Discrete Mathematics and Theoretical Computer Science DMTCS vol. **12:5**, 2010, 83–96

Crucial abelian k-power-free words

Amy Glen[1†] Bjarni V. Halldórsson[2‡] Sergey Kitaev[2§]

[1]*Department of Mathematics and Statistics, School of Chemical and Mathematical Sciences, Murdoch University, Perth, WA 6150, AUSTRALIA*
[2]*Reykjavík University, Menntavegur 1, 101 Reykjavík, ICELAND*

received 6th July 2009, revised 31st October 2010, accepted 19th November 2010.

In 1961, Erdős asked whether or not there exist words of arbitrary length over a fixed finite alphabet that avoid patterns of the form XX' where X' is a permutation of X (called *abelian squares*). This problem has since been solved in the affirmative in a series of papers from 1968 to 1992. Much less is known in the case of *abelian k-th powers*, i.e., words of the form $X_1X_2 \cdots X_k$ where X_i is a permutation of X_1 for $2 \leq i \leq k$.

In this paper, we consider *crucial words* for abelian k-th powers, i.e., finite words that avoid abelian k-th powers, but which cannot be extended to the right by any letter of their own alphabets without creating an abelian k-th power. More specifically, we consider the problem of determining the minimal length of a crucial word avoiding abelian k-th powers. This problem has already been solved for abelian squares by Evdokimov and Kitaev (2004), who showed that a minimal crucial word over an n-letter alphabet $\mathcal{A}_n = \{1, 2, \ldots, n\}$ avoiding abelian squares has length $4n-7$ for $n \geq 3$. Extending this result, we prove that a minimal crucial word over $\mathcal{A}_n$ avoiding abelian cubes has length $9n - 13$ for $n \geq 5$, and it has length 2, 5, 11, and 20 for $n = 1, 2, 3,$ and 4, respectively. Moreover, for $n \geq 4$ and $k \geq 2$, we give a construction of length $k^2(n-1) - k - 1$ of a crucial word over $\mathcal{A}_n$ avoiding abelian k-th powers. This construction gives the minimal length for $k = 2$ and $k = 3$. For $k \geq 4$ and $n \geq 5$, we provide a lower bound for the length of crucial words over $\mathcal{A}_n$ avoiding abelian k-th powers.

MSC (2010): 05D99; 68R05; 68R15

Keywords: pattern avoidance; abelian square-free word; abelian cube-free word; abelian power; crucial word; Zimin word

1 Introduction

Let $\mathcal{A}_n = \{1, 2, \ldots, n\}$ be an n-letter alphabet and let $k \geq 2$ be an integer. A word W over $\mathcal{A}_n$ contains a *k-th power* if W has a factor of the form $X^k = XX \cdots X$ (k times) for some non-empty word X. A k-th power is *trivial* if X is a single letter. For example, the word $V = 13243232323243$ over $\mathcal{A}_4$ contains the (non-trivial) 4-th power $(32)^4 = 32323232$. A word W contains an *abelian k-th power* if W has a factor of the form $X_1X_2 \cdots X_k$ where X_i is a permutation of X_1 for $2 \leq i \leq k$. The cases $k = 2$ and $k = 3$

†Email: `amy.glen@gmail.com`
‡Email: `bjarnivh@ru.is`
§Email:`sergey@ru.is`

give us (*abelian*) *squares* and *cubes*, respectively. For instance, the preceding word V contains the abelian square $43232\,32324$ and the word $123\,312\,213$ is an abelian cube. A word W is (*abelian*) k-*power-free* if W *avoids* (abelian) k-th powers, i.e., if W does not contain any (abelian) k-th powers. For example, the word 1234324 is abelian cube-free, but not abelian square-free since it contains the abelian square $234\,324$.

A word W is *crucial* with respect to a given set of *prohibited words* (or simply *prohibitions*) if W avoids the prohibitions, but Wx does not avoid the prohibitions for any letter x occurring in W. A *minimal crucial word* is a crucial word of the shortest length. For example, the word $W = 21211$ (of length 5) is crucial with respect to abelian cubes since it is abelian cube-free and the words $W1$ and $W2$ end with the abelian cubes 111 and $21\,21\,12$, respectively. Actually, W is a minimal crucial word over $\{1, 2\}$ with respect to abelian cubes. Indeed, one can easily verify that there do not exist any crucial abelian cube-free words on two letters of length less than 5.

Abelian squares were first introduced by Erdős [4], who asked whether or not there exist words of arbitrary length over a fixed finite alphabet that avoid patterns of the form XX' where X' is a permutation of X (i.e., abelian squares). This question has since been solved in the affirmative in a series of papers from 1968 to 1992 (see [5, 9, 7] and also [2]). Problems of this type were also considered by Zimin [10], who used the following sequence of words as a key tool.

The *Zimin word* Z_n over $\mathcal{A}_n$ is defined recursively as follows: $Z_1 = 1$ and $Z_n = Z_{n-1}nZ_{n-1}$ for $n \geq 2$. The first four Zimin words are:

$$Z_1 = 1,\ Z_2 = 121,\ Z_3 = 1213121,\ Z_4 = 121312141213121.$$

The k-*generalized Zimin word* $Z_{n,k} = X_n$ is defined as

$$X_1 = 1^{k-1} = 11\cdots 1,\ X_n = (X_{n-1}n)^{k-1}X_{n-1} = X_{n-1}nX_{n-1}n\cdots nX_{n-1}$$

where the number of 1's, as well as the number of n's, is $k-1$. Thus $Z_n = Z_{n,2}$. It is easy to see (by induction) that each $Z_{n,k}$ avoids (abelian) k-th powers and has length $k^n - 1$. Moreover, it is known that $Z_{n,k}$ gives the length of a minimal crucial word avoiding k-th powers.

However, much less is known in the case of abelian powers. Crucial abelian square-free words (also called *right maximal abelian square-free words*) of exponential length are given in [3] and [6], and it is shown in [6] that a minimal crucial abelian square-free word over an n-letter alphabet has length $4n-7$ for $n \geq 3$.

In this paper, we extend the study of crucial abelian k-power-free words to the case of $k > 2$. In particular, we provide a complete solution to the problem of determining the length of a minimal crucial abelian cube-free word (the case $k = 3$) and we conjecture a solution in the general case. More precisely, we show that a minimal crucial word over $\mathcal{A}_n$ avoiding abelian cubes has length $9n - 13$ for $n \geq 5$ (Corollary 9), and it has length 2, 5, 11, and 20 for $n = 1, 2, 3,$ and 4, respectively. For $n \geq 4$ and $k \geq 2$, we give a construction of length $k^2(n-1)-k-1$ of a crucial word over $\mathcal{A}_n$ avoiding abelian k-th powers (see Theorem 11). This construction gives the minimal length for $k = 2$ and $k = 3$, and we conjecture that this is also true for any $k \geq 4$ and sufficiently large n. We also provide a rough lower bound for the length of minimal crucial words over $\mathcal{A}_n$ avoiding abelian k-th powers, for $n \geq 5$ and $k \geq 4$ (see Theorem 12).

For a crucial word X over $\mathcal{A}_n$, we let $X = X_i\Delta_i$ where Δ_i is the factor of minimal length such that $\Delta_i i$ is a prohibition for $i \in \mathcal{A}_n$. Note that we can rename letters, if needed, so we can assume that for any minimal crucial word X, one has

$$\Delta_1 \subset \Delta_2 \subset \cdots \subset \Delta_n = X$$

where "$\subset$" means (proper) *right factor* (or *suffix*). In other words, for each $i = 2, 3, \ldots, n$, we have $\Delta_i = Y_i \Delta_{i-1}$ for some non-empty Y_i. In what follows we will use X_i and Y_i as stated above. We note that the definitions imply:

$$X = X_i \Delta_i = X_i Y_i \Delta_{i-1} = X_{n-1} Y_{n-1} Y_{n-2} \cdots Y_2 \Delta_1,$$

for any $i = 2, 3, \ldots, n-1$. Furthermore, in the case of crucial words avoiding abelian k-th powers, we write $\Delta_i i = \Omega_{i,1} \Omega_{i,2} \cdots \Omega_{i,k}$, where the *k blocks* $\Omega_{i,j}$ are equal up to permutation, and we denote by $\Omega'_{i,k}$ the block $\Omega_{i,k}$ without the rightmost i.

Hereafter, we let $\ell_k(n)$ denote the length of a minimal crucial word over $\mathcal{A}_n$ avoiding abelian k-th powers. The length of a word W is denoted by $|W|$, and we denote by $|W|_x$ the number of occurrences of a letter x in W. The *Parikh vector* of a word W over $\mathcal{A}_n$ is defined by

$$\mathcal{P}(W) := (|W|_1, |W|_2, \ldots, |W|_n).$$

Clearly, if W is an abelian k-th power, then $|W|_x \equiv 0 \pmod{k}$ for all letters x occurring in W.

2 Crucial words for abelian cubes

2.1 An upper bound for $\ell_3(n)$

The fact that the 3-generalized Zimin word $Z_{n,3}$ is crucial with respect to abelian cubes already gives us an upper bound of $3^n - 1$ for $\ell_3(n)$. In Theorem 1 (below) we improve this upper bound to $3 \cdot 2^{n-1} - 1$. We then discuss a further improvement of the bound using a greedy algorithm, which gives asymptotically $O((\sqrt{3})^n)$. This greedy algorithm provides minimal crucial abelian cube-free words over $\mathcal{A}_n$ for $n = 3$ and $n = 4$, while the construction in Theorem 1 is optimal only for $n = 3$. We also provide a construction of a crucial word of length $9n - 10$ which exceeds our lower bound by only 3 letters, for $n \geq 5$. Finally, we end this section with a construction of a crucial abelian cube-free word over $\mathcal{A}_n$ of length $9n - 13$, which coincides with the lower bound given in Theorem 8 of Section 2.2 for $n \geq 5$.

Theorem 1 *One has that* $\ell_3(n) \leq 3 \cdot 2^{n-1} - 1$.

Proof: We construct a crucial abelian cube-free word $X = X_n$ iteratively as follows. Set $X_1 = 11$ and assume X_{n-1} has been constructed. Then do the following:

1. Increase all letters of X_{n-1} by 1 to obtain X'_{n-1}.
2. Insert 1 after (to the right of) each letter of X'_{n-1} and adjoin one extra 1 to the right of the resulting word to get X_n.

For example, $X_2 = 21211$, $X_3 = 31213121211$, etc. It is easy to verify that $|X_n| = 3 \cdot 2^{n-1} - 1$. We show by induction that X_n avoids abelian cubes, whereas $X_n x$ does not avoid abelian cubes for any $x \in \mathcal{A}_n$. Both claims are trivially true for $n = 1$. Now take $n \geq 2$. If X_n contains an abelian cube, then removing all 1's from it, we would deduce that X_{n-1} must also contain an abelian cube, contradicting the fact that X_{n-1} contains no abelian cubes.

It remains to show that extending X_n to the right by any letter x from $\mathcal{A}_n$ creates an abelian cube. If $x = 1$ then we get 111 from the construction of X_n. On the other hand, if $x > 1$ then we swap the rightmost 1 with the rightmost x in Xx, thus obtaining a word where every other letter is 1; removing all 1's and decreasing each of the remaining letters by 1, we have $X_{n-1}(x-1)$, which contains an abelian cube (by the induction hypothesis). □

As a further improvement of Theorem 1 we sketch here, without providing all the details, the work of a *greedy algorithm*. Here, "greedy" means that we assume (by induction) that an "optimal" crucial abelian cube-free word X_{n-1} over $\mathcal{A}_{n-1}$ has been constructed. Next (for the greedy step) we add just two n's (the minimum possible), and then we add as few of the other letters as possible to build a crucial abelian cube-free word X_n over $\mathcal{A}_n$. More precisely, we set $\Delta_1 = 11$ and assuming that Δ_{i-1} is built, we consider the minimum list T_i of letters we must add (forming Y_i and possibly updating Δ_{i-1} by permuting letters in Y_{i-1}) to build Δ_i, for $i = 2, 3, \ldots, n$, which can then be turned into a construction of a crucial abelian cube-free word X_n over $\mathcal{A}_n$. It is easy to see from the definitions of an abelian cube and a crucial word that T_2 must contain at least two 2's and at least one 1. Then T_3 must contain at least two 3's, at least one 2, and at least three 1's (the last statement follows from the fact that the two 3's in Δ_3 are supposed to be accompanied by at least one 1 in $\Omega_{3,1}$ and $\Omega_{3,2}$ in $\Delta_3 3 = \Omega_{3,1}\Omega_{3,2}\Omega_{3,3}$, but without extra 1's, whose number of occurrences must be divisible by 3, we cannot manage it). Running this type of argument, one can come up with the following T_i for initial values of i:

$$T_2 = \{1,2,2\}, T_3 = \{1,1,1,2,3,3\}, T_4 = \{1,1,1,2,2,2,3,4,4\},$$

$$T_5 = \{1,1,1,1,1,1,1,1,1,2,2,2,3,3,3,4,5,5\},$$

$$T_6 = \{1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,3,3,3,4,4,4,5,6,6\}, \text{etc.}$$

In particular, we observe that $(|T_n|)_{n\geq 1} = 2, 3, 6, 9, 18, 27, \ldots$ where

$$|T_{2i+1}| = 2 \cdot 3^i \quad \text{and} \quad |T_{2i+2}| = 3^i \quad \text{for all } i \geq 0.$$

Hence, these considerations lead to a crucial abelian cube-free word over an n-letter alphabet of length

$$\sum_{j=0}^{\lfloor \frac{n-1}{2} \rfloor} 2 \cdot 3^j + \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} 3^j.$$

Initial values for the lengths are 2, 5, 11, 20, 38, 65, ... Furthermore, by observing that $\sum_{j=0}^{m} 3^j = \frac{3^{m+1}-1}{2}$ for any integer $m \geq 0$ (which can be easily proved by induction), we deduce that the length of a crucial abelian cube-free word obtained in this way is asymptotically $O((\sqrt{3})^n)$. Even though this length is exponential, we *do* in fact obtain optimal values for $n \leq 4$. Below we list optimal abelian cube-free crucial words over $\mathcal{A}_n$ for $n = 1, 2, 3, 4$ of lengths 2,5,11, and 20, respectively:

$$11;$$
$$21211;$$
$$11231321211;$$
$$42131214231211321211.$$

A construction giving the best possible upper bound for $n \geq 5$ can be easily described by examples, and we do this below (for $n = 4, 5, 6, 7$; the construction does not work for $n \leq 3$). We also provide a general description. The pattern in the construction is easy to recognize.

An almost optimal construction for crucial abelian cube-free words.

An almost optimal construction of a crucial abelian cube-free word W_n over $\mathcal{A}_n$ for $n = 4, 5, 6, 7$ is shown below. We use spaces to separate the blocks $\Omega_{n,1}$, $\Omega_{n,2}$, and $\Omega'_{n,3}$ in $W_n = \Delta_n$.

$$W_4 = 344233122\ 433221432\ 32122334$$

$$W_5 = 455344233122\ 544332215432\ 43212233445$$

$$W_6 = 566455344233122\ 655443322165432\ 54321223344556$$

$$W_7 = 677566455344233122\ 766554433221765432\ 65432122334455667$$

In general, the block $\Omega_{n,1}$ in $W_n = \Delta_n = \Omega_{n,1}\Omega_{n,2}\Omega'_{n,3}$ is built by adjoining the factors $i(i+1)(i+1)$ for $i = n-1, n-2, \ldots, 1$. The block $\Omega_{n,2}$ is built by adjoining together the following factors: n, xx for $n-1 \geq x \geq 2$, and $n(n-1)\ldots 2$. Finally, the block $\Omega'_{n,3}$ is built by adjoining the factors $(n-1)(n-2)\ldots 1$, xx for $2 \leq x \leq n-1$, and finally n.

It is easy to see that $|W_n| = 9n - 10$. We omit the details of showing that W_n is crucial with respect to abelian cubes since this can be shown in a similar manner for the optimal construction described below (see the proof of Theorem 11).

An optimal construction for crucial abelian cube-free words.

An optimal construction of a crucial abelian cube-free word E_n over $\mathcal{A}_n$ for $n = 4, 5, 6, 7$ works as follows. As above, we use spaces to separate the blocks $\Omega_{n,1}$, $\Omega_{n,2}$, and $\Omega'_{n,3}$ in $E_n = \Delta_n$.

$$E_4 = 34423311\ 34231134\ 3233411$$

$$E_5 = 45534423311\ 45342311345\ 4323344511$$

$$E_6 = 56645534423311\ 56453423113456\ 5432334455611$$

$$E_7 = 67756645534423311\ 67564534231134567\ 6543233445566711$$

In general, the block $\Omega_{n,1}$ in $E_n = \Delta_n = \Omega_{n,1}\Omega_{n,2}\Omega'_{n,3}$ is built by adjoining the factors $i(i+1)(i+1)$ for $i = n-1, n-2, \ldots, 2$, followed by two 1's. The block $\Omega_{n,2}$ is built by adjoining the following factors: $i(i+1)$ for $i = n-1, n-2, \ldots, 2$, followed by 11, and then the factor $34\cdots(n-1)n$. Finally, the block $\Omega'_{n,3}$ is built by adjoining the factors $(n-1)(n-2)\cdots 32$, then xx for $3 \leq x \leq n-1$, followed by n, and finally two 1's.

We have $E_n = \Omega_{n,1}\Omega_{n,2}\Omega'_{n,3}$ where $\Omega_{n,3} = \Omega'_{n,3}n$, and by construction each $\Omega_{n,i}$ contains two 1's, one 2, two n's, and three x's for $x = 3, \ldots, n-1$. That is, for each $i = 1, 2, 3$, the Parikh vector of the block $\Omega_{n,i}$ is given by $\mathcal{P}(\Omega_{n,i}) = (2, 1, 3, 3, \ldots, 3, 2)$. Hence, $\mathcal{P}(E_n) = (6, 3, 9, 9, \ldots, 9, 5)$, and therefore $|E_n| = 6 + 3 + 9(n-3) + 5 = 9n - 13$. Moreover, for all $n \geq 4$, the word E_n is crucial with respect to abelian cubes. We omit the proof of this latter fact since it is very similar to the proof of Theorem 11 (to follow), from which the fact can actually be deduced by setting $k = 3$ (in view of Remark 3, later). Thus, a minimal crucial word avoiding abelian cubes has length at most $9n - 13$ for $n \geq 4$. That is:

Theorem 2 *For $n \geq 4$, we have $\ell_3(n) \leq 9n - 13$.*

Proof: See the proof of Theorem 11 where one needs to set $k = 3$ (in view of Remark 3, later). □

2.2 A lower bound for $\ell_3(n)$

If $X = \Delta_n$ is a crucial word over $\mathcal{A}_n$ with respect to abelian cubes, then clearly the number of occurrences of each letter except n must be divisible by 3, whereas the number of occurrences of n is 2 modulo 3. We sort in non-decreasing order the number of occurrences of the letters $1, 2, \ldots, n-1$ in X to get a non-decreasing sequence of numbers $(a_1 \leq a_2 \leq \cdots \leq a_{n-1})$. Notice that a_i does not necessarily correspond to the letter i. We denote by a_0 the number of occurrences of the letter n. Also note that a_0 can be either larger or smaller than a_1. By definitions, $|X| = \sum_{i=0}^{n-1} a_i$.

For example, the abelian cube-free crucial word E_n of length $9n - 13$ in Sec. 2.1 has the following sequence of a_i's: $(a_0, a_1, \ldots, a_{n-1}) = (5, 3, 6, 9, \ldots, 9)$. In this subsection, we prove that this sequence cannot be improved for $n \geq 5$, meaning that, e.g., 5 cannot be replaced by 2, and/or 6 cannot be replaced by 3, and/or 9('s) cannot be replaced by 3('s) or 6('s), no matter what construction we use to form a crucial word. This is a direct consequence of Lemmas 4–7 (below) and is recorded in Theorem 8. In the rest of this section we use, without explanation, the following two facts that are easy to see from the definitions. For any letter x in a crucial abelian cube-free word X over $\mathcal{A}_n$:

1. $|\Delta_x|_x \equiv 2 \pmod 3$ and $|\Delta_x|_y \equiv 0 \pmod 3$ for any other letter y occurring in X.

2. If $x + 1$ occurs in X, then we have $\Delta_{x+1} = Y_{x+1}\Delta_x$ where $|Y_{x+1}|_{x+1} \equiv 2 \pmod 3$, $|Y_{x+1}|_x \equiv 1 \pmod 3$, and $|Y_{x+1}|_y \equiv 0 \pmod 3$ for any other letter y occurring in X.

The following fact will also be useful.

Lemma 3 *Suppose X is a crucial abelian cube-free word over $\mathcal{A}_n$ containing letters x and y such that $x < y < n$ and $|X|_x = |X|_y = 6$. Then Δ_x cannot contain 5 occurrences of the letter x.*

Proof: Suppose to the contrary that (under the hypotheses of the lemma) Δ_x contains 5 occurrences of the letter x. Let $A_1 = Y_n Y_{n-1} \cdots Y_{y+1}$ and $A_2 = Y_y Y_{y-1} \cdots Y_{x+1}$ so that $X = A_1 A_2 \Delta_x$. Then $|A_1 A_2|_x = 1$ and $|A_1 A_2|_y \geq 3$, contradicting the fact that each of the blocks $\Omega_{n,1}$, $\Omega_{n,2}$, and $\Omega'_{n,3}$ in $X = \Delta_n = \Omega_{n,1}\Omega_{n,2}\Omega'_{n,3}$ must each receive two x's and two y's. □

Lemma 4 *For a crucial abelian cube-free word X over $\mathcal{A}_n$, the sequence of a_i's cannot contain 3, 3. That is, $(a_1, a_2) \neq (3, 3)$.*

Proof: Suppose that x and y are letters such that $x < y < n$ and $|X|_x = |X|_y = 3$. Let $A_1 = Y_n Y_{n-1} \cdots Y_{y+1}$ and $A_2 = Y_y Y_{y-1} \cdots Y_{x+1}$ so that we have $X = A_1 A_2 \Delta_x$. Then we must have the following distribution of x's and y's in X: $|A_1|_y = 1$, $|A_2|_y = 2$, $|A_2|_x = 1$, and $|\Delta_x|_x = 2$. However, we get a contradiction, since each of the blocks $\Omega_{n,2}$ and $\Omega'_{n,3}$ in $X = \Delta_n = \Omega_{n,1}\Omega_{n,2}\Omega'_{n,3}$ must receive one copy of x and one copy of y, which is impossible (no x can exist between the two rightmost y's). □

Lemma 5 *For a crucial abelian cube-free word X over $\mathcal{A}_n$, the sequence of a_i's cannot contain 6, 6, 6.*

Proof: Suppose that x, y, z are three letters such that $x < y < z < n$ and $|X|_x = |X|_y = |X|_z = 6$. Let $A_1 = Y_nY_{n-1}\cdots Y_{z+1}$, $A_2 = Y_zY_{z-1}\cdots Y_{y+1}$, and $A_3 = Y_yY_{y-1}\cdots Y_{x+1}$ so that $X = A_1A_2A_3\Delta_x$. Then the minimal requirements on the A_i are as follows: $|A_1|_z \geq 1$, $|A_2|_z \geq 2$, $|A_2|_y \geq 1$, $|A_3|_y \geq 2$, and $|A_3|_x \geq 1$. Moreover, applying Lemma 3 to x and y, we have $|\Delta_x|_2 = 2$. And applying the same lemma to the letters y and z guarantees that A_1A_2 contains 4 y's (in particular, Δ_x does not contain any y's).

Looking at $X = \Delta_n = \Omega_{n,1}\Omega_{n,2}\Omega'_{n,3}$, we see that for each $i = 1, 2, 3$, $|\Omega_{n,i}|_x = |\Omega_{n,i}|_y = |\Omega_{n,i}|_z = 2$. Thus, in A_3, we must have the following order of letters: x, y, y and the boundary between $\Omega_{n,2}$ and $\Omega'_{n,3}$ must be between x and y in A_3. But then Δ_x entirely belongs to $\Omega'_{n,3}$, so it cannot contain any z's (if it would do so, Δ_x would then contain 3 z's which is impossible). On the other hand, we must have $|A_3|_z = 3$ for $\Omega'_{n,3}$ to receive 2 z's. Thus, Δ_y contains 2 y's, 3 z's, and 3 x's, which is impossible by Lemma 4 applied to the word Δ_y with two letters occurring exactly 3 times each. (Alternatively, one can see, due to the considerations above, that no z can be between the two rightmost x's, contradicting the structure of Δ_y). □

Lemma 6 *For a crucial abelian cube-free word X over $\mathcal{A}_n$, the sequence of a_i's cannot contain* $3, 6, 6$.

Proof: Suppose that x, y, and z are letters such that $|X|_x = 3$ and $|X|_y = |X|_z = 6$. We consider three cases covering all the possibilities up to renaming y and z.

Case 1: $z < y < x < n$. One can see that Δ_y does not contain x, but it contains at least 3 z's contradicting the fact that each of the blocks $\Omega_{n,1}$, $\Omega_{n,2}$, and $\Omega'_{n,3}$ must receive 1 x and 2 z's.

Case 2: $x < z < y < n$. We let $A = Y_nY_{n-1}\cdots Y_{z+1}$ so that $X = A\Delta_z$. All three x's must be in Δ_z, while A must contain at least 3 y's contradicting the fact that each of the blocks $\Omega_{n,1}$, $\Omega_{n,2}$, and $\Omega'_{n,3}$ must receive 1 x and 2 y's.

Case 3: $z < x < y < n$. We let $A_1 = Y_nY_{n-1}\cdots Y_{y+1}$, $A_2 = Y_yY_{y-1}\cdots Y_{x+1}$, and $A_3 = Y_xY_{x-1}\cdots Y_{z+1}$ so that $X = A_1A_2A_3\Delta_z$. The minimal requirements on the A_i and Δ_z are as follows: $|A_1|_y \geq 1$, $|A_2|_y \geq 2$, $|A_2|_x = 1$, $|A_3|_x = 2$, $|A_3|_z \geq 1$, and $|\Delta_z|_z \geq 2$. Now Δ_z cannot contain 3 y's, for otherwise, considering the structure of Δ_x, it would not be possible to distribute x's and y's in a proper way. However, if A_3 contains 3 y's then, so as not to contradict the structure of Δ_x (no proper distribution of y's and z's would exist), Δ_x must contain 3 z's, which contradicts to the structure of $X = \Delta_n = \Omega_{n,1}\Omega_{n,2}\Omega'_{n,3}$ (no proper distribution of y's and z's would exist among the blocks $\Omega_{n,1}$, $\Omega_{n,2}$, and $\Omega'_{n,3}$, each of which is supposed to contain exactly 2 occurrences of y and 2 occurrences of z). Thus, there are no y's in Δ_x, contradicting the structure of Δ_n (no proper distribution of y's and x's would exist among the blocks $\Omega_{n,1}$,$\Omega_{n,2}$, and $\Omega'_{n,3}$).

□

Lemma 7 *For a crucial abelian cube-free word X over $\mathcal{A}_n$,*

$$(a_0, a_1, a_2, a_3, a_4) \neq (2, 3, 6, 9, 9).$$

Proof: Suppose $|X|_n = 2$ and assume that $|X|_t = 3$ for some other letter t. If $t \neq n-1$, then all three occurrences of t are in Δ_{n-1}, whereas the two occurrences of n are in Y_n (recall that $X = \Delta_n = Y_n\Delta_{n-1}$). This contradicts the fact that $|\Omega_{n,1}|_n = |\Omega_{n,1}|_t = 1$. Thus, $t = n-1$ and $|X|_{n-1} = 3$.

Now, assuming x, y, and z are three letters, with $x < y < z < n-1$, occurring $\{6, 9, 9\}$ times in X (we do not specify which letter occurs how many times). Then, as in the proof of Lemma 5, we deduce that Δ_z belongs entirely to the block $\Omega'_{n,3}$. Moreover, the block $\Omega'_{n,3}$ has $\{2, 3, 3\}$ occurrences of letters x, y, z (in some order). However, if x or y occur twice in $\Omega'_{n,3}$, they occur twice in Δ_z, contradicting the structure of Δ_z. Thus z must occur twice in $\Omega'_{n,3}$, and the letters x and y each occur 3 times in $\Omega'_{n,3}$. But then it is clear that x and y must each occur 3 times in Δ_z, contradicting the fact that x and z should be distributed properly in Δ_z, by Lemma 4. □

Theorem 8 *For $n \geq 5$, we have $\ell_3(n) \geq 9n - 13$.*

Proof: This is a direct consequence of the preceding four lemmas, which tell us that any attempt to decrease numbers in the sequence $(5, 3, 6, 9, 9, \ldots, 9)$ corresponding to E_n will lead to a prohibited configuration. □

Corollary 9 *For $n \geq 5$, we have $\ell_3(n) = 9n - 13$.*

Proof: The result follows immediately from Theorems 2 and 8. □

Remark 1 *Recall from Sec. 2.1 that $\ell_3(n) = 2, 5, 11, 20$ for $n = 1, 2, 3, 4$, respectively. For instance, the word* 42131214231211321211 *is a minimal crucial abelian cube-free word over $\mathcal{A}_4$ of length* 20 *($= 2 + 3 + 6 + 9$). This can be proved using similar arguments as in the proofs of the Lemmas 4–7.*

3 Crucial words for abelian k-th powers

3.1 An upper bound for $\ell_k(n)$ and a conjecture

The following theorem is a direct generalization of Theorem 1 and is a natural approach to obtaining an upper bound that improves $k^n - 1$ given by the k-generalized Zimin word $Z_{n,k}$.

Theorem 10 *For $k \geq 3$, we have $\ell_k(n) \leq k \cdot (k-1)^{n-1} - 1$.*

Proof: We proceed as in the proof of Theorem 1, with the only difference being that we begin with $X_1 = 1^{k-1}$ and put $(k-2)$ 1's to the right of each letter except for the last $(k-2)$ letters, after which we put $(k-1)$ 1's instead. □

We skip here the analysis of the work of a greedy algorithm, and proceed directly with the construction of a crucial abelian k-power-free word $W_{n,k}$ over $\mathcal{A}_n$ before describing the construction of a similar word $D_{n,k}$ that we believe to be optimal.

A construction of a crucial abelian k-power-free word $W_{n,k}$, where $n \geq 4$ and $k \geq 3$.

We illustrate each step of the algorithm by example, letting $k = n = 4$.

The construction can be explained directly, but we introduce it recursively, obtaining, for $n \geq 4$, $W_{n,k}$ from $W_{n,k-1}$ and using the abelian cube case $W_{n,3} = W_n$ as the basis. For $n = 4$,

$$W_{4,3} = \Omega_{4,1}\Omega_{4,2}\Omega'_{4,3} = 344233122\,433221432\,32122334.$$

Assume that $W_{n,k-1} = \Omega_{n,1}\Omega_{n,2}\ldots\Omega'_{n,k-1}$ is constructed and implement the following steps to obtain $W_{n,k}$:

1. Duplicate $\Omega_{n,1}$ in $W_{n,k-1}$ to obtain the word

$$W'_{n,k-1} = \Omega_{n,1}\Omega_{n,1}\Omega_{n,2}\ldots\Omega'_{n,k-1}.$$

For $n = k = 4$,

$$W'_{4,3} = 344233122\,344233122\,433221432\,32122334.$$

2. Append to the second $\Omega_{n,1}$ in $W'_{n,k}$ the factor $n(n-1)\ldots2$ (in our example, 432) to obtain $\Omega_{n,2}$ in $W_{n,k}$, and in each of the remaining blocks $\Omega_{n,i}$ in $W'_{n,k-1}$, duplicate the rightmost occurrence of each letter x, where $2 \leq x \leq n$, to obtain $W_{n,k}$. For $n = k = 4$,

$$W_{4,4} = 344423331222\,344233122432\,433221443322\,32122233344.$$

We provide two more examples here, namely $W_{5,4}$ and $W_{4,5}$, respectively, so that readers can check their understanding of the construction:

$$4555344423331222\ 4553442331225432\ 5443322155443322\ 432122233344455;$$

$$3444423333122222\ 344423331222432\ 344233122443322\ 433221444333222\ 32122223333444.$$

It is easy to see that $|W_{n,k}| = k^2(n-1) - 1$. We omit the proof that $W_{n,k}$ is crucial with respect to abelian k-th powers since it is similar to the proof for the following word $D_{n,k}$, which is k letters shorter than $W_{n,k}$ (see Theorem 11).

A construction of a crucial abelian k-power-free word $D_{n,k}$, where $n \geq 4$ and $k \geq 2$.

As we shall see, the following construction of the word $D_{n,k}$ is optimal for $k = 2, 3$. We believe that it is also optimal for any $k \geq 4$ and sufficiently large n (see Conjecture 1).

As our basis for the construction of the word $D_{n,k}$, we use the following word D_n, which is constructed as follows, for $n = 4, 5, 6, 7$. (As previously, we use spaces to separate the blocks $\Omega_{n,1}$ and $\Omega'_{n,2}$ in $D_n = \Delta_n$.)

$$D_4 = 34231\,3231$$

$$D_5 = 4534231\,432341$$

$$D_6 = 564534231\,54323451$$

$$D_7 = 67564534231\,6543234561$$

In general, the first block $\Omega_{n,1}$ in $D_n = \Delta_n = \Omega_{n,1}\Omega'_{n,2}$ is built by adjoining the factors $i(i+1)$ for $i = n-1, n-2, \ldots, 2$, followed by the letter 1. The second block $\Omega'_{n,2}$ is built by adjoining the factors $(n-1)(n-2)\cdots 432$, then $34\cdots(n-2)(n-1)$, and finally the letter 1.

Remark 2 *The above construction coincides with the construction given in [6, Theorem 5] for a minimal crucial abelian square-free word over $\mathcal{A}_n$ of length $4n-7$. In fact, the word D_n can be obtained from the minimal crucial abelian cube-free word E_n (defined in Sec. 2.1) by removing the second block in E_n and deleting the rightmost copy of each letter except 2 in the first and third blocks of E_n.*

Now we illustrate each step of the construction for the word $D_{n,k}$ by example, letting $n = 4$ and $k = 3$. The construction is very similar to that of $W_{n,k}$ and can be explained directly, but we introduce it recursively, obtaining $D_{n,k}$ from $D_{n,k-1}$ for $n \geq 4$, and using the crucial abelian square-free word $D_{n,2} := D_n$ as the basis. For $n = 4$, we have

$$D_{4,2} = \Omega_{4,1}\Omega'_{4,2} = 34231\,3231.$$

Assume that $D_{n,k-1} = \Omega_{n,1}\Omega_{n,2}\cdots\Omega'_{n,k-1}$ is constructed and implement the following steps to obtain $D_{n,k}$:

1. Duplicate $\Omega_{n,1}$ in $D_{n,k-1}$ to obtain the word

$$D'_{n,k-1} = \Omega_{n,1}\Omega_{n,1}\Omega_{n,2}\cdots\Omega'_{n,k-1}.$$

For $n = 4$ and $k = 3$, $D'_{4,2} = 34231\,34231\,3231$.

2. Append to the second $\Omega_{n,1}$ in $D'_{n,k-1}$ the factor $134\cdots n$ (in our example, 134; in fact, any permutation of $\{1, 3, 4, \ldots, n\}$ would work at this place) to obtain $\Omega_{n,2}$ in $D_{n,k}$. In each of the remaining blocks $\Omega_{n,i}$ in $D'_{n,k-1}$, duplicate the rightmost occurrence of each letter x, where $1 \leq x \leq n$ and $x \neq 2$. Finally, in the last block of $D'_{n,k}$ insert the letter n immediately before the leftmost 1 to obtain the word $D_{n,k}$. For $n = 4$ and $k = 3$, we have

$$D_{4,3} = \mathbf{34}4\mathbf{23}3\mathbf{1}1\,\mathbf{34231}134\,\mathbf{323}34\mathbf{1}1,$$

where the bold letters form the word $D'_{4,2}$ from which $D_{4,3}$ is derived.

We provide five more examples here, namely $D_{5,3}$, $D_{5,4}$, $D_{4,4}$, $D_{4,5}$, and $D_{6,4}$, respectively, so that readers can check their understanding of the construction:

$$45534423311\ 45342311345\ 4323344511;$$

$$455534442333111\ 455344233111345\ 453423111334455\ 43233344455111;$$

$$34442333111\ 34423311134\ 34231113344\ 3233344111;$$

$$34444233331111\ 344423331111134\ 34423311113344\ 34231111333444\ 3233334441111;$$

$$5666455534442333111\ 5664553442331113456\ 5645342311133445566\ 543233344455566111.$$

Remark 3 *By construction, $D_{n,3} = E_n$ for all $n \geq 4$.*

Theorem 11 *For $n \geq 4$ and $k \geq 2$, we have $\ell_k(n) \leq k^2(n-1) - k - 1$.*

Proof: Fix $n \geq 4$ and $k \geq 2$. We have

$$D_{n,k} = \Omega_{n,1}\Omega_{n,2}\cdots\Omega_{n,k-1}\Omega'_{n,k}$$

where $\Omega_{n,k} = \Omega'_{n,k}n$, and by construction each $\Omega_{n,i}$ contains $(k-1)$ occurrences of the letter 1, one occurrence of the letter 2, $(k-1)$ occurrences of the letter n, and k occurrences of the letter x for $x = 3, 4, \ldots, n-1$. That is, for each $i = 1, 2, \ldots, k$, the Parikh vector of the block $\Omega_{n,i}$ is given by

$$\mathcal{P}(\Omega_{n,i}) = (k-1, 1, k, k, \ldots, k, k-1), \tag{1}$$

and hence $\mathcal{P}(D_{n,k}) = (k(k-1), k, k^2, k^2, \ldots, k^2, k(k-1)-1)$. Thus,

$$|D_{n,k}| = k(k-1) + k + k^2(n-3) + k(k-1) - 1 = k^2(n-1) - k - 1.$$

We will now prove that $D_{n,k}$ is crucial with respect to abelian k-th powers; whence the result. The following facts, which are easily verified from the construction of $D_{n,k}$, will be useful in the proof.

Fact 1. In every block $\Omega_{n,i}$, the letter 3 has occurrences before and after the single occurrence of the letter 2.

Fact 2. In every block $\Omega_{n,i}$, all $(k-1)$ of the 1's occur after the single occurrence of the letter 2 (as the factor 1^{k-1}).

Fact 3. For all i with $2 \leq i \leq k-1$, the block $\Omega_{n,i}$ ends with n^{i-1} and the other $(k-1-i+1)$ n's occur (together as a string) before the single occurrence of the letter 2 in $\Omega_{n,i}$. In particular, there are exactly $k-2$ occurrences of the letter n between successive 2's in $D_{n,k}$.

Freeness: First we prove that $D_{n,k}$ is abelian k-power-free. Obviously, by construction, $D_{n,k}$ is not an abelian k-th power (as the number of occurrences of the letter n is not a multiple of k) and $D_{n,k}$ does not contain any *trivial* k-th powers, i.e., k-th powers of the form x^k for some letter x. Moreover, each block $\Omega_{n,i}$ is abelian k-power-free. For if not, then according to the Parikh vector of $\Omega_{n,i}$ (see (1)), at least one of the $\Omega_{n,i}$ must contain an abelian k-th power consisting of exactly k occurrences of the letter x for all $x = 3, 4, \ldots, n-1$, and no occurrences of the letters 1, 2, and n. But, by construction, this is impossible because, for instance, the letter 3 has occurrences before and after the letter 2 in each of the blocks $\Omega_{n,i}$ in $D_{n,k}$ (by Fact 1).

Now suppose, by way of contradiction, that $D_{n,k}$ contains a non-trivial abelian k-th power, say P. Then it follows from the preceding paragraph that P overlaps at least two of the blocks $\Omega_{n,i}$ in $D_{n,k}$. We first show that P cannot overlap three or more of the blocks in $D_{n,k}$. For if so, then P must contain at least one of the blocks, and hence P must also contain all k of the 2's. Furthermore, all of the 1's in each block occur after the letter 2 (by Fact 2), so there are $(k-1)^2 = k^2 - 2k + 1$ occurrences of the letter 1 between the leftmost and rightmost 2's in $D_{n,k}$. Thus, P must contain all $k(k-1) = k^2 - k$ of the 1's. Hence, since $\Omega'_{n,k}$ ends with 1^{k-1}, we deduce that P must end with the word

$$W = 23^{k-1}1^{k-1}\Omega_{n,2}\cdots\Omega_{n,k-1}\Omega'_{n,k},$$

where $|W|_n = k$, $|W|_3 = k(k-1) + (k-1) = k^2 - 1$, and $|W|_x = k(k-1)$ for $x = 4, \ldots, n-1$. It follows that P must contain all k^2 of the 3's. But then, since

$$D_{n,k} = (n-1)n^{k-1}\cdots 34^{k-1}W$$

(by construction), we deduce that P must contain all k^2 of the 4's that occur in $D_{n,k}$, and hence all k^2 of the 5's, and so on. That is, P must contain all k^2 occurrences of the letter x for $x = 3, \ldots, n-1$; whence, since $D_{n,k}$ begins with the letter $n-1$, we have $P = \Omega_{n,1}\Omega_{n,2}\cdots\Omega_{n,k} = D_{n,k}$, a contradiction.

Thus, P overlaps exactly two adjacent blocks in $D_{n,k}$, in which case P cannot contain the letter 2; otherwise P would contain all k of the 2's, and hence would overlap all of the blocks in $D_{n,k}$, which is impossible (by the preceding arguments). Hence, P lies strictly between two successive occurrences of the letter 2 in $D_{n,k}$. But then P cannot contain the letter n as there are exactly $k-2$ occurrences of the letter n between successive 2's in $D_{n,k}$ (by Fact 3). Therefore, since the blocks $\Omega_{n,i}$ with $2 \leq i \leq k-1$ end with the letter n, it follows that P overlaps the blocks $\Omega_{n,1}$ and $\Omega_{n,2}$. Now, by construction, $\Omega_{n,1}$ ends with 1^{k-1}, and hence P contains k of the $2(k-1) = 2k-2$ occurrences of the letter 1 in $\Omega_{n,1}\Omega_{n,2}$. But then P must contain the letter 2 because $\Omega_{n,1}$ contains exactly $(k-1)$ occurrences of the letter 1 (as a suffix) and all $(k-1)$ of the 1's in $\Omega_{n,2}$ occur after the letter 2 (by Fact 2); a contradiction.

We have now shown that $D_{n,k}$ is abelian k-power-free. It remains to show that $D_{n,k}x$ ends with an abelian k-th power for each letter $x = 1, 2, \ldots, n$.

Cruciality: By construction, $D_{n,k}n$ is clearly an abelian k-th power. It is also easy to see that $D_{n,k}1$ ends with the (abelian) k-th power $\Delta_1 1 := 1^k$. Furthermore, for all $m = n, n-1, \ldots, 4$, we deduce from the construction that

$$\begin{aligned}
\Omega_{m,1} &= (m-1)m^{k-1}\Omega_{m-1,1}, \\
\Omega_{m,2} &= (m-1)m^{k-2}\Omega_{m-1,2}m, \\
&\vdots \\
\Omega_{m,k-2} &= (m-1)m^{2}\Omega_{m-1,k-2}m^{k-3}, \\
\Omega_{m,k-1} &= (m-1)m\Omega_{m-1,k-1}m^{k-2}, \\
\Omega'_{m,k} &= (m-1)\Omega'_{m-1,k}[1^{k-1}]^{-1}(m-1)m^{k-2}1^{k-1},
\end{aligned}$$

where $\Omega'_{m-1,k}[1^{k-1}]^{-1}$ indicates the deletion of the suffix 1^{k-1} of $\Omega'_{m-1,k}$.

Consequently, for $x = n-1, n-2, \ldots, 3$, the word $D_{n,k}x$ ends with the abelian k-th power $\Delta_x x$ where Δ_x is such that

$$\Delta_{x+1} = x(x+1)^{k-1}\Delta_x \quad \text{with } \Delta_n := D_{n,k}.$$

□

Observe that $|D_{n,2}| = 4n-7$ and $|D_{n,3}| = 9n-13$. Hence, since $D_{n,k}$ is a crucial abelian k-power-free word (by the proof of Theorem 11), it follows from [6, Theorem 5] and Corollary 9 that the words $D_{n,2}$ and $D_{n,3}$ are minimal crucial words over $\mathcal{A}_n$ avoiding abelian squares and abelian cubes, respectively. That is, for $k = 2, 3$, the word $D_{n,k}$ gives the length of a minimal crucial word over $\mathcal{A}_n$ avoiding abelian k-th powers. In the case of $k \geq 4$, we make the following conjecture.

Conjecture 1 *For $k \geq 4$ and sufficiently large n, the length of a minimal crucial word over $\mathcal{A}_n$ avoiding abelian k-th powers is given by $k^2(n-1) - k - 1$.*

3.2 A Lower Bound for $\ell_k(n)$

A trivial lower bound for $\ell_k(n)$ is $nk-1$ as all letters except n must occur at least k times, whereas n must occur at least $k-1$ times. We give here the following slight improvement of the trivial lower bound, which must be rather imprecise though.

Theorem 12 *For $n \geq 5$ and $k \geq 4$, we have $\ell_k(n) \geq k(3n-4)-1$.*

Proof: Assuming that X is a crucial word over the n-letter alphabet $\mathcal{A}_n$ with respect to abelian k-th powers ($k \geq 4$), we see that adjoining any letter from $\mathcal{A}_n$ to the right of X must create a cube as a factor from the right. In particular, adjoining n from the right side leads to creating a cube of length at least $9n-13$ (by Lemmas 4–7). This cube will be $\Omega_{n,k-2}\Omega_{n,k-1}\Omega'_{n,k}$ in X and thus $\Omega_{n,i}$, for $1 \leq i \leq k-1$, will have length at least $3n-4$, whereas $\Omega'_{n,k}$ has length at least $3n-5$, which yields the result. □

4 Further Research

1. Prove or disprove Conjecture 1. Notice that the general construction uses a greedy algorithm for going from $k-1$ to k, which does not work for going from $n-1$ to n for a fixed k. However, we believe that the conjecture is true.

2. A word W over $\mathcal{A}_n$ is *maximal* with respect to a given set of prohibitions if W avoids the prohibitions, but xW and Wx do not avoid the prohibitions for any letter $x \in \mathcal{A}_n$. For example, the word 323121 is a maximal abelian square-free word over $\{1,2,3\}$ of minimal length. Clearly, the length of a minimal crucial word with respect to a given set of prohibitions is at most the length of a shortest maximal word. Thus, obtaining the length of a minimal crucial word we get a lower bound for the length of a shortest maximal word.

   Can we use our approach to tackle the problem of finding maximal words of minimal length? In particular, Korn [8] proved that the length $\ell(n)$ of a shortest maximal abelian square-free word over $\mathcal{A}_n$ satisfies $4n-7 \leq \ell(n) \leq 6n-10$ for $n \geq 6$, while Bullock [1] refined Korn's methods to show that $6n-29 \leq \ell(n) \leq 6n-12$ for $n \geq 8$. Can our approach improve Bullock's result (probably too much to ask when taking into account how small the gap is), or can it provide an alternative solution?

Acknowledgements

The authors would like to thank an anonymous referee for helpful comments and suggestions. The first and third authors were supported by the Icelandic Research Fund (grant nos. 09003801/1 and 060005012/3).

References

[1] E.M. Bullock, Improved bounds on the length of maximal abelian square-free words, *Electron. J. Combin.* 11 (2004) #R17.

[2] A. Carpi, On the number of abelian square-free words on four letters, *Discrete Appl. Math.* 81 (1998) 155–167.

[3] L.J. Cummings, M. Mays, A one-sided Zimin construction, *Electron. J. Combin.* 8 (2001), #R27.

[4] P. Erdős, Some unsolved problems, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 6 (1961) 221–254.

[5] A.A. Evdokimov, Strongly asymmetric sequences generated by a finite number of symbols, *Dokl. Akad. Nauk SSSR* 179 (1968) 1268–1271; *Soviet Math. Dokl.* 9 (1968) 536–539.

[6] A. Evdokimov, S. Kitaev, Crucial words and the complexity of some extremal problems for sets of prohibited words, *J. Combin. Theory Ser. A* 105 (2004) 273–289.

[7] V. Keränen, Abelian squares are avoidable on 4 letters, in: *Automata, languages and programming* (Vienna, 1992), 41–52; *Lecture Notes in Comput. Sci.*, vol. 623, Springer, Berlin, 1992.

[8] M. Korn, Maximal abelian square-free words of short length, *J. Combin. Theory Ser. A* 102 (2003) 207–211.

[9] P.A.B. Pleasants, Non-repetitive sequences, *Proc. Cambridge Philos. Soc.* 68 (1970) 267–274.

[10] A. Zimin, Blocking sets of terms, *Mat. Sb. (N.S.)* 119 (161) (1982) 363–375, 447; *Math. USSR Sbornik* 47 (1984) 353–364.