ON MAXIMAL PATHS AND CIRCUITS OF GRAPHS

By

P. ERDŐS (Budapest), corresponding member of the Academy,
and T. GALLAI (Budapest)

Introduction

In 1940 TURÁN raised the following question: if the number of nodes, n, of a graph[1] is prescribed and if l is an integer $\leqq n$, what is the number of edges which the graph has to contain in order to ensure that it necessarily contains a complete l-graph? TURÁN gave a precise answer to this question by determining the smallest number depending on n and l, with the property that a graph with n nodes and with more edges than this number necessarily contains a complete l-graph ([9], [10]). More generally, the question can be posed, as was done by TURÁN: given a graph with a prescribed number of nodes, what is the minimum number of edges which ensures that the graph necessarily contains a "sufficiently large" subgraph of a certain prescribed type? An alternative formulation of this question is as follows: the number of nodes being fixed, we seek the maximum value of μ, μ being such that there exists a graph with μ edges which does not contain a subgraph of the type in question with more than a certain given number of nodes. In our paper we are concerned with this problem for the case in which the types of graphs considered are paths, circuits and independent edges. (These terms are defined in § 1.)

Our results are not exhaustive, because, in general, we only give an estimate of the extremal values, only in isolated cases — for certain special values of the number of the nodes — do we succeed in determining the extreme values and the "extreme" graphs completely. Here are some of our results capable of simple formulation:

Every graph with n nodes and more than $(n-1)l/2$ *edges* $(l \geqq 2)$ *contains a circuit with more than* l *edges.* The value $(n-1)l/2$ is exact if and only if $n = q(l-1)+1$, then there exists a graph having n nodes and

[1] The graphs considered in this paper are all finite, every edge has two distinct end-nodes, and any two nodes are joined by at most one edge.

l. c. letters always denote non-negative integers, n always denotes an integer $\geqq 1$. A *complete l-graph* is a graph with l nodes, every pair of distinct nodes joined by an edge. A graph is said to contain its subgraphs. (See § 1 of this paper and [6], pp. 1—3.)

$(n-1)l/2$ edges which contains no circuit with more than l edges. (Theorem (2. 7).)

For all $n \geqq (k+1)^3/2$, $k \geqq 1$, *every graph with* n *nodes and more than* $nk-k(k+1)/2$ *edges contains a path or a circuit with more than* $2k$ *edges.* The value $nk-k(k+1)/2$ is exact. (Theorem (3. 6).)

For our proofs we need a group of theorems different from the above, which are of interest in their own right. In these it is not the number of edges, but the fact that every node has a sufficiently high degree (the degree of a node is the number of edges incident with the node), which implies the existence of a "sufficiently large" subgraph of a prescribed type. This class of problems was first considered by ZARANKIEWICZ [11] and DIRAC [3]. From among these older results we require two theorems due to DIRAC ([3], Theorems 3 and 4), for which we give new simple proofs in § 1. (A simple proof of Theorem 3 can be found in [8].) We also prove some new theorems of the type just now discussed.

In § 1 we present the necessary preliminary notions and some lemmas, and we prove the theorems pertaining to the ZARANKIEWICZ—DIRAC field of problems. In § 2 we carry out the estimations connected with problems of the TURÁN type. In § 3 we determine two extremal values exactly for a sufficiently large number of nodes. In § 4 we determine the maximum number of edges in a graph of n nodes and at most k independent edges. We distinguish our more important results from the less interesting assertions leading to them by the appellation "Theorem".

§ 1

(1. 1) Let $M=\{P_1, \ldots, P_n\}$ be a finite non-empty set and let the set of all unordered pairs of distinct elements $\overline{P_iP_j}=\overline{P_jP_i}$ $(i \neq j)$ of M be denoted by N. (If $n=1$, then N is the empty set.) The elements of M are called *nodes*, the elements of N are called *edges*, and the edge $\overline{P_iP_j}$ is said to be incident with the nodes P_i and P_j. Let N_1 be an arbitrary subset of N, M and N_1 are said to define a *graph* $\Gamma=(M, N_1)$. The elements of M and N_1, respectively, are the nodes and edges of Γ. If $\overline{P_iP_j} \in N_1$, then P_i and P_j are said to be *joined* (in Γ), or we say that the edge $\overline{P_iP_j}$ exists (in Γ).

The graph $\Gamma=(M, N)$ is called *complete*, more exactly a *complete n-graph*. The graph $\overline{\Gamma}=(M, N-N_1)$ is the *complement* of the graph $\Gamma=(M, N_1)$.

Let N' denote the set of pairs of elements of the finite set M' and let $N_1' \subseteq N'$. The graph $\Gamma'=(M', N_1')$ is called a *subgraph* of the graph $\Gamma=(M, N_1)$ if $M' \subseteq M$ and $N_1' \subseteq N_1$. We also say that Γ *contains* Γ' and that Γ' is in Γ. If Γ' is a subgraph of Γ, then the graph $[\Gamma'] = (M', N' \cap N_1)$ is called the subgraph (of Γ) *spanned by Γ' or by M' in Γ*. If P is a node of the graph Γ, then $\Gamma - P$ denotes the graph obtained by deleting P and all edges incident with P from Γ.

(1.2) If a (1—1) correspondence can be established between the nodes of the graphs Γ_1 and Γ_2 so that nodes joined in one graph correspond to nodes which are joined in the other and, conversely, then the two graphs are regarded as identical, and this is expressed in symbols by $\Gamma_1 = \Gamma_2$.

The number of nodes and edges in the graph Γ is denoted by $\pi(\Gamma)$ and $\nu(\Gamma)$, respectively.

The number of edges incident with the node P in the graph Γ is called the *degree of P in Γ*. If there is no room for misunderstanding, then we speak of the degree of P for short, and denote it by $\varrho(P)$. If $\varrho(P) = 0$, then we call P an *isolated* node of Γ, if $\varrho(P) = 1$, we call P a *terminal* node of Γ.

A graph L is called a *loop*, more accurately a *P-loop*, if a series[2] $P_1, \ldots, P_n, P_{n+1}$ $(n \geqq 1)$ can be constructed from the nodes of L so that every node of L appears in the series, $P = P_1$, the nodes $P_1, \ldots, P_n$ are all distinct, $P_{n+1} \neq P_n$, and if $n > 1$, then $P_{n+1} \neq P_{n-1}$, and the set of edges of L consists of $\overline{P_i P_{i+1}}$ $(i = 1, \ldots, n)$. It is easy to see that we can form in at most two ways from the nodes of a P-loop a sequence of the required properties and that P_{n+1} is uniquely determined. P is the *initial node* of the loop and P_{n+1} the *final node*. The loop is also said to *start from* P and to *lead to* P_{n+1}. We call the P-loop *directed* if one of the above-mentioned sequences is made to correspond to it.

If P_{n+1} is different from $P_1, \ldots, P_n$, then the loop is called a *path*, more accurately a P_1P_{n+1}-path, if $P_{n+1} = P_1$, then it is called a *circuit*. Paths and circuits will be designated by the common term *arc*. If $P_{n+1} = P_j$ $(1 \leqq j \leqq n-2)$, then the nodes $P_j, P_{j+1}, \ldots, P_{n+1}$ and the edges $\overline{P_i P_{i+1}}$ $(i = j, \ldots, n)$ together form a circuit which is called the *circuit of the loop L*. The number of edges of L is called the *length* of L. Paths, circuits and arcs of length l are called *l-paths, l-gons* and *l-arcs*, respectively. The equation

$$L = (P_1, \ldots, P_n, P_{n+1} = P_j) \qquad (1 \leqq j \leqq n-2)$$

states that the graph L is a P_1-loop wich is composed of the nodes $P_1, \ldots, P_n$ and the edges $\overline{P_i P_{i+1}}$ $(i = 1, \ldots, n)$. The equation

$$W = (P_1, \ldots, P_{n+1})$$

[2] If j and g are natural numbers and $j < g$ or $j > g$, then $P_j, \ldots, P_g$ denotes the set of nodes P_i where i runs through the natural numbers from j to g. If $j = g$, then $P_j, \ldots, P_g$ means P_j by itself.

states that W is a P_1P_{n+1}-path composed of the nodes $P_1, \dots, P_{n+1}$ and the edges $\overline{P_iP_{i+1}}$ $(i = 1, \dots, n)$.

A set of edges of Γ is called *independent* if no two of them have a node in common. We shall say that the maximum number of independent edges is k, if there exists a set of k such edges and there does not exist a set of $k+1$ such edges.

(1.3) The graph Γ is *connected* if it consists of a single node or if corresponding to any two distinct nodes P and Q Γ contains a PQ-path. The "maximal" connected subgraphs of Γ are called its *components*. The subgraph Γ' of the graph Γ is *maximal* with respect to some property if Γ' contains no subgraph with this property of which Γ' is a proper subgraph. If Γ' is a component of Γ, then $\Gamma - \Gamma'$ denotes the graph obtained from Γ by deleting Γ'.

The nodes $P_1, \dots, P_j$ $(j \geqq 1)$ are said to *separate* the two (distinct) nodes A and B in the connected graph Γ if $P_i \neq A$, $P_i \neq B$ $(i = 1, \dots, j)$ and every AB-path in Γ contains at least one of the nodes $P_1, \dots, P_j$. The nodes $P_1, \dots, P_j$ *divide* the connected graph Γ if Γ contains two nodes which they separate.

The graph Γ is *n-fold connected* $(n \geqq 2)$ if it is connected and if no set of fewer than n nodes divides it. A complete l-graph is said to be n-fold connected for all n.

The maximal twofold connected subgraphs of the connected graph Γ are called the *members* of Γ. Every edge of Γ is an edge of some member and every member, except for the graph consisting of a single node, contains more than one node. If Γ is connected but not twofold connected, then it has more than one member, and it may be verified ([6], pp. 224—231) that two of its members have at most one node in common and that such a node divides Γ.[3] Furthermore, it may easily be verified that Γ has at least two members containing only one cut-node each. Such members are called *terminal members* of Γ. If Γ' is a terminal member of Γ, then $\Gamma - \Gamma'$ denotes the graph obtained from Γ by deleting all edges of Γ' and all nodes of Γ' except its cut-node.

(1.4) *If in the graph Γ every node except the single node A has degree $\geqq k$ $(k \geqq 2)$ and $\pi(\Gamma) \leqq 2k$, then, if $\varrho(A) \geqq 2$, Γ is twofold connected and if $\varrho(A) = 1$, then $\Gamma - A$ is twofold connected and Γ is connected.*

PROOF. Suppose first that $\varrho(A) \geqq 2$. Then it follows from our assumptions that every component of Γ and every terminal member of Γ (if any) contains at least 3 nodes, and at least two of these have degree $\geqq k$. If Γ

[3] Such a node is called a *cut-node* of the graph.

is not twofold connected, then it has two components or two terminal members. At least one of these components or terminal members has not more than k nodes. The nodes in it with at most one exception have degree $<k$, which is a contradiction.

If $\varrho(A)=1$, then let the edge incident with A be denoted by $\overline{AA'}$. If $k \geqq 3$, then the degree of A' in $\Gamma - A$ is clearly at least 2. If $k=2$, then only $\pi(\Gamma)=4$ is possible, in which case the degree of A' in $\Gamma - A$ is 2. In both cases the theorem with $\varrho(A) \geqq 2$ can be applied to $\Gamma - A$. Γ is obviously connected.

(1.5) A circuit of the graph Γ which contains all nodes of Γ is called a *Hamiltonian line* of Γ, *H-line* for short. A path of Γ which contains all nodes of Γ is called an *open H-line* of Γ. Two distinct nodes P and Q of Γ are said to be *H-independent* in Γ if Γ contains no open H-line starting in P and ending in Q.

Our later reasoning is based on the following lemma:

LEMMA (1.6) *If the circuit $C=(P_1, \ldots, P_n, P_{n+1}=P_1)$ is an H-line of the graph Γ and if P_i and P_j $(i,j \neq n+1)$ are H-independent nodes of Γ, then*

$$\varrho(P_{i+1})+\varrho(P_{j+1}) \leqq \pi(C)=n;$$

$\varrho(P)$ denotes the degree of the node P in Γ.

PROOF. It may be assumed that $i=1$ and $1<j \leqq n$. Because two neighbouring nodes of C are not H-independent, $3 \leqq j \leqq n-1$. Accordingly, $n \geqq 4$ and the nodes P_1, P_2, P_j, P_{j+1} are distinct.

If Γ contains the edge $\overline{P_2 P_g}$ $(3 \leqq g \leqq j)$, then it does not contain the edge $\overline{P_{j+1} P_{g-1}}$. For if this edge belonged to Γ, then the path

$$W=(P_1, P_n, \ldots, P_{j+1}, P_{g-1}, \ldots, P_2, P_g, \ldots, P_j)$$

would be an open H-line of Γ starting in P_1 and ending in P_j.

It follows that $\overline{P_2 P_{j+1}}$ does not exist in Γ, since $\overline{P_2 P_3}$ does.

If the edge $\overline{P_2 P_l}$ $(j+2 \leqq l \leqq n)$ exists in Γ, then the edge $\overline{P_{j+1} P_{l+1}}$ cannot exist in Γ. For if this edge existed, then the path

$$W'=(P_j, \ldots, P_2, P_l, \ldots, P_{j+1}, P_{l+1}, \ldots, P_{n+1})$$

would be an open H-line of Γ starting in P_j and ending in $P_{n+1}=P_1$.

Accordingly, with every node, other than P_1, joined to P_2 in Γ there can be associated a node not joined to P_{j+1} in such a way that these associated nodes are all distinct. It follows that the number of nodes not joined to P_{j+1} is at least $\varrho(P_2)-1$, so that $\varrho(P_{j+1}) \leqq (n-1)-(\varrho(P_2)-1)=n-\varrho(P_2)$.

NOTE. It follows from Lemma (1. 6) and its proof that, under the hypotheses of the lemma, $\varrho(P_{i-1})+\varrho(P_{j-1})\leqq n$ and the edges $\overline{P_{i+1}P_{j+1}}$ and $\overline{P_{i-1}P_{j-1}}$ do not exist in Γ ($P_0=P_n$).

Let the loop $L=(P_1,\ldots,P_n,P_{n+1}=P_m)$ $(1\leqq m\leqq n-2)$ be a subgraph of the graph Γ. A node P_j belonging to the circuit $C=(P_m,\ldots,P_n,P_{n+1}=P_m)$ of the loop L and different from the terminal node P_m of L is called *H-node of the loop* L (with respect to Γ) if Γ contains a P_jP_m-path containing all the nodes of C and no other nodes, or, otherwise expressed, if the graph $[C]$ spanned by C in Γ contains an open H-line starting in P_j and ending in P_m.

(1. 7) *Let* $L=(P_1,\ldots,P_n,P_{n+1}=P_m)$ $(1\leqq m\leqq n-2)$ *be a loop of the graph* Γ *and let the circuit of* L *be denoted by* C. *If every H-node of* L *(with respect to* Γ*) has degree* $\geqq k$ $(k\geqq 2)$ *in* $[C]$ *and if* $\pi(C)\leqq 2k-1$, *then every node of* C *different from* P_m *is an H-node of* L.

PROOF. P_{m+1} and P_n are clearly H-nodes. Our theorem is established if we prove the following assertion: If P_{j+1} $(m<j<n)$ is an H-node, then so is P_j. To see that this is so, let it be assumed that P_{j+1} $(m<j<n)$ is an H-node and P_j is not. Then P_m and P_j are H-independent in $[C]$, and, since C is an H-line of the graph $[C]$, it follows from Lemma (1. 6) that $\varrho'(P_{m+1})+$ $+\varrho'(P_{j+1})\leqq\pi(C)$, where $\varrho'(P_i)$ denotes the degree of the node P_i in $[C]$. But it was assumed that $\varrho'(P_{j+1})\geqq k$ and $\varrho'(P_{m+1})\geqq k$. We have a contradiction!

Let A denote a node of degree $\geqq 1$ of the graph Γ and let the degree of all nodes of Γ other than A be $\geqq 2$. If $W=(P_1,\ldots,P_n)$ $(P_1=A)$ is any path of Γ which starts from A, then, the degree of P_n being $\geqq 2$, Γ contains an edge $\overline{P_nP_{n+1}}$ incident with P_n and different from $\overline{P_nP_{n-1}}$. This edge and W together form an A-loop which is longer than W. Thus to every path W starting in A there exists an A-loop longer than W, hence the longest A-loops of Γ are not path (i. e. they contain circuits).

These longest A-loops of Γ which possess the longest circuits will be called *maximal A-loops*.

(1. 8) *Let* A *be a node of the graph* Γ *with degree* $\geqq 1$ *and suppose that the degree of every node of* Γ *other than* A *is* $\geqq k$ *where* $k\geqq 2$. *Further, let* L *be a maximal A-loop and let the terminal node of* L *be denoted by* B *and its circuit by* C. *Then if* $\pi(C)\leqq 2k-1$, *it follows that* $[C]$ *is either a component or a terminal member of* Γ *and in the latter case the cut-node of* $[C]$ *is* B. *In both cases every node of* $[C]$ *distinct from* B *is connected to* B *by an open H-line of* $[C]$ *and* $\pi(C)\geqq k+1$.

PROOF. It follows from the assumptions made that Γ contains a maximal A-loop. Let $L=(P_1,\ldots,P_n,P_{n+1}=P_m)$ $(1\leqq m\leqq n-2, P_1=A, P_m=B)$

be the maximal A-loop concerned. If P_j $(m<j\leqq n)$ is any H-node of L, then P_j is not joined by an edge in Γ to any node which is not in C. For if $\overline{P_jP}$ is an edge of Γ and if W denotes an open H-line of $[C]$ leading from P_j to P_m, then if we add the edge $\overline{P_jP}$ and the path $(P_1, \dots, P_m)$ to W the result is an A-loop which is as long as L and which is a path if $P\neq P_i$ $(i=1,\dots,n)$ and which has a longer circuit than C if $P=P_g$ $(1\leqq g<m)$. A contradiction with the maximal nature of L is therefore avoided only if $P=P_g$ $(m\leqq g\leqq n)$. It follows that the degree of every H-node of L in $[C]$ is $\geqq k$, therefore if $\pi(C)\leqq 2k-1$, then according to (1.7) every node of C other than B is an H-node of L. There follows firstly the existence of the open H-lines asserted in the theorem and secondly that every node of C other than B is joined exclusively to nodes of C. The latter fact implies that $[C]$ is a component or a terminal member with cut-node B. ($[C]$ is obviously twofold connected because of C.) $\pi(C)\geqq k+1$, because $\varrho(P_n)\geqq k$ and P_n is joined exclusively to nodes of C.

If the graph Γ of Theorem (1.8) has at most $2k-1$ nodes, then it follows from Theorem (1.4) that $[C]=\Gamma$ or $[C]=\Gamma-A$ according as $\varrho(A)\geqq 2$ or $\varrho(A)=1$ and that, if $\pi(\Gamma)=2k$ and $\varrho(A)\geqq 2$, then $\pi(C)=2k$.

The following two theorems can be deduced:

THEOREM (1.9) *If the node A of the graph Γ is not isolated and the degree of every node of Γ distinct from A is $\geqq k$ $(k\geqq 2)$ and if $\pi(\Gamma)\leqq 2k-1$, then A is connected by open H-lines to every node of Γ.*

THEOREM (1.10) (DIRAC) *If every node of the graph Γ has degree $\geqq k$ $(k\geqq 2)$, and if $\pi(\Gamma)\leqq 2k$, then Γ has an H-line.*

(1.9) obviously implies the following theorem:

THEOREM (1.11) *If every node of the graph Γ has degree $\geqq k$ $(k\geqq 2)$, and if $\pi(\Gamma)\leqq 2k-1$, then any two distinct nodes are connected by an open H-line.*

If the graph Γ of Theorem (1.8) is twofold connected and if $\pi(\Gamma)\geqq$ $\geqq 2k$, then $\pi(C)\geqq 2k$. From this it follows that the node A of Theorem (1.8) is the initial node of a path of length $\geqq 2k-1$.

The following two theorems can be deduced:

THEOREM (1.12) *If Γ is a twofold connected graph and every node with the exception of one single node A has degree $\geqq k$ $(k\geqq 1)$ and if in addition $\pi(\Gamma)\geqq 2k$, then Γ contains a path with at least $2k-1$ edges which starts from A.*

(This theorem is trivial for $k=1$.)

THEOREM (1. 13) (DIRAC) *If the degree of every node of the twofold connected graph* Γ *is* $\geqq k$ $(k \geqq 2)$ *and if* $\pi(\Gamma) \geqq 2k$, *then* Γ *contains a circuit with at least* $2k$ *edges.*

REMARK. It follows from the above considerations that the assertions in Theorems (1. 10) and (1. 13) remain true if the degree of every node except one node A is at least k $(k \geqq 2)$ and $\varrho(A) \geqq 2$. If two nodes have degree $< k$, then these theorems are not generally true.

THEOREM (1. 14) *If* Γ *is a connected graph and the degree of every one of its nodes is* $\geqq k$ $(k \geqq 1)$ *and if* $\pi(\Gamma) \geqq 2k+1$, *then* Γ *contains a path with* $2k$ *or more edges.*[4]

PROOF. If $k = 1$, the theorem is trivial. In what follows it will be assumed that $k \geqq 2$.

If Γ is twofold connected, then, by (1. 13), Γ contains a circuit $C = (P_1, \ldots, P_m, P_{m+1} = P_1)$ $(m \geqq 2k)$. If $m > 2k$, $W = (P_1, \ldots, P_m)$ is a path of the required kind. If $m = 2k$, then, by our assumptions, Γ contains a node P which is not in C and which is joined to a node of C, say P_1. Then $W = (P, P_1, \ldots, P_m)$ is a path of the required kind.

If Γ is not twofold connected, then let Γ_1 and Γ_2 denote two terminal members of Γ, and A_1 and A_2 their cut-nodes. Γ_1 and Γ_2 are twofold connected and apart from A_1 and A_2 their nodes have degree $\geqq k$ in Γ_1 and Γ_2, respectively. This is possible only if $\pi(\Gamma_1) \geqq k+1$ and $\pi(\Gamma_2) \geqq k+1$. From this and from (1. 9) and (1. 12) it follows that Γ_1 contains a path of length $\geqq k$ which starts in A_1 and Γ_2 contains a path of length $\geqq k$ which starts in A_2. If $A_1 = A_2$, then these two paths together constitute a path with at least $2k$ edges, and if $A_1 \neq A_2$, then these two paths together with an A_1A_2-path of $\Gamma - \Gamma_1 - \Gamma_2$ constitute a path with more than $2k$ edges.

REMARK. Theorem (1. 14) can be proved easily whithout using the preceding theorems.

(1. 15) The "accuracy" of Theorems (1. 13) and (1. 14) is demonstrated by the following graph Γ:

Γ consists of the nodes $P_1, \ldots, P_k, Q_1, \ldots, Q_{n-k}$ $(2 \leqq k \leqq n-k)$ and of all edges $\overline{P_iQ_j}$ $(i = 1, \ldots, k; j = 1, \ldots, n-k)$. Γ is an *even* graph ([6], p. 170). The degree of every node of Γ is $\geqq k$ and it may easily be verified that Γ is k-fold connected, further that the longest circuits and paths in Γ have $2k$ edges.

It is seen from the example of this graph that in the theorems in ques-

[4] This result was obtained independently by G. A. DIRAC.

tion it is not possible to assert the existence of longer paths and circuits than those proved to exist, even if the connectivity is assumed to be higher.

(1. 16) Using an altered form of Theorem (1. 8) and MENGER's well-known "n-path theorem" the following theorems can be established:

If the degree of every node of the twofold connected graph Γ is $\geqq k$ $(k \geqq 2)$ and if $\pi(\Gamma) \geqq 2k$, then through each node of Γ there passes a circuit having at least $2k$ edges.

If the degree of every node of the twofold connected graph Γ with the exception of the two nodes A and B is $\geqq k$ $(k \geqq 2)$, then every node of Γ lies on an AB-path having at least k edges.

These theorems are not proved in this paper.

§ 2

(2. 1) Let the classes of all graphs containing exactly n nodes and containing, respectively, no paths, circuits, arcs with more than l edges $(l \geqq 1)$ be denoted by $F(n, l)$, $G(n, l)$, $H(n, l)$. The graphs in each class which contain the most edges are called the *extreme* graphs of the class concerned, and the number of edges in these graphs will be denoted by $f(n, l)$, $g(n, l)$ and $h(n, l)$, respectively. So if the graph Γ is a member of, respectively, $F(n, l)$, $G(n, l)$, $H(n, l)$, then the following inequalities hold:

$$(*) \qquad \nu(\Gamma) \leqq f(n, l), \quad \nu(\Gamma) \leqq g(n, l), \quad \nu(\Gamma) \leqq h(n, l),$$

and if Γ is an extreme graph of the class concerned, then equality holds under $(*)$.

We wish to estimate or determine f, g and h and to find the extreme graphs.

Clearly, if $n \leqq l$ $(l \geqq 2)$, the only extreme graphs of $F(n, l-1)$, $G(n, l)$, $H(n, l)$ are the complete n-graphs.

(2. 2) Our method of estimating the values in question from below is to construct graphs belonging to the classes F, G, H and containing as many edges as possible.

In this paper Γ_n^{2k} $(1 \leqq k < n)$ denotes the graph which consists of the nodes $P_1, \ldots, P_k, Q_1, \ldots, Q_{n-k}$ together with all the edges $\overline{P_i P_j}$ $(i, j = 1, \ldots, k;$ $i \neq j)$ and all the edges $\overline{P_i Q_j}$ $(i = 1, \ldots, k; j = 1, \ldots, n-k)$.

If $1 \leqq k < n-1$, then Γ_n^{2k+1} denotes the graph obtained from the graph Γ_n^{2k} by the addition of the edge $\overline{Q_1 Q_2}$. Accordingly, the graph Γ_n^l is defined only for $l \geqq 2$ and $n > [(l+1)/2]$. In what follows the graphs Γ_n^2 will be called *stars* and the graphs Γ_n^3 will be called *Δ-stars* for all values of n.

With the notation

$$\varphi(n, 2k) = \binom{k}{2} + (n-k)k = nk - \binom{k+1}{2},$$
$$\varphi(n, 2k+1) = \varphi(n, 2k) + 1$$

we have $\nu(\Gamma_n^l) = \varphi(n, l)$ for $l \geqq 2$ and $n > [(l+1)/2]$.

It may easily be verified that the graph Γ_n^l has the following properties: it is $[l/2]$-fold connected, it contains no path and no circuit with more than l edges, if $n \geqq l \geqq 3$, then it contains an l-gon, and if $n > l$, then it contains an l-path. Hence Γ_n^l is a member of each of the classes $F(n, l)$, $G(n, l)$ and $H(n, l)$, and so — having regard to the remarks concerning the case $n \leqq l$ in (2. 1) —

$$(2.\ 3) \qquad f(n, l) \geqq \varphi(n, l), \quad g(n, l) \geqq \varphi(n, l), \quad h(n, l) \geqq \varphi(n, l) \qquad (l \geqq 1).$$

(2. 4) The graph $\Gamma_{n,j}$ $(j \geqq 1)$ is defined as follows: Let $n = qj + r$ where $r < j$. $\Gamma_{n,j}$ has exactly n nodes, and if $r = 0$, it consists of q components, while if $r > 0$, it consists of $q+1$ components, if $r = 0$, all its components are complete j-graphs, and if $r > 0$, then q of its components are complete j-graphs and the remaining component is a complete r-graph. ($\Gamma_{1,1}$ consists of a single node.)

$G^*(n, j)$ $(j \geqq 1)$ denotes the following class of graphs: Let $n = q(j-1) + r$ where $1 \leqq r \leqq j-1$. $G^*(n, j)$ is the class of connected graphs containing exactly n nodes which have q members if $n > 1$ and $r = 1$, and $q+1$ members if $n > 1$ and $r > 1$, every member is a complete j-graph if $n > 1$ and $r = 1$, and q members are complete j-graphs and one member is a complete r-graph if $n > 1$ and $r > 1$. $G^*(1, j) = \{\Gamma_{1,1}\}$ for all $j \geqq 1$.

$\Gamma_{n,j}^*$ is to denote that element of $G^*(n, j)$ which contains at most one cut-node.

In the notation

$$\psi(n, j, r) = \frac{1}{2} nj - \frac{1}{2} r(j+1-r)$$

it is found that

$$\nu(\Gamma_{n,j+1}) = \psi(n, j, r) \quad \text{where} \quad n = q(j+1) + r \qquad (r < j+1)$$

and that

$$\text{if } \Gamma \in G^*(n, j), \text{ then } \nu(\Gamma) = \psi(n, j, r) \text{ where } n = q(j-1) + r \quad (1 \leqq r \leqq j-1).$$

The following statements may easily be verified:

$$\Gamma_{n,l+1} \in F(n, l), \quad G^*(n, l) \subset G(n, l), \quad \Gamma_{n,l} \in H(n, l) \qquad (l \geqq 1).$$

Hence

$$(2.5)\quad \begin{cases} f(n,l)\geqq\psi(n,l,r) & (n=q(l+1)+r;\ r<l+1), \\ g(n,l)\geqq\psi(n,l,r) & (n=q(l-1)+r;\ 1\leqq r\leqq l-1), \qquad (l\geqq 1). \\ h(n,l)\geqq\psi(n,l-1,r) & (n=ql+r;\ r<l) \end{cases}$$

The statements below follow from a simple calculation:

(1) If $l=2k$ $(k\geqq 1)$, then $\varphi(n,2k)\leqq\psi(n,2k,r)$ and equality holds only if $r=k$ or $r=k+1$. So (2.5) gives a better estimate of f and g unless $r=k$ or $r=k+1$, in which cases (2.3) and (2.5) are equally good.

(2) If $l=2k$ $(k\geqq 1)$ and $n>k(k+1)$, then $\varphi(n,2k)>\psi(n,2k-1,r)$. Here (2.3) gives a better estimate of h.

(3) If $l=2k+1$ $(k\geqq 1)$ and $n>k+3$, then $\varphi(n,2k+1)<\psi(n,2k+1,r)$. For this case (2.5) gives a better estimate of f and g.

(4) If $l=2k+1$, then, according as (a) $r=k$ or $r=k+1$, or (b) $r=k-1$ or $r=k+2$, or (c) $r<k-1$ or $r>k+2$, $\varphi(n,2k+1)>,=,$ or $<$ $\psi(n,2k,r)$ and, accordingly, (2.3) gives a better or equally good or worse estimate of h than (2.5).

In order to estimate f, g and h from above we need the theorems of § 1.

THEOREM (2.6)

$$f(n,l)\leqq\frac{1}{2}nl \qquad (l\geqq 1),$$

equality holds only if $n=q(l+1)$, *in which case* $\Gamma_{n,l+1}$ *is the only extreme graph of the class* $F(n,l)$

PROOF. If $n\leqq l+1$, then a graph Γ having exactly n nodes cannot contain a path with more than l edges, so

$$\nu(\Gamma)\leqq\frac{n(n-1)}{2}\leqq\frac{nl}{2}$$

and equality holds here only if Γ is a complete $(l+1)$-graph. So the theorem is true if $n\leqq l+1$.

Now let $n'>l+1$ and suppose that the theorem is true for all n such that $n<n'$. We prove that in that case the theorem is also true for n'. Let $\Gamma\in F(n',l)$.

(1) If Γ is not *connected*, then let its components be $\Gamma_1,\ldots,\Gamma_p$ $(p\geqq 2)$, $\pi(\Gamma_i)=n_i$ $(i=1,\ldots,p)$. Then $n_1+\cdots+n_p=n'$, $n_i<n'$ and $\Gamma_i\in F(n_i,l)$ $(i=1,\ldots,p)$. So by hypothesis

$$\nu(\Gamma)=\nu(\Gamma_1)+\cdots+\nu(\Gamma_p)\leqq n_1\frac{l}{2}+\cdots+n_p\frac{l}{2}=n'\frac{l}{2},$$

and equality holds only if $\nu(\Gamma_i)=n_i\frac{l}{2}$ $(i=1,\dots,p)$, i.e. if $\Gamma_1,\dots,\Gamma_p$ are all complete $(l+1)$-graphs. In this case our theorem is therefore true for n'.

(2) Suppose that Γ is *connected.* We show that there exists a node P' whose degree in Γ, $\varrho(P')$, is at most $l/2$. For if no such node P' exists, then, whether $l=2k$ or $l=2k+1$, $\varrho(P)\geqq k+1$ for every node P. If $l=2k$, then $\pi(\Gamma)>2k+1$, and if $\pi(\Gamma)>2k+2$, then, by (1. 14), Γ contains a path having at least $2k+2$ edges, while if $\pi(\Gamma)=2k+2$, then, by Theorem (1. 10), Γ contains a path having $2k+1$ edges. If $l=2k+1$, then $\pi(\Gamma)>2k+2$ and so, by (1. 14), Γ contains a path having $2k+2$ edges. In every case we have a contradiction.

Suppose therefore that $\varrho(P')\leqq l/2$ and let $\Gamma''=\Gamma-P'$. Then $\Gamma''\in\Gamma(n'-1,l)$. Γ'' cannot contain a complete $(l+1)$-graph because if it did, then Γ would contain an $(l+1)$-path. So by our induction hypothesis $\nu(\Gamma'')<(n'-1)\,l/2$, and therefore

$$\nu(\Gamma)=\varrho(P')+\nu(\Gamma'')<\frac{l}{2}+(n'-1)\frac{l}{2}=n'\frac{l}{2}.$$

THEOREM (2. 7)

$$g(n,l)\leqq\frac{1}{2}(n-1)l\qquad(l\geqq 2),$$

equality holds only if $n=q(l-1)+1$, in which case the extreme graphs of the class $G(n,l)$ are the elements of the class $G^*(n,l)$

PROOF. The theorem is trivially true for $n=1$. If $1<n\leqq l$ and $\Gamma\in G(n,l)$, then

$$\nu(\Gamma)\leqq\frac{n(n-1)}{2}\leqq\frac{(n-1)l}{2},$$

and equality here can hold only if Γ is a complete l-graph. The theorem is therefore true for $n\leqq l$.

Suppose that $n'>l$ and suppose that the theorem is true for all n if $n<n'$. We show that it is then true also for n'. Let $\Gamma\in G(n',l)$.

(1) If Γ is *not connected,* let its components be $\Gamma_1,\dots,\Gamma_p$ $(p\geqq 2)$ and let $\pi(\Gamma_i)=n_i$. Then $n_1+\dots+n_p=n', n_i<n'$ and $\Gamma_i\in G(n_i,l)$ $(i=1,\dots,p)$. By our hypothesis therefore

$$\nu(\Gamma)=\nu(\Gamma_1)+\dots+\nu(\Gamma_p)\leqq(n_1-1)\,l/2+\dots+$$
$$+(n_p-1)l/2=(n'-p)l/2<(n'-1)l/2.$$

(2) If Γ is *connected but not twofold connected,* then let Γ_1 denote a terminal member of Γ and let $\Gamma_2=\Gamma-\Gamma_1$. $\pi(\Gamma_1)=n_1$, $\pi(\Gamma_2)=n_2$. Then

$n'=n_1+n_2-1$, $n_1<n'$, $n_2<n'$, $\Gamma_1\in G(n_1,l)$, $\Gamma_2\in G(n_2,l)$, and so, according to our hypothesis,

$$\nu(\Gamma)=\nu(\Gamma_1)+\nu(\Gamma_2)\leqq(n_1-1)l/2+(n_2-1)l/2=(n'-1)l/2,$$

and equality can hold only if $\Gamma_1\in G^*(n_1,l)$ and $\Gamma_2\in G^*(n_2,l)$. But then $\Gamma\in G^*(n'\,l)$.

(3) Let Γ be *twofold connected*. We show that Γ then has a node P' of degree $\leqq l/2$. For if no such node exists, then $\varrho(P)\geqq k+1$ for every node P both if $l=2k$ and if $l=2k+1$. If $l=2k$, then $\pi(\Gamma)>2k$, in which case if $\pi(\Gamma)\geqq 2k+2$, then by (1. 13) Γ contains an m-gon with $m\geqq 2k+2$ while if $\pi(\Gamma)=2k+1$, then by (1. 10) Γ has an H-line and therefore contains a $(2k+1)$-gon. If $l=2k+1$, then $\pi(\Gamma)>2k+1$ and so Γ contains an m-gon with $m\geqq 2k+2$ by (1. 13). We have obtained a contradiction in every case.

So assume that $\varrho(P')\leqq l/2$ and let $\Gamma'=\Gamma-P'$. Then $\Gamma'\in G(n'-1,l)$. Γ' is connected and $\varrho(P')\geqq 2$. Γ' is not an element of $G^*(n'-1,l)$. For if this were the case, then P' would lie on a circuit with more than l edges. From our induction hypothesis it follows that

$$\nu(\Gamma)=\varrho(P')+\nu(\Gamma')<l/2+(n'-2)l/2=(n'-1)l/2.$$

This proves the theorem.

For the investigation of $h(n,l)$ it is useful to consider the cases $l=2k$ and $l=2k+1$ separately.

THEOREM (2. 8)

$$h(n,2k)\leqq(n-1)k\qquad(k\geqq 1),$$

if $n=1$, then equality holds for all k and $\Gamma_{1,1}$ is the extreme graph of $H(1,2k)$, if $n>1$ and $k=1$, then equality holds for all n and the star with n nodes Γ_n^2 is the only extreme graph of the class $H(n,2)$, finally if $n>1$ and $k>1$, then equality holds only if $n=2k$, and the complete $(2k)$-graph is the only extreme graph of $H(2k,2k)$.

PROOF. Because $H(n,l)\subseteq G(n,l)$, we have that $h(n,2k)\leqq g(n,2k)$. By (2. 7), $g(n,2k)\leqq(n-1)k$, so

$$h(n,2k)\leqq(n-1)k.$$

Equality can hold only if $n=q(2k-1)+1$ and if $H(n,2k)$ contains an element of $G^*(n,2k)$. But an element of $G^*(n,2k)$ belongs to $H(n,2k)$ only if it contains no path with more than $2k$ edges. This holds only for the graphs described in Theorem (2. 8), whether $k=1$ or $k>1$.

If $l=2k+1$, then, since $H(n,1)=F(n,1)$, we need only consider the case $k\geqq 1$.

THEOREM (2.9)

$$h(n,2k+1)\leqq nk \qquad (k\geqq 1),$$

and here equality does not hold if $k=1$ *and* $n=1$ *or* $n=2$, *and it does hold if* $k=1$ *and* $n>2$, *in which case the extreme graphs are those of which all the components are* Δ*-stars; if* $k>1$, *then equality holds only if* $n=q(2k+1)$ $(q>0)$, *and the only extreme graph is* $\Gamma_{n,2k+1}$.

The proof which is similar to the proof of (2.6) will be left to the reader.

§ 3

In this paragraph we determine the graphs with the most edges among the *connected* graphs of the class $F(n,2k)$ and the extreme graphs of the class $H(n,2k)$ for sufficiently large values of n.

We denote the classes of the *connected* graphs in $F(n,l)$, $G(n,l)$, $H(n,l)$, respectively, by $\tilde{F}(n,l)$, $\tilde{G}(n,l)$, $\tilde{H}(n,l)$, and the number of edges in the extreme graphs of these classes (the graphs with the most edges) by $\tilde{f}(n,l)$, $\tilde{g}(n,l)$, $\tilde{h}(n,l)$, respectively. From the fact that Γ_n^l is connected it follows that each of these maximal numbers of edges is $\geqq\varphi(n,l)$.

(3.1) *If* Γ *is an extreme graph of any one of the classes* $\tilde{F}(n,2k)$, $\tilde{G}(n,2k)$, $\tilde{H}(n,2k)$ $(k\geqq 2)$ *and if* $n>k^2-k+1$, *then* Γ *contains a* $2k$*-gon.*

PROOF. If Γ contains no $2k$-gon, then $\Gamma\in G(n,2k-1)$, and so by (2.7) $\nu(\Gamma)\leqq(n-1)\left(k-\frac{1}{2}\right)$. On the other hand, it follows from the extremal property of Γ and from the remark on the graphs Γ_n^l made above, that $\nu(\Gamma)\geqq\nu(\Gamma_n^{2k})=nk-k(k+1)/2$. Accordingly,

$$(n-1)\left(k-\frac{1}{2}\right)\geqq nk-k(k+1)/2,$$

from which it follows that $n\leqq k^2-k+1$. This contradicts our hypothesis.

From Lemma (1.6) we deduce the following lemma:

LEMMA (3.2) *Let the graph* Γ *have an H-line and let* $\pi(\Gamma)=m\geqq 4$. *If* Γ *contains* p *mutually H-independent nodes* $S_1,\ldots,S_p$, *then* $p\leqq m/2$ *and*

$$\nu(\Gamma)\leqq\binom{m}{2}-\frac{mp}{2}+\binom{p+1}{2}, \tag{1}$$

and if $p=m/2$, *then equality in* (1) *holds only if* $\Gamma=\Gamma_m^{2p}$ *and the nodes* $S_1,\dots,S_p$ *span a complete* p*-graph in* Γ.

PROOF. The theorem is trivially true for $m=4$. In what follows it will be assumed that $m>4$. Let $C=(P_1,\dots,P_m,P_{m+1}=P_1)$ be an H-line of Γ and let the nodes $S_i=P_{\sigma_i}$ $(i=1,\dots,p)$ be mutually H-independent. It may be assumed that $1=\sigma_1<\dots<\sigma_p\leqq m$. Further let $R_i=P_{\sigma_i-1}$, $T_i=P_{\sigma_i+1}$, $\varrho(T_i)=\varrho_i$ $(i=1,\dots,p;P_0=P_m)$ and let the elements of the sets $\{R_1,\dots,R_p\}$, $\{S_1,\dots,S_p\}$ and $\{T_1,\dots,T_p\}$ be named R-, S-, T-nodes, respectively.

Because the S-nodes are H-independent, two S-nodes cannot be neighbours on C, and therefore this is true of the R- and T-nodes also, so that an S-node cannot coincide with a T-node or with an R-node. It follows that $p\leqq m/2$.

By (1.6)

$$\varrho_i+\varrho_j\leqq m \qquad (i,j=1,\dots,p;i\neq j). \tag{2}$$

Adding these inequalities together

$$\varrho_1+\dots+\varrho_p\leqq pm/2. \tag{3}$$

Let $\bar{\Gamma}$ denote the complement of the graph Γ and let the degree of T_i in $\bar{\Gamma}$ be denoted by $\bar{\varrho}_i$ $(i=1,\dots,p)$. Then

$$\varrho_i+\bar{\varrho}_i=m-1 \qquad (i=1,\dots,p). \tag{4}$$

By the remark after Lemma (1.6) neither two R-nodes nor two T-nodes can be joined in Γ. Consequently, all the edges $\overline{T_iT_j}$ $(i,j=1,\dots,p;i\neq j)$ are in $\bar{\Gamma}$, and so the number of edges of $\bar{\Gamma}$ at least one end-node of which is a T-node is $\bar{\varrho}_1+\dots+\bar{\varrho}_p-\binom{p}{2}$.

Accordingly,

$$\nu(\bar{\Gamma})\geqq\bar{\varrho}_1+\dots+\bar{\varrho}_p-\binom{p}{2}, \tag{5}$$

and so, having regard to (4) and (3),

$$\nu(\bar{\Gamma})\geqq p(m-1)-(\varrho_1+\dots+\varrho_p)-\binom{p}{2}\geqq p(m-1)-$$
$$-\frac{mp}{2}-\binom{p}{2}=\frac{pm}{2}-\binom{p+1}{2}.$$

Hence

$$\nu(\Gamma)=\binom{m}{2}-\nu(\bar{\Gamma})\leqq\binom{m}{2}-\frac{mp}{2}+\binom{p+1}{2}. \tag{6}$$

Equality holds here only if it holds in (2) and in (5) throughout. If $p=m/2$, then this is the case only if $\varrho_1=\dots=\varrho_p=m/2$ and every edge of $\bar{\Gamma}$ is incident with some T-node, that is to say any two nodes which are not T-nodes are joined by an edge in Γ. In this case, however, every R-node is also a T-node; and these are all joined to every S-node, further any two S-nodes are joined to each other. This implies that $\Gamma = \Gamma_{2p}^{2p}$ and that the S-nodes span a complete p-graph in Γ.

(3. 3) *If the connected graph Γ contains a $2k$-gon ($k \geqq 2$) but does not contain any path having more than $2k$ edges and if $n = \pi(\Gamma) \geqq 3k+2$, then $\nu(\Gamma) \leqq \varphi(n, 2k)$, and equality holds only for $\Gamma = \Gamma_n^{2k}$.*

PROOF. Let $C = (P_1, \ldots, P_{2k}, P_{2k+1} = P_1)$ be a $2k$-gon of Γ. The nodes of C will be called P-nodes and the remaining nodes of Γ will be called Q-nodes. Let the Q-nodes be denoted by $Q_1, \ldots, Q_q$ where $q = n - 2k \geqq k+2$.

Not two Q-nodes are joined in Γ. For if the edge $\overline{Q_1 Q_2}$ exists, for example, then, since Γ is connected, there is a path W in Γ which starts in Q_1 or Q_2 and ends in a P-node, say P_1, and does not contain any P-node other than P_1, and contains only one of the nodes Q_1, Q_2 — say Q_1. Then the edge $\overline{Q_1 Q_2}$ and the paths W and $(P_1, \ldots, P_{2k})$ together constitute a path with at least $2k+1$ edges. This contradicts our hypotheses.

From this and from the fact that Γ is connected it follows that every Q-node is joined to some P-node.

If P_i and P_j are distinct P-nodes and if there are two *distinct* nodes Q_g and Q_h such that the edges $\overline{P_i Q_g}$ and $\overline{P_j Q_h}$ exist, then P_i and P_j are H-independent in $[C]$. For such an open H-line leading from P_i to P_j in $[C]$ would, together with the edges $\overline{P_i Q_g}$ and $\overline{P_j Q_h}$, constitute a $(2k+1)$-path.

We divide the P-nodes into three classes. A P-node will be called an *α-node* if it is joined to at least two Q-nodes, it will be called a *β-node* if it is joined to exactly one Q-node and it will be called a *γ-node* if it is not joined to any Q-node. The number of α-, β- and γ-nodes will be denoted by p_α, p_β and p_γ, respectively. $p_\alpha + p_\beta + p_\gamma = 2k$.

According to the above, any pair of α-nodes are H-independent in $[C]$, and so are any α-node and any β-node. Since two neighbouring nodes of C are not H-independent in $[C]$, the neighbours on C of an α-node can only be γ-nodes. It follows from this that $p_\alpha \leqq p_\gamma$, and so $p_\alpha \leqq k$ and

$$p_\beta \leqq 2k - 2p_\alpha. \tag{1}$$

C is an H-line of the graph $[C]$, so Lemma (3. 2) applies to the α-nodes. Accordingly,

$$\nu([C]) \leqq \binom{2k}{2} - kp_\alpha + \binom{p_\alpha + 1}{2}. \tag{2}$$

The number of edges which join a P-node to a Q-node is

$$\nu_{PQ} \leqq q p_\alpha + p_\beta, \tag{3}$$

and so it follows from (1), (2) and (3) that

$$\nu(\Gamma) = \nu([C]) + \nu_{PQ} \leqq \binom{2k}{2} + 2k + (q-k-2)p_\alpha + \binom{p_\alpha+1}{2}.$$

Because $q \geqq k+2$, the expression on the right attains its maximum value in the range $0 \leqq p_\alpha \leqq k$ only if $p_\alpha = k$, and a simple calculation shows that this maximum is equal to $nk - \binom{k+1}{2} = \varphi(n, 2k)$. So $\nu(\Gamma) \leqq \varphi(n, 2k)$ and equality holds only if $p_\alpha = k$ and equality holds in (1), (2) and (3). But if $p_\alpha = k$, then $p_\beta = 0$ and equality holds in (1), and further, by (3. 2), equality then holds in (2) only if $[C] = \Gamma_{2k}^{2k}$ and the α-nodes span a complete k-graph in $[C]$. Finally, because $p_\beta = 0$, equality holds in (3) when every Q-node is joined to every α-node. From the properties which have been enumerated it follows that $\Gamma = \Gamma_n^{2k}$.

THEOREM (3. 4) *If* $n > k^2 - k + 6$ $(k \geqq 1)$, *then* $\tilde{f}(n, 2k) = \varphi(n, 2k)$, *and the only extreme graph of the class* $\tilde{F}(n, 2k)$ *is* Γ_n^{2k}.

PROOF (1) First assume that $k = 1$. By (2. 6) $f(n, 2) \leqq n$ with equality holding only if $n = 3q$, and then $\Gamma_{n,3}$ is the only extreme graph. From this and from $\tilde{f}(n, 2) \leqq f(n, 2)$ it is seen that $\tilde{f}(n, 2) \leqq n-1 = \varphi(n, 2)$ except if $n = 3$. But if Γ is connected and $\pi(\Gamma) = n$, $\nu(\Gamma) = n-1$, then Γ is a tree ([6], p. 47), and therefore does not contain a path with more than 2 edges only if $\Gamma = \Gamma_n^2$. The theorem is therefore true if $k = 1$ and $n > 3$.

(2) Assume that $k \geqq 2$. Then, because $n > k^2 - k + 6$, by (3. 1) any extreme graph Γ of $\tilde{F}(n, 2k)$ contains a $2k$-gon, and since $k^2 - k + 6 \geqq$ $\geqq 3k + 2$, it follows from Theorem (3. 3) that $\nu(\Gamma) \leqq \varphi(n, 2k)$, equality holding only if $\Gamma = \Gamma_n^{2k}$. Our theorem follows from this and from $\Gamma_n^{2k} \in \tilde{F}(n, 2k)$.

(3. 5). *If* Γ *is an extreme graph of the class* $H(n, 2k)$, *then* Γ *has not more than* $(k+1)/2$ *components.*

PROOF. Suppose that the components of Γ are $\Gamma_1, \ldots, \Gamma_p$ and $\pi(\Gamma_i) = n_i$ $(i = 1, \ldots, p)$. Then $n_1 + \cdots + n_p = n$ and Γ_i is an extreme graph of the class $H(n_i, k)$ for $i = 1, \ldots, p$. Then, according to Theorem (2. 8), $\nu(\Gamma_i) \leqq$ $\leqq (n_i - 1)k$ $(i = 1, \ldots, p)$, and so $\nu(\Gamma) = \nu(\Gamma_1) + \cdots + \nu(\Gamma_p) \leqq (n-p)k$. On the other hand, $\nu(\Gamma) \geqq \varphi(n, 2k)$ by (2. 3), therefore

$$(n-p)k \geqq nk - k(k+1)/2,$$

and hence $p \leqq (k+1)/2$.

THEOREM (3. 6) *If* $n > \frac{1}{2}(k+1)^3$, *then* $h(n, 2k) = \varphi(n, 2k)$ *and* Γ_n^{2k} *is the only extreme graph of the class* $H(n, 2k)$.

PROOF. For $k=1$ the assertion of our theorem is contained in (2. 8). Suppose that $k \geqq 2$, $n > \frac{1}{2}(k+1)^3$, and Γ is an extreme graph of the class $H(n, 2k)$. According to (3. 5), Γ then has a component Γ' such that $n' = \pi(\Gamma) > (k+1)^2$. Γ' is an extreme graph of the class $\tilde{H}(n', 2k)$ and therefore contains a $2k$-gon by (3. 1). $(k+1)^2 > 3k+2$ because $k \geqq 2$, so by (3. 3) $\nu(\Gamma') \leqq \varphi(n', 2k)$, equality holding only if $\Gamma' = \Gamma_{n'}^{2k}$. But $\Gamma_{n'}^{2k} \in \tilde{H}(n', 2k)$, and so $\Gamma' = \Gamma_{n'}^{2k}$ and $\nu(\Gamma') = \varphi(n', 2k)$. We show that $\Gamma' = \Gamma$. For suppose that $\Gamma' \neq \Gamma$ and let $\Gamma'' = \Gamma - \Gamma'$. $n'' = \nu(\Gamma'') = n - n'$ and $\Gamma'' \in H(n'', 2k)$. By (2. 8) $\nu(\Gamma'') \leqq (n''-1)k$, and so $\nu(\Gamma) = \nu(\Gamma') + \nu(\Gamma'') \leqq n'k - \binom{k+1}{2} + (n''-1)k = \varphi(n, 2k) - k$. But this contradicts (2. 3).

CONJECTURES. We conjecture from the above that all extreme graphs of the classes occurring in § 2 and § 3 can be found among the graphs Γ_n^l, $\Gamma_{n,l}$ and the members of the class $G^*(n, l)$. Among the twofold connected graphs Γ_n^l is probably in every case the only extreme graph if $n > l+1$.

§ 4

We are going to prove the following

THEOREM (4. 1) *Let* $\pi(\Gamma) = n$. *Assume further the maximum number of independent edges is* k $(k \geqq 1)$. *Then*

$$\nu(\Gamma) \leqq \max\left(\binom{2k+1}{2},\ k(n-k) + \binom{k}{2}\right).$$

Equality can occur only if $\Gamma = \Gamma_n^{2k}$, *or if one component of* Γ *is a complete* $(2k+1)$*-graph and the other components are isolated nodes.*

PROOF. We can clearly assume $n > 2k$. Choose k independent edges and call these *α'-edges* and the remaining edges *β-edges*. The $n-2k$ nodes of the graph which are not incident with α'-edges we call *unsaturated*. Following BERGE ([2], p. 176) we add a node U to Γ and connect U with every unsaturated node by an edge. The new graph we call Γ' and the new edges and the old α'-edges we call *α-edges* (U is incident with $n-2k$ α-edges, every other node is incident with exactly one α-edge).

Let L be a directed U-loop. We call L *alternating* if its edges are alternatingly α-edges and β-edges in the ordering determined by L. A node

P of Γ' is called *α-accessible* if there exists an alternating U-loop of final node P whose last edge is an α-edge and it is called *β-accessible* if there exists an alternating U-loop of final node P whose last edge is a β-edge. Further, by definition, U is called β-accessible. The nodes which are β-accessible but not α-accessible are called *β-nodes.* Denote the number of β-nodes by $\mu+1$.

It is easy to see that U is not α-accessible, thus *U is a β*-node ([1], [2], p. 176; [5], p. 140).

Denote the components of the graph obtained from Γ' by omitting the β-nodes and the edges incident to it by $\Gamma_1, \ldots, \Gamma_m$ $(m \geqq 1)$. Γ_i is called *odd* or *even* if $\pi(\Gamma_i)$ is odd or even. If Γ_i is odd, put $\pi(\Gamma_i) = 2a_i + 1$, if Γ_i is even, put $\pi(\Gamma_i) = 2a_i$. An α-edge one node of which is a β-node and the other node of which belongs to one of the Γ_i we call an *entering edge* of Γ_i.

The following facts are well known ([1], [2], pp. 169—170; [5], pp. 141—142).

Every α-edge incident to a β-node is an entering edge of some odd Γ_i and every odd Γ_i has exactly one entering edge.

From this it follows that every α-edge is either an entering edge (of some odd Γ_i) or is an edge of some Γ_i. Further Γ_i contains exactly a_i α-edges.

The α-edges in Γ_i $(1 \leqq i \leqq m)$ are clearly α'-edges. We obtain their number by subtracting from k the number of α'-edges incident to the β-nodes, i. e. their number is $k-\mu$. Thus

$$\sum_{i=1}^{m} a_i = k-\mu.$$

If Γ_i is even, then $\nu(\Gamma_i) \leqq \binom{2a_i}{2}$. If Γ_i is odd, then

$$\nu(\Gamma_i) \leqq \binom{2a_i+1}{2} = \binom{2a_i}{2} + 2a_i.$$

Thus

$$\sum_{i=1}^{m} \nu(\Gamma_i) \leqq \sum_{i=1}^{m} \binom{2a_i}{2} + \sum_{i=1}^{m} 2a_i \leqq \binom{2k-2\mu}{2} + 2k-2\mu = \binom{2k-2\mu+1}{2}.$$

Equality is only possible if every Γ_i is odd and is a complete graph and $a_i = 0$ for all $i \geqq 2$.

The number of edges in Γ incident to the β-nodes is less than or equal to $(n-\mu)\mu + \binom{\mu}{2}$.

Since every edge of Γ either belongs to one of the Γ_i or is incident to one of the β-nodes, we obtain that

$$\nu(\Gamma) \leqq \binom{2k-2\mu+1}{2} + (n-\mu)\mu + \binom{\mu}{2} = f(\mu).$$

Since $f(\mu)$ is a convex function of μ and $0 \leqq \mu \leqq k$, we obtain

$$\nu(\Gamma) \leqq \max\,(f(0), f(k)) = \max\left(\binom{2k+1}{2},\quad (n-k)k + \binom{k}{2}\right).$$

Equality is only possible if $\mu=0$ or $\mu=k$, if $\mu=0$, one of the Γ_i must be a complete $(2k+1)$-graph and the other Γ_i must be isolated nodes. If $\mu=k$, all the Γ_i must be isolated nodes and every β-node must be connected with all the nodes of Γ, i. e. $\Gamma=\Gamma_n^{2k}$.

(Received 24 June 1959)

Bibliography

[1] H. B. Belck, Reguläre Faktoren von Graphen, *Journal f. reine u. angew. Math.*, **188** (1950), pp. 228—252.

[2] C. Berge, *Theorie des graphes et ses applications* (Paris, 1958).

[3] G. A. Dirac, Some theorems on abstract graphs, *Proc. London Math. Soc.* (3), **2** (1952), pp. 69—81.

[4] P. Erdős and A. H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.*, **52** (1946), pp. 1087—1091.

[5] T. Gallai, On factorisation of graphs, *Acta Math. Acad. Sci. Hung.*, **1** (1950), pp. 133—153.

[6] D. König, *Theorie der endlichen und unendlichen Graphen* (Leipzig, 1936).

[7] T. Kővári, Vera T. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloquium Math.*, 3 (1954), pp. 50—57.

[8] D. J. Newman, A problem in graph theory, *Amer. Math. Monthly*, **65** (1958), p. 611.

[9] P. Turán, Egy gráfelméleti szélsőértékfeladatról, *Mat. és Fiz. Lapok*, **48** (1941), pp. 436—452.

[10] P. Turán, On the theory of graphs, *Colloquium Math.*, **3** (1954), pp. 19—30.

[11] K. Zarankiewicz, Sur les relations symétriques dans l'ensemble fini, *Colloquium Math.*, 1 (1947), pp. 10—14.