Extremal Problems in Graph Theory*

Paul Erdős

Received 1964

Without any doubt, Paul Erdős is one of the most famous mathematicians in the world. His areas of interest include number theory, set theory, probability, analysis, and graph theory. He has probably written more papers for mathematical journals than any other living mathematician. He is very well known for his fondness of travel and has surely lectured in more universities than anyone else. Erdős usually writes a joint paper with one or more of the mathematicians at each university he visits. So it was not entirely unnatural that a fantastic (false) rumor was spread about him to the effect that he even wrote a joint paper with a railroad conductor while traveling from one university to another. After his home base at the Mathematical Institute of the Hungarian Academy of Science in Budapest, his favorite locations are the Technion in Haifa, Israel, and the University College, London.

Everyone who has met him knows the classical Erdős terminology. An epsilon is a child. A married couple consists of one boss and one slave, the wife and husband respectively. When a couple is married, the boss is said to have captured the slave. After a divorce, the slave is liberated. If the man should remarry, he is said to be recaptured, and so on.

As indicated in the lecture itself, the subject of extremal problems in graph theory was initiated by another Hungarian mathematician, Paul Turán. Presently, it is the most popular area of graph theory in Hungary with papers on the subject written by Erdős jointly with each of the mathematicians, Andrásfai, Bollobás, Gallai, Hajnal, and Pósa (who was an epsilon of only thirteen when his prodigious talent was discovered by the lecturer).

F. H.

The starting point for many extremal problems in graph theory is the work of Turán [13] and [14] who initiated this topic in 1940 while in a labor camp. He showed that every graph with n points and $1 + [n^2/4]$ lines contains a triangle. Throughout this lecture, $G(n;m)$ will denote an arbitrary graph with n points and m lines. Also, $K(m,n)$ will denote the complete bicolored graph with m

* Another article with the same title appeared in *Theory of graphs and its applications*, edited by M. Fiedler. Prague, 1964, pp. 29–36. All results given without references are unpublished.

points of one color and n of the other. Similarly, $K(n_1, n_2, \ldots, n_k)$ will denote the complete k-colored graph with n_i points of the ith color.

Turán's result asserts that every graph $G(n; 1 + [n^2/4])$ contains the complete graph K_3. This is the best possible result because there exists a graph, namely $K([(n+1)/2], [nl2])$, with $[n^2/4]$ lines but no triangles. These facts are illustrated in Fig. 8.1 where $n = 5$; graph G_1 has seven lines and a triangle, whereas G_2 has six lines and no triangle. More generally, for each integer $p \le n$, Turán determined the least integer $m(n,p)$ such that any graph $G(n; m(n,p))$ contains the complete graph K_p. He also showed that the only graph $G(n; m(n,p) - 1)$ which does not contain K_p is $K(n_l, n_2, \ldots, n_{p-1})$, where the n_i are as nearly equal as possible. Dirac [5] showed that in addition to containing K_p, any graph $G(n; m(n,p))$ contains $K_{p+1} - x$, the graph obtained from K_{p+1} by deleting one line.

The graphs $G(n; 1 + [n^2/4])$ have other interesting properties. When $n \ge 4$, each such graph contains the first graph shown in Fig. 8.2, and when n is sufficiently large, it contains all of the graphs with one cycle as shown in Fig. 8.1. Furthermore, it contains any of the four possible graphs $G(5;6)$ with five points and six lines, all shown in Fig. 8.3. In addition, for n large enough, any graph $G(n; I + [n^2/4])$ contains both of the graphs with five points and seven lines shown in Fig. 8.4. These two graphs are special cases of the following result. For any integer r, there is an integer $n_0(r)$, such that when $n > n_0(r)$, each graph $G(n; 1 + [n^2/4])$ contains a subgraph of the form

$$K([(r+1)/2], [r/2]) + x,$$

obtained by adding a line to the complete bipartite graph. Similarly, for n large enough, every graph $G(n; 1 + [n^2/4])$ contains the cycle C_{2k+1}.

A variation of this problem is the determination of the value of $f(n,k,r)$, the smallest integer m such that every graph $G(n;m)$ contains some graph with k points and r lines. As indicated above, $f(n,3,3)$, $f(n,4,4)$, $f(n,4,5)$, $f(n,5,5)$, and $f(n,5,6)$ all have the same value, $1 + [n^2/4]$, for n sufficiently large. In general, not much is known about $f(n,k,r)$, but we do have the following results.

If $r \le k/2$, then $f(n,k,r) = r$.

If $k/2 < r < k$, then $f(n,k,r) = f(n, 2r + 2 - k, 2r + 1 - k)$.

Lastly, $f(n,k,k-1) = 1 + \left[\frac{(k-2)n}{k-1}\right]$.

Questions remain concerning $f(n,k,k)$. There are only two graphs $G(4;4)$, the first graph of Fig. 8.2 and C_4. The extremal graphs for the former have been found, but for the latter, the problem remains unsolved. In 1938 Erdős proved the existence of a constant c such that every graph $G(n; [cn^{3/2}])$ contains a cycle of length 4 for n large enough. Reiman [12] showed that for arbitrarily small ϵ and sufficiently large n,

$$(1-\epsilon)\frac{n^{3/2}}{2\sqrt{2}} < f(n,4,4) < (1+\epsilon)\frac{n^{3/2}}{2},$$

and it has just been shown by Brown, Rényi, Sós, and Erdős that

$$\lim_{n\to\infty} \frac{f(n,4,4)}{n^{3/2}} = \frac{1}{2}.$$

Kővári, Sós, and Turán [10] have shown that for some constant c, every graph $G(n;[cn^{2-1/k}])$ contains $K(p,p)$, which is a special case of a problem of Zarankiewicz [15]. Except when $p = 2$, however, it is not even known if this order of magnitude is best possible. Another open problem is the determination of how many lines a graph must have to ensure that the graph $K_l + K(p,p)$ is a subgraph. Again, Dirac and Erdős have settled this independently when $p = 2$. When $p = 3$, it is conjectured that the extremal graph is the join of two cycles $C_{[(n+1)/2]} + C_{[n/2]}$.

In 1941 Rademacher (see Erdős [6]) showed that every graph $G(2n;n^2+1)$ contains n triangles. Generalizing this, Erdős [7] proved the existence of a constant c such that when $k < cn$, every graph $G(2n;n^2+k)$ contains kn triangles. This is false when $k = n$ since the graph $\bar{K}_{n-1} + C_{n+1}$ contains n^2+n lines but only n^2-1 triangles. In proving this, several interesting lemmas were required. For example, there exists a constant $c \le 1/3$, such that every graph $G(2n;n^2+1)$ contains a line belonging to at least $[cn]$ triangles. It has also been shown that every graph $G(3n;3n^2+1)$ contains n^2 cycles of length 4 and that the result is best possible.

Ore [11] proved that every graph $G(n;2+n(n+1)/2)$ contains a hamiltonian cycle, clearly another best possible result. Dirac [3] has shown that any graph in which all points have degree at least $n/2$ is also hamiltonian. We have proved that every graph $G(n;r)$ is hamiltonian if every point has degree at least k and $r > \binom{n-t}{2} + t^2$ for every t satisfying $k \le t < n/2$. This result is also best possible. The proof uses the following theorem of Pósa which generalizes Dirac's result: A graph with n points is hamiltonian if whenever $k < n/2$, there are at most $k-1$ points with degree less than $k+1$.

It has been conjectured that any graph with rm points, all of degree at least $m(r-1)$, contains mK_r. Its validity when $r = 2$ follows from Dirac's result [3] on hamiltonian cycles; Corrádi and Hajnal [2] proved it when $r = 3$. Erdős and Gallai [8] have shown that every graph $G(n;r)$ contains m independent lines if

$$r > \max\left\{\binom{2m-1}{2}, n(m-1) - (m-1)^2 + \binom{m-2}{2}\right\}.$$

In this problem the extremal/graphs are K_{2m-1} and $K_{m-1} + \bar{K}_{n-m+1}$.

It was shown by Dirac [4] (see also Erdős and Pósa [9]) that when $n \ge 4$, every graph $G(n;2n-2)$ contains a subgraph homeomorphic to K_4, still another best possible result. It has also been conjectured that when $n \ge 5$, every graph $G(n;3n-5)$ contains a subgraph homeomorphic to K_5.

Bollobás and Erdős [1] proved that every graph $G(n;[(3n-1)/2])$ contains a cycle and another point adjacent to two points of the cycle, and thus every such graph contains two points which are joined by three linedisjoint paths. These results are best possible. It was conjectured that every graph $G\left(1+n(m-1); 1+n\binom{m}{2}\right)$ contains two points which are joined by m disjoint paths. The graph K_1+nK_m shows that, if true, this is best possible. Bollobás proved this for $m=4$; in fact, he showed that every graph $G(n;2n-1)$ contains two points which are joined by four line-disjoint paths and this is best possible. Perhaps every graph $G(n;2n-2)$ contains a cycle and another point adjacent to three points of the cycle. If true, this would strengthen Dirac's result mentioned above because in the subgraph homeomorphic to K_4, the three paths incident with one point would consist of a single line. It is obvious that every graph $G(n;n)$ contains a cycle. A complicated result due to Erdős and Pósa [9] is that every graph $G(n;(2m-1)n-2m^2+m+1)$ contains m disjoint cycles if $n>24m$. The result still holds when a line is removed if doing so does not yield the graph $K_{2m-1}+\bar{K}_{n-2m+1}$. The condition that $n>24m$ can be weakened somewhat. The contribution made by Pósa, who was thirteen years old at the time, was the following ingenious argument, which can be generalized, to show that every graph $G(n;3n-5)$ contains two disjoint cycles if $n\geq 6$.

That this result holds when $n=6$ may be verified by considering the various possibilities. Assume that it is valid when $n\leq t-1$ and consider any graph $G(t;3t-5)$. There is some point, say w_0, in which the degree $d(w_0)\leq 5$. Suppose that w_0 is adjacent only to the five points w_1, w_2, w_3, w_4, w_5. If the subgraph induced by these six points contains at least 13 lines, the result holds, since it has already been established when $n=6$. If not, then there is a point w_1 which is not adjacent to at least two of the other points, say w_2 and w_3. Add the new lines w_1w_2 and w_1w_3, and remove w_0 and the five lines incident with it from the original graph. There remains a graph $G(t-1;3t-8)$ which contains two disjoint cycles by the induction hypothesis. At least one of these two cycles does not contain either of the lines w_1w_2 or w_1w_3. In any case, it is easily seen, by including w_0 in a cycle if necessary, that the original graph must contain two disjoint cycles if $d(w_0)=5$.

Next, consider the case where w_0 is adjacent only to the four points w_1, w_2, w_3, w_4. It may be assumed that the subgraph generated by $W=\{w_0,w_1,w_2,w_3,w_4\}$ is K_5 since otherwise an argument similar to that used above yields the required result. Let $H=G(t;3t-5)-W$, the graph obtained from the original graph $G(t;3t-5)$ by removing the points of W. If any point in H is adjacent to two or more points in W, the result is clearly true. So it may be supposed that no point in H is adjacent to more than one point in W. Remove w_0, w_1, and w_2 from the original graph. The remaining graph has $n-3$ points and at least $(3n-5)-(n-5)-9=2n-9$ lines. Since $2n-9\geq n-3$ if $n\geq 6$, this remaining graph has at least one cycle. This cycle and $w_0w_lw_2w_0$ form two disjoint cycles in the original graph.

When the degree $d(w_0)\leq 3$, the result is shown by applying the induction hypothesis to the graph $G(t;3t-5)-w_0$. This suffices to complete the proof by induction.

We close with a proof of the following result by Pósa. Every graph $G(n;n+4)$ contains two line-disjoint cycles. The result is proved for multigraphs. This clearly holds when $n = 1$. Assume that it holds when $n = t - 1$ and consider any graph $G(t; t+4)$. If the graph contains a cycle C of length 3 or 4 the result certainly holds since removing the lines of C yields a graph with a cycle line-disjoint from C. Hence it may be assumed that every cycle has length greater than 4. If there is a point of degree 1, the result follows by removing this point and applying the induction hypothesis to the remaining graph. If some point v_0 is adjacent to only two points v_l and v_2, then the induction hypothesis can be applied to the graph obtained by adding the new line v_1v_2 and removing v_0. If one of the two line-disjoint cycles in this graph contains the line v_1v_2, then this line is replaced by the lines v_1v_0 and v_0v_2 in the original graph. Hence, it may be assumed that the degree of every point in the graph $G(t; t+4)$ is at least 3, which implies that $3t/2 \leq t + 4$, that is, $t < 8$. But it is not difficult to see that there are no multigraphs with fewer than nine points in which every point has degree at least 3 and every cycle has length greater than 4. Therefore, the result holds when $n = t$ and hence in general by induction.

References

[1] B. Bollobás and P. Erdős, On extremal problems in graph theory (in Hungarian). *Mat. Lapok.* **13** (1962) 143–152.

[2] K. Corrádi and A. Hajnal, On the maximal number of independent circuits in a graph. *Acta Math. Acad. Sci. Hung.* **14** (1963) 423–439.

[3] G. A. Dirac, Some theorems on abstract graphs. *Proc. London Math. Soc.* (3) **2** (1952) 69–81.

[4] G. A. Dirac, In abstrakten Graphen vorhandene vollständige 4-Graphen and ihre Unterteilungen. *Math. Nachr.* **22** (1960) 61–85.

[5] G. A. Dirac, Extensions of Turán's theorem on graphs. *Acta Math. Acad. Sci. Hung.* **14** (1963) 417–422.

[6] P. Erdős, Some theorems on graphs. *Riveon Lematematika* **9** (1955) 13–17.

[7] P. Erdős, On a theorem of Rademacher-Turán. *Ill. J. Math.* **6** (1962) 122–127.

[8] P. Erdős and T. Gallai, On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hung.* **10** (1959) 337–356.

[9] P. Erdős and L. Pósa, On the maximal number of disjoint circuits of a graph. *Publ. Math. Debrecen* **9** (1962) 3–12.

[10] T. Kővári, V. T. Sós, and P. Turán, On a problem of K. Zarankiewicz. *Colloq. Math.* **3** (1954) 50–57.

[11] O. Ore, Arc coverings of graphs. *Ann. Mat. Pura Appl.* **55** (1961) 315–322.

[12] I. Reiman, Über ein Problem von K. Zarankiewicz. *Acta Math. Acad. Sci. Hung.* **9** (1958) 269–279.

[13] P. Turán, Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok.* **48** (1941) 436–452.

[14] P. Turán, On the theory of graphs. *Colloq. Math.* **3** (1954) 19–30.

[15] K. Zarankiewicz, Problem 101. Colloq. Math. **2** (1951) 301.