COLLOQUIA MATHEMATICA SOCIETATIS JÁNOS BOLYAI
4. COMBINATORIAL THEORY AND ITS APPLICATIONS, BALATONFÜRED (HUNGARY), 1969.

Nonaveraging sets II

by

P. Erdős and E. G. Straus
Budapest, Hungary Los Angeles, USA

1. INTRODUCTION

We wish to consider sets of integers $A = \{a_1, \dots, a_n\}$ so that $0 \le a_1 < a_2 < \dots < a_n \le x$ and no a_i is the arithmetic mean of any subset of A consisting of two or more elements. In a previous paper [3] one of us has initiated the study of the maximal number of elements in nonaveraging sets and sets which satisfy related conditions.

Using the notation of [3] we define $f(x)$ as the maximal number of elements in a nonaveraging set; $h(x)$ as the number of elements of a maximal set of integers in the interval $[0, x]$ such that no two distinct subsets have the same arithmetic mean; and $h^*(x)$ as the number of elements of a maximal set of integers in $[0, x]$ such that no two subsets with a relatively prime number of elements have the same arithmetic mean. In [3] we proved $(\log_r x = \log x / \log r)$:

(1.1) $$\log_2 f(x) > \sqrt{2 \log_2 x} + \frac{1}{2} + O(1/\sqrt{\log x})$$

$$(1.2)\qquad (1+o(1))\log x/\log\log x < h(x) < \log_2 x + O(\log\log x)$$

$$(1.3)\qquad \log_2 h^*(x) \geq \sqrt{\log_2 x} - 1 + O(1/\sqrt{\log x})$$

In the present note we prove in § 2 that (1.2) can be replaced by

$$(1.4)\qquad -1 + \log_4 x \leq h(x) < \log_2 x + O(\log\log x)$$

Next, in § 3, we prove that even if we ease the restriction on our sets so that only subsets with different numbers of elements must have different averages then the maximal number, $h^{**}(x)$, of elements satisfies

$$(1.5)\qquad h^{**}(x) < c(\log x)^2 \qquad \text{for some constant } c.$$

Finally in § 4 we get an upper bound for

$$(1.6)\qquad f(x) < cx^{3/4}$$

2. SETS FOR WHICH DIFFERENT SUBSETS HAVE DIFFERENT AVERAGES

The new lower bound (1.4) for $h(x)$ is obtained inductively as follows. Assume that $h(x) = k$ and that $\{a_1, \dots, a_k\}$ is a set with the desired property. A $(k+1)$-st element a_{k+1} must satisfy $< 4^{k+1}$ inequalities

$$(2.1)\qquad t(a_{i_1} + \dots + a_{i_s}) \neq s(a_{j_1} + \dots + a_{j_t})$$

for each pair of subsets $\{a_{i_1}, \dots, a_{i_s}\}$ and $\{a_{j_1}, \dots, a_{j_t}\}$ of $\{a_1, \dots, a_{k+1}\}$ (where at least one of the elements $a_{i_1}, \dots, a_{i_s}, a_{j_1}, \dots, a_{j_s}$ must be a_{k+1}).

Hence it is possible to choose a_{k+1} in the interval $[0, 4^{k+1}]$ so that $h(4^{k+1}) \geq k+1$ and in general $h(x) \geq [\log_4 x] > \log_4 x - 1$.

The conjecture has been communicated to us that the sequence $\{3^k\}$ has the property discussed in this section. On the other hand it is easy to see that the sequence $\{2^k\}$ does not. Nevertheless it seems likely to us that correct asymptotic value is

$$h(x) = (1 + o(1)) \log_2 x. \tag{2.2}$$

3. SETS FOR WHICH SUBSETS OF DIFFERENT CARDINALITY HAVE DIFFERENT AVERAGES

Let $h^{**}(x) = n$ and let $\{a_1, \dots, a_n\}$ be a set with the desired property. We form all possible subsets $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ and consider the $\binom{n}{k}$ sums $a_{i_1} + \dots + a_{i_k}$. Since all these sums are in the interval $\{0, kx\}$ there are at least

$$\binom{n}{k} / kx = N \tag{3.1}$$

such sums which have a common value t.

Now there cannot exist three subsets $\{a_{i_1}, \dots, a_{i_k}\}$, $\{a_{j_1}, \dots, a_{j_k}\}, \{a_{\ell_1}, \dots, a_{\ell_k}\}$ whose sums are t and whose index sets have pairwise equal intersections. For otherwise, deleting the common part from all three sets, the average of the first set will equal the average of the union of the second and third set.

However, according to a theorem of P. Erdős and R. Rado [1, Theorem III] it follows that whenever

$$N \geq k!\, 3^{k+1} \tag{3.2}$$

there do exist three sets with pairwise equal intersections. Combining (3.1) and (3.2) we get

$$\binom{n}{k} < k \cdot k!\, 3^{k+1} x \tag{3.3}$$

and hence, if we choose $k = \frac{1}{2} \log x$ we have

$$\frac{(n-k)^k}{k!} < \binom{n}{k} < \left((1+o(1))\, \frac{3k}{c}\right)^k x$$

or

$$n-k < (1+o(1))\, \frac{3k^2}{c^2}\, x^{1/k} < \left(\frac{3}{4} + o(1)\right) (\log x)^2$$

so that

$$h^{**}(x) < \left(\frac{3}{4} + o(1)\right)(\log x)^2. \tag{3.4}$$

It was conjectured in [1] that Theorem III could be improved so that the lower bound in (3.2) could be replaced by c^k for a suitable c. In that case it would follow that $h^{**}(x) < c_1 \log x$ for suitable c_1 and this would, according to § 2, give the correct order of magnitude.

4. UPPER BOUNDS FOR THE NUMBER OF ELEMENTS IN A NONAVERAGING SET

As was pointed out in [3] we can find upper bounds for $f(x)$ by finding upper bounds for the maximum number, $F(x)$ of elements in two sets of integers A, B in $[0, x]$ so that A and B have the same number of elements and the sums of elements of nonempty subsets of A and B are distinct. It was conjectured there that $F(x) \leq \sqrt{2x}$ and it was observed that $f(x) \leq 2F(x)+1$.

In this section we want to obtain upper bounds for $F(x)$. Assume $n = F(x) \geq cx^{3/4}+1$. Then there are at least $\frac{1}{2}c^2x^{3/2}$ sums $a_i + a_j$, $a_i, a_j \in A$ with $1 \leq i < j \leq n$ all lying in the interval $(0, 2x)$. Thus there exists an integer with more than $\frac{1}{4}c^2x^{1/2}$ representations of the form a_i+a_j and similarly there exists an integer with more than $\frac{1}{4}c^2x^{1/2}$ representations of the form $a_i + a_j$.

Now let M be the maximal integer which has at least $\frac{1}{8}c^2x^{1/2}$ representations either as a_i+a_j or as $b_i + b_j$. Without loss of generality we may assume that it has representations of the form $b_i + b_j$. There can be no more than $\frac{1}{2}cx^{3/4}$ elements of A which exceed M since otherwise more than $\frac{1}{4}c^2x^{3/2}$ sums a_i+a_j would exceed M and therefore some integer greater than M would have more than $\frac{1}{8}c^2x^{1/2}$ representations as a_i+a_j contrary to hypothesis. Thus there are more than $\frac{1}{2}cx^{3/4}$ elements of A below M.

According to a theorem of Szemerédi [4] which will appear in Acta Arithmetica (for a slightly weaker result see Ryavec [2]) it follows that if

$$\frac{1}{2}cx^{3/4} > c_1M^{1/2} \tag{4.1}$$

then there exists a sum of distinct elements of A (all less than M) which is divisible by M. Here c_1 is an absolute constant.

Now if

$$a_{i_1} + \cdots + a_{i_m} = LM < mM < c_1M^{3/2} \tag{4.2}$$

then $L < c_1M^{1/2} < 2c_1x^{1/2}$ (since $M < 2x$). Thus, if

$$\frac{1}{8}c^2x^{1/2} \geq 2c_1x^{1/2} > L \tag{4.3}$$

then LM can be represented as a sum of different elements of A contrary to hypothesis when (4.2) is satisfied. Thus we get

$$c \leq \max\{2c_1, 4c_1^{1/2}\}$$

and $F(x) < cx^{3/4} + 1, \quad f(x) < 2cx^{3/4} + 3$.

Assume that the following result holds: Let $a_1, \dots, a_k$, $k > c_1 M^{1/2}$ be k distinct residues $\bmod M$. Then there are ℓ distinct a's, satisfying

(4.4) $$a_{i_1} + \dots + a_{i_\ell} \equiv 0 \pmod M, \quad \ell < c_2 M/k$$

This result combined with the above proof gives immediately $F(x) < c_3 x^{2/3}$. Szemerédi just informed us that his proof gives (4.4) without any difficulty.

As we conjectured in [3] it is probable that

$$f(x) = \exp(c\sqrt{\log x}) = o(x^{\varepsilon}) .$$

The method of estimating $F(x)$ can of course yield nothing better than $F(x) = O(x^{1/2})$ and hence $f(x) = O(x^{1/2})$.

REFERENCES

[1] P. ERDŐS and R. RADO, Intersection theorems for systems of sets, J. London Math. Soc. 35 (1960), 85-90.

[2] C. RYAVEC, The addition of residue classes modulo n, Pacific J. of Math. 26 (1968), 367-373.

[3] E.G. STRAUS, Nonaveraging sets, Proc. Symposia in Pure Math. A.M.S. (1967)

[4] E. SZEMERÉDI, On a conjecture of Erdős and Heillronn, Acta Arithmetica