Can. J. Math., Vol. XXXI, No. 5, 1979, pp. 897–910

ON INFINITE-DIFFERENCE SETS

C. L. STEWART AND R. TIJDEMAN

1. Introduction. Let A be a sequence; throughout this paper sequences are understood to be infinite, strictly increasing and composed of non-negative integers. We define D, the infinite-difference set of A, to be the set of those non-negative integers which occur infinitely often as the difference of two terms of A. Plainly D has no positive terms if and only if $a_{i+1} - a_i \to \infty$ as $i \to \infty$. Note that D contains zero. We shall be interested in the case when $\bar{d}(A) > 0$. Then D certainly contains more than one term. In fact, see Corollary 1, §2, $\underline{d}(D) \geqq \bar{d}(A)$ in this case. Here $\bar{d}$ and $\underline{d}$ denote the (natural asymptotic) upper and lower density respectively.

Let h be a positive integer and let $A_1, \ldots, A_h$ be sequences with positive upper densities $\epsilon_1, \ldots, \epsilon_h$ respectively. Erdös asked whether $D_1 \cap \ldots \cap D_h$, the intersection of the associated infinite-difference sets, necessarily contains positive terms. We shall show that in fact the intersection has positive lower density. We put

$$(1) \quad C_1 = \epsilon_1 \quad \text{and} \quad C_h = \prod_{i=1}^{h} (\epsilon_i/5 \log (h+1)) \quad \text{for } h \geqq 2,$$

and we prove

THEOREM 1. *If $\bar{d}(A_i) \geqq \epsilon_i$ for $i = 1, \ldots, h$ then there exists a sequence A with $\underline{d}(A) \geqq C_h$ such that*

$$D \subseteq D_1 \cap \ldots \cap D_h.$$

In fact it follows from Theorem 3 that Theorem 1 remains true even with the stronger conclusion $D = D_1 \cap \ldots \cap D_h$.

By Corollary 1 we have $\underline{d}(D) \geqq \bar{d}(A)$ and thus we see from the above theorem that

$$\underline{d}(D_1 \cap \ldots \cap D_h) \geqq C_h.$$

Apart from the factor $5 \log(h + 1)$, which appears in the definition of C_h, Theorem 1 is best possible. For let $n_1, n_2, \ldots, n_h$ be positive integers and put $A_1 = \{a \mid a \geqq 0 \text{ and } a \equiv 0 \pmod{n_1}\}$ and $A_i = \{a \mid a \geqq 0 \text{ and } a \equiv 0, 1, \ldots, n_1 \ldots n_{i-1} - 1 \pmod{n_1 \ldots n_i}\}$ for $i = 2, \ldots, h$. We then have $d(A_i) = 1/n_i$ for $i = 1, \ldots, h$. Furthermore $D_1 = \{a \mid a \geqq 0 \text{ and } a \equiv 0 \pmod{n_1}\}$ while $D_i = \{a \mid a \geqq 0 \text{ and } a \equiv 0, \pm 1, \pm 2, \ldots, \pm (n_1 \ldots n_{i-1} - 1) \bmod (n_1 \ldots n_i)\}$ for $i = 2, \ldots, h$. An easy induction shows that

$$D_1 \cap \ldots \cap D_h = \{a \mid a \geqq 0 \text{ and } a \equiv 0 \pmod{n_1 \ldots n_h}\}.$$

Received September 19, 1977 and in revised form August 3, 1978

Therefore

$$d(D_1 \cap \ldots \cap D_h) = (\prod_{i=1}^h n_i)^{-1} = \prod_{i=1}^h d(A_i) = \prod_{i=1}^h \epsilon_i.$$

One might ask whether $D_1 \cap \ldots \cap D_h$ can contain gaps of arbitrary length. It will follow as a consequence of our next theorem that this is not possible. Independently Prikry [**7**] has obtained this result by means of a theorem of Hindman [**5**]. Further his proof remains valid if D_i is replaced by

$$\{x \mid \bar{d}(A_t \cap A_i + x) > 0\}$$

for $i = 1, \ldots, h$; here $A + k$ is the set $\{a + k \mid a \in A\}$. From Theorem 1 we see that it is sufficient to show that the difference set of a sequence of positive upper density does not contain arbitrarily long gaps. We denote the non-negative integers by $\mathbf{N}_0$.

THEOREM 2. *Let A be a sequence with $\bar{d}(A) = \epsilon > 0$. Then there exist r integers $k_1, \ldots, k_r$ such that*

$$\bigcup_{j=1}^r (D + k_j) \supseteq \mathbf{N}_0.$$

with $r \leqq \epsilon^{-(\log 3)/\log 2}$.

It follows from Theorem 2 that D cannot contain gaps of size larger than twice the maximum in absolute value of the k_j's. For if there was a larger gap the integers closest to the middle of the gap would not be in the union of the sets $D + k_j$ contradicting Theorem 2. We observe that it is vain to hope for an estimate for $\max_j |k_j|$ in terms of ϵ. For let A denote the set of integers of the form $3nt + i$ for $i = 1, \ldots, t$ and $n = 0, 1, 2, \ldots$. Then D consists of the non-negative integers of the form $3nt \pm i$ for $i = 0, \ldots, t$ and $n = 0, 1, 2, \ldots$ and so contains infinitely many gaps of length t. On the other hand $d(A) = 1/3$.

Theorems 1 and 2 show that infinite-difference sets possess a certain regularity. This might suggest that every infinite-difference set associated with a sequence of positive upper density has a density. However this is certainly not the case since we have

THEOREM 3. *Let D be the infinite-difference set of a sequence A. Let E be a set of non-negative integers with $D \subseteq E$. Then there exists a sequence B with $\bar{d}(B) = \bar{d}(A)$ and $\underline{d}(B) = \underline{d}(A)$ whose infinite-difference set is E.*

An immediate consequence of this result is that there exist sequences A with $\bar{d}(A) = \underline{d}(A) > 0$ for which $\bar{d}(D) > \underline{d}(D)$. Further, Theorem 3 is a step in the proof of the following theorem concerning $\mathbf{D}$, the collection of infinite-difference sets associated with sequences of positive upper density. Let $\mathscr{P}(\mathbf{N}_0)$ denote the set of all subsets of $\mathbf{N}_0$. We have

THEOREM 4. $\mathbf{D}$ *is a filter of* $\mathscr{P}(\mathbf{N}_0)$. *Furthermore all cofinite subsets of* $\mathbf{N}_0$ *which contain zero are in* $\mathbf{D}$.

D is not an ultrafilter, for there exist disjoint sets B_1 and B_2 satisfying $B_1 \cup B_2 = \mathbf{N}_0$ and $\underline{d}(B_1) = \underline{d}(B_2) = 0$; by Corollary 1 every infinite-difference set associated with a sequence of positive upper density has a positive lower density and thus neither B_1 nor B_2 is in **D**.

We define the difference set of a finite or infinite sequence A to be the set of those non-negative integers which occur as the difference of two elements of A and we denote this set by $\mathscr{D}(A)$. It is interesting to note that the collection of all difference sets associated with sequences of positive upper density does not form a filter. First, the collection does not satisfy the superset property. Observe that while $\mathscr{D}(E) = E$, where E denotes the non-negative even integers there exists no sequence A with $\mathscr{D}(A) = E \cup \{1\}$. Second, the collection does not satisfy the intersection property as the following example shows. Put $A = \{a \mid a \geqq 0 \text{ and } a \equiv 0 \pmod{10}\} \cup \{7\}$ and $B = \{b \mid b \geqq 0 \text{ and } b \equiv 7 \pmod{10}\} \cup \{0\}$; it is readily checked that $\mathscr{D}(A) \cap \mathscr{D}(B) = A$ and that there is no sequence C of positive upper density with $\mathscr{D}(C) = A$. It would be desirable to explicitly describe those sets which are infinite-difference sets or difference sets of sequences of positive upper density. A first attempt for the case of difference sets has been made by Ruzsa [**9**].

Obviously one always has $D \subseteq \mathscr{D}(A)$. On the other hand we have

THEOREM 5. *Given a sequence A with positive upper density there exists a sequence A' with $\bar{d}(A) \leqq \underline{d}(A')$ such that $\mathscr{D}(A') \subseteq D$.*

It follows from the above theorem that we may replace D by $\mathscr{D}(A)$ in the statement of Theorem 1; hence plainly the analogous statement of Theorem 1 holds with difference sets in place of infinite-difference sets.

An infinite difference set need not contain an infinite arithmetical progression. In fact we shall show that for every α with $0 < \alpha < 1$ there exist sequences A with density α for which the intersection of $\mathscr{D}(A)$ with any infinite arithmetical progression of difference v is a set of density at most $2\alpha/v$. Let $|X|$ be the cardinality of a set X and denote the set $\{0, 1, \ldots, n-1\}$ by $\hat{n}$. We have

THEOREM 6. *Let θ be an irrational number and let α be a number between* 0 *and* 1. *There exists a sequence A with density α for which*

$$\limsup_{n \to \infty} \frac{|\mathscr{D}(A) \cap E \cap \hat{n}|}{|E \cap \hat{n}|} \leqq 2\alpha$$

for every sequence $E = \{e_1, e_2, \ldots\}$ such that $\{\theta e_k\}_{k=1}^{\infty}$ is uniformly distributed modulo 1.

It is well known (see e.g. [**6**] Ch. 1, Theorem 4.1) that for any sequence $E = \{e_1, e_2, \ldots\}$ the sequence $\{\eta e_k\}_{k=1}^{\infty}$ is uniformly distributed modulo 1 for almost all real numbers η. Hence, given countably many sequences $E^{(i)} = \{e_k^{(i)}\}$ we can find an irrational number θ for which $\{\theta e_k^{(i)}\}$ is uniformly distributed modulo one for all i. In particular it follows from Theorem 6 that

https://doi.org/10.4153/CJM-1979-085-6 Published online by Cambridge University Press

for every α with $0 < \alpha < 1$ there exists a sequence A with density α such that

$$\limsup_{n\to\infty} \frac{|\mathscr{D}(A) \cap E \cap \hat{n}|}{|E \cap \hat{n}|} \leqq 2\alpha$$

for every arithmetical progression $\{ak + b\}_{k=1}^{\infty}$ with $a, b \in \mathbf{N}_0$, $a > 0$, for every geometrical progression $\{ab^k\}_{k=1}^{\infty}$ with $a, b \in \mathbf{N}_0$, $a > 0$, $b > 1$, and for every sequence $\{P(k)\}_{k=1}^{\infty}$, where $P(x)$ is a non-constant polynomial mapping $\mathbf{N}_0$ into $\mathbf{N}_0$.

Theorem 7 concerns sequences which have a non-empty intersection with every infinite-difference set D associated with a sequence A of positive upper density. We prove that there are arbitrarily thin sequences of positive integers with this property.

THEOREM 7. *For every sequence $f_1, f_2, \ldots$ there exists a sequence $E = \{e_1, e_2 \ldots\}$ with $e_j \geqq f_j$ for all j such that for every sequence A*

$$\liminf_{n\to\infty} \frac{|D \cap E \cap \hat{n}|}{|E \cap \hat{n}|} \geqq \bar{d}(A).$$

The sequence E constructed for the proof of Theorem 7 has the property that for all positive integers h,

$$\liminf_{j\to\infty} e_{j+h}/e_j = 1.$$

This condition is critical. On the one hand we have

THEOREM 8. *If $k_1, k_2, \ldots$ is a sequence of positive integers satisfying*

$$\liminf_{j\to\infty} k_{j+h}/k_j = 1$$

for every positive integer h, then there exists a sequence $E = \{e_1, e_2, \ldots\}$ with $e_{j+1}/e_j \geqq k_{j+1}/k_j$ for $j = 1, 2, \ldots$ such that $D \cap E \neq \emptyset$ for every sequence A of positive upper density.

On the other hand, if $k_1, k_2, \ldots$ is a sequence satisfying

$$\liminf_{j\to\infty} k_{j+h}/k_j > 1$$

for some h, then for every sequence $E = \{e_1, e_2, \ldots\}$ with $e_{j+1}/e_j \geqq k_{j+1}/k_j$ for $j = 1, 2, \ldots$ there exists a sequence A with $\underline{d}(A) > 0$ for which $D \cap E = \emptyset$. This result is a consequence of Theorem 9.

THEOREM 9. *Let $k_1, k_2, \ldots$ be a sequence of positive integers. If, for a positive integer h and for real numbers $c_1, \ldots, c_h$ larger than 2, we have*

$$k_{(j+1)h+i}/k_{jh+i} \geqq c_i,$$

for $i = 1, \ldots, h$ and $j = 0, 1, 2, \ldots$, *then there exists a sequence* A, *having a density, with*

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

for which $k_j \notin \mathscr{D}(A)$ *for* $j = 1, 2, \ldots$.

Observe that if $k_{j+h}/k_j \geqq \alpha > 1$ for $j = 1, 2, \ldots$, and if g is an integer with $g \geqq (\log 3)/\log \alpha$, then $k_{j+gh}/k_j \geqq 3$ for $j = 1, 2, \ldots$ since

$$\frac{k_{j+gh}}{k_j} = \frac{k_{j+gh}}{k_{j+(g-1)h}} \cdots \frac{k_{j+h}}{k_j} \geqq \alpha^g \geqq 3.$$

Further, if $\liminf_{j\to\infty} k_{j+l}/k_j > 1$ for some positive integer l then there exists a real number α with $\alpha > 1$ such that $k_{j+l}/k_j \geqq \alpha$ for $j = 1, 2, \ldots$. Thus we may apply Theorem 9 with $h = gl$ and $c_1 = c_2 = \ldots = c_h = 3$ to conclude that there exists a sequence A having a positive density with $k_j \notin \mathscr{D}(A)$ for $j = 1, 2, \ldots$ as was asserted previously.

To illustrate Theorem 9 we show that there exists a sequence A with $d(A) \geqq 2/11$ which does not have a factorial as the difference of two terms. This follows on putting $k_1 = 1!$, $k_2 = 2!$, ... and applying the theorem with $h = 2$, $c_1 = 6$ and $c_2 = 12$.

Theorem 9 is related to a general problem of Motzkin who asked how dense a sequence A can be if $\mathscr{D}(A)$ does not contain any elements from a given set K. Cantor and Gordon [**1**] and more recently Haralambis [**4**], have obtained some results in this connection, mainly for finite sets K. Sarközy [**10**], [**11**] and [**12**] considered the case of some interesting infinite sets K. He obtained results like: if A is a sequence with positive upper density then two distinct elements of A differ by a square. Furstenberg [**3**], using the methods of ergodic theory, has also proved this result. Let $K = \{k_1, k_2, \ldots\}$ be a sequence of positive integers for which $k_{i+1} - k_i \to \infty$ as $i \to \infty$. In response to a question of Erdös and Hartman, Rotenberg [**8**] showed that every infinite sequence A possesses an infinite subsequence A' for which $\mathscr{D}(A') \cap K = \emptyset$. In conclusion we should like to thank M. Best and P. Erdös for some helpful comments.

2. Preliminary lemmas. For any subset T of $\hat{n}$ and any integer a we put $T(a) = T + a \cap \hat{n}$ where $T + a$ denotes the set of numbers $t + a$ with $t \in T$. We prove

LEMMA 1. *Let* δ *and* ϵ *satisfy* $0 < \delta < 1$, $0 < \epsilon < 1$. *If* T *is a subset of* $\hat{n}$ *with* $|T| \geqq \epsilon n$ *then there exist integers* $k, a_1, \ldots, a_k$ *and a set* E *with* $|E| \leqq \delta n$ *such that*

$$T \cup T(a_1) \cup \ldots \cup T(a_k) = \hat{n} \backslash E$$

and such that

$$k \leqq 2[(\log \delta)/\log(1 - \epsilon)].$$

Proof. We first observe that $C(a) = T(a) \cup T(a - n)$ is a cyclic shift of T for $a = 0, \ldots, n - 1$ and hence $|C(a)| \geqq \epsilon n$. Further, given any subset G of $\hat{n}$ with θn terms for $0 \leqq \theta \leqq 1$ we may find an integer b for which $|C(b) \cap G| \geqq \epsilon\theta n$. To see this note that each integer from $\hat{n}$ is contained in at least ϵn of the cyclic shifts $C(0), \ldots, C(n - 1)$. Thus

$$\sum_{a=0}^{n-1} |C(a) \cap G| \geqq \epsilon\theta n^2$$

and as a consequence $|C(b) \cap G| \geqq \epsilon\theta n$ for some integer b, as required.

Now set $G_1 = \hat{n}\backslash T$. We have $|G_1| = \theta_1 n$ where $\theta_1 \leqq 1 - \epsilon$ since $|T| \geqq \epsilon n$. By the above paragraph we may find an integer b_1 such that $|C(b_1) \cap G| \geqq \epsilon\theta_1 n$ and thus $G_2 = \hat{n}\backslash\{T \cup C(b_1)\}$ satisfies $|G_2| = \theta_2 n$ for $\theta_2 \leqq \theta_1 - \epsilon\theta_1 \leqq (1 - \epsilon)^2$. Iterating this argument $l - 1$ times yields integers $b_1, \ldots, b_{l-1}$ and a set $G_l = \hat{n}\backslash\{T \cup C(b_1) \cup \ldots \cup C(b_{l-1})\}$ satisfying $|G_l| \leqq (1 - \epsilon)^l n$. On recalling that $C(b_i) = T(b_i) \cup T(b_i - n)$ we see that if $l - 1 = [\log \delta/\log(1 - \epsilon)]$ then $T \cup T(b_1) \cup T(b_1 - n) \cup \ldots \cup T(b_{l-1}) \cup T(b_{l-1} - n) = \hat{n}\backslash G_l$ where $|G_l| \leqq \delta n$. Putting $2(l - 1) = k$, $b_i = a_{2i} - 1$ and $b_i - n = a_{2i}$ for $i = 1, \ldots, l - 1$ and $G_l = E$ the lemma follows.

LEMMA 2. *Let A be a sequence with $\bar{d}(A) = \epsilon$. For any positive integers b and r there are at least $[\epsilon r]$ of the integers $b, 2b, \ldots, rb$ in D.*

Proof. Split A into b subsequences $A_j = A \cap \{ib + j\}_{i=0}^{\infty}$ for $j = 0, 1, \ldots, b - 1$. At least one of the sequences A_j satisfies $\bar{d}(A_j) \geqq \epsilon/b$. We define the sequence B by $i \in B$ if and only if $ib + j \in A_j$ for this particular value of j. Let D_0 be the infinite difference set of B. It is clear that if $d \in D_0$ then $bd \in D_j \subseteq D$. Hence, it suffices to prove that at least $[\epsilon r]$ of the integers $1, 2, \ldots, r$ belong to D_0. Plainly we may assume that $\epsilon > 0$.

Since $\bar{d}(B) \geqq \epsilon$, there are infinitely many integers m_i such that $|B \cap [m_i, m_i + r]| > \epsilon r$. By the box principle there is a set of $[\epsilon r] + 1$ integers $b_0, \ldots, b_{[\epsilon r]}$ with $0 \leqq b_0 < b_1 < \ldots < b_{[\epsilon r]} \leqq r$ such that for infinitely many integers m_i one has $m_i + b_k \in B$ for $k = 0, 1, \ldots, [\epsilon r]$. It follows that $b_k - b_0$ $(k = 1, \ldots, [\epsilon r])$ are $[\epsilon r]$ differences which occur in D_0. This proves our assertion whence the lemma follows.

COROLLARY 1. *For any sequence A we have $\underline{d}(D) \geqq \bar{d}(A)$.*

Proof. The result follows on taking $b = 1$ in Lemma 2.

Let $||x||$ denote the distance from x to the nearest integer.

LEMMA 3. *Let c be a real number, larger than 2, and let $k_1, k_2, \ldots$ be a sequence of positive integers with*

$$k_{j+1}/k_j \geqq c,$$

for $j = 1, 2, \ldots$. *Then there exists a real number* θ *such that*

$$||k_j\theta|| \geqq (c-2)/2(c-1),$$

for $j = 1, 2, \ldots$.

Proof. See Lemma 1 of [**2**].

3. Proof of theorem 1. Let $A_1, \ldots, A_h$ be sequences with $\bar{d}(A_i) \geq \epsilon_i > 0$ for $i = 1, \ldots, h$. We first observe that if $\epsilon_i > \frac{1}{2}$ for some integer i then $D_i = \mathbf{N}_0$. For if there is a positive integer k which is not in D_i then the sequence $A_i' = A_i \cup A_i + k$ satisfies $\bar{d}(A_i') \geqq 2\bar{d}(A_i)$ which is plainly impossible for $\epsilon_i > \frac{1}{2}$. To see that $\bar{d}(A_i') \geqq 2\bar{d}(A_i)$ note that $|A_i \cap A_i + k| = l < \infty$ by assumption and thus for all $n > 0$,

$$|A_i' \cap \hat{n}| \geqq |A_i \cap \hat{n}| + |A_i + k \cap \hat{n}| - l \geqq 2|A_i \cap \hat{n}| - k - l$$

from which the conclusion follows. Accordingly we may assume that the ϵ_i's are all at most $\frac{1}{2}$.

We shall construct, for each positive integer n, a set $W_n = W$ with

(2) $W \subseteq \hat{n}$, $|W| \geqq C_h n$ and $\mathscr{D}(W) \subseteq D_1 \cap \ldots \cap D_h$.

Since $\bar{d}(A_i) \geqq \epsilon_i$, for each positive integer n there are infinitely many integers k for which $(A_i - k) \cap \hat{n}$ contains at least $\epsilon_i n$ terms. By the pigeon hole principle there exists an infinite subsequence $k_1, k_2, \ldots$ of the k's for which $(A_i - k_1) \cap \hat{n} = (A_i - k_2) \cap \hat{n} = \ldots$. Set $T_i = (A_i - k_1) \cap \hat{n}$. Plainly we have $T_i \subseteq \hat{n}$, $|T_i| \geqq \epsilon_i n$ and

(3) $\mathscr{D}(T_i) \subseteq D_i$.

Thus if $h = 1$ we may take $W = T_1$ and (2) holds.

Assume that $h \geqq 2$. On setting $\delta = (h+1)^{-1}$, $\epsilon = \epsilon_i$ and $T = T_i$ in Lemma 1 we conclude that there exist integers $k(i)$, $a_{i,1}, \ldots, a_{i,k(i)}$ and a set E_i with $|E_i| \leqq n/(h+1)$ such that

(4) $T_i \cup T_i(a_{i,1}) \cup \ldots \cup T_i(a_{i,k(i)}) = \hat{n} \backslash E_i$

and such that

(5) $k(i) \leqq 2[-\log(h+1)/\log(1-\epsilon_i)]$,

for $i = 1, \ldots, h$. Put

$$F = \bigcap_{i=1}^{h} (\hat{n} \backslash E_i).$$

By construction, $|F| \geqq n/(h+1)$. Setting $a_{i,0} = 0$, so that $T_i = T_i(a_{i,0})$, we find from (4) that

$$F \subseteq \bigcup \bigcap_{i=1}^{h} T_i(a_{i,j(i)}),$$

where the union is taken over the $(k(1)+1)\dots(k(h)+1)$ h-tuples $(j(1),\dots,j(h))$ with $0 \leqq j(i) \leqq k(i)$. Thus for at least one h-tuple the set

$$W = T_1(a_{1,j(1)}) \cap \dots \cap T_h(a_{h,j(h)})$$

contains at least

(6) $\quad w = n/(h+1)(k(1)+1)\dots(k(h)+1)$

terms. Clearly $\mathscr{D}(W) \subseteq \bigcap_{i=1}^h \mathscr{D}(T_i(a_{i,j(i)}))$ and therefore, since $\mathscr{D}(T_i(a)) \subseteq \mathscr{D}(T_i)$ for all integers a, $\mathscr{D}(W) \subseteq \bigcap_{i=1}^h \mathscr{D}(T_i)$. Thus from (3),

$$\mathscr{D}(W) \subseteq D_1 \cap \dots \cap D_h.$$

Since W is plainly contained in $\hat{n}$ we need only show that $|W| \geqq w \geqq C_h n$. We have from (6) and (5) that

$$w \geqq \frac{n}{h+1} \prod_{i=1}^h \left(\frac{2\log(h+1)}{-\log(1-\epsilon_i)} + 1\right)^{-1}$$

which, since $0 \leqq \epsilon_i \leqq \frac{1}{2}$ for $i = 1,\dots,h$, gives

$$w \geqq \frac{n}{h+1} \prod_{i=1}^h \left(\frac{-\log(1-\epsilon_i)}{2\log(h+1)+\log 2}\right).$$

We may now use the inequality $-\log(1-x) \geqq x$, which holds for $0 \leqq x \leqq \frac{1}{2}$, and the fact that $h \geqq 2$ to deduce that

$$w \geqq n \prod_{i=1}^h \left(\frac{\epsilon_i}{\sqrt{3}(2+\log 2/\log 3)\log(h+1)}\right).$$

It is easily checked that $w \geqq C_h n$. Thus (2) is seen to hold for $h \geqq 2$ as well as for $h = 1$.

We construct the sequence A from the sets W_n in the following way. For $n = 1, 2, 3, \dots$ we set $Q(n) = \binom{n}{2}$ and put $A \cap [Q(n), Q(n+1)) = W_{n-[\log n]} + Q(n)$. The sequence A is well defined since, for $n \geqq 1$, $W_{n-[\log n]} + Q(n) \subseteq [Q(n), Q(n)+n]$ and $Q(n) + n = Q(n+1)$. We now show that $\underline{d}(A) \geqq C_h$. Given a positive integer m we define k by the inequalities $Q(k) \leqq m \leqq Q(k+1)$. From (2) we have $|W_n| \geqq C_h n$ for $n = 1, 2, \dots$ and thus

$$|A \cap \hat{m}| \geqq \sum_{n=1}^{k-1} |W_{n-[\log n]}| \geqq \sum_{n=1}^{k-1} C_h(n - [\log n]) \geqq C_h(Q(k) - k[\log k]).$$

Therefore

$$\frac{|A \cap \hat{m}|}{m} \geqq \frac{C_h Q(k) - k[\log k]}{Q(k+1)} \geqq C_h - \frac{2(1+[\log k])}{k}.$$

Letting m and hence k tend to infinity we see that $\underline{d}(A) \geqq C_h$.

Finally we show that $D \subseteq D_1 \cap \dots \cap D_h$. The terms of A in the interval $[Q(n), Q(n+1)]$ differ, by construction, by at least $[\log n]$ from the terms of

https://doi.org/10.4153/CJM-1979-085-6 Published online by Cambridge University Press

the interval $[Q(n+1), Q(n+2)]$. Thus if a difference occurs infinitely often in A it must occur as the difference of two elements from the interval $[Q(m), Q(m+1)]$ for some positive integer m and so it must be contained in $\mathscr{D}(W_{m-[\log m]})$. From (2) we see that the difference is contained in $D_1 \cap \ldots \cap D_h$ as required. This completes the proof.

4. Proof of theorem 2. We set $D^0 = D$ and $l_0 = 0$. For $i \geqq 1$ we define l_1 and D^i whenever $D^{i-1} \not\supseteq \mathbf{N}_0$ by the following inductive process: set l_i equal to the smallest positive integer which is not in D^{i-1} and put $D^i = D^{i-1} \cup D^{i-1} + l_i \cup D^{i-1} - l$. We shall prove that $D^s \supseteq \mathbf{N}_0$ for some positive integer s satisfying $s \leqq [-(\log \epsilon)/\log 2]$ where C_h is defined as in (1). This will establish the theorem since

$$D^s = \bigcup_{i=1}^{r} (D^0 + k_i),$$

where the k_i are the $r = 3^s$ finite sums of the form $a_1 + \ldots + a_s$ with a_i one of 0, l_i or $-l_i$ for $i = 1, \ldots, s$.

As in the proof of Theorem 1, for any positive integer n there exists by the pigeon hole principle; a set $T \subseteq \hat{n}$ with $|T| \geqq \epsilon n$ satisfying $\mathscr{D}(T) \subseteq D$. Assume that D^i, hence also l_i, has been defined for $i = 0, \ldots s$. Set $T = T^0$ and define T^i to be $T^{i-1} \cup T^{i-1} + l_i$, for $i = 1, \ldots, s$. For any set of integers A and any integer l it is readily checked that $\mathscr{D}(A \cup A + l) \subseteq \mathscr{D}(A) \cup \mathscr{D}(A) + l \cup \mathscr{D}(A) - l$. From the definition of D^i and the fact that $\mathscr{D}(T^0) \subseteq D^0$ we conclude that $\mathscr{D}(T^i) \subseteq D^i$ for $i = 0, \ldots, s$. Therefore l_{i+1} does not occur as the difference of two terms in T^i since by assumption l_{i+1} is not in D^i. Accordingly $T^i \cap T^i + l_{i+1} = \emptyset$ so that $|T^{i+1}| = 2|T^i|$ and thus $|T^s| = 2^s|T^0| \geqq 2^s\epsilon n$. On the other hand $T^s \subseteq [0, n + l_1 + \ldots + l_s]$ and therefore

$$2^s\epsilon n \leqq n + l_1 + l_2 + \ldots + l_s + 1.$$

Dividing by n and letting n tend to infinity we see that $2^s\epsilon \leqq 1$ whence

$$s \leqq [-(\log \epsilon)/\log 2]$$

as required.

5. Proof of theorem 3. Let $E = \{e_1, e_2, \ldots\}$. We construct a sequence $F = \{f_1, f_2, \ldots\}$ by setting, for $n = 1, 2, \ldots : f_n = e_j$ where j is the unique integer satisfying both $n = \binom{m}{2} + j$ and $1 \leqq j \leqq m$ for some positive integer m. Note that every element of E occurs infinitely often as a term of F.

We now construct B. The terms of B are the integers $3^{e_n} + e_n$ and $3^{e_n} + e_n + f_n$ and those integers of A which do not lie in the intervals $[3^{e_n}, 3^{e_n} + 3e_n]$ for $n = 1, 2, \ldots$. Since E is an increasing sequence of non-negative integers,

$e_n \geqq n - 1$ and $\lim_{n\to\infty} (3ne_n)/3^{e_n} = 0$, whence B differs from A only on a set of density zero. Thus $\bar{d}(B) = \bar{d}(A)$ and $\underline{d}(B) = \underline{d}(A)$.

The intervals $[3^{e_n}, 3^{e_n} + 3e_n]$ are disjoint for $n = 1, 2, \ldots$. Further, $f_n \leqq e_n$ for all $n > 0$ and thus the difference of an element of B from the interval $[3^{e_n}, 3^{e_n} + 3e_n]$ with one not from this interval is $\geqq e_n$. Since $e_n \to \infty$ as $n \to \infty$ the infinite difference set of B is equal to the union of those integers which occur infinitely often as the difference of two terms of B neither of which is in $\bigcup_{n=1}^{\infty} [3^{e_n}, 3^{e_n} + 3e_n]$ with those integers which occur as the difference of two terms of B in $[3^{e_n}, 3^{e_n} + 3e_n]$ for infinitely many integers n. The former set is plainly contained in $D \subseteq E$ while the latter set is exactly E since $3^{e_n} + e_n + f_n - (3^{e_n} + e_n) = f_n$ is the only positive integer which occurs as the difference of two terms of B from $[3^{e_n}, 3^{e_n} + 3e_n]$ and since every element of E occurs infinitely often as a term of F. This completes the proof.

6. Proof of theorem 4. To prove that $\mathbf{D}$ is a filter of $\mathscr{P}(\mathbf{N}_0)$ we must show that (i) $\mathbf{D} \neq \emptyset$, (ii) $D \neq \emptyset$ for $D \in \mathbf{D}$, (iii) $D \in \mathbf{D}$ and $D \subseteq E \subseteq \mathbf{N}_0$ then $E \in \mathbf{D}$, (iv) $D_1 \cap D_2 \in \mathbf{D}$ for $D_1, D_2 \in \mathbf{D}$. Properties (i) and (ii) are readily seen to hold. Property (iii) follows from Theorem 3. Property (iv) follows from property (iii) and Theorem 1. Therefore $\mathbf{D}$ is a filter of $\mathscr{P}(\mathbf{N}_0)$.

Further we must show that every cofinite subset of $\mathbf{N}_0$ which contains zero is in $\mathbf{D}$. Given a set of positive integers $n_1 < n_2 < \ldots < n_k$ we consider the set of positive multiples of $n_k + 1$. This has an infinite-difference set which does not contain $n_1, \ldots, n_k$ and so by the superset property (iii) we can find a D which is exactly $\mathbf{N}_0 \backslash \{n_1, \ldots, n_k\}$. This completes the proof.

7. Proof of theorem 5. Put $\epsilon = \bar{d}(A)$ and let n be any positive integer. We prove first that there exist infinitely many integers m such that $|A \cap [m, m + k)| \geqq \epsilon k$ for $k = 1, \ldots, n$. Suppose this statement is false. Then for every $m \geqq m_0$ there exists a k_m with $1 \leqq k_m \leqq n$ such that $|A \cap [m, m + k_m)| < \epsilon k_m$. Put

$$\epsilon' = \{\max i/k \mid i, k \in \mathbf{N}_0, 1 \leqq k \leqq n, i/k < \epsilon\}.$$

Note that $\epsilon' < \epsilon$ and that for every $m \geqq m_0$ we have $|A \cap [m, m + k_m)| \leqq \epsilon' k_m$. Define the sequence $m_0, m_1, m_2, \ldots$ inductively by putting $m_{j+1} = m_j + k_j$. Let x be at least m_0 and define J by the inequalities $m_J \leqq x \leqq m_{J+1}$. Since for every positive integer j we have $|A \cap [m_j, m_{j+1})| \leqq \epsilon'(m_{j+1} - m_j)$ the number of elements of A less than x is at most

$$m_0 + \epsilon'(m_J - m_0) + x - m_J \leqq \epsilon' x + m_0 + n.$$

Thus $\bar{d}(A) \leqq \epsilon'$ which is a contradiction.

Let $r_1^{(n)}, r_2^{(n)}, \ldots$ be a sequence such that $|A \cap [r_j^{(n)}, r_j^{(n)} + k)| \geqq \epsilon k$ for $j = 1, 2, \ldots$ and $k = 1, \ldots, n$. We consider the sets $(A - r_j^{(n)}) \cap \hat{n}$. By the pigeon hole principle there exists on infinite subsequence $\{s_j^{(n)}\}$ of $\{r_j^{(n)}\}$ such that $(A - s_j^{(n)}) \cap \hat{n}$ is the same set $S^{(n)}$ for every j. We obtain in this way a set $S^{(n)}$, for every positive integer n, such that for $k = 1, \ldots, n$ the number of elements less than k is at least ϵk. We now construct the sequence A' by induction. Suppose $A' \cap \hat{n}$ has been constructed in such a way that there are infinitely many integers ν with $S^{(\nu)} \cap \hat{n} = A' \cap \hat{n}$. We put $n \in A'$ if and only if there are infinitely many integers ν' among these integers ν with $n \in S^{(\nu')}$. It follows that there are infinitely many integers ν with $S^{(\nu)} \cap (\widehat{n+1}) = A' \cap (\widehat{n+1})$. By construction the number of elements of A' less than n is equal to the number of elements of $S^{(\nu)}$ less than n for some $\nu > n$ and hence is at least ϵn. Thus $\underline{d}(A') \geqq \epsilon$. Let a_1' and a_2' be any two elements of A' with $a_1' < a_2'$. Then a_1' and a_2' are in $S^{(\nu)}$ for some integer ν. Therefore $a_i' + s_j^{(\nu)} \in A$ and $a_2' + s_j^{(\nu)} \in A$ for $j = 1, 2, \ldots$ whence $a_2' - a_1' \in D$. This completes the proof.

Note that we have even proved that the Schnirelmann density of $A' + 1$ is at least ϵ since $|A' \cap [0, n)| \geqq \epsilon n$ for every positive integer n.

8. Proof of theorem 6. Let θ be an irrational number and let α be a number between 0 and 1. Define A to be the sequence composed of those non-negative integers n for which an integer m exists with $n\theta - m \in (0, \alpha)$. Since $\{n\theta\}_{n=1}^{\infty}$ is uniformly distributed modulo 1, $d(A) = \alpha$. If $n \in \mathscr{D}(A)$, then $n = n_1 - n_2$ with $n_1, n_2 \in A$, and there exist m_1, m_2 such that $0 < n_1\theta - m_1 < \alpha$, $0 < n_2\theta - m_2 < \alpha$ and hence $-\alpha < (n_1 - n_2)\theta - (m_1 - m_2) < \alpha$. Thus $\mathscr{D}(A)$ consists of non-negative integers n for which an integer m exists with $n\theta - m \in (-\alpha, \alpha)$. Therefore, if $E = \{e_1, e_2, \ldots\}$ is a sequence for which $\{\theta e_k\}$ is uniformly distributed modulo 1, we have

$$\limsup_{n\to\infty} \frac{|\mathscr{D}(A) \cap E \cap \hat{n}|}{|E \cap \hat{n}|} \leqq 2\alpha.$$

9. Proof of theorem 7. Clearly we may assume, for all positive integers n, that

(7) $f_{n+1}/f_n > n$, for all n.

Put $Q(m) = \binom{m}{2}$ for $m = 1, 2, 3, \ldots$. We define the sequence E by setting

$$e_{Q(m-1)+j} = j f_{Q(m)}$$

for $j = 1, 2, \ldots, m - 1$ and $m = 2, 3, \ldots$. It follows that

$$e_{Q(m-1)+j} \geqq f_{Q(m)} = f_{Q(m-1)+m} \geqq f_{Q(m-1)+j}$$

for these values of j and m. Thus $e_j \geqq f_j$ for all j. Further, by (7), for $m \geqq 2$,

$$e_{Q(m)} = (m-1)f_{Q(m)} \leqq Q(m)f_{Q(m)} < f_{Q(m+1)} = e_{Q(m)+1}.$$

It follows that the sequence E is strictly increasing and further that the elements of E in the interval $[f_{Q(m)}, f_{Q(m+1)})$ are $jf_{Q(m)}$ for $j = 1, 2, \ldots, m - 1$.

Let A be a sequence with $d(A) = \alpha$. By Lemma 2, the number of elements of $D \cap E$ in the interval $[f_{Q(m)}, f_{Q(m+1)})$ is at least $[(m-1)\alpha]$. Let n be an integer larger than f_1. Take m such that $f_{Q(m)} \leqq n < f_{Q(m+1)}$. Then, as the numbers e_j are distinct,

$$\frac{|D \cap E \cap \hat{n}|}{|E \cap \hat{n}|} \geqq \frac{[\alpha] + [2\alpha] + \ldots + [(m-2)\alpha]}{1 + 2 + \ldots + (m-1)} \geqq \frac{\binom{m-1}{2}\alpha - m}{\binom{m}{2}}.$$

Hence,

$$\liminf_{n\to\infty} \frac{|D \cap E \cap \hat{n}|}{|E \cap \hat{n}|} \geqq \lim_{m\to\infty} \frac{\binom{m}{2}\alpha - 2m}{\binom{m}{2}} = \alpha.$$

10. Proof of theorem 8. Let $j(1), j(2), j(3), \ldots$ be a sequence with $j(h+1) > j(h) + h$ for $h = 1, 2, \ldots$ such that

$$k_{j(h)+h}/k_{j(h)} < 1 + 1/h.$$

Such a sequence exists since

$$\liminf_{j\to\infty} k_{j+h}/k_j = 1$$

for every positive integer h. We define the sequence E inductively. Put $e_1 = 1$. If i is not of the form $j(h) + l$ for some h and l with $2 \leqq l \leqq h$, then we choose e_i to be the smallest integer with $e_i/e_{i-1} \geqq k_i/k_{i-1}$. On the other hand, if i is of the form $j(h) + l$ with $2 \leqq l \leqq h$, we put $e_{j(h)+l} = le_{j(h)+1}$; since $j(h+1) > j(h) + h$, both h and l are uniquely determined and the sequence E is well defined. By construction

$$e_{j(h)+l}/e_{j(h)+l-1} \geqq 1 + h^{-1} > k_{j(h)+l}/k_{j(h)+l-1}$$

for $l = 2, 3, \ldots, h$ and hence $e_i/e_{i-1} \geqq k_i/k_{i-1}$ for all i. By Lemma 2 (compare with the proof of Theorem 7) the subsequence $e_{j(1)+1}$, $e_{j(2)+1}$, $e_{j(2)+2}$, $e_{j(3)+1}$, $e_{j(3)+2}$, $e_{j(3)+3}$, $e_{j(4)+1}, \ldots$ of E has a non-empty intersection with D for every sequence A of positive upper density. Hence, $D \cap E \neq \emptyset$, as required.

11. Proof of theorem 9. By Lemma 3 there exists a real number θ_i, for $i = 1, \ldots, h$ satisfying

(8) $\quad ||k_{jh+i}\theta_i|| \geqq (c_i - 2)/2(c_i - 1),$

for $j = 0, 1, 2, \ldots$. Put

(9) $\quad g_i = (c_i - 2)/2(c_i - 1),$

for $i = 1, \ldots, h$. Let $\{x\}$ denote the fractional part of x. Given i with $1 \leqq$

$i \leqq h$ and an integer l_i we put $\lambda_i = \{l_i g_i\}$ and $\gamma_i = \{(l_i + 1)g_i\}$. We then define the set A_{i,l_i} in the following manner: if $\gamma_i \geqq \lambda_i$ then

$$A_{i,l_i} = \{n \mid \lambda_i \leqq \{n\theta_i\} < \gamma_i\},$$

while if $\gamma_i < \lambda_i$ then

$$A_{i,l_i} = \{n \mid \text{either } \lambda_i \leqq \{n\theta_i\} < 1 \text{ or } 0 \leqq \{n\theta_i\} < \gamma_i\}.$$

We have $\mathscr{D}(A_{i,l_i}) \subseteq \{n \mid ||n\theta_i|| < g_i\}$ for $i = 1, \ldots, h$. We now set

$$A_{l_1,\ldots,l_h} = \{n \mid n \in A_{i,l_i} \text{ for } i = 1, \ldots, h\},$$

so that

$$\mathscr{D}(A_{l_1,\ldots,l_h}) \subseteq \{n \mid ||n\theta_i|| < g_i \text{ for } i = 1, \ldots, h\}.$$

It follows from (8) and (9) that $k_j \notin \mathscr{D}(A_{l_1,\ldots,l_h})$ for $j = 1, 2, \ldots$. Therefore it suffices to show that for some choice of $l_1, \ldots, l_h$, the sequence $A_{l_1,\ldots,l_h}$ has a density which is at least $g_1 g_2 \ldots g_h$.

We observe that for any real numbers $\lambda_1, \ldots, \lambda_h$ and $\gamma_1, \ldots, \gamma_h$ the set $\{n \mid \lambda_i \leqq \{n\theta_i\} < \gamma_i\}$ for $i = 1, \ldots, h$ possesses a density; this is a consequence of the uniform distribution of the points $(\{n\theta_1\}, \ldots, \{n\theta_h\})$ in the maximal linearly independent subspace generated by $\theta_1, \ldots, \theta_h$ in the h-dimensional unit cube and may be deduced from Weyl's criterion, (see p. 48 of [**6**]). Furthermore if A and B are disjoint sets possessing densities then $d(A \cup B) = d(A) + d(B)$. Using these observations we see, after taking an appropriate partition of the unit cube in blocks and a corresponding decomposition and regrouping of sets $A_{l_1,\ldots,l_h}$, that for every positive integer L

$$\sum_{l_1=0}^{L-1} \ldots \sum_{l_h=0}^{L-1} d(A_{l_1,\ldots,l_h}) \geqq [Lg_1][Lg_2] \ldots [Lg_h] d(\{n \mid 0 \leqq \{n\theta_i\} < 1 \text{ for } i = 1, \ldots, h\}).$$

But $d(\{n \mid 0 \leqq \{n\theta_i\} < 1 \text{ for } i = 1, \ldots, h\}) = d(\mathbf{N}_0) = 1$ and therefore

$$\liminf_{L\to\infty} \frac{1}{L^n} \sum_{l_1=0}^{L-1} \ldots \sum_{l_h=0}^{L-1} d(A_{l_1,\ldots,l_h}) \geqq g_1 g_2 \ldots g_h,$$

whence for some choice of $l_1, \ldots, l_h$ we have

$$d(A_{l_1,\ldots,l_h}) \geqq g_1 g_2 \ldots g_h.$$

The result now follows.

Remark. As the referee has pointed out, by considering the cartesian product of measure preserving systems associated with sequences of positive upper density as elaborated by Furstenberg in [**3**] it is possible to remove the factor $5 \log (h + 1)$ in Theorem 1. Recently Y. Katznelson and I. Ruzsa have found elementary proofs of this fact. Furthermore Kamae and Mendes France have used Fourier analysis to obtain some of the results of Sárközy referred to in §1, (see *Van der Corput's difference theorem*, Israel J. Math., *31* (1978), 335–342).

REFERENCES

1. D. G. Cantor and B. Gordon, *Sequences of integers with missing differences*, J. Comb. Th. (A) *14* (1973), 281–287.

2. J. W. S. Cassels, *The inhomogeneous minimum of binary quadratic, ternary cubic and quaternary quartic forms*, Proc. Camb. Phil. Soc. *48* (1952), 72–86.

3. H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. *31* (1977), 204–256.

4. N. M. Haralambis, *Sets of integers with missing differences*, J. Comb. Th. (A) *23* (1977), 22–33.

5. N. Hindman, *Finite sums from sequences within cells of a partition of* N, J. Comb. Th. (A) *17* (1974), 1–11.

6. L. Kuipers and H. Niederreiter, *Uniform distribution of sequences* (Wiley, New York, 1974).

7. K. Prikry, private communication.

8. D. Rotenburg, *Sur une classe de parties évitables*, Colloq. Math. *20* (1969), 67–68.

9. I. Z. Ruzsa, *On difference-sequences*, Acta Arith. *25* (1974), 151–157.

10. A. Sárközy, *On difference sets of sequences of integers I*, Acta Math. Acad. Sci. Hungar., *31* (1978), 125–149.

11. ——— *On difference sets of sequences of integers II*, Annales Univ. Sci. Budapest. Eötvös (to appear).

12. ——— *On difference sets of sequences of integers III*, Acta Math. Acad. Sci. Hungar., *31* (1978), 355–386.

University of Waterloo,
Waterloo, Ontario;
Mathematical Institute,
Wassenaarseweg 80,
Leiden, Netherlands

https://doi.org/10.4153/CJM-1979-085-6 Published online by Cambridge University Press