ON CIRCUITS AND SUBGRAPHS OF CHROMATIC GRAPHS

P. Erdös

A graph is said to be k-chromatic if its vertices can be split into k classes so that two vertices of the same class are not connected (by an edge) and such a splitting is not possible for $k-1$ classes. Tutte was the first to show that for every k there is a k-chromatic graph which contains no triangle [**1**].

The lower girth of a graph is defined as the smallest integer t so that our graph has a circuit of t edges. J. B. Kelly and L. M. Kelly [**2**] showed that there exist graphs of arbitrarily high chromatic number and lower girth 6. I proved [**3**] that for every t and k there is a graph of chromatic number k and lower girth t. In fact I showed the following sharper result: To every k there is an ϵ so that for $n > n_0(\epsilon, k)$ there is a $G^{(n)}$ ($G^{(n)}$ will denote a graph of n vertices, $G_l^{(n)}$ will denote a graph with n vertices and l edges) of chromatic number k and lower girth $\geqslant \epsilon \log n$. We shall show that in some sense this result is best possible. First we introduce some notations. $f(m, k; n)$ denotes the maximum of the chromatic number of all graphs $G^{(n)}$, every subgraph of m vertices of which has chromatic number not exceeding k; $g_k(n)$ is the largest integer for which there is a $G^{(n)}$ of chromatic number k and lower girth $g_k(n)$. Clearly $g_3(n)$ is the largest odd integer not exceeding n (since every odd circuit has chromatic number 3). For $k > 3$ the determination of $g_k(n)$ seems very difficult. In [**3**] I proved† ($c_1, c_2, \ldots$ will denote suitable positive constants)

$$g_k(n) > c_1 \frac{\log n}{\log k}. \tag{1}$$

Now I shall prove

Theorem 1. *For $k \geqslant 4$ we have*

$$g_k(n) \leqslant \frac{2 \log n}{\log (k-2)} + 1.$$

Theorem 1 and (1) shows that for $k \geqslant 4$ the order of magnitude of $g_k(n)$ is $\log n$ (it would be easy to replace (1) by an explicit inequality). It seems likely that for $k > 3$

$$\lim_{n \to \infty} g_k(n)/\log n$$

exists, but I have not been able to prove this.

Theorem 1 shows that the chromatic number can be "large" only if the lower girth is $\leqslant \epsilon \log n$. Theorem 1 further implies that every $G^{(n)}$

† In [**3**], (1) is proved in a slightly different form.

[Mathematika 9 (1962), 170–175]

which is 4-chromatic must contain a circuit of length $\leqslant 1+2\log_2 n$. I thought that every 4-chromatic $G^{(n)}$ must also contain an odd circuit of length $< c_2 \log n$. In other words, I conjectured that for a sufficiently large constant c_2 we have $f([c_2 \log n], 2; n) = 3$ (a graph all of whose circuits are even is 2-chromatic). T. Gallai (not knowing of my conjecture) constructed a 4-chromatic $G^{(n)}$ the smallest odd circuit of which has length $[n^{\frac{1}{2}}]$. Gallai's example is not yet published. Gallai and I then conjectured that the largest value of m for which $f(m, 2; n) = k$ is of the order of magnitude $n^{1/(k+2)}$, but we have not even been able to prove that for every $\epsilon > 0$ and $n > n_0(\epsilon)$, $f([\epsilon n], 2; n) = 3$.

The situation seems to change quite radically if we consider $f(m, 3, n)$ instead of $f(m, 2, n)$. In fact I shall prove

THEOREM 2. *To every k there is an $\epsilon > 0$ so that if $n > n_0(\epsilon, k)$ there exists a k-chromatic $G^{(n)}$ every subgraph of which having $[\epsilon n]$ vertices is at most 3 chromatic.*

Instead of Theorem 2 we shall prove the following stronger

THEOREM 3. *For $m > 3$ we have*

$$f(m, 3; n) > c_3 \left(\frac{n}{m}\right)^{1/3} \left(\log \frac{n}{2m}\right)^{-1}. \tag{1'}$$

For $f(m, k; n)$ at present we only can show a trivial upper bound:

$$f(m, k; n) \leqslant \left[\frac{n}{m} + 1\right] k. \tag{2}$$

(2) is indeed trivial since we can split the vertices of $G^{(n)}$ into at most $[n/m]+1$ sets each having $\leqslant m$ elements, and by assumption the graphs spanned by these vertices are at most k-chromatic.

(2) is certainly very far from being best possible. It is easy to deduce from a result of Szekeres and myself [**4**] that for $m > k$ [$f(m, k, n)$ in fact is meaningful only for $m > k$]

$$f(m, k; n) \leqslant f(k+1, k; n) < c_4 n^{1-(1/k)}. \tag{3}$$

The deduction of (3) from [**4**] is easy and can be left to the reader (to simplify his task we only remark that if every subgraph of $k+1$ vertices of $G^{(n)}$ is at most k-chromatic then $G^{(n)}$ cannot contain a complete $(k+1)$-gon $G^{(k+1)}_{\binom{k+1}{2}}$).

I further proved that [**5**]

$$f(3, 2; n) > c_5 n^{\frac{1}{2}} / \log n. \tag{4}$$

It seems probable that

$$f(k+1, k; n) > n^{1-(1/k)-\epsilon},$$

for every $\epsilon > 0$ if $n > n_0(\epsilon, k)$. I do not know to what extent the exponent $\frac{1}{3}$ in Theorem 3 can be improved for all values of m.

Proof of Theorem 1. A simple induction argument shows that every k-chromatic $G^{(n)}$ contains a subgraph $G^{(m)}$ every vertex of which has valency $\geqslant k-1$ (the valency, or order, of a vertex is the number of edges incident to it). Assume now that $G^{(n)}$ is k-chromatic and is of lower girth t. Let $G^{(m)}$ be a subgraph of $G^{(n)}$ every vertex of which has valency $\geqslant k-1$ and let X_0 be any vertex of $G^{(m)}$. Consider the set of vertices of $G^{(m)}$ which can be reached from X_0 by a path of $[(t-1)/2]$ or fewer edges. Clearly every such vertex can be reached by only one such path (for otherwise $G^{(m)}$, and therefore $G^{(n)}$, would contain a circuit of fewer than t edges). Since, further, every vertex of $G^{(m)}$ has valency $\geqslant k-1$, we obtain by a simple argument that there are more than $(k-2)^{[(t-1)/2]}$ vertices which can be reached from X_0 by a path of $[(t-1)/2]$ or fewer edges. Hence

$$(k-2)^{(t-2)/2} \leqslant (k-2)^{[(t-1)/2]} \leqslant m \leqslant n,$$

which proves Theorem 1†.

The proof of Theorem 3 will use simple probabilistic arguments and will be similar to previous proofs used by Renyi and the author [**5**]. First we need two Lemmas which are of independent interest. Denote by $G_l^{(n)}$ a graph having n vertices and l edges. If the vertices are labelled then the number of different graphs $G_l^{(n)}$ clearly equals $\binom{\binom{n}{2}}{l}$. A set of vertices of $G_l^{(n)}$ is said to be independent if no two of them are connected by an edge.

LEMMA 1. *Let $l = [rn]$, $r > c_6$: then for all except possibly $\frac{1}{10}\binom{\binom{n}{2}}{l}$ graphs $G_l^{(n)}$ the maximum number of independent vertices is less than $(n/r)\log r$.*

Let $x_1, \ldots, x_n$ be the vertices of $G_l^{(n)}$. The number of graphs $G^{(n)}$ for which $x_{i_1}, \ldots, x_{i_u}$ is an independent set is clearly

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

Since the vertices can be chosen in $\binom{n}{u}$ ways, the number of graphs $G_l^{(n)}$ for which the maximum number of independent points is $\geqslant u$ is not greater than

$$\binom{n}{u}\binom{\binom{n}{2}-\binom{u}{2}}{l} < \frac{n^u e^u}{u^u}\binom{\binom{n}{2}-\binom{u}{2}}{l} < \left(\frac{en}{u}\right)^u \left(1-\frac{\binom{u}{2}}{\binom{n}{2}}\right)^l \binom{\binom{n}{2}}{l} \tag{5}$$

$$< \binom{\binom{n}{2}}{l}\left(\frac{en}{u}\right)^u e^{-lu^2/n^2}.$$

† This idea is used in [3] and also in Lemma 3 of P. Erdös and L. Pósa. "On the maximal number of disjoint circuits of a graph", *Publ. Math. Debrecen*, 9 (1962), 3–12.

By (5) the proof of our lemma will be complete, if we show that, for $u \geqslant (n/r)\log r$, $r > c_6$, we have

$$\left(\frac{en}{u}\right)^u e^{-lu^2/n^2} < \tfrac{1}{10}. \tag{6}$$

(6) can be shown by a simple computation and is left to the reader.

It would be easy to drop the condition $r > c_6$, but then $(n/r)\log r$ would have to be replaced by, say,

$$\frac{n\log(r+2)}{r+c_7}.$$

It seems that the order of magnitude $(n/r)\log r$ is not far from being best possible at least for certain ranges of r.

COROLLARY. *Let $l = [rn]$, $r > c_6$. Then for all except $\frac{1}{10}\binom{\binom{n}{2}}{l}$ graphs $G_l^{(n)}$ the chromatic number of $G_l^{(n)}$ is greater than $r/\log r$.*

The corollary immediately follows from Lemma 1, since if $G^{(n)}$ is k-chromatic the maximum number of independent vertices must be $\geqslant n/k$ (since the n vertices can be split into k independent sets).

LEMMA 2. *Let $l = [rn] \leqslant \dfrac{n^{4/3}}{4^{1/3}.100}$. Then for all but $\frac{1}{10}\binom{\binom{n}{2}}{l}$ graphs $G_l^{(n)}$ every subgraph spanned by u of its vertices, $4 \leqslant u \leqslant 10^{-6}nr^{-3}$, contains fewer than $\frac{3}{2}u$ edges.*

In particular the lemma implies that these $G_l^{(n)}$ contain no complete quadrilateral. This result is contained in my paper with Rényi quoted in [**6**].

Denote by $N(u, t)$, $4 \leqslant u \leqslant 10^{-6}nr^{-3}$, $\frac{3}{2}u \leqslant t \leqslant \min\left(\binom{u}{2}, l\right)$ the number of graphs $G_l^{(n)}$ which contain a subgraph $G_t^{(u)}$. To prove our lemma we have to show that

$$\sum_u \sum_t N(u, t) < \tfrac{1}{10}\binom{\binom{n}{2}}{l}, \tag{7}$$

where the summation is extended over $4 \leqslant u \leqslant 10^{-6}nr^{-3}$,

$$\tfrac{3}{2}u \leqslant t \leqslant \min\left(\binom{u}{2}, l\right).$$

First we estimate $N(u, t)$. Let $x_{i_1}, \ldots, x_{i_u}$ be any u vertices of $G_l^{(n)}$. The number of graphs $G_l^{(n)}$ for which the subgraph spanned by $x_{i_1}, \ldots, x_{i_u}$ contains t edges clearly equals

$$\binom{\binom{u}{2}}{t}\binom{\binom{n}{2}-\binom{u}{2}}{l-t}.$$

Since the vertices $x_{i_1}, \ldots, x_{i_u}$ can be chosen in $\binom{n}{u}$ ways, we evidently have

$$N(u, t) = \binom{n}{u}\binom{\binom{u}{2}}{t}\binom{\binom{n}{2}-\binom{u}{2}}{l-t}. \tag{8}$$

From (8) we obtain by a simple computation

$$(t \geqslant \tfrac{3}{2}u,\ l = [rn],\ u \leqslant 10^{-6}nr^{-3}),$$

$$N(u,t)\binom{\binom{n}{2}}{l}^{-1} < \frac{n^u e^u}{u^u}\cdot\frac{u^{2t}e^t}{t^t 2^t}\left(\frac{3l}{n^2}\right)^t < \left\{\left(\frac{en}{u}\right)^{2/3}\frac{eul}{n^2}\right\}^t$$

$$< \left(\frac{10u^{1/3}l}{n^{4/3}}\right)^t \leqslant \left(\frac{10rn(10^{-6}nr^{-3})^{1/3}}{n^{4/3}}\right)^t = 10^{-t}. \qquad (9)$$

From (9) we easily obtain by $u \geqslant 4$, $t \geqslant \frac{3}{2}u$, that (7) holds and hence our lemma is proved. $\left(r \leqslant \frac{n^{4/3}}{4^{1/3}.100}\right.$ was needed to make sure that $10^{-6}nr^{-3} \geqslant 4$ should be true; in other words, that the range for u should not be empty.$\Big)$

COROLLARY. *Let $l = [rn] \leqslant \frac{n^{4/3}}{4^{1/3}.100}$, then for all but $\frac{1}{10}\binom{\binom{n}{2}}{l}$ graphs $G_l^{(n)}$ every subgraph spanned by u of its vertices $u \leqslant 10^{-6}nr^{-3}$ is at most 3-chromatic.*

As stated previously a simple induction argument shows that every $G^{(u)}$ of chromatic number $\geqslant 4$ contains a subgraph $G^{(v)}$ every vertex of which has valency $\geqslant 3$. Thus $G^{(v)}$ has at least $\frac{3}{2}v$ edges and the corollary follows from Lemma 2.

The constant 10^{-6} could easily be replaced by a larger one and the exponent -3 in $10^{-6}nr^{-3}$ could also be slightly increased, but I do not pursue these investigations since the corollary is sharp enough to deduce Theorems 2 and 3 and at present I cannot obtain best possible estimations, or even estimations which are likely to be anywhere near being best possible.

Now we can prove Theorem 3. Put $r = \frac{1}{100}(n/m)^{1/3}$, $l = [rn]$. By the corollary to Lemma 1 we first of all obtain that for all but $\frac{9}{10}\binom{\binom{n}{2}}{l}$ graphs $G_l^{(n)}$ their chromatic number is greater than

$$\frac{r}{\log r} > c_3\left(\frac{n}{m}\right)^{1/3}\left(\log\frac{2n}{m}\right)^{-1}, \qquad (10)$$

if c_3 is sufficiently small. (Lemma 1 applies since we can assume that $r > c_6$, for if not then $m \geqslant 10^{-6}nc_6^{-3}$ and for sufficiently small c_3 (1′) becomes trivial.)

Secondly, by the corollary to Lemma 2 $\left(\text{since } m \geqslant 4,\ r \leqslant \frac{n^{1/3}}{4^{1/3}.100}\right.$ and Lemma 2 applies$\Big)$ for all but $\frac{9}{10}\binom{\binom{n}{2}}{l}$ graphs $G_l^{(n)}$ the chromatic number of all their subgraphs having at most u vertices is $\leqslant 3$ for

$$u \leqslant 10^{-6}nr^{-3} = m, \qquad (11)$$

(10) and (11) implies that for $m \geqslant m_0$ at least $\frac{4}{5}\binom{\binom{n}{2}}{l}$ of the graphs satisfies (1′), which completes the proof of Theorems 3 and 2.

To conclude I just wish to remark that from (4) one can deduce a much stronger result than is obtained by putting $m = 4$ in Theorem 3.

References.

1. Blanche Descartes, " A three colour problem ", *Eureka* (April 1947). Solution March 1948. See also J. Mycielski, " Sur le colorage des graphs ", *Colloquium Math.*, 3 (1955), 161–162.

2. J. B. Kelly and L. M. Kelly, " Paths and circuits in critical graphs ", *American J. of Math.*, 76 (1954), 786–792.

3. P. Erdös, " Graph theory and probability ", *Canadian J. of Math.*, 11 (1959), 34–38.

4. ——— and G. Szekeres, " A combinatorial problem in geometry ", *Comp. Math.*, 2 (1935), 463–470.

5. ———, " Graph theory and probability (II) ", *Canadian J. of Math.*, 13 (1961), 346–352.

6. ——— and A. Rényi, " On the evolution of random graphs ", *Pub. Mat. Inst. Hung. Acad.*, 5 (1960), 17–61. See also the paper quoted in [**3**] and [**5**].

University College,
London, W.C.1.

(*Received on the* 13*th of November*, 1962.)