Avoiding Large Squares in Partial Words

F. Blanchet-Sadri1    Ilkyoo Choi2    Robert Mercaş3

April 5, 2011

Abstract

Well-known results on the avoidance of large squares in (full) words include: (1) Fraenkel and Simpson showed that we can construct an infinite binary word containing at most three distinct squares; (2) Enterring, Jackson and Schatz that there exists an infinite binary word avoiding all squares of the form $xx$ such that $|x| \geq 3$, and that the bound 3 is optimal; (3) Dekking that there exists an infinite cube-free binary word that avoids all squares $xx$ with $|x| \geq 4$, and that the bound of 4 is best possible. In this paper, we investigate these avoidance results in the context of partial words, or sequences that may have some undefined symbols called holes. Here, a square has the form $uv$ with $u$ and $v$ compatible, and consequently, such square is compatible with a number of full words that are squares over the given alphabet. We show that (1) holds for partial words with at most two holes. We prove that (2) extends to partial words having infinitely many holes. Regarding (3), we show that there exist binary partial words with infinitely many holes that avoid cubes and have only eleven full word squares compatible with factors of it. Moreover, this number is optimal, and all such squares $xx$ satisfy $|x| \leq 4$.

Keywords: Combinatorics on words; Partial words; Squares; Cubes.

*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. A World Wide Web server interface has been established at www.uncg.edu/cmp/research/squares2 for automated use of the program. We thank the referees of a preliminary version of this paper for their very valuable comments and suggestions.

1Department of Computer Science, University of North Carolina, P.O. Box 26170, Greensboro, NC 27402-6170, USA, blanchet@uncg.edu
2Department of Mathematical Sciences, Carnegie Mellon University, 5032 Forbes Ave., Pittsburgh, PA 15289, USA
3GRLMC, Universitat Rovira i Virgili, Departament de Filologies Românicas, Av. Catalunya 35, Tarragona, 43002, Spain, robertmercas@gmail.com
1 Introduction

The existence of square-free words over a three-letter alphabet has been proven several times in various ways, first by Thue in [14], in which he also showed the avoidance of the triples $aca$ and $bcb$. According to Currie, [3], “One reason for this sequence of rediscoveries is that non-repetitive sequences have been used to construct counterexamples in many areas of mathematics: ergodic theory, formal language theory, universal algebra and group theory, for example ....”

For full words, the question of how many distinct squares a binary word contains was first raised in [6]. There, Entringer, Jackson and Schatz give a construction for an infinite binary word that avoids all squares $xx$ with $|x| \geq 3$, and show that the bound 3 is the best possible. More precisely, using the notation from [7], they obtain $g(5) = \infty$, where $g(n)$ denotes the maximum length of a binary word containing no more than $n$ squares. In [7], Fraenkel and Simpson improve the result by showing that there actually exists an infinite binary word containing only the squares $a^2$, $b^2$ and $(ab)^2$. Their construction, involving several steps and non-uniform morphisms, was simplified by Rampersad, Shallit and Wang in [13], where two uniform morphisms are used. We also mention the nice and short proof given in [9]. Moreover in [5], Dekking shows that there exists an infinite cube-free binary word that avoids all squares $xx$ such that $|x| \geq 4$. He also proves that the bound 4 is best possible.

In this paper, we investigate the avoidance of large squares in partial words, sequences over a finite alphabet that may contain some “holes” denoted by ♦’s (here ♦ is compatible with, or matches, any symbol of the alphabet). In this context, a square has the form $uv$ where $u$ is compatible with $v$, and consequently, such square is compatible with a number of full words that are squares over the alphabet. Square and cube-freeness of partial words have been studied for the first time by Manea and Mercaş [11]. See also the more recent results in [2] and in [8]. One of our goals is to compute, whenever it exists, $g_h(n)$, the maximum length of a binary partial word with $h$ holes and having no more than $n$ distinct full squares compatible with factors of it. We also investigate the avoidance of large squares in cube-free binary partial words.

The contents of our paper are as follows: In Section 2, we compute all the $g_h(n)$’s, extending the results from [7]. In particular, we show that all infinite binary partial words whose factors are compatible with no more than three distinct full squares have at most two holes. Moreover, we show that there exist arbitrarily long binary partial words having three or four holes.
and at most three distinct full squares compatible with factors of them. We also construct a binary word with infinitely many holes such that the only full squares compatible with factors of it are $aa$, $bb$, $abab$ and $bbbb$. Hence, we extend the result of Entringer, Jackson and Schatz to partial words with infinitely many holes. In Section 3, we extend to partial words the result of Dekking. We show that for at most two holes there exist infinite cube-free binary partial words avoiding all squares $xx$ such that $|x| \geq 4$, and that the bound four given by Dekking, regarding the length of the squares allowed, is exceeded for more than two holes. In Section 4, we investigate the possibility of obtaining less squares by increasing the alphabet size. We build an infinite partial word with one hole over the ternary alphabet $\{a, b, c\}$ such that the only full square compatible with factors of it is $aa$. It turns out that the position of the hole must be the first.

We end this section with an overview of basic concepts of combinatorics on partial words. The reader is referred to [1] for more background material.

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 $A \cup \{\diamond\}$, the alphabet $A$ being augmented with the “hole” symbol $\diamond$ (a full word is a partial word without holes). We will denote by $u(i)$ the symbol at position $i$ of the partial word $u$, $i \geq 0$. The set containing all full words over $A$ is denoted by $A^*$, while the set of all partial words over $A$ is denoted by $A^* \diamond$.

The length of a partial word $u$ 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 $\varepsilon$. 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 $u^\omega$.

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. If $u, v, w$ are non-empty partial words, then $uvw$ is called a cube if there exists a partial word $x$ such that $u \subset x, v \subset x$ and $w \subset x$. Whenever we refer to a square $uv$, it will imply that $u \uparrow v$. A partial word is square-free (respectively, cube-free) if it contains no squares (respectively, cubes).

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 $u \neq v$), and $u$ is an internal factor of $v$ if $x, y$ are non-empty. We say that $u$ is a prefix of $v$ if $x = \varepsilon$ and
a suffix of \( v \) if \( y = \varepsilon \). For partial words \( u, v, w \) with \( w = uv \), we will also denote \( v \) by \( u^{-1}w \), and \( u \) by \( wv^{-1} \).

2 Avoiding Large Squares in Binary Partial Words

Let \( g_h(n) \) be the length of a longest binary partial word with \( h \) holes and having at most \( n \) distinct full squares compatible with factors of it if such words exist, and \( g_h(n) \) is undefined in the case where there exist arbitrarily long binary partial words but there does not exist any infinite partial word satisfying the above conditions. In this section, we show how the sequence \( \{g_h(n)\} \) behaves. The case of \( h = 0 \) appears in [7]. There, it is shown that \( g_0(1) = 7 \) (\( aaabaaa, abaaaba \) and their complements, \( abbbabb \) and its reverse and their complements), while \( g_0(2) = 18 \) (the only word having this property is \( abaabbaaabbbaabbab \) and its complement which is also its reverse). It is also shown that there exists an infinite binary word which has only three different squares, that is, \( g_0(n) = \infty \) for all integers \( n \geq 3 \). Simpler proofs of this latter result appear in [13, 9].

Using a computer program we found that for one hole, \( g_1(0) = 1 \) (\( \diamond \)), \( g_1(1) = 5 \) (\( \diamond aaba \)), \( g_1(2) = 16 \) (\( \diamond abbaaabbaabbab \)); two holes, \( g_2(0) = 0 \), \( g_2(1) = 3 \) (\( \diamond o \diamond \)), \( g_2(2) = 14 \) (\( \diamond abbaaabbaaabbb \)); three holes, \( g_3(0) = 0 \), \( g_3(1) = 0 \), \( g_3(2) = 9 \), and the only words containing two distinct squares are \( \diamond o o \diamond \), \( \diamond o o \diamond \diamond \), \( \diamond o o \diamond \diamond \), their reverses and complements. Moreover, for more than three holes no word exists containing less than three squares. So for \( h \geq 3 \), \( g_h(n) = 0 \) for all non-negative integers \( n \leq 0 \). We refer the reader to Table 1 that lists the \( g_h(n) \)’s that will be derived in this paper.

In order to prove our results we make use of some morphisms given by Rampersad, Shallit and Wang in [13]. Let us first recall these, together with some results obtained there. Let \( \alpha : \{a, b, c, d, e\}^* \to \{a, b, c, d, e\}^* \) (respectively, \( \beta : \{a, b, c, d, e\}^* \to \{a, b\}^* \)) be the five-letter 24-uniform morphism (respectively, the 6-uniform morphism) defined as follows:

\[
\begin{align*}
a & \mapsto abcdcbabcdeabcbabcbcdce \\
b & \mapsto abcbedeacdedeacdedeacdede \\
c & \mapsto abcbedeacdedeacdedeacdede \\
d & \mapsto abcbedeacdedeacdedeacdede \\
e & \mapsto abcbedeacdedeacdedeacdede \\
\end{align*}
\]

and
Let us assume that there are more than three full distinct squares showing that $aa$, $bb$ and $abab$, showing as mentioned before that $g_0(n) = \infty$ for all integers $n \geq 3$.

We now construct infinite binary partial words with one or two holes such that $aa$, $bb$ and $abab$ are the only full squares compatible with their factors, showing that $g_1(n) = g_2(n) = \infty$ for all integers $n \geq 3$.

**Theorem 1.** The only full squares compatible with factors of $w_1 = \circ \beta(\alpha^\omega(a))$, an infinite binary partial word with one hole, are $aa$, $bb$ and $abab$.

**Proof.** Let us assume that there are more than three full distinct squares compatible with factors of $w_1$. Since $w_0$ contains only $aa$, $bb$ and $abab$ as squares, there exists in $w_1$ a square of the form $\circ u a_0 u'$, with $\circ u \uparrow a_0 u'$ and $u a_0 u' a_1$ prefix of $w_0$, where $u = u'$ are non-empty words, and $a_0$, $a_1$ are letters from $\{a, b\}$. Moreover, we can express $|u| = 6q + r$, for some integers $q, r$ with $0 \leq r < 6$, where $q$ represents the number of blocks of length 6 created by the images of the morphism $\beta$. Computer programs showed that actually $q \geq 2$. If $r = 0$, that is $u = \beta(x)$, for some word $x \in \{a, b, c, d, e\}^+$, since $u = u'$ is a prefix of $w_0$ and $|u| = 6q$, it follows that for some letter $f \in \{a, b, c, d, e\}$, $\beta(f)$ is a prefix of $a_0 \beta(a)$. This is impossible since we do not have a letter whose morphism through $\beta$ is $aabba$ or $babba$. If $r = 5$, then for some word $x \in \{a, b, c, d, e\}^+$, $u a_0 = \beta(x)$. But since the images of all letters after applying the morphism $\beta$ have the first five characters different and $u = u'$, it must be the case that $a_1 = a_0$. This implies that $w_0$ contains the square $u a_0 u' a_1$, which is a contradiction with the fact that $w_0$ has factors compatible with only the full squares $aa$, $bb$ and $abab$.

Hence, $0 < r < 5$. Since $\beta(ab)$ is a prefix of $u$, it follows that there exist letters $b_0, b_1, b_2$, such that $\beta(ab)$ is an internal factor of $\beta(b_0 b_1 b_2)$ (where $\beta(b_0)$ is the block starting at position $|u| - r$ in $u$ and $\beta(b_1 b_2)$ are the first two blocks starting at position $5 - r$ in $u'$). Note that $b_0 \neq b_1$ and $b_1 \neq b_2$, since two consecutive letters in $\alpha^\omega(a)$ are always different. Checking all options, the only possibility for this to happen is when $b_0 = e$, $b_1 = d$ and $b_2 = c$. Yet, according to the definition of $\alpha$, $\alpha^\omega(a)$ does not contain $ede$ as a factor, so we get a contradiction.

Since all cases lead to contradiction, we conclude that the only full squares compatible with factors of $w_1$ are $aa$, $bb$, $abab$. \qed
Theorem 2. The only full squares compatible with factors of the word \(w_2 = \omega \beta(\alpha^\omega(a))\), an infinite binary partial word with two holes, are \(aa\), \(bb\) and \(abab\).

Proof. Let us assume that there are more than three full distinct squares compatible with factors of \(w_2\). According to Theorem 1, it follows that there exists in \(w_2\) a square of the form \(\diamond \alpha \omega(a)\omega'\), with \(u_0a_1u'a_2a_3\) prefix of \(w_0\), where \(u = u'\) are non-empty words, and \(a_0\), \(a_1\), \(a_2\) and \(a_3\) are letters from \(\{a, b\}\). Let us refer to the notation used in the proof of Theorem 1. Using a computer we found that \(q > 5\). If \(r = 0\), we have that for some letter \(f \in \{a, b, c, d, e\}\), \(\beta(f)\) is a prefix of \(a_0a_1\beta(a)\). Again we reach a contradiction since \(abbb\) is not a suffix of any of the images of \(\beta\). If \(r = 5\), then for some letter \(f \in \{a, b, c, d, e\}\), \(\beta(f)\) is a prefix of \(a_1\beta(a)\), which is a contradiction.

If \(r = 4\), then either \(a_0a_1 = ab\) or \(a_0a_1 = ba\). In order to avoid getting a square in \(w_0\), it must be the case that \(a_2a_3 = ba\) or \(a_2a_3 = ab\) respectively. Hence, if the block starting in \(u\) at position \(|u| - r\) is \(\beta(d)\), respectively \(\beta(e)\), then the block starting in \(u'\) at position \(|u'| - r\) is \(\beta(e)\), respectively \(\beta(d)\). By the definition of \(\alpha\), we know that \(\beta(e)\) is always preceded by \(\beta(d)\). This implies that either \(u\) or \(u'\) will end in \(\beta(dd)\), which is a contradiction since \(w_0\) only contains the squares \(aa\), \(bb\), \(abab\).

Since the last case of \(0 < r < 4\) is similar to the one in the proof of Theorem 1, we conclude that the only squares compatible with factors of \(w_2\) are \(aa\), \(bb\) and \(abab\).\qed

We next construct arbitrarily long binary partial words with three or four holes such that \(aa\), \(bb\) and \(abab\) are the only full squares compatible with factors of them. We also show that no infinite binary partial word exists with three or four holes and with at most three distinct full squares compatible with factors of it, showing that \(g_3(3), g_4(3)\) are undefined. We first prove two lemmas.

Lemma 1. Let \(m\) be an arbitrarily large positive integer. The only full squares compatible with factors of \(w_3 = \beta(\alpha^m(a))\), a binary partial word with one hole, are \(aa\), \(bb\) and \(abab\).

Proof. Considering that \(\beta(a) = \text{rev}(s(e))\) and \(\beta(b) = \text{rev}(s(d))\), where \(\text{rev}\) denotes the reversal of words and \(s(a) = b\) and \(s(b) = a\), and the fact that only the first two letters of an \(\alpha\) image are used, gives quite straightforwardly that Lemma 1 is just a symmetric to Theorem 1.\qed
Lemma 2. Let $m$ be an arbitrarily large positive integer. The only full squares compatible with factors of $w_4 = \beta(\alpha^m(a))\circ\circ$, a binary partial word with two holes, are $aa$, $bb$ and $abab$.

Proof. The proof is similar to that of Lemma 1, the result here being the symmetric of Theorem 2.

Theorem 3. There exist arbitrarily long binary partial words with three or four holes, where $aa$, $bb$ and $abab$ are the only full squares compatible with factors of them.

Proof. Let $m$ be an arbitrarily large positive integer. We will show that $aa$, $bb$ and $abab$ are the only full squares compatible with factors of $\circ\circ w_4 = \circ\circ \beta(\alpha^m(a))\circ\circ$ (for three holes, take $\diamond w_4$). For the sake of contradiction, let us assume that $\circ\circ \beta(\alpha^m(a))\circ\circ$ has a factor $v_0v_1$ that is compatible with a full square other than $aa$, $bb$ and $abab$. By Theorems 1, 2 and Lemmas 1 and 2, $v_0v_1$ has at least one hole at both ends. Since $\alpha$ is a 24-uniform morphism and $\beta$ is a 6-uniform morphism, $\beta(\alpha^m(a))$ is of length $6 \times 24^m$. Therefore, $v_0v_1$ must have the same number of holes at both ends since squares are of even length. Let $u$ and $u'$ denote words of equal length in the following two cases (note that the length of $u$ is divisible by 6, that is, $u$ and $u'$ consist of full blocks only):

If $v_0v_1 = \circ\circ uu'\circ\circ$, then $\circ\circ u \uparrow u'\circ\circ$. Since $\beta(a)$ is a prefix of $u$ but none of the images of $\beta$ have $abbb$ as a suffix, this is a contradiction.

If $v_0v_1 = \circ uu'\circ$, then $\circ u \uparrow u'\circ$. Since $\beta(a)$ is a prefix of $u$ but none of the images of $\beta$ have $abbaa$ as a suffix, this is a contradiction. We conclude that $aa$, $bb$, $abab$ are the only full squares compatible with factors of $\circ\circ w_4$.

Using a computer program, we found that all binary partial words of the form $u_0u_1$ with $|u_0| = |u_1| = 8$ have more than three distinct squares compatible with factors of them. This implies that all arbitrarily long partial words with at most three distinct full squares compatible with factors of them and at least three holes should have the holes placed within the first or the last eight positions. Again, using a computer program, we found that all binary partial words with at least three holes within the first eight symbols and no more than three distinct full squares compatible with factors of them have maximum length thirteen. This gives us the next two straightforward results.

Proposition 1. All infinite binary partial words with more than two holes have more than three distinct full squares compatible with factors of them.
Proposition 2. All binary partial words with more than four holes have more than three distinct full squares compatible with factors of them.

Thus for \( h > 4 \), \( g_h(n) = 0 \) for all non-negative integers \( n \leq 3 \).

We actually found a construction for a binary word with infinitely many holes such that the only full squares compatible with factors of it are \( aa, bb, abab \) and \( bbbb \). This shows that when \( n \geq 4 \), \( g_h(n) = \infty \) for all integers \( h \geq 3 \).

We use the morphism \( \alpha \) and a slightly modified version of the morphism \( \hat{\beta} : \{a, b, c, d, e\}^* \to \{a, b, \Diamond\}^* \), defined as follows:

\[
\begin{align*}
a &\mapsto \Diamond bbbbaa \\
b &\mapsto babbbaa \\
c &\mapsto bbbaaa \\
d &\mapsto bbaaba \\
e &\mapsto bbaaab
\end{align*}
\]

Theorem 4. The only full squares compatible with factors of \( w_5 = \hat{\beta}(\alpha^w(a)) \), a binary partial word with infinitely many holes, are \( aa, bb, abab \) and \( bbbb \).

Proof. Let \( uu' \) be a factor of \( w_5 \) such that \( u \uparrow u' \). Using a computer program, we checked that no other full squares of length less than 30 are compatible with factors of \( w_5 \). In order to prove that no full squares besides \( aa, bb, abab \) and \( bbbb \) are compatible with factors of \( w_5 \), it suffices to show that for each hole in \( u \), the corresponding character in \( u' \) is \( a \), and vice versa. In other words, replacing the holes with \( a \)'s would transform \( w_5 \) into \( w_0 \), which we know has only the three full squares \( aa, bb, abab \). Assume that a hole in \( u \) is replaced by \( b \) in \( u' \) (the case of a hole in \( u' \) being identical).

If there exists a hole at position \( i \) in \( uu' \), for some integer \( i < |u| - 6 \), then \( u \) has a factor \( \Diamond bbbbaab \), since \( \hat{\beta}(a) \) cannot be followed by another \( \hat{\beta}(a) \). Hence, there is a corresponding factor \( bbbbaab \) in \( u' \). Since this can be obtained only through \( \hat{\beta}(ec) \), a contradiction follows by the definition of \( \alpha \).

If there exists a hole at position \( i \) in \( uu' \), for some integer \( i \geq |u| - 6 \geq 9 \), then the hole is preceded by \( v = aaabbbbaa \) or \( v' = ababbaaab \), which are suffixes of \( \hat{\beta}(cb) \) or \( \hat{\beta}(de) \) respectively, the latter images always preceding \( \hat{\beta}(a) \). Since a factor compatible with \( v \) or \( v' \) is followed by a \( \Diamond \) or an \( a \), we get the desired result.

Note that Theorem 4 extends to partial words with infinitely many holes the surprising result of Entringer, Jackson and Schatz stating that there exists an infinite binary word that avoids all squares \( xx \) such that \( |x| \geq 3 \), and that the bound 3 is best possible [6].
3 Avoiding Large Squares in Cube-Free Binary Partial Words

In [5], Dekking shows that there exists an infinite cube-free binary word that avoids all squares $xx$ such that $|x| \geq 4$. Dekking also proves that the bound 4 is best possible. Using a computer program, we found that all words $w$ with the property above and with length $|w| \geq 114$ must indeed contain all squares $|x| \leq 3$ except $(aa)^2$, $(bb)^2$, $(aaa)^2$ and $(bbb)^2$. Note that a word containing $(aa)^2$, $(bb)^2$, $(aaa)^2$ or $(bbb)^2$ is not cube-free.

Remark 1. All cube-free binary words of length greater than 113 avoiding all squares $xx$ such that $|x| \geq 4$ must contain all of the following ten squares: $a^2$, $b^2$, $(ab)^2$, $(ba)^2$, $(aab)^2$, $(aba)^2$, $(abb)^2$, $(baa)^2$, $(bab)^2$ and $(bba)^2$.

In [13], it is shown that the infinite full binary word $v_0 = \delta(\gamma^\omega(a))$ is cube-free and avoids all squares $xx$ such that $|x| \geq 4$, where the 10-uniform morphism $\gamma : \{a, b, c, d\}^* \to \{a, b, c, d\}^*$ is defined by:

- $a \mapsto adbacabacd$
- $b \mapsto adbacdabac$
- $c \mapsto acabadbacd$
- $d \mapsto acadabacab$

and the 6-uniform morphism $\delta : \{a, b, c, d\}^* \to \{a, b\}^*$ by:

- $a \mapsto aboaabb$
- $b \mapsto ababba$
- $c \mapsto abbaab$
- $d \mapsto ababa$

Note that $v_0$ is prolongable on $a$ (therefore $\delta(a) = aboaabb$ is a prefix of $v_0$). Using the morphisms $\gamma$ and $\delta$, we will prove that there exist cube-free binary partial words that avoid all squares $xx$ such that $|x| \geq 4$.

Theorem 5. The only full squares compatible with factors of the word $v_1 = \diamond(\delta(a))^{-1}v_0$, an infinite cube-free binary partial word with one hole, are the ten squares in Remark 1.

Proof. Let us assume there are more than ten full distinct squares compatible with factors of $v_1$. Since $v_0$ only contains the ten squares in Remark 1, there must be a square of the form $\diamond u a_0 u'$ with $\diamond u \uparrow a_0 u'$, and $\diamond u a_0 u' a_1$ a prefix of $v_1$, where $u = u'$ are non-empty words and $a_0$, $a_1$ are letters from $\{a, b\}$. Moreover, by the construction of $v_1$, we can express $|u| = 6q + r$, where
The only full squares compatible with factors of $v$ are prefixes of $v$ that are cube-free and avoid all squares with the square $abaaba$ must be a prefix of $u$. If $r \neq 5$, this implies there exist letters $b_0, b_1, b_2$ such that $\delta(b_0b_1b_2) = y_0\delta(db)y_1$ where $y_0, y_1$ are words with $|y_0| \neq 0$ (mod 6). This is impossible since none of the images of $\delta$ has $b$ or $aa$ as a prefix, nor is it $abaaba$. Therefore, $r = 5$. But, since all of the images of $\delta$ have unique 5-letter prefixes, this implies $a_0 = a_1$. This is a contradiction since now $v_0$ has the square $ua_0u'a_1$.

Since $v_0$ is cube-free and avoids all squares $xx$ such that $|x| \geq 4$, to show that $v_1$ is cube-free it suffices to search for small cubes of length at most 18, which are prefixes of $v_1$.

**Theorem 6.** The only full squares compatible with factors of $v_2 = \delta baabv_1$, an infinite cube-free binary partial word with two holes, are the ten squares in Remark 1.

**Proof.** Let us assume there are more than ten full distinct squares compatible with factors of $v_2$. Since, by Theorem 5, the only full squares compatible with factors of $v_1$ are the ten squares in Remark 1, there must be a square of the form $zu_0u_0u'$ with $zu' \uparrow z_0a_0u'$ where $a_0$ is a letter and $u, u', z, z_0$ are words such that, $u = u', z \uparrow z_0$, and $z$ is a suffix of $obaab$. Let $obaab\delta u_0u_0u'z_1a_1$ be the prefix of $v_2$, where $|z| = |z_1|$ and $a_1$ is a letter. Note that $0 < |z| \leq 5$, since $|z| = 0$ falls into the case of Theorem 5. Using the exact same notation and reasoning from the proof of Theorem 5, it must be that $r = 5 - |z|$. This implies that $z_0a_0$ is a suffix of an image of $\delta$. If $|z| = 1$, then $r = 4$, meaning the last image of $\delta$ that starts in $u$ has four symbols in $u$. Since all of the images of $\delta$ have unique 4-letter suffixes, this implies $z_0a_0 = z_1a_1$. So we have the square $u_0u_0u'z_1a_1$ in $v_1$. If $|z| = 2$, then $r = 3$, which implies that $aba_0$ must be a suffix of an image of $\delta$ (recall that $z_0 \uparrow z$). Since $v_2 \uparrow v_0$, we cannot have that $a_0 = b$, since this implies that, there exists a square $zbu_0bu'$ in $v_0$. Thus, $a_0 = a$, and therefore the last image of $\delta$ starting in $u$ must be $ababa$. But, this is also a prefix of $u'$, so we obtain the square $\delta(dd)$ in $v_1$. If $3 \leq |z| \leq 5$, then according to the previous reasoning, $a_0 \neq b$, indicating one of the blocks ends in $aaba$. There exists no image of $\delta$ having $aaba$ as a suffix.

Since $v_1$ is cube-free and avoids all squares $xx$ such that $|x| \geq 4$, to show that $v_2$ is cube-free it is enough to check the small cubes of length at most 12, which are prefixes of $v_2$. 

\[ \square \]
Theorem 7. Let \( m \) be an arbitrarily large positive integer. The only full squares compatible with factors of \( v_3 = \delta(\gamma^{3m+1}(a)(acd)^{-1})\circ baab\circ \delta(a) \), a cube-free binary partial word with two holes, are the ten squares in Remark 1.

Proof. Let us assume there are more than ten full distinct squares compatible with factors of \( v_3 \). Since \( v_0 \) only contains the ten squares in Remark 1 and since \( v_3 \) is compatible with a prefix of \( v_0 \), there must be a square of the form \( ua_1z_1u'oz \) with \( ua_1z_1 \uparrow u'oz \) and \( a_0z_0ua_1z_1u'oz \) compatible with a factor of \( v_0 \), where \( a_0, a_1 \) are letters, and \( u, u', z, z_0, z_1 \) are words such that \( u = u' \), \( z \uparrow z_1 \), \( |z| = |z_0| = |z_1| \), and \( z \) is a prefix of \( baab \). Note that \( 0 \leq |z| \leq 5 \). Using a similar notation and reasoning as in the proof of Theorem 5, it must be that \( r = 5 - |z| \). In this case we have that \( a_1z_1 \) is a prefix of an image of \( \delta \). If \( 0 \leq |z| \leq 4 \), since \( a_1 \) is the first letter of an image of \( \delta \), we must have that \( a_1 = a \). But, \( a_1 \neq a \) since this implies \( v_0 \) has the square \( ua_2z_1u'az \). If \( |z| = 5 \), this implies \( a_1z_1 \uparrow \circ baab \circ \delta \). The only image of \( \delta \) that has \( baab \) as an internal factor is \( \delta(a) = abaabb \). Hence, \( a_1z_1 = \delta(a) \). This is impossible since \( v_0 \) now has the square \( u\delta(a)u'\delta(a) \).

Since \( v_3 \) avoids large squares, it must avoid large cubes. One can easily check \( v_3 \) avoids small cubes by looking at the morphisms that constructed \( v_3 \).

Let us now look at the case of more than two holes. First, note that in a cube-free binary partial word, any letters adjacent to a hole must be different, and between any two holes there must be at least two different letters. Now, using a computer program, we found that there are at least eleven distinct full squares compatible with factors of any cube-free binary partial word of the form \( u_0u_1 \) with \( |u_0| = |u_1| = 9 \). This implies that any cube-free binary partial word that has at most ten distinct full squares compatible with factors of it must have all the holes in the first or last nine positions. It is then easy to check that it is impossible to avoid a cube in a word of length eleven with three holes in the first nine positions.

Proposition 3. \( \bullet \) At least eleven distinct full squares are compatible with factors of any infinite cube-free binary partial word containing more than two holes.

\( \bullet \) At least eleven distinct full squares are compatible with factors of any cube-free binary partial word with more than four holes.

Theorem 8. Let \( m \) be an arbitrarily large positive integer. The only full squares compatible with factors of \( v_4 = \circ baab\circ \delta(a)^{-1}\gamma^{3m+1}(a)(acd)^{-1})\circ baab\circ \delta(a)^{-1} \),
a cube-free binary partial word with four holes, are the ten squares in Remark 1.

Proof. Let us assume there are more than ten full distinct squares compatible with factors of \( v_4 \). By Theorems 6 and 7, there must exist a square of the form \( zuu'z' \) with \( zu \uparrow u'z' \), for \( z \) a non-empty suffix of \( \delta baab \), \( z' \) a non-empty prefix of \( \delta baab \) and, \( uu' = \delta(a^{-1}i3m+1(a)(acd))^{-1} \). If it is not the case that \( z = z' = \delta baab \), since \( \delta(db) \) is a prefix of \( u \) and a suffix of \( u' \), this implies that there exist letters \( b_0, b_1, b_2 \) such that \( \delta(b_0b_1b_2) = y_0\delta(db)y_1 \) for \( |y_0| \neq 0 \) (mod 6). This is impossible since no image of \( \delta \) has \( b \) or \( aa \) as a prefix, nor is it \( abaaba \). If \( z = z' = \delta baab \), since \( \delta(a) \) is the only image of \( \delta \) that has \( baab \) as an internal factor, \( \delta(a) \) is a prefix of \( u' \) and a suffix of \( u \). We get a contradiction since the square \( \delta(aa) \) is now a factor of \( uu' \), which is a factor of \( v_0 \).

Next, we extend the result of Dekking by presenting a construction for a cube-free binary partial word with infinitely many holes that only has eleven distinct full squares compatible with factors of it. The squares are of the form \( xx \), where \( |x| \leq 4 \) (the ten squares stated in Remark 1 plus \((abba)^2\)). Recall that a word is \((r^+, n)\)-free if it contains no repetition \( u^s \) with \( |u| \geq n \) and \( s > r \), where \( n \) is an integer, \( r \) is a rational number, and \( s \) is a real number. A word is \( r^+\)-free if it is \((r^+, 1)\)-free. In [12], Ochem gives a non-iterative morphism \( \tau : \{a, b, c\}^* \rightarrow \{a, b\}^* \) that, for any \( \frac{2}{4}^+ \)-free word \( w \in \{a, b, c\}^* \), \( \tau(w) \) is \( \frac{5}{2}^+ \)-free, \( \left(\frac{2}{3}^+, 3\right)\)-free, and \( \left(\frac{23}{142}^+, 4\right)\)-free. In order to make use of this morphism we will use a construction given in [4], in the context of repetition threshold. Here Dejean gives an iterative morphism on \( \{a, b, c\}^* \) that preserves \( \frac{2}{4}^+ \)-freeness:

\[
\begin{align*}
a & \mapsto abacbacabcaabcaabcaabaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaab
In the following, let us define by \( v_5 \) the fixed point of Dejean’s morphism. The infinite ternary word \( v_5 \) is square-free (and thus cube-free). Moreover, the infinite binary word \( \tau(v_5) \) is cube-free and does not have squares of the form \( xx \), where \( |x| \geq 4 \). We will prove that the only extra square \( \hat{\tau}(v_5) \) has is \((abba)^2\).

**Theorem 9.** There exists a cube-free binary partial word with infinitely many holes such that the only full squares compatible with factors of it are the ten squares in Remark 1 as well as the square \((abba)^2\).

**Proof.** Assume for the sake of contradiction that there are more than eleven full distinct squares compatible with factors of \( \hat{\tau}(v_5) \). Since \( \hat{\tau}(v_5) \) has been obtained from \( \tau(v_5) \) by replacing some \( a \)'s by \( \diamond \)'s, any square in \( \hat{\tau}(v_5) \) not compatible with any of the ten above mentioned full squares must be of the form \( uu'vbv' \) for some partial words \( u, u', v, v' \) such that \( uu' \uparrow vbv' \), \(|u| = |v| \) and \(|u'| = |v'|\).

If \(|u'| \geq 12\), then \( u' \) must have the factor of length 12 following the \( \diamond \) in \( \hat{\tau}(b) \), \( baababbabaa \), as a prefix. Thus \( v' \) must have a prefix compatible with \( baababbabaa \). One can check that this pattern is compatible with only factors following either an \( a \) or a \( \diamond \). If \(|u| \geq 12\), then \( u \) must have \( ababaabababbaa \) as a suffix (thus \( v \) must have a suffix compatible with \( ababaabababbaa \)). Again, it can be checked that this pattern is compatible with only factors preceding either an \( a \) or a \( \diamond \).

Finally, we can check that if \(|u| < 12\) and \(|u'| < 12\), the only square in \( \hat{\tau}(v_5) \) not compatible with any of the ten squares in Remark 1 is when \( u = v = a \) and \( u' = v' = ba \). \( \square \)

### 4 Conclusion

In this paper, we have given, in particular, a complete answer to the question of how the sequence \( \{g_h(n)\} \) behaves. To summarize these results, we provide Table 1 for the \( g_h(n) \)'s where the undefined cases are denoted by \( - \).

What if we increase the size of the alphabet? Is it possible to get fewer squares? An upper bound of two squares is given in the following theorem.

**Theorem 10.** [2, 8] There exist infinitely many partial words with infinitely many holes over the three-letter alphabet \( \{a, b, c\} \) such that the only full squares compatible with factors of it are \( bb \) and \( cc \).

Is it possible to reduce the bound to one square? In fact, any infinite partial word with a hole after the first position is such that at least two
Table 1: The $g_h(n)$’s

<table>
<thead>
<tr>
<th>$n \setminus h$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>16</td>
<td>14</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>3</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$-$</td>
<td>$-$</td>
<td>0</td>
<td>0</td>
<td>...</td>
</tr>
<tr>
<td>4</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>...</td>
</tr>
<tr>
<td>5</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>$\infty$</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

distinct full squares are compatible with factors of it (hence any infinite partial word with more than one hole satisfies the property). To see this, assuming that there exists an infinite partial word $u$ with a hole after the first position such that only one full square is compatible with factors of $u$, it must be the case that $u$ has a factor of the form $\circ \circ$ or $a \circ b$ for some distinct letters $a, b$ (yielding in both cases the full squares $aa$ and $bb$), or $a \circ a$ for some letter $a$ (yielding the squares $aa$ and $(ab)^2$, where $b$ is a letter not necessarily distinct from $a$).

We now build an infinite partial word with one hole over the ternary alphabet $\{a, b, c\}$ such that the only full square compatible with factors of it is $aa$. By the above, the position of the hole must be the first. Let us first recall the morphism $\phi : \{a, b, c\}^* \to \{a, b, c\}^*$ with $\phi(a) = abc$, $\phi(b) = ac$ and $\phi(c) = b$, whose fixed point $\sigma = \phi^\omega(a)$ is square-free [10]. In addition, $\sigma$ does not contain the factors $aba$ and $cbc$.

**Theorem 11.** The only full square compatible with factors of $\sigma$, an infinite partial word with one hole over the three-letter alphabet $\{a, b, c\}$, is $aa$.

**Proof.** Let us assume that there exists a square in $\sigma$ other than $\sigma a$. It is easy to check that no such square occurs in the prefix of length 10 of $\sigma$. It follows that this square should be of the form $\circ aueau'$, for some letter $e \in \{a, b, c\}$ and some non-empty words $u = u'$. Since $\sigma$ is square-free, it follows that either $e = b$ or $e = c$. We know that $aueau'$ is a prefix of $\phi^m(a)$, for some integer $m$. It follows that after another $m$ iterations we will get the prefix $\phi^m(au)\phi^m(e)\phi^m(a)$. Since $e = b$ or $e = c$, it follows that if $m$ is even, then $\phi^m(e)$ ends in $e$. Hence, we get that $\sigma$ contains the factor $e \phi^m(a)$. But since $aueau'$ is a prefix of $\phi^m(a)$, we get a contradiction with the fact that $\sigma$ is square-free.
Thus, $m$ is odd. Since $\phi$ is prolongable on $a$ we have that $\phi^{m+1}(a) = \phi^m(a)\phi(v)$, for some word $v$. Moreover, a later iteration of our factor will appear somewhere in the word. Hence, from

$$\phi(\phi^m(e)\phi^m(a)) = \phi^{m+1}(e)\phi^{m+1}(a) = \phi^{m+1}(e)\phi^m(a)\phi(v)$$

we get that $\phi^m(a)$ is preceded by the last symbol of $\phi^{m+1}(e)$, which is $e$ since $m$ is odd. We get a contradiction since $eaueau'$ cannot be a factor of $\sigma$. 

References


