives us a lower bound for the avoidability index of a binary pattern in the context of partial words. Indeed, since a full word is a partial word without holes, the avoidability index of a binary pattern in full words is not greater than the avoidability index of that pattern in partial words. So all the binary patterns in Theorem 1(1) have avoidability index ∞ in partial words.

In Section 3, we prove that the patterns $\alpha\alpha\beta\alpha\beta\beta$ and $\alpha\beta\alpha\beta\beta\alpha$ are 2-avoidable by iterated morphisms, while in Section 4 that $\alpha\beta\alpha\alpha\beta$ and $\alpha\alpha\beta\beta\alpha$ are 2-avoidable by non-iterated morphisms. Similar results were shown in [5, 10] in the context of full words (in fact, in full words, the patterns $\alpha\beta\alpha\alpha\beta$ and $\alpha\alpha\beta\beta\alpha$ are 2-unavoidable by iterated morphisms, but are 2-avoidable). In Section 5, we show that the pattern $\alpha\beta\alpha\beta\alpha$ is non-trivially 2-avoidable by iterated morphisms.

3 Binary Patterns 2-Avoidable by Iterated Morphisms

Let ν be the morphism that maps a to aab and b to bba. Define the sequence produced by ν as $t_0 = a$, and $t_n = \nu(t_{n-1})$. Recall that $t = \nu^\omega(a)$ avoids $\alpha\beta\alpha\beta\beta\alpha$ and $\alpha\alpha\beta\alpha\beta\beta$ [8].

Lemma 1. *For any $n \geq 0$, $t_{n+1} = t_n t_n \overline{t_n}$, where t_i is the ith iteration of the sequence produced by ν.*

Theorem 2. *Over a binary alphabet there exist infinitely many partial words, containing infinitely many holes, that avoid the pattern $\alpha\beta\alpha\beta\beta\alpha$.*

Proof. Let $p = \alpha\beta\alpha\beta\beta\alpha$, and let $t = \nu^\omega(a)$ be the fixed point of the morphism ν. Denote by t'_5 the word obtained by replacing the letter b at position 58 in t_5 by $\diamond$, and by t' the word where infinitely many non-overlapping occurrences of the factor t_5 have been replaced by t'_5:

$$\begin{aligned} t'_5 = {} & aabaabbbaaabaabbbabbabbaaabaabaabbbaaabaabbbabbabbaaabbbab\underline{\diamond}aaabb \\ & babbaaabaabaabbbaaabaabbbaaabaabbbabbabbaaabaabaabbbaaabaabbbab \\ & babbaaabbbabbaaabbbabbaaabaabaabbbabbabbaaabbbabbaaabaabaabbbabb \\ & abbaaabbbabbaaabaabaabbbaaabaabbbaaabaabbbabbabbaaab \end{aligned}$$

Assume, to get a contradiction, that the pattern $\alpha\beta\alpha\beta\beta\alpha$ occurs somewhere in t'. So there exists a factor $x_1y_1x_2y_2y_3x_3$ in t' that contains h holes, with the x_i's and the y_i's pairwise compatible for all $i \in \{1,2,3\}$, and no occurrence of p with less than h holes exists. Let $x_1, x_2, x_3 \subset x$ and $y_1, y_2, y_3 \subset y$ for some non-empty x, y. It is not hard to verify with a computer program that $|x| \geq 9$ or $|y| \geq 9$, since a hole is more than 58 positions far from either end of t'_5.

If $|x| > 4$ and there exists a hole in x_i, for $0 < i \leq 3$, then there exists x_j, with $j \neq i$, that has a factor that is compatible with a word from $\{\underline{a}aaab, b\underline{a}aaa, ab\underline{a}aa, bab\underline{a}a, bbab\underline{a}\}$ (note that the underlined letter is the one that corresponds to the hole in x_i). Note that it is impossible to have a hole at another position than the underlined one, in any of the previously mentioned factors. We conclude that $x_i = x_j$ since, otherwise, we have that t contains one of the factors $aaaa$, $abaaa$ or $baba$, which is a contradiction. The same proof works for y_i, where $0 < i \leq 3$. It follows that either t does not avoid the pattern or there exists in t' an occurrence of the pattern with less than h holes, which are both contradictions.

Thus, either $|x| \leq 4$ and y_i contains no holes for $0 < i \leq 3$, or $|y| \leq 4$ and x_i contains no holes for $0 < i \leq 3$ (otherwise we have that $|x_1y_1x_2y_2y_3x_3| \leq 24$ contains more than two holes, which is a contradiction since between each two holes there are at least 72 symbols according to our construction).

Let us first assume that a hole is in one of x_1, x_2 or x_3. If x_1 contains a hole then, since $|x| \leq 4$ and y_1 contains no hole, looking at the factor following $\diamond$ we conclude that the corresponding position in x_2 must also contain a hole. Now, if x_2 contains a hole then, it follows from the previous observation that x_1 has to contain a hole, and moreover, since y_1 and y_3 contain no holes, looking at the factor preceding the hole, we get that x_3 has a hole at the corresponding position. Finally, if x_3 contains a hole, according to the previous observation, x_2 has a hole. We conclude that if x_i has a hole, then $x_i = x_j$, for all $i, j \in \{1,2,3\}$. Hence, there exists an occurrence of p having no holes, a contradiction.

Since the case when a hole is in one of y_1, y_2 or y_3 is similar, we conclude that t' does not contain any occurrence of the pattern $\alpha\beta\alpha\beta\beta\alpha$. $\square$

Remark 2. *If uu is a factor of the fixed point of the morphism ν, for some word u with $|u| > 3$, it must be that $|u| \equiv 0 \bmod 3$. Moreover, for all different occurrences of the same factor v with $|v| > 3$, there exist unique words x, y, z such that $v = x\nu(y)z$, with $|x|, |z| < 3$. In other words, all occurrences of the same factor start at the same position of an iteration of ν.*

Theorem 3. *Over a binary alphabet there exist infinitely many partial words, containing infinitely many holes, that avoid the pattern $\alpha\alpha\beta\alpha\beta\beta$.*

Proof. Let $p = \alpha\alpha\beta\alpha\beta\beta$, t be again the fixed point of the morphism ν, and denote by t'_5 the word obtained from t_5 by replacing position 57 by $\diamond$:

$t'_5 = aabaabbbaaabaabbbabbabbaaabaabaabbbaaabaabbbabbabbaaabbba\underline{\diamond}baaabb$
$babbaaabaabaabbbaaabaabbbaaabaabbbabbabbaaabaabaabbbaaabaabbbab$
$babbaaabbbabbaaabbbabbaaabaabaabbbabbabbaaabbbabbaaabaabaabbbabb$
$abbaaabbbabbaaabaabaabbbaaabaabbbaaabaabbbabbabbaaab$

Moreover, denote by t' the word obtained by replacing in t infinitely many non-overlapping occurrences of t_5 by t'_5.

We proceed by contradiction. Assume there does exist an occurrence $x_1x_2y_1x_3y_2y_3$ of $\alpha\alpha\beta\alpha\beta\beta$ in t' containing $h > 0$ holes, and none exists with less than h holes. Let $x_1, x_2, x_3 \subset x$ and $y_1, y_2, y_3 \subset y$ for some non-empty $x, x_1, x_2, x_3, y, y_1, y_2, y_3$. Note that it is enough to consider the case when some x_i contains a hole, the case of some y_i containing a hole is symmetrical. By putting a hole at position 57 of t_5, we get that $|x_1x_2y_1x_3y_2y_3| \geq 54$ (it is not hard to verify this with a computer program since we have put a hole more than 54 positions from either end of t_5 to create t'_5). The only cases that need to be checked are those where either $|x| \geq 9$ or $|y| \geq 9$.

Assume that for some $0 < i \leq 3$, x_i contains a hole. If the hole is not at the last position, and $|x| > 6$ then, we get that x_j, for all $j \neq i$ and $j \in \{1, 2, 3\}$, contains a hole at the same position as x_i, and we get a contradiction with the fact that an instance of the pattern having less than h holes exists (this is easily done by looking at all possible factors).

Let us now assume that $|x| \leq 6$. We have that either no y_j has a hole, or if one does, then the hole is at the last position and $|y| > 73$ (otherwise $x_1x_2y_1x_3y_2y_3$ has length at most 36 and contains at least two holes, a contradiction with the way t' is constructed).

If either x_1 or x_2 has a hole, we can look for all the factors containing a hole in t', and having length at most 12. It can be checked that the only squares that appear are $bba\ \underline{b}ba$ and

$$\begin{array}{l} bbabbaaab\ bba\underline{b}baaab \\ babbaaabb\ ba\underline{b}baaabb \\ abbaaabbb\ a\underline{b}baaabbb \\ bbaaabbba\ \underline{b}baaabbba \\ baaabbba\underline{b}\ baaabbbab \\ aaabbba\underline{b}b\ aaabbbabb \\ aabbba\underline{b}ba\ aabbbabba \\ abbba\underline{b}baa\ abbbabbaa \end{array}$$

where the underlined symbol is represented by the hole. It follows that there exists an occurrence of p containing less than h holes, a contradiction (not changing the b into a hole makes no difference). If x_3 has a hole, then, according to the above, neither x_1 nor x_2 has holes. Since both prefixes of length 5 of y_1 and y_2 consist of full words, the hole in x_3 is followed by $baaa$. Thus, the last position in x_2 is followed by $baaa$, and so it has a b. It follows that the hole in x_3 is not needed and so we get a contradiction with h being the minimum number of holes an occurrence of p has.

The last case that needs to be considered is when $|x| > 6$ and there are holes at the last positions of some of the x_i's. If $|y| = 1$, since all x_i's start with $baaa$, for $0 < i \leq 3$, it follows that $y = b$, and we get the factor $x_1x_2bx_3bb$. Moreover, the x_i's differ only in the last position so we have $x'c_1x'c_2bx'c_3bb$, for some word $|x'| > 70$ and some $c_1, c_2, c_3 \in \{a, \diamond\}$. In this case using Remark 2, we reach a contradiction.

Now let $|y| > 1$. If x_1 has a hole it follows that all x_i's start with $baaa$, for $0 < i \leq 3$. Thus, y_i ends in b, for all $0 < i \leq 3$. If any of the y_i's end in $\diamond$, then an occurrence of the pattern with less than h holes exists, a contradiction. Furthermore, all y_i's start with $baaa$. But this implies that both x_2 and x_3 end in b or $\diamond$, a contradiction with the fact that there does not exist an occurrence of p with less than h holes.

If x_2 or x_3 has a hole, it follows that all y_i's start with $baaa$ and all x_i's start with ab, where $0 < i \leq 3$. Moreover, it follows that y_2 ends in b or $\diamond$. But, since x_3 starts with ab, and so y_1 ends in a, we get that y_2 ends in $\diamond$. Thus,

$$x_1x_2y_1x_3y_2y_3 = abx'a\ abx'\diamond\ baaay'a\ abx'\diamond\ baaay'\diamond\ baaay'c$$

with $c \in \{a, \diamond\}$. But this implies that x' represents an image of ν, and so, x_1 is preceded by an a. Thus, t_5 has as a factor

$$aabx'\ aabx'\ \underline{b}baaay'\ aabx'\ \underline{b}baaay'\ \underline{b}baaay'$$

with the underlined b's standing for the ones changed into $\diamond$'s in t'. This is a contradiction since we know that t_5 avoids the pattern $\alpha\alpha\beta\alpha\beta\beta$. The cases when the y_i's end in $\diamond$'s are similar. □

4 Binary Patterns 2-Avoidable by Non-Iterated Morphisms

Let us now look at the pattern $\alpha\alpha\beta\beta\alpha$. Let $A = \{a, b\}$ and $A' = A \cup \{c\}$, and let $\psi : A'^* \to A'^*$ be the morphism defined by $\psi(a) = abc$, $\psi(b) = ac$ and $\psi(c) = b$. It is well known that $\psi^\omega(a)$ avoids $\alpha\alpha$, in other words it is square-free [8].

Furthermore, define the morphism $\chi : A'^* \to A^*$ with $\chi(a) = aa$, $\chi(b) = aba$, and $\chi(c) = abbb$. If $w = \chi(\psi^\omega(a))$, then we know from [5] that w does not contain any occurrence of $\alpha\alpha\beta\beta\alpha$. Moreover, let $\chi_4 = \chi(\psi^4(a))$. Since $\psi^4(a)$ occurs infinitely often as a factor of $\psi^\omega(a)$, it follows that χ_4 occurs infinitely often as a factor of w. Hence, we can write $w = w_0\chi_4 w_1 \chi_4 w_2 \chi_4 \cdots$, for some words w_i with $|w_i| > 1$, for all i.

Now, let us replace the a in position 23 of χ_4 by $\diamond$ and denote the new partial word by χ_4':

$$\chi_4' = aaabaabbbaaabbbabaaaaba\underline{\diamond}bbbabaaaabbbaaabaabbbaaabbb\\ abaaaabbbaaabaabbbaba$$

Lemma 2. *Let u_1u_2 denote a factor of $w' = w_0\chi_4' w_1 \chi_4' w_2 \chi_4' \cdots$, that is obtained by inserting holes in v_1v_2, a factor of w with $u_1 \subset v_1$ and $u_2 \subset v_2$. If $u_1 \uparrow u_2$, but $v_1 \neq v_2$, then $|u_1| \leq 4$, more specifically, either u_1 or u_2 is in $\{\diamond, \diamond b, \diamond bb, a\diamond, a\diamond b, a\diamond bb, ba\diamond\}$.*

Proof. Obviously, if $u_1 \uparrow u_2$ but $v_1 \neq v_2$, then a hole appears in u_1 and there is no hole at the corresponding position in u_2, or vice versa. Without loss of generality we can assume that a hole appears in u_1. Assume that $u_1 \notin \{\diamond, \diamond b, \diamond bb, a\diamond, a\diamond b, a\diamond bb, ba\diamond\}$. It follows that u_1 has as a factor $\diamond bbb$, $aba\diamond$ or $ba\diamond b$. Moreover, note that the only occurrence of bbb in w' is in the factors $abbba$ and $\diamond bbba$. Similarly, the only time aba appears as a factor of w' is as a factor of $abaa$ and $aba\diamond$, and the only time a word x compatible with $ba\diamond b$ appears in w' is when $x = ba\diamond b$ or $x = baab$. We see that in all of these cases the corresponding factor in u_2 must be $abbba$, $abaa$ or $baab$, a contradiction since we always have $v_1 = v_2$. □

Lemma 3. *There exists no factor uu of $w = \chi(\psi^\omega(a))$, such that either $aaab$ or aba is a prefix of u.*

Proof. Assume there exists an u with prefix $aaab$ so that $uu = w(i) \cdots w(i+2l-1)$, for some integers i, l with $l > 3$. Note that $aaab$ only appears as a prefix of $\chi(x)$, for some word $x \in \{a, b, c\}^+$. Moreover, since the second u also starts with $aaab$, we have that $u = \chi(x)$. Hence, uu is actually $\chi(xx)$, for some word $x \in \{a, b, c\}^+$. It follows that $\phi^\omega(a)$ contains the square xx, which is a contradiction with the nature of the ψ morphism. Similarly, aba only appears as an image of $\chi(b)$. □

Theorem 4. *Over a binary alphabet there exist infinitely many partial words, containing infinitely many holes, that avoid the pattern $\alpha\alpha\beta\beta\alpha$.*

Proof. Let $p = \alpha\alpha\beta\beta\alpha$. To prove this claim, we assume that the partial word w' is not p-free and get a contradiction. Let $x'_1x'_2y'_1y'_2x'_3$ be an occurrence of p in w', and denote by $x_1x_2y_1y_2x_3$ the factor of the full word w in which holes were inserted to get p. Note that if $x_1 = x_2 = x_3$ and $y_1 = y_2$, then we have an occurrence of p in w, which is a contradiction. Therefore one of the equalities fails. Also, note that if $x_i \neq x_j$ then either x'_i or x'_j contains a hole, where $i, j \in \{1, 2, 3\}$, while if $y_1 \neq y_2$ then either y'_1 or y'_2 contains a hole. Moreover, if $x_1 \neq x_2$ or $y_1 \neq y_2$, according to Lemma 2, x'_1 or x'_2 or, y'_1 or y'_2 is in $\{\diamond, \diamond b, \diamond bb, a\diamond, a\diamond b, a\diamond bb, ba\diamond\}$. By looking at the factor χ'_4, it is easy to check that the only possibilities are for x'_1 or y'_1 to be in $\{\diamond, \diamond b\}$ and for x'_2 or y'_2 to be in $\{\diamond, a\diamond, a\diamond b\}$.

If $y'_1 = \diamond b$, it is easy to check that x'_3 must start with aba. According to Lemma 3, we cannot have that $x_1 = x_2$. It follows, according to Lemma 2, that $|x'_1| \leq 4$, a contradiction with the factor preceding the hole. The proof is identical for the case when $x'_1 = \diamond b$. If $y'_1 = \diamond$, then x'_3 starts with bba. Thus, x'_2 starts with bba and x'_1 ends in ab or $\diamond b$. It follows that x'_2 ends in ab or $\diamond b$, thus, y'_1 is preceded by ab, which is a contradiction. If $x'_1 = \diamond$, then y'_1 and y'_2 start with bba and x'_3 is a b. From Lemma 2, $y_1 = y_2$. Thus, y_1y_2bb is a factor of w. It follows that for some word $z \in \{a, b\}^+$, $y_1y_2bb = bb\chi(zc)\chi(zc)$ is a factor of w. We get a contradiction with the fact that $\psi^\omega(a)$ is square-free.

If x'_2 or y'_2 is in $\{\diamond, a\diamond, a\diamond b\}$, then either y'_1 or x'_3 is a factor of a word in $\{bbba, bba\}$. A contradiction is reached again with the help of Lemma 2 and the fact that $\psi^\omega(a)$ is square-free.

The final case that needs to be analyzed is when $x_1 = x_2$ and $y_1 = y_2$, and $x'_3 \uparrow x'_1$ and $x_3 \neq x_1$. Let us denote $x = x_1 = x_2$ and $y = y_1 = y_2$. We get that w has $xxyyx_3$ as a factor and there exists at least one hole in x_3' that corresponds to b's in x_1' and x_2'.

If $|x| > 4$ it follows that x has $aba\underline{b}$, $ba\underline{b}b$, $\underline{b}bbb$ as a factor (the underlined b represents the letter corresponding to the hole in x_3'). Since none of these are possible factors of w, we conclude that it is impossible. Hence, $x_3' \in \{\diamond, \diamond b, \diamond bb, a\diamond, a\diamond b, a\diamond bb, ba\diamond\}$. It follows that

$$xx \in \{b^2, (bb)^2, (bbb)^2, (ab)^2, (abb)^2, (abbb)^2, (bab)^2\}$$

But, only bb is a possible factor of w. It is easy to check that in this case $|y| > 6$ and we conclude that y has either aaa, aba, $baaa$ or $baba$ as a prefix. In all of these cases, using Lemma 3 we reach a contradiction.

Since all cases lead to contradictions we conclude that $\alpha\alpha\beta\beta\alpha$ is avoidable over a binary alphabet. □

Finally, let us look at the pattern $\alpha\beta\alpha\alpha\beta$. According to [8, Lemma 3.3.2], $\gamma(\psi^\omega(a))$ avoids $\alpha\beta\alpha\alpha\beta$, where $\gamma : \{a, b, c\}^* \to \{a, b\}^*$ with $\gamma(a) = aaa$, $\gamma(b) = bbb$, and $\gamma(c) = ababab$. Moreover, the only squares that occur in $w = \gamma(\psi^\omega(a))$ are $a^2, b^2, (aa)^2, (ab)^2, (ba)^2, (bb)^2$ and $(baba)^2$. Finally, note that $\psi^\omega(a)$ does not contain any of the factors aba or cbc.

Let us replace the b in position 84 of $\gamma_5 = \gamma(\psi^5(a))$ by $\diamond$ and denote the new partial word by γ_5':

$$\begin{aligned}\gamma_5' = & aaabbbabababaaaabababbbbaaabbbabababbbbaaaabababaaabbbabababaaaa\\ & bababbbbaaaabababaaa\underline{\diamond}bbabababbbbaaabbbabababaaaabababbbbaaabbbab\\ & ababbbbaaaabababaaabbbabababbbbaaabbbabababaaaabababbbbaaaababab\end{aligned}$$

Moreover, let us denote by w' the word obtained from w after the insertion of a hole at position 84 of an occurrence of γ_5.

Proposition 1. *Over a binary alphabet there exist infinitely many infinite partial words, containing exactly one hole, that avoid the pattern $\alpha\beta\alpha\alpha\beta$.*

Proof. Let us assume, to get a contradiction, that there exists an occurrence of $p = \alpha\beta\alpha\alpha\beta$ in w', and denote this occurrence by $x_1y_1x_2x_3y_2$, for $x_i, y_j \in \{a, b, \diamond\}^+$, $0 < i \leq 3$ and $0 < j < 3$. Either $x_1 = x_2 = x_3 = x$ or $y_1 = y_2 = y$, for some words x and y, but not both. Note that if the variable containing the hole has length greater than 3 then the hole is among the last three symbols. Otherwise the corresponding variable has as a factor $\underline{a}bba$ (the underlined a stands for the hole position in the first variable), which is a contradiction.

Case 1. The factor with the hole has length greater than three.

Here, $x_1 = x$ (otherwise w has the factor xx with $|x| > 3$, which is a contradiction).

If the hole is in y_1, then x must be one of the squares aforementioned. Moreover, since the hole is among the last three symbols of y_1, by looking at the symbols following the hole, $x \in \{b, ab\}$. If $x = b$, then $by'\diamond bby'a$ with $y_1 = y'\diamond$ for some word y'. But, since $|y| > 3$, we get that y ends in aaa, and it follows that $y = \gamma(cz)$, for some word $z \in \{a, b, c\}^*$. We get that $\psi^\omega(a)$ either contains $cczbczc$ or $bczbczc$ as a factor. Since both of these contain squares and $\psi^\omega(a)$ is square-free, we reach a contradiction.

If the hole is in y_2, then for some prefix y' of y we get

$$xy_1xxy_2 \in \{xy'axxy'\diamond, xy'abxxy'\diamond b, xy'abbxxy'\diamond bb\}$$

If $y_1 = y'a$, since $|y| > 3$, y' ends in aaa. It follows that $x = ba$. Hence, we have the factor $bay'ababay'\diamond$. In this case $y' = b\gamma(z)$, for some word z, and so $\psi^\omega(a)$ has $c\gamma(z)c\gamma(z)$ as a factor, a contradiction with the fact that it is square-free. If $y_1 = y'abb$, then $x \in \{ba, baba\}$. It is easy to check that in this case we must have $|y'| > 2$. We get that w has $aaaabb$ as a factor, a contradiction. Finally, if $y_1 = y'ab$, then y' ends in aa. We obtain $x = ab$. But then we have that $xy'abxxy'\diamond b = aby'abababy'\diamond b$. Hence, we reach a contradiction with the fact that $\psi^\omega(a)$ cannot have $czcz$ as a factor, where $y' = \gamma(z)$ for some word z.

If the hole is in x_3, then since $x_1y \uparrow x_3y$, the hole must be within the last two symbols of x_3 (otherwise we once again have in w the factor $abba$). It follows that $x_1 = x'aaax''$, for some words $x', x'' \in \{a, b\}^*$. We get that $xyxx_3y = x'aaax''yx'aaax''x'aa\diamond x''y$. It follows that $|x''x'| > 3$, and since $|x''| < 2$, x' has aba as a suffix. Moreover, if $|x''y| > 1$, then $x''y$ has bb as a prefix, and so w has $aaaabb$ as a factor, a contradiction. Thus, we conclude that $x'' = \varepsilon$ and $y = b$. We get the factor $x'aaabx'aaax'aa\diamond b$. But since ba is a suffix of x', ba is a prefix of x'. We also get that $abba$ is a factor of w, a contradiction.

If the hole is in x_2, then we have the factor xyx_2xy. Denoting $x = x'ax''$, for some words $x', x'' \in \{a, b\}^*$, we obtain the factor $x'ax''yx'\diamond x''x'ax''y$, where $|x'x''| > 2$. It can be checked that in this case x' ends in aaa. It follows that $ax''yx'$ and $bx''x'$ are both the image of some words through γ. We get that $x'' = \varepsilon$, $ay = \gamma(z_1)b$ and $bx' = \gamma(z_2)$, for some words $z_1, z_2 \in \{a, b, c\}$. In addition, $bxyx_2xy$ is obtained by introducing a hole in $\gamma(z_2)\gamma(z_1)\gamma(z_2)\gamma(z_2)\gamma(z_1)$. This is a contradiction since w avoids the pattern $\alpha\beta\alpha\alpha\beta$.

Case 2. The factor with the hole has length at most three.

It is easy to check that in this case we cannot have a hole in y_1 or y_2, and we reach contradictions whenever we consider a hole in x_2 or x_3. If the hole is in x_1, then $x \in \{a, aa, ab\}$. In all the cases we get contradictions because of possible forms of y (we get contradictions because y_1 starts either with ba or bba). □

Theorem 5. *Over a binary alphabet there exist infinitely many partial words, containing infinitely many holes, that avoid the pattern $\alpha\beta\alpha\alpha\beta$.*

Proof. Let us denote by w' the word obtained from w after infinitely many non-overlapping occurrences of γ_5 starting at an even position have been replaced by γ_5'. Furthermore, let us assume, to get a contradiction, that the pattern $p = \alpha\beta\alpha\alpha\beta$ is unavoidable and denote by $x_1y_1x_2x_3y_2$ an occurrence of p containing $h > 1$ holes, such that no occurrence of the pattern p having less than h holes appears in w'. Here, we assume that $x_1, x_2, x_3 \subset x$ and $y_1, y_2 \subset y$ for some x, y.

Since, according to Proposition 1, $h > 1$ and the distance between every two holes is at least 170, it follows that $|xy| > 85$. Thus, there exist $z, z' \in \{x_1, x_2, x_3, y_1, y_2\}$, $z \neq z'$, $z \uparrow z'$, $z = z_1 \diamond z_2$ and $z' = z_1' a z_2'$, for some words z_i, z_i', with $z_i \uparrow z_i'$ for $i = 1, 2$. If $|z_2| > 2$, then z_2' has a prefix compatible with bba. Since the only factor in w' compatible with bba is bba we conclude that az_2' has $abba$ as a prefix, which is a contradiction with the fact that $abba$ is a valid factor of w. It follows that $|z_2| < 3$.

Moreover, if $z \in \{x_1, x_3\}$, since y_1 follows x_1 and y_2 follows x_3, we get $x_1y_1 \uparrow x_3y_3$. If $|y| > 2$ a conclusion similar to the previous one is reached. It follows that $0 < |z_2y_1| < 3$. In this case, $|z_1| > 82$ and the hole is preceded by $abaaa$. If $z_2 \neq \varepsilon$ or $|y_1| > 1$, then w' has $baaaabb$ as a factor, a contradiction. It must be that $z = z_1\diamond$ and $y = b$. Since $abaaa$ is a suffix of z_1, by looking at x_2, we get that z_1 has as a prefix a factor compatible with $bababb$. So $\diamond$ is followed by a factor compatible with $bbababb$, a contradiction.

If $x_2 = z$, we get the factor $z_1az_2y_1z_1'\diamond z_2'z_1az_2y_2$. Note that here we denote $x_1 = z_1az_2 = x_3$, since otherwise we fall in the previous case. Moreover, it is easy to check that in this case $|z_1| > 10$ (there is no square that contains a in the first part and has length less than 2×13). It follows that $aababababaaa$ is a suffix of z_1, and since it is the only word compatible with a factor of w' it is also a suffix of z_1'. It follows that z_2y_1 and z_2y_2 are both preceded by $aababababaaaa$. Thus, $z_2y_1z_1$ starts with $baba$, and so, w has $\gamma(acac)$ as a factor, a contradiction with the fact that $\psi^\omega(a)$ is square-free.

It follows that $x_1 = x_2 = x_3$, and x_1, x_2, x_3 are compatible with words from the set $\{a, b, aa, ab, ba, bb, baba\}$. Since the number of holes is minimal, we can only consider the factors $xz_1\diamond z_2xxz'_1az'_2$ and $xz'_1az'_2xxz_1\diamond z_2$, where $|z_2| < 3$.

If the factor is $xz_1\diamond z_2xxz'_1az'_2$, looking at the letters following the hole and the possible values of x, we conclude that we have either $bz_1\diamond bbz'_1a$, $baz_1\diamond bbabaz'_1ab$ or $abz_1\diamond bbababz'_1abb$. In the second case, it follows that z'_1, and respectively z_1, starts with ab, giving us the prefix $baab$, which is not a valid factor of w'. In the last case, we get that z_1, and respectively z'_1, ends in aaa, giving us the suffix $aaaabb$, which is not a valid factor of w'. Finally, in the case of $bz_1\diamond bbz'_1a$, we get that z_1 is preceded by a factor compatible with bbb. It follows that the factor that we have before changing the b following z_1 into a hole is $bbbz_1bbbz'_1a$. In this case, taking $\alpha = b$ and $\beta = bz_1$, we get an occurrence of the pattern $\alpha\beta\alpha\alpha\beta$, for less than h holes, a contradiction.

Finally, let us assume that we have the factor $xz'_1az'_2xxz_1\diamond z_2$ with $|z_2| < 3$. Since in this case z_1 ends in $abaaa$, it follows that z'_1a ends in $abaaaa$ and so, the only possibilities for x are ab and ba. If $x = ab$, it follows that $z'_2 = b$ and we have the factor $abz'_1abababz_1\diamond b$. We get that z_1 ends in $abaaa$, and moreover, has a prefix compatible with bbb. It follows that $z'_1, z_1 \subset bbb\psi(z)abababaaa$, for some word $z \in \{a, b, c\}^*$, and we have a factor compatible with $abababbbbzabaaaababab bbbzabaaa\diamond b$. We see that in this case, $\psi^\omega(a)$ has $cbzca\ cbzca\ b$ as a factor, a contradiciton with the fact that $\psi^\omega(a)$ is square-free.

If $x = ba$, it follows that $z'_2 = \varepsilon$ and we have the factor $abz'_1ababaz_1\diamond$. Since z'_1 has a suffix compatible with $abaaaa$, we get that z_1 has a prefix compatible with $bbbb$. Once more, it follows that $z'_1, z_1 \subset bbbb\psi(z)abababaaa$, for some word $z \in \{a, b, c\}^*$. We have that $\psi^\omega(a)$ has $cbzca\ cbzca\ b$ as a factor, a contradiciton with the fact that $\psi^\omega(a)$ is square-free. $\square$

Using Theorem 5, since $\alpha\beta\alpha\alpha\beta \mid \alpha\beta\alpha\beta\beta\alpha$, Theorem 2 can be easily obtained in a non-iterative way.

5 Non-trivially Avoidable Binary Patterns

First let us recall that the pattern $\alpha\beta\alpha\beta\alpha$ is not 2-avoidable in partial words.

Theorem 6 ([3, 6])**.** *All infinite binary partial words with more than one hole meet the pattern $\alpha\beta\alpha\beta\alpha$.*

In [3], it is also shown that $\alpha\beta\alpha\beta\alpha$ is 3-avoidable in partial words. Furthermore, since $a\diamond$ and $\diamond a$ represent occurrences of the pattern $\alpha\alpha$, we conclude that $\alpha\alpha$ is unavoidable in the context of partial words. What restriction do we need to impose on the concept of avoidability in order to construct infinitely many binary partial words with infinitely many holes that avoid the pattern $\alpha\alpha$ (the pattern $\alpha\beta\alpha\beta\alpha$)?

Definition 1. *Let $p = \alpha_0 \cdots \alpha_m$, where $\alpha_i \in E$ for $i = 0, \ldots, m$. Define an occurrence $u_0 \cdots u_m$ of p in a partial word w as* non-trivial *if $u_i \neq \diamond$ for all $i \in \{0, \ldots, m\}$. We call w* non-trivially p-free *if it contains no non-trivial occurrence of p.*

In [3], it was shown that there exist infinitely many partial words with infinitely many holes over a 3-letter alphabet that non-trivially avoid $\alpha\alpha$, and so the non-trivial avoidability index of $\alpha\alpha$ in partial words is 3. Since in full words all binary patterns with avoidability index 3 meet $\alpha\alpha$, using Remark 1 we conclude that all binary patterns in Theorem 1(2) also have non-trivial avoidability index 3 in the context of partial words. Moreover, in [9], the case of patterns of the form α^m, $m \geq 3$ was considered, the avoidability index in partial words being 2. In Theorems 2, 3, 4 and 5, we have proved that the patterns $\alpha\beta\alpha\beta\beta\alpha$, $\alpha\alpha\beta\alpha\beta\beta$, $\alpha\alpha\beta\beta\alpha$ and $\alpha\beta\alpha\alpha\beta$ are 2-avoidable. Thus, according to [8], if we can find the non-trivial avoidability index in partial words of $\alpha\beta\alpha\beta\alpha$, then we will have completed the classification of the patterns in Theorem 1(3) in terms of non-trivial avoidability in partial words.

In this section, we prove in particular the following result.

Theorem 7. *With respect to non-trivial avoidability in partial words, the avoidability index of a binary pattern is the same as in Theorem 1.*

Let us recall the iterative Thue-Morse morphism ϕ such that $\phi(a) = ab$ and $\phi(b) = ba$. It is well known that $\phi^\omega(a)$ avoids $\alpha\beta\alpha\beta\alpha$ [7].

Proposition 2. *Over a binary alphabet there exist infinitely many infinite partial words, containing exactly one hole, that non-trivially avoid the pattern $\alpha\beta\alpha\beta\alpha$.*

Proof. Let $p = \alpha\beta\alpha\beta\alpha$, and $t = \phi^\omega(a)$ be the fixed point of the Thue-Morse morphism. We show that there exist infinitely many positions in t in which one can replace the letter at that position with a hole and obtain a new word t' that is still non-trivially p-free. Also, since all factors of the infinite Thue-Morse word t avoid p, it follows that any occurrence of p in t' must contain the hole.

Let $x_1, x_2, x_3 \subset x$ and $y_1, y_2 \subset y$, for some $x, x_1, x_2, x_3, y, y_1, y_2$ such that $|x|, |y| \geq 1$. We start by proving that there does not exist a non-trivial occurrence of p, $x_1y_1x_2y_2x_3$, in t' such that $|x| \geq 8$ or $|y| \geq 8$. We proceed by contradiction. We analyze the following three cases based on the possible positions of the hole.

Case 1. The hole is in x_1.

Note that this case is symmetrical to when the hole is in x_3 (we are using the fact that if w is a factor of the Thue-Morse word t, then so is $\mathrm{rev}(w)$). Since t is overlap-free, the only possibility is to have in t a factor of the form $x'cx''yx'\bar{c}x''yx'\bar{c}x''$, with $c \in \{a, b\}$, and $x_1 = x'\diamond x''$, $x_2 = x_3 = x'\bar{c}x'' = x$ and $y_1, y_2 = y$, for some words $x', x'' \in \{a, b\}^*$ with $|x'x''| \geq 7$ or $|y| > 7$. Moreover, x' is non-empty since, otherwise, t contains the factor $x''y\bar{c}x''y\bar{c}x''$ which is impossible since t is p-free. Looking at the symbols that precede and follow $\diamond$ in x_1 and $\bar{c}$ in x_2, we get that if $|x'x''| \geq 7$ either $\bar{c}c\bar{c}c\bar{c}$ is a factor of $x_2 = x'\bar{c}x''$ when c is preceded by c in $x'cx''$, or $c\bar{c}c\bar{c}c$ is a factor of $x'cx''$ when c is preceded by $\bar{c}$ in $x'cx''$. Finally, if $|y| > 7$ either $c\bar{c}c\bar{c}c$ is a factor of $x'cx''$ when $\diamond$ is preceded by c in x_1, or $\bar{c}c\bar{c}c\bar{c}$ is a factor of $x_2 = x'\bar{c}x''$ when $\diamond$ is preceded by $\bar{c}$ in x_1. All cases lead to contradiction with the fact that t is overlap-free.

Let us illustrate by an example how this works. Let us consider the case when c is preceded by a c, $|x'| = 1$ and $|y| > 7$. We look at the factors x_1y_1 and x_2y_2 that differ at only one position. We have that y_1 starts with $\bar{c}$, so that ccc is not a prefix of our factor. It follows that $x''y_2$ starts with $\bar{c}c$ so we do not get the cube $\overline{ccc}$ in t. But, in x_1y_1 we have the factor $cc\bar{c}c$, which must be followed by $\bar{c}$. Again looking at the prefix of $x''y_2$, we get $\bar{c}c\bar{c}c$. It follows that $ccx''y$ has as prefix $cc\bar{c}c\bar{c}c$ which contains an overlap, a contradiction with the fact that t is overlap-free.

Case 2. The hole is in y_1.

Note that this case is symmetrical to when the hole is in y_2. Since t is overlap-free, the only possibility is to have in t a factor of the form $xy'cy''xy'\bar{c}y''x$, with $c \in \{a, b\}$, and $x_1 = x_2 = x_3 = x$, $y_1 = y'\diamond y''$, $y_2 = y'\bar{c}y''$, for some words $y', y'' \in \{a, b\}^*$ with $|y'| + |y''| \geq 1$. Since at least one of y' and y'' is non-empty, and either $|x| > 7$ or $|y'y''| \geq 7$, the proof follows from the previous case.

Case 3. The hole is in x_2.

Since t is overlap-free, the only possibility is to have in t a factor of the form $x'cx''yx'\bar{c}x''yx'cx''$, with $c \in \{a, b\}$, and $x_1 = x_3 = x'cx''$, $x_2 = x'\diamond x''$ and $y_1 = y_2 = y$, for some words $x', x'' \in \{a, b\}^*$ with $|x'| + |x''| \geq 1$. Again, since at least one of x' and x'' is non-empty, and either $|y| > 7$ or $|x'x''| \geq 7$, the proof follows from the first case.

We have thus shown that if $|x| \geq 8$ or $|y| \geq 8$, there are no non-trivial occurrences $x_1y_1x_2y_2x_3$, where $x_1, x_2, x_3 \subset x$ and $y_1, y_2 \subset y$, of p in t', therefore any such occurrence in t' must satisfy $|x_1y_1x_2y_2x_3| < 36$. We now put a hole at position 47 of $t_7 = \phi^7(a)$ to create t_7':

$$
\begin{aligned}
t_7' = & abbabaabbaababbabaababbaabbabaabbaababbaabbabaa\underline{\diamond}abbabaabbaababba \\
& baababbaabbabaababbabaabbaababbaabbabaabbaababbabaababbaabbabaab
\end{aligned}
$$

It is not hard to verify (using a computer program) that there are no occurrences of p in t_7'. Furthermore, we have put the hole more than 36 positions from either end of t_7. Note that t_7 occurs as a factor of t infinitely often. We can choose any arbitrary occurrence of the factor t_7 in t, and replace it with t_7' in order to obtain an infinite word with one hole that non-trivially avoids p. □

Theorem 8. *Over a binary alphabet there exist infinitely many partial words, containing infinitely many holes, that non-trivially avoid the pattern* $\alpha\beta\alpha\beta\alpha$.

Proof. Let $p = \alpha\beta\alpha\beta\alpha$, and let t be the fixed point of the Thue-Morse morphism. Note, there are infinitely many non-overlapping occurrences of t_7 starting at an even position in t [9]. For each such occurrence, we put a hole at position 47 of the factor t_7. We claim that the resulting word is non-trivially p-free.

We proceed by induction. Let t' denote a word obtained by replacing any h positions in t with holes in the manner described above. We have already proven the $h = 1$ case in Proposition 2, so assume that t'', a word obtained by replacing any $h-1$ positions in t with holes also in the manner described above, non-trivially avoids p. We show that t' non-trivially avoids p. Assume the contrary. Thus, there is some factor $x_1y_1x_2y_2x_3$ of t' such that $x_1, x_2, x_3 \subset u$ and $y_1, y_2 \subset v$ for some words $u, v \in \{a, b\}^*$. Observe that by the way we insert holes, there are at least 127 letters in between any two holes. Also, $x_1y_1x_2y_2x_3$ must contain all h holes (otherwise there is an occurrence of p in some t''). Hence, $|x_1y_1x_2y_2x_3| \geq 129$ when $h \geq 2$.

There are several cases to consider: the cases where there exist at least one hole in x_1, or in x_3, or in y_1, or in y_2, and the case where all holes are in x_2.

Case 1. At least one hole is in x_1.

Note that the case when x_3 contains at least one hole is similar. We have $x_1 = w_{11}c_1w_{12}$, $x_2 = w_{21}c_2w_{22}$, $x_3 = w_{31}c_3w_{32}$ where $w_{ij} \subset u_j$ for any $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$, $c_1 = \diamond$ and $u_1 \diamond u_2 \subset u$. In order for the inductive hypothesis to hold, we have either $c_2 = a$ or $c_3 = a$. Additionally, note that we cannot have that both $c_2 \neq b$ and $c_3 \neq b$ since, otherwise, the hole at c_1 is unnecessary to create a non-trivial occurrence of p. Moreover, since we are dealing with a non-trivial occurrence of p and a hole is in x_1, we have $|x_1| \geq 2$. Thus, either $|u_1| > 0$ or $|u_2| > 0$.

If $|u_2| > 1$, it follows that t'' has $w_{12}\ y_1 w_{21} c_2 w_{22}\ y_2 w_{31} c_3 w_{32}$ as a factor. Here, taking u_2 as image for α and vu_1c as image for β, with $c_2, c_3 \subset c$, we get a contradiction with the inductive hypothesis.

If $|u_1| > 0$ and $|u_2| = 1$, it follows that c_1 is preceded by a. It is easy to check that $|u_2 v| > 2$ (otherwise we get that t contains overlaps as factors). Thus, c_1 is also followed by ab. It follows that c_2 is also preceded by a factor x' compatible with a and followed by a factor x'' that is compatible with ab, such that $x' c_2 x''$ is compatible with $aaab$ and contains at most one hole. Since the hole is always followed by ab, it follows that $c_2 = a$. But, since $|u_1| > 0$ and $|u_2| = 1$, c_3 must be an a, and the only possibility for a hole here is the following position (the first position of w_{32}). Since this hole is an odd number of positions apart from $c_1 = \diamond$, and the holes are always put at odd positions, this is a contradiction.

If $|u_2| = 1$ and $|u_1| = 0$, since the hole is always put at an odd position, it follows that $w_{32} = a$ is at an even position and is followed by b (one of the images of ϕ). Moreover, y_1 starts with b and y_2 has a prefix compatible with b. It follows that t'' has a factor compatible with $aby'a\ a\diamond y'a\ ab$, where $y_1, y_2 \subset by'$. Taking ab for image of α and $y'a$ as image for β, p occurs in t'', which is a contradiction with our inductive hypothesis.

If $|u_2| = 0$, it follows that the hole is preceded by an a. Since c_3 is an odd number of positions apart from c_1, it must be the case that $c_3 = a$ (if $c_3 = \diamond$ then we have a hole at an even position, and if $c_3 = b$ we have an occurrence of p in t''). Moreover, since c_3 is at an odd position, it follows that c_3 is preceded by b. This is a contradiction since c_1 is preceded by a.

Case 2. At least one hole is in y_1.

Here, x_1, x_3 are full (the case when at least one hole is in y_2 is similar). Let $y_1 = w_{11} c_1 w_{12}$ and $y_2 = w_{21} c_2 w_{22}$ where $w_{ij} \subset v_j$ for any $i, j \in \{1, 2\}$, $c_1 = \diamond$ and $v_1 \diamond v_2 \subset v$. In order for the inductive hypothesis to hold, we need to have $c_2 = a$.

Since $|x_1 y_1 x_2 y_2 x_3| \geq 129$, either c_1 is preceded by aa and followed by a or c_1 is preceded by a and followed by ab. Moreover, since $x_1 y_1 x_2 \uparrow x_2 y_2 x_3$, $c_2 = a$ is also preceded by factors compatible with aa or a and is followed by factors compatible with a or ab. Since a hole is always followed by ab and replaces a b, we see that none of these other positions contains one, giving us that t contains a cube aaa, which is a contradiction.

Case 3. All the holes are in x_2.

Let $x_1 = w_{11}cw_{12}$, $x_2 = w_{21}\diamond w_{22}$, $x_3 = w_{31}cw_{32}$ where $w_{ij} \subset u_j$ for every $i \in \{1,2,3\}$ and $j \in \{1,2\}$, and $c = a$.

In the case $|u_1| \geq 1$, it follows that the c in x_1 is preceded and followed by a's. We get that t contains aaa as a factor, which is a contradiction. Otherwise, it follows that $|y_2u_2| \geq 2$. Hence, since the hole is preceded by aa, the c in x_3 is preceded by aa. Again we get a contradiction with the fact that aaa is not a factor of t.

We have thus shown that we can insert infinitely many holes in t and non-trivially avoid p. Moreover, we can choose an arbitrary number of positions where we have holes, replace them back with b's, and obtain infinitely many partial words with infinitely many holes that non-trivially avoid p. □

6 Avoidable Binary Patterns

In the previous sections we have proved that all binary patterns that are k-avoidable in full words are also k-avoidable in partial words, with $k \in \{2,3\}$, when considering non-trivial avoidability. A question that arises is how the classification of the binary patterns looks like when considering general avoidability in partial words (not restricted to non-trivial avoidability).

With respect to general avoidability in partial words, the binary patterns $\alpha\alpha\alpha$, $\alpha\beta\alpha\beta\beta\alpha$, $\alpha\alpha\beta\alpha\beta\beta$, $\alpha\beta\alpha\alpha\beta$ and $\alpha\alpha\beta\beta\alpha$ have avoidability index 2, the pattern $\alpha\beta\alpha\beta\alpha$ has avoidability index 3, and the patterns $\varepsilon, \alpha, \alpha\beta, \alpha\beta\alpha$ and $\alpha\alpha$, and their complements, are all unavoidable (or have avoidability index ∞).

The following lemmas are straightforward.

Lemma 4. *The binary patterns $\alpha\alpha\beta$, $\alpha\alpha\beta\alpha$ and $\alpha\alpha\beta\alpha\alpha$ are unavoidable in partial words.*

Proof. Since in a partial word with an infinity of holes, there must exist a symbol $a \in A$, that has at least two occurrences preceded or followed by $\diamond$, it follows easily that there always exist factors of the form $\diamond ax$, $\diamond aya$ and $\diamond az\diamond a$, for some $x, y, z \in A_\diamond^*$. We see that these factors represent occurrences of the aforementioned patterns. □

Lemma 5. *The patterns $\alpha\alpha\beta\beta$, $\alpha\beta\alpha\beta$ and $\alpha\alpha\beta\alpha\beta$ have avoidability index 3 in partial words.*

Proof. For the last two patterns we see that both meet the pattern $\alpha\alpha$ which is non-trivially avoidable. Hence, the conclusion follows. For the pattern, $\alpha\alpha\beta\beta$, let us consider a word w that non-trivially avoids $\alpha\alpha$. In [3], such a word is provided over a three-letter alphabet, having the property that each two holes are at least 10 positions apart. If our pattern has an occurrence, say $x_1x_2y_1y_2$ with $x_1 \uparrow x_2$ and $y_1 \uparrow y_2$, then $|x_1| = |y_1| = 1$ and the factor contains two holes. This is a contradiction with the construction of w. $\square$

We are left considering the pattern $\alpha\beta\beta\alpha$ and all patterns from Theorem 1(3) (other than $\alpha\beta\alpha\beta\alpha$).

By looking at all possible factors of a binary pattern of length up to 7, we conclude the following.

Remark 3. *All patterns of length seven or more are 2-avoidable.*

Thus, the only patterns that we have to consider in order to characterize all binary patterns, are $\alpha\beta\beta\alpha$ and $\alpha\beta\alpha\beta\alpha\alpha$, their reverses, and complements. Moreover, since $\alpha\beta\alpha\beta\alpha$ is 3-avoidable, and both $\alpha\beta\alpha\beta\alpha\alpha\alpha$ and $\alpha\beta\alpha\beta\alpha\alpha\beta$ are 2-avoidable, we conclude that $\alpha\beta\alpha\beta\alpha\alpha$ is either 2- or 3-avoidable.

Theorem 9. *Over a binary alphabet, there exist infinitely many partial words, containing infinitely many holes, that avoid the pattern $\alpha\beta\alpha\beta\alpha\alpha$.*

Proof. Let us consider the Thue-Morse word $t = \phi^\omega(a)$. In Theorem 8 it was proved that t', the word obtained after arbirarily many holes were inserted in t, non-trivially avoids $\alpha\beta\alpha\beta\alpha$. It follows that if our pattern occurs in t, since $\alpha\beta\alpha\beta\alpha$ meets our pattern, either the image corresponding to α or the one corresponding to β has length one.

Let us assume towards a contradiction that there exists an occurrence of $\alpha\beta\alpha\beta\alpha\alpha$ in t'. It follows that there exists in t' a factor $x_1y_1x_2y_2x_3x_4$ with $x_i \uparrow x_j$ and $y_1 \uparrow y_2$, for all $i, j \in \{1, 2, 3, 4\}$. Let us first consider the case where $x_1y_1x_2y_2x_3x_4$ contains only one hole. If the hole is in x_1 it follows that our factor is of the form $\diamond yayaa$, where $x_1 = \diamond$, $y_1 = y_2 = y$ and $x_2 = x_3 = x_4 = a$. This is due to the fact that the occurrence of the pattern $\alpha\beta\alpha\beta\alpha$ must be trivial, and moreover, the $\diamond$ replaces an a in t, so it should correspond to b's in our occurrence of the pattern. But, since the $\diamond$ is always followed by an a in t, we get that a is a prefix of y, and so $yayaa$ is an overlap that occurs in t. This is a contradiction with the fact that t is overlap-free.

If the hole is in x_2 we get $ay\diamond yaa$. It is easy to check that in this case it must be that $|y| > 2$. Hence, y has baa as a suffix. It follows that for some word y' with $y = y'baa$ we have the factor $ay'baa\diamond y'baaaa$, a contradiction with the fact that t is overlap-free and so aaa is not a factor of it. If the hole is in x_3 we get $ayay\diamond a$, with baa a suffix of y. A contradiction follows as in the previous case. Note that we cannot have $x_4 = \diamond$, since otherwise we have that t is not overlap-free.

When $y_1 = \diamond$, and the occurrence of the factor is $x\diamond xaxx$, we get the non-trivial overlap $x\diamond xax$. Finally, when $y_2 = \diamond$, and the occurrence of the factor is $xax\diamond xx$, we get that x has aa as a suffix. It follows that $a\diamond xx$ is a non-trivial occurrence of the pattern $\alpha\beta\alpha\beta\alpha$ in t', again a contradiction.

Now let us assume that there exist more than one hole in $x_1y_1x_2y_2x_3x_4$, and at least one of the images of our variables is $\diamond$. In this case we use the fact that each two holes in t' are at least 129 positions apart. It follows that if the $\diamond$ corresponds to α, the y_i's have length at least 120, and if the $\diamond$ corresponds to β, the x_i's have length at least 120. Using the fact that the $\diamond$ is always preceded by baa and so the symbol corresponding to it in the other occurrences of the same variable is either $\diamond$ or b, and the reasoning from the one hole case, the conclusion follows. □

The proof for the reverse of the pattern is almost identical. Hence, the only pattern left to classify is $\alpha\beta\beta\alpha$. Since $\alpha\beta\beta\alpha$ is 3-avoidable in full words, we conclude that the pattern is 2-unavoidable in partial words. The next natural question though is if $\alpha\beta\beta\alpha$ is 3-avoidable or not in partial words.

In order to find an upper bound for the avoidability index of $\alpha\beta\beta\alpha$, let us again consider an infinite partial word $w \in \{a, b, c\}^*$ that avoids all squares except $f\diamond$ and $\diamond f$, for some $f \in \{a, b, c\}$. According to [3], such a word exists. In this word let us replace all occurrences of $\diamond$ by $\diamond d$, for a new symbol d, and get a new partial word w' that has infinitely many holes. We prove that w' avoids $\alpha\beta\beta\alpha$.

Proposition 3. *The pattern $\alpha\beta\beta\alpha$ is 4-avoidable.*

Proof. It is quite straigthforward to see that since w avoids non-trivial squares so does w'. Hence, if an occurrence of $\alpha\beta\beta\alpha$ exists in w', then the image of β has length 1. Let us assume that such an occurrence exists and denote by $x_1a_1a_2x_2$ the factor corresponding to our pattern, where $x_1 \uparrow x_2$ and a_1 or a_2 is a hole. It follows that either $a_2 = d$ or d is a prefix of x_2. We consider the later case, the former one being quite similar. It follows that x_1 has either d or $\diamond$ as a prefix. If x_1 has $\diamond$ as a prefix, since $a_1 = \diamond$ and the holes cannot be too close together according to the construction of w, then x_2 has dd as a prefix, a contradiction with the way w' has been constructed. Hence, it must be the case that x_1 starts with d. But, in this case, it follows that before the d's were inserted, the factors corresponding to x_1 and x_2 were still compatible. Since, $\diamond$ is compatible with any letter of the alphabet it follows that w has a non-trivial square, which is a contradiction. □

Using the same strategy as in the case of factors of length six or more, starting with the factors $a\diamond a$ and $a\diamond b$ we get the following.

Remark 4. *If a ternary word meeting the pattern $\alpha\beta\beta\alpha$ exists, it must be that the word is not square-free after replacing the holes with any symbol.*

A question that arises is if there exist iterated morphisms for which, replacing some of the letters of the fixed point with holes, gives us a ternary word that avoids the pattern $\alpha\beta\beta\alpha$. In particular, note that since the word cannot be square-free, it must be that factors of the form xx exist. Hence, we get arbitrarily long such factors after iterated applications of the morphism on this factor. Moreover, it follows that applications of the morphism on the word formed from the symbol following the pattern and the one preceding the factor, will always be unbordered.

We can now reformulate Theorem 1 for partial words.

Theorem 10. *For partial words, binary patterns fall into four categories:*

1. *The binary patterns ε, α, $\alpha\alpha$, $\alpha\alpha\beta$, $\alpha\alpha\beta\alpha$, $\alpha\alpha\beta\alpha\alpha$, $\alpha\beta$, $\alpha\beta\alpha$, and their complements, are unavoidable (or have avoidability index ∞).*

2. *The binary patterns $\alpha\alpha\beta\alpha\beta$, $\alpha\alpha\beta\beta$, $\alpha\beta\alpha\beta$ and $\alpha\beta\alpha\beta\alpha$, their reverses, and complements, have avoidability index 3.*

3. *The binary pattern $\alpha\beta\beta\alpha$ has avoidability index 3 or 4.*

4. *All other binary patterns, and in particular all binary patterns of length six or more, have avoidability index 2.*

References

[1] D. R. Bean, A. Ehrenfeucht, and G. McNulty. Avoidable patterns in strings of symbols. *Pacific Journal of Mathematics*, 85:261–294, 1979.

[2] F. Blanchet-Sadri. *Algorithmic Combinatorics on Partial Words.* Chapman & Hall/CRC Press, Boca Raton, FL, 2008.

[3] F. Blanchet-Sadri, R. Mercaş, and G. Scott. A generalization of Thue freeness for partial words. *Theoretical Computer Science*, 410(8-10):793–800, 2009. (`www.uncg.edu/cmp/research/freeness`).

[4] F. Blanchet-Sadri, R. Mercaş, S. Simmons, and E. Weissenstein. Avoidable binary patterns in partial words. In A.-H. Dediu,

H. Fernau, and C. Martín-Vide, editors, *LATA 2010, 4th International Conference on Language and Automata Theory and Applications, Trier, Germany*, volume 6031 of *Lecture Notes in Computer Science*, pages 106–117, Berlin, Heidelberg, 2010. Springer-Verlag. (`www.uncg.edu/cmp/research/unavoidablesets4`).

[5] J. Cassaigne. Unavoidable binary patterns. *Acta Informatica*, 30:385–395, 1993.

[6] V. Halava, T. Harju, T. Kärki, and P. Séébold. Overlap-freeness in infinite partial words. *Theoretical Computer Science*, 410(8-10):943–948, 2009.

[7] M. Lothaire. *Combinatorics on Words*. Cambridge University Press, Cambridge, 1997.

[8] M. Lothaire. *Algebraic Combinatorics on Words*. Cambridge University Press, Cambridge, 2002.

[9] F. Manea and R. Mercaş. Freeness of partial words. *Theoretical Computer Science*, 389(1–2):265–277, 2007.

[10] P. Roth. Every binary pattern of length six is avoidable on the two-letter alphabet. *Acta Informatica*, 29:95–106, 1992.

[11] U. Schmidt. Motifs inévitables dans les mots. Technical Report LITP 86–63 Thèse Paris VI, 1986.

[12] U. Schmidt. Avoidable patterns on two letters. *Theoretical Computer Science*, 63:1–17, 1989.

[13] A. Thue. Über unendliche Zeichenreihen. *Norske Vid. Selsk. Skr. I, Mat. Nat. Kl. Christiana*, 7:1–22, 1906. (Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo, Norway (1977), pp. 139–158).

[14] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. *Norske Vid. Selsk. Skr. I, Mat. Nat. Kl. Christiana*, 1:1–67, 1912. (Reprinted in *Selected Mathematical Papers of Axel Thue*, T. Nagell, editor, Universitetsforlaget, Oslo, Norway (1977), pp. 413–478).

[15] A. I. Zimin. Blocking sets of terms. *Mathematics of the USSR-Sbornik*, 47(2):353–364, 1984.