Small Sets Which Meet All the $k(n)$-Term Arithmetic Progressions in the Interval $[1,n]$

Tom C. Brown and Allen R. Freedman

Citation data: T.C. Brown and A.R. Freedman, *Small sets which meet every $f(n)$-term arithmetic progressions in the interval* $[1,n]$, J. Combin. Theory Ser. A **51** (1989), 244–249.

Abstract

For given n,k, the minimum cardinal of any subset B of $[1,n]$ which meets all of the k-term arithmetic progressions contained in $[1,n]$ is denoted by $f(n,k)$. We show, answering questions raised by Professor P. Erdős, that $f(n,n^{\varepsilon}) < C\cdot n^{1-\varepsilon}$ for some constant C (where C depends on ε) and that $f(n,\log n) = o(n)$. We also discuss the behavior of $f(p^2,p)$, where p is a prime, and we give a simple lower bound for the function associated with Szemerédi's theorem.

1 Introduction

Let n,k be positive integers. We define $f(n,k)$ to be the minimum cardinal of any subset B of $[1,n]$ which meets all of the k-term arithmetic progressions contained in $[1,n]$. For example, $f(9,3)=4$, since the set $B=\{2,5,6,7\}$ meets every 3-term arithmetic progression contained in $[1,9]$, and no smaller subset B of $[1,9]$ has this property. Professor Erdős [2] has asked whether $f(n,n^{\varepsilon}) \le C\cdot n^{1-\varepsilon}$ for some constant $C=C(\varepsilon)$, and whether $f(n,\log n)=o(n)$. We answer these questions below, in the affirmative. (Here we are considering $[n^{\varepsilon}]$-term arithmetic progressions and $[\log n]$-term arithmetic progressions, respectively.)

Note that $[n/k] \le f(n,k)$ for all n and k, since $[1,n]$ contains $[n,k]$ pairwise disjoint blocks of k consecutive integers.

If we regard k as a constant, then Szemerédi's theorem [3] gives a definitive statement about the behavior of $f(n,k)$ for large n, namely that $f(n,k)=n-o(n)$. However, if $k(n)$ is a function of n which increases sufficiently rapidly with n, then it can happen that

$$[n/k(n)] \le f(n,k(n)) \le Cn/k(n) \quad \text{for all } n,$$

where C is a constant.

We will show, for example, that for any fixed ε, $0<\varepsilon<1$,

$$n^{1-\varepsilon} \le f(n,n^{\varepsilon}) \le (12/\varepsilon)\cdot n^{1-\varepsilon}, \quad \text{for all } n.$$

On the other hand, it is not hard to constuct (using Szemerédi's theorem) a function $k(n)$ which goes to infinity with n but which increases so slowly that $(1/n)\cdot f(n,k(n))$ approaches 1 as n approaches infinity.

(Define $n_2 < n_3 < \cdots$ by setting $n_2 = 1$ and choosing n_k so that $f(n,k) > (1-1/k)\cdot n$ for all $n > n_k$. Then , for each $k \geq 2$, set $k(n) = k$ for $n_k \leq n < n_{k+1}$.)

We will show that $f(n,\log n) = o(n)$, but we do not know if $f(n,\log n) = O(n/\log n)$. The most slowly growing functions $k(n)$ for which we can show $f(n,k(n)) = o(n)$ are the functions $k(n) = (\log n)/(\log\log n)^{\varepsilon}$, for $\varepsilon < 1$.

We also discuss the behavior of $f(p^2,p)$ where p is a prime, and we give a lower bound for the function, analogous to the van der Waerden numbers, associated with the "finite form" of Szemerédi's theorem.

2 Asymptotic results

Lemma. *If p is prime, $p \geq 3$, $t \geq 0$, and $p^t \leq n < p^{t+1}$, then*

$$f(n,p) \leq 3tn/p.$$

Proof. First we consider the case $p^t \leq n < p^{t+1} - p^t$, where $t \geq 1$. (The case $t = 0$ is trivial.) For each j, $0 \leq j \leq t-1$, let

$$B_j = \{x \in [1,n] : x \equiv i \pmod{p^{j+1}}, 1 \leq i \leq p^j\}.$$

Now let $a + dp^j x$, $0 \leq x \leq p-1$, $(d,p) = 1$, be a p-term arithmetic progression contained in $[1,n]$. Then $j \leq t-1$, since otherwise the largest term of the progression, $a + dp^j(p-1)$, will fall outside the interval $[1,n]$.

We will show that this progression meets the set B_j. Choose i, $1 \leq i \leq p^j$, so that $a \equiv i \pmod{p^j}$, say $a - i = sp^j$. Next choose x_0, $0 \leq x_0 \leq p-1$, so that $s + dx_0 \equiv 0 \pmod{p}$. Then $a + dp^j x_0 \equiv i \pmod{p^{j+1}}$, which means that $a + dp^j x_0$ is in B_j.

We now know that $B_0 \cup B_1 \cup \cdots \cup B_{t-1}$ meets every p-term arithemtic progression contained in $[1,n]$. From

$$|B_j| \leq p^j([n/p^{j+1}]+1) \leq (n/p) + p^j \leq 2n/p,$$

we get

$$f(n,p) \leq |B_0| + |B_1| + \cdots + |B_{t-1}| \leq 2tn/p.$$

Note that for the special case $n = p^t$, we have $|B_j| = n/p = p^{t-1}$, so that

$$f(p^t,p) \leq tp^{t-1}.$$

The remaining case is $p^{t+1} - p^t \leq n < p^{t+1}$ ($t \geq 1$). Here, we use the preceding remark to get

$$f(n,p) \leq f(p^{t+1},p) \leq (t+1)p^t \leq 2tp^t \leq 3(1-(1/p))tp^t \leq 3tn/p. \qquad \square$$

Theorem 1. *Let $k(n)$ be any function. Then, whenever $k(n) \geq 4$, we have*

$$f(n,k(n)) \leq \frac{12n\log n}{k(n)\log k(n)}.$$

Proof. For $k(n) \geq 4$, there is a prime p and a non-negative integer t such that, using Bertrand's postulate,

$$3 \leq p \leq k(n) \leq 2p \quad \text{and} \quad p^t \leq n < p^{t+1}.$$

By the lemma, $(n,k(n)) \leq f(n,p) \leq 3tn/p$. Now $t \leq (\log n)/(\log p)$, $1/p \leq 2/k(n)$ and $1/\log p \leq 1/(\log k(n) - \log 2) \leq 2/\log k(n)$. The result follows. □

Corollary 1. *If* $\log n = o(k(n)\log k(n))$, *then* $f(n,k(n)) = o(n)$.

Applications. (a) Let $k(n) = n^\varepsilon$, $0 < \varepsilon < 1$. Then $f(n,n^\varepsilon) \leq (12/\varepsilon)n^{1-\varepsilon}$, for all n (Note $(12/\varepsilon)n^{1-\varepsilon} \leq n$ implies $\log n \geq 4$.)

(b) When $k(n) = \log n$, $f(n,\log n) \leq 12n/\log\log n$, for all n. (Note $12/\log\log n \leq n$ implies $\log\log n \geq 4$.)

(c) Letting $k(n) = (\log n)/(\log\log\log n)$ or the smaller function $(\log n)/(\log\log n)^\varepsilon$ for $0 < \varepsilon < 1$, we get functions $k(n) = o(\log n)$ such that $f(n,k(n)) = o(n)$. Note the corollary does not apply to $k(n) = (\log n)/(\log\log n)$.

3 Other results

Theorem 2. *For every odd prime p,*

$$f(p^2,p) \leq 2p - 2.$$

For every constant C,

$$p + C \leq f(p^2,p)$$

for infinitely many primes p.

Proof. For an odd prime p, let

$$B = \{kp : 1 \leq k \leq p-2\} \cup [p^2 - p - 1, p^2 - 2].$$

Then $|B| = 2p-2$ and B meets every p-term arithmetic progression in $[1,p^2]$. Indeed, there is only one such progression with common difference $p+1$ and it contains the element p^2-p-1. Every progression with common difference p meets the interval p^2-p-1, p^2-2. Finally, every progression of common difference less than p must contain an element congruent to 0 mod p. If this element happens to be p^2 or p^2-p, then the given progression meets the interval $[p^2-p-1, p^2-2]$ since $p \geq 3$. Otherwise it meets the set $\{kp : 1 \leq k \leq p-2\}$. This proves the first assertion.

To prove the second assertion, let C be a fixed positive integer. We suppose that for all large primes p there is a set $A \subseteq [1,p^2]$ such that $|A| \leq p+C$, and A meets every p-term arithmetic progression in $[1,p^2]$. Consider the blocks $B_i = [ip+1, (i+1)p]$ for $i = 0,1,\ldots,p-1$. Each B_i contains at least one element of A. Also, each residue mod p is congruent to at least one member of A. Call a block B_i "good" if $B_i \cap A$ is a singleton $\{a\}$ and the residue of a mod p is unique (i.e., for all $a' \in A - \{a\}$, $a \not\equiv a' \bmod p$). An easy count shows that the number of good blocks is not less that $p - 3C$ and so there must be a consecutive string of good blocks, $B_{u+1}, B_{u+2}, \ldots, B_{u+t}$ of length $t \geq (p-3C)/(3C+1)$. Let $M = 2(3C+1)$ and consider the primes $p \equiv -1 \bmod (M+1)!$. Note that $t \geq M+1$ (for p sufficiently large). Let $B_{u+i} \cap A = \{a_i\}$

and denote the $t-1$ "jumps" by $j_i = a_{i+1} - a_i$. We claim that each j_i is less than $p-M$. Write $j = j_i$. If $j \geq p+1$, then there are p consecutive integers which do not meet A. If $j = p$, then a_i and a_{i+1} are congruent mod p. If $j = p-r$, for $1 \leq r \leq M$, then $j \equiv 0 \bmod (r+1)$ and there will thus be a missing residue mod $(r+1)$ among the elements a_k in a consecutive string of $r+1$ good blocks which contains the blocks B_{u+i} and B_{u+i+1}. This implies the existence of a p-term arithmetic progression (with common difference $r+1$) which does not meet A.

The proof is concluded with the following contradiction: We have $(t-2p)p < a_t - a_1 = j_1 + j_2 + \cdots + j_{t-1} < (t-1)(p-M)$ which reduces to $tM < p+M$. This implies $((p-3C)/(3C+1))M = 2p - 6C < p+M$, which is false for $p \geq 12C+2$. □

Theorem 3. *For each ε, $0 < \varepsilon < 1$, and each positive integer k, let $g(k,\varepsilon)$ denote the smallest positive integer such that if $m \geq g(k,\varepsilon)$, $[1,m] \supseteq A$ and $|A| > \varepsilon m$, then A must contain a k-term arithmetic progression. (Thus $g(k,\varepsilon)$ is the number whose existence is asserted by Szemerédi's theorem.) Then for every prime p and every ε, $0 < \varepsilon < 1$,*

$$g(p,\varepsilon) > p^{[(p-1)\log(1/\varepsilon)]}.$$

Also, if $\varepsilon < 1/e$ then $g(p,\varepsilon) > p^p$ for sufficiently large p. In particular,

$$g(p,1/3) > p^p \quad \text{for all } p \geq 7.$$

(This means: for every prime $p \geq 7$, there is a subset A of $[1,p^p]$ such that $|A| > \frac{1}{3}p^p$ and A contains no p-term arithmetic progression.)

Proof. For a given positive integer n, let A be the set of all integers x in $[0,p^n-1]$ such that when x is expressed as an n-digit p-ary number, none of the n digits is 0. Then A contains no p-term arithmetic progression. (By considering the first non-zero digit in the p-ary form of the common difference of a given p-term arithemtic progression, one easily sees that some term of the progression contains a zero in p-ary form.) Clearly $|A| = (p-1)^n$. Thus by the definition of $g(k,\varepsilon)$, if $(p-1)^n > \varepsilon p^n$, then $g(p,\varepsilon) > p^n$. Now if $n \leq (p-1)\log(1/\varepsilon)$, then $n\log(1+1/(p-1)) < n/(p-1) \leq \log(1/\varepsilon)$, so $n\log(p/(p-1)) + \log\varepsilon < 0$, or $\varepsilon p^n < (p-1)^n$, so that $g(k,\varepsilon) > p^n$. Taking $n = [(p-1)\log(1/\varepsilon)]$ we get

$$g(p,\varepsilon) > p^{[(p-1)\log(1/\varepsilon)]}.$$

Finally, if $\varepsilon < 1/e$ then $\varepsilon < ((p-1)/p)^p$ for large p, so that $(p-1)^p > \varepsilon p^p$ and $g(p,\varepsilon) > p^p$. For $\varepsilon = 1/3$, the inequalities hold for all $p \geq 7$. (In the same way, if $\varepsilon < 1/e^k$ then $g(p,\varepsilon) > p^{kp}$ for large p.) □

4 Remarks

1. Theorem 1 shows that the functions $k(n) = (\log n)/g(n)$, where $g(n) = o(\log\log n)$, grow rapidly enough that $f(n,k(n)) = o(n)$. One naturally would like to find the boundary between those functions $k(n)$ for which $f(n,k(n)) = o(n)$ and those functions $k(n)$ for which $f(n,k(n))$ is not $o(n)$. In

particular, one would like to know whether or not $f(n,(\log n)/(\log\log n)) = o(n)$ and whether or not $f(n,\log\log n) = o(n)$. Naturally, if $k(n) \le h(n)$ and $f(n,k(n)) = o(n)$, then $f(n,h(n)) = o(n)$, since $f(n,h(n)) \le f(n,k(n))$.

However, the statement that $f(n,\log\log n)$ is not $o(n)$ is stronger that Szemerédi's theorem. In fact, given any function $k(n)$ which goes to infinity with n, the statement that $f(n,k(n))$ is not $o(n)$ is stronger than Szemerédi's theorem. This is a consequence of Behrend's theorem [1]. Indeed, if $f(n,k(n)) \ne o(n)$, then there exists an $\varepsilon > 0$ such that for infinitely many n,

$$[B \subseteq [1,n], |B| < \varepsilon n] \Rightarrow [B \text{ does not meet some } k(n)\text{-term A.P.}].$$

Then for the same set of (infinitely many) n,

$$[A \subseteq [1,n], |A| > (1-\varepsilon)n] \Rightarrow [A \text{ contains a } k(n)\text{-term A.P.}].$$

Now let k be an arbitrary positive integer. Choose n_0 so that the preceding implication holds for $n = n_0$ and such that $k(n_0) \ge k$. It easily follows that for all $n \ge 2n_0/\varepsilon$,

$$[A \subseteq [1,n], |A| > (1-\varepsilon/2)n] \Rightarrow [A \text{ contains a } k\text{-term A.P.}].$$

This is exactly the hypothesis of Behrend's theorem, and Szemerédi's theorem is the conclusion.

2. The constant "12" which appears in Theorem 1 can be decreased to "$2+\varepsilon$" (at the cost of replacing "whenever $k(n) \ge 4$" by "for all sufficiently large $k(n)$") by noting that in the Lemma we have $f(n,p) \le (2+\varepsilon)tn/p$ for sufficiently large p, by using $1/(\log k(n) - \log 2) \le (1+\varepsilon)\log k(n)$ for sufficiently large $k(n)$, and by using the Prime Number Theorem instead of Bertrand's postulate. Then one obtains

$$f(n,k(n)) \le \frac{(2+\varepsilon)n\log n}{k(n)\log k(n)},$$

for all sufficiently large $k(n)$.

On the other hand, the method of Theorem 1 also give $f(n,k(n)) \le 18n\log n/k(n)\log k(n)$, whenever $k(n) \ge 3$.

Note added in proof. Professor John Truss has improved Theorem 2 to $f(n^2,n) > n + n^{1/2}/2$, for all n.

References

[1] F. Behrend, *On sequences of integers containing no arithmetic progression*, Časopis Mat. Fys. Praha (Část Mat.) **67** (1938), 235–239.

[2] Paul Erdős, *Personal communication*.

[3] E. Szemerédi, *On sets of integers containing no k elements in an arithmetic progression*, Acta. Arith. **27** (1975), 199–245, Collection of articles in memory of Jurii Vladimirovic Linnik.