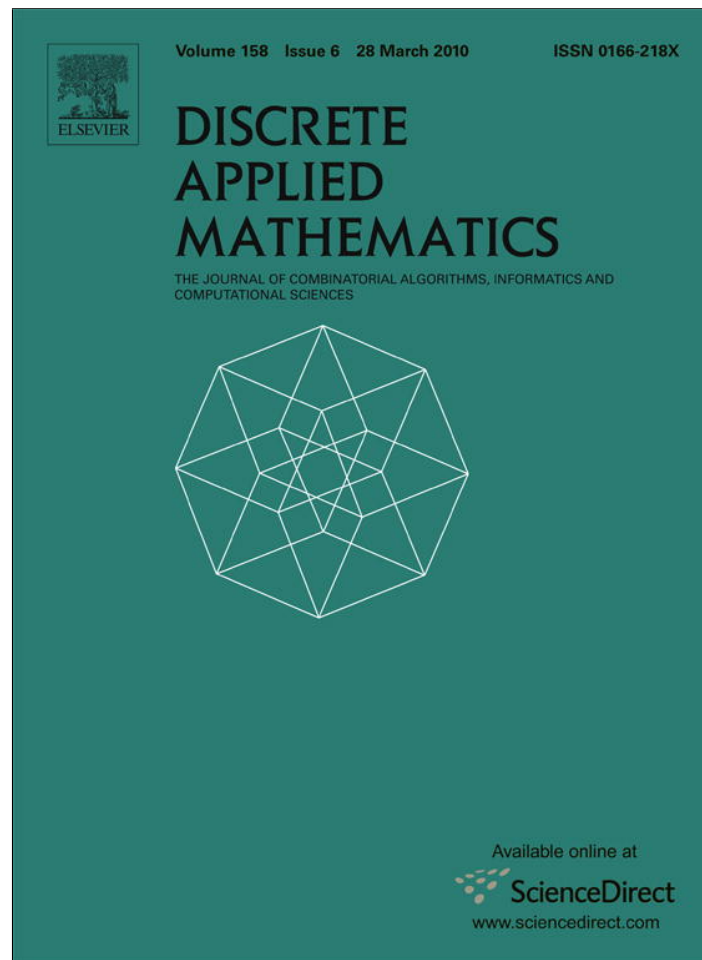


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

On shortest crucial words avoiding abelian powers

Sergey Avgustinovich^a, Amy Glen^{b,1}, Bjarni V. Halldórsson^c, Sergey Kitaev^{d,*}^a Sobolev Institute of Mathematics, Novosibirsk, Russia^b Department of Mathematics and Statistics, School of Chemical and Mathematical Sciences, Murdoch University, Perth, WA, 6150, Australia^c The Mathematics Institute, School of Science and Engineering, Reykjavík University, Kringlan 1, IS-103 Reykjavík, Iceland^d The Mathematics Institute, School of Computer Science, Reykjavík University, Kringlan 1, IS-103 Reykjavík, Iceland

ARTICLE INFO

Article history:

Received 22 June 2009

Received in revised form 6 November 2009

Accepted 24 November 2009

Available online 6 December 2009

Keywords:

Pattern avoidance

Abelian power

Crucial word

ABSTRACT

Let $k \geq 2$ be an integer. An *abelian k th power* is a word of the form $X_1X_2 \cdots X_k$ where X_i is a permutation of X_1 for $2 \leq i \leq k$. A word W is said to be *crucial* with respect to abelian k th powers if W avoids abelian k th powers, but Wx ends with an abelian k th power for any letter x occurring in W .

Evdokimov and Kitaev (2004) [2] have shown that the shortest length of a crucial word on n letters avoiding abelian squares is $4n - 7$ for $n \geq 3$. Furthermore, Glen et al. (2009) [3] proved that this length for abelian cubes is $9n - 13$ for $n \geq 5$. They have also conjectured that for any $k \geq 4$ and sufficiently large n , the shortest length of a crucial word on n letters avoiding abelian k th powers, denoted by $\ell_k(n)$, is $k^2n - (k^2 + k + 1)$. This is currently the best known upper bound for $\ell_k(n)$, and the best known lower bound, provided in Glen et al., is $3kn - (4k + 1)$ for $n \geq 5$ and $k \geq 4$. In this note, we improve this lower bound by proving that for $n \geq 2k - 1$, $\ell_k(n) \geq k^2n - (2k^3 - 3k^2 + k + 1)$; thus showing that the aforementioned conjecture is true asymptotically (up to a constant term) for growing n .

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Let $\mathcal{A}_n = \{1, 2, \dots, n\}$ 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 . For example, the word $V = 13243232323243$ over \mathcal{A}_4 contains the (non-trivial) fourth 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 $k = 3$ give us *abelian squares* (introduced by Erdős [1]) and *abelian 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 $1234 324$ 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.

* Corresponding author.

E-mail addresses: avgust@math.nsc.ru (S. Avgustinovich), amy.glen@gmail.com (A. Glen), bjarnivh@ru.is (B.V. Halldórsson), sergey@ru.is (S. Kitaev).¹ Current address: The Mathematics Institute, School of Computer Science, Reykjavík University, Kringlan 1, IS-103 Reykjavík, Iceland.

Let $\ell_k(n)$ denote the length of a minimal crucial word on n letters avoiding abelian k th powers. Evdokimov and Kitaev [2] have shown that $\ell_2(n) = 4n - 7$ for $n \geq 3$. Furthermore, Glen et al. [3] proved that $\ell_3(n) = 9n - 13$ for $n \geq 5$, and $\ell_3(n) = 2, 5, 11$, and 20 for $n = 1, 2, 3$, and 4 , respectively. The latter authors also showed that $\ell_k(n) \leq k^2n - (k^2 + k + 1)$ for $n \geq 4$ and $k \geq 3$, and they made the following conjecture.

Conjecture 1 ([3]). For any $k \geq 4$ and sufficiently large n , we have $\ell_k(n) = k^2n - (k^2 + k + 1)$.

The best known lower bound for $\ell_k(n)$, provided in [3], is $3kn - (4k + 1)$ for $n \geq 5$ and $k \geq 4$. The main result of this short note is the following improvement to the foregoing lower bound.

Theorem 1. For $k \geq 2$ and $n \geq 2k - 1$, $\ell_k(n) \geq k^2n - (2k^3 - 3k^2 + k + 1)$.

The above theorem shows in particular that **Conjecture 1** is true asymptotically (up to a constant term) for growing n . Note that for $k = 2$ the lower bound in **Theorem 1** cannot be improved as it gives the exact value of $\ell_2(n)$.

2. Proof of Theorem 1

For a crucial word X (with respect to abelian k th powers) on the n -letter alphabet \mathcal{A}_n , we let $X = X_i\Delta_i$ where Δ_i is the factor of minimal length such that Δ_i is an abelian k th power 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

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

where “ \subset ” means (proper) right factor (or suffix). In other words, for each $i = 2, 3, \dots, n$, 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:

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

for any $i = 2, 3, \dots, n - 1$. Furthermore, in the case of crucial words avoiding abelian k th powers, we write $\Delta_i = \Omega_{i,1}\Omega_{i,2} \dots \Omega_{i,k}$, 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 .

Lemma 2. Let $X = \Omega_{n,1}\Omega_{n,2} \dots \Omega'_{n,k} = Y_nY_{n-1} \dots Y_1$ be a crucial word on \mathcal{A}_n avoiding abelian k th powers ($k \geq 2$). For any i , $1 \leq i \leq n - 1$, if the factor $Y_{i+1}Y_i$ entirely belongs to the block $\Omega_{n,j}$ for some j (possibly $\Omega'_{n,k}$), then the letter i occurs at least k^2 times in X .

Proof. Since $\Delta_{i-1}(i - 1)$ and $\Delta_{i+1}(i + 1)$ are abelian k th powers (by definition), we deduce that the number of occurrences of the letter i in each of the factors Δ_{i-1} and Δ_{i+1} is divisible by k . Moreover, since Δ_i is an abelian k th power, the number of occurrences of the letter i in Δ_i is $(k - 1)$ modulo k , where $\Delta_i = Y_i\Delta_{i-1}$. For the preceding statements to hold, the number of occurrences of i in Y_i must be at least $k - 1$, and hence $(k - 1)$ modulo k . Similarly, the number of occurrences of i in Y_{i+1} must be at least 1 , and hence 1 modulo k . Therefore, the factor $Y_{i+1}Y_i$ contains at least k occurrences of the letter i . Finally, since $Y_{i+1}Y_i$ is a factor of the block $\Omega_{n,j}$ for some j (by hypothesis) and $Xn = \Omega_{n,1}\Omega_{n,2} \dots \Omega'_{n,k}n$ is an abelian k th power, we conclude that i occurs at least k^2 times in X . \square

Proof of Theorem 1. Let $X = \Omega_{n,1}\Omega_{n,2} \dots \Omega'_{n,k} = Y_nY_{n-1} \dots Y_1$ be a crucial word on \mathcal{A}_n avoiding abelian k th powers ($k \geq 2$). As an obvious corollary to **Lemma 2**, one has that if the factor $Y_{i+1}Y_i \dots Y_t$ entirely belongs to the block $\Omega_{n,j}$ for some j , then each of the letters $t, t + 1, \dots, i$ must occur at least k^2 times in X . Thus, out of all Y_i 's entirely belonging to a single block $\Omega_{n,j}$ only one of them can correspond to a letter occurring less than k^2 times in X , which gives in total at most k such letters. Additionally, at most $k - 1$ Y_i 's cannot be inside a single block $\Omega_{n,j}$, and theoretically the letters corresponding to such Y_i 's could occur less than k^2 times in X (we have no information on such Y_i 's). To summarize, X would have the minimal possible number of letters if X was of the form

$$X = Y_nZ_nY_{i_1}Y_{i_1-1}Z_{n-1}Y_{i_2}Y_{i_2-1} \dots Y_{i_{k-1}}Y_{i_{k-1}-1}Z_1,$$

where factor Z_i entirely belongs to one of $k - 1$ blocks $\Omega_{n,j}$, or the block $\Omega'_{n,k}$. Moreover, $Y_n, Y_{i_1-1}, Y_{i_2-1}, \dots, Y_{i_{k-1}-1}$ are the leftmost factors entirely inside the corresponding blocks $\Omega_{n,j}$ or $\Omega'_{n,k}$; $Y_{i_1}, Y_{i_2}, \dots, Y_{i_{k-2}}$ are factors having letters in two adjacent blocks $\Omega_{n,j}$ and $\Omega_{n,j-1}$ for some j ; and the factor $Y_{i_{k-1}}$ has letters in both $\Omega_{n,k-1}$ and $\Omega'_{n,k}$. All of the $2k - 1$ factors Y_i in the above decomposition correspond to letters occurring k times, whilst any other letter must occur at least k^2 times. Thus, for $n \geq 2k - 1$, the length of X is bounded from below by

$$k^2(n - (2k - 1)) + (2k - 1)k - 1 = k^2n - (2k^3 - 3k^2 + k + 1),$$

where the term $(2k - 1)k - 1$ comes from the fact that each letter out of those occurring less than k^2 times in X occurs at least k times there, except for the letter n which may occur only $k - 1$ times. \square

3. Further discussion

Even though the goal of this short paper was to prove an “asymptotic version” of [Conjecture 1](#) (rather than attempting to prove the conjecture itself or improving our lower bound as much as possible), we would like to share some ideas that may be helpful for further study of crucial words avoiding abelian powers, particularly with regards to proving/disproving [Conjecture 1](#) or improving the lower bound in [Theorem 1](#).

In our considerations here (and in [3]), we do not use the fact that crucial words avoid abelian powers. Thus, it is natural to study what we call *weakly crucial words* with respect to abelian powers (or other prohibitions); namely words that may contain abelian powers, but extending these words to the right by any letter x occurring within them creates an abelian power involving the rightmost x . For weakly crucial words, one has a nice hereditary property that erasing any letter in such a word gives a weakly crucial word. In particular, to prove lower bounds for the length of (regular) crucial words with respect to abelian k th powers, one could start by erasing all the letters occurring more than k^2 times, thus obtaining a weakly crucial word, and then work only with letters occurring less than k^2 times (the number of such letters, due to the proof of [Theorem 1](#), is at most $2k - 1$). It is likely that the best lower bounds for the lengths of weakly crucial words and regular crucial words with respect to abelian k th powers coincide.

Our final remark is as follows. Assuming the following conjecture is true, the lower bound in [Theorem 1](#) would be significantly improved to $k^2(n - 3) + 3k - 1 = k^2n - (3k^2 - 3k + 1)$.

Conjecture 2. For a (weakly) crucial word $X = Y_n Y_{n-1} \cdots Y_1$ on \mathcal{A}_n avoiding abelian k th powers ($k \geq 2$), there cannot exist 3 letters less than n that occur less than k^2 times in X .

Acknowledgements

The first author was partially supported by RFBR Grant No. 07-01-00248-a. The second and fourth authors were supported by the Icelandic Research Fund (Grant Nos. 09003801/1 and 060005012/3).

References

- [1] P. Erdős, Some unsolved problems, Magyar Tud. Akad. Mat. Kut. Int. Kozl. 6 (1961) 221–254.
- [2] 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.
- [3] A. Glen, B.V. Halldórsson, S. Kitaev, Crucial words for abelian powers, in: Proceedings of the 13th International Conference on Developments in Language Theory, DLT 2009, Stuttgart, Germany, June 30–July 3, 2009.