# SOME NEW PROBLEMS AND RESULTS IN GRAPH THEORY AND OTHER BRANCHES OF COMBINATORIAL MATHEMATICS

PAUL ERDÖS
The Hungarian Academy of Sciences
Budapest, Hungary

## § 0.

Recently I published several papers on finite and infinite combinatorial problems. I will try to make the overlap with this paper as small as possible; as a result I have to omit some of my most interesting problems, but first of all I give some references to my older papers where these questions have been discussed :

P. Erdös, Old and new problems in combinatorial analysis and graph theory, Second International Conference on Combinatorial Mathematics, New York Academy of Sciences, Vol. 319 (1979), 177-187.

P. Erdös, Problems and results on finite and infinite Combinatorial Analysis I and II, Coll. Math. Soc. J. Bólyai, 10 : Infinite and finite sets, Kenthely, Hungary (1973), 403-424, II will appear in Enseignement Math. in 1981.

P. Erdös, Some old and new problems in various branches of Combinatorics, Proc. Tenth Conference in Combinatorics, Graph Theory and Computing (1979) (Boca-Raton Conference). This paper contains extensive references to my previous papers.

P. Erdös, Combinatorial problems which I would most like to see solved, will soon appear in the new Hungarian periodical Combinatorica.

For applications of probabilistic methods to combinatorial analysis see our book, P. Erdös and J. Spencer : Probabilistic methods in Combinatorics, Acad. Press and Hung. Acad. Sci. (1974).

## § 1.

First I discuss problems connected with Ramsey's theorem and its generalisations, here I of course can not avoid overlap with previous papers. $r(n_1,\dots,n_k)$ is the smallest integer for which if one colors the edges of $K(r(n_1,\dots,n_k))$ by $k$ colors $(K(t)$

is the complete graph of $t$ vertices) then there is always an $i$, $1 \leq i \leq k$, so that there is a $K(n_i)$ all of whose edges are of the $i$-th color. Very interesting problems arise if $k$ tends to infinity, but we will not discuss these in great detail. We just mention that it is not even known how fast $r(n_1,\ldots,n_k)$ tends to infinity if all the $n_i$ are 3. This problem goes back essentially to I. Schur who proved

$$r_k(C_3) = r_k(3,\ldots,3) < e.k!\,.$$

It is not yet known if

$$r_k(C_3) < C^k \qquad \ldots \qquad (1)$$

holds for all $k$ if $C$ is a sufficiently large absolute constant. More generally it is quite possible that there is an absolute constant $C$ so that

$$r(n_1,\ldots,n_k) < C^{n_1+\ldots+n_k} \qquad \ldots \qquad (2)$$

It is easy to show by induction that

$$r(n_1,\ldots,n_k) < k^{\sum_{i=1}^{k}(n_i-2)}\,.$$

The proof or disproof of (1) and (2) seem to be very interesting questions.

Let us now restrict ourselves to $k = 2$. It is well known that

$$c_1\, n^{\frac{1}{2}}.2^{\frac{n}{2}} < r(n,n) < c_2 \binom{n}{[\frac{n}{2}]} \frac{\log\log n}{\log n} \qquad \ldots \qquad (3)$$

I offered and offer 1000 rupees (or an equivalent in Swiss Francs) for a proof or disproof of

$$\lim_{n\to\infty} r(n,n)^{\frac{1}{n}} = C \qquad \ldots \qquad (4)$$

I offer another 1000 rupees for the value of $C$. I think that perhaps the proof for the existence of $C$ will not be difficult (though I can not do it), but I do not think the determination of $C$ will be easy and I have no idea what its value will be.

It is known that

$$\frac{c_1.n^2}{(\log n)^2} < r(3,n) < c_2\,\frac{n^2}{\log n}. \qquad \ldots \qquad (5)$$

The lower bound is due to me. The upper bound was proved very recently by Ajtai, Komlós and Szemerédi who improved the previous bound $\frac{c\, n^2 \log\log n}{\log n}$ of Graver and

Yackel.

Very likely for every $k$ and $\varepsilon > 0$, if $n \to \infty$

$$r(k,n) > n^{k-1-\varepsilon}. \qquad \ldots \qquad (6)$$

In fact probably

$$r(k,n) > c_1 \frac{n^{k-1}}{(\log n)^{c_2}}. \qquad \ldots \qquad (6')$$

All our attempts to prove (6) and (6') - even for $k = 4$ - failed completely. It is not impossible that the difficulties are only technical. Both the upper and lower bounds of (4) are obtained by probabilistic methods and probably (6) will have to be attacked similarly.

Almost nothing is known about the local growth properties of $r(n,m)$. S. Burr and I conjectured that

$$r(n+1,n) > (1+c)\, r(n,n), \qquad \ldots \qquad (7)$$

but at the moment (7) is intractable. Faudree, Schelp, Rousseau and I needed recently a lemma stating

$$\lim_{n\to\infty} \frac{r(n+1,n) - r(n,n)}{n} = \infty. \qquad \ldots \qquad (8)$$

We could prove (8) without much difficulty, but could not prove that $r(n+1,n) - r(n,n)$ increases faster than any polynomial of $n$. We of course expect

$$\lim_{n\to\infty} \frac{r(n+1,n)}{r(n,n)} = C^{\frac{1}{2}}$$

where $C = \lim\limits_{n\to\infty} r(n,n)^{\frac{1}{n}}$.

V.T. Sós and I recently needed the following results.

$$r(n+1,3) - r(n,3) \to \infty, \qquad \ldots \qquad (9)$$

and

$$r([n(1+c_1)],3) > (1+c_2)\, r(n,3). \qquad \ldots \qquad (9')$$

Both (9) and (9') must certainly be true but we could certainly not prove them. Probably

$$(r(n+1,3) - r(n,3))/n^{\frac{1}{2}} \to 0. \qquad \ldots \qquad (10)$$

All these results would easily follow if one could get a good asymptotic formula with a good error term for $r(n,3)$, but needless to say this is nowhere in sight.

One of the reasons for our inability to prove such simple results may be that we lack constructive methods for giving good lower bounds for $r(m,n)$. I offer 1000 rupees for a constructive proof of $r(n,n) > (1+c)^n$. The currently known sharpest constructive proof is due to P. Frankl who proved that

$$\lim_{n\to\infty} \frac{r(n,n)}{n^k} = \infty,$$

for every $k$.

Denote by $r(C_{2n+1},k)$ the largest integer $t$ for which the edges of $K(t)$ can be colored by $k$ colors so that there should be no monochromatic $C_{2n+1}$. Graham and I conjectured that

$$\lim_{n\to\infty} r(C_{2n+1},k)/r(C_3,k) = 0 \qquad \ldots \quad (11)$$

(11) is open even for $n = 2$. Perhaps the proof of $r(C_5,k) < C^k$ will not be too hard. Another problem of Graham and myself states : It is well known and easy to see that the edges of $K(2^r)$ can be colored by $r$ colors so that each color graph is bipartite and that such a decomposition does not exist for $K(2^r+1)$. Let now $f(r)$ be the smallest integer so that every coloring of the edges of $K(2^r+1)$ by $r$ colors contains a $C_{2f(r)+1}$. Estimate $f(r)$ as well as possible.

Now I discuss the so called generalised Ramsey numbers. The systematic formulation of the problems was due to Harary and Cockayne. Let $G_1,\ldots,G_k$ be $k$ graphs, then $r(G_1,\ldots,G_k)$ is the smallest integer $n$ for which if we color the edges of $K(n)$ by $k$ colors, then there is an $i$ so that the edges of the $i$-th color contain $G_i$ as a subgraph. Chvátal and Harary proved using the method of the proof of (3) that if $G$ is $t$-chromatic, then

$$r(G,G) > (1+c)^t. \qquad \ldots \quad (12)$$

After learning of (12) I conjectured that

$$\min_G r(G,G) = r(t,t). \qquad \ldots \quad (13)$$

In other words, if $G$ runs through the family of $t$-chromatic graphs, then the minimum of $r(G,G)$ is assumed for the complete graph $K(t)$ and further I conjecture that the minimum is assumed only for this graph. This is trivial for $t = 3$, but $t = 4$ alread seems to present considerable difficulties. Let, in particular, $G$ be the pentagonal wheel. The conjecture for $t = 4$ would follow if we could prove

$$r(G,G) > r(4,4) = 18. \qquad \dots \quad (14)$$

Perhaps (14) could be proved in cooperation with a computer, the best results so far are due to Chvátal and Schwenk and they proved that $17 \leq r(G,G) \leq 21$.

Burr recently published two excellent survey papers on the generalised Ramsey numbers, also Burr, Faudree, Rousseau, Schelp and I published several papers on this subject and several more of our papers will be published soon - here I give a short summary of some of our results and open problems.

Following some preliminary results of Bondy and myself, V. Rosta and independently Faudree and Schelp determined $r(C_n,C_m)$ for every $n$ and $m$. Bondy and I conjectured

$$r(C_n,C_n,C_n) \leq 4n - 3, \qquad \dots \quad (15)$$

which is still open. For odd $n$, (15), if true, is best possible.

Denote by $G(n)$ a graph of $n$ vertices. $G(n)$ is said to have edge density $\leq C$ if for every subgraph $G(m)$ of $G(n)$, we have $e(G(m)) < C.m$, where $e(G)$ denotes the number of edges of $G$. Burr and I conjectured that if $G(n)$ has edge density $< C$, then

$$r(G(n),G(n)) < f(C).n. \qquad \dots \quad (16)$$

The proof of this very attractive conjecture is nowhere in sight. Denote by $G_c(n)$ the graph determined by the edges of the $n$-dimentional cube; $G_c(n)$ has $2^n$ vertices and $n\,2^{n-1}$ edges. We could not decide whether for some absolute constant $c_1$

$$r(G_c(n),G_c(n)) < c_1.2^n \qquad \dots \quad (16')$$

is true. (16) and (16') seem to me to be two very attractive problems. Burr and I expected (16) to be true and (16') to be false.

Now I state some of the problems and results of our work with Faudree, Rousseau and Schelp. We are fairly certain that $r(K(n),C_4) < n^{2-\varepsilon}$ holds for a certain $\varepsilon$ and all $n > n_o(\varepsilon)$, but all we could prove is that for $r \geq 5$, $r(K(n),C_r) < n^{2-\varepsilon}$.

Our most striking and original problem states the following : Denote by $\hat{r}(G_1,G_2)$ the smallest integer $m$ for which there is a graph $G$ of $m$ edges so that if we color the edges of $G$ by two colors, then either color I contains $G_1$ or color II contains $G_2$. We called $\hat{r}(G_1,G_2)$ the size Ramsey number of $G_1$ and $G_2$. Let $P_n$ be the path of length $n$. Is it true that

$$\frac{\hat{r}(P_n,P_n)}{n^2} \to 0 \quad \text{but} \quad \frac{\hat{r}(P_n,P_n)}{n} \to \infty \ ? \qquad \ldots \quad (17)$$

Of course one really would like to determine $\hat{r}(P_n,P_n)$ exactly or at least to get an asymptotic formula for it; but in fact we could not make any progress with (17).

Harary recently asked the following question : Let $G_n$ be a graph of $n$ edges. What is the smallest possible value of $r(G_n,G_n)$ ? We are far from being able to give a complete solution but could prove that there is a $G_n$ for which

$$r(G_n,G_n) < c.n^{2/3}. \qquad \ldots \quad (18)$$

Perhaps 2/3 is the best exponent, we only know that 2/3 can not be replaced by an exponent less than 3/5. We further conjectured that

$$r(G_n,3) \leq 2n + 1, \qquad \ldots \quad (19)$$

but could prove it only with $3\text{-}c$ instead of $2n + 1$. (19), if true, is best possible.

P. Erdös and J. Spencer, Probabilistic methods in Combinatorics; Acad. Press and Hung. Acad. Sci. (1974).

S.A. Burr, Generalised Ramsey theory for graphs, in Graphs and Combinatorics, R. Bari and F. Harary (eds.), Springer Verlag, Berlin 52-75.

P. Frankl, A constructive lower bound for some Ramsey numbers, Ars Combinatoria 3(1977), 297-302.

P. Erdös and R.L. Graham, On partition theorems for finite graphs, Coll. Math. Soc. J. Bólyai 10 : Infinite and finite sets, Keszthely, Hungary (1973), 515-527.

S. Burr and P. Erdös, On the magnitude of generalised Ramsey numbers for graphs, ibid, 215-240.

V. Rosta, On a Ramsey type problem of J.A. Bondy and P. Erdös, Journal of Comb. Theory Ser. B, 15(1973), 94-120.

P. Erdős, R.J. Faudree, C.C. Rousseau and R.H. Schelp, The size Ramsey number, Periodica Math., Vol.9.

S.A. Burr and us, Ramsey minimal graphs for multiple copies, Indagationes Math., 40(1978), 187-195.

§ 2.

Now I discuss some recent problems on number theory and geometry. Let $1 \le a_1 < \dots < a_k \le x$. Assume $a_i + a_j \nmid a_i a_j$. Put $f(x) = \max k$. I first assumed that it will be easy to show that $f(x) = (\frac{1}{2} + o(1))x$. The odd numbers show that $f(x) > \frac{x}{2}$. I am no longer sure that my conjecture is correct. Odlyzko found with the aid of a computer that $f(1000) \ge 717$ and now I am no longer sure what happens. A related question states : Let $1 \le b_1 < \dots < b_\ell \le x$, $(b_i + b_j) \nmid 2b_i b_j$. Put $g(x) = \max \ell$. Is it true that $g(x) = o(x)$ ? Clearly if $a_i + a_j \nmid a_i a_j$ then $3x$ and $6x$ can not be both a's. I am sure that there is a sequence $u_1 < u_2 < \dots$, $\Sigma \frac{1}{u_i} < \infty$ so that the set of integers not divisible by any of the u's satisfies $(a_i + a_j) \nmid a_i a_j$ and that this sequence will give $\lim f(x)/x$. The details are not quite clear.

Silverman and I some time ago asked the following questions. Define a graph whose vertices are the integers, as follows : Join $i$ to $j$ if $i + j$ is a square. Is it true that this graph has infinite chromatic number? We also asked : Let $1 \le u_1 < \dots < u_t \le x$ and assume that $u_i + u_j$ is never a square. Put $\max t = h(x)$. $h(x) > x/3$ is trivial - take the $u_i \equiv 1 \pmod 3$. Is it true that $h(x) = (\frac{1}{3} + o(1))x$ ? A weaker conjecture states : Let $v_1, \dots, v_\ell$ be residue classes (mod d). Assume that $x^2 \not\equiv (v_i + v_j) \pmod d$ for $1 \le i \le j \le \ell$. Is it true that $\ell \le \frac{d}{3}$ ? If not - how large can $\ell$ be? None of these questions has been investigated very carefully and I have to ask the indulgence of the reader if the answer turns out to be trivial.

One of my oldest questions in number theory states : Is it true that the density of integers $n$ which have two divisors $d_1 < d_2 < 2d_1$ is one, i.e., almost all integers have two divisors which are close together? I proved long ago that the density of these integers exists but could never prove that it is 1.

Denote by $d(n)$ the number of divisors of $n$ and by $d^+(n)$ the number of integers $k$ for which $n$ has a divisor in $(2^k, 2^{k+1})$. If my conjecture is correct then for almost all integers $n$, $d^+(n) < d(n)$. I conjectured that in fact for almost all integers $d^+(n)/d(n) \to 0$. Tenenbaum and I last summer at the number theory meeting in Durham disproved this conjecture and recently Tenenbaum obtained an inequality on the density of the integers $n$ for which $d^+(n) < C.d(n)$, but I still believe that for almost all $n$, $d^+(n) < d(n)$ is true.

Denote by $d_t(n)$ the number of divisors of $n$ in $(t,2t)$ and put $D(n) = \max_t d_t(n)$. As far as I know Hooley was the first to investigate $D(n)$. He proved

$$\sum_{n=1}^{x} D(n) < x(\log x),$$

and I proved

$$\frac{1}{x}\sum_{n=1}^{x} D(n) \to \infty .$$

Hooley asked : Is it true that for every $\varepsilon > 0$

$$\sum_{n=1}^{x} D(n) = o(x(\log x)^{\varepsilon}) \ ? \qquad \dots \quad (20)$$

I several times tried to prove (20) but I was unsuccessful so far.

Now I state a few problems in Combinatorial Geometry. I published several survey papers on this subject and G. Purdy and I hope to write a book on this subject, here I only state problems which are at least partially new. Let $x_1,\dots,x_n$ be $n$ distinct points in $k$-dimensional space. Join two of the $x_i$ if their distance is one and denote this graph by $G(x_1,\dots,x_n)$. The chromatic number $\chi(E^{(k)})$ is defined as the upper bound of $\chi(G(x_1,\dots,x_n))$ where $n$ and the $\{x_i\}$ are both variable. Hadwiger and Nelson first asked for the determination of $\chi(E^{(2)})$. It is now known that $4 \le \chi(E^{(2)}) \le 7$ and it seems likely that $\chi(E^{(2)}) > 4$. Probably

$$\chi(E^{(k)}) > (1+c)^k, \qquad \dots \quad (21)$$

but we are very far from being able to prove (21), the sharpest result so far is due to P. Frankl who proved $\chi(E^{(k)}) > k^{\ell}$ for every fixed $\ell$ if $k > k_o(\ell)$; he sharpened previous results of Larman and Rogers.

Simonovits and I define the essential chromatic number $t = \chi_e(M)$ of a metric space $M$ as follows : $t$ is the smallest integer so that for every $G(x_1,\dots,x_n)$ we can omit $o(n^2)$ edges from $G(x_1,\dots,x_n)$ so that the resulting graph has chromatic number $\leq t$. Simonovits and I prove $\chi_e(E^4) = 2$, $\chi_e(E^k) \geq k-2$. In fact we conjecture $\chi_e(E^k) > (1+c)^k$. $\chi_e(E^2) = \chi_e(E^3) = 1$ simply means that the number of edges of $G(x_1,\dots,x_n)$ is $o(n^2)$. Our paper on this subject will appear in Ars Combinatoria soon.

Recently the following question in elementary geometry occured to me : Is it true that for every $n$ there are $n$ distinct points in the plane in general position (i.e., no three on a line and no four on a circle) so that these points determine exactly $n-1$ distinct distances where further the $i$-th distance occurs $i$ times. The existence of such a set is trivial for $n = 3$ and $n = 4$ (an isosceles triangle with the centre of its circumscribed circle shows this). I thought that such a set does not exist for $n = 5$ but Pomerance gave a simple example for such a system : $x_1,x_2,x_3$ are the vertices of an equilateral triangle, $x_4$ is the centre of its circumscribed circle, $x_5$ is the point of intersection of the perpendicular bisector of $(x_3,x_4)$ with the circle with centre $x_1$ and radius $d(x_1,x_2)$ where $d(x_1,x_2)$ is the distance of $x_1$ to $x_2$. It is easy to see that these points are in general position and the $i$-th distance occurs $i$ times. I believe no such system exists for $n \geq 6$. Perhaps, if we also require that no circle whose centre is one of our points should contain three of our points, then such a system can not exist for $n = 5$.

P. Erdös, On some problems of elementary geometry. Annali Math. Pure Apl., 103(1975), 99-108.

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

P. Erdös, Combinatorial problems in geometry and number theory, Proc. Symp. Pure Math., Amer. Math. Soc., 34(1979), 149-162.

C. Hooley, On a new technique and its applications to the theory of numbers, Proc. London Math. Soc., 38(1979), 115-151 (see p. 125-128).

For further problems and results on sequences of integers see the excellent book of Hallerstam and Roth, "Sequences", 1966, Oxford.