Generalized Davenport–Schinzel sequences: results, problems, and applications

Martin Klazar*

Abstract

We survey in detail extremal results on Davenport–Schinzel sequences and their generalizations, from the seminal papers of H. Davenport and A. Schinzel in 1965 to present. Geometric and enumerative applications, generalizations to colored trees, and generalizations to hypergraphs are discussed. Eleven illustrative examples are given with proofs. Eighteen open problems are posed.

1 Introduction

DS sequences. Why a survey on Davenport–Schinzel sequences? (We shall abbreviate this term as DS sequences.) Two combinatorially orientated survey articles have appeared, Stanton and Dirksen [51] and Klazar [27]. Both are now outdated. Sharir and Agarwal [48] and Agarwal and Sharir [3] focus on geometric applications. Survey and historic sections can be found also in [25], [33], and [57] but the main focus of these works is different. Here we give a more detailed and concise treatment, pose many open problems, and present combinatorially interesting and often unexplored generalizations of the original problem. We concentrate on its extremal side but we do discuss related enumerative aspects.

In Section 1 the classical Davenport–Schinzel's extremal function $N_s(n)$ is introduced and several simple bounds on it are proved. Section 2 surveys extremal results on DS sequences obtained in the early period, before the

*Supported by the project LN00A056 of the Ministry of Education of the Czech Republic.

superlinearity of $N_5(n)$ was discovered in [19]. Section 3 explains superlinear bounds on lengths of DS sequences. Section 4 deals with the generalization of DS sequences to any forbidden sequence introduced in [1]. Section 5 describes applications of generalized DS sequences in combinatorial geometry and enumerative combinatorics, and their further generalization to colored trees and hypergraphs.

A *DS sequence* corresponding to the parameter s or s-DS sequence for short, where s is a positive integer, is any finite sequence $u = a_1a_2\dots a_l$ over a fixed infinite alphabet A satisfying two conditions:

1. $a_i \neq a_{i+1}$ for every $i = 1, 2, \dots, l-1$, that is, u contains no immediate repetition, and
2. there do not exist s indices $1 \le i_1 < i_2 < \dots < i_s \le l$ such that $a_{i_1} = a_{i_3} = a_{i_5} = \dots = a$, $a_{i_2} = a_{i_4} = a_{i_6} = \dots = b$, and $a \neq b$, that is, u contains no alternating subsequence of length s.

We write DS_s to denote the set of s-DS sequences. What are the elements of A is not important. We assume that we have in A all positive integers $1, 2, \dots$, letters $a, b, c, d, \dots$, and perhaps some other symbols. The set A^* contains all finite sequences over A. Two sequences from A^* of the same length differing only by an injective renaming of symbols, for example 121331 and $2c2aa2$, are called *isomorphic*. For our purposes isomorphic sequences are identical. Every sequence is isomorphic to a unique *normal* sequence u which uses as symbols all integers $1, 2, \dots, n$ for some n and the first occurrences of $1, 2, \dots, n$ in u, if we scan u from left to right, come in this order.

Example 1. There exist exactly ten normal 4-DS sequences u using at most 3 symbols:

$$u = \emptyset, 1, 12, 121, 1213, 12131, 123, 1231, 1232, \text{and } 12321.$$

$\square$

$\mathbf{N}$ and $\mathbf{N}_0$ denote the sets $\{1, 2, \dots\}$ and $\{0, 1, 2, \dots\}$. We write $[n]$, $n \in \mathbf{N}$, for the set $\{1, 2, \dots, n\}$ and $[a, b]$, $a, b \in \mathbf{N}$, $a \le b$, for the set $\{a, a+1, \dots, b\}$. For two functions $f, g : \mathbf{N} \to \mathbf{R}$, the asymptotic notation $f \ll g$, which is synonymous to the $f = O(g)$ notation, means that $|f(n)| < c|g(n)|$ holds for every $n > n_0$ and a constant $c > 0$. The subscripts, such as $f \ll_k g$, indicate that c depends only on the mentioned parameters. Notation $f = o(g)$ means that $f(n)/g(n)$ tends to 0 if $n \to \infty$.

Extremal function $\mathbf{N_s(n)}$. For a sequence $u = a_1 a_2 \dots a_l$ over A, we use $|u|$ to refer to its lentgh l. $S(u) = \{a_1, a_2, \dots, a_l\}$ is the set of symbols used in u and $\|u\| = |S(u)|$ is their number. Obviously, always $|u| \ge \|u\|$. We define, for $s, n \in \mathbf{N}$,

$$N_s(n) = \max\{|u| :\ u \in \mathrm{DS}_s \ \&\ \|u\| \le n\}. \tag{1}$$

The function $N_s(n)$ measures the maximum length of s-DS sequences using at most n symbols. It is not hard to see that, for every n and s, $N_s(n) < \infty$, $N_1(n) = 0$, $N_2(n) = 1$, $N_3(n) = n$, $N_s(1) = 1$ (for $s > 1$), and $N_s(2) = s - 1$.

Example 2 ([12]). We bound $N_s(n)$ in a rough way, determine $N_4(n)$ precisely, and bound $N_5(n)$ in a finer way.

We begin with the bound

$$N_s(n) \le \binom{n}{2}(s-2) + 1 \tag{2}$$

which holds for all n and $s \ge 2$. Suppose a sequence $u = a_1 a_2 \dots a_l$ satisfies condition 1 (no immediate repetition) and $\|u\| \le n$, but its length l exceeds the bound. Then among the $l - 1 \ge \binom{n}{2}(s-2) + 1$ sets $\{a_i, a_{i+1}\}$ some $s - 1$ must coincide (by the pigeonhole principle), which produces an alternating subsequence of length s in u. This proves (2).

Let us prove that

$$N_4(n) = 2n - 1.$$

The sequences $u = 1\ 2\ \dots\ (n-1)\ n\ (n-1)\ \dots\ 2\ 1$ show that $N_4(n) \ge 2n-1$. The opposite inequality follows by induction on n. Certainly $N_4(1) = 1$. In every $u \in \mathrm{DS}_4$ with $\|u\| \le n$ and $|u| = N_4(n)$ a symbol x appears only once: any of the symbols appearing inbetween the closest repetition has this property. Delete x and, if necessary, one of its neighbours (to preserve condition 1). The sequence v obtained is a 4-DS sequence and $\|v\| \le n - 1$. By induction, $N_4(n) = |u| \le |v| + 2 \le N_4(n-1) + 2 = 2(n-1) - 1 + 2 = 2n - 1$.

We prove that

$$N_5(n) \ll n \log n.$$

Let $u \in \mathrm{DS}_5$ with $\|u\| \le n$ and $|u| = N_5(n)$. Note that the maximum length implies $\|u\| = n$. For $x \in S(u)$ we denote $k(x)$ the number of appearances of

x in u. Only the first and last one may have equal neighbours, because for middle appearances of x they would create the forbidden 5-term alternating subsequence. By deleting at most $k(x)+2$ elements from u we get rid of all appearances of x and preserve condition 1. The sequence v obtained is a 5-DS sequence and $\|v\| \le n-1$. Thus $N_5(n) = |u| \le |v| + k(x) + 2 \le N_5(n-1) + k(x) + 2$. Summing these inequalities over all $x \in S(u)$, we obtain the inequality $nN_5(n) \le nN_5(n-1) + N_5(n) + 2n$, which we rewrite as

$$\frac{N_5(n)}{n} - \frac{N_5(n-1)}{n-1} \le \frac{2}{n-1}.$$

Summing these inequalities for $2, 3, \ldots, n$ leads to the bound $N_5(n) \le n(1 + 2(1^{-1} + 2^{-1} + \cdots + (n-1)^{-1})) \ll n \log n$. $\square$

2 The early period

Geometric origin of $\mathbf{N_s(n)}$. Davenport and Schinzel introduced the sequences which now bear their names in 1965 in [12]. They were led to $N_s(n)$ by the following geometric problem. Suppose $f_1, \ldots, f_n : \mathbf{R} \to \mathbf{R}$ are n continous functions such that the equation $f_i(x) = f_j(x)$ has for $i \ne j$ at most s solutions $x \in \mathbf{R}$. In other words, the graphs of any two functions intersect in at most s points. The real line then splits uniquely into l nonempty open intervals $I_1 = (-\infty, a_1), I_2 = (a_1, a_2), I_3 = (a_2, a_3), \ldots, I_l = (a_{l-1}, \infty)$ so that the pointwise minimum function $f(x) = \min_{j=1\ldots n} f_j(x)$ coincides on each I_i with a unique function $f_{j(I_i)}$, $1 \le j(I_i) \le n$, and $j(I_i) \ne j(I_{i+1})$. See Figure 1 in the next section for a very similar situation. The problem is how large the number l can be. It is easy to prove that the sequence $u = j(I_1)j(I_2)\ldots j(I_l)$ is a $(s+2)$-DS sequence. Thus the number $|u| = l$ of distinct portions of the graph of f can be bounded from above by $N_{s+2}(n)$.

A natural system of functions f_i with bounded number of solutions of $f_i = f_j$ is, for example, any system of distinct polynomials of bounded degree or, as was the case in [12], any system of distinct solutions of a homogeneous linear differential equation with constant coefficients. The problem to determine or bound the maximum number of portions of the graph of f originated in control theory and was communicated to Davenport and Schinzel by K. Malanowski ([12]). They reduced geometry to combinatorics and asked about the values of $N_s(n)$. (In [12] and some other articles the subscript in $N_s(n)$

is smaller by 1 compared to our notation.) In [12] they proved that

$$N_s(n) \leq n(n-1)(s-1)+1 \tag{3}$$
$$N_4(n) = 2n-1 \tag{4}$$
$$N_5(n) \ll n\log n \tag{5}$$
$$N_s(n) \ll_s n\cdot\exp(10(s\log s)^{1/2}(\log n)^{1/2}). \tag{6}$$

Their proofs of (3)–(5) are reproduced in Example 2. (We have slightly corrected the proof of (3) to obtain (2) instead. Of the two proofs of (4) in [12] we present the second one, based on "an observation, made to us by Mrs. Turan that (...) one of the integers (...) occurs only once".) They proved further that $N_s(n) \geq (s^2-6s+8)n - C(s)$ ($s > 4$ is even) and $N_s(n) \geq (s^2-7s+14)n - C(s)$ ($s \geq 5$ is odd). In particular, $N_5(n) > 5n - c$.

Davenport's results. In the posthumously published paper [11] (edited by Schinzel), Davenport improved (5) to $N_5(n) \ll n\log n/\log\log n$. He noted that the limit of the ratio $N_5(n)/n$ exists, because $N_5(m+n) \geq N_5(m)+N_5(n)$ for every m and n (easy to see from definitions). He proved more specificly that

$$\lim_{n\to\infty}\frac{N_5(n)}{n} \geq 8.$$

The third result of Davenport is the inequality $N_5(lm+1) \geq 6lm-m-5l+2$ which was "found in collaboration with J. H. Conway". It implies that $N_5(n) \geq 5n-8$, with strict inequality for odd $n \geq 13$ and even $n \geq 18$. Note added in proof (apparently by Schinzel) says that Z. Kołba proved that $N_5(2m) \geq 11m-13$ (this lowers 18 to 12).

Results of Roselle and Stanton. Roselle and Stanton proved in [52] that $N_s(3) = 3s-7$ (for odd $s > 4$) and $N_s(3) = 3s-8$ (for even $s > 4$), in [45] that $N_s(4) = 6s-19$ (for odd $s > 5$) and $N_s(4) = 6s-20$ (for even $s > 5$), and in [44] that $N_s(5) = 10s-37$ (for odd $s > 7$; case $s = 7$ contains an error) and $N_s(5) = 10s-39$ (for even $s > 6$). It is shown that for 4 symbols and even s, and 5 symbols and odd s there are two (normal) longest DS sequences. In the remaining cases the longest DS sequence is unique. In [44] the bound $N_5(n) \geq 5n-8$ is proved for $n \geq 4$. In [45] Roselle and Stanton gave the general bound ($s > n+1$)

$$N_s(n) \geq \left\{\begin{array}{ll} \binom{n}{2}(s-1) - F(n) & s \text{ is odd} \\ & \\ \binom{n}{2}(s-1) - F(n) - \lfloor\frac{n-1}{2}\rfloor & s \text{ is even}\end{array}\right\} \tag{7}$$

where $F(n) = (2n^3 + 9n^2 - 32n + 9)/12$ for odd $n \geq 3$ and $F(n) = (2n^3 + 9n^2 - 32n + 12)/12$ for even n. For $n = 3, 4, 5$ these bounds are sharp.

If $n = o(s)$, bounds (2) and (7) yield the asymptotics $N_s(n) = (1 + o(1))\binom{n}{2}s$. But what if n is bigger?

Problem 1. For $n = s$ bounds (2) and (7) give

$$\frac{n^3(1 + o(1))}{3} < N_n(n) < \frac{n^3(1 + o(1))}{2}.$$

What is the precise asymptotics of $N_n(n)$? □

Further results. Peterkin [40] determined by computer the value $N_6(6) = 29$, found all 35 corresponding longest DS sequences, corrected the value $N_7(5)$ of Roselle and Stanton to 34 (they had the incorrect value 33), and proved that $N_6(n) \geq 7n - 13$ $(n > 5)$ and $N_7(n) \geq 13n - 32$ $(n > 5)$.

Burkowski and Ecklund [10] found for small values n, r, d the maximum lengths of d-DS sequences over n symbols in which no symbol appears more than r times. MR reviewer N. G. de Bruijn writes: "The following question was raised by D. J. Newman: Is there a word S in some $\Phi_{n,4}$ [5-DS sequences over n symbols] that contains each symbol at least 5 times? The authors give an affirmative answer (but the proof seems to be incomplete)."

Dobson and Macdonald [14] obtained a slight improvement of (7). We state one of their bounds: if n is even and r is odd then $N_{n+r}(n) \geq \frac{1}{6}(2n^3 + 3n^2(r - 3) - 2n(3r - 8) + 6r - 6)$. For $n > 2r + 4$ this improves (7). Their other bounds are similar.

Rennie and Dobson [42] derived inequality

$$\left(n - 2 + \frac{1}{s - 4}\right) \cdot N_s(n) \leq n \cdot N_s(n - 1) + \frac{2n - s + 3}{s - 4}. \qquad (8)$$

From it they could obtain good upper bounds on $N_s(n)$ for small values of s and n.

The next table of bounds for $N_s(n)$ when $6 \leq s \leq 13$ and $5 \leq n \leq 12$ is taken from Rennie and Dobson [42]. The upper bound is obtained from (8). The lower bound is the larger of the lower bounds given by Dobson and Macdonald or (shown in italic) by Roselle and Stanton in (7).

s	6	7	8	9
n				
5	22	34	41	53
6	29	46–47	56–58	72–76
7	36–37	59–62	72–77	96–102
8	43–46	72–78	89–99	120–131
9	50–56	85–96	106–123	145–164
10	57–66	98–115	123–149	170–200
11	64–77	111–136	140–177	195–239
12	71–89	124–158	157–207	220–281

s	10	11	12	13
n				
5	61	73	81	93
6	*85*–88	*102*–105	*115*–117	*132*–135
7	110–119	134–143	*152*–159	*176*–184
8	140–154	170–186	192–207	*223*–240
9	170–193	210–234	236–261	276–303
10	201–236	250–287	284–321	332–373
11	232–283	291–345	332–387	392–450
12	263–334	332–408	381–458	452–534

Mills in [37] proved inequalities $N_5(k^2+5-j) \geq 6k^2-2k+16-6j$ and $N_5(k^2+k+5-j) \geq 6k^2+4k+15-6j$ where $0 \leq j < k$. In [37] and [38] he determined the values of $N_5(n)$ for $n \leq 21$:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$N_5(n)$	1	4	8	12	17	22	27	32	37	42	47	53	58	64

n	15	16	17	18	19	20	21
$N_5(n)$	69	75	81	86	92	98	104

The formula $N_5(n) = 5n-8$, valid for $4 \leq n \leq 11$, breaks down later, as noted already by Davenport.

Szemerédi's general bound. In 1974 Szemerédi [53] published a remarkable result with a complicated proof:

$$N_s(n) \ll_s n \log^* n \tag{9}$$

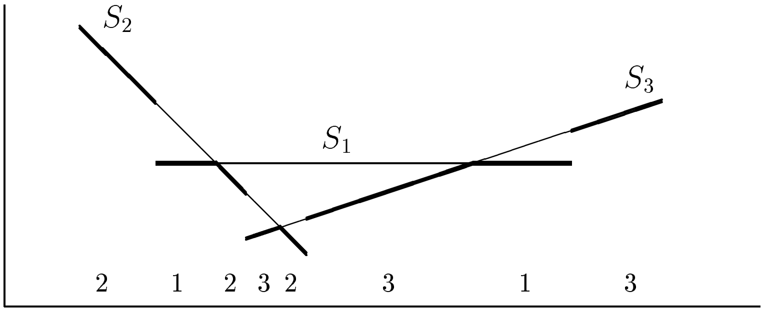


Figure 1: Lower envelope of plane segments.

where $\log^* n$ is the smallest number of iterations of the exponential function e^x, starting from $x = 1$, giving a number $\geq n$. In the proof the key role plays a decomposition lemma based on the doubly exponential upper bound in the classical Ramsey theorem for triples and 2 colors. Another lemma of Szemerédi, today called Szemerédi's regularity lemma, does not appear in the article. (9) improves considerably both (6) and (5).

Mills article [38] and Stanton and Dirksen's survey [51], both published in 1976, mark the end of early investigations of $N_s(n)$. DS sequences were dormant for the next 10 years.

3 Superlinearity

Bounds on $\mathbf{N_5(n)}$: enigma solved. In the mid of 1980s importance of DS sequences for combinatorial and computational geometry was discovered, first by Atallah [7]. Or rather rediscovered, since the geometric motivation was in background from the very beginning, only computational geometry did not exist in the times of [12]. $N_5(n)$ and more generally $N_s(n)$ for $s > 4$ remained mysterious. Despite the effort invested in the proofs of the upper bounds (6) and (9), $O(n)$ bounds were not in sight and the correct growth magnitudes of these functions were unclear. Stanton and Dirksen conjectured in [51] that $N_5(n)/n$ tends to infinity.

Example 3 ([48]). We illustrate the role of DS sequences in combinatorial geometry by a classical example, very similar to the problem in [12] but more recent. Let $S_1, \ldots, S_n$ be n straight segments in the plane, none of them vertical and no two of them overlapping. We regard them as graphs of

n real functions (now defined only on intervals) and define in the same way as before their pointwise minimum function f, the corresponding splitting of $\mathbf{R}$ into intervals $I_1, \ldots, I_l$ (now f is undefined on some of them, certainly on I_1 and I_l), and the sequence u over $\{1, 2, \ldots, n\}$, $|u| \le l - 2$, which records participation of the segments S_i on the graph of f (the intervals where f is undefined are omitted in u). In Figure 1 $n = 3$, $l = 10$, and $u = 21232313$. The graph of f is the *lower envelope* of the system $\{S_1, \ldots, S_n\}$ and the length of u is its complexity. From the fact that two (nonoverlapping) plane segments intersect in at most one point it follows that u is a 5-DS sequence. Hence the complexity $|u|$ of the lower envelope of n plane segments is bounded from above by $N_5(n)$. □

Another geometric connection was recently discussed by Alon and Onn [6]: the partitions of a set of n points lying on the moment curve in $\mathbf{R}^d$ into p parts whose convex hulls are mutually disjoint correspond to $(d+2)$-DS sequences (immediate repetitions are now allowed) over $\{1, 2, \ldots, p\}$ with length n.

Light on the behaviour of $N_5(n)$ was shed by Hart and Sharir in 1986 in the breakthrough article [19]. They proved that

$$n\alpha(n) \ll N_5(n) \ll n\alpha(n) \tag{10}$$

where $\alpha(n)$ is the inverse Ackermann function. It is defined as follows. Ackermann function $A(n)$ is the diagonal function $A(n) = F_n(n)$ of the hierarchy of functions $F_i : \mathbf{N} \to \mathbf{N}$, $i \in \mathbf{N}$, beginning with $F_1(n) = 2n$ and continuing by the rule $F_{i+1}(n) = F_i(F_i(\ldots F_i(1)\ldots))$ with n applications of F_i. The inverse Ackermann function is $\alpha(n) = \min\{m \in \mathbf{N} :\ A(m) \ge n\}$. It grows to infinity much more slowly than $\log^* n$. (For information about the role of very fast and very slow functions in combinatorics and computer science see Loebl and Nešetřil [36].) Of course, (10) was not only an improvement upon (9) for $s = 5$ but it settled almost completely the 20 years old riddle of Davenport and Schinzel about the growth rate of $N_5(n)$.

Hart and Sharir in [19] translated 5-DS sequences to certain tree objects motivated by data structures algorithms, called path compression schemes, derived the upper and lower bound for them, and translated results back to 5-DS sequences. They were inspired by some ideas and techniques of Tarjan [54] who introduced $\alpha(n)$ in computer science.

Their method gave bounds both for 5-DS sequences and for path compression schemes but it was technically complicated. It turned out that the translation is not really necessary and that one can work directly with sequences. This approach is adopted in all subsequent articles. (For explanation of path compression schemes and their relation to 5-DS sequences, see also the book of Sharir and Agarwal [48].) Komjáth [34] proved the superlinear lower bound $N_5(n) \gg n\alpha(n)$ by a construction purely in terms of sequences. Wiernik and Sharir [60] gave a simpler construction and, more importantly and remarkably, they proved that the 5-DS sequences u produced by it can be realized by plane segments in the way described in Example 3. Thus there exist systems of n plane segments whose lower envelopes have $\gg n\alpha(n)$ portions. We come to the natural but open

Problem 2. Can every 5-DS sequence be realized by lower envelope of plane segments? $\square$

In [48] the authors express opinion that the correct answer is negative.

Example 4 ([60]). We decribe the construction of [60] proving $N_5(n) \gg n\alpha(n)$. It is explained in [48] too. One defines, by double induction, a two-dimensional array $S : \mathbf{N} \times \mathbf{N} \to A^*$ of sequences. Before giving the precise inductive definition, we have to say that the sequences $S(k, m)$ have no immediate repetition and are of the form

$$S(k, m) = u_1 v_1 u_2 v_2 \ldots u_N v_N$$

where every u_i is a sequence of length m which contains m distinct symbols and v_i are possibly empty intermediate sequences. Sequences u_i are called *fans* or *m-fans* and v_i are called *separating sequences*. The key property of fans is this: every symbol of $S(k, m)$ appears in exactly one fan and this is its leftmost appearance in $S(k, m)$. The number $N = N(k, m)$ will be defined inductively in the construction. Sequences u_i and v_i depend on k and m as well, of course, but to avoid cumbersome notation we do not mark this dependence.

The first row $k = 1$ consists of the sequences $S(1, m) = u_1 = 12 \ldots m$ and $N(1, m) = 1$. If row $k \geq 1$ is already defined, we define $S(k+1, 1)$ to be identical with $S(k, 2)$, except that every 2-fan in $S(k, 2)$ is now regarded as two neighbouring 1-fans in $S(k+1, 1)$. Thus $N(k+1, 1) = 2N(k, 2)$.

Let now row $k \geq 1$ be already defined, as well as the sequences in row $k+1$ up to the column $m \geq 1$; the same for the values of $N(x, y)$. We define $S(k+1, m+1)$. Take the sequence $w_0 = S(k, N(k+1, m))$. Set $M =$

$N(k, N(k+1, m))$ and create M copies $w_1, w_2, \dots, w_M$ of the sequence $S(k+1, m)$, renaming symbols so that no two of the $M+1$ sequences $w_0, w_1, \dots, w_M$ share a symbol. We have as many copies of $S(k+1, m)$ as fans in w_0 and any fan in w_0 has as many elements as $S(k+1, m)$ has fans.

By duplicating the last term in every fan in every $w_i, i = 0, 1, \dots, m$, create sequences w_i'. Set

$$S(k+1, m+1) = w_1^* z_1 w_2^* z_2 \dots w_M^* z_M.$$

Here w_i^* is obtained from w_i' by inserting between the twins on the end of the j-th (counted from the left) m-fan of w_i' the j-th term of the i-th $N(k+1, m)$-fan of w_0' (everything is counted from the left). Sequence z_i consists of the second member of the twins which were on the end of the i-th fan of w_0' followed by the separating sequence of w_0' lying between the i-th and $(i+1)$-th fan. Said more briefly but less precisely, the two sequences w_0' and $w_1' w_2' \dots w_M'$ are merged together so that the order of terms is preserved and all immediate repetitions created by duplications are canceled. As for $(m+1)$-fans in $S(k+1, m+1)$, they consist of the old m-fans in $w_1', \dots, w_M'$ enlarged by 1 by the elements from the fans in w_0'. Hence

$$N(k+1, m+1) = N(k+1, m) \cdot N(k, N(k+1, m)).$$

One can easily check that the key property of fans is preserved during the construction.

Note that $S(k, m)$ uses exactly $m \cdot N(k, m)$ symbols. Using the key property of fans, it is easy to show by induction that every $S(k, m)$ is a 5-DS sequence. One can define, for details consult [60], a sequence of numbers $1 \le m_1 < m_2 < \cdots$ such that, writing n_k for $\|S(k, m_k)\| = m_k \cdot N(k, m_k)$, the inequality $|S(k, m_k)| \ge n_k \alpha(n_k) - 3n_k$ holds. (We owe superlinearity to duplications.) Hence

$$N_5(n_k) \ge n_k \alpha(n_k) - 3n_k \tag{11}$$

for every $k \in \mathbf{N}$. A simple interpolation argument of [48] shows that

$$N_5(n) \ge \tfrac{1}{2} n \alpha(n) - 2n$$

holds for all $n \in \mathbf{N}$. □

Hart and Sharir [19] proved the lower bound in (10) with the constant $\frac{1}{4} + o(1)$. The constants achieved in the upper bound were $52 + o(1)$ in [19] and $68 + o(1)$

in [48] but the objective of these works was not really to obtain the best constants. Klazar [24] achieved $4 + o(1)$ and in [29] he proved that

$$N_5(n) < 2n\alpha(n) + O(n\alpha(n)^{1/2}). \tag{12}$$

Problem 3. Does the limit

$$\lim_{n\to\infty} \frac{N_5(n)}{n\alpha(n)}$$

exist? □

If it exists, by (11) and (12) it lies in the interval $[1, 2]$. (And not only in $[\frac{1}{2}, 2]$, as we stated sloppily in [29].)

We answer the question from the MR review of Burkowski and Ecklund [10] that we quoted in the previous section. We know that $N_5(n) > cn$ for large n for every constant c. We deduce from it that for every $k \in \mathbf{N}$ there exists a 5-DS sequence v in which every symbol appears at least k times. Let $v \in \mathrm{DS}_5$ be such that $|v| \geq (k+1)\|v\|$ and $\|v\|$ is as small as possible. If an $a \in S(v)$ occurs in v less than k times, we eliminate all a-occurrences by deleting at most $k - 1 + 2 = k + 1$ terms from v and obtain a subsequence $w \in \mathrm{DS}_5$ such that $|w| \geq (k+1)\|w\|$ and $\|w\| \leq \|v\| - 1$ (see the proof in Example 2). But this contradicts the minimality of $\|v\|$. Therefore v has the stated property. Note that for $k = 5$ sequence v must use at least 22 symbols because by Mills table in Section 2 $N_5(n) < 5n$ for $n < 22$.

Bounds on $\mathbf{N_s(n)}$ for $\mathbf{s > 5}$. The next obvious step was to extend new techniques to $N_s(n)$ for $s > 5$. Sharir in [46] proved the upper bound

$$N_s(n) \ll n\alpha(n)^{c_s\alpha(n)^{s-5}}$$

and in [47] the lower bound

$$N_{2s+1}(n) \gg_s n\alpha(n)^{s-1}.$$

Of course, $N_{2s}(n) \geq N_{2s-1}(n) \geq N_{2s-2}(n)$.

This line of research culminated in the long and technical work [4] of Agarwal, Sharir and Shor published in 1989. For $s = 6$ they proved estimate

$$n2^{\alpha(n)} \ll N_6(n) \ll n2^{\alpha(n)}. \tag{13}$$

For $s > 6$ they obtained strong bounds as well but they could not match completely the precision of (10) and (13). Their lower bound says that

$$N_{2s}(n) \gg_s n2^{c_s\alpha(n)^{s-2}+Q_s(n)} \tag{14}$$

where $c_s = 1/(s-2)!$ and $Q_s(n)$ is a polynomial in $\alpha(n)$ of degree at most $s-3$. As for the upper bound, they proved that

$$N_{2s+1}(n) \le n2^{\alpha(n)^{s-2}\log(\alpha(n))+C_{2s+1}(n)} \quad \text{and} \quad N_{2s}(n) \le n2^{\alpha(n)^{s-2}+C_{2s}(n)} \tag{15}$$

where $C_k(n)$ equals 6 and 11 for k equal to 5 and 6, respectively, $C_{2s+1}(n) = O(\alpha(n)^{s-2})$, and $C_{2s}(n) = O(\alpha(n)^{s-3}\log(\alpha(n)))$. We must remark that in these bounds (and the whole [4]) $\log n$ denotes the *binary logarithm* with base 2, whereas in Example 2 and (5) we have natural logarithm.

Let us summarize the current best bounds on $N_s(n)$. Case $s \le 3$ is trivial. $N_4(n) = 2n - 1$ was proved by Davenport and Schinzel in [12], see Example 2. $N_5(n)$ and $N_6(n)$ grow up to multiplicative constants as $n\alpha(n)$ and $n2^{\alpha(n)}$, respectively, as proved by Hart and Sharir [19] and Agarwal, Sharir and Shor [4]. Bounds (14) and (15) of [4] estimate $N_s(n)$ for $s > 6$.

Problem 4. What is the exact speed of growth of $N_7(n)$ and $N_8(n)$? And, more generally, of $N_s(n)$ for $s > 6$? □

By (14) and (15),

$$n2^{\alpha(n)} \ll N_7(n) \ll n2^{(1+o(1))\alpha(n)\log(\alpha(n))}$$

and

$$n2^{(1+o(1))\alpha(n)^2/2} \ll N_8(n) \ll n2^{(1+o(1))\alpha(n)^2}.$$

Bounds on $N_s(n)$ found many applications in problems and algorithms of computational geometry. We suggest to the interested reader works [3] and [48] of Agarwal and Sharir for detailed information and many references. We remark that the "Web of science" listed in the end of the year 2000 more than 110 citations of [19], which documents big impact of this work.

4 Generalizations of $N_s(n)$

Containment of sequences. We generalize $N_s(n)$ and define an extremal function for any forbidden sequence, not only for alternating one. Recall that

A is an infinite alphabet such that $A \supset \mathbf{N}$ and $a, b, c, d, \ldots \in A$, and that our sequences are finite and over A. Recall that two sequences $u = a_1 a_2 \ldots a_l$ and $v = b_1 b_2 \ldots b_l$ of the same length are isomorphic if for some injection $f : A \to A$ we have $a_i = f(b_i)$, $i = 1, 2, \ldots, l$. This is an equivalence relation and each class of isomorphic sequences contains exactly one normal sequence (normality is defined before Example 1). We shall use the convention that $a, b, c, d, \ldots$ refer to elements of A and $u, v, w, \ldots$ to sequences over A.

Let u and v be two sequences. We say that u *contains* v and write $u \supset v$, if u has a subsequence isomorphic to v. Containment is a nonstrict partial order on classes of isomorphic sequences. For example, $u = a_1 a_2 \ldots a_l$ contains $abccba$ iff there are six indices $1 \le i_1 < \cdots < i_6 \le l$ such that $a_{i_1} = a_{i_6}$, $a_{i_2} = a_{i_5}$, $a_{i_3} = a_{i_4}$, and these are the only equality relations among $a_{i_1}, \ldots, a_{i_6}$. If u does not contain v we say that u is *v-free*.

Extremal function $\mathbf{Ex(v, n)}$. A sequence $u = a_1 a_2 \ldots a_l$ is called *k-sparse* if $a_i = a_j, i > j$, implies $i - j \ge k$. In other words, in every interval in u of length at most k all terms are distinct. For $k = 2$ we get condition 1 from the definition of DS sequences. Recall that $|u|$ is the length of a sequence and $\|u\|$ the number of symbols used in u.

Let v be a sequence and $n \in \mathbf{N}$. We associate with v the extremal function

$$\mathrm{Ex}(v, n) = \max\{|u| :\ u \not\supset v \ \&\ u \text{ is } \|v\|\text{-sparse} \ \&\ \|u\| \le n\}. \tag{16}$$

It clearly extends $N_s(n)$ because, if al_s denotes the alternating sequence $abab\ldots$ of length s, $N_s(n) = \mathrm{Ex}(\mathrm{al}_s, n)$. On known properties of $\mathrm{Ex}(v, n)$ we are going to report.

$\mathrm{Ex}(v, n)$ was introduced, albeit in a different notation, by Adamec, Klazar and Valtr in [1] in 1992. Notice that it is always defined because a modification of the argument from Example 2 proving (2) gives that

$$\mathrm{Ex}(v, n) < \|v\|(\binom{n}{\|v\|}(|v| - 1) + 1). \tag{17}$$

Before we proceed to further general properties, we derive two specific bounds to give the reader flavour of arguments used to handle $\mathrm{Ex}(v, n)$, and we present a historical remark.

Example 5 ([24]). We determine $\mathrm{Ex}(abba, n)$ exactly and prove a linear upper bound on $\mathrm{Ex}(a_1 a_2 \ldots a_k a_1 a_2 \ldots a_k, n)$, where a_i are k distinct symbols.

We prove that, for all $n \in \mathbf{N}$,

$$\mathrm{Ex}(abba, n) = 3n - 2.$$

The sequences

$$u = 1\ 2\ 1\ 2\ 3\ 2\ 3\ 4\ 3\ 4\ 5\ 4\ \ldots (n-2)\ (n-1)\ (n-2)\ (n-1)\ n\ (n-1)\ n$$

show that $\mathrm{Ex}(abba, n) \geq 3n-2$. We prove the opposite inequality. For $n = 1$ it is true. For $n > 1$ we use induction on n. Let $u = a_1 a_2 \ldots a_l$ be 2-sparse, $abba$-free, and $\|u\| \leq n$. Suppose first that a symbol $a \in S(u)$ appears in u at least four times. Then we select four indices $1 \leq i_1 < \cdots < i_4 \leq l$ such that $a_{i_1} = a_{i_2} = a_{i_3} = a_{i_4} = a$ and $a_j = a$ for no other j, $i_1 \leq j \leq i_4$. A moment of thought reveals that the symbol $b = a_{i_2+1}$ is distinct from a and appears in u only once. Deleting a_{i_2+1} and also a_{i_2} if $i_3 = i_2 + 2$, we decrease $\|u\|$ by one and obtain by induction the even stronger bound $|u| \leq 3(n-1) - 2 + 2 = 3n - 3$. Thus every symbol appears in u at most three times and we have at once that $|u| \leq 3n$. But $a = a_1$ cannot appear in u three times because otherwise the deletion argument still works for $b = a_2$. The same for $a = a_l$. If $a_1 = a_l$, clearly $|u| \leq 2n-1$. Thus we can assume in addition that $a_1 \neq a_l$ and we obtain the small improvement $|u| \leq 3n - 2$.

Let $v_k = a_1 a_2 \ldots a_k a_1 a_2 \ldots a_k$. We prove that

$$\mathrm{Ex}(v_k, n) \ll_k n.$$

Notice that $\mathrm{Ex}(v_1, n) = n$ and $\mathrm{Ex}(v_2, n) = N_4(n) = 2n - 1$. Let $k, n \in \mathbf{N}$ where k is fixed. We set $K = (k-1)^4 + 1$ and $L = \mathrm{Ex}(v_k, K-1) + 1$. The number L exists by the rough general bound (17). Let u be a k-sparse sequence ($\|v_k\| = k$) such that $\|u\| \leq n$. We assume that $|u| \geq (2n+1)L$ and we will derive from it that $u \supset v_k$. So we can split u into $2n+1$ disjoint intervals of length at least L each. One of these intervals, let us call it I, contains neither the first nor the last appearance of any $a \in S(u)$ because these are only at most $2n$ in number. By the definition of L, either $I \supset v_k$ and we are done or $\|I\| \geq K$. So $|S(I)| \geq K$. By the property of I, every $a \in S(I)$ appears before I, in I, and after I. Applying twice the classical Erdős–Szekeres lemma saying that every sequence of $(k-1)^2 + 1$ numbers contains a monotonous subsequence of length k, we see that there is a subset $Y \subset S(I)$ of k elements, $Y = \{y_1, y_2, \ldots, y_k\}$, such that $y_1, y_2, \ldots, y_k$ appear before I in this order, in I in this or in the opposite oder, and after I also in this or in the opposite order. But then in two of the three cases the orders agree and again $u \supset v_k$ which was to be proved. We conclude that $\mathrm{Ex}(v_k, n) < (2n+1)L$. $\square$

Historical remark. The author of this survey wondered from time to time why for so long (until the apperance of [1]) nobody tried to generalize $N_s(n)$ and everybody followed so faithfully the original formulation of the problem forbidding only alternating sequences. The extension (16) of (1) is straightforward and Example 5 shows that one can obtain for $\mathrm{Ex}(v, n)$ results of certain interest with only a modest effort, while to prove (10), say, is quite difficult. Of course, retrospect views are often dubious. We prefer to leave meditations about possible psychological causes to the reader but we want to share with her/him a little historical discovery of ours.

Surprisingly, revisionists appeared already in the very beginning and it was nobody else but Davenport and Schinzel who in 1965, besides the famous [12], published also [13] hinting on a generalization of $N_s(n)$. The latter forgotten note is missing in all bibliographies of DS sequences we know of ([3], [18, problem E20], [24], [48], [51], ...) and probably is not refered to anywhere. It is accesible in Davenport [8] where we found it. Davenport and Schinzel derive in [13] an inequality on lentghs of subsequences of a 2-sparse sequence. In the last paragraph they say:

> The inequality is of some interest in connection with sequences which, in addition to having no immediate repetition, satisfy some prescribed "hereditary" conditions, that is, some condition which if valid for a sequence is necessarily valid for every subsequence. Take as an illustration the condition that the sequence contains no subsequence
>
> $$\ldots,\ a, \ldots,\ b, \ldots,\ b, \ldots,\ a,\ \ldots\ (b \neq a)\ .$$
>
> Then the length of any such a sequence is at most $2n(n-1)$; for we can apply (1) [they refer to the inequality] with $m = 2$, in which case $M \leq 4$. (Actually in this particular case the maximum length is $3n - 2$.)

Nobody followed the hint then.

Almost linear bound on $\mathbf{Ex(v, n)}$. As for strengthening of (17), Klazar [21] showed by a simple combinatorial argument that $\mathrm{Ex}(v, n) \ll_v n^2$. The main result of [21] is that if v is a sequence with $\|v\| = k \geq 2$ and $|v| = l \geq 5$, then for every $n \in \mathbf{N}$

$$\mathrm{Ex}(v, n) \leq n \cdot k2^{l-3} \cdot (10k)^{2\alpha(n)^{l-4} + 8\alpha(n)^{l-5}}. \tag{18}$$

For $k = 1$ or $l \leq 4$ it is easy to show that $\mathrm{Ex}(v, n) \ll_v n$. This general bound was derived by adaptation of techniques from Sharir [46].

Slight generalization: $\mathbf{Ex(v, k, n)}$. One can investigate the more general extremal function $\mathrm{Ex}(v, k, n)$ defined as the maximum length of a v-free and k-sparse sequence u with $\|u\| \leq n$. Obviously, $\mathrm{Ex}(v, \|v\|, n) = \mathrm{Ex}(v, n)$ and $\mathrm{Ex}(v, k, n) = \infty$ for $k < \|v\|$. One has to take $k \geq \|v\|$. For asymptotics it brings nothing new because, as proved in [1],

$$\mathrm{Ex}(v, n) \ll_{v,k} \mathrm{Ex}(v, k, n) \leq \mathrm{Ex}(v, n) \tag{19}$$

for every sequence v and every $k \geq \|v\|$, the upper bound being trivial. As for the precise values, in Klazar [26] it was proved that for $n \geq k \geq 2$

$$\mathrm{Ex}(abab, k, n) = 2n - k + 1 \quad \text{and} \quad \mathrm{Ex}(abba, k, n) = 2n + \left\lfloor \frac{n-1}{k-1} \right\rfloor - 1.$$

For $n \leq k - 1$ both functions equal to n.

As one expects, the asymptotic order of $\mathrm{Ex}(v, n)$ respects containment of sequences:

$$u \subset v \Rightarrow \mathrm{Ex}(u, n) \ll_{u,v} \mathrm{Ex}(v, n). \tag{20}$$

For $\|u\| = \|v\|$ this is triviality, since then even $\mathrm{Ex}(u, n) \leq \mathrm{Ex}(v, n)$. For $\|u\| < \|v\|$ this follows immediately from (19) (and $\ll$ cannot in general be replaced by $\leq$).

Blow-ups. We write a^i for the sequence $aa\ldots a$ of i a's and call u a *chain* if $\|u\| = |u|$, that is, u has no repetition whatsoever. Obviously, $\mathrm{Ex}(a^i, n) = (i - 1)n$ and $\mathrm{Ex}(v, n) = \min(|v| - 1, n)$ for any chain v.

The ultimate goal in our theory would be to determine function $\mathrm{Ex}(v, n)$ exactly for every sequence v, or at least to determine its asymptotics. At the present state of our knowledge this goal is inaccesible and we can give only very partial answers. The most succesful approach so far turned out to be one when from a sequence v, or from two sequences u and v, a new sequence w is created in a specific way and $\mathrm{Ex}(w, n)$ is bounded in terms of $\mathrm{Ex}(v, n)$, or of $\mathrm{Ex}(u, n)$ and $\mathrm{Ex}(v, n)$. We have found three such operations. We begin with the simplest one, with the blow-up.

Every sequence u can be expressed uniquely in the form $u = a_1^{i_1} a_2^{i_2} \dots a_r^{i_r}$ where $a_j \in A$, $a_j \neq a_{j+1}$, and $i_j \in \mathbf{N}$. The *blow-up* of u is any sequence isomorphic to $a_1^{k_1} a_2^{k_2} \dots a_r^{k_r}$ where $k_1 \geq i_1$, $k_2 \geq i_2, \dots, k_r \geq i_r$. For example, 1221111 and $a^3b^3a = aaabbba$ are blow-ups of $\mathrm{al}_3 = aba$. For every nonchain v (that is, if $|v| > \|v\|$) we have $\mathrm{Ex}(v, n) \geq |1\ 2\ \dots\ n| = n$. Thus the extremal function of every proper blow-up of a chain u grows substantially faster than $\mathrm{Ex}(u, n)$ because $\mathrm{Ex}(u, n)$ is eventually constant. There are reasons to believe that, apart of this trivial situation, blowing up a sequence does not change the asymptotics of its extremal function.

Problem 5. Prove (or disprove) that if u is not a chain and v is a blow-up of u, then

$$\mathrm{Ex}(v, n) \ll_{u,v} \mathrm{Ex}(u, n).$$

□

The lower bound $\mathrm{Ex}(v, n) \geq \mathrm{Ex}(u, n)$ is trivial.

Adamec, Klazar and Valtr [1] proved that the bound in Problem 5 is often true, namely, if a is a symbol and u, v sequences, then

$$\mathrm{Ex}(aau, n) = \mathrm{Ex}(au, n) + O_{au}(n) \quad \text{and} \quad \mathrm{Ex}(uaaav, n) \ll_{uaav} \mathrm{Ex}(uaav, n). \tag{21}$$

From symmetry, also $\mathrm{Ex}(uaa, n) = \mathrm{Ex}(ua, n) + O_{ua}(n)$. Iterating (21) we see that if $v = a_1^{k_1} a_2^{k_2} \dots a_r^{k_r}$ is any blow-up of $u = a_1^{i_1} a_2^{i_2} \dots a_r^{i_r}$ such that $k_j > i_j$ implies $i_j \geq 2$ or $j = 1$ or $j = r$, then $\mathrm{Ex}(v, n) \ll \mathrm{Ex}(u, n)$ and thus both functions are of the same asymptotic order. For example, $a^4bc^5ab^2$ and abc^2ab have extremal functions with the same asymptotic order.

The blow-up of uav to $uaav$ seems to be a qualitatively different matter and whether it changes asymptotics is unknown. Surprisingly, this happens for tree generalizations of $\mathrm{Ex}(v, n)$ as shown by Valtr in [57]. This is a contrary evidence showing that perhaps blow-up may change asymptotics.

Example 6 ([1]). We prove the first bounds in (19) and (21). We begin with (19). Let v be a fixed sequence, $k \geq \|v\|$ a number, and u a $\|v\|$-sparse and v-free sequence. It suffices to show that u has a k-sparse subsequence w such that $|w| \gg_{v,k} |u|$. We set w to be the maximum k-sparse subsequence of u. Let I be any of the intervals in u that are disjoint to w and into which u splits after the removal of w. Clearly, $\|I\| \leq 2k - 2$ because else some $a \in S(I)$ would differ from the $k - 1$ terms of w preceding and following I

and one appearance of a in I could be added to w, in contradiction with the maximality of w. Hence $|I| \le \mathrm{Ex}(v, 2k-2)$ and $|w| \ge 1 + |u|/(1+\mathrm{Ex}(v, 2k-2))$.

We prove that $\mathrm{Ex}(aau, n) = \mathrm{Ex}(au, n) + O_{au}(n)$. Let $k = \|aau\|$ and v be a k-sparse and aau-free sequence with $\|v\| \le n$. First, we show that there is a constant $c > 0$ depending only on aau with the following property. For every term x of v there is a k-sparse subsequence w of v avoiding x and of length $|w| \ge |v| - c$. In other words, x plus some other $O(1)$ terms of v can be deleted so that the k-sparseness is preserved. We fix arbitrary interval I in v containing x and of length $|I| = \mathrm{Ex}(aau, 3k-3)+1$. So $\|I\| \ge 3k-2$ and there is a subset $Y \subset S(I)$, $|Y| = k-1$, such that every $y \in Y$ is distinct from x and from the $k-1$ terms preceding and following I. We fix in I one appearance for every $y \in Y$ and delete from v the rest of I. The resulting sequence w is clearly k-regular, x was deleted, and $|w| \ge |v| - (\mathrm{Ex}(aau, 3k-3) + 2 - k)$. We can set $c = \mathrm{Ex}(aau, 3k-3) + 2 - k$.

Now we delete from v in this way first appearance of every $x \in S(v)$. Thus at most cn elements were deleted and the resulting subsequence w is k-sparse. Clearly, $w \not\supset au$ for otherwise $v \supset aau$ would be forced. Thus $|v| \le |w| + cn \le \mathrm{Ex}(au, n) + cn$ and $\mathrm{Ex}(aau, n) \le \mathrm{Ex}(au, n) + O_{au}(n)$. Trivially, $\mathrm{Ex}(aau, n) \ge \mathrm{Ex}(au, n)$. □

Two-letter theorem. If u is such that $\|u\| \le 2$ then either $u \supset ababa$ and $\mathrm{Ex}(u, n) \ge N_5(n) \gg n\alpha(n)$ or u is contained in a blow-up of $abab$. From the previous discussion it is clear that $\mathrm{Ex}(abbaab, n) \ll n$ implies $\mathrm{Ex}(u, n) \ll_u n$ for every u contained in a blow-up of $abab$. That $\mathrm{Ex}(abbaab, n) \ll n$ and consequently $\mathrm{Ex}(a^i b^j a^k b^l, n) \ll_m n$ (here $m = \max(i, j, k, l)$) for every $i, j, k, l \in \mathbf{N}_0$ is the main result of [1]. We state it as in [1] (Theorem 5) in the equivalent form: if $\|u\| \le 2$ then

$$\mathrm{Ex}(u, n) \ll_u n \iff u \not\supset ababa. \tag{22}$$

In [57] this characterization is called *two-letter theorem*.

More results and errors on blow-ups. In [33] the first author wrote on p. 467 : "However, it may be checked that the method [of Hart and Sharir] (...) works for $a^i b^i a^i b^i a^i$ as well and so $Ex(a^i b^i a^i b^i a^i, n) = \Theta(n\alpha(n))$.". However, after some time he realized that no matter how he tried he could not recollect the proof anymore and thus we have the following problem.

Problem 6. Prove (or disprove) that $\mathrm{Ex}(v, n) \ll_v n\alpha(n)$ for every blow-up v of $ababa$. That is, prove (or disprove) that

$$\mathrm{Ex}(abbaabba, n) \ll n\alpha(n).$$

$\square$

It is a special case of Problem 5. We are not done with the forbidden 5-term alternating subsequence yet! The applications of the conjectural $\mathrm{Ex}(a^ib^ia^ib^ia^i, n) \ll_i n\alpha(n)$ in 3.2 of [33] must be considered as unproved.

A simpler proof of the two-letter theorem was given in Klazar [25]. In particular, he proved that

$$7n - 9 \le \mathrm{Ex}(abbaab, n) \le 8n - 7$$

and $\mathrm{Ex}(a^ib^ia^ib^i, n) < (1 + o(1))32i^2n$. The method of [25] was extended in Klazar [23] to the blow-ups of the sequences v_k of Example 5. In [23] he proved that, for every $i \in \mathbf{N}$ and k distinct symbols $a_1, a_2, \ldots, a_k$,

$$\mathrm{Ex}(a_1^ia_2^i \ldots a_k^ia_1^ia_2^i \ldots a_k^i, n) \ll_{i,k} n. \tag{23}$$

Two general questions on the growth of $\mathbf{Ex(v, n)}$. Is equivalence (22) valid for sequences using more than two symbols? Klazar and Valtr realized that the answer was "no". A specific example is given in Klazar [22]: $u_1 = abcbadadbcd \not\succ ababa$ but $\mathrm{Ex}(u_1, n) \gg n\alpha(n)$. A more extremal example in this direction is given in [29]:

$$u_2 = abcbadadbecfcfedef \not\succ ababa \text{ but } \mathrm{Ex}(u_2, n) \gg n2^{\alpha(n)}.$$

Thus the containment of $\mathrm{al}_5 = ababa$ is not the sole cause of superlinearity. These results are obtained simply by showing that the constructions of [60] and [4] produce sequences avoiding not only $ababa$ and $ababab$, respectively, but even u_1 and u_2. Some new constructions might help to solve the following problem.

Problem 7. We conjecture that for every constant $c > 0$ there exists a sequence u such that

$$u \not\succ ababa \text{ but } \mathrm{Ex}(u, n) \gg n2^{\alpha(n)^c}.$$

$\square$

To put it differently, in view of (18), we conjecture that the magnitudes of extremal functions of *ababa*-free sequences are cofinal among all extremal functions. Below we will see that $u \not\succ abab$ implies $\mathrm{Ex}(u,n) \ll_u n$.

Another attractive but at present hopeless question is whether $n\alpha(n)$ is the smallest superlinear asymptotics of extremal functions.

Problem 8. Is it true that for every sequence u either $\mathrm{Ex}(u,n) \ll n$ or $\mathrm{Ex}(u,n) \gg n\alpha(n)$? $\Box$

We do not know any $\mathrm{Ex}(v,n)$ that would have different asymptotics on two different infinite subsequences of $\mathbf{N}$. Valtr [57] (Proposition 3) shows that for every sequence u,

$$\limsup_{n\to\infty} \frac{\mathrm{Ex}(u,n)}{n} = \infty \Longrightarrow \liminf_{n\to\infty} \frac{\mathrm{Ex}(u,n)}{n} = \infty.$$

If u is irreducible, which means that there is no nontrivial decomposition $u = u_1u_2$ with $S(u_1) \cap S(u_2) = \emptyset$, it is easy to prove the stronger result that $\lim_{n\to\infty} \mathrm{Ex}(u,n)/n$ exists. Also, recall that $\mathrm{Ex}(u,n)$ is eventually constant iff u is a chain, and if u is not a chain then $\mathrm{Ex}(u,n) \geq n$.

Insertions and intertwinings. We proceed to the remaining two operations for deriving upper bounds; they constitute the main result of Klazar and Valtr [33]. Suppose $u = u_1aau_2$ and v are two sequences such that $S(u) \cap S(v) = \emptyset$ (u_i are sequences and a is a symbol). *Insertion* of v into u gives the sequence $w = u(v) = u_1avau_2$. *Intertwining* of a sequence $u = u_1aau_2a$ and a symbol b where $b \notin S(u)$ (again u_i are sequences and a is a symbol) gives the sequence $w = u(b) = u_1abbau_2ab$. Notice that v can be inserted in u only if v and u do not share symbols and similarly the symbol b to be intertwined with u must not appear in u.

Let $u(v)$ arise by inserting v in u. If v is a chain it is easy to see that $\mathrm{Ex}(u(v),n) \leq \mathrm{Ex}(u,n)$. If v is not a chain then, as proved in [33],

$$\mathrm{Ex}(u(v),n) = \mathrm{Ex}(u_1avau_2,n) \ll_{u,v} \mathrm{Ex}(v,\mathrm{Ex}(u,n)). \tag{24}$$

Let $u(b)$ arise by intertwining b with u. Then, as proved in [33],

$$\mathrm{Ex}(u(b),n) = \mathrm{Ex}(u_1abbau_2ab,n) \ll_u \mathrm{Ex}(u,n). \tag{25}$$

From symmetry, the analogous statement holds for bau_1abbau_2. The lower bounds $\mathrm{Ex}(u(v),n) \gg_{u,v} \max(\mathrm{Ex}(u,n),\mathrm{Ex}(v,n))$ and $\mathrm{Ex}(u(b),n) \gg_u \mathrm{Ex}(u,n)$

are immediate from (20). Simpler proofs of (24) and (25) are given in Valtr [57]. Applying blow-ups on $u(b)$, we obtain that $\mathrm{Ex}(u_1ab^iau_2ab^j, n) \ll_{u,i,j} \mathrm{Ex}(u, n)$ for every $i, j \in \mathbf{N}$.

Linear sequences. It is natural (but difficult) to investigate the class of *linear sequences*

$$\mathrm{Lin} = \{v: \ \mathrm{Ex}(v, n) \ll n\}.$$

We have seen already several members of it: $abab$ (Example 2), $abba$ and v_k (Example 5), a^i and chains (trivial). On the other hand, $ababa \notin \mathrm{Lin}$ by (10).

Problem 9. What are the elements of Lin? $\Box$

The only way to generate linear sequences we know of is to start with a^i and to apply repeatedly blow-ups, insertions, and intertwinings. More precisely, let L be the smallest class of sequences closed on isomorphism and such that, for every $i \in \mathbf{N}$, (i) $a^i \in L$, (ii) $au \in L \Rightarrow a^iu \in L$, $ua \in L \Rightarrow ua^i \in L$, (iii) $uaav \in L \Rightarrow ua^iv \in L$, (iv) $u = u_1aau_2 \in L$ & $v \in L$ & $S(u) \cap S(v) = \emptyset \Rightarrow u_1avau_2 \in L$, (v) $u = u_1aau_2a \in L$ & $b \notin S(u) \Rightarrow u_1abbau_2ab \in L$, and $u = au_1aau_2 \in L$ & $b \notin S(u) \Rightarrow bau_1abbau_2 \in L$. Then ([33])

$$L \subset \mathrm{Lin}.$$

This follows from (21), (24), and (25). How big $\mathrm{Lin}\backslash L$ is we do not know.

Problem 10. Decide if there is a sequence $u \in \mathrm{Lin}\backslash L$. $\Box$

By (20), $u \subset v \in \mathrm{Lin}$ implies that $u \in \mathrm{Lin}$. Hence we can characterize Lin by means of the class of minimal nonlinear sequences

$$B = \{u: \ u \notin \mathrm{Lin} \text{ but } v \in \mathrm{Lin} \text{ whenever } v \subset u \ \&\ |v| < |u|\}.$$

Namely, $u \in \mathrm{Lin} \iff \forall v \in B\ [v \not\subset u]$. Knowing (effectively) B would bring the knowledge of Lin. But at present we know only two significant properties of B: (i) $ababa \in B$ and (ii) B has at least two elements (the other one being contained in $u_1 = abcbadadbcd$, viz the above discussion on u_1 and u_2). Might B be infinite?

Problem 11. Is the set B of minimal nonlinear sequences finite or infinite? $\Box$

If $(A^*, \subset)$ were well quasiordering (no infinite antichains), B would be finite (since it is an antichain in $\subset$) and there would (ineffectively) exist an algorithm for recognizing Lin. However, one can easily find infinite antichains in $(A^*, \subset)$ and B still might be infinite. An infinite antichain is presented in Klazar [22] where certain well quasiordered subsets of A^* are determined.

Interesting sequences in the class L. Let us return to the class L of sequences whose linearity we can prove. It is rich enough to contain several interesting sequences. First we prove by induction on $|u|$ that every $abab$-free sequence u falls in L. If $abab \not\subset u$ then u is isomorphic to $au_1au_2a\ldots au_k$ where $k \geq 1$ and u_i are possibly empty sequences, $S(u_i)$ are mutually disjoint, and $a \notin S(u_i)$ for every i. By induction, $u_i \in L$ for every i. Starting with a^{k+1} and inserting in it sequences u_i, we obtain that $u \in L$. Thus $\mathrm{Ex}(u, n) \ll_u n$ for every $abab$-free sequence u. In particular,

$$\mathrm{DS}_4 \subset \mathrm{Lin}.$$

Recall that $u_1 = abcbadadbcd \notin \mathrm{Lin}$. Thus $\mathrm{DS}_5 \not\subset \mathrm{Lin}$.

Intertwining and blowing up yield an immediate proof of the two-letter theorem: $a^3 \in L$, $ab^2a^2b \in L$ by intertwining, and $a^ib^ia^ib^i \in L$ $(i \geq 2)$ by blowing up. Hence every $a^{i_1}b^{i_2}a^{i_3}b^{i_4}$ $(i_j \geq 0)$ is linear.

Intertwining brings in L more complicated sequences. Repeated intertwinings and blow-ups show that the sequence

$$u' = 1\ 2\ 1\ 2\ 3\ 2\ 3\ 4\ 3\ 4\ 5\ 4\ \ \ldots (k-2)\ (k-1)\ (k-2)\ (k-1)\ k\ (k-1)\ k$$

of Example 4 lies in L. Hence this longest $abba$-free 2-sparse sequence is linear for every k (as well as its every blow-up). On the other hand, $abcadbcd \not\supset abba$ but $abcadbcd \notin L$. At present we are not able to prove linearity of all $abba$-free sequences.

Similarly, intertwinings and blow ups show that for every k and i the N-shaped sequence

$$u_N(k, i) = 1^i2^i \ldots (k-1)^ik^i(k-1)^i \ldots 2^i1^i2^i \ldots (k-1)^ik^i$$

belongs to L and thus is linear. The bound

$$\mathrm{Ex}(u_N(k, i), n) \ll_{i,k} n \tag{26}$$

is an important result of [33] which strengthens considerably (23).

It follows by an easy case analysis that every *ababa*-free sequence u with $\|u\| \leq 3$ is contained in a blow-up of one of the three sequences $u_1 = ababcbc$, $u_2 = abcbabc$, and $u_3 = abacacbc$. All blow-ups of u_1 and u_2 are in L and thus are linear (u_1 is the $k = 3$ instance of the above u' and $u_2 = u_N(3, 1)$). But $u_3 = abacacbc \notin L$.

Problem 12. Is it true that $\mathrm{Ex}(abacacbc, n) \ll n$? And what about the blow-ups of *abacacbc*? □

In fact, the minimal subsequences of *abacacbc* lying outside L are *abacabc*, *abcacbc*, *abacacb*, and *bacacbc*. The first two are, up to isomorphism, reversals of one another and hence have equal extremal functions. The same holds for the last two sequences. Writing $\overline{u}$ for the reversal of u, we have this incomplete characterization of linear sequences over three symbols: for $\|u\| \leq 3$,

$$u \text{ and } \overline{u} \text{ contain none of } ababa,\ abacabc, \text{ and } abacacb \Rightarrow \mathrm{Ex}(u, n) \ll_u n.$$

In the opposite direction we know only that $ababa \subset u$ implies nonlinearity of u.

5 Geometric graphs, colored trees, permutations, partitions, and hypergraphs

Geometric graphs. DS sequences found interesting applications in combinatorics of *geometric graphs*. These are realizations of graphs with vertices being represented by points in the plane in general position and edges by possibly crossing straight segments. Two edges in a geometric graph *cross* if their relative interiors intersect and they are *parallel* if they form opposite sides of a convex quadrilateral.

Katchalski and Last [20] proved, using bound (4), that a geometric graph with n vertices and no two parallel edges has at most $2n - 1$ edges. Valtr [56] lowered this bound to $2n - 2$, which proves the conjecture of Y. S. Kupitz from 1979; Figure 2 shows a scheme of geometric graphs attaining this value. (Another nice application of (4) in combinatorial geometry was given by Edelsbrunner and Sharir in article [15] with a self-explaining title.) Applying the N-sequence bound (26), Valtr proved in [56] more generally that every geometric graph with n vertices and no k pairwise parallel edges

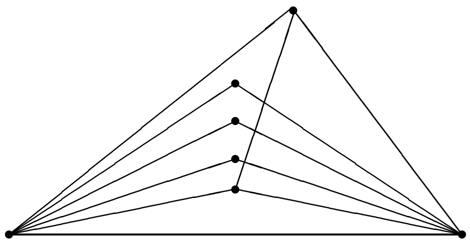


Figure 2: Geometric graphs with n vertices and $2n-2$ edges, no two of them "parallel".

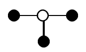
has in total $\ll_k n$ edges. From this he derived that if k pairwise crossing edges are forbidden, the number of edges is $\ll_k n \log n$. (For the proofs see also [57].) This improves the previous bound $\ll_k n \log^{2k-6} n$ (for $k > 2$) of Agarwal et al. [2]. Whether the bound $\ll_k n$ holds is open. It is known to hold only for $k = 2$ (classical case of planar graphs) and $k = 3$ ([2]).

Precisely speaking, the bounds $\ll_k n \log^{2k-6} n$ of [2] and $\ll_k n \log n$ of [56] are in a sense incomparable. The former bound holds, as stated in [2], for the more general representation of edges by curves (details are in [2] given only for the case of segments). The latter stronger bound applies on more special representations by segments. In [55] Valtr extended it to the situation when edges are represented by x-monotone curves but the case of general curves seems out of reach of his method.

For further applications of DS sequences in computational and combinatorial geometry see [3] and [48].

Colored trees. Looking through the glasses of graph theory, we can view sequences as undirected colored paths, colors being the symbols used. For example, $abcabc$ is the path of six vertices $v_1v_2\dots v_6$ where v_1 and v_4 are colored a, v_2 and v_5 are colored b, and v_3 and v_6 are colored c. To work with new and interesting objects, we consider colored paths to be just a special case of colored trees. Can one extend in a reasonable way definition (16) to colored trees? Can one prove then an analogue of, say, $N_4(n) = \mathrm{Ex}(abab, n) = 2n-1$? We begin with the latter question.

Let $T(abab, n)$ be the maximum number of vertices in a tree T that can be vertex-colored by at most n colors so that

1. the coloring is proper, which means that no edge is monochromatic,
2. no subgraph of T is a subdivision of the properly 2-colored 4-vertex path , and
3. no subgraph of T is a subdivision of the properly 2-colored 4-vertex star .

Condition 1 is the analogy of 2-sparseness and 2 forbids the color pattern *abab* in T. For colored trees conditions 1 and 2 do not suffice to bound the number of vertices, as shown by arbitrary large properly colored stars. Therefore we add condition 3. It says, in other words, that for every color c in the tree T_c, which is the smallest subtree of T containing all c-colored vertices, all vertices with degrees at least 3 must be colored c. Since *abab* is isomorphic to its reversal, for colored paths T the function $T(abab, n)$ coincides with $N_4(n)$.

Example 7 ([27]). We prove that

$$T(abab, n) = 2n - 1. \tag{27}$$

Even if all colored trees are taken into account, $T(abab, n)$ still equals $N_4(n)$.

$T(abab, n) \geq 2n - 1$ is achieved already on colored paths but it is easy to color any tree with $2n - 1$ vertices by n colors so that conditions 1, 2, and 3 are satisfied. We prove the opposite inequality $T(abab, n) \leq 2n - 1$ by induction on the number of colors and branching vertices.

Let $T = (V, E)$ be a tree and $f : V \to \{1, 2, \ldots, n\}$ a coloring satisfying conditions 1, 2, and 3. We assume that in the sum $n + \#\{v \in V :\ \deg_T(v) \geq 3\}$ the second summand is positive (else T is a path and the inequality holds) and $n > 1$. We suppose that for every tree with coloring satisfying conditions 1–3 and with smaller sum the inequality holds. By *smoothing* a vertex u of degree 2 or 1 in T we mean its deletion and (if it is not a leaf) subsequent connecting its neighbours by an edge.

There is a vertex v in T with degree at least 3 and such that $T - v = P_1 \cup P_2 \cup \ldots \cup P_l \cup C$ where $l \geq 2$, every P_i is a path, and C is a tree which may not be a path. First we show that f can be assumed injective on v and the vertices in $P_1, \ldots, P_l$. If $f(u_1) = f(u_2)$ for two distinct vertices such that

$u_1 = v$ or $u_2 = v$ or $u_1, u_2 \in V(P_i)$ for an i, there must be a vertex u_3 between u_1 and u_2 whose color does not appear elsewhere in T. As in Example 2, we smooth u_3 and, if necessary (not to create monochromatic edge), one of its neighbours. Number of colors has decreased by 1 and, by induction, $|V| \le 2(n-1) - 1 + 2 = 2n - 1$. Thus, in particular, we may assume that the color $f(v)$ does not appear in any P_i. The other case: $u_1 \in V(P_i)$ and $u_2 \in V(P_j)$ with $f(u_1) = f(u_2) = c$ and $i \ne j$. The color c does not appear elsewhere because otherwise we would have the previous case or condition 3 would be violated. If u_1 has two neighbours, they must have distinct colors for else we would have the $abab$ pattern. The same holds for u_2. We get rid of c by smoothing u_1 and u_2 and use induction on n as before.

So we assume no color repeats on the vertices in $T - C$. We form from the colored vertices in $P_1, \ldots, P_l$ a single path P which we join to v. The new colored tree (T^*, f^*) has fewer branch vertices ($\deg_{T^*}(v) = 2$) and satisfies conditions 1 and 3. To apply induction, we need only to show that the order of vertices in P can be chosen so that $abab$ is not created in (T^*, f^*).

To this end we define a binary relation R on the set of colors in P by setting aRb iff $a \ne b$ and there is a path $Q = (v_0, \ldots, v_k)$, $v_0 = v$, in $C \cup \{v\}$ such that, for some $i < j$, $f(v_i) = a$ and $f(v_j) = b$. We show that R is a strict partial ordering. We prove that R is antisymmetric. Suppose for the contradiction that aRb, witnessed by a path Q_1, and at the same time bRa, witnessed by a path Q_2. No matter where the merging point of Q_1 and Q_2 is, the vertices colored a and b in Q_1, Q_2, and the paths P_i form an $abab$ pattern in (T, f), which is a contradiction. We finish the proof by showing that R is transitive. Let aRb, witnessed by the path $Q_1 = (v_0, \ldots, v_k)$, $v = v_0$, and bRc, witnessed by the path $Q_2 = (w_0, \ldots, w_l)$, $v = w_0$. Let Q_1 and Q_2 merge at $v_i = w_i$, $i > 0$. Let $f(v_j) = a$ and let first $i < j$. If $f(w_m) = b$ and $m < i$ then the colors a and b realize, with the help of a in some P_i, $abab$ pattern in (T, f). If $m \ge i$ then, because of condition 3, it must be $f(w_i) = b$ and we arrive at the same contradiction. So $j \le i$ and, going from v, the colors a, b, and c appear on Q_2 in this order and therefore aRc.

Thus R is a partial ordering. Any possible occurrence of the pattern $abab$ in (T^*, f^*) would use two vertices of C and two vertices of P. We order the vertices in P in a linear extension of R so that if aRb then the vertex in P colored a is closer to v then the one colored b. Then no $abab$ pattern can appear. By induction, $|V(T)| = |V(T^*)| \le 2n - 1$. □

For a general sequence u, we define $T(u, n)$ simply by keeping condition 3

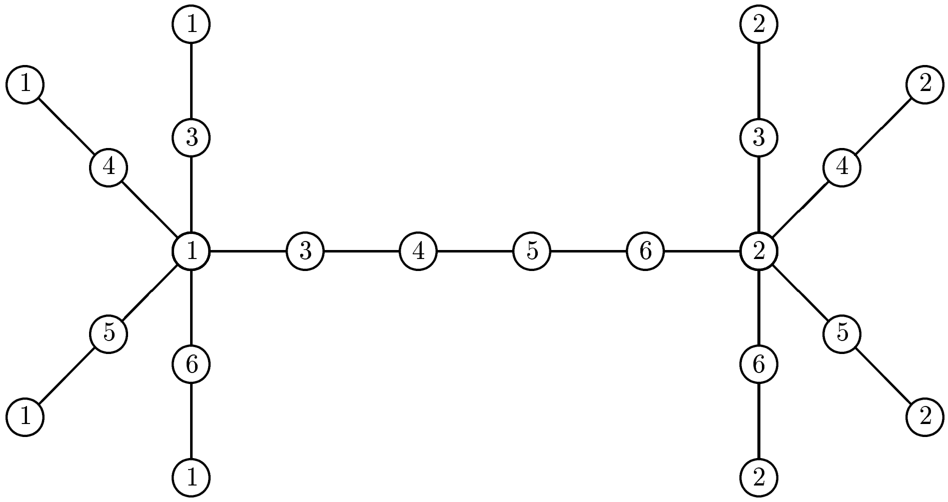


Figure 3: The bound $T(abba,n) \geq 5n-8$ for $n=6$.

and modifying condition 1 as to require every two distinct vertices in T with the same color to be at distance at least $\|u\|$ and condition 2 as to forbid in T subdivisions of the colored path u. For $\|u\|>1$ this is fine, $T(u,n)<\infty$, but for $u=a^i$ with $i \geq 4$ still $T(a^i,n)=\infty$ (consider monochromatic stars). Thus for $u=a^i$ we require the coloring to be proper. We must not forget that after replacing sequences by paths we loose the unique left-right order of terms. Forbidding u we forbid also its reversal $\overline{u}$.

The generalization of $\mathrm{Ex}(u,n)$ to colored trees, $T(u,n)$, was considered first in [24]. In the proof in Example 6 we have fixed a minor mistake of [27]—the paths P_i cannot in general be preserved in P as intervals. Various results on $T(u,n)$ were proved in Klazar [28], among others an extension of the general bound (18). One should not be led by (27) to thinking that always $T(u,n)=\mathrm{Ex}(u,n)$:

Example 8 ([58]). We show that $T(abba,n) \geq 5n-8$, in contrast with $\mathrm{Ex}(abba,n)=3n-2$. (Since again $abba$ is isomorphic to its reversal, $T(abba,n)$ really generalizes $\mathrm{Ex}(abba,n)$.) We hope that the construction of (T,f), independently discovered by Ch. Vogt, is clear enough from its instance $n=6$ visualized in Figure 3.

The coloring is proper, contains no subdivision of $abba$, and meets condition 3. In general the horizontal path has n vertices and the rays contribute

$4(n-2)$ vertices. Together we have $5n-8$ vertices. $\square$

In [24] and [27] we posed as a problem to show that $T(abba, n) \ll n$. This was accomplished in [58] by Valtr who proved more generally that

$$T(a^i b^i a^i, n) \leq 24in.$$

Thus $5n-8 \leq T(abba, n) \leq 48n$.

Problem 13. Improve these bounds. Or, better, determine the value of $T(abba, n)$ exactly. $\square$

The theory of tree extremal functions was much advanced by Valtr in [57] and [59]. He introduced other extensions of $\mathrm{Ex}(v, n)$ to colored trees which are better to work with than $T(v, n)$ and proved analogues of most of the results we described in the previous section. In particular, he extended the operations of blowing up, insertion, and intertwining to colored trees. To discuss properly his results would mean to write another Section 4; the reader is refered for details to [57] and [59]. Here we only say that, contrary to my original expectations, the behaviour of tree extremal functions often turns out to be much different compared to $\mathrm{Ex}(v, n)$. For example, the two-letter theorem for colored trees ([57], [59]) is: if $\|u\| \leq 2$,

$$T(u, n) \ll_u n \iff u \not\succ ababa \ \&\ u \not\succ ab^2a^2b.$$

In fact ([57], [59]), $T(ab^2a^2b, n) \gg n\alpha(n)$. Comparing this with (27), we see that for colored trees blow-ups change asymptotics. Unlike for sequences, for colored trees the patterns $abab$ and $abba$ are only just so linear.

Permutations. Recall that $[n] = \{1, 2, \ldots, n\}$. Alon and Friedgut [5] applied generalized DS sequences to a problem in enumerative combinatorics. Let $p = a_1 a_2 \ldots a_m$ and $q = b_1 b_2 \ldots b_n$ be permutations of $[m]$ and $[n]$, respectively. We say that q contains p if q has a subsequence $b_{i_1} b_{i_2} \ldots b_{i_m}$ such that $b_{i_r} < b_{i_s} \iff a_r < a_s$ for every r and s. Else we say that q *avoids* p. Let $S_n(p)$ be the number of permutations of $[n]$ avoiding p. The *Stanley–Wilf conjecture* (posed in 1992) asserts that for every fixed p we have

$$S_n(p) < c^n \tag{28}$$

with a constant c depending only on p. Using the general bound (18), Alon and Friedgut proved for every p the upper bound

$$S_n(p) < \beta_p(n)^n, \tag{29}$$

where $\beta_p(n)$ is an extremely slowly growing function defined in terms of $\alpha(n)$. Using the N-sequence bound (26), they proved that (28) holds if p is unimodal (first decreases and then increases or vice versa). Bóna [9] proved that (28) holds for all p of the form $p = a_1 a_2 \dots a_m = s_1 s_2 \dots s_k$ where s_i are decreasing sequences and, for every i, all terms of s_i are smaller than those of s_{i+1}.

In general these enumerative problems are difficult and an exact formula for the numbers $S_n(p)$ or for their generating function or even the asymptotics is known only in few cases. One of them is when p is a permutation of $\{1, 2, 3\}$. Then ([50])

$$S_n(p) = \frac{1}{n+1}\binom{2n}{n}.$$

The numbers $S_n(p)$, defined more generally with a set R of forbidden permutations instead of p, were investigated systematicly first by Simion and Schmidt [50].

Partitions. Looking through the glasses of algebraic combinatorics, we can view sequences as set partitions. For example, *abcabc* is the partition of $\{1, 2, 3, 4, 5, 6\}$ into the parts $\{1, 4\}$, $\{2, 5\}$, and $\{3, 6\}$. In general, to $u = a_1 a_2 \dots a_l$ corresponds partition of $[l]$ in which i and j are in the same part iff $a_i = a_j$. The blocks of mutually isomorphic sequences of length l are in 1-1 correspondence with partitions of $[l]$. Representing partitions by their equivalence relations, we can formulate containment of sequences $\subset$ in this way: for two partitions $u = ([k], \sim_u)$ and $v = ([l], \sim_v)$, the relation $u \subset v$ means that there is an *increasing* injection $f : [k] \to [l]$ such that, for every $x, y \in [k]$, $x \sim_u y \iff f(x) \sim_v f(y)$.

Partitions not containing partition (sequence) *abab* are standardly called *noncrossing*. Simion wrote an interesting survey [49] (with over 170 references) of results on noncrossing partitions and related topics. The seminal work dealing with them was of Kreweras [35], followed shortly by Poupard [41]. In the same year 1972 Mullin and Stanton published independently article [39] on enumeration of 4-DS sequences. These are noncrossing partitions in a mild disguise. Other enumerative works on 4-DS sequences are Roselle [43], Gardy and Gouyou-Beauchamps [17], and Klazar [26].

It's time to criticize also others a little bit. Although the curious MR review of [29] opens by "The author improves previous results to show that the number $N_5(n)$ of finite sequences ..." (compare the title!), unfortunately, to my knowledge, no enumerative results on 5-DS sequences are known. Some should be discovered! (Bounds on the lengths of DS sequnces are used to obtain bounds in an *enumerative* problem in [6] where Alon and Onn, to bound numbers of separable partitions on the moment curve, apply bounds (2) and (7).)

Problem 14. What are the numbers a_l of *ababa*-free partitions of $[l]$? (In other words, of normal *ababa*-free sequences u, $|u| = l$.) What is the asymptotics of a_l when $l \to \infty$? $\Box$

Note that a_l grow superexponentially: since every partition of $[l]$ into two-element parts (suppose l is even) avoids *ababa*, we have $a_l \ge 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (l-1)$.

With the variety of enumerative formulas for noncrossing partitions in mind, we can state the following problem.

Problem 15. What are the numbers b_l of *abcabc*-free partitions of $[l]$? What is the asymptotics of b_l when $l \to \infty$? $\Box$

It follows from the second bound in (36) (proved in [30]) that $b_l < c^l$ for a constant $c > 1$. (For this particular bound one can probably devise a more direct and simpler argument.) Since $b_{k+l} \ge b_k b_l$ for every k and l, $b_l^{1/l}$ tends to a finite limit.

Hypergraphs. The containment of partitions is a special case of the hypergraph containment introduced in Klazar [30]. We survey some of the results and omissions of this article.

A *hypergraph* $\mathcal{H} = (H_i : \ i \in I)$ is a finite list of finite nonempty subsets H_i of $\mathbf{N} = \{1, 2, \ldots\}$, called *edges*. *Simple* hypergraphs have no repeated edges. Two edges H_1 and H_2 of $\mathcal{H}$ are *separated* if $x < y$ holds for every $x \in H_1$ and $x \in H_2$. The elements of $\bigcup \mathcal{H} = \bigcup_{i=1}^{n} H_i \subset \mathbf{N}$ are called *vertices*. We say that $\mathcal{H} = (H_i : \ i \in I)$ and $\mathcal{H}' = (H_i' : \ i \in I')$ are *isomorphic* if there is an *increasing* (with respect to the standard linear ordering of $\mathbf{N}$) bijection $F : \bigcup \mathcal{H} \to \bigcup \mathcal{H}'$ and a bijection $f : I \to I'$ such that, for every $i \in I$, $F(H_i) = H'_{f(i)}$. We say that $\mathcal{H}'$ is a *reduction* of $\mathcal{H}$ if $I' \subset I$ and, for every $i \in I'$, $H_i' \subset H_i$. Finally, if $\mathcal{H}$ has a reduction isomorphic to $\mathcal{H}'$, we say that $\mathcal{H}$ *contains* $\mathcal{H}'$ and write $\mathcal{H} \supset \mathcal{H}'$. Else we say that $\mathcal{H}$ is $\mathcal{H}'$-*free*. For example, $\mathcal{H}$ is $(\{1\}_1, \{1\}_2)$-free iff its edges are pairwise disjoint, that is, $\mathcal{H}$ is a partition.

This setup is general enough to include as particular cases several important classes of structures. If the edges of $\mathcal{H}$ are pairwise disjoint (in particular, $\mathcal{H}$ is simple), we obtain *partitions*. Partitions in which every edge has two elements are *matchings*. Even more specificly, if $\mathcal{H}$ is a matching and has no two separated edges, we obtain *permutations*. Indeed, then a unique permutation $p = a_1a_2 \dots a_k$ of $[k]$ exists so that $\mathcal{H}$ is isomorphic to the *permutation hypergraph*

$$\mathcal{H}_p = (\{i, k + a_i\} : \ i = 1, \dots, k).$$

For example, $\mathcal{H}_{132} = (\{1, 4\}, \{2, 6\}, \{3, 5\})$. It is easy to check that the hypergraph containment reduces for partitions to the partition containment, that is, to the containment of sequences. Also, for permutation hypergraphs it reduces to the containment of permutations discussed above. Simple hypergraphs having two-element edges with allowed intersections constitute the important class of *graphs*. If simplicity is dropped, we obtain *multigraphs*.

We measure the size of $\mathcal{H} = (H_i : \ i \in I)$ in three ways: by the number of vertices $v(\mathcal{H}) = |\bigcup \mathcal{H}|$, by the number of edges $e(\mathcal{H}) = |I|$, and by the number of incidences between vertices and edges $i(\mathcal{H}) = \sum_{i \in I} |H_i|$.

For a fixed hypergraph $\mathcal{F}$, one can ask two questions about $\mathcal{F}$-free hypergraphs $\mathcal{H}$: (i) how many such $\mathcal{H}$s are there and (ii) how big such $\mathcal{H}$s may be.

Let us begin with the enumerative question. Given a hypergraph $\mathcal{F}$ and a number $n \in \mathbf{N}$, what is the number

$$a(\mathcal{F}, n) = \#\{\mathcal{H} : \ \mathcal{H} \text{ is simple } \& \bigcup \mathcal{H} = [n] \ \& \ \mathcal{H} \not\succ \mathcal{F}\},$$

that is, of nonisomorphic simple $\mathcal{F}$-free hypergraphs $\mathcal{H}$ with n vertices? Recalling the commentary following Problem 14, we are not surprised that for most $\mathcal{F}$ numbers $a(\mathcal{F}, n)$ grow superexponentially. If $\mathcal{F}$ has an edge with at least three elements or two intersecting edges or two separated two-element edges, then $a(\mathcal{F}, n) \geq (\lfloor n/2 \rfloor)! = \mathrm{e}^{(1+o(1))\frac{1}{2}n \log n}$ because every permutation hypergraph $\mathcal{H}_p$ with n vertices (or $\mathcal{H}_p$ on $n-1$ vertices plus a singleton edge if n is odd) is then $\mathcal{F}$-free. It is clear that $\mathcal{F}$ escapes this disjunction if and only if it is isomorphic to $\mathcal{H}_p \cup \mathcal{H}'$ where p is a permutation, $\mathcal{H}'$ has only singleton edges and is simple, and $\bigcup \mathcal{H}_q \cap \bigcup \mathcal{H}' = \emptyset$. We say briefly that $\mathcal{F}$ is of the form $\mathcal{H}_p$ *with singletons*.

In [30] we proved that for every permutation p there exists an extremely slowly growing function $\beta_p(n)$, defined in terms of the inverse Ackermann function $\alpha(n)$, such that the number of simple $\mathcal{H}_p$-free hypergraphs with n vertices satisfies

$$a(\mathcal{H}_p, n) < \beta_p(n)^n. \tag{30}$$

This strengthens (29) because now many more $\mathcal{H}_p$-free objects are included in the count. Since every $\mathcal{H}_p$ with singletons is contained in an $\mathcal{H}_q$, bound (30) holds for $\mathcal{H}_p$ with singletons as well. Hence in the enumerative problem we have this alternative:

$$a(\mathcal{F}, n) \left\{ \begin{array}{lll} < \beta_p(n)^n & \ldots & \mathcal{F} \subset \mathcal{H}_p \text{ for a } p \\ & & \\ > \mathrm{e}^{(1+o(1))\frac{1}{2} n \log n} & \ldots & \mathcal{F} \not\subset \mathcal{H}_p \text{ for every } p. \end{array} \right\} \tag{31}$$

We look now at the extremal problem. For any fixed hypergraph $\mathcal{F}$ we introduce two extremal functions

$$\begin{aligned} H_e(\mathcal{F}, n) &= \max\{e(\mathcal{H}) : \ \mathcal{H} \not\supset \mathcal{F} \ \& \ \mathcal{H} \text{ is simple } \& \ v(\mathcal{H}) \le n\} \\ H_i(\mathcal{F}, n) &= \max\{i(\mathcal{H}) : \ \mathcal{H} \not\supset \mathcal{F} \ \& \ \mathcal{H} \text{ is simple } \& \ v(\mathcal{H}) \le n\}. \end{aligned}$$

Let us compare this with definition (16). $H_e(\mathcal{F}, n)$ and $H_i(\mathcal{F}, n)$ are defined in the "traditional" way of extremal set theory: the number of edges should be bounded in terms of the number of vertices. In (16), contrarywise, the number of vertices should be bounded in terms of the number of edges (i.e. parts of a partition). Since vertices are linearly ordered but edges are not, the simple incidence duality, which interchanges vertices and edges, does not apply here. We investigate $H_e(\mathcal{F}, n)$ and $H_i(\mathcal{F}, n)$ in [32].

In [30] we proved, by means of (18), an analogue of (30): for every permutation p,

$$H_i(\mathcal{H}_p, n) < n \cdot \gamma_p(n), \tag{32}$$

where again $\gamma_p(n)$ is defined in terms of $\alpha(n)$ and thus grows to infinity extremely slowly.

Results of Füredi and Hajnal. Here we must mention important work [16] of Füredi and Hajnal on extremal problems closely related to ours. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two simple bipartite graphs with linearly ordered parts $[m], [n]'$ and $[k], [l]'$, respectively; here $[n]' = \{1', 2', \ldots, n'\}$ and $[n]'$ is ordered in the obvious way. If there are increasing injections $F : [k] \to [m]$ and $G : [l]' \to [n]'$ mapping every edge $\{i, j'\}$ of $\mathcal{G}_2$ to an edge $\{F(i), G(j')\}$ of $\mathcal{G}_1$, we say that $\mathcal{G}_1$ *contains* $\mathcal{G}_2$ and write $\mathcal{G}_1 \supset \mathcal{G}_2$. In [16] the extremal function

$$f(m, n; \mathcal{G}_2) = \max\{e(\mathcal{G}_1) : \ \mathcal{G}_1 \text{ has parts } [m], [n]' \text{ and } \mathcal{G}_1 \not\supset \mathcal{G}_2\}$$

is investigated; Füredi and Hajnal use the equivalent language of 0-1 matrices. Besides many other results, they prove bounds

$$n\alpha(n) \ll f(n, n; (\{1, 1'\}, \{1, 3'\}, \{2, 2'\}, \{2, 4'\})) \ll n\alpha(n). \tag{33}$$

The upper bound is obtained from (10) by an easy reduction to 5-DS sequences. For the lower bound they give a construction of their own. In the end of [16] they ask whether $f(n,n;\mathcal{G}) \ll_{\mathcal{G}} n$ holds for all permutation bipartite graphs $\mathcal{G}$ (these are $\mathcal{G}$ with pairwise disjoint edges).

Back to hypergraphs. Trivially, $H_e(\mathcal{F},n) \le H_i(\mathcal{F},n)$ for every $\mathcal{F}$ and n. We show that in most cases, certainly for every $\mathcal{F} = \mathcal{H}_p$, the opposite inequality holds in a weaker form as well.

Example 9. Suppose $\mathcal{F}$ does not have two separated edges, $p = v(\mathcal{F})$, and $q = e(\mathcal{F}) > 1$ (case $q = 1$ is trivial). Then for every $n \in \mathbf{N}$,

$$H_i(\mathcal{F},n) \le (2p-1)(q-1)H_e(\mathcal{F},n). \tag{34}$$

For the proof suppose that $\mathcal{H}$ is a simple and $\mathcal{F}$-free hypergraph with $v(\mathcal{H}) \le n$. We transform $\mathcal{H}$ in a new hypergraph $\mathcal{H}'$ by keeping all edges with less than p vertices and replacing every edge $H = \{v_1, v_2, \ldots, v_s\}$ of $\mathcal{H}$ with $s \ge p$, where $v_1 < v_2 < \cdots < v_s$, by $t = \lfloor |H|/p \rfloor$ new p-element edges $\{v_1, \ldots, v_p\}$, $\{v_{p+1}, \ldots, v_{2p}\}, \ldots,$ $\{v_{(t-1)p+1}, \ldots, v_{tp}\}$. $\mathcal{H}'$ may not be simple and therefore we define a simple hypergraph $\mathcal{H}''$ by keeping from every family of repeated edges of $\mathcal{H}'$ only one edge. Two observations: (i) no edge of $\mathcal{H}'$ repeats q or more times and (ii) $\mathcal{H}''$ is $\mathcal{F}$-free. If (i) were false, there would be q distinct edges $H_1, \ldots, H_q$ in $\mathcal{H}$ such that $|\bigcap_{i=1}^q H_i| \ge p$. But this implies the contradiction $\mathcal{F} \subset \mathcal{H}$. As for (ii), realize that any $\mathcal{F}$-copy in $\mathcal{H}''$ may use for every $H \in \mathcal{H}$ only at most one new edge born from H, so it is also an $\mathcal{F}$-copy in $\mathcal{H}$. The observations and definitions of $\mathcal{H}'$ and $\mathcal{H}''$ imply

$$\begin{aligned} i(\mathcal{H}) &\le \frac{(2p-1)i(\mathcal{H}')}{p} \le \frac{(2p-1)(q-1)i(\mathcal{H}'')}{p} \\ &\le (2p-1)(q-1)e(\mathcal{H}'') \\ &\le (2p-1)(q-1)H_e(\mathcal{F},n). \end{aligned}$$

□

Note that $H_e((\{1\},\{2\}),n) = 1$ but $H_i((\{1\},\{2\}),n) = n$. In the same vein, for $\mathcal{F} = (\{1\},\{2\},\{3\})$, $H_e(\mathcal{F},n) = 3$ for $n > 1$ (curiously, value 3 is attained only by the two-vertex hypergraph $\mathcal{H} = (\{1,2\},\{1\},\{2\})$) but $H_i(\mathcal{F},n) = 2n-1$ for $n \ne 2$ and $= 4$ for $n = 2$. In general it may happen that $H_i(\mathcal{F},n)$ is a substantially larger function than $H_e(\mathcal{F},n)$. But these examples are certainly somewhat singular.

Problem 16. What are the hypergraphs $\mathcal{F}$ such that $H_i(\mathcal{F}, n) \ll H_e(\mathcal{F}, n)$ is false? Is there an $\mathcal{F}$ with this property for which $H_e(\mathcal{F}, n)$ is not eventually constant? $\square$

We missed inequality (34) of Example 9 in [30] and made the proof of (32) more complicated than needed. In particular, (34) renders Theorem 2.4 of [30] and its proof completely unnecessary.

It turns out that if we are content only with upper bounds, the extremal problem is more fundamental than the enumerative one. In [30] first (32) is proved and from it, as a corollary, (30) is derived, by means of an inductive argument. This simple but powerful argument was discovered by Alon and Friedgut in [5] and it was further simplified and made more transparent in Klazar [31]. Indeed, as we show in the next example, the argument is almost trivial once it is discovered and formulated properly.

Example 10 ([31]). Recall that $S_n(p)$ is the number of permutations of $[n]$ avoiding a fixed permutation p of $[k]$. To p corresponds in an obvious way a bipartite graph, in fact a matching, $\mathcal{G}_p$ on parts $[k], [k]'$. Let $G_p(n)$ be the number of all simple bipartite graphs $\mathcal{G}$ on $[n], [n]'$ such that $\mathcal{G} \not\succ \mathcal{G}_p$. (We use here the setting from the above paragraph on results of [16].) Clearly,

$$S_n(p) \le G_p(n)$$

because now we count many more p-free objects than just permutations (which is the trick). Let $m = \lceil n/2 \rceil$. Suppose we know that $e(\mathcal{G}') < m\beta(m)$ holds for every $\mathcal{G}'$ counted by $G_p(m)$, where $\beta(m)$ is a nondecreasing function. We claim that this implies the inequality

$$G_p(n) < 15^{m\beta(m)} \cdot G_p(m). \tag{35}$$

To see this, associate with every $\mathcal{G}$ counted by $G_p(n)$ a $\mathcal{G}'$ counted by $G_p(m)$: $\{i, j'\}$, where $i, j \in [m]$, is an edge of $\mathcal{G}'$ iff in $\mathcal{G}$ there is at least one edge connecting sets $\{2i-1, 2i\}$ and $\{2j'-1, 2j'\}$. In other words, we identify in $\mathcal{G}$ the vertices in each of the pairs $(1,2), (3,4), \ldots; (1', 2'), (3', 4'), \ldots$ and replace arising multiple edges by simple edges. Observe that $\mathcal{G} \not\succ \mathcal{G}_p$ implies $\mathcal{G}' \not\succ \mathcal{G}_p$. If $\mathcal{G}'$ has e edges, there are at most 15^e graphs $\mathcal{G}$ that transform to $\mathcal{G}'$ because every edge of $\mathcal{G}'$ can arise from $\mathcal{G}$ in at most $2^{2\cdot 2} - 1 = 15$ ways. Thus we obtain (35).

Iterating (35) until $m = 1$, we obtain the bound

$$S_n(p) \le G_p(n) < 15^{2n\beta(n)} = 225^{n\beta(n)}.$$

If $\beta(n)$ is almost constant we obtain an almost exponential bound on $S_n(p)$. Such a bound on $\beta(n)$ can be obtained without difficulties from (18); see [31] for more details. Of course, if a constant bound on $\beta(n)$ could be obtained, we would have an exponential bound for $S_n(p)$. $\square$

Is it possible to prove for the extremal problem an alternative analogous to (31)? If $\mathcal{F}$ has an edge with at least three elements or two separated two-element edges or a 2-path isomorphic to $(\{1,2\},\{2,3\})$, then the complete bipartite graph with parts $\{1,2,\ldots,\lfloor n/2\rfloor\}$ and $\{\lfloor n/2\rfloor+1,\ldots,n-1,n\}$ shows that $H_e(\mathcal{F},n)\geq(\lfloor n/2\rfloor)^2$. However, already

$$H_e((\{1\}_1,\{1\}_2),n)=H_i((\{1\}_1,\{1\}_2),n)=n$$

shows that intersections in $\mathcal{F}$ may be compatible with linear (or almost linear) growth of $H_e(\mathcal{F},n)$ and $H_i(\mathcal{F},n)$. A slightly more complicated example: show that for $\mathcal{F}=(\{1,2\},\{1,3\})$ we have $H_e(\mathcal{F},n)=2n-1$ and $H_i(\mathcal{F},n)=3n-2$.

Problem 17. Identify the class of hypergraphs Ψ such that

$$H_e(\mathcal{F},n),H_i(\mathcal{F},n)\left\{\begin{array}{lll}<n\cdot\gamma(n) & \ldots & \mathcal{F}\in\Psi\\ \\ >n\cdot\delta(n) & \ldots & \mathcal{F}\not\in\Psi\end{array}\right.$$

holds, where $\gamma(n)\to\infty$ but is almost constant, defined in terms of $\alpha(n)$, and $\delta(n)\to\infty$ much faster than $\gamma(n)$; both functions depend on $\mathcal{F}$ as well. $\square$

Note that the analogous class in (31) consists exactly of all permutation hypergraphs $\mathcal{H}_p$ with singletons.

In [30] we conjecture that in (30) and (32) functions $\beta_p(n)$ and $\gamma_p(n)$ may be replaced by constants depending only on p.

Problem 18. Prove or disprove that for every permutation p we have $H_i(\mathcal{H}_p,n)\ll_p n$. $\square$

If this is true, $a(\mathcal{H}_p,n)<c_p^n$ for every p follows, by the argument of [30], strengthening the Stanley–Wilf conjecture. In actuality alternative (31) is probably sharper, with a constant c_p in the place of $\beta_p(n)$.

In the extremal problem, for the class Ψ, situation is more complicated. Consider the graph $\mathcal{G}=(\{1,3\},\{1,5\},\{2,4\},\{2,6\})$. By the lower bound in

(33), $H_e(\mathcal{G}, n) \gg n\alpha(n)$. On the other hand, $\mathcal{G} \in \Psi$ ([32]). Hence Ψ does contain superlinear hypergraphs.

In [30] the conjectured linear and exponential bounds

$$H_i(\mathcal{H}_p, n) \ll_p n \ \text{ and } \ a(\mathcal{H}_p, n) < c_p^n \tag{36}$$

were proved, by means of the N-sequence bound (26), for a special class of permutations when $p = a_1 a_2 \dots a_k$ first increases and then decreases or the opposite pattern is realized by p^{-1}.

We conclude our survey by return to the pattern *abab* alias ●–○–●–○ alias $(\{1,3\},\{2,4\})$.

Example 11 ([30]). We prove that for $\mathcal{F} = (\{1,3\},\{2,4\})$ and every $n > 1$,

$$H_e(\mathcal{F}, n) = 4n - 5 \ \text{ and } \ H_i(\mathcal{F}, n) = 8n - 12.$$

Let $\mathcal{G}$ be first a simple $(\{1,3\},\{2,4\})$-free graph with vertex set $[n]$. We prove by induction on n that $e(\mathcal{G}) \le 2n - 3$. For $n = 2$ this is true. Let $n > 2$. We split $[n]$ into two overlapping intervals $I_1 = [1, k]$ and $I_2 = [k, n]$, where k, $k < n$, is the last vertex connected by an edge to 1 if $\{1, n\} \notin \mathcal{G}$ and k is the second last such vertex otherwise. The restrictions of $\mathcal{G}$ to I_1 and I_2 are simple $(\{1,3\},\{2,4\})$-free graphs and every edge of $\mathcal{G}$ except possibly $\{1, n\}$ lies in I_1 or I_2. By induction,

$$e(\mathcal{G}) \le 1 + 2|I_1| - 3 + 2|I_2| - 3 = 2(|I_1| + |I_2|) - 5 = 2n - 3.$$

Thus $e(\mathcal{G}) \le 2v(\mathcal{G}) - 3$ in the graph case.

Let now $\mathcal{H}$ be a simple $(\{1,3\},\{2,4\})$-free hypergraph with vertex set $[n]$. We look on the *big* edges of $\mathcal{H}$ having 3 and more vertices. We claim that after deleting from each of them its first and last vertex, the resulting sets (i) lie all in $[2, n-1]$ and (ii) are mutually disjoint. The first claim is clear. If (ii) were not true, we would have two distinct edges $H_1, H_2 \in \mathcal{H}$ and five not necessarily distinct vertices $v_1, \dots, v_5 \in \bigcup \mathcal{H}$ such that $v_2 < v_3 < v_4$, $v_1 < v_3 < v_5$, $\{v_1, v_5\} \subset H_1$, $\{v_2, v_4\} \subset H_2$, and $v_3 \in H_1 \cap H_2$. Moreover, we may assume that $v_1 \ne v_2$ or $v_4 \ne v_5$ because $H_1 \ne H_2$ ($\mathcal{H}$ is simple). But then $\mathcal{H} \supset (\{1,3\},\{2,4\})$, a contradiction. Thus the resulting sets must be disjoint and their number is at most $n - 2$, which bounds the number of big edges in $\mathcal{H}$.

Not forgetting about singleton edges, we conclude that $\mathcal{H}$ has at most $n + (2n-3) + (n-2) = 4n - 5$ edges. Sum of their cardinalities is at most $n + 2(2n-3) + (n-2) + 2(n-2) = 8n - 12$. The hypergraphs

$$(\{i\}, \{j, j+1\}, \{1, k\}, \{1, j, j+1\} : \ i \in [n], j \in [2, n-1], k \in [2, n])$$

show that these bounds are tight. $\square$

Interestingly, the previous argument works for the *abba* pattern $(\{1,4\},\{2,3\})$ as well. The only adjustment must be made in the graph part, where the upper bound remains $2n-3$ but for a different reason. The second part, bounding the number of big edges, needs no changes. The hypergraphs

$$(\{i\}, \{1, n\}, \{1, j\}, \{j, n\}, \{1, j, n\} : \ i \in [n], j \in [2, n-1])$$

show that the upper bounds $4n-5$ on $e(\mathcal{H})$ and $8n-12$ on $i(\mathcal{H})$ are again tight. Hence $\mathcal{F} = (\{1,4\},\{2,3\})$ and $\mathcal{F} = (\{1,3\},\{2,4\})$ have the same extremal functions

$$H_e(\mathcal{F}, n) = 4n - 5 \quad \text{and} \quad H_i(\mathcal{F}, n) = 8n - 12.$$

This contrasts with the situation for sequences and colored trees where extremal functions of *abab* and *abba* differ: $\mathrm{Ex}(abab, n) = 2n-1$, $\mathrm{Ex}(abba, n) = 3n-2$, $T(abab, n) = 2n-1$, and $5n-8 \le T(abba, n) \le 48n$ (Examples 2, 5, 7, and 8).

Acknowledgments. I am indebted to Jiří Matoušek for reading the manuscript and for his valuable comments.

References

[1] R. Adamec, M. Klazar and P. Valtr, Generalized Davenport–Schinzel sequences with linear upper bound, *Discrete Mathematics,* **108** (1992), 219–229. [Zentralblatt für Mathematik und ihre Grenzgebiete 768.05007 and Mathematical Reviews 93k:05015.]

[2] P. K. Agarwal, B. Aronov, J. Pach, R. Pollack and M. Sharir, Quasi-planar graphs have a linear number of edges, *Combinatorica,* **17** (1997), 1–9. [Zbl 880.05050 and MR 98f:05046.]

[3] P. K. Agarwal and M. Sharir, Davenport–Schinzel sequences and their geometric applications, in: J.-R. Sacks and J. Urrutia (editor(s)), *Handbook of Computational Geometry,* Elsevier, Amsterdam, 2000, pp. 1–47. [Zbl 991.24846 and MR CMP 1 746 674.]

[4] P. K. Agarwal, M. Sharir and P. Shor, Sharp upper and lower bounds on the length of general Davenport–Schinzel sequences, *Journal of Combinatorial Theory, Series A,* **52** (1989), 228–274. [Zbl 697.05003 and MR 90m:11034.]

[5] N. Alon and E. Friedgut, On the number of permutations avoiding a given pattern, *J. Comb. Theory, Ser. A,* **89** (2000), 133–140. [Zbl 991.47660 and MR 2000i:05007.]

[6] N. Alon and S. Onn, Separable partitions, *Discrete Applied Mathematics,* **91** (1999), 39–51. [Zbl 926.05011 and MR 2000j:05010.]

[7] M. I. Atallah, Some dynamic computational geometry problems, *Computers and Mathematics with Applications,* **11** (1985), 1171–1181. [Zbl 586.68085 and MR 87b:68095.]

[8] B. J. Birch, H. Halberstam and C. A. Rogers (editors), *The Collected Works of Harold Davenport. Volume IV,* Academic Press, London, 1977. [Zbl 383.01022 and MR 58#27141.]

[9] M. Bóna, The solution of a conjecture of Stanley and Wilf for all layered patterns, *J. Comb. Theory, Ser. A,* **85** (1999), 133–140. [Zbl 919.05002 and MR 99i:05005.]

[10] F. J. Burkowski and E. F. Ecklund, Jr., Exclusion sequences with regularity conditions, in: F. Hoffman et al. (editor(s)), *The Fifth Southeastern Conference on Combinatorics, Graph Theory and Computing, Boca Raton 1974 (USA), Congressus Numerantium X,* Utilitas Mathematica Publishing, Inc., Winnipeg, 1974, pp. 261–266. [Zbl 313.05004 and MR 52#2901.]

[11] H. Davenport, A combinatorial problem connected with differential equations II, *Acta Arithmetica,* **17** (1971), 363–372. [Zbl 216.30204 and MR 44#2619.]

[12] H. Davenport and A. Schinzel, A combinatorial problem connected with differential equations, *American Journal of Mathematics,* **87** (1965), 684–694. [Zbl 132.00601 and MR 32#7426.]

[13] H. Davenport and A. Schinzel, A note on sequences and subsequences, *Elemente der Mathematik,* **20** (1965), 63–64. [Zbl 132.25101 and MR — not reviewed.]

[14] A. J. Dobson and S. O. Macdonald, Lower bounds for the lengths of Davenport–Schinzel sequences, *Utilitas Mathematica,* **6** (1974), 251–257. [Zbl 299.05010 and MR 50#9781.]

[15] H. Edelsbrunner and M. Sharir, The maximum number of ways to stab n convex nonintersecting sets in the plane is $2n-2$, *Discrete and Computational Geometry,* **5** (1990), 35–42. [Zbl 712.52009 and MR 90k:52019.]

[16] Z. Füredi and P. Hajnal, Davenport–Schinzel theory of matrices, *Discrete Math.,* **103** (1992), 233–251. [Zbl 776.05024 and MR 93g:05026.]

[17] D. Gardy and D. Gouyou-Beauchamps, Enumerating Davenport-Schinzel sequences, *RAIRO. Informatique Théorique et Applications,* **26** (1992), 387–402. [Zbl 769.05007 and MR 93k:05011.]

[18] R. K. Guy, *Unsolved Problems in Number Theory,* Springer, Berlin, 1981. [Zbl 474.10001 and MR 83k:10002.] (Second edition appeared in 1994.)

[19] S. Hart and M. Sharir, Nonlinearity of Davenport–Schinzel sequences and of generalized path compression schemes, *Combinatorica,* **6** (1986), 151–177. [Zbl 636.05003 and MR 88f:05004.]

[20] M. Katchalski and H. Last, On geometric graphs with no two edges in convex position, *Discrete Comput. Geom.,* **19** (1998), 399–404. [Zbl 908.05033 and MR 99d:52015.]

[21] M. Klazar, A general upper bound in extremal theory of sequences, *Commentationes Mathematicae Universitatis Carolinae,* **33** (1992), 737–746. [Zbl 781.05049 and MR 94g:05095.]

[22] M. Klazar, Two results on a partial ordering of finite sequences, *Commentat. Math. Univ. Carol.,* **34** (1993), 697–705. [Zbl 789.05090 and MR 95i:05108.]

[23] M. Klazar, A linear upper bound in extremal theory of sequences, *J. Comb. Theory, Ser. A,* **68** (1994), 454–464. [Zbl 808.05096 and MR 95m:11033.]

[24] M. Klazar, *Combinatorial Aspects of Davenport–Schinzel Sequences*, Ph.D. thesis, Charles University, Prague, 1995.

[25] M. Klazar, Extremal functions for sequences, *Discrete Math.,* **150** (1996), 195–203. [Zbl 851.05093 and MR 97c:05019.]

[26] M. Klazar, On *abab*-free and *abba*-free set partitions, *European Journal of Combinatorics,* **17** (1996), 53–68. [Zbl 840.05004 and MR 96j:05019.]

[27] M. Klazar, Combinatorial aspects of Davenport–Schinzel sequences, *Discrete Math.,* **165-166** (1997), 431–445. [Zbl 948.05008 and MR 98d:05149.]

[28] M. Klazar, Extremal problems for colored trees and Davenport-Schinzel sequence, *Discrete Math.,* **197-198** (1999), 469–482. [Zbl 936.05060 and MR 99k:05004.]

[29] M. Klazar, On the maximum lengths of Davenport–Schinzel sequences, in: R. L. Graham et al. (editor(s)), *Contemporary Trends in Discrete Mathematics, Štiřín castle 1997 (Czech Republic),* American Mathematical Society, Providence RI, 1999, pp. 169–178. [Zbl 938.05004 and MR 2000h:11023.]

[30] M. Klazar, Counting pattern-free set partitions. II: Noncrossing and other hypergraphs, *Electronic Journal of Combinatorics,* **7** (2000), R34. [Zbl 944.05002 and MR CMP 1 769 065.]

[31] M. Klazar, The Füredi–Hajnal conjecture implies the Stanley–Wilf conjecture, in: D. Krob, A. A. Mikhalev and A. V. Mikhalev (editor(s)), *Formal Power Series and Algebraic Combinatorics, Moscow 2000 (Russia),* Springer, Berlin, 2000, pp. 250–255. [Zbl 954.05048 and MR CMP 1 798 218.]

[32] M. Klazar, Extremal problems for hypergraphs with linearly ordered vertex sets, in preparation.

[33] M. Klazar and P. Valtr, Generalized Davenport–Schinzel sequences, *Combinatorica,* **14** (1994), 463–476. [Zbl 812.05067 and MR 96h:05206.]

[34] P. Komjáth, A simplified construction of nonlinear Davenport–Schinzel sequences, *J. Comb. Theory, Ser. A,* **49** (1988), 262–267. [Zbl 673.05001 and MR 89m:11019.]

[35] G. Kreweras, Sur les partitions non croisées d'un cycle, *Discrete Math.,* **1** (1972), 333–350. [Zbl 231.05014 and MR 46#8852.]

[36] M. Loebl and J. Nešetřil, Fast and slow growing (a combinatorial study of unprovability), in: A. D. Keedwell (editor(s)), *Surveys in Combinatorics, 1991,* Cambridge University Press, Cambridge, UK, 1991, pp. 119–160. [Zbl 735.03017 and MR 93d:03061.] (London Mathematical Society Lecture Note Series, 166.)

[37] W. H. Mills, Some Davenport–Schinzel sequences, in: R. S. D. Thomas and H. C. Williams (editor(s)), *The 3rd Manitoba Conference on Numerical Mathematics, Winnipeg 1973 (USA), Congressus Numerantium IX,* Utilitas Mathematica Publishing Inc., Winnipeg, 1974, pp. 307–314. [Zbl 321.05006 and MR 50#135.]

[38] W. H. Mills, On Davenport–Schinzel sequences, *Utilitas Math.,* **9** (1976), 87–112. [Zbl 337.05006 and MR 53#10703.]

[39] R. C. Mullin and R. G. Stanton, A map-theoretic approach to Davenport–Schinzel sequences, *Pacific Journal of Mathematics,* **40** (1972), 167–172. [Zbl 233.05013 and MR 46#1745.]

[40] C. R. Peterkin, Some results on Davenport–Schinzel sequences, in: R. S. D. Thomas and H. C. Williams (editor(s)), *The 3rd Manitoba Conference on Numerical Mathematics, Winnipeg 1973 (USA), Congressus Numerantium IX,* Utilitas Mathematica Publishing Inc., Winnipeg, 1974, pp. 337–344. [Zbl 324.05008 and MR 50#136.]

[41] Y. Poupard, Étude et dénombrement parallèles des partitions non croisées d'un cycle et des découpages d'un polygone convexe, *Discrete Math.,* **2** (1972), 279–288. [Zbl 252.05005 and MR 46#3369.]

[42] B. C. Rennie and A. J. Dobson, Upper bounds for the lengths of Davenport–Schinzel sequences, *Utilitas Math.,* **8** (1975), 181-185. [Zbl 364.05004 and MR 52#13624.]

[43] D. P. Roselle, An algorithmic approach to Davenport–Schinzel sequences, *Utilitas Math.,* **6** (1974), 91–93. [Zbl 306.05010 and MR 50#9780.]

[44] D. P. Roselle and R. G. Stanton, Results on Davenport–Schinzel sequences, in: R. Mullin, K. B. Reid and D. P. Roselle (editor(s)), *Louisiana Conference on Combinatorics, Graph Theory and Computing, Baton Rouge 1970 (USA),* Louisiana State University, Baton Rouge, 1970, pp. 249–267. [Zbl 239.05004 and MR 43#68.]

[45] D. P. Roselle and R. G. Stanton, Some properties of Davenport–Schinzel sequences, *Acta Arith.,* **17** (1971), 355–362. [Zbl 215.33203 and MR 44#1641.]

[46] H. Sharir, Almost linear upper bounds on the length of general Davenport–Schinzel sequences, *Combinatorica,* **7** (1987), 131–143. [Zbl 636.05004 and MR 89f:11037.]

[47] H. Sharir, Improved lower bounds on the length of Davenport–Schinzel sequences, *Combinatorica,* **8** (1988), 117–124. [Zbl 672.05015 and MR 89j:11024.]

[48] M. Sharir and P. K. Agarwal, *Davenport–Schinzel Sequences and Their Geometric Applications,* Cambridge University Press, Cambridge, UK, 1995. [Zbl 834.68113 and MR 96e:11034.]

[49] R. Simion, Noncrossing partitions, *Discrete Math.,* **217** (2000), 367–409. [Zbl pre992.06522 and MR CMP 1 766 277.]

[50] R. Simion and F. Schmidt, Restricted permutations, *Eur. J. Comb.,* **6** (1985), 383–406. [Zbl 615.05002 and MR 88a:05006.]

[51] R. G. Stanton and P. H. Dirksen, Davenport–Schinzel sequences, *Ars Combinatoria,* **1** (1976), 43–51. [Zbl 335.05001 and MR 53#13106.]

[52] R. G. Stanton and D. P. Roselle, A result on Davenport–Schinzel sequences, in: P. Erdős, A. Rényi and V. T. Sós (editor(s)), *Combinatorial Theory and Its Applications, Balatonfüred 1969 (Hungary), Colloquia Mathematica Societatis János Bolyai 4, Vol. II,* Nort-Holland, Amsterdam, 1970, pp. 1023–1027. [Zbl 215.33202 and MR 46#3324.]

[53] E. Szemerédi, On a problem of Davenport and Schinzel, *Acta Arith.,* **25** (1974), 213–224. [Zbl 291.05003 and MR 49#244.]

[54] R. E. TARJAN, Efficiency of a good but not linear set union algorithm, *Journal of the Association for Computing Machinery,* **22** (1975), 215–225. [Zbl 307.68029 and MR 56#17194.]

[55] P. VALTR, Graph drawing with no k pairwise crossing edges, in: G. DiBattista (editor(s)), *Graph Drawing'97, Rome 1997 (Italy),* Springer, Berlin, 1997, pp. 205–218. [Zbl — not reviewed? and MR 99e:68166.] (Lecture Notes in Computer Science, 1353.)

[56] P. VALTR, On geometric graphs with no k pairwise parallel edges, *Discrete Comput. Geom.,* **19** (1998), 461-469. [Zbl 908.05035 and MR 99a:05138.]

[57] P. VALTR, Generalizations of Davenport–Schinzel sequences, in: R. L. Graham et al. (editor(s)), *Contemporary Trends in Discrete Mathematics, Štiřín castle 1997 (Czech Republic),* American Mathematical Society, Providence RI, 1999, pp. 349–389. [Zbl 942.68104 and MR 2001d:68185.]

[58] P. VALTR, On an extremal problem for colored trees, *Eur. J. Comb.,* **20** (1999), 115–121. [Zbl 924.05040 and MR 99k:05061.]

[59] P. VALTR, Davenport–Schinzel trees, submitted.

[60] A. WIERNIK AND M. SHARIR, Planar realizations of nonlinear Davenport–Schinzel sequences by segments, *Discrete Comput. Geom.,* **3** (1988), 15-47. [Zbl 636.68043 and MR 89h:11088.]

DEPARTMENT OF APPLIED MATHEMATICS (KAM)
CHARLES UNIVERSITY
AND
INSTITUTE FOR THEORETICAL COMPUTER SCIENCE (ITI)
MALOSTRANSKÉ NÁMĚSTÍ 25
118 00 PRAHA
CZECH REPUBLIC
e-mail: `klazar@kam.ms.mff.cuni.cz`