If $k+1$ is prime then $\Gamma^*(k) = k^2+1$ and if $k+1$ is composite and k can be written in the form $p(p-1)$ for prime p then $\Gamma^*(k) = \frac{1}{2}k(p^2-1)+1$.

Acknowledgements. I would like to thank the Science Research Council for my Maintenance Grant while this research was being done.

References

[1] I. Chowla, *A theorem on the addition of residue classes*, Proc. Indian Acad. Sci. 2 (1935), pp. 242–243.
[2] H. Davenport and D. J. Lewis, *Homogeneous additive equations*, Proc. Roy. Soc. A 274 (1963), pp. 443–460.
[3] M. M. Dodson, *Homogeneous additive congruences*, Philos. Trans. Roy. Soc. London Ser. A 261 (1967), pp. 163–210.
[4] H. Halberstam and K. Roth, *Sequences*, Vol. 1, Oxford 1966.
[5] K. K. Norton, *On homogeneous additive congruences of odd degree*, Technical Report No. 16, Office of Naval Research, Washington, D. C., 1966.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF YORK

Received on 22. 8. 1972 (314)

ACTA ARITHMETICA
XXV (1974)

On difference-sequences

by

I. Z. Ruzsa (Budapest)

1. Introduction. Throughout this paper capitals will be used for denoting (strictly increasing) sequences of non-negative integers; if $A, B, \ldots, A', A^{(1)}, \ldots$ denotes some sequence, then $a_n, b_n, \ldots, a'_n, a_n^{(1)}, \ldots$ is its nth element, and $A(x), B(x), \ldots, A'(x), A^{(1)}(x), \ldots$ the number of its elements less than or equal to x.

Let $A+B$ denote the sequence of all numbers which can be written in the form a_i+b_j. P. Erdös ([1], page 30, problem 15) asked the following: if $A^{(2)} = A+A$, how much denser will $A^{(2)}$ be than A. In general, we can state nothing, as is seen from the example of the multiples of some fixed number d; Erdös conjectured that $A(x) = o(x)$ implies

$$\limsup A^{(2)}(x)/A(x) \geqslant 3.$$

(In all the limits, if not stated differently, we mean $x \to \infty$.) Freyman in his book ([2], page 120) has proven this (in fact, a little more).

Also Erdös asked the analogous question for the sequence of differences (that is, the numbers which can be represented in the form $a_i - a_j$, $i \geqslant j$; this sequence will be called the *difference-sequence of* A). If A, A', $A^{(1)}, \ldots$ is a sequence, its difference-sequence will be denoted by D, D', $D^{(1)}, \ldots$ Erdös conjectured that if $A(x) = o(x)$, then $D(x)/A(x) \to \infty$ (unlike the case of sums, where $\limsup A^{(2)}(x)/A(x) \geqslant 3$ is best possible), and with Sárközy he proved that if the upper density of A is positive, then $d_{n+1} - d_n = O(1)$. In Section 2 we prove Erdös's conjecture, and in Section 3 we determine exactly what can be asserted about the lower density of D, knowing the upper density of A. Finally, in Section 4 we shall mention some more problems and some results, only outlining the proofs.

2. The case $A(x) = o(x)$.

Theorem I. *If A is an infinite sequence for which $A(x) = o(x)$, then*

$$\lim D(x)/A(x) = \infty.$$

In fact we shall prove the following stronger

THEOREM I*. *If A is an arbitrary infinite sequence, then either*

$$\lim D(x)/A(x) = \infty, \quad \text{or} \quad d_{n+1} - d_n = O(1).$$

In case A is of density 0, Theorem I* yields Theorem I. In case A is of positive upper density Theorem I* yields the Erdös–Sárközy result mentioned above. Later on for the second we give an improved lower estimation for the lower density of D.

Proof. Let us suppose that $\liminf D(x)/A(x) = c < \infty$; then we can choose a sequence X for which

$$\lim_{k\to\infty} D(x_k)/A(x_k) = c. \tag{1}$$

Now consider the following sequences:

$$A^{(n)} = (A - a_n) \setminus \bigcup_{i=1}^{n-1} (A - a_i) \qquad (A^{(1)} = A - a_1).$$

($A - a$ means the sequence $\{a_i - a\}$, beginning from the suffix i_0 satisfying $a_{i_0} \geqslant a > a_{i_0} - 1$.)

Certainly $D = \bigcup_{i=1}^{\infty} A^{(i)}$, and the $A^{(i)}$'s are disjoint, thus

$$\sum_{i=1}^{\infty} A^{(i)}(x) = D(x).$$

It follows from (1) that

$$\limsup_{k\to\infty} \frac{\sum_{i=1}^{n} A^{(i)}(x_k)}{A(x_k)} \leqslant c$$

holds for every n. Therefore

$$\lim_{n\to\infty} \limsup_{k\to\infty} \frac{\sum_{i=1}^{n} A^{(i)}(x_k)}{A(x_k)} = c_1 \leqslant c.$$

Let us choose the suffix n_0 so that

$$\limsup_{k\to\infty} \frac{\sum_{i=1}^{n_0} A^{(i)}(x_k)}{A(x_k)} = c_2 > c_1 - \frac{1}{3}. \tag{2}$$

We now choose a sequence $Y \subseteq X$ such that

$$\lim_{k\to\infty} \frac{\sum_{i=1}^{n_0} A^{(i)}(y_k)}{A(y_k)} = c_2 \tag{3}$$

hold.

Let $G = \{a_1, \ldots, a_{n_0}\}$, and repeat the above series of operations for the sequence $A' = A \setminus G$ (it is possible, because $\frac{D'(y_k)}{A'(y_k)} \leqslant \frac{D(y_k)}{A'(y_k)} = \frac{D(y_k)}{A(y_k) + o(1)} \to c$). In this way we get a subsequence Z of Y and a suffix n_1 such that

$$\lim_{k\to\infty} \frac{\sum_{i=1}^{n_1} A'^{(i)}(z_k)}{A(z_k)} = c_4 > c_3 - \frac{1}{3}, \tag{4}$$

where

$$c_3 = \lim_{n\to\infty} \limsup_{k\to\infty} \frac{\sum_{i=1}^{n} A'^{(i)}(y_k)}{A(y_k)}. \tag{5}$$

(Obviously we can replace $A'(x)$ by $A(x)$ in the denominators.) Similarly let $G' = \{a'_1, \ldots, a'_{n_1}\} = \{a_{n_0+1}, \ldots, a_{n_0+n_1}\}$, $A'' = A' \setminus G' = \{a_{n_0+n_1+1}, \ldots\}$. Now we shall prove $d''_{n+1} - d''_n = O(1)$, from which the theorem will evidently follow because $D'' \subseteq D$.

Let $d''_n \in D''$ be arbitrary. Then

$$d''_n = a_p - a_q, \qquad p > q > n_0 + n_1.$$

For a k large enough because of (2), (3), (4) and (5) we have

$$\sum_{i=n_0+1}^{p} A^{(i)}(z_k) < \frac{1}{3} A(z_k) \tag{6}$$

and

$$\sum_{j=n_1+1}^{q-n_0} A'^{(j)}(z_k) < \frac{1}{3} A(z_k). \tag{7}$$

According to (6) the number of integers a_j satisfying $a_p + g'_{n_1} < a_j < z_k$ and $a_j - a_p \notin A - g_s$ for all s is less than $\frac{1}{3} A(z_k)$ since the condition $a_j - a_p \notin A - g_s$ for all s is equivalent to $a_j - a_p \in \bigcup_{i=n_0+1}^{p} A^{(i)}$. The same estimation is valid for the number of those a_j for which $a_j - a_q \notin A - g'_t$ for all t, since $a_q = a'_{q-n_0}$ and so the condition $a_j - a_q \notin A - g'_t$ for all t is equivalent to $a_j - a_q \in \bigcup_{i=n_1+1}^{q-n_0} A'^{(i)}$.

Hence it follows that there exists a j such that

$$a_j - a_p = a_u - g_s,$$
$$a_j - a_q = a_v - g'_t$$

from which it results that

$$a_v - a_u = a_p - a_q + g'_t - g_s = b''_n + (g'_t - g_s).$$

On the other hand, $a_v > a_u > g'_{n_1}$ for $a_j > a_p + g'_{n_1}$, hence $a_v, a_u \in A''$, $a_v - a_u \in D''$, that is, since $0 < g'_t - g_s < g'_{n_1}$, $d''_{n+1} - d''_n < g'_{n_1}$, q.e.d.

3. The case $A(x) \neq o(x)$. Theorem I* implies that under this condition $\liminf D(x)/x > 0$. First we show a simple improvement of this.

THEOREM II. $\inf \frac{D(x)}{x} \geqslant \limsup \frac{A(x)}{x}$ *for an arbitrary infinite sequence* A.

This will be proven in the following, slightly stronger form:

THEOREM II*. *For arbitrary natural numbers* x, y, x_0

$$A(x+y) - A(y) \leqslant \left[\frac{x-1}{x_0} + 1\right] D(x_0).$$

Making $y = 0$ we get

$$\frac{A(x)}{x} \leqslant \left(1 + \frac{x_0}{x}\right) \frac{D(x_0)}{x_0};$$

by $x \to \infty$ for all x_0

$$\limsup \frac{A(x)}{x} \leqslant \frac{D(x_0)}{x_0},$$

which is Theorem II.

Considering that the interval $(y, x+y]$ can be covered by $\left[\frac{x-1}{x_0} + 1\right]$ ones of length x_0, it is sufficient to see that $A(x+x_0) - A(x) \leqslant D(x_0)$. But this is almost trivial; if $(x, x+x_0]$ contains no a_i at all, we are ready. If it does, let a_j be the least of them; then all they are elements of $D + a_j$ ($a_i - a_j \in D$ because of the definition of D), thus indeed their number is less than or equal to $D(x_0)$.

Still weakening the statement of Theorem II we get $\liminf D(x)/x \geqslant \limsup A(x)/x$. The next theorem is a best possible improvement of this inequality.

THEOREM III. *If* $\limsup A(x)/x = c$, *then*

$$\liminf \frac{D(x)}{x} \geqslant \frac{1}{[1/c]}.$$

A suitable subsequence of the multiples of $[1/c]$ shows that there is no more connection between $\limsup A(x)/x$ and $\liminf D(x)/x$.

Theorem III will be shown in the following form:

THEOREM III*. *Let* d *be any natural number. If*

$$\liminf \left(D(x) - \frac{x}{d}\right) = -\infty,$$

then

$$A(x) < \frac{x}{d+1} + O(1).$$

Theorem III follows with $d = [1/c]$.

Proof. One can select $d+1$ natural numbers $c_1 < c_2 < \ldots < c_{d+1}$ such that $c_i - c_j \notin D$. Namely if we have k ($\leqslant d$) of them, then we can choose one more. We have to satisfy $c_{k+1} - c_1 \notin D, \ldots, c_{k+1} - c_k \notin D$, that is $c_{k+1} \notin E = \bigcup_{i=1}^{k} (D + c_i)$. But according to our condition for suitable x

$$E(x) \leqslant kD(x) \leqslant dD(x) \leqslant x - c_k - 1,$$

therefore there must be a number beyond c_k not contained in it.

But if we have chosen such $c_1, \ldots, c_{d+1}$, then the sequences $A + c_i$ are disjoint ($a_u + c_i = a_v + c_j$ would just imply the contradiction $c_i - c_j = a_v - a_u \notin D$), from which regarding the a_i's non exceeding x

$$(d+1) A(x) \leqslant x + c_{d+1},$$

q.e.d.

4. Problems and results.

1. The question arises whether Theorems I* and II have a common generalization, that is whether one can estimate $d_{n+1} - d_n$ from above knowing $\limsup A(x)/x$. The answer is no, even in the simplest cases, as shown by the following example of Szemerédi: let $a \in A$ if $a \equiv i \pmod d$, $0 \leqslant i < d(\frac{1}{2} - \varepsilon)$ with some large fixed d. Hence for arbitrary large c and small $\varepsilon > 0$ A can have a density larger than $\frac{1}{2} - \varepsilon$, while $d_{n+1} - d_n > c$ infinitely many times.

2. Does D have a density if A has a positive one? Not necessarily, as can be seen from the following example of Sárközy: A consist of 0, and of

$$4i+1 \quad \text{if} \quad (2n-1)! \leqslant i < (2n)!,$$
$$4i+2 \quad \text{if} \quad (2n)! \leqslant i < (2n+1)!.$$

3. Under what conditions will a sequence be the difference-sequence of another? A simple sufficient condition given by Sárközy: if it contains intervals of arbitrary length (as an interesting special case, all the sequences having upper density 1). To see this let the sequence be B. We select a sequence C such that all the differences of $A = \{c_n, c_n+b_n\}$ fall into B; certainly its difference-sequence will be B. But it is easily satisfied; having chosen $c_1, \ldots, c_n$ we have only to choose c_{n+1} so that

$$[c_{n+1}-c_n, c_{n+1}+b_{n+1}] \subseteq B,$$

which is solvable by our conditions.

An interesting consequence of this condition is that being given a sequence B and an arbitrary arithmetical function $\omega(n) \geqslant 0$, $\omega(n) \to \infty$ one can find an A such that $D \supseteq B$, $D(x)-B(x) \leqslant \omega(x)$.

Clear necessary conditions are that every d_n be representable in the form $d_i - d_j$ infinitely many ways, and that for all m, if B denotes the sequence of the differences divisible by m, then $b_{n+1}-b_n$ do not tend to the infinity. But these conditions are by no means sufficient. These conditions if fulfilled for two sequences, hold for their union as well, while the union of two difference-sequences does not have to be a difference-sequence. If $a'_n = 2^{n-1}$, $a''_n = 2n$, then $D' \cup D''$ is not a difference-sequence. If for some A D were equal to $D' \cup D''$, then A would not consists of elements of the same parity; let a_i be even and a_j odd. Since all the odd elements of D are of the form 2^k-1, all the odd elements of A greater than a_i should have form a_i+2^k-1 and the even ones $\geqslant a_j$ form a_j+2^k-1, hence all even numbers in D, that is all even numbers would be of the form 2^k-2^n, $2^k-1+a_i-a_s$, $2^k-1+a_j-a_t$, or a_s-a_t where $s, t \leqslant \max(i, j)$, which is impossible.

One would expect that the problem can be "finitarized" in the following way: if for all n exist such numbers $c_1^{(n)}, \ldots, c_n^{(n)}$ that for $A^{(n)} = \{c_i^{(n)}, c_i^{(n)}+b_i\}$ $D^{(n)} \subseteq B$, then B is a difference-sequence. (In the definition of $A^{(n)}$ we meant those numbers ordered increasingly.) Unfortunately this is not true, as is shown by the sequence containing all the non-negative integers of the form 2^i-2^j and 2^i+2^j. Here suits $c_k^{(n)} = 2^n - 2^i$ if $b_k = 2^i \pm 2^j$. That B is not a difference-sequence, can be seen similarly to the previous.

4. Given D, the above sufficient condition (in the beginning of 3) gives infinitely many constructions of the sequence A. The question arises whether this is inevitable. The answer is no, since it is easy to see that if $a_{n+1} \geqslant 2a_n$ for all n, then $A' \neq A+c$ implies $D' \neq D$. What other conditions imply the uniqueness of A?

5. Can Theorem I* be extended to the difference of two distinct sequences? If A and B are two sequences, then by $A-B$ we mean the sequence of all $a_i - b_j \geqslant 0$. I have found examples for the following two statements, showing that both parts of Theorem I* fail to be valid in this case.

a. There exist such sequences A and B of positive density, that for $F = (A-B) \cup (B-A)$

$$f_{n+1}-f_n \neq O(1).$$

b. There exist such sequences A and B of density 0, that for $F = (A-B) \cup (B-A)$

$$F(x) = O\big(\min(A(x), B(x))\big).$$

The constructions are complicated.

References

[1] P. Erdös, *Some unsolved problems*, Publ. Math. Inst. Hung. Acad. Sci. 6 (1961), pp. 221–254.

[2] G. A. Freyman, *The foundations of the structural theory of addition of sets* (in Russian), Kazan 1966.

Received on 1. 10. 1972 (331)