ACTA ARITHMETICA
LIV (1989)

Sumsets containing long arithmetic progressions and powers of 2

by

MELVYN B. NATHANSON* (New York) and ANDRÁS SÁRKÖZY (Budapest)

1. Introduction. Let A be a set of nonnegative integers. The cardinality of A is denoted by $|A|$. The counting function $A(n)$ of the set A is defined by $A(n) = |A \cap \{1, 2, \ldots, n\}|$. We denote by hA the set of all sums of h elements of A, with repetitions allowed. We denote the set of all sums of h distinct elements of A by $h^\wedge A$.

P. Erdös and R. Freud [2] conjectured that if A satisfies

(1) $$A \subseteq \{1, 2, \ldots, 3n\} \quad \text{and} \quad |A| \geqslant n+1$$

then there is a power of 2 that can be written as a sum of distinct elements of A. They also conjectured that if B satisfies

(2) $$B \subseteq \{1, 2, \ldots, 4n\} \quad \text{and} \quad |B| \geqslant n+1$$

then there is a square-free number that can be written as a sum of distinct elements of B. Recently, G. Freiman [3] solved both problems. His results, however, are not entirely satisfactory, since they require at least $c \cdot \log n$ distinct summands from the set A in order to represent the power of 2, and also at least $c \cdot \log n$ distinct summands from the set B in order to represent the square-free number, while one might like to bound the number of summands by an absolute constant independent of n.

In order to obtain such an upper bound, Erdös, Nathanson, and Sárközy [4] first studied the infinite analogue of these problems. By deriving and applying some consequences of Kneser's theorem on the asymptotic density of sumsets [7], they proved that if the lower asymptotic density of an infinite set A is at least 1/3 and if $3 \nmid a$ for some integer $a \in A$, then there are infinitely many powers of 2 that can be written as sums of at most five distinct elements of A. They also proved that if the lower asymptotic density of B is at least 1/4 and if $4 \nmid a$ for some $a \in B$, then there are infinitely many square-free integers that can be written as sums of at most six distinct elements of B.

* The research of M. B. Nathanson was supported in part by grant number 667337 from the PSC-CUNY Research Award Program of the City University of New York.

In this paper we return to the more difficult finite case. One of the problems here is that Kneser's theorem cannot be used. Instead, we apply results of Mann [8] and Dyson [1] to prove a theorem of independent interest on long arithmetic progressions contained in sums of finite sequences. Using this theorem and its consequences, we prove that if n is large and if A satisfies (1), then there is a power of 2 that can be written as a sum of at most 30961 distinct elements of A. Using the same method, we can also prove that if B satisfies (2), then there is a square-free number that can be written as a sum of at most 21 distinct elements of B. We do not include this result, because Filaseta [5] has recently found a direct and elementary argument that shows that 2 summands suffice for n sufficiently large. Nathanson [9] has extended this to the case of k-free numbers.

2. Long arithmetic progressions. In this section we prove that a bounded sum of a sufficiently dense finite set of integers will contain a long arithmetic progression with bounded difference. We use the following result of Dyson ([1], [6]), which generalizes a famous theorem of Mann [8].

DYSON'S THEOREM. *Let h and n be positive integers. Let B be a subset of $\{0, 1, \ldots, n\}$ with $0 \in B$. Let $C = hB$. If δ is a positive real number such that $B(m) \geqslant \delta m$ for $m = 1, 2, \ldots, n$, then $C(m) \geqslant (\min(1, h\delta))m$ for $m = 1, 2, \ldots, n$.*

THEOREM 1. *Let N and k be positive integers. Let A be a subset of $\{1, 2, \ldots, N\}$ such that*

(3) $$|A| \geqslant N/k+1.$$

Then there exists an integer d with

(4) $$1 \leqslant d \leqslant k-1$$

such that if h and z are any positive integers satisfying the inequality

(5) $$N/h+zd \leqslant |A|$$

then the sumset $(2h)A$ contains an arithmetic progression with z terms and difference d.

Proof. We denote the elements of the set A by $a_1, \ldots, a_s$, where $|A| = s$ and $a_1 < a_2 < \ldots < a_s$. Define the integer d by

$$d = \min\{a_{i+1}-a_i \mid i = 1, 2, \ldots, s-1\}.$$

It follows from (3) that

$$N \geqslant a_s = \sum_{i=1}^{s-1} (a_{i+1}-a_i)+a_1 \geqslant d(s-1)+1 \geqslant d(N/k)+1 > dN/k,$$

and so $1 \leqslant d \leqslant k-1$. Thus, d satisfies inequality (4). Moreover, there exists an integer $a^* \in A$ such that $a^*+d \in A$. We shall show that the integer d satisfies the conditions of Theorem 1. For $i = 1, 2, \ldots, d$, let $A_i = \{n \mid d(n-1)+i \in A\}$. Then

(6) $$A_i \subseteq \{1, 2, \ldots, [(N+d-i)/d]\}$$

and

(7) $$|A| = \sum_{i=1}^{d} |A_i|.$$

Choose r so that $|A_r| = \max\{|A_i| \mid i = 1, 2, \ldots, d\}$. It follows from (7) that

(8) $$|A_r| \geqslant (1/d) \sum_{i=1}^{d} |A_i| = |A|/d.$$

Let h and z be positive integers satisfying (5). Note that Theorem 1 is trivial if $z = 1$, and so we can assume that $z \geqslant 2$. We shall show that there exists an integer u such that

(9) $$1 \leqslant u < u+z-1 \leqslant (N+d-r)/d$$

and

(10) $$A_r(u+m)-A_r(u) \geqslant m/h \quad \text{for } m = 1, 2, \ldots, z-1.$$

Assume that u satisfies (9) and (10). Define the set B by

(11) $$B = \{0\} \cup \{b \mid 1 \leqslant b \leqslant z-1 \text{ and } b+u \in A_r\}.$$

Then (10) implies that $B(m) \geqslant m/h$ for $m = 1, \ldots, z-1$. Let $C = hB$. Dyson's Theorem implies that $C(m) = m$ for $m = 1, \ldots, z-1$, and so

(12) $$\{0, 1, \ldots, z-1\} \subseteq C = hB.$$

Next we show that

(13) $$b \in B \quad \text{implies} \quad db+du+r+a^* \in 2A.$$

Inequality (10) implies $u+1 \in A_r$, hence $du+r \in A$. Since $a^* \in A$, it follows that $du+r+a^* \in 2A$. Thus, (13) holds for $b = 0$.

Let $b \in B$ and $b > 0$. Then $b+u \in A_r$, hence $d(b+u-1)+r \in A$. Since also $a^*+d \in A$, it follows that

$$db+du+r+a^* = (d(b+u-1)+r)+(a^*+d) \in 2A.$$

This proves (13).

Define $q = h(du+r+a^*)$. Let $n \in \{1, \ldots, z-1\}$. By (12), $n = b_1+\ldots+b_h$, where $b_i \in B$ for $i = 1, \ldots, h$. It follows from (13) that $db_i+du+r+a^* \in 2A$, and so

$$dn+q = d(b_1+\ldots+b_h)+h(du+r+a^*) \in 2hA.$$

Thus, $2hA$ contains the arithmetic progression $dn+q$ for $n = 1, \ldots, z-1$.

It remains only to prove the existence of an integer u satisfying (9) and (10). Assume that u does not exist. We shall then construct a finite sequence of nonnegative integers $n_0 < n_1 < \ldots < n_t \leqslant (N+d-r)/d$ such that

$$A_r(n_j)-A_r(n_{j-1}) < (n_j-n_{j-1})/h \quad \text{for } j = 1, \ldots, t.$$

Let $n_0 = 0$. Assume that $n_0, \ldots, n_{j-1}$ have been determined. If $n_{j-1}+z-1 \leqslant (N+d-r)/d$, then the indirect assumption implies that there exists an integer $m \in \{1, \ldots, z-1\}$ such that

$$A_r(n_{j-1}+m) - A_r(n_{j-1}) < m/h.$$

Let $n_j = n_{j-1}+m$.

If $n_{j-1}+z-1 > (N+d-r)/d$, then we end the construction of the finite sequence, that is, we set $t = j-1$.

It now follows from (5), (6), and (8) that

$$\begin{aligned}|A_r| &= \sum_{j=1}^{t} \big(A_r(n_j) - A_r(n_{j-1})\big) + \big(A_r((N+d-r)/d) - A_r(n_t)\big)\\ &< \sum_{j=1}^{t} (n_j - n_{j-1})/h + \big((N+d-r)/d - n_t\big)\\ &< n_t/h + z - 1 \leqslant (N+d-r)/hd + z - 1\\ &< N/hd + z = (N/h+zd)/d \leqslant |A|/d \leqslant |A_r|.\end{aligned}$$

This contradiction proves the existence of an integer u satisfying (9) and (10), and completes the proof of Theorem 1.

COROLLARY 1. *Let N and k be positive integers. Let A be a subset of $\{1, \ldots, N\}$ satisfying* (3). *Then there exists an integer d satisfying* (4) *such that $4kA$ contains an arithmetic progression with difference d and length $[N/2kd] \geqslant [N/2(k-1)k]$.*

Proof. Apply Theorem 1 with $h = 2k$ and $z = [N/2kd]$.

To obtain a refinement of Theorem 1 in the case of distinct summands, we shall need the following three lemmas.

LEMMA 1. *Let t be a positive integer and let δ be a positive real number. There exists a number $N_0(\delta, t)$ such that if $N > N_0$ and $A \subseteq \{1, 2, \ldots, N\}$, and if we define $A'_t \subseteq \{1, 2, \ldots, N\}$ by*

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

then $|A \setminus A'_t| < \delta N$.

Proof. If $|A \setminus A'_t| \geqslant \delta N$ and $N > N_0$, then Szemerédi's theorem [10] implies that $A \setminus A'_t$ contains an arithmetic progression of length $2t-1$, the middle term of which would belong to A'_t, which is absurd. Therefore, $|A \setminus A'_t| < \delta N$.

LEMMA 2. *Let A be a finite or infinite set of integers. Let $h \geqslant 1$. Define A'_h by* (14) *with $t = h$. Then $hA'_h \subseteq h^\wedge A$.*

Proof. This is Lemma 2 in [4].

LEMMA 3. *Let M, a, d, z, K be positive integers with $z > 1$ and $K > 1$. Let A be a subset of $\{0, 1, \ldots, M\}$ such that*

(15) $$0 \in A$$

and

(16) $$\{a, a+d, \ldots, a+(z-1)d\} \subseteq A.$$

Let $u = [KM/d(z-1)] + K$. Then there exist positive integers r, s such that

(17) $$s/r > K,$$

(18) $$\{rd, (r+1)d, \ldots, sd\} \subseteq (ud)A,$$

(19) $$s > KM.$$

Proof. Define $u' = [M/d(z-1)]+1$. If $u' \leqslant h < u$, then $a \in A \subseteq \{0, 1, \ldots, M\}$ implies that

$$h \geqslant u' = [M/d(z-1)]+1 > M/d(z-1) \geqslant a/d(z-1),$$

and so

$$hd(z-1) \geqslant a.$$

Therefore,

$$hda + hd(z-1)d \geqslant (h+1)da,$$

and so the following two arithmetic progressions overlap:

$$(hd)\{a, a+d, \ldots, a+(z-1)d\} = \{hda, \ldots, hda+hd(z-1)d\}$$

and

$$((h+1)d)\{a, a+d, \ldots, a+(z-1)d\} = \{(h+1)da, \ldots, (h+1)d(a+(z-1)d)\}.$$

It follows from (15) and (16) that

(20) $$\begin{aligned}(ud)A &\supseteq \bigcup_{h=u'}^{u} (hd)\{a, a+d, \ldots, a+(z-1)d\}\\ &= \{u'da, u'da+d, \ldots, uda+ud(z-1)d\}.\end{aligned}$$

Let us denote the first and last terms of this arithmetic progression by rd and sd, respectively. Then (20) implies (18), and (17) also holds, since

$$\begin{aligned}s/r &= \big(ua+ud(z-1)\big)/u'a > u/u'\\ &= \big([KM/d(z-1)]+K\big)/\big([M/d(z-1)]+1\big) \geqslant K.\end{aligned}$$

Finally, (19) follows from

$$s = ua+ud(z-1) > ud(z-1) = [KM/d(z-1)+K]d(z-1) > KM.$$

This completes the proof of the lemma.

THEOREM 2. *Let δ be a positive real number, and let k be a positive integer. If $N > N_0(\delta, k)$ and A is a subset of $\{1, 2, \ldots, N\}$ with*

(21) $$|A| \geqslant (1/k+\delta)N$$

then there exists an integer d satisfying $1 \leqslant d \leqslant k-1$ such that if h and z are positive integers satisfying the inequality

(22) $$N/h+zd \leqslant (1-\delta)|A|,$$

then there exists an arithmetic progression of length z and difference d, each of whose terms can be written as the sum of exactly 2h distinct elements of A.

Proof. Define A'_{2h} by (14) with $2h$ in place of t. By Lemma 1, for N sufficiently large we have

$$|A \backslash A'_{2h}| < (\delta/2k)N. \tag{23}$$

It follows from (21), (22), and (23) that conditions (3) and (5) in Theorem 1 are satisfied with A'_{2h} in place of A. Thus, $(2h)A'_{2h}$ contains an arithmetic progression of length z and difference d. By Lemma 2, this progression is contained in $(2h)^\wedge A$. This completes the proof of the Theorem.

COROLLARY 2. *Let $\delta > 0$, and let k be a positive integer. If $N > N_0(\delta, k)$ and A is a subset of $\{1, 2, \ldots, N\}$ with $|A| \geqslant (1/k+\delta)N$ then there exists an integer d satisfying $1 \leqslant d \leqslant k-1$ such that $4kA$ contains an arithmetic progression of length $[N/2kd] \geqslant [N/2k(k-1)]$ and difference d, each of whose terms can be written as the sum of exactly $4k$ distinct elements of A.*

Proof. Use $h = 2k$ and $z = [N/2kd]$ in inequality (22) of Theorem 2.

THEOREM 3. *Let N, k, K be positive integers with $K > 1$ and*

$$N > 64k^4K. \tag{24}$$

Let A be a subset of $\{1, 2, \ldots, N\}$ such that

$$|A| \geqslant N/k+1. \tag{25}$$

Then there exist positive integers d, r, and s such that $d \leqslant k-1$, $s/r > K$, $s > 4kKN$, and each term of the arithmetic progression $\{rd, (r+1)d, \ldots, sd\}$ can be written as the sum of at most $4dkK(8k^2+1)$ elements of A.

Proof. By (25), we can use Corollary 1 to obtain an integer d satisfying (4) such that the sumset $4kA$ contains an arithmetic progression $\{a, a+d, \ldots, a+(z-1)d\}$ of length $z = [N/2kd]$. Let $B = \{0\} \cup 4kA$. Then (15) and (16) hold with $4kN$ and B in place of M and A, respectively, and so we can apply Lemma 3 to obtain an arithmetic progression

$$\{rd, (r+1)d, \ldots, sd\} \subseteq udB = ud(\{0\} \cup 4kA) \tag{26}$$

where $s/r > K$ and $s > KM = 4kKN$. Since $1/(1-x) < 1+2x$ for $0 < x < 1/2$, it follows from (24) that

$$\begin{aligned} u &= [KM/d(z-1)]+K \\ &= [4kKN/d(z-1)]+K \leqslant 4kKN/d([N/2kd]-1)+K \\ &\leqslant 4kKN/d((N/2kd)-2)+K = 8k^2K/(1-(4kd/N))+K \\ &< 8k^2K(1+8kd/N)+K < (8k^2+1)K+64k^4K/N < (8k^2+1)K+1. \end{aligned} \tag{27}$$

Then (26) and (27) imply that each term of the progression in (26) can be written as the sum of at most $4dku \leqslant 4dkK(8k^2+1)$ elements of A. This completes the proof of Theorem 3.

THEOREM 4. *Let δ be a positive real number, and let k and K be integers greater than 1. If $N > N_0(\delta, k, K)$ and if A is a subset of $\{1, 2, \ldots, N\}$ with*

$$|A| \geqslant (1/k+\delta)N, \tag{28}$$

then there exist positive integers d, r, and s satisfying $d \leqslant k-1$, $s/r > K$, and $s > 4kKN$ such that each term of the arithmetic progression $\{rd, (r+1)d, \ldots, sd\}$ can be written as the sum of at most $4dkK(8k^2+1)$ distinct elements of A.

Proof. Let $t = 4dkK(8k^2+1)$, and define the set A'_t by (14). For N sufficiently large, Lemma 1 implies that

$$|A'_t| = |A| - |A \backslash A'_t| > (1/k+\delta)N - (\delta/2)N = (1/k+\delta/2)N,$$

so that (25) holds with A'_t in place of A. Applying Theorem 3 for large N, we obtain the existence of integers d, r, s such that $d < k$, $s/r > K$, $s > 4kKN$, and each term of the arithmetic progression $\{rd, (r+1)d, \ldots, sd\}$ can be written as a sum of at most t elements of A'_t. By Lemma 2, each term of this progression can be written as the sum of the same number of distinct elements of A. This completes the proof.

3. Powers of 2. We now apply the results in the preceding section to solve the Erdös–Freud problem on powers of 2.

THEOREM 5. *Let $n > 2^7 3^3 = 3\,456$. If $A \subseteq \{1, 2, \ldots, 3n\}$ and $|A| \geqslant n+1$, then there is a power of 2 that can be written as the sum of at most 3 504 elements of A.*

Proof. With $N = 3n$, $k = 3$, and $K = 2$, conditions (24) and (25) of Theorem 3 are satisfied, and so there exist positive integers d, r, and s such that $d = 1$ or 2, $s > 2r$, and each term of the arithmetic progression $\{rd, (r+1)d, \ldots, sd\}$ can be written as the sum of at most $4dkK(8k^2+1) \leqslant 3\,504$ elements of A. Since $s > 2r$, there exists an integer m with $r < 2^m \leqslant s$. Then $2^m d$ is a power of 2 that can be written as the sum of at most 3 504 elements of A.

THEOREM 6. *For n sufficiently large, if $A \subseteq \{1, 2, \ldots, 3n\}$ and $|A| \geqslant n+1$, then there is a power of 2 that can be written as the sum of at most 30 961 distinct elements of A.*

Proof. Since $|A| > n$, there exists $a^* \in A$ such that $3 \nmid a^*$. Then (28) holds with $A \backslash \{a^*\}$ in place of A, and with $N = 3n$, $k = 4$, and $K = 5$. By Theorem 4, there exist integers d, r, and s satisfying $d \leqslant 3$, $s > 5r$, and $s > 240n$ such that each term of the arithmetic progression $\{rd, (r+1)d, \ldots, sd\}$ can be written as the sum of at most $4dkK(8k^2+1) \leqslant 30\,960$ distinct elements of $A \backslash \{a^*\}$.

If $d = 1$ or 2, then there is an integer m with $r < 2^m < s$, and so $2^m d$ is a power of 2 that can be written as the sum of at most 30 960 distinct elements of A.

If $d = 3$, then each term of the arithmetic progression

$$\{rd, (r+1)d, \ldots, sd\} + \{a^*\} = \{3r+a^*, 3(r+1)+a^*, \ldots, 3s+a^*\}$$

is a sum of at most 30 961 distinct elements of A. The quotient of the greatest and least elements of this set is

$$(3s+a^*)/(3r+a^*) > (3s+a^*)/(3(s/5)+a^*) = (15s+5a^*)/(3s+5a^*)$$
$$= 4+(3s-15a^*)/(3s+5a^*) > 4+(720n-45n)/(3s+5a^*) > 4.$$

It follows that there exists an integer m such that

$$3r+a^* \leqslant 2^m < 2^{m+1} \leqslant 3s+a^*.$$

Since $3 \nmid a^*$, either 2^m or 2^{m+1} is congruent to a^* modulo 3, hence belongs to the arithmetic progression above, and so can be written as the sum of at most 30 961 distinct elements of A. This completes the proof.

References

[1] F. Dyson, *A theorem on the densities of sets of integers*, J. London Math. Soc. 20 (1945), 8–14.

[2] P. Erdős, *Some problems and results on combinatorial number theory*, in: *Proc. First China–U.S.A. Conference on Graph Theory and its Applications (Jinan, 1986)*, Annals New York Acad. Sci., to appear.

[3] P. Erdős and G. Freiman, *On two additive problems*, J. Number Theory, to appear.

[4] P. Erdős, M. B. Nathanson, and A. Sárközy, *Sumsets containing infinite arithmetic progressions*, J. Number Theory 28 (1988), 159–166.

[5] M. Filaseta, *Sets with elements summing to square-free numbers*, C. R. Math. Rep. Acad. Sci. Canada 9 (1987), 243–246.

[6] H. Halberstam and K. F. Roth, *Sequences*, Springer-Verlag, Berlin 1983.

[7] M. Kneser, *Abschätzungen der asymptotischen Dichte von Summenmengen*, Math. Z. 58 (1953), 459–484.

[8] H. B. Mann, *A proof of the fundamental theorem on the density of sums of sets of positive integers*, Ann. Math. 43 (1942), 523–527.

[9] M. B. Nathanson, *Sumsets containing k-free integers*, to appear.

[10] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199–245.

OFFICE OF THE PROVOST AND
VICE PRESIDENT FOR ACADEMIC AFFAIRS
LEHMAN COLLEGE (CUNY)
Bronx, New York 10468 USA

MATHEMATICAL INSTITUTE OF THE
HUNGARIAN ACADEMY OF SCIENCES
Budapest, Hungary

Received on 11.2.1988
and in revised form on 6.7.1988 (1788)

ACTA ARITHMETICA
LIV (1989)

Notions relatives de régulateurs et de hauteurs

par

A.-M. Bergé et J. Martinet (Talence)

1. Introduction. Soit L/K une extension de corps de nombres. L'étude faite dans [1] des minorations géométriques de régulateurs suggère la définition suivante du régulateur de L/K :

(1.1) Définition. Le *régulateur relatif* de L/K est: $R_{L/K} = Q_{L/K} R_L/R_K$, où $Q_{L/K}$ ("l'indice de Hasse" de L/K) est l'ordre du sous-groupe de torsion du quotient $E_L/\mu_L E_K$, les notations R_M, μ_M, E_M désignant respectivement le régulateur, le groupe des racines de l'unité et le groupe des unités d'un corps de nombres M.

Dans le cas d'une extension L/K primitive (c'est-à-dire sans sous-extension intermédiaire), on trouve dans [1] une démonstration d'une inégalité de la forme

$$R_{L/K} \geqslant \frac{1}{C_2}\left[\operatorname{Log}\frac{N_{K/\mathbf{Q}}(\mathfrak{d}_{L/K})}{C_3}\right]^{C_1} \quad (\mathfrak{d}_{L/K} \text{ est le discriminant relatif}),$$

où C_1, C_2, C_3 sont des constantes dépendant seulement des signatures de K et L; comme constante C_1, on peut prendre la différence $r_L - r_K$ des rangs des groupes d'unités de L et de K (on suppose implicitement que la norme du discriminant relatif est $> C_3$). Cette inégalité est une généralisation du résultat classique de Remak [9] sur les corps primitifs, résultat que l'on retrouve en faisant $K = \mathbf{Q}$ et qui est basé sur une minoration de la norme euclidienne, dans le réseau des unités de L, en fonction du discriminant.

Dans le cas d'une extension $L/\mathbf{Q}$ imprimitive, la recherche d'une bonne constante C_1 nécessite en outre un argument de géométrie diophantienne sur la minoration de la hauteur d'un nombre algébrique en fonction de son seul degré: la hauteur logarithmique est en effet une norme dans le réseau des unités de L. Rappelons à ce propos une définition des hauteurs (c'est bien celle que donne Lang dans [7], ch. 3, §1, même si les degrés locaux n'y figurent pas explicitement).

(1.2) Définition. Soit d un entier > 0 et soit $x = (x_0, \ldots, x_d)$ un point de l'espace projectif $\mathbf{P}^d(\bar{\mathbf{Q}})$. La *hauteur* de x est

$$H(x) = \Big(\prod_w \operatorname{Max}_i |x_i|_w\Big)^{1/[L:\mathbf{Q}]},$$