Studia Scientiarum Mathematicarum Hungarica 10 (1975) 131—140.

THE HAMMING-SPHERE HAS MINIMUM BOUNDARY

by

G. O. H. KATONA

Introduction

In their information-theoretical investigations [6] AHLSWEDE, GÁCS and KÖRNER needed the solution of the following problem. Let $\mathscr{A}$ be a subset of the space of 0—1 sequences of length n. The Hamming-distance $\varrho(a, b)$ of the sequences a and b is the number of places where they differ. $\delta(\mathscr{A})$ is the set of sequences which have Hamming-distance $\leqq 1$ at least from one element of $\mathscr{A}$. The question: what is the minimum of $|\delta(\mathscr{A})|$ ($|X|$ means the number of elements of X) if $|\mathscr{A}|$ is given. To determine the minimum of $|\delta(\mathscr{A})-\mathscr{A}|$ is an equivalent question. They have found an asymptotical solution in a paper of MARGULIS [5], but the problem of determining the exact minimum remained open*. The aim of this paper is to give the exact minimum.

If $|\mathscr{A}|$ allows, the optimal $\mathscr{A}$ is a Hamming-sphere. If $|\mathscr{A}|$ is different, then we have to choose some additional points in a suitable way.

The proof seems to be quite complicated, but it is very easy after knowing the technique of a similar combinatorial question described below (see also Theorem 1): Let $\mathscr{A}$ be a family of k-tuples of an n-element set (0—1 sequences with k 1's). Determine the minimal number of $(k-1)$-tuples which are contained at least in one k-tuple of $\mathscr{A}$ (that is, "lower" Hamming-boundary). This question was solved first by KRUSKAL [1], later (but independently) by the author [2]. The technique is used in the proofs of HANSEL [3] and ECKHOFF and WEGNER [4]. This last proof is the shortest variant of this type. (For other ways of proofs see [7] and [8].) We did not succeed in reducing our problem to this one, but we use the methods. We use heavily an inequality (see Lemma 2) which appears in different forms in [2], [3] and [4].

There is a natural correspondence between the 0—1 sequences of length n and the subsets of an n-element set. We use both terms alternately.

Summary of the used earlier results

LEMMA 1. *If m and k are given non-negative integers, then there is a unique representation*

$$(1) \qquad m = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \dots + \binom{a_t}{t},$$

where $\qquad a_k > a_{k-1} > \dots > a_t \geqq t \geqq 1.$

---

* In the paper of Margulis it is slightly differently formulated. $\partial(\mathscr{A})$ consists of the sequences x belonging to $\mathscr{A}$ and having a sequence $y \notin \mathscr{A}$ with Hamming distance $\partial(x, y)=1$. However, it is easy to see that $\partial(\mathscr{A})=\partial(\bar{\mathscr{A}})-\bar{\mathscr{A}}$.

The proof can be find in [1]–[2].

(1) is called the *k-canonical representation* of m. We define $(k, m>0)$

$$F(k, m) = \binom{a_k}{k-1} + \binom{a_{k-1}}{k-2} + \dots + \binom{a_t}{t-1}, \tag{2}$$

$$F(k, 0) = 0.$$

LEMMA 2. *If* $k>0$, $m_1, m_2 \geqq 0$, *then*

$$F(k+1, m_1+m_2) \leqq \max\big(m_2, F(k+1, m_1)\big) + F(k, m_2). \tag{3}$$

This inequality appears in a modified form in [2], and [3]. This form and the shortest proof can be found in [4].

If $\mathscr{A}$ is a family of k-element subsets of a set of n elements, then $\delta_L(\mathscr{A})$ means the family of $k-1$-element subsets which are subsets of a k-element set $\in \mathscr{A}$.

THEOREM 1. *If* $\mathscr{A}$ *consists of different k-element subsets of an n-element set and* $0 \leqq m = |\mathscr{A}| \leqq \binom{n}{k}$, *then*

$$|\delta_L(\mathscr{A})| \geqq F(k, m)$$

and this is the best lower bound.

This theorem can be found in [1], [2], [3], [4], [7] and [8], and it is an easy consequence of Lemma 2. It is easy to see, that $|\delta_L(\mathscr{A})| = F(k, m)$ if we choose the first m 0—1 sequences with k 1's in the lexicographic order.

LEMMA 3. *If* $0<k$, $0 \leqq m_1, m_2$ *then*

$$F(k, m_1+m_2) \leqq F(k, m_1) + F(k, m_2), \tag{4}$$

PROOF. It can be found in [2]. However (4) is an easy consequence of Theorem 1, thus we give here the proof. Take two disjoint sets S_1 and S_2 of n_1 $\left(m_1 \leqq \binom{n_1}{k}\right)$ and n_2 elements $\left(m_2 \leqq \binom{n_2}{k}\right)$, respectively. Construct an optimal family of n_1 k-tuples on S_1 which contains $F(k, m_1)$ $(k-1)$-tuples. Take the same for S_2. It means, that the family on $S_1 \cup S_2$ contains exactly $F(k, m_1) + F(k, m_2)$ $(k-1)$-tuples. By theorem 1 this family must contain at least $F(k, m_1+m_2)$ $(k-1)$-tuples. This gives (4).

LEMMA 4. *If* $0<k$; $0 \leqq m_1 \leqq m_2 \leqq \binom{n}{k}$, $\binom{n}{k} \leqq m_1 + m_2$, *then*

$$\binom{n}{k-1} + F\left(k, m_1+m_2-\binom{n}{k}\right) \leqq F(k, m_1) + F(k, m_2). \tag{5}$$

PROOF. If $m_1 = \binom{n}{k}$ or $m_2 = \binom{n}{k}$, equality holds in (5). Thus we may assume $m_1, m_2 < \binom{n}{k}$, that is,

$$F\left(k+1, \binom{n}{k+1} + m_1\right) = \binom{n}{k} + F(k, m_1) > m_2.$$

On the other hand,

$$F\left(k+1, \binom{n+1}{k+1}+m_1+m_2-\binom{n}{k}\right) = \binom{n+1}{k}+F\left(k, m_1+m_2-\binom{n}{k}\right),$$

using, that $m_1+m_2-\binom{n}{k}<\binom{n}{k}$. Now we shall use Lemma 2 with the numbers $\binom{n}{k+1}+m_1$ and m_2:

$$F\left(k+1, \binom{n}{k+1}+m_1+m_2\right) = F\left(k+1, \binom{n+1}{k+1}+m_1+m_2-\binom{n}{k}\right) =$$

$$= \binom{n+1}{k}+F\left(k, m_1+m_2-\binom{n}{k}\right) \leqq F\left(k+1, \binom{n}{k+1}+m_1\right)+F(k, m_2) =$$

$$= \binom{n}{k}+F(k, m_1)+F(k, m_2).$$

Thus we obtained an inequality which is equivalent to (5).

A consequence. Theorem 2 in [2] (which has a complicated proof in [2]) is an easy consequence of this lemma. The theorem says (in a slightly more general form), that if $\mathscr{A}$ is a family of different k-tuples on a set $S_1 \cup S_2$ $(S_1 \cap S_2=\varnothing)$, where $|S_2| \leqq$ $\leqq |S_1|=n$, $\binom{n}{k} \leqq |\mathscr{A}| \leqq \binom{n}{k}+\binom{|S_2|}{k}$ and at most one of the relations $A \cap S_1 \neq \varnothing$ $A \cap S_2 \neq \varnothing$ $(A \in \mathscr{A})$ holds, then

$$|\delta_L(\mathscr{A})| \geqq \binom{n}{k-1}+F\left(k, |\mathscr{A}|-\binom{n}{k}\right).$$

That is, the best arrangement is, if we choose all the k-tuples from S_1 and the remainder from S_2.

Proof. Let m_1 and m_2 denote the number of the subsets $(\in \mathscr{A})$ contained by S_1 and S_2, respectively. The minimum of $(k-1)$-tuples "contained by $\mathscr{A}$" in S_1 is $F(k, m_1)$ by Theorem 1, and $F(k, m_2)$ in S_2. Thus Lemma 4 gives the result.

The results

We start with an analogue of lemma 1.

LEMMA 5. *If u and n are given non-negative integers* $(u<2^n)$ *then there is a unique representation (called n-bounded canonical representation)*

(6) $$u = \binom{a_n}{n}+\binom{a_{n-1}}{n-1}+\ldots+\binom{a_t}{t},$$

where $n=a_n=a_{n-1}=\ldots=a_{k+1}>a_k>a_{k-1}>\ldots>a_t \geqq t \leqq 1$ *for some* $k\,(t-1 \leqq k<n)$.

May be, it is better to write

$$u=\binom{n}{n}+\binom{n}{n-1}+\dots+\binom{n}{k+1}+\binom{a_k}{k}+\dots+\binom{a_t}{t} \tag{7}$$

$$(n>a_k>a_{k-1}>\dots>a_t\geqq t\geqq 1),$$

with the remark that the part of a's can completely vanish.

Now we are able to introduce the following notation

$$G(n,u)=\binom{n}{n}+\binom{n}{n-1}+\dots+\binom{n}{k+1}+\binom{n}{k}+\binom{a_k}{k-1}+\dots+\binom{a_t}{t-1} \tag{8}$$

if $u>0$, and $G(n,0)=0$.

LEMMA 6. *If* $0\leqq u_1\leqq u_2$, *then*

$$G(n,u_1+u_2)\leqq \max(u_2, G(n-1,u_1))+G(n-1,u_2). \tag{9}$$

If $\mathscr{A}=\{A_1,\dots,A_u\}$ is a family of different subsets of an n-element set S, then $\delta(\mathscr{A})$ denotes the family of subsets B of S, the Hamming-distance $\varrho(B,A_i)$ of which is $\leqq 1$ at least for one member A_i of $\mathscr{A}$.

THEOREM 2. *If* $\mathscr{A}$ *is a system of different subsets of an n-element set S and* $|\mathscr{A}|=u<2^n$, *then*

$$|\delta(\mathscr{A})|\geqq G(n,u)$$

and this is the best possible bound.

In general, $\delta_d(\mathscr{A})$ is defined in the following way:

$$\delta_d(\mathscr{A})=\{B:\exists A\in\mathscr{A}, \varrho(B,A)\leqq d\}.$$

Similarly, we need the generalization of (7):

$$G_d(n,u)=$$

$$=\binom{n}{n}+\binom{n}{n-1}+\dots+\binom{n}{k+1}+\binom{n}{k}+\dots+\binom{n}{k-d+1}+\binom{a_k}{k-d}+\dots+\binom{a_t}{t-d}.$$

The following theorem is a more general form of Theorem 2:

THEOREM 3. *If* $|\mathscr{A}|=u<2^n$ *then*

$$|\delta_d(\mathscr{A})|\geqq G_d(n,u).$$

PROOF of Lemma 5. (Warning: it is easier to prove than read!) First we prove there is a representation of form (6). Take the minimal k satisfying

$$u\geqq\binom{n}{n}+\dots+\binom{n}{k+1}=v.$$

Then, applying lemma 1 for $u-v$ we obtain a representation of form (6). It remains only to prove that $n>a_k$. In the contrary case

$$u\geqq\binom{n}{n}+\dots+\binom{n}{k+1}+\binom{a_k}{k}\geqq\binom{n}{n}+\dots+\binom{n}{k+1}+\binom{n}{k}$$

holds, thus k was not the minimum, in contradiction with our suppositions.

We have to prove that (6) is unique. Suppose, the contrary case holds, there are two representations. If the k's are the same in both, then $u-v$ has two different representations of form (1) contradicting lemma 1. We can suppose that the k's are different ($k>k'$). Let the other representation be

$$u=\binom{n}{n}+\binom{n}{n-1}+\ldots+\binom{n}{k'+1}+\binom{b_{k'}}{k'}+\ldots+\binom{b_{t'}}{t'}. \tag{10}$$

Using a well-known formula

$$\binom{a_k}{k}+\ldots+\binom{a_t}{t}<\binom{n-1}{k}+\binom{n-2}{k-1}+\ldots+\binom{n-k+t-1}{t}+$$
$$+\binom{n-k+t-2}{t-1}+\ldots+\binom{n-k-1}{0}=\binom{n}{k}.$$

Thus,

$$u<\binom{n}{n}+\binom{n}{n-1}+\ldots+\binom{n}{k+1}+\binom{n}{k},$$

from (6), and

$$u\geqq\binom{n}{n}+\binom{n}{n-1}+\ldots+\binom{n}{k+1}+\binom{n}{k}$$

from (10). These two statements contradict each other. The lemma is proved.

PROOF of theorem 2. First we reduce the theorem to Lemma 6 which will be proved afterwards.

We use induction over n. If $n=1$, then $u=1=\binom{1}{1}$, $G(1,1)=\binom{1}{1}+\binom{1}{0}=2$, and $|\delta(\mathscr{A})|$ is always 2. Assume the theorem is proved for $n-1$, and prove it for n.

Fix an element x of S, and divide $\mathscr{A}$ into two families. $\mathscr{A}_1$ or $\mathscr{A}_2$ consists of the subsets which contain or do not contain x, respectively. The operation * on a family of subsets means that x is added to the members of the family which do not contain it and it is omitted from the members which do. Denote $|\mathscr{A}_1|$ and $|\mathscr{A}_2|$ by z_1 and z_2, respectively. Obviously,

$$z=|\mathscr{A}|=z_1+z_2.$$

We distinguish several cases.

Case 1. $z_1\leqq z_2\leqq G(n-1,z_1)$. We have, by the induction hypothesis,

$$|\delta(\mathscr{A}_2)|\geqq G(n-1,z_2) \tag{11}$$

and

$$|\delta(\mathscr{A}_1^*)|\geqq G(n-1,z_1) \tag{12}$$

as $|\mathscr{A}_1^*|=|\mathscr{A}_1|=z_1$. (Here δ is taken for the $(n-1)$-element set $S-\{x\}$). Similarly,

$$|\delta(\mathscr{A}_1^*)^*|\geqq G(n-1,z_1) \tag{13}$$

follows from (12). As $\delta(\mathscr{A}_2)\subset\delta(\mathscr{A})$, $\delta(\mathscr{A}_1^*)^*\subset\delta(\mathscr{A})$ and they are disjoint,

$$|\delta(\mathscr{A})|\geqq|\delta(\mathscr{A}_2)|+|\delta(\mathscr{A}_1^*)^*|\geqq$$
$$\geqq G(n-1,z_2)+G(n-1,z_1)$$

from (11) and (12). However, this is at least

$$G(n, z_1+z_2) = G(n, z)$$

by Lemma 6 and the suppositions of this case. Case 1 is settled.

Case 2. $z_1 \leqq z_2$, $G(n-1, z_1) \leqq z_2$. Now, we use $\delta(\mathscr{A}) \supset \delta(\mathscr{A}_2)$, $\delta(\mathscr{A}) \supset \mathscr{A}_2^*$ and $\delta(\mathscr{A}_2) \cap \mathscr{A}_2^* = \emptyset$. These fact result in

$$|\delta(\mathscr{A})| \geqq |\delta(\mathscr{A}_2)| + |\mathscr{A}_2^*|.$$

Here $|\mathscr{A}_2^*| = z_2$, and by the inductional hypothesis $|\delta(\mathscr{A}_2)| \geqq G(n-1, z_2)$, thus

$$|\delta(\mathscr{A})| \geqq z_2 + G(n-1, z_2).$$

The right hand side is at least $G(n, z_1+z_2) = G(n, z)$ by lemma 6 and the suppositions of this case. This case in settled, too.

Case 3. $z_2 \leqq z_1 \leqq G(n-1, z_2)$. We can repeat the proof of Case 1. The only difference, that in lemma 6 we have to write z_2 in place of u_1 and z_1 in place of u_2.

Case 4. $z_2 \leqq z_1$, $G(n-1, z_2) \leqq z_1$. Now, we use $\delta(\mathscr{A}) \supset \delta(\mathscr{A}_1^*)$, $\delta(\mathscr{A}) \supset \mathscr{A}_1$ and $\delta(\mathscr{A}_1^*) \cap \mathscr{A}_1 = \emptyset$. These facts result in

$$|\delta(\mathscr{A})| \geqq |\delta(\mathscr{A}_1^*)| + |\mathscr{A}_1|.$$

Here $|\mathscr{A}_1| = z_1$, and by the inductional hypothesis $|\delta(\mathscr{A}_1^*)| \geqq G(n-1, z_1)$, thus

$$|\delta(\mathscr{A})| \geqq z_1 + G(n-1, z_1).$$

The right-hand side is at least $G(n, z_1+z_2) = G(n, z)$ by lemma 6 and the suppositions of this case. The inequality of theorem 2 is proved.

We have to construct an $\mathscr{A}$ showing that the inequality is the best possible. Let $\mathscr{A}$ consist of all the subsets having at least $k+1$ elements and of the first $u - \binom{n}{n} - \dots - \binom{n}{k+1}$ k-tuples in the lexicographic order. It is easy to see that $\delta(\mathscr{A})$ contains all the subsets having at least k elements and $\binom{a_k}{k-1} + \dots + \binom{a_t}{t-1}$ $(k-1)$-tuples according to Theorem 1. The proof is completed.

PROOF of Lemma 6. *Case* 1. $G(n-1, u_1) \leqq u_2$. It is easy to see that $G(n, u)$ is monotonically increasing in u. We have to prove

$$G(n, u_1+u_2) \leqq u_2 + G(n-1, u_2). \tag{14}$$

By the monotonity it is enough to prove this inequality for the maximal possible u_1 satisfying $G(n-1, u_1) \leqq u_2$. Suppose, u_2 has the $(n-1)$-bounded representation

$$u_2 = \binom{n-1}{n-1} + \dots + \binom{n-1}{\gamma+1} + \binom{c_\gamma}{\gamma} + \binom{c_{\gamma-1}}{\gamma-1} + \dots + \binom{c_s}{s}, \tag{15}$$

and μ is the smallest index satisfying $c_\mu > \mu$. Then

$$U_1 = \binom{n-1}{n-1} + \dots + \binom{n-1}{\gamma+2} + \binom{c_\gamma}{\gamma+1} + \dots + \binom{c_\mu}{\mu+1} \tag{16}$$

satisfies $G(n-1, U_1) \leqq u_2$. But U_1+1 does not satisfy it. This is trivial if $\mu = s$.

If $\mu > s$ then U_1+1 has an additional term $\binom{c_{\mu-1}}{\mu}=\binom{\mu}{\mu}=1$ and

$$G(n-1, U_1+1) =$$
$$=\binom{n-1}{n-1}+\ldots+\binom{n-1}{\gamma+1}+\binom{c_\gamma}{\gamma}+\ldots+\binom{c_\mu}{\mu}+\binom{\mu}{\mu-1}>$$
$$>\binom{n-1}{n-1}+\ldots+\binom{n-1}{\gamma+1}+\binom{c_\gamma}{\gamma}+\ldots+\binom{c_\mu}{\mu}+\mu-s.$$

Thus, we really can consider U_1 for u_1 in (14): Here, from (16)

$$U_1+u_2=\binom{n}{n}+\ldots+\binom{n}{\gamma+2}+\binom{c_\gamma+1}{\gamma+1}+\ldots+\binom{c_\mu+1}{\mu+1}+\binom{\mu}{\mu}+\ldots+\binom{s+1}{s+1},$$

and

$$G(n, U_1+u_2) =$$
$$=\binom{n}{n}+\ldots+\binom{n}{\gamma+1}+\binom{c_\gamma+1}{\gamma}+\ldots+\binom{c_\mu+1}{\mu}+\binom{\mu}{\mu-1}+\ldots+\binom{s+1}{s}=$$
$$=\binom{n-1}{n-1}+\ldots+\binom{n-1}{\gamma+1}+\binom{c_\gamma}{\gamma}+\ldots+\binom{c_\mu}{\mu}+\binom{\mu-1}{\mu-1}+\ldots+\binom{s}{s}+$$
$$+\binom{n-1}{n-1}+\ldots+\binom{n-1}{\gamma}+\binom{c_\gamma}{\gamma-1}+\ldots+\binom{c_\mu}{\mu-1}+\binom{\mu-1}{\mu-2}+\ldots+\binom{s}{s-1}=$$
$$=u_2+G(n-1, u_2).$$

The case is settled.

Case 2. $u_2 < G(n-1, u_1)$. Let u_1 have the form

(17) $$u_1=\binom{n-1}{n-1}+\ldots+\binom{n-1}{\beta+1}+\binom{b_\beta}{\beta}+\ldots+\binom{b_r}{r}.$$

From the inequality $u_1 \leqq u_2$ (see (15)) it follows $\beta \geqq \gamma$. On the other hand, from $u_2 < G(n-1, u_1)$ $\beta \leqq \gamma+1$ follows and if $\beta=\gamma+1$, then

(18) $$v_2=\binom{c_\gamma}{\gamma}+\ldots+\binom{c_s}{s}<\binom{b_{\gamma+1}}{\gamma}+\ldots+\binom{b_r}{r-1}.$$

Summarizing, β can be γ or $\gamma+1$. These two cases will be distinguished.

Case 2a. $\beta=\gamma$. Let us introduce the following notations

$$v_1=\binom{b_\beta}{\beta}+\ldots+\binom{b_r}{r}$$

$$v_2=\binom{c_\beta}{\beta}+\ldots+\binom{c_s}{s}.$$

Then

(19) $$u_1+u_2=\binom{n}{n}+\ldots+\binom{n}{\gamma+2}+\left[\binom{n-1}{\gamma+2}+v_1+v_2\right].$$

Unfortunately, it is not a perfect form for taking $G(n, u_1+u_2)$. However, if

Case 2aa. $v_1+v_2<\binom{n-1}{\gamma}$, then $\binom{n-1}{\gamma+1}+v_1+v_2<\binom{n}{\gamma+1}$ that is, the bracket does not disturb the part $\binom{n}{n}+\ldots+\binom{n}{\gamma+2}$ in (19). Thus

$$G(n, u_1+u_2)=\binom{n}{n}+\ldots+\binom{n}{\gamma+1}+F\left(\gamma+1, \binom{n-1}{\gamma+1}+v_1+v_2\right)= \tag{20}$$
$$=\binom{n}{n}+\ldots+\binom{n}{\gamma+1}+\binom{n-1}{\gamma}+F(\gamma, v_1+v_2).$$

On the other hand

$$G(n-1, u_1)=\binom{n-1}{n-1}+\ldots+\binom{n-1}{\gamma}+F(\gamma, v_1) \tag{21}$$

and

$$G(n-1, u_2)=\binom{n-1}{n-1}+\ldots+\binom{n-1}{\gamma}+F(\gamma, v_2). \tag{22}$$

From (20), (21) and (22) it is easy to see that (9) is reduced to $F(\gamma, v_1+v_2)\leqq F(\gamma, v_1)+$ $+F(\gamma, v_2)$ which is lemma 3. We can turn to the next case.

Case 2ab. $v_1+v_2\geqq\binom{n-1}{\gamma}$. In this case we use a modified form of (19):

$$u_1+u_2=\binom{n}{n}+\ldots+\binom{n}{\gamma+2}+\binom{n}{\gamma+1}+\left[v_1+v_2-\binom{n-1}{\gamma}\right].$$

Here $v_1+v_2-\binom{n-1}{\gamma}<\binom{n-1}{\gamma}$ and the last term in bracket can not disturb the previous terms.

$$G(n, u_1+u_2)=\binom{n}{n}+\ldots+\binom{n}{\gamma}+F\left(\gamma, v_1+v_2-\binom{n-1}{\gamma}\right). \tag{23}$$

(23), (21) and (22) lead to

$$\binom{n-1}{\gamma-1}+F\left(\gamma, v_1+v_2-\binom{n-1}{\gamma}\right)\leqq F(\gamma, v_1)+F(\gamma, v_2).$$

This is true by lemma 4. Case 2a is proved.

Case 2b. $\beta=\gamma+1$. Now

$$u_1+u_2=\binom{n}{n}+\ldots+\binom{n}{\gamma+2}+(v_1+v_2)$$

holds, where $v_2<\binom{n-1}{\gamma}$, $v_1<\binom{n-1}{\gamma+1}$, thus $v_1+v_2<\binom{n}{\gamma+1}$. The term v_1+v_2 does not disturb the previous ones. Hence

$$G(n, u_1+u_2)=\binom{n}{n}+\ldots+\binom{n}{\gamma+1}+F(\gamma+1, v_1+v_2). \tag{24}$$

Furthermore the terms on the right-hand side of (9) have the forms

$$G(n-1, u_1) = \binom{n-1}{n-1} + \dots + \binom{n-1}{\gamma+1} + F(\gamma+1, v_1) \tag{25}$$

$$G(n-1, u_2) = \binom{n-1}{n-1} + \dots + \binom{n-1}{\gamma} + F(\gamma, v_2). \tag{26}$$

Comparing (24), (25) and (26), (9) reduces to

$$F(\gamma+1, v_1+v_2) \leqq F(\gamma+1, v_1) + F(\gamma, v_2).$$

This is true by lemma 2 if

$$v_2 \leqq F(\gamma+1, v_1) = \binom{b_{\gamma+1}}{\gamma} + \dots + \binom{b_r}{r-1}.$$

But this is (18) which always holds when $\beta=\gamma+1$. The proof of the lemma is completed.

Proof of theorem 3. We prove the theorem by induction over d. For $d=1$ it is theorem 2. Suppose $d>1$ and the statement is proved for smaller values. Observe that $\delta_d(\mathscr{A})=\delta(\delta_{d-1}(\mathscr{A}))$ and hence by the inductional hypothesis

$$|\delta_d(\mathscr{A})| \geqq G(n, G_{d-1}(n, u)).$$

We have only to prove

$$G(n, G_{d-1}(n, u)) = G_d(n, u). \tag{27}$$

This is trivial, if

$$u = \binom{n}{n} + \dots + \binom{n}{k+1} + \dots + \binom{a_t}{t}$$

and $t \geqq d$ or $k+1 \leqq d$. Otherwise, if $t<d<k+1$, then

$$G_{d-1}(n, u) = \binom{n}{n} + \dots + \binom{n}{k-d+2} + \binom{a_k}{k-d+1} + \dots + \binom{a_t}{t-d+1} =$$

$$= \binom{n}{n} + \dots + \binom{n}{k-d+2} + \binom{a_k}{k-d+1} + \dots + \binom{a_{d-1}}{0}.$$

Let μ be minimal index such that $a_\mu < a_{\mu+1}-1$ $(d-1<\mu \leqq k)$. Then

$$G_{d-1}(n, u) = \binom{n}{n} + \dots + \binom{n}{k-d+2} + \binom{a_k}{k-d+1} + \dots + \binom{a_{\mu+1}}{\mu-d+2} + \binom{a_\mu+1}{\mu-d+1}$$

and

$$G(n, G_{d-1}(n, u)) =$$

$$= \binom{n}{n} + \dots + \binom{n}{k-d+2} + \binom{n}{k-d+1} + \binom{a_k}{k-d} + \dots + \binom{a_{\mu+1}}{\mu-d+1} + \binom{a_\mu+1}{\mu-d}.$$

On the other hand,

$$G_d(n, u) = \binom{n}{n} + \dots + \binom{n}{k-d+1} + \binom{a_k}{k-d} + \dots + \binom{a_d}{0} =$$

$$= \binom{n}{n} + \dots + \binom{n}{k-d+1} + \binom{a_k}{k-d} + \dots + \binom{a_{\mu+1}}{\mu-d+1} + \binom{a_\mu + 1}{\mu - d}.$$

This shows (27) and the theorem. The proof is completed.

Remarks and an open problem

Theorem 2 gives a formula for the min $|\delta(\mathscr{A})|$. It is easy to derive formulas from it for min $|\delta(\mathscr{A})-\mathscr{A}|$ and min $|\partial(\mathscr{A})|$:

$$\min |\delta(\mathscr{A})-\mathscr{A}| = \min |\delta(\mathscr{A})| - |\mathscr{A}| =$$

$$= \binom{n}{k} + \binom{a_k}{k-1} + \dots + \binom{a_t}{t-1} - \binom{a_k}{k} - \dots - \binom{a_t}{t}.$$

On the other hand

$$\min |\partial(\mathscr{A})| = \min |\delta(\overline{\mathscr{A}}) - \overline{\mathscr{A}}| = \min |\delta(\overline{\mathscr{A}})| - \overline{\mathscr{A}}.$$

Thus, we have to write $2^n - |\mathscr{A}|$ into the n-bounded canonical representation, and $G(n, 2^n - |\mathscr{A}|)$ gives the minimum.

An open question: What is the minimum of $|\delta(\mathscr{A})|$ if $|\mathscr{A}|$ is fixed and $|A|=k$ for $A \in \mathscr{A}$?

REFERENCES

[1] KRUSKAL, J. B.: The number of simplicies in a complex, *Mathematical Optimization Techniques*, University of Calif. Press, Berkeley and Los Angeles, 1963, pp. 251—278.

[2] KATONA, G.: A theorem of finite sets, Theory of Graphs, Proc. Coll. held at Tihany 1966, Akadémiai Kiadó, 1968, pp. 187—207.

[3] HANSEL, G.: Complexes et décompositions binomiales, *J. Combinatorial Th.* **7** (1969) 230—238.

[4] ECKHOFF, J. und WEGNER, G.: Über einen Satz von Kruskal (to appear in Periodica Math. Hung.).

[5] MARGULIS, A. A.: Probabilistic properties of graphs with large connectivity, *Probl. Peredachi i Informacii*, **10** (1974) (2) 101—108.

[6] AHLSWEDE, R., GÁCS P. and KÖRNER, J.: Bounds on conditional probabilities with applications in multi-user communication, *Zeitschrift f. Wahrscheinlichkeitsth. verw. Geb.* (To appear)

[7] CLEMENTS, G. F. and LINDSTRÖM, B.: A generalization of a combinatorial theorem of Macaulay, *J. Combinatorial Th.* **7** (1969) 230—238.

[8] DAYKIN, D. E.: A simple proof of the Kruskal-Katona theorem, *J. Combinatorial The. A.* **17** (1974) 252—253.

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received December 10, 1974)