ON DIFFERENCE SETS OF SETS OF INTEGERS

by Cam L. STEWART

1. Introduction.

Let N_0 denote the set of non-negative integers. Let A be a subset of N_0 , and let d be any integer. Put $A-d=\{a-d\;;\;a\in A\}$ and, for convenience, denote $A\cap A-d$ by A[d]. We define the ordinary-difference set O(A) of A, by

$$Q(A) = \{d \in N_O ; A[d] \neq \emptyset\}$$
.

Thus, the ordinary-difference set of A is the set of all non-negative integers which can be written as the difference of two elements of A . Until recently very little was known about this set, and this contrasted with the situation, see [3] or [5], for the sum set S(A) of A . S(A) is defined as the set of all non-negative integers which can be written as the sum of two elements of A . In the last few years, however, several papers have appeared on the subject of ordinary-difference sets. In this survey, we shall review some of this work, and we shall also discuss some results which have been obtained about the related infinite-difference and density-difference sets.

Let |A| denote the number of elements of A. We define the infinite-difference set $O_{\infty}(A)$ of A by

$$O_{\infty}(A) = \{d \in N_{O} ; |A[d]| = \infty\}.$$

Next let $|A|_{X}$ be the number of elements of A which are less than x. The upper density of A is given by $d^-(A) = \lim\sup_{X \to \infty} (|A|_{X}/x)$, and the lower density of A by $d_-(A) = \lim\inf_{X \to \infty} (|A|_{X}/x)$. If $d^-(A) = d_-(A)$, then this limit value is the density d(A) of A. We define the density-difference set $\theta_0(A)$ of A by

$$Q_{O}(A) = \{d \in N_{O} ; d^{-}(A[d]) > 0\}$$

In this article, we shall restrict our attention to sets A of positive upper density. One reason for doing this is that given any subset K of N_0 containing O, it is a straightforward task (see theorem 3 of [14]) to construct a set A with $d^-(A) = 0$ for which $O_\infty(A) = K$. Thus we can't say anything non-trivial about $O_\infty(A)$ knowing only that $d^-(A) = 0$. Further, if $d^-(A) = 0$, then $O_0(A)$ is easily seen to be empty. When $d^-(A)$ is positive, however, structure is imposed on O(A), $O_\infty(A)$ and $O_0(A)$, and the three types of difference set then have several common properties.

We remark that interesting problems do arise concerning O(A) when d(A) = 0. RUZSA [9], for example, has shown that, if A is an infinite set with d(A) = 0, then $\lim_{x\to\infty} (|O(A)|_X/|A|_X) = \infty$. Suprisingly, this is in contrast to the situation

for the sum set S(A). FREIMAN [3] resolved a conjecture of Erdős by proving that, if A is an infinite set with $d^{-}(A)=0$, then

(1)
$$\lim \sup_{X\to\infty} \frac{|S(A)|_X}{|A|_X} \geqslant 3,$$

and further that there exists sots A as above for which the inequality in (1) is actually an equality.

2. The structure of difference sets.

It would be desirable to characterise those sets which are difference sets of sets of integers. No simple characterisation is known for any of the three types of difference set. It is known, however, that, if we iterate the operation of forming the ordinary-difference set of a set of positive upper density that we eventually obtain the set of all multiples of a fixed number. Put $0^1(A) = 0(A)$, and $0^k(A) = 0(0^{k-1}(A))$, for k = 2, 3, ... The following result is due to STEWART and TIJDEMAN [15].

THEOREM 1. - Let A have positive upper density ε . Then, there exists an integer k, with $1 < k < \varepsilon^{-1}$, such that $D^{\mathbf{r}}(A) = \{jk\}_{j=0}^{\infty}$, for all integers r, with $r > 2[(\log \varepsilon^{-1})/\log 2]$.

We would conjecture that theorem 1 applies with the lower bound for r sharpened to $r > [(\log \varepsilon^{-1})/\log 2] + 1$ and with the operation of forming the ordinary—difference set replaced by any one of the three operations of forming a difference set. The example of the set of integers

$$A_h = \{a; a > 0 \text{ and } a \equiv 0 \text{ or } 1 \text{ (mod h)}\},$$

with h=6, say, shows that the lower bound for r cannot be replaced by $[(\log \epsilon^{-1})/\log 2]$ for any of the three types of difference set.

According to RUZSA [9], ERDOS and SARKÖZY proved that, if d(A) is positive, then O(A) does not have arbitrarily large gaps. In other words, if the elements of O(A) are ordered according to size, then the difference between consecutive terms is bounded. RUZSA [10], by refining work of STEWART and TIJDEMAN [14], obtained the following improvement of this result.

THEOREM 2. - Let A have positive upper density ϵ . Then, there exist r integers k_1 , ..., k_r such that

$$\bigcup_{j=1}^{r} (O_{O}(A) + k_{j}) \supseteq N_{O}, \underline{\text{with}} \quad r \leqslant \varepsilon^{-1}.$$

This result is best possible as the example $A_{\ell} = \{a \; ; \; a > 0 \; \text{ and } \; a \equiv 0 \pmod{\ell}\}$ shows, since in this case $\mathcal{O}_0(A_{\ell}) = A_{\ell}$, and plainly ℓ shifts of $\mathcal{O}_0(A_{\ell})$ are necessary to cover all of N_0 . While the number of shifts of $\mathcal{O}_0(A)$ required to cover N_0 is bounded in terms of ϵ , it is not the case that the $\max_j |k_j|$ is necessarily bounded in terms of ϵ . For let A_t consist of the integers of the form

3nt + i , for i = 1 , ... , t and n = 0 , 1 , 2 , ... In this case, $\Omega_0(A)$ is the set of non-negative integers of the form 3nt \pm i , for i = 0 , ... , t and n = 0 , 1 , 2 , ... , and so contains infinitely many gaps of length t . Thus $\max_i |k_i| > [t/2]$. On the other hand, $d(A_t) = \varepsilon = 1/3$.

Plainly, we have $\mathbb{Q}(A) \supset \mathbb{Q}_{\infty}(A) \supset \mathbb{Q}_{0}(A)$, and it is an immediate consequence of theorem 2 that, if $d^{-}(A) = \varepsilon$, then

$$d_{-}(\mathcal{D}_{O}(A)) \geq \left[e^{-1}\right]^{-1}.$$

Thus, all three difference sets of A have lower density at least the upper density of A. The following theorem, see [15], is often useful for translating results about one type of difference set of that of another.

THEOREM 3. - Given a set $A \subseteq \mathbb{N}_0$, there exists a set $B \subseteq \mathbb{N}_0$, with $d_{-}(B) \geqslant d^{-}(A)$ such that $\mathbb{O}(B) \subseteq \mathbb{O}_0(A)$.

The above results suggest that difference sets possess a great deal of regularity. Theorems 1 and 2 might lead one to suppose that, if d (A) is positive, O(A) contains an infinite arithmetical progression. The next theorem, which is a consequence of theorem 6 of [14], shows that this is not the case.

THEOREM 4. - Let ε be any countable set of infinite sets of positive integers, and let α be any number between 0 and 1. There exists a set Λ , with density α , for which

$$\lim\sup_{\mathbf{x}\to\infty}\frac{\left|\mathbb{D}\left(\mathbf{A}\right)\,\cap\,\mathbf{E}\right|_{\mathbf{x}}}{\left|\mathbf{E}\right|_{\mathbf{x}}}\leq2\alpha\;,\;\underbrace{\text{for every}}_{\mathbf{E}}\;\in\;\boldsymbol{\epsilon}\;.$$

On taking & to be the set of all infinite arithmetical progressions and α to be any number between 0 and 1/2, we find that there exists a set Λ of density α whose difference set contains no infinite arithmetical progression.

3. The union and intersection of difference sets.

Neither the union nor the intersection of two ordinary-difference sets need be an ordinary-difference set. For example, on putting

$$A = \{a; a \ge 0 \text{ and } a \equiv 0 \pmod{10}\} \cup \{7\}$$

and

$$B = \{b : b > 0 \text{ and } b \equiv 7 \pmod{10}\} \cup \{0\}$$
,

we readily check that $O(\Lambda) \cap O(B) = \Lambda$, and that there is no set C, with $O(C) = \Lambda$. Similarly, on putting

$$\Lambda = \{a : a > 0 \text{ and } a \equiv 0 \pmod{10}\} \cup \{2\}$$

and

$$B = \{b : b > 0 \text{ and } b \equiv 0 \pmod{10}\} \cup \{9\}$$

we see that there is no set C, with $O(C) = O(A) \cup O(B)$. This is not the case with infinite-difference sets as the next theorem (see [14]) shows. Let D denote the collection of all infinite-difference sets associated with sets of positive upper density.

THEOREM 5. - D is a filter of the set of all subsets of $N_{\rm O}$.

D is not an ultrafilter since there exist disjoint sets possessing arbitrarily large gaps whose union is $N_{\rm O}$; and by theorem 2, the infinite-difference set of positive upper density has only bounded gaps. Since D is a filter both the union and the intersection of two infinite-difference sets associated with sets of positive upper density is again an infinite-difference set associated with a set of positive upper density. Further, if A has positive upper density, and if $E \subseteq N_{\rm O}$ and $O_{\rm o}(A) \subseteq B$, then there exists a set C of positive upper density, with $O_{\rm o}(C) = B$. Neither the collection of ordinary-difference sets nor that of density-difference sets has the above superset property. Let E denote the non-negative even integers. We have $O(E) = O_{\rm o}(E) = E$. Plainly $E \cup \{1\}$ is not the ordinary-difference set of any set. Similarly, see [15], there is no set A such that $O_{\rm o}(A) = E \cup \{1\}$.

Both the union and the intersection, see [15], of two density-difference sets is again a density-difference set. It is a consequence of the next theorem, see [15], that the intersection of any two difference sets associated with sets of positive upper density must itself have a large positive upper density.

THEOREM 6. - If A and B are subsets of N₀, then there exists a set $C \subseteq N_0$ such that $O_0(C) = O_0(A) \cap O_0(B)$ and such that $d_0(C[d]) \ge d_0(A[d]) \cdot d_0(B[d])$, for every $d \in N_0$.

On taking d=0 in theorem 6 and recalling (2), we see that $d_{-}(\mathbb{Q}_{\cap}(A)\cap\mathbb{Q}_{\cap}(B))>[(d^{-}(A)\ d^{-}(B))^{-1}]^{-1}.$

This inequality, which is best possible, see [14], has also been proved by RUZSA [10].

4. Lacunary sequences.

In the next two sections, we shall discuss the following problem. For what sequences of positive integers K does there exist a set of positive upper density having no terms of K in its ordinary-difference set? In this section, we shall discuss a condition on the rate of growth of K which ensures the existence of a suitable set A . If $K = \{k_j\}_{j=1}^{\infty}$ is lawurary, then A exists. More precisely, if, for some positive integer h , we have $\lim\inf_{j\to\infty} (k_{j+h}/k_j) > 1$, then there exists a set A , with positive upper density for which $k_j \neq 0$ (A) , for j=1, 2, ... This condition is critical, see theorem 8 of [14], since, if K is an increasing sequence of positive integers satisfying $\lim\inf_{j\to\infty} ((k_{j+h}/k_j) = 1$, for every positive integer h , then there exists an increasing sequence of positive integers

 $E = \{e_j\}_{j=1}^{\infty}$, with $(e_{j+1}/e_j) \ge (k_{j+1}/k_j)$, for j=1, 2, ... such that, for every set A of positive upper density, the ordinary-difference set of A contains terms of E. The following theorem, see [14], puts the above condition in a more quantitative form.

for i=1, ..., h and j=0, 1, 2, ..., then there exists a set A, having a density, with

$$d(A) \ge \prod_{i=1}^{h} \left(\frac{c_i - 2}{2(c_i - 1)}\right),$$

for which $k_j \not\in O(A)$, for $j = 1, 2, \cdots$

Observe that if $(k_{j+\ell}/k_j) \ge \alpha > 1$, for j=1, 2, ..., and g is an integer with $g \ge (\log 3)/\log \alpha$, then $(k_{j+g\ell}/k_j) \ge 3$, for j=1, 2, ..., and we may apply theorem 7, with $h=g\ell$, and $c_1=c_2=\cdots=c_h=3$ to obtain the condition mentioned previously.

Let, for example, k_1 , k_2 , ..., be the sequence of factorials 1!, 2!, ... It follows from theorem 7, on choosing h=2, $c_1=6$ and $c_2=12$, that there exists a set A with density at least 2/11 which does not have a factorial as the difference of two terms.

The proof of theorem 7 proceeds in two stages. First, by means of a construction of nested intervals, a real number θ_i is found, for i=1, ..., h, satisfying

(3)
$$\|\mathbf{k}_{jh+i} \ \theta_i\| \ge \frac{c_i - 2}{2(c_i - 1)}$$
,

for j=0, 1, 2, ...; Here ||x|| denotes the distance from x to the nearest integer. Secondly, by means of an averaging argument and Weyl's criterion for uniform distribution (see [14]) an appropriate set A is shown to exist. A has the form

{n; $\lambda_{i} \leq \{n\theta_{i}\} < \lambda_{i} + g_{i} \pmod{1}$, for i = 1, ..., h},

where λ_1 , ..., λ_h are real numbers, $g_i=(c_i-2)/2(c_i-1)$, for $i=1,\ldots,h$, and $\{x\}$ denotes the fractional part of x. We then have

$$Q(A) \subseteq \{n ; ||n\theta_i|| < g_i, \text{ for } i = 1, \dots, h\},$$

and, by our choice of $\theta_{\bf j}$, recall (3), $k_{\bf j} \notin \Omega(A)$, for ${\bf j}=1$, 2, ..., as required. As an alternative second stage, we can define $A_{\bf i}=\{n\;;\;0\in\{n\theta_{\bf i}\}<\theta_{\bf i}\}$, for ${\bf i}=1$, ..., ${\bf h}$. In this case, $d(A_{\bf i})>g_{\bf i}$, and $\Omega(A_{\bf i})\subseteq\{n\;;\;\|n\theta_{\bf i}\|<\theta_{\bf i}\}$, and we may deduce from theorems 3 and 6 that there exists a set B with $d_{\bf i}(B)>g_{\bf i}$, with $\Omega(B)\subseteq\bigcap_{{\bf i}=1}^h\Omega(A_{\bf i})$. Apart from the fact that B might not have a density, this gives another proof of theorem 7.

ERDŐS and SARKÖZY [2] have proved a similar result to theorem 7. They showed that, if $(k_{j+1}/k_j) \geqslant \Delta > 1$, for j=1, 2, ..., then there exists a set Δ with

(4)
$$d_{\Delta}(A) > 24^{-((\log 3/\log \Delta)+1)},$$

for which $k_j \notin O(A)$, and $k_j \notin S(A)$, for j=1, 2, ... We remark that a slight modification of the proof of theorem 7 allows one to conclude, with the same assumptions on K as in the statement of theorem 7, that there exists a set B, with

$$d_{-}(B) > \prod_{i=1}^{h} \left(\frac{c_{i}-2}{4(c_{i}-1)}\right),$$

for which $k_j \notin \mathcal{Q}(B)$, and $k_j \notin S(B)$, for j=1, 2, ... This improves the lower bound given by (4). We construct a set A as before by means of Weyl's criterion and an averaging argument, but with g_i replaced by $g_i/2$. We then have $d(A) > g_1 \cdots g_h 2^{-h}$, and $\mathcal{Q}(A) \subseteq \{n: \|n\theta_i\| < g_i/2$, for i=1, ..., h}. On putting $B = \mathcal{Q}(A)$, we see that both $\mathcal{Q}(B)$ and S(B) are contained in $\{n: \|n\theta_i\| < g_i$, for i=1, ..., h}, and, by (2), $d_i(B) > \prod_{i=1}^h (g_i/2)$ as required.

5. Sets which intersect difference sets.

Theorem 7 gives a sufficient condition on K for the existence of a set A of positive upper density having no terms of K in its ordinary-difference set. Clearly, this condition is not a necessary one since we may take K to be the odd integers and A to be the even integers. Indeed more generally, if K is any set with no terms divisible by some fixed integer q, then there exists a set \mathbf{A}_q of positive upper density with $\mathcal{O}(\mathbf{A}_q) \cap \mathbf{K} = \emptyset$; we just take

$$A_q = \{n ; n > 0 \text{ and } n \equiv 0 \pmod{q}\}$$
.

There are several interesting sets K which are not dealt with by theorem 7 or the above congruence condition. For instance, we can put K = P + 1, P - 1 or the set of squares; here P denotes the set of prime numbers. ŚARKÖZY [11], [12], [13] has shown, by means of the Hardy-Littlewood circle method, that, for all three of the above sets, $c(A) \cap K \neq \emptyset$ whenever A has positive upper density. In fact, he has proved that there exist positive constants c_1 and c_2 such that if

$$|A|_{x} > c_1 \frac{x(\log \log \log x)^3 \log \log \log \log x}{(\log \log x)^2}$$

then there are two elements of A less than x whose difference is a prime minus one, and, if

$$|A|_{x} > c_{2} \frac{x(\log \log x)^{2/3}}{(\log x)^{1/3}}$$

then there are two elements of A less than x whose difference is the square of a positive integer. FURSTENBERG [4], using ergodic theory, has also shown that, if K is the set of positive squares, then $\mathcal{Q}(A) \cap K \neq \emptyset$ whenever $d^{\bullet}(A)$ is positive.

Let K be a set of positive integers, and define $K_{\mathbf{q}}$, for every positive integer q , by

$$K_{q} = \{k \in K ; k \equiv 0 \pmod{q}\},\$$

and let $\{k_{1,q}\}_{i=1}^{\infty}$ be the sequence formed by ordering the elements of K_q according to size. KAMAE and MENDES FRANCE, see example 3 and theorem 2 of [7], using techniques from Fourier analysis, have proved that, if the sequence $\{k_{1,q},q\}_{i=1}^{\infty}$ is uniformly distributed modulo 1, for every positive integer q and every irrational number θ , then $\theta(A) \cap K \neq \emptyset$, for every set A of positive upper density. The above criterion is very useful. For let P(x) be a palynomial of degree at least 2 with integer coefficients and leading coefficient positive, and let θ be a positive irrational number. The sequence $\{P(n),\theta\}_{n=1}^{\infty}$ and, since P(x) has degree at least 2, the sequences $\{P(n+h)\theta-P(n)\theta\}_{n=1}^{\infty}$, h=1, 2, ..., are u. d. mod 1 (see theorem 3.2, p. 27 of [8]). Therefore the sequence $\{P(qn+r)\theta\}_{n=1}^{\infty}$ is u. d. mod 1, for every positive integer q and non-negative integer r (see theorem 2.1, p. 238 of [8]). Since the leading coefficient of P(x) is positive the sequence formed by ordering the elements of the set $\{P(qn+r)\theta\}_{n=1}^{\infty}$ is also u. d. mod 1. Furthermore, since the set

$$K_q = \{P(n); n > 0, P(n) > 0 \text{ and } P(n) \equiv 0 \pmod{q}\}$$

is the union of finitely many sets of the form $\{P(qn+r):n>0$ and $P(qn+r)>0\}$ the sequence $\{k_{i,q} \ \theta\}_{i=1}^{\infty}$ formed from K_q is u. No. mod 1, for every irrational θ whenever K_q is non-empty. Note that since the leading coefficient of P(x) is positive if there exists an integer m with $q \mid P(m)$, then there exist infinitely many such integers and thus K_q is non-empty. We may now apply the congruence condition mentioned at the start of this section in conjunction with the criterion of KAMAE and MENDES FRANCE to deduce the following theorem. We remark that the result is obvious for polynomials of degree 1.

THEOREM 8. - Let $P(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$ be a non-constant polynomial with integer coefficients and with a_m positive. Put

 $K = \{P(n) ; n \text{ and } P(n) \text{ positive integers} \}$

Then $K \cap O(A) \neq \emptyset$, for every set A of positive upper density if, and only if, for every integer q, there exists an integer m such that $q \mid P(m)$.

Thus, if P(x) is a monic polynomial with an integer root, then the set $K = \{P(n); n > 0 \text{ and } P(n) > 0\}$ has the property that $K \cap O(A) \neq \emptyset$, for every set A of positive upper density whence, in particular, the set K of k-th powers has the intersection property whenever k is a positive integer. There also exist reductible polynomials which do not have an integer root and yet which have the above intersection property. For instance, the polynomial $(x^2 - a)(x^2 - b)(x^2 - ab)$, where a, b and ab are integers which are not squares, has a linear factor modulo q, for every positive integer q, since a, b and ab cannot all be quasitative.

dratic non-residues modulo $\,q\,$. On the other hand, if $\,P(x)\,$ is an irreducible polynomial of degree at least 2, then there are sets A of positive upper density whose intersection with the associated set K is empty since in this case, by a theorem of Frobenius, see theorem 9 of [6], there is a prime number $\,q\,$ for which $\,P(x)\,$ does not have a linear factor modulo $\,q\,$.

Lastly, we remark that theorem 3 shows that O(A) may be replaced by $O_O(A)$ in the statement of theorem 8.

REFERENCES

- [1] ERDÖS (P.) and SARKÖZY (A.). On differences and sums of integers, I, J. Number Theory, to appear.
- [2] ERDÖS (P.) and SARKÖZY (A.). On differences and sums of integers, II, to appear.
- [3] FREIMAN (G. A.). Foundations of a structural theory of set addition. Providence, American mathematical Society, 1973 (Translations of mathematical Monographs, 37).
- [4] FURSTENBERG (H.). Ergodic behavior of diagonal measures and a theorem of Szeméredi on arithmetic progressions, J. Analyse math., Jérusalem, t. 31, 1977, p. 204-256.
- [5] HALBERSTAM (H.) and ROTH (K.). Sequences. Oxford, at the Clarendon Press, 1966.
- [6] HEILBRONN (H.). Zeta-functions and L-functions, "Algebraic number theory [1965. Brighton]", p. 204-230. London, Academic Press, 1967.
- [7] KAMME (T.) and MENDES FRANCE (M.). Van der Corput's difference theorem, Israel J. Math., to appear.
- [8] KUIPERS (L.) and NIEDERREITER (H.). Uniform distribution of sequences. New York, J. Wiley and Sons, 1974 (Pure and applied Mathematics, Wiley-Interscience).
- [9] RUZSA (I. Z.). On difference sequences, Acta Arithm., Warszawa, t. 25, 1974, p. 151-157.
- [10] RUZSA (I. Z.). On difference sets, to appear.
- [11] SARKÖZY (A.). On difference sets of sequence of integers, I, Acta math. Acad. Sc. Hung., to appear.
- [12] SARKÖZY (A.). On difference sets of sequences of integers, II, Annales Univ. Sc. Budapest, Eötvös, to appear.
- [13] SARKÖZY (A.). On difference sets of sequences of integers, III, Acta math. Acad. Sc. Hung., to appear.
- [14] STEWART (C. L.) and TIJDEMAN (R.). On infinite-difference sets of sequences of positive integers, Canadian J. Math., to appear.
- [15] STEWART (C. L.) and TIJDEMAN (R.). On density-difference sets of sequences of integers, to appear.
- [16] TIJDEMAN (R.). Distance sets of sequences of integers, Proceedings of a bicentennial conference Wiskundig Genootschap, Math. Centre, Amsterdam, to appear.

(Texte reçu le 8 septembre 1978)

Cam L. STEWART, Dept of pure Math., University, WATERLOO, Ont. NZL 3GL (Canada).