Discrete Mathematics 257 (2002) 385–403

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

On the maximum size of (p, Q)-free families

Zoltán Füredi[a,b,1], András Gyárfás[c,2], Miklós Ruszinkó[c,*,3]

[a]Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[b]Rényi Institute of Mathematics of the Hungarian Academy of Sciences, P.O. Box 127, Budapest 1364, Hungary

[c]Computer and Automation Research Institute, Hungarian Academy of Sciences, P.O. Box 63, Budapest 1518, Hungary

Received 24 May 2001; received in revised form 24 July 2001; accepted 11 September 2001

Dedicated to Daniel J. Kleitman on his 65th birthday

Abstract

Let p be a positive integer and let Q be a subset of $\{0, 1, \dots, p\}$. Call p sets $A_1, A_2, \dots, A_p$ of a ground set X a (p, Q)-system if the number of sets A_i containing x is in Q for every $x \in X$. In hypergraph terminology, a (p, Q)-system is a hypergraph with p edges such that each vertex x has degree $d(x) \in Q$. A family of sets $\mathscr{F}$ with ground set X is called (p, Q)-free if no p sets of $\mathscr{F}$ form a (p, Q)-system on X. We address the Turán-type problem for (p, Q)-systems: $f(n, p, Q)$ is defined as $\max|\mathscr{F}|$ over all (p, Q)-free families on the ground set $[n] = \{1, 2, \dots, n\}$. We study the behavior of $f(n, p, Q)$ when p and Q are fixed (allowing 2^{p+1} choices for Q) while n tends to infinity. The new results of this paper mostly relate to the middle zone where $2^{n-1} \leqslant f(n, p, Q) \leqslant (1 - c)2^n$ (in this upper bound c depends only on p). This direction was initiated by Paul Erdős who asked about the behavior of $f(n, 4, \{0, 3\})$. In addition, we give a brief survey on results and methods (old and recent) in the low zone (where $f(n, p, Q) = o(2^n)$) and in the high zone (where $2^n - (2 - c)^n < f(n, p, Q)$).

MSC: 05C65; 05D05

Keywords: Hypergraph; Degree; Regular

* Corresponding author.

E-mail addresses: z-furedi@math.uiuc.edu, furedi@renyi.hu (Z. Füredi), gyarfas@sztaki.hu (A. Gyárfás), ruszinko@lutra.sztaki.hu (M. Ruszinkó).

[1] Research supported in part by the Hungarian National Science Foundation under grant OTKA T032452, and by the National Science Foundation under Grant 1-5-29066 NSF DMS 00-70312.

[2] Research supported in part by OTKA Grant T029074.

[3] Research supported in part by OTKA Grants T029074, T030059, FKFP 0607 and KFKI-ISYS Datamining Lab.

PII: S0012-365X(02)00437-5

1. Introduction

The starting point of this paper was a question of Paul Erdős. Unfortunately, we had no opportunity to continue the conversations on the subject with him which started during the Summer of 1996. Following his well-known habit, he started with a special case: "how many subsets of $[n]=\{1,2,\ldots,n\}$ can you give if no four of them cover their union exactly three times? Can you give more than 2^{n-1} or only significantly less than 2^n?" He explained that it seems to be the first interesting special case of $f(n,p,q)$, the maximum number of sets of $[n]$ such that there are no p sets covering each element of their union exactly q times. He gave some comments as well: "we proved with Vera (T. Sós) that $f(n,3,2)=2^{n-1}+1$ and $f(n,5,3)$ are related to a number theory problem with Vera and Sárközy". Here, Paul have had referred to their paper on product representation [13].

We shall prove here that $f(n,p,q)/2^n \leqslant 1-c_{p,q}$ if $p/2\leqslant q$ and give a construction showing that $f(n,p,p-1)\geqslant 2^{n-1}+cn^{p-3}$ (Section 6). Probably $c_{p,q}$ tends to 1/2 for $p/2\leqslant q$. This seems to be a difficult problem, we could prove it only in a few special cases (in fact by Theorem 14 this is true for most choices of q in this zone). Section 9 is related to the original question of Erdős about $f(n,4,3)$.

We explore what is known and unknown about the following generalization of the initial problem. Assume that Q is an arbitrary subset of $\{0,1,\ldots,p\}$. Call p sets $A_1, A_2,\ldots,A_p$ of $[n]$ a (p,Q)-system if the number of sets A_i containing x is in Q for every, $x\in[n]$. Using hypergraph terminology,

Definition 1. A (p,Q)-system is a hypergraph with p edges such that each vertex x has degree $d(x)\in Q$. A family of sets $\mathscr{F}$ is (p,Q)-free if it does not contain any (p,Q)-system. Then $f(n,p,Q)$ is defined as $\max|\mathscr{F}|$ over all (p,Q)-free families $\mathscr{F}$ on $[n]$.

Notice that $f(n,p,q)=f(n,p,\{0,q\})$.

Sets with multiplicities are excluded, therefore, $f(n,p,Q)\leqslant 2^n$. We study the behavior of $f(n,p,Q)$, when p and Q are fixed (allowing 2^{p+1} choices for Q), while n tends to infinity. Many important results and problems of extremal set theory fit into this general Turán-type question (for various choices of p and Q). Also observe that, e.g., $f(n,p,\{0,1,p\})$, is the maximum size of a hypergraph containing no *Δ-system* with p edges, which is a classical (and still unsolved) problem, too. Parts of this paper give a survey of such results and methods.

2. Summary, classification of zones

In this section the 2^{p+1} possible choices of Q are classified into three categories, low, middle and high zones (plus some trivial cases). Also some simple but important examples are given. The names of the zones come from the facts that $f(n,p,Q)=o(2^n)$ in the low zone, $2^{n-1}\leqslant f(n,p,Q)\leqslant(1-c)2^n$ in the middle zone and $2^n - (2-c_1)^n < f(n,p,Q) < 2^n - (2-c_2)^n$ in the high zone. Details, proofs are subject of the following sections.

Trivial cases: We always suppose that $p \geqslant 2$. If $Q = \{0, 1, \ldots, p\}$, then every configuration is forbidden, $f(n,p,Q) = p-1$. It is easy to see (Claim 4) that in all other cases f is exponential in n

$$f(n,p,Q) > (1+\varepsilon)^n \quad \text{for every } Q \text{ with } Q \neq \{0,1,2,\ldots,p\}, \tag{1}$$

where $\varepsilon > 0$ depends only on p. Moreover, if $p > 2^n$ or Q is one of the sets $\emptyset$, $\{0\}$, $\{p\}$, or $\{0,p\}$, then nothing is forbidden, $f(n,p,Q) = 2^n$. We shall exclude these trivial cases.

Complementation: Considering the complements of a family $\mathscr{F}$, denoted by $\mathscr{F}^{\mathrm{C}}$ and defined as $\{[n] \backslash F\colon\ F \in \mathscr{F}\}$, one can see that $\mathscr{F}$ is (p,Q)-free if and only if $\mathscr{F}^{\mathrm{C}}$ is $(p,\bar{Q})$-free, where $\bar{Q} := \{p-q : q \in Q\}$. This implies that

$$f(n,p,Q) = f(n,p,\bar{Q}). \tag{2}$$

Therefore, we shall usually assume that Q has nonempty intersection with the interval $[0, p/2]$.

Monotonicity:

$$f(n,p,Q_1) \geqslant f(n,p,Q_2) \quad \text{for } Q_1 \subseteq Q_2. \tag{3}$$

2.1. Low zone

The low zone is defined by the sets Q for which $\{0,p\} \subseteq Q$. (In this zone $|Q| \geqslant 3$ since the trivial cases are excluded.) The name comes from the facts that by Theorem 5, we have

$$f(n,p,Q) = \mathrm{o}(2^n) \quad \text{for } \{0,p\} \subseteq Q, \tag{4}$$

while in all other cases $f \geqslant 2^{n-1}$. Indeed, considering the family of sets containing a fixed element x, $\mathscr{F}[x] := \{F\colon\ x \in F \subseteq [n]\}$, or the family of the sets avoiding x, $\mathscr{F}[x]^{\mathrm{C}} := \{F\colon\ x \notin F \subseteq [n]\}$ one obtains that

$$f(n,p,Q) \geqslant 2^{n-1} \quad \text{for } \{0,p\} \not\subseteq Q. \tag{5}$$

2.2. Middle zone

The middle zone is defined by the sets Q not in the low zone (i.e., $\{0,p\} \not\subseteq Q$) with a Q meeting both the lower and upper halves of the interval $[0,p]$,

$$[0, p/2] \cap Q \neq \emptyset \quad \text{and} \quad [p/2, p] \cap Q \neq \emptyset. \tag{6}$$

We shall prove (Theorem 15) that (6) implies

$$f(n,p,Q) \leqslant \left(1 - \frac{1}{2p} + \mathrm{o}(1)\right) 2^n. \tag{7}$$

We conjecture that more is true, in the middle zone extremal families have sizes $(1+o(1))2^{n-1}$:

Conjecture 2. In the middle zone

$$\lim_{n\to\infty} \frac{f(n,p,Q)}{2^n} = \frac{1}{2}.$$

In Sections 7–10 the reader can find several results supporting this conjecture.

2.3. High zone

In the rest of the cases Q meets only one of the open half-intervals. The high zone is defined by the sets Q for which Q is contained in either the initial or the final open half-interval,

$$Q \subseteq [0, p/2) \quad \text{or} \quad Q \subseteq (p/2, p].$$

It is easy to see, that in the high zone a (p,Q)-free family may contain almost all subsets of n. Suppose, say, that $q = \max Q < p/2$.

Proposition 3. *If* $q < p/2$ *then* $f(n,p,Q) \geqslant f(n,p,\{0,1,2,\dots,q\}) = (1-o(1))2^n$.

Proof. Let $\mathscr{F}_{q/p}$ consist of all subsets of size $\geqslant \lfloor qn/p \rfloor + 1$.

$$\mathscr{F}_{q/p} := \left\{ F \subseteq [n] : |F| > \frac{q}{p} n \right\}. \tag{8}$$

This family is $(p,\{0,1,\dots,q\})$-free. Indeed, take arbitrary $A_1,\dots,A_p$ members of $\mathscr{F}$. Since $|A_i| > qn/p$, the sum of the degrees of the elements in the collection $A_1,\dots A_p$ exceeds qn. Therefore—by the pigeonhole principle—there exists some $x \in [n]$ which is contained in at least $q+1$ members of the collection $A_1,\dots A_p$, which shows that $\mathscr{F}$ is $(p,\{0,1,\dots,q\})$-free. We have that there exists an $\varepsilon = \varepsilon(p) > 0$ such that

$$\sum_{i \leqslant (q/p)n} \binom{n}{i} \leqslant \sum_{i/n \leqslant 1/2 - 1/2p} \binom{n}{i} < (2-\varepsilon)^n.$$

This implies

$$f(n,p,Q) \geqslant 2^n - (2-\varepsilon)^n. \qquad \square$$

We shall see, that a result of Kleitman implies that there exists a $c := c(Q) > 0$ that

$$f(n,p,Q) \leqslant 2^n - (2-c)^n$$

holds, too.

Details and further refinements of the zones are given in subsequent sections. Most of our new results (and the problem of Erdős we have started from) regard the middle zone.

3. Low zone, exponential lower bounds from random choice

First, we show (1) that $f(n,p,Q)$ is at least exponential in n for all nontrivial cases.

Claim 4. *Let $Q\subset[p]$, $Q\neq[p]$, and suppose that $p\geqslant 2$. Then $f(n,p,Q)>(1+\varepsilon)^n$ where $\varepsilon>0$ depends only on p.*

Proof. One can obtain Claim 4 as a corollary of the estimates of the sizes of p-independent families. A family of sets $\mathscr{F}$ on the set $[n]$ is called p-independent, if every p members $F_1,\dots,F_p\in\mathscr{F}$ have 2^p atoms, i.e., for every $I\subseteq[p]$ there exists an element $x\in[n]$ such that $x\in F_i$ holds for $i\in I$, but $x\notin F_j$ for $j\in[p]\setminus I$. Clearly, if $\mathscr{F}$ is p-independent then it is (p,Q)-free for arbitrary $Q=[p]\setminus\{t\}$ $(0\leqslant t\leqslant p)$. The size of the largest p-independent family is denoted by $i(n,p)$. For 2-independent families, we obtain $i(n,2)=\binom{n-1}{\lfloor (n-2)/2\rfloor}$ from Milner's [31] result on intersecting Sperner families. In general, Kleitman and Spencer [28] showed that $i(n,p)$ is exponential in n for every fixed p. They proved with a simple probabilistic argument that $\log i(n,p)\geqslant n/(2p2^p)$. (Alon [1], using algebraic geometry codes gave a weaker, but still exponential constructive lower bound on $i(n,p)$.) This implies that Claim 4 with $\varepsilon>1/(2p2^p)$ holds. □

We note that a routine random choice argument, a standard example of the alteration method gives a better ε (polynomial in p) in Claim 4, namely

$$\varepsilon>\frac{1}{4p\sqrt{p}}.$$

Indeed, for $Q=[p]\setminus\{t\}$, $0<t<p$, one can take a random 0–1 matrix M of size $2m\times n$ with $\mathrm{Prob}(M_{ij}=1)=t/p$. Here the expected number of (p,Q)-subfamilies is

$$\binom{2m}{p}\left(1-\binom{p}{t}\left(\frac{t}{p}\right)^t\left(\frac{p-t}{p}\right)^{p-t}\right)^n.$$

This is at most m for $m\leqslant(1+\varepsilon)^n$ with

$$1+\varepsilon=\left(1-\binom{p}{t}\frac{t^t(p-t)^{p-t}}{p^p}\right)^{-1/(p-1)}>1+\frac{1}{4p\sqrt{p}}.$$

4. Low zone, Δ-systems and Boolean algebras

The aim of this section is to show the upper bound (4). It is not obvious at all that $f(n,p,Q)=o(2^n)$ holds in the low zone. However, it is easy to obtain this from known results concerning Δ-systems and Boolean lattices.

Theorem 5. *For fixed p and $\{0,p\}\subseteq Q$ as $n\to\infty$ one has*

$$f(n,p,Q)=O(2^n n^{-1/2^p})=o(2^n). \tag{9}$$

By monotonicity (3), it is enough to show this for $Q=\{0,i,p\}$ for arbitrary $i\in[p-1]$.

4.1. Δ-systems

The $(p,\{0,1,p\})$-systems are usually called Δ-systems and have been introduced in a famous two-part paper of Erdős and Rado [12]. They showed that if $\mathscr{D}$ is a family of at most k-element sets, and $|\mathscr{D}|>k!(p-1)^k$, then $\mathscr{D}$ contains a *Δ-system* of size p, i.e., there are members $D_1,\dots,D_p\in\mathscr{D}$ such that $D_i\cap D_j=\bigcap_{1\leqslant\alpha\leqslant p}D_\alpha$ holds for every $1\leqslant i<j\leqslant p$. Thus $\{D_1,\dots,D_p\}$ is a $(p,\{0,1,p\})$-system.

The case $p=3$, $i=1$ is settled by Erdős and Szemerédi [14]. They showed that $f(n,3,\{0,1,3\})\leqslant 2^{n-\sqrt{n}/10}=o(2^n)$. Improving and extending this result Deuber et al. [9] showed that for each p and all sufficiently large n the following upper bound holds:

$$f(n,p,\{0,1,p\})<2^{n-(n\log\log n)^{1/2}/\log\log\log n}=o(2^n). \tag{10}$$

4.2. Boolean subalgebras

To show that $f(n,p,\{0,i,p\})=o(2^n)$ is always true, one can apply a result of Rödl on Boolean algebras of sets. A p-dimensional *Boolean algebra* on the vertex set $[n]$ is defined by considering the pairwise disjoint subsets $A,S_1,\dots,S_p\subseteq[n]$ where only A can be empty and then taking all 2^p distinct sets of the form $A\cup S_I$ where $I\subseteq[p]$ and

$$S_I=\bigcup_{i\in I}S_i.$$

Let $b(n,p)$ denote the maximum number of subsets of $[n]$ containing no p-dimensional Boolean algebra with ground set $X\subseteq[n]$. Observe that $b(n,1)=\binom{n}{\lfloor n/2\rfloor}$ is Sperner's [35] well-known theorem about antichains. Erdős and Kleitman [10] determined $b(n,2)$ up to a constant factor

$$b(n,2)=\Theta(2^n n^{-1/4}). \tag{11}$$

Since a p-dimensional Boolean algebra obviously contains $(p,\{0,i,p\})$-systems for every $i\in[p-1]$, $f(n,p,\{0,i,p\})\leqslant b(n,p)=o(2^n)$ follows from the following result.

Theorem 6 (Rödl [32]). *For every fixed* p, $b(n,p) = o(2^n)$.

In fact, Rödl [32] promised that the sharper result $b(n,p) = O(2^n n^{-c})$, where c depends only on p: "…will appear together with other related problems elsewhere". He kept his promise, the following very recent result of Gunderson et al. [23] is a strengthening of (11):

$$b(n,p) \leqslant c2^n n^{-1/2^p}. \tag{12}$$

Our final note here is that perhaps the low zone coincides with the very low zone to be treated in the next section. This would follow from the affirmative answer to the following conjecture.

Conjecture 7. For every $1 \leqslant i \leqslant p-1$ one has $f(n,p,\{0,i,p\}) \leqslant (2-c)^n$ (where c depends only on p).

5. Very low zone, string quartets performed by a trio

This zone is defined by the pairs (p,Q) for which $f(n,p,Q) \leqslant (2-\varepsilon)^n$ with some ε which depends only on p. Many results here are motivated by coding theory and handled in the language of binary strings, i.e., the edges of hypergraphs are considered as characteristic vectors.

Theorem 8 (Lindström [29]). $2^{n/2} \leqslant f(n,4,\{0,2,4\}) \leqslant 2^{(n+1)/2} + 1$.

The lower bound is an explicit construction. Define the binary addition of two sets as the set whose characteristic vector is obtained as mod 2 addition regarding the characteristic vectors of the sets. One can observe that in a $(4,\{0,2,4\})$-free family $\mathscr{F}$, the map $F_1 + F_2$ (with binary addition) is one-to-one from the pairs $F_1, F_2 \in \mathscr{F}$ to subsets of $[n]$. Thus $\binom{|\mathscr{F}|}{2} \leqslant 2^n$ which implies the upper bound.

If the binary addition is replaced by integer addition, then the same map can send only disjoint pairs to the same sequence of $\{0,1,2\}^n$. This easily implies the following.

Proposition 9. *For any even* p, $f(n,p,\{0,p/2,p\}) = O(3^{n/2})$.

The details are omitted because we have found the same observation in the manuscript of Alon et al. [3] string quartets in binary. The next two results are extracted from the same paper. The first result comes from a very clever application of a well-known lemma of Sauer, Shelah-Perles, Vapnik-Chervonenkis.

Theorem 10 (Alon et al. [3]). *For every* $0 < i < p$, $f(n,p,[p]\backslash\{i\}) \leqslant p2^{0.78n}$.

The most spectacular application of this result is the case $p=4$, $i=2$ since it was unknown whether $Q = \{0,1,3,4\}$ belongs to the very low zone (this question was asked by Vera T. Sós in 1987).

Finally, we mention the following result whose proof applies Körner's method based on subadditivity of graph entropy. It shows that $f(n,4,\{0,2,3,4\})$ can be separated from $f(n,4,\{0,2,4\})$. In fact the exponent in the upper bound $f(n,4,\{0,2,3,4\}<2^{(1/2-\varepsilon)n}$ by Alon et al. [3] was subsequently improved by Alon et al. [2] to 0.4968. The best bound is due to Fachini et al. [17].

Theorem 11 (Fachini et al. [17]). *For n sufficiently large* $f(n,4,\{0,2,3,4\})<2^{0.4561n}$.

6. Middle zone, lower bounds

Since either p or 0 is missing from Q in the middle zone, as we have seen in (5), $f(n,p,Q)\geqslant 2^{n-1}$ is obvious. We can improve this slightly, our best improvement is for the case $Q=\{0,p-1\}$, i.e., for Erdős's original question.

Theorem 12. *For* $p\geqslant 3$ *there exists a* $c=c(p)>0$ *such that*

$$f(n,p,\{0,p-1\})\geqslant 2^{n-1}+cn^{p-3}.$$

Proof. More precisely, it will be proved that

$$f(n,p,\{0,p-1\})\geqslant 2^{n-1}+\sum_{i=o}^{p-3}\binom{n-1}{i},$$

from which Theorem 12 clearly follows. Let $\mathscr{F}$ contain all subsets of $[n]$ containing the element 1 plus all sets of size $\leqslant p-3$. We claim that $\mathscr{F}$ is $(p,\{0,p-1\})$-free.

Assume that $A_1,A_2,\ldots,A_p\in\mathscr{F}$ form a $(p,\{0,p-1\})$-system. Let $U=:\bigcup_{1\leqslant i\leqslant p}A_i$ and $U_i:=U\backslash A_i$. As every element of U is covered exactly $p-1$ times we have that $U_i\subseteq A_j$ for all $j\neq i$. This implies that the sets $U_1,\ldots,U_p$ are pairwise disjoint.

All but at most one of the U_i's are nonempty. Indeed, $U_1=U_2=\emptyset$ implies $A_1=A_2=U$, a contradiction. Each A_i is a union of $p-1$ U_j's and at least $p-2$ of those are nonempty. We obtain that $|A_i|=\sum_{j\neq i}|U_j|\geqslant p-2$. Thus, all A_i must contain element 1, a contradiction. □

Remark 13. The above construction is *maximal*, i.e., adding an arbitrary subset A of $[n]$ to it will result in a $(p,\{0,p-1\})$-system.

Proof. Indeed, $|A|\geqslant p-2$ thus there exists a partition $A=A_1\cup A_2\cup\cdots\cup A_{p-2}$ such that $A_i\neq\emptyset$ for all i. Let $A_i{}'=A\backslash A_i$ $(i=1,\ldots,p-2)$. Then the p-collection A, $A\cup\{1\}$, $A_i{}'\cup\{1\}$ $(i=1,\ldots,p-2)$ is a $(p,\{0,p-1\})$-system. □

7. Middle zone, asymptotics

For every $p\geqslant 2$ there are $2^{p-1}-1$ choices of Q belonging to the low zone, $(1-o(1))3\times 2^{p-1}$ choices of the middle zone and $\Theta(2^{p/2})$ choices of the high zone.

The aim of this section is to settle Conjecture 2 for all but $O(3^{p/2})$ choices of the middle zone.

Theorem 14. *Suppose that $Q \subset \{0,1,\dots,p\}$ is such that $\{0,p\} \not\subset Q$ and there exists a $q \in Q$ with $1 \leqslant q \leqslant p/2$ and $(p-q) \in Q$. Then $f(n,p,Q) = (1+o(1))2^{n-1}$.*

Proof. The lower bound follows from (5). To prove the upper bound consider a family $\mathscr{F}$ of subsets of $[n]$ of size

$$|\mathscr{F}| > 2^{n-1} + b(n,p)/2,$$

where $b(n,p)$ is defined in Section 4 (cf. (12)). We show that $\mathscr{F}$ contains a $(p,\{q,p-q\})$-system.

Note that $\mathscr{F}$ contains more than $b(n,p)/2$ complementary pairs $\{F,[n]\backslash F\} \subseteq \mathscr{F}$. Apply (12) to the subsystem $\mathscr{F}' = \{F_1,\dots,F_m\}$ containing all the members of complementary pairs. Since $m > b(n,p)$ we obtain the pairwise disjoint, nonempty sets $S_1,\dots,S_p \subset [n]$ and a set $A \subseteq [n]\backslash(S_1 \cup \cdots \cup S_p)$ such that $A \bigcup_{i\in I} S_i \in \mathscr{F}'$ for all $I \subseteq [p]$. Since $\mathscr{F}'$ consists of complementary pairs, $B \bigcup_{i\in I} S_i \in \mathscr{F}'$, too, where $B = [n]\backslash (A \bigcup_{1\leqslant i\leqslant p} S_i)$. Take q sets of the first kind and $p-q$ of the second kind in such a way that the degrees are q on $S_1,\dots,S_p$, then we obtain a desired $(p,\{q,p-q\})$-system. Say, we can take the sets $A \cup \{S_i \cup S_{i+1} \cup \cdots \cup S_{i+q-1}\}$ for $1 \leqslant i \leqslant q$ and the sets $B \cup \{S_i \cup S_{i+1} \cup \cdots \cup S_{i+q-1}\}$ for $q < i \leqslant p$, where one has to take the indices of the S_i's modulo p. □

8. Middle zone, upper bounds

Our aim here is to prove an upper bound (7) separating the middle zone from the high zone.

Theorem 15. *In the middle zone $f(n,p,Q) \leqslant (1 - 1/2p + o(1))2^n$.*

First, we prove a proposition which helps to reduce the general problem to the cases: either $Q = \{0,q\}$, $p/2 \leqslant q < p$ or $Q = \{q,p-q\}$, $0 < q \leqslant p/2$.

Proposition 16. *Suppose that $q_1,q_2 \in Q \subset [p]$ such that*

$$q_1 \leqslant \frac{p}{2} \leqslant q_2, \quad q_1 \neq q_2 \quad \text{and} \quad \{q_1,q_2\} \neq \{0,p\}.$$

Let t be a positive integer $t \leqslant q_1$, $t \leqslant p - q_2$ but $t < \max\{q_1, p-q_2\}$. Then

$$f(n,p,Q) \leqslant f(n,p-2t,\{q_1-t,q_2-t\}) + 2t. \tag{13}$$

Proof. Note that $p-2t \geqslant 2$. Moreover, $0 \leqslant q_1 - t \leqslant (p-2t)/2 \leqslant q_2 - t \leqslant (p-2t)$ but $\{q_1-t,q_2-t\} \neq \{0,p-2t\}$ thus $f(n,p-2t,\{q_1-t,q_2-t\})$ belongs to the middle zone, too. Thus the right-hand side of (13) is at least $2^{n-1}+2t$ by (5), proving (13) if $f(n,p,Q)<2^{n-1}+t$. Otherwise there is a (p,Q)-free family $\mathscr{F}$ with $|\mathscr{F}|=f(n,p,Q)\geqslant 2^{n-1}+t$. It contains at least t complementary pairs, a subfamily $\mathscr{P}:=\{F_1,[n]\backslash F_1,\dots,F_t,[n]\backslash F_t\}\subseteq\mathscr{F}$. Leave these t pairs out from $\mathscr{F}$. Then $\mathscr{F}\backslash\mathscr{P}$ is $(p-2t,\{q_1-t,q_2-t\})$-free, otherwise such a subsystem $\mathscr{R}\subseteq\mathscr{F}\backslash\mathscr{P}$ together with $\mathscr{P}$ form a $(p,\{q_1,q_2\})$-system. Therefore, $|\mathscr{F}\backslash\mathscr{P}|=f(n,p,Q)-2t\leqslant f(n,p-2t,\{q_1-t,q_2-t\})$. □

Proof of Theorem 15. If $Q\cap\bar{Q}\neq\emptyset$, (where again $\bar{Q}=\{p-q:q\in Q\}$), then $\lim_{n\to\infty} f(n,p,Q)2^{-n}=1/2$ follows from Theorem 14. Otherwise, using the reductions of the above Proposition 16 and taking complementation if necessary (see (2)), one can obtain for some $0<p'/2<q'<p'$ that $f(n,p,Q)\leqslant f(n,p',\{0,q'\})+\mathrm{O}(p)$. So we can reduce the proof to the case originally proposed by Erdős, i.e., $Q=\{0,q\},\ p/2<q<p$.

Let $\mathscr{F}$ be a $(p,\{0,q\})$-free family on $[n]$ and denote by f_k the size of $\mathscr{F}_k$ (i.e., the size of $\{F\in\mathscr{F}:\ |F|=k\}$). Let $\mathscr{H}$ be a $(p,\{0,q\})$-system on $[n]$, $K=\{|H|:\ H\in\mathscr{H}\}$, $\alpha_k=|\{H\in\mathscr{H}:\ |H|=k\}|$. We claim that

$$\sum_{k\in K}\frac{\alpha_k f_k}{\binom{n}{k}}\leqslant p-1. \tag{14}$$

Indeed, consider a permutation π of $[n]$ and apply it to $\mathscr{H}$ and consider $\pi(\mathscr{H})\cap\mathscr{F}$. It consists of at most $p-1$ hyperedges, so we get

$$\sum_{\pi\in S_n}|\pi(\mathscr{H})\cap\mathscr{F}|\leqslant(p-1)n!$$

On the other hand, every edge $E\in\mathscr{H}$ appears exactly $f_k|E|!(n-|E|)!$ times on the left-hand side. We obtain

$$\sum_{E\in\mathscr{H}}f_k|E|!(n-|E|)!=\sum_{k\in K}\alpha_k f_k k!(n-k)!\leqslant(p-1)n!$$

Rearranging we get (14).

Now suppose that $p\leqslant v\leqslant n$. We explicitly construct some $(p,\{0,q\})$-system $\mathscr{H}_v$ on vertex set $[v]$ and then will apply (14) to it. The edges of $\mathscr{H}_v$ $\{E_1,\dots,E_p\}$ are defined as follows. E_i meets $[p]$ in q vertices, $E_i\cap[p]=\{i,i+1,\dots,i+q-1\}$ (we have to take the elements here modulo p), and for $p<x\leqslant v$ the element x belongs to the edges E_{qx+i} for $1\leqslant i\leqslant q$ (again indices are taken modulo p). Then $\mathscr{H}_v$ consists of edges of sizes $\lfloor vq/p\rfloor$ and $\lceil vq/p\rceil$ only. For every k in the range $q\leqslant k<\lfloor qn/p\rfloor$ there exists a v such that $k<vq/p<k+1$, therefore $\mathscr{H}_v$ has edges of sizes k and $k+1$. It follows from (14) that

$$\text{either}\quad f_k\Big/\binom{n}{k}\leqslant(p-1)/p\quad\text{or}\quad f_{k+1}\Big/\binom{n}{k+1}\leqslant(p-1)/p\,.$$

Let

$$I=\left\{i\in[n]:\ f_i\leqslant\frac{p-1}{p}\binom{n}{i}\right\},$$

we have that I has no large gap, $I \cap \{k, k+1\} \neq \emptyset$ for every $q \leqslant k < \lfloor qn/p \rfloor$. Then,

$$
\begin{aligned}
|\mathscr{F}| = \sum_{i=0}^{n} f_i &\leqslant \sum_{i=0}^{n} \binom{n}{i} - \frac{1}{p} \sum_{i \in I} \binom{n}{i} \\
&\leqslant 2^n - \frac{1}{p} \left(\sum_{i \text{ is even}} \binom{n}{i} - \sum_{i \leqslant q} \binom{n}{i} - \sum_{i \geqslant nq/p} \binom{n}{i} \right) \\
&= \left(1 - \frac{1}{2p} + \mathrm{o}(1)\right) 2^n. \qquad \square
\end{aligned}
$$

8.1. Decomposition method

Another possible approach to prove an upper bound is the *decomposition method*, i.e., to decompose all subsets of an $[n]$ to as many $(p, \{0, q\})$-systems as it is possible, and obtain an upper bound by deleting a set from each decomposed part. Using this we may easily get the following, slightly weaker upper bound than (7). For arbitrary $p/2 < q < p$,

$$
f(n, p, \{0, q\}) \leqslant (1 + \mathrm{o}(1)) \left(1 - \frac{1}{pq}\right) 2^n. \tag{15}
$$

To get a desired decomposition in our case (and in many other applications) a very useful tool is the following form of Baranyai's theorem [4], see [5].

Theorem 17 (Baranyai [4]). *Let $\mathscr{F} = (V, E)$ be the r-uniform complete hypergraph on k vertices (i.e., the family containing all $\binom{k}{r}$ r-subsets of a k-set), and let $e_1, e_2, \ldots, e_s$ be nonnegative integers such that $\sum_{i=1}^{s} e_i = \binom{k}{r}$. Then the edge set E of $\mathscr{F}$ can be partitioned into s sets: $E = \bigcup_{i=1}^{s} E_i$ such that $|E_i| = e_i$ and every $\mathscr{F}_i = (V, E_i)$ is almost regular.*

Proof sketch of (15). Let t be an arbitrary positive integer such that $tp \leqslant n$, and take an arbitrary tp-subset V of $[n]$. Let $\mathscr{F}_V$ be the complete tq-uniform hypergraph with vertex set V. Choose

$$
s = \left\lceil \binom{tp}{tq} \Big/ p \right\rceil.
$$

If $\binom{tp}{tq}/p$ is an integer, let $\forall i$: $1 \leqslant i \leqslant s = \binom{tp}{tq}/p$ $e_i = p$, otherwise let $\forall i$: $1 \leqslant i \leqslant s-1$ $e_i = p$ and $e_s = \binom{tp}{tq} - (s-1)p$. Now apply Theorem 17 to $\mathscr{F}_V$ choosing the parameters as above. We get a decomposition of F_V into $(p, \{0, q\})$-systems. But a $(p, \{0, q\})$-free family does not contain at least one member from every decomposed part. We get the desired upper bound using this fact to all sets V of suitable sizes t avoiding repetitions. $\square$

9. Middle zone, strong $(4,\{0,3\})$-systems

Here we demonstrate the difficulty of Conjecture 2 for the smallest unknown case, formulating a subconjecture. A $(4,\{0,3\})$-free family is called *strong* if it does not contain two disjoint sets with their union. In particular, a strong $(4,\{0,3\})$-free family cannot contain the empty set. The following proposition (the easy proof is left to the reader) shows a useful extension property for strong $(4,\{0,3\})$-free families.

Proposition 18. *Suppose that* $\mathscr{F}$ *is a strong* $(4,\{0,3\})$*-free family on the vertex set* $[n]$. *Then*

$$2^{[n+1]} \supseteq \mathscr{F}' = \mathscr{F} \cup \{\{n+1\} \cup X\colon\ X \in \mathscr{F}\}$$

is also a strong $(4,\{0,3\})$*-free family.*

Assume that there exists a strong $(4,\{0,3\})$-free family with more than 2^{n_0-1} sets on $[n_0]$. Then, using Proposition 18 repeatedly, one can construct a (strong) $(4,\{0,3\})$-free family with $(1+\varepsilon)2^{n-1}$ sets, provided that n is large. Therefore, the following conjecture is weaker than Conjecture 2.

Conjecture 19. A strong $(4,\{0,3\})$-free family on $[n]$ has at most 2^{n-1} sets.

We know only the following five minimal strong $(4,\{0,3\})$-free families on $[n]$ with 2^{n-1} sets. Here minimal means that they cannot be generated through Proposition 18:

1. $\{1\}$ on [1],
2. $\{1\},\{2\}$ on [2],
3. $\{1\},\{1,2\},\{1,3\},\{2,3\}$ on [3],
4. $\{1\},\{2\},\{3\},\{1,2,3\}$ on [3],
5. $\{1\},\{1,2\},\{1,3\},\{1,4\},\{2,4\},\{3,4\},\{1,2,3\},\{2,3,4\}$ on [4].

10. Middle zone, sharp results

This section is devoted to the few special choices of p and Q for which we could determine precisely the value of $f(n,p,Q)$. We start with an unpublished result of Erdős and Sós.

Theorem 20 (Erdős-Sós [15]). $f(n,3,\{0,2\}) = 2^{n-1}+1$.

Proof. Let $\mathscr{F}$ be the family of subsets of $[n]=\{1,\ldots,n\}$ containing a fixed element plus the empty set. This $\mathscr{F}$ is $(3,\{0,2\})$-free, hence $f(n,3,\{0,2\}) \geqslant 2^{n-1}+1$. To prove the reverse inequality observe that a $(3,\{0,2\})$-free family cannot contain three sets A, B and $A \triangle B = (A\backslash B)\cup(B\backslash A)$, since they would form a $(3,\{0,2\})$-system. Now let $\mathscr{F}$ be an arbitrary $(3,\{0,2\})$-free family and $A \neq \emptyset$ be an arbitrary member of it. By the previous observation, if $B \in \mathscr{F}\backslash\{A,\emptyset\}$ then $A \triangle B \notin \mathscr{F}\backslash\{A,\emptyset\}$. Therefore, to each

$B \in \mathscr{F}\setminus\{A,\emptyset\}$ one can assign $B' = A \triangle B \notin \mathscr{F}\setminus\{A,\emptyset\}$. Since symmetric differences of distinct sets with a fixed set A are distinct, $|\mathscr{F}\setminus\{A,\emptyset\}| \leqslant 2^{n-1} - 1$. □

The following proposition is an easy generalization of a folklore remark (in fact, its origin is [11]): an intersecting family on $[n]$ has at most 2^{n-1} sets. Notice that an intersecting family is a $(2,\{0,1\})$-system. In fact, the folklore proof gives more: $f(n,2,\{1\}) = 2^{n-1}$. An easy generalization is the following.

Proposition 21. *Assume that* $p \geqslant 2$ *is even and* $p/2$ *is the largest element of* Q. *Then*

$$f(n,p,Q) = 2^{n-1} + (p/2) - 1.$$

Proof. Let $\mathscr{F}$ be a $(p,\{p/2\})$-free family on n vertices. Considering the 2^{n-1} complementary pairs $\{X,[n]\setminus X\}$ one can see that at most $(p/2)-1$ can be contained entirely in $\mathscr{F}$. This gives $|\mathscr{F}| \leqslant 2^{n-1} + (p/2) - 1$. Adding arbitrarily $(p/2) - 1$ new sets to $\mathscr{F}[x]$ shows that equality is possible. □

Proposition 22. *Assume that* p *is odd and* $(p+1)/2$ *is the largest element of* Q *and* $\{(p-1)/2,(p+1)/2\} \subseteq Q$. *Then*

$$f(n,p,Q) = 2^{n-1} + \frac{p-3}{2}.$$

Proof. Like in the previous proposition suppose that $|\mathscr{F}| > 2^{n-1} + (p-3)/2$. Then it contains $(p-1)/2$ complementary pairs. Adding one more edge to these pairs one obtains a $(p,\{(p-1)/2,(p+1)/2\})$-system. Thus $f(n,p,\{(p-1)/2,(p+1)/2\}) \leqslant 2^{n-1} + (p-3)/2$. Adding arbitrarily $(p-3)/2$ new sets to $\mathscr{F}[x]$ shows that equality is possible. □

Perhaps the folklore remark above can be also extended in another direction:

Conjecture 23. For every p, $f(n,p,\{1,p-1\}) = 2^{n-1}$ if n is large.

We can prove this only for $p \leqslant 4$. If $p = 2$ or 3 it is true for every n. (The former statement is the folklore remark, and the latter is also obvious since two complementary sets with any other set form a triple with degrees in $\{1,2\}$.) For $p = 4$ the one and two element sets of [3] give six sets with no four with degrees in $\{1,3\}$, so $f(n,4,\{1,3\}) = 2^{n-1} + 2$ for $n = 3$. Probably $f(n,4,\{1,3\}) = 2^{n-1}$ holds already for $n = 4,5$, the next theorem proves this for $n \geqslant 6$.

Theorem 24. *For every* $n \geqslant 6$, $f(n,4,\{1,3\}) = 2^{n-1}$.

Proof. Assume that $\mathscr{F}$ is a family of subsets of $[n]$ with $2^{n-1} + 1$ members. We show that a set H of four members of $\mathscr{F}$ form a $(4,\{1,3\})$-system. Two sets of $\mathscr{F}$ are called a *complementary pair* if they complement each other with respect to $[n]$. Assume first that $\mathscr{F}$ has at least 2^{n-2} complementary pairs. Using the upper bound from Theorem 8, $f(n,4,\{0,2,4\}) \leqslant 2^{(n+1)/2}+1<2^{n-2}$ if $n\geqslant 6$. Therefore, selecting one set from each complementary pair, we get a 2^{n-2}-element subset of $\mathscr{F}$ which contains four sets forming a $(4,\{0,2,4\})$-system. Then the required H is obtained by replacing one of the four sets with its complement (which is also in $\mathscr{F}$ and distinct from the other three).

On the other hand, if $\mathscr{F}$ has less than 2^{n-2} complementary pairs, we can select two sets, A_1, A_2 from $\mathscr{F}$, so that neither is the complement of any set in $\mathscr{F}$. This condition ensures that the map g which sends $A\in\mathscr{F}$ to $A+A_1+A_2+[n]$ (where $+$ denotes binary addition) has the property: for every $A\in(\mathscr{F}\setminus\{A_1,A_2\})$, the quadruple $H(A)=(A,g(A),A_1,A_2)$ has four distinct sets. Indeed, $A=g(A)$ is equivalent to $A_1=\overline{A_2}$ and $g(A)=A_1$ ($g(A)=A_2$) is equivalent to $A=\overline{A_2}$ ($A=\overline{A_1}$). From the definition of g, for every $A\in(\mathscr{F}\setminus\{A_1,A_2\})$, the four distinct sets of $H(A)$ form a $(4,\{1,3\})$-system. Observe that g is a one-to-one map from $\mathscr{F}$ to the subsets of $[n]$ which sends A_1 (A_2) to $\overline{A_2}\notin\mathscr{F}$ ($\overline{A_1}\notin\mathscr{F}$). Therefore, $g(A)\notin\mathscr{F}$ holds for at most $2^{n-1}-3$ sets of $\mathscr{F}\setminus\{A_1,A_2\}$. Since $\mathscr{F}\setminus\{A_1,A_2\}$ has $2^{n-1}-1$ sets, $g(A)\in\mathscr{F}$ must hold for some $A\in(\mathscr{F}\setminus\{A_1,A_2\})$. For this A, the quadruple $H(A)$ is the required $(4,\{1,3\})$-system of $\mathscr{F}$. □

Corollary 25. *Assume that $p\geqslant 4$ is even, $n\geqslant 6$, $p/2+1$ is the largest element of Q and $\{p/2-1, p/2+1\}\subseteq Q$. Then $f(n,p,Q)=2^{n-1}+p/2-2$.*

Proof. All sets through a fixed element plus $p/2-2$ other sets show that equality is possible. The upper bound follows by induction, the case $p=4$ is Theorem 24. Selecting a complementary pair from $\mathscr{F}$ of $2^{n-1}+p/2-1$ sets the remaining $2^{n-1}+(p-2)/2-2$ contains by induction $p-2$ sets with degrees in $\{(p-2)/2-1, (p-2)/2+1\}$. Adding the complementary pairs we have the required p sets. □

We conclude this section with the summary of cases when exact results are known to us for $p\leqslant 4$. The trivial cases are omitted and from complementary pairs of subsets of [2,3,4] only one is mentioned (the smaller in lexicographic order).

10.1. Half is the best

The statement $f(n,p,Q)=2^{n-1}$ holds for $p=2$, $Q=\{1\}$ or $Q=\{0,1\}$; for $p=3$, $Q=\{1,2\}$ or $Q=\{0,1,2\}$; for $p=4$, $Q=\{1,2,3\}$ or $Q=\{0,1,2,3\}$ or $Q=\{1,3\}$ (here $n\geqslant 6$). In these cases $p\notin Q$, therefore sets through a fixed vertex show that equality is possible. The upper bounds follow either directly or immediately from previous results. There is one further case:

Proposition 26. *For $n\geqslant n_0$, $f(n,4,\{0,2,3\})=2^{n-1}$.*

Proof. Assume that $\mathscr{F}$ is a family of subsets of $[n]$ with $2^{n-1}+1$ members. We show that a set H of four members of $\mathscr{F}$ forms a $(4,\{0,2,3\})$-system. If $\mathscr{F}$ has more than one complementary pair then two such pairs give H. Otherwise $\mathscr{F}$ contains precisely one complementary pair A,B and all other complementary pairs are split by $\mathscr{F}$.

Case 1. $\emptyset\notin\mathscr{F}$. This implies that $[n]\in\mathscr{F}$. Then take the four sets $H=\{A,B,[n],C\}$ with an arbitrary $C\in\mathscr{F}$ (distinct from the other three).

Case 2. $A_0=\emptyset\in\mathscr{F}$. Then, using (10), $f(n,3,\{0,2,3\})=f(n,3,\{0,1,3\})=o(2^n)\leqslant 2^{n-1}$ if n is large enough. Thus, H can be defined as three sets forming a $(3,\{0,2,3\})$-system together with the empty set. □

10.2. Half plus one is the best

The statement $f(n,p,Q)=2^{n-1}+1$ holds for $p=3$, $Q=\{0,2\}$; for $p=4$, $Q=\{2\}$ or $\{0,2\}$ or $\{1,2\}$. In these cases neither p, nor $p-1$ is in Q therefore the empty set plus all sets through a fixed vertex show equality. The upper bounds follow from previously stated results.

11. High zone, Kleitman-type results

We have seen (Proposition 3) that in the high zone, a maximal (p,Q)-free family contains almost all subsets of $[n]$. In this section we review some results of Kleitman and others yielding sharper bounds.

11.1. No p pairwise disjoint sets, $Q=\{0,1\}$

The family $F_1,\ldots,F_p$ is a $(p,\{0,1\})$-system if and only if its members are pairwise disjoint. To avoid trivialities, we deal with the case $p\geqslant 3$ only. If each F_i is a subset of $[n]$, then one of them has size at most n/p. Thus

$$\mathscr{F}_p^1:=\{F\subseteq[n]:\ |F|>n/p\}$$

provides a $(p,\{0,1\})$-free family. For $p\geqslant 3$ this construction contains almost all sets, $|\mathscr{F}_p^1|=(1-o(1))2^n$. Kleitman proved that

Theorem 27 (Kleitman [25]). *For* $n\equiv -1 \pmod p$ $f(n,p,\{0,1\})=|\mathscr{F}_p^1|$.

However, in the case $n\not\equiv -1 \pmod p$, $\mathscr{F}_p^1$ is not maximal, one can add a few more sets. Kleitman also solved the case $n=mp$, the other cases are still open.

11.2. No partition into p parts, $Q=\{1\}$

A family $A_1,A_2,\ldots,A_p$ is a $(p,\{1\})$-system if and only if they form a partition on $[n]$. Here $f(n,p,\{1\})$ can be estimated by applying results on $d(n,p)$, the maximum size of a family which does not contain $p-1$ pairwise disjoint sets and their union. Such families are called $(p-1)$-*disjoint-union free*. For $p\geqslant 3$ such a family is

$$\mathscr{F}_p^2:=\left\{F\subseteq[n]:\ n/p<|F|<\frac{p-1}{p}n\right\},$$

and it contains almost all sets, $|\mathscr{F}_p^2|=(1-o(1))2^n$.

Theorem 28 (Kleitman [26] (for $p=3$), Frankl [18] (for all $p\geqslant 3$)). *If* $n\equiv -1 \pmod p$ *then* $d(n,p)=|\mathscr{F}_p^2|$.

The cases $n\neq mp-1$ are unsolved. There are slightly larger families than $\mathscr{F}_p^2$, it is not even maximal, one can add a few more sets. Theorem 28 has the following consequence.

Corollary 29. *Suppose that* $n\equiv -1 \pmod p$ *and* $p\geqslant 3$. *Then*

$$f(n,p,\{1\})=|\mathscr{F}_p^1|=2^n-(1+o(1))\,\frac{p-1}{p-2}\binom{n}{\lfloor n/p\rfloor}.$$

Proof. Suppose that $\mathscr{F}$ is a maximal $(p,\{1\})$-free family, we have $|\mathscr{F}|\geqslant|\mathscr{F}_p^1|$. Then $\mathscr{F}\cap\mathscr{F}^{\mathrm{C}}$ (i.e, the sets having complements in $\mathscr{F}$) is $(p-1)$-disjoint-union free. Thus

$$|\mathscr{F}_p^1|\leqslant|\mathscr{F}|=\tfrac{1}{2}(2^n+|\mathscr{F}\cap\mathscr{F}^{\mathrm{C}}|)\leqslant\tfrac{1}{2}(2^n+|\mathscr{F}_p^2|)=|\mathscr{F}_p^1|. \qquad\square$$

Similarly, $f(n,p,\{0,1\})\leqslant f(n,p,\{1\})\leqslant\frac{1}{2}(2^n+d(n,p))$, and very likely equality holds here.

11.3. An estimate for the general case

Conjecture 30. The construction $\mathscr{F}_{q/p}$ defined in Section 2.3 is (almost) optimal.

The smallest open case is $p=5$, max $Q=2$. The following result, based on Corollary 29, shows this at least for the case $q|p$.

Theorem 31. *If* max $Q=q<p/2$ *then* $f(n,p,Q\})\leqslant f(n,\lceil p/q\rceil,\{1\})+p-\lceil p/q\rceil$.

Proof. It is sufficient to show that $f(n,p,\{q\})\leqslant f(n,\lceil p/q\rceil,\{1\})+p-\lceil p/q\rceil$. Take a family $\mathscr{F}$ of size larger than the right-hand side here. Write p in the form $p=a_1+a_2+\cdots+a_q$ where $\lfloor p/q\rfloor\leqslant a_1\leqslant a_2\leqslant\cdots\leqslant a_q\leqslant\lceil p/q\rceil$. Then $|\mathscr{F}|>f(n,a_1,\{1\})$, thus one can find a subfamily $\mathscr{A}_1$ of size a_1 forming a partition (i.e., a $(a_1,\{1\})$-family). Repeat this for $\mathscr{F}\backslash\mathscr{A}_1$ and find a subfamily of size a_2, and so on. The obtained partitions $\mathscr{A}_1\cup\cdots\cup\mathscr{A}_q$ cover each vertex exactly q times. $\square$

Corollary 32.

$$2^n-O\left(\binom{n}{\lfloor n/(p/q)\rfloor}\right)\leqslant|\mathscr{F}_{q/p}|\leqslant f(n,p,Q)\leqslant 2^n-\Theta\left(\binom{n}{n/\lceil p/q\rceil}\right)$$

12. Conclusion

We think—and hope the reader agrees—that many (old and new) problems, results and conjectures relate immediately to the initial question of Erdős on $f(n,4,\{0,3\})$.

We conclude this paper by returning to its origin for the k-uniform case, i.e., when only k-element sets are allowed. Let $f(n,k,4,\{0,3\})$ denote the maximum size of a k-uniform $(4,\{0,3\})$-free family. For convenience we shall assume that $k|n$. Obviously, for $k \not\equiv 0 \pmod 3$ $(4,\{0,3\})$-systems do not exist, so in this case $f(n,k,4,\{0,3\}) = \binom{n}{k}$. On the other hand, the case $k \equiv 0 \pmod 3$ is certainly difficult, $f(n,3,4,\{0,3\}) = \mathrm{ex}(n,K_3^4)$ is a famous Turán number. For history of Turán numbers see [7] and for general Turán-type problems [22]. In the following theorem the magnitude of $f(n,k,4,\{0,3\})(k \equiv 0 \pmod 3)$ is determined. Let K_3^4 denote the complete 3-uniform hypergraph on 4 vertices and

$$\pi(K_3^4) = \lim_{n\to\infty} \frac{\mathrm{ex}(n,K_3^4)}{\binom{n}{3}},$$

where $\mathrm{ex}(n,K_3^4)$ is the maximum number of edges of a 3-uniform hypergraph containing no K_3^4. This limit is known to exist (see [24]), but is not determined yet. The construction of Turán gives $5/9 \leqslant \pi(K_3^4)$, which is conjectured to be optimal. The best upper bound, due to Chung and Lu [8] yields $\pi(K_3^4) \leqslant (3+\sqrt{17})/12$.

Theorem 33. *For* $k \equiv 0 \pmod 3$

$$(1-\mathrm{o}(1))\,\frac{1}{e} \leqslant \frac{f(n,k,4,\{0,3\})}{\binom{n}{k}} \leqslant \pi(K_3^4) \leqslant \frac{3+\sqrt{17}}{12} = 0.59359\ldots\ .$$

Proof. The upper bound follows from the fact that

$$\frac{f(n,k,4,\{0,3\})}{\binom{n}{k}} \leqslant \pi(K_3^4).$$

This follows from a well-known blow-up technique of Sidorenko [33,34] and Frankl [20] (also see [21]).

To see the lower bound, take a partition of $[n]$ into parts X and $[n]\backslash X$ of size n/k and $n-n/k$, respectively. Let $\mathscr{F}$ consist of all k-subsets of $[n]$ containing exactly one element from X. We claim that $\mathscr{F}$ is $(4,\{0,3\})$-free. Indeed, take an arbitrary collection of four sets from $\mathscr{F}$. Then the chosen four sets either contain the same element from X, or at least two different ones. In the first case the given element is covered four times while in the second case there is an element (in the union of the chosen four sets) which is covered at most twice. Hence $\mathscr{F}$ is $(4,\{0,3\})$-free. We obtain $|\mathscr{F}| = (n/k)\binom{n-n/k}{k-1}$, from which the desired result follows by a straightforward computation. □

Remark 34. One can get a somewhat larger constant in the lower bound than $1/e$ defining $\mathscr{F} = \{F \in \binom{[n]}{k}\colon\ |F \cap X| \equiv 1 \pmod 3\}$. But since this improvement probably does not give a tight result, we omit the detailed computation.

Acknowledgements

We appreciate the fruitful discussions with János Körner, Gábor Simonyi and Vera Sós. We also thank the referees for their careful work.

References

[1] N. Alon, Explicit constructions of exponential sized families of k-independent sets, Discrete Math. 58 (1986) 191–193.
[2] N. Alon, E. Fachini, J. Körner, Locally thin set families, Combin. Probab. Comput. 9 (2000) 481–488.
[3] N. Alon, J. Körner, A. Monti, String quartets in binary, Combin. Probab. Comput. 9 (2000) 381–390.
[4] Zs. Baranyai, On the factorization of the complete uniform hypergraph, in: A. Hajnal, R. Rado, V.T. Sós (Eds.), Infinite and Finite Sets, North-Holland, Amsterdam, pp. 91–108.
[5] B. Bollobás, Combinatorics, Cambridge University Press, Cambridge, 1986.
[6] T.C. Brown, J.P. Buhler. Lines imply spaces in density Ramsey theory, J. Combin. Theory Ser. A 36 (1984) 214–220.
[7] D. de Caen, The current status of Turán's problem on hypergraphs, in: Extremal Problems for Finite Sets, Bolyai Society of Mathematical Studies, Vol. 3, Visegrád, 1991, János Bolyai Math. Soc., Budapest, 1994, pp. 187–197.
[8] Fan Chung, Linyuan Lu, An upper bound for the Turán number $t_3(n,4)$, J. Combin. Theory Ser. A 87 (1999) 381–389.
[9] W.A. Deuber, P. Erdős, D.S. Gunderson, A.V. Kostochka, A.G. Meyer, Intersection statements for systems of sets, J. Combin. Theory Ser. A 79 (1997) 118–132.
[10] P. Erdős, D.J. Kleitman, On collections of subsets containing no 4-member Boolean algebra, Proc. Amer. Math. Soc. 28 (1971) 87–90.
[11] P. Erdős, C. Ko, R. Rado, Intersection theorems for systems of finite sets, Quart. J. Math. Oxford 12 (1961) 313–320.
[12] P. Erdős, R. Rado, Intersection theorems for systems of sets I and II, J. London Math. Soc. 35 (1960) 85–90; P. Erdős, R. Rado, Intersection theorems for systems of sets I and II, J. London Math. Soc. 44 (1969) 467–479.
[13] P. Erdős, A. Sárközy, V. Sós, On product representations of powers I, European J. Combin. 16 (1995) 567–588.
[14] P. Erdős, E. Szemerédi, Combinatorial properties of systems of sets, J. Combin. Theory Ser. A 24 (1978) 308–313.
[15] Erdős, Private communication, 1996.
[16] E. Fachini, J. Körner, A. Monti, Self-similarity bounds for locally thin set families, Combin. Probab. Comput. 10 (2001) 309–315.
[17] E. Fachini, J. Körner, A. Monti, A better bound for locally thin set families, J. Combin. Theory Ser. A 95 (2001) 209–218.
[18] P. Frankl, Families of finite sets containing no k disjoint sets and their union, Period. Math. Hungar. 8 (1977) 29–31.
[19] P. Frankl, Intersection and containment problems without size restriction, in: Algebraic, Extremal and Metric Combinatorics, London Mathematical Society, Lecture Note Series, Vol. 131, Montreal, 1986, Cambridge University Press, Cambridge, 1988, pp. 62–111.
[20] P. Frankl, Asymptotic solution of a Turán-type problem, Graphs Combin. 6 (1990) 223–227.
[21] P. Frankl, Z. Füredi, Extremal problems and the Lagrange function for hypergraphs, Bull. Inst. Math. Acad. Sinica 16 (1988) 305–313.
[22] Z. Füredi, Turán type problems, in: Surveys in Combinatorics, 1991 London Mathematical Society, Lecture Note Series, Vol. 166, Guildford, 1991, Cambridge University Press, Cambridge, 1991, pp. 253–300.
[23] D.S. Gunderson, V. Rödl, A. Sidorenko, Extremal problems for sets forming Boolean algebras and complete partite hypergraphs, J. Combin. Theory Ser. A 88 (1999) 342–367.

[24] G. Katona, T. Nemetz, M. Simonovits, On a problem of Turán in the theory of graphs, Mat. Lapok 15 (1964) 228–238 (Hungarian, English summary).
[25] D.J. Kleitman, Maximal number of subsets of a finite set no k of which are pairwise disjoint, J. Combin. Theory Ser. A 5 (1968) 157–163.
[26] D.J. Kleitman, On families of subsets of a finite set containing no two disjoint sets and their union, J. Combin. Theory Ser. A 5 (1968) 235–237.
[27] D.J. Kleitman, Extremal properties of collections of subsets containing no two sets and their union, J. Combin. Theory Ser. A 20 (1976) 390–392.
[28] D.J. Kleitman, J. Spencer, Families of k-independent sets, Discrete Math. 6 (1973) 141–147.
[29] B. Lindström, Determination of two vectors from the sum, J. Combin. Theory Ser. A 6 (1969) 402–407.
[30] B. Lindström, On B_2-sequences of vectors, J. Number Theory 4 (1972) 261–265.
[31] E.C. Milner, A combinatorial theorem on systems of sets, J. London Math. Soc. 43 (1968) 204–206.
[32] V. Rödl, Note on finite Boolean algebras, Acta Polytech. (1982) 47–49.
[33] A.F. Sidorenko, Extremal combinatorial problems in spaces with continuous measure, Issledovanie Operatchiy i ASU, No. 34, Kiev, 1989, pp. 34–40, 101 (Russian).
[34] A.F. Sidorenko, An analytic approach to extremal problems for graphs and hypergraphs. in: Extremal Problems for Finite Sets, Bolyai Society of Mathematical Study, Vol. 3, Visegrád, 1991, János Bolyai Mathematical Society, Budapest, 1994, pp. 423–455.
[35] E. Sperner, Ein Satz über Untermengen einer endlichen Menge, Math. Z. 27 (1928) 544–548.