PROBLEMS AND RESULTS IN COMBINATORIAL ANALYSIS

P. Erdös

I gave many lectures by this and similar titles, many in fact in these conferences and I hope in my lecture in 1978 I will give a survey of the old problems and describe what happened to them. In the first part of this paper I discuss some new problems in Ramsey theory and in the second part I discuss some miscellaneous old and new problems.

My paper: "Problems and results in graph theory and combinatorial analysis" in Proc. Fifth Brit. Combinatorial Conference, Univ. of Aberdeen, 1975 (Congressus Numerantium XV, Utilitas Math., 169-192) contains many problems and a fairly complete list of my combinatorial problem papers.

I.

Denote by $f_k^{(r)}(n,\alpha)$ the smallest integer for which it is possible to split the r-tuples of a set S, $|S| = n$ into k classes, so that for every $S_1 \subset S$, $|S_1| \geq f_k^{(r)}(n,\alpha)$ every class contains more than $\alpha \binom{S_1}{r}$ r-tuples of S_1 $(0 \leq \alpha < \frac{1}{k})$. The reader I hope will forgive the somewhat clumsy notation - perhaps mistakenly I feel that this is the best way to state the results and problems. For $\alpha=0$ $f_2^{(2)}(n,0)$ is the familiar Ramsey function, the smallest integer $\ell=f_2^{(2)}(n,0)$ for which there is a graph on n vertices which does not contain a complete subgraph or independent set of size ℓ (I realise that this is a somewhat unusual way of formulating the well-known Ramsey theorems). It is well-known that

$$(1) \qquad \frac{\log 2}{2} \log n < f_2^{(2)}(n,0) < 2\log 2 \log n$$

and more generally

$$(2) \qquad c_1(\alpha)\log n < f_2^{(2)}(n,\alpha) < c_2(\alpha)\log n$$

The lower bound in (1) and (2) is given by the probabilistic method (see P. Erdös and J. Spencer [1]).

Thus no great mysteries remain for the case $r=2$ (there are similar formulas for $k>2$). It would of course be very desirable to obtain an

PROC. 8TH S-E CONF. COMBINATORICS, GRAPH
THEORY, AND COMPUTING, pp. 3-12.

asymptotic formula for $f_2^{(2)}(n,\alpha)$, but no doubt this is very difficult - even $\alpha=0$ seems hopeless at present. (This is the Ramsey case and I often offered and still offer 100 dollars for the proof of the existence of the limit $f_2^{(2)}(n,0)/\log n$ and another 100 for the value of the limit, and also offer 100 dollars for a constructive proof of the lower bound of (1) and (2).) It is not difficult to obtain inequalities for $c_1(\alpha)$ and $c_2(\alpha)$, both tend to infinity if $\alpha \to \frac{1}{2}$ and it would not be too difficult to obtain rough estimates how fast they tend to infinity. The important thing is to observe that the order of magnitude of $f_k^{(2)}(n,\alpha)$ is $\log n$ for every k and every $0 \le \alpha < \frac{1}{k}$.

This situation changes radically for $r > 2$. Denote by $\log_r n$ the r times iterated logarithm. Hajnal,Rado,and I proved [2]

(3) $$c_1(r,k)\log_{r-1} n < f_k^{(r)}(n,0) < c_2(r,k)\log_{(r-2)} n.$$

The probability method only gives

(4) $$f_k^{(r)}(n,0) < c_2(r,k)(\log n)^{1/r-1}.$$

We are quite sure that in (3) the lower bound gives the correct order of magnitude. In fact Hajnal proved this for $k \ge 4$.

A completely new situation develops if α is close to $\frac{1}{2}$ (respectively $\frac{1}{k}$). For simplicity we mostly restrict ourselves to $r=3$, $k=2$ wherever possible. Let $G^{(3)}(n,[\alpha n^3])$ be a three-graph (uniform hypergraph with $r=3$) of n vertices and $[\alpha n^3]$ triples. I proved [3] that it always contains a $k_3(t,t,t)$ with $t=[c_3(\log n)^{1/2}]$ where $k_3(t,t,t)$ has $3t$ vertices which are divided into three disjoint sets $|A| = |B| = |C| = t$ and its t^3 edges are (x,y,z) $x\epsilon A$, $y\epsilon B$, $z\epsilon C$. The probability method easily gives that this theorem is best possible apart from the value of c_3 and in fact the method easily gives that one can divide the triples of $G^{(3)}\left(n,\binom{n}{3}\right)$ into two classes so that every $k_3(t,t,t)$ for $t > c_3(\log n)^{1/2}$ contains a triple from both classes.

Further I proved [4] that if we divide the triples of a $G^{(3)}\left(n,\binom{n}{3}\right)$ into two classes there always are two sets $|A|=|B|=[c_4(\log n)^{1/2}]$ so that all the triples (x,y,z), $x\epsilon A$, $y\epsilon A$, $z\epsilon B$ are in the same class. Apart from the value of c_4 the result is best possible.

Observe that a $G^{(3)}(n,[\alpha n^3])$ $\alpha < \frac{1}{6}$ does not have to contain such a configuration - here the splitting of all the triples into two classes was strongly used. I am unable to decide the order of magnitude of the value of the largest t for which there are two sets $|A|=|B|=t$ so that all the triples of $A \cup B$ which meet both A and B are in the same class. It is quite possible that the right order of magnitude of t here is $\log\log n$.

Both these results easily imply that there is an absolute constant $\delta < \frac{1}{2}$ so that for every $\alpha > \delta$

$$c_5(\alpha)(\log n)^{1/2} < f_2^{(3)}(n,\alpha) < c_4(\alpha)(\log n)^{1/2} \tag{5}$$

The lower bound is (3), the upper bound was obtained long ago by Spencer and myself [5]. To see this observe that the existence of a $k_3(t,t,t)$ all whose triples are in class, say I, implies that if $|A|=|B|=|C|=t$ are the vertices of our $k_3(t,t,t)$ the distribution of the triples cannot be "too uniform" in all 7 sets A,B,C, $A \cup B$, $A \cup C$, $B \cup C$, $A \cup B \cup C$. The simple verification can be left to the reader. This proves (5). Further a similar argument easily gives $(\delta < \frac{1}{k}, \alpha > \delta)$.

$$c_5'(\delta,k)(\log n)^{1/2} < f_k^{(3)}(n,\alpha) < c_4'(\delta,k)(\log n)^{1/2} \tag{6}$$

Let us now assume for the moment that

$$c_6 \log\log n < f_2^{(3)}(n,0) < c_7 \log\log n \tag{7}$$

has already been proved (in fact at the moment by Hajnal this is known only for $k \geq 4$ thus what we say at the moment only applies for $k \geq 4$). $f_2^{(3)}(n,\alpha)$ is for $\alpha=0$ of the order of magnitude $\log\log n$. For some $\alpha=\delta$, $0 < \delta < \frac{1}{2}$ it becomes of order of magnitude $(\log n)^{1/2}$ and it follows from the probability method that for $\delta < \alpha < \frac{1}{2}$ this is the correct order of magnitude. Clearly $f_2^{(3)}(n,\alpha)$ is for fixed n a non-decreasing function of α and it would be very interesting to determine where the jump occurs (from $\log\log n$ to $(\log n)^{1/2}$) and if the jump occurs in several stages. If I can hazard a guess - completely unsupported by evidence - I am afraid that the jump occurs all in one step at 0 and for $0 < \alpha < \frac{1}{2}$, $f_2^{(3)}(n,\alpha)$ grows continuously. It would of course be more interesting if several jumps would occur, perhaps for $r > 3$ where $f_2^{(r)}(n,\alpha)$ has to grow from $\log_{r-1} n$ to $(\log n)^{1/r-1}$ there is more chance for this but I know nothing and hope one of my readers will be more successful. In any case I offer 300 dollars for the clearing up of this mystery, or for any substantial progress in this direction.

There is another older problem of mine on hypergraphs which also shows the increase of complication from $r=2$ to $r > 2$ and which I now restate. Let $G^{(r)}(n_i;m_i)$, $n_1 < n_2 < \ldots$ be an infinite sequence of r-uniform hypergraphs. The edge density of these hypergraphs is said to be α if α is the largest number for which for every T there is an n_i which contains a subgraph $G^{(r)}(u_i;v_i)$ $u_i > T$, $v_i > (\alpha+o(1))\binom{u_i}{r}$. The theorem of Stone and myself [6] shows that for $r=2$ the only possible values of α are $1 - \frac{1}{k}$, $k=1,2,\ldots,\infty$. I proved in [3] that for $r > 2$ if $\alpha > 0$ then $\alpha \geq \frac{r!}{r^r}$. First of all I conjecture that there is a constant $c_r < 0$ so that if $\alpha > r!/r^r$ then $\alpha > \frac{r!}{r^r} + c_r$. Next try to determine for every r all possible values for the density of a family of r-uniform hypergraphs. It seems certain that this set is countable but our work with Brown and Simonovits ([7] most of which is still unpublished) seems to indicate that a complete answer to this question will not be easy.

Some of these questions could be investigated for the subgraphs of other graphs than the complete graphs e.g. it is easy to see that the edge density of subgraphs of the complete bipartite graphs is either 0 or 1 but perhaps there are other classes of graphs (or hypergraphs) where non trivial statements can be made about the edge density of subgraphs. The graph of the n-dimensional cube seems a natural candidate - unfortunately I do not believe that there are interesting and non trivial theorems here, perhaps a reader will be more successful than I. The n-dimensional cube has 2^n vertices and $n2^{n-1}$ edges. Let G be a subgraph having $\varepsilon\, n\, 2^n$ edges - is there an α so that there is a k-dimensional subgraph $(k=k(n) \to \infty$ as $n \to \infty)$ so that this subcube has $> \alpha\, k\, 2^k$ edges of G^2. I very much doubt that this is true.

A similar situation may prevail with Van der Waerden's Theorem: Denote by $F(k,k,\alpha)$ the largest integer so that we can split the integers not exceeding $F(k,k,\alpha)$ into two classes so that both classes contain at least αk $(\alpha < \frac{1}{2})$ terms in every arithmetic progression of k terms. It is not hard to prove by the probabilistic method that $F(k,k,\alpha) > (1+c_\alpha)^k$, *and for* $\alpha=0$ obtain the classical Van der Waerden theorem.

It is possible that the order of magnitude of $F(k,k,\alpha)$ depends significantly on α but nothing at all is known about this.

1. P. Erdös and G. Szekeres, A combinatorial problem in geometry, Comp. Math. 2 (1935), 463-470; P. Erdös, Some remarks on the theory of graphs, Bull. Amer. Math. Soc. 53 (1947), 292-294. For the probability method see P. Erdös and J. Spencer, Probabilistic methods in combinatorics, Publishing House Hung. Acad. Sci. and Academic Press (Prob. and Math. stat. 17), 1974.
2. P. Erdös and R. Rado, Combinatorial theorems on classifications of subsets of a given set, Proc. London Math. Soc. (3) 2 (1952), 417-439; P. Erdös, A. Hajnal and R. Rado, Partition relations for cardinal numbers, Acta Math. Acad. Sci. Hung. 16 (1965), 93-196 see p. 139-140.
3. P. Erdös, On extremal problems of graphs and generalised graphs, Israel J. Math. 2 (1965), 183-190.
4. P. Erdös, On some extremal problems on r-graphs, Discrete Math. 1 (1971), 1-6.
5. P. Erdös and J. Spencer, Imbalances in k-colorations, Networks 1 (1972), 379-385.
6. P. Erdös and A.H. Stone, On the structure of linear graphs, Bull. Amer. Math. Soc. 52 (1946), 1087-1091. For further results see P. Erdös and M. Simonovits, A limit theorem in graph theory, Studia Sci. Math. Hung. Acad. 1 (1966), 51-57, and B. Bollobás, P. Erdös and M. Simonovits, On the structure of edge graphs II, J. London Math. Soc. 12 (1976), 219-224.
7. W.G. Brown, P. Erdös and M. Simonovits, Extremal problems for directed graphs, J. Comb. Theory (B) 15 (1973), 77-93.

II

I now state a few miscellaneous problems and results which occupied me and my coworkers recently all of which have a combinatorial flavor.

1. Kneser defined a graph as follows: Let $|S|=2n+k$. The vertices of our graphs are the $\binom{2n+k}{n}$ subsets of size n of S. Two vertices are joined if the corresponding subsets are disjoint. Kneser conjectured that the chromatic number of this graph is $k+2$. It is easy to see that

the chromatic number is $\le k+2$, the whole difficulty was to show that it is not less than $k+2$. Szemerédi proved some time ago that the chromatic number is $> f(k)$ where $f(k) \to \infty$ as $k \to \infty$. A few weeks ago Lovász proved Kneser's conjecture. His proof will appear soon. Just now (1977 IV 7) Báránz found a simpler proof of Kneser's conjecture).

As far as I know the following stronger conjecture (which can be considered as an extension of a theorem of Ko, Rado and myself [1]) is still open:

Let $|S|=2n+k$ and T_n the family of its $\binom{2n+k}{n}$ n-tuples. Let further $F_1,\dots,F_r$, $r \le k$ be disjoint subsets of T_n where for every i, $1 \le i \le r$ any two n-tuples of F_i have an element in common. (In other words in our graph the F_i are independent.) Put

$$f(n;r) = \max \sum_{i=1}^{r} |F_i|.$$

The well known theorem of Ko, Rado and myself implies $f(n;1) = \binom{2n+k-1}{n-1}$. Determine $f(n;r)$ for $r > 1$. This question was raised by Hajnal and myself independently several years ago. Probably for $n > n_0(k,r)$

(1) $$f(n;r) = \sum_{i=1}^{r} \binom{2n+k-i}{n-1}$$

Hilton pointed out that (1) does not always hold for $r \ge 2$. The conjecture of Kneser would follow from

(2) $$f(n;r) < \sum_{i=1}^{r+1} \binom{2n+k-i}{n-1}$$

I hope that (1) holds with very few exceptions and that it will be possible to determine all the exceptional cases.

P. Erdös, C. Ko and R. Rado, Intersection theorems for systems of finite sets, Quarterly J. Math. (2) 12 (1961), 313-320.

2. V.T. Sós and I conjectured that if $|S|=n$, $n > n_0(k)$ and $A_i \subset S$, $|A_i|=k$, $1 \le i \le t$, $t > \binom{n-2}{k-2}$ then for some $1 \le i_1 < i_2 \le t$, $|A_{i_1} \cap A_{i_2}|=1$. That $t > \binom{n-2}{k-2}$ is needed is shown by the sets containing the same two elements of S. We observed that for $k=3$ t must be $n+1$. Katona (unpublished) proved our conjecture for $k=4$ and recently Frankl

proved our conjecture for all $k > 4$; his proof will be published soon (Bull. Australian Math. Soc.).

The following related conjecture of mine is still open: Denote by $F(n,r)$ the smallest integer for which if $A_i \subset S$, $1 \leq i \leq F(n,r)$ is any family of $F(n,r)$ subsets of the set S of size n there are always two of them A_{i_1} and A_{i_2} with $|A_{i_1} \cap A_{i_2}| = r$. I conjecture that to every $\eta > 0$ there is an $\varepsilon = \varepsilon(\eta)$ so that

$$\max_{\eta n < r < (\frac{1}{2} - \eta) n} F(n,r) < (2-\varepsilon)^n. \tag{3}$$

I am very far from being able to prove (3). It would of course be of interest to determine $F(n,r)$ explicitly for every r and n. (3) would have immediate applications in n-dimensional geometry.

D.G. Larman and C.A. Rogers, The realisation of distances within sets in Euclidean space, Mathematics 19 (1972) 1-24.

3. Euclidean Ramsey problems. We [1] conjectured that if S is a set in the plane, no two points of S are at distance 1, then the complement of S contains the vertices of a unit square - more generally: what sets of points can be imbedded congruently in S?

Ms. R. Juhász just proved that every set of four points can be embedded in S, but that 12 points on the line at distance 1 cannot in general be embedded; her proof will appear in the Journal of Combinatorial Theory.

I conjectured and R.L. Graham proved a few weeks ago that if we divide the plane into a finite number of sets $\bigcup_{i=1}^{k} S_i$, then for some i S_i contains three points $X;Y;Z$ so that the triangle (X,Y,Z) has area 1. More generally he shows that for some i S_i contains the vertices of triangles of all areas. I further conjecture that if the plane is divided into infinitely many sets S_i, $1 \leq i < \infty$, then for some i there is an S_i so that S_i contains all triangles of area $< \varepsilon$.

The following old problem of mine is as far as I know still open: Is it true that there is a c so that if S is a measurable set in the plane of plane measure $> c$ then S contains the vertices of a triangle of area 1?

Recently in the Journal of Recreational Mathematics Silverman stated several interesting questions about decomposing sets of simple structure into a finite number of sets none of which contains three points determining a right angle. This prompted me to conjecture: Is it true that a set of n^2 points in the plane always contains $2n-2$ points which do not determine a right angle? The lattice points (x,y), $0 \le x < n$, $0 \le y < n$ show that the conjecture if true is best possible.

Perhaps the conjecture is too optimistic - if a reader finds a counterexample $2n-2$ should be replaced by cn. I can only prove $cn^{2/3}$ instead of cn.

Consider the n^3 lattice points (x,y,z), $0 \le x,y,z < n$ in three dimensional space. Determine the largest subset which does not contain the vertices of a right angled triangle. I could not solve this problem but perhaps it is easy and I overlook something obvious.

A few years ago Fajtlowicz asked: Can one decompose the plane into $\aleph_0$ sets S_i, $1 \le i < \infty$ so that none of the S_i contain the three vertices of a right angled triangle. I proved that this is possible if and only if $c = \aleph_1$.

P. Erdös, R.L. Graham, P. Montgomery, B.L. Rothschild, J. Spencer, E.G. Straus, Euclidean Ramsey theorems I, II, III, J. Comb. Theory, Ser. A 14 (1973), 341-363, Coll. Math. Soc. J. Bolyai 10, Infinite and Finite Sets (Hungary) 1973 vol. 1 529-557 and 559-583.

4. Silverman and I conjectured that if G_r is a graph whose vertices are the integers and we join i and j if $i+j$ is an r-th power, then the chromatic number of G_r is infinite for every r. We could not even prove this for $r=2$. This is really a problem in number theory: Is it true that if one divides the set of integers into k classes then there are always two distinct integers of the same class whose sum is an r-th power? If sum is replaced by difference this is a theorem of Sárközy and Furstenberg. They in fact prove that if $a_1 < a_2 < \dots < a_k \le X$ is such that none of the differences $a_i - a_j$ is an r-th power than $k=o(X)$. Nothing like this is true for $a_i + a_j$ e.g. the sum of two integers $\equiv 1 \pmod 3$ is never a square. Denote in fact by $f_r(n)$ the largest integer k for which there is a sequence $1 \le a_1 < \dots < a_k \le n$ so that no two sums $a_i + a_j$ are the r-th powers. Determine or estimate $f_r(n)$. The same question can

of course be asked if the a's are elements of a group. Sárközy's proof will soon appear.

5. Let $G(n)$ be a regular graph of n vertices. Berge conjectured that $G(n)$ always has two disjoint maximal independent sets i.e. if $X_1,\dots,X_n$ are the set of vertices of $G(n)$ there are two disjoint independent subsets $Y_1,\dots,Y_k$ and $Z_1,\dots,Z_\ell$ so that every $Y \neq Y_i$, $1 \le i \le k$ is joined to one of the Y's and every $X \neq Z_j$, $1 \le j \le \ell$ is joined to one of the Z's.

Hobbs and I proved a few weeks ago that there is an $\varepsilon > 0$ so that if every vertex of $G(n)$ has valency (or degree) not less than $n-(2+\varepsilon)n^{1/2}$ then $G(n)$ certainly has two disjoint maximal independent sets. Also we showed that our theorem certainly fails if $n-(2+\varepsilon)n^{1/2}$ is replaced by $n-cn^{2/3}$. It would be of some interest to determine or estimate the largest $f(n)$ for which every $G(n)$, each vertex of which has degree $> n-f(n)$, contains two disjoint maximal independent sets. It is very doubtful though whether this would throw any light on the conjecture of Berge. Berge by the way informs me that his conjecture has been proved if the valency of $G(n)$ is not more than 7.

6. Blanchard considered the following problem: Let S be a set of n elements. Denote by $f(n)$ the maximum number of pairs chosen from the elements of S so that the union of any two of the pairs is different. He proves (Bull. Assoc. Proc. Math. No. 300 (1975), p. 538) $c_1 n^{3/2} < f(n) < c_2 n^{3/2}$, and he asks: is it true that $\lim f(n)/n^{3/2} = c$ exists and determine its value.

Bollobás and I considered the following generalisation: Denote by $f_r(n)$ the maximum number of r-tuples chosen from $|S|=n$ so that the union of any two of the pairs is different. We proved

$$c_1' n^2 < f_3(n) < c_2' n^2.$$

In fact we proved that if $|A_i|=3$, $1 \le i \le c_2'' n^2$ then one can always find four A's $A_{i_1}, A_{i_2}, A_{j_1}, A_{j_2}$ satisfying $A_{i_1} \cap A_{i_2} = A_{j_1} \cap A_{j_2} = \emptyset$ and $A_{i_1} \cup A_{i_2} = A_{j_1} \cup A_{j_2}$.

In graph theoretic language Blanchard's problem can be stated as follows: Denote by $f(n)$ the largest integer for which there is a graph

$G(n;f(n))$ (n vertices and $f(n)$ edges) which contains no c_3 and no c_4. It is known that

$$\left(\frac{1}{2\sqrt{2}} + o(1)\right)n^2 < f(n) < \left(\frac{1}{2} + o(1)\right)n^2.$$

W.G. Brown, On graphs that do not contain a Thomsen graph, Bull. Canad. Math. Soc. 9 (1966), 281-288, P. Erdös, A. Rényi and V.T. Sós, On a problem of graph theory, Studia Math. Hung. Acad. Sci. 1 (1966), 215-235, I. Reiman, Über ein Problem von Zaranhiewicz, Acta. Math. Acad. Sci. Hungar 9 (1958), 269-278.