ARITHMETIC PROGRESSIONS IN FINITE SETS OF REAL NUMBERS

by W. KLOTZ

(Received 13 September, 1971)

In this paper we investigate the structure of a set of n reals that contains a maximal number of l-term arithmetic progressions. This problem has been indicated by J. Riddell. Let l and n be positive integers with $2 \leqq l \leqq n$. By $F_l(n)$ we denote the maximal number of l-term arithmetic progressions that a set of n reals can contain. A set of n reals containing $F_l(n)$ l-progressions will be called an $F_l(n)$-set.

THEOREM. *Let* $[x]$ *be the integral part of* x. *Then*

$$F_l(n) = \tfrac{1}{2}(l-1)\left[\frac{n-1}{l-1}\right]\left(\left[\frac{n-1}{l-1}\right]-1\right)+\left[\frac{n-1}{l-1}\right]\left(n-(l-1)\left[\frac{n-1}{l-1}\right]\right). \tag{1}$$

Every arithmetic progression of n terms is an $F_l(n)$*-set. Let* $n-1 \equiv r \bmod (l-1)$, $0 \leqq r \leqq l-2$. *For* $l \geqq 3$ *there are exactly the following types of* $F_l(n)$*-sets besides arithmetic progressions. (If* $n \geqq l \geqq 3$ *then every* $F_l(n)$*-set that is not a progression corresponds under a linear transformation to one of the sets listed below. The digit* 1 *stands for an element of the set,* 0 *for a gap.)*

(i) $l = 3$.

(a) $r = 1$.

1 1 1 ... 1 1 1 0 1,
1 0 1 1 1 ... 1 1 1 0 1 0 1,
. . . .
$\underbrace{1\,0\,1\,0 \ldots 1\,0}_{\frac{n-4}{2}}\;1\,1\,1\;\underbrace{0\,1\,0\,1 \ldots 0\,1\,0\,1}_{\frac{n-2}{2}}$. *elements*

(b) $r = 0$.

1 0 1 1 1 ... 1 1 1 0 1,
1 0 1 0 1 1 1 ... 1 1 1 0 1 0 1,
. . . .
$\underbrace{1\,0\,1\,0\,1\,0 \ldots 1\,0}_{\frac{n-3}{2}}\;1\,1\,1\;\underbrace{0\,1 \ldots 0\,1\,0\,1\,0\,1}_{\frac{n-3}{2}}$. *elements*

A

https://doi.org/10.1017/S0017089500001828 Published online by Cambridge University Press

(ii) $l \geqq 4$.

(a) $r = l-2$.

1 1 1 ... 1 1 1 0 1.

(b) $r \leqq l-3$.

Only progressions.

Proof. Let $P_l(n)$ denote the number of l-progressions contained in $\{1, 2, \ldots, n\}$. Then $P_l(n) \leqq F_l(n)$. We shall prove equality. For practical reasons we put $P_1(n) = F_1(n) = 0$ for every $n \geqq 0$ and $P_l(n) = F_l(n) = 0$ for $n < l$. Clearly $P_2(n) = F_2(n) = \binom{n}{2}$, if $n \geqq 2$. Assume that $P_{l-2}(n) = F_{l-2}(n)$ for every $n \geqq 0$ has already been shown. We count the l-progressions in an arbitrary $F_l(n)$-set $S = \{a_1, \ldots, a_n\}$ with $a_1 < a_2 < \ldots < a_n$ and $l \leqq n$. First we count all progressions with a second term a_i, with $1 < i \leqq 1+(n-1)/(l-1)$. Then we take the progressions having a last but one term a_j, with $n-(n-1)/(l-1) < j < n$. Deleting the first and the last term of every l-progression not yet counted gives a $1-1$ mapping into the set of $(l-2)$-progressions contained in

$$S' = \left\{a_i \mid a_i \in S,\ 1+\frac{n-1}{l-1} < i \leqq n-\frac{n-1}{l-1}\right\}.$$

Thus we see that

$$F_l(n) \leqq \sum_{1 < i \leqq 1+(n-1)/(l-1)} (i-1) + \sum_{n-(n-1)/(l-1) < i < n} (n-i) + F_{l-2}(|S'|). \tag{2}$$

We recall that $F_{l-2}(|S'|) = P_{l-2}(|S'|)$. If S is an n-progression, the right-hand side of (2) counts every l-progression exactly once. This shows that $F_l(n) = P_l(n)$.

We can now compute $F_l(n)$ as the number of l-progressions in $\{1, 2, \ldots, n\}$. We find the number of l-progressions having the first term j, and sum over j.

$$F_l(n) = \sum_{j=1}^{n} \left[\frac{n-j}{l-1}\right] = \left[\frac{n-1}{l-1}\right]\left(n-(l-1)\left[\frac{n-1}{l-1}\right]\right) + (l-1) \sum_{k=1}^{\left[\frac{n-1}{l-1}\right]-1} k.$$

This proves (1).

For the following we consider only $F_l(n)$-sets S. We keep in mind that in (2) equality holds. Therefore S' is an $F_{l-2}(|S'|)$-set. Moreover, every $(l-2)$-progression in S' can be extended to an l-progression in S by adding a new first and a new last element. For every $a_i \in S$, with $1 < i \leqq 1+(n-1)/(l-1)$ (resp. $n-(n-1)/(l-1) < i < n$), and any $a_j < a_i$ (resp. $a_i < a_j$), there must be an l-progression in S containing a_i and a_j such that a_i is its second (resp. its last but one) term.

(i) $l = 3$.

(a) $r = 1$. Consider the elements $a_{i_0}, a_{i_0+1} \in S$ ($i_0 = n/2$). Every $a_i < a_{i_0}$ and every $a_j > a_{i_0+1}$ can be reflected at a_{i_0} and a_{i_0+1} resp. to an element of S. From this, and because there are no elements between a_{i_0} and a_{i_0+1}, it follows that the elements of S can only take positions on the infinite arithmetic progression defined by a_{i_0} and a_{i_0+1}. Thus we can represent

S by zeros and ones on this progression. Recognizing that every gap $< a_{i_0}$ (resp. $> a_{i_0+1}$) can be reflected at a_{i_0+1} (resp. at a_{i_0}) to another gap we find out all the sets listed in part (i)(a) of the theorem. From the construction of these sets it is clear that they contain as many 3-progressions as counted on the right-hand side of (2). So they are indeed $F_3(n)$-sets.

(b) $r = 0$. The proof is analogous to that for (a).

(ii) $l \geqq 4$.

It might happen that $|S'| < l-2$. For $r = 0$, this would mean that

$$|S'| = n - \frac{n-1}{l-1} - \frac{n-1}{l-1} - 1 \leqq l-3$$

and so $n \leqq l$; while, for $r \geqq 1$,

$$|S'| = n - \frac{n-1-r}{l-1} - \frac{n-1-r}{l-1} - 2 \leqq l-3$$

and so $n \leqq l+1+\dfrac{2(1-r)}{l-3} \leqq l+1$. But for $4 \leqq l \leqq n \leqq l+1$ the arithmetic progression of n terms is the only $F_l(n)$-set.

Let $|S'| \geqq l-2$. As S' is an $F_{l-2}(|S'|)$-set, we may assume for $l \geqq 5$ that S' is either a progression or homothetic to one of the sets listed in the theorem. But every $(l-2)$-progression in S' has an extension to an l-progression in S. Therefore S' is a progression itself. This is also true for $l = 4$.

We represent S' by the positive integers $i+1, i+2, \ldots, j-1$ (and j, if $r \neq 0$), and we show that, if $r \neq 0$, then S is of the form

$$S = \{1-g, \ldots, 0, 1, \ldots, i, i+1, \ldots, j-1, j, \ldots, n, n+1, \ldots, n+h\},$$

with

$$g, h \geqq 0,\ i = \frac{n-1-r}{l-1} + 1,\ j = n - \frac{n-1-r}{l-1},\ a_t = t \quad \text{for} \quad i+1 \leqq t \leqq j-1.$$

$i \in S$, because it belongs to the extension of the $(l-2)$-progression $i+1, \ldots, i+l-2$ of S'. Clearly $i \leqq a_i$. If $i < a_i$, then it follows from $a_i + (a_i - i) \in S$ that $a_i = i + \frac{1}{2}$. But as $l \geqq 4$ we must also have $a_i + 2(a_i - i) \in S$, which is not true. Therefore $a_i = i$ must hold. Similarly we see that every element of S is an integer. In particular, we find that $a_j = j$. Let $a_1 = 1-g$ and $a_n = n+h$ $(g, h \geqq 0)$. The l-progression having $n+h$ as its last term and j as its last but one term is contained in S. This is also true for $r = 0$. Let its first term be denoted by $(n+h)'$. Then

$$(n+h)' = n - \frac{n-1-r}{l-1} - (l-2)\left(n+h-n+\frac{n-1-r}{l-1}\right) = r+1-h(l-2).$$

If $(n+h)' < 1-h$, then the l-progression starting with $(n+h)'$ and i would demand the existence of an element of S greater than $n+h$. Therefore we conclude that $r+1-h(l-2) \geqq 1-h$,

$$h \leqq r/(l-3), \quad \text{and similarly} \quad g \leqq r/(l-3) \quad \text{for} \quad r = 0, 1, \ldots, l-2. \tag{3}$$

If $r \leqq l-4$, then $g = h = 0$, and S is a progression.

https://doi.org/10.1017/S0017089500001828 Published online by Cambridge University Press

Suppose that $0 \in S$ and $1 \in S$; then every integer up to i must belong to S, which is impossible, because S has only i elements not greater than i. Thus we see that not both 0 and 1 or analogously n and $n+1$ can belong to S. Now let $r = l-2$, $h = 1$. We have

$$(n+1)' = l-2+1-(l-2) = 1$$

and so $S = \{1, 2, \ldots, n-1, n+1\}$. This is an $F_l(n)$-set. According to (3), for $l=4$ we need consider only the following cases:

(α) $r = l-2$, $h = 2$. We have

$$(n+2)' = l-2+1-2(l-2) = -1.$$

Here the elements $-1, i, j, n+2$ form a 4-progression. This means that there is a progression counted twice on the right-hand side of (2). Therefore S cannot be an $F_4(n)$-set.

(β) $r = l-3$, $h = 1$. We have

$$(n+1)' = l-3+1-(l-2) = 0$$

and so $S = \{0, 2, 3, \ldots, n-1, n+1\}$. But now $0, i, 2i, \ldots, j = (l-2)i, n+1 = (l-1)i$ form an l-progression. For the same reason as before S cannot be an $F_l(n)$-set.

Finally we pose the following problem. Given a system of linear equations

$$\sum_{j=1}^{r} a_{ij} x_j = 0, \sum_{j=1}^{r} a_{ij} = 0, \quad \text{for} \quad i = 1, 2, \ldots, s,$$

with integer coefficients a_{ij}. Denote by $F(n)$ the maximal number of solutions in a set of n reals and by $P(n)$ the number of solutions in $\{1, 2, \ldots, n\}$. Is $F(n) \sim P(n)$ always true?

REFERENCE

1. J. Riddell, On sets of numbers containing no l terms in arithmetic progression, *Nieuw Arch. Wisk.* (3) **17** (1969), 204–209.

Institut für Mathematik
Technische Universität
3392 Clausthal, Germany

https://doi.org/10.1017/S0017089500001828 Published online by Cambridge University Press