Advances in Applied Mathematics 48 (2012) 194–214

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

Abelian repetitions in partial words ☆

F. Blanchet-Sadri [a,*], Sean Simmons [b], Dimin Xu [c]

[a] Department of Computer Science, University of North Carolina, P.O. Box 26170, Greensboro, NC 27402-6170, USA
[b] Department of Mathematics, The University of Texas at Austin, 1 University Station C1200, Austin, TX 78712-0233, USA
[c] Department of Mathematics, Bard College, 30 Campus Road, Annandale-on-Hudson, NY 12504, USA

ARTICLE INFO

Article history:
Received 12 November 2010
Revised 10 May 2011
Accepted 22 June 2011
Available online 26 August 2011

MSC:
68R15
05A05

Keywords:
Combinatorics on words
Pattern avoidance
Partial words
Abelian powers
Morphisms
Parikh vectors

ABSTRACT

We study abelian repetitions in partial words, or sequences that may contain some unknown positions or holes. First, we look at the avoidance of abelian pth powers in infinite partial words, where $p > 2$, extending recent results regarding the case where $p = 2$. We investigate, for a given p, the smallest alphabet size needed to construct an infinite partial word with finitely or infinitely many holes that avoids abelian pth powers. We construct in particular an infinite binary partial word with infinitely many holes that avoids 6th powers. Then we show, in a number of cases, that the number of abelian p-free partial words of length n with h holes over a given alphabet grows exponentially as n increases. Finally, we prove that we cannot avoid abelian pth powers under arbitrary insertion of holes in an infinite word.

1. Introduction

The study of word structures, or combinatorics on words, plays an important role in theoretical computer science. At the beginning of the twentieth century, Thue discovered in [25] that the consecutive repetitions of non-empty factors, also called squares, can be avoided by an infinite word over a

☆ 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 will be presented at DLT'11 (Blanchet-Sadri and Simmons, 2011 [8]). A research assignment from the University of North Carolina at Greensboro for the first author is gratefully acknowledged. Some of this assignment was spent at the LIAFA: Laboratoire d'Informatique Algorithmique: Fondements et Applications of Université Paris 7-Denis Diderot, Paris, France. We thank the referee of a preliminary version of this paper for his/her very valuable comments and suggestions.
* Corresponding author.
E-mail address: blanchet@uncg.edu (F. Blanchet-Sadri).

0196-8858/$ – see front matter
doi:10.1016/j.aam.2011.06.006

ternary alphabet. Such word is said to be *2-free* or *square-free*. More generally, a word u is *p-free* if it does not contain x^p as a factor, where x is any non-empty factor of u.

The study of abelian repetitions was started by Erdös in [15]. He raised the question whether there exist infinite abelian square-free words over a given alphabet (words in which no two consecutive factors are permutations of each other). For example, the word *abcbac* is an abelian square since *abc* and *bac* are permutations of each other. It can easily be checked that no infinite abelian square-free word exists over a three-letter alphabet because every word of length eight over three letters contains an abelian square. Evdokimov in [16] showed that abelian squares are avoidable over 25 letters. Pleasants in [22] showed that there exists an infinite abelian square-free word over five letters using a uniform morphism of size 5×15. Keränen in [18] reduced the alphabet size to four by introducing a uniform abelian square-free morphism of size 4×85. Other constructions over four letters appear in [10,19]. Computer-assisted music analysis is an application area of abelian square-free words [20]. Cryptography is another application area [3,23]. For instance, Rivest presented a new way of dithering for iterated hash functions based on such square-free words. Abelian square-free words have also been used in the study of free partially commutative monoids [11,14].

An *abelian pth power* consists of p consecutive factors which are permutations of each other, and a word is *abelian p-free* if it does not contain any abelian pth power. Results for $p > 2$ include: Justin in [17] showed that there exists an infinite abelian 5-free word over two letters with a uniform morphism of size 2×5. Dekking in [13] proved that over a ternary alphabet, there exists an abelian cube-free infinite word, and using a non-uniform morphism that abelian 4th powers are avoidable over a binary alphabet (abelian cubes cannot be avoided over a binary alphabet). Aberkane and Currie in [1] improved the latter result by introducing a uniform morphism having an abelian 4-free fixed point.

The problem of avoiding abelian squares in partial words was initiated by Blanchet-Sadri et al. in [6]. The idea of partial words was first introduced by Berstel and Boasson in [4] and the one of freeness in partial words by Manea and Mercaş in [21]. Partial words are sequences over a finite alphabet that may have some unknown positions or holes, represented by $\diamond$'s, which are *compatible* with, or match, each letter of the alphabet. For instance, $ab\diamond b\diamond c$ is a partial word with two holes over the ternary alphabet $\{a, b, c\}$. A partial word u over an alphabet A is an *abelian square* if it is possible to substitute letters from A for each hole in such a way that u becomes an abelian square that is a full word (or a partial word without holes). The partial word u is *abelian square-free* if it does not have any full or partial abelian square, except those of the form $\diamond a$ or $a\diamond$, where $a \in A$. For example, the full word *abacaba* is abelian square-free, while the partial word $abcdad\diamond ada$ is not (it contains the abelian square $cdad\diamond a$ compatible with *cdadca*). In particular, in [6], lower and upper bounds were given for the number of letters needed to construct infinite abelian square-free partial words with finitely or infinitely many holes. It was proved that there exists an infinite abelian square-free partial word with one hole over a four-letter alphabet, and none exists with more than one hole. It was also showed that the minimal alphabet size for the existence of an abelian square-free partial word with more than one hole is five. Several of the constructions are based on iterating morphisms.

In this paper, we build on previous work by studying abelian repetitions in partial words. A partial word u is *p-free* if for every non-empty factor $u_0 \cdots u_{p-1}$ of u, there does not exist a full word v compatible with u_i, for all $i \in \{0, \dots, p-1\}$. A partial word u is *abelian p-free* if for every non-empty factor $u_0 \cdots u_{p-1}$ of u, there does not exist a full word v compatible with some permutation of u_i, for all $i \in \{0, \dots, p-1\}$. We are interested in the following three problems: (1) Study avoidance of abelian pth powers, for $p > 2$. In the context of partial words, which abelian powers can be avoided over a given alphabet? (2) Investigate the number of abelian p-free partial words of a fixed length over an alphabet of a given size. The number of abelian 2-free full words has been studied in [9] and [19] for an alphabet of size four (work has also been done for higher powers [1,2,12]). (3) Can we construct an infinite full word that remains abelian p-free even after arbitrarily many positions are replaced with holes? In [21], a 3-free infinite full word over an optimal alphabet size of four letters was given that remains 3-free after such replacements, while in [7] a 2-free infinite full word over an optimal size of eight letters was given that remains 2-free after such replacements. Here, among other things, we prove that abelian 6th powers can be avoided over a binary alphabet, while such is not the case for abelian 3rd powers. We also show that there is a partial word with infinitely many holes over a three-letter (resp., three-letter, four-letter) alphabet that avoids abelian 5th powers (resp., 4th powers, 3rd powers).

The contents of our paper is as follows. In Section 2, we give some preliminaries on partial words. In Section 3, we prove that there exists a partial word with infinitely many holes that avoids abelian 4th powers over a ternary alphabet, and a partial word over a four-letter alphabet with infinitely many holes which is abelian cube-free. For powers greater than four, we show that there exists a binary partial word with infinitely many holes that avoids abelian 6th or higher powers. There, we also deal with the case of finitely many holes. We study the existence of infinite partial words with finitely many holes that avoid abelian 3rd or 4th powers. In particular, we construct an infinite abelian 4-free partial word with one hole over a binary alphabet, and an infinite abelian cube-free partial word with one hole over a ternary alphabet. In Section 4, we show, in cases related to the above results, that the number of abelian p-free partial words of length n with h holes grows exponentially as n increases. In Section 5, we are concerned with the problem of whether an infinite word can remain abelian p-free after an arbitrary insertion of holes, that is, the problem of replacing an arbitrary collection of positions' letters with holes. There, we show that every infinite abelian p-free word will contain an abelian pth power after arbitrarily inserting infinitely many holes. In Section 6, we conclude with some remarks and directions for future work.

2. Preliminaries

For more background information on partial words, we refer the reader to [5].

Let A be an *alphabet*, or a non-empty finite set of symbols, where each element of A is called a *letter*. A sequence of letters from A is called a *(full) word* over A. A *partial word* over A is a sequence of symbols from $A_\diamond = A \cup \{\diamond\}$, where $\diamond$ is the symbol representing an unknown position. Any occurrence of $\diamond$ in a partial word is called a *hole* (a full word is a partial word without holes). The *length* of a partial word u, denoted by $|u|$, represents the number of symbols in u, while $u(i)$ represents the symbol at position i of u, where $0 \leqslant i < |u|$. The *empty word* is the sequence of length zero and is denoted by ε. The set of distinct letters in u, or the *alphabet* of u, is denoted by $\alpha(u)$. For instance, the partial word $u = ab\diamond bba\diamond$, where a, b are distinct letters of the alphabet A, satisfies $\alpha(u) = \{a, b\}$. The set of all words over A is denoted by A^*, while the set of all partial words over A by $A_\diamond^*$. A (right) (resp., two-sided) infinite partial word is a function $u : \mathbb{N} \to A_\diamond$ (resp., $u : \mathbb{Z} \to A_\diamond$).

Let u and v be partial words of equal length. Then u is *contained* in v, denoted by $u \subset v$, if $u(i) = v(i)$, for all i such that $u(i) \in A$. The partial words u and v are *compatible*, denoted by $u \uparrow v$, if there exists a partial word w such that $u \subset w$ and $v \subset w$. For example, $a\diamond bb\diamond \uparrow ab\diamond b\diamond$, since $a\diamond bb\diamond \subset abbb\diamond$ and $ab\diamond b\diamond \subset abbb\diamond$.

A partial word u is a *factor* or *subword* 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 $y = \varepsilon$. The notation $v[i..j)$ represents the factor $v(i) \cdots v(j-1)$, while $v[i..j]$ the factor $v(i) \cdots v(j)$. The *reversal* of the partial word u is the word written backwards, denoted by $\mathrm{rev}(u)$. The powers of u are defined recursively as follows: $u^0 = \varepsilon$, and for any integer $p > 0$, $u^p = uu^{p-1}$.

Given $a \in A$, $|u|_a$ denotes the number of occurrences of a in a partial word u over alphabet A. If we write $A = \{a_0, \ldots, a_{k-1}\}$, where the cardinality of A, $\|A\|$, is k, then the *Parikh vector* of u, $P(u)$, is $P(u) = \langle |u|_{a_0}, \ldots, |u|_{a_{k-1}} \rangle$. For a positive integer p, a non-empty word $u \in A^*$ is an *abelian pth power* if we can write $u = u_0 \cdots u_{p-1}$ so that $P(u_0) = \cdots = P(u_{p-1})$. This is the same as saying that each u_i is a permutation of u_0. A word u is *abelian p-free* if no factor of u is an abelian pth power. A non-empty partial word $u \in A_\diamond^*$ is an *abelian pth power* if it is possible to substitute letters from A for each hole in such a way that u becomes a full word that is an abelian pth power. In other words, u is an abelian pth power if there exists a full word v, compatible with u, that is an abelian pth power. For example, the partial word $abca\diamond b$ is an abelian 2nd power, or abelian square, since we can replace the $\diamond$ with letter c and form $abcacb$. Whenever we refer to an abelian pth power $u_0 \cdots u_{p-1}$, it implies that there exists a non-empty word w such that for each i, $0 \leqslant i < p$, some permutation of u_i is compatible with w. The partial word u is *abelian p-free* if it does not have any full or partial abelian pth power.

Let A and B be alphabets, and 2^{B^*} be the set of all subsets of B^*. A morphism $\phi : A^* \to B^*$ is called *abelian p-free* if $\phi(u)$ is abelian p-free whenever u is abelian p-free. A *multi-valued substitution* is a function $\theta : A^* \to 2^{B^*}$, with the property that if u_1 and u_2 are words in A^*, then $\theta(u_1u_2) = \theta(u_1)\theta(u_2) = \{v_1v_2 \mid v_i \in \theta(u_i)\}$. The multi-valued substitution θ is said to be abelian p-free if, whenever $u \in A^*$ is abelian p-free then so are all of the words in $\theta(u)$. See [12] for more information on multi-valued substitutions.

Finally, inserting a hole in a word u is defined as replacing a position's letter in u with a $\diamond$. An arbitrary insertion of holes is replacing any collection of positions' letters in u with $\diamond$'s, so that every pair of consecutive $\diamond$'s is separated by at least two letters. This restriction is to avoid some trivial occurrences of powers (a pth power $u_0 \cdots u_{p-1}$, $|u_0| = \cdots = |u_{p-1}|$, is called *trivial* if $u_i = \diamond$ for some i).

3. Avoiding abelian powers greater than two

In this section, we study avoidance of abelian powers greater than two in the context of partial words. From [13], we can construct an infinite full word over a binary alphabet that avoids abelian 4th powers. Here, we investigate the minimal alphabet size for the construction of an infinite partial word that avoids abelian pth powers, where $p > 2$. In Section 3.1, we consider partial words with infinitely many holes, while in Section 3.2, partial words with finitely many holes.

3.1. The case with infinitely many holes

It is already known that abelian squares can be avoided by a partial word with infinitely many holes over five letters, where the alphabet size is minimal [6]. So we move on to higher abelian powers, that is, we investigate the minimal alphabet size needed to construct partial words with infinitely many holes that avoid abelian powers greater than two.

We begin by proving a result about abelian 4th powers.

Theorem 1. *There exists a partial word with infinitely many holes over a three-letter alphabet that avoids abelian 4th powers.*

Proof. By [13], there exists an infinite word w over $A = \{a, b\}$ which avoids abelian 4th powers. Let c be any letter not in A, and write $B = \{a, b, c\}$. Let $k_i = 5 \times 6^i$. Then we can define an infinite partial word w' over B by

$$w'(j) = \begin{cases} \diamond & \text{if } j = k_i \text{ for some } i \\ c & \text{if } j = k_i + 1 \text{ or } j = k_i - 1 \text{ for some } i \\ w(j) & \text{otherwise} \end{cases}$$

We want to show that w' contains no abelian 4th powers. To do this, we assume that w' contains an abelian 4th power $u_0u_1u_2u_3$, where $u_j = w[i_j..i_{j+1} - 1]$ for each j.

Assume that none of the u_i's contains the letter c. This implies that none of the u_i's contains a hole since each occurrence of $\diamond$ in w' is part of the subword $c\diamond c$. By definition of w', however, this implies that $u_0u_1u_2u_3$ is in fact a subword of w. This contradicts the fact that w is abelian 4-free. Therefore, we must conclude that some u_i contains the letter c.

Since at least one u_i contains a c, each u_j must contain a c or a $\diamond$. In particular, this implies that $|u_i| > 1$ for all i. To see this, note that if $|u_0| = \cdots = |u_3| = 1$, then since each u_i contains c or $\diamond$ it follows that either $u_i = \diamond$ or $u_i = c$. Therefore $u_0u_1u_2u_3$ is a partial word of length four over the alphabet $\{c\}$. However, by construction, no such factor exists in w'.

Since each u_j contains $\diamond$ or c, there exists some k_{r_j} so that either $i_j \leqslant k_{r_j} < i_{j+1}$ or $i_j \leqslant k_{r_j} - 1 < i_{j+1}$ or $i_j \leqslant k_{r_j} + 1 < i_{j+1}$. In particular, this implies that $k_{r_1} \leqslant i_2$ and $i_3 - 1 \leqslant k_{r_3}$. Since $|u_2| > 1$, $i_2 + 1 < i_3$, we get that $k_{r_1} \leqslant i_2 < i_3 - 1 \leqslant k_{r_3}$. Then note that $|u_3| = |u_2| = |u_1| = |u_0| = i_1 - i_0 \leqslant i_1 \leqslant k_{r_1} + 1$ and that, if we let $v = w[0..i_0 - 1]u_0$, then $|v| = i_1 \leqslant k_{r_1} + 1$. Therefore, we get

$$k_{r_3} \leqslant i_4 = \left|w[0..i_4 - 1]\right| = |vu_1u_2u_3| = |v| + |u_1| + |u_2| + |u_3| \leqslant 4(k_{r_1} + 1) < k_{r_1+1} \leqslant k_{r_3}$$

This implies that $k_{r_3} < k_{r_3}$, a contradiction, so w' is abelian 4-free. $\square$

Note that the basic idea of the proof of Theorem 1 is to place the holes so far apart that any abelian 4th power could contain only a few holes. This idea will be used repeatedly in the sequel, and so it is made more precise in the following lemma.

Lemma 1. *Let w be an infinite partial word over a finite alphabet, and let $p > 1$ be an integer. Assume that*

$$u_0 \cdots u_{p-1} = w[i_0..i_1 - 1] \cdots w[i_{p-1}..i_p - 1]$$

is an abelian pth power in w. Let $k_0 < k_1 < \cdots$ be a sequence of integers so that $k_0 > 0$, and $k_i > pk_{i-1}$ for all $i \geqslant 1$. Then if $q > 0$, there exists at most one j such that $i_q \leqslant k_j \leqslant i_{q+1} - 1$. Furthermore, if $q_1 < q_2$, and if there exist j_1, j_2 such that $i_{q_1} \leqslant k_{j_1} \leqslant i_{q_1+1} - 1$ and $i_{q_2} \leqslant k_{j_2} \leqslant i_{q_2+1} - 1$, then $q_1 = 0$.

Proof. Let w and $u_0 \cdots u_{p-1}$ be as above. Moreover, assume that $0 < q_1 < q_2$ is such that there exist j_1, j_2 so that $i_{q_1} \leqslant k_{j_1} \leqslant i_{q_1+1} - 1$ and $i_{q_2} \leqslant k_{j_2} \leqslant i_{q_2+1} - 1$. Since $q_1 > 0$, we get that $i_1 \leqslant i_{q_1} \leqslant k_{j_1} < k_{j_2} \leqslant i_{q_2+1} - 1 \leqslant i_p - 1$. This implies that

$$k_{j_1} \geqslant i_1 \geqslant i_1 - i_0 = |u_0| = \cdots = |u_{p-1}|$$

Furthermore, if $v = w[0..i_0 - 1]u_0$ then $|v| = i_1 \leqslant k_{j_1}$. This yields

$$k_{j_2} < i_p = \left|w[0..i_p - 1]\right| = |vu_1 \cdots u_{p-1}| \leqslant pk_{j_1} < k_{j_1+1} \leqslant k_{j_2}$$

Consequently, $k_{j_2} < k_{j_2}$, which is a contradiction.

The other half of the lemma is proved almost identically. $\square$

Corollary 1. *Let w be an infinite partial word. Let k_0 be the smallest integer such that $w(k_0)$ is a hole, k_1 the next smallest integer such that $w(k_1)$ is a hole, and so on. Assume that $k_0 > 0$ and that $k_i > pk_{i-1}$, for some integer $p > 1$. Also assume that $u_0 \cdots u_{p-1}$ is an abelian pth power in w. If $j > 0$, then u_j contains at most one hole. Furthermore, if $i < j$ and both u_i and u_j contain holes, then $i = 0$.*

Then we move on to abelian cubes.

Theorem 2. *There exists a partial word with infinitely many holes over a four-letter alphabet that avoids abelian cubes.*

Proof. Let $A = \{a, b, c\}$ and $B = \{a, b, c, d\}$. Consider the morphism $\phi : A^* \to A^*$ defined as $\phi(a) = aabc$, $\phi(b) = bbc$ and $\phi(c) = acc$. Let w be the fixed point of ϕ. It is known that w is abelian cube-free (see [13]).

Begin by noting that $accaccbbc$ occurs infinitely often in w. We can thus find a sequence $k_0 < k_1 < \cdots$ so that $k_0 > 20$; for each i, $k_{i+1} > 3k_i$; and for each i, $w[k_i - 2..k_i + 6] = accaccbbc$. Then we can define an infinite partial word w' over B by

$$w'(j) = \begin{cases} \diamond & \text{if } j = k_i \text{ for some } i, i \equiv 0 \bmod 19 \\ d & \text{if } j \in \{k_i - 1, k_i + 2, k_i + 3\} \text{ for some } i, i \equiv 0 \bmod 19 \\ d & \text{if } j = k_i \text{ for some } i, i \not\equiv 0 \bmod 19 \\ w(j) & \text{otherwise} \end{cases}$$

Note that w' is well defined. We want to show that w' is abelian cube-free.

We begin by assuming that $u_0u_1u_2$ is an abelian cube in w'. Write $u_j = w'[i_j..i_{j+1}-1]$ for each j. Note that some u_i must contain a hole. Otherwise, if we produce u_i' by replacing each d in u_i with c, then $u_0'u_1'u_2'$ occurs in w and is an abelian cube, contradicting the fact that w is abelian cube-free.

Note that $|u_i| > 3$ for all i. To see this, assume $|u_i| \leqslant 3$. Note, since some u_i contains a hole, that $u_0u_1u_2$ must be contained in a subword of w' of length 18 containing a hole. We then simply check the possibilities with a computer to realize that this cannot happen.

Each u_i contains a d. To see this, assume that some u_i does not contain a d. This implies that u_i contains a hole. However, every $\diamond$ in w' is contained in a word of the form $d\diamond ad$, so since $|u_i| > 3$, we get that u_i contains a d.

The abelian cube $u_0u_1u_2$ contains at most one hole. To see this, assume that it contains at least two holes. Then note that $u_0u_1u_2$ must contain at least 21 d's (this is the number of d's between each pair of consecutive $\diamond$'s). Therefore one of u_0, u_1 or u_2 must contain at least 7 d's. It follows that there exist at least seven j's in $\mathbb{Z}$ so that $w(j) = d$ or $w(j) = \diamond$ and $i_2 \leqslant j < i_3$. So by construction of w', there exists an i so that $i_1 < i_2 \leqslant k_i < k_{i+1} < i_3$. Then

$$i_3 = \left|w[0..i_3-1]\right| = \left|w[0..i_1-1]\right| + |u_1| + |u_2| \leqslant i_1 + i_1 + i_1 < 3k_i < k_{i+1} < i_3$$

which is a contradiction. Thus $u_0u_1u_2$ contains at most one hole.

Each u_i has the same number of occurrences of the letter d. To see this, note that since $u_0u_1u_2$ is an abelian cube, we must be able to replace the hole with some letter to get a full word $u_0'u_1'u_2'$ that is an abelian cube. If some u_i and u_j have a different number of d's, then the letter we replace the $\diamond$ with must be a d. However, it follows that if we replace each of the d's in $u_0'u_1'u_2'$ by c, we get a word that is an abelian cube, but is also a subword of w. This contradicts the fact that w is abelian cube-free, so each u_i must contain the same number of d's.

Some, and thus all, of the u_i's must contain at least two d's. To see this, assume that each u_i contains one d. This implies that u_0 does not contain a hole. If it does, the hole must occur in a subword of the form $d\diamond add$. Since $|u_i| > 3$ for all i, either u_0 contains $d\diamond ad$ or u_2 contains dd. Since either case is impossible, u_0 cannot contain a hole. By a similar logic u_1 cannot contain a hole. Therefore u_2 must contain a hole. So there is some j, $w'(k_j) = \diamond$, where $i_2 \leqslant k_j < i_3$. Note that $i_2 \leqslant k_j - 1$, since otherwise $i_2 < k_j + 2 < k_j + 3 < i_3$, implying that u_2 contains dd, which is impossible. Since u_1 contains a d, there must be some s so that $i_1 \leqslant k_s < i_2$. However, this contradicts Lemma 1. Thus each u_i contains at least two d's.

The above implies that there exist r_1, r_2, r_3, r_4 so that $i_1 \leqslant r_1 < r_2 < i_2 \leqslant r_3 < r_4 < i_3$, where $w'(r_i) = d$ for each i. Thus we can write $r_i = k_{s_i} + \lambda_i$ for some $\lambda_i \in \{-1, 0, 2, 3\}$. Note that if $s_1 = s_4$, then $r_1 = k_{s_1} - 1, r_2 = k_{s_1}, r_3 = k_{s_1} + 2$ and $r_4 = k_{s_1} + 3$. Thus $ddedd$ is a subword of w' for some $e \in \{\diamond, a, b, c, d\}$, which is not the case. Therefore, $s_1 < s_4$. This implies that $i_1 \leqslant k_{s_1}$ so we get

$$k_{s_4} \leqslant i_3 = \left|w[0..i_3-1]\right| = \left|w[0..i_1-1]\right| + |u_1| + |u_2| \leqslant 3i_1 \leqslant 3k_{s_1} < k_{s_1+1} \leqslant k_{s_4}$$

This is a contradiction, so the result follows. □

To study larger powers, we first prove two lemmas.

Lemma 2. *Let w be an infinite word over a finite alphabet, and let $p > 1$ be an integer. Assume that there exist infinitely many i's so that, if w' is defined by*

$$w'(j) = \begin{cases} \diamond & \text{if } j = i \\ w(j) & \text{otherwise} \end{cases}$$

then w' is abelian p-free. Then we can insert infinitely many holes in w so that the resulting partial word is abelian $(p+1)$-free.

Proof. Assume w is as above. Then we can choose a sequence $k_0, k_1, \ldots$ so that $k_0 > 0$; for each i, $k_{i+1} > pk_i$; and for each i, if we define w' so that

$$w'(j) = \begin{cases} \diamond & \text{if } j = k_i \\ w(j) & \text{otherwise} \end{cases}$$

then w' is abelian p-free. Note that we can construct such a sequence since there are infinitely many k_i's that meet the last requirement. We can define v by

$$v(j) = \begin{cases} \diamond & \text{if } j = k_i \text{ for some } i \\ w(j) & \text{otherwise} \end{cases}$$

Then assume that $u_0 \cdots u_p$ is an abelian $(p+1)$st power that occurs in v. By Corollary 1, there is at most one $i > 0$ so that u_i contains a hole, and that u_i contains at most one hole. This implies that $u_1 \cdots u_p = v[l..l']$ contains at most one hole. Clearly it must contain at least one hole, since otherwise $u_1 \cdots u_p$ would occur as an abelian pth power in w, which is impossible. Therefore there exists some k_i so that $l \leqslant k_i \leqslant l'$. If we define w' by

$$w'(j) = \begin{cases} \diamond & \text{if } j = k_i \\ w(j) & \text{otherwise} \end{cases}$$

then clearly $u_1 \cdots u_p$ occurs in w'. This, however, contradicts the fact that w' is abelian p-free. Therefore v must be abelian $(p+1)$-free. □

Lemma 3. *There exists an infinite word over a two-letter alphabet such that, if we insert a hole anywhere in the word, the resulting partial word is abelian* 8*-free.*

Proof. There is an infinite binary word w that avoids abelian 4th powers [13]. Then consider any integer $i > 0$. Define w' so that

$$w'(j) = \begin{cases} \diamond & \text{if } j = i \\ w(j) & \text{otherwise} \end{cases}$$

Assume that $u_0 \cdots u_7$ is an abelian 8th power in w'. At most one of the u_i's contains a hole. This implies that there exists some j so that $u_j u_{j+1} u_{j+2} u_{j+3}$ is full. Therefore $u_j u_{j+1} u_{j+2} u_{j+3}$ is a factor of w, which is an abelian 4th power. However, this contradicts the fact that w is abelian 4-free. Thus w' is abelian 8-free. □

Using Lemmas 2 and 3, we get the following result.

Theorem 3. *There exists a partial word with infinitely many holes over a two-letter alphabet that avoids abelian* 9*th powers.*

Proof. Let w, w' be as in Lemma 3. As shown there, w' is abelian 8-free. The theorem then follows from Lemma 2. □

Furthermore, abelian 6th powers can be avoided over a binary alphabet. However, the construction is more complicated than in the case above, and is not the result of simply inserting holes to a word that is produced by a morphism.

Theorem 4. *There exists a partial word with infinitely many holes over a two-letter alphabet that avoids abelian* 6*th powers.*

Proof. Let $A = \{a, b\}$. As mentioned before, there is an infinite binary word w that avoids abelian 4th powers (see [13]). More specifically, it is the fixed point of the morphism $\phi : A^* \to A^*$ defined by $\phi(a) = aaab$ and $\phi(b) = bab$. We also define a morphism $\psi : A^* \to A^*$ by $\psi(a) = b$ and $\psi(b) = a$. Let $v_0 = w[0..119]$ and $v_1 = w[121..240]$. Note that $w(120) = a$. In other words $v_0 a v_1$ is the prefix of w where $|v_0 a v_1| = 241$. Note that we can check with a computer that $\psi(v_0)\diamond v_1$ is abelian 6-free. Similarly, we can check that if u is a subword of w, $30 \leqslant |u| < 60$, then $|u|_a > |u|_b$. This implies that if $|u| \geqslant 30$ and u is a subword of w then $|u|_a > |u|_b$, since we can write $u = u_0 \cdots u_n$ for some n, where $30 \leqslant |u_i| < 60$.

Since $v_0 a v_1$ appears once in w, it must appear infinitely often. Therefore we can define a sequence $k_0 < k_1 < \cdots$ so that $k_0 > 120$, $w[k_i - 120..k_i + 120] = v_0 a v_1$, and $k_i > 7k_{i-1}$. Let $k_{-1} = -1$. We can then define a partial word w' so that

$$w'(j) = \begin{cases} \diamond & \text{if } j = k_i \text{ for some } i \\ w(j) & \text{if } k_i < j < k_{i+1} \text{ for some } i, i \text{ odd} \\ \psi(w(j)) & \text{otherwise} \end{cases}$$

Our goal is to show that w' is abelian 6-free. To see this, assume that $u_0u_1u_2u_3u_4u_5$ is an abelian 6th power in w'. From Corollary 1, at most one of u_1, u_2, u_3, u_4, or u_5 contains a hole. This implies that $u_1u_2u_3u_4u_5$ occurs as a factor of w of the form $\psi(w_0v_0)\diamond v_1w_1$ or $w_0v_0\diamond\psi(v_1w_1)$, where w_0 and w_1 are subwords of w, since otherwise $u_1u_2u_3u_4u_5$ would have to contain more than one hole. We assume it occurs in a word of the form $\psi(w_0v_0)\diamond v_1w_1$, the other case being similar. Note that either u_2, u_3 or u_4 must contain a hole, otherwise we would have that either $u_1u_2u_3u_4$ or $u_2u_3u_4u_5$ is full, so $\diamond$ must occur as a factor in either w or $\psi(w)$. However, both w and $\psi(w)$ are abelian 4-free, so this is impossible. Then consider the possibility that $|u_0| < 30$. However, this implies, since either u_2, u_3 or u_4 contains a hole, that $u_0u_1u_2u_3u_4u_5$ occurs in $\psi(v_0)\diamond v_1$. As mentioned above, from computer testing $\psi(v_0)\diamond v_1$ does not contain any abelian 6th powers. Therefore we must assume that $|u_0| \geqslant 30$. We know that $u_1u_2u_3u_4u_5$ occurs as a subword of $\psi(w_0v_0)\diamond v_1w_1$. Moreover, since the hole is contained in either u_2, u_3 or u_4, we get that u_5 must be contained in v_1w_1, while u_1 must be contained in $\psi(w_0v_0)$. Since u_1 and u_5 are full and $u_1 \uparrow u_5$, we get that $|u_1|_a = |u_5|_a$ and $|u_1|_b = |u_5|_b$. On the other hand, since u_5 is a subword of w and $|u_5| \geqslant 30$ we know that $|u_5|_a > |u_5|_b$, while $|u_1|_a < |u_1|_b$ because u_1 is a subword of $\psi(w)$ and $|u_1| \geqslant 30$. This is a contradiction, so w' does not contain any abelian 6th powers. □

3.2. The case with finitely many holes

The proof of Theorem 4 produces the following corollary related to the finitely many hole case.

Corollary 2. *There exists a two-sided infinite partial word with one hole over a two-letter alphabet that avoids abelian* 5*th powers.*

Proof. Let a, b be distinct letters. Let w and ψ be as in the proof of Theorem 4. We claim that the partial word $v = \mathrm{rev}(\psi(w))\diamond w$ avoids abelian 5th powers. To see this, assume that $u_0u_1u_2u_3u_4$ is an abelian 5th power in v. Note that either u_1, u_2 or u_3 must contain the hole, since otherwise we either have that $u_0u_1u_2u_3$ occurs as an abelian 4th power in $\mathrm{rev}(\psi(w))$, or $u_1u_2u_3u_4$ occurs as an abelian 4th power in w, both of which are impossible. Therefore u_0 is a subword of $\mathrm{rev}(\psi(w))$ and u_4 is a subword of w. Then consider the case where $|u_0| < 30$. Then $u_0u_1u_2u_3u_4$ must be contained in $v[-120..120]$, but a simple computer check tells us that $v[-120..120]$ does not contain any abelian 5th power. On the other hand, if $|u_0| \geqslant 30$, then we know from the proof of Theorem 4 that since u_4 is a subword of w and $|u_4| \geqslant 30$, $|u_4|_a > |u_4|_b$. Similarly, $|u_0|_a < |u_0|_b$ since u_0 is a subword of length at least 30 of $\mathrm{rev}(\psi(w))$. However, as in the proof of Theorem 4, this leads to a contradiction, so the corollary follows. □

It is already known that abelian squares can be avoided by a partial word with one hole over four letters, where the alphabet size is minimal [6]. So we move on to higher abelian powers.

Proposition 1. *There exists an infinite partial word with one hole over a two-letter alphabet that avoids abelian* 4*th powers.*

Proof. The proof is similar to the one in [6]. We use the morphism $\phi : \{a, b\}^* \to \{a, b\}^*$ provided by Aberkane and Currie in [1], defined by

$$\phi(a) = aabaaababbbabaaababbaaaba$$

$$\phi(b) = bbabbbabaaababbbabaabbbab$$

The length of each image is 25. The Parikh vectors are $P(\phi(a)) = \langle 15, 10\rangle$ and $P(\phi(b)) = \langle 10, 15\rangle$. We want to show that $\diamond\phi^n(a)$ is abelian 4-free for all $n \geqslant 0, n \in \mathbb{Z}$. Since $\phi^n(a)$ is abelian 4-free, it is sufficient to check if we have abelian 4th powers that start with the hole.

Assume some prefix of $\diamond\phi^n(a)$ is an abelian 4th power, denoted by $u_0u_1u_2u_3$, where $|u_0| = |u_1| = |u_2| = |u_3|$. We can write $u_0u_1u_2u_3 = \diamond\phi(w_0)\phi(e)\phi(w_1)xu_2u_3$, where $e \in A = \{a, b\}$, $w_0, w_1, x \in A^*$ are such that $\diamond\phi(w_0)$ is a prefix of u_0, u_0 is a proper prefix of $\diamond\phi(w_0e)$, and x is a proper prefix of either $\phi(a)$ or $\phi(b)$. If we delete the same number of occurrences of any given image of ϕ present in both $\phi(w_0)$ and $\phi(w_1)$, we claim we only need to check the case when $0 \leqslant |w_0| - |w_1| < 2$. To see this, if $|w_1| > |w_0|$, then $|u_1| > |u_0|$, a contradiction. If $|w_0| - |w_1| \geqslant 2$, then $|u_1| < |u_0|$, a contradiction.

Denote by u_0' the word obtained from u_0 after replacing the hole by a letter in A so that $P(u_0') = P(u_1)$. Since w_0 and w_1 do not share any letter, w_0 does not contain both a's and b's. We prove the case where w_0 consists of a number of a's and w_1 of a number of b's (the other case follows similarly).

We can build a system of equations since the number of a's (resp., b's) in u_0' is the same as the number of a's (resp., b's) in u_1:

$$15\big|\phi(w_0)\big|_{\phi(a)} = 10\big|\phi(w_1)\big|_{\phi(b)} + \lambda_a$$

$$10\big|\phi(w_0)\big|_{\phi(a)} = 15\big|\phi(w_1)\big|_{\phi(b)} + \lambda_b$$

Here, the parameter λ_i relates to the number of occurrences of letter i created by $\diamond$, $\phi(e)$ and x. More precisely, there exist $y_0, y_1 \in A^*$, $y_1 \neq \varepsilon$, such that $\phi(e) = y_0y_1$, $u_0 = \diamond\phi(w_0)y_0$ and $u_1 = y_1\phi(w_1)x$. Since $|u_0'|_a = |u_1|_a$, the following equality

$$\diamond_a + 15\big|\phi(w_0)\big|_{\phi(a)} + |y_0|_a = |y_1|_a + 10\big|\phi(w_1)\big|_{\phi(b)} + |x|_a$$

holds, where $\diamond_a = 1$ if $\diamond$ is replaced by a and $\diamond_a = 0$ otherwise. So $\lambda_a = |y_1|_a + |x|_a - \diamond_a - |y_0|_a$. The parameter λ_b is similarly defined.

It can be checked that the only scenarios that lead to non-negative integer solutions for $\diamond_i$, $|y_0|_i$, $|y_1|_i$, $|x|_i$, $i \in \{a, b\}$, are when $w_0 = w_1 = \varepsilon$, or $w_0 = a$ and $w_1 = \varepsilon$ (note that because of the hole at the beginning, there is one $\diamond_i$ that must be 1 while the other must be 0). For the first case, $u_0u_1u_2u_3 = \diamond\phi(e)xu_2u_3$, while for the second case, $u_0u_1u_2u_3 = \diamond\phi(a)\phi(e)xu_2u_3$. It is easy to verify that neither of them is an abelian 4th power. $\square$

Corollary 3. *There exists a two-sided infinite partial word with two holes over a three-letter alphabet that avoids abelian* 4*th powers.*

Proof. Let $A = \{a, b\}$. For $w \in A^*$, let $\phi'(w) = \mathrm{rev}(\phi(w))$, with $\phi : A^* \to A^*$ being the morphism from the proof of Proposition 1. Hence, $\phi'(w)$ is abelian 4-free for all abelian 4-free words $w \in A^*$, and $\phi'^n(a)\diamond$ is abelian 4-free for all $n \geqslant 0$, $n \in \mathbb{Z}$. The partial word $\phi'^n(a)\diamond c\diamond\phi^n(a)$, where $n \geqslant 0$, $n \in \mathbb{Z}$, is abelian 4-free. To see this, suppose to the contrary that it contains an abelian 4th power, which must have the c. However, since there are at most two holes in the abelian 4th power, at least one of the four segments has neither symbol c nor $\diamond$. $\square$

Using similar techniques, we can prove results about other powers as well.

Proposition 2. *There exists an infinite partial word with one hole over a three-letter alphabet A that avoids abelian cubes, except for the trivial abelian cube $\diamond aa$, where $a \in A$.*

Proof. We use the abelian cube-free morphism over a ternary alphabet of Dekking [13]. Let $A = \{a, b, c\}$, and $\phi : A \to A^*$ be the non-uniform morphism defined by $\phi(a) = aabc$, $\phi(b) = bbc$ and $\phi(c) = acc$. We show that $\diamond\phi^n(a)$ is abelian cube-free for all $n \geqslant 0$, $n \in \mathbb{Z}$, except for the trivial abelian cube $\diamond aa$. We assume there exists an abelian cube and reach a contradiction. For the same reason as in Proposition 1, the abelian cube has to contain the hole.

Let the abelian cube be $u_0u_1u_2 = \diamond\phi(w_0)\phi(e)\phi(w_1)xu_2$, where $e \in A$, w_0, $w_1 \in A^*$, $\diamond\phi(w_0)$ is a prefix of u_0, u_0 is a proper prefix of $\diamond\phi(w_0e)$, and x is a proper prefix of some image of ϕ. Delete the same number of occurrences of any given image of ϕ present in both $\phi(w_0)$ and $\phi(w_1)$. After replacing the hole with some letter to get u_0' from u_0, we construct a system of equations based on the fact that the number of occurrences of each letter in u_0' has to be equal to that in u_1:

$$2\left|\phi(w_0)\right|_{\phi(a)} + \left|\phi(w_0)\right|_{\phi(c)} = 2\left|\phi(w_1)\right|_{\phi(a)} + \left|\phi(w_1)\right|_{\phi(c)} + \lambda_a$$
$$\left|\phi(w_0)\right|_{\phi(a)} + 2\left|\phi(w_0)\right|_{\phi(b)} = \left|\phi(w_1)\right|_{\phi(a)} + 2\left|\phi(w_1)\right|_{\phi(b)} + \lambda_b$$
$$\left|\phi(w_0)\right|_{\phi(a)} + \left|\phi(w_0)\right|_{\phi(b)} + 2\left|\phi(w_0)\right|_{\phi(c)} = \left|\phi(w_1)\right|_{\phi(a)} + \left|\phi(w_1)\right|_{\phi(b)} + 2\left|\phi(w_1)\right|_{\phi(c)} + \lambda_c$$

As in the proof of Proposition 1, λ_i is related to the number of occurrences of letter i created by $\diamond$, $\phi(e)$ and x. Again, let $y_0, y_1 \in A^*$, $y_1 \neq \varepsilon$, be such that $\phi(e) = y_0y_1$, $u_0 = \diamond\phi(w_0)y_0$ and $u_1 = y_1\phi(w_1)x$. It can be checked that there are no non-negative integer solutions for $\diamond_i$, $|y_0|_i$, $|y_1|_i$, $|x|_i$, $i \in \{a, b, c\}$, unless $|w_0| = 0$ or $|w_1| = 0$.

When $|w_0| = 0$, if $|w_1| \geqslant 2$, then $|u_0| < |u_1|$, which leads to a contradiction. When $|w_1| = 0$, if $|w_0| \geqslant 3$, then $|u_1| < |u_0|$. Therefore, the ordered pair $(|w_0|, |w_1|)$ can only be $(0, 0)$, $(0, 1)$, $(1, 0)$ or $(2, 0)$.

When $|w_0| = |w_1| = 0$, the abelian cube becomes $\diamond\phi(e)xu_2$, which can be checked to be abelian cube-free. When $|w_0| = 2$ and $|w_1| = 0$, the abelian cube becomes $\diamond\phi(fg)\phi(e)xu_2$, for some $f, g \in A$. Then the only way to have $|u_0| = |u_1|$ is to let $|\phi(fg)| = 6$ and $|\phi(e)x| = 7$. Thus $\phi(e)x = \phi(a)aab$ since $|x| < 4$. It is easy to verify that no combination creates such abelian cube. When $|w_0| = 0$ and $|w_1| = 1$, the abelian cube becomes $\diamond\phi(e)\phi(f)xu_2$, for some $f \in A$, and the only way to have $|u_0| = |u_1|$ is if $\phi(e) = \phi(a)$ and $|\phi(f)| = 3$, which can be verified to be impossible. Therefore, $|w_0| = 1$ and $|w_1| = 0$. However, we can check using a computer program that this situation is also impossible. $\square$

4. Counting abelian p-free partial words

In this section, we investigate the number of abelian p-free partial words of length n with h holes over a given alphabet. We begin by noting that Theorem 4 allows us to give a lower bound on the number of binary words of length n with h holes that avoid abelian 6th powers.

Theorem 5. *Let $h > 0$ be an integer. Let $c_{n,h,6}$ denote the number of words over a two-letter alphabet of length n with h holes that avoid abelian 6th powers. Then there exist an integer $N > 0$ and real numbers $r > 1$, $\beta > 0$ such that for all $n > N$, $c_{n,h,6} \geqslant \beta r^n$.*

Proof. Assume that h is even, the other case being similar. Let ϕ, ψ and w be as in Theorem 4. Then by the proof of Theorem 4, there exists a word of the form $v = w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond w_h$, where each w_i is a subword of w, w_h is infinite, and v avoids abelian 6th powers. There exist an infinite subword u of w and a finite word x, $|x| > 5|w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond|$, such that if u' is any prefix of u then $x\phi(u')$ is a prefix of w_h.

We can then define the multi-valued substitution $\theta : A^* \to 2^{A^*}$ so that $\theta(a) = \{aaab\}$ and $\theta(b) = \{bab, abb\}$. From [12], θ is abelian 4-free. Therefore let $N = |w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond x| + 16$. Consider $n > N$ and denote the prefix of v of length n as v_n. We can write $w_h = x\phi(u')y$ for some prefix u' of u and some word y. Note that since $|\phi(\alpha)| \leqslant 4$ for all $\alpha \in A$, we can choose u' so that

$$\left|u'\right| \geqslant \left\lfloor \frac{|v_n| - |w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond x|}{4} \right\rfloor = \left\lfloor \frac{n-N+16}{4} \right\rfloor \geqslant \frac{n-N+12}{4}$$

First, we claim that if $\gamma \in \theta(u')$, then $w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond x\gamma y$ is abelian 6-free. To see this, note that if $w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond x\gamma y$ contains an abelian 6th power $u_0u_1u_2u_3u_4u_5$, then one of the u_i's must overlap with γ, since $w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond x$ and y are known to be abelian 6-free. However, since $|x| > 5|w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond|$, the only u_i that can intersect with $w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond$ is u_0, so $u_1u_2u_3u_4$ is contained in $x\gamma y$. By construction $x\phi(u')y$ is a factor of w, so $x\gamma y$ is a factor of some word in $\theta(w)$. However, the fact that w is abelian 4-free implies that all elements in $\theta(w)$ are as well, so $x\gamma y$ must be abelian 4-free. It follows that $w_0\diamond\psi(w_1)\diamond\cdots\diamond\psi(w_{h-1})\diamond x\gamma y$ is abelian 6-free. Note that this implies that $|\theta(u')| \leqslant c_{n,h,6}$.

Moreover, since u' is a factor of w, it does not have $aaaa$ as a subword. Therefore $|u'|_b \geqslant \lfloor\frac{|u'|}{4}\rfloor \geqslant \frac{|u'|-3}{4}$. By construction, $|\theta(u')| = 2^{|u'|_b}$. Using the same analysis as in [12], this gives us that $\theta(u')$ contains at least $2^{-\frac{15}{16}}2^{\frac{1}{16}(n-N+16)}$ words. Therefore, setting $r = 2^{\frac{1}{16}}$ and $\beta = 2^{-\frac{15}{16}-\frac{N}{16}+1}$, we get that $\beta r^n \leqslant |\theta(u')| \leqslant c_{n,h,6}$, so the claim follows. $\square$

Similar methods as in the above theorem can be used to count words avoiding other abelian powers.

Theorem 6. *Let $h > 0$ be an integer. Let $c_{n,h,4}$ denote the number of words over a three-letter alphabet of length n with h holes that avoid abelian 4th powers. Then there exist an integer $N > 0$ and real numbers $r > 1$, $\beta > 0$ such that for all $n > N$, $c_{n,h,4} \geqslant \beta r^n$.*

Proof. This is similar to the proof of Theorem 5, except that we use the tools of Theorem 1. $\square$

In the above we counted, for some p, partial words of length n with h holes that avoid abelian pth powers over an alphabet A, where A is the smallest alphabet known to admit partial words with infinitely many holes that avoid abelian pth powers. We have, however, proved in Section 3.2 that in some cases we can create a partial word with one hole of the form $\diamond w$, for some word w over an alphabet B that is strictly smaller than A, that avoids abelian pth powers. We now show, in some of these cases, that the number of such partial words grows exponentially with length.

Theorem 7. *The number of partial words over a three-letter alphabet of length n with one hole, of the form $\diamond w$, that avoid non-trivial abelian cubes grows exponentially with n (a* trivial *abelian cube has the form $\diamond aa$ or $a\diamond a$ or $aa\diamond$ for some letter a).*

Proof. We use the substitution σ provided by Aberkane et al. in [2], which is defined on $A = \{a, b, c\}$, where $\sigma(a) = \{aabaac\}$, $\sigma(b) = \{bbabbc\}$, and $\sigma(c) = \{aaccbc, bccaac\}$. They prove that the set

$$S = \left\{w \;\middle|\; uwv \in \sigma^n(a) \text{ for some integer } n > 0, \text{ and some words } u, v\right\}$$

is abelian cube-free. We can prove that if $w \in S$, then $\diamond w$ is non-trivial abelian cube-free using the same method as in Proposition 2.

Let w be the prefix of length $n-1$ of a word in $\sigma^\omega(a)$. Then $w \in \sigma(v)y$, where y, $|y| \leqslant 5$, is a prefix of a word in $\sigma^\omega(a)$. Since $|v|_c \sim |v|/4$ and $\sigma(v)y$ contains at least $2^{|v|_c}$ words (shown in [2]), and $n = |\diamond w| = |\sigma(v)y| + 1 \leqslant 6|v| + 6$, the number of partial words over a three-letter alphabet of length n with one hole, of the form $\diamond w$, that avoid non-trivial abelian cubes is at least $2^{|v|/4} \geqslant 2^{(n-6)/24} = 2^{-6/24}2^{n/24} > r^n$, where $r = 2^{1/24}$. $\square$

Theorem 8. *The number of partial words over a two-letter alphabet of length n with one hole, of the form $\diamond w$, that avoid abelian 4th powers grows exponentially with n.*

Proof. Let c_n be the number of binary partial words of length n that avoid abelian 4th powers and that have exactly one hole at the beginning. We show that c_n grows exponentially with n. Our technique is similar to that in [12]. Let $A = \{a, b\}$, and let σ be the abelian 4-free multi-valued substitution on A^* so that $\sigma(a) = \{aaab\}$ and $\sigma(b) = \{abb, bab\}$. Moreover, let S be the set of words that are prefixes of either $aaab$, bab, or abb. Also, consider the matrix

$$M = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$$

Then if u and v are words, $u \in \sigma(v)$, we get that $P(u) = MP(v)$.

Consider $n > 0$. We claim that if $w \in \sigma^n(b)$ and $w(0) = b$ we can replace the letter of w at position 0 by $\diamond$ and have the resulting partial word, w', avoid abelian 4th powers.

In order to see this, assume that w' contains an abelian 4th power $u_0u_1u_2u_3$. Assume that it is the shortest such power. Moreover, $|u_0| > 4$ by computer testing. Note that u_0 must start with $\diamond$. We can write $u_i = w'[il..(i+1)l)$ for some $l > 0$. Let u'_0, u'_1, u'_2 and u'_3 be u_0, u_1, u_2 and u_3 with the hole replaced by a. Similarly, let u''_0, u''_1, u''_2 and u''_3 be u_0, u_1, u_2 and u_3 with the $\diamond$ replaced by b. Since $u_0u_1u_2u_3$ is an abelian 4th power, either $u'_0u'_1u'_2u'_3$ or $u''_0u''_1u''_2u''_3$ must be one. Since $u''_0u''_1u''_2u''_3$ is a subword of w, it follows that it is not an abelian 4th power, so $u'_0u'_1u'_2u'_3$ must be one.

Now we can define a morphism $f : A^* \to \mathbb{Z}_5$ by $f(a) = 1$ and $f(b) = 2$. The fact that $u'_0u'_1u'_2u'_3$ is an abelian 4th power implies that $f(u'_0) = f(u'_1) = f(u'_2) = f(u'_3)$. Moreover, note that $f(u''_i) = f(u'_i)$ for $i > 0$, while $f(u'_0) = f(u''_0) - f(b) + f(a) = f(u''_0) - 1$ so that $r = f(u''_0) - 1 = f(u''_1) = f(u''_2) = f(u''_3)$. It is also worth noting that for any word v, if $u \in \sigma(v)$ then $f(u) = 0$. Note that $f(S) = \{0, 1, 2, 3\}$.

Since $|u_0| > 4$, there exist $e_0, e_1, e_2, e_3, e_4 \in A$ and $x_0, x_1, x_2, x_3 \in A^*$ so that $e_0x_0e_1x_1e_2x_2e_3x_3e_4$ is a prefix of w and so that there exist $y_0, y_1, y_2, y_3, v_0, v_1, v_2, v_3, v_4, v'_0, v'_1, v'_2, v'_3, v'_4 \in A^*$ satisfying $v'_i \neq \varepsilon$, $y_i \in \sigma(x_i)$, $v_iv'_i \in \sigma(e_i)$ and $v'_iy_iv_{i+1} = u''_i$. Note that for all i, $f(y_i) = 0$, and $f(v_iv'_i) = 0$ so $f(v_i) = -f(v'_i)$.

Note that $e_0 = b$, $v'_0 = bab$ and $v_0 = \varepsilon$, so $f(v'_0) = 0$. Moreover, note that $r + 1 = f(u''_0) = f(v'_0) + f(y_0) + f(v_1) = f(v_1)$. Furthermore, for $i > 0$ we get by induction that $r = f(u''_i) = f(v'_i) + f(y_i) + f(v_{i+1}) = -f(v_i) + f(y_i) + f(v_{i+1}) = -ir - 1 + f(v_{i+1})$, so $f(v_{i+1}) = (i+1)r + 1$. By definition we have that $f(v_i) \in f(S) = \{0, 1, 2, 3\}$ for all $i > 0$. Thus $r + 1, 2r + 1, 3r + 1, 4r + 1 \in f(S)$. However, it is easy to check that this implies that $r = 0$. Therefore $f(v_i) = 1$ for all $i > 0$. Since $v_i \in S$, we can easily check that $v_i = a$ if $i > 0$. Therefore either $e_i = a$ and $v_iv'_i = aaab$, or $e_i = b$ and $v_iv'_i = abb$.

Consider $i, j > 0$. Then $u''_i = v'_iy_ia$ is a permutation of $u''_j = v'_jy_ja$. This implies that $av'_iy_i = v_iv'_iy_i$ is a permutation of $av'_jy_j = v_jv'_jy_j$. However, note that since $v_iv'_iy_i \in \sigma(e_ix_i)$, $v_jv'_jy_j \in \sigma(e_jx_j)$ and M is invertible,

$$P(e_ix_i) = M^{-1}P\left(v_iv'_iy_i\right) = M^{-1}P\left(v_jv'_jy_j\right) = P(e_jx_j)$$

A similar argument can be used to show that e_ix_i is a permutation of ax_0 for all $i > 0$, so that $ax_0e_1x_1e_2x_2e_3x_3$ is an abelian 4th power.

Note that since $e_0x_0e_1x_1e_2x_2e_3x_3e_4$ is a prefix of w, it follows that $\diamond x_0e_1x_1e_2x_2e_3x_3e_4$ is a prefix of w'. However, we know from above that $ax_0e_1x_1e_2x_2e_3x_3$ is an abelian 4th power, so $\diamond x_0e_1x_1e_2x_2e_3x_3$ is an abelian 4th power. Moreover, it is strictly shorter than $u_0u_1u_2u_3$. This is a contradiction, so it follows that w' is indeed abelian 4-free.

This implies that for $n > 0$, we have

$$c_n \geqslant \left|\left\{|v| = n \;\middle|\; v \text{ is the prefix of some word in } \sigma^{\omega}(b) \text{ such that } v(0) = b\right\}\right|$$

However, we know from [12] that

$$\left|\left\{|v| = n \;\middle|\; v \text{ is the prefix of some word in } \sigma^{\omega}(b)\right\}\right|$$

grows exponentially with respect to n, which implies that

$$\left|\left\{|v| = n \;\middle|\; v \text{ is the prefix of some word in } \sigma^{\omega}(b) \text{ such that } v(0) = b\right\}\right|$$

does as well, so the theorem follows. (Note that Currie actually proves exponential growth for prefixes of elements in $\sigma^{\omega}(a)$ instead of $\sigma^{\omega}(b)$, but the argument is exactly the same.) $\square$

5. Inserting arbitrarily many holes

In this section, we investigate whether infinite words can be constructed that remain abelian p-free after an arbitrary insertion of holes. We begin by inserting finitely many holes. In Section 3.1, we saw that there exists an infinite word over a two-letter alphabet such that, if we insert a hole anywhere in the word, the resulting partial word is abelian 8-free.

Proposition 3. *There exists an infinite abelian cube-free word over a nine-letter alphabet such that, if we insert a hole anywhere in the word, the resulting partial word is non-trivial abelian cube-free.*

Proof. Consider words w_0 over $A_0 = \{a_0, b_0, c_0\}$, w_1 over $A_1 = \{a_1, b_1, c_1\}$, and w_2 over $A_2 = \{a_2, b_2, c_2\}$ that are abelian cube-free (see [13]). Then we claim that the word $w = w_0(0)w_1(0)w_2(0)w_0(1)w_1(1)w_2(1)\cdots$ is non-trivial abelian cube-free after inserting one arbitrary hole. Suppose w contains the factor $u_0u_1u_2$, which becomes a non-trivial abelian cube after inserting a hole. If $|u_0| \equiv 1 \bmod 3$, each of the u_i's contains one more letter from a different alphabet (for example, if u_0 contains one more letter in A_0, one of u_1 and u_2 will contain one more letter in A_1 and the other in A_2). Then it is impossible to match the number of letters of three different alphabets by inserting one hole. Similarly, if $|u_0| \equiv 2 \bmod 3$, one of the u_i's contains two more letters, one in A_0 and one in A_1. The other two u_i's either contain two more letters in A_0 and A_2, or in A_1 and A_2. Thus one hole cannot produce an abelian cube. Therefore, $|u_0| \equiv 0 \bmod 3$, which means that each u_i contains the same number of letters from each of A_0, A_1 and A_2. This means the hole can only replace a letter by another one in the same alphabet, let us say, A_0. Since the replacement of the hole does not affect letters in A_1 and A_2, either the abelian cube does not contain any letters in A_1 or A_2, which leads to a trivial abelian cube, or the letters in A_1 or A_2 originally build an abelian cube, which becomes a subword of w_1 or w_2, contradicting the fact that both of them are abelian cube-free. $\square$

Using similar methods, we can show the following.

Proposition 4. *There exists an infinite abelian 4-free word over a four-letter alphabet such that, if we insert a hole anywhere in the word, the resulting partial word is non-trivial abelian 4-free.*

Proof. Consider $w = w_0(0)w_1(0)w_0(1)w_1(1)\cdots$, where w_0 is an abelian 4-free word over $A_0 = \{a_0, b_0\}$ and w_1 is another one over $A_1 = \{a_1, b_1\}$. Suppose that $u_0u_1u_2u_3$, where $|u_0| = |u_1| = |u_2| = |u_3| = l$, is a factor of w that becomes an abelian 4th power after inserting a hole. If l is odd, then u_0 and u_2 contain one more letter from one alphabet, while u_1 and u_3 contain one more letter from the other alphabet, which is impossible. Thus l is even, and each u_i contains $\frac{l}{2}$ letters from A_0 and $\frac{l}{2}$ letters from A_1. Since the hole replaces a letter from some alphabet, say A_0, we can remove all the letters from A_0 that are in $u_0u_1u_2u_3$, and the remaining word is an abelian 4th power in w_1, which is a contradiction. $\square$

Our goal is now to prove that we cannot avoid abelian pth powers under arbitrary insertion of holes. Moreover, we can make the abelian pth powers we get as large as we want. We can also look at the following results as saying that every infinite word has a subword that is almost an abelian pth power. We begin by proving a lemma giving a necessary and sufficient condition for the existence of abelian pth powers.

Lemma 4. *Let $p > 1$ be an integer, and let $v_0 \cdots v_{p-1}$ be a partial word over a k-letter alphabet $A = \{a_0, \ldots, a_{k-1}\}$ such that $|v_i| = |v_0|$, for all i. Let $m_i = \max_j\{|v_j|_{a_i}\}$, for $0 \leqslant i < k$. Then $v_0 \cdots v_{p-1}$ is an abelian pth power if and only if $m_0 + \cdots + m_{k-1} \leqslant |v_0|$.*

Proof. First, assume that $m_0 + \cdots + m_{k-1} \leqslant |v_0| = \cdots = |v_{p-1}|$. Then v_i has $|v_i| - (|v_i|_{a_0} + \cdots + |v_i|_{a_{k-1}}) \geqslant (m_0 + \cdots + m_{k-1}) - (|v_i|_{a_0} + \cdots + |v_i|_{a_{k-1}}) \geqslant 0$ holes. We can then produce v_i' from v_i by replacing $m_0 - |v_i|_{a_0}$ holes with the letter a_0, $m_1 - |v_i|_{a_1}$ holes with a_1, and so on. Replace the remaining holes with the letter a_0. Note that the above procedure is possible due to the fact that v_i has at least $m_0 - |v_i|_{a_0} + m_1 - |v_i|_{a_1} + \cdots + m_{k-1} - |v_i|_{a_{k-1}} = m_0 + \cdots + m_{k-1} - (|v_i|_{a_0} + \cdots + |v_i|_{a_{k-1}})$ holes. Then $|v_i'|_{a_j} = m_j = |v_0'|_{a_j}$ for all $j > 0$, while $|v_i'|_{a_0} = |v_0| - m_1 - \cdots - m_{k-1} = |v_0'|_{a_0}$. Therefore each v_i' is a permutation of v_0', so $v_0' \cdots v_{p-1}'$ is an abelian pth power. Therefore $v_0 \cdots v_{p-1}$ is an abelian pth power.

On the other hand, assume that $v_0 \cdots v_{p-1}$ is an abelian pth power. This implies that we can fill in the holes to get a full word $v_0' \cdots v_{p-1}'$ that is an abelian pth power. However, this implies that $|v_0'|_{a_i} = |v_j'|_{a_i} \geqslant |v_j|_{a_i}$ for every i, j, so $|v_0'|_{a_i} \geqslant m_i$. It follows that $|v_0| = |v_0'| = |v_0'|_{a_0} + \cdots + |v_0'|_{a_{k-1}} \geqslant m_0 + \cdots + m_{k-1}$. $\square$

We now consider a generalization of arbitrary insertion. In our definition of arbitrary insertion, we require two arbitrarily inserted holes to be separated by at least two letters. Below, however, we consider an arbitrary positive integer m and analyze what happens when we require each pair of consecutive holes to be separated by at least m letters. It actually turns out that this generalization is very useful for our purposes.

The next technical lemma says, in some sense, that if an infinite word contains subwords that are arbitrarily close to being abelian pth powers, then it must contain an abelian pth power under arbitrary insertion.

Lemma 5. *Let w be an infinite word over a finite alphabet, and let $p > 1, m > 0$ be integers. Assume that for every integer $l > 0$ and real number $\epsilon > 0$, there exists a subword $u_0 \cdots u_{p-1}$ of w so that $|u_i| > l$ for all i, and so that, if $a \in A$ then $||u_i|_a - |u_{i'}|_a| < \epsilon |u_j|$ for all i, i', j. Then we can insert holes in w so that each pair of consecutive holes is separated by at least m letters, and so that the resulting partial word contains an abelian pth power $v_0 \cdots v_{p-1}$, where $|v_i| \geqslant l$ for all $0 \leqslant i < p$.*

Proof. Let p, m and w be as above, and choose any $l > 0$. Let $A = \{a_0, \ldots, a_{k-1}\}$. Consider $\epsilon > 0$ such that $2k\epsilon(k(m+1)+1)(2(2p-1)k+1) < 1$. Similarly, consider $L > l$ such that $L > 4m(k+1)$. Then there exists a subword $u_0 \cdots u_{p-1}$ of w so that $|u_i| > L$ for all i and so that, if $a \in A$, then $||u_i|_a - |u_{i'}|_a| < \epsilon |u_j|$ for all i, i' and j. This implies that $||u_i| - |u_{i'}|| < k\epsilon |u_j|$ for all i, i' and j.

Let $\mu = \min |u_i| > L$. Then it is clear that $||u_i|_a - |u_{i'}|_a| < \epsilon\mu$, $||u_i| - |u_{i'}|| < k\epsilon\mu$, and $||u_i| - \mu| < k\epsilon\mu$ for all i, i'. Let v_0 be the word consisting of the first μ letters of $u_0 \cdots u_{p-1}$, v_1 the next μ, and so on up to v_{p-1}. Note that the length of the factor of u_0 that does not overlap with v_0 is $||u_0| - |v_0|| = ||u_0| - \mu| < k\epsilon\mu$, the length of the factor of v_1 that does not overlap with u_1 plus the length of the factor of u_1 that does not overlap with v_1 is at most $||u_0| - |v_0|| + ||u_0u_1| - |v_0v_1|| = ||u_0| - \mu| + ||u_0| - \mu + |u_1| - \mu| \leqslant 2||u_0| - \mu| + ||u_1| - \mu| < 3k\epsilon\mu$, the length of the factor of v_2 that does not overlap with u_2 plus the length of the factor of u_2 that does not overlap with v_2 is at most

$||u_0u_1| - |v_0v_1|| + ||u_0u_1u_2| - |v_0v_1v_2|| \leqslant 2||u_0| - \mu| + 2||u_1| - \mu| + ||u_2| - \mu| < 5k\epsilon\mu$, and so on. In particular,

$$\big||v_0|_a - |u_0|_a\big| < k\epsilon\mu, \quad \big||v_1|_a - |u_1|_a\big| < 3k\epsilon\mu, \quad \ldots, \quad \big||v_{p-1}|_a - |u_{p-1}|_a\big| < (2p-1)k\epsilon\mu$$

which implies that

$$\big||v_i|_a - |v_{i'}|_a\big| \leqslant \big||v_i|_a - |u_i|_a\big| + \big||u_i|_a - |u_{i'}|_a\big| + \big||u_{i'}|_a - |v_{i'}|_a\big| < \big(2(2p-1)k+1\big)\epsilon\mu$$

for all $a \in A$, and all i, i'. Also $|v_i| = \mu > L$ for all i.

Write $v_0 \cdots v_{p-1} = w[i_0..i_1 - 1] \cdots w[i_{p-1}..i_p - 1]$ where $v_j = w[i_j..i_{j+1})$ for all j. We can assume that $|v_i|_{a_0} > \frac{\mu}{k} - (2(2p-1)k+1)\epsilon\mu$ for all i (to see this, by the pigeonhole principle there exists an i such that $|v_0|_{a_i} \geqslant \frac{\mu}{k}$; we assume that $i = 0$ without loss of generality, and using the fact that $||v_i|_a - |v_{i'}|_a| < (2(2p-1)k+1)\epsilon\mu$ for all $a \in A$ and all i, i', we can get that the assumption is safe to make). Let $\chi = \min |v_j|_{a_0}$.

Let $t = \lfloor \frac{\chi - 2m}{m+1} \rfloor$. Note that, since $\mu > L > 4m(k+1)$, we get

$$\begin{aligned} t &> \frac{\frac{\mu}{k} - (2(2p-1)k+1)\epsilon\mu - 2m}{m+1} \\ &= \frac{\mu}{2k(m+1)} + \frac{\mu - 4mk}{2k(m+1)} - \frac{2(2p-1)k+1}{m+1}\epsilon\mu \\ &> \frac{\mu}{2k(m+1)} - \frac{2(2p-1)k+1}{m+1}\epsilon\mu \end{aligned}$$

Consider $0 \leqslant j < p$. Then we can define $k_j^{(0)}$ to be the position of the first occurrence of a_0 in v_j, $k_j^{(1)}$ the position of the $(m+2)$nd occurrence, and in general $k_j^{(\alpha)}$ to be the position of the $(\alpha m + \alpha + 1)$th occurrence. Assume that β_j is the largest integer where $k_j^{(\beta_j)}$ is defined. Then we produce w' by inserting a hole in w at position $k_j^{(\alpha)}$ for each j, and each $0 \leqslant \alpha < \beta_j$. Let $v'_j = w'[i_j..i_{j+1})$. Each pair of consecutive holes is separated by at least m letters, and each v'_i contains at least t holes.

Let $m_i = \max |v'_j|_{a_i}$. Then clearly $m_i < |v_0|_{a_i} + (2(2p-1)k+1)\epsilon\mu$, where $0 < i < k$, while $m_0 < |v_0|_{a_0} + (2(2p-1)k+1)\epsilon\mu - t$. Then we get

$$\begin{aligned} m_0 + \cdots + m_{k-1} &\leqslant |v_0|_{a_0} + \big(2(2p-1)k+1\big)\epsilon\mu - t \\ &\quad + |v_0|_{a_1} + \big(2(2p-1)k+1\big)\epsilon\mu + \cdots \\ &\quad + |v_0|_{a_{k-1}} + \big(2(2p-1)k+1\big)\epsilon\mu \\ &= k\big(2(2p-1)k+1\big)\epsilon\mu + \mu - t \\ &< \mu + k\big(2(2p-1)k+1\big)\epsilon\mu - \left(\frac{\mu}{2k(m+1)} - \frac{2(2p-1)k+1}{m+1}\epsilon\mu\right) \\ &= \mu + \frac{2k\epsilon(k(m+1)+1)(2(2p-1)k+1) - 1}{2k(m+1)}\mu < \mu = \big|v'_0\big| \end{aligned}$$

So it follows from Lemma 4 that $v'_0 \cdots v'_{p-1}$ is an abelian pth power, since $|v'_0| > l$, the claim is proved. □

Finally, we look at what happens in general. We show that, no matter how large p is or how many letters we require between each pair of consecutive inserted holes, it is always possible to insert holes in an infinite word so that the resulting partial word contains arbitrarily long abelian pth powers.

Theorem 9. *Let w be an infinite word over a finite alphabet, and let $p > 1$, $m > 0$, $l > 0$ be integers. Then we can insert holes in w so that each pair of consecutive holes is separated by at least m letters, and so that the resulting partial word contains an abelian pth power $u_0 \cdots u_{p-1}$, where $|u_i| \geqslant l$ for all $0 \leqslant i < p$.*

Proof. Our ultimate goal here is to use Lemma 5 to prove this theorem. To do that, however, we need to take a small detour through some topological arguments (a good reference for the basic results and definitions about the topology of metric spaces that we are using is [24]). We set $A = \{a_0, \ldots, a_{k-1}\}$.

We begin by defining some set $S \subseteq \mathbb{R}^k$. If $s = (s_0, \ldots, s_{k-1}) \in \mathbb{R}^k$, then $s \in S$ if and only if there exists a sequence of non-empty words $v_0, v_1, \ldots$ so that $\lim_{n\to\infty} |v_n| = \infty$, each v_i is a subword of w, and

$$\lim_{n\to\infty} \frac{1}{|v_n|} P(v_n) = \lim_{n\to\infty} \left(\frac{|v_n|_{a_0}}{|v_n|}, \ldots, \frac{|v_n|_{a_{k-1}}}{|v_n|} \right) = s$$

Note that $S \neq \emptyset$. To see this, consider any sequence $v_0, v_1, \ldots$ of non-empty subwords of w such that $|v_n| = n$. Then note that $\frac{1}{|v_n|} P(v_n) \in [0,1]^k$. Since $[0,1]^k$ is compact, there exists $s \in [0,1]^k$ so that s is the limit of some subsequence of $\frac{1}{|v_0|} P(v_0), \frac{1}{|v_1|} P(v_1), \ldots$. Clearly this implies $s \in S$, so $S \neq \emptyset$.

Furthermore, S is closed. To see this, assume that $t_0, t_1, \ldots \in S$ is a sequence with limit $t \in \mathbb{R}^k$. Then by definition of S, for any choice of i we can find a word v_i so that $|v_i| > i$, v_i is a subword of w, and $|t_i - \frac{1}{|v_i|} P(v_i)| < (\frac{1}{2})^i$. Then the sequence $\frac{1}{|v_0|} P(v_0), \frac{1}{|v_1|} P(v_1), \ldots$ approaches t, so it follows that $t \in S$. Thus S is closed by definition. Moreover, since S is a closed and bounded subset of $\mathbb{R}^k$ it follows from the Heine–Borel theorem that S is in fact compact.

If $s \in S$, then $s = (s_0, \ldots, s_{k-1})$. We can then define r_j so that $r_0 = \sup_{s\in S} s_0$, and for $j > 0$:

$$r_j = \sup_{s\in S,\, s_i = r_i \text{ for } i<j} s_j$$

In order for this definition to make sense, we require that for each j there is $s \in S$ so that $s_i = r_i$ for $i < j$. This, however, follows from the fact that S is compact. Similarly, if we let $r = (r_0, \ldots, r_{k-1})$, it is easy to see that since S is compact we have that $r \in S$.

It is worth noting the following fact about r. Assume that $j \geqslant 0$ and $\delta > 0$. Then there exist $N > 0$ and $\alpha_0, \ldots, \alpha_{j-1} > 0$ so that, if v is a subword of w, $|v| > N$, and if for every i, $0 \leqslant i < j$, we get that $|\frac{|v|_{a_i}}{|v|} - r_i| < \alpha_i$, then it follows that $\frac{|v|_{a_j}}{|v|} < r_j + \delta$. If this was not the case, it would be possible to construct a sequence $v_0, v_1, \ldots$ so that each v_i is a non-empty subword of w, $\lim_{n\to\infty} |v_n| = \infty$, and $\lim_{n\to\infty} \frac{1}{|v_n|} P(v_n) = s$, where $s_i = r_i$ for $i < j$, and $s_j \geqslant r_j + \delta > r_j$. However, since $s \in S$, this contradicts with the fact that

$$r_j = \sup_{s\in S,\, s_i = r_i \text{ for } i<j} s_j$$

Thus such N and α_i's must exist for every δ.

Now consider any $\epsilon > 0$ and $l > 0$. We want to choose $\delta > 0$ so that

$$p\delta < \epsilon - \frac{p}{\frac{p}{\epsilon} + 1}$$

Note that such a δ exists due to the fact that the right-hand side of the above equation is larger than 0. Let $\beta = \frac{p}{\frac{p}{\epsilon}+1} + (p-1)\delta$, and note that $\delta + \beta < \epsilon$.

We now define certain numbers $\delta_i > 0$, $\beta_i > 0$, $N_i > 0$ and $\alpha_{i,j} > 0$ for all i, j such that $i > j$. To do this, we begin with $i = k-1$ and go downwards. Set $\delta_{k-1} = \delta$, $\beta_{k-1} = \beta$. Then we can choose $N_{k-1} > \frac{p}{\epsilon}$ and $\alpha_{k-1,0}, \ldots, \alpha_{k-1,k-2} > 0$ so that, if v is a subword of w with $|v| > N_{k-1}$ and $|\frac{|v|_{a_j}}{|v|} - r_j| < \alpha_{k-1,j}$ for $0 \leqslant j < k-1$, then $\frac{|v|_{a_{k-1}}}{|v|} < r_{k-1} + \delta_{k-1}$.

We then proceed inductively, from $i = k-1$ downwards. Assume that δ_j, β_j, N_j, and $\alpha_{j,0}, \ldots, \alpha_{j,j-1}$ are defined for $j > i$. Then we begin by choosing $M_i > N_{i+1}$ so that $\frac{p}{M_i+1} < \alpha_{j,i}$ for all $j > i$. We can then choose $\delta_i > 0$ so that $\delta_i < \delta$, and so that $\frac{p}{M_i+1} + p\delta_i < \alpha_{j,i}$ for all $j > i$. Let $\beta_i = \frac{p}{M_i+1} + (p-1)\delta_i$. Note that $\beta_i \leqslant \beta$. Finally, choose $N_i > M_i$ and $\alpha_{i,0}, \ldots, \alpha_{i,i-1} > 0$ so that, if v is a subword of w with $|v| > N_i$ and $|\frac{|v|_{a_j}}{|v|} - r_j| < \alpha_{i,j}$ for $0 \leqslant j < i$ then $\frac{|v|_{a_i}}{|v|} < r_i + \delta_i$.

Let N be such that $N > \max\{N_0, l\}$. Moreover for all i, let $\gamma_i > 0$ be such that $\beta_i > \frac{p}{N+1} + (p-1)\delta_i + p\gamma_i$ (we can do this since $N > N_i > M_i$). Since $r \in S$, there exists a word v that is a subword of w, $|v| > p(N+1)$ and so that $|\frac{|v|_{a_i}}{|v|} - r_i| < \gamma_i$ for all i. Let $\mu = \lfloor \frac{|v|}{p} \rfloor > N$. Note that we can write $|v| = p\mu + \mu'$ for some $0 \leqslant \mu' < p$. Therefore we can write $v = v_0 \cdots v_{p-1}x$, where $|v_i| = \mu$ and $|x| = \mu' < p$.

We want to show that $|\frac{|v_q|_{a_j}}{|v_q|} - r_j| < \beta_j < \alpha_{t,j}$ for all q, j and t. Assume that $|\frac{|v_q|_{a_i}}{|v_q|} - r_i| < \beta_i < \alpha_{t,i}$ for all q and t when $i < j$, $i < t$. Then this implies that, since $|v_s| > N > N_j$ and $|\frac{|v_s|_{a_i}}{|v_s|} - r_i| < \alpha_{j,i}$ for all $i < j$ and all s, it follows that $\frac{|v_s|_{a_j}}{|v_s|} < r_j + \delta_j$.

Then note that

$$\begin{aligned} r_j - \gamma_j &< \frac{|v|_{a_j}}{|v|} \\ &= \frac{|v_0|_{a_j} + \cdots + |v_{k-1}|_{a_j} + |x|_{a_j}}{|v|} \\ &< \frac{1}{p}\left(\frac{|v_q|_{a_j}}{|v_q|} + (p-1)(r_j + \delta_j)\right) + \frac{p}{|v|} \end{aligned}$$

We can rearrange this to get

$$r_j - \beta_j < r_j - p\gamma_j - \frac{p}{N+1} - (p-1)\delta_j < r_j - p\gamma_j - \frac{p^2}{|v|} - (p-1)\delta_j < \frac{|v_q|_{a_j}}{|v_q|}$$

Therefore $r_j - \beta_j < \frac{|v_q|_{a_j}}{|v_q|} < r_j + \gamma_j$. Since $\beta_j > \gamma_j$, it follows that $|r_j - \frac{|v_q|_{a_j}}{|v_q|}| < \beta_j$, which is just what we want.

Given any i, j, s and t, we get

$$\left|\frac{|v_i|_{a_s} - |v_j|_{a_s}}{|v_t|}\right| = \left|\frac{|v_i|_{a_s}}{|v_i|} - \frac{|v_j|_{a_s}}{|v_j|}\right| < |r_s + \delta_s - r_s + \beta_s| = \delta_s + \beta_s \leqslant \delta + \beta < \epsilon$$

Thus $||v_i|_{a_s} - |v_j|_{a_s}| < \epsilon|v_t|$, and since $|v_i| > l$ for all i, the theorem follows from Lemma 5. □

The next corollary relates back to partial words with infinitely many holes that avoid abelian pth powers. In particular, it says that in such a word the holes cannot be too close together.

Corollary 4. *Let w be a partial word with infinitely many holes over a finite alphabet, and let $p > 1$, $\mu > 0$ be integers. Assume there are fewer than μ letters between each pair of consecutive holes in w. Then w contains an abelian pth power.*

Proof. Let w be a partial word with infinitely many holes over the alphabet $A = \{a_0, \ldots, a_{k-1}\}$, so that there are fewer than μ letters between each pair of consecutive holes in w. We can assume without loss of generality that w begins with a hole, since otherwise we can write $w = w_0 \diamond w_1$, where $\diamond w_1$ meets all the requirements of the corollary, and clearly if $\diamond w_1$ contains an abelian pth power so does w. Note that this implies that every factor of w of length μ contains a hole. Produce the full word v over $A' = \{a_0, \ldots, a_k\}$, $a_k \notin A$, by taking w and replacing all the $\diamond$'s in w with a_k's. Let $m = 16k\mu$ and $l = 2m$. By Theorem 9, there exists a factor $u_0 \cdots u_{p-1}$ of v, so that we can insert holes in $u_0 \cdots u_{p-1}$ to get a partial word $u'_0 \cdots u'_{p-1}$ that is an abelian pth power, where $|u'_i| \geqslant l$ for all i, and so that each pair of consecutive holes is separated by at least m letters. Since each pair of consecutive holes is separated by at least m letters, each u'_i contains at most $\lceil \frac{|u'_i|}{m} \rceil < \frac{|u'_i|}{m} + 1$ holes. Since $u'_0 \cdots u'_{p-1}$ is an abelian pth power, it follows that $||u'_i|_a - |u'_j|_a| < 2(\frac{|u'_i|}{m} + 1)$ for all $a \in A$ and i, j. Moreover, since u_i differs from u'_i only in the positions where holes are inserted, $||u_i|_a - |u'_i|_a| < \frac{|u'_i|}{m} + 1$, for all $a \in A$ and i. Therefore

$$\big||u_i|_a - |u_j|_a\big| \leqslant \big||u_i|_a - \big|u'_i\big|_a\big| + \big|\big|u'_i\big|_a - \big|u'_j\big|_a\big| + \big||u_j|_a - \big|u'_j\big|_a\big| < 4\left(\frac{|u'_i|}{m} + 1\right)$$

It is worth noting that, since $\frac{|u'_i|}{m} > 1$, we have that $\frac{|u'_i|}{m} + 1 < 2\frac{|u'_i|}{m} = 2\frac{|u_i|}{m}$. Thus, if $0 \leqslant i < k$ and $m_i = \max |u_j|_{a_i}$, then $m_i < |u_0|_{a_i} + 4(\frac{|u'_0|}{m} + 1) < |u_0|_{a_i} + 8\frac{|u_0|}{m}$. Then we can produce $v_0, \ldots, v_{p-1}$ by replacing the a_k's in $u_0, \ldots, u_{p-1}$ with $\diamond$'s. Note that $v_0 \cdots v_{p-1}$ is a factor of w. Moreover, if $a_i \in A$ then $|u_j|_{a_i} = |v_j|_{a_i}$. This implies that $m_i = \max |v_j|_{a_i}$. Since at least one of every μ symbols in w is a $\diamond$, v_0 contains at least $\lfloor \frac{|v_0|}{\mu} \rfloor > \frac{|v_0|}{\mu} - 1 > \frac{|v_0|}{2\mu}$ holes, where the last inequality follows since $1 < \frac{|v_0|}{2\mu}$. Let h be the number of holes in v_0, then we get

$$\begin{aligned}
m_0 + \cdots + m_{k-1} &< |u_0|_{a_0} + 4\left(\frac{|u_0|}{m} + 1\right) + \cdots + |u_0|_{a_{k-1}} + 4\left(\frac{|u_0|}{m} + 1\right) \\
&= |v_0|_{a_0} + \cdots + |v_0|_{a_{k-1}} + 4k\left(\frac{|v_0|}{m} + 1\right) \\
&= |v_0| - h + 4k\left(\frac{|v_0|}{m} + 1\right) \\
&< |v_0| - \frac{|v_0|}{2\mu} + 8k\frac{|v_0|}{m} \\
&= |v_0| - \frac{|v_0|}{2\mu} + 8k\frac{|v_0|}{16k\mu} \\
&= |v_0|
\end{aligned}$$

Therefore by Lemma 4 it follows that $v_0 \cdots v_{p-1}$ is an abelian pth power. $\square$

Corollary 4 leads to the question of how close infinitely many holes can be inserted in an abelian p-free word so that the resulting partial word is abelian p-free. We see that they cannot be separated by a constant distance. On the other hand, we know from previous results that there are many cases in which a partial word can be constructed with exponential spacing. We do not know whether it is possible to do this with less separation, perhaps so that the distance between each pair of consecutive holes be bounded by a polynomial. Finally, note that Theorem 9 has the following number theoretic corollary.

Corollary 5. *Let S be a finite set of real numbers. Consider a sequence $s_0, s_1, \dots$ so that $s_i \in S$. Then, for any integers $p > 1$, $l > 0$ and real number $\epsilon > 0$, there exist integers $n > l$ and m so that, if $0 \leqslant t < t' < m$ and $0 \leqslant s < m$ then*

$$\left| \left(\sum_{i=m+t'n}^{m+(t'+1)n-1} s_i \right) - \left(\sum_{i=m+tn}^{m+(t+1)n-1} s_i \right) \right| < \epsilon \left(\sum_{i=m+sn}^{m+(s+1)n-1} s_i \right) \leqslant \epsilon n \max |s_i|$$

Proof. This follows easily by considering the alphabet $A = S$ in Theorem 9. $\square$

We end this section by inserting an arbitrary hole in p-generalized Zimin words, $Z_{n,p}$, which are known to be abelian p-free. They were introduced in [26] in the context of blocking sets.

Let $A = \{a_0, \dots, a_{k-1}\}$ be a k-letter alphabet, and let $p > 1$ be an integer. The p-generalized Zimin words $Z_{n,p}$ are defined by $Z_{1,p} = a_0^{p-1}$ for $n = 1$, and $Z_{n,p} = (Z_{n-1,p}a_{n-1})^{p-1} Z_{n-1,p}$ for $1 < n \leqslant k$.

Theorem 10. *Let $n, p > 1$ be integers. If we insert a hole anywhere in $Z_{n,p}$, the resulting partial word contains an abelian pth power.*

Proof. We proceed by induction on n. First, consider the case $n = 2$. If we replace an a_1 with a $\diamond$ the resulting word contains a subword of the form $\diamond a_0^{p-1}$, which is an abelian pth power. Therefore assume the $\diamond$ replaces an a_0. Then the resulting partial word can be written as

$$\left(a_0^{p-1}a_1\right)^{r_1} a_0^{s_1} \diamond a_0^{s_2} \left(a_1 a_0^{p-1}\right)^{r_2}$$

where $r_1 + r_2 = p - 1$ and $s_1 + s_2 = p - 2$. Then the above partial word has as a factor

$$\begin{gathered} \left(a_0^{p-1-r_1} a_1 a_0^{r_1-1}\right) \cdots \left(a_0^{p-3} a_1 a_0\right)\left(a_0^{p-2} a_1\right) \\ \left(a_0^{s_1} \diamond a_0^{s_2}\right) \\ \left(a_1 a_0^{p-2}\right)\left(a_0 a_1 a_0^{p-3}\right) \cdots \left(a_0^{r_2-1} a_1 a_0^{p-1-r_2}\right) \end{gathered}$$

Note that there are p blocks, each compatible with a permutation of $a_1 a_0^{p-2}$, so we have an abelian pth power. Thus the claim holds when $n = 2$.

Assume the claim holds for n, then consider $Z_{n+1,p} = (Z_{n,p}a_n)^{p-1} Z_{n,p}$. If the hole is inserted in one of the $Z_{n,p}$'s, then by induction it contains an abelian pth power, so $Z_{n+1,p}$ contains an abelian pth power. On the other hand, assume the $\diamond$ replaces an a_n. Then $\diamond Z_{n,p}$ occurs in the resulting word, where $\diamond Z_{n,p} = \diamond(Z_{n-1,p}a_{n-1})^{p-1} Z_{n-1,p} = \diamond Z_{n-1,p}(a_{n-1}Z_{n-1,p})^{p-1}$. This is an abelian pth power. Therefore the claim follows. $\square$

6. Conclusion

In summary, we investigated the question whether there exist infinite abelian p-free partial words over a given alphabet, that is, words in which we can replace the holes with letters from the alphabet in such a way that no p consecutive factors are permutations of each other. In previous work, infinite abelian 2-free partial words with one hole were constructed over a minimal alphabet size of four, while the minimal size needed for more than one hole was shown to be five [6]. In this paper, we gave lower and upper bounds for the number of letters needed to construct infinite abelian p-free partial words with finitely or infinitely many holes, for any $p > 2$. In the case of infinitely many holes, we proved that the minimal alphabet size for 6th or higher powers is two, while for 5th, 4th, and 3rd powers it is at most three, three, and four respectively. We also investigated, in particular, the number of partial words of length n with a fixed number of holes over a binary alphabet that avoid abelian 6th powers and showed that this number grows exponentially with n. In addition, we showed that we cannot avoid abelian pth powers under arbitrary insertion of holes.

More specifically, for the problem of avoiding abelian powers in the infinitely many hole case, for any given $p > 2$, we gave lower and upper bounds on the minimal alphabet size so that there exists a word with infinitely many holes over that upper bounded alphabet which avoids abelian pth powers. The following table provides the power p to be avoided, a lower bound on the minimal alphabet size, and an upper bound, all based on results in this paper for $p > 2$ and results in [6] for $p = 2$:

p	LB	UB
2	5	5
3	3	4
4	2	3
5	2	3
$\geqslant 6$	2	2

How many letters do we need to construct a partial word with infinitely many holes that avoids abelian 3rd powers (resp., 4th powers, 5th powers)? For instance, we have proved that there exists a partial word with infinitely many holes over an alphabet of size four that avoids abelian 3rd powers, and none exists over an alphabet of size two. Whether or not three is the minimal alphabet size remains open. We also need to investigate whether three is the minimal alphabet size for abelian 4th powers and 5th powers.

In the finitely many hole case, for any given $p \geqslant 2$, Part 1 of the following table provides upper and lower bounds on the minimal alphabet size needed to build a (right) infinite partial word with one hole that avoids pth powers:

p	LB-1	UB-1	LB-2	UB-2
2	4	4	5	5
3	3	3	3	4
4	2	2	2	3
$\geqslant 5$	2	2	2	2

Note that in Part 1, the hole is put at the beginning of the word. Part 2 gives data for the construction of a two-sided infinite partial word with one hole that avoids pth powers.

A World Wide Web server interface has been established at

`www.uncg.edu/cmp/research/abelianrepetitions2`

for automated use of a program that given as input an integer p and a partial word w, checks if the given partial word contains any abelian pth powers for $p > 2$. If it does contain an abelian pth power, the program then outputs exactly which abelian pth powers it contains.

References

[1] A. Aberkane, J. Currie, A cyclic binary morphism avoiding abelian fourth powers, Theoret. Comput. Sci. 410 (2009) 44–52.
[2] A. Aberkane, J. Currie, N. Rampersad, The number of ternary words avoiding abelian cubes grows exponentially, J. Integer Seq. 7 (2004), Article 04.2.7, 13 pp.
[3] E. Andreeva, C. Bouillaguet, P.-A. Fouque, J. Hoch, J. Kelsey, A. Shamir, S. Zimmer, Second preimage attacks on dithered hash functions, in: N. Smart (Ed.), Advances in Cryptology, in: Lecture Notes in Comput. Sci., vol. 4965, Springer, Berlin, 2008, pp. 270–288.
[4] J. Berstel, L. Boasson, Partial words and a theorem of Fine and Wilf, Theoret. Comput. Sci. 218 (1999) 135–141.
[5] F. Blanchet-Sadri, Algorithmic Combinatorics on Partial Words, Chapman & Hall/CRC Press, Boca Raton, FL, 2008.
[6] F. Blanchet-Sadri, J.I. Kim, R. Mercaş, W. Severa, S. Simmons, Abelian square-free partial words, in: A.-H. Dediu, H. Fernau, C. Martín-Vide (Eds.), Language and Automata Theory and Applications, in: Lecture Notes in Comput. Sci., vol. 6031, Springer, Berlin, 2010, pp. 94–105.

[7] F. Blanchet-Sadri, R. Mercaş, G. Scott, A generalization of Thue freeness for partial words, Theoret. Comput. Sci. 410 (2009) 793–800.
[8] F. Blanchet-Sadri, S. Simmons, Avoiding abelian powers in partial words, in: G. Mauri, A. Leporati (Eds.), Developments in Language Theory, in: Lecture Notes in Comput. Sci., vol. 6795, Springer, Berlin, 2011, pp. 70–81.
[9] A. Carpi, On the number of abelian square-free words on four letters, Discrete Appl. Math. 81 (1998) 155–167.
[10] A. Carpi, On abelian squares and substitutions, Theoret. Comput. Sci. 218 (1999) 61–81.
[11] R. Cori, M. Formisano, Partially abelian square-free words, Theor. Inform. Appl. 24 (1990) 509–520.
[12] J. Currie, The number of binary words avoiding abelian fourth powers grows exponentially, Theoret. Comput. Sci. 319 (2004) 41–446.
[13] F.M. Dekking, Strongly non-repetitive sequences and progression-free sets, J. Combin. Theory Ser. A 27 (1979) 181–185.
[14] V. Diekert, Research topics in the theory of free partially commutative monoids, Bull. Eur. Assoc. Theor. Comput. Sci. 40 (1990) 479–491.
[15] P. Erdös, Some unsolved problems, Magyar Tud. Akad. Mat. Kut. Int. Közl. 6 (1961) 221–254.
[16] A.A. Evdokimov, Strongly asymmetric sequences generated by a finite number of symbols, Soviet Math. Dokl. 9 (1968) 536–539.
[17] J. Justin, Characterization of the repetitive commutative semigroups, J. Algebra 21 (1972) 87–90.
[18] V. Keränen, Abelian squares are avoidable on 4 letters, in: W. Kuich (Ed.), Automata, Languages and Programming, in: Lecture Notes in Comput. Sci., vol. 623, Springer, Berlin, 1992, pp. 41–52.
[19] V. Keränen, A powerful abelian square-free substitution over 4 letters, Theoret. Comput. Sci. 410 (2009) 3893–3900.
[20] T. Laakso, Musical rendering of an infinite repetition-free string, in: C. Gefwert, P. Orponen, J. Seppänen (Eds.), Logic, Mathematics and the Computer, in: Finnish Artificial Intelligence Society, vol. 14, Hakapaino, Helsinki, 1996, pp. 292–297.
[21] F. Manea, R. Mercaş, Freeness of partial words, Theoret. Comput. Sci. 389 (2007) 265–277.
[22] P.A.B. Pleasants, Non repetitive sequences, Proc. Cambridge Philos. Soc. 68 (1970) 267–274.
[23] R.L. Rivest, Abelian Square-Free Dithering for Iterated Hash Functions, MIT, 2005, http://people.csail.mit.edu/rivest/publications.html.
[24] W. Rudin, Principles of Mathematical Analysis, McGraw–Hill, New York, NY, 1976.
[25] A. Thue, Über unendliche Zeichenreihen, Norske Vid. Selsk. Skr. I, Mat. Nat. Kl. Christiana 7 (1906) 1–22.
[26] A.I. Zimin, Blocking sets of terms, Math. USSR-Sbornik 47 (1984) 353–364.