JOURNAL OF COMBINATORIAL THEORY 6, 402–407 (1969)

Determination of Two Vectors from the Sum

BERNT LINDSTRÖM

Department of Mathematics, University of Stockholm, Sweden

Communicated by N. G. de Bruijn

Received March 6, 1968

ABSTRACT

Let S_m be the set of all vectors of dimension m with all components 0 or 1. Let $\varphi(m)$ be the maximum of $|A + B|$ for pairs A, B of subsets of S_m such that the sums $\mathbf{a} + \mathbf{b}$ are different for different pairs $(\mathbf{a}, \mathbf{b})$, $\mathbf{a} \in A$, $\mathbf{b} \in B$. Let $\lambda(m)$ be the maximum of $|A|$, $A \subset S_m$, such that the sums $\mathbf{a}_1 + \mathbf{a}_2$ are different for different subsets $\{\mathbf{a}_1, \mathbf{a}_2\}$ in A. Let $\nu(m)$ be the maximum of $|A|$, $A \subset S_m$, for A such that the sums $\mathbf{a}_1 + \mathbf{a}_2$ are different modulo 2 for different subsets $\{\mathbf{a}_1, \mathbf{a}_2\}$ in A, $\mathbf{a}_1 \neq \mathbf{a}_2$. The problem is to estimate $\varphi(m)^{1/m}$, $\lambda(m)^{1/m}$ and $\nu(m)^{1/m}$ as $m \to \infty$.

1. INTRODUCTION AND STATEMENT OF RESULTS

Let S_m be the set of all m-dimensional vectors with all components 0 or 1. For any two subsets A and B of S_m define $A + B$ as the set of all sums $\mathbf{a} + \mathbf{b}$, where $\mathbf{a} \in A$ and $\mathbf{b} \in B$. If X is a finite set let $|X|$ be the number of elements in X.

We shall consider pairs A, B of subsets of S_m such that $|A + B| = |A||B|$, i.e., such that the sums $\mathbf{a} + \mathbf{b}$ are distinct for distinct pairs $(\mathbf{a}, \mathbf{b})$ with $\mathbf{a} \in A$ and $\mathbf{b} \in B$. Let $\varphi(m)$ be the maximum of $|A + B|$ for pairs A, B of subsets of S_m with this property. We shall prove the estimates

THEOREM 1.

$$6^{1/2} \leqslant \lim_{m \to \infty} \varphi(m)^{1/m} \leqslant 8^{1/2}.$$

The upper bound will be proved with the aid of information theory. A similar method is used by Katona in [3, see p. 190] and by the author in [5].

Let $\nu(m)$ denote the maximum of $|A|$ for subsets A of S_m such that $\mathbf{a}_1 + \mathbf{a}_2 \equiv \mathbf{a}_3 + \mathbf{a}_4 \pmod 2$ for $\mathbf{a}_1 \neq \mathbf{a}_2$, $\mathbf{a}_3 \neq \mathbf{a}_4$ in A implies $\{\mathbf{a}_1, \mathbf{a}_2\} = \{\mathbf{a}_3, \mathbf{a}_4\}$. We shall prove

THEOREM 2.

$$\lim_{m \to \infty} \nu(m)^{1/m} = 2^{1/2}.$$

Let $\lambda(m)$ be the maximum of $|A|$ for subsets A of S_m such that $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{a}_3 + \mathbf{a}_4$ implies $\{\mathbf{a}_1, \mathbf{a}_2\} = \{\mathbf{a}_3, \mathbf{a}_4\}$ when $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and $\mathbf{a}_4$ are elements in A. We shall prove the estimates

THEOREM 3.

$$2^{1/2} \leqslant \underline{\lim}_{m\to\infty} \lambda(m)^{1/m}, \qquad \overline{\lim}_{m\to\infty} \lambda(m)^{1/m} \leqslant 2^{2/3}.$$

The problem to determine $\lambda(m)$ is analogous to a problem in additive number theory for B_2-sequences (see [2, p. 85]). The lower bound in Theorem 3 can be deduced from a theorem for B_2-sequences by Bose (Theorem 2, p. 81, in [2]). Since $\lambda(m) \geqslant \nu(m)$ the lower bound is a consequence of Theorem 2 above.

The upper bound in Theorem 3 will be proved with a combinatorial argument related to a problem by Zarankiewicz [1]. With the aid of information theory I can now prove

$$\overline{\lim}_{m\to\infty} \lambda(m)^{1/m} \leqslant 2^{3/5} \text{ (added in proof).}$$

We shall not examine generalizations to sums with more than two terms. I shall only mention one result of this kind. Let A be a subset of S_m such that distinct subsets of A have distinct sums. Then if $F(m)$ is the maximum of $|A|$, we have asymptotically (with the aid of Theorem 1 and (1.8) in [6])

$$F(m) \sim \tfrac{1}{2} m \log_2 m, \qquad \text{as} \qquad m \to \infty.$$

2. PROOF OF THEOREM 1

First I prove the existence of $\lim_{m\to\infty} \varphi(m)^{1/m}$. Let m_1 and m_2 be positive integers and assume $|A_i + B_i| = |A_i||B_i| = \varphi(m_i)$ for two pairs of subsets A_i, B_i of $S_{m_i} (i = 1, 2)$. The direct products $A_1 \times A_2$ and $B_1 \times B_2$ are subsets of $S_{m_1+m_2}$. If $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2) \in A_1 \times A_2$ and $\mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2) \in B_1 \times B_2$ then $\mathbf{a} + \mathbf{b} = (\mathbf{a}_1 + \mathbf{b}_1, \mathbf{a}_2 + \mathbf{b}_2)$ determines $\{\mathbf{a}_i, \mathbf{b}_i\}$ $(i = 1, 2)$ uniquely and then $\{\mathbf{a}, \mathbf{b}\}$ is determined by the sum $\mathbf{a} + \mathbf{b}$. It follows

$$\varphi(m_1 + m_2) \geqslant \varphi(m_1)\,\varphi(m_2), \qquad m_1, m_2 \geqslant 1. \tag{2.1}$$

If $\mathbf{a}, \mathbf{b} \in S_m$ all components of $\mathbf{a} + \mathbf{b}$ are 0, 1, or 2. Then we easily obtain the inequality $\varphi(m) \leqslant 3^m$. If we write $a_m = \log 3^m - \log \varphi(m)$ (natural logarithms), then $a_m \geqslant 0$ and $a_{m+n} \leqslant a_m + a_n$ for $m, n \geqslant 1$ by

(2.1). We then conclude that $\lim_{m\to\infty} a_m/m$ exists (see [7, p. 17, Problem 98]) and the existence of $\lim_{m\to\infty} \varphi(m)^{1/m}$ follows easily.

Let $A = \{(0, 0), (1, 1)\}$ and $B = \{(0, 0), (0, 1), (1, 0)\}$. Then $|A + B| = |A||B|$ and we have $\varphi(2) \geqslant 6$ (one easily proves $\varphi(2) = 6$, but we shall not need this). With the aid of (2.1) we then obtain $\varphi(2k) \geqslant 6^k$ and the lower bound in Theorem 1 follows.

The upper bound will be proved with the aid of some information theory (see, e.g., [4]).

Assume $|A + B| = |A||B| = N$ for two subsets A and B of S_m. In the direct product space $A \times B$ we give the probability N^{-1} to each point and get a sample space U with the entropy $H(U) = \log N$ (natural logarithms throughout).

The i-th component of $\mathbf{a} + \mathbf{b}$, where $(\mathbf{a}, \mathbf{b}) \in A \times B$, is a random variable y_i, $i = 1,..., m$. Since $|A + B| = |A||B|$ there is a one-one correspondence between U and the joint distribution space for the sequence $y_1,..., y_m$. By a well-known inequality [4, p. 36], we have

$$\log N = H(U) = H(y_1,..., y_m) \leqslant H(y_1) + \cdots + H(y_m). \tag{2.2}$$

Next we shall prove

$$H(y_i) \leqslant (3/2)\log 2. \tag{2.3}$$

Let x and y be the frequencies of 0's in the i-th component of the vectors in A and B, respectively. The random variable y_i takes the values 0, 1, and 2 with the probabilities

$$\Pr(y_i = 0) = xy, \quad \Pr(y_i = 1) = x + y - 2xy, \quad \Pr(y_i = 2) = (1 - x)(1 - y).$$

We introduce the new variables u and v defined by

$$u = xy, \qquad v = (1 - x)(1 - y). \tag{2.4}$$

The region in the $x - y$ plane defined by $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ is mapped onto a region in the $u - v$ plane defined by the inequalities

$$u \geqslant 0, \qquad v \geqslant 0, \qquad u^{1/2} + v^{1/2} \leqslant 1. \tag{2.5}$$

The last inequality in (2.5) follows easily from (2.4) with the aid of the geometric-arithmetic inequality. Conversely, if (2.5) holds one can determine x and y such that $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ and (2.4) holds.

The entropy of y_i is (put $0 \log 0 = 0$)

$$H(y_i) = -u \log u - v \log v - (1 - u - v)\log(1 - u - v) = f(u, v). \tag{2.6}$$

We shall determine the maximum of $f(u, v)$ in the region defined by (2.5). This maximum is attained on the boundary of the region. For if $f'_u(u, v) = f'_v(u, v) = 0$ then $u = v = 1/3$, which violates the last inequality in (2.5).

If $u = 0$ or $v = 0$ we easily obtain $f(u, v) \leqslant \log 2$. If $u^{1/2} + v^{1/2} = 1$ put $u^{1/2} = t$ and define $g(t)$ for $0 \leqslant t \leqslant 1$ by

$$g(t) = f(t^2, (1-t)^2) = -2t\log t - 2(1-t)\log(1-t) - 2t(1-t)\log 2.$$

Since

$$g(0) = g(1) = 0, \quad g(\tfrac{1}{2}) = 0, \quad \text{and} \quad g''(t) = 4\log 2 - \frac{2}{x(1-x)} < 0,$$

for $0 < t < 1$, it follows that the maximum of $g(t)$ is attained for $t = \frac{1}{2}$, which gives $g(\frac{1}{2}) = (\frac{3}{2})\log 2$. Hence the maximum of $f(u, v)$ in the region defined by (2.5) is $(\frac{3}{2})\log 2$, i.e. (2.3).

From (2.2) and (2.3) it follows that $N \leqslant 2^{3m/2}$ and, since

$$N = |A + B| = |A|\,|B|,$$

we have the upper bound of Theorem 1.

3. Proof of Theorems 2 and 3

Let A be a subset of S_m such that $\mathbf{a}_1 + \mathbf{a}_2 \equiv \mathbf{a}_3 + \mathbf{a}_4 \pmod 2$ implies $\{\mathbf{a}_1, \mathbf{a}_2\} = \{\mathbf{a}_3, \mathbf{a}_4\}$ for $\mathbf{a}_1 \neq \mathbf{a}_2$, $\mathbf{a}_3 \neq \mathbf{a}_4$ in A. If $|A| = n$ it follows easily $\frac{1}{2}n(n-1) \leqslant 2^m$ and

$$\overline{\lim_{m\to\infty}}\, \nu(m)^{1/m} \leqslant 2^{1/2}. \tag{3.1}$$

Next we shall prove

$$\nu(2r) \geqslant 2^r, \qquad r \geqslant 1. \tag{3.2}$$

Let A be the set of all vectors (x, x^3) with $x \in GF(2^r)$. If $(x, x^3) + (y, y^3) = (u, v)$ for two elements $x \neq y$ in $GF(2^r)$, we easily obtain $x + y = u \neq 0$ and $xy = (v/u) - u^2$. Since an equation of the second degree cannot have more than two roots in the field, we find that $\{x, y\}$ is uniquely determined by (u, v). The elements in $GF(2^r)$ are vectors of dimension r with components 0 and 1, which are added modulo 2. Hence A is a set of vectors of dimension $2r$ with components 0 and 1, which are added modulo 2. A has the required property and (3.2) follows.

Since $\nu(m)$ is non-decreasing, we have

$$\varlimsup_{m\to\infty} \nu(m)^{1/m} \geqslant 2^{1/2}.$$

If we take (3.1) into account, Theorem 2 follows. Since $\lambda(m) \geqslant \nu(m)$, we also obtain the lower estimate in Theorem 3.

The upper estimate in Theorem 3 will follow from

$$\lambda(3r) < 2^{2r+1}, \qquad r \geqslant 1. \tag{3.3}$$

But first we shall prove the lemma

LEMMA. *Let $M_1, \ldots, M_k$ be subsets of S with $|S| = h$ and $|M_i \cap M_j| \leqslant 1$ for $i \neq j$. Then it follows*

$$n \leqslant h + \tfrac{1}{2}k(k-1), \qquad \text{where} \quad n = \sum_{i=1}^{k} |M_i|. \tag{3.4}$$

PROOF:

$$h \geqslant \Big|\bigcup_{i=1}^{k} M_i\Big| \geqslant n - \sum_{1\leqslant i<j\leqslant k} |M_i \cap M_j| \geqslant n - \tfrac{1}{2}k(k-1).$$

The second inequality follows from $|M \cup N| = |M| + |N| - |M \cap N|$ by induction over k, the number of sets.

This result and generalizations can be found in [1].

Now we shall prove (3.3). For $m = 3r$, $r \geqslant 1$, let A be a subset of S_m with $|A| = \lambda(3r)$ and such that any set $\{\mathbf{a}, \mathbf{a}'\}$ of elements in A is uniquely determined by $\mathbf{a} + \mathbf{a}'$. We split each $\mathbf{a} \in A$ in two vectors $\mathbf{b}$ and $\mathbf{c}$, where $\mathbf{b}$ consists of the first r components of $\mathbf{a}$ and $\mathbf{c}$ consists of the last $2r$ components of $\mathbf{a}$. Write $\mathbf{a} = (\mathbf{b}, \mathbf{c})$. Let $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$ be an enumeration of all distinct $\mathbf{b}$'s which occur in the $\mathbf{a}$'s of A. Let M_i be the set of all $\mathbf{c}$'s such that $(\mathbf{b}_i, \mathbf{c}) \in A$.

We shall prove that $|M_i \cap M_j| \leqslant 1$ if $i \neq j$. Assume $\mathbf{c}, \mathbf{c}' \in M_i \cap M_j$. Then $(\mathbf{b}_i, \mathbf{c})$, $(\mathbf{b}_i, \mathbf{c}')$, $(\mathbf{b}_j, \mathbf{c})$, and $(\mathbf{b}_j, \mathbf{c}')$ are vectors in A and

$$(\mathbf{b}_i, \mathbf{c}) + (\mathbf{b}_j, \mathbf{c}') = (\mathbf{b}_i, \mathbf{c}') + (\mathbf{b}_j, \mathbf{c}).$$

This is possible only if $\mathbf{c} = \mathbf{c}'$, for $a + a'$ determines $\{a, a'\}$ uniquely in A. Hence $|M_i \cap M_j| \leqslant 1$ if $i \neq j$.

Each M_i is a subset of S_{2r}. The number of elements in S_{2r} is $h = 2^{2r}$. The number of elements in A is $n = |M_1| + \cdots + |M_k|$, and k is at most 2^r. With the aid of the lemma we have $\lambda(3r) = n < 2^{2r+1}$, i.e., (3.3).

The upper estimate in Theorem 3 follows from (3.3) since $\lambda(m)$ is non-decreasing.

ACKNOWLEDGMENT

Thanks are due to the referee for many improvements on my manuscript.

REFERENCES

1. K. ČULIK, Teilweise Lösung eines verallgemeinerten Problems von K. Zarankiewicz, *Ann. Polon. Math.* **3** (1956), 165–168.
2. H. HALBERSTAM AND K. F. ROTH, *Sequences*, Vol. I, Oxford University Press, Oxford and New York, 1966.
3. G. KATONA, On Separating Systems of a Finite Set, *J. Combinatorial Theory* **1** (1966), 174–194.
4. A. I. KHINCHIN, *Mathematical Foundations of Information Theory*, Dover, New York, 1957.
5. B. LINDSTRÖM, On a Combinatory Detection Problem, I, *Publ. Math. Inst. Hungar. Acad. Aci.* **9** (1964), 195–206.
6. B. LINDSTRÖM, On a Combinatorial Problem in Number Theory, *Canad. Math. Bull.* **8** (1965), 477–490.
7. G. PÓLYA AND G. SZEGÖ, *Aufgaben und Lehrsätze aus der Analysis*, Vol. I, 2nd ed., Springer-Verlag, Berlin, 1954.