ON SETS OF INTEGERS NOT CONTAINING ARITHMETIC PROGRESSIONS OF PRESCRIBED LENGTH

H. L. ABBOTT, A. C. LIU, and J. RIDDELL

(Received 11 April 1972)

Communicated by G. Szekeres

Let m, n and l be positive integers satisfying $m \geqq n \geqq l \geqq 3$. Denote by $h(m,n,l)$ the largest integer with the property that from every n-subset of $\{1,2,\cdots,m\}$ one can select $h(m,n,l)$ integers no l of which are in arithmetic progression. Let $f(n,l) = h(n,n,l)$ and let $g(n,l) = \min_m h(m,n,l)$. In what follows, by a P_l-free set we shall mean a set of integers not containing an arithmetic progression of length l.

It has been conjectured that $f(n,l) = o(n)$ for each fixed l, but this has been proved only in the cases $l = 3$ and 4, by Roth [8] and Szemerédi [8] respectively. Szekeres had conjectured (see [3; p. 223]) that to each l there corresponds a number $\alpha_l < 1$ such that $f(n,l) = O(n^{\alpha_l})$. This however was proved false by Salem and Spencer [9] who proved that $f(n,3) > n^{1-(\log 2+\varepsilon)/\log\log n}$ for every $\varepsilon > 0$ provided n is large enough. Improvements, refinements and extensions of this result were obtained by Behrend [2], Moser [5] and Rankin [6]. Rankin proved that if

(1) $$2^s < l \leqq 2^{s+1} \text{ and } c(s,\varepsilon) = (s+1)2^{s/2}(\log 2)^{s/(s+1)}(1+\varepsilon)$$

then

(2) $$f(n,l) > n^{1-c(s,\varepsilon)/(\log n)^{s/(s+1)}}$$

provided n is sufficiently large.

As far as the function g is concerned, Riddell [7] proved that $g(n,l) > cn^{1-2/l}$ and that $g(n,3) > cn^{1/2}$. Erdös has informed us that Szemerédi has recently proved $g(n,3) > n^{1-\varepsilon}$ for every $\varepsilon > 0$ provided n is sufficiently large. Szemerédi's proof has not yet been published. We observe that while $g(n,l) \leqq f(n,l)$, it is only in the case $l = 3$, with $n = 5$ or 14, that strict inequality is known to hold. The sets 1, 3, 4, 5, 7 and 1, 3, 4, 5, 7, 8, 9, 11, 12, 13, 15, 16, 17, 19 illustrate that $g(5,3) = 3 < f(5,3) = 4$ and $g(14,3) \leqq 7 < f(14,3) = 8$. The values of $f(n,3)$ for $n \leqq 52$ have been computed by Wagstaff [11].

With regard to the function h, Riddell [7] proved that if $a \geqq 1$,

$$h([n^a], n, 3) > n^{1-3\sqrt{2a\log 2/\log n} - 3\log 2/\log n}. \tag{3}$$

In [7] Riddell proves also that if $m \geqq n^3$ then almost all n-subsets of $\{1, 2, \cdots, m\}$ contain a P_3-free subset of cardinality at least $n^{1-(3\sqrt{6\log 2}+\varepsilon)/\sqrt{\log n}}$.

The arguments used in [7] do not generalize to the case $l > 3$. It is to this question that we turn our attention in this paper. We prove the following two results which improve and extend the results in [7].

THEOREM 1. *Let $l \geqq 3$ be given and let s and $c(s, \varepsilon)$ be defined by* (1). *Then for $m \geqq m_0(s, \varepsilon)$*

$$h(m, n, l) > nm^{-c(s,\varepsilon)/(\log m)^{s/(s+1)}}. \tag{4}$$

THEOREM 2. *Almost all sets of n integers from $\{1, 2, \cdots, m\}$ contain a P_l-free subset of cardinality at least $n^{1-c(s,\varepsilon)/(\log n)^{s/(s+1)}}$ provided $m \geqq n^{1+\varepsilon}$ and n is sufficiently large.*

PROOF OF THEOREM 1. Let A be a P_l-free subset of $\{1, 2, \cdots, m\}$. We assume that A is maximal so that $|A| = f(m, l)$. If λ is an integer then by $A + \lambda$ is meant $\{a + \lambda \mid a \in A\}$. Let $\lambda_1 = 0$, and after numbers $\lambda_1, \lambda_2, \cdots, \lambda_r$ have been defined, select λ_{r+1} so that $A + \lambda_{r+1}$ contains the largest number of elements in $\{1, 2, \cdots, m\}$ that do not belong to $A + \lambda_j$ for $j = 1, 2, \cdots, r$. Let k be the first integer such that $\bigcup_{j=1}^{k} A + \lambda_j \supseteq \{1, 2, \cdots, m\}$. It was proved in [1], using a modification of an argument of Lorentz [4], that

$$k \leqq \frac{2m+1}{f(m,l)} \sum_{j=1}^{f(m,l)} \frac{1}{j}. \tag{5}$$

Since the argument is not long we present the proof of (5) here. Let $M = \{1, 2, \cdots, m\}$ and let $A_\lambda = M \cap (A + \lambda)$. Let $B_1 = A_{\lambda_1}$, and for $r \geqq 2$ let $B_r = A_{\lambda_r} - \bigcup_{i=1}^{r-1} A_{\lambda_i}$. Let $z = f(m, l)$ and define numbers $t(z)$, $t(z-1), \cdots, t(1)$, $t(0)$ recursively as follows: $t(z)$ is the largest integer such that $|B_\mu| = z$ for $\mu = 1, 2, \cdots, t(z)$. After the numbers $t(z)$, $t(z-1), \cdots, t(u+1)$ have been defined ($u \geqq 1$) let $t(u)$ be the largest positive integer for which $|B_\mu| = u$ for

$$\sum_{i=u+1}^{z} t(i) < \mu \leqq \sum_{i=u}^{z} t(i),$$

provided such a positive integer exists. If there is no such positive integer, put $t(u) = 0$. Finally put $t(0) = 0$. It is clear that

$$k = \sum_{u=1}^{z} t(u). \tag{6}$$

Now define a sequence of subsets $M_z, M_{z-1}, \cdots, M_1, M_0$ of M as follows: $M_z = M$ and for $1 \leqq u \leqq z - 1$,

$$M_u = \left\{ a \,\middle|\, a \in M, a \notin A_{\lambda_i} \text{ for } i = 1, 2, \cdots, \sum_{j=u+1}^{z} t(j). \right\}$$

Let M_0 be the empty set. Then clearly, for $1 \leqq u \leqq z$,

$$|M_{u-1}| = |M_u| - ut(u).$$

Equivalently,

$$t(u) = \frac{1}{u}(|M_u| - |M_{u-1}|). \tag{7}$$

From (6) and (7) we get

$$k = \sum_{u=1}^{z} \frac{1}{u}(|M_u| - |M_{u-1}|) = \sum_{u=1}^{z-1} \frac{|M_u|}{u(u+1)} + \frac{|M_z|}{z}. \tag{8}$$

For each λ, $|\lambda| \leqq m$, we have $|A_\lambda \cap M_u| \leqq u$ and hence

$$\sum_{\lambda=-m}^{m} |A_\lambda \cap M_u| \leqq (2m+1)u. \tag{9}$$

Since each $r \in M_u$ belongs to exactly z of the sets A_λ we have

$$\sum_{\lambda=-m}^{m} |A_\lambda \cap M_u| = z|M_u|. \tag{10}$$

From (9) and (10) we get

$$\frac{|M_u|}{u} \leqq \frac{2m+1}{z} \tag{11}$$

and from (8) and (11) it follows that

$$k \leqq \frac{2m+1}{z} \sum_{u=1}^{z} \frac{1}{u}$$

which is (5).

Now let $S \subseteq M$, $|S| = n$. Then for some j, $1 \leqq j \leqq k$ we must have $|S \cap (A + \lambda_j)| \geqq n/k$. Since arithmetic progressions are invariant under translation $S \cap (A + \lambda_j)$ is P_l-free. Hence we have

$$h(m, n, l) \geqq n/k. \tag{12}$$

It now follows from (2), (5) and (12) that (4) holds and hence Theorem 1 is proved.

REMARK 1. If we take $m = [n^a]$ and $l = 3$ in (4) we get

$$h([n^a], n, 3) > n^{1 - 2\sqrt{2a \log 2(1+\varepsilon)/\log n}},$$

which is an improvement over (3).

https://doi.org/10.1017/S144678870001990X Published online by Cambridge University Press

REMARK 2. One can also ask questions of the following type: What is the size of a maximal P_l-free set that can be chosen from some set of integers that arises in a "natural way"? We mention only one example. It follows from our theorem that one can select from the set of the first n primes a P_l-free subset of cardinality at least $n^{1-c(s,\varepsilon)/(\log n)^{s/(s+1)}}$. This can be seen by taking m to be the nth prime and appealing to the prime number theorem.

Before proving Theorem 2 we shall need to prove some lemmas which are extensions of results given in [7]. By a $P^{(l)}$ set of intervals we shall mean a set of intervals

$$X_j = (u+(x_j-1)v,\ u+x_jv],\ j = 1, 2, \cdots, r$$

where $\{x_1, x_2, \cdots, x_r\}$ is a P_l-free set of integers and where u and v are real numbers, $v > 0$. Put $\underline{X}_j = (u+(x_j-1)v,\ u+(x_j-\frac{1}{2})v]$ and put $\bar{X}_j = (u+(x_j-\frac{1}{2})v,\ u+x_jv]$.

LEMMA 1. *$\bigcup_{j=1}^r \underline{X}_j$ contains no l-term arithmetic progression with terms in different intervals; similarly for $\bigcup_{j=1}^r \bar{X}_j$.*

PROOF. The proof in the case $l = 3$ is given in [7]. The argument for $l \geqq 3$ is similar. Suppose $\bigcup_{j=1}^r \underline{X}_j$ contains an l-term arithmetic progression. If two terms of this arithmetic progression lie in an interval $\underline{X}_j$ then all of the terms must belong to $\underline{X}_j$ since the common difference of the arithmetic progression is less than the distance between intervals. The only other possibility is that the l terms of the arithmetic progression are in l different intervals, say $\underline{X}_{j_1}, \underline{X}_{j_2}, \cdots, \underline{X}_{j_l}$. However, this implies that $x_{j_1}, x_{j_2}, \cdots, x_{j_l}$ form an arithmetic progression and this is a contradiction. The same argument applies to $\bigcup_{j=1}^r \bar{X}_j$.

LEMMA 2. *If a set of numbers has elements in each interval of a $P^{(l)}$ set of r intervals, then it contains a P_l-free subset of cardinality at least $[(r+1)/2]$.*

PROOF. This follows easily from Lemma 1.

LEMMA 3. *Let t be an integer, $t \leqq m$. Let $w = mt^{-1}$. Let $b(k,n)$ be the number of n-subsets of $\{1, 2, \cdots, m\}$ that have elements appearing in fewer than k of the intervals*

$$(0, w],\ (w, 2w],\ (2w, 3w],\ \cdots,\ ((t-1)w, tw]. \tag{13}$$

Then

$$b(k,n) < \frac{(w+1)^n t^k k^{n+1}}{n!}. \tag{14}$$

PROOF. Denote by $f(j)$ the number of n-subsets of $\{1, 2, \cdots, m\}$ which have elements in exactly j of the intervals (13). Then

$$f(j) \leqq \binom{t}{j} \sum_{b_1+b_2+\cdots+b_j=n} \binom{[w+1]}{b_1}\binom{[w+1]}{b_2}\cdots\binom{[w+1]}{b_j}$$

where the summation is over all compositions of n into exactly j parts. From the inequality $\binom{a}{b} \leqq \frac{a^b}{b!} \leqq a^b$ and the multinomial theorem we get

$$f(j) \leqq t^j \sum_{b_1+b_2+\cdots+b_j=n} \frac{(w+1)^n}{b_1!b_2!\cdots b_j!} \leqq \frac{(w+1)^n t^j j^n}{n!}$$

Now we can estimate $b(k,n)$. We have

$$b(k,n) \leqq \sum_{j=1}^{k-1} f(j) \leqq \frac{(w+1)^n}{n!} \sum_{j=1}^{k-1} t^j j^n < \frac{(w+1)^n t^k k^{n+1}}{n!}$$

and this is (14). Hence Lemma 3 is proved.

If we take $t = [n^{1+\varepsilon}]$ in Lemma 3, put $k = [\varepsilon n/(1+\varepsilon)]$ and impose the condition $0 < \varepsilon < 1/(2e-1)$, we get, after some routine calculations,

$$b(k,n) = o\left(\binom{m}{n}\right).$$

Thus we have a further lemma.

LEMMA 4. *Let* $0 < \varepsilon < 1/(2e-1)$. *Let* $k = [\varepsilon n/(1+\varepsilon)]$, $t = [n^{1+\varepsilon}]$, $m \geqq t$ *and* $w = mt^{-1}$. *Then almost all n-subsets of* $\{1, 2, \cdots, m\}$ *have elements occurring in at least k of the intervals* (13).

PROOF OF THEOREM 2. Let ε, m, t, n and k satisfy the conditions of Lemma 4. Let S be an n-subset of $\{1, 2, \cdots, m\}$, and suppose S has elements in at least k of the intervals (13). By Lemma 4, almost all n-subsets S will have the latter property. At least $h(t,k,l)$ ofthese t intervals form a $P^{(l)}$ set of intervals, and hence, by Lemma 2, S contains a P_l-free set of cardinality at least $[\{h(t,k,l)+1\}/2]$. It follows from this and Theorem 1 that S contains a P_l-free set of cardinality at least $(k/2)t^{-c(s,\varepsilon)/(\log t)^{s/(s+1)}}$. This implies Theorem 2.

REMARK. 3. If we take $l = 3$ we find that almost all n-subsets of $\{1, 2, \cdots, m\}$ contain a P_3-free set of cardinality at least $n^{1-2\sqrt{2\log 2(1+\varepsilon)/\log n}}$ provided $m \geqq n^{1+\varepsilon}$ and n is sufficiently large. This result is sharper than the corresponding result in [7].

References

[1] H. L. Abbott and A. C. Liu, 'On partitioning integers into progression free sets', *J. Comb. T.* (to appear).

[2] F. A. Behrenα, 'On sets of integers which contain no three terms in arithmetical progression', *Proc. Nat. Acad. Sci. U. S. A.* 32 (1946), 331–332.

[3] P. Erdös, 'Some recent advances and current problems in number theory', *Lectures on Modern Mathematics*, Vol. 3, T. L. Saaty, ed. (Wiley, New York, 1963).

[4] G. G. Lorentz, 'On a problem of additive number theory', *Proc. Amer. Math. Soc.* 5 (1954), 838–841.

[5] L. Moser, 'On non-averaging sets of integers', *Can. J. Math.* 5 (1953), 245–252.

[6] R. A. Rankin, 'Sets of integers containing not more than a given number of terms in arithmetical progression', *Proc. Roy. Soc. Edin, A*, 65 (1962) 332–344.

[7] J. Riddell, 'On sets of integers containing no *l* terms in Arithmetic progression', *Neiuw Arch. voor Wisk.* (3), 17 (1969), 204–209.

[8] K. F. Roth, 'Sur quelques ensembles d'entiers', *C. R. Acad. Sci. Paris.* 234 (1952), 388-390.

[9] R. Salem and D. C. Spencer, 'On sets of integers which contain no three terms in arithmetical progression', *Proc. Nat. Acad. Sci. U. S.A.* 28 (1942), 561–563.

[10] E. Szemerédi, 'On sets of integers containing no four terms in arithmetic progression', *Acta. Math. Acad. Sci. Hung.* 20 (1969), 89–104.

[11] S. S. Wagstaff, Jr., 'On sequences of integers with no 4, or no 5 numbers in arithmetical progression', *Math. Comp.* 21 (1967), 695–699.

University of Alberta, Edmonton, Alberta

University of Victoria, Victoria, British Columbia