A Linear Upper Bound in Extremal Theory of Sequences

Martin Klazar
Department of Applied Mathematics
Charles University
Malostranské nám. 25
118 00 Praha 1
Czech republic

Abstract

An extremal problem considering sequences related to Davenport-Schinzel sequences is investigated in this paper. We prove that $f(x_1^i x_2^i..x_k^i x_1^i x_2^i..x_k^i, n) = O(n)$ where the quantity on the left side is defined as the maximum length m of the sequence $u = a_1a_2..a_m$ of integers such that 1) $1 \leq a_r \leq n$, 2) $a_r = a_s, r \neq s$ implies $|r - s| \geq k$ and 3) u contains no subsequence of the type $x_1^i...x_k^i x_1^i...x_k^i$ (x^i stands for $xx..x$ i-times).

Introduction

In this paper we shall deal with finite sequences consisting of some symbols. $S(u)$ denotes the set of all symbols occuring in the sequence $u = a_1a_2...a_m$, $|u|$ stands for its length ($|u| = m$) and $\|u\|$ stands for the cardinality of $S(u)$. If $a_i = a \in S(u)$ then a_i is called *a-letter*. That a_i *precedes* a_j (in u) means that $i < j$. We write $u \leq v$ and say that the sequence v *contains* the sequence u if some subsequence w of v differs from u only in names of symbols (in particular $|w| = |u|$ and $\|w\| = \|u\|$). Example: $u_1 = 1232454$ contains $v_1 = xxyy$ (here x, y were renamed to 2,4). The *k-regularity* of $u = a_1a_2...a_m$ means that $a_i = a_j, i \neq j$ implies $|i - j| \geq k$. Example: v_1 above is not 2-regular, u_1 is but is not 3-regular. The maximum length of sequences not containing a given (forbidden) sequence u is measured by the function

$$f(u, n) = \max\{|v| \mid u \not\leq v, \|v\| \leq n, v \text{ is } \|u\|\text{-regular}\}.$$

We shall show below that the maximum is defined correctly.

The first problem considering $f(u, n)$ was posed by Davenport and Schinzel [DS] in 1965 when they asked about the asymptotic growth of $F = f(ababa, n)$ and in general of $f(ababab\ldots, n)$. They proved $F = O(n \log n / \log\log n)$. This was later improved by Szemerédi [Sz] to $O(n \log^* n)$ (for any of those functions, $\log^* n$ is the minimum number of 2's in $2^{2^{\cdot^{\cdot^{2}}}}$ making this tower greater or equal to n) but whether $F = O(n)$ remained unclear. Hart and Sharir [HS] answered this question negatively: $F = \Theta(n\alpha(n))$ where $\alpha(n)$ is the functional inverse to the Ackermann function and goes to infinity but very slowly. Recently both sharp upper and lower bounds on the functions $f(ababab\ldots, n)$ were found [ASS], [S].

The aim of this paper is to give (linear) upper bounds for extremal functions of forbidden sequences $a(i, k) = x_1^i x_2^i ... x_k^i x_1^i x_2^i ... x_k^i$. Here x_j are k distinct symbols and x^i stands for $xx..x$ i times. The main result is the estimate $f(a(i, k), n) = O(n)$.The sequences with a linear upper bound form the set

$$Lin = \{u \mid f(u, n) = O(n)\}$$

and our result may be reformulated as $x_1^i x_2^i ... x_k^i x_1^i x_2^i ... x_k^i \in Lin$. It generalizes the result $a^i b^i a^i b^i \in Lin$ achieved in [AKV]. Finding all elements of Lin seems to be an interesting and not an easy problem (see concluding remarks).

The linearity of $a(i, k)$ is derived from two statements—Theorem A and Theorem B—which are perhaps of some independent interest.

1)The symbols a, b are called *l-good* in the sequence u if at most l b-letters lie between the first a-letter and the last a-letter or vice versa. A k-regular sequence u is called l-mixed if there are two l-good symbols a, b among every k elements of $S(u)$.Theorem A states $|u| = O(\|u\|)$ for such u. In fact we prove something stronger.

2) We write $u \leq\leq v$ if v contains u in all possible ways, i.e. if $u \leq w$ for any w obtained from v by restricting v on some $\|u\|$ symbols. Suppose i, k are given. Theorem B says that $a(i, k) \leq u$ whenever $\|u\|$ is large and $a^l b^l a^l b^l \leq\leq u$ for large l.

The plan of the paper is as follows. In the first section we recall definitions and introduce several new ones. Then we derive $a(i, k) \in Lin$ from 1) and 2) and prove several auxiliary but useful lemmas. In the second section results related to 1) are proved. The proof of 2), which is technical, may be found in the third section.

Section 1

Definition 1.1 *Suppose that there are two* l-good *symbols* a, b *among every* m *symbols of a sequence* u. *This means that at most* l b*-letters occur between the first* a*-letter and the last* a*-letter or vice versa. In such a situation* u *is called* (m, l)-mixed $(m \geq 2, l \geq 0)$.

Definition 1.2 *A sequence* u *is called* weakly (m, l)-mixed $(m \geq 2, l \geq 0)$ *if the sequence* u^* *is* (m, l)*-mixed. The sequence* u^* *is obtained from* u *by deleting all* l-outer *letters. A letter in* u *is* l-outer *if belongs to the first* l *or to the last* l x*-letters for some* $x \in S(u)$.

In case u is k-regular and $m = k$ we shall say simply that u is (weakly) l-mixed.

Definition 1.3 *We say that a letter* a_i *may be* c*-deleted from a* k*-regular sequence* $u = a_1 a_2 ... a_m$ *if it is possible to delete* a_i *with at most* $c - 1$ *other letters in such a way that the remaining sequence is still* k*-regular.*

It may be easily seen that any letter may be 2-deleted from any 2-regular sequence. It is not the case for three- and more regular sequences: in the sequence

$$...xyzxyzxyzayxzyxzyxz...$$

which is 3-regular it is impossible to delete the single a-letter and to preserve 3-regularity without deleting many x, y, z-letters. We shall see below that under the condition of not containing a forbidden sequence c-deleting is possible for general k-regularity.

Definition 1.4 *We define a greedy algorithm* $A(k)$ *that choses from a given sequence* $u = a_1 a_2 ... a_m$ *a* k*-regular subsequence* v *in the following way. At first* $v = a_1$ *and* $i = 1$. *Let* j *be a minimum integer with respect to* $j > i, va_j$ *is* k*-regular. If such* j *exists then we put* $v = va_j$, $i = j$ *and repeat. Otherwise the algorithm terminates. Observe that* $\|I\| \leq k - 1$ *for any interval* I *in* u *such that* $I \cap v = \emptyset$.

Theorem A *Let* $m \geq k \geq 2, l \geq 0$ *be integers and let* $u, k \geq \|u\|$ *be a sequence. Suppose the sequence* v *is* k*-regular, weakly* (m, l)*-mixed and* $u \not\leq v$. *Then* $|v| \leq c\|v\|$ *where the constant* $c = c(k, l, m, u)$ *depends on the indicated parameters.*

Recall that $u \leq\leq v$ means that for any $S \subset S(v), |S| = \|u\|$ there is a u-copy $\bar{u}$ in v such that $S(\bar{u}) = S$.

Theorem B *For all positive integers* i *and* k *there exist integers* n *and* l *such that* $a(i, k) \leq u$ *whenever* $\|u\| \geq n$ *and* $x^l y^l x^l y^l \leq\leq u$.

To derive the main result from Theorem A and Theorem B one more lemma is needed.

Lemma 1.5
a) If $x^{2l+1} y^{2l+1} x^{2l+1} y^{2l+1} \leq\leq u$ *then* u *is not weakly* $(\|u\|, l)$*-mixed.*
b) If u *is not weakly* $(\|u\|, l)$*-mixed then* $x^m y^m x^m y^m \leq\leq u$ *where* $m = [(l + 1)/2]$.
c) If $a(2l + 1, k) \leq u$ *then* u *is not weakly* (k, l)*-mixed.*

Proof: Both a) and c) follow immediately from definitions. We prove b). Suppose u is not weakly $(\|u\|, l)$-mixed and $a, b \in S(u)$ are two distinct symbols. The violation of the weak mixness on a, b means that there are two intervals v and w in u such that v starts and finishes with an a-letter, w starts and finishes with an b-letter, the first letter of v precedes the first letter of w, v contains at least $l + 1$ b-letters, w contains at least $l + 1$ a-letters, neither the first nor the last letter of v belongs to the l-outer a-letters and similarily for w. We split $v = v_1 v_2$ in such a way that either v_i contains m b-letters. If there are at least m a-letters in the intersection $v_1 \cap w$ then l b-letters lie before w, m a-letters lie in $v_1 \cap w$, m b-letters lie in v_2 and finally

l a-letters lie after v—$b^m a^m b^m a^m \leq u$. Otherwise there are l a-letters before v, m b-letters lie in v_1, the remaining at least m a-letters lie in w and finally l b-letters lie after w—$a^m b^m a^m b^m \leq u$.

♡

Theorem 1.6 (main result) $a(i,k) \in Lin$.

Proof: Let v be a k-regular sequence not containing $a(i,k)$. We show that v satisfies the hypothesis of Theorem A for parameters $m = n(i,k), k = k, l = 2l(i,k), u = a(i,k)$ where $l(i,k)$ and $n(i,k)$ are the integers of Theorem B. Suppose on the contrary that v is not weakly (m,l)-mixed. Then according to Lemma 1.5 b) there exists a subsequence w of v such that $\|w\| = n(i,k)$ and $x^{l(i,k)}y^{l(i,k)}x^{l(i,k)}y^{l(i,k)} \leq\leq w$. Theorem B yields $a(i,k) \leq w$ which is a contradiction. Thus, according to Theorem A, $|v| \leq c\|v\|, c = c(i,k)$.

♡

In the rest of Section 1 we prove four auxiliary lemmas which will be needed in the following sections.

Lemma 1.7 *Consider a generalization of the function* $f(u,n)$

$$f(u,n,l) = max\{|v| \mid u \not\leq v, \|v\| \leq n, v \text{ is } l\text{-regular}\}.$$

a) $f(u,n,l)$ *is finite for any sequence* u *and all integers* $n \geq 1, l \geq \|u\|$.
b) $f(u,n,l) \leq f(u,n,k) \leq (1 + f(u,l-1,k))f(u,n,l)$ *for any sequence* u *and all integers* $n \geq 1, l > k \geq \|u\|$.

Proof of a): If $1 \leq n < l$ then trivially $f(u,n,l) \leq n$. Let $n \geq l$. We prove $u \leq v$ whenever v is l-regular, $\|v\| \leq n$ and $|v| \geq \|u\|(1 + \binom{n}{\|u\|})(|u| - 1))$. Let $v = v_1 v_2 .. v_c w$ where $|v_i| = \|v_i\| = \|u\|$, $c = \binom{n}{\|u\|}(|u| - 1) + 1$. The Dirichlet principle implies $S(v_{i_1}) = S(v_{i_2}) = ... = S(v_{i_{|u|}})$ for some $|u|$ indices $1 \leq i_1 < i_2 < ... < i_{|u|} \leq c$. Thus $u \leq v_{i_1} v_{i_2} ... v_{i_{|u|}}$.
Proof of b): The first inequality is obvious. We prove the second one. Suppose $v = a_1 a_2 ... a_m$ is k-regular, $u \not\leq v$ and $\|v\| \leq n$. It suffices to apply the greedy algorithm $A(l)$ on v. We obtain an l-regular subsequence v^* of v such that $|v| \leq |v^*|(1 + f(u,l-1,k))$ because $\|I\| \leq l - 1$ for any interval I in v omitted by $A(l)$. Hence $|v| \leq (1 + f(u,l-1,k))f(u,n,l)$.

♡

Lemma 1.8 *Suppose* v *is* k*-regular,* $v \not\geq u$ *and* $k \geq \|u\|$*. Then any letter may be* $c = c(k,u)$*-deleted from* v*.*

Proof: One can assume $|v| \geq 2k - 1 + f(u,3k-3,k)$. Consider the partition $v = v_1 v_2 v_3 v_4 v_5$ where $|v_2| = \|v_2\| = |v_4| = \|v_4\| = k-1$, the letter a_i choosen to be deleted occurs in v_3 and $|v_3| = f(u,3k-3,k)+1$. Hence $\|v_3\| \geq 3k - 2$ and there are $k - 1$ symbols $S \subset S(v_3)$ such that $S \cap (\{a_i\} \cup S(v_2) \cup S(v_4)) = \emptyset$. We choose such $k - 1$ letters $b_1, b_2, ..., b_{k-1}$ in v_3 that $\{b_1, b_2, ..., b_{k-1}\} = S$ and delete from v_3 all other letters (i.e. we delete exactly $f(u,3k-3,k) + 2 - k$ letters). What remains is still a k-regular sequence.

♡

Lemma 1.9 *Suppose* v *is* k*-regular and weakly* l*-mixed. Then any letter may be* $c = c(k,l)$*-deleted from* v*.*

Proof: According to Lemma 1.5 c) $a(2l+1,k) \not\leq v$ and the previous lemma applies.

♡

Lemma 1.10 *Let* $k, l \geq 2$ *be integers and let* u *be a* k*-regular sequence. Then there exists a subsequence* v *of* u *such that*
1) v *is* k*-regular,*
2) *between any two* x*-letters in* v *there are at least* $l - 1$ x*-letters in* u *and*
3) $|v| \geq |u| \frac{1}{k^2 l(l-1) + kl}$.

Proof: Let $a_1^x, a_2^x, ...$ be all x-letters in u numerated from left to right for all $x \in S(u)$. The sequence u^* is defined as consisting of those a_i^x that $i \equiv 1(mod\ l)$. The desired sequence v is obtained from u^* by the greedy algorithm $A(k)$. The sequence v posesses obviously properties 1) and 2). It remains to prove that v is sufficiently long. We define S as the set of all intervals in u^* into which v divides u^*. Let $I \in S$. We decompose $I = J_I K_I = I_1 I_2 ... I_p K_I, |I_i| = k, |K_I| \leq k-1$. The definition of $A(k)$ yields $\|I\| \leq k-1$. Thus in any I_i some symbol repeats. The construction of u^* implies that there are another $l-1$ letters of that symbol between those two letters in u. But u is k-regular so together there are at least $p(kl-1-(k-2)) = p(k(l-1)+1)$ letters in $u \backslash u^*$ between the first and the last letter of J_I. If we denote the set of those letters as $R_I \subset (u \backslash u^*)$ then

$$|J_I| = pk \leq |R_I| \frac{k}{k(l-1)+1}.$$

The union L of all J_I and the union M of all K_I, $I \in S$ form a partition $u^* = v \cup L \cup M$. Obviously $|u^*| \geq \frac{1}{l}|u|$ and $|u \backslash u^*| \leq \frac{l-1}{l}|u|$. Thus

$$|L| \leq \frac{k}{k(l-1)+1} \Sigma_{I \in S} |R_I| \leq \frac{k}{k(l-1)+1} |u \backslash u^*| \leq \frac{k}{k(l-1)+1} \frac{l-1}{l} |u|.$$

Further

$$|M \cup v| = |u^*| - |L| \geq \frac{1}{l}|u| - \frac{k}{k(l-1)+1} \frac{l-1}{l}|u| = \frac{1}{kl(l-1)+l}|u|.$$

The mapping that maps K_I on the precedessor (in u^*) of the first letter of I is an injection from $\{K_I \mid I \in S\}$ to v and $|K_I| \leq k-1$ for all I. Therefore $k|v| \geq |M \cup v|$ and

$$|v| \geq |u| \frac{1}{k^2 l(l-1) + kl}.$$

♡

Lemma 1.8 and Lemma 1.10 have interesting consequences for the structure of *Lin*—see the concluding remarks.

Section 2

In this section Theorem A will be proved.

Theorem 2.1 *Any k-regular and 1-mixed sequence w satisfies $|w| \leq (k^2+k)\|w\| + k$.*

Proof: Let us take such a w. Suppose the pair of symbols $a, b \in S(w)$ is 1-good. There are exactly five possible configurations in which all a-letters and b-letters in w may lie (we restrict w to $\{a, b\}$):

a) $a...ab...b$ b) $a...ab...ba...a$ c) $a...ab...bab...ba...a$ d) $a...ab...bab...b$ and e) $a...aba...ab...b$.

Here $a...a$ stands for $a^i, i \geq 1$ and the first a-letter is supposed to precede the first b-letter. The situation c) is denoted as $a > b$ and the central a-letter as $a(b)$. In order to estrimate $|w|$ we estimate the number p in the splitting $w = w_1 w_2 ... w_p v$ where $|w_i| = k$ and $|v| \leq k-1$. Some two symbols $x, y \in S(w_i)$ must be 1-good for any w_i. Thus (see the configurations a)–e)) any w_i contains the first letter of some symbol or contains the last letter of some symbol or contains an element of the set $M = \{a(b) \mid a, b \in S(w), a > b\}$ (in the worst c) case). Obviously

$$|w| \leq pk + k - 1 < k(p+1) \leq k(2\|w\| + |M| + 1)$$

and it suffices to estimate the size of M.

For this purpose we define a mapping F that maps M into the set of all last letters in w. We put $F(a_0) =$ the last element of $\{b_0 \mid b_0$ is the last b-letter$, a > b, a_0 = a(b)\}$. We prove $|F^{-1}(b_0)| \le k-1$ for any last b-letter b_0, $b \in S(w)$. If it is done we conclude —

$$|w| \le k(2\|w\| + |M| + 1) \le k(2\|w\| + (k-1)\|w\| + 1) = k(k+1)\|w\| + k.$$

Suppose on the contrary that $F(\bar{x}_1) = F(\bar{x}_2) = ... = F(\bar{x}_k) = \bar{y}$ where $\bar{x}_i$ is an x_i-letter, $\bar{y}$ is the last y-letter, x_i and y are $k+1$ different symbols, $x_i > y$ and $\bar{x}_i = x_i(y)$. Some two symbols of $x_1, x_2, ..., x_k$ must be 1-good. Thus $x_r > x_s$ for some r, s because only c) can occur. It may be easily checked that $\bar{x}_r = x_r(y) = x_r(x_s)$ which is a contradiction with the definition of F because $\bar{y}$ precedes the last x_s-letter.

♡

The estimate in Lemma 2.1 may be slightly improved:
a) Even $|M| \le (k-1)(\|w\| - 1)$ because F^{-1}(the end of w) $= \emptyset$. Using this idea we obtain $|M| \le (k-1)\|w\| - ((k-1) + (k-2) + ... + 1) = (k-1)(\|w\| - \frac{k}{2})$.
b) $2\|w\|$ in $|w| \le k(2\|w\| + |M|) + |v|$ may be replaced by $2\|w\| - |v|$ (any letter of v is the last letter of some symbol).
Thus one can do better a bit: $|w| \le k(k+1)\|w\| - \frac{1}{2}k^2(k-1)$, $(\|w\| \ge k)$.

Theorem 2.2
a) Any k-regular and l-mixed sequence w satisfies $|w| \le k^2(k+1)l(k(l-1)+1)\|w\|$.
b) Any k-regular and weakly l-mixed sequence w satisfies $|w| \le c\|w\|$ where $c = c(k,l)$.

Proof of a): We apply on w Lemma 1.10. The obtained subsequence v is k-regular and according to 2) of Lemma 1.10 also 1-mixed. The upper bound on $|w|$ is consequence of the previous theorem and of 3) of Lemma 1.10.
Proof of b): According to Lemma 1.9 any l-outer letter of w can be $d = d(k,l)$-deleted. The remaining sequence v is k-regular and l-mixed. According to a) $|w| \le 2dl\|w\| + |v| \le 2dl\|w\| + k^2(k+1)l(k(l-1)+1)\|v\|$ and b) follows.

♡

It is possible (but it costs some effort) to prove Theorem 2.2 a) like Theorem 2.1 and to avoid using Lemma 1.10. A substantially better constant in the linear upper bound on $|w|$ is obtained in this way — $O(k^2 l)$ instead of $O(k^4 l^2)$.

Theorem A *Let $m \ge k \ge 2$, $l \ge 0$ be integers and let u be a sequence such that $k \ge \|u\|$. Suppose v is weakly (m,l)-mixed, k-regular and $u \not\le v$. Then $|v| \le c\|v\|$ where $c = c(k,l,m,u)$.*

Proof: According to Lemma 1.7 b) (now $u \not\le v$ is used) it does not matter (up to the constant in O) what regularity number k is choosen on beginning. For an m-regular sequence v the statement reduces to the previous theorem b).

♡

Section 3

Here Theorem B will be proved.

Denote by $R(r,n)$ the minimal N such that any colouring of two–term subsets of an N–term set S by r colours yields a monochromatic n–term subset of S (see [GRS]).

Lemma 3.1 *Suppose that $x^l y^l x^l y^l$ is a subsequence of $w = w_1 w_2 w_3$. Then y^l is a subsequence of w_1 or x^l is a subsequence of w_2 or y^l is a subsequence of w_3.*

Proof: Obvious. ♡

Lemma 3.2 *Suppose w contains $m+1$ x-letters and $m+1$ y-letters $x_0, x_1, ..., x_m, y_0, y_1, ..., y_m$ enumerated from left to right and x_0 precedes y_0. Let $m = 2i(k-1)$. Then one of a)–d) holds.*
a) $x_{(a+1)i}$ precedes y_{ai} for some $a \in \{0, 1, .., 2k-2\}$
b) $y_{(a+1)i}$ precedes x_{ai} for some $a \in \{0, 1, .., 2k-2\}$
c) x_{a_1} precedes y_{a_1} precedes $x_{a_2}...x_{a_k}$ precedes y_{a_k} for some indices $0 \le a_1 < a_2 < \ldots < a_k \le 2k-2$
d) y_{a_1} precedes x_{a_1} precedes $y_{a_2}...y_{a_k}$ precedes x_{a_k} for some indices $0 \le a_1 < a_2 < \ldots < a_k \le 2k-2$

Proof: If neither a) nor b) holds for any $a \in \{0, 1, ..., 2k-2\}$ then x_0, y_0 precede x_i, y_i precede...precede $x_{(2k-2)i}, y_{(2k-2)i}$ and c) or d) occurs for some k indices.

♡

The number of all mutual configurations (with respect to preceding) of the x-letters and y-letters in the situation in Lemma 3.2 is equal to $p_m = \binom{2m+1}{m+1}$. Suppose that there are k symbols ${}^1x, {}^2x, ..., {}^kx$ in Lemma 3.2 instead of two and that $m = 2(i-1)(ik-1)$. Suppose further that ${}^1x_0$ precedes ${}^2x_0$ ...precedes ${}^kx_0$ and that the configuration of ${}^ax$ and ${}^bx$ is the same for all $\binom{k}{2}$ pairs $1 \le a < b \le k$. Then it is not difficult to see that either ${}^1x^i\ {}^2x^i ... {}^kx^i$ or ${}^kx^i\ {}^{k-1}x^i ... {}^1x^i$ is a subsequence of w (apply on that configuration Lemma 3.2).

Theorem B *For any positive integers i and k there exist integers n and l such that $a(i,k) \le w$ whenever $\|w\| \ge n$ and $x^l y^l x^l y^l \le\le w$.*

Proof: It suffices to put $n = R(2p_m^2, h), l = m+1 = 1+2(i-1)(hi-1), h = 6R(2p_r, k)-3, r = 2(i-1)(ki-1)$. We split $w = w_1 w_2$ so that any $x \in S(w)$ has l letters in both w_1, w_2 (we split w in the last l-th letter). The $\binom{n}{2}$ pairs $\{x, y\} \subset S(w)$ are coloured by $2p_m^2$ colours according to their configurations in w_1 and w_2 and according to the position of their first letters in w_1 and w_2. The h-term monochromatic subset $S \subset S(w)$, $S = \{{}^1x, {}^2x, ..., {}^hx\}$, ${}^1x_0$ precedes ${}^2x_0$...precedes ${}^hx_0$ (in w_1, in w_2 they may be in the opposite order) ensured by Ramsey theorem implies (according to the consideration above) that ${}^1x^i\ {}^2x^i ... {}^hx^i$ or ${}^hx^i\ {}^{h-1}x^i ... {}^1x^i$ is a subsequence in both w_1, w_2. One can suppose ${}^1x^i\ {}^2x^i ... {}^hx^i\ {}^hx^i ... {}^2x^i\ {}^1x^i$ is a subsequence in w (one of the remaining three cases is treated similarily and in remaining two we are done). We denote that subsequence as K and define $g = \frac{h-1}{2}$. We split w on three parts $w = v_1 v_2 v_3$ so that the beginning of v_2 is one of the ${}^{g+1}x$-letters in the first half of K and the end of v_2 is one of the ${}^{g+1}x$-letters in the second half of K. Now Lemma 3.1 on every pair of symbols $({}^1x, {}^hx), ({}^2x, {}^{h-1}x), ..., ({}^gx, {}^{g+2}x)$ is applied (${}^ix$ with small left upper indices play the role of x in Lemma 3.1 and those with large ones play the role of y). One of the cases of Lemma 3.1 occurs for at least $\lceil \frac{g}{3} \rceil = R(2p_r, k)$ pairs $({}^ix, {}^{h-i+1}x), i \in I \subset \{1, ..., g\}$. One can suppose it is the second one (remaining two are treated analogously). It means that ${}^ix^{r+1}$ (even ${}^ix^l$) is a subsequence of v_2 for any $i \in I, |I| \ge R(2p_r, k)$. We colour pairs of $\{\ {}^ix \mid i \in I\}$ by $2p_r$ colours (configuration and position of their first letters) and apply Ramsey theorem again. According to the above consideration applied on the k-term monochromatic set of symbols ${}^{i_1}x^i\ {}^{i_2}x^i ... {}^{i_k}x^i$ or ${}^{i_k}x^i\ {}^{i_{k-1}}x^i ... {}^{i_1}x^i$ is a subsequence of v_2 for some (k symbols) $\{i_1 < i_2 < \ldots < i_k\} \subset I$. These letters together with the first and with the last fourth of K (which lie in v_1 and in v_3) give $a(i,k) \le w$.

♡

Concluding remarks

Our proof of $a(i,k) \in Lin$ is just a step in finding all elements of Lin. For two–symbol sequences the results $a^i b^i a^i b^i \in Lin$ ([AKV]) and $ababa \notin Lin$ ([HS]) yields the equivalence $u \in Lin$ iff $ababa \not\le u$. This is not the case for a general sequence u because the construction in [WS] realizing the lower bound $f(ababa, n) = \Omega(n.\alpha(n))$ by segments in the plane proves also implicitly $u_1 = abcbadadbcd \notin Lin$ (and $ababa \not\le u_1$). Define the function

$$f(n) = \max\{|v| \mid v \in Lin, \|v\| \le n, v \text{ is 2-regular}\}.$$

Clearly $f(n) \leq f(ababa, n) = O(n\alpha(n))$.

Problem 1 Does $f(n) = O(n)$ hold?

A simple consequence of Lemma 1.8 is that $ua \in Lin$ inplies $ua^k \in Lin$ for any sequence u, symbol a and natural number k (similarily for au). In the same manner we can easily derive from Lemma 1.10 that $ua^2v \in Lin$ implies $ua^kv \in Lin$ for any sequences u, v, natural number k and symbol a. Analogously if we consider general sequences, not only linear—the change of exponents in the described manner does not change the growth rate of $f(u, n)$.

Problem 2 Does $uav \in Lin$ imply $ua^2v \in Lin$ in any case? Does in general $f(ua^2v, n) = O(f(uav, n))$ hold? Except, of course, for uav without repetitions when $f(uav, n)$ is constant.

Problem 3 Does $babcbcac \in Lin$ hold?

Problem 4 How many minimum (to $\leq$) nonlinear sequences there are? From the above comment it follows that beside $ababa$ there is at least one such an element.

Acknowledgments

The autor is grateful to P.Valtr, M.Loebl and J. Nešetřil for their useful consultations.

References

[AKV] R.ADAMEC, M.KLAZAR and P.VALTR, Generalized Davenport-Schinzel sequences with linear upper bound, *Discrete Math.* 108 (1992), 219–229

[ASS] P.AGARWAL, M.SHARIR and P.SHOR Sharp upper and lower bounds on the length of general Davenport-Schinzel sequences, *J. of Comb. Th.* A 52 (1989), 228–274

[DS] H.DAVENPORT and A.SCHINZEL A combinatorial problem connected with differential equations I and II, *Amer. J. Math.* 87 (1965), 683-694 and *Acta Arithmetica*, 17 (1971), 363–372

[GRS] R.L.GRAHAM, B.L.ROTHSCHILD and J.H.SPENCER Ramsey Theory, New York 1980, John Wiley

[HS] S.HART and M.SHARIR Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes *Combinatorica* 6 (1986), 151–177

[S] M.SHARIR Almost linear upper bounds on the length on general Davenport-Schinzel sequences, *Combinatorica* 7 (1987), 131–143

[Sz] E.SZEMERÉDI On a problem by Davenport and Schinzel, *Acta Arithmetica* 15 (1974), 213–224

[WS] A.WIERNIK and M.SHARIR Planar realization of nonlinear Davenport-Schinzel sequences by segments, *Discrete Comp. Geom.* 3 (1988), 15–47