ON THE NUMBER OF COMPLETE SUBGRAPHS AND CIRCUITS CONTAINED IN GRAPHS

P. ERDÖS, Budapest

(Received January 10, 1968)

Dedicated to V. JARNÍK *on the occasion of his 70-th birthday.*

Denote by $\mathcal{G}(n; k)$ a graph of n vertices and k edges. Put for $n \equiv r \pmod{p-1}$

$$m(n, p) = \frac{p-2}{2(p-1)}(n^2 - r^2) + \binom{r}{2}, \quad 0 \leqq n \leqq p - 1$$

and denote by K_p the complete graph of p vertices. A well known theorem of TURÁN [6] states that every $\mathcal{G}(n; m(n, p) + 1)$ contains a K_p and that this result is best possible. Thus in particular every $\mathcal{G}(2n; n^2 + 1)$ contains a triangle. Denote by $f_n(p; l)$ the largest integer so that every $\mathcal{G}(n; m(n, p) + l)$ contains at least $f_n(p; l)$ distinct K_p's. RADEMACHER proved that $f_n(3; 1) = [n/2]$ and I proved [1] that there exists a constant $0 < c < \frac{1}{2}$ so that for every

$$(1) \qquad l < cn\,, \quad f_n(3; l) = l\left[\frac{n}{2}\right]$$

and I conjectured that (1) holds for every $l < [n/2]$. We are very far from being able to determine $f_n(p; l)$ in general, the problem is unsolved even for $p = 3$ (though W. BROWN has certain plausible unpublished conjectures). NORDHAUS and STEWART [4] conjectured that

$$\lim_{n=\infty} \min_{l} \frac{f_n(3; l)}{\frac{1}{2} l n} = \frac{8}{9}, \quad 0 < l \leqq \binom{n}{2} - \left[\frac{n^2}{4}\right]$$

I proved that for $l = o(n^2)$

$$(2) \qquad f_n(3; l) = (1 + o(1))\, l \frac{n}{2}\,.$$

I do not give the proof of (2) in this paper.

Theorem 1. *Let* $n > n_0(p)$. *Then*

$$f_n(p; 1) = \prod_{i=0}^{p-3} \left[\frac{n+i}{p-1}\right]. \tag{3}$$

The special case

$$f_{3n}(4; 1) = n^2$$

was stated without proof in [1]. It is possible that the condition $n > n_0(p)$ can be omitted and that (3) holds for every n.

Instead of Theorem 1 we prove the following more general

Theorem 2. *Let* $n > n_0(p)$ $(l_1 < \varepsilon_p n,\ \varepsilon_p > 0)$ *be a sufficiently small constant. Then*

$$f_n(p; l_1) = l_1 \prod_{i=0}^{p-3} \left[\frac{n+i}{p-3}\right].$$

In the case $p = 3$ the proof of Theorem 1 is much simpler than that of Theorem 2, [2], but for the general case I have no simpler proof for Theorem 1 than for Theorem 2.

Our principal tool for the proof of Theorems 1 and 2 will be

Theorem 3. *Let* $n > n_0(p)$, $l_2 < n/200p^4$. *Let there be given a* $\mathcal{G}(n; m(n, p) - l_2)$ *which contains a* K_p. *Then it has an edge which is contained in* $n^{p-2}/(10p)^{6p}$ K_p*'s of our graph.*

By Turáns theorem every $\mathcal{G}(n; m(n, p) + 1)$ contains a K_p. Thus Theorem 3 implies the following corollary of independent interest.

Theorem 3'. *Every* $\mathcal{G}(n; m(n, p) + 1)$ *has an edge which is contained in* $n^{p-2}/(10p)^{6p}$ K_p*'s of our graph.*

For $p = 3$ all our Theorems are known [1]. In fact I can show that every $\mathcal{G}(n; [n^2/4] + 1)$ has an edge which is contained in at least $(n/6) + O(1)$ triangles and that $n/6$ is best possible. For $p > 3$. I have not succeeded in determining the best possible constant in Theorem 3'. The constants in all our Theorems are very far from being best possible.

To prove Theorem 3 we need two Lemmas, but first we have to introduce some notations. $\mathcal{G}_m$ will denote a graph of m vertices. $\mathcal{G}(y_1, \ldots, y_l)$ will denote the subgraph of $\mathcal{G}$ spanned by the vertices $y_1, \ldots, y_l$. $\mathcal{G} - x_1 - \ldots - x_r$ denotes the subgraph of $\mathcal{G}$ from which the vertices $x_1, \ldots, x_r$ and all edges incident to them have been omitted. Let $e_1, \ldots, e_r$ be edges of $\mathcal{G}$. $\mathcal{G} - e_1 - \ldots - e_r$ denotes the subgraph of $\mathcal{G}$ from which the edges $e_1, \ldots, e_r$ have been omitted. $e(\mathcal{G})$ will denote the number of edges of $\mathcal{G}$, $v(x)$ the valency of the vertex x is the number of edges of $\mathcal{G}$ incident to x. $K(u_1, \ldots, u_p)$ denotes the complete p − chromatic graph, with u_i vertices of the i-th color and where any two vertices of different color are joined by an edge. If $\mathscr{S}$ is a set $|\mathscr{S}|$ denotes the number of its elements and if $A \subset \mathscr{S}$, $\bar{A}$ is the complement of A in $\mathscr{S}$.

We always assume $p \geqq 4$, since our Theorems are all known for $p = 3$.

Lemma 1. *Let* $|\mathscr{S}| = n$ *and* $A_i \subset \mathscr{S}$, $1 \leqq i \leqq p$. *Assume*

$$(4) \qquad |A_i| > n\left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right), \quad 1 \leqq i \leqq p\,.$$

Then there are values $1 \leqq i < j \leqq p$ *so that*

$$(5) \qquad |A_i \cap A_j| > n\left(\frac{p-3}{p-1} + \frac{1}{10p^3}\right).$$

(5) is not best possible, but suffices for our purpose. From (4) and $|\mathscr{S}| = n$ it follows that if (5) fails to hold for every $1 \leqq i < j \leqq p$, then

$$(6) \qquad |A_i| \leqq n\left(\frac{p-2}{p-1} + \frac{1}{10p^3} + \frac{1}{100p^4}\right).$$

From (6) we have

$$(7) \qquad |\bar{A}_i| \geqq n\left(\frac{1}{p-1} - \frac{1}{10p^3} - \frac{1}{100p^4}\right).$$

Further clearly

$$(8) \qquad |A_i \cap A_j| = |A_i| + |A_j| - n + |\bar{A}_i \cap \bar{A}_j|\,.$$

Thus if (5) never holds we have from (4) and (8) that for every $1 \leqq i < j \leqq p$

$$(9) \qquad |\bar{A}_i \cap \bar{A}_j| \leqq n\left(\frac{1}{50p^4} + \frac{1}{10p^3}\right).$$

It is easy to see that (7) and (9) lead to a contradiction. We evidently have

$$(10) \qquad n = |\mathscr{S}| \geqq \sum_{i=1}^{p} |\bar{A}_i| - \sum_{1 \leqq i < j \leqq p} |\bar{A}_i \cap \bar{A}_j|\,.$$

Thus from (7) and (10)

$$\max_{1 \leqq i < j \leqq p} |\bar{A}_i \cap \bar{A}_j| \geqq \frac{1}{\binom{p}{2}}\, n\left(\frac{1}{p-1} - \frac{1}{10p^2} - \frac{1}{100p^3}\right)$$

which contradicts (9) and hence proves the Lemma.

Lemma 2. *Let $\mathscr{G}(n; m(n, p) - l_2) = \mathscr{G}$, $l_2 < n/200p^4$ be a graph which contains a K_p. Then it has a subgraph $\mathscr{G}_N$, $N > n/100p^2$ which also contains a K_p and each vertex of which has (in $\mathscr{G}_N$) valency*

$$v(x) > N\left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right). \tag{11}$$

If our $\mathscr{G}$ satisfies (11) our Lemma is proved. If not let $x_1, \ldots$ be a sequence of vertices of our $\mathscr{G}$ so that the valency of x_i in $\mathscr{G} - x_1 - \ldots - x_{i-1}$ satisfies

$$v(x_i) \leqq (n - i)\left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right). \tag{12}$$

Suppose this process stops in k steps, in other words every vertex of $\mathscr{G} - x_1 - \ldots - x_k$ has valency greater than

$$(n - k)\left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right). \tag{13}$$

But then by (12) and by the fact that $e(\mathscr{G} - x_1 - \ldots - x_k) \leqq \binom{n-k}{2}$ a simple argument shows that

(14)

$$e(\mathscr{G}) = m(n, p) - l_2 = \frac{p-2}{p-1}\binom{n}{2} + O(n) < \left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right)\binom{n}{2} + \binom{n-k}{2}.$$

(14) clearly leads to a contradiction if $n > n_0(p)$ and $n - k \leqq n/100p^2$. Thus $n - k > n/100p^2$. Put $\mathscr{G}_N = \mathscr{G} - x_1 - \ldots - x_k$. By (13) $\mathscr{G}_N$ satisfies (11), it clearly satisfies $N > n/100p^2$. Finally by (12) and $k \geqq 1$ we obtain by a simple computation

$$e(\mathscr{G}_N) \geqq e(\mathscr{G}) - \sum_{i=0}^{k-1}(n - i)\left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right) > \tag{15}$$

$$> m(n, p) - \frac{n}{200p^4} - \sum_{i=0}^{k-1}(n - i)\left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right) > m(n - k, p) = m(N, p).$$

(15) implies by Turáns theorem that our $\mathscr{G}_N$ contains a K_p, which completes the proof of Lemma 2.

Now we are ready to prove Theorem 3. Our $\mathscr{G}(n; m(n, p) - l_2)$ contains by Lemma 2 a $\mathscr{G}_N$, $N > n/100p^2$ the valency of each vertex of which satisfies (11) and it contains a K_p say $(x_1, \ldots, x_p)$. Denote by A_i the set of vertices in $\mathscr{G}_N$ joined to x_i. By (11) we can apply Lemma 1 and obtain that there are two vertices x_i and x_j, $1 \leqq i < j \leqq p$ both of which are joined to ($y_1, \ldots, y_t$ are vertices of $\mathscr{G}_N$)

$$y_1, \ldots, y_t, \quad t > N\left(\frac{p-3}{p-1} + \frac{1}{10p^3}\right), \quad N > n/100p^2. \tag{16}$$

Consider now the graph $\mathscr{G}_N(y_1, \ldots, y_t)$. By (11) and (16) we have for every i

$$(17)\quad v(y_i) > N\left(\frac{p-2}{p-1} - \frac{1}{100p^4}\right) - N + t = t - N\left(\frac{1}{p-1} + \frac{1}{100p^4}\right) >$$

$$> t\left(1 - \frac{\dfrac{1}{p-1} + \dfrac{1}{100p^4}}{\dfrac{p-3}{p-1} + \dfrac{1}{10p^3}}\right) > t\left(\frac{p-4}{p-3} + \frac{1}{20p^3}\right).$$

In (17) $v(y_i)$ of course denotes valency in $\mathscr{G}_N(y_1, \ldots, y_t)$. Denote by B_i the set of y's joined to y_i. It immediately follows from (17) that for every $i_1, \ldots, i_r$, $r \leqq p - 3$

$$(18)\qquad |B_{i_1} \cap \ldots \cap B_{i_r}| > \frac{t}{20p^3},$$

(for $r < p - 3$ (17) could of course be considerably improved).

For (18) and (15) we immediately obtain that $\mathscr{G}_N(y_1, \ldots, y_t)$ contains at least $(t > (p-3)\,N/(p-1) > n/300p^2)$

$$(19)\qquad \frac{1}{(p-2)!}\,\frac{t^{p-2}}{(20p^3)^{p-2}} > \frac{1}{(p-2)!}\,\frac{n^{p-2}}{(10p)^{5(p-2)}} > \frac{n^{p-2}}{(10p)^{6p}}$$

K_{p-2}'s. (19) follows from the fact that by (18) we have for each r at least $t/20p^3$ choices for the r-th vertex of our K_{p-2}. Each of these K_{p-2}'s form together with the edge (x_i, x_j) a K_p of our $\mathscr{G}(n; m(n, p) - l_2)$ each of which contain the edge (x_i, x_j), and this completes the proof of Theorem 3.

Now we prove Theorem 2. The proof is very similar to [1]. We use the following theorem of Simonovits [5]:

To every p there is a δ_p so that if $l < \delta_p n$ and the graph $\mathscr{G}(n; m(n, p) - l)$ does not contain a K_p then it is $(p-1)$-chromatic, in other words it is a subgraph of some $K(u_1, \ldots, u_{p-1})$ with $\sum_{i=1}^{p-1} u_i = n$.

Now we are ready to prove Theorem 2. Consider Turáns graph

$$K(u_1, \ldots, u_{p-1}), \quad u_i = \left[\frac{n+i-1}{p-1}\right], \quad 1 \leqq i \leqq p-1,$$

having the vertices $x_j^{(i)}$, $1 \leqq j \leqq [(n+i-1)/(p-1)]$, $1 \leqq i \leqq p-1$. Add the l_1 edges $(x_1^{(p-1)}, x_j^{(p-1)})$, $2 \leqq j \leqq l_1 + 1$. This $\mathscr{G}(n; m(n, p) + l_1)$ clearly has $l_1 \prod_{i=0}^{p-3} [(n+i)/(p-1)]\, K_p$'s. Thus to prove Theorem 2 we only have to show

$$(20)\qquad f_n(p, l) \geqq l_1 \prod_{i=0}^{p-3} \left[\frac{n+i}{p-1}\right].$$

To prove (20) observe that by Turáns theorem our $\mathcal{G}(n; m(n, p) + l_1)$ contains a K_p, let r be the smallest integer so that $\mathcal{G} - e_1 - \ldots - e_r$ contains no K_p. By Turáns theorem we have $r \geqq l_1$. Assume first $r \geqq (10p)^{6p} l_1$. From Theorem 3 (and from the proof of Theorem 3) we obtain that if $\varepsilon_p < 1/2.10^8 p^{6p+2}$, $(l_1 < \varepsilon_p n)$ then each of the edges e_i, $1 \leqq i \leqq (10p)^{6p} . l_1$ are contained in at least $n^{p-2}/(10p)^{6p}$ K_p's of $\mathcal{G} - e_1 - \ldots - e_{i-1}$. These K_p's are clearly all different. Thus $\mathcal{G}$ contains at least

$$l_1 n^{p-2} > l_1 \prod_{i=0}^{p-3} [(n + i)/(p - 1)]$$

K_p's which proves (20) in this case.

Assume next $r < (10p)^{6p} l_1$. Let $\varepsilon_p < \delta_p/(10p)^{6p}$. We have by assumption $l_1 < \varepsilon_p n$. Then by the theorem of Simonovits $\mathcal{G} - e_1 - \ldots - e_r$ must be contained in a $K(u_1, \ldots, u_{p-1})$, $\sum_{i=1}^{p-1} u_i = n$. Now we assume $p \geqq 4$. We then easily obtain

$$(21) \qquad u_i = \left[\frac{n + i - 1}{p - 1}\right], \quad 1 \leqq i \leqq p - 1\,.$$

To see this observe that if $p \geqq 4$ and $\sum_{i=1}^{p-1} u_i = n$ and (21) is not satisfied for all i we would have by a simple computation for sufficiently small δ_p

$$m(n, p) - r < e(\mathcal{G} - e_1 - \ldots - e_r) \leqq \prod_{i=1}^{p-1} u_i < m(n, p) - \delta_p n$$

an evident contradiction since $r < \delta_p n$.

Observe now that (since δ_p is small) the edges e_i, $1 \leqq i \leqq r$ must join vertices of the same color of our $K(u_1, \ldots, u_n)$. By (21) we observe by a simple argument that each e_i, $1 \leqq i \leqq r$ is contained in at least $(r - l_1 = r_1)$

$$\left(\left[\frac{n}{p - 1}\right] - r_1\right) \prod_{i=1}^{p-3} \left[\frac{n + 1}{p - 1}\right]$$

K_p's and these K_p's are clearly, all different, or our graph contains at least

$$(22) \qquad r\left(\left[\frac{n}{p - 1}\right] - r_1\right) \prod_{i=1}^{p-3} \left[\frac{n + i}{p - 1}\right]$$

K_p's. From $r < \delta_p n$ it follows for sufficiently small δ_p that $r([n/(p - 1)] - r_1)$ is minimal if r_1 is as small as possible, in other words if $r = l_1$, $r_1 = 0$. Thus by (22) our $\mathcal{G}$ contains at least

$$l_1 \prod_{i=0}^{p-3} \left[\frac{n + i}{p - 1}\right]$$

K_p's, which completes the proof of (20) and Theorem 2.

With considerably greater care we could prove the following further results:

Theorem 4. *Let* $n > n_0(p)$

$$(23)\quad l = \sum_{i=0}^{j}\left(\left[\frac{n+i}{p-1}\right]-1\right)+t,\quad 0 \leqq t < \left[\frac{n+j+1}{p-1}\right],\quad -1 \leqq j \leqq p-3\,.$$

Then every $\mathscr{G}(n; m(n,p)+1-l)$ *which contains a* K_p *contains at least*

$$(24)\qquad \left(\left[\frac{n+j+1}{p-1}\right]-t\right)\prod_{j+1}^{p-3}\left[\frac{n+i}{p-1}\right] = g(n,p,l)$$

K_p*'s. Further every* $\mathscr{G}(n; m(n,p)+1-l)$ *satisfying* (23), *which contains a* K_p *has an edge which is contained in* $e_p\, g(n,p,l)\, K_p$*'s.*

The proof of Theorem 4 is quite complicated, it uses methods of [1] and will not be given here. It is quite easy to see though that (24) is best possible. It suffices to consider a Turán graph $K(u_1, \ldots, u_{p-1})$, $u_i = [(n+i-1)/(p-1)]$, $1 \leqq i \leqq$ $\leqq p-1$ having vertices $x_j^{(i)}$, $1 \leqq j \leqq [(n+i-1)/(p-1)]$, $1 \leqq i \leqq p-1$. Add the edge $(x_1^{(p-1)}, x_2^{(p-1)})$ and omit l suitable edges emunating from $x_1^{(p-1)}$. The details can be left to the reader.

By the methods of this paper we can prove the following

Theorem 5. *Every* $\mathscr{G}(2n; n^2+1)$ *contains at least* $n(n-1)(n-2)$ *pentagons.*

$K(n,n)$ with one edge added shows that Theorem 5 is best possible. Theorem 5 could be generalised for $(2r+1)$-gons but we will return to these questions at another occasion.

References

[1] *P. Erdös:* On a theorem of Rademacher-Turán, Illinois J. Math. *6* (1962), 122—127.

[2] *P. Erdös:* Some theorems on graphs, Riveon lematematika, *10* (1955), 13—16 (in Hebrew).

[3] *P. Erdös:* Some recent results on extremal problems in graph theory. Theory of graphs, International Symposium, Rome 1966, p. 117—130.

[4] *E. A. Nordhaus* and *B. M. Stewart:* Priangles in an ordinary graph, Canad. J. Math. *15* (1963), 33—41.

[5] *M. Simonovits:* A method for solving extremal problems in graph theory. Stability problems, Theory of Graphs, Proc. Colloquium held at Tihany, Hungary, Acad. Press and Akad. Kiadó 1968, 279—334.

[6] *P. Turán:* Eine Extremalaufgabe aus der Graphentheorie, Mat. és Fiz. Lapok, *48* (1941), 436—452 (written in Hungarian). See also P. Turán. On the theory of graphs, Coll. Math. *3* (1954), 19—30.

Author's address: Mathematical Institute, Hungarian Academy of Sciences, Budapest, Hungary.