Acta Mathematica Academiae Scientiarum Hungaricae
Tomus 32 (1—2), (1978), 97—120.

GRAPHS WITH MAXIMAL NUMBER OF ADJACENT PAIRS OF EDGES

By

R. AHLSWEDE[1] (Bielefeld) and G. O. H. KATONA[2] (Budapest)

1. Introduction and statement of the results

Let G_n^N denote an undirected graph (without loops and multiple edges) with n vertices and N edges. $P(G_n^N)$ shall count the number of pairs of different edges which have a common vertex and finally $f(n, N)$ is defined by

$$f(n, N) = \max p(G_n^N), \tag{1.1}$$

where the maximum is taken over all possible graphs G_n^N.

In information theory the problem came up to determine $f(n, N)$ for certain hypergraphs. We give here a solution for graphs and for bipartite graphs. As Vera T. Sós kindly informed us, this problem has been solved by MOSHE KATZ [1] for "nice" N's.

In order to state our results we need the concepts of a quasi-complete graph and of a quasi-star. Suppose the vertices of the graph are denoted by $1, 2, \ldots, n$. We define the *quasi-complete* graph C_n^N with N edges in the following way: i and j are connected for $i, j \leqq a$ $(i \neq j)$ and $a+1$ is connected with $1, 2, \ldots, b$, where a and b are determined by the unique representation

$$N = \binom{a}{2} + b, \quad 0 \leqq b < a. \tag{1.2}$$

A *quasi-star* S_n^N with N edges is defined as follows: use the unique representation

$$\binom{n}{2} - N = \binom{c}{2} + d, \quad 0 \leqq d < c \tag{1.3}$$

and connect the first $n-c-1$ vertices with every other, connect the vertex $n-c$ with the first $n-d$ vertices.

It is easy to see that $S_n^{\binom{n}{2}-N}$ is the complement graph of C_n^N if we change the order of the vertices. We use the abbreviations

$$C(n, N) = p\binom{N}{n}, \quad S(n, N) = p(S_n^N). \tag{1.4}$$

Let $G_{l,m}^N$ denote an aritrary bipartite graph with N edges and $l+m$ vertices, where l vertices are coloured red and m pink.

[1] Research of this author was supported by the Deutsche Forschungsgemeinschaft.
[2] This work was done while the author visited the Institute of Mathematical Statistics, University of Göttingen.

THEOREM 1. *Suppose* $1 \leqq m$, $N = qm + r$, $0 \leqq r < m$. *Then* $\max p(G^N_{l,m})$ *is assumed for the graph in which* q *red vertices are all connected with all the pink ones and one more red vertex is connected with* r *pink ones.*

THEOREM 2. $f(n, N) = \max[C(n, N),\ S(n, N)]$.

THEOREM 3.

$$f(n, N) = \begin{cases} S(n, N) & \text{if} \quad 0 \leqq N < \frac{1}{2}\binom{n}{2} - \frac{n}{2} \\ C(n, N) & \text{if} \quad \frac{1}{2}\binom{n}{2} + \frac{n}{2} < N \leqq \binom{n}{2}. \end{cases}$$

Moreover, there are infinitely many n's for which $f(n, N) = S(n, N)$ for *all* $N < \frac{1}{2}\binom{n}{2}$ and $f(n, N) = C(n, N)$ for all $N > \frac{1}{2}\binom{n}{2}$. On the other hand there are infinitely many n's for which this is not true. (For further details see Lemma 8.)

We give two different proofs of Theorem 2. The first is more elegant and is based on Theorem 1.

The second proof is more elaborate, however, it uses techniques which we also use in the proof of Theorem 3, and it is worthwhile knowing that both Theorems can be proved by the same approach. The first proof might be more suited for generalizations of Theorem 2.

2. Proof of Theorem 1

The present proof and also the first proof of Theorem 2 are formulated in terms of vertex-vertex incidence matrices.

For the bipartite graph $G^N_{l,m}$ with l red and m pink vertices and with N edges the matrix $J(G^N_{l,m})$ is defined by

$$J(G^N_{l,m}) = (a_{ij})^{i=1,\dots,l}_{j=1,\dots,m}$$

where

$$a_{ij} = \begin{cases} 1 & \text{if the } i\text{-th red and } j\text{-th pink vertices are connected in } G^N_{l,m} \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

If $J^N_{l,m}$ is an $l \times m$ $0-1$-matrix with N 1's $q(J^N_{l,m})$ denotes

$$\sum_{i=1}^{l} r_i^2 + \sum_{i=1}^{m} s_i^2,$$

where $r_1, \dots, r_l$ and, $s_1, \dots, s_m$ are the numbers of 1's in the rows and columns, respectively. Since

$$q(J(G^N_{l,m})) = 2p(G^N_{l,m}) + 2N, \tag{2.2}$$

for our purposes it suffices to maximize the quadratic form $q(J^N_{l,m})$. We need

LEMMA 1. $\max q(J^N_{l,m})$ *is assumed for a matrix with the property: if* $a_{ij} = 1$ *and* $i' \leqq i$, $j' \leqq j$, *then* $a_{i'j'} = 1$.

PROOF. Suppose $J^N_{l,m}$ maximizes and its rows and columns are ordered such that $r_1 \geqq r_2 \geqq \ldots \geqq r_l$ and $s_1 \geqq s_2 \geqq \ldots \geqq s_m$. Suppose that there are entries $a_{ij}=1$, $a_{ij'}=0$ for $j'<j$ (or $a_{i'j}=0$ for $i'<i$). By exchanging this 1 and 0, $q(J^N_{l,m})$ changes by

$$-s_{j'}^2+(s_j-1)^2+(s_{j'}+1)^2-s_{j'}^2 = 2+2(s_{j'}-s_j) > 0,$$

a contradiction and the lemma is proved.

PROOF OF THEOREM 1. We prove it by induction on $n=l+m$. If $l+m=2$ the statement is trivial. Suppose that $l+m>2$ and also that $J^N_{l,m}$ has the properties described in Lemma 1. We can also assume that $r_1 \geqq s_1$, because otherwise we can exchange the role of rows and columns without changing the total number of rows and columns, because $r_1<s_1\leqq m$, and we can have 1's only in the first r_2 columns. (We write 0's in the undefined places.)

Notice that we have now 1's only in the submatrix J^* determined by the first r_1 columns and first s_1 rows, and by Lemma 1 the first row of this submatrix contains only 1's. Denoting by $J^{N-r_1}_{s_1-1,r_1}$ the $(s_1-1)\times r_1$-matrix, which is derived form J^* by omitting its first row, we can establish the recursion formula

$$(2.3) \qquad q(J^N_{l,m}) = \sum_{i=1}^{m} s_i^2 + \sum_{i=1}^{l} r_i^2 = \sum_{i=1}^{r_1} s_i^2 + \sum_{i=1}^{s_1} r_i^2 =$$

$$= \left(\sum_{i=1}^{r_1}(s_i-1)^2+\sum_{i=2}^{c_1} r_i^2\right)+\sum_{i=1}^{r_1}(2s_i-1)+r_1^2 = q(J^{N-r_1}_{s_1-1,r_1})+2N-r_1+r_1^2.$$

This means that if we want to maximize $q(J^N_{l,m})$ with fixed r_1 and c_1 then we have to maximize $q(J^{N-r_1}_{s_1-1,r_1})$. Here $s_1-1+r_1<l+m$ and we can use the induction hypothesis. Since $c_1-1<r_1$, $\max q(J^{N-r_1}_{s_1-1,r_1})$ is assumed for a matrix in which the first $u-1$ rows are full with 1's and in the u-th row the first r entries are 1's, where

$$(2.4) \qquad N-r_1 = (u-1)r_1+v, \quad 0 \leqq v < r_1.$$

All the other entries are 0's.

We have thus proved that it is sufficient to consider the matrices $J^N_{l,m}$ which have 1's in the first r_1 places of the first u rows and in the first v places in the $(u+1)$-th row, where $u\leqq r_1 \leqq m$ and $N=ur_1+v$, $0\leqq v<r_1$. Denote these matrices by $J(N, r_1)$. We have only to prove that

$$(2.5) \qquad q\big(J(N, r_1+1)\big) \geqq q\big(J(N, r_1)\big).$$

For this we use the equation

$$(2.6) \qquad q\big(J(N, r_1)\big) = u_1^2+v^2+v(u+1)^2+(r_1-v)u^2.$$

We distinguish two cases:

1. *Case* $v\geqq u$. Then

$$N = u(r_1+1)+v-u, \quad 0 \leqq v-u < r$$

and by (2.6)

$$(2.7) \quad q\big(j(N, r_1+1)\big) = u(r_1+1)^2+(v-u)^2+(v-u)(u+1)^2+(r_1-v+u+1)u^2.$$

An elementary calculation yields

$$q\big(J(N, r_1+1)\big)-q\big(J(N, r_1)\big) = 2r_1u-2vu.$$

Since $r_1>v$ this proves (2.5) in the first case.

2. *Case* $v<u$. Then $N=(u-1)(r_1+1)+(v-u+r_1+1)$, where $0\leqq v-u+r_1+1<$ $<r_1+1$ follows from $u\leqq r_1$ and $v<u$. We can again use (2.6) and get

$$\text{(2.8)} \quad q\big(J(N, r_1+1)\big) = (u-1)(r_1+1)^2+(v-u+r_1+1)^2+(v-u+r_1+1)u^2+ +(u-v)(u-1)^2.$$

Again by an elementary calculation

$$q\big(J(N, r_1+1)\big)-q\big(J(N, r_1)\big) = 2r_1v-2uv,$$

which is non-negative because $r_1\geqq u$. We thus have proved (2.5) in both cases and $q\big(J(N, r_1)\big)$ is maximal if we choose $r_1=m$. The Theorem is proved.

3. First proof of Theorem 2

Denote the vertex-vertex incidence matrix of G_n^N by

$$I(G_n^N) = (a_{ij})_{i=1,\dots,n}^{j=1,\dots,n};$$

where

$$a_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected} \\ 0 & \text{otherwise.} \end{cases}$$

The $n\times n$-matrix $I(G_n^N)$ is symmetric, has $2N$ 1's and 0's in the diagonal. Also every $0-1$-matrix I_n^N is the incidence matrix of a graph G_n^N. Define

$$\text{(3.1)} \qquad Q(I_n^N) = \sum_{i=1}^{n} c_i^2,$$

where c_i counts the number of 1's in the i-th row or i-th column. Clearly,

$$\text{(3.2)} \qquad Q\big(I(G_n^N)\big) = 2p\big(I(G_n^N)\big)+2N$$

and it suffices to maximize the expression in (3.1). Again we need an auxiliary result

LEMMA 2. $\max Q(I_n^N)$ *is assumed for a matrix with the property: if* $a_{ij}=1$ *and* $i'\leqq i$, $j'\leqq j$, $i'\neq j'$, *then* $a_{i'j'}=1$.

PROOF. Let the rows and columns of an optimal I_n^N be numbered by $1, 2, \dots, n$ such that $c_1\geqq c_2\geqq\dots\geqq c_n$. Now suppose that $a_{ij}=1$, $i<j$, but $a_{i'j}=0$ for $i'<i$. By exchanging this 1 and 0 (also symmetrically the corresponding term) we get

$$c_{i'}^* = c_{i'}+1, \quad c_i^* = c_i-1, \quad c_t^* = c_t \quad \text{for} \quad t\neq i', i,$$

and

$$\sum_{t=1}^{n} c_t^{*2} - \sum_{t=1}^{n} c_t^2 = 2c_i' - 2c_i + 2 > 0,$$

because $c_i' \geqq c_i$. This contradiction proves the lemma.

We begin now the proof of Theorem 2. Suppose I_n^N has the form described in Lemma 2. Denote the largest index i satisfying $a_{i,i-1}=1$ by $w=w(I_n^N)$. Then

$$\begin{cases} a_{w+1,w} = a_{w,w+1} = 0, \quad a_{ij} = 0, \quad \text{for} \quad w+1 \leqq i, j. \\ a_{ij} = 1 \quad \text{for} \quad 1 \leqq i, j \leqq w \quad \text{with} \quad i \neq j. \end{cases} \tag{3.3}$$

In the class of matrices I_n^N maximizing $Q(I_n^N)$, let the minimal and maximal value of $w(I_n^N)$ be denoted by w_1 and w_2, respectively. We distinguish two cases:

1. *Case* $w_2 \geqq \frac{n}{2}$. Cut the matrix I_n^N, with $w_2 = w(I_n^N)$, after the w_2-th row and w_2-th column. It is separated now into four submatrices A, B, C and D. Here A is a $w_2 \times w_2$ matrix with 0's only in the diagonal. B and C are $w_2 \times (n-w_2)$ and $(n-w_2)\times w_2$ matrices (symmetrical images of each other), respectively, D is a 0-matrix.

Consider the expression

$$Q(I_n^N) = \sum_{i=j}^{n} c_i^2 = \sum_{i=1}^{w_2} (c_i - (w_2-1))^2 + 2\sum_{i=1}^{w_2} c_i(w_2-1) - \sum_{i=1}^{w_2} (w_2-1)^2 + \sum_{i=w_2+1}^{n} c_i^2 = \tag{3.4}$$

$$= \sum_{i=1}^{w_2} (c_i - (w_2-1))^2 + \sum_{i=w_2+1}^{n} c_i^2 + 2N(w_2-1) - 2w_2(w_2-1)^2.$$

B and C are incidence matrices of bipartite graphs and

$$q(B) = q(C) = \sum_{i=1}^{w_2} (c_i - (w_2-1))^2 + \sum_{i=w_2+1}^{n} c_i^2. \tag{3.5}$$

To maximize $Q(I_n^N)$ for fixed w_2 means therefore simply to maximize $q(B)$. Since $w_2 \geqq n - w_2$ it follows from Theorem 1 that the first column of B contains as many 1's as the number of 1's in B permits. This leads to a contradiction if the first column is full, because $a_{w_2, w_2+1} = 1$ contradicts the definition of w_2. It follows that B contains 1's only in the first column and that this number is smaller than w_2.

C is symmetrical to B. The matrix I_n^N, which consists of A, B, C and D is the incidence matrix of a *quasi-complete graph.*

2. *Case* $w_2 < \frac{n}{2}$. In this case also $w_1 < \frac{n}{2}$. Cut now the matrix I_n^N with $w_1 = w(I_n^N)$ after the first $w_1 - 1$ rows and $w_1 - 1$ columns. We obtain four matrices: A, B, C and D. A is a $(w_1-1)\times(w_1-1)$ matrix with 0's only in the diagonal. B is a $(w_1-1)\times(n-w_1+1)$ matrix, C is the symmetrical image of B. D is a 0-matrix since $a_{w_1, w_1+1} = 0$. We can use (3.4) with $w_1 - 1$ instead of w_2. Then $Q(I_n^N)$ can be maximized by maximizing $q(B)$. Since $w_1 - 1 < n - w_1 + 1$ by Theorem 1, $q(B)$ is maximal if we choose B full with 1's in the first rows. However, by definition of w_1, $a_{w_1-1, w_1} = 1$, that is, the first element of the last row of B is 1. It follows that

the first w_1-2 rows of B are full with 1's, because otherwise we could find a B (and I_n^N) which maximizes $Q(I_n^N)$ and has $w(I_n^N)<w_1$.

Thus we have found an optimal matrix which has 1's in the first w_1-2 rows and columns everywhere except the diagonal, furthermore it has 1's at the beginning of the (w_1-1)-th row and column. The other entries are 0. This is the incidence matrix of a *quasi-star graph.* Case 1 and 2 yield

$$f(n, N) = \max\big(C(n, N),\ S(n, N)\big).$$

Notice that the proof gives also an interesting property of $S(n, N)$:

$$S(n+1, N) \geqq S\{n, N) \quad \text{if} \quad n \geqq 4 \quad \text{and} \quad N \leqq \frac{1}{2}\binom{n}{2}. \tag{3.6}$$

To see this it suffices to check that $w(S_n^N)<\dfrac{n+1}{2}$. Suppose the contrary, then

$$2N \geqq w(w-1)+2(n+1-w)(w-2) = -w^2+(2n+5)w-4(n+1) \geqq$$

$$\geqq -\left(\frac{n+1}{2}\right)^2+(2n+5)\frac{n+1}{2}-4(n+1) = \frac{(n+1)(3n-7)}{4} > \frac{n(n-1)}{2}.$$

This contradiction for $n \geqq 4$ establishes (3.6).

4. Comparison of $C(n, N)$ and $S(n, N)$, the proof of Theorem 3

At the first moment one might think that it should be easy to compare $C(n, N)$ and $S(n, N)$ for given n and N. However, the functions are given only in an implicit way by number-theoretical-combinatorial expressions. Also around $N=\dfrac{1}{2}\dbinom{n}{2}$ they are very close to each other. Of course it is quite easy to make the comparison if $N \ll \dfrac{1}{2}\dbinom{n}{2}$ or $N \gg \dfrac{1}{2}\dbinom{n}{2}$, but we would like to consider values of N around $\dfrac{1}{2}\dbinom{n}{2}$ as well. We shall need several lemmas, which we now state and prove.

LEMMA 3.

a) $S(n, N) = C\left(n, \dbinom{n}{2}-N\right)-n\dbinom{n-1}{2}-4N+2nN$

b) $S(n, N)-C(n, N) = C\left(n, \dbinom{n}{2}-N\right)-S\left(n, \dbinom{n}{2}-N\right).$

PROOF. Since the quasi-star with N edges and the quasi-complete graph with $\dbinom{n}{2}-N$ edges are complementary to each other (if we change the order of the vertices), it suffices to prove the statements for any pair of complementary graphs. If we denote by $c_1, \dots, c_n$ the valencies of the first graph, then $\sum\limits_{i=1}^{n} c_i=2N$, $\sum\limits_{i=1}^{n}\dbinom{c_i}{2}$

is the number of adjacencies and $\sum_{i=1}^{n}\binom{n-c_i-1}{2}$ is the corresponding number for the complement graph. Now

$$\sum_{i=1}^{n}\binom{n-c_i-1}{2}=\frac{1}{2}\sum_{i=1}^{n}(n-c_i-1)(n-c_i-2)=$$

$$=\frac{1}{2}\sum_{i=1}^{n}n^2-\sum_{i=1}^{n}nc_i-\frac{3}{2}\sum_{i=1}^{n}n+2\sum_{i=1}^{n}c_i+\sum_{i=1}^{n}1+\frac{1}{2}\sum_{i=1}^{n}(c_i^2-c_i)=$$

$$=\frac{n^3}{2}-2Nn-\frac{3}{2}n^2+4N+n+\sum_{i=1}^{n}\binom{c_i}{2}=\sum_{i=1}^{n}\binom{c_i}{2}+n\binom{n-1}{2}+4N-2nN.$$

b) follows easily from a).

After we know now that one of the functions can be expressed in terms of the other one, we express now $C(n, N)$ as partial sum of an infinite sequence. Define β_{ij} by

(4.1) $$\beta_{ij}=i+j,\quad 0\leqq j\leqq i.$$

The k-th element of the sequence $\beta_{00}, \beta_{10}, \beta_{20}, \beta_{21}, \beta_{22}, \dots$ is denoted by α_k.

LEMMA 4. $C(n, N)=\sum_{k=1}^{N}\alpha_k$.

PROOF. We proceed by induction on N. The statement clearly holds for $N=1$. Use the expansion

$$N=\binom{a}{2}+b,\quad 0\leqq b<a.$$

It is easy to see that $\alpha_N=\beta_{a-1,b-1}$. Recall that the quasi-complete graph with N edges is composed out of a complete graph of a vertices and an additional $(a+1)$-st vertex, which is connected with the first b vertices.

Suppose that $b+1<a$, then the quasi-complete graph with $N+1=\binom{a}{2}+b+1$ $(0\leqq b+1<a)$ has one more edge from the $(a+1)$-st vertex. The number of new adjacencies is $a-1+b$. This is, indeed, equal to the new term $\alpha_{N+1}=\beta_{a-1,b}$. If $b+1=a$, then the quasi-complete graph with $N+1=\binom{a+1}{2}$ edges is simply a complete graph with $a+1$ vertices. The number of new adjacencies is $2a-2=$ $=\beta_{a-1,a-1}$ which again equals α_{N+1}. The proof is complete.

LEMMA 5. a) $C(n, N)\leqq S(n, N)$ *is equivalent to*

(4.1) $$\sum_{k=1}^{N}(\alpha_k+\alpha_{\binom{n}{2}-k+1})\leqq N(2n-4).$$

b) *If* $C(n, N)\leqq S(n, N)$ *for some* $n(\geqq 4)$, $N\left(\leqq\frac{1}{2}\binom{n}{2}\right)$ *then it is also true for* $n+1, N$.

PROOF. By Lemma 3a), $C(n, N) \leqq S(n, N)$ is equivalent to

$$C(n, N)+n\binom{n-1}{2}-C\left(n, \binom{n}{2}-N\right) \leqq N(2n-4). \tag{4.2}$$

Since

$$\sum_{i=1}^{n-2} \sum_{j=0}^{i} \beta_{ij} = n\binom{n-1}{2} \tag{4.3}$$

we get from Lemma 4

$$\binom{n-1}{2}-C\left(n, \binom{n}{2}-N\right) = \sum_{i=1}^{n-2} \sum_{j=0}^{i} \beta_{ij} - \sum_{k=1}^{\binom{n}{2}-N} \alpha_k = \sum_{k=1}^{N} \alpha_{\binom{n}{2}-k+1}$$

and that (4.1) is equivalent to

$$\sum_{k=1}^{N} (\alpha_k + \alpha_{\binom{n}{2}-k+1}) \leqq N(2n-4). \tag{4.4}$$

Obviously b) follows from (3.6) and the fact that $C(n+1, N) = C(n, N)$ for $N \leqq \binom{n}{2}$.

We give now a proof which uses only results of the present section. It is far more complicated than the above argument, but it also shows how the new techniques work, which we need later anyhow.

It is clear from a) that it suffices to prove the inequality

$$\sum_{k=1}^{N} (\alpha_k + \alpha_{\binom{n+1}{2}-k+1}) \leqq N(2n-2)$$

or equivalently that

$$\sum_{k=1}^{N} (\alpha_{\binom{n+1}{2}-k+1} - \alpha_{\binom{n}{2}-k+1}) \leqq 2N. \tag{4.5}$$

We prove it first for some special N's. Suppose

$$N = \binom{n}{2} - \binom{l}{2} \quad \left(\left\{\frac{n+1}{2}\right\} \leqq l \leqq n\right)$$

($\{x\}$ is the smallest integer $\geqq x$). Then

$$\sum_{k=1}^{N} \alpha_{\binom{n}{2}-k+1} = \sum_{i=l-1}^{n-2} \sum_{j=0}^{i} \beta_{ij} = \sum_{i=0}^{n-2} \sum_{j=0}^{i} \beta_{ij} - \sum_{i=0}^{l-2} \sum_{j=0}^{i} \beta_{ij} = n\binom{n-1}{2} - l\binom{l-1}{2},$$

and

$$\sum_{k=1}^{N} \alpha_{\binom{n+1}{2}-k+1} = \sum_{i=l+1}^{n-1} \sum_{j=0}^{i} \beta_{ij} + \sum_{j=n-l}^{l} \beta_{l,j} = \sum_{i=0}^{n-1} \sum_{j=0}^{i} \beta_{ij} - \sum_{i=0}^{l} \sum_{j=0}^{i} \beta_{ij} + \sum_{j=n-1}^{l} \beta_{l,j} =$$

$$= (n+1)\binom{n}{2} - (l+2)\binom{l+1}{2} + (2l-n+1)l + (2l-n+1)\frac{n}{2}.$$

Therefore the left hand side of (4.5) is

$$(n+1)\binom{n}{2}-(l+2)\binom{l+1}{2}+(2l-n+1)\left(l+\frac{n}{2}\right)-n\binom{n-1}{2}+l\binom{l-1}{2},$$

which equals

$$2\left(\binom{n}{2}-\binom{l}{2}\right)=2N.$$

We proved that in the case

$$N=\binom{n}{2}-\binom{l}{2}\qquad\left(\left\{\frac{n+1}{2}\right\}\leqq l\leqq n\right)$$

exact equality holds in (4.5).

We shall now verify that the difference of the sides (right-left) of (4.5) is an increasing function of N in the interval

$$\binom{n}{2}-\binom{l}{2}\leqq N\leqq\binom{n+1}{2}-\binom{l+1}{2} \tag{4.6}$$

and it is decreasing in

$$\binom{n+1}{2}-\binom{l+1}{2}\leqq\binom{n}{2}-\binom{l-1}{2}. \tag{4.7}$$

In the last interval (4.6) the last terms of the left hand side of (4.5) are

$$\alpha_{\binom{n+1}{2}-N+1}=\beta_{l,\binom{n+1}{2}-\binom{l+1}{2}-N}$$

and

$$\alpha_{\binom{n}{2}-N+1}=\beta_{l-2,l-1-N+\binom{n}{2}-\binom{l}{2}}\quad\left(\beta_{l-1,0}\quad\text{if}\quad N=\binom{n}{2}-\binom{l}{2}\right).$$

Thus for $N+1$, the new terms are

$$\beta_{l,\binom{n+1}{2}-\binom{l+1}{2}-N-1}\quad\text{and}\quad\beta_{l-2,l-2-N+\binom{n}{2}-\binom{l}{2}}.$$

Their difference is

$$l+\binom{n+1}{2}-\binom{l+1}{2}-N-1-l+2-l+2+N-\binom{n}{2}+\binom{l}{2}=n-2l+3.$$

The right hand side of (4.5) is increased by 2, so the change of the difference of the sides is $2l-n-1$, which is non-negative by the supposition $\left\{\frac{n+1}{2}\right\}\leqq l$. Similarly if we are in the interval (4.7), the last terms in (4.5) are

$$\alpha_{\binom{n+1}{2}-N+1}=\beta_{l-1,l-N+\binom{n+1}{2}-\binom{l+1}{2}}\quad\left(\beta_{l,0}\quad\text{if}\quad N=\binom{n+1}{2}-\binom{l+1}{2}\right)$$

and

$$\alpha_{\binom{n}{2}-N+1}=\beta_{l-2,l-1-N+\binom{n}{2}-\binom{l}{2}}.$$

Thus going to $N+1$, the new terms are

$$\beta_{l-1,l-N+\binom{n+1}{2}-\binom{l+1}{2}+1}\quad\text{and}\quad\beta_{l-2,l-N+\binom{n}{2}-\binom{l}{2}}.$$

Their difference is

$$l-1+l-N+\binom{n+1}{2}-\binom{l+1}{2}+1-l+2-l+N-\binom{n}{2}+\binom{l}{2}=n-l+2.$$

The change of the difference of the sides of (4.5) is $l-n$, non-positive because of $l\leqq n$. This proves the statement of our lemma, if

$$N\leqq\binom{n}{2}-\binom{\left\{\frac{n+1}{2}\right\}}{2}$$

but we need it for $N\leqq\frac{1}{2}\binom{n}{2}$. However

$$\frac{1}{2}\binom{n}{2}\leqq\binom{n}{2}-\binom{\left\{\frac{n+1}{2}\right\}}{2} \quad \text{holds, when} \quad 0\leqq\frac{1}{2}\binom{n}{2}-\binom{\frac{n+2}{2}}{2},$$

that is

$$0\leqq 2n^2-2n-n^2-2n=n^2-4n,$$

which is true for $n\geqq 4$. The lemma is proved.

In particular we know that if $C(n, N)\leqq S(n, N)$ holds for $N\leqq\frac{1}{2}\binom{n}{2}-\frac{n}{2}$, then it is also true for $n+1$. However, in order to prove Theorem 3 by induction on n we have to prove the inequality for $N\leqq\frac{1}{2}\binom{n+1}{2}-\frac{n+1}{2}$. Also in some cases it is true up to $\frac{1}{2}\binom{n+1}{2}$ and we want to consider those cases as well. This is done in 3 more lemmas. The last one, Lemma 8, gives the complete solution.

Instead of $C(n, N)\leqq S(n, N)$ we shall use a further modified version of (4.1). Since

$$\sum_{k=1}^{\binom{n}{2}}\alpha_k=n\binom{n-1}{2},$$

the left hand side of (4.1) is equal to

$$n\binom{n-1}{2}-\sum_{k=N+1}^{\binom{n}{2}-N}\alpha_k.$$

Introducing the notation $r=\frac{1}{2}\binom{n}{2}-N$ we obtain from (4.1) necessarily

$$\text{(4.8)} \qquad \sum_{k=\frac{1}{2}\binom{n}{2}-r+1}^{r-\frac{1}{2}\binom{n}{2}}\alpha_k\geqq r(2n-4),$$

where r is not necessarily an integer, but there are always $2r$ terms on the left hand side.

One more form of (4.8) is

$$\frac{\sum \alpha_k}{r} \geqq 2n-4. \tag{4.9}$$

Let e und f be defined by

$$\frac{1}{2}\binom{n}{2} = \binom{e}{f} + f \quad (0 \leqq f < e). \tag{4.10}$$

(It is easy to see that this is unique).

LEMMA 6. *Suppose* $f < \frac{e}{2}$ *(in (4.10)) if* $2e+2f \geqq 2n-1$. *Then* $C(n, N) \leqq S(n, N)$ *holds when* $r = \frac{1}{2}\binom{n}{2} - N \leqq e+f-1$.

PROOF. *Case A:* $f > \frac{e}{2}+1$. We want to prove (4.8). The sequence of numbers, we are investigating (when f is an integer) is (starting from left to right and from the middle going backward in the next row).

$$\begin{cases} |e-2, \ldots, 2f-5, \; 2f-4, \ldots, 2e-4| \; e-1, \ldots, \; e+f-2 \\ \underbrace{2f-2, \ldots, e+1}_{-e+2f-2} \; |\underbrace{2e, \ldots, 2f,}_{2e-2f+1} \; \underbrace{2f-1, \ldots, e}_{2f-e}| \; \underbrace{2e-2, \ldots, e+f-1}_{e-f}. \end{cases} \tag{4.11}$$

If f is not an integer, the only difference is, that on the right hand side one number, $e+f-\frac{3}{2}$ stands instead of $e+f-2$ and $e+f-1$.

Observe that in one interval in table (4.11) the sum of the numbers standing under each other is constant. The sums are

$$\underbrace{e+2f-4, \ldots,}_{-e+2f-2} \; \underbrace{2e+2f-4, \ldots,}_{2e-2f+1} \; \underbrace{e+2f-2, \ldots,}_{-e+2f} \; \underbrace{2e+2f-3, \ldots}_{e-f}. \tag{4.12}$$

Subcase 1: Prove the statement for $r = \frac{1}{2}\binom{n}{2} - N \leqq e-f$. Using (4.8) we have to prove that the sum of the last $2r$ terms in (4.11) is not smaller than $r(2n-4)$. The sum of the last $2r$ terms of (4.11) is exactly the sum of the last r terms in (4.12). Since $r \leqq e-f$, they are constant. It is enough to prove that

$$2e+2f-3 \geqq 2n-4. \tag{4.13}$$

Later we need the inequality

$$2e+2f-4 \geqq 2n-4 \tag{4.14}$$

as well, thus we make the calculations here together. From the definition (4.10) of e and f we have

$$n(n-1) = 2e^2-2e+4f. \tag{4.15}$$

Suppose, the converse of (4.13) ((4.14)) is true:

$$n > e+f+\frac{1}{2} \quad (n > e+f).$$

Substituting into (4.15) we obtain

$$2e^2-2e+4f>\left(e+f+\frac{1}{2}\right)\left(e+f-\frac{1}{2}\right)$$

$$\left(2e^2-2e+4f>(e+f)(e+f-1)\right).$$

Reordering it:

$$(4.16)\qquad e^2-2e-f(2e+f-4)+\frac{1}{4}>0\quad \left(e^2-e-f(2e+f-5)>0\right).$$

If $e\geqq 2$ then $2e+f-4$ (if $e\geqq 3$ then $2e+f-5$) is non negative, $f(2e+f-4)$ $\left(f(2e+f-5)\right)$ is an increasing function of f, so (4.16) remains true if we use the inequality $f\geqq\frac{e}{2}$ (supposition of our lemma):

$$(4.17)\qquad \begin{cases} e^2-2e-\dfrac{e}{2}\left(2e+\dfrac{e}{2}-4\right)+\dfrac{1}{4}=-\dfrac{e^2}{4}+\dfrac{1}{4}>0\\ \left(e^2-e-\dfrac{e}{2}\left(2e+\dfrac{e}{2}-5\right)=-\dfrac{e^2}{4}+\dfrac{3e}{2}>0\right). \end{cases}$$

However, these inequalities do not hold when $e\geqq 2$ $(e\geqq 6)$. If $n\geqq 3$ then $e\geqq 2$, thus (4.13) is proved for $n\geqq 3$. The inequality (4.14) is proved only for $e\geqq 6$, that is $n\geqq 9$. For the values $n\leqq 3<9$ it is easy to check in Table 1 that the supposition $f\geqq\frac{e}{2}$ of the lemma is not satisfied for $n=3, 4, 6, 7$ and for the remaining values $n=5, 8, 9$ (4.14) holds.

n	$\frac{1}{2}\binom{n}{2}$	e	f	n	$\frac{1}{2}\binom{n}{2}$	e	f
3	1,5	2	0,5	23	126,5	16	6,5
4	3	3	0	24	138	17	2
5	5	3	2	25	150	17	14
6	7,5	4	1,5	26	162,5	18	9,5
7	10,5	5	0,5	27	175,5	19	4,5
8	14	5	4	28	189	19	18
9	18	6	3	29	203	20	13
10	22,5	7	1,5	30	217,5	21	7,5
11	27,5	7	6,5	31	232,5	22	1,5
12	33	8	5	32	248	22	17
13	39	9	3	33	264	23	11
14	45,5	10	0,5	34	280,5	24	4,5
15	52,5	10	7,5	35	297,5	24	21,5
16	60	11	5	36	315	25	15
17	68	12	2	37	333	26	8
18	76,5	12	10,5	38	351,5	27	0,5
19	85,5	13	7,5	39	370,5	27	19,5
20	95	14	4	40	390	28	12
21	105	15	0	41	410	29	4
22	115,5	15	10,5	42	430,5	29	24,5

Table 1

Subcase 2: Prove the statement for $e-f<r\leqq f$. The second interval in (4.12) contains the value $e+2f-2$. If $e+2f-2\geqq 2n-4$, we do not have to prove anything. Otherwise the average $\frac{1}{r}\sum\alpha_k$ (see (4.9)) decreases when we use larger values of r. So, it is sufficient to prove (4.8) for $r=f$. The left hand side of (4.8) is

$$(-e+2f)(e+2f-2)+(e-f)(2e+2f-3) = e^2+2f^2-e-f.$$

(4.8) is equivalent to

$$e^2+2f^2-e-f \geqq f(2n-4).$$

Using $e^2-e=\binom{n}{2}-2f$ we obtain

$$\binom{n}{2}+2f^2-3f \geqq f(2n-4)$$

and this is equivalent to

$$2\left(f-\frac{n}{2}\right)\left(f-\frac{n-1}{2}\right) \geqq 0.$$

This inequality always holds, because f is a half of an integer, it cannot satisfy $\frac{n-1}{2}<f<\frac{n}{2}$.

Subcase 3: $f<r\leqq 2e-f+1$. The new term (see (4.12)) in this interval is $2e+$ $+2f-4$. We have proved (4.14), thus the average $\frac{1}{r}\sum\alpha_k$ is increasing in this interval, consequently (4.9) and (4.8) hold.

Subcase 4: $2e-f+1<r\leqq e+f-1$. If the new term $e+2f-4$ is $\geqq 2n-4$, we are done. In the contrary case the average is decreasing, so it is sufficient to prove it for $r=e+f-1$. The statement means that the average of α_k' s for the 4 intervals is $\geqq 2n-4$ (see (4.9)). We have proved it for 2 intervals, so it is sufficient to prove the same for the two new intervals:

$$(-e+2f-2)(e+2f-4)+(2e+2f-4)(2e-2f+1) \geqq (e-1)(2n-4).$$

This is equivalent to

(4.18) $$3e^2-2f \geqq 2n(e-1).$$

Since $f<e$, we can prove (4.18) with $3e$ instead of $2f$:

(4.19) $$3e \geqq 2n.$$

We know $n(n-1)=2e^2-2e+2f$ and using again $f<e$ we have

(4.20) $$n(n-1) < 2e^2+2e.$$

Suppose (4.19) does not hold: $e<\frac{2n}{3}$, and substitute into (4.20):

$$n(n-1) < \frac{8n^2}{9}+\frac{4n}{3}, \quad \text{or} \quad n-21 < 0,$$

which is a contradiction for $n \geqq 21$. For smaller n's it is easy to check from Table 1 that (4.19) holds for $n=3, 4, 6, 7, 9, 10, 12, 13, \ldots, 20$. For $n=5, 8$ and 11 (4.19) is not true, but (4.18) holds. Case A is proved.

Case B: If $\frac{e}{2} \leqq f \leqq \frac{e}{2}+1$, then (4.11) and (4.12) have a slightly different form:

$$\left\{\begin{array}{l} |e-2, \ldots, 2e-4|\ e-1, \ldots, e+f-2 \\ |2e, \ldots, e+2f-2, \ldots,\ 2f,\ 2f-1, \ldots, e|\ 2e-2, \ldots, e+f-1, \end{array}\right. \tag{4.11'}$$

$$\underbrace{2e+2f-4, \ldots,}_{e-1}\ \underbrace{e+2f-2, \ldots,}_{-e+2f}\ \underbrace{2e+2f-3, \ldots}_{e-f}. \tag{4.12'}$$

So, in this case we consider only 3 intervals. The first two are exactly like in Case A. The length of the third is different, but we did not use it. The lemma is proved.

LEMMA 7. *Suppose* $f<\frac{e}{2}$ *(in (4.10)). If* $2e+2f \geqq 2n-1$, *then* $C(n, N) \leqq S(n, N)$ *holds when*

$$r=\frac{1}{2}\binom{n}{2}-N \leqq 2e-f+1.$$

If $2e+2f<2n-1$, *then* $C(n, N) \leqq S(n, N)$ *holds when*

$$\min\left(\frac{n}{2}, e-f\right) \leqq r \leqq \begin{cases} 2e+f-3 & (f<2) \\ 2e-f+1 & (f \geqq 2) \end{cases}$$

and the opposite inequality holds when $0<r=f$. *The inequality* $C(n, N) \leqq S(n, N)$ *changes its direction only once in this interval.*

PROOF. *Case A:* $f \geqq 2$. The sequence of numbers, we are investigating now, have the following form:

$$\left\{\begin{array}{l} \ldots, e+2f-7, \ldots, 2e-6\,|\,e-2, \ldots, e+2f-3, \ldots, 2e-4\,|\,e-1, \ldots, e+f-2 \\ |\underbrace{2e, \ldots, e+2f-1,}_{e-2f+2}\ \underbrace{e+2f-2, \ldots, e|}_{2f-1}\ \underbrace{2e-2, \ldots, e+2f-1,}_{e-2f}\ \underbrace{e+2f-2, \ldots, e+f-1.}_{f} \end{array}\right. \tag{4.21}$$

Observe that in one interval in table (4.21) the sum of the numbers standing under each other is constant. The sums are

$$\underbrace{3e+2f-7, \ldots,}_{e-2f+2}\ \underbrace{2e+2f-4, \ldots,}_{2f-1}\ \underbrace{3e+2f-5,}_{e-2f}\ \underbrace{\ldots, 2e+2f-3, \ldots}_{f}. \tag{4.22}$$

Subcase 1. If $2e+2f-3 \geqq 2n-4$, then (4.8) holds for $0<r \leqq f$, otherwise it does not hold. This proves the lemma for the first interval.

Subcase 2. Consider now the interval $f<r \leqq e-f$. If $2e+2f-3 \geqq 2n-4$ is true, $3e+2f-5 \geqq 2n-4$ is also true whenever $e \geqq 2$ (that is, $n \geqq 3$). Then (4.8) holds. If $2e+2f-3<2n-4$ holds, then first we prove that (4.8) holds for $r=e-f$. We have to prove

$$(e-2f)(3e+2f-5)+f(2e+2f-3) \geqq (e-f)(2n-4),$$

that is,

(4.23)
$$3e^2-2f^2-2ef-e+3f \geqq 2n(e-f).$$

We use $e^2=\binom{n}{2}+e-2f$:

$$3\binom{n}{2}-2f^2-2ef+2e-3f \geqq 2n(e-f).$$

By the suppositions $2e+2f \leqq 2n-2$, so it is enough to prove this last inequality after substituting $-f(2e+2f)$ by $-f(2n-2)$:

$$3\binom{n}{2}-2nf+2f+2e-3f \geqq 2n(e-f)$$

or

$$3\binom{n}{2}+2e-f \geqq 2ne.$$

Since $f<e$, we can write

(4.24)
$$3\binom{n}{2}+e \geqq 2ne.$$

We prove it in an indirect way. Suppose the contrary, i.e.

(4.25)
$$e>\frac{3\binom{n}{2}}{2n-1}$$

and use $\binom{n}{2}=e^2-e+2f \geqq e^2-e$. As the right hand side of this inequality is increasing in e, it follows

$$\binom{n}{2} > \frac{3\binom{n}{2}}{2n-1}\left(\frac{3\binom{n}{2}}{2n-1}-1\right)$$

by (4.25). Reordering it, we obtain $0>n^2-13n+4$. This is a contradiction for $n \geqq 13$. For smaller n's: (4.26) holds (see Table 1) when $n=5, 6, 8, 9, 10, 11, 12$. For $n=3, 4, 7$ (4.24) is not true, but (4.23) is true.

Subcase 2a. If $\frac{n}{2}<e-f$, we have to prove that the inequality (4.8) changes its direction earlier than $e-f$, at $\frac{n}{2}$. It is enough to prove that it holds for $r=\frac{n}{2}$. In other words:

$$\left(\frac{n}{2}-f\right)(3e+2f-5)+f(2e+2f-3) \geqq (2n-4)\frac{n}{2}$$

has to be proved, or equivalently

(4.26)
$$3ne+2nf-2ef+4f \geqq 2n^2+n.$$

It is sufficient to prove $3ne \geqq 2n^2+n$ (we omitted a positive number, as $n \geqq e$), or

(4.27)
$$3e \geqq 2n+1.$$

We verify it in an indirect way. Suppose $e<\frac{2n+1}{3}$, and substitute into the inequality

$$n^2-n = 2e^2-2e+4f < 2e^2+2e = 2e(e+1),$$

$$n^2-n < 2\frac{2n+1}{3}\cdot\frac{2n+4}{3},$$

or equivalently $n^2-29n-8<0$ which is a contradiction if $n\geqq 30$. For smaller n's: (4.27) holds (see Table 1) when $n=4, 7, 10, 13, 14, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29$. The remaining n's do not belong to this case $\left(f<\frac{e}{2},\ \frac{n}{2}<e-f\right)$.

Subcase 3: $e-f<r\leqq e+f-1$. If $2e+2f-4\geqq 2n-4$, the averages do not decrease in this interval, and we are done. If $2e+2f-4<2n-4$, it is enough to investigate $r=e+f-1$, the average will be the smallest here. We have to prove that the sum in the last 3 small intervals is $\geqq(e+f-1)\,(2n-4)$. The 3^d interval gives

$$(2f-1)(2e+2f-4) = 4ef+4f^2-2e-10f+4.$$

Adding to the sum of the 2 previous ones (see (4.23)):

$$3e^2+2f^2+2ef-7e-3f+4 \geqq (e+f-1)(2n-4)$$

or equivalently

$$3e^2-3e \geqq 2en-2n+f(2n-2f-2e-1). \tag{4.28}$$

It is easy to see, that

$$f(2n-2f-2e-1) = \frac{1}{2}2f(2n-2f-2e-1) \leqq \frac{1}{2}\left(\frac{2n-2e-1}{2}\right)^2 =$$

$$= \frac{n^2}{2}+\frac{e^2}{2}+\frac{1}{8}-ne-\frac{n}{2}+\frac{e}{2}.$$

Substitute it into (4.28)

$$6e^2-6e \geqq 4en-4n+n^2+e^2+\frac{1}{4}-2ne-n+e$$

or equivalently

$$-n^2+5e^2-2en-7e+5n-\frac{1}{4} \geqq 0.$$

Use $e^2=\binom{n}{2}+e-2f$:

$$3n^2-4en-4e+5n-20f-\frac{1}{2} \geqq 0. \tag{4.29}$$

Since $2f<e$:

$$3n^2-4en-14e+5n-\frac{1}{2} \geqq 0. \tag{4.30}$$

From $\binom{n}{2}=e^2-e+2f$ we obtain

$$e=\frac{1+\sqrt{1-8f+4\binom{n}{2}}}{2}<\frac{1+\sqrt{2n^2}}{2}=\frac{1+n\sqrt{2}}{2}.$$

Substitute it into

$$(4.30)\qquad (3-2\sqrt{2})n^2-(7\sqrt{2}-3)n-7-\frac{1}{2}\geqq 0.$$

This holds for $n\geqq 42$. An ugly compution shows, that for $3\leqq n\leqq 41$, $n\neq 6$, either (4.29) holds or $f\geqq\frac{e}{2}$, that is, not our case. For $n=6$ (4.28) holds.

Subcase 4: $e+f-1<r\leqq 2e-f+1$. If

$$(4.31)\qquad 3e+2f-7\geqq 2n-4,$$

then the new terms do not decrease the average $\frac{1}{r}\sum\alpha_k$. Use the assumption $f\geqq 2$ of Case A:

$$(4.32)\qquad 3e+1\geqq 2n.$$

This is weaker than (4.29), which was proved for $n\geqq 21$ and was checked for $n=3, 4, 6, 7, 9, 10, 12, 13, \dots, 20$. For $n=5, 8$, and 11 (4.32) holds.

Case B: $0<f<2$. (4.21) and (4.22) have a slightly modified form

$$(4.21')\qquad \begin{cases}\dots|e-3,\dots,2e-6|\ e-2,\dots,e+2f-4,\ e+2f-3-4|e-1,\dots,e+f-2\\ |\underbrace{2e,\dots,e+2f-1}_{e-2},\ \underbrace{e+2f-2,\dots,e}_{2f-1}|\ \underbrace{2e-2,\dots,e+2f-1}_{e-2f},\ \underbrace{e+2f-2,\dots,e+f-1}_{f}\end{cases}$$

$$(4.22')\qquad \underbrace{3e+2f-7,\dots}_{e-2},\ \underbrace{2e+2f-4,\dots}_{2f-1},\ \underbrace{3e+2f-5,\dots}_{e-2f},\ \underbrace{2e+2f-3,\dots}_{f}.$$

The only change that the 4th interval is shorter, but we did not use its length. The cases $n=3$ and 7 can be done by an easy computation.

Case C: $f=0$.

$$(4.21'')\qquad \begin{cases}\dots,2e-8|\ e-3,\dots,2e-7,\ 2e-6|\ e-2,\dots,2e-4|\\ |2e,\ 2e-1,\ 2e-2,\ 2e-3,\ \underbrace{2e-4,\dots,e}_{e-3}|\ \underbrace{2e-2,}_{1}\ \underbrace{2e-3,\dots,e-1}_{e-1}\end{cases}$$

$$(4.22'')\qquad \underbrace{3e-7,\dots,}_{e-3}\ \underbrace{4e-8,\dots,}_{1}\ \underbrace{3e-5,\dots}_{e-1}.$$

We prove that all these numbers are $\geqq 2n-4$. It is enough to prove that $3e-7\geqq$ $\geqq 2n-4$, that is $3e-3\geqq 2n$. We prove it in an indirect way. Suppose $e<\frac{2n+3}{3}$

and use $n(n-1)=2e(e-1)$:

$$n(n-1) < 2\frac{2n+3}{3}\cdot\frac{2n}{3},$$

or equivalently $n-21<0$. This is a contradiction if $n\geqq 21$. There is only one case, when $n<21$ and $f=0$ (see Table 1), namely when $n=4$. It is easy to check, that the statement holds for $n=4$.

LEMMA 8. *If* $f\geqq\frac{e}{2}$ *in* (4.10), *then*

$$C(n, N) \leqq S(n, N) \quad \textit{for} \quad 0 \leqq N \leqq \frac{1}{2}\binom{n}{2} \tag{4.33}$$

and

$$C(n, N) \geqq S(n, N) \quad \textit{for} \quad \frac{1}{2}\binom{n}{2} \leqq N \leqq \binom{n}{2}. \tag{4.34}$$

If $f<\frac{e}{2}$, *but* $2e+2f\geqq 2n-1$, *then* (4.33) *and* (4.34) *hold, again. If* $f<\frac{e}{2}$, $2e+2f<$ $<2n-1$ *then there is an* R *such that* $f\leqq R\leqq\min\left(\frac{n}{2},\ e-f\right)$, *and*

$$C(n, N) \leqq S(n, N) \quad \textit{for} \quad 0 \leqq N \leqq \frac{1}{2}\binom{n}{2} - R$$

$$C(n, N) \geqq S(n, N) \quad \textit{for} \quad \frac{1}{2}\binom{n}{2} - R \leqq N \leqq \frac{1}{2}\binom{n}{2}$$

$$C(n, N) \leqq S(n, N) \quad \textit{for} \quad \frac{1}{2}\binom{n}{2} \leqq N \leqq \frac{1}{2}\binom{n}{2} + R$$

$$C(n, N) \geqq S(n, N) \quad \textit{for} \quad \frac{1}{2}\binom{n}{2} + R \leqq N \leqq \binom{n}{2}.$$

PROOF. We use induction over n. For $n=3$ the statements are true. Let $n>3$. Suppose the statement is true for $n-1$, we want to prove it for n. According to the induction hypothesis $C(n-1, N)\leqq S(n, N)$ if

$$0 \leqq N \leqq \frac{1}{2}\binom{n-1}{2} - \frac{n-1}{2}. \tag{4.35}$$

By Lemma 5

$$C(n, N) \leqq S(n, N) \tag{4.36}$$

follows under the condition (4.35). From Lemma 6 we know that (4.36) holds when

$$\frac{1}{2}\binom{n}{2} - z \leqq N \leqq \frac{1}{2}\binom{n}{2} - R, \tag{4.37}$$

where

$$z = \begin{cases} e+f-1 & \text{if } \frac{e}{2} \leqq f \\ 2e-f+1 & \text{if } 2 \leqq f < \frac{e}{2} \\ 2e+f-3 & \text{if } 0 \leqq f < 2 \end{cases} \tag{4.38}$$

and $R=0$ if $\frac{e}{2} \leqq f$,

$$f \leqq R \leqq \min\left(\frac{n}{2}, e-f\right) \quad \text{if} \quad \frac{e}{2} > f.$$

If we are able to prove that

$$\frac{1}{2}\binom{n}{2} - z \leqq \frac{1}{2}\binom{n-1}{2} - \frac{n-1}{2} \tag{4.39}$$

(see (4.35) and (4.37)), then (4.36) holds under the condition

$$0 \leqq N \leqq \frac{1}{2}\binom{n}{2} - R.$$

(4.39) is equivalent to

$$n-1 \leqq z. \tag{4.40}$$

Case A: $\frac{e}{2} \leqq f$. Then $z=e+f-1$. Assume the opposite of (4.40) holds: $n-1>e+f-1$ or $n>e+f$. Substitute it into the equality $n(n-1)=2e^2-2e+4f$:

$$(e+f)(e+f-1) < 2e^2-2e+4f,$$

which is equivalent to

$$e^2-f^2-2ef-e+5f > 0. \tag{4.41}$$

If $f+2e-5 \geqq 0$ (if $e \geqq 2$), then

$$f(f+2e-5) \geqq \frac{e}{2}(f+2e-5) \geqq \frac{e}{2}\left(\frac{e}{2}+2e-5\right) = \frac{5e^2}{4} - \frac{5e}{2},$$

(4.41) results in $-e^2+6e>0$. This is a contradiction if $e \geqq 6$, that is, if $n \geqq 9$. For smaller n's either $\frac{e}{2}>f$ or (4.40) holds.

Case B: $2 \leqq f < \frac{e}{2}$. We have to prove $n-1 \leqq 2e-f+1$ or $2n \leqq 3e+4$. Use

$$\frac{n-1}{2} < \sqrt{\binom{n}{2}} = \sqrt{e^2-e+2f} \leqq e:$$

$$2n \leqq \frac{3n}{\sqrt{2}} - \frac{3}{\sqrt{2}} + 4,$$

which holds for $n \geqq 0$.

Case C: $0 \leqq f < 2$. We have to prove $n-1 \leqq 2e+f-3$ or

$$n \leqq 2e-2. \tag{4.42}$$

Use $\frac{n-1}{\sqrt{2}} \leqq e$: $n \leqq \frac{2n-2}{\sqrt{2}}-2$. This holds for $n \geqq 9$. For $n=4, 6$ and 7 ($n=5$ and 8 do not belong to this case) (4.42) holds. The lemma is proved. Lemma 8 completely proves our Theorem 3.

5. The second proof of Theorem 2

We use the results of Section 3, in particular Lemma 8, and its methods and notation. First we have to prove another inequality:

LEMMA 9. $C(n, N+n)-C(n, N) \geqq 2N+\binom{n}{2}$ *if*

$$\binom{\left[\frac{n}{2}\right]}{2} \leqq N \leqq \binom{n}{2}.$$

PROOF. We start with the case $N=\binom{l}{2}$ $\left(\left[\frac{n}{2}\right] \leqq l \leqq n\right)$. Then the statement holds with equality. The difference $C(n, N+n)-C(n, N)$ is the sum of the following terms (see Lemma 4):

$$\beta_{l-1,0}, \ldots, \beta_{l-1,l-1}, \beta_{l,0}, \ldots, \beta_{l,n-l-1}$$

(here we used $l \leqq n$ and $n \leqq 2l+1$). The sum of these terms is

$$\sum_{i=l-1}^{2l-2} i+\sum_{i=l}^{n-1} i = 3\binom{l}{2}+(n-l)\frac{n+l-1}{2} = 2\binom{l}{2}+\binom{n}{2}.$$

The desired equality is proved. Now we prove that the function

$$C(n, N+n)-C(n, N)-2N-\binom{n}{2} \tag{5.1}$$

is increasing in the interval

$$\binom{l}{2} \leqq N \leqq \binom{l}{2}+2l+1-n$$

and it is decreasing in

$$\binom{l}{2}+2l+1-n \leqq N \leqq \binom{l+1}{2} \quad \left(\left[\frac{n}{2}\right] \leqq l < n\right).$$

This proves the statement, since (5.1) is 0 for $N=\binom{l}{2}$ and $N=\binom{l+1}{2}$. If

$$\binom{l}{2} \leqq m < \binom{l}{2}+2l+1-n,$$

then the difference $C(n, N+n)-C(n, N)$ is the sum of the terms

$$\beta_{l-1,N-\binom{l}{2}}, \ldots, \beta_{l-1,l-1}, \beta_{l,0}, \ldots, \beta_{l,n-l-1+N-\binom{l}{2}}.$$

Here

$$0 \leqq n-1+N-\binom{l}{2} \quad \left(\text{because of } l<n, \binom{l}{2} \leqq N\right)$$

and

$$n-l-1+N-\binom{l}{2}<l.$$

If we change N to $N+1$, a new term comes in: $\beta_{l,n-l+N\binom{l}{2}}$, and $\beta_{l-1,N-\binom{l}{2}}$ wil be omitted. The change is

$$n+N-\binom{l}{2}-l+1-N+\binom{l}{2}=n-l+1.$$

However, in (5.1) $2N$ changes to $2(N+1)$. So, the total change is $n-l-1\geqq 0$, and the function (5.1) is increasing. On the other hand, if

$$\binom{l}{2}+2l+1-n \leqq N<\binom{l+1}{2},$$

then the difference $C(n, N+n)-C(n, N)$ is the sum of the terms

$$\beta_{l-1,N-\binom{l}{2}}, \ldots, \beta_{l-1,l-1}, \beta_{l,0}, \ldots, \beta_{l,l}, \beta_{l+1,0}, \ldots, \beta_{l+1,n+2l+N-\binom{l}{2}},$$

where

$$-1 \leqq n-2l-2+N-\binom{l}{2}<l+1$$

(using $N<\left(\frac{l+1}{2}\right)$ and $n\leqq 2l+1$, if $=-1$, $b_{l,l}$ is the last term). Changing N into $N+1$, the sum changes with

$$\beta_{l+1,n-2l-1+N-\binom{l}{2}}-\beta_{l-1,N-\binom{l}{2}}=n-2l+1,$$

and (5.1) changes with $n-2l-1\leqq 0$. The function is decreasing in this interval. The lemma is proved.

LEMMA 10.

(5.2) $$f(n, N)=\max\left(f(n-1, N), f(n-1, N-n+1)+\binom{n-1}{2}+2(N-n+1)\right)$$

if $n-1\leqq N\leqq\binom{n-1}{2}$. *(Otherwise only the defined term is considered on the right hand side.)*

PROOF. Consider an optimal graph. There are two possibilities: a) either each vertex is contained in at least one edge b) or not.

In case b) our graph must be optimal also for $n-1$ vertices, so $f(n, N)=$ $=f(n-1, N)$. In Case a) there is an edge containing the last vertex, thus by Lemma 1, the first vertex is connected with every vertex. Consequently, at the first point there are $\binom{n-1}{2}$ adjacencies. The remaining $N-n+1$ edges have one adjacency at their both ends with edges going to the first vertex. This is $2(N-n+1)$. What remained is the number of adjacencies among edges not containing the first vertex. They must form an optimal configuration on $n-1$ points. Thus in case a) the number of adjacencies is

$$f(n-1, N-n+1)+\binom{n-1}{2}+2(N-n+1).$$

The lemma is proved.

PROOF OF THEOREM 3. We prove that $f(n, N)=C(n, N)$ or $S(n, N)$ by induction on n. The exact result is given by Lemma 8. This is true for $n=2, 3$. Suppose it is also true for $n-1$ (with any N) and let us prove it for n.

Case A. Suppose $f(n-1, N)$ and $f(n-1, N-n+1)$ are both assumed for quasi-complete graphs. Since $N \leqq \binom{n}{2}$, also $N-n+1 \leqq \binom{n-1}{2}$.

Subcase 1. If additionally

$$\binom{\left[\frac{n-1}{2}\right]}{2} \leqq N-n+1 \leqq \binom{n-1}{2},$$

then we can use Lemma 9:

$$C(n-1, N)-C(n-1, N-n+1) \geqq \binom{n-1}{2}+2(N-n+1).$$

That is the first term under the max in (5.2) is $\geqq$ than the other one. Consequently $f(n, N)$ is assumed for the quasi-complete graph.

Subcase 2: $N-n+1<\binom{\left[\frac{n-1}{2}\right]}{2}$. In this case we shall prove that $C(n-1, N-n+1) \leqq S(n-1, N-n+1)$. Indeed, it follows from Lemma 8, when

$$\binom{\left[\frac{n-1}{2}\right]}{2}-1 \leqq \frac{1}{2}\binom{n-1}{2}-\frac{n-1}{2},$$

that is always for $n \geqq 4$. It means that in Subcase 2, $f(n-1, N-n+1)$ is assumed by the quasi-star and it belongs to the next case.

Case B: Suppose $f(n-1, N)$ is assumed for the quasi-complete graph and $f(n-1, N-n+1)$ for the quasi-star. It is easy to see that

$$S(n-1, N-n+1)+\binom{n-1}{2}+2(N-n+1) = S(n, N).$$

In (5.2) we have $\max(C(n-1, N), S(n, m))$, we obtain always either a quasi-star or a quasi-complete graph.

Case C: $f(n-1, N)$ is assumed strictly for the quasi-star and $f(n-1, N-n+1)$ strictly for the quasi-complete graph. It means that

$$C(n-1, N) < S(n-1, N) \tag{5.3}$$

and

$$C(n-1, N-n+1) > S(n-1, N-n+1). \tag{5.4}$$

By Lemma 8 (5.3) results in

$$N < \frac{1}{2}\binom{n-1}{2} + \frac{n-1}{2}, \tag{5.5}$$

and (5.4) results in

$$\frac{1}{2}\binom{n-1}{2} - \frac{n-1}{2} < N-n+1. \tag{5.6}$$

But (5.5) and (5.6) contradict each other.

Case D: $f(n-1, N)$ is assumed strictly for the quasi-star and $f(n-1, N-n+1)$ (not necessarily strictly) also for the quasi-star. We have again (5.3) and (5.5). We shall prove that in (5.2) the second term under the maximum is larger, that is, $S(n-1, N) \leqq S(n, N)$. Rewrite this inequality using Lemma 3:

$$C\left(n-1, \binom{n-1}{2} - N\right) - (n-1)\binom{n-2}{2} - 4N + 2(n-1)N \leqq$$

$$\leqq C\left(n, \binom{n}{2} - N\right) - n\binom{n-1}{2} - 4N + 2nN,$$

or

$$C\left(n-1, \binom{n}{2} - N\right) - C\left(n-1, \binom{n-1}{2} - N\right) \geqq n\binom{n-1}{2} - (n-1)\binom{n-2}{2} - 2N =$$

$$= \binom{n-1}{2} + 2\left(\binom{n-1}{2} - N\right).$$

By Lemma 10 this holds when

$$\binom{\left[\frac{n-1}{2}\right]}{2} \leqq \binom{n-1}{2} - m \leqq \binom{n-1}{2}$$

or

$$0 \leqq m \leqq \binom{n-1}{2} - \binom{\left[\frac{n-1}{2}\right]}{2}.$$

However, this follows from (1.2) when

$$\frac{1}{2}\binom{n-1}{2} + \frac{n-1}{2} \leqq \binom{n-1}{2} - \binom{\left[\frac{n-1}{2}\right]}{2} + 1$$

holds, that is always if $n \geqq 4$. The theorem is finally proved.

6. Open questions

1. Some strange number-theoretical combinatorial questions arise. What is the relative density of the numbers n for which $R=0$ (see Lemma 8)? What is the distribution of $\frac{R}{n}$?

2. We started to think about the next problem. N $\left(0 \leqq N \leqq \binom{n}{3}\right)$ different 3-types are given on an n-set. What is the maximal number of pairs of 3-tuples having 2 elements in common. By easy symmetry-arguments it is enough to consider the case $0 \leqq N \leqq \frac{1}{2}\binom{n}{3}$. We conjecture that in this interval the maximum is assumed for a quasi-complete 3-graph, or a quasi-3-star on some $m \leqq n$ elements. Recently Vera T. Sós and M. Simonovits have some results in connection with this problem.

Reference

[1] MOSHE KATZ, Rearrangements of (0—1) matrices, *Israel J. Math.*, **9** (1971). 53—72.

(Received January 17, 1977)

MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
1053 BUDAPEST, REÁLTANODA U. 13—15.

78-227 — Szegedi Nyomda — F. v.: Dobó József