ON SOME SEQUENCES OF INTEGERS

PAUL ERDÖS *and* PAUL TURÁN*.

[*Extracted from the Journal of the London Mathematical Society, Vol.* 11, 1936.]

Consider a sequence of integers $a_1 < a_2 < \ldots \leqslant N$ containing no three terms for which $a_i - a_l = a_l - a_s$, *i.e.* a sequence containing no three consecutive members of an arithmetic progression. Such sequences we call A sequences belonging to N, or simply A sequences. We consider those with the maximum number of elements, and denote by $r = r(N)$

* Received 6 June, 1936; read 18 June, 1936.

the number of elements of such maximum sequences. In this paper we estimate $r(N)$.

THEOREM I. $r(2N) \leqslant N$ *if* $N \geqslant 8$.

Remark. It is interesting to observe that, as we shall see, the theorem is true for $N = 4, 5, 6$, but not for $N = 7$.

Proof. First we observe that, if $a_1 < a_2 < \ldots < a_r$ represents an A sequence belonging to N, then

$$N+1-a_r < N+1-a_{r-1} < \ldots < N+1-a_1 \tag{1}$$

is also an A sequence.

The same holds for

$$a_1-k < a_2-k < \ldots < a_r-k, \tag{2}$$

for any integer $k < a_1$.

Hence, evidently,

$$r(m+n) \leqslant r(m)+r(n). \tag{3}$$

We prove Theorem I by induction. Consider first the case $N = 4$. If we have $r(8) = 5$, then, in consequence of (1) and (2), we may suppose that 1 and two other integers less than or equal to 4 occur in the maximum sequence. Hence the sequence contains either 1, 2, 4 or 1, 3, 4. But it is evident that neither of these sequences leads to $r(8) = 5$. Hence $r(8) \leqslant 4$, and, since 1, 2, 4, 5 is an A sequence, $r(8) = 4$.

Consider now $r(10)$. If $r(10) = 6$, then, in consequence of $r(8) = 4$ and (2), 1, 2, 9, 10 occurs in the sequence. But then 3, 5, 6, and 8 cannot occur. Thus the only possibility is 1, 2, 4, 7, 9, 10; this is impossible because it contains 1, 4, 7. Hence $r(10) \leqslant 5$, and, since 1, 2, 4, 9, 10 is an A sequence, $r(10) = 5$*.

Now we consider $r(12)$. If $r(12) = 7$, by the above argument 1, 2, 11, 12 occurs in our sequence. In consequence of $r(8) = 4$ and (2), 4 and 9 must occur, too. Hence the sequence contains 1, 2, 4, 9, 11, 12; but it cannot contain any other integers. Thus $r(12) = 6$. Since 1, 2, 4, 5, 10, 11, 13, 14 is an A sequence, $r(14) = 8$ and $r(13) = 7$. In consequence of (3), we have $r(16) \leqslant 8$, $r(18) \leqslant 9$, $r(20) \leqslant 10$, $r(22) \leqslant 11$.

From these results we now easily deduce the general theorem.

* $r(9) = 5$ and $r(11) = 6$, since 1, 2, 4, 8, 9 and 1, 2, 4, 8, 9, 11 are A sequences.

Suppose that the theorem holds for $2N-8$. Then, by (3),

$$r(2N) \leqslant r(2N-8)+r(8) < N-3+4 = N+1,$$

i.e. the theorem is proved, for we have established it for the special cases 16, 18, 20, 22.

For sufficiently large N, we have a better estimate by

THEOREM II. *For* $\epsilon > 0$ *and* $N > N_0(\epsilon)$,

$$r(N) < (\tfrac{4}{9}+\epsilon)N.$$

First we prove that $r(17) = 8$. Since $r(14) = 8$, it is evident that $r(17) \geqslant 8$. In the case $r(17) = 9$, the numbers 1 and 17 must occur, since $r(14) = 8$. But then 9 cannot occur, and so, by (2), $r(17) \leqslant r(8)+r(8) = 8$. Thus $r(34) \leqslant 16$. Further, $r(35) \leqslant 16$. For, if $r(35) \geqslant 17$, then, by $r(34) \leqslant 16$, the integers 1 and 35 must occur; but then 18 cannot occur, since the sequence would contain 1, 18, 35. Hence, as previously, $r(35) \leqslant 16$.

Similarly $r(71) \leqslant 32$, ..., $r(2^k+2^{k-3}-1) \leqslant 2^{k-1}$. Hence the result.

By a similar but very much longer argument we find that

$$r(18) = r(19) = r(20) = 8.$$

On the other hand, $r(21) = 9$, since 1, 3, 4, 8, 9, 16, 18, 19, 21 is an A sequence; further,

$$r(22) = r(23) = 9.$$

Hence, as previously, we find that, for sufficiently large $N > N(\epsilon)$,

$$r(N) < (\tfrac{3}{8}+\epsilon)N.$$

At present this is the best result for $r(N)$. It is probable that

$$r(N) = o(N).$$

It may be noted that, from $r(20) = 8$, $r(41) \leqslant 16$. On the other hand, $r(41) = 16$, since 1, 2, 4, 5, 10, 11, 13, 14, 28, 29, 31, 32, 37, 38, 40, 41 is an A sequence. G. Szekeres has conjectured that $r\{\frac{1}{2}(3^k+1)\} = 2^k$. This is proved* for $k = 1, 2, 3, 4$.

More generally, he has conjectured that, if we denote by $r_l(N)$ the maximum number of integers less than or equal to N such that no l of

* It is easily seen that $r\{\frac{1}{2}(3^k+1)\} \geqslant 2^k$; for, if $u \leqslant \frac{1}{2}(3^k-1)$ is any integer not containing the digit 2 in the ternary scale, then the integers $u+1$ form an A sequence.

them form an arithmetic progression, then, for any k, and any prime p,

$$r_p\left(\frac{(p-2)p^k+1}{p-1}\right)=(p-1)^k.$$

An immediate and very interesting consequence of this conjecture would be that for every k there is an infinity of k combinations of primes forming an arithmetic progression.

Another consequence of it would be a new proof of a theorem of van der Waerden which would give much better limits than any of the previous proofs. Namely, it would follow from the conjecture that, if we denote by $N=f(k, l)$ the least integer such that, if we split the integers up to N into l classes, at least one of them contains an arithmetic progression of k terms, then

$$f(k, l) < k^{ck \log l}.$$

The University, Manchester;
Budapest VI, Andrássy út 50.

Printed by C. F. Hodgson & Son, Ltd., Newton St., London, W.C.2.