RAIRO-Theor. Inf. Appl. 48 (2014) 315–339
DOI: 10.1051/ita/2014014

Available online at:
www.rairo-ita.org

ABELIAN PATTERN AVOIDANCE IN PARTIAL WORDS*

F. Blanchet-Sadri[1], Benjamin De Winkle[2] and Sean Simmons[3]

Abstract. Pattern avoidance is an important topic in combinatorics on words which dates back to the beginning of the twentieth century when Thue constructed an infinite word over a ternary alphabet that avoids squares, *i.e.*, a word with no two adjacent identical factors. This result finds applications in various algebraic contexts where more general patterns than squares are considered. On the other hand, Erdős raised the question as to whether there exists an infinite word that avoids abelian squares, *i.e.*, a word with no two adjacent factors being permutations of one another. Although this question was answered affirmately years later, knowledge of abelian pattern avoidance is rather limited. Recently, (abelian) pattern avoidance was initiated in the more general framework of partial words, which allow for undefined positions called holes. In this paper, we show that any pattern p with $n > 3$ distinct variables of length at least 2^n is abelian avoidable by a partial word with infinitely many holes, the bound on the length of p being tight. We complete the classification of all the binary and ternary patterns with respect to non-trivial abelian avoidability, in which no variable can be substituted by only one hole. We also investigate the abelian avoidability indices of the binary and ternary patterns.

Mathematics Subject Classification. 68R15.

Keywords and phrases. Combinatorics on words, partial words, abelian powers, patterns, abelian patterns, avoidable patterns, avoidability index.

* *This material is based upon work supported by the National Science Foundation under Grant Nos. DMS–0754154 and DMS–1060775. The Department of Defense is gratefully acknowledged. Part of this paper was presented at MFCS'12 [3]. We thank the referees of preliminary versions of this paper for their very valuable comments and suggestions.*

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

[2] Department of Mathematics, Pomona College, 610 North College Avenue, Claremont, CA 91711–4411, USA.

[3] Department of Mathematics, Massachusetts Institute of Technology, Building 2, Room 236, 77 Massachusetts Avenue, Cambridge, MA 02139–4307, USA.

Article published by EDP Sciences

© EDP Sciences 2014

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

1. Introduction

Combinatorics on words, or sequences of letters over a finite alphabet, goes back to the work of the mathematician, Thue, at the beginning of the twentieth century [18]. The interest in this topic has been increasing since it finds applications in various research areas of mathematics, computer science, biology, and physics where the data can be easily represented as words over some alphabet. Motivated by molecular biology of nucleic acids, *partial words*, or sequences that may have some "do not know symbols", also called "holes", were introduced by Berstel and Boasson in [1], and have been extensively studied since (see [2] for instance). For example, $cca\diamond ctcg\diamond ccctc$ is a partial word with two holes, represented by the two $\diamond$'s, over the DNA alphabet $\{a, c, g, t\}$ (the $\diamond$'s are *compatible* with, or match, every letter of the alphabet). Partial words are interesting from a theoretical point of view as they approximate full words (those without holes), but also from a practical point of view. Such practical uses occur for instance in bio-inspired computing where they have been considered for identifying good encodings for DNA computations [15]. More specifically, a basic issue in DNA computing is to find strands, which are used as codewords, that should not form so-called secondary structures (a sequence has such a structure if there is some kind of repetition in it). The combinatorial concept of repetition-freeness, or repetition avoidance, has been proposed to exclude such formation. Requiring a big Hamming distance between the strands has also been proposed. So partial words can be used as codewords, the latter requirement being provided through the compatibility relation. They offer a more powerful and realistic model than full words due to the errors caused by the evolutionary processes of deletion, insertion, and mutation in biological sequences. A deeper understanding of avoidance of patterns such as repetitions in the framework of partial words can thus be exploited.

In the context of full words, Thue showed that arbitrarily long non-repetitive words, or those that avoid repetitions, can be constructed with only three letters. The importance of this result has been seen through several rediscoveries that come from applications in ergodic theory, formal language theory, group theory, universal algebra, *etc.* In a recent paper, Currie [8] reviews results on pattern avoidance, and discusses a number of open problems on words avoiding squares and more general patterns which are words over an alphabet of *variables*, denoted by $\alpha, \beta, \gamma, \ldots$ Squares, for instance, are represented by the unary pattern $\alpha^2 = \alpha\alpha$, which is a power of a single variable α.

Erdős in 1961 initiated abelian pattern avoidance by raising the question whether abelian squares can be avoided [13]. A word contains an abelian square if it has a factor uv, where u is a permutation of v or u is a rearrangement of the letters of v. For example, $w = babcacb$ contains the abelian square $abcacb$ even though w is square-free. More generally, if $p = \alpha_0 \ldots \alpha_{n-1}$, where the α_i's are variables, a word w *meets* p *in the abelian sense* if w contains a factor $u_0 \ldots u_{n-1}$ where u_i is a permutation of u_j, whenever $\alpha_i = \alpha_j$; otherwise, w *avoids* p *in the abelian sense*. In 1979, Dekking showed that abelian cubes can be avoided over

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

three letters, and two letters are enough to avoid abelian 4th powers [12]. Thirteen years later, Keränen proved that abelian squares are avoidable over four letters, settling the problem of classifying all the unary patterns in the abelian sense [14]: α is abelian unavoidable, $\alpha\alpha$ is abelian 4-avoidable but not abelian 3-avoidable, $\alpha\alpha\alpha$ is abelian 3-avoidable but not abelian 2-avoidable, and α^n, $n \geq 4$, is abelian 2-avoidable (where a pattern p is *k-avoidable* if there is an infinite abelian p-free word, *i.e.*, not containing any occurrence of p, over a k-letter alphabet). For more results on abelian pattern avoidance, see [9–11].

In [4], the investigation of abelian avoidability in partial words was initiated. A partial word is *abelian square-free* if it does not contain any factor that results in an abelian square after filling in the holes. For example, $abc\diamond cb$ is an abelian square because we can replace the $\diamond$ by letter a and get $abcacb$. Blanchet−Sadri *et al.* constructed a partial word with infinitely many holes over five letters that is abelian square-free (except for trivial squares of the form $a\diamond$ or $\diamond a$, where a is a letter), and proved that none exists over four letters. In [7], Blanchet−Sadri *et al.* looked at the abelian avoidance of the patterns α^n in infinite partial words, where $n > 2$. They investigated, for a given n, the smallest alphabet size needed to construct an infinite partial word with finitely or infinitely many holes that avoids α^n in the abelian sense. They constructed in particular a binary partial word with infinitely many holes that avoids α^n, $n \geq 6$ in the abelian sense. Then, they proved that one cannot avoid α^n in the abelian sense under arbitrary insertion of holes in an infinite full word.

In this paper, we provide a more general study of abelian pattern avoidance in the context of partial words. We construct, given k-abelian avoidable patterns p in full words satisfying some conditions, abelian avoiding partial words with infinitely many holes over alphabet sizes that depend on k, as well as abelian avoiding infinite partial words with h holes over alphabet sizes that depend on k and h, for every integer $h > 0$. We construct, given infinite full words avoiding pattern p in the abelian sense, infinite abelian avoiding full words with the property that any of their positions can be changed to a hole while still avoiding p in the abelian sense. We also show that a pattern p with $n > 3$ distinct variables such that $|p| \geq 2^n$ is abelian avoidable by a partial word with infinitely many holes. The bound on $|p|$ turns out to be tight. We end up completing the classification of the binary and ternary patterns with respect to non-trivial abelian avoidability, in which no variable can be substituted by only one hole. Finally, we investigate the minimal size alphabet for avoiding binary and ternary patterns in the abelian sense.

2. Preliminaries on partial words and patterns

First, we recall some basic concepts of combinatorics on partial words; for more information, see [2]. A *partial word* over a finite alphabet A is a sequence of symbols from $A_\diamond = A \cup \{\diamond\}$, the alphabet A being augmented with the hole symbol $\diamond$; a *full word* is a partial word without holes. The special symbol $\diamond$ is interpreted as an undefined position, or hole, and *matches* or is *compatible* with every letter of A.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

If u, v are partial words of equal length, then u is compatible with v, denoted $u \uparrow v$, if $u[i] = v[i]$ for all i such that $u[i], v[i] \in A$. We use $u \,\updownarrow\!\!\!\!\!\!\phi\, v$ to denote u is *abelian compatible* with v, that is, some permutation of u is compatible with v.

We denote by $u[i]$ the symbol at position i of a partial word u (labelling starts at 0). We denote by $|u|_a$ the number of occurrences of letter a in u. The *length* of u, denoted by $|u|$, represents the number of symbols in u. The *empty word* is the sequence of length zero and is denoted by ε. The set of all full words over A, the free monoid generated by A, is denoted by A^*. Also, A^n denotes the set of all words over A of length n. A partial word u is a *factor* of a partial word v if there exist x, y such that $v = xuy$; the set of factors of v is denoted by $\mathrm{fac}(v)$. We denote by $v[i \dots j]$ (resp., $v[i \dots j)$) the factor $v[i] \dots v[j]$ (resp., $v[i] \dots v[j-1]$). The powers of u are defined by $u^0 = \varepsilon$ and for $n \geq 1$, $u^n = uu^{n-1}$.

Now, let us take a look at some concepts regarding patterns. Let E be a pattern alphabet and let $p = \alpha_0 \dots \alpha_{n-1}$, where $\alpha_i \in E$. Define an *abelian occurrence* of p in a partial word w as a factor $u_0 \dots u_{n-1}$ of w such that there exists a full word $v_0 \dots v_{n-1}$, where for all i, $u_i \neq \varepsilon$ and $u_i \uparrow v_i$, and where for all i, j, if $\alpha_i = \alpha_j$ then v_i is a permutation of v_j.

The partial word w *meets* the pattern p in the abelian sense, or p *occurs* in w in the abelian sense, if for some factorization $w = xuy$, we have that u is an abelian occurrence of p in w; otherwise, w *avoids* p in the abelian sense (or w is *abelian p-free*). For instance, $ab\diamond ba\diamond baa$ meets $\alpha\beta\beta\alpha$ in the abelian sense (take $ab\underline{\diamond b}\ \underline{a}\ \underline{\diamond}\ \underline{ba}a$). An abelian occurrence of p is *trivial* if $u_i = \diamond$ for some i; otherwise, it is *non-trivial*. We call w *non-trivially abelian p-free* if it contains no non-trivial abelian occurrences of p. These definitions also apply to full words over A as well as infinite partial words over A which are functions from $\mathbb{N}$ to $A_\diamond$.

We also make use of the concept of *isomorphic* patterns, introduced in [11]. Let E and F be pattern alphabets, and let $\phi : E \to F$ be a bijection. We can extend ϕ to an isomorphism from E^* to F^*. Then pattern $p \in E^*$ is isomorphic to $\phi(p) \in F^*$. Further, a partial word meets p in the abelian sense if and only if it meets $\phi(p)$ in the abelian sense.

Remark 2.1. Let p be a pattern over alphabet $E = \{\alpha_0, \alpha_1, \dots, \alpha_{\|E\|-1}\}$. Then there exists some pattern, p', over E isomorphic to p and $p'[0] = \alpha_0$.

Based on this, we make the following simplifying assumptions for the rest of the paper (unless otherwise stated): Any pattern p over alphabet $E = \{\alpha_0, \alpha_1, \dots, \alpha_{\|E\|-1}\}$ begins with α_0. Similarly, all binary (resp., ternary) patterns are over the alphabet $\{\alpha, \beta\}$ (resp., $\{\alpha, \beta, \gamma\}$) and begin with α.

A pattern p is *k-abelian avoidable (resp., non-trivially k-abelian avoidable) in partial words* if there is a word with infinitely many holes over an alphabet of size k that avoids p in the abelian sense (resp., avoids non-trivial abelian occurrences of p); otherwise, p is *k-abelian unavoidable*. A pattern which is k-abelian avoidable (resp., k-abelian unavoidable) for some k (resp., every k) is called *abelian avoidable* (resp., *abelian unavoidable*). The *abelian avoidability index* of p is the smallest integer k such that p is k-abelian avoidable, or is ∞ if p is abelian unavoidable.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Note that k-abelian avoidability implies $(k+1)$-abelian avoidability. For clarity or emphasis, we may say a pattern is *completely* abelian avoidable (resp., unavoidable) in partial words if it can (resp., cannot) be avoided both trivially and non-trivially in the abelian sense.

Proposition 2.2. *Let p be an abelian unavoidable pattern in full words. Then every partial word non-trivially meets p in the abelian sense.*

Proof. For the sake of a contradiction, suppose a partial word w over an alphabet A exists that non-trivially avoids p in the abelian sense. Consider the alphabet $B = \{ab \mid a, b \in A_\diamond\}$. Define a full word v over B so that $v[i] = w[2i]w[2i+1]$. Then there is an abelian occurrence of p in v, call it u. However, if $\varphi : B^* \to A_\diamond^*$ maps ab treated as a letter in B to ab treated as a partial word in $A_\diamond^*$, then $\varphi(u)$ is a non-trivial abelian occurrence of p in w. □

3. Avoiding partial words with infinitely many holes

We first address the problem of avoiding patterns in the abelian sense by partial words with infinitely many holes. Before we start trying to avoid patterns, we investigate how far apart holes in abelian pattern avoiding partial words must be and how the hole spacing restricts abelian occurrences of a pattern within an infinite partial word.

We generalize some results from [7] about hole spacing in infinite abelian pattern avoiding partial words. The first proposition essentially states that in an abelian pattern avoiding partial word, the holes cannot be spaced at a constant distance, *i.e.*, they have to be far apart.

Proposition 3.1. *Let w be an infinite partial word with infinitely many holes, let p be a pattern, and let μ be an integer. Assume there are fewer than μ letters between each pair of consecutive holes in w. Then w meets p.*

Proof. By ([7], Cor. 4) w contains arbitrarily high abelian powers. Thus, w meets p. □

The next two lemmas are useful in many proofs later in this paper. They are two different ways to generalize ([7], Lem. 1).

Lemma 3.2. *Let $u = u_0 \dots u_{n-1}$ be a factor of an infinite partial word w, and write $u_j = w[i_j \dots i_{j+1})$ for all j. Suppose $0 < k_0 < k_1 < \dots$ is a sequence of integers such that $k_{j+1} > nk_j$ for all j. Let M be the smallest integer such that $|u_M| = \max\{|u_i| \mid i \in [0 \dots n-1]\}$. Then there is at most one j such that $i_{M+1} \le k_j < i_n$.*

Proof. Assume, towards a contradiction, that there is some j such that $i_{M+1} \le k_j < k_{j+1} < i_n$. Note that $|u_j| \le i_{M+1} - i_0$ for all j. Hence

$$|u| \le n(i_{M+1} - i_0) \le n(k_j - i_0) < k_{j+1} - ni_0 < i_n - i_0 = |u|,$$

which yields a contradiction. □

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

This lemma has the following immediate consequence: If $u = u_0 \dots u_{n-1}$ is an abelian occurrence of a pattern $p = \alpha_0 \dots \alpha_{n-1}$ in an infinite partial word w, k_j and u_j are as above, and M is an integer such that all variables in p occur in $\alpha_0 \dots \alpha_M$, then there is at most one j such that $i_{M+1} \leq k_j < i_n$.

Lemma 3.3. *Let $u = u_0 u_1 \dots u_{n-1}$ be a factor of an infinite partial word w, and write $u_j = w[i_j \dots i_{j+1})$ for all j. Suppose $0 < k_0 < k_1 < \dots$ is a sequence of integers such that $k_{j+1} > 2k_j$ for all j, and let $l < l'$ be integers such that $|u_l| = |u_{l'}|$. Then there is at most one j such that $i_{l'} \leq k_j < i_{l'+1}$.*

Proof. Assume, towards a contradiction, that there exists some j such that $i_{l'} \leq k_j < k_{j+1} < i_{l'+1}$. Note $i_{l+1} \leq i_{l'}$, so

$$|u_l| = i_{l+1} - i_l \leq i_{l+1} \leq i_{l'} \leq k_j < k_{j+1} - k_j < i_{l'+1} - i_{l'} = |u_{l'}| = |u_l|,$$

which is a contradiction. □

Both of these lemmas allow us to limit the number of holes in an abelian occurrence of a pattern within an infinite partial word if we space the holes exponentially far apart.

3.1. Binary patterns

Next we consider binary patterns. The following lemma is very simple, but useful in several upcoming proofs.

Lemma 3.4. *If p is a pattern which contains an abelian cube, then p is 4-abelian avoidable in partial words.*

Proof. There exists an infinite partial word, w, with infinitely many holes over four letters that avoids abelian cubes [7]. Hence w avoids p. □

Using this lemma, we can classify all binary patterns with respect to abelian avoidability in partial words (both trivial and non-trivial), and give upper limits on the abelian avoidability index.

Theorem 3.5. *Let p be a binary pattern. Then the following enumerate all possibilities for p:*

1. *If p is not isomorphic to a subpattern of $\alpha\alpha\beta\alpha\alpha\alpha$, then p is completely 5-abelian avoidable in partial words.*
2. *If $p \in S_0 = \{\alpha\alpha, \alpha\alpha\beta, \alpha\beta\beta, \alpha\alpha\beta\alpha, \alpha\beta\alpha\alpha, \alpha\alpha\beta\alpha\alpha\}$, then p is trivially abelian unavoidable, but has non-trivial abelian avoidability index 5.*
3. *if $p \in S_1 = \{\alpha, \alpha\beta, \alpha\beta\alpha\}$, then p is completely abelian unavoidable in partial words.*

Furthermore, any trivially or non-trivially abelian avoidable binary pattern in partial words can be avoided with only five letters.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Proof. We first prove Statement 1. If $|p| \geq 10$, then p contains an abelian cube and the result follows from Lemma 3.4. Hence we may assume $|p| < 10$, and by hypothesis p is not isomorphic to a subpattern of $\alpha\alpha\beta\alpha\alpha$. Suppose p contains either an abelian cube or an abelian square, u_0u_1, where $|u_0| = |u_1| \geq 2$. In the former case, Lemma 3.4 again proves the result. In the latter case, consider the infinite partial word over five letters, w, from [4] which has infinitely many holes and avoids non-trivial abelian squares. If w meets p, it must contain a non-trivial abelian square. Hence w avoids p. We can check that the only pattern not handled above is $p = \alpha\alpha\beta\beta$; however, we can show w avoids this pattern as well. Assume, by way of contradiction, that $v = v_0v_1v_2v_3$ is an abelian occurrence of p in w. Since there are no non-trivial occurrences of abelian squares in w, it must be that either $v_0 = \diamond$ or $v_1 = \diamond$. Similarly, either $v_2 = \diamond$ or $v_3 = \diamond$. However, this implies $v_0v_1 \not\Phi v_2v_3$, so v is a non-trivial abelian square, which contradicts our choice of w. Thus w avoids p.

Next we prove Statement 2. Let w be a partial word with infinitely many holes over A. Then there exists some $a \in A_\diamond$ so that $a\diamond$ occurs infinitely often in w. Thus there is a non-empty partial word u so that $a\diamond ua\diamond$ is a factor of w. However, $a\diamond ua\diamond$ is an occurrence of $\alpha\alpha\beta\alpha\alpha$, so there is no partial word with infinitely many holes that avoids $\alpha\alpha\beta\alpha\alpha$ in the abelian sense. Notice that all the patterns in S_0 contain a square, so the 5-letter partial word from above non-trivially avoids these, showing the non-trivial abelian avoidability index is less than or equal to 5. Now let w' be a partial word on four letters with infinitely many holes. In [4] it is stated that any partial word of the form $u\diamond v$ where $|u| = |v| = 12$ contains a non-trivial abelian square. Note there is some factor, x, of w' of this form that occurs infinitely often. By considering two occurrences of x, we can find a factor of w' which we can write as $u_0u_1vu_0u_1$, where u_0u_1 is a non-trivial abelian square in x and v is a non-empty factor of w'. Hence w' meets the pattern $\alpha\alpha\beta\alpha\alpha$. Since all patterns in S_0 are subpatterns of this pattern (up to isomorphism), w' meets every pattern in S_0. Thus every pattern in S_0 is non-trivially abelian 4-unavoidable in partial words.

Finally we prove Statement 3. The pattern $\alpha\beta\alpha$ is isomorphic to the second Zimin pattern, Z_2, which is unavoidable in full words (see Chapter 3 of [16]), hence completely abelian unavoidable in partial words. □

However, five is only an upper bound on the abelian avoidability index of any trivially or non-trivially abelian avoidable binary pattern in partial words, and we wish to classify which patterns have lower indices. Since it is shown in [7] that there exists a partial word over a ternary alphabet with infinitely many holes that is abelian α^4-free, all binary patterns meeting abelian 4th powers are abelian 3-avoidable. Meanwhile, Currie *et al.* in [10] give the first example of a binary pattern which is abelian 2-avoidable, but which contains no abelian 4th-powers. We can show that this pattern is also 3-abelian avoidable in the context of partial words.

Proposition 3.6. *There exists a partial word with infinitely many holes over a ternary alphabet that avoids* $\alpha\alpha\alpha\beta\alpha\alpha\alpha\beta\beta\beta\alpha\alpha\alpha$ *in the abelian sense.*

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Proof. Let $p = \alpha\alpha\alpha\beta\alpha\alpha\alpha\beta\beta\beta\alpha\alpha\alpha$. Currie *et al.* define in [10] a family of morphisms $f_n : \{a,b\}^* \to \{a,b\}^*$ generated by $f_n(a) = a^{n+1}b$, $f_n(b) = ab^n$. Then word w_n is defined by $w_n = \lim_{m\to\infty} f_n^m(a)$. They have shown that p can be avoided in the abelian sense by w_2. Note that Dekking [12] found (essentially) w_2.

Let c be any letter not in $A = \{a,b\}$ and write $B = \{a,b,c\}$. Define a sequence $1 < k_0 < k_1 < \ldots$ such that $k_{i+1} > 13k_i = |p|k_i$. Then we can define an infinite partial word w' over B as follows:

$$w'[j] = \begin{cases} \diamond, & \text{if } j = k_i \text{ for some } i; \\ c, & \text{if } j = k_i + 1 \text{ or } j = k_i - 1 \text{ for some } i; \\ w_2[j], & \text{otherwise.} \end{cases}$$

We show that w' avoids p in the abelian sense by contradiction. Assume there exists an abelian occurrence of p in w', denoted as

$$u = u_0u_1u_2v_0u_3u_4u_5v_1v_2v_3u_6u_7u_8,$$

where each u_i is abelian compatible with each u_j and each v_i is abelian compatible with each v_j.

By Lemma 3.2, $u_3u_4u_5v_1v_2v_3u_6u_7u_8$ contains at most one hole. This implies some u_i for $i \in [3\ldots8]$ does not contain a c or a $\diamond$. Consequently no u_i can contain a c, and since all $\diamond$'s come in factors of the form $c\diamond c$, no u_i contains a $\diamond$. If $|v_0| \geq 2$, then since all v_i have the same length, some v_i for $i \in [1\ldots3]$ does not contain a c or a $\diamond$. By the same argument as above, no v_i contains a c or a $\diamond$, so u is an abelian occurrence of p in w, which is a contradiction. So $|v_0| = 1$, which implies v_0 does not have a c or a $\diamond$ (else some u_i would have a c), again this implies that no v_i has a c or a $\diamond$, which is a contradiction. □

We also give the following bounds.

Proposition 3.7. *Let p be a binary pattern. If $|p| \geq 363$ (resp., $|p| \geq 119$), then a partial word over a binary (resp., ternary) alphabet with infinitely many holes exists that is abelian p-free.*

Proof. Let $p = \alpha_0 \ldots \alpha_{n-1}$ be a pattern over $E = \{\alpha,\beta\}$ such that $|p| \geq 363$. Without loss of generality we can assume that $\alpha_0 = \alpha$. Note that $\alpha_i = \beta$ for some $i < 6$, since otherwise p contains α^6 as a subpattern, which can be abelian avoided by a binary partial word with infinitely many holes [7]. Then write $p = qq_0q_1q_2$ where $|q_i| = 119$ for each i, and $|q| \geq 6$. We can do this since $|p| \geq 363 = 3\times119+6$. From [11], Lemma 3.2, there exist infinite binary words w_2 and w_3 so that every binary pattern of length at least 119 is avoided in the abelian sense by either w_2 or w_3. By the pigeonhole principle there exist $l < l'$, where q_l and $q_{l'}$ are either both avoided in the abelian sense by w_2 or both avoided in the abelian sense by w_3. Let w be the infinite binary word that avoids q_l and $q_{l'}$ in the abelian sense. Define a sequence $0 < k_0 < k_1 < k_2 < \ldots$ so that $k_{i+1} > |p|k_i$ for all i. We can then define the partial word w' by $w'[j] = \diamond$ if $j = k_i$ for some i, and $w'[j] = w[j]$ otherwise.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

For the sake of a contradiction, suppose that $uu_0u_1u_2$ is an abelian occurrence of p in w', where u is an abelian occurrence of q, u_0 is an abelian occurrence of q_0, *etc.* By Lemma 3.2, there is at most one hole in $u_0u_1u_2$. Thus, at least one of u_l and $u_{l'}$ is full, and is an occurrence of q_l or $q_{l'}$ in w, which contradicts our choice of w.

Now, let $p = \alpha_0 \dots \alpha_{n-1}$ be a pattern over $E = \{\alpha, \beta\}$ such that $|p| \geq 119$. As above, we can assume that $\alpha_0 = \alpha$. Note that $\alpha_i = \beta$ for some $i < 6$, since otherwise p contains α^6 as a subpattern, which can be avoided by a binary partial word with infinitely many holes in the abelian sense. From [11], Lemma 3.2, there exists an infinite word w over $A = \{a, b\}$ that avoids abelian occurrences of p. Moreover, bab occurs infinitely often as a factor of w (this can be seen by considering the words w_2 and w_3 in [11]). Define a sequence $3 < k_0 < k_1 < k_2 < \dots$ so that $|p|k_i < k_{i+1}$ and $w[k_i - 1 \dots k_i + 1] = bab$, for all i. Consider $c \notin A$. We can then define a partial word w' as follows: $w'[j] = \diamond$ if $j = k_i$ for some i, $w'[j] = c$ if $j = k_i - 1$ or $j = k_i + 1$ for some i, and $w'[j] = w[j]$ otherwise.

For the sake of a contradiction, suppose that $u_0 \dots u_{n-1}$ is an abelian occurrence of p in w'. Since w avoids abelian occurrences of p, some u_j must contain a hole. However, since each $\diamond$ in w' occurs in a factor of the form $c\diamond c$, some u_i must contain c. Assume that $\alpha_i = \beta$, the other case being similar. By Lemma 3.2, there is at most one j so that $j \geq 6$ and so that u_j contains a hole. Then one variable in $\alpha_7 \dots \alpha_{13}$ must be β. Let α_l be this variable. Then u_l contains either a c or a $\diamond$. Since each c occurs next to a $\diamond$, $u_{l-1}u_lu_{l+1}$ contains a hole, where $l - 1 \geq 6$ and $l+1 \leq 14$. Similarly, there exists an l' so that $23 \geq l' \geq 17$, and where $u_{l'-1}u_{l'}u_{l'+1}$ contains a hole. However, this contradicts the fact that there exists at most one j, $j \geq 6$, such that u_j contains a hole. Thus w' avoids p in the abelian sense. □

3.2. General patterns

Now we move on to general patterns.

Our next goal is to find a bound so that if a pattern has length at least this bound, then it is abelian avoidable in partial words. The following proposition helps us in this.

Proposition 3.8. *Let p be a pattern over an alphabet E such that $p \neq \alpha\alpha$, for any $\alpha \in E$. If each variable in p occurs at least twice, then p can be avoided in the abelian sense by a partial word with infinitely many holes. In particular, if p is abelian k-avoidable, then there exists a partial word with infinitely many holes over an alphabet of size $k + 4$ that is abelian p-free.*

Proof. Since each variable in $p = \alpha_0 \dots \alpha_{m-1}$ occurs at least twice, by [9], Lemma 7, p can be avoided in the abelian sense by an infinite full word w over some alphabet A of size k. Let $B = A \cup \{a, b, c, d\}$ where $a, b, c, d \notin A$. Since there is some factor v, $|v| = 5$, that occurs infinitely often in w, consider a sequence $5 < k_0 < k_1 < k_2 < \dots$ so that $w[k_i - 2 \dots k_i + 2] = v$ and $2k_i < k_{i+1}$, for all i. Then define w' by

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

$$w'[j] = \begin{cases} a, & \text{if } j = k_i - 2 \text{ for some } i,\ i \equiv 0 \bmod 4|p|; \\ b, & \text{if } j = k_i - 1 \text{ for some } i,\ i \equiv 0 \bmod 4|p|; \\ \diamond, & \text{if } j = k_i \quad \text{for some } i,\ i \equiv 0 \bmod 4|p|; \\ c, & \text{if } j = k_i + 1 \text{ for some } i,\ i \equiv 0 \bmod 4|p|; \\ a, & \text{if } j = k_i + 2 \text{ for some } i,\ i \equiv 0 \bmod 4|p|; \\ d, & \text{if } j = k_i \quad \text{for some } i,\ i \not\equiv 0 \bmod 4|p|; \\ w[j], & \text{otherwise.} \end{cases}$$

For the sake of contradiction, suppose w' contains $u = u_0u_1 \dots u_{m-1}$ as an abelian occurrence of p and write $u_j = w'[i_j \dots i_{j+1})$. Note that $u_0u_1 \dots u_{m-1}$ contains a hole since otherwise we can replace the b's in u with $v[1]$'s *etc.*, to get a word that is still an abelian occurrence of p, but that is also a factor of w, which is impossible. Moreover, u contains at least $4|p| - 1$ d's. To see this, assume that u_j contains a hole. Then there are two cases to consider. If $u_j \neq \diamond$, then u_j contains b or c, since every $\diamond$ in w' occurs in a factor of the form $ab\diamond ca$. Without loss of generality, assume u_j has $b\diamond$ as a factor. Since every variable in p occurs at least twice, there is some j', distinct from j, so that $u_{j'}$ is compatible with a permutation of u_j. Thus $u_{j'}$ contains either b or $\diamond$. Suppose $j' > j$, the case $j' < j$ being similar. Then there exists some s such that $i_j \leq k_s < i_{j+1}$, and some t such that $i_{j'} < k_t \leq i_{j'+1}$ or $i_{j'} \leq k_t < i_{j'+1}$, where $s < t$, $s \equiv t \equiv 0 \bmod 4|p|$. Thus $i_0 \leq k_s < k_{s+4|p|} \leq k_t \leq i_m$. So there are $4|p| - 1$ l's so that $k_s < k_l < k_{s+4|p|}$. Since for each $l \not\equiv 0 \bmod 4|p|$, $w'(k_l) = d$, there are $4|p| - 1$ d's in u. If $u_j = \diamond$, then a similar reasoning works.

Since u contains at least $4|p| - 1 = 4m - 1$ d's, it follows from the pigeonhole principle that some u_j contains at least two d's. So there is an s such that $i_j \leq k_s < k_{s+1} < i_{j+1}$. Since there is a j' such that $u_{j'}$ is compatible with a permutation of u_j, let $j' > j$ (the other case is similar). Here, $u_{j'}$ contains either a d or a $\diamond$ for each d in u_j. This implies there is some t such that $i_{j'} \leq k_t < k_{t+1} < i_{j'+1}$, where $t > s$, which contradicts Lemma 3.3. □

Now we consider one bound that implies abelian avoidability in partial words. We will later improve upon it, but we need it in order to prove our tight bound, so we include it here.

Theorem 3.9. *Let p be a pattern with n distinct variables. If $|p| \geq 3 \times 2^{n-1}$, then there exists a partial word with infinitely many holes that is abelian p-free.*

Proof. We proceed by induction on n. If $n = 1$, then $p = \alpha^m$ for some $\alpha \in E$ and $m \geq 3$, whence p is abelian avoidable in partial words [7]. Assume the claim holds up to $n - 1$ distinct variables. If each variable in p occurs at least twice, then the result follows from Proposition 3.8. So we may assume there is some variable that occurs in p exactly once. Then p has a factor q with at most $n-1$ distinct variables with $|q| \geq 3 \times 2^{n-2}$, so by the inductive hypothesis a word with infinitely many holes exists that avoids q, and thus p, in the abelian sense. □

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

By considering the pattern $p = \alpha\alpha$, we can see that the bound in Theorem 3.9 is tight over a unary alphabet of variables. Moreover, Theorem 3.5 shows that it is tight over a binary alphabet of variables. Our next step is to strengthen Theorem 3.9 for pattern alphabets of at least three variables.

Theorem 3.10. *Let p be a pattern over an alphabet E. Then there exists a partial word with infinitely many holes that is abelian p-free if one of the following holds: (1) $\|E\| = 3$ and $|p| \geq 9$; (2) $\|E\| > 3$ and $|p| \geq 2^{\|E\|}$.*

Proof. For Statement 1, let $p = \alpha_0 \dots \alpha_{m-1} \in E^m$ be a pattern over the alphabet $E = \{\alpha, \beta, \gamma\}$. It is sufficient to consider the case where $|p| = 9$. For the sake of a contradiction, suppose that there is no partial word with infinitely many holes that avoids p in the abelian sense. At least one variable in p occurs exactly once (the case when each variable in p occurs at least twice follows from Prop. 3.8). Without loss of generality, we can assume that this variable is γ. Therefore we can write $q_0\gamma q_1 = p$, where q_0 and q_1 are patterns over $\{\alpha, \beta\}$. There exists no partial word with infinitely many holes that avoids q_0 or q_1 in the abelian sense, because that word should also avoid p in the abelian sense. By Theorem 3.9, $|q_i| < 6$ for both i. Since $|q_0| + |q_1| = 9 - 1 = 8$, $(|q_0|, |q_1|) \in \{(5,3), (3,5), (4,4)\}$.

There exists an infinite word w over an alphabet A that avoids abelian squares. Assume that $a_0a_1a_2 \in A^3$ occurs infinitely often in w, then consider any sequence $4 < k_0 < k_1 < \dots$ where $k_{i+1} > 9k_i$ and $w[k_i - 1 \dots k_i + 1] = a_0a_1a_2$, for all i. Moreover, consider $a, b, c \notin A$, where a, b and c are distinct letters. Let $B = A \cup \{a, b, c\}$. We can then define a partial word w' over B as follows:

$$w'(j) = \begin{cases} a, & \text{if } j = k_i - 1 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ \diamond, & \text{if } j = k_i \quad\ \text{for some } i,\ i \equiv 0 \bmod 3; \\ b, & \text{if } j = k_i + 1 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ c, & \text{if } j = k_i \quad\ \text{for some } i,\ i \not\equiv 0 \bmod 3; \\ w(j), & \text{otherwise.} \end{cases}$$

The partial word w' avoids non-trivial abelian squares, following the same argument as in the proof of ([4], Thm. 4.4). We want to show that w' avoids p in the abelian sense. For the sake of a contradiction, suppose that $u_0 \dots u_8$ is an abelian occurrence of p. Set $u_j = w'[i_j \dots i_{j+1})$ for each j.

First, consider the case $|q_0| = 3$, $|q_1| = 5$. Since $|q_1| > 4$ we know that q_1 is avoidable in full words. By Theorem 3.5, $q_1 = \alpha\alpha\beta\alpha\alpha$ or $q_1 = \beta\beta\alpha\beta\beta$. Without loss of generality we assume that $q_1 = \alpha\alpha\beta\alpha\alpha$. Then $u_4u_5u_6u_7u_8$ is an abelian occurrence of q_1. Since u_4u_5 and u_7u_8 are both abelian squares and w' avoids non-trivial abelian squares, we get that u_4u_5 and u_7u_8 both contain a hole. Thus there is some s such that $i_4 \leq k_s < k_{s+1} < i_9$. Moreover, β occurs in q_0, since otherwise $q_0 = \alpha\alpha\alpha$, so this contradicts Lemma 3.2.

Next, consider the case where $|q_0| \geq 4$. This implies that q_0 is a subpattern of $\alpha\alpha\beta\alpha\alpha$. Suppose that $q_0 = \alpha\alpha\beta\alpha$, the other cases being similar. Since w' avoids non-trivial abelian squares, we know $|u_0| = |u_1| = |u_3| = 1$. Also, u_0 and u_1 must

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

be elements of $\{a, b, \diamond\}$, which implies u_3 must also be an element of that set. So there are some integers, s and t, such that $i_0 \leq k_s < i_3 \leq k_{s+3t} \leq i_4 + 1$. Since $i_2 = i_0 + 2$ and $i_4 = i_3 + 1$, it must be that $i_2 \leq k_{s+1} < k_{s+2} < i_3$. This implies u_2 contains at least two c's. Since $|q_1| \geq 3$, there is at least one β in q_1. Thus, there is some $j \geq 5$ such that u_j contains two letters in $\{c, \diamond\}$, whence there is some s' such that $i_j \leq k_{s'} < k_{s'+1} < i_{j+1}$. This contradicts Lemma 3.3.

For Statement 2, let p be a pattern over the alphabet $E = \{\alpha, \beta, \gamma, \delta\}$. It suffices to consider the case $|p| = 16$. For the sake of a contradiction, suppose that no partial word with infinitely many holes avoids p in the abelian sense. We know that p can be avoided in the abelian sense by an infinite full word. Then there is some variable in p that occurs exactly once, assume it is δ. We can write $p = q_1 \delta q_0$, where each q_i is a pattern over $\{\alpha, \beta, \gamma\}$ that cannot be avoided by a partial word with infinitely many holes in the abelian sense. Furthermore, note that this implies $|q_i| < 9$. Since $|q_0| + |q_1| = 16 - 1 = 15$, either $|q_0| = 7$ and $|q_1| = 8$ or *vice versa*. This means that each of α, β and γ actually appear in each q_i at least once, since otherwise q_i would be a pattern over a binary alphabet of length greater than 6, and by Theorem 3.9 there would exist a partial word with infinitely many holes that avoids it in the abelian sense.

There exists an infinite word w over an alphabet A that avoids abelian squares. Assume that $a_0a_1a_2 \in A^3$ occurs infinitely often in w. Letting $a, b, c \notin A$, construct a sequence $3 < k_0 < k_1 < \ldots$ so that $k_{i+1} > 2k_i$ and $w[k_i - 1 \ldots k_i + 1] = a_0a_1a_2$, for all i, and define w' as in Statement 1. We already know that w' avoids non-trivial abelian squares. For the sake of a contradiction, suppose w' does not avoid p in the abelian sense, and let $u_0 \ldots u_{15}$ be an abelian occurrence of p. Consider the case where $|q_1| = 8$ and $|q_0| = 7$ (the argument is similar when we consider the case where $|q_1| = 7$ and $|q_0| = 8$). Some variable must occur in q_1 exactly once, since otherwise q_1 could be avoided by a partial word with infinitely many holes by Proposition 3.8. Let us assume this variable is γ. Then we can write $q_1 = q_0'\gamma q_1'$. Since $|q_1| = 8$ we have that $4 \leq |q_i'|$ for some i, and q_i' must be a subpattern of $\alpha\alpha\beta\alpha\alpha$ by Theorem 3.5. Since there is a β in q_1, this leads to a contradiction, similar to above.

Now, let p be a pattern over an alphabet E. Let $n = \|E\| > 4$ and $|p| \geq 2^n$. We will show that p is abelian avoidable in partial words by induction on n, using $n = 4$ as the base case. By Proposition 3.8 we may assume some variable, α, occurs only once in p. Then we can write $p = q_0\alpha q_1$, where q_0 and q_1 have at most $n - 1$ distinct variables and some q_i has length at least 2^{n-1}. Then q_i is abelian avoidable in partial words by our inductive hypothesis, so p is also abelian avoidable in partial words. □

The bounds in Theorem 3.10 are tight. For (1), $\alpha\alpha\beta\alpha\alpha\gamma\alpha\alpha$ cannot be avoided by any partial word with infinitely many holes (since there must be some $a \in A_\diamond$ such that $a\diamond$ occurs infinitely often). For (2), the nth Zimin pattern, Z_n, is abelian unavoidable and is of length $2^n - 1$ (see Chapt. 3 of [16]). More specifically, if $E = \{\alpha_0, \ldots, \alpha_{n-1}\}$, then $Z_1 = \alpha_0$, and $Z_m = Z_{m-1}\alpha_{m-1}Z_{m-1}$ if $1 < m \leq n$. Meanwhile, we can consider the case of non-trivial avoidance.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Theorem 3.11. *Let p be a pattern with each of its variables occurring at least twice that is abelian k-avoidable in full words. Then a partial word with infinitely many holes over an alphabet of size $k+2$ exists that is non-trivially abelian p-free.*

Proof. Suppose w is a word over an alphabet A of size k that avoids p in the abelian sense. Then we define a sequence of k_i's, where $k_{i+1} > 5k_i$ such that $w[k_i] = a$ for some $a \in A$. Let $B = A \cup \{c, d\}$, where c, $d \notin A$. Then define a partial word, w', as follows:

$$w'[i] = \begin{cases} \diamond, & \text{if } i = k_j \qquad\qquad\qquad\quad \text{for some } j,\ j \equiv 0 \bmod 6|p|; \\ c, & \text{if } i = k_j - 1 \text{ or } i = k_j + 1 \text{ for some } j,\ j \equiv 0 \bmod 6|p|; \\ d, & \text{if } i = k_j \qquad\qquad\qquad\quad \text{for some } j,\ j \not\equiv 0 \bmod 6|p|; \\ w[i], & \text{otherwise.} \end{cases}$$

Then w' contains infinitely many holes. Set $p = \alpha_0 \dots \alpha_{n-1} \in E^n$, and assume that $u = u_0 \dots u_{n-1} = w'[i_0 \dots i_1) \dots w'[i_{n-1} \dots i_n)$, where $u_0 = w'[i_0 \dots i_1), \dots, u_{n-1} = w'[i_{n-1} \dots i_n)$, is a non-trivial abelian occurrence of p in w', *i.e.*, if $\alpha_i = \alpha_j$ then u_i is compatible with a permutation of u_j, and $u_i \neq \diamond$, for $0 \leq i < n$.

Let α_l and $\alpha_{l'}$ be two occurrences of any variable α in p, where $l < l'$. Let $J_1 = \{j \mid i_l \leq k_j < i_{l+1}\}$ and $J_2 = \{j \mid i_{l'} \leq k_j < i_{l'+1}\}$. Then $\|J_2\| < 2$ by Lemma 3.3. If $\|J_1\| \geq 3$ then there are at least 2 d's in u_l, which means that $u_{l'}$ has to contain at least 2 d's or $\diamond$'s. Then J_2 contains at least 2 j's, which is a contradiction. Therefore $\|J_1\| < 3$. Since for each α in p, $\|J_1 \cup J_2\| \leq 3$, there are at most $3|p|$ integer j's such that $i_0 \leq k_j \leq i_n$.

Now we show that no u_l contains a hole. Suppose some u_l contains one. Then c must also occur in u_l since $|u_l| > 1$. Then either $\diamond c$ or $c\diamond$ occurs in $u_{l'}$, which is compatible with a permutation of u_l. Suppose $\diamond c$ occurs in $u_{l'}$, the other case is similar. Then between the occurrence of $\diamond$ in u_l and $\diamond$ in $u_{l'}$ there are at least $6|p|-2$ k_j's. This contradicts the fact that there are at most $3|p|$ k_j's from position i_0 to position i_n. Therefore u contains no holes, and similarly u contains no c's. Then we can replace all the occurrences of d by a, and we get an abelian occurrence of p in w. However, this contradicts the fact that p can be avoided in the abelian sense by infinite full words over k letters. Therefore w' non-trivially avoids p in the abelian sense. □

Corollary 3.12. *Let p be a pattern with n distinct variables. If $|p| \geq 2^n$, then there exists a finite alphabet A and a partial word with infinitely many holes over A that is non-trivially abelian p-free.*

Proof. We proceed by induction on n. If $n = 1$, then p can be written as $p = \alpha^{|p|}$ for some $\alpha \in E$. However, since $|p| \geq 2^n = 2$, p can be non-trivially avoided by a partial word with infinitely many holes in the abelian sense. Assume the claim holds up to $n-1$ distinct variables. Let p be a pattern with n distinct variables, $|p| \geq 2^n$. By ([9], Lem. 7), p can be avoided in the abelian sense by an infinite full word. If every variable in p occurs at least twice, then the result follows by

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Theorem 3.11. On the other hand, assume that there is a variable $\alpha \in E$ that occurs in p, but occurs only once. This implies that p contains a factor q that does not contain α, $|q| \geq 2^{n-1}$. Since q does not contain α, q has at most $n-1$ distinct variables, so by the inductive hypothesis there is a partial word with infinitely many holes that avoids q non-trivially in the abelian sense, and thus avoids p non-trivially in the abelian sense. □

We can strengthen Theorem 3.11 and Proposition 3.8, by adding another hypothesis.

Theorem 3.13. *Let $p = \alpha_0 \dots \alpha_{n-1}$ be an abelian k-avoidable pattern in full words such that for some m, all of the variables in p are present in both $\alpha_0 \dots \alpha_{m-1}$ and $\alpha_m \dots \alpha_{n-1}$. If $\alpha_{m-1} \neq \alpha_m$ or there exists some $j \notin \{m-1, m\}$ such that $\alpha_j = \alpha_m$, then p is abelian $(k+2)$-avoidable in partial words.*

Proof. By assumption, there exists some infinite word, w, over a k-letter alphabet, A, which avoids p in the abelian sense. Let $a \in A$ be a letter that occurs infinitely often in w. Define a sequence, $1 < k_0 < k_1 < \dots$, such that $k_i > (n+1)k_{i-1}$ for all i and $w[k_i] = a$. Let $b, c \notin A$ be letters, let $\mu = 4m+4$, and define w' as follows:

- If $i \equiv 0 \bmod \mu$, then $w'[k_i - 1 \dots k_i + 1] = b \diamond c$,
- If $i \equiv 1, 2, \dots, 2m+1 \bmod \mu$, then $w'[k_i] = b$,
- If $i \equiv 2m+2 \bmod \mu$, then $w'[k_i - 1 \dots k_i + 1] = c \diamond b$,
- If $i \equiv 2m+3, 2m+4, \dots 4m+3 \bmod \mu$, then $w'[k_i] = c$,
- unless specified above, $w'[j] = w[j]$.

Note that we are creating w' from w by replacing letters.

Assume, by way of contradiction, that $u = u_0 \dots u_{n-1}$ is an abelian occurrence of p in w'. Let $v_0 = u_0 \dots u_{m-1}$ and $v_1 = u_m \dots u_{n-1}$ and write $v_0 = w'[i_0 \dots i_1)$ and $v_1 = w'[i_1 \dots i_2)$. Notice that for each u_j in v_0, there is some $u_{j'}$ in v_1 such that $u_j \updownarrow u_{j'}$. We claim that there exists some k_i such that $i_0 \leq k_{i-1} < i_1 \leq k_i < i_2 < k_{i+1}$.

Assume there is at most one k_i such that $i_0 \leq k_i < i_2$. There must be at least one b or c in u, otherwise $u \in \mathrm{fac}(w)$, which contradicts our choice of w. Hence there is some k_i, such that $k_{i-1} \leq i_0 - 1 \leq k_i \leq i_2 \leq k_{i+1}$ and $i \equiv 0 \bmod \frac{\mu}{2}$, or $k_{i-1} < i_0 \leq k_i < i_2 \leq k_{i+1}$ and $i \not\equiv 0 \bmod \frac{\mu}{2}$. In the latter case, u is full and exactly one u_j contains a b or a c. However, there is some $u_{j'}$ that is abelian compatible with u_j, but $u_{j'}$ contains only letters in A, a contradiction. In the former case, if $k_i + 1 < i_1$ or $k_i - 1 \geq i_1$, then we arrive at a similar contradiction. So k_i must be either $i_1 - 1$ or i_1. Suppose $k_i = i_1 - 1$ and $i \equiv 0 \bmod \mu$ (the other cases being similar). Then some u_j in v_0 contains a b, but there is no $u_{j'}$ in v_1 that contains either a b or a $\diamond$, whence there is no $u_{j'}$ in v_1 abelian compatible with u_j, a contradiction.

Thus, we may assume there is some k_i such that $i_0 \leq k_{i-1} < k_i < i_2$. Suppose there is no k_j such that $i_1 \leq k_j < i_2$ and i is the largest integer satisfying $k_i < i_1$. Then $i_0 \leq k_{i-1} < k_i < i_1$. Suppose $i \equiv 0, 1, \dots,$ or $2m+1 \bmod \mu$ (the other case

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

being similar). Then, some u_j in v_0 contains a b, but there are no b's or $\diamond$'s in v_1. Hence there is no $u_{j'}$ in v_1 abelian compatible with u_j, a contradiction. The claim follows from Lemma 3.2.

Let k_i be such that $i_0 \leq k_{i-1} < i_1 \leq k_i < i_2$. Suppose $i \equiv 0 \bmod \mu$ and $k_i = i_1$. Then u_{m-1} contains a b, and the only factor in v_1 that contains a $\diamond$ or a b is u_m, so it must be that $\alpha_{m-1} = \alpha_m$. By hypothesis, there must be some $\alpha_{j_1} = \alpha_m$ for $j_1 \notin \{m-1, m\}$. Then $u_{j_1} \not\Updownarrow u_{m-1}$, so u_{j_1} contains either a $\diamond$ or a b. This implies u_{j_1} is not in v_1, and if u_{j_1} is in v_0, then $i_0 - 1 \leq k_{i-(2m+2)} < \ldots < k_{i-1} < k_i = i_1$. Hence there are $2m+1$ c's in v_0 and by the pigeonhole principle, at least one factor, u_{j_2}, in v_0 contains three c's. However, there is only one $\diamond$ and one c in v_1, so there is no factor in v_1 abelian compatible with u_{j_2}. If $k_i > i_1$, then there is a b in v_1, forcing there to be a $\diamond$ or a b in v_0 and we arrive at the same contradiction. The case when $i \equiv 2m+2 \bmod \mu$ proceeds similarly.

Now, suppose $i \equiv 1, 2, \ldots,$ or $2m+1 \bmod \mu$, and let $i' < i$ be the largest integer satisfying $i' \equiv 0 \bmod \mu$. If $i_0 - 1 \geq k_{i'}$, then there is a c in v_0, so some factor of v_0 does not have an abelian compatible factor in v_1, because v_1 contains no $\diamond$'s and no c's. So if $i_0 \leq k_j < i_2$, then $j \equiv 1, 2, \ldots,$ or $2m+1 \bmod \mu$, and moreover, the only letters in u not in A must be at position k_j for some j. Thus we can replace each of these with an a to find an instance of p in w, a contradiction. The case where $i \equiv 2m+3, 2m+4, \ldots,$ or $4m+3 \bmod \mu$ is almost identical. Therefore, in all cases, we reach a contradiction. □

Our next goal is to give an avoidability criterion. We are also able to come up with an alternative definition of an abelian occurrence of p.

Proposition 3.14. *Let E be an alphabet, $p = \alpha_0 \ldots \alpha_{n-1}$ be a pattern with $\alpha_i \in E$. Moreover, let $A = \{a_0, \ldots, a_{k-1}\}$ be an alphabet, $u_0 \ldots u_{n-1}$ be a partial word over A such that $|u_i| = |u_j|$ whenever $\alpha_i = \alpha_j$. For $\alpha \in E$, let $m_{\alpha,i} = \max_{\alpha_j = \alpha} |u_j|_{a_i}$. Then $u_0 \ldots u_{n-1}$ is an abelian occurrence of p if and only if $|u_i| \geq m_{\alpha_i,0} + \ldots + m_{\alpha_i,k-1}$ for all i.*

Proof. First, assume that $u_0 \ldots u_{n-1}$ is an abelian occurrence of p. Thus there exists a corresponding full word $v_0 \ldots v_{n-1}$ that is an abelian occurrence of p, where v_i is compatible with u_i for each i. This implies that if $\alpha_i = \alpha_j$ and $a \in A$ then $|v_i|_a = |v_j|_a \geq |u_j|_a$. Thus we get that $|v_i|_{a_j} \geq m_{\alpha_i,j}$ for all j, so that

$$|u_i| = |v_i| = |v_i|_{a_0} + \ldots + |v|_{a_{k-1}} \geq m_{\alpha_i,0} + \ldots + m_{\alpha_i,k-1}$$

just as we wanted.

On the other hand, we can begin by assuming that $|u_i| \geq m_{\alpha_i,0} + \ldots + m_{\alpha_i,k-1}$ for all i. Consider any $\alpha \in E$ that occurs in p. Let $\{i_0, \ldots, i_{m-1}\} = \{i \mid \alpha_i = \alpha\}$, then for any j define v_{i_j} in the way that we replace $m_{\alpha,l} - |u_i|_{a_l}$ of the holes in u_{i_j} with a_l for $l > 0$, and replace the rest of the holes with a_0. We can do this since $|u_{i_j}| \geq m_{\alpha,0} + \ldots + m_{\alpha,k-1}$ for all j's. Moreover, note that v_{i_j} is compatible with u_{i_j}, while $|v_{i_j}|_{a_l} = m_{\alpha,l} = |v_{i_0}|_{a_l}$ for $l \neq 0$, and $|v_{i_j}|_{a_0} = |v_{i_j}| - (m_{\alpha,1} + \ldots + m_{\alpha,k-1}) = |v_{i_0}| - (m_{\alpha,1} + \ldots + m_{\alpha,k-1}) = |v_{i_0}|_{a_0}$.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Thus each v_{i_j} is a rearrangement of v_{i_0}. Then it follows that $v_0 \dots v_{n-1}$ is an abelian occurrence of p. □

Our next result is very general.

Theorem 3.15. *Let $p = \alpha_0 \dots \alpha_{n-1}$ be an abelian avoidable pattern in full words. Then there exists a partial word with infinitely many holes so that, if $u_0 \dots u_{n-1}$ is an abelian occurrence of p then $|u_i| < |p| + 2$ for some i.*

Proof. Let p be as above, and note that $|p| = n$. Since p is abelian avoidable there exist alphabets $A_0, \dots, A_{n+1}$, each disjoint from one another, so that there exists an infinite word w_i over A_i that avoids p in the abelian sense. We can define

$$w = w_0[0] \dots w_{n+1}[0] w_0[1] \dots w_{n+1}[1] w_0[2] \dots w_{n+1}[2] \dots$$

which is an infinite word over $A = A_0 \cup A_1 \cup \dots \cup A_{n+1}$. Consider $b \notin A$. Then consider a sequence $n + 2 = |p| + 2 < k_0 < k_1 < k_2 < \dots$ so that $w[k_i] \in A_0$, and so that $k_{i+1} > |p| k_i$. We can define a partial word v as follows:

$$v[j] = \begin{cases} \diamond, & \text{if } j = k_i \text{ for some } i,\ i \equiv 0 \bmod 3; \\ b, & \text{if } j = k_i \text{ for some } i,\ i \not\equiv 0 \bmod 3; \\ w[j], & \text{otherwise.} \end{cases}$$

Assume that $u_0 \dots u_{n-1}$ is an abelian occurrence of p in v. Let $u_j = v[i_j \dots i_{j+1})$. Assume that $\alpha_j = \alpha_l$ for $j < l$. Then we get that u_l contains at most one $\diamond$ by Lemma 3.3. Similarly, u_j can contain at most one $\diamond$. This is because between each pair of $\diamond$'s lie two b's. However, for each b contained in u_j we have that u_l contains either a b or a $\diamond$, so if u_j contains two $\diamond$'s then u_l contains two symbols from the set $\{b, \diamond\}$. So there is a t so that $i_l \leq k_t < k_{t+1} < i_{l+1}$, a contradiction with Lemma 3.3.

Since $u_0 \dots u_{n-1}$ is an abelian occurrence of p, there is a full word $v_0 \dots v_{n-1}$ so that $v_i \uparrow u_i$ for all i, and $v_0 \dots v_{n-1}$ is an abelian occurrence of p. If α_i occurs exactly once in p, we can replace v_i with any non-empty string and the result is still an abelian occurrence of p; therefore we assume that v_i is constructed from u_i by replacing each hole in u_i with a letter from A_0. Otherwise, α_i occurs more than once in p, in which case u_i contains at most one $\diamond$, and in order to form v_i from u_i, we need to fill in at most one $\diamond$ with a letter from A. Therefore there are at most $|p|$ $\diamond$'s that are filled in with letters not from the alphabet A_0. By the pigeonhole principle, there exists an alphabet A_j so that none of the $\diamond$'s are filled with letters from A_j. Then define a morphism $\phi : (A \cup \{b\})^* \to A_j^*$ so that, given $a \in A \cup \{b\}$ we have $\phi(a) = a$, if $a \in A_j$, and $\phi(a) = \varepsilon$, otherwise.

Then note that, by construction, $\phi(v_0 \dots v_{n-1})$ is a factor of w_j. This implies that $\phi(v_0 \dots v_{n-1})$ is not an abelian occurrence of p, since w_j avoids p in the abelian sense. However, the only way that $\phi(v_0 \dots v_{n-1})$ cannot be an abelian occurrence of p is if $\phi(v_i) = \varepsilon$ for some i. This implies that v_i does not contain any letters from A_j, so u_i does not contain any letters from A_j. However, u_i is a factor of v, and every factor of v of length at least $|p| + 2$ contains a letter from A_j. Thus $|u_i| < |p| + 2$. □

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

3.3. Ternary patterns

Next, we consider ternary patterns in partial words. By Corollary 3.12, all ternary patterns of length ≥ 8 are non-trivially abelian avoidable by partial words with infinitely many holes. Moreover, a word with infinitely many holes over five letters exists that avoids non-trivial abelian squares [4]. Therefore, patterns containing abelian squares are non-trivially abelian avoidable as well, and we only need to examine patterns of length at most 7 without any abelian squares. Currie *et al.* in [9] characterize these remaining six patterns: $\alpha\beta\alpha\gamma\alpha\beta\alpha$, $\alpha\beta\gamma\alpha\beta\alpha$ and $\alpha\beta\gamma\alpha\gamma$ are abelian unavoidable, while $\alpha\beta\alpha\gamma\beta\alpha\beta$, $\alpha\beta\alpha\gamma\beta\gamma$ and $\alpha\beta\gamma\beta\alpha\beta\gamma$ are abelian avoidable. By Proposition 2.2, there is no word with arbitrarily many holes that non-trivially avoids $\alpha\beta\alpha\gamma\alpha\beta\alpha$, $\alpha\beta\gamma\alpha\beta\alpha$ or $\alpha\beta\gamma\alpha\gamma$ in the abelian sense. By Proposition 3.8, since each variable occurs at least twice in $\alpha\beta\alpha\gamma\beta\gamma$ and $\alpha\beta\gamma\beta\alpha\beta\gamma$, partial words with infinitely many holes exist that avoid them in the abelian sense. The last pattern, $\alpha\beta\alpha\gamma\beta\alpha\beta$, can be shown to be abelian avoidable by a word with infinitely many holes with a proof similar to that of [9], Lemma 8. Therefore, the non-trivial abelian avoidability of the ternary patterns in the context of partial words is complete.

We now consider the case of complete abelian avoidability.

Proposition 3.16. *There exists a partial word with infinitely many holes that avoids $\alpha\beta\alpha\gamma\beta\alpha\beta$ in the abelian sense.*

Proof. We know that $\alpha\beta\alpha\gamma\beta\alpha\beta$ can be avoided in the abelian sense by some infinite word w over some alphabet A. Assume $a, b, c, d, e \notin A$, and let u, $|u| = 5$, be a factor of w that occurs infinitely often. Then define $3 < k_0 < k_1 < \ldots$ so that $2k_i < k_{i+1}$ and $w[k_i - 2 \ldots k_i + 2] = u$ for all i. We can then define w' as follows:

$$w'[j] = \begin{cases} a, & \text{if } j = k_i - 2 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ b, & \text{if } j = k_i - 1 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ \diamond, & \text{if } j = k_i \quad \text{for some } i,\ i \equiv 0 \bmod 3; \\ d, & \text{if } j = k_i + 1 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ e, & \text{if } j = k_i + 2 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ c, & \text{if } j = k_i \quad \text{for some } i,\ i \not\equiv 0 \bmod 3; \\ w[j], & \text{otherwise.} \end{cases}$$

Then using similar methods to those we have used throughout the paper, we can show that w' avoids $\alpha\beta\alpha\gamma\beta\alpha\beta$ in the abelian sense. □

Proposition 3.16 implies the following corollary.

Corollary 3.17. *If a ternary pattern, p, does not contain $\omega\omega$ as a subpattern for any $\omega \in E = \{\alpha, \beta, \gamma\}$, then p is avoided by a partial word with infinitely many holes in the abelian sense if and only if p is abelian avoidable in full words.*

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Proof. Assume that p is abelian avoidable, but not abelian avoidable by any partial word with infinitely many holes. Since $p \notin \{\alpha\alpha, \beta\beta, \gamma\gamma\}$, this implies by Proposition 3.8 that there is some variable in p that occurs exactly once. Without loss of generality, we can assume it is γ, so $p = p_0\gamma p_1$ where each p_i is a pattern over $\{\alpha, \beta\}$. Note neither p_0 nor p_1 can be avoided by a partial word with infinitely many holes, so by Theorem 3.5, since p does not contain $\omega\omega$ as a subpattern for any $\omega \in E$, the possibilities for p_0, p_1 are (up to isomorphism) $\varepsilon, \alpha, \alpha\beta, \alpha\beta\alpha$. However, the only such pattern that is abelian avoidable is $\alpha\beta\alpha\gamma\beta\alpha\beta$. By Proposition 3.16, this can be avoided by a partial word with infinitely many holes in the abelian sense, a contradiction. □

Proposition 3.18. *Let p be an avoidable ternary pattern in full words. Then p can be avoided in the abelian sense by a partial word with infinitely many holes if and only if it is not a subpattern of any pattern in*

$$S = \{\alpha\alpha\beta\alpha\alpha\gamma\alpha\alpha\alpha, \alpha\alpha\beta\alpha\gamma\alpha\alpha\alpha, \alpha\alpha\beta\alpha\gamma\gamma, \alpha\alpha\beta\gamma\alpha\alpha\beta, \alpha\alpha\beta\gamma\alpha\beta\beta, \alpha\alpha\beta\gamma\beta\beta, \alpha\beta\beta\gamma\alpha\alpha\beta, \alpha\beta\beta\gamma\alpha\beta\beta\}.$$

Proof. Let w be an infinite partial word with infinitely many holes. Then some factor of w of the form $a\diamond b$ occurs infinitely often. By considering this, we see that the patterns in S cannot be avoided by a partial word with infinitely many holes in the abelian sense, so that direction of the above statement follows.

Therefore assume that p cannot be avoided in the abelian sense by any partial word with infinitely many holes, and that p is not a subpattern of any pattern in S. Since p is avoidable, p contains $\omega\omega$ as a factor for some $\omega \in E$. We can assume that $p \neq \omega\omega$, since in that case p cannot be avoided in the abelian sense by a partial word with infinitely many holes, and p is up to isomorphism a subpattern of a pattern in S. We can assume without loss of generality that $\omega = \alpha$. More, assume that there is a $\delta \in E$, $\delta \neq \alpha$, so that $\delta\delta$ is a subpattern of p. We can assume without loss of generality that $\delta = \beta$. From Proposition 3.8, there must be some variable in p that occurs exactly once. Since α and β each occurs more than once, this variable must be γ. Thus we can write $p = q_0\gamma q_1$, where q_0 and q_1 are both patterns over $\{\alpha, \beta\}$ that cannot be avoided by a partial word with infinitely many holes. For each i either q_i is a factor of $\alpha\alpha\beta\alpha\alpha$ or $\beta\beta\alpha\beta\beta$. Note that q_0 contains either $\alpha\alpha$ or $\beta\beta$ as a subpattern, since otherwise q_1 would have to contain both as subpatterns, which is impossible.

Assume q_0 contains $\alpha\alpha$, the other case being similar. Then q_1 contains $\beta\beta$. This implies that q_0 is a factor of $\alpha\alpha\beta\alpha\alpha$ and q_1 is a factor of $\beta\beta\alpha\beta\beta$. First, consider the case that $|q_0| < 4$, $|q_1| < 4$. Then since p is not a subpattern of any pattern in S the only possibilities are $\alpha\alpha\beta\gamma\beta\beta\alpha$ and $\beta\alpha\alpha\gamma\alpha\beta\beta$.

Consider $\alpha\alpha\beta\gamma\beta\beta\alpha$, the other case being similar. Then let w be an infinite word that avoids abelian squares over A. There exists u so that u occurs infinitely often in w, $|u| = 3$. Consider $2 < k_0 < k_1 < k_2 < \ldots$ so that $w[k_i - 1 \ldots k_i + 1] = u$ for

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

all i and $2k_i < k_{i+1}$. Then consider $a, b, c \notin A$, and define w' by

$$w'[j] = \begin{cases} a, & \text{if } j = k_i - 1 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ \diamond, & \text{if } j = k_i \quad \text{for some } i,\ i \equiv 0 \bmod 3; \\ b, & \text{if } j = k_i + 1 \text{ for some } i,\ i \equiv 0 \bmod 3; \\ c, & \text{if } j = k_i \quad \text{for some } i,\ i \not\equiv 0 \bmod 3; \\ w[j], & \text{otherwise.} \end{cases}$$

We know from earlier that w' avoids non-trivial abelian squares. Moreover, we claim w' avoids $\alpha\alpha\beta\gamma\beta\beta\alpha$ in the abelian sense. To see this, assume that $v_0v_1v_2v_3v_4v_5v_6$ is an abelian occurrence. Then v_0v_1 is an abelian square and must be trivial. Thus v_0v_1 contains a hole and $|v_0| = |v_1| = 1$, so v_0v_1 is a factor of $a\diamond b$. Similarly, v_4v_5 is a factor of $a\diamond b$, and since v_2 corresponds to the same letters as v_4 and v_5 it follows that either $v_2 = a$, $v_2 = b$, or $v_2 = \diamond$. Thus $v_0v_1v_2$ is a factor of length 3 over $\{a, b, \diamond\}$. However, the only such word is $a\diamond b$. This implies $v_0 = a$. By a similar logic $v_4v_5v_6 = a\diamond b$, so $v_6 = b$. However, v_0 and v_6 correspond to the same variable in p, which implies that $v_0 = a$ is a rearrangement of $v_6 = b$, which is not the case.

The other possibility is that $\alpha\alpha$ appears as an abelian square in p but $\beta\beta$ and $\gamma\gamma$ do not. This case is similar to the previous one. □

We finally wish to put upper bounds on the abelian avoidability indices of ternary patterns.

Proposition 3.19. *Let p be any pattern. If p contains an abelian square, then it is non-trivially 5-abelian avoidable in partial words.*

Proof. As in the proof of Theorem 3.5, let w be an infinite partial word over 5 letters with infinitely many holes that avoids non-trivial abelian squares. Then w avoids p, so p is non-trivially abelian 5-avoidable. □

This proposition encompasses all of the non-trivially avoidable ternary patterns, except $\alpha\beta\alpha\gamma\beta\gamma$, $\alpha\beta\gamma\beta\alpha\beta\gamma$, and $\alpha\beta\alpha\gamma\beta\alpha\beta$, so only these three remain to classify. In fact, the patterns $\alpha\alpha\beta\gamma\alpha\beta\alpha$ and $\alpha\beta\alpha\gamma\alpha\alpha\beta$ demonstrate that the bound in Proposition 3.19 is tight with respect to ternary patterns. Both of these patterns have ordinary (non-abelian) avoidability index 5 in partial words [5, 6], and since any ordinary occurrence of a pattern is also an abelian occurrence, this implies the abelian avoidability indices of these patterns in partial words are at least 5.

Now we address the pattern $\alpha\beta\gamma\beta\alpha\beta\gamma$. We first deal with avoiding the pattern $\alpha\beta\gamma\beta\alpha\beta\gamma$ in full words using a morphism given in [17]. Let $A = \{a, b, c, d, e\}$ be an alphabet and define two morphisms, ψ and ϕ, on A^* as follows: ψ is the cyclic morphism $\psi(a) = b$, $\psi(b) = c$, *etc.*, and $\phi(a) = bacaeacadaeadab$, $\phi(b) = \psi(\phi(a))$, $\phi(c) = \psi^2(\phi(a))$, *etc.* The morphism ϕ has several nice properties, specifically that the image of any letter has length 15, and contains 7 instances of that letter and 2 instances of every other letter. It is also an abelian square-free morphism, that is, if $w \in A^*$ is abelian square-free, then $\phi(w)$ is also abelian square-free [17].

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Proposition 3.20. *Let $w \in A^*$, and assume $\phi(w)$ meets $\alpha\beta\gamma\beta\alpha\beta\gamma$ in the abelian sense. Then w meets $\alpha\beta\gamma\beta\alpha\beta\gamma$ in the abelian sense, or w contains an abelian square. Furthermore, the word $\phi^\omega(a)$ avoids $\alpha\beta\gamma\beta\alpha\beta\gamma$ in the abelian sense.*

Proof. Let $p = \alpha\beta\gamma\beta\alpha\beta\gamma$. We will use a method of proof similar to the one used in [17] to prove that ϕ is abelian square-free. By assumption, there exists some factor, $u = u_0u_1 \dots u_6$, of $\phi(w)$ such that $u_0 \,\phi\, u_4$, $u_1 \,\phi\, u_3 \,\phi\, u_5$, and $u_2 \,\phi\, u_6$. Let w_i be the largest factor of w such that $\phi(w_i) \in \mathrm{fac}(u_i)$. If w_i is non-empty, then we can write $u_i = x_i\phi(w_i)y_i$, where x_i is a suffix of $\phi(\ell_i)$ for some $\ell_i \in A$, y_i is a prefix of $\phi(\hat{\ell}_i)$ for some $\hat{\ell}_i \in A$, and $|x_i|, |y_i| \leq 14$. It is obvious in this case that

$$15|w_i| \leq |u_i| = 15|w_i| + |x_i| + |y_i| \leq 15|w_i| + 28. \tag{1}$$

Moreover, in the case that w_i is empty, u_i can be at most 28 letters long, otherwise u_i will contain $\phi(\ell)$ for some ℓ. So in either case, Inequality (1) holds.

Now let w_0' and w_4' be the respective results (up to permutation) of canceling as many shared letters from w_0 and w_4 as possible. For example, if $w_0 = bac$ and $w_4 = adab$, then $w_0' = c$ and $w_4' = ad$. Define w_2' and w_6' similarly. Finally, we compute w_1', w_3', and w_5' (up to permutation) by canceling letters common to all three factors, w_1, w_3, and w_5. Then we construct u_i' by replacing $\phi(w_i)$ in u_i with $\phi(w_i')$. Since we have removed the same letters from both u_0 and u_4, we have that $u_0' \,\phi\, u_4'$, and in the same way $u_1' \,\phi\, u_3' \,\phi\, u_5'$ and $u_2' \,\phi\, u_6'$. So $u' = u_0'u_1' \dots u_6'$ is an abelian instance of p.

We claim that $|w_0'|, |w_4'| \leq 2$. Without loss of generality, assume that $|w_0'| \leq |w_4'|$. Further assume, by way of contradiction, that $|w_4'| = 3$. Since w_4' is non-empty, we can write $u_4' = x_4\phi(w_4')y_4$, where x_4 and y_4 are as defined above. Recall from Inequality (1) that

$$15|w_0| \leq |u_0| \leq 15|w_0| + 28, \text{ and}$$
$$15|w_4| \leq |u_4| \leq 15|w_4| + 28.$$

Since $|u_0| = |u_4|$, these inequalities imply that the difference between $|w_0|$ and $|w_4|$, which is the same as the difference between $|w_0'|$ and $|w_4'|$, can be at most 1. Then w_0' is non-empty, so w_0 is as well. Hence, we can write $u_0' = x_0\phi(w_0')y_0$, where x_0 and y_0 are as defined above, and consider the three following cases.

Case 1. Suppose $w_4' \,\phi\, \ell\ell\ell$ for some $\ell \in A$. Then $|u_4'|_\ell \geq |\phi(w_4')|_\ell = 21$, but $|\phi(w_0')|_\ell \leq 6$ (since $|w_0'| \leq 3$ and w_0' cannot have an ℓ). Since $|u_0'|_\ell = |u_4'|_\ell$, it must be that $|x_0y_0|_\ell \geq 15$. However, $|x_0|_\ell$ and $|y_0|_\ell$ are at most 7, so this is a contradiction.

Case 2. Suppose $w_4' \,\phi\, \ell_1\ell_1\ell_2$ for some distinct $\ell_1, \ell_2 \in A$. Then $|u_4'|_{\ell_1} \geq |\phi(w_4')|_{\ell_1} = 16$, and $|\phi(w_0')|_{\ell_1} \leq 6$. Since $|u_0'|_{\ell_1} = |u_4'|_{\ell_1}$, it must be that $|x_0y_0|_{\ell_1} \geq 10$. The only way this is possible is if x_0 and y_0 are both affixes of $\phi(\ell_1)$. Moreover, $|u_4'|_{\ell_2} \geq |\phi(w_4')|_{\ell_2} = 11$, but $|\phi(w_0')|_{\ell_2} \leq 6$. This implies $|x_0y_0|_{\ell_2} \geq 5$. However, this is only possible if either x_0 or y_0 is an affix of $\phi(\ell_2)$, which contradicts our above conclusion.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Case 3. Suppose $w'_4 \Updownarrow \ell_1\ell_2\ell_3$ for some distinct $\ell_1, \ell_2, \ell_3 \in A$. For each $i \in [1 \dots 3]$, we have that $|u'_4|_{\ell_i} \geq |\phi(w'_4)|_{\ell_i} = 11$, but $|\phi(w'_0)|_{\ell_i} \leq 6$. Since $|u'_0|_{\ell_i} = |u'_4|_{\ell_i}$, this implies $|x_0 y_0|_{\ell_i} \geq 5$, which is only possible if either x_0 or y_0 is an affix of $\phi(\ell_i)$. Without loss of generality, we may assume that x_0 is a suffix of $\phi(\ell_1)$. Then it must be that y_0 is a prefix of $\phi(\ell_2)$, because x_0 cannot be a suffix of $\phi(\ell_2)$. However, this is a contradiction, because neither x_0 nor y_0 can be an affix of $\phi(\ell_3)$.

Since w'_4 must fit into one of these cases and each case leads to a contradiction, we know that $|w'_4| \neq 3$. Moreover, if $|w'_4| > 3$, we can find these same contradictions by considering a length 3 factor of w'_4 and augmenting all of the counts, $|\phi(w'_0)|_\ell, |\phi(w'_4)|_\ell$, *etc.*, by $2(|w'_4| - 3)$. This proves the claim.

Note that we can apply this same argument to show that $|w'_2|, |w'_6| \leq 2$. Moreover, we can apply this pairwise to w'_1, w'_3, w'_5 to conclude that there are at most 2 letters in w_1 not in w_3, at most 2 letters in w_1 not in w_5, at most 2 letters in w_3 not in w_5, and *vice versa*. This shows that $|w'_1|, |w'_3|, |w'_5| \leq 4$.

Observe that u' must be a factor of $\phi(w')$ for some $w' \in A^*$. Let w' be the smallest such word. Suppose we can write $u'_i = x_i\phi(w'_i)y_i$ and let m_i be the position of $\phi(w')$ which is the first position of $\phi(w'_i)$, or the first position of y_i if $w'_i = \varepsilon$. If both w'_i and y_i are empty, then let m_i be the first position after u'_i. Now let $j_i = \frac{m_i}{15}$, and note that $w'_i = w'[j_i \dots j_i + |w'_i|)$.

Now suppose $w'_i \neq \varepsilon$ for all i, and suppose $v' = v'_0v'_1 \dots v'_6$ is an abelian instance of p in w' and write $v'_i = w'[k_i \dots k_{i+1})$. Further suppose

$$k_i \leq j_i \leq j_i + |w'_i| \leq k_{i+1}. \tag{2}$$

Note that j_i is defined because $w'_i \neq \varepsilon$. Then $w'_i \in \mathrm{fac}(v'_i)$. Construct v_i by replacing w'_i in v'_i with w_i. Since we removed the same letters from w_0 and w_4, it follows from this construction that $v_0 \Updownarrow v_4$ if and only if $v'_0 \Updownarrow v'_4$, and an analogous statement holds for v_1, v_3, v_5 and v_2, v_6, so $v = v_0v_1 \dots v_6$ is an abelian instance of p. Moreover, $v \in \mathrm{fac}(w)$, so w meets p in the abelian sense.

Now suppose $w'_1 = w'_3 = w'_5 = \varepsilon$, and suppose $v'_0v'_2v'_4v'_6$ is an abelian instance of $\alpha\gamma\alpha\gamma$ in w', where $v'_i = w'[k_i \dots k_{i+1})$, and such that k_i satisfies Inequality (2) for each $i \in \{0, 2, 4, 6\}$. Construct $v = v_0v_2v_4v_6$ as above, and note that v is an abelian instance of $\alpha\gamma\alpha\gamma$ in w. Furthermore, v is an abelian square, so w contains an abelian square. We can consider when $w'_0 = w'_4 = \varepsilon$ or $w'_2 = w'_6 = \varepsilon$ and find similar results.

We conclude that if there exists some $v' = v'_0v'_1 \dots v'_6$ which is an abelian instance of p in w', where $v'_i = w'[k_i \dots k_{i+1})$, such that k_i satisfies Inequality (2) whenever j_i is defined and we allow v'_i to be empty (so long as $v' \neq \varepsilon$), then w either meets p in the abelian sense, or w contains an abelian square. Thus, it remains to show that for any u', w' must contain such an abelian instance of p. Note that there are only finitely many u'_i where $w'_i = \varepsilon$. Moreover, if $w'_i \neq \varepsilon$ and u'_i can be written as $x_i\phi(w'_i)y_i$, then there are only finitely many x_i and y_i. Finally, because $|w'_i|$ is bounded, there are only finitely many w'_i. Thus there are finitely many u' which we need to consider. Using a computer, we can check that for all possible u', w' must contain an abelian instance of p as described above.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

Therefore, because the word a is abelian p-free and square-free, ϕ is known to be an abelian square-free morphism, and any instance of p in $\phi^\omega(a)$ must occur in some $\phi^n(a)$ where $n \in \mathbb{N}$, it follows by induction that $\phi^\omega(a)$ avoids p in the abelian sense. □

Corollary 3.21. *The pattern $\alpha\beta\gamma\beta\alpha\beta\gamma$ is 5-abelian avoidable in full words.*

The next corollary follows trivially from the above and Theorem 3.13.

Corollary 3.22. *The pattern $\alpha\beta\gamma\beta\alpha\beta\gamma$ is 7-abelian avoidable in partial words.*

4. Avoiding partial words with finitely many holes

Now, we give constructions for avoiding words with finitely many holes.

Theorem 4.1. *If p is an abelian k-avoidable pattern, then for every integer $h \geq 0$ there exists an infinite word with h holes over an alphabet of size $k + 2h$ that is non-trivially abelian p-free.*

Proof. Since p is abelian avoidable, $|p| \geq 2$ and we write $p = \alpha\beta q$, where q is a word and α, β are variables. If $\alpha = \beta$, p contains a square, and from [4], a word with infinitely many holes over a five-letter alphabet can be constructed that avoids non-trivial abelian squares. If we only keep h holes and fill in the other holes with any letter of the alphabet in that construction, we obtain an infinite word with h holes that avoids non-trivial abelian squares and thus non-trivially avoids p in the abelian sense. Therefore we only need to consider when $\alpha \neq \beta$.

First consider the case where α and β are both contained in q. Since p is abelian k-avoidable, there exists an infinite word w over a k-letter alphabet A such that w avoids p in the abelian sense. Let $C = \{a_0, \ldots, a_{h-1}\} \cup \{b_0, \ldots, b_{h-1}\}$, where $A \cap C = \emptyset$. Define $A' = A \cup C$, so $\|A'\| = k + 2h$. Then an infinite partial word w' over A' is defined as follows:

$$w'[i] = \begin{cases} a_j, & \text{if } i = 4j \quad \text{for some } j,\ 0 \leq j < h; \\ a_j, & \text{if } i = 4j+1 \text{ for some } j,\ 0 \leq j < h; \\ \diamond, & \text{if } i = 4j+2 \text{ for some } j,\ 0 \leq j < h; \\ b_j, & \text{if } i = 4j+3 \text{ for some } j,\ 0 \leq j < h; \\ w[i], & \text{otherwise.} \end{cases}$$

Then w' is of the form $a_0a_0\diamond b_0a_1a_1\diamond b_1 \ldots a_{h-1}a_{h-1}\diamond b_{h-1}w(4h)w(4h+1)\ldots$

Let $p = \alpha_0 \ldots \alpha_{m-1}$, $\alpha_i \in E$. Note that w' has h holes and each a_j only appears twice and each b_j only appears once. Suppose there exists $u = u_0u_1 \ldots u_{m-1}$ that is a non-trivial abelian occurrence of p in w'. Notice that u_0 must contain at least one a_i or b_i, else u occurs in w. Since α is in q, there is some $l > 1$ such that $u_l \updownarrow u_0$. It is easy to see that this forces u_l to contain a hole. Hence u_0 is a factor of $w'[0 \ldots 4h-1]$. The only possibilities for u_0, that is the only factors of $w'[0 \ldots 4h-1]$ that have an abelian compatible factor at a later position in w'

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

starting at least one position after u_0, are of the forms a_i, $a_i\diamond$, or $\diamond b_i$ for some i. If $u_0 = a_i$, then $u_l = \diamond$, so u is a trivial occurrence. Hence u_0 is of the form $a_i\diamond$ or $\diamond b_i$ for some i. Since β appears again in q, using the same logic as above, u_1 must also be of the form $a_j\diamond$ or $\diamond b_j$ for some j. However, such u_0u_1 cannot be a factor of w'.

We prove the rest of the claim by induction on the length of p. The case of $|p| = 1$ is trivial. When α is not contained in q, if $\alpha\beta q$ is abelian k-avoidable, so is βq. Since $|\beta q| < |\alpha\beta q|$, there exists an infinite partial word with h holes over a $(k + 2h)$-letter alphabet that non-trivially avoids βq in the abelian sense, thus non-trivially avoiding $\alpha\beta q$ in the abelian sense. Similarly, if β does not occur in q, then q is non-trivially abelian avoidable over an alphabet of size $k + 2h$, and so are βq and $\alpha\beta q$, which is p. □

Proposition 4.2. *Let $p \neq \alpha\alpha$ be an abelian k-avoidable pattern in full words with each of its variables occurring at least twice. Then for every integer $h \geq 0$ there exists an infinite partial word with h holes over an alphabet of size $k + 2h$ that is abelian p-free.*

Proof.
The infinite partial word with h holes is exactly the same as the one we provided in Theorem 4.1, *i.e.* w'. We have already shown that it is non-trivially abelian p-free. Moreover, we prove that it avoids p in the abelian sense even trivially by induction on the number of holes, if each variable of p occurs at least twice. Trivial abelian squares are not avoidable, so we do not consider them. In other cases, suppose w' contains a trivial occurrence of p, denoted by u. Then u contains at least one hole. When u contains only one hole, it is a factor of $b_{i-1}a_{i-1}a_i\diamond b_ia_ia_{i+1}$, which obviously does not meet p in the abelian sense. Assume any factor u' containing $j - 1$ holes, denoted by $u' = b_{i-1}a_{i-1}a_i\diamond b_ia_i \ldots \diamond b_{i+j-2}a_{i+j-2}a_{i+j-1}$, where $1 \leq j - 1 < h$, is abelian p-free. Then consider $u'' = u'\diamond b_{i+j-1}a_{i+j-1}a_{i+j}$ or $u'' = b_{i-2}a_{i-2}a_{i-1}\diamond u'$. Let u'' be the former one (the latter one is handled similarly). If a trivial abelian occurrence of p is in u'', the hole before b_{i+j-1} can only be replaced by either the letter a_{i-1} or b_{i-1}, since otherwise the sequence contains $j-1$ holes and is trivially abelian p-free. Suppose it is replaced by a_{i-1}. Then from the first occurrence of $\diamond$ in u'' to that occurrence of a_{i-1} there are only $j - 1$ holes and at least $2(j - 1)$ distinct letters. It is impossible that each of those letters be paired with one of the holes. The claim follows by induction. □

The following result is concerned with an arbitrary insertion of a hole in an infinite word that avoids a pattern in the abelian sense.

Theorem 4.3. *If p is an abelian avoidable pattern, then there exists an infinite abelian avoiding full word so that we can insert a hole into any position and get a partial word that is non-trivially abelian p-free.*

Proof. Assume $p = \alpha_0 \ldots \alpha_{n-1}$ is abelian avoided by an infinite word w over alphabet A. Let $A' = A^2 \times \{0, 1\}$, and define an infinite word v as follows. Let

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

$j \in \{0,1\}$. For integer $i \geq 0$, if $i \equiv j \bmod 2$, let $v[i] = (w[i-j]w[i-j+1], j) \in A'$. Suppose towards a contradiction that v is not abelian p-free. Thus, there exist i_1 and i_2 such that $\alpha_{i_1} = \alpha_{i_2}$, $w[i_1 \dots i_1 + l]$ and $w[i_2 \dots i_2 + l]$ are not permutations of each other, while $v[i_1 \dots i_1 + l]$ and $v[i_2 \dots i_2 + l]$ are. This means $w[i_1 - j]w[i_1 - j + 1] \dots w[i_1 + l - j]w[i_1 + l - j + 1]$ is a permutation of $w[i_2 - j]w[i_2 - j + 1] \dots w[i_2 + l - j]w[i_2 + l - j + 1]$. Since $j = 0$ or 1, this contradicts the fact that w avoids p in the abelian sense.

Assume we replace one letter in v with a $\diamond$ to get v', and $u = u_0 u_1 \dots u_{n-1}$ is a non-trivial abelian occurrence of p in v'. Note that u must contain the hole. Suppose u_s contains the hole and u_s is compatible with a permutation of u_t, for some s, t. Then since $|u_s| > 1$, either $(w[i_1 - j_1]w[i_1 - j_1 + 1], j_1)\diamond$ or $\diamond(w[i_1 - j_1]w[i_1 - j_1 + 1], j_1)$ is a factor of u_s, for some integers i_1, j_1. This means u_t contains a letter $(w[i_2 - j_2]w[i_2 - j_2 + 1], j_2)$ where $j_2 = j_1$, $w[i_2 - j_2] = w[i_1 - j_1]$, and $w[i_2 - j_2 + 1] = w[i_1 - j_1 + 1]$. Since j_1, j_2 is either 0 or 1, $w[i_1] = w[i_2]$ for either case. This contradicts w avoiding p in the abelian sense. □

5. Conclusion

In this paper, we have generalized results from abelian powers in partial words given by Blanchet−Sadri *et al.* [4,7] and abelian patterns in full words given by Currie *et al.* [9–11]. Theorem 3.10 gives tight bounds on the lengths of patterns, over any pattern alphabet of at least three variables, for abelian pattern avoidance by partial words with infinitely many holes. Our bound in Theorem 3.9 is tight for unary and binary pattern alphabets.

Our current bound, given in Proposition 3.7, for abelian 2-avoidable (resp., 3-avoidable) binary patterns p is $|p| \geq 363$ (resp., $|p| \geq 119$). A topic for future research would be to improve on those two bounds.

We have also given upper limits for the abelian avoidability indices of all binary and ternary patterns in partial words, which we have done for all binary patterns and, non-trivially, for all but two ternary patterns. Referring to Section 3, an upper bound for the abelian avoidability indices of $\alpha\beta\alpha\gamma\beta\alpha\beta$ and $\alpha\beta\alpha\gamma\beta\gamma$ remains open. The avoidability of patterns over four variables has been studied by Currie *et al.* [9], but has not been completed.

Acknowledgements. The authors thank Dimin Xu for valuable comments and suggestions. 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.

References

[1] J. Berstel and L. Boasson, Partial words and a theorem of Fine and Wilf. *Theoret. Comput. Sci.* **218** (1999) 135–141.

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

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press

[3] F. Blanchet-Sadri and S. Simmons, Abelian pattern avoidance in partial words, in *MFCS 2012, 37th International Symposium on Mathematical Foundations of Computer Science*. Edited by B. Rovan, V. Sassone and P. Widmayer. Vol. 7464 of *Lect. Notes Comput. Sci.* Springer-Verlag, Berlin (2012) 210–221.
[4] F. Blanchet-Sadri, J.I. Kim, R. Mercaş, W. Severa, S. Simmons and D. Xu, Avoiding abelian squares in partial words. *J. Combin. Theory, Ser. A* **119** (2012) 257–270.
[5] F. Blanchet-Sadri, A. Lohr and S. Scott, Computing the partial word avoidability indices of binary patterns. *J. Discrete Algorithms* **23** (2013) 113–118
[6] F. Blanchet-Sadri, A. Lohr and S. Scott. Computing the partial word avoidability indices of ternary patterns. *J. Discrete Algorithms* **23** (2013) 119–142
[7] F. Blanchet-Sadri, S. Simmons and D. Xu, Abelian repetitions in partial words. *Adv. Appl. Math.* **48** (2012) 194–214.
[8] J.D. Currie, Pattern avoidance: themes and variations. *Theoret. Comput. Sci.* **339** (2005) 7–18.
[9] J. Currie and V. Linek, Avoiding patterns in the abelian sense. *Can. J. Math.* **53** (2001) 696–714.
[10] J. Currie and T. Visentin, On abelian 2-avoidable binary patterns. *Acta Informatica* **43** (2007) 521–533.
[11] J. Currie and T. Visentin, Long binary patterns are abelian 2-avoidable. *Theoret. Comput. Sci.* **409** (2008) 432–437.
[12] F.M. Dekking, Strongly non-repetitive sequences and progression-free sets. *J. Combin. Theory, Ser. A* **27** (1979) 181–185.
[13] P. Erdős, Some unsolved problems. *Magyar Tudományos Akadémia Matematikai Kutató Intézete Közl.* **6** (1961) 221–254.
[14] V. Keränen, Abelian squares are avoidable on 4 letters, in *ICALP 1992, 19th International Colloquium on Automata, Languages and Programming*. Edited by W. Kuich, vol. 623 of *Lect. Notes Comput. Sci.* Springer-Verlag, Berlin (1992) 41–52.
[15] P. Leupold, Partial words for DNA coding, in *10th International Workshop on DNA Computing*. Edited by G. Rozenberg, P. Yin, E. Winfree, J.H. Reif, B.-T. Zhang, M.H. Garzon, M. Cavaliere, M.J. Pérez-Jiménez, L. Kari and S. Sahu. Vol. 3384 of *Lect. Notes Comput. Sci.* Springer-Verlag, Berlin (2005) 224–234.
[16] M. Lothaire, *Algebraic Combinatorics on Words*. Cambridge University Press, Cambridge (2002).
[17] P.A.B. Pleasants, Non repetitive sequences. *Proc. Cambridge Philosophical Soc.* **68** (1970) 267–274.
[18] A. Thue, Über unendliche Zeichenreihen. *Norske Vid. Selsk. Skr. I, Mat. Nat. Kl. Christiana* **7** (1906) 1–22.

Communicated by Ch. Choffrut.
Received March 14, 2014. Accepted March 17, 2014.

https://doi.org/10.1051/ita/2014014 Published online by Cambridge University Press