THE STRUCTURE OF HIGHER SUMSETS

VSEVOLOD F. LEV

Abstract. Merging together a result of Nathanson from the early 70s and a recent result of Granville and Walker, we show that for any finite set A of integers with $\min(A) = 0$ and $\gcd(A) = 1$ there exist two sets, the "head" and the "tail", such that if $m \geq \max(A) - |A| + 2$, then the m-fold sumset mA consists of the union of the head, the appropriately shifted tail, and a long block of consecutive integers separating them. We give sharp estimates for the length of the block, and find all those sets A for which the bound $\max(A) - |A| + 2$ cannot be substantially improved.

1. Background, motivation, and summary of results.

Let A be a finite set of integers with $\min(A) = 0$ and $\gcd(A) = 1$. It is a basic fact dating back to Frobenius and Sylvester that the additive semigroup generated by A contains all positive integers from some point on. The largest integer that does not belong to this semigroup is called the *Frobenius number* of A. It is well-known that the Frobenius number of the three-element set $A = \{0, a, b\}$ is $(a-1)(b-1)-1$, but no simple explicit formula seems to exist for sets with four or more elements. The problem of finding the Frobenius number is known as the *linear diophantine problem of Frobenius*; see [RA05] for a comprehensive account.

The semigroup $\mathcal{S}(A)$ generated by A can be written as the infinite union $\mathcal{S}(A) = \{0\} \cup A \cup 2A \cup 3A \cup \ldots$, where $mA := \{a_1 + \cdots + a_m \colon a_1, \ldots, a_m \in A\}$ are the *sumsets* of A; that is, mA is the set of all possible sums of m elements of A, not necessarily distinct. Notice, that as a result of $0 \in A$, the sumsets satisfy $\{0\} \subseteq A \subseteq 2A \subseteq 3A \subseteq \ldots$

In the light of the Frobenius-Sylvester observation, it is interesting to investigate the structure of the individual sumsets mA. In this direction, Nathanson has obtained the following nice result.

Theorem 1 (Nathanson [N72]). *Suppose that A is a finite set of $n := |A| \geq 3$ integers with $\min(A) = 0$ and $\gcd(A) = 1$, and let $l := \max(A)$. Then there exist nonnegative integers h and t and finite integer sets $H \subseteq [0, h-2]$ and $T \subseteq [0, t-2]$ depending only on A such that for any integer $m \geq l^2(n-1)$ we have*

$$mA = H \cup [h, ml - t] \cup (ml - T). \tag{1}$$

2020 *Mathematics Subject Classification.* Primary 11B13; Secondary 11D07, 11A99.
Key words and phrases. Sumsets, Frobenius problem.

Loosely speaking, if m is sufficiently large, then the sumset mA consists of a "head", an appropriately shifted "tail" (both head and tail depending only on A but not on m), and an interval separating them. The fact that $h \le ml - t$, meaning that the interval is nonempty, is not stated explicitly in [N72], but is implicit in the proof.

Both the Frobenius-Sylvester observation and the theorem of Nathanson manifest the same phenomenon, which is that high-multiplicity sumsets of properly normalized sets contain long blocks of consecutive integers.

Let $\mathcal{E}(A)$ be the complement of the semigroup $\mathcal{S}(A)$ in the set of all positive integers. Following [GW21], we call the set $\mathcal{E}(A)$ the *exceptional set* of A. (This set is also called the *set of gaps* of the semigroup $\mathcal{S}(A)$.) The exceptional set is, therefore, the set of all positive integers not representable as a nonnegative linear combination of the elements of A, and its largest element $\max(\mathcal{E}(A))$ is the Frobenius number of A.

For brevity, we will write $\mathcal{S}$ and $\mathcal{E}$ instead of $\mathcal{S}(A)$ and $\mathcal{E}(A)$.

Keeping the notation of Theorem 1, as an immediate consequence of the theorem, the semigroup $\mathcal{S}$ contains all integers larger than h, while an integer $z \in [0, h]$ belongs to $\mathcal{S}$ if and only if it belongs to H; that is, $H = [0, h] \setminus \mathcal{E}$. Furthermore, from (1) we get $m(l - A) = ml - mA = T \cup [t, ml - h] \cup (ml - H)$. Therefore, denoting by $\mathcal{E}'$ the exceptional set of $l - A$, and repeating the argument above with A replaced by $l - A$, we obtain $T = [0, t] \setminus \mathcal{E}'$, and hence $ml - T = [ml - t, ml] \setminus (ml - \mathcal{E}')$. Thus, Theorem 1 says that if $m \ge l^2(n - 1)$, then $mA = [0, ml] \setminus (\mathcal{E} \cup (ml - \mathcal{E}'))$.

The bound $m \ge l^2(n - 1)$ has been improved in a number of subsequent papers, such as [WCC11] or [GS20]. Recently, Granville and Walker have shown that $m \ge l - n + 2$ suffices.

Theorem 2 (Granville-Walker [GW21, Theorem 1])**.** *Suppose that A is a set of $n := |A| \ge 3$ integers with $\min(A) = 0$ and $\gcd(A) = 1$, and let $l := \max(A)$. Then for any integer $m \ge M := l - n + 2$ we have*

$$mA = [0, ml] \setminus \big(\mathcal{E} \cup (ml - \mathcal{E}')\big), \tag{2}$$

where $\mathcal{E}$ and $\mathcal{E}'$ are the exceptional sets of A and $l - A$, respectively.

As shown in [GS20] and [GW21], the bound $m \ge l - n + 2$ is best possible. However, a natural question remains unanswered: can the sets $\mathcal{E}$ and $ml - \mathcal{E}'$ overlap? Does the assumption $m \ge l - n + 2$ guarantee the existence of an interval separating these sets, and if so, what is the length of the interval?

Our first goal here is to answer this question and give the Granville-Walker result an alternative, surprisingly short proof. We show not only that (2) holds, but also that $\mathcal{E}$ and $ml - \mathcal{E}'$ are disjoint and, moreover, there is a block of at least $(m - M + 1)l + 1$ consecutive integers contained in mA and separating $\mathcal{E}$ from $ml - \mathcal{E}'$.

Theorem 3. *Suppose that A is a set of $n := |A| \geq 3$ integers with $\min(A) = 0$ and $\gcd(A) = 1$, and let $l := \max(A)$. Then for any integer $m \geq M := l - n + 2$ we have*

$$mA = [0, ml] \setminus \big(\mathcal{E} \cup (ml - \mathcal{E}')\big),$$

where $\mathcal{E}$ and $\mathcal{E}'$ are the exceptional sets of A and $l - A$, respectively. Moreover, the interval $[\max(\mathcal{E})+1, \min(ml-\mathcal{E}')-1]$ is nonempty and contained in mA; indeed, writing $l - 1 = k(n-2) + r$ with integers $k \geq 1$ and $r \in [0, n-3]$, we have

$$\min(ml - \mathcal{E}') - \max(\mathcal{E}) \geq (m - M + 1)l + \Delta,$$

where

$$\Delta = l(k-1)(n-3) + rk(n-3) + r^2 + k + 1 \geq 2.$$

We adopt the convention that if $\mathcal{E} = \varnothing$, then $\max(\mathcal{E}) = -1$, and similarly for $\max(\mathcal{E}')$.

Notice that the quantity Δ is "normally" much larger than 2; for instance, it is easy to show that $\Delta > \left(1 - \frac{1}{k} - \frac{1}{n-2}\right) l^2$.

The lower bound $(m-M+1)l+\Delta$ established in the second part of the theorem is best possible; equality is obtained, for instance, for the sets of the form $A = \{0, d, 2d, \ldots, l\} \cup \{l-1\}$, where d is a nontrivial divisor of l. For these sets we have $n = \frac{l}{d} + 2$ whence $l = (d-1)(n-2) + (n-3) + 1$, so that $k = d - 1$ and $r = n - 3$. Furthermore, the semigroup $\mathcal{S}$ generated by A is identical to that generated by the three-element set $\{0, d, l-1\}$, from which, recalling the formula at the very beginning of this section, we derive that $\max(\mathcal{E}) = (d-1)(l-2) - 1 = k(l-2) - 1$. Also, $\mathcal{E}' = \varnothing$ since $1 \in (l - A)$. Consequently, $\max(\mathcal{E}) + \max(\mathcal{E}') = k(l-2) - 2$ and therefore

$$\min(ml - \mathcal{E}') - \max(\mathcal{E}) = ml - k(l-2) + 2 = (m - M + 1)l + \Delta,$$

where the second equality can be verified by cancelling out the common summand ml and expressing the rest in terms of the parameters n and d.

More generally, one can consider sets of the form

$$A := \{0, d, 2d, \ldots, sd\} \cup \{sd - 1 - td, sd - 1 - (t-1)d, \ldots, sd - 1\},$$

where s, d and t are positive integers satisfying $t < s$ and $2 \leq d < \frac{s}{t} + 1$. We have $l = sd$, $n = s+t+2$, and the exceptional set of A is the same as that of the set $\{0, d, sd-1-td\}$. It follows that $\max(\mathcal{E}) = (d-1)((s-t)d - 2) - 1$, and consequently,

$$\min(ml - \mathcal{E}') - \max(\mathcal{E}) = ml - (d-1)((s-t)d - 2) + 2 = (m - M + 1)l + \Delta.$$

(The last equality can be verified by observing that $k = d - 1$, as it follows from the assumption $d < \frac{s}{t} + 1$.)

We prove Theorem 3 in Section 3. The proof is a further elaboration on the ideas from [GS20, GW21].

Our second goal is to investigate the corresponding stability problem and determine those sets A such that for (2) to hold, one needs m to be almost as large as $l-n+2$. Some results in this direction are obtained in [GW21] where the sets requiring $m \ge l-n$ are fully described (under the technical assumptions $l \ge 9$ and $n \ge 5$). Citing from [GW21],

> *Indeed, our proofs are sufficiently flexible that one can go on and prove that $mA = [0, ml] \setminus \big(\mathcal{E} \cup (ml - \mathcal{E}')\big)$ holds for all $m \ge \max\{1, l-n-C\}$ for even larger values of C, except in some explicit finite set of families of sets A, though the number of cases seems to grow prohibitively with C.*

We prove the following result showing that the "tough" sets are, essentially, dense subsets of the set $\{0,1\} \cup [m+2, l]$ or of its mirror reflection $[0, l-(m+2)] \cup \{l-1, l\}$.

Theorem 4. *Let A be a set of $n := |A| \ge 6$ integers with $\min(A) = 0$ and $\gcd(A) = 1$. Write $l := \max(A)$, and let $\mathcal{E}$ and $\mathcal{E}'$ be the exceptional sets of A and $l - A$, respectively. Then for any integer m with*

$$m \ge \max\left\{l - \frac{3}{2}\,n + \frac{9}{2},\ \frac{2}{3}\,(l-n+2)\right\} \tag{3}$$

we have

$$mA = [0, ml] \setminus \big(\mathcal{E} \cup (ml - \mathcal{E}')\big), \tag{4}$$

except if either $\{0,1\} \subseteq A \subseteq \{0,1\} \cup [m+2, l]$, or $\{l-1, l\} \subseteq A \subseteq [0, l-(m+2)] \cup \{l-1, l\}$, in which cases (4) does not hold.

Notice that each of the sets $\{0,1\} \cup [m+2, l]$ and $[0, l-(m+2)] \cup \{l-1, l\}$ has size $l-m+1 < \frac{3}{2}\,n$, the inequality following from (3); this shows that their n-element subsets are contained therein with density exceeding $2/3$.

We also remark that for $n \ge 6$, the maximum in the right-hand side of (3) is always smaller than $l-n+2$, the bound of Theorem 3.

The proof of Theorem 4 is presented in Section 4.

2. The toolbox

In this section we collect various results used in the proofs of Theorems 3 and 4.

We use the standard set addition notation: if B and C are subsets of an additively written group, then the sumset $B + C$ is defined to be the set $\{b + c \colon b \in B,\ c \in C\}$.

Theorem 5 (Olson [O84, Theorem 1])**.** *Suppose that A and B are finite, nonempty subsets of a group. If $0 \in A$, then either A is contained in the subgroup of all those group elements z satisfying $(A+B) + z = A + B$, or $|A+B| \ge |B| + \frac{1}{2}|A|$.*

Note that if A is not contained in a proper subgroup, then the first alternative is ruled out unless $A+B$ is the whole group. With this observation in mind, arguing inductively we obtain the following corollary.

Corollary 1. *Suppose that A is a finite subset of an abelian group G, and $m \geq 1$ is an integer. If A is not contained in a proper coset, then either $mA = G$, or $|mA| \geq \frac{m+1}{2}|A|$.*

Here is yet another immediate corollary of Olson's theorem.

Corollary 2. *If A and B are finite, nonempty subsets of an abelian group G with $|A+B| \leq |B|+1$, $A+B \neq G$, and $|A| \geq 3$, then A is contained in a coset of a proper subgroup of G.*

Theorem 6 (Scherk [S55])**.** *Suppose that B and C are finite subsets of an abelian group with $0 \in B \cap C$. If 0 has a unique representation in $B+C$ (which is $0 = 0+0$), then $|B+C| \geq |B|+|C|-1$.*

Iterating Theorem 6, we get

Corollary 3 (Alon [A87, Corollary 2.3])**.** *Suppose that $m \geq 1$ is an integer, and A is a finite subset of an abelian group. If $0 \notin mA$, then $|A \cup (2A) \cup \cdots \cup (mA))| \geq m|A|$.*

Next, we need a result often referred to as *Freiman's* $(3n-3)$*-theorem.*

Theorem 7 (Freiman [F62])**.** *Suppose that A is a set of $n := |A| \geq 3$ integers with $\min(A) = 0$ and $\gcd(A) = 1$, and let $l := \max(A)$. Then*

$$|2A| \geq \min\{l, 2n-3\} + n.$$

A sequence of elements of an abelian group is called *zero-sum-free* if it does not have a finite, nonempty subsequence with the zero sum of its terms. The following theorem describing the structure of long zero-sum-free sequences in finite cyclic groups was proved by Savchev and Chen and, simultaneously and independently, by Yuan.

Theorem 8 (Savchev-Chen [SC07, Theorem 8], Yuan [Y07, Theorem 3.1])**.** *Suppose that u and l are positive integers, and that $(a_1, \ldots, a_u)$ is a zero-sum-free sequence of elements of the cyclic group of order l. If $u > l/2$, then there exist positive integers $x_1, \ldots, x_u$ with $x_1 + \cdots + x_u < l$ and a group element a of order l such that $a_i = x_i a$ for any $i \in [1, n]$.*

Finally, we list two theorems by the present author.

Theorem 9 ([L99, Proposition 1])**.** *Suppose that $X = (x_1, \ldots, x_u)$ is a nonempty sequence of positive integers written in an increasing order: $1 \leq x_1 \leq \cdots \leq x_u$. If X has fewer than $2u$ distinct subsequence sums, then $x_2, \ldots, x_u$ are all divisible by x_1, and $x_{i+1} \leq x_1 + \cdots + x_i$ for each $i \in [1, u-1]$.*

Theorem 10 ([L97, Theorems 1 and 3 (ii)]). *Suppose that A is a set of $n := |A| \ge 3$ integers with $\min(A) = 0$ and $\gcd(A) = 1$. Let $l := \max(A)$ and write $l-1 = k(n-2)+r$ with $k \ge 1$ and $r \in [0, n-3]$ integers.*

(i) *If $m \ge 2k$, then $[kl - k(n-1-r), (m-k)l + k(n-1-r)] \subseteq mA$;*
(ii) *if $m \ge 3k$, then indeed mA contains a block of at least $(m-k)l + k(n-1-r) + 1$ consecutive integers.*

An immediate corollary of Theorem 10 is an estimate for the Frobenius number of a set of given size and "diameter", originally proved by Dixmier.

Corollary 4 (Dixmier [D90, Theorem 3]). *Suppose that A is a set of $n := |A| \ge 3$ integers with $\min(A) = 0$ and $\gcd(A) = 1$. Let $l := \max(A)$ and write $l - 1 = k(n-2) + r$ with integer k and $0 \le r < n-2$. Then $\max(\mathcal{E}) \le k(l - n + r + 1) - 1$, where $\mathcal{E}$ is the exceptional set of A.*

3. Proof of Theorem 3

Recall that we have defined integers $k \ge 1$ and $r \in [0, n-3]$ by $l - 1 = k(n-2) + r$.

Our first goal is to show that if $n \ge 4$, then $[0, kl] \setminus \mathcal{E} \subseteq mA$. (We remark that the assumption $n \ge 4$ is essential and cannot be dropped; say, if $A = \{0, 1, l\}$, then $M = k = l - 1$, $\mathcal{E} = \varnothing$, and $[0, kl] \nsubseteq MA$ as, for instance, $l(l-1) - 1 \notin MA$.)

Suppose, for a contradiction, that $g \in [0, kl] \setminus \mathcal{E}$ is an integer not representable as a sum of m or fewer elements of A. If g is a multiple of l, then we can write $g = vl$ with an integer v satisfying $v > m$ in view of $g \notin mA$ and $l \in A$. On the other hand, $v \le k$ since $g \le kl$. Hence,

$$k \ge v > m \ge M = l - (n-2) \ge (k(n-2)+1) - (n-2) = (k-1)(n-2) + 1 \ge k,$$

a contradiction. Thus, g is not a multiple of l.

Since $g \notin \mathcal{E}$, we can write

$$g = a_1 + \cdots + a_u + lv \tag{5}$$

where $u, v \ge 0$ and $a_1, \ldots, a_u \in A \setminus \{0, l\}$ are integers; indeed, since g is shown above not to be divisible by l, we have $u \ge 1$ and $v \le k - 1$. Suppose that the representation (5) has the smallest value of the parameter u possible, among all representations of g in this form. We show that in this case $u + v \le l - (v+1)(n-2)$; since the quantity in the right-hand side is at most $l - n + 2 = M$, this will imply $g \in MA \subseteq mA$, contradicting the choice of g.

We notice that, in view of $u+v > m \geq M = l-n+2$ (which follows from the assumption $g \notin mA$) and $v \leq k-1$,

$$u \geq l-n+3-v \geq (k(n-2)+r+1)-(n-2)+1-(k-1) \\ \geq (k-1)(n-3)+2 \geq k+1 \geq v+2. \quad (6)$$

An important observation originating from [GS20, GW21] is that the sequence $(a_1, \ldots, a_u)$ reduced modulo l is zero-sum-free; that is, does not have any nonempty subsequences with the sum of their elements divisible by l. Indeed, if we had, say, $a_1+\cdots+a_s = wl$ with $s \in [1,u]$ and w integers, this would lead to $g = a_{s+1}+\cdots+a_u+(v+w)l$, contradicting minimality of u.

As a result of $(a_1, \ldots, a_u)$ being zero-sum-free, the $u-v-1$ sums

$$\sigma_s := a_1+\cdots+a_s, \quad v+2 \leq s \leq u$$

are pairwise distinct modulo l, cf. (6). Following [GW21], we claim that, moreover, these sums are also distinct modulo l from all elements of the sumset $(v+1)A$. To see this, suppose that $\sigma_s = b_1+\cdots+b_t+wl$ with integers $v+2 \leq s \leq u$, $1 \leq t \leq v+1$, and w, and elements $b_1, \ldots, b_t \in A \setminus \{0, l\}$. Then $\sigma_s > 0$ implies $w \geq -(t-1) \geq -v$, showing that $v+w \geq 0$. Consequently, $g = b_1+\cdots+b_t+a_{s+1}+\cdots+a_u+(v+w)l$ is a representation of g contradicting minimality of u in view of $t \leq v+1 < v+2 \leq s$.

Let $\overline{A}$ denote the canonical image of A in the quotient group $\mathbb{Z}/l\mathbb{Z}$. As we have just shown, the sumset $(v+1)\overline{A}$ is disjoint from the set of all $u-v-1$ sums σ_s, $s \in [v+2, u]$, taken modulo l. It follows that

$$|(v+1)\overline{A}| \leq l-(u-v-1). \quad (7)$$

Since $u+v \geq m+1 \geq M+1 = l-n+3$, this gives

$$|(v+1)\overline{A}| \leq l+v+1-u \leq (l+v+1)-(l-n+3-v) = 2v+n-2 = |\overline{A}|+2v-1.$$

As a result, $v \geq 1$, and there exists an integer $v_0 \in [1, v]$ with $|(v_0+1)\overline{A}| \leq |v_0\overline{A}|+1$. By Corollary 2, the set $\overline{A}$ is contained in a coset of a proper subgroup. Therefore, the original set A is contained in an arithmetic progression with the endpoints 0 and l and the difference larger than 1, contradicting the assumption $\gcd(A) = 1$.

We have thus shown that if $n \geq 4$, then

$$[0, kl] \setminus \mathcal{E} \subseteq mA.$$

Switching the roles of A and $l-A$, we conclude that $[0, kl] \setminus \mathcal{E}' \subseteq m(l-A)$; equivalently,

$$[(m-k)l, ml] \setminus (ml-\mathcal{E}') \subseteq mA.$$

If $m \le 2k$, then the intervals $[0, kl]$ and $[(m-k)l, ml]$ jointly cover the whole interval $[0, ml]$, and we immediately obtain (2). If $m > 2k$, then by Theorem 10 (i) we have $[kl, (m-k)l] \subseteq mA$ which, again, leads to (2).

This proves the first assertion of the theorem in the case where $n \ge 4$. Since the case $n = 3$ has received a nice and simple dedicated treatment in [GS20, Theorem 4], we ignore it here and proceed to the second assertion (without assuming $n \ge 4$ any longer).

Write $e := \max(\mathcal{E})$ and $e' := \max(\mathcal{E}')$. By Theorem 10 (ii), the sumset $(3k)A$ contains a block, say B, of $|B| = 2kl + k(n-1-r) + 1$ consecutive integers. Since $|B| > l$, while the difference between any two consecutive elements of $\mathcal{E}$ is easily seen not to exceed l (and indeed, not to exceed the smallest nonzero element of A) and similarly for $\mathcal{E}'$, we have $B \subseteq [e+1, 3kl - e' - 1]$. Therefore $3kl - e' - e - 1 \ge |B| = 2kl + k(n-1-r) + 1$ whence $kl - e - e' - 1 \ge k(n-1-r) + 1$. It follows that

$$\begin{aligned}\min(ml - \mathcal{E}') - \max(\mathcal{E}) &= ml - e - e' \\ &\ge ml - kl + k(n-1-r) + 2 \\ &= (m - M + 1)l + \Delta\end{aligned}$$

(verifying the last equality is tedious, but straightforward, and we omit the computation).

4. Proof of Theorem 4

We keep writing $l - 1 = k(n-2) + r$ with integers $k \ge 1$ and $r \in [0, n-3]$. Notice that if $k \ge 2$, then from (3) and the assumption $n \ge 6$, we have

$$m \ge l - \frac{3}{2}n + \frac{9}{2} \ge k(n-2) - \frac{3}{2}(n-2) + \frac{5}{2} = \left(k - \frac{3}{2}\right)(n-2) + \frac{5}{2} \ge 4k - \frac{7}{2}$$

whence

$$m \ge 4k - 3 \ge k. \tag{8}$$

Clearly, this resulting estimate remains valid also if $k = 1$.

We split the proof into several parts.

4.1. **Sufficiency.** First, we show that $\{0,1\} \subseteq A \subseteq \{0,1\} \cup [m+2, l]$ implies $mA \ne [0, ml] \setminus (\mathcal{E} \cup (ml - \mathcal{E}'))$; by symmetry, the same conclusion follows also from $\{l-1, l\} \subseteq A \subseteq [0, l-(m+2)] \cup \{l-1, l\}$. We actually show that $m+1 \notin \mathcal{E}$ and $m+1 \notin (ml - \mathcal{E}')$; this proves the assertion since, clearly, $m+1 \notin mA$ while $m+1 \in [0, ml]$.

The relation $m+1 \notin \mathcal{E}$ is, indeed, trivial since $\mathcal{E} = \varnothing$ in view of $1 \in A$. To prove that $m+1 \notin (ml - \mathcal{E}')$ we rewrite this as $ml - m - 1 \notin \mathcal{E}'$ and notice that by Corollary 4, we have $\max(\mathcal{E}') \le k(l-n+r+1) - 1$. Thus, it suffices to show that $ml - m - 1 > k(l-n+r+1) - 1$; that is, $m(l-1) > k(l-n+r+1)$. However, the last inequality is clearly true in view of (8) and since $r \le n-3$.

The rest of the argument deals with necessity; we want to show that, under the stated assumptions, we have $mA = [0, ml] \setminus \big(\mathcal{E} \cup (ml - \mathcal{E}')\big)$, unless either $\{0, 1\} \subseteq A \subseteq \{0, 1\} \cup [m + 2, l]$, or $\{l - 1, l\} \subseteq A \subseteq [0, l - (m + 2)] \cup \{l - 1, l\}$.

4.2. **The general setup.** The case where $l \leq n$ is easy to analyze, and we assume below that $l > n$.

We show that if m satisfies (3), then either $1 \in A \subseteq \{0, 1\} \cup [m + 2, l]$, or

$$[0, kl] \setminus \mathcal{E} \subseteq mA.$$

Once this is established, the proof can be easily completed. Namely, applying the assertion to the set $l - A$, we conclude that either $(l - 1) \in A \subseteq [0, l - (m + 2)] \cup \{l - 1, l\}$, or

$$[(m - k)l, ml] \setminus (ml - \mathcal{E}') \subseteq mA.$$

Taking into account that, by Theorem 10 (i), if $m \geq 2k$, then $[kl, (m - k)l] \subseteq mA$, we conclude that

$$[0, ml] \setminus \big(\mathcal{E} \cup (ml - \mathcal{E}')\big) \subseteq mA,$$

and the converse inclusion is trivial.

We thus assume that $g \in [0, kl] \setminus \mathcal{E}$ is an integer with $g \notin mA$, and show that $A \cap [2, m + 1] = \varnothing$ and that $1 \in A$.

4.3. **Preliminaries.** From (8) we have $kA \subseteq mA$. Therefore g is not a multiple of l, as if we had $l \mid g$, then $g/l \leq k$ and $l \in A$ would lead to $g = (g/l)l \in kA \subseteq mA$.

Since $g \notin \mathcal{E}$, we can write

$$g = a_1 + \cdots + a_u + vl \tag{9}$$

where $u, v \geq 0$ and $a_1, \ldots, a_u \in A \setminus \{0, l\}$ are integers. Speaking about *representations* of g we always mean representations of this particular form. In view of $g \notin mA$, for any representation we have

$$u + v \geq m + 1. \tag{10}$$

Since g is not a multiple of l, we also have $u > 0$ and $v \leq k - 1$. Hence, recalling (8),

$$u + k - 1 \geq u + v \geq m + 1 \geq 4k - 2;$$

therefore, $u \geq 3k - 1 \geq v + 2k$ and

$$u + v \geq 4k - 2. \tag{11}$$

Without loss of generality, we assume that the representation (9) has the smallest possible value of the parameter u among all representations of g; this assumption will be referred to as *minimality* (of u).

4.4. **The zero-sum-free property.** Repeating literally the argument from the proof of Theorem 3, we conclude that the sequence $(a_1, \dots, a_u)$ is zero-sum-free modulo l, the sums

$$\sigma_s := a_1 + \cdots + a_s, \quad v + 2 \le s \le u,$$

taken modulo l are distinct from each other and from all elements of the sumset $(v+1)\overline{A}$, where $\overline{A} := A \pmod{l}$, and as a result, (7) holds true.

On the other hand, by Corollary 1, since $\overline{A}$ is not contained in a proper subgroup (as it follows from $\gcd(A) = 1$), we have $|(v+1)\overline{A}| \ge \frac{v+2}{2}|\overline{A}|$. Therefore,

$$l - u + v + 1 \ge \frac{v+2}{2}(n-1);$$

consequently, since $u + v \ge m + 1 \ge l - \frac{3}{2}n + \frac{11}{2}$ by (10) and (3),

$$\begin{aligned} l &\ge u - v - 1 + \frac{v+2}{2}(n-1) \\ &= (u+v) + \frac{v}{2}(n-5) + n - 2 \\ &\ge l - \frac{3}{2}n + \frac{11}{2} + \frac{v}{2}(n-5) + n - 2 \\ &= l + \frac{1}{2}(v-1)(n-5) + 1 \end{aligned}$$

which is wrong whenever $v \ge 1$ and $n \ge 6$. Thus, $v = 0$. As a result, by (10) and (3),

$$u \ge l - \frac{3}{2}n + \frac{11}{2} \text{ and } u \ge \frac{2}{3}(l - n + 2) + 1, \tag{12}$$

and by (11)

$$u \ge 4k - 2. \tag{13}$$

On the other hand, we have $u < l$ due to the basic fact that a zero-sum-free sequence in a cyclic group has length smaller than the size of the group.

We investigate separately three possible cases: $g < l$, $l < g < 2l$, and $g > 2l$.

4.5. **The case where $g < l$.** Consider the sets

$$D_j := (jA \setminus (j-1)A) \cap [0, l], \quad j \in [1, u],$$

where $0A = \{0\}$ is assumed; thus, $D_1 = A \setminus \{0\}$, $g \in D_u$, and $D_j \subseteq jA$. The last relation shows that every element $d \in D_j$ can be written as $d = a + b$ with $a \in A$ and $b \in (j-1)A$; moreover, if $j \ge 2$, then $b \notin (j-2)A$ as otherwise we would have $d = a + b \in (j-1)A$. Consequently, $b \in D_{j-1}$ and we conclude that $D_j \subseteq D_{j-1} + A$. This shows, in particular, that if $D_{j-1} = \varnothing$, then also $D_j = \varnothing$; therefore, from $g \in D_u$, all sets $D_1, \dots, D_u$ are nonempty. Since these sets are pairwise disjoint and contained in

the interval $[0, l]$, using (12) and the assumption $n < l$ (made at the very beginning of the proof) we obtain $u \geq 3$, and then, using (12) once again,

$$|D_2| + \cdots + |D_{u-1}| \leq l + 1 - (1 + |D_1| + |D_u|) \leq l - n < 2(u - 2).$$

As a result, there is an integer $j \in [2, u - 1]$ with $|D_j| = 1$.

We now observe that for any subset $I \subseteq [1, u]$ of size $h := |I|$ we have $\sum_{i \in I} a_i \in D_h$ as if the sum were lying in $hA \setminus D_h$, then it would be representable as a sum of fewer than h elements of A, leading to a representation of g with fewer than u summands.

It follows that all sums $\sum_{i \in I} a_i$ with $|I| = j$ are equal to the same number, which is the element of D_j. As a result, all elements $a_1, \ldots, a_u$ are equal to each other, and we denote by a their common value; thus, $g = ua \notin (u - 1)A$ by the minimality of u.

By the box principle, from $g \notin 2A$ it follows that $|A \cap [0, g]| \leq \lceil g/2 \rceil$ whence

$$n = |A| \leq |A \cap [0, g]| + (l - g) \leq l - \lfloor g/2 \rfloor .$$

Hence, $\lfloor g/2 \rfloor \leq l - n$ showing that $au = g \leq 2(l - n) + 1 < 3u$, as it follows from (12). Therefore, $a < 3$; that is, $a = 1$ or $a = 2$.

If $a = 1$, then $g = u \geq m + 1$ and $A \cap [2, u] = \varnothing$ in view of $1 = a \in A$ and by the minimality of u, completing the proof. If $a = 2$, then $g = 2u$ and $2 = a \in A$ along with the minimality of u is easily seen to imply $A \cap [3, u] = \varnothing$. Moreover, A is disjoint from the set of all even integers in the range $[u + 1, 2u]$. A simple counting now gives

$$|A| \leq 3 + \left\lfloor \frac{u}{2} \right\rfloor + (l - 2u) \leq l - \frac{3}{2}\, u + 3 < n$$

(following from (12)), a contradiction.

4.6. **The case where $l < g < 2l$.** In this case in view of $g < kl$ we have $k \geq 2$; consequently, $u \geq 4k - 2 \geq 6$ by (13).

Modifying slightly the argument employed in the case $g < l$, this time we define

$$D_j := (jA \setminus (j - 1)A) \cap [0, 2l], \quad j \in [1, u].$$

Since $k \geq 2$, we have $l \geq 2n - 3$; hence, $|2A| \geq 3n - 3$ by Theorem 7. Therefore, arguing as in the case $g < l$,

$$\begin{aligned} |D_3| + \cdots + |D_{u-1}| &\leq 2l + 1 - (1 + |D_1| + |D_2|) - |D_u| \\ &\leq 2l + 1 - |2A| - 1 \\ &\leq 2l - (3n - 3) \\ &< 2(u - 3), \end{aligned}$$

the last inequality following from (12). Consequently, there is an integer $j \in [3, u - 1]$ with $|D_j| = 1$. As above, this leads to $a_1 = \cdots = a_u$, and denoting this common value by a, we have $g = ua \notin (u - 1)A$.

Since $u \ge 6$, we have $g \notin 4A$. By the box principle, $|2A \cap [0,g]| \le \lceil g/2 \rceil$ implying $|2A| \le \lceil g/2 \rceil + (2l-g) = 2l - \lfloor g/2 \rfloor$ and then, by Theorem 7 and (12),

$$\begin{aligned} \lfloor g/2 \rfloor &\le 2l - 3n + 3, \\ g - 1 &\le 2(2l - 3n + 3), \\ g &\le 4\left(l - \frac{3}{2}n + \frac{7}{4}\right) < 4u. \end{aligned}$$

Since $g = ua$, we conclude that $a \le 3$. On the other hand, $a \ge 2$ in view of $al > ua = g > l$.

If $a = 2$, then $g = 2u$, $2 = a \in A$, and $l/2 < u < l$. Consequently, from the minimality of u we derive that $A \cap [3,u] = \varnothing$, and A does not contain any even integers in the range $[u+1, l]$. Therefore

$$|A| \le 3 + \left\lfloor \frac{l-1}{2} \right\rfloor - \left\lfloor \frac{u}{2} \right\rfloor \le 3 + \frac{l}{2} - \frac{u}{2} = \frac{1}{2}(l - u + 6) < n,$$

a contradiction.

Assume now that $a = 3$. If A contained an element in the interval $[4,u]$, then denoting this element by b, we could replace b summands in the representation $g = 3 + \cdots + 3$ with just three summands $b + b + b$, contradicting minimality of u; therefore $A \cap [4,u] = \varnothing$.

Let

$$B_i := \{b \in A \colon u + 1 \le b \le l \text{ and } b \equiv i \pmod 3\}, \ i \in [0,2].$$

By minimality of u we have $B_0 = \varnothing$; in particular, $l \not\equiv 0 \pmod 3$. Furthermore, $|A \cap [0,3]| \le 3$: otherwise considering $j \in \{1,2\} \subseteq A$ such that $l + j \equiv 0 \pmod 3$ we get a contradiction. Next, if one of B_1 and B_2 is empty, then by (12)

$$n = |A| \le 3 + \frac{1}{3}(l - u + 2) \le 3 + \frac{1}{3}\left(\frac{3}{2}n - \frac{7}{2}\right) < n,$$

a contradiction. Finally, if both B_1 and B_2 are nonempty, then the sumset $B_1 + B_2$ consists of multiples of 3, and hence its smallest element exceeds $g = 3u$. Since $\min(B_i) \le l - 3(|B_i| - 1)$, $i \in \{1,2\}$, with at least one of these two inequalities strict, and since $B_1, B_2 \ne \varnothing$ implies $A \cap \{1,2\} = \varnothing$, we have in this case

$$\begin{aligned} 3u &< (l - 3(|B_1| - 1)) + (l - 3(|B_2| - 1)) - 1 \\ &= 2l - 3(|B_1| + |B_2|) + 5 = 2l - 3(n-2) + 5 \end{aligned}$$

whence $3u \le 2(l-n) + 5 - (n-6) \le 2(l-n) + 5$, contradicting (12).

4.7. **The case where $g > 2l$.** Since $g < kl$, in this case we have $k \geq 3$.

As shown above, the sequence $(a_1, \ldots, a_u)$ reduced modulo l is zero-sum-free. From (12), and since $k \geq 3$, the length of this sequence is

$$u \geq l - \frac{3}{2}(n-2) + \frac{5}{2} \geq l - \frac{3}{2}\frac{l-1}{k} + \frac{5}{2} \geq \frac{l}{2} + 3. \tag{14}$$

Consequently, by Theorem 8, there are integers $a \in [1, l-1]$ and $x_1, \ldots, x_u \geq 1$ such that $\gcd(a, l) = 1$, $x_1 + \cdots + x_u < l$, and $a_i \equiv ax_i \pmod{l}$ for each $i \in [1, u]$. As an immediate corollary, $u < l$. Renumbering, we can assume that $x_1 \leq \cdots \leq x_u$.

For $j \in \{u-1, u\}$, let P_j denote the set of all nonempty subsequence sums of the sequence $(a_i)_{i \in [1,u] \setminus \{j\}}$. The elements of the set $a_j + P_j$ are distinct modulo l from the elements of A: if, say, $a_j + \sum_{i \in I} a_i = b + wl$ with a nonempty set $I \subseteq [1, u] \setminus \{j\}$, and with integers $b \in A \setminus \{0, l\}$ and w, then $b + wl > 0$ whence $w \geq 0$ and $g = b + \sum_{i \in [1,u] \setminus (I \cup \{j\})} a_i + wl$, contradicting the minimality of u. Recalling the notation $\overline{A} = A \pmod{l}$, and writing $\overline{P}_j := P_j \pmod{l}$, we conclude that

$$|\overline{P}_j| \leq l - |\overline{A}| = l - n + 1 \leq 2(u-1) - 2,$$

the last inequality being a consequence of (12). Since $x_1 + \cdots + x_u < l$, reducing the set P_j modulo l does not affect its size. Therefore, the subsequence sum set of the integer sequence $(x_i)_{i \in [1,u] \setminus \{j\}}$ has size $|P_j| + 1 = |\overline{P}_j| + 1 \leq 2(u-1) - 1$. Applying Theorem 9 twice, first time with $j = u - 1$, and the second time with $j = u$, we conclude that all integers $x_1, \ldots, x_u$ are divisible by x_1, and that $x_{i+1} \leq x_1 + \cdots + x_i$ for each $i \in [1, u-1]$.

Assuming for a moment that $x_1, \ldots, x_u$ are not all equal to each other, let $s \in [1, u-1]$ be the smallest integer such that $x_{s+1} > x_s$. Then the subsequence sum set of the truncated sequence $(x_1, \ldots, x_s)$ is the arithmetic progression $\{0, x_1, 2x_1, \ldots, sx_1\}$; therefore, there exists an index set $I \subseteq [1, s]$ with $|I| \geq 2$ such that $x_{s+1} = \sum_{i \in I} x_i$. It follows that

$$\sum_{i \in I} a_i = a_{s+1} + wl$$

where w is an integer and $w \geq 0$ in view of $wl + a_{s+1} > 0$. This, however, is easily seen to contradict the minimality of u.

We have thus shown that $x_1, \ldots, x_u$ are all equal to each other; hence, $a_1, \ldots, a_u$ are equal to each other, too, and we denote their common value by a. We notice that

$$3 \leq a \leq 2k - 1 : \tag{15}$$

the lower bound follows from $2l < g = ua < al$, and the upper bound from $kl > g = ua > \frac{1}{2}al$, cf. (14).

Comparing (15) and (13), we conclude that $u \geq 2a$. Furthermore, we have $A \cap [a+1, u] = \varnothing$, as if there existed an integer $b \in A \cap [a+1, u]$, then we could reduce the number of summands in the representation (9) by replacing the b-term sum $a + \cdots + a$ with the

a-term sum $b+\cdots+b$. We thus can partition A as $A=A_0\cup A_1$ where $A_0:=A\cap[0,a]$ and $A_1=A\cap[u+1,l]$.

Let $H:=\lceil g/l\rceil-1$, and suppose that $1\le h\le H$. If $za\in A_0+hA_1$ with an integer z, then $z\ge h+2$ in view of

$$(h+1)a\le\frac{1}{2}\,(h+1)u\le uh<\min(A_0+hA_1),$$

and similarly, $z\le u$ in view of

$$(u+1)a=a+g>a+Hl\ge a+hl=\max(A_0+hA_1).$$

Therefore, $h+2\le z\le u$, which is easily seen to contradict the minimality of u (replace z summands of the representation $g=a+\cdots+a$ with the $(h+1)$-summand representation of za). Consequently, the sumset A_0+hA_1 does not contain multiples of a. Since this holds for any $h=1,\dots,H$, letting $T:=A_1\cup\cdots\cup HA_1$, we conclude that, indeed, A_0+T does not contain multiples of a; that is, $0\notin\widetilde{A}_0+\widetilde{T}$ where we let $\widetilde{A}_i:=A_i \pmod a$ $(i=0,1)$ and $\widetilde{T}:=T \pmod a$.

By the box principle, from $0\notin\widetilde{A}_0+\widetilde{T}$ it follows that $|\widetilde{T}|\le a-|\widetilde{A}_0|=a-|A_0|+1$. The fact that $0\notin\widetilde{A}_0+\widetilde{T}$ also implies $0\notin\widetilde{T}$; consequently, $|\widetilde{T}|\ge H|\widetilde{A}_1|$ by Corollary 3. Comparing these estimates, we obtain

$$H|\widetilde{A}_1|\le a-|A_0|+1. \tag{16}$$

On the other hand, by (14),

$$H>\frac{g}{l}-1=\frac{u}{l}\,a-1>\frac{1}{2}\,a-1 \tag{17}$$

whence, using (15),

$$3H\ge\frac{3}{2}\,a-\frac{3}{2}>a-|A_0|+1. \tag{18}$$

From (16) and (18) we obtain $|\widetilde{A}_1|\le 2$.

If $|\widetilde{A}_1|=1$, then A_1 is contained in one single residue class modulo a. Hence, by (12) and in view of (15),

$$|A_1|\le\frac{l-u-1}{a}+1\le\frac{1}{a}\left(\frac{3}{2}\,n-\frac{13}{2}\right)+1\le\frac{1}{2}\,n-\frac{7}{6}.$$

On the other hand, from (16) and (14),

$$a-|A_0|+1\ge H>\frac{g}{l}-1=\frac{u}{l}\,a-1>\left(1-\frac{3}{2k}\right)a-1$$

whence, by (15),

$$|A_0|<\frac{3}{2k}\,a+2<5.$$

Combining the upper bounds, we get

$$n=|A_0|+|A_1|\le 4+\frac{1}{2}\,n-\frac{7}{6}$$

which is wrong for $n \geq 6$.

Thus $|\widetilde{A}_1| = 2$. Substituting back to (16) and using (17), we get

$$|A_0| \leq a + 1 - 2H < a + 1 - (a - 2) = 3.$$

Hence, $|A_0| = 2$; that is, $A_0 = \{0, a\}$. Moreover, A_1 is contained in a union of $|\widetilde{A}_1| = 2$ residue classes modulo a. Therefore, from $a \geq 3$,

$$n = |A_0| + |A_1| \leq 2 + 2\left(\frac{l - u - 1}{a} + 1\right) \leq 4 + \frac{2}{3}(l - u - 1);$$

that is, $u \leq l - \frac{3}{2}n + 5$, contradicting (12).

This completes the proof of Theorem 4.

Acknowledgement

The author is grateful to the anonymous referee for the very careful reading of the manuscript and a number of remarks and suggestions.

References

[A87] N. Alon, Subset Sums, *Journal of Number Theory* **27** (1987), 196–205.

[F62] G.A. Freiman, Inverse problems in additive number theory, VI. On the addition of finite sets, III, Izv. Vyssh. Uchebn. Zaved. Mat. 3 (1962), 151–157 (Russian).

[D90] J. Dixmier, Proof of a conjecture by Erdős and Graham concerning the problem of Frobenius, *Journal of Number Theory* **34** (1990), 198–209.

[GS20] A. Granville and G. Shakan, The Frobenius postage stamp problem, and beyond, *Acta Math. Hungar.* **161** (2) (2020), 700–718.

[GW21] A. Granville and A. Walker, A tight structure theorem for sumsets, *Proc. Amer. Math. Soc.* **149** (2021), no. 10, 4073–4082.

[L97] V.F. Lev, Optimal representations by sumsets and subset sums, *Journal of Number Theory* **62** (1997), 127–143.

[L99] ———, The structure of multisets with a small number of subset sums, Astèrisque **258** (1999), 179–186.

[N72] M. Nathanson, Sums of finite sets of integers, *The American Mathematical Monthly* **79** (9) (1972), 1010–1012.

[O84] J. Olson, On the sum of two sets in a group, *Journal of Number Theory* **18** (1984), 110–120.

[RA05] J.L. Ramírez Alfonsín, *The Diophantine Frobenius problem*, Oxford Lecture Series in Mathematics and its Applications **30**. Oxford University Press, Oxford, 2005. xvi+243 pp.

[SC07] S. Savchev and F. Chen, Long zero-free sequences in finite cyclic groups, *Discrete Math.* **307** (22) (2007), 2671–2679.

[S55] P. Scherk, *Amer. Math. Monthly* **62** (1955), 46–47.

[WCC11] J.D. Wu, F.J. Chen, and Y.G. Chen, On the structure of the sumsets, *Discrete Math.* **311** (6) (2011), 408–412.

[Y07] P. Yuan, On the index of minimal zero-sum sequences over finite cyclic groups, *Journal of Combinatorial Theory, Series A* **114** (2007), 1545–1551.

Email address: seva@math.haifa.ac.il

Department of Mathematics, The University of Haifa at Oranim, Tivon 36006, Israel