A LIMIT THEOREM IN GRAPH THEORY

by

P. ERDŐS and M. SIMONOVITS[1]

In this paper $G(n; l)$ will denote a graph of n vertices and l edges, K_p will denote the complete graph of p vertices $G\left(p; \binom{p}{2}\right)$ and $K_r(p_1, \ldots, p_r)$ will denote the r-chromatic graph with p_i vertices of the i-th colour, in which every two vertices of different colour are adjacent. $\pi(G)$ will denote the number of vertices of G and $\nu(G)$ denotes the number of edges of G. $\bar{G}(n; l)$ denotes the complementary graph of $G(n; l)$ i. e. $\bar{G}(n; l)$ is the $G\left(n; \binom{n}{2} - l\right)$ which has the same vertices as $G(n; l)$ and in which two vertices are joined with an edge if and only if they aren't joined in $G(n; l)$. $\bar{K}(p_1, \ldots, p_r)$ thus denotes the union of the disjoint graphs K_{p_i} $(i = 1,2, \ldots, r)$.

In 1940 TURÁN [8] posed and solved the following question. Determine the smallest integer $m(n, p)$ so that every $G(n; m(n, p))$ contains a K_p. TURÁN in fact showed that the only $G(n; m(n, p) - 1)$ which contains no K_p is $K_{p-1}(m_1, \ldots, m_{p-1})$ where $\sum_{i=1}^{p-1} m_i = n$ and the m_i are all as nearly equal as possible.

A simple computation shows that

$$\lim_{n \to \infty} \frac{m(n, p)}{\binom{n}{2}} = 1 - \frac{1}{p-1} \qquad (p > 1). \tag{1}$$

Recently several more extremal problems in graph theory have been investigated and in this paper we continue some of these investigations [4]. First of all we prove the following general

THEOREM 1. *Let $G_1, \ldots, G_l$ be l given graphs and denote by $f(n; G_1, \ldots, G_l)$ the smallest integer so that every $G(n; f(n; G_1, \ldots, G_l))$ contains one of the graphs $G_1, \ldots, G_l$ as subgraphs. We have*

$$\lim_{n \to \infty} \frac{f(n; G_1, \ldots, G_l)}{\binom{n}{2}} = 1 - \frac{1}{r}$$

where $r \geqq 1$ is an integer which depends on the graphs G_i $(1 \leqq i \leqq l)$.

Theorem 1 easily follows from the following known result [3]:

[1] Eötvös L. University, Department of Mathematics Budapest.

For every $p>1$, $r>1$, $\varepsilon>0$ and $n>n_0(p, r, \varepsilon)$

$$\binom{n}{2}\left(1-\frac{1}{r-1}+o(1)\right)<f(n; K_r(p, \ldots, p))<\binom{n}{2}\left(1-\frac{1}{r-1}+\varepsilon\right). \tag{2}$$

Denote by $\varkappa(G)$ the chromatic number of G and put

$$\min_{1\le i\le l}\varkappa(G_i) = r+1. \tag{3}$$

Without loss of generality assume that $\varkappa(G_1)=r+1$.

Turán's graph $K_r(m_1, \ldots, m_r)$ where $\sum_{i=1}^{r} m_i=n$ and the m_i are as nearly equal as possible is clearly r-chromatic thus by (3) can not contain any of the G_i, $1\le i\le l$. A simple computation shows

$$v(K_r(m_1, \ldots, m_r)) = \binom{n}{2}\left(1-\frac{1}{r}+o(1)\right). \tag{4}$$

Put $\pi(G_1)=t$ and let $\varepsilon>0$ be arbitrary and let $n>n_0(t, r+1, \varepsilon)$. Then by (2) every $G\left(n; \binom{n}{2}\left(1-\frac{1}{r}+\varepsilon\right)\right)$ contains a $K_{r+1}(t, \ldots, t)$ which by $\pi(G_1)=t$ clearly contains G_1. This together with (4) completes the proof of Theorem 1.

An unpublished result of P. Erdős states that

$$f(n; K_r(p, \ldots, p)) = \binom{n}{2}\left(1-\frac{1}{r-1}\right)+O(n^{2-c}) \tag{5}$$

where c depends only on p and r. (5) easily implies that

$$f(n; G_1, \ldots, G_l)<\binom{n}{2}\left(1-\frac{1}{r}\right)+O(n^{2-c})$$

where $r+1=\min\varkappa(G_i)$ and c depends only on the graphs $G_1, \ldots, G_l$. Now we prove

THEOREM 1′. *Let k be an integer and $H_1, \ldots, H_m$ with $v(H_j)\le k$ given graphs. Denote by $h(n; H_1, \ldots, H_m; k)$ the smallest integer for which there is a graph $G(n; h(n; H_1, \ldots, H_m; k))$ every subgraph of which spanned by any k vertices of our graph $G(n; h(n; H_1, \ldots, H_m; k))$ contains one of the graphs $H_1, \ldots, H_m$. Then*

$$\frac{\lim_{n\to\infty} h(n; H_1, \ldots, H_m; k)}{\binom{n}{2}} = \frac{1}{t}$$

where $t\ge 1$ is an integer, or $t=\infty$.

Theorem 1′ could also be deduced easily from (2), but we show, that it follows from theorem *I*. In fact we shall show, that the two theorems are equivalent.

a) First we show that if there are given graphs $L_1, \ldots, L_\mu$ with $\pi(L_i)\le k$, then there exist graphs $M_1, \ldots, M_\nu$ so, that a graph G of k vertices contains at least one of $L_1, \ldots, L_\mu$ if and only if $\bar{G}$ contains none of $M_1, \ldots, M_\nu$.

From this of course follows that a graph G of k vertices contains none of $L_1, \dots, L_\mu$ if and only if $\bar{G}$ contains at least one of $M_1, \dots, M_\nu$ which shows the symmetricity between $L_1, \dots, L_\mu$ and $M_1, \dots, M_\nu$.

To prove our statement we define the graphs M_j: Let M_j be those graphs, for which $\pi(M_j)=k$ and $\overline{M}_j$ contains none of $L_1, \dots, L_\mu$. A very important property of the set of graphs $M_1, \dots, M_\nu$ is that if $H\supset M_j$ and $\pi(H)=k$ then H occurs among $M_1, \dots, M_\nu$ because $\overline{M}_j\subset\overline{H}$, further $\overline{M}_j$ contains none of $L_1, \dots, L_\mu$, and so $\overline{H}$ does not contain any of $L_1, \dots, L_\mu$.

Now, if $G\supset L_i$, then $\bar{G}$ does not occur among $M_1, \dots, M_\nu$ so $\bar{G}$ does not contain any of $M_1, \dots, M_\mu$. On the other hand, if G does not contain any of $L_1, \dots, L_\mu$, then $\bar{G}$ occurs among $M_1, \dots, M_\mu$, and this proves the second half of our statement.

If we have a graph F, which has $f(n, L_1, \dots, L_\mu)-1$ edges and does not contain any of the graphs $L_1, \dots, L_\mu$ then each subgraph spanned by its k vertices contains none of $L_1, \dots, L_\mu$, so each subgraph of $\bar{F}$ spanned by its k vertices contain at least one of those $M_1, \dots, M_\nu$ which we have defined in a), moreover $\bar{F}$ has the minimal number of edges among the graphs, each subgraph of which spanned by its k vertices contain at least one of $M_1, \dots, M_\nu$:

$$v(\bar{F}) = h(n; M_1, \dots, M_\nu) = \binom{n}{2} - f(n; L_1, \dots, L_\mu)+1.$$

So we can investigate a problem of the second type instead of a problem of the first type.

b) On the other hand, if there are given $M_1, \dots, M_\nu$, with $\pi(M_j)\leqq k$, we know, that there exist $L_1, \dots, L_\mu$ so, that a graph G of k vertices contains at least one of $M_1, \dots, M_\nu$ if and only if $\bar{G}$ contains none of $L_1, \dots, L_\mu$, or (what is equivalent with this) a graph G of k vertices contains none of $M_1, \dots, M_\nu$ if and only if $\bar{G}$ contains at least one of $L_1, \dots, L_\mu$. Now, if H is a graph, which has $h(n; M_1, \dots, M_\nu)$ edges, and each of its subgraph, spanned by its k vertices contains at least one of $M_1, \dots, M_\nu$, then each subgraph of $\overline{H}$ spanned by its k vertices contains none of $L_1, \dots, L_\mu$, moreover has the maximal number of edges among the graphs each subgraph of which spanned by its k vertices contain none of $L_1, \dots, L_\mu$:

$$v(\overline{H}) = f(n; L_1, \dots, L_\mu)-1 = \binom{n}{2} - h(n; M_1, \dots, M_\nu).$$

This proves in particular Theorem 1'.

Now we return to the study of our function $f(n; G_1, \dots, G_l)$. The proof of Theorem 1 shows that the order of magnitude of $f(n; G_1, \dots, G_l)$ depends only on $\min \varkappa(G_i)$. Nevertheless we show that the graphs G_i of higher chromatic number and in fact the structure of all the G_i, $(1\leqq i\leqq l)$ also have an influence on $f(n; G_1, \dots, G_l)$. To see this let G_1 be the graph consisting of a quadrilateral and a fifth vertex which is joined to all four vertices of the quadrilateral. It is known that [4] for $n>n_0$

$$f(n; G_1) = \left[\frac{n^2}{4}\right] + \left[\frac{n}{4}\right] + \left[\frac{n+1}{4}\right] + 1. \tag{6}$$

But on the other hand it is easy to show that for $n > n_0$

$$f(n; G_1, K_4) = \left[\frac{n^2}{4}\right] + \left[\frac{n+1}{4}\right] + 1. \tag{7}$$

Both (6) and (7) are easy to prove by induction and can be left to the reader.

Observe that $f(n; G_1) > f(n; G_1, K_4)$, G_1 is three-chromatic and K_4 is four-chromatic.

In every case we know the structure of the „extremal graphs" i. e. those

$$G(n; f(n; G_1, \ldots, G_l) - 1) \tag{8}$$

which do not contain any of the graphs G_i $(1 \leqq i \leqq l)$ these graphs are Turán graphs $K_{p-1}(m_1, \ldots, m_{p-1})$ for some p to which perhaps $o(n^2)$ further edges are added. Perhaps this is true in the general case, or at least the extremal graphs (8) contain a very large Turán graph (with say cn vertices). At present we are unable to attack this conjecture. The simplest case where we do not know anything about the structure of the extreme graph is the case of $K_3(2, 2, 2)$. It is known [4] that

$$\frac{n^2}{4} + c_2 n^{3/2} < f(n; K_3(2, 2, 2)) < \frac{n^2}{4} + c_1 n^{3/2}$$

but we don't know whether the extreme graphs contain a „large" Turán graph

Let $\binom{u}{2} \leqq l < \binom{u+1}{2}$, $u \geqq 2$. We now prove ([4])

THEOREM 2. *Let n be sufficiently large. Then*

$$f(n; G(k; l)) \leqq f\left(n; G\left(u; \binom{u}{2}\right)\right) = m(n, u). \tag{9}$$

Equality only if either $G(k; l)$ contains a K_u or if $u = 3$ and $G(k; l)$ is a pentagon.

First we prove the following

LEMMA 1. Let $l < \binom{u+1}{2}$. Then either $\varkappa(G(k; l)) < u$ or $G(k; l)$ has an edge e so that $\varkappa(G(k; l) - e) < u$. G—e is the graph from which the edge e has been omitted.*

We use induction with respect to u. It is easy to see that the Lemma holds for $u = 3$. Assume that it holds for $u - 1$ and we prove it for u. If G has a vertex x of valency $\geqq u$, let G^* be the graph which we obtain from G by omitting x and all edges incident to x. G^* has fewer than $\binom{u}{2}$ edges. Hence by the induction hypothesis there is an edge e so that $\varkappa(G^* - e) \leqq u - 2$, or $\varkappa(G - e) \leqq u - 1$.

We can therefore assume that all vertices of G have valency exactly $u - 1$, (since the vertices of valency $< u - 1$ could simply be omitted.) Since G has at most $\binom{u+1}{2} - 1$ edges, we obtain that it has at most $u + 2$ vertices and for these graphs our Lemma can be proved by simple inspection.

* Our original proof was more complicated. This simple proof we owe to V. T. Sós.

Now we can prove Theorem 2. 1) First assume $u>3$. It is known [5] that for every r and $n>n_0(r)$ every $G(n; m(n, u))$ contains a $K_u(r, \ldots, r)$ and an extra edge joining two vertices of the first r-tuple, and by our Lemma it is easy to see that for $r \geqq k$ our $G(k; l)$ is a subgraph of this graph.

2) If $u=3$, $m(G)<\binom{4}{2}=6$ and G contains no triangle then $\varkappa(G) \leqq 3$ and $\varkappa(G)=3$ if and only if G is a pentagon and it is known [4] that in this case

$$f(n; G)=\left[\frac{n^2}{4}\right]+1=m(n, 3). \tag{10}$$

3) Lastly, the case when $u=3$ and G contains a triangle was discussed in [4].

The equality in (9) if $u>3$ holds if and only if G contains a K_u. This can be obtained by a simple discussion, which we leave to the reader.

Finally we investigate $h(n; G; k)$ for some special graphs G. Let G be the graph which consists of l disjoint edges and assume $k>2l$. We outline the proof of the following

THEOREM 3. *Let $n>n_0(k, l)$. Then*

$$h(n; G_l; k)=\binom{n}{2}-m(n, k-2l+2)+1$$

and the only graph $G(n; h(n; G_l; k)$ for which every subgraph spanned by k of its vertices contains a G_l is $\overline{K}_{k-2l+1}$ $(m_1, \ldots, m_{k-2l+1})$ where $\sum_{i=1}^{k-2l+1} m_i=n$ and the m_i are all as nearly equal, as possible.

First of all it is easy to see that the subgraph spanned by any k vertices of $\overline{K}_{k-2l+1}(m_1, \ldots, m_{k-2l+1})$ contains a G_l. We leave the simple verification to the reader. This shows

$$h(n; G_l; k) \leqq \binom{n}{2}-m(n, k-2l+2)+1.$$

To complete the proof of Theorem 3 we now have to show the opposite inequality in other words we have to show that if $G\left(n; \binom{n}{2}-m(n, k-2l+2)\right)$ is any graph then there are k vertices $x_1, \ldots, x_k$ so that the subgraph of $G\left(n; \binom{n}{2}-m(n, k-2l+2)\right)$ spanned by these k vertices does not contain a G_l and further that the only $G\left(n; \binom{n}{2}-m(n, k-2l+2)+1\right)$ which does not have this property is $\overline{K}_{k-2l+1}(m_1, \ldots, m_{k-2l+1})$. These statements will follow immediately from the following

LEMMA 2. *There is a constant $c_r>0$, independent of n, so, that every $G(n; m(n, r+1))$, or a $G(n; m(n, r+1)-1)$, which is not a $K_r(m_1, \ldots, m_r)$ (where $\Sigma m_i=n$ and m_i are all as nearly equal as possible) contains a K_r and $c_r n$ other vertices, each of which is joined to every vertex of an K_r.*

REMARKS. TURÁN's theorem implies that every $G(n; m(n, r+1))$ contains a K_{r+1} and it is known [2], [4] that every such graph contains a K_{r+2} from which one edge is perhaps missing i. e. it contains a K_r and two vertices each of which is joined to every vertex of our K_r. Our Lemma is sharpening of this result.

We supress the proof of our Lemma since it is very similar to the case when $r=2$ which is known [6].

Let now n be sufficiently large and $G(n; e)$ be any graph, for which

$$e \leqq v(\overline{K}_{k-2l+1}(m_1, \ldots, m_{k-2l+1})) = \binom{n}{2} - m(n, k-2l+2)+1$$

and which is not a $\overline{K}_{k-2l+1}(m_1, \ldots, m_{k-2l+1})$ (where $\Sigma m_i = n$; and m_i are all nearly equal, as possible).

By our lemma, the complementary graph of our $G(n, e)$ contains a K_{k-2l+1} and $2l-1$ vertices, each of which is joined to every vertex of our K_{k-2l+1}, i. e. there are k vertices, which span in our $G(n; e)$ a subgraph, which consist of $k-2l+1$ isolated, vertices and a graph of $2l-1$ vertices and hence it can not contain l independent edges. This completes the proof of Theorem 3.

It is easy to see that if $k=2l$ the extreme graphs are no longer TURÁN's graphs, it is easy to see that in this case

$$h(n; G_l; 2l) = \binom{n}{2} - (l-1)n. \tag{11}$$

To see this observe that if one vertex of G_l is not joined to $2l-1$ vertices these $2l$ vertices can not contain independent edges. This proves

$$h(n; G_l; 2l) \geqq \binom{n}{2} - (l-1)n.$$

On the other hand, the following example shows, that

$$h(n; G_l; 2l) \leqq \binom{n}{2} - (l-1)n.$$

Let the vertices of G_l^* be the n-th roots of unity, two such vertices are joined if their distance — on the circle $|z|=1$ — is greater, then $(l-1)\frac{2\pi}{n}$.

In this case, if the vertices of our graph are $P_1, \ldots, P_n$ and $A_1, \ldots, A_k$ are k vertices of them enumerated, as they are on the circle, then A_i and $A_{i+l}, (i=1, 2, \ldots, l)$ will be connected, so there will be l independent edges in the subgraph, spanned by $A_1, \ldots, A_k$.

We do not investigate the question of the unicity of the extremal graphs.

Denote by $G_l^{(3)}$ the graph consisting of l independent triangles. We outline the proof of the following

THEOREM 4. *Let $n > n_0(l)$. Then*

$$h(n; G_l^{(3)}; 3l+2) = \binom{n}{2} - m(n, 3)+1 = \binom{\left[\frac{n}{2}\right]}{2} + \binom{\left[\frac{n+1}{2}\right]}{2}$$

and the only extreme graph is $\overline{K}_2\left(\left[\frac{n}{2}\right]; \left[\frac{n+1}{2}\right]\right)$.

On the other hand, if $l>1$

$$(12)\qquad \binom{n}{2}-C_l n^{3/2}<h(n;G_l^{(3)};3l+1)<\binom{n}{2}-n^{1+\varepsilon_l}\qquad (\varepsilon_l>0).$$

The structure of the extreme graph (or graphs) is unknown. It is easy to see that $h(n;G_l^{(3)};4)=\binom{n-1}{2}$ and K_{n-1} is the only extreme graph.

First of all it is easy to see that every subgraph spanned by $3l+2$ vertices of $\overline{K}_2\left(\left[\frac{n}{2}\right],\left[\frac{n+1}{2}\right]\right)$ contains a $G_l^{(3)}$.

Let G be any graph for which

$$\pi(G)=n,\quad \nu(G)\geqq\left[\frac{n^2}{4}\right]$$

and if $\nu(G)=\left[\frac{n^2}{4}\right]$ then G is not $K_2\left(\left[\frac{n}{2}\right],\left[\frac{n+1}{2}\right]\right)$. If $n>n_0(l)$ then it is known [5] that G contains a subgraph of $3l+2$ vertices $x_1,x_2,x_3;y_1,\dots,y_{3l-1}$ where all the edges

$$(x_1,x_2);\ (x_i,y_j)\quad 1\leqq i\leqq 3\quad 1\leqq j\leqq 3l-1$$

are in G. Clearly these $3l+2$ vertices span a subgraph of $\overline{G}$ which does not contain $G_l^{(3)}$. This completes the proof of the first half of Theorem 4.

To prove the second half we observe that it is known that every $G(n;[c_l n^{3/2}])$ contains a $K_2(2,3l-1)$ hence the subgraph of $\overline{G}(n;[c_l n^{3/2}])$ spanned by the vertices of our $K_2(2,3l-1)$ clearly contains no $G_l^{(3)}$ this proves the left side inequality of (12).

Now we outline the proof of the right hand side of (12). First of all it is known [7] that there exists a graph $G(n;n^{1+\varepsilon_l})$ which contains no circuit having $\leqq 3l+1$ edges. Thus our proof will be complete if we can show that if $l>1$ and $G(3l+1;p)$ is any graph of $3l+1$ vertices which contains no circuit then $\overline{G}(3l+1;p)$ contains l independent triangles. This can be shown easily by induction with respect to l and can be left to the reader. Thus the proof of Theorem 4 is complete.

(Received October 19, 1964.)

REFERENCES

[1] BROOKS, R. L.: On colouring the nodes of a network, *Proc. Cambridge Phil. Soc.* **37** (1941) 194—197.
[2] DIRAC, G.: Extensions of Turán's theorem on graphs, *Acta Hung. Acad. Sci.* **14** (1963) 417—422, see also [4].
[3] ERDŐS, P. and STONE, A. H.: On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946) 1087—1091.
[4] See e. g. ERDŐS, P.: Extremal problems in graph theory, *Proc. Symposium, Theory of Graphs and its applications* Smolenice (1963), 29—36.
[5] ERDŐS, P.: On the structure of linear graphs, *Israel J. of Math.* **1** (1963), 156—160. In this paper the theorem is proved in detail only for $r=2$.
[6] ERDŐS, P.: On the theorem of Rademacher—Turán, *Illinois J. of. Math.* **6** (1962) No. 1. March.
[7] ERDŐS, P.: Graph theory and probability, *Canad. J. Math.* **11** (1959) 34—38.
[8] TURÁN, P.: *Mat. és Fiz. Lapok* **48** (1941) 436—452 (in Hungarian), see also „On the theory of graphs" *Coll. Math.* **3** (1954) 19—30.