ACTA ARITHMETICA
LXXIV.3 (1996)

Sure monochromatic subset sums

by

NOGA ALON (Tel Aviv) and PAUL ERDŐS (Budapest)

1. Introduction. For an integer $n > 1$ let $f(n)$ denote the smallest integer f such that one can color the integers $\{1, \ldots, n-1\}$ by f colors so that there is no monochromatic subset the sum of whose elements is n. Paul Erdős [2] asked if for every positive ε, $f(n) > n^{1/3-\varepsilon}$ for all $n > n_0(\varepsilon)$. In this note we prove that this is indeed the case, in the following more precise form.

THEOREM 1.1. *There exist positive constants c_1, c_2 so that*
$$c_1 \frac{n^{1/3}}{\log^{4/3} n} \le f(n) \le c_2 \frac{n^{1/3}(\log\log n)^{1/3}}{\log^{1/3} n}$$
for all $n > 1$.

We suspect that the upper bound is closer to the actual value of $f(n)$ than the lower bound but this remains open. The (simple) proof of the upper bound is described in Section 2. The lower bound is established in Section 3.

To simplify the presentation, we omit all floor and ceiling signs, whenever these are not essential. We make no attempt to optimize the absolute constants throughout the paper. For a set of integers A, let A^* denote the set of all sums of subsets of A.

2. The upper bound. Given n, we prove that
$$f(n) \ll \frac{n^{1/3}(\log\log n)^{1/3}}{\log^{1/3} n}$$
by exhibiting an explicit family of subsets of $N = \{1, \ldots, n-1\}$ whose union covers N, so that $n \notin A^*$ for each subset A in the family. Define
$$s = \frac{n^{1/3}(\log\log n)^{1/3}}{\log^{1/3} n}.$$
For each integer k satisfying $1 \le k \le s$, let $A_k = \{i \in N : n/(k+1) \le i < n/k\}$. Note that $n \notin A_k^*$, since the sum of any set of at most k members

of A_k is less than n whereas the sum of any set of at least $k+1$ members of A_k exceeds n. For each prime $p \le s$ that does not divide n define $B_p = \{i \in N : p \mid i\}$. Since all members of B_p^* are divisible by p it follows that $n \notin B_p^*$. It is well known (see, e.g., [4]) that Brun's sieve method gives that for any set P of primes which are all at most m, the number of integers between 1 and m which are not divisible by any member of P does not exceed $O(m \prod_{p \in P}(1-1/p))$. It follows that there is an absolute constant c so that the number, call it S, of integers in N not covered by the union of all sets A_k and B_p above satisfies

$$S \le c \frac{n}{s \log n \prod_{p\mid n,\, p \le s}(1-1/p)}.$$

(Note that all these integers are smaller than n/s.) However, it is easy to check that

$$\prod_{p\mid n,\, p\le s}(1-1/p) \gg \frac{1}{\log\log n},$$

showing that

$$S \ll \frac{n^{2/3}(\log\log n)^{2/3}}{\log^{2/3} n}.$$

We can now split the set of these remaining integers arbitrarily into $\lceil S/s \rceil$ sets C_j of size at most s each. Since each member of C_j is at most n/s, $n \notin C_j^*$ for any C_j. The sets A_k, B_p and C_j together cover N, and their total number is at most

$$O\left(\frac{n^{1/3}(\log\log n)^{1/3}}{\log^{1/3} n}\right),$$

completing the proof of the upper bound in Theorem 1.1. ∎

3. The lower bound. The proof of the lower bound is based on the following result of Sárközy [5] (see also [3] and [1] for similar results).

Theorem 3.1 ([5], Theorem 4). *Let $m > 2500$ be an integer, and let A be a subset of $\{1,\dots,m\}$ of cardinality $|A| = 1000(m\log m)^{1/2}$. Then there are integers d, y, z such that $1 \le d \le 10m^{1/2}/\log^{1/2} m$, $z > 10m\log m$, and $y < z/(10\log m)$, such that $\{yd, (y+1)d, (y+2)d, \dots, zd\} \subset A^*$.*

We also need the following simple lemma.

Lemma 3.2. *Let d be a positive integer, and let B be a set of $d-1$ positive integers, all relatively prime to d. Then for any integer x, B^* contains a member congruent to x modulo d.*

Proof. Let $B = \{b_1, \dots, b_{d-1}\}$ and define $b_i' = b_i \pmod d$, $B_i = \{b_1', \dots, b_i'\}$. Then B_i is a subset of the cyclic group Z_d. Let B_i^* denote the set of all sums of subsets of B_i, computed in Z_d. Our objective is to prove that $B^*_{d-1} = Z_d$. Note that $B^*_1 = \{0, b'_1\}$ and $B^*_i = B^*_{i-1} \cup (B^*_{i-1} + b'_i)$, where the sum is computed in Z_d. If for some i, $|B^*_i| = |B^*_{i-1}|$, then for every $b \in B^*_{i-1}$, $b + b'_i$ is also in B^*_{i-1}, and since $0 \in B^*_{i-1}$ and b'_i generates Z_d, $B^*_{i-1} = Z_d$, as needed. Otherwise, $|B^*_i| > |B^*_{i-1}|$ for all i, and hence $B^*_{d-1} = Z_d$, completing the proof. ∎

COROLLARY 3.3. *Let $C \subset \{1, \ldots, m\}$ be a set of* ***primes*** *of cardinality*

$$|C| = 1000(m \log m)^{1/2} + 20\frac{m^{1/2}}{\log^{1/2} m} + k,$$

where $m > 2500$. Let S denote the sum of the largest k members of C. Then any integer t satisfying $200m^{3/2}/\log^{1/2} m \le t \le S$ lies in C^*.

Proof. Let A denote the set of the $1000(m \log m)^{1/2}$ smallest members of C. By Theorem 3.1 there are d, y, z as in the theorem, so that $yd, (y+1)d, \ldots, zd$ are all in A^*. Thus, in particular,

$$zd \le m|A| \le 1000m^{3/2} \log^{1/2} m. \tag{1}$$

Let B be the set of the $20m^{1/2}/\log^{1/2} m$ smallest members of $C - A$.

CLAIM. *Every integer x satisfying $yd + md \le x \le zd$ lies in $(A \cup B)^*$.*

Proof. B contains at least $d-1$ elements larger than d, and all of them are relatively prime to d. Therefore, by Lemma 3.2, there is a number x' which is the sum of at most $d-1$ members of B and $x' \equiv x \pmod d$. Clearly $x' \le md$ and thus $zd \ge x \ge x - x' \ge yd$. Since $x - x'$ is divisible by d it lies in A^*, implying that $x \in B^* + A^* = (A \cup B)^*$, as needed.

Returning to the proof of the corollary let I denote the interval of all integers between $yd + md$ and zd, and let $x_1, \ldots, x_k$ be all elements in $C - (A \cup B)$. Then the length of I is at least $zd/2 \ge 5m \log m > m$ and all the $k+1$ intervals $I, I + x_1, I + (x_1 + x_2), \ldots, I + (x_1 + \ldots + x_k)$ lie in C^*. The union of these intervals contains all the integers t satisfying $yd + md \le t \le S + zd$, and the desired result follows from (1), since

$$yd \le \frac{zd}{10 \log m} \le 100\frac{m^{3/2}}{\log^{1/2} m},$$

and $md \le 10m^{3/2}/\log^{1/2} m$. ∎

COROLLARY 3.4. *For all sufficiently large n, and for any set C of at least $200n^{1/3} \log^{2/3} n$* ***primes*** *between $n^{2/3} \log^{1/3} n/200$ and $n^{2/3} \log^{1/3} n/100$, the number n lies in C^*.*

Proof. Apply the previous corollary with $m = n^{2/3}\log^{1/3} n/100$. Here

$$k > 50n^{1/3}\log^{2/3} n,\ \ 200m^{3/2}/\log^{1/2} m < n \ \text{ and } \ S > k\frac{n^{2/3}\log^{1/3} n}{200} > n,$$

implying that indeed $n \in C^*$. ∎

Proof of Theorem 1.1 (lower bound). Clearly we may assume that n is sufficiently large, by an appropriate choice of c_1. Given a large n, and a coloring of $\{1,\ldots,n-1\}$ by $f = f(n)$ colors without a monochromatic subset whose sum is n, there is, by the prime number theorem, a monochromatic set containing at least

$$(1+o(1))\frac{3n^{2/3}}{2f\cdot 200\log^{2/3} n}$$

primes between $n^{2/3}\log^{1/3} n/200$ and $n^{2/3}\log^{1/3} n/100$. By the last corollary, this number cannot exceed

$$200n^{1/3}(\log n)^{2/3},$$

implying the assertion of the theorem. ∎

References

[1] N. Alon and G. Freiman, *On sums of subsets of a set of integers*, Combinatorica 8 (1988), 297–306.

[2] P. Erdős, *Some of my recent problems in Combinatorial Number Theory, Geometry and Combinatorics*, in: Graph Theory, Combinatorics and Applications, Proceedings of the Seventh Quadrennial International Conference on the Theory and Application of Graphs, Y. Alavi and A. Schwenk (eds.), Wiley Interscience, New York, 1995, 335–350.

[3] G. Freiman, *New analytical results in subset-sum problem*, Discrete Math. 114 (1993), 205–218; Erratum, ibid. 126 (1994), 447.

[4] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press, 1974.

[5] A. Sárközy, *Finite addition theorems, II*, J. Number Theory 48 (1994), 197–218.

DEPARTMENT OF MATHEMATICS
RAYMOND AND BEVERLY SACKLER FACULTY
OF EXACT SCIENCES
TEL AVIV UNIVERSITY
TEL AVIV, ISRAEL
E-mail: NOGA@MATH.TAU.AC.IL

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
P.O.B. 127
H-1364 BUDAPEST
HUNGARY

Received on 15.5.1995 (2795)