SOME APPLICATIONS OF GRAPH THEORY TO NUMBER THEORY

Paul Erdös, Hungarian Academy of Science

The problems which we will discuss in this paper deal with sequences of integers; they are all of a combinatorial nature and graph theoretic results can be applied to some of them.

First we define the concept of an r-graph (for $r = 2$ we obtain the ordinary graphs). The elements of the r-graph are its vertices some of whose r-tuples belong to our r-graph. $G_r(n, t)$ denotes an r-graph of n vertices and t r-tuples. $G_r(n; \binom{n}{r})$ denotes the complete r-graph $K_r(n)$ and $K_r(p_1, \ldots, p_r)$ denotes the r-graph of $(p_1 + \ldots + p_r)$ vertices with p_i vertices of the i-th class where each r-tuple all whose vertices are in different classes belongs to our $K_r(p_1, \ldots, p_r)$. *Throughout this paper, "graph" will indicate a* 2-graph. We will denote 2-graphs by G (i.e., in G_2 the index 2 will be omitted).

It is well known and easy to see that if $a_1 < \ldots < a_k < n$ and no a_i divides any other a_j then $\max k = [\frac{n+1}{2}]$. Also if we assume that no a_i divides the product of all the other a_j s we can easily show that $\max k = \Pi(n)$. The same result holds if we assume that all the products $\prod_{i=1}^{k} a_i^{\alpha_i}$ are distinct (the α_i's being non-negative integers).

Let us now assume that our sequence has the property that no a_i divides the product of two other a_j's. I proved [3] that in this case

$$(1) \qquad \Pi(x) + c_1\, x^{2/3}/(\log x)^2 < \max k < \Pi(x) + c_2\, x^{2/3}/(\log x)^2.$$

We outline the proof of the upper bound of (1). A simple lemma states that every integer $m \le x$ can be written in the form $u \cdot v$ where u is either a prime or is less than $x^{2/3}$ and v is less than $x^{2/3}$. Corresponding to the sequence $a_1 < \ldots < a_k$ we form a graph as follows: The vertices of our graph are the integers $< x^{2/3}$ and the primes p, $x^{2/3} < p \le x$. Put $a_i = u_i v_i$ by our lemma and let a_i correspond to the edge joining the vertices u_i and v_i. Our graph contains no path of length three (since no a_i divides the product of two other a_j's); thus our graph is a tree and thus has fewer edges than vertices or $k < \Pi(x) + x^{2/3}$. The inequality

$k < \Pi(x) + c_2\, x^{2/3}/(\log x)^2$ can be obtained by an improvement of the lemma (not all the integers $< x^{2/3}$ are needed in the representation $m = uv$).

The lower bound in (1) uses Steiner triples. It would be interesting to sharpen (1) and prove that for a certain absolute constant c

$$(2) \qquad \max k = \Pi(x) + c\, x^{2/3}/(\log x)^2 + o\left(\frac{x^{2/3}}{(\log x)^2}\right)$$

I have not been able to prove (2).

A generalization of the method which we used in the proof of (1) leads to the following more general result: Let $a_1 < \ldots < a_k \le x$ be a sequence of integers where no a_i divides the product of r other a_j's. Then

$$(3) \quad \Pi(x) + c_1^{(r)}\, x^{2/(r+1)}/(\log x)^2 < \max k < \Pi(x) + c_2^{(r)}\, x^{2/(r+1)}/(\log x)^2.$$

Assume now that our sequence $a_1 < \ldots < a_k \le x$ is such that the products $a_i a_j$ are all distinct. Then [3] [4]

$$(4) \quad \Pi(x) + c_4\, x^{3/4}/(\log x)^{3/2} < \max k < \Pi(x) + c_3\, x^{3/4}/(\log x)^{3/2}.$$

The proof of (4) again uses the lemma used in the proof of (1) and the graph theoretic representation of the sequence $a_1 < \ldots < a_k$. The fact that the products $a_i a_j$ are all distinct implies that the graph corresponding to the sequence $a_1 < \ldots < a_k$ contains no 4-cycle. The upper bound in (4) follows from the fact that every $G(n; [c_5\, n^{3/2}])$ contains a rectangle. The lower bound is due to Miss E. Klein and myself and is easy to obtain using finite geometries [3].

Here I would like to mention a problem in graph theory which is not yet completely solved. Denote by $f(n)$ the smallest integer for which every $G(n; f(n))$ contains a 4-cycle. W. Brown and V.T. Sós, Rényi and I ([2], [5]) proved that

$$(5) \qquad f(n) = (1/2 + o(1))n^{3/2}.$$

We are unable to give an exact formula for $f(n)$ and are far from being able to determine the structure of the extremal graphs, i.e. we do not know the structure of the graphs $G(n; f(n) - 1)$ which do not contain a rectangle.

Let $a_1 < \ldots < a_k \le x$ and assume that the product of any r a_i's are different (or that the product of any r or fewer a_i's are different). I am not able to give a very satisfactory estimation for $\max k - \Pi(x)$ if $r > 2$. Perhaps the answer depends essentially on the fact whether we only require that the product of r or fewer distinct a_i's are all different or whether we permit repetitions. Here we only state one result: Let $a_1 < \dots < a_k \le x$ be such that all products $\prod_{i=1}^{k} a_i^{\varepsilon_i}$, $\varepsilon_i = 0$ or 1, are distinct. Then [6]

$$(6) \qquad \max k < \Pi(x) + c_6\, x^{1/2}/\log x.$$

The proof of (6) is not graph theoretical and will not be discussed here. Perhaps (6) can be improved to

$$(7) \quad \max k < \Pi(x) + \Pi(x^{1/2}) + o\left(\frac{x^{1/2}}{\log x}\right) = \Pi(x) + \frac{(2 + o(1))x^{1/2}}{\log x}.$$

The inequality (7), if true, is best possible. To see this, let the a_i's be the primes and their squares.

An old and difficult conjecture of Turán and myself can be stated as follows: Let $a_1 < \dots$ be an infinite sequence of integers and denote by $f(n)$ the number of solutions of $n = a_i + a_j$. Then $f(n) > 0$ for $n > n_o$ implies $\lim\limits_{m} \sup f(n) = \infty$. A more general conjecture which is perhaps more amenable to attack goes as follows: Let $a_k < c\,k^2$, then $\lim\limits_{n} \sup f(n) = \infty$. I could only prove that $a_k < c\,k^2$ implies that the sums $a_i + a_j$ cannot all be different [14]. We come to very interesting problems if we restrict ourselves to finite sequences. Let $A(n, r)$ be the largest integer so that there is a sequence $a_1 < \dots < a_k \le n$, $k = A(n, r)$ for which all sums of r or fewer a_i's are distinct. It is known that [7]

$$(8) \qquad (1 + o(1))n^{1/2} < A(n, 2) < n^{1/2} + n^{1/4} + 1.$$

I conjecture that $A(n, 2) = n^{1/2} + O(1)$. Bose and Chowla proved [1]

$$A(n, r) \ge (1 + o(1))n^{1/r}$$

and they conjectured $A(n, r) = (1 + o(1))n^{1/r}$.

Let $a_1 < \dots < a_k \le n$ be a sequence of integers so that all the sums $\sum_{i=1}^{k} \varepsilon_i a_i$, $\varepsilon_i = 0$ or 1, are distinct. An old conjecture of mine states that

$$\max k = \frac{\log n}{\log 2} + O(1).$$

Moser and I [8] proved

$$\max k \le \frac{\log x}{\log 2} + \frac{\log\log x}{2 \log 2} + O(1).$$

Conway and Guy proved (unpublished) that if $n = 2^r$ is sufficiently large then $\max 2^r \ge r + 2$.

These problems perhaps have nothing to do with graph theory, but often their multiplicative analogue can be settled by graph theoretic methods. In fact I proved the following theorem [9]. Let $a_1 < \ldots$ be an infinite sequence of integers. Denote by $g(n)$ the number of solutions of $n = a_i a_j$. Then if for $n > n_o$, $g(n) > 0$ we have $\lim\limits_{n} \sup g(n) = \infty$, and in fact $g(n) > (\log n)^{c_7}$ for infinitely many n. This latest result cannot be improved very much since it fails to hold if c_7 is replaced by a sufficiently large constant c_8.

Denote by $u_p(n)$ the smallest integer so that if $a_1 < \ldots < a_k \le n$, $k = u_p(n)$, is any sequence of integers then for some $m, g(m) \ge p$. We have for $2^{r-1} < p \le 2^r$ [9],

$$u_p(n) = (1 + o(1))\, n(\log\log n)^{r-1}/(r-1)!\, \log n \tag{9}$$

$$= (1 + o(1))\, \Pi_r(n) ,$$

where $\Pi_r(n)$ denotes the number of integers not exceeding n having r distinct prime factors.

For $p > 2$ I cannot at present get a result which is as sharp as (4). I just want to state without proof a special result in this direction, namely

$$\frac{n\log\log n}{\log n} + c_9\, n/(\log n)^2 < u_3(n) < \frac{n\log\log n}{\log n} + c_{10}\, n/(\log n)^2. \tag{10}$$

It is not clear whether (10) can be sharpened.

The basic lemma needed for the proof of all these theorems is the following result on r-graphs: To every k and r there is an $\varepsilon_{k,r}$ so that every $G_r(n; c_{11}\, n^{r-\varepsilon_{k,r}})$ contains a $K_r(k, \ldots, k)$. For $r = 2$, $k = 2$, (5) shows that $\varepsilon_{k,r} = 1/2$. A result of Kövári, V.T. Sós and Turán [13] shows that $\varepsilon_{k,r} \ge 1/k$. In fact probably $\varepsilon_{k,2} = 1/k$ is the best value for $\varepsilon_{k,2}$. For $k = 3$ this is a result of W. Brown [2], but the cases $k > 3$ are still open. For $r > 3$ the best values of $\varepsilon_{k,r}$ are not known.

These extremal problems for r-graphs are usually much simpler for $r = 2$ (i.e. for the ordinary graphs). To illustrate this difficulty denote by $f(n, r, s)$ the smallest integer for which every $G_r(n; f(n, r, s))$ contains a $K_r(s)$. Turán determined $f(n, 2, s)$ for every n and s (e.g. $f(n, 2, 3) = [\frac{n^2}{4}] + 1$) and he posed the problem for $r > 2$ but as far as I know there are only inequalities and conjectures for $r > 2$. Turán conjectured that $f(2n, 3, 5) = n^2(n - 1) + 1$. It is easy to show that

$$\lim_{n\to\infty} f(n, r, s)/n^r = \delta_{r,s}$$

always exists and Turán proved $\delta_{2,s} = 1/2 - 1/2s$, but the value of

$\delta_{r,s}$ is unknown for every $s > r > 2$.

I would like to state one further conjecture for r-graphs: Every $G_3(3n; n^3 + 1)$ contains either a $G_3(4;3)$ or a $G_3(5;7)$.

Now I state a problem in number theory which can be reduced to a combinatorial problem:

Denote by $f(r, n)$ the smallest integer so that if $a_1 < \dots < a_k \le n$, $k = f(r, n)$ then there are r a_j's which pairwise have the same greatest common divisor. Using a combinatorial result of Rado and myself [11], I proved [12] that for every fixed r

$$e^{c_r \log n/\log\log n} < f(r, n) < n^{3/4+\varepsilon} . \tag{11}$$

It seems that the lower bound in (10) gives the correct order of magnitude. This would follow (11) from the following conjecture of Rado and myself: There is a constant α_r so that if $A_1, \dots A_s$, $s > \alpha_r^k$, are sets all having k elements, then there are always r of them, $A_{i_1}, \dots, A_{i_r}$ which pairwise have the same intersection.

Finally, I would like to mention a few problems in combinatorial number theory: Let $a_1 < \dots$ be an infinite sequence of integers, and assume that if

$$\prod_{r=1}^{q_1} a_{i_r} = \prod_{r=1}^{q_2} a_{j_r} , \quad \text{then } q_1 = q_2. \tag{12}$$

Is it true that for ε, there exists such a sequence of density $> 1 - \varepsilon$? Trivially, the a_i's can have density $1/4$. To see this, let the a_i's be the integers $\equiv 2 \pmod 4$. Selfridge showed that to every ε there is a sequence of density $> 1/e - \varepsilon$ satisfying (12). To see this let A be large and $A < p_1 < \dots < p_k$ the sequence of consecutive primes satisfying

$$\sum_{i=1}^{k} 1/p_i < 1 < \sum_{i=1}^{k+1} 1/p_i .$$

The a_i's are the integers divisible by precisely one of the p_i's, $1 \le i \le k$. It is easy to see that for sufficiently large A, the a_i's have the required properties.

We come to non-trivial questions if we restrict ourselves to finite sequences. Let $a_1 < \dots < a_k \le n$ be a sequence of integers satisfying (12). How large can $\max k$ be? Is it true that $\max k = n + o(n)$? I have no good upper or lower bounds for k. Trivially, $\max k > n(\log 2 - o(1))$. To see this, consider the integers not exceeding n having a prime factor $> \sqrt{n}$. I can slightly improve the constant $\log 2$ but cannot prove $\max k = n + o(n)$.

Let $a_1 < \dots < a_k \le n$; $b_1 < \dots < b_q \le n$ be two sequences of integers and assume that the products $a_i b_j$ are all distinct. Is it true that $kq < c\, n^2/\log n$?

Finally many of these problems can be modified as follows: Let $a_1 < \dots < a_k$ be a sequence of real numbers. Assume that any two of the numbers $\Pi a_i^{\alpha_i}$ differ by at least one. Is it true that $\max k = \Pi(n)$?

REFERENCES

1. R.C. Bose and S. Chowla, Theorems in the additive theory of numbers, Comm. Math. Helv. 37 (1962-63), 141-147.
2. W.G. Brown, On graphs that do not contain a Thomsen graph, Canad. Math. Bull. 9 (1966), 281-285.
3. P. Erdös, On sequences of integers no one of which divides the product of two others, Izr. Inst. Math. and Mech. Univ. Tomsk 2 (1938), 74-82.
4. P. Erdös, On some applications of graph theory to number theoretic problems, Publ. Ramanujan Inst. (to appear).
5. P. Erdös, A. Rényi, and V.T. Sós, On a problem of graph theory, Studia Sci. Math. Hung. 1 (1966), 215-235.
6. P. Erdös, Extremal problems in number theory II, Mat. Lapok. 17 (1966), 135-155.
7. P. Erdös and P. Turán, On a problem of Lidon in additive number theory and on related problems, J. London Math. Soc. 16 (1941), 212-216.
8. P. Erdös, Problems and results in additive number theory, Coll. Théorie des Nombres, Brussels (1955), pp. 127-137.
9. P. Erdös, On the multiplicative representation of integers, Israel J. Math. 2 (1964), 251-261
10. P. Erdös, On extremal problems of graphs and generalized graphs, Israel J. Math. 2 (1964), 183-190.
11. P. Erdös and R. Rado, Intersection theorems for systems of sets, J. London Math. Soc. 35 (1960), 85-90.
12. P. Erdös, On a problem in elementary number theory and a combinatorial problem, Math. of Computation 18 (1964), 644-646.
13. T. Kővári, V.T. Sós, and P. Turán, On a problem of K. Zarankiewicz, Colloq. Math. 3 (1955), 50-57.
14. A. Stöhr, Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe I. 194 (1955), 40-65; II 194 (1955), 111-140.