Astérisque

ANDRÁS SÁRKŐZY

On finite addition theorems

Astérisque, tome 258 (1999), p. 109-127

<http://www.numdam.org/item?id=AST_1999__258__109_0>

Astérisque **258**, 1999, p. 109–127

ON FINITE ADDITION THEOREMS

by

András Sárkőzy

Abstract. — If a finite set A of integers included in $\{1,\dots,N\}$ has more than N/k elements, one may expect that the set ℓA of sums of ℓ elements of A, contains, when ℓ is comparable to k, a rather long arithmetic progression (which can be required to be homogeneous or not). After presenting the state of the art, we show that some of the results cannot be improved as far as it would be thought possible in view of the known results in the infinite case. The paper ends with lower and upper bounds for the order, as asymptotic bases, of the subsequences of the primes which have a positive relative density.

1. Throughout this paper we use the following notations: $c_1, c_2 \dots$ denote positive absolute constants. If $f(n) = O(g(n))$, then we write $f(n) \ll g(n)$. The cardinality of the finite set S is denoted by $|S|$. The set of the integers, non-negative integers, resp. positive integers is denoted by $\mathbb{Z}$, $\mathbb{N}_0$ and $\mathbb{N}$. $\mathcal{A}$, $\mathcal{B}$ ... denote (finite or infinite) subsets of $\mathbb{N}_0$, and the counting functions of their positive parts are denoted by $A(n)$, $B(n)$, ..., so that, *e.g.*, $A(n) = |\mathcal{A} \cap \{1, 2, \dots n\}|$. The Schnirelmann density of the set $\mathcal{A} \subset \mathbb{N}_0$ is denoted by $\sigma(\mathcal{A})$, while the asymptotic density, asymptotic lower density, resp. asymptotic upper density of it is denoted by $d(\mathcal{A})$, $\underline{d}(\mathcal{A})$ and $\overline{d}(\mathcal{A})$ (see [16] for the definition of these density concepts). $\mathcal{A}_1 + \mathcal{A}_2 + \dots + \mathcal{A}_k$ denotes the set of the integers that can be represented in the form $a_1 + a_2 + \dots + a_k$ with $a_1 \in \mathcal{A}_1$, $a_2 \in \mathcal{A}_2$, ..., $a_k \in \mathcal{A}_k$; in particular, we write

$$\mathcal{A} + \mathcal{A} = 2\mathcal{A} = S(\mathcal{A}),$$

$$k\mathcal{A} = \mathcal{A} + (k-1)\mathcal{A} \quad \text{for} \quad k = 3, 4, \dots,$$

and

$$0\mathcal{A} = \{0\}, \quad 1\mathcal{A} = \mathcal{A}.$$

1991 ***Mathematics Subject Classification.*** — 11B13, 11B25, 11B05.

Key words and phrases. — additive number theory, density, additive bases, structure theory of set addition.

Research partially supported by Hungarian National Foundation for Scientific Research, Grant No. 1811.

If $\mathcal{A} \subset \mathbb{N}$ then $\mathcal{P}(\mathcal{A})$ denotes the set of the distinct positive integers n that can be represented in the form $n = \sum_{a \in \mathcal{A}} \varepsilon_a a$ where $\varepsilon_a = 0$ or 1 for all a and, if $\mathcal{A}$ is infinite, then all but finitely many of the ε's are equal to 0. (This notation will be used only in Section 3, while later the letter $\mathcal{P}$ will be reserved for denoting sets of primes.) An arithmetic progression is said to be *homogeneous* if it consists of the consecutive multiples of a non-zero number, i.e., it is of the form $kd, (k+1)d, \dots, \ell d$ (where $d \neq 0$).

2. The classical Schnirelmann-Mann-Kneser-Folkman theory of the set addition studies sums of *infinite* sets (the density and, in case of Kneser's theorem, the structure of the sum set). However, in many applications we are dealing with *finite* sets; in such a case, we cannot use this classical set addition theory or, in the best case, we have difficulties in applying it. Thus recently I have worked out a theory of addition of *finite* sets (partly jointly with Erdős, resp. Nathanson) which is more or less analogous to the case of infinite sets, and several conclusions and applications of this theory are close to the ones obtained by Freiman using a completely different approach. A considerable part of this work was inspired by a paper of Erdős and Freiman [5]. In this paper, first I will give a brief survey of my papers written on this subject. In the second half of the paper two further related problems will be studied.

3. Nathanson and I [20] proved that if we take "many" integers up to N, and we add the set obtained in this way sufficiently many times, then the sum set contains a long arithmetic progression:

Theorem 1. — *If $N \in \mathbb{N}$, $k \in \mathbb{N}$, $\mathcal{A} \subset \{1, 2, \dots, N\}$ and*

$$|\mathcal{A}| \geq \frac{N}{k} + 1, \tag{3.1}$$

then there exists an integer d with

$$1 \leq d \leq k - 1 \tag{3.2}$$

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

$$\frac{N}{h} + zd \leq |\mathcal{A}|,$$

then the sum set $(2h)\mathcal{A}$ contains an arithmetic progression with z terms and difference d.

Choosing here $h = 2k$ and $z = [N/2kd]$, we obtain

Corollary 1. — *If $N \in \mathbb{N}$, $k \in \mathbb{N}$, $\mathcal{A} \subset \{1, 2, \dots, N\}$ and $\mathcal{A}$ satisfies (3.1), then there exists an integer d satisfying (3.2) such that $4k\mathcal{A}$ contains an arithmetic progression with difference d and length $[N/2kd] \geq [N/2(k-1)k]$.*

The proof of Theorem 1 was based on Dyson's theorem [3] (which slightly generalizes Mann's theorem [19]). We used Theorem 1 to study a problem of Erdős and Freud on the solvability of the equation

$$a_1 + a_2 + \cdots + a_x = 2^y, \quad a_1, a_2, \dots, a_x \in \mathcal{A} \tag{3.3}$$

in "large" subsets $\mathcal{A}$ of $\{1, 2, \ldots, N\}$ (in sets $\mathcal{A}$ with $|\mathcal{A}| > [N/3]$). Indeed, we improved on a result of Erdős and Freiman [5]. Later Freiman [14] found another ingenious approach and he improved further on the result.

Corollary 1 was sufficient to study equation (3.3), however, it is not sharp in the sense that it guarantees an arithmetic progression of length only $\gg N/k^2$ in the sum set while one would expect a longer arithmetic progression and, indeed, later I needed a sharper result of this type. In fact, I proved [21] that having the same assumptions as in Corollary 1, one can guarantee a much longer *homogeneous* arithmetic progression in a sum set $\ell\mathcal{A}$ with $\ell \ll k$ (in many applications, we need the existence of a *homogeneous* arithmetic progression in the sum set, and this fact causes certain difficulties):

Theorem 2. — *If $N \in \mathbb{N}$, $k \in \mathbb{N}$, $\mathcal{A} \subset \{1, 2, \ldots, N\}$ and (3.1) holds, then there are integers d, ℓ, m such that (3.2) holds, moreover we have*

$$1 \leq \ell < 118k \tag{3.4}$$

and

$$\{(m+1)d, (m+2)d, \ldots, (m+N)d\} \subset \ell\mathcal{A}. \tag{3.5}$$

It is easy to see that this theorem is the best possible apart from the constant factor 118 in (3.4). This result can be considered as the finite analog of Kneser's theorem [18] (see Lemma 2 below). The proof of Theorem 2 is complicated, it uses both Dyson's theorem and Kneser's theorem.

One might like to sharpen this result by showing that all the elements of the arithmetic progression in (3.5) can be represented as the sum of possibly few *distinct* elements of $\mathcal{A}$; see [20] and Alon [1] for results of this type. The case when the number of distinct summands is unlimited will be studied later (Theorem 4 below).

Before the famous $\alpha + \beta$ conjecture was proved by Mann [19], Khintchin [17] had settled that most important special case of the conjecture when sum sets of the form $k\mathcal{A}$ are considered; indeed, he proved that

$$\sigma(k\mathcal{A}) \geq \min(1, k\sigma(\mathcal{A})). \tag{3.6}$$

In [23] I proved the following finite analog of this result:

Theorem 3. — *If $N \in \mathbb{N}$, $k \in \mathbb{N}$, $\mathcal{A} \subset \{1, 2, \ldots, N\}$ and $|\mathcal{A}| \geq 2$, then there are m, d such that $m \in \mathbb{Z}, d \in \mathbb{N}$,*

$$d < 2\frac{N}{|\mathcal{A}|} \tag{3.7}$$

and

$$|\{m+d, m+2d, \ldots, m+Nd\} \cap k\mathcal{A}| \geq \left(\min(1, \frac{1}{800}k\frac{|\mathcal{A}|}{N})\right) N. \tag{3.8}$$

The proof is similar to the proof of Theorem 2, although also further ideas are needed. Again, it is easy to see that this theorem is the best possible apart from the constants 2 in (3.7) and, mostly, $\frac{1}{800}$ in (3.8) (we will return to this question in

section 4). Note that an easy consideration shows that here we have to give up the requirement that the arithmetic progression in (3.8) should be homogeneous.

An infinite set $\mathcal{A} \subset \mathbb{N}$ is said to be *subcomplete* if it contains an infinite arithmetic progression. Improving on a result of Erdős [4], Folkman [11] proved the following remarkable theorem: if $\mathcal{A} \subset \mathbb{N}$ is an infinite set such that there are $\varepsilon > 0$ and N_0 with

$$A(N) > N^{1/2+\varepsilon} \text{ for } N > N_0,$$

then $\mathcal{P}(\mathcal{A})$ is subcomplete. Improving on a result of Alon and Freiman [2], I proved [22] the following finite analogue of Folkman's theorem:

Theorem 4. — *If $N \in \mathbb{N}$, $N > 2500$, $\mathcal{A} \subset \{1, 2, \ldots, N\}$ and*

$$|\mathcal{A}| > 200(N \log N)^{1/2}, \tag{3.9}$$

then there are integers d, y, z such that

$$1 \le d < 10^4 \frac{N}{|\mathcal{A}|},$$

$$z > 7^{-1}10^{-4}|\mathcal{A}|^2,$$

$$y < 7 \cdot 10^4 N z |\mathcal{A}|^{-2} \tag{3.10}$$

and

$$\{yd, (y+1)d, \ldots, zd\} \subset \mathcal{P}(\mathcal{A}).$$

Previously Alon and Freiman had proved a similar result with $N^{2/3+\varepsilon}$ on the right hand side of (3.9) and a slightly weaker inequality in place of (3.10). Moreover, independently and nearly simultaneously Freiman [13] proved a result essentially equivalent to Theorem 4 above. I derived Theorem 4 from Theorem 2; this part of the proof is easier, than the proof of Theorem 2. Freiman's proof is also complicated; he combines methods from the geometry of numbers and exponential sums in the manner of his book [12].

Again, Theorem 4 is the best possible apart from the constant factors and, perhaps, the factor $(\log N)^{1/2}$ on the right hand side of (3.9). Probably this logarithmic factor (or, at least, some of it) is unnecessary, although it is quite interesting and unexpected that exactly the same factor appears also in Freiman's result (obtained by a completely different method).

Theorem 4 has many applications. Alon and Freiman [2] found the first applications of a result of this type. Several further applications are discussed in my paper [22]. Papers [6], [7], [8] and [10] contain further applications.

Erdős and I [9] studied the following problem: what happens, if we replace assumption (3.9) by a slightly weaker one so that $|\mathcal{A}|$ drops below $N^{1/2}$? It turns out that there is a sharp drop in the length of the maximal arithmetic progression that we can guarantee in $\mathcal{P}(\mathcal{A})$, however, still it must contain quite a long one. Indeed, let $u = F(N, t)$ denote the greatest integer u such that for every $\mathcal{A} \subset \{1, 2, \ldots, N\}$ with $|\mathcal{A}| = t$, the set $\mathcal{P}(\mathcal{A})$ contains u consecutive multiples of a positive integer d:

$$\{(x+1)d, (x+2)d, \ldots, (x+u)d\} \subset \mathcal{P}(\mathcal{A})$$