GRAPH THEORY AND PROBABILITY

P. ERDÖS

A well-known theorem of Ramsay **(8; 9)** states that to every n there exists a smallest integer $g(n)$ so that every graph of $g(n)$ vertices contains either a set of n independent points or a complete graph of order n, but there exists a graph of $g(n) - 1$ vertices which does not contain a complete subgraph of n vertices and also does not contain a set of n independent points. (A graph is called complete if every two of its vertices are connected by an edge; a set of points is called independent if no two of its points are connected by an edge.) The determination of $g(n)$ seems a very difficult problem; the best inequalities for $g(n)$ are **(3)**

$$(1) \qquad 2^{\frac{1}{2}n} < g(n) \leqslant \binom{2n-2}{n-1}.$$

It is not even known that $g(n)^{1/n}$ tends to a limit. The lower bound in (1) has been obtained by combinatorial and probabilistic arguments without an explicit construction.

In our paper **(5)** with Szekeres $f(k, l)$ is defined as the least integer so that every graph having $f(k, l)$ vertices contains either a complete graph of order k or a set of l independent points ($f(k, k) = g(k)$). Szekeres proved

$$(2) \qquad f(k, l) \leqslant \binom{k+l-2}{k-1}.$$

Thus for

$$k = 3, f(3, l) \leqslant \binom{l+1}{2}.$$

I recently proved by an explicit construction that $f(3, l) > l^{1+c_1}$ **(4)**. By probabilistic arguments I can prove that for $k > 3$

$$(3) \qquad f(k, l) > l\binom{k+l-2}{k-1}^{c_2},$$

which shows that (2) is not very far from being best possible.

Define now $h(k, l)$ as the least integer so that every graph of $h(k, l)$ vertices contains either a closed circuit of k or fewer lines, or that the graph contains a set of l independent points. Clearly $h(3, l) = f(3, l)$.

By probabilistic arguments we are going to prove that for fixed k and sufficiently large l

$$(4) \qquad h(k, l) > l^{1+1/2k}.$$

Further we shall prove that

Received December 13, 1957.

https://doi.org/10.4153/CJM-1959-003-9 Published online by Cambridge University Press

(5) $$h(2k+1, l) < c_3\, l^{1+1/k}, \quad h(2k+2, l) < c_3\, l^{1+1/k}.$$

A graph is called r chromatic if its vertices can be coloured by r colours so that no two vertices of the same colour are connected; also its vertices cannot be coloured in this way by $r-1$ colours. Tutte (1, 2) first showed that for every r there exists an r chromatic graph which contains no triangle and Kelly (6) showed that for every r there exists an r chromatic graph which contains no k-gon for $k \leqslant 5$. (Tutte's result was rediscovered several times, for instance, by Mycielski (7). It was asked if such graphs exist for every k.) Now (4) clearly shows that this holds for every k and in fact that there exists a graph of n vertices of chromatic number $> n^{\epsilon}$ which contains no closed circuit of fewer than k edges.

Now we prove (4). Let n be a large number,

$$0 < \epsilon < \frac{1}{k}$$

is arbitrary. Put $m = [n^{1+\epsilon}]$ ($[x]$ denotes the integral part of x, that is, the greatest integer not exceeding x), $p = [n^{1-\eta}]$ where $0 < \eta < \epsilon/2$ is arbitrary. Let $\mathfrak{G}^{(n)}$ be the complete graph of n vertices $x_1, x_2, \ldots, x_n$ and $\mathfrak{G}^{(p)}$ any of its complete subgraphs having p vertices. Clearly we can choose $\mathfrak{G}^{(p)}$ in $\binom{n}{p}$ ways. Let

$$\mathfrak{G}_\alpha^{(n)}, \quad 1 \leqslant \alpha \leqslant \binom{\binom{n}{2}}{m}$$

be an arbitrary subgraph of $\mathfrak{G}^{(n)}$ having m edges (the number of possible choices of α is clearly as indicated).

First of all we show that for almost all α $\mathfrak{G}_\alpha^{(n)}$ has the property that it has more than n common edges with every $\mathfrak{G}^{(p)}$. Almost all here means: for all α's except for

$$o\binom{\binom{n}{2}}{m}.$$

Let the vertices of $\mathfrak{G}^{(p)}$ be $x_1, x_2, \ldots, x_p$. The number of graphs $\mathfrak{G}_\alpha^{(n)}$ containing not more than n of the edges (x_i, x_j), $1 \leqslant i < j \leqslant p$ equals by a simple combinatorial reasoning

$$\sum_{l=0}^{n} \binom{\binom{p}{2}}{l}\binom{\binom{n}{2}-\binom{p}{2}}{m-l} < (n+1)\binom{\binom{p}{2}}{n}\binom{\binom{n}{2}-\binom{p}{2}}{m}$$

$$< p^{2n}\binom{\binom{n}{2}-\binom{m}{2}}{m} < \binom{\binom{n}{2}}{m} p^{2n}\left(1-\frac{\binom{p}{2}}{\binom{n}{2}}\right)^m < \binom{\binom{n}{2}}{m} p^{2n}\left(1-\frac{p^2}{n^2}\right)^m$$

$$< \binom{\binom{n}{2}}{m} p^{2n} \exp\left(-\frac{mp^2}{n^2}\right).$$

https://doi.org/10.4153/CJM-1959-003-9 Published online by Cambridge University Press

Now the number of possible choices for $\mathfrak{G}^{(p)}$ is

$$\binom{n}{p} < n^p < p^n.$$

Thus the number of α's for which there exists a $\mathfrak{G}^{(p)}$ so that $\mathfrak{G}^{(p)} \cap \mathfrak{G}_\alpha^{(n)}$ has not more than n^ϵedges is less than $(\eta < \epsilon/2)$

$$\binom{\binom{n}{2}}{m} p^{3n} \exp(- n^{1+\epsilon-2\eta}) = o\binom{\binom{n}{2}}{m}$$

as stated.

Unfortunately almost all of these graphs $\mathfrak{G}_\alpha^{(n)}$ contain closed circuits of length not exceeding k (in fact almost all of them contain triangles). But we shall now prove that almost all $\mathfrak{G}_\alpha^{(n)}$ contain fewer than n/k closed circuits of length not exceeding k.

The number of graphs $\mathfrak{G}_\alpha^{(n)}$ which contain a given closed circuit (x_1, x_2), $(x_2, x_3), \ldots, (x_l, x_1)$ clearly equals

$$\binom{\binom{n}{2} - l}{m - l}.$$

The circuit is determined by its vertices and their order—thus there are $n(n-1)\ldots(n-l+1)$ such circuits. Therefore the expected number of closed circuits of length not exceeding k equals

$$\binom{\binom{n}{2}}{m}^{-1} \sum_{l=3}^{k} l! \binom{n}{l} \binom{\binom{n}{2} - l}{m - l} < (1 + \sigma(1)) \sum_{l=3}^{k} n^l \left(\frac{m}{\binom{n}{2}}\right)^l$$

$$< (1 + o(1))\, n^k \frac{(2m)^k}{n^{2k}} = \sigma(n)$$

since $\epsilon < 1/k$. Therefore, by a simple and well-known argument, the number of the α's for which $\mathfrak{G}_\alpha^{(n)}$ contains n/k or more closed paths of length not exceeding k is

$$o\binom{\binom{n}{2}}{m},$$

as stated.

Thus we see that for almost all α $\mathfrak{G}_\alpha^{(n)}$ has the following properties: in every $\mathfrak{G}^{(p)}$ it has more than n edges and the number of its closed circuits having k or fewer edges is less than n/k. Omit from $\mathfrak{G}_\alpha^{(n)}$ all the edges contained in a closed circuit of k or fewer edges. By what has just been said we omit fewer than n edges. Thus we obtain a new graph $\mathfrak{G}_\alpha'^{(n)}$ which by construction does not contain a closed circuit of k or fewer edges. Also clearly $\mathfrak{G}_\alpha'^{(n)} \cap \mathfrak{G}^{(p)}$

is not empty for every $\mathfrak{G}^{(p)}$. Thus the maximum number of independent points in $\mathfrak{G}_\alpha'^{(n)}$ is less than $p = [n^{1-\eta}]$, or

$$h(k, [n^{1-\eta}]) > n$$

which proves (4).

By more complicated arguments one can improve (4) considerably; thus for $k = 3$ I can show that for every $\epsilon > 0$ and sufficiently large l

$$f(3, l) = h(3, l) > l^{2-\epsilon},$$

which by (2) is very close to the right order of magnitude.

At the moment I am unable to replace the above "existence proof" by a direct construction.

By using a little more care I can prove by the above method the following result: there exists a (sufficiently small) constant c_4 so that for every k and l

$$h(k, l) > c_4\, l^{1+\frac{1}{3k}}. \tag{6}$$

(If $k > c \log l$ (6) is trivial since $h(k, l) \geqslant l$.)

From (6) it is easy to deduce that to every r there exists a c_5 so that for $n > n_0(r, c_5)$ there exists an r chromatic graph of n vertices which does not contain a closed circuit of fewer than $[c_5 \log n]$ edges. I am not sure if this result is best possible.

We do not give the details of the proof of (3) since it is simpler than that of (4). For $k = 3$ (3) follows from (4). If $k > 3$, put

$$m = c_6[n^{2-\frac{2}{k-1}}]$$

and denote by $\mathfrak{G}_\alpha^{(n)}$ the "random" graph of m edges. By a simple computation it follows that for sufficiently small c_6, $\mathfrak{G}_\alpha^{(n)}$ does not contain a complete graph of order k for more than

$$0\cdot 9\binom{\binom{n}{2}}{m}$$

values of α, and that for more than this number of values of α $\mathfrak{G}_\alpha^{(n)}$ does not contain a set of $c_7 n^{2/k-1} \log n$ independent points ($c_7 = c_7(c_6)$ is sufficiently large). Thus

$$f(k, c_7 n^{2/k-1} \log n) > n,$$

which implies (3) by a simple computation.

Now we prove (5). It will clearly suffice to prove the first inequality of (5). We use induction on l. Let there be given a graph $\mathfrak{G}$ having $h(2k + 1, l) - 1$ vertices which does not contain a closed circuit of $2k + 1$ or fewer edges and for which the maximum number of independent points is less than l. If every point of $\mathfrak{G}$ has order at least $[l^{1/k}] + 2$ (the order of a vertex is the number of edges emanating from it) then, starting from an arbitrary point, we reach in k steps at least l points, which must be all distinct since otherwise $\mathfrak{G}$ would have to contain a closed circuit of at most $2k$ edges. The endpoints thus obtained must be independent, for if two were connected by an edge $\mathfrak{G}$ would contain a closed circuit of $2k+1$ edges. Thus $\mathfrak{G}$ would have a set of at least l independent points, which is false.

Thus $\mathfrak{G}$ must have a vertex x_1 of order at most $[l^{1/k}]+1$. Omit the vertex x_1 and all the vertices connected with it. Thus we obtain the graph $\mathfrak{G}'$ and x_1 is not connected with any point of $\mathfrak{G}'$, thus the maximum number of independent points of $\mathfrak{G}'$ is $l-1$, or $\mathfrak{G}'$ has at most $h(2k+1, l-1)-1$ vertices, hence

$$h(2k+1, l) \leqslant h(2k+1, l-1) + [l^{1/k}] + 2$$

which proves (5).

References

1. Blanche Descartes, *A three colour problem*, Eureka (April, 1947). (Solution March, 1948.)
2. ——— *Solution to Advanced Problem no. 4526*, Amer. Math. Monthly, *61* (1954), 352.
3. P. Erdös, *Some remarks on the theory of graphs*, B.A.M.S. *53*, (1947), 292–4.
4. ——— *Remarks on a theorem of Ramsey*, Bull. Research Council of Israel, Section F, *7* (1957).
5. P. Erdös and G. Szekeres, *A combinatorial problem in geometry*, Compositio Math. *2* (1935) 463–70.
6. J. B. Kelly and L. M. Kelly, *Paths and circuits in critical graphs*, Amer. J. Math., *76* (1954), 786–92.
7. J. Mycielski, *Sur le colorage des graphs*, Colloquium Math. *3* (1955), 161–2.
8. F. P. Ramsay, *Collected papers*, 82–111.
9. T. Skolem, *Ein kombinatorischer Satz mit Anwendung auf ein logisches Entscheidungs problem*, Fund. Math. *20* (1933), 254–61.

University of Toronto
and
Technion, Haifa