ON THE NUMBER OF POPULAR DIFFERENCES

SERGEI V. KONYAGIN AND VSEVOLOD F. LEV

Abstract. We prove that there exists an absolute constant $c > 0$ such that for any finite set $A \subseteq \mathbb{Z}$ with $|A| \ge 2$ and any positive integer $m < c|A|/\ln|A|$ there are at most m integers $b > 0$ satisfying $|(A+b) \setminus A| \le m$; equivalently, there are at most m positive integers possessing $|A| - m$ (or more) representations as a difference of two elements of A.

This is best possible in the sense that for each positive integer m there exists a finite set $A \subseteq \mathbb{Z}$ with $|A| > m \log_2(m/2)$ such that $|(A + b) \setminus A| \le m$ holds for $b = 1, \dots, m + 1$.

1. Introduction

For a finite subset A of an abelian group and a group element b let

$$\Delta_A(b) := |(A + b) \setminus A|.$$

Thus, $|A| - \Delta_A(b)$ is the number of representations of b as a difference of two elements of A. Also, if b is of infinite order, then $\Delta_A(b)$ is the smallest number of arithmetic progressions with difference b into which A can be partitioned.

The function Δ_A was used, for instance, by Olson in [O68] and (in a somewhat implicit form) by Erdős and Heilbronn in [EH64]. Its basic properties are as follows:

(i) $\Delta_A(0) = 0$;

(ii) $\Delta_A(-b) = \Delta_A(b)$ for any group element b;

(iii) $\Delta_A(b_1 + b_2) \le \Delta_A(b_1) + \Delta_A(b_2)$ for any group elements b_1 and b_2, and consequently $\Delta_A(hb) \le h\Delta_A(b)$ for any group element b and integer $h \ge 1$.

Property (i) is trivial, property (ii) is almost immediate from the definition, property (iii) is not difficult to prove and the reader can either regard this as an exercise, or check any of [EH64, O68].

What one normally seeks in connection with the function Δ_A is to show that any sufficiently large subset of the group contains an element b with $\Delta_A(b)$ relatively large; that is, to show that Δ_A does not assume "too many small values". We investigate this problem in the case where the underlying group is the group of integers, for which we use the standard notation $\mathbb{Z}$. By $\mathbb{N}$ we denote the set of all *positive* integers.

2000 *Mathematics Subject Classification.* Primary: 11B75; Secondary: 11B25, 11P70.

Key words and phrases. Popular differences, set addition, additive combinatorics.

If A is a block of consecutive integers, then for every $m \in [1, |A|-1]$ there is exactly one positive integer b with $\Delta_A(b) = m$; consequently, there are exactly m positive integers b with $\Delta_A(b) \le m$. A theorem of Gabriel [G32, Theorem 2] implies that this is the worst case "in average": if $A \subseteq \mathbb{Z}$ and $B \subseteq \mathbb{N}$ are finite, $\mathcal{A}$ is a block of consecutive integers with $|\mathcal{A}| = |A|$, and $\mathcal{B} = [1, |B|]$, then

$$\sum_{b \in B} \Delta_A(b) \ge \sum_{b \in \mathcal{B}} \Delta_{\mathcal{A}}(b). \tag{1}$$

On a historical note we mention that Gabriel's theorem extends an earlier result of Hardy and Littlewood [HL28]; see also [HLP88, Theorem 374] or [L98, Theorem C]. Alternatively, (1) can be derived from a theorem of Pollard [P74].

We observe that under the extra assumption $|B| \le |A|$, from (1) it is easy to deduce

$$\sum_{b \in B} \Delta_A(b) \ge \frac{1}{2}\,|B|(|B|+1);$$

consecutively, if $0 < |B| \le |A|$, then there exists $b \in B$ with

$$\Delta_A(b) \ge \frac{1}{2}\,(|B|+1). \tag{2}$$

In the absence of evident counter-examples one can expect that, in fact, the consecutive integers case is critical not only in average, but also "pointwise"; that is, for any finite sets $A \subseteq \mathbb{Z}$ and $B \subseteq \mathbb{N}$ with $|B| \le |A|$ there exists $b \in B$ such that $\Delta_A(b) \ge |B|$. In other words, for any $m \in [1, |A|]$ the function Δ_A assumes on $\mathbb{N}$ at most $m-1$ values, smaller than m. Clearly, the assumption $|B| \le |A|$ cannot be dropped here: if $|B| > |A|$, then $\Delta_A(b) \ge |B|$ does not hold as the values of Δ_A never exceed $|A|$. It turns out that this assumption is actually *too weak*: in the Appendix we prove that if m is a positive integer,

$$A = \bigcup_{0 \le k < \log_2 m} \big[km, (k+1)m - 2^k\big),$$

and $B = [1, m]$, then $\Delta_A(b) \le m-1$ holds for each $b \in B$, while it is easy to check that $|B| < (\ln 2 + o(1))|A|/\ln|A|$ as $m \to \infty$. The goal of this paper is to show that no such examples exist if $|B| < c|A|/\ln|A|$ with an appropriate absolute constant $c > 0$.

For finite subsets A and B of an abelian group we write

$$\mu_A(B) := \max\{\Delta_A(b) \colon b \in B\},$$

subject to the agreement that the maximum of the empty set is 0.

Theorem 1. *There is an absolute constant $c > 0$ such that $\mu_A(B) \ge |B|$ holds for all finite sets $A \subseteq \mathbb{Z}$, $B \subseteq \mathbb{N}$ with $|A| > 1$ and $|B| < c|A|/\ln|A|$.*

The constant c of Theorem 1 can be computed explicitly from our argument, but the value we can obtain is very small. For this reason, and also to exhibit a surprising connection with the famous Graham g.c.d. conjecture, we also prove the following asymptotically weaker result.

Theorem 2. *We have $\mu_A(B) \geq |B|$ for all finite sets $A \subseteq \mathbb{Z}$, $B \subseteq \mathbb{N}$ with $|B| \leq \sqrt{|A|}$.*

Theorem 2 is proved in Section 2. In Section 3 we reduce Theorem 1 to the special case where $B = [1, |B|]$; this case, which will be separately stated as the Main Lemma, is treated in Section 4.

2. Proof of Theorem 2

Suppose that $A \subseteq \mathbb{Z}$ is a finite set and that $B = \{b_1, \dots, b_m\} \subseteq \mathbb{N}$, where $1 \leq m \leq \sqrt{|A|}$; we want to show that there exists $j \in [1, m]$ with $\Delta_A(b_j) \geq m$. Assuming this is wrong, fix arbitrarily $i, j \in [1, m]$. By the assumption, A is a union of $\Delta_A(b_i) \leq m-1$ arithmetic progressions with difference b_i. At least one of these progressions has m or more terms in view of $|A| > (m-1)^2$; accordingly, suppose that $a + kb_i \in A$ for $k = 1, \dots, m$. Since, on the other hand, A is a union of $\Delta_A(b_j) \leq m-1$ arithmetic progressions with difference b_j, there are $1 \leq k_1 < k_2 \leq m$ satisfying $a + k_2 b_i \equiv a + k_1 b_i \pmod{b_j}$. Letting $k := k_2 - k_1$ we get $b_j \mid kb_i$ whence $b_j/(b_i, b_j)$ is a divisor of k and therefore $b_j/(b_i, b_j) < m$. This, however, contradicts a theorem of Balasubramanian and Soundararajan [BS96] which, confirming a conjecture of Graham [G70], says that $\max\{s'/(s', s'') \colon s', s'' \in S\} \geq |S|$ for any finite set $S \subseteq \mathbb{N}$.

This proves Theorem 2.

3. The Main Lemma

Given an integer $m \geq 1$, we say that the finite subset $A \subseteq \mathbb{Z}$ is *m-coverable* if $\mu_A([1, m]) < m$; that is, for any $d \in [1, m]$ the set A is a union of at most $m-1$ arithmetic progressions with difference d. Notice that by Theorem 2 we have

$$|A| < m^2 \tag{3}$$

for any m-coverable set A.

We derive Theorem 1 from the following assertion.

Main Lemma. *There is an absolute constant $C \geq 1$ such that if $m \geq 2$ is an integer and A is an m-coverable set, then $|A| < Cm \ln m$.*

Notice that the Main Lemma is essentially a particular case of Theorem 1, obtained for $B = [1, m]$.

Postponing the (quite involved) proof of the Main Lemma to Section 4, we show in the remainder of the present section how this lemma implies the assertion of Theorem 1.

For an integer $h \ge 1$ and a subset S of an abelian group by hS we denote the h-fold sumset of S:

$$hS := \{s_1 + \cdots + s_h \colon s_1, \ldots, s_h \in S\}.$$

For further references we record in terms of the quantity $\mu_A(B)$ two observations which have already appeared above.

Lemma 1. *Let A and B be finite subsets of an abelian group. If h is a positive integer, then*

$$\mu_A(hB) \le h\mu_A(B). \tag{4}$$

Furthermore, if the underlying group is the group of integers and $0 < |B| \le |A|$, then

$$\mu_A(B) \ge \frac{1}{2}\,(|B| + 1). \tag{5}$$

Proof. The first estimate follows from the property (iii) at the beginning of the Introduction, for the second estimate see (2). □

We need the following result on the rate of growth of the sumsets hB.

Theorem 3 ([L96, Corollary 1])**.** *Let S be a finite set of integers, not contained in an arithmetic progression with difference, larger than 1. Write $n := |S|$ and $l := \max S - \min S$ and suppose that κ is an integer, satisfying $\kappa(n-2)+1 \le l \le (\kappa+1)(n-2)+1$. Then for any integer $h \ge 1$ we have*

$$|hS| \ge \begin{cases} \frac{h(h+1)}{2}\,(n-2) + h + 1 & \text{if } h \le \kappa, \\ \frac{\kappa(\kappa+1)}{2}\,(n-2) + \kappa + 1 + (h-\kappa)l & \text{if } h \ge \kappa. \end{cases}$$

Corollary 1. *Let S be a finite set of integers, not contained in an arithmetic progression with difference, larger than 1. If $\max S - \min S \ge 3|S| - 5$, then*

$$|3S| \ge 6|S| - 8.$$

Proof. If $|S| = 2$, then S consists of two consecutive integers and the assertion is immediate. If $|S| \ge 3$, set $l := \max S - \min S$ and $\kappa := \lfloor (l-1)/(|S|-2) \rfloor$ and apply Theorem 3 with $h = 3$, observing that $\kappa \ge 3$. □

The following lemma shows that if S is a dense set of integers, then the difference set

$$S - S := \{s' - s'' \colon s', s'' \in S\}$$

contains long blocks of consecutive integers.

Lemma 2 ([L06, Lemma 3])**.** *Let S be a finite, non-empty set of integers. If $\max S - \min S < \frac{2k-1}{k}|S| - 1$ with an integer $k \geq 2$, then $S - S$ contains all integers from the interval $(-|S|/(k-1), |S|/(k-1))$.*

We are now prepared for the main task of this section.

Deduction of Theorem 1 from the Main Lemma. Let $A \subseteq \mathbb{Z}$ and $B \subseteq \mathbb{N}$ be finite sets with $|A| > 1$ and $|B| < (36C)^{-1}|A|/\ln|A|$, where C is the constant of the Main Lemma. Assuming that

$$\mu_A(B) \leq |B| - 1 \tag{6}$$

we obtain a contradiction.

The cases where $B = \varnothing$ or $|A| = 2$ are immediate; suppose therefore that $B \neq \varnothing$ and $|A| \geq 3$. Write $d := \gcd B$, $B^* := \{b/d \colon b \in B\}$, and

$$A_j := \{(a-j)/d \colon a \in A,\ a \equiv j \pmod d\}; \quad j \in [0, d-1].$$

Fix an integer N so that the sets $A_j + jN$ are pairwise disjoint and let $A^* := \cup_{j=0}^{d-1}(A_j + jN)$. Evidently, we have $\Delta_A(b) \geq \Delta_{A^*}(b/d)$ for every $b \in B$, and thus it follows from (6) that $\mu_{A^*}(B^*) \leq \mu_A(B) \leq |B^*| - 1$. Since $|A^*| = |A|$, $|B^*| = |B|$, and $\gcd B^* = 1$, passing from our original sets A and B to the sets A^* and B^*, we ensure that $\gcd B = 1$.

Set $B^{\pm} := (-B) \cup \{0\} \cup B$ so that $\mu_A(B^{\pm}) = \mu_A(B)$ by properties (i) and (ii) at the beginning of the Introduction, and $B^{\pm}$ is not contained in an arithmetic progression with difference, larger than 1, in view of $\gcd B = 1$.

We have

$$\sum_{b \in 3B^{\pm}} \Delta_A(b) = \sum_{b \in 3B^{\pm}} \big(|A| - |A \cap (A+b)|\big) \geq |A||3B^{\pm}| - |A|^2,$$

whence

$$\mu_A(3B^{\pm}) \geq |A| \left(1 - \frac{|A|}{|3B^{\pm}|}\right).$$

Consequently, if $|3B^{\pm}| \geq 2|A|$, then $\mu_A(3B^{\pm}) \geq |A|/2$ and using (4) we obtain

$$\mu_A(B) = \mu_A(B^{\pm}) \geq \frac{1}{3}\,\mu_A(3B^{\pm}) \geq \frac{1}{6}\,|A| \geq |B|,$$

as wanted; accordingly, we assume $|3B^{\pm}| < 2|A|$. This allows us to apply (5) to the set $(3B^{\pm})_+$ of all positive elements of $3B^{\pm}$; using (4) and (6) we get then

$$\begin{aligned}\frac{1}{2}\,(|3B^{\pm}| - 1) = |(3B^{\pm})_+| &< 2\mu_A((3B^{\pm})_+) \\ &= 2\mu_A(3B^{\pm}) \leq 6\mu_A(B^{\pm}) = 6\mu_A(B) \leq 6(|B| - 1) = 3(|B^{\pm}| - 3)\end{aligned}$$

and hence

$$|3B^{\pm}| \le 6|B^{\pm}| - 18. \tag{7}$$

Let $l := \max(B^{\pm}) - \min(B^{\pm})$ and $\kappa := \lfloor (l-1)/(|B^{\pm}|-2) \rfloor$. By (7) and Corollary 1, we have $l \le 3|B^{\pm}| - 6$, and consequently $\kappa \le 2$; hence

$$|6B^{\pm}| - 1 \le 6l \le 6(\kappa+1)(|B^{\pm}|-2) \le 18(|B^{\pm}|-2) < 36|B| < C^{-1}|A|/\ln|A|.$$

It follows that $|A| > |6B^{\pm}| - 1$ whence, indeed,

$$|A| > C(|6B^{\pm}| - 1)\ln|A| > C(|6B^{\pm}| - 1)\ln(|6B^{\pm}| - 1). \tag{8}$$

On the other hand, applying Theorem 3 with $S = B^{\pm}$ and $h = 6$ and recalling that $\kappa \le 2$ we get

$$|6B^{\pm}| > 1 + (6-\kappa)l \ge 4l + 1 > \frac{2}{3}(6l+1)$$

and therefore

$$\max(6B^{\pm}) - \min(6B^{\pm}) = 6l < \frac{3}{2}|6B^{\pm}| - 1.$$

By Lemma 2 (applied with $S = 6B^{\pm}$ and $k = 2$) we have

$$[1, |6B^{\pm}| - 1] \subseteq 6B^{\pm} - 6B^{\pm} = 12B^{\pm}. \tag{9}$$

Since the function μ_A is monotonic in the sense that $B_1 \subseteq B_2$ implies $\mu_A(B_1) \le \mu_A(B_2)$, using (4) and the Main Lemma (which is applicable in view of (8)), from (9) we derive that

$$\begin{aligned} 12\mu_A(B) = 12\mu_A(B^{\pm}) \ge \mu_A(12B^{\pm}) &\ge \mu_A([1, |6B^{\pm}| - 1]) \\ &\ge |6B^{\pm}| - 1 \ge 6(|B^{\pm}| - 1) = 12|B|, \end{aligned}$$

contradicting (6). □

To establish Theorem 1 it remains to prove the Main Lemma.

4. Proof of the Main Lemma

For finite sets $A, S \subseteq \mathbb{Z}$ we write $\mathfrak{g}_A(S) := |S \setminus A|$ (to be interpreted as *the number of gaps in* S). Assuming that $m \in \mathbb{N}$ and $A \subseteq \mathbb{Z}$ are implicitly defined by the context, by a *problem* we mean a pair of the form $(a, a+d)$ with $a \in A$, $a + d \notin A$, and $d \in [1, m]$; we say that this problem is created by a. Evidently, if A is m-coverable, then for each fixed $d \in [1, m]$ there are at most $m - 1$ problems of the form $(a, a+d)$, hence at most $m(m-1)$ problems totally.

We split the proof of the Main Lemma into a number of auxiliary statements.

Lemma 3. *Let $m \in \mathbb{N}$ and let A be an m-coverable set. If J is a block of consecutive integers with $|J| \le \lceil m/2 \rceil$, then each of the $\lceil m/4 \rceil$ largest elements of the set $A \cap (-\infty, \min J)$ creates at least $\min\{m/4, \mathfrak{g}_A(J)\}$ problems.*

Proof. Write $J = [u+1, u+v]$, so that $v \le \lceil m/2 \rceil$, and suppose that a is one of the $\lceil m/4 \rceil$ largest elements of the set $A \cap (-\infty, u]$. If $a \ge u + v - m$, then (a, b) is a problem for every $b \in J \setminus A$, and hence a creates at least $\mathfrak{g}_A(J)$ problems. If $a < u + v - m$, then (a, b) is a problem for every $b \in [a+1, u] \setminus A$ with $b - a \le m$. Consequently, if $u - a \le m$, then the assertion follows from

$$|[a+1, u] \setminus A| \ge u - a - (\lceil m/4 \rceil - 1) \ge m - v - \lceil m/4 \rceil + 2 \\ \ge m - \lceil m/2 \rceil - \lceil m/4 \rceil + 2 \ge m/4,$$

and if $u - a > m$, then it follows from

$$|[a+1, a+m] \setminus A| \ge m - (\lceil m/4 \rceil - 1) \ge m/4.$$

□

To use Lemma 3 efficiently we have to show that there are many blocks of consecutive integers with large number of gaps. This constitutes the major difficulty and we postpone the corresponding part of the argument, demonstrating first how the proof of the Main Lemma is completed once this is done.

Proposition 1. *Let $m \in \mathbb{N}$ and let A be an m-coverable set. Suppose that $0 < \varepsilon \le 1/2$ and $L \ge m$ are real numbers such that for any integer u there is an integer w with $|w - u| \le L$, satisfying $\mathfrak{g}_A([w+1, w+m]) \ge \varepsilon m$. Then $|A| < 30\varepsilon^{-1}L$.*

Proof. We set $I_1 := [\min A - m, \min A - 1]$ and inductively construct blocks $I_2, I_3, \ldots$ of consecutive integers as follows. Assume that I_k has been constructed for some $k \in \mathbb{N}$. Applying the assumption of the proposition with $u := \max I_k + \lfloor L \rfloor$, we find integer $w \in [\max I_k, \max I_k + 2L]$ with $\mathfrak{g}_A([w+1, w+m]) \ge \varepsilon m$, and set $I_{k+1} := [w+1, w+m]$. We continue the process until we hit for the first time a block I_t with $\min I_t > \max A$. By the construction, the blocks $I_1, \ldots, I_t$ satisfy

$$\mathfrak{g}_A(I_k) \ge \varepsilon m; \quad k \in [1, t]$$

and

$$0 < \min I_{k+1} - \max I_k \le 2L + 1; \quad k \in [1, t-1].$$

In each block I_k we find a sub-block J_k with $|J_k| = \lceil m/2 \rceil$ and $\mathfrak{g}_A(J_k) \ge 0.5\varepsilon m$ and let

$$F_k := [\min J_k, \min J_{k+1}); \quad k = 1, \ldots, t-1.$$

Notice that $A \subseteq F_1 \cup \cdots \cup F_{t-1}$, and that

$$|F_k| = \min J_{k+1} - \min J_k \le (\min I_{k+1} + \lfloor m/2 \rfloor) - (\max I_k - m + 1) \le 2L + 3m/2$$

for each $k \in [1, t-1]$.

Letting $n_k := \min\{\lceil m/4\rceil, |F_k \cap A|\}$ for $k = 1, \dots, t-1$, by Lemma 3 we conclude that the elements of $F_k \cap A$ create at least $0.5\varepsilon m n_k$ problems. Since the total number of problems is at most $m(m-1)$, the number of those $k \in [1, t-1]$ with $n_k = \lceil m/4\rceil$ is less than $8\varepsilon^{-1}$, and hence the number of elements of A, lying in the corresponding sets F_k, is less than $8\varepsilon^{-1}(2L + 3m/2)$. On the other hand, the number of elements of A, lying in the sets F_k corresponding to those $k \in [1, t-1]$ with $n_k = |F_k \cap A|$, does not exceed $n_1 + \dots + n_{t-1}$, which is at most $2\varepsilon^{-1}m$ in view of $0.5\varepsilon m(n_1 + \dots + n_{t-1}) \le m(m-1)$. It follows that

$$|A| < 8\varepsilon^{-1}(2L + 3m/2) + 2\varepsilon^{-1}m < 30\varepsilon^{-1}L.$$

□

To prove the Main Lemma it suffices to show that there exists an absolute constant $\varepsilon > 0$ such that for any m-coverable set A, the assumption of Proposition 1 holds with $L = O(m \ln m)$. This is achieved in the series of lemmas that follow.

For $x \in \mathbb{Z}$ and $d \in \mathbb{N}$ set

$$P_d^-(x) := \{x, x-d, x-2d, \dots\},\ P_d^+(x) := \{x, x+d, x+2d, \dots\},$$

and

$$P_d(x) := P_d^-(x) \cup P_d^+(x).$$

Lemma 4. *Let $m \in \mathbb{N}$ and let A be an m-coverable set. Suppose that $d \in [1, m]$ is an integer, $g > 0$ is a real number, and I is a block of consecutive integers. If $\mathfrak{g}_A(I) \ge g$ and $m - g/5 \le |I| \le d$, then either there are at least $g/5$ integers $x \in I$ with $P_d^-(x) \cap A = \varnothing$, or there are at least $g/5$ integers $x \in I$ with $P_d^+(x) \cap A = \varnothing$.*

Proof. Set $G := I \setminus A$ and $G_0 := \{x \in G\colon P_d^-(x) \cap A \ne \varnothing,\ P_d^-(x) \cap A \ne \varnothing\}$; it suffices to show that $|G \setminus G_0| \ge 2g/5$.

We notice that for any $x \in I \cap A$ there is an element $a \in P_d(x) \cap A$ with $a + d \notin A$, and for any $x \in G_0$ there are at least *two* elements $a \in P_d(x) \cap A$ with $a + d \notin A$. Moreover, if x_1 and x_2 are distinct elements of I, then the progressions $P_d(x_1)$ and $P_d(x_2)$ are disjoint in view of $|I| \le d$. Since the number of problems of the form $(a, a+d)$ is at most $m - 1$, we have

$$|I \cap A| + 2|G_0| \le m - 1$$

whence

$$\begin{aligned} |G \setminus G_0| = |G| - |G_0| &\ge |I \setminus A| - \frac{1}{2}\big(m - 1 - |I \cap A|\big) \\ &= \frac{1}{2}\big(|I| + 1\big) + \frac{1}{2}\,\mathfrak{g}_A(I) - \frac{1}{2}\,m > \frac{1}{2}\,\mathfrak{g}_A(I) - \frac{1}{10}\,g \ge \frac{2}{5}\,g, \end{aligned}$$

as required. □

Lemma 5. *Let $m \in \mathbb{N}$ and let A be an m-coverable set. Suppose that d is an integer, $g > 0$ is a real number, and I is a block of m consecutive integers. If $\mathfrak{g}_A(I) \geq g$ and $m - g/10 \leq d \leq m$, then either there are at least $g/10$ integers $x \in I$ with $P_d^-(x) \cap A = \varnothing$, or there are at least $g/10$ integers $x \in I$ with $P_d^+(x) \cap A = \varnothing$.*

Proof. Since $d \geq m - g/10 > m/2$, we can represent I as a union of two (intersecting) blocks, consisting of d consecutive integers each. At least one of these two blocks, say I', satisfies $\mathfrak{g}_A(I') \geq g/2$. It remains to apply Lemma 4 with I replaced by I' and g replaced by $g/2$. □

Our next lemma, along with Proposition 1, is the key ingredient in the proof of the Main Lemma.

Lemma 6. *There exists an integer $K \geq 2$ with the following property: if $m \in \mathbb{N}$ and A is an m-coverable set, then for any integer u with $K \leq \mathfrak{g}_A([u+1, u+m]) \leq m/K$ there is an integer w such that $|w - u| \leq Km$ and*

$$\mathfrak{g}_A([w+1, w+m]) > 2\mathfrak{g}_A([u+1, u+m]).$$

Proof. Suppose that K and m are positive integers, A is an m-coverable set, and u is an integer with $K \leq \mathfrak{g}_A([u+1, u+m]) \leq m/K$. We want to show that if K is large enough (where "enough" is independent of m, A, and u), then there exists an integer w as in the statement of the lemma.

Write $I := [u+1, u+m]$ and $g := \mathfrak{g}_A(I)$. By Lemma 5, to any $d \in [m - g/10, m]$ there corresponds a subset $I_d \subseteq I \setminus A$ with $|I_d| \geq g/10$ such that either $P_d^-(x) \cap A = \varnothing$ for each $x \in I_d$, or $P_d^+(x) \cap A = \varnothing$ for each $x \in I_d$. We assume for definiteness that

$$P_d^+(x) \cap A = \varnothing;\ x \in I_d$$

holds for at least $g/20$ values of d, and accordingly there are at least $g^2/200$ pairs (d, x) with $d \in [m - g/10, m]$ and $x \in I \setminus A$, satisfying $P_d^+(x) \cap A = \varnothing$. Since there are at most g options for the second component of these pairs, there exists $x_0 \in I \setminus A$ such that $P_d^+(x_0) \cap A = \varnothing$ holds for at least $g/200$ values of d. We denote the set of these values by D; thus,

$$D \subseteq [m - g/10, m],\ |D| \geq g/200,\ \text{and } P_d^+(x_0) \cap A = \varnothing \text{ for every } d \in D.$$

For $k \in [1, K]$ write $W_k := \{(m-d)k \colon d \in D\}$ and let $W := W_1 \cup \cdots \cup W_K$, so that

$$W \subseteq [0, m-1]. \tag{10}$$

We have $|W_k| \geq g/200$ for $k = 1, \ldots, K$ and consequently, for every integer $C \in [1, K]$,

$$|W| \geq \left| \bigcup_{k=K-C+1}^{K} W_k \right| \geq \frac{Cg}{200} - \frac{1}{2} C^2 \max_{K-C<k<l\leq K} |W_k \cap W_l|. \tag{11}$$

We observe that $|W_k \cap W_l|$ is the number of solutions of the equation

$$kd_1 - ld_2 = (k-l)m$$

in the variables $d_1, d_2 \in D$. Since this equation uniquely determines the residue class of $d_1 \in [m - g/10, m]$ modulo $l/\gcd(k,l)$, its number of solutions is at most

$$\frac{g/10}{l/\gcd(k,l)} + 1 \le \frac{g}{10}\,\frac{l-k}{l} + 1 < \frac{g}{10}\,\frac{C}{K-C} + 1 < \frac{Cg}{K}$$

provided that, say, $10 \le C \le K/2$. Hence, for $C = 2000$ and K large enough, (11) gives

$$|W| > 5g. \tag{12}$$

Next, we notice that there are $m - g$ progressions $P_m(a)$ with $a \in I \cap A$, and every such progression contains a problem of the form $(a', a' + m)$ with some $a' \in A$. Since the total number of problems of this form is at most $m - 1$, there are at most $g - 1$ infinite arithmetic progressions with difference m, containing two or more problems. On the other hand, by the definition of the set W, for each $v \in W$ there exists $k \in [1, K]$ such that $x_0 - v + km \notin A$; consequently, if for some $v \in W$ we have $x_0 - v \in A$ and $x_0 - v + Km \in A$, then $P_m(x_0 - v)$ contains at least two problems. Combining these observations and recalling (10) we obtain

$$|\{v \in W \colon x_0 - v \in A,\, x_0 - v + Km \in A\} < g$$

and it follows that

$$|(x_0 - W) \setminus A| + |(x_0 + Km - W) \setminus A| > |W| - g.$$

Using (10) and (12) we see that the conclusion of the lemma holds true with either $w = x_0 - m$, or $w = x_0 + (K-1)m$. □

To satisfy the assumption $\mathfrak{g}_A([u+1, u+m]) \ge K$ of Lemma 6 we need some "seed gaps"; these are supplied by our last lemma.

Lemma 7. *Let $m \in \mathbb{N}$ and let A be an m-coverable set. Then for any integer u and $1 \le g \le m/2$ there is an integer w with $|w-u| < gm$ such that $\mathfrak{g}_A([w+1, w+m]) \ge g$.*

Proof. Since the number of problems of the form $(a, a+m)$ does not exceed $m-1$, there is a residue class modulo m, not represented in A. Consequently, there exists $b \in \mathbb{Z}$ with $|u - b| \le m/2$ such that $P_m(b) \cap A = \varnothing$. To simplify the notation we assume that $b = 0$; this does not restrict generality since one can replace A by $A - b$ and u by $u - b$. Thus, A does not contain multiples of m, and $|u| \le m/2$.

Set

$$J^- := \{j \in [-g+1, g-1] \colon P^-_{m-1}(jm) \cap A \neq \varnothing\},$$
$$J^+ := \{j \in [-g+1, g-1] \colon P^+_{m-1}(jm) \cap A \neq \varnothing\},$$

and

$$J := \{j \in [g, m-g-1] \colon P_{m-1}(jm) \cap A \neq \varnothing\}.$$

We notice that the progressions $P_{m-1}(jm)$ with distinct $j \in [-g+1, m-g-1]$ are pairwise disjoint. Furthermore, for any $j \in J^-$ there is $a \in P_{m-1}(jm) \cap A$ such that $a + (m-1) \notin A$, and $a < jm$; next, for any $j \in J^+$ there is $a \in P_{m-1}(jm) \cap A$ such that $a + (m-1) \notin A$, and $a > jm$; finally, for any $j \in J$ there is $a \in P_{m-1}(jm) \cap A$ such that $a + (m-1) \notin A$. Since there are at most $m-1$ problems of the form $(a, a+(m-1))$, it follows that

$$|J^+| + |J^-| + |J| < m,$$

whence either $|J^-| + |J|/2 < m/2$, or $|J^+| + |J|/2 < m/2$ holds. Suppose, for definiteness, that the latter of the two inequalities is true, so that

$$|J^+| + |J| < \frac{1}{2}\, m + \frac{1}{2}\, |J| \leq m - g,$$

and let $w := (g-1)m$. Then for each $j \in [-g+1, g-1] \setminus J^+$ the interval $[w+1, w+m]$ has a common element with the progression $P^+_{m-1}(jm)$, and for each $j \in [g, m-g-1] \setminus J$ it has a common element with the progression $P_{m-1}(jm)$. Since all these progressions are disjoint with A, we conclude that

$$\mathfrak{g}_A([w+1, w+m]) \geq (2g - 1 - |J^+|) + (m - 2g - |J|) \geq g,$$

and the result follows in view of $|w-u| \leq (g-1)m + m/2 < gm$. □

Eventually we are ready to prove the Main Lemma.

Proof of the Main Lemma. Suppose that $m \geq 2$ is an integer and A is an m-coverable set; we want to show that $|A| < Cm \ln m$ with an absolute constant C.

Let K be an integer, satisfying the conclusion of Lemma 6. We assume that $m \geq 20K$ as otherwise $|A| < 400K^2$ by (3) and the assertion is immediate.

Given an integer u, we apply Lemma 7 with $g = K$ to find an integer u_0 with $|u_0 - u| < Km$ and $\mathfrak{g}_A([u_0+1, u_0+m]) \geq K$. If $\mathfrak{g}_A([u_0+1, u_0+m]) \leq m/K$, then by Lemma 6 there exists an integer u_1 with $|u_1 - u_0| \leq Km$ such that $\mathfrak{g}_A([u_1+1, u_1+m]) > 2K$. We continue in this way finding subsequently $u_2, u_3, \ldots$ until we reach some u_t

satisfying $\mathfrak{g}_A([u_t+1, u_t+m]) > m/K$, and we let $w = u_t$. By the construction, we have $|w-u| < (t+1)Km$, whereas

$$m/K \ge \mathfrak{g}_A([u_{t-1}+1, u_{t-1}+m]) \ge 2^{t-1}K,$$

unless $t = 0$. Hence $t < \log_2(m/K^2) + 1 \le \log_2 m - 1$, implying $|w-u| < 2Km \ln m$. This shows that the assumptions of Proposition 1 hold with $\varepsilon = K^{-1}$ and $L = 2Km \ln m$, and therefore the assertion of the Main Lemma holds with $C = 60K^2$. $\square$

Appendix. Large sets A with $\mu_A([1,m]) < m$

We prove here that if $m \ge 2$ is an integer and

$$A = \bigcup_{0 \le k < \log_2 m} \big[km, (k+1)m - 2^k\big),$$

then for every $d \in [1,m]$ the set A is a union of at most $m-1$ arithmetic progressions with difference d, so that $\Delta_A(d) \le m-1$. For $d = m$ this is immediate, and we assume below that $d < m$. Recalling the terminology of Section 4, we say that the pair $(a, a+d)$ is a problem (created by a) if $a \in A$ and $a+d \notin A$; we want to show that to every fixed $d \in [1, m-1]$ there correspond at most $m-1$ problems. To this end for each $0 \le k < \log_2 m$ we let $I_k := \big[km, (k+1)m - 2^k\big)$ and count problems, created by the elements of I_k.

Let K denote the smallest integer with $2^K \ge m - d$, so that $0 \le K < \log_2 m + 1$. If $0 \le k < K$ and $a \in I_k$, then

$$a + d < (k+1)m - 2^k + d < (k+2)m - 2^{k+1},$$

and it follows that the number of problems, created by all elements of I_k, is at most

$$\min\{m - |I_k|, |I_k|\} = \min\{2^k, m - 2^k\}.$$

Consequently, if $K \ge \log_2 m$, then the total number of problems is at most

$$1 + 2 + \cdots + 2^{K-2} + (m - 2^{K-1}) = m - 1,$$

and the proof is over.

Suppose now that $K < \log_2 m$. If $K \le k < \log_2 m - 1$ and $a \in I_k$, then

$$(k+1)m - 2^k + d \ge (k+2)m - 2^{k+1}$$

whence

$$I_k + d \supseteq I_{k+1}.$$

It follows that the total number of problems, created together by all elements of I_k for all $k \ge K$, is at most

$$(|I_K| - |I_{K+1}|) + (|I_{K+1}| - |I_{K+2}|) + \cdots = |I_K| = m - 2^K.$$

On the other hand, from the estimate above it follows that the number of problems, created by the elements of I_k for all $0 \le k < K$, is at most

$$1 + 2 + \cdots + 2^{K-1} = 2^K - 1.$$

Therefore, the total number of all problems is at most

$$(m - 2^K) + (2^K - 1) = m - 1$$

and the assertion follows.

Acknowledgements

We are grateful to Yuri Rabinovich for bringing to our attention the problem, studied in this paper.

The research was carried out while the first author was visiting the Institute for Advanced Study. It is his pleasure to thank the IAS for its hospitality and excellent working conditions and the Oswald Veblen Fund for its generous support of the visit.

References

[BS96] R. Balasubramanian and K. Soundararajan, On a conjecture of R.L. Graham, *Acta Arith.* **75** (1) (1996), 1–38.

[EH64] P. Erdős and H. Heilbronn, On the addition of residue classes mod p, *Acta Arith.* **9** (1964), 149–159.

[G32] R.M. Gabriel, The rearrangement of positive Fourier coefficients, *Proc. London Math. Soc. (2)* **33** (1932), 32–51.

[G70] R. L. Graham, Problem 5749, *Amer. Math. Monthly* **77** (2) (1970), 775.

[HL28] G.H. Hardy and J.E. Littlewood, Notes on the theory of series (VIII): an inequality, *J. London Math. Soc.* **3** (1928), 105–110.

[HLP88] G.H. Hardy, J.E. Littlewood, and G. Polya, *Inequalities* 2d ed., Camb. Univer. Press, 1988.

[L96] V.F. Lev, Structure theorem for multiple addition and the Frobenius problem, *J. Number Theory* **58** (1) (1996), 79–88.

[L98] ———, Linear equations over $\mathbb{F}_p$ and moments of exponential sums, *Duke Math. J.* **107** (2) (2001), 239–263.

[L06] ———, Large sum-free sets in $\mathbb{Z}/p\mathbb{Z}$, *Israel J. Math.* **154** (2006), 221–233.

[O68] J.E. Olson, An addition theorem modulo p, *J. Comb. Theory* **5** (1968), 45–52.

[P74] J.M. Pollard, A generalization of the theorem of Cauchy and Davenport, *J. London Math. Soc. (2)* **8** (1974), 460–462.

Department of Mechanics and Mathematics, Moscow State University, Moscow, Russia

E-mail address: konyagin@ok.ru

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

E-mail address: seva@math.haifa.ac.il