

THE SUM OF THE RECIPROCAL OF A SET OF INTEGERS WITH NO ARITHMETIC PROGRESSION OF k TERMS¹

JOSEPH L. GERVER

ABSTRACT. It is shown that for each integer $k > 3$, there exists a set S_k of positive integers containing no arithmetic progression of k terms, such that $\sum_{n \in S_k} 1/n > (1 - \varepsilon)k \log k$, with a finite number of exceptional k for each real $\varepsilon > 0$. This result is shown to be superior to that attainable with other sets in the literature, in particular Rankin's sets $\mathcal{A}(k)$, which have the highest known asymptotic density for sets of positive integers containing no arithmetic progression of k terms.

Let S_k be any set of positive integers which contains no arithmetic progression of k terms. Erdős and Davenport [1] proved that for any such S_k ,

$$\liminf_{n \rightarrow \infty} |S_k \cap [1, n]|/n = 0,$$

where $|X|$ denotes the cardinality of X , and $[1, n]$ the set of integers from 1 to n inclusive. More recently, Szemerédi [2] proved that

$$\lim_{n \rightarrow \infty} |S_k \cap [1, n]|/n = 0.$$

On the other hand, it has been shown by Behrend [3] and Moser [4] in the case $k = 3$, and by Rankin [5] for all $k > 3$, that there exist sets S_k , with no arithmetic progression of k terms, such that, for all positive integers n ,

$$|S_k \cap [1, n]| > n \exp[-c(\log n)^b]$$

where b and c are positive numbers which depend on k but not on n .

These results have led Erdős [6] to conjecture that $\sum_{n \in S_k} 1/n$ must converge. If this conjecture is true, then for each $k \geq 3$, there exists

$$A_k = \sup_{\text{all } S_k} \sum_{n \in S_k} 1/n.$$

For suppose A_k did not exist for some k . Let $S_k(1)$ be any set of positive integers containing no arithmetic progression of k terms, and, for each integer $m \geq 1$, let $a_k(m)$ be the least integer such that

Received by the editors March 22, 1976.

AMS (MOS) subject classifications (1970). Primary 10L10; Secondary 10H20.

Key words and phrases. Arithmetic progression, sum of reciprocals, asymptotic density, Szemerédi's theorem, Rankin's sets.

¹The above research was supported in part by a NATO postdoctoral fellowship and in part by a grant of one hundred francs from Paul Erdős, who also kindly provided me with some references to previous work in this field.

© American Mathematical Society 1977

$$\sum_{n \in S_k(m) \cap [1, a_k(m)]} 1/n \geq \sum_{n \in S_k(m)} (1/n) - 1$$

and let $S_k(m + 1)$ be any set of positive integers with no arithmetic progression of k terms such that

$$\sum_{n \in S_k(m+1)} 1/n \geq \log a_k(m) + \log 2 + \gamma + 2.$$

Then $U_{k+1} = \cup_{m=1}^{\infty} S_k(m) \cap [1, a_k(m)]$ has no arithmetic progression of $k + 1$ terms, yet $\sum_{n \in U_{k+1}} 1/n$ diverges.

When Erdős made this conjecture, the best lower bound known to him for A_k was $\frac{1}{2} k \log 2$, based on the theorem that 2^k consecutive integers can be partitioned into two sets, neither of which contains an arithmetic progression of k terms [7]. We will prove here that $A_k > (1 - \epsilon)k \log k$, for every $\epsilon > 0$, with a finite number of exceptional k for each ϵ .

THEOREM. *For every $\epsilon > 0$, there exist, for all but a finite number of integers $k \geq 3$, sets S_k of positive integers, containing no arithmetic progression of k terms, such that $\sum_{n \in S_k} 1/n > (1 - \epsilon)k \log k$.*

PROOF. Let Z^+ be the set of positive integers. For any prime p , let

$$S_p = Z^+ - \{n: \exists i, j \in Z^+, jp^i - p^{i-1} + 1 \leq n \leq jp^i\}.$$

Then S_p contains no integers divisible by p , so by the Chinese remainder theorem, it contains no arithmetic progression with p terms, provided the intervals between the terms are not divisible by p . For any positive integers i and m , such that $m \leq p^i$, we have $S_p \cap \{n: n \equiv m \pmod{p^i}\}$ equal to $\{p^i(n - 1) + m: n \in S_p\}$ if m is an element of S_p , and equal to the null set if m is not an element of S_p . Therefore, one can prove, by induction on i , that S_p contains no arithmetic progression with p terms even if the intervals are divisible by p . Now

$$\sum_{n \in S_p} \frac{1}{n} > 1 + \sum_{i=0}^{\infty} \left(\frac{p-1}{p}\right)^i \sum_{n=p^{i+1}}^{p^{i+1}-p^i} \frac{1}{n}.$$

But

$$\sum_{n=p^{i+1}}^{p^{i+1}-p^i} \frac{1}{n} = \sum_{n=p^i}^{p^{i+1}-1} \frac{1}{n} - \frac{1}{p^i} - \sum_{n=p^{i+1}-p^{i+1}}^{p^{i+1}-1} \frac{1}{n} > \log p - \frac{1}{p^i} - \frac{1}{p-1},$$

so

$$\begin{aligned} \sum_{n \in S_p} \frac{1}{n} &> 1 + \log p \sum_{i=0}^{\infty} \left(\frac{p-1}{p}\right)^i - \frac{1}{p-1} \sum_{i=0}^{\infty} \left(\frac{p-1}{p}\right)^i - \sum_{i=0}^{\infty} \left(\frac{p-1}{p^2}\right)^i \\ &= 1 + p \log p - \frac{p}{p-1} - \frac{p-1}{p^2-p+1} > p \log p - \frac{2}{p-1}. \end{aligned}$$

By a theorem of Huxley [8], there always exists a prime between k and $k - k^{5/8}$. Therefore, if p is the greatest prime less than k then $\sum_{n \in S_p} 1/n >$

$(1 - \epsilon)k \log k$, with only a finite number of exceptional k for each $\epsilon > 0$. Q.E.D.

One might think that this result could be improved by using Rankin's sets. However, that is not the case, because Rankin's constant c depends so strongly on k . Indeed $c > \frac{1}{3} \sqrt{k} \log k$, and $b > \log 2 / (\log k + \log 2)$. It is not difficult to show that if R_k is a set of positive integers, and

$$|R_k \cap [1, n]| \leq n \exp\left[-\frac{1}{3} \sqrt{k} \log k (\log n)^{\log 2 / (\log k + \log 2)}\right]$$

for all n , then $\sum_{n \in R_k} 1/n$ is bounded from above by a constant which does not depend on k . For

$$\begin{aligned} \sum_{n \in R_k} 1/n &< 1 + \int_1^\infty x^{-1} \exp\left[-\frac{1}{3} \sqrt{k} \log k (\log x)^{\log 2 / (\log k + \log 2)}\right] dx \\ &= 1 + \int_0^\infty \exp\left[-\frac{1}{3} \sqrt{k} (\log k) u^{\log 2 / (\log k + \log 2)}\right] du \\ &= 1 + \left(\frac{1}{3} \sqrt{k} \log k\right)^{-(\log k + \log 2) / \log 2} \int_0^\infty \exp\left[-u^{\log 2 / (\log k + \log 2)}\right] du. \end{aligned}$$

But $\int_0^1 \exp\left[-u^{\log 2 / (\log k + \log 2)}\right] du < 1$, and

$$\int_1^\infty \exp\left[-u^{\log 2 / (\log k + \log 2)}\right] du < \int_1^\infty \exp\left[-2 \log u + a\right] du = e^a,$$

where a is a constant, depending only on k . We can solve for a by finding the maximum value of $2 \log u - u^{\log 2 / (\log k + \log 2)}$. It turns out that

$$a = \frac{2(\log k + \log 2)}{\log 2} \left[\log\left(\frac{\log k + \log 2}{\log 2}\right) + \log 2 - 1 \right].$$

Therefore,

$$\begin{aligned} \sum_{n \in R_k} 1/n &< 1 + \left(\frac{1}{3} \sqrt{k} \log k\right)^{-(\log k + \log 2) / \log 2} \\ &\quad + \left[\frac{1}{12} \sqrt{k} (\log k) e^2 \left(\frac{\log 2}{\log k + \log 2}\right)^2 \right]^{-(\log k + \log 2) / \log 2}. \end{aligned}$$

If one considers this expression as a function of a real variable k , then it is continuous for $k \geq 3$, and tends to 1 for large k , so it is bounded from above.

It is evident that the problem of maximizing $\sum_{n \in S_k} 1/n$ (or, since it is not known that this sum always converges, maximizing $\sum_{n \in S_k \cap [1, m]} 1/n$ for large m) is by no means equivalent to that of maximizing the asymptotic density of S , which is to say, maximizing $\sum_{n \in S_k \cap [1, m]} 1$ for large m . It might turn out that it is easier to find an upper bound on $\sum_{n \in S_k \cap [1, m]} 1/n$ than on $\sum_{n \in S_k \cap [1, m]} 1$ directly; that is, the best way to prove that $\sum_{n \in S_k} 1/n$ converges might be to prove that it is less than a specific value. A possible approach might be as follows: The set $S_p = Z^+ - \{n: \exists i, j, jp^i - p^{i-1} + 1 \leq n \leq jp^i\}$ can be constructed inductively; $n \in S_p$ if and only if the union of $\{n\}$

with $S_p \cap [1, n - 1]$ contains no arithmetic progression of p terms. This is a poor way to construct sets with high asymptotic density, since no account is taken of the possibility that by omitting an element at some point, it might be possible to include two extra elements later. In the case of sets S_k with high $\sum_{n \in S_k} 1/n$, however, there is a heavy penalty for postponing the inclusion of an element, so the set S_p defined above may indeed be maximal when k is prime. Indeed, I have been unable to improve on it even marginally. Clearly this set does maximize $\sum_{n \in S_p} 1/2^n$, over all S_p , and it might be possible to extend this result to $\sum_{n \in S_p} f(n)$ for other $f(n)$, even if not to $\sum_{n \in S_p} 1/n$. If one could prove that $\sum_{n \in S_k} 1/n \log \log n$ always converges, then this would suffice to show that the primes contain arithmetic progressions with an arbitrary number of terms.

REFERENCES

1. H. Davenport and P. Erdős, *On sequences of positive integers*, Acta Arith. **2** (1937), 147–151.
2. E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245. MR **51** #5547.
3. F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332. MR **8**, 317.
4. L. Moser, *On non-averaging sets of integers*, Canad. J. Math. **5** (1953), 245–252. MR **14**, 726; erratum p. 1278.
5. R. A. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, Proc. Roy. Soc. Edinburgh Sect. A **65** (1960/61), 332–344. MR **26** #95.
6. Unpublished lecture, Faculté des Sciences, Paris, December 4, 1975.
7. E. R. Berlekamp, *A construction for partitions which avoid long arithmetic progressions*, Canad. Math. Bull. **11** (1968), 409–414. MR **38** #1066.
8. M. N. Huxley, *The distribution of prime numbers*, Oxford Math. Monograph, 1972, p. 119.

MATHEMATIQUES, BATIMENT 425, UNIVERSITE DE PARIS-SUD, 91405 ORSAY, FRANCE

Current address: Department of Mathematics, University of Hawaii, Honolulu, Hawaii 96822