

**PROBLEMS AND RESULTS  
ON INTERSECTIONS OF SET SYSTEMS  
OF STRUCTURAL TYPE**

P. ERDÖS AND V.T. SÓS

The study of intersection theorems for set systems was started in the papers of [BE 1] and [EKR 1].

In the last twenty years a fairly extended theory has developed, several survey papers [EK 1], [K] were written. There is also a forthcoming book [FFK 1] about extremal problems of set systems.

Intersection theorems of structural type started in the papers of [SS 1] were followed by several others [SS 2] [SS 3] [SS 4] [R 1] [CFGs 1] [FSS 1]. The structures considered are mostly graphs or subsets of integers and the intersection properties are given in graph theoretical or arithmetical terms.

The general problem is the following:

Let  $S$  be an  $n$ -element set, and  $\underline{J}$  be a family of subsets of  $S$ .  $\underline{J}$  will be called the intersection family.

### I. Strong intersection problem.

Let  $\Lambda = \{A_1, \dots, A_m\}$  be a family of subsets of  $S$  satisfying

$$A_i \cap A_j \in \underline{J} \quad \text{for } 1 \leq i < j \leq m. \quad (1)$$

For fixed  $n$  and  $\underline{J}$  let  $g(n, \underline{J})$  denote the cardinality of the largest family  $\underline{A}$  satisfying (1). Determine  $g(n, \underline{J})$ .

An important subcase is the following.

### II. Weak intersection problem.

Using the same notation as above, let  $\underline{A} = \{A_1, \dots, A_m\}$  be a family of subsets of  $S$  satisfying the intersection conditions: for any  $1 \leq i < j \leq m$ ,

$$I \subseteq A_i \cap A_j \quad \text{for some } I \subseteq \underline{J}. \quad (2)$$

Let  $f(n, \underline{J})$  denote the cardinality of the largest family  $\underline{A}$  satisfying (2). Determine  $f(n, \underline{J})$ .

First we mention a few results of type I.

### Subsets of integers.

**THEOREM [SS [ ]].** Let  $P_k$  denote the set of arithmetic progressions of length  $\geq k$  and  $S = \{1, \dots, n\}$ . Then

$$g(n, P_k) = \left( \frac{\pi^2}{24} + o(1) \right) n^2 \quad \text{if } k \geq 2$$

and

$$g(n; P_0) = \binom{n}{3} + \binom{n}{2} + n + 1.$$

**Remark.** Curiously enough for  $k = 1$  not even the asymptotic value of  $f(n, P_1)$  is known. A plausible guess in [SS 1] is

**CONJECTURE 1.**

$$g(n; P_1) = \binom{n}{2} + 1.$$

For results on  $g(n; P_1)$  see [SS 1].

### Graph intersection theorems.

Let  $\underline{G}$  and  $H$  be graphs on the same vertex set  $V$ . Their intersection  $\underline{G} \cap H$  is the graph whose edge-set  $E(\underline{G} \cap H) = E(\underline{G}) \cap E(H)$  and whose vertices are the elements of  $V$  incident to some edges of  $E(\underline{G} \cap H)$ .

Given a family  $\mathbf{L}$  of graphs,  $h(n; \mathbf{L})$  denotes the maximum number of graphs  $\underline{G}_1, \dots, \underline{G}_N$  defined on the same  $n$ -element vertex set  $V$  for which

$$\underline{G}_i \cap \underline{G}_j \in \mathbf{L}.$$

Here we mention some results for the case when the intersection family is the set of all cycles.

**THEOREM [SSL 4].** If  $n \geq 4$  and  $C$  denotes the set of all cycles; then

$$h(n; C) = \binom{n}{2} - 2$$

and the only extremal system, that is the only  $C$ -intersection system  $\underline{G}_1, \dots, \underline{G}_N$  for  $N = f(n; C)$  is the following one:

$E(\underline{G}_1)$  forms a triangle and  $E(\underline{G}_i)$  contains  $E(\underline{G}_1)$  and exactly one additional edge for  $i = 2, \dots, \binom{n}{2} - 2$ .

**REMARK 1:** A family  $\underline{G}_1, \dots, \underline{G}_N$  is called a strong  $\Delta$ -system. The extremal system is a very stable one in the following sense.

**THEOREM [SS 4].** If  $\underline{G}_1, \dots, \underline{G}_N$  is a family of graphs on  $n$  vertices which satisfy

$$\underline{G}_i \cap \underline{G}_j \in C \quad \text{for } 1 \leq i < j \leq N$$

and which is not a strong  $\Delta$ -system

$$N \leq \frac{1}{\sqrt{6}}n^2 + n. \quad (3)$$

In the same paper we formulate the

**CONJECTURE [SS [ ]].** Let  $\underline{G}_1, \dots, \underline{G}_N$  be a family of graphs which satisfies the conditions of the previous theorem. Then

$$N \leq \frac{1}{6}n^2 + 0(n).$$

The above conjecture is sharp if true.

[CFGs 1] and [FSS [ ]] the following weak intersection theorem is proved:

**THEOREM A.** Let  $S = \{1, 2, \dots, n\}$  and  $\underline{J}_k$  be the family of sets  $\{a + 1, \dots, a + k\}$  where  $k$  is a fixed positive integer  $a = 0, 1, \dots, n$  and  $a + j$  is taken mod  $n$ . Then

$$f(n, \underline{J}_k) = 2^{n-k}. \quad (4)$$

**REMARK: 2a.** It is trivial that  $f(n; \underline{J}_k) \geq 2^{n-k}$ : take all subsets of  $S$  which contain  $\{1, 2, \dots, k\}$ .

The same trivial lower bound holds whenever  $\underline{J}$  contains a  $k$ -element set.

**b.** If  $\underline{J}_k$  is the family of sets  $\{a + 1, \dots, a + k\}$ , where  $a + k \leq n$ , then it is simpler to prove that  $f(n; \underline{J}_k) = 2^{n-k}$ .

The following intersection lemma seems to help in many intersection problems and was used to prove the theorem above.

**INTERSECTION-LEMMA [FSS 1], [CFGs 1].** Let  $\underline{A} = \{A_1, \dots, A_m\}$  be a family of subsets of  $S$ ,  $|S| = n$ . Suppose  $\underline{A}$  satisfies the following intersection property.

There exists a partition  $S = \cup_{v=1}^k S_v$  such that

$$s(A_i \cap A_j) =: |\{v | A_i \cap A_j \cap S_v \neq \phi\}| \geq r. \quad (5)$$

Then

$$m \leq \frac{c(k, r)}{2^k} \cdot 2^n \quad (6)$$

where

$$c(k, r) = \begin{cases} \sum_{i=0}^{\ell} \binom{k}{i} & \text{for } k - r = 2\ell \\ \sum_{i=0}^{\ell} \binom{k}{i} + \binom{k-1}{\ell} & \text{for } k - r = 2\ell + 1. \end{cases} \quad (7)$$

**REMARK 3:** In fact  $c(k, r)$  is the maximum number of 0-1 sequences of length  $k$  having the property that the Hamming distance of any two is at most  $k - r$ . The value of  $c(k, r)$  is given e.g., in [KA 1].

In some cases the extremal family  $\underline{A}$  is a so called kernel-system which in the weak intersection case means that

$$\bigcap_{i=1}^m A_i \in \underline{J} \quad \text{for } 1 \leq i < j \leq m. \quad (8)$$

Obviously this condition implies the intersection property. In the general case (strong intersection problem) (8) does not automatically imply the intersection property, but it gives enough information to get the extremal system.

### Hypergraph intersection problem.

In this paper we shall investigate set intersection problems where the intersection family is defined as an  $r$ -uniform hypergraph. The motivation comes from Theorem A. The problem there can be formulated as follows:

Let  $H(V)$  be a fixed Hamiltonian cycle on the vertex set  $V$ ;  $|V| = n$ . How large can the family  $\underline{A} = \{A_1, \dots, A_m\}$  be,  $A_i \subseteq V$ , if  $A_i \cap A_j$  spans at least one edge of  $H$  for every  $1 \leq i < j \leq m$ ?

This suggests the following setting of a general problem, which is a subcase of the weak intersection problem. Let  $\underline{G}^r(V; \underline{J})$  be an  $r$ -uniform hypergraph. The intersection family  $\underline{J}$  is the edge-set of  $\underline{G}^r$ . Let  $\underline{A} = \{A_1, \dots, A_m\}$  be a family of subsets of an  $n$ -element set  $S \supseteq V$ . Suppose  $\underline{A}$  satisfies the weak intersection condition (2).

$f(n; \underline{G}^r)$  (or  $f(n; \underline{J})$ ) denotes the cardinality of the largest family which satisfies (2). Determine  $f(n; \underline{G}^r)$ .

The idea behind using different settings of the general structural intersection problems is that we focus on different aspects, on different structural properties of  $\underline{J}$  which determine the order of magnitude of  $f(n; \underline{J})$ .

In this setting we are interested in the graph theoretical properties (of the hypergraph  $\underline{G}^r(V; \underline{J})$  relevant to the value of  $f(n; \underline{J})$ .

One would think that the size of  $\underline{J}$  (i.e., how rich the family  $\underline{J}$  is) has a strong effect on  $f(n; \underline{J})$ . This is not entirely the case. Our result

below shows that e.g., the chromatic number, the size of the largest independent set of  $\underline{G}^r(V; \underline{J})$  are more relevant parameters. Our results in this paper are unfortunately far from being complete, many unsolved problems are left and we are far from the complete understanding of the situation.

For an arbitrary  $\underline{G}^r(n; \underline{J})$  we have the trivial

FACT 1.

$$\frac{1}{2^r} 2^n \leq f(n; \underline{G}^r) \leq \frac{1}{2} 2^n.$$

The problems considered here actually refer to the determination of  $\lim 2^{-n} f(n; \underline{G})$  which trivially exists. First we consider the case  $\underline{G} = \underline{G}^2$  (i.e., the ordinary graphs).

PROPOSITION 1. Let  $\chi(\underline{G})$  denote the chromatic number of  $\underline{G}(V; \underline{J})$ . If  $\chi(\underline{G}) \leq k$ , then

$$f(n; \underline{G}) \leq \frac{c(k; 2)}{2^k} \cdot 2^n. \quad (9)$$

For  $\chi(\underline{G}) = 2$  or 3 then (9) is sharp, i.e.,

$$f(n; \underline{G}) = \frac{1}{4} 2^n.$$

PROOF: Let  $T_k(n_1, \dots, n_k)$  be a complete  $k$ -chromatic graph on  $n_1 + \dots + n_k$  vertices,  $V = \cup_{i=1}^k V_i$ ,  $|V_i| = n_i$ ,  $E(T_k) = \cup_{i \neq j} V_i \times V_j$ . Since  $\underline{G} \subset T_k(n_1, \dots, n_k)$  for some  $n_1, \dots, n_k$ , obviously

$$f(n; \underline{G}) \leq f(n; T_k(n_1, \dots, n_k)).$$

Now we can apply the intersection-lemma.

Suppose  $\underline{A} = \{A_1, \dots, A_m\} \subseteq 2^V$  satisfies (2). Then for any  $1 \leq i < j \leq m$

$$s(A_i \cap A_j) = |\{\ell | A_i \cap A_j \cap V_\ell \neq \phi, 1 \leq \ell \leq k\}| \geq 2.$$

By the intersection lemma this proves (9).

How relevant is the number of edges of  $\underline{G}^2$ ? For this we have

PROPOSITION 2. Let  $e(\underline{G})$  denote the number of edges of  $\underline{G}$ . Suppose

$$f(n; \underline{G}) \leq \frac{c(k, 2)}{2^k} \cdot 2^n.$$

Then

$$e(\underline{G}) \leq e(T_k(n)). \quad (10)$$

For  $T_k = T_k(n_1, \dots, n_k)$

$$f(n; T_k) = f(n; K_k) = \frac{c(k, 2)}{2^k} 2^n. \quad (11)$$

(Here  $T_k(n)$  is the Turán-graph, the complete  $k$ -chromatic graph  $T_k(n_1, \dots, n_k)$  with  $n = n_1 + \dots + n_k$ ,  $|n_i - n_j| \leq 1$ .)

To prove (10), observe that if  $e(\underline{G}) > e(T_k(n))$ , then by Turán's theorem T[1],  $K_{k+1} \subset \underline{G}$ . Let  $V^* \subseteq V$  be the vertex set of a  $K_{k+1}$  contained in  $\underline{G}$ . We define our family as follows:

$$A \subseteq \underline{A} \quad \text{if} \quad |A \cap V^*| \geq \frac{k+3}{2}.$$

Now if  $A, A' \in \underline{A}$ , then

$$|A \cap A' \cap V^*| \geq 2,$$

hence  $A \cap A'$  contains an edge of  $\underline{G}$ . For this  $\underline{A}$  we have

$$|\underline{A}| = \sum_{i \geq \frac{k+3}{2}}^{k+1} \binom{k+1}{i} \frac{1}{2^{k+1}} 2^n = \frac{c(k+1, 2)}{2^{k+1}} 2^n > \frac{c(k, 2)}{2^k} 2^n. \quad (12)$$

REMARK 4: Though the chromatic number is a relevant parameter it is not the only one. This is shown by the following fact.

Let  $W_5$  be the pentagonal wheel.  $\chi(W_5) = 4$  and one can show that

$$f(n; W_5) = \frac{1}{4} 2^n,$$

while

$$f(n; K_4) = \frac{5}{16} 2^n, \quad f(n; K_3) = \frac{1}{4} 2^n.$$

The size of the largest independent set,  $\alpha(\underline{G})$  also has an important effect on  $f(n; \underline{G})$ . We have the following simple

PROPOSITION 3. For any  $\epsilon > 0$  there is a  $\delta(\epsilon) > 0$  with the following property:

Suppose  $\underline{G}$  contains a subgraph  $\underline{G}'(V', E')$  of  $m = |V'|$  vertices for which  $\alpha(\underline{G}') < \delta\sqrt{m}$ . Then

$$f(n; \underline{G}) > \left(\frac{1}{2} - \epsilon\right) 2^n.$$

PROOF: Take all subsets of  $V$  containing at least  $\frac{m+\delta\sqrt{m}}{2}$  vertices of  $V'$ . Then the intersection of any two sets contains at least  $\delta\sqrt{m}$  vertices, but it spans at least one edge of  $\underline{G}$ .

Remark 5. It is well known that there exists a graph  $\underline{G}$  of  $n$  vertices ( $n > n_0$ ) which does not contain a  $K_4$  and  $\alpha(\underline{G}) < \delta\sqrt{n}$ . Hence we get

**COROLLARY 1.** For any  $\epsilon > 0$  there is a graph  $\underline{G}$  containing no  $K_4$  and for which

$$f(n; \underline{G}) > \left(\frac{1}{2} - \epsilon\right) 2^n.$$

We have very few results for  $r \geq 3$ . We mention only the obvious.

**PROPOSITION 4.**

$$f(n; K_{r+1}^r) = \frac{1}{2^r} 2^n$$

$$f(n; K_{r+2}^r) \geq \frac{r+3}{r+2} 2^n.$$

**REMARK:** Let  $h(n; K_3)$  denote the maximum cardinality of a family  $\underline{G}$  of graphs satisfying the intersection condition, that

$$\underline{G}_i \cap \underline{G}_j \supseteq K_3, \quad 1 \leq i < j \leq m.$$

**CONJECTURE 2.** Simonovits-Sós formulated

$$h(n; K_3) = \frac{1}{8} 2^n.$$

In [CFGS 1] it is proved that

$$h(n; K_3) < \frac{1}{4} 2^n. \quad (13)$$

If we reformulate this result in the present terminology, we arrive at the following.

**PROPOSITION 5.** For 3-uniform hypergraphs  $\underline{G}^3$  a large chromatic number does not imply that  $f(n; \underline{G}^3)$  is close to  $\frac{1}{2} 2^n$ , not even that it is larger than  $\frac{1}{4} 2^n$ .

To see this let  $\underline{G}^3$  be the 3-uniform hypergraph with vertex-set  $[V]^2$  and the edge-set  $\underline{F}(\underline{G}^3)$  the triples of  $[V]^2$  forming a triangle in  $K_n$  on the vertex-set  $V$ . The chromatic number of  $\underline{G}^3$  is  $> \log n (\log \log n)^{-1}$ . This follows from Schur's theorem implying that the value of the Ramsey-number  $R(3_1, \dots, 3_r)$  is less than  $e r!$  However (13) gives (with  $m = \binom{n}{2}$ )

$$f(m; \underline{G}^3) < \frac{1}{4} 2^m.$$

## PROBLEMS

**PROBLEM 1:** Is it true that for every 4-chromatic graph  $\underline{G}$  not containing  $K_4$  we have

$$f(n; \underline{G}) = \frac{1}{4} 2^n?$$

Or perhaps the opposite is true, the set  $\{f(n; \underline{G}) | \chi(\underline{G}) = 4, K_4 \subseteq \underline{G}\}$  is everywhere dense in  $[\frac{1}{4}, \frac{5}{16}]$ .

By Corollary 1 this is not true in general. It might be of some interest to find the "smallest" graph (i.e., smallest chromatic number or smallest number of edges) which contain no  $K_4$  and for which  $f(n; \underline{G}) > \frac{1}{4}2^n$ .

**PROBLEM 2:** Does there exist a graph  $\underline{G}$  which contains no  $K_3$  and for which

$$f(n; \underline{G}) > \frac{1}{4}2^n?$$

Perhaps the following stronger result holds. Is it true that for every  $\epsilon > 0$  and  $r$  there is a graph  $\underline{G}$  not containing a  $C_\ell$  for  $\ell \leq r$  and for which

$$f(n; \underline{G}) > \left(\frac{1}{2} - \epsilon\right) 2^n?$$

**PROBLEM 3:** What is the set of limit points of  $f(n; \underline{G}^r)2^{-n}$ ? Is it dense in  $[\frac{1}{2^r}, \frac{1}{2}]$ ?

**PROBLEM 4:** Call a graph  $\underline{G}$  critical from below resp. from above if the deletion resp. the addition of any edge decreases resp. increases  $f(n; \underline{G})$ . Obviously for  $k \geq 3$  all complete  $k$ -chromatic graphs different from  $K_k$  are critical from above. However the complete bipartite graphs are not critical.

The complete graphs  $k > 4$  are critical from below.

Can one characterise all critical graphs from above resp. from below or the graphs which are critical both from above and from below?

For  $r = 2$  we do not know any graph different from the above ones which are critical. We know by Corollary 1 that such graphs exist. It might be of some interest to find the smallest one. For  $r > 2$  we have such graphs, see Proposition 5 and Remark 7.

**PROBLEM 5:** What is the smallest number of edges of an  $r$ -uniform hypergraph  $\underline{G}^r$  for which  $f(n; \underline{G}^r) > \frac{1}{2^r}2^n$ ?

We know that  $f(n; K_4^3) = \frac{1}{8}2^n$  and  $f(n; K_5^3) = \frac{6}{2^5}2^n = \frac{3}{16}2^n$ . If we omit an edge of  $K_5^3$  we get  $\frac{1}{8}2^n$  and surely every  $\underline{H}^{(3)}$  with 9 triples has  $f(n; \underline{G}^1) = \frac{1}{8}2^n$ .

**PROBLEM 6:** Which are the  $r$ -uniform hypergraphs on  $n$  vertices with the largest number of edges for which  $f(n; \underline{G}^r) = \frac{1}{2^r}2^n$ ? Perhaps this is the  $r$  graph of maximal number of edges not containing a  $K^{(r)}(r+2)$ .

**PROBLEM 7:** Is it true that for every  $\epsilon > 0$  there is a  $\delta(\epsilon) > 0$  such that if  $\underline{G}$  has the property that for  $k > k_0$  every  $V' \subset V$  of size  $k$  contains an



independent set in  $\underline{G}$  of size  $> \epsilon\sqrt{k}$  then

$$f(n; \underline{G}) < \left(\frac{1}{2} - \delta\right) 2^n?$$

(See also Proposition 3.) It is well known that every graph not containing a  $k(3)$  satisfies the condition of Problem 7. Thus the answer to Problem 4 and the stronger form of Problem 2 can not both be affirmative.

PROBLEM 8: Let  $\underline{A} = \{A_1, \dots, A_m\}$  be a family of subsets of  $\{1, \dots, n\}$ .

- a) How large can  $m$  be if for  $1 \leq i < j \leq m$   $A_i \cap A_j$  contains a solution of ( $x, y, u, v$  are distinct positive numbers)

$$x + y = u + v, \quad (14)$$

a simple computation gives that  $m > \frac{1}{8} 2^n$ , thus the extremal system is not a kernel system.

- b) How large can  $m$  be, if for  $1 \leq i < j \leq m$   $A_i \cap A_j$  contains a solution of

$$x + y = v? \quad (15)$$

REMARKS: If instead of the above conditions we require a solution of

$$x + y + z = u + v + w \quad (16)$$

then the answer is  $(\frac{1}{2} - o(1)) 2^n$ .

This follows from the known result that given at least  $cn^{1/3}$  integers, not greater than  $n$ , then there must be a solution of (16). Hence if we take all subsets of  $\{1, \dots, n\}$  of size at least  $\frac{n}{2} + cn^{1/3}$ , the intersection of any two will be large enough to ensure a solution of (16). Perhaps the following little problem in combinatorial number theory is of some interest: Let  $1 \leq a_1 < \dots < a_t \leq n$   $t > c n^{1/r}$ . Is it true that if  $c > c_0(r)$  then there is always a set of  $2r$   $a$ 's  $a_{i_1}, \dots, a_{i_r}; a_{j_1}, \dots, a_{j_r}$  for which  $a_{i_1} + \dots + a_{i_r} = a_{j_1} + \dots + a_{j_r}$ , but all other subset sums formed from  $a_{i_1}, \dots, a_{i_r}; a_{j_1}, \dots, a_{j_r}$  are distinct?

This method does not work for (14) since there we would have to take all sets of size at least  $\frac{n}{2} + (\frac{1}{2} + \epsilon)\sqrt{n}$  and the number of these sets is  $< (\frac{1}{2} - \epsilon) 2^n$ .

Analogous reasoning can not be used e.g., for (15) since there is no density theorem for  $x + y = v$ , namely there exist sequences of integers of positive density, e.g., the odd numbers, without containing a solution of (15).

Observe that the above example gives a  $\underline{G}^3$  which contains no  $K_7^3$  and despite this property for every  $\epsilon > 0$   $f(n; \underline{G}) > (\frac{1}{2} - \epsilon) 2^n$  if  $n > n_0(\epsilon)$ .

## REFERENCES

- AK 1. Alswede and G.O.H. Katona.
- BE 1. N.G. de Bruijn and P. Erdős, *On a combinatorial problem*, Proc. Konink Nederl. Akad. Wetensch. **41** (1948), 421-423, Amsterdam.
- CFGS 1. F.K. Chung, P. Frankl, R.L. Graham and F. Shearer, *Some intersection theorems for ordered sets and graphs*.
- EKR 1. P. Erdős, Chao Ko and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. **2** (1961), 313-320.
- F SS 1. R. Faudree, R. Schelp and V.T. Sós, *Intersection theorems for set-systems and functions*, Combinatorica.
- GSS 1. R.L. Graham, M. Simonovits and V.T. Sós, *A note on the intersection properties of subsets of integers*, J. Comb. Th.(A) **28** (1980), 106-110.
- SS 2. M. Simonovits and V.T. Sós, *Graph intersection theorems*, Proc. Colloq. Combinatorics and Graph Theory (1976), 389-391, Orsay, Paris.
- SS 1. M. Simonovits and V.T. Sós, *Intersection theorems for subsets of integers*, European J. of Comb..
- SS 3. M. Simonovits and V.T. Sós, *Intersections on structures*, Combinatorial Math., Optimal design and their applications, Ann. Discrete Math. **6** (1980), 301-314.
- SS 4. M. Simonovits and V.T. Sós, *Intersection theorem for graphs II*, Coll. Math. Soc. J. Bolyai, Combinatorics **18** (1976), 1017-1029, Keszthely, Hungary.