Reprinted from JOURNAL OF NUMBER THEORY
All Rights Reserved by Academic Press, New York and London

Vol. 28, No. 2, February 1988
Printed in Belgium

Sumsets Containing Infinite Arithmetic Progressions

PAUL ERDÖS

Mathematics Institute, Hungarian Academy of Sciences, Budapest, Hungary

MELVYN B. NATHANSON

Office of the Provost and Vice President for Academic Affairs, Lehman College (CUNY), Bronx, New York 10468

AND

ANDRÁS SÁRKÖZY

Department of Mathematics, Baruch College (CUNY), New York, New York 10010; and Mathematics Institute, Hungarian Academy of Sciences, Budapest, Hungary

Communicated by the Editors

Received March 9, 1987

Let A be a set of nonnegative integers such that $d_L(A) = w > 0$. Let k be the least integer satisfying $k \geqslant 1/w$. It is proved that there is an infinite arithmetic progression with difference at most $k+1$ such that every term of the progression can be written as a sum of exactly $k^2 - k$ distinct terms of A, and there is an infinite arithmetic progression with difference at most $k^2 - k$ such that every term of the progression can be written as a sum of exactly $k+1$ distinct terms of A. A solution is also obtained to the infinite analog of a problem of Erdös and Freud on powers of 2 and on square-free numbers that can be represented as bounded sums of distinct elements chosen from a set A with positive density.

1. Introduction

If a set A of nonnegative integers has positive upper asymptotic density, then A contains arbitrarily long finite arithmetic progressions (Szemerédi [3]). It is not true, however, that a set of positive upper asymptotic density must contain an infinite arithmetic progression. In fact, it is easy to construct a set of positive lower asymptotic density that does not contain an infinite arithmetic progression. For example, if $0 < w < 1$ and if x is real and irrational, let A consist of all nonnegative integers a

0022-314X/88 $3.00

such that $0 < \{ax\} < w$, where $\{ax\}$ denotes the fractional part of ax. Then A has density w, but contains no infinite arithmetic progression. Indeed, if $hw < 1$, then the h-fold sumset hA contains no infinite arithmetic progression.

In this paper we investigate infinite arithmetic progressions, each term of which can be represented as a sum of a bounded number of integers belonging to a fixed set of positive density. If a set A has positive upper asymptotic density, then it is not true that there must exist a positive integer h such that the sumset hA contains an infinite arithmetical progression. For example, let (t_n) be a strictly increasing sequence of positive integers such that t_{n+1}/t_n tends to infinity, and let the set A be the union of the intervals $[t_{2n}+1, t_{2n+1}]$. Then A has upper asymptotic density $d_U(A) = 1$ and lower asymptotic density $d_L(A) = 0$. For fixed h and all sufficiently large n, the sumset hA is disjoint from the interval $[ht_{2n-1}+1, t_{2n}]$. Thus, hA contains arbitrarily long gaps, and so cannot contain an infinite arithmetic progression.

We shall prove that if A has positive lower asymptotic density, then some sumset hA does contain an infinite arithmetic progression, and we can bound both the number h of summands and the difference g of the arithmetic progression in terms of the density of A. In addition, we show that each term of the arithmetic progression can be represented as a sum of h pairwise distinct elements of A.

These results are related to two problems of P. Erdös and R. Freud. They conjectured that if S is a finite set of integers contained in $[1, 3n]$ and $\text{card}(S) > n$, then there is a power of 2 that can be written as a sum of distinct elements of S. Also, they conjectured that if T is a finite set of integers contained in $[1, 4n]$ and $\text{card}(T) > n$, then there is a square-free number that can be written as a sum of distinct elements of T. Recently, G. Freiman [1] has solved both these problems. His proof, however, does not yield a uniform bound for the number of distinct summands needed to represent the power of 2 or the square-free number; it shows only that $\log n$ summands suffice. In a subsequent paper we shall give a solution to the Erdös–Freud problem with a uniform bound on the number of summands.

In this paper we give an infinite analog of these results. We show that if $d_L(a) \geqslant \frac{1}{3}$ and $3 \nmid a$ for some $a \in A$, then at least half the powers of 2 can be written as sums of at most five distinct elements of A. We also prove that if $d_L(A) \geqslant \frac{1}{4}$ and $4 \nmid a$ for some $a \in A$, then infinitely many square-free integers can be written as sums of at most six distinct elements of A.

Notation. For any set A of nonnegative integers, the counting function $A(x)$ denotes the number of positive elements of A not exceeding x. The lower asymptotic density of A is defined by $d_L(A) = \liminf A(x)/x$. The upper asymptotic density of A is defined by $d_U(A) = \limsup A(x)/x$. If $d_L(A) = d_U(A)$, then A has asymptotic density $d(A) = d_L(A)$. For $g \geqslant 1$, define $A^{(g)} = \{a' \geqslant 0 \mid a' \equiv a \pmod{g}$ for some $a \in A\}$. We write $A \sim B$ if the sets A and B coincide for all sufficiently large integers. The h-fold sumset of A, denoted hA, is the set of all sums of h elements of A, with repetitions allowed. Denote by $h^\wedge A$ the set of all sums of h *distinct* elements of A. The set A is an asymptotic basis of order h if $hA \sim N$, where N denotes the set of all nonnegative integers.

For any real number w, let $\langle w \rangle$ denote the smallest integer n such that $n \geqslant w$. Let $\{w\}$ denote the fractional part of w, and let $\|w\| = \min(\{w\}, 1 - \{w\})$ denote the distance to the nearest integer.

2. Arithmetic Progressions

In this section we obtain quantitative results on infinite arithmetic progressions contained in sumsets of sets of positive lower asymptotic density. If $d_L(A) > \frac{1}{2}$, then an elementary counting argument shows that A is an asymptotic basis of order 2, and so $2A$ contains an infinite arithmetic progression with difference 1. Therefore, it is sufficient to consider only sets A such that $d_L(A) \leqslant \frac{1}{2}$.

An essential tool in this paper is Kneser's theorem [2] in the following form: Let A be a set of nonnegative integers. Then either (i) $d_L(hA) \geqslant h\, d_L(A)$ or (ii) there exists a minimal integer $g \geqslant 1$ such that $hA \sim hA^{(g)}$ and $d_L(hA) \geqslant h\, d_L(A) - (h-1)/g$.

Lemma 1. *Let $t > 0$. Let A be a set of nonnegative integers. Define the set $A' \subseteq A$ by*

$$A' = \{a \in A \mid a + id \in A \text{ for some } d > 0 \text{ and all } |i| < t\}. \qquad (*)$$

Then $d(A \backslash A') = 0$. In particular, $d_L(A) = d_L(A')$ and $d_U(A) = d_U(A')$.

Proof. If $d_U(A \backslash A') > 0$, then Szemerédi's theorem implies that $A \backslash A'$ contains an arithmetic progression of length $2t - 1$, but this is impossible, since the middle term of this arithmetic progression would belong to A'. Therefore, $d(A \backslash A') = 0$.

Lemma 2. *Let A be a finite or infinite set of integers. Let $h \geqslant 1$. Define A' by $(*)$ with $t = h$. Then $hA' \subseteq h^\wedge A$.*

Proof. Let $n = a_1 + \cdots + a_h \in hA'$. Let F be a maximal subset of the summands a_j whose elements are pairwise distinct. If $\text{card}(F) = h$, then $n \in h^\wedge A$. If $\text{card}(F) < h$, choose $a_k \notin F$. There exists $j \neq k$ and $a_j \in F$ with $a_j = a_k$. Since $a_k \in A'$, it follows that there exists $d > 0$ such that $a_k + id \in A$

for all $|i| < h$. Choose $i > 0$ such that $a_k + id \notin F$ and $a_j - id \notin F$, and replace a_k and a_j with $a_k + id$ and $a_j - id$, respectively. This gives a new representation of n as a sum of h elements of A. Define $F' = (F \setminus \{a_j\}) \cup \{a_k + id, a_j - id\}$. The elements of F' are pairwise distinct, and $\text{card}(F') > \text{card}(F)$. Let G be a maximal subset of the summands in the new representation of n such that $G \supseteq F'$ and the elements of G are pairwise distinct. The summands in the new representation of n that do not belong to G are all elements of A'. Repeating the argument inductively gives a representation of n as a sum of h distinct elements of A. This proves the lemma.

THEOREM 1. *Let A be a set of nonnegative integers such that $d_L(A) = w \in (0, \frac{1}{2}]$. Define $k = \langle 1/w \rangle$. Then*

(i) *there exists $g \leqslant k^2 - k$ such that $(k+1)^\wedge A$ contains an infinite arithmetic progression with difference g;*

(ii) *there exists $g \leqslant k + 1$ such that $(k^2 - k)^\wedge A$ contains an infinite arithmetic progression with difference g.*

Proof. Let $s \geqslant 1$. Then $d_L((k+s) A) \leqslant 1 \leqslant kw < (k+s) d_L(A)$. Therefore, the second case of Kneser's theorem holds, and there exists a minimal integer g such that $(k+s) A \sim (k+s) A^{(g)}$. Then $(k+s) A$ contains an infinite arithmetic progression with difference g.

If $g = 1$, then $(k+s) A$ contains all sufficiently large integers. If $g > 1$, then $(k+s) A \sim (k+s) A^{(g)}$ and

$$(k+s) w - (k+s-1)/g \leqslant d_L((k+s) A) \leqslant 1 - (1/g).$$

Since $w \geqslant 1/k$, it follows that

$$1 < g \leqslant k + (k^2 - 2k)/s.$$

For $s = 1$, this inequality implies that $(k+1) A$ contains an infinite arithmetic progression with difference g for some $g \leqslant k^2 - k$. For $s = k^2 - 2k$, the inequality implies that $(k^2 - k) A$ contains an infinite arithmetic progression with difference g for some $g \leqslant k + 1$.

The only property of A that has been used in the proof thus far is $d_L(A) = w > 0$. Define A' by $(*)$ with $t = k^2 - k$. Then Lemma 1 shows that $d_L(A) = d_L(A')$. Apply the results above to A' instead of to A. Lemma 2 implies that sums of $k+1$ (resp. $k^2 - k$) elements of A' with repetitions allowed can be replaced by sums of $k+1$ (resp. $k^2 - k$) distinct elements of A. This proves the theorem.

COROLLARY. *Let A be a set of nonnegative integers with $d_L(A) = w > 0$. Let $k = \langle 1/w \rangle$, and let m be the least common multiple of the integers*

1, 2, 3, ..., $k+1$. *Then* $m(k^2-k)A$ *contains all sufficiently large multiples of* m.

Proof. Theorem 1 implies that $(k^2-k)A$ contains an infinite arithmetic progression with difference g for some $g \leqslant k+1$, and so $(k^2-k)A$ contains an infinite arithmetic progression with difference m. Hence, $m(k^2-k)A$ contains all sufficiently large multiplies of m.

Remark. Theorem 1 is best possible in the sense that for every $k \geqslant 1$ there exist sets A such that $d_L(A)=1/k$, but the sumset kA does not contain an infinite arithmetic progression. For example, let $\{t_n\}$ be a strictly increasing sequence of positive integers such that t_{n+1}/t_n tends to infinity, and let A be the set of integers in the intervals $[t_{n-1}, (t_n/k)-\sqrt{t_n}]$. Then $d_L(A)=1/k$ and $d_U(A)=1$, and the sumset kA is adjoint from the interval $(t_n - k\sqrt{t_n}, t_n)$ for all large n. Since kA contains arbitrarily long gaps, it cannot contain an infinite arithmetic progression.

There also exist sequences A with asymptotic density exactly $1/k$ such that kA does not contain an infinite arithmetic progression. The following example uses the theory of continued fractions.

LEMMA 3. *There exists an irrational number* α *such that the set* $\{q_n\}$ *of denominators of the convergents of the continued fraction of* α *contains infinitely many terms of every infinite arithmetic progression.*

Proof. Let (u_1, v_1), (u_2, v_2), ... be an infinite sequence of ordered pairs of positive integers such that every ordered pair occurs infinitely often in the sequence. We shall construct α by defining the sequence of its partial quotients a_k inductively. Recall the following two properties of the denominators of the convergents of a continued fraction:

(i) $q_n = a_n q_{n-1} + q_{n-2}$ for $n=2, 3, \ldots,$

(ii) $(q_{n-1}, q_n) = 1$ for $n=2, 3, \ldots.$

Let $a_0=0$ and $a_1=1$. Suppose that the partial quotients $a_0, a_1, \ldots, a_{2k-1}$ have been defined. Then $q_0, \ldots, q_{2k-1}$ are determined by (i). Since $(q_{2k-2}, q_{2k-1})=1$ by property (ii), there exist positive integers a such that $(aq_{2k-1}+q_{2k-2}, v_k)=1$. (For example, let a be the product of the primes that divide v_k but not q_{2k-2}.) Let a_{2k} be a positive integer with this property. By (i), we have

$$q_{2k} = a_{2k} q_{2k-1} + q_{2k-2}.$$

Since $(q_{2k}, v_k)=1$, there exist positive integers a such that

$$aq_{2k} + q_{2k-1} \equiv u_k \pmod{v_k}.$$

Let a_{2k+1} be a positive integer with this property. By (i),

$$q_{2k+1} = a_{2k+1} q_{2k} + q_{2k-1} \equiv u_k \pmod{v_k}.$$

Let α be the real number defined by the sequence of partial quotients $a_0, a_1, a_2, \ldots$. Then for every pair of positive integers (u, v), the sequence $\{q_k\}$ contains infinitely many terms such that $q \equiv u \pmod{v}$.

THEOREM 2. *For every positive integer $k \geqslant 2$ there exists a set A with asymptotic density $d(A) = 1/k$ such that kA does not contain an infinite arithmetic progression.*

Proof. Let α be an irrational number satisfying the condition of Lemma 3. Define the set A by

$$A = \{a \mid 1/a^{1/2} < \{a\alpha\} < 1/k - 1/a^{1/2}\}. \tag{$**$}$$

Since the sequence $\{\alpha\}$, $\{2\alpha\}$, $\{3\alpha\}$, ... is uniformly distributed modulo 1, it follows that $d(A) = 1/k$. Moreover, if $q = a_1 + \cdots + a_k \in kA$, then $(**)$ implies that

$$\begin{aligned} \|q\alpha\| &> 1/a_1^{1/2} + \cdots + 1/a_k^{1/2} \\ &> 1/(\min(a_1, \ldots, a_k))^{1/k} \\ &\geqslant 1/((a_1 + \cdots + a_k)/k)^{1/2} \\ &= (k/q)^{1/2} \\ &> q^{1/2}. \end{aligned}$$

Suppose the set kA contains an infinite arithmetic progression. Then infinitely many terms q of this progression are elements of the set $\{q_k\}$ of denominators of the convergents in the continued fraction expansion of α. By the theory of continued fractions, every denominator q satisfies

$$\|q\alpha\| < 1/q,$$

but this contradicts the previous inequality.

Problem. If $d_L(A) = 1/k$, then $(k+1)A$ contains an infinite arithmetic progression with difference at most $k^2 - k$. We do not know if $(k+1)A$ must contain an infinite arithmetic progression with difference at most $O(k)$.

3. Powers of 2 and Square-free Numbers

In this section we solve the Erdös–Freud problems in the infinite case.

Theorem 3. *Let B be a set of nonnegative integers such that $d_L(B) \geqslant \frac{1}{3}$ and $3 \nmid b^*$ for some $b^* \in B$. Then infinitely many powers of 2 can be written as sums of either four or five distinct elements of B.*

Proof. Note that the even powers of 2 belong to the congruence class 1 (mod 3) and the odd powers of 2 belong to the congruence class 2 (mod 3).

Let $A = B \setminus \{b^*\}$. Define A' by $(*)$ with $t = 4$. Applying Kneser's theorem to the sumset $4A'$, we obtain an integer $g \leqslant 6$ such that $4A' \sim 4((A')^{(g)})$ and $d_L(4A') \geqslant 4/3 - 3/g$.

If $g = 1$, then every large integer belongs to $4A'$.

If $g = 2$, then $4A'$ contains all large even integers.

If $g = 4$, then $4A'$ contains all large multiples of 4.

If $g = 5$, then $d_L(A') \geqslant \frac{1}{3}$ implies that A' contains representatives of at least two congruence classes modulo 5, and so $4A'$ contains all large numbers.

If $g = 6$, then $d_L(4A') \geqslant \frac{5}{6}$, and so $4A'$ contains all sufficiently large elements of five congruence classes modulo 6. In particular, $4A'$ contains all sufficiently large integers in a nonzero congruence class modulo 3.

In these five cases, $4A'$ contains infinitely many powers of 2, each of which is, by Lemma 2, a sum of four distinct elements of $B \setminus \{b^*\}$.

Finally, let $g = 3$. If $4A'$ contains an integer not divisible by 3, then it contains all sufficiently large elements of a nonzero congruence class modulo 3, and we are done. If $4A'$ consists of all large multiples of 3, each of which is then a sum of four distinct elements of $B \setminus \{b^*\}$, then each sufficiently large integer in the nonzero congruence class b^* (mod 3) is a sum of five distinct elements of B. This concludes the proof.

Lemma 4. *Let $g \geqslant 2$ and $r \geqslant 0$. Let $d = (g, r)$. There are infinitely many square-free numbers q such that $q \equiv r \pmod{g}$ if and only if d is square-free.*

Proof. If p^2 divides d for some prime p, then p^2 divides every element of the congruence class $r \pmod{g}$. Now let d be square-free. Then $(g/d, r/d) = 1$, and there are infinitely many primes p such that $p \equiv r/d \pmod{g/d}$ and $(p, d) = 1$. Then pd is square-free and $pd \equiv r \pmod{g}$.

Theorem 4. *Let B be a set of nonnegative integers such that $d_L(B) \geqslant \frac{1}{4}$ and $4 \nmid b^*$ for some $b^* \in B$. Then there are infinitely many square-free num-*

bers that can be represented as sums of either five or six distinct elements of B.

Proof. Let $A = B \backslash \{b^*\}$. Define A' by $(*)$ with $t=5$. Applying Kneser's theorem to $5A'$, we obtain $g \geqslant 1$ such that $5((A')^{(g)}) \sim 5A'$ and $d_L(5A') \geqslant 5/4 - 4/g$.

The square-free numbers have asymptotic density $6/\pi^2$. If $g > 4$, then $d_L(5A') > 1 - 6/\pi^2$, and so $5A'$ contains a set of square-free numbers of positive density. Lemma 2 implies that each of these numbers is a sum of 5 distinct elements of B. Therefore, it suffices to consider only $g \leqslant 4$.

Let $g < 4$. Then g is square free. Since $5A'$ contains an arithmetic progression with difference g, Lemma 3 implies that $5A'$ contains infinitely many square-free numbers, each of which is a sum of 5 distinct elements of $B \backslash \{b^*\}$.

Let $g = 4$. Then $5A'$ contains an arithmetic progression of the form $r + 4i$, each element of which is a sum of 5 distinct elements of $B \backslash \{b^*\}$. If $4 \nmid r$, then $(r, 4)$ is square free and we are done. If $4 \mid r$, then each element of the arithmetic progression $b^* + 4i$ is a sum of 6 distinct elements of B. Since $4 \nmid b^*$, this progression contains infinitely many square-free integers. This completes the proof.

References

1. G. Freiman, On two additive problems, preprint.
2. M. Kneser, Abschätzung der asymptotischen Dichte von Summenmengen, *Math. Z.* **58** (1953), 459–484.
3. E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium