SOME EXTREMAL PROBLEMS ON INFINITE GRAPHS

by

J. CZIPSZER, P. ERDŐS and A. HAJNAL

1. A well known theorem of TURÁN ([1]) states that every graph $G^{(n)}_{f(k,n)+1}$ of n vertices and $f(k, n) + 1$ edges where

$$f(k, n) = \frac{k-2}{2(k-1)}(n^2 - r^2) + \binom{r}{2}, \quad n = (k-1)\,t + r, \ 0 \leq r < k - 1$$

contains a complete k-gon and that this theorem is best possible since there are graphs $G^{(n)}_{f(k.n)}$ not containing complete k-gons, and in fact the structure of these graphs is uniquely determined.

Some problems in measure and set theory led us to consider the following problems. Let the vertices of the infinite graph $G^{(\infty)}$ be the integers $1, 2, \ldots, n, \ldots$. (In what follows $G^{(\infty)}$ will always denote such a graph.) Denote by $G^{(n)}$ the subgraph of $G^{(\infty)}$ spanned by the vertices $1, 2, \ldots, n$ and by $g(n)$ the number of edges of $G^{(n)}$. At first thought it seemed possible that if $g(n)$ is "large" for all $n > n_0$ then this will imply that $G^{(\infty)}$ contains a complete k-gon even though $g(n)$ does not have to be as large as $f(k, n)$. But it is easy to see that no such theorem can hold. To see this let the edges of $G^{(\infty)}$ be (i, j): i odd, j even. Clearly $g(n) = f(3, n)$ for every n and nevertheless $G^{(\infty)}$ does not contain a triangle. Nevertheless it will be possible to obtain using our function $g(n)$ some results which do not seem uninteresting to us. First some definitions: By an I_k-path (increasing path of length k) of $G^{(\infty)}$ (or of a finite graph with vertices $1, 2, \ldots, n$) we shall mean a path $i_1 i_2 \ldots i_k i_{k+1}$ $(i_1 < i_2 < \ldots < i_k < i_{k+1})$. A path of length k will denote an ordinary path of k edges. Clearly if $G^{(\infty)}$ contains a complete graph of $k + 1$ vertices it also contains an I_k-path, but the converse is not true.

By an I_∞-path of $G^{(\infty)}$ we shall mean an infinite path $i_1 i_2 \ldots i_n \ldots$ where $i_1 < i_2 < \ldots < i_n < \ldots$.

ERDŐS and GALLAI [1] found nearly best possible estimates for the smallest integer $h_k(n)$ for which every $G^{(n)}_{h_k(n)+1}$ will contain a path of length k, but these results will not concern us here. It is easy to see that there is a graph with vertices $1, 2, \ldots, n$ and with $f(k + 1, n)$ edges which does not contain an I_k-path. To see this it suffices to consider TURÁN's well known graph $G^{(n)}_{f(k+1,n)}$ and enumerate its vertices in an obvious way. Nevertheless the situation changes completely if we assume a suitable lower bound for $g(n)$ which holds for all sufficiently large n. In fact we shall prove

Theorem I. *Let $G^{(\infty)}$ be a graph and assume that for all $n > n_0$ and an $\varepsilon > 0$*

$$g(n) > \left(\frac{1}{4} - \frac{1}{4k} + \varepsilon\right) n^2 \tag{1.1}$$

where $k = 2$ or $k = 3$.
Then $G^{(\infty)}$ contains infinitely many I_k-paths.

The theorem holds perhaps for $k > 3$ also, but at present we can not decide this question.

Remarks. Since $f(k+1, n) = \frac{1}{2}\left(1 - \frac{1}{k}\right) n^2 + O(1)$ for $n \to \infty$ our theorem implies that if $k = 2$ or $3, g(n) > \left(\frac{1}{2} + \varepsilon\right) f(k+1, n)$ for all $n < n_0$ then $G^{(\infty)}$ must already contain infinitely many I_k-paths.

It is easy to see that our theorem is best possible.

To see this define $G^{(\infty)}$ as follows: Let m_1 and m_2 ($m_1 < m_2$) be two vertices of $G^{(\infty)}$. m_1 and m_2 are connected if and only if $1 \leqq i_1 < i_2 \leqq k$ where $m_1 \equiv i_1 \pmod{k}$ and $m_2 \equiv i_2 \pmod{k}$. It is easy to see that for our $G^{(\infty)}$

$$g(n) = \frac{1}{4}\left(1 - \frac{1}{k}\right) n^2 + O(n)$$

and it clearly does not contain an I_k-path. In fact we shall prove the following sharper

Theorem II. *Let $G^{(\infty)}$ be a graph for which*

$$g(n) > \frac{n^2}{8} + \left(\frac{1}{32} + \varepsilon\right) \frac{n^2}{\log^2 n} \quad \text{if} \quad n > n_0.$$

Then $G^{(\infty)}$ contains infinitely many I_2-paths. The result is best possible since there exists a $G^{(\infty)}$ for which

$$g(n) > \frac{n^2}{8} + \frac{1}{32} \frac{n^2}{\log^2 n} + o\left(\frac{n^2}{\log^2 n}\right)$$

and which does not contain any I_2-path.

By the same method as used in the proof of Theorem II we can prove the following theorem: Assume that for $n > n_0$

$$g(n) > \frac{n^2}{8} + \left(\frac{1}{32} + \varepsilon\right) \frac{n^2}{\log^2 n}$$

Then $G^{(\infty)}$ contains infinitely many pairs of I_2-paths whose first and last endpoints coincide, i.e. it contains infinitely many quadruplets $i_1 < i_2 < i_4$, $i_1 < i_3 < i_4$ $(i_2 \neq i_3)$ and the edges $(i_1, i_2), (i_1, i_3), (i_2, i_4), (i_3, i_4)$. We do not discuss the proof. By induction we can easily prove the following Turánian theorem (see [1]): If G is a graph with vertices $1, 2, \ldots, n$ and $\left[\frac{n^2}{4}\right] + \left[\frac{n-1}{2}\right] + 1$ edges then G contains two I_2-paths whose first and last endpoints coincide. The estimation for the number of edges is best possible.

Theorem III. *Let* $G^{(\infty)}$ *be such that*

$$g(n) \geqq \frac{1}{4} n^2 - Cn . \tag{1.2}$$

Then $G^{(\infty)}$ *contains an infinite path. This result is best possible in the sense that* C *can not be replaced by* $A(n)$ *where* $A(n) \to \infty$.

It seemed to us likely that $g(n) > \left(\frac{1}{4} + \varepsilon\right) n^2$ will also imply the existence of an I_∞-path. But this is not the case. In fact we have

Theorem IV. *There exists a* $G^{(\infty)}$ *with*

$$\liminf \frac{g(n)}{n^2} > \frac{1}{4}$$

which does not contain an I_∞*-path. But there exists a constant* $\alpha > 0$ *such that every* $G^{(\infty)}$ *with*

$$\liminf \frac{g(n)}{n^2} > \frac{1}{2} - \alpha$$

contains an I_∞*-path.*

VERA T. SÓS asked the question: What condition on $g(n)$ will imply that $G^{(\infty)}$ should contain an infinite complete subgraph? We prove

Theorem V. *If* $g(n) > \frac{n^2}{2} - Cn$ *for infinitely many* n *then* $G^{(\infty)}$ *contains an infinite complete subgraph. But if we only assume that*

$$g(n) > \frac{n^2}{2} - f(n)\, n \tag{1.3}$$

for all n *where* $f(n)$ *tends to infinity as slowly as we please then* $G^{(\infty)}$ *does not have to contain an infinite complete graph.*

At present we can not answer the following question: Let G be any infinite graph every vertex of which is incident only to a finite number of edges what has to be assumed about $g(n)$ to make sure that $G^{(\infty)}$ should contain a subgraph isomorphic to G? In fact we get two problems here depending whether we require the vertices of G to be ordered or not. Our example used in the proof of the negative part of Theorem V (cf. § 5.) shows that we have to assume that every vertex of G has finite valency.

Without proof we state a few results connected with Theorem V. Assume that $g(n) > \binom{n}{2} - (1 - \alpha)n$ for every $n > n_0$ and some $\alpha > 0$, then $G^{(\infty)}$ contains an infinite complete graph whose vertices form a sequence of positive lower density. If we only assume $g(n) > \binom{n}{2} - n + o(n)$ it is easy to see that this does not have to remain true. If we only assume that $g(n) > \binom{n}{2} - Cn$ for infinitely many n and some C then $G^{(\infty)}$ contains an infinite complete graph whose vertices form a sequence of positive upper density. In fact the following stronger result holds: To every $\varepsilon > 0$ there exists a k so that $G^{(\infty)}$ contains k complete graphs the union of the set of their vertices forms a sequence of upper density $> 1 - \varepsilon$. Finally if we assume that $g(n) > \binom{n}{2} - Cn$ for every n and some C then to every $\varepsilon > 0$ there exists a k so that $G^{(\infty)}$ contains k complete graphs, the union of the set of their vertices forms a sequence of lower density $> 1 - \varepsilon$. We leave the proof of these statements to the reader.

2. Proof of Theorem I for $k = 2$. We shall show that if $G^{(\infty)}$ contains only finitely many I_2-paths then

$$\liminf \frac{g(n)}{n^2} \leqq \frac{1}{8} \tag{2.1}$$

which contradicts (1.1) for $k = 2$. Omitting a finite number of edges we can assume that $G^{(\infty)}$ does not contain any I_2-path. Then if a vertex u is the upper endpoint of an edge it can not be the lower endpoint of another edge. Denote by $u_1 < u_2 < \ldots$ the vertices which are not lower endpoints of any edge. Clearly the u_n sequence is infinite and two u_n are never connected. Hence

$$g(n) \leqq \sum_{u_k \leqq n} (u_k - k).$$

Now we establish a lemma which belongs to the theory of series and which clearly implies (2.1).

Lemma. *If* $u_1, u_2, \ldots$ *is a sequence with positive terms then*

$$\liminf \frac{u_1 + \ldots + u_n - \frac{n^2}{2}}{u_n^2} \leqq \frac{1}{8}. \tag{2.2}$$

Put $\limsup \frac{u_n}{n} = c$. If $c = 0$, (2.2) obviously holds. If $0 < c < \infty$ we choose a sequence u_{n_s} for which

$$\lim \frac{u_{n_s}}{n_s} = c.$$

Then

$$\sum_1^{n_s} u_k - \frac{n_s^2}{2} \leqq \sum_1^{n_s} ck - \frac{n_s^2}{2} + o(n_s^2) = (c-1)\frac{n_s^2}{2} + o(n_s^2),$$

and

$$\frac{\sum_1^{n_s} u_k - \frac{n_s^2}{2}}{u_{n_s}^2} \leqq \frac{c-1}{2c^2} + o(1) \leqq \frac{1}{8} + o(1), \tag{2.3}$$

which proves (2.2).

Finally if $c = \infty$, we choose a sequence u_{n_s} such that

$$\frac{u_k}{k} \leqq \frac{u_{n_s}}{n_s} \quad (1 \leqq k \leqq n_s) \text{ and } \frac{u_{n_s}}{n_s} \to \infty.$$

Putting $\frac{u_{n_s}}{n_s} = c$ (2.3) holds with c_s instead of c and consequently (2.2) holds also.

Remark. It is clear that in (2.2) there is strict inequality unless $\limsup \frac{u_n}{n} = 2$. This statement can be inverted in the following sense:

If $u_1, u_2, \ldots$ is an increasing sequence of positive numbers and

$$\liminf \frac{u_1 + \ldots + u_n - \frac{n^2}{2}}{u_n^2} \geqq \frac{1}{8} \tag{2.4}$$

then $\frac{u_n}{n} \to 2$.

In the proof of (2.4) we can suppose $\limsup \frac{u_n}{n} = 2$ so that

$$u_n \leqq 2n + o(n).$$

Put $\liminf \frac{u_n}{n} = \alpha$ and suppose $\alpha < 2$. We choose two numbers β and ε such that

$$\alpha < \beta < 2 - \varepsilon < 2 \quad \text{and} \quad \frac{\varepsilon}{2} < \left(1 - \frac{\beta}{2}\right)^2$$

and two sequences of integers $m_1, m_2, \ldots$ and $n_1, n_2, \ldots$ such that

$$n_1 < n_2 < \ldots, \quad m_\nu < n_\nu, \quad \frac{u_{m_\nu}}{m_\nu} \leqq \beta,$$

$$\frac{u_n}{n} < 2 - \varepsilon \text{ for } m_\nu \leqq n < n_\nu \quad \text{and} \quad \frac{u_{n_\nu}}{n_\nu} \geqq 2 - \varepsilon .$$

Hence for any $1 \leqq l \leqq m_\nu$ we have

$$\sum_1^{n_\nu} u_k \leqq \sum_1^{l} (2k + o(k)) + \sum_{l+1}^{m_\nu} m_\nu \beta + \sum_{m_\nu+1}^{n_\nu} (2-\varepsilon) k + O(1) \leqq$$

$$\leqq l^2 + o(l^2) + (m_\nu - l) m_\nu \beta + (2-\varepsilon) \frac{n_\nu^2 - m_\nu^2}{2} + o(u_{n_\nu}^2).$$

Putting $l = \left[\frac{m_\nu \beta}{2}\right]$ we obtain from here

$$\sum_1^{n_\nu} u_k \leqq m_\nu^2 \left(\beta - \frac{\beta^2}{4}\right) + (2-\varepsilon)\frac{n_\nu^2 - m_\nu^2}{2} + o(u_{n_\nu}^2) =$$

$$= -\left(\left(1 - \frac{\beta}{2}\right)^2 - \frac{\varepsilon}{2}\right) m_\nu^2 + \left(1 - \frac{\varepsilon}{2}\right) n_\nu^2 + o(u_{n_\nu}^2) \leqq$$

$$\leqq \frac{n_\nu^2}{2} + \frac{(2-\varepsilon)^2}{8} n_\nu^2 - \frac{\varepsilon^2}{8} n_\nu^2 + o(u_{n_\nu}^2) \leqq$$

$$\leqq \frac{n_\nu^2}{2} + \frac{u_{n_\nu}^2}{8} - \frac{\varepsilon^2}{32} u_{n_\nu}^2 + o(u_{n_\nu}^2)$$

which contradicts (2.4).

3. Proof of Theorem I for $k = 3$. We suppose that $G^{(\infty)}$ does not contain infinitely many I_3-paths. We shall then show that

$$\liminf \frac{g(n)}{n^2} \leq \frac{1}{6} \tag{3.1}$$

which contradicts (1.1) for $k = 3$. Moreover we can suppose that $G^{(\infty)}$ does not contain any I_3-path since the omission of a finite number of edges of $G^{(\infty)}$ does not alter the validity of (3.1). We denote by N the set of natural numbers and by C the set of those numbers which are not lower endpoints of any edge. Analogously let B be the set of natural numbers which belong to $N - C$ and which are not connected with any greater number in $N - C$. Putting $A = N - (B \cup C)$, it is clear that if two numbers m, n $(m < n)$ are connected then $m \in A, n \in B \cup C$ or $m \in B, n \in C$. It is also clear that C is infinite since otherwise $G^{(\infty)}$ would contain an I_∞-path.

Let $u_1, u_2, \ldots$ be an enumeration of the elements of $B \cup C$ in increasing order. Let $v_1 < v_2 < \ldots < v_t < \ldots$ be those indices for which $u_{v_t} \in C$. Since C is infinite, the u and v series are also infinite. For every u_k the number of elements of A less than u_k is $u_k - k$ and for every u_{v_l} the number of elements of B less than u_{v_l} is $v_l - l$, consequently

$$g(n) \leq u_1 + \ldots + u_s - \binom{s+1}{2} + v_1 + \ldots + v_t - \binom{t+1}{2}$$

where

$$u_s \leq n < u_{s+1}, \quad v_t \leq s < v_{t+1}.$$

For every natural number k we denote by w_k the number of v_l less than k. By an elementary computation we get

$$v_1 + \ldots + v_t = st - (w_1 + \ldots + w_s).$$

Hence

$$g(n) \leq \sum_{1}^{s} (u_k - w_k) - \binom{s+1}{2} - \binom{t+1}{2} + st. \tag{3.2}$$

Since $u_k \geq k > w_k$, we have $u_k - w_k > 0$.

We can select a sequence $s_1, s_2, \ldots$ $(s_\nu \to \infty)$ so that

$$\sum_{1}^{s_\nu} (u_k - w_k) \leq \frac{s_\nu}{2}(u_{s_\nu} - w_{s_\nu}) + o(u_{s_\nu}^2). \tag{3.3}$$

If $\limsup \frac{u_s - w_s}{s} = \infty$ then we choose the s_ν-s in such a way that $\frac{u_k - w_k}{k} \leq \frac{u_{s_\nu} - w_{s_\nu}}{s_\nu}$ should hold for $1 \leq k \leq s_\nu$ and for this sequence (3.3) is clearly satisfied. If $\limsup \frac{u_s - w_s}{s} = \varrho < \infty$ (3.3) will hold for any sequence s_ν for which $\frac{u_{s_\nu} - w_{s_\nu}}{s_\nu} \to \varrho$.

From (3.2) and (3.3) we have for $s = s_\nu$ and $n = u_s$

$$\frac{g(u_s)}{u_s^2} \leqq \frac{s}{2u_s^2}(u_s - w_s) - \frac{s^2}{2u_s^2} - \frac{t^2}{2u_s^2} + \frac{st}{u_s^2} + o(1)\,.$$

Considering that $w_s \geqq i - 1$ we deduce from here

$$\frac{g(u_s)}{u_s^2} \leqq \frac{1}{2}\,\frac{s}{u_s}\left(1 - \frac{s}{u_s}\right) + \frac{1}{2}\,\frac{s^2}{u_s^2}\,\frac{t}{s}\left(1 - \frac{t}{s}\right) + o(1) \leqq$$

$$\leqq \frac{1}{2}\,\frac{s}{u_s}\left(1 - \frac{s}{u_s}\right) + \frac{1}{8}\,\frac{s^2}{u_s^2} + o(1) = \frac{1}{6} - \frac{3}{8}\left(\frac{s}{u_s} - \frac{2}{3}\right)^2 + o(1) \leqq \frac{1}{6} + o(1)$$

for $s = s_\nu$. This proves (3.1).

4. Proof of Theorem II. First we show that there exists a $G^{(\infty)}$ with

$$g(n) = \frac{n^2}{8} + \left(\frac{1}{32} + o(1)\right)\frac{n^2}{\log^2 n} \tag{4.1}$$

which does not contain any I_2-path. To see this put

$$u_1 = 1, \quad u_2 = 2, \quad u_k = 2\,k + \left[\frac{k}{\log k}\right] \qquad (k = 3, 4, \ldots)$$

and consider the graph $G^{(\infty)}$ in which m and n ($m < n$) are connected if and only if n is an u_k and m is not an u_k. Clearly $G^{(\infty)}$ has no I_2-path. A simple computation shows that if $u_\nu \leqq n < u_{\nu+1}$ then

$$g(n) = \sum_1^\nu (u_k - k) = \frac{\nu^2}{2} + \sum_2^\nu \frac{k}{\log k} + O(\nu) =$$

$$= \frac{\nu^2}{2} + \frac{\nu^2}{2\log \nu} + \frac{\nu^2}{4\log^2 \nu} + o\left(\frac{\nu^2}{\log^2 \nu}\right) =$$

$$= \frac{u_\nu^2}{8} + \frac{1}{32}\,\frac{u_\nu^2}{\log^2 u_\nu} + o\left(\frac{\nu^2}{\log^2 \nu}\right) = \frac{n^2}{8} + \frac{1}{32}\,\frac{n^2}{\log^2 n} + o\left(\frac{n^2}{\log^2 n}\right)$$

which proves (4.1).

Theorem II is clearly implied by the following lemma which is essentially a refinement of the Lemma in § 2 and which may deserve interest for its own.

Lemma. *If $u_1, u_2, \ldots$ is a sequence with terms > 1 for $n > n_0$ then for any $\varepsilon > 0$ the inequality*

$$u_1 + u_2 + \ldots + u_n - \frac{n^2}{2} \leqq \frac{u_n^2}{8} + \left(\frac{1}{32} + \varepsilon\right)\frac{u_n^2}{\log^2 u_n}$$

holds for infinitely many n.

To prove this we shall distinguish several cases.

Case A. For infinitely many n $u_n > 2n$.

Case A. 1. $\limsup \frac{\log n}{n}(u_n - 2n) = 0$.

Put
$$u_n = 2n + A_n \frac{n}{\log^2 n} \qquad (n = 2, 3, \dots)$$
then
$$A_n \leqq o(\log n).$$

If $\limsup A_n = \infty$ for infinitely many n the relation
$$A_n \geqq A_m \qquad (m = 1, 2, \dots, n)$$
holds and for these n
$$\sum_1^n u_k - \frac{n^2}{2} \leqq \frac{n^2}{2} + A_n \sum_2^n \frac{k}{\log^2 k} + O(n) =$$
$$= \frac{n^2}{2} + A_n \frac{n^2}{2\log^2 n} + o\left(\frac{n^2}{\log^2 n}\right) \leqq \frac{u_n^2}{8} + o\left(\frac{u_n^2}{\log^2 u_n}\right).$$

If $\limsup A_n = c < \infty$ then for a suitable subsequence of the u_n we have
$$u_n = 2n + (c + o(1)) \frac{n}{\log^2 n}$$
and for these u_n
$$\sum_1^n u_k - \frac{n^2}{2} \leqq \frac{n^2}{2} + (c + o(1)) \sum_1^n \frac{k}{\log^2 k} + O(n) =$$
$$= \frac{n^2}{2} + c\frac{n^2}{2\log^2 n} + o\left(\frac{n^2}{\log^2 n}\right) = \frac{u_n^2}{8} + o\left(\frac{n^2}{\log^2 n}\right) \leqq \frac{u_n^2}{8} + o\left(\frac{u_n^2}{\log^2 u_n}\right).$$

Case A. 2. $0 < \limsup \frac{\log n}{n}(u_n - 2n) = c < \infty$.

Put
$$u_n = 2n + B_n \frac{n}{\log n \log\log n} \qquad (n > e^e).$$

We can choose a subsequence B_{n_ν} so that $B_m \leqq B_{n_\nu}$ if $m \leqq n_\nu$ and
$$\frac{B_{n_\nu}}{\log\log n_\nu} = c + o(1).$$

We have by a simple computation for $n = n_\nu$
$$\sum_1^n u_k - \frac{n^2}{2} \leqq \frac{n^2}{2} + B_n \sum_{27}^n \frac{k}{\log k \log\log k} + O(n) =$$
$$= \frac{n^2}{2} + B_n \frac{n^2}{2\log n \log\log n} + B_n \frac{n^2}{4\log^2 n \log\log n} + o\left(\frac{n^2}{\log^2 n}\right)$$

and

$$\frac{u_n^2}{8} + \left(\frac{1}{32} + \varepsilon\right)\frac{u_n^2}{\log^2 u_n} = \frac{n^2}{2} + B_n \frac{n^2}{2\log n \log\log n} +$$

$$+ \frac{B_n^2}{8}\,\frac{n^2}{\log^2 n(\log\log n)^2} + \left(\frac{1}{8} + 4\,\varepsilon\right)\frac{n^2}{\log^2 n} + o\left(\frac{n^2}{\log^2 n}\right).$$

We have to show that

$$\frac{B_n n^2}{4\log^2 n \log\log n} \leqq \frac{n^2}{8\log^2 n(\log\log n)^2} + \left(\frac{1}{8} + o(1)\right)\frac{n^2}{\log^2 n},$$

i.e.

$$\frac{B_n}{4\log\log n} \leqq \frac{B_n^2}{8(\log\log n)^2} + \frac{1}{8} + o(1)$$

for sufficiently large $n = n_\nu$ but this amounts to

$$\frac{c}{4} \leqq \frac{c^2}{8} + \frac{1}{8}$$

which is true.

Case A. 3. $\limsup \frac{\log n}{n}(u_n - 2\,n) = \infty$.

Put

$$u_n = 2\,n + C_n \frac{n}{\log n}.$$

For a suitable subsequence C_{n_ν} we have $C_m \leqq C_{n_\nu}$ if $m \leqq n_\nu$ and $C_{n_\nu} \to \infty$ Hence

$$\sum_1^n u_k - \frac{n^2}{2} \leqq \frac{n^2}{2} + C_n \sum_2^n \frac{k}{\log k} + O(n) =$$

$$= \frac{n^2}{2} + C_n \frac{n^2}{2\log n} + C_n \frac{n^2}{4\log^2 n} + C_n\, o\left(\frac{n^2}{\log^2 n}\right) \leqq \frac{u_n^2}{8}$$

if $n = n_\nu$ and ν is sufficiently large.

Case B. $u_n \leqq 2n$ for $n \geqq n_1$. If $\limsup \frac{u_n}{n} = 0$ then the statement of the lemma is evidently true. If $0 \leqq \limsup \frac{u_n}{n} < 2$ the lemma directly follows from (2.3). So we can suppose $\limsup \frac{u_n}{n} = 2$. Putting $u_n = 2n - D_n n$ we have $\liminf D_n = 0$ and for a suitable n_2 and infinitely many n $D_n \leqq D_m$ if $n_2 \leqq m \leqq n$. For these n

$$\sum_1^n u_k - \frac{n^2}{2} \leqq \frac{n^2}{2}(1 - D_n) + O(n) \leqq \frac{u_n^2}{8} + o\left(\frac{u_n^2}{\log^2 u_n}\right).$$

This concludes the proof of the lemma.

5. Proof of Theorem V. First we prove the second statement of the theorem. Let $h(n)$ tend to infinity sufficiently fast and connect n with all the m for which either $n < m \leqq h(n)$ or $m < n \leqq h(n)$. Clearly our $G^{(\infty)}$ does not contain an infinite complete subgraph since in fact every vertex has finite valency and if $h(n)$ tends to infinity sufficiently fast (1.3) is clearly satisfied.

Now we prove the positive part of Theorem V.

If $G^{(\infty)}$ does not contain an infinite complete graph we can construct by induction a sequence $1 = i_0 < i_1 < i_2 < \ldots$ so that if $i_k \leqq y$ then y is not connected with at least one vertex lying in $[i_{k-1}, i_k)$. Now if k is fixed and $n \geqq i_k$ then for every $i_k \leqq y \leqq n$ there are at least k vertices to the left of y which are not connected with y. Hence $g(n) < \binom{n}{2} - (n - i_k)\, k \leqq \binom{n}{2} - \frac{k}{2} n$ if $n \geqq 2i_k$. If $k > 2C$, this is a contradiction which proves the theorem.

6. Proof of Theorem III. We can assume that $G^{(\infty)}$ does not contain any infinite complete subgraph. The proof will be based on the following lemma.

Lemma. *Let us say that an infinite graph G whose vertices are natural numbers, has property $\mathcal{P}$ if*

a) *G has no infinite complete subgraph,*

b) *Denoting by $v(n)$ the number of vertices $\leqq n$ and by $g(n)$ the number of edges connecting vertices $\leqq n$ the inequality*

$$g(n) \geqq \frac{1}{4} v^2(n) - Cv(n) \tag{6.1}$$

holds for some C and for every n. If G has property $\mathcal{P}$ G has an infinite component who has also property $\mathcal{P}$.

First we deduce the theorem from the lemma. Applying the lemma to $G^{(\infty)}$ we get an infinite component G_1' of $G^{(\infty)}$, with property $\mathcal{P}$. Omitting an arbitrary vertex i_1 of G_1 we get a graph G_1 which clearly also verifies $\mathcal{P}$. Hence G_1 has also an infinite component G_2' with property $\mathcal{P}$ and because of the connectedness of G_1', G_2' contains a vertex i_2 which is connected with i_1. Putting $G_2 = G_2' - \{i_2\}$ G_2 has also property $\mathcal{P}$. Repeating this construction ad infinitum we get a sequence $i_1, i_2, \ldots$ of distinct vertices which form an infinite path.

In the proof of the Lemma we can assume that G is a $G^{(\infty)}$-graph that is $v(n) = n$ and $g(n)$ has the usual meaning. Let us denote by $G_1, G_2, \ldots$ the components of G. $v_k(n)$ and $g_k(n)$ denote the number of vertices $\leqq n$ of G_k respectively the number of edges of G_k which connect vertices $\leqq n$.

First we prove that

(6.2) *There exists a subscript k_0 such that the function $v_{k_0}(n)$ majorizes the functions $v_k(n)$ for every $n > n_0$ and for every k.*

The negation of (6.2) would clearly imply the existence of an infinite sequence $n_1 < n_2 < \ldots$ satisfying the following condition.

For every ν there exist numbers k_ν', k_ν'' for which $k_\nu' < k_\nu''$ and for every k $v_k(n_\nu) \leqq v_{k_\nu'}(n_\nu) = v_{k_\nu''}(n_\nu)$.

Putting

$$\gamma_\nu = v_{k_\nu'}(n_\nu) = v_{k_\nu''}(n_\nu) \tag{6.3}$$

it follows

$$\gamma_\nu \leqq \frac{n_\nu}{2} \tag{6.4}$$

and

$$g(n_\nu) \leqq g_{k'_\nu}(n_\nu) + \frac{1}{2}\gamma_\nu^2 + \frac{1}{2}\sum_{k \neq k'_\nu} v_k^2(n_\nu) \leqq$$

$$\leqq g_{k'_\nu}(n_\nu) + \frac{1}{2}(n_\nu - \gamma_\nu)\gamma_\nu.^1$$

Combining this with (6.1) we obtain

$$g_{k'_\nu}(n_\nu) \geqq \frac{1}{2}\gamma_\nu^2 + \frac{n_\nu}{2}\left(\frac{n_\nu}{2} - \gamma_\nu - 2C\right). \tag{6.5}$$

Considering that $g_{k'_\nu}(n_\nu) \leqq \frac{1}{2}\gamma_\nu^2$, we have

$$\gamma_\nu \geqq \frac{n_\nu}{2} - 2C. \tag{6.6}$$

From (6.5), (6.4) and (6.6) it follows

$$g_{k'_\nu}(n_\nu) \geqq \frac{1}{2}\gamma_\nu^2 - 2C\gamma_\nu - 4C^2. \tag{6.7}$$

Considering (6.3) it follows from (6.6) that

$$v_k(n_\nu) \leqq 4C \quad \text{if} \quad k'_\nu \neq k \neq k''_\nu. \tag{6.8}$$

Let ν_0 be an integer for which

$$n_{\nu_0} > 12C.$$

In view of (6.6))

$$\gamma_{\nu_0} > 4C.$$

Thus for $\nu > \nu_0$

$$v_{k'_{\nu_0}}(n_\nu) > 4C \quad \text{and} \quad v_{k''_{\nu_0}}(n_\nu) > 4C$$

and so in view of (6.8) we have

$$k'_\nu = k'_{\nu_0}, \quad k''_\nu = k''_{\nu_0}.$$

Hence (6.7) means that $G_{k'_{\nu_0}}$ satisfies the hypothesis of Theorem V and so it must contain an infinite complete subgraph which contradicts our assumption on G. Thus (6.2) is proved.

We can suppose $k_0 = 1$ i.e.

$$v_k(n) \leqq v_1(n) \quad \text{for} \quad n > n_0.$$

[1] Deducing the second inequality we used the fact that if $\Sigma x_i = a$, $x_i \leqq b$ then $\Sigma x_i \leqq ab$, all numbers occuring being supposed nonnegative.

We have then

$$g(n) \leqq g_1(n) + \frac{1}{2} \sum_{k>1} v_k^2(n) \leqq g_1(n) + \frac{1}{2}(n - v_1(n))\, v_1(n)$$

(cf.p. 447, footnote[1]). In view of (6.1) we get from here

$$g_1(n) \geqq \frac{1}{4} n^2 - \frac{1}{2}(n - v_1(n))\, v_1(n) - Cn = \tag{6.9}$$

$$= \left(\frac{1}{2} n - \left(\frac{1}{2} v_1(n) + C\right)\right)^2 + \frac{1}{2} v_1^2(n) - \left(\frac{1}{2} v_1(n) + C\right)^2$$

and finally

$$g_1(n) \geqq \frac{1}{4} v_1^2(n) - C v_1(n) - C^2. \tag{6.10}$$

It is evident from (6.9) that G_1 is infinite; this together with (6.10) means that G_1 has property $\mathcal{P}$ and the lemma is proved with $G' = G_1$.

To show that our theorem is best possible we have only to choose a sequence $n_1 < n_2 < \ldots$ of positive integers and consider the graph $G^{(\infty)}$ in which two vertices are connectid if and only if they belong to the same interval $[n_k, n_{k+1})$. Clearly $G^{(\infty)}$ does not contain any infinite path and if $A(n) \to \infty$ is given and the sequence is chosen to increase sufficiently fast then we clearly have $g(n) \geqq \dfrac{n^2}{4} - A(n)\, n$.

7. Proof of the first part of Theorem IV. We choose a sequence $l_0, l_1, l_2, \ldots$ of integers such that

$$l_0 = 0, \quad 5\, l_\nu < 2\, l_{\nu+1}, \quad \frac{l_\nu}{l_{\nu+1}} \to 0.$$

We put

$$\varphi_\nu(n) = \begin{cases} 0 & \text{if} \quad 1 \leqq n \leqq 2\, l_\nu, \\ 1 & \text{if} \quad 2\, l_\nu < n \leqq 3\, l_\nu, \\ 0 & \text{if} \quad 3\, l_\nu < n \leqq 5\, l_\nu, \\ 1 & \text{if} \quad 5\, l_\nu < n, \end{cases}$$

$$\varphi(n) = \sum_{\nu=1}^{\infty} \varphi_\nu(n),$$

and consider the graph $G^{(\infty)}$ in which two edges n_1 and n_2 $(n_1 < n_2)$ are connected if and only if $\varphi(n_1) \geqq \varphi(n_2)$.

Since $\varphi(n) > \nu$ if $n > 5 l_\nu$ we have $\lim \varphi(n) = \infty$. Consequently $G^{(\infty)}$ can not contain an I_∞-path. We shall show that on the other hand

$$\liminf \frac{g(n)}{n^2} \geqq \frac{1}{4} + \frac{1}{36}.$$

We estimate $g(n)$ from below if $2 l_\nu < n \leqq 2 l_{\nu+1}$. First we have $g(n) \geqq$ $\geqq g_\nu(n)$ where $g_\nu(n)$ is the number of the edges of $G^{(\infty)}$ whose endpoints belong to the interval $(5 l_{\nu-1}, 2 l_{\nu+1}]$. Now in this interval all functions $\varphi_\mu(n)$ except for $\mu = \nu$ are constant (namely $\varphi_\mu(n) = 1$ if $\mu < \nu$ and $\varphi_\mu(n) = 0$ if $\nu < \mu$) so that for $5l_{\nu-1} < n_1 < n_2 \leqq 2l_{\nu+1}$ n_1 and n_2 are connected if and only if $\varphi_\nu(n_1) \geqq \varphi_\nu(n_2)$. Using this remark we easily obtain

$$(7.1) \qquad g_\nu(n) = \begin{cases} 2\,l_\nu^2 + \frac{1}{2}(n - 2\,l_\nu)^2 + o(n^2) & (2\,l_\nu < n \leqq 3\,l_\nu)\,, \\ \frac{5}{2}\,l_\nu^2 + \frac{1}{2}(n - 3\,l_\nu)^2 + 3\,l_\nu(n - 3l_\nu) + o(n^2) & (3\,l_\nu < n \leqq 5\,l_\nu)\,, \\ \frac{21}{2}\,l_\nu^2 + \frac{1}{2}(n - 5\,l_\nu) + l_\nu(n - 5\,l_\nu) + o(n^2) & (5\,l_\nu < n \leqq 2\,l_{\nu+1})\,. \end{cases}$$

In these relations ν should be considered as function of n defined by the inequalities $2l_\nu < n \leqq 2l_{\nu+1}$. We obtain by a simple and elementary computation that

$$\frac{g(n)}{n^2} \geqq \frac{g_\nu(n)}{n^2} \geqq \min_{2l_\nu < m \leqq 2l_{\nu+1}} \frac{g_\nu(m)}{m^2} = \frac{5}{18} + o(1)$$

which completes our proof.

$\varphi(n)$ is an integer valued function which assumes each value on a finite number of places. We shall somewhat modify $\varphi(n)$ by introducing a function $\varphi'(n)$ in the following way: If k is any value of $\varphi(n)$ assumed for $n_1, n_2, \ldots, n_\varrho$ $(n_1 < n_2 < \ldots < n_\varrho)$ then we put

$$\varphi'(n_\nu) = k + \frac{1}{\nu} \qquad (\nu = 1, 2, \ldots, \varrho)\,.$$

It is clear that φ' is schlicht and for any two positive integers n' and n'' $\varphi(n') \geqq \varphi(n'')$ is equivalent to $\varphi'(n') > \varphi'(n'')$. The range of φ' is an infinite set of positive numbers without a limit point, hence it can be mapped by a strictly increasing function ψ onto the set of natural numbers. Thus $\varkappa = \psi \circ \varphi'$ is a permutation of the set of natural numbers for which n' and n'' are connected in $G^{(\infty)}$ if and only if $(n' - n'')\,(\varkappa(n') - \varkappa(n'')) < 0$. So we can state the somewhat paradoxical fact that the positive integers can be rearranged in a series $k_1, k_2, \ldots$ in such a drastic way that the number of inversions devided by the number of all unordered pairs formed by $k_1, k_2, \ldots, k_n$ is more than a half plus a fix positive number for all sufficiently large n.

8. Proof of the second part of Theorem IV. We suppose that $G^{(\infty)}$ does not contain any I_∞-path and we prove that then

$$(8.1) \qquad \liminf \frac{g(n)}{n^2} \leqq \frac{1}{2} - \frac{1}{16}\,.$$

This means that the second part of Theorem IV is valid with $\alpha = \frac{1}{16}$ (although we do not know the greatest possible value of α).

Omitting from $G^{(\infty)}$ all edges (n, m) with $n^2 < m$ we get a subgraph $\bar{G}^{(\infty)}$ every vertex of which has finite valency[2] and for which $g(n) \geqq \bar{g}(n) \geqq \geqq g(n) - Cn^{3/2}$, consequently

$$\liminf \frac{\bar{g}(n)}{n^2} = \liminf \frac{g(n)}{n^2}.$$

This means that we can suppose without loss of generality that every vertex of $G^{(\infty)}$ has finite valency.

We define by induction the sets $A_0, A_1, A_2, \ldots$ requiring that $A_0 = 0$ and for $k > 0$ A_k is the set of those $n \in N - \bigcup_{l=0}^{k-1} A_k$ which are not connected with any m if $n < m$ and $m \in N - \bigcup_{l=0}^{k-1} A_l$. ($N$ is the set of natural numbers.) The sets A_k exhaust N:

$$\bigcup_{0}^{\infty} A_k = N. \tag{8.1}$$

To prove this suppose that for some n_1 $n_1 \in N - \bigcup_{0}^{\infty} A_k$. It is clear from the definition of the sets A_k that for every $k > 0$ n_1 is the starting point of an I_k-path. Since n_1 has finite valency an infinite number of these I_k-paths must have the same second vertex that is there is an edge (n_1, n_2) where $n_2 > n_1$ and n_2 is the starting point of I_k-paths for arbitrarily large k's. The repetition of this argument clearly yields an I_∞-path $n_1 n_2 n_3 \ldots$ against our assumption which proves (8.2).

Put

$$B_k = \bigcup_{l=0}^{k} A_l, \quad B_k(n) = B_k \cap [1, n]$$

and denote by β_k the upper density of B_k that is

$$\beta_k = \limsup \frac{b_k(n)}{n}$$

where $b_k(n)$ is the number of elements of $B_k(n)$. Suppose first that

$$\textit{for some } k \ \beta_k \geqq \frac{1}{2}. \tag{8.3}$$

Denoting by k_0 the least of these k we have

$$\beta_{k_0} \geqq \frac{1}{2}, \quad k_0 > 0$$

(since $\beta_0 = 0$) and

$$\beta_{k_0 - 1} < \frac{1}{2}. \tag{8.4}$$

Given a natural number n_0 and $\varepsilon > 0$ we can choose an $n > n_0$ such that

$$b_{k_0}(n) > \left(\frac{1}{2} - \varepsilon\right) n, \quad b_{k_0 - 1}(2n) < n. \tag{8.5}$$

[2] The *valency* of a vertex is the number of edges emanating from this vertex.

In view of (8.4) we can not have for every integer $\nu \geqq 0$

$$b_{k_0-1}(2^{\nu+1} n) - b_{k_0-1}(2^\nu n) \geqq 2^{\nu-1} n\,,$$

since this would imply by addition

$$b_{k_0-1}(2^{\nu+1} n) - b_{k_0-1}(n) \geqq \left(2^\nu - \frac{1}{2}\right) n \qquad (\nu = 0, 1, 2, \ldots)$$

which would give $\beta_{k_0-1} \geqq \frac{1}{2}$. Hence there is an integer $\nu_0 \geqq 0$ such that

$$b_{k_0-1}(2^{\nu_0+1} n) - b_{k_0-1}(2^{\nu_0} n) < 2^{\nu_0-1} n \tag{8.6}$$

and

$$b_{k_0-1}(2^{\nu+1} n) - b_{k_0-1}(2^\nu n) \geqq 2^{\nu-1} n \text{ for } 0 \leqq \nu < \nu_0\,. \tag{8.7}$$

Putting $2^{\nu_0} n = m$ we get from (8.5), (8.6) and (8.7)

$$\begin{aligned} b_{k_0}(m) &= b_{k_0}(2^{\nu_0} n) = b_{k_0}(n) + \sum_{\nu=0}^{\nu_0-1} \left(b_{k_0}(2^{\nu+1} n) - b_{k_0}(2^\nu n)\right) > \\ &> \left(\frac{1}{2} - \varepsilon\right) n + n \sum_{\nu=0}^{\nu_0-1} 2^{\nu-1} = (2^{\nu_0-1} - \varepsilon)\, n \geqq \left(\frac{1}{2} - \varepsilon\right) m \end{aligned} \tag{8.8}$$

and

$$b_{k_0-1}(2\,m) - b_{k_0-1}(m) < \frac{1}{2}\, m\,. \tag{8.9}$$

(If $\nu_0 = 0$ then the sums figuring in (8.8) are void.)

It is clear from the definition of the sets A_k that if $u \in A_k$, $v \in A_l$ and $u < v$, $k \leqq l$ then u and v are not connected in $G^{(\infty)}$. Consequently there is no edge connecting a member of $B_{k_0}(m)$ with a member of $(m, 2m] - B_{k_0-1}$. Now the first of these sets has $b_{k_0}(m)$ members and the second one $m - (b_{k_0-1}(2m) - b_{k_0-1}(m))$ members hence using (8.8) and (8.9)

$$\begin{aligned} g(2\,m) &\leqq \binom{m}{2} - b_{k_0}(m)\left(m - \left(b_{k_0-1}(2\,m) - b_{k_0-1}(m)\right)\right) < \\ &< \frac{4\,m^2}{2} - \left(\frac{1}{2} - \varepsilon\right)\frac{1}{2}\, m^2 = \frac{4\,m^2}{2} - \frac{1}{16}\, 4\, m^2 + \frac{\varepsilon}{4}\, m^2\,. \end{aligned}$$

Since $2m > m \geqq n > n_0$ and n_0 and ε have been chosen arbitrary, (8.1) is proved under the assumption (8.3)

Next we consider the case that

$$\beta_k < \frac{1}{2} \quad \text{for} \quad k = 0, 1, 2, \ldots\,. \tag{8.10}$$

Given $n_0 > 0$ we can choose k_0 so that $B_{k_0}(n_0) = [1, n_0]$ (cf. (8.2)) i.e.

$$b_{k_0}(n_0) = n_0. \tag{8.11}$$

Let us denote by ν_0 the least non negative integer for which

$$b_{k_0}(2^{\nu_0} n_0) \leqq 2^{\nu_0 - 1} n_0 .$$

ν_0 exists since otherwise we would have $\beta_{k_0} \geqq \frac{1}{2}$ which would contradict (8.10), (8.11) implies that $v_0 > 0$, hence

$$b_k(2^{\nu_0-1} n_0) > 2^{\nu_0-2} n_0 .$$

Putting $2^{\nu_0-1} n_0 = m$, we have

$$b_{k_0}(2m) \leqq m \quad \text{and} \quad b_{k_0}(m) > \frac{1}{2} m . \tag{8.9}$$

The members of $B_{k_0}(m)$ are not connected with those of $(m, 2m] - - B_{k_0}(m)$. Using (8.9) it follows

$$g(2m) \leqq \binom{2m}{2} - b_{k_0}(m)\left(m - \left(b_{k_0}(2m) - b_{k_0}(m)\right)\right) \leqq$$

$$\leqq \frac{4m^2}{2} - \frac{1}{2} m\left(m - \left(m - \frac{1}{2} m\right)\right) = \frac{4m^2}{2} - \frac{4m^2}{16} .$$

Since $m \geqq n_0$ (8.1) is proved under the assumption (8.10) also.

(Received October 12, 1962)

REFERENCES

[1] Erdős, P. and Gallai, T.: "On maximal paths and circuits of graphs." *Acta Math Academiae Sci. Hung.* **10** (1959) 337–356.
[2] Turán, P.: "On the theory of graphs." *Colloquium Math.* **3** (1954) 19–30.

ЭКСТРЕМАЛЬНЫЕ ПРОБЛЕМЫ ОТНОСИТЕЛЬНО БЕСКОНЕЧНЫХ ГРАФОВ

J. CZIPSZER, P. ERDŐS, A. HAJNAL

Резюме

Пусть $G^{(\infty)}$ есть граф, вершины которого суть натуральные числа $1, 2, \ldots, n, \ldots$. Для каждого n обозначим через $g(n)$ число тех ребер графа $G^{(\infty)}$, вершины которых находятся среди чисел $1, 2, \ldots, n$. Лежащий в $G^{(\infty)}$ путь $n_1 n_2 \ldots n_k n_{k+1}$ называется монотонным путем длины k или I_k-путем, если $n_1 < n_2 < \ldots < n_k < n_{k+1}$. Лежащий в $G^{(\infty)}$ бесконечный путь $n_1 n_2 \ldots n_k \ldots$ называется бесконечным монотонным путем или I_∞-путем, если $n_1 < n_2 < \ldots < n_k < \ldots$. Нижеследующие теоремы показывают, как можно с помощью условый относительно порядка роста $g(n)$ гарантировать существование в $G^{(\infty)}$ бесконечного числа I_2-путей, бесконечного числа I_3-путей, бес конечного пути, I_∞-пути или бесконечного полного подграфа.

Теорема 1. *Если* $k = 2$ *или* 3 *и для некоторого* $\varepsilon > 0$ *и достаточно больших* n

$$g(n) > \left(\frac{1}{4} - \frac{1}{4k} + \varepsilon\right) n^2,$$

то $G^{(\infty)}$ *содержит бесконечно много* I_k*-путей. Это утверждение точно в том смысле, что* εn^2 *не может быть заменено на* $O(n)$.

Пока не известно, имеет ли место теорема и при $k > 3$.

Теорема 2. *Если для некоторого* $\varepsilon > 0$ *и достаточно больших* n

$$g(n) > \frac{n^2}{8} + \left(\frac{1}{32} + \varepsilon\right) \frac{n^2}{\log^2 n}$$

то $G^{(\infty)}$ *содержит бесконечно много* I_2*-путей. Здесь* $\frac{1}{32}$ *не может быть уменьшено.*

Теорема 3. *Если для всех* n

$$g(n) > \frac{1}{4} n^2 - Cn$$

то $G^{(\infty)}$ *содержит бесконечный путь. Здесь* C *не может быть заменено на* A_n, *если* $A_n \to \infty$.

Теорема 4. *Из того, что*

$$\liminf \frac{g(n)}{n^2} > \frac{1}{4},$$

еще не следует, что $G^{(\infty)}$ *содержит* I_∞*-путь. Но существует такая постоянная* $0 < \alpha < \frac{1}{4}$, что если

$$\liminf \frac{g(n)}{n^2} > \frac{1}{2} - \alpha,$$

то $G^{(\infty)}$ *содержит* I_∞*-путь.*

Теорема 5. *Если для бесконечно многих* n

$$g(n) > \frac{n^2}{2} - Cn,$$

то $G^{(\infty)}$ *содержит бесконечный полный подграф. Это утверждение точное в том смысле, что вместо* C *нельзя писать* A_n, *если* $A_n \to \infty$.