ON THE NUMBER OF COMPLETE SUBGRAPHS CONTAINED IN CERTAIN GRAPHS

by

P. ERDŐS

$G^{(n)}$ will denote a graph of n vertices, G_l a graph of l edges and $G_l^{(n)}$ a graph of n vertices and l edges. Loops will not be permitted and two vertices can be connected by at most one edge. In the complete graph $G^{(n)}_{\binom{n}{2}}$ of n vertices, every two vertices are connected by an edge. A complete graph $G_3^{(3)}$ of three vertices is called a triangle. The complementary graph $\bar{G}_l^{(n)}$ of $G_l^{(n)}$ is defined as follows: The two graphs have the same vertices and two vertices are connected by an edge in $\bar{G}_l^{(n)}$ if and only if they are not connected by an edge in $G_l^{(n)}$. In other words a graph $G_l^{(n)}$ and its complementary $\bar{G}_l^{(n)}$ gives a splitting of the edges of the complete graph $G^{(n)}_{\binom{n}{2}}$ into two disjoint classes. $\bar{G}_l^{(n)}$ can be written as $G^{(n)}_{\binom{n}{2}-l}$, but of course this in general does not determine its structure uniquely since the number of vertices and edges does not determine the structure of the graph.

The vertices of G will be denoted by $x, x_1, \ldots, y_1, \ldots$. The graph $(G - x_1 - \ldots - x_r)$ will denote the graph from which the vertices $x_1, \ldots, x_r$ and all the edges incident to them have been omitted. $G(x_1, \ldots, x_k)$ will denote the subgraph of G spanned by the vertices $x_1, \ldots, x_k$. The valency $v(x)$ of x is the number of edges incident to it. $\nu(G)$ will denote the number of edges of G, and $\pi(G)$ the number of its vertices.

$C_k(G)$ will denote the number of complete subgraphs $G^{(k)}_{\binom{k}{2}}$ of G. Recently A. GOODMAN [1] proved that

$$\min\left(C_3(G^{(n)}) + C_3(\bar{G}^{(n)})\right) = \begin{cases} 2\binom{u}{3} & \text{if } n = 2u \\ \frac{2}{3}u(u-1)(4u+1) & \text{if } n = 4u+1 \\ \frac{2}{3}u(u+1)(4u-1) & \text{if } n = 4u+3 \end{cases} \tag{1}$$

where the minimum is to be taken over all graphs $G^{(n)}$ having n vertices.

A simpler proof of (1) was later given by A. SAUVÉ [2].

GOODMAN asked if the sign of equality in (1) can hold if $C_3(\bar{G}^{(n)}) = 0$, i.e. if $\bar{G}^{(n)}$ contains no triangle. His answer was affirmative for even n. For odd n I showed [2] that the answer is negative for $n > 7$ and it is easily seen to be affirmative for $n \leq 7$.

G. LORDEN [3] proved the following stronger result:
Assume that $C_3(\bar{G}^{(n)}) = 0$. Then for all even n and odd $n > 9$

$$\min C_3(G^{(n)}) = \binom{\left[\frac{n}{2}\right]}{3} + \binom{\left[\frac{n+1}{2}\right]}{3}, \tag{2}$$

(i.e. $G^{(n)}$ runs through all graphs whose complement contains no triangle).
LORDEN further determined all cases where there is equality in (2).
GOODMAN also raised the problem of determining

$$\min\,(C_k(G^{(n)}) + C_k(\bar{G}^{(n)})),$$

but this seems difficult even for $k = 4$.
I will prove by probabilistic arguments the following
Theorem 1. *For every $k \geqq 3$ and every n*

$$\min\,(C_k(G^{(n)}) + C_k(\bar{G}^{(n)})) < \frac{2\binom{n}{k}}{2^{\binom{k}{2}}}.$$

It is surprising that a crude probabilistic argument gives a result which for $k = 3$ is so close the correct one. This phenomenon can often be observed in this subject [4]. Theorem 1 seems to show that Goodman's problem will be much more difficult for $k > 3$ then for $k = 3$, since it does not seem easy to find graphs which give values of $C_4(G^{(n)}) + C_4(\bar{G}^{(n)})$ which are as small as $\binom{n}{4}\Big/32$. The construction analogous to the one of GOODMAN gives only $3\binom{\left[\frac{n}{3}\right]}{4}$ which is much bigger. It seems likely that

$$\lim_{n=\infty} \min \frac{C_k(G^{(n)}) + C_k(\bar{G}^{(n)})}{\binom{n}{k}} = \frac{1}{2^{\binom{k}{2}-1}}. \tag{3}$$

(3) follows from (1) for $k = 3$. I can not prove it for $k > 3$. I will only outline the proof of the crude estimate $\left(\binom{2k-2}{k-1} = t\right)$

$$\lim_{n=\infty} \min \frac{C_k(G^{(n)}) + C_k(\bar{G}^{(n)})}{\binom{n}{k}} \geq \frac{k!}{t(t-1)\ldots(t-k+1)}. \tag{4}$$

The following further problems might be of interest. Determine

$$\min C_k(G^{(n)}) = f(n, k, l)$$

where $G^{(n)}$ runs through all graphs of n vertices for which $\bar{G}^{(n)}$ does not contain a complete graph of l vertices.
The result of LORDEN gives that for all even n and for odd $n > 9$

$$f(n, 3, 3) = \binom{\left[\frac{n}{2}\right]}{3} + \binom{\left[\frac{n+1}{2}\right]}{3}.$$

I can not at present determine $f(n, k, l)$ for any other values of k and l. Perhaps for $n > n_0(k, l)$

$$f(n, k, l) = \sum_{i=0}^{l-2} \binom{\left[\frac{n+i}{l-1}\right]}{k}. \tag{5}$$

The simplest special case which I can not do is $f(3n, 3, 4) = 3\binom{n}{3}$. HANANI and I proved the following

Theorem 2. *Let* $l = \binom{t}{2} + r$, $0 < r \leq t$. *Then* (the maximum is to be taken over all graphs having l edges)

$$\max C_k(G_l) = \binom{t}{k} + \binom{r}{k-1} = g(l). \tag{6}$$

Finally we prove

Theorem 3. *Let* $l > k$. *We have*

$$\max C_k(G^{(n)}) = \sum_{0 \leq i_1 < \ldots < i_k \leq l-2} \prod_{r=1}^{k} \left[\frac{n+i_r}{l-1}\right] = h(n, l, k) \tag{7}$$

where the maximum is taken over all graphs having n vertices which do not contain a complete l-gon (i.e. a $G^{(l)}_{\binom{l}{2}}$).

Theorem 3 is probably connected with the conjecture (5). (See [8].)

Proof of Theorem 1. The number of graphs $G^{(n)}$ having the labelled vertices $x_1, \ldots, x_n$ clearly equals $2^{\binom{n}{2}}$. A simple argument shows that the number of graphs $G^{(n)}$ for which either $G^{(n)}$ or $\bar{G}^{(n)}$ contains the complete subgraph having the vertices $x_{i_1}, \ldots, x_{i_k}$ is $2 \cdot 2^{\binom{n}{2} - \binom{k}{2}}$. Thus summing over all the $\binom{n}{k}$ k-tuples

$$\sum (C_k(G^{(n)}) + C_k(\bar{G}^{(n)})) = \binom{n}{k} 2^{\binom{n}{2} - \binom{k}{2} + 1} \tag{8}$$

where the summation is extended over all the $2^{\binom{n}{2}}$ graphs $G^{(n)}$. (8) immediately implies

$$\min (C_k(G^{(n)}) + C_k(\bar{G}^{(n)})) \leq \frac{2\binom{n}{k}}{2^{\binom{k}{2}}}. \tag{9}$$

The sign of inequality in (9) follows if we observe that if $G^{(n)}$ is the complete graph of n vertices, then

$$C_k(G^{(n)}_{\binom{n}{2}}) = \binom{n}{k} > \frac{2\binom{n}{k}}{2^{\binom{k}{2}}}$$

($\bar{G}^{(n)}_{\binom{n}{2}}$ is the graph without edges). Thus for at least one of the $2^{\binom{n}{2}}$ summands (8) we have the inequality sign in (9), which completes the proof of Theorem 1.

Now we prove (4). A well known theorem of RAMSEY [5] asserts that for $t = \binom{2k-2}{k-1}$

$$C_k(G^{(t)}) + C_k(\bar{G}^{(t)}) \geqq 1\,. \tag{10}$$

(10) implies that if $x_{i_1}, \ldots, x_{i_t}$ are any t vertices of $G^{(n)}$ then

$$C_k(G(x_{i_1}, \ldots, x_{i_t})) + C_k(\bar{G}(x_{i_1}, \ldots, x_{i_t})) \geqq 1\,. \tag{11}$$

From (11) we have by a simple argument (every t-tuple gives at least one complete k-gon of $G^{(n)}$ or $\bar{G}^{(n)}$ and the same k-tuple occurs in exactly $\binom{n-k}{t-k}$ t-tuples)

$$C_k(G^{(n)}) + C_k(\bar{G}^{(n)}) \geqq \frac{\binom{n}{t}}{\binom{n-k}{t-k}} = \frac{n(n-1)\ \ldots\ (n-k+1)}{t(t-1)\ \ldots\ (t-k+1)}$$

which easily implies (4).

It would not be difficult to show that

$$\lim_{n=\infty} \min \frac{C_k(G^{(n)}) + C_k(\bar{G}^{(n)})}{\binom{n}{k}}$$

exists, but I can not determine it (its value was conjectured in (3)).

Now we *outline* the proof of Theorem 2. $\max C_k(G_l) \geqq g(l)$ is trivial. It suffices to consider the complete graph $G^{(t)}_{\binom{t}{2}}$ and an extra vertex connected with r vertices of $G^{(t)}_{\binom{t}{2}}$. The Theorem is trivial for $l \leqq \binom{k}{2}$. For $l < \binom{k}{2}$ both sides of (6) are 0, and for $l = \binom{k}{2}$ both sides are 1. We shall now use induction and assume that Theorem 2 holds for all $l' < l$ and then prove it for $l = \binom{t}{2} + r$, $0 < r \leqq t$, $k \leqq t$. We clearly must have $\pi(G_l) \geqq t + 1$ $\left(\text{since } l > \binom{t}{2}\right)$. Assume first that G_l has a vertex x_1 of valency $< t$. Clearly G_l contains at most $\binom{v(x_1)}{k-1}$ complete k-graphs one vertex of which is x_1. Thus clearly

$$C_k(G_l) \leqq \binom{v(x_1)}{k-1} + C_k(G_l - x_1) \text{ and } \nu(G_l - x_1) = l - v(x_1)\,.$$

Hence by our induction hypothesis and a simple computation $(v(x_1) < t)$

$$C_k(G_l) \leqq \binom{v(x_1)}{k-1} + g(l - v(x_1)) \leqq g(l)\,.$$

If all vertices of G_l have valency $\geqq t$, then from $l \leqq \binom{t+1}{2}$, $\pi(G_l) \geqq \geqq t+1$ we easily obtain

$$l = \binom{t+1}{2}, \quad \pi(G_l) = t+1 .$$

But then $C_k(G_l) = \binom{t+1}{k} = g(l)$, which completes the proof of Theorem 2.

We prove Theorem 3 by induction with respect to n. (7) holds for all k if $n \leqq k$ (for $n < k$ both sides of (7) are 0 and for $n = k$ they are both 1). Assume that (7) holds for every $m < n$ and every k. Since $G^{(n)}$ does not contain a $G^{(l)}_{\binom{l}{2}}$ by a theorem of ZARANKIEWICZ [6] it must contain a vertex x of valency not greater than

$$n - \left[\frac{n+l-2}{l-1}\right] = N .$$

By our induction hypothesis

$$C_k(G^{(n)} - x) \leqq h(n-1, l, k) . \tag{12}$$

Denote by $y_1, \ldots, y_t; t = v(x) \leqq N$ the vertices of $G^{(n)}$ connected to x by an edge. Clearly the graph $G(y_1, \ldots, y_t)$ contains no $G^{(l-1)}_{\binom{l-1}{2}}$, thus by our induction hypothesis it contains at most $h(t, l-1, k-1)$ subgraphs $G^{(k-1)}_{\binom{k-1}{2}}$. Hence the number of subgraphs $G^{(k)}_{\binom{k}{2}}$ of $G^{(n)}$ one vertex of which is x is at most

$$h(t, l-1, k-1) \leqq h(N, l-1, k-1) . \tag{13}$$

From (12 and (13) we easily obtain by a simple argument

$$\max C_k(G^{(n)}) \leqq h(n-1, l, k) + h(N, l-1, k-1) = h(n, l, k) . \tag{14}$$

To show that in (14) the sign of equality holds it suffices to consider the graph of TURÁN [7] where the vertices are split into $l-1$ classes, the i-th class has $\left[\frac{n+i-1}{l-1}\right]$ vertices and no two vertices of the same class are connected, but every two vertices of different class are connected by an edge. Thus the proof of (7) and Theorem 3 is completed.

Finally I would like to state the following conjecture which is a sharpening of (7): Put

$$F(n, l) = \sum_{0 \leqq i_1 < i_2 \leqq l-2} \left[\frac{n+i_1}{l-1}\right]\left[\frac{n+i_2}{l-1}\right].$$

$F(n, l)$ is the number of edges of Turán's graph, by his theorem [7] for every $G^{(n)}_{F(n,l)+1}$ contains a $G^{(l)}_{\binom{l}{2}}$. I believe that

$$\max C_k(G_{F(n,l)}) = h(n, l, k) \tag{15}$$

where the maximum is taken over all $G_{F(n,l)}$ which do not contain a $G^{(l)}_{\binom{l}{2}}$. (15) would imply (7) since by the theorem of TURÁN just stated a graph $G^{(n)}$ which contains no $G^{(l)}_{\binom{l}{2}}$ has $\leq F(n, l)$ edges.

(Received October 8, 1962)

REFERENCES

[1] GOODMAN A. W.: "On sets of acquaintances and strangers at any party." *Amer. Math. Monthly* **66** (1959) 778–783.
[2] SAUVÉ, L.: "On chromatic graphs." *Amer. Math. Monthly* **68** (1961) 107–111.
[3] LORDEN, G.: "Blue-empty chromatic graphs." *Amer. Math. Monthly* **6** (1962) 114–120.
[4] See ERDŐS, P.: "Some remarks on the theory of graphs." *Bull. Amer. Math. Soc.* **53** (1947) 292–294 and also "Graph theory and probability I and II." *Canadian Journal of Math.* **11** (1959) 34–38 and **13** (1961) 346–352.
[5] ERDŐS, P. and SZEKERES, G.: "A combinatorial problem in geometry." *Compositio Math.* **2** (1935) 463–470.
[6] ZARANKIEWICZ, K.: "Sur les relations symmétriques dans les ensembles fini." *Coll. Math.* **1** (1947) 10–14.
[7] TURÁN, P.: *Matematikai és Fizikai Lapok* **48** (1941) 436–452 (in Hungarian), see also "On the theory of graphs." *Coll. Math.* **3** (1954) 19–30.
[8] MOON, I. W. — MOSER, L.: "On a problem of Turán." *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* **7** (1962) A, 311 — 314.

О ЧИСЛЕ ПОЛНЫХ ГРАФОВ НАХОДЯЩИХСЯ В НЕКОТОРЫХ ГРАФАХ

P. ERDŐS

Резюме

Для графов, не содержащих рёбер-петель и многократных рёбер, имеют силу следующие теоремы:

Теорема 1. *Для всякий n и $k \geq 3$ существует такой граф G с n вершинами, что сумма чисел полных графов с k вершинами, находящихся в G и в дополнительном графе от G меньше* $2\binom{n}{k}\Big/2^{\binom{k}{2}}$.

Теорема 2. *Пусть* $l=\binom{t}{2}+r,\quad 0<r\leq t$. *Тогда граф, имеющий l рёбер может максимально содержать* $\binom{t}{k}+\binom{r}{k-1}$ *полных графов с k вершинами.*

Теорема 3. *Пусть $k<l$. Тогда граф с n вершинами, не содержащий полного графа с l вершинами, может содержать максимально*

$$\sum_{0\leq i_1<\ldots<i_k\leq l-2}\ \prod_{r=1}^{k}\left[\frac{n+i_r}{l-1}\right]$$

полных графов с k вершинами.

A kiadásért felel az Akadémiai kiadó igazgatója Műszaki szerkesztő: Vidosa László
Kézirat érkezett: 1962. XI. 19. — Példányszám: 800 — Terjedelem 19,9 (A/5) ív

63.56324 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György