CLASSIFICATION THEOREMS FOR SUMSETS MODULO A PRIME

HOI H. NGUYEN AND VAN H. VU

ABSTRACT. Let $\mathbf{Z}_p$ be the finite field of prime order p and A be a subsequence of $\mathbf{Z}_p$. We prove several classification results about the following questions:
(1) When can one represent zero as a sum of some elements of A ?
(2) When can one represent every element of $\mathbf{Z}_p$ as a sum of some elements of A ?
(3) When can one represent every element of $\mathbf{Z}_p$ as a sum of l elements of A ?

1. INTRODUCTION.

Let G be an additive group and A be a sequence of (not necessarily different) elements of G. We denote by S_A the collection of partial sums of A

$$S_A := \left\{ \sum_{x \in B} x | B \subset A, |B| < \infty \right\}.$$

For a positive integer $l \leq |A|$ we denote by l^*A the collection of partial sums of l elements of A,

$$l^*A := \left\{ \sum_{x \in B} x | B \subset A, |B| = l \right\}.$$

Example. If $G = \mathbf{Z}_{11}, A = \{1, 1, 7\}$ then $S_A = \{1, 2, 7, 8, 9\}$ and $2^*A = \{2, 8\}$.

The following questions are among the most popular in additive combinatorics

Question 1.1. *When is $0 \in S_A$ and when is $S_A = G$?*

Question 1.2. *For a given l when is $0 \in l^*A$ and when is $l^*A = G$?*

This work was written while the first author was supported by a DIMACS summer research fellowship.

The second author is an A. Sloan Fellow and is supported by an NSF Career Grant.

There is a vast amount of results concerning these questions ([8]), including classical results such as Olson's theorem and the Erdos-Ginzburg-Ziv theorem.

If $0 \notin S_A$ (or, respectively, $0 \notin l^*A$), then we say that A is *zero-sum-free* (or, respectively, *l-zero-sum-free*). If $S_A = G$ (or, respectively, $l^*A = G$), then we say that A is *complete* (or, respectively, *l-complete*); and otherwise we say that A is *incomplete* (*l-incomplete*).

We will focus on the case $G = \mathbf{Z}_p$, the cyclic group of order p, where p is a large prime. The main goal of this paper is to give a strong classification for zero-sum-free, incomplete and l-incomplete sequences of $\mathbf{Z}_p$. These classifications refine and extend an implicit result in [9]. Together they support the following general phenomenon:

The main reason for a sequence to be zero-sum-free or incomplete is that its elements have small norm.

For instance, if the elements of a sequence (viewed as positive integers between 0 and $p-1$) add up to a number less than p, then the sequence is clearly zero-sum-free. One of our results, Theorem 2.2, shows that any zero-sum-free sequence in $\mathbf{Z}_p$ can be brought into this form after a dilation and after truncation of a negligible subset.

Our results have many applications (see Sections 3,4,5 and 6). In particular, we will prove a refinement of the well-known Erdős-Ginzburg-Ziv theorem (see Section 6). The common theme of these applications is the following.

Any long zero-sum-free or incomplete sequence is a subsequence of a unique extremal sequence (after a proper linear transformation and a possible truncation of a negligible subsequence).

In the rest of this section, we introduce our notation. The remaining sections are organized as follows. In Section 2, we present our classification theorems. Sections 3,4,5,6 are devoted to applications. Section 7 contains the main lemmas needed for the proofs. The proofs of the classification theorems come in Sections 8,9 and 10.

Notation.

We will use $\mathbf{Z}$ to denote the set of integers and $\mathbf{Q}$ to denote the set of rational numbers. Also $\mathbf{Z}_D$ will denote the congruence group modulo D.

For sequences A and B, define $A + B := \{a + b | a \in A, b \in B\}$.

For an element $b \in \mathbf{Z}_p$ and a sequence A, define $b \cdot A := \{ba | a \in A\}$.

A good way to present a sequence A is to write $A := \{a_1^{[m_1]}, \dots, a_k^{[m_k]}\}$, where m_a is the multiplicity of a in A (sometime we use the notation $m_a(A)$ to emphasize the role of A), and $a_1, \dots, a_k$ are the different elements of A.

The maximum multiplicity of A is $m(A) := \max_{a \in \mathbf{Z}_p} m_a(A)$. We will always assume that $m(A) \le p$, for every sequence A in the paper.

We say A is decomposed into subsequences $A_1, \dots, A_k$ and write $A = \bigcup_{i=1}^{*k} A_i$ if $m_a(A) = \sum_{i=1}^{k} m_a(A_i)$ for every $a \in \mathbf{Z}_p$.

Asymptotic notation will be used under the assumption that $p \to \infty$. For $x \in \mathbf{Z}_p$, $\|x\|$ (the norm of x) is the distance from x to 0. (For example, the norm of $p-1$ is 1).

A subset X of $\mathbf{Z}_p$ is called a K-net if for any $n \in \mathbf{Z}_p$ there exists $x \in X$ such that $n \in [x, x+K]$. It is clear that if X is a K-net, then $X + T = \mathbf{Z}_p$ for any interval T of length K in $\mathbf{Z}_p$. We will use the same notion over $\mathbf{Z}$ and $\mathbf{Q}$ as well.

For a finite set X of real numbers we use $\min(X)(or, respectively, \max(X))$ to denote the minimum (respectively, maximum) element of X.

2. The classifications.

In order to make the statements of the theorems less technical, we define

$$f(p,m) := \left\lfloor (pm)^{6/13} \log^2 p \right\rfloor.$$

2.1. **Zero-sum-free sequences.** View the elements of $\mathbf{Z}_p$ as integers between 0 and $p-1$. The most natural way to construct a zero-sum-free sequence is to select non-zero elements whose sum is less than p. Our first theorem shows that this is essentially the only way.

Theorem 2.2. *There is a positive constant c_1 such that the following holds. Let $1 \le m \le p$ be a positive integer and A be a zero-sum-free sequence of $\mathbf{Z}_p$ satisfying $m(A) \le m$. Then there is a non-zero residue b and a subsequence $A^\flat \subset A$ of cardinality at most $c_1 f(p,m)$ such that*

$$\sum_{a \in b \cdot (A \setminus A^\flat)} a < p.$$

Notice that zero-sum-freeness and incompleteness are preserved under dilation. This explains the presence of the element b in the theorem. Another issue one needs to address is the cardinality of the exceptional sequence $A^\flat$. It is known (and not hard to prove) that most zero-sum-free sequences with maximum multiplicity

m in $\mathbf{Z}_p$ have cardinality $\Theta((pm)^{1/2})$. Thus, in most cases, the cardinality of $A^\flat$ (which is at most $(pm)^{6/13+o(1)}$) is negligible compared to that of $|A|$. (The same will apply for later results.) Exceptional sequences cannot be avoided (see Sections 3,4 and also [5]).

By setting $m = 1$, we have the following corollary for the case when A is a set.

Corollary 2.3. *There is an absolute positive constant c_1 such that the following holds. For any zero-sum-free subset A of $\mathbf{Z}_p$ there is a non-zero residue b and a set $A^\flat \subset A$ of cardinality at most $c_1 f(p,1)$ such that*

$$\sum_{a \in b\cdot(A\setminus A^\flat)} a < p.$$

2.4. **Incomplete sequences.** The easiest way to construct an incomplete sequence is to select elements with small norms. Clearly, if A is a sequence where $\sum_{a\in A}\|a\| < p-1$ then A is *incomplete*. Our second theorem shows that this trivial construction is essentially the only possibility.

Theorem 2.5. *There is a positive constant c_2 such that the following holds. Let $1 \le m \le p$ be a positive integer and A be an incomplete sequence in $\mathbf{Z}_p$ satisfying $m(A) \le m$. Then there is a non-zero element $b \in \mathbf{Z}_p$ and a subsequence $A^\flat \subset A$ of cardinality at most $c_2 f(p,m)$ such that*

$$\sum_{a \in b\cdot(A\setminus A^\flat)} \|a\| < p.$$

By setting $m = 1$, we have

Corollary 2.6. *There is a positive constant c_2 such that the following holds. For any incomplete subset A of $\mathbf{Z}_p$ there is a non-zero residue b and a set $A^\flat \subset A$ of cardinality at most $c_2 f(p,1)$ such that*

$$\sum_{a \in b\cdot(A\setminus A^\flat)} \|a\| < p.$$

2.7. **l-incomplete sequences.** View A as a sequence of integers in the interval $[-(p-1)/2, (p-1)/2]$. Our classification in this subsection is a little bit different from the previous two. We are going to classify the structure of l^*A instead of that of A. The reason is that this classification is natural and easy to state. Furthermore, it is also easy to derive information about A using the classification of l^*A.

If all l-sums of A belong to an interval of length less than p in $\mathbf{Z}$, then A is l-incomplete in $\mathbf{Z}_p$. Of course, the converse is not true. However, our third theorem says that the reversed statement can be obtained at the cost of a small modification (in the spirit of the previous theorems).

Theorem 2.8. *There is a positive constant c_3 such that the following holds. Let $1 \leq m \leq p$ be a positive integer, let A be a sequence in $\mathbf{Z}_p$, and let l be an integer satisfying $c_3 f(p,m) \leq l \leq |A| - c_3 f(p,m)$. Assume furthermore that A is l-incomplete and $m(A) \leq m$. Then there exist*

- *residues $b, c \in \mathbf{Z}_p$ with $b \neq 0$,*
- *a sequence $A^\flat \subset A$ of cardinality less than $c_3 f(p,m)$, and*
- *an integer $l_1 \geq l - 2f(p,m)$*

such that the union $\bigcup_{l_1 \leq l' \leq l_1 + (pm)^{3/13}} l'^* A'$ is contained in an interval of length less than p, where $A' := b \cdot (A \backslash A^\flat) + c$ is considered as a sequence of integers in $[-(p-1)/2, (p-1)/2]$.

The property l-incompleteness is preserved under linear transforms. This explains why we need two parameters b and c in the theorem. The reader is invited to state a corollary for the case when A is a set.

3. Structure of long zero-sum-free sequences.

Let $1 \leq m \leq p$ be a positive integer and A be a zero-sum-free sequence of $\mathbf{Z}_p$ with maximum multiplicity $m(A) \leq m$. Trying to make A as long as possible, we come up with the following natural candidate

$$A_1^m := \{1^{[m]}, 2^{[m]}, \dots, (n-1)^{[m]}, n^{[k]}\}$$

where k and n are the unique integers satisfying $1 \leq k \leq m$ and

$$m(1 + 2 + \cdots + n - 1) + kn < p \leq m(1 + 2 + \cdots + n - 1) + (k+1)n.$$

As a consequence of Theorem 2.2, one can show that any zero-sum free sequence with $m(A) \leq m$ and cardinality close to $|A_1^m|$ is almost a subsequence of A_1^m, after a proper dilation.

Theorem 3.1. *Let $6/13 < \alpha < 1/2$ be a fixed constant. Assume that A is a zero-sum-free sequence of $\mathbf{Z}_p$ with maximum multiplicity $m(A) \leq m$ and cardinality $|A_1^m| - O((pm)^\alpha)$. Then there is a non-zero element $b \in \mathbf{Z}_p$ and a subsequence $A^\flat \subset A$ of cardinality $O((pm)^{(\alpha+1/2)/2})$ such that $b \cdot (A \backslash A^\flat) \subset A_1^m$.*

We can go further by showing not only that $|A \backslash A_1^m|$ is small, but also that the sum of the norm of the elements in this sequence is small. An example is given by Theorem 1.9 of [5], which we restate below.

Theorem 3.2. [5] *Let A be a zero-sum-free subset of $\mathbf{Z}_p$ of size at least $.99\sqrt{2p}$. Then there is some non-zero element $b \in \mathbf{Z}_p$ such that*

$$\sum_{a \in b \cdot A, a < p/2} \|a\| \leq p + O(p^{1/2})$$

and

$$\sum_{a \in b \cdot A, a > p/2} \|a\| = O(p^{1/2}).$$

The bound $O(p^{1/2})$ is sharp.

Now assume that the cardinality of A differs from that of the extreme example A_1^m by a constant. In this case, we can tell exactly what A is.

Let $n(p)$ denote the largest integer n such that

$$\sum_{i=1}^{n-1} i < p.$$

Theorem 3.3. [5] *There is a constant C such that the following holds for all primes $p \geq C$.*

- *If $p \neq \frac{n(p)(n(p)+1)}{2} - 1$, and A is a subset of $\mathbf{Z}_p$ with $n(p)$ elements, then $0 \in S_A$.*
- *If $p = \frac{n(p)(n(p)+1)}{2} - 1$, and A is a subset of $\mathbf{Z}_p$ with $n(p)+1$ elements, then $0 \in S_A$. Furthermore, up to a dilation, the only zero-sum-free set with $n(p)$ elements is $\{-2, 1, 3, 4, \dots, n(p)\}$.*

We sketch the proof of Theorem 3.1.

Proof (Proof of Theorem 3.1.) Theorem 2.2 implies that there is a non-zero residue b and a subsequence $A^\flat \subset A$ of cardinality less than $c_1 f(p,m)$ such that $\sum_{a \in A'} a < p$, where $A' = b \cdot (A \backslash A^\flat)$ is viewed as sequence of integers in $[1, p-1]$.

Notice that $|A'| = |A_1^m| - O((pm)^\alpha) - c_1 f(p,m) = |A_m^1| - O((pm)^\alpha)$. For short put $t = |A' \backslash A_1^m|$. It follows from the inequality $n + \sum_{a \in A_1^m} a \geq p \geq \sum_{a \in A'} a$ that

$$\sum_{a \in A' \backslash A_1^m} a \leq n + \sum_{a \in A_1^m \backslash A'} a. \tag{1}$$

Let A_1' be the any subsequence of cardinality t in $A_1^m \backslash A'$ and let $A_1'' = A_1^m \backslash (A' \cup A_1')$. Note that

$$|A_1''| = |A_1^m| - |A'| = O(pm)^\alpha \text{ and } a \leq n \leq (2p/m)^{1/2} + 1$$

for any $a \in A_1''$. Thus

$$n + \sum_{a \in A_1''} a = O(pm)^\alpha (p/m)^{1/2}. \tag{2}$$

On the other hand, by definition, every element of $A' \backslash A_1^m$ is strictly greater than every element of A_1'. Additionally, since the maximum multiplicity is m, we have

$$\sum_{a \in A' \backslash A_1^m} a - \sum_{a \in A_1'} a \geq 1 + \cdots + 1 + 2 + \cdots + 2 + 3 + \cdots + 3 + \cdots,$$

where on the right hand side all numbers (with the possible exception of the last) appear exactly m times and the total number of summands is t. It is clear that such a sum is greater than $t^2/3m$; thus

$$\sum_{a \in A' \backslash A_1^m} a - \sum_{a \in A_1'} a \geq t^2/3m. \tag{3}$$

(1),(2),(3) together give

$$t^2/3m \leq \sum_{a \in A' \backslash A_1^m} a - \sum_{a \in A_1'} a \leq n + \sum_{a \in A_1''} a = O(pm)^\alpha (p/m)^{1/2}.$$

In other words, $t = O((pm)^{(\alpha+1/2)/2})$.

■

4. Structure of long incomplete sequence.

Let $1 \leq m \leq p$ be a positive integer and A be an incomplete sequence of $\mathbf{Z}_p$ with maximum multiplicity $m(A) \leq m$. Trying to make A as large as possible, we come up with the following example,

$$A_2^m = \{-n^{[k]}, -(n-1)^{[m]}, \ldots, -1^{[m]}, 0^{[m]}, 1^{[m]}, \ldots, (n-1)^{[m]}, n^{[k]}\}$$

where $1 \leq k \leq m$ and n are the unique integers satisfying

$$2m(1 + 2 + \cdots + n - 1) + 2kn < p \leq 2m(1 + 2 + \cdots + n - 1) + 2(k+1)n.$$

Using Theorem 2.5, we can prove the following.

Theorem 4.1. *Let $6/13 < \alpha < 1/2$ be a fixed constant. Assume that A is an incomplete sequence of $\mathbf{Z}_p$ with maximum multiplicity m and cardinality $|A| = |A_2^m| - O((pm)^\alpha)$. Then there is a non-zero element $b \in \mathbf{Z}_p$ and a subsequence $A^\flat \subset A$ of cardinality $O((pm)^{(\alpha+1/2)/2})$ such that $b \cdot (A \backslash A^\flat) \subset A_2^m$.*

The proof is similar to that of Theorem 3.1 and is omitted.

As an analogue of Theorem 3.2, we have

Theorem 4.2. [5] *Let A be an incomplete subset of $\mathbf{Z}_p$ of size at least $1.99p^{1/2}$. Then there is some non-zero element $b \in \mathbf{Z}_p$ such that*

$$\sum_{a \in b \cdot A} \|a\| \leq p + O(p^{1/2}).$$

(Again, the error term $O(p^{1/2})$ is sharp.)

A well-known theorem of J. E. Olson [7] gives a sharp estimate for the maximum cardinality of an incomplete set.

Theorem 4.3. *Let A be a subset of $\mathbf{Z}_p$ of cardinality more than $(4p-3)^{1/2}$. Then A is complete.*

5. The number of zero-sum-free and incomplete sequences.

In this section we apply Theorems 2.2, 2.5 to count the number of zero-sum-free sequences and incomplete sequences.

We fix m. The following theorem is well known in theory of partitions (a corollary of a theorem of G. Meinardus, [1, Theorem 6.2]).

Theorem 5.1. *Let $p_m(n)$ be the number of partitions of n in which each positive integer appears at most m-times. Then*

$$p_m(n) = \exp\left((\sqrt{(1 - \frac{1}{m+1})\frac{2}{3}}\pi + o(1))\sqrt{n}\right).$$

By Theorem 2.2, the main part of zero-sum-free sequences (after a proper dilation) corresponds to a partition of a number less than p. Thus, using Theorem 5.1, we infer the following.

Theorem 5.2. *Let N_1^m be the number of zero-sum-free sequences A satisfying $m(A) \leq m$. Then*

$$N_1^m = \exp((\sqrt{(1 - \frac{1}{m+1})\frac{2}{3}}\pi + o(1))\sqrt{p}).$$

Corollary 5.3. *The number of zero-sum-free sets is* $\exp((\sqrt{\frac{1}{3}}\pi + o(1))\sqrt{p})$.

By Theorem 2.5, the main part of incomplete sequences (after a proper dilation) can be split into two parts, each of which corresponds to a partition of a number less than p. Thus we obtain the following.

Theorem 5.4. *Let* N_2^m *be the number of incomplete sequences* A *satisfying* $m(A) \leq m$. *Then*

$$N_2^m = \exp((\sqrt{(1-\frac{1}{m+1})\frac{4}{3}}\pi + o(1))\sqrt{p}).$$

Corollary 5.5. *The number of incomplete sets is* $\exp((\sqrt{\frac{2}{3}}\pi + o(1))\sqrt{p})$.

Proof (Proof of Theorem 5.2) The lower bound for N_1^m is obvious, any partition of $p-1$ in which each number appears at most m-times gives a zero-sum-free sequence of maximum multiplicity bounded by m.

For the upper bound, we apply Theorem 2.2. First, the number of choice for $A^\flat$ is $\sum_{n\leq (pm)^{6/13+o(1)}} \binom{pm}{n} = \exp(o(\sqrt{p}))$. Second, the elements of $A' := b(A\backslash A^\flat)$ forms a partition of $\sum_{a\in A'} a$ (which is a positive integer less than p) in which each positive integer appears at most m-times. Hence, the number of choice for $A^\sharp$ is at most

$$\sum_{n\leq p-1} p_m(n) \leq p\exp((\sqrt{(1-\frac{1}{m+1})\frac{2}{3}}\pi + o(1))\sqrt{p}).$$

Finally, together with dilations, the number of zero-sum-free sequences is bounded by

$$p^2\exp((\sqrt{(1-\frac{1}{m+1})\frac{2}{3}}\pi + o(1))\sqrt{p}) = \exp((\sqrt{(1-\frac{1}{m+1})\frac{2}{3}}\pi + o(1))\sqrt{p}).$$

■

Proof (Proof of Theorem 5.4) The lower bound for N_2^m is again obvious, any two partitions of $(p-3)/2$ in which each number appears at most m-times give two nonnegative sequences. We then take the union of one sequence with the negative of the other sequence. It is not hard to check that the formed sequence A is incomplete and $m(A) \leq m$. Thus

$$N_2^m \geq (p_m((p-1)/2))^2 = \exp((\sqrt{(1-\frac{1}{m+1})\frac{4}{3}}\pi + o(1))\sqrt{p}).$$

For the upper bound we use Theorem 2.5. Argue similarly as in the proof of Theorem 5.2, we infer that the number of exceptional sequences $A^\flat$ is at most

$e^{o(\sqrt{p})}$. Write $A' := b(A \backslash A^\flat) = A^+ \cup A^-$, the decomposition of A' into sequences of nonnegative and negative elements respectively. The elements of A^+ form a partition of $\sum_{a \in A^+} a$ in which each positive integer appears at most m-times. The elements of A^- corresponds to a partition of $\sum_{a \in A^-}(-a)$ in which each (negative) number appears at most m-times. Thus the number of choice for A' is at most

$$\sum_{k+l<p} p_m(k)p_m(l) \leq p^2 \exp((\sqrt{(1-\frac{1}{m+1})\frac{4}{3}}\pi + o(1))\sqrt{p}).$$

Putting everything together, we obtain an upper bound for N_2^m,

$$\begin{aligned} N_2^m &\leq pe^{o(\sqrt{p})}p^2 \exp((\sqrt{(1-\frac{1}{m+1})\frac{4}{3}}\pi + o(1))\sqrt{p}) \\ &\leq \exp((\sqrt{(1-\frac{1}{m+1})\frac{4}{3}}\pi + o(1))\sqrt{p}). \end{aligned}$$

■

6. l-INCOMPLETE SEQUENCES

Assume that A, l, m satisfy conditions of Theorem 2.8. Trying to make A as large as possible, we come up with the following example,

$$A_3^m = \{-n^{[k]}, -(n-1)^{[m]}, \dots, -1^{[m]}, 0^{[m]}, 1^{[m]}, \dots, (n-1)^{[m]}, n^{[k]}\}$$

where k and n are the optimal integers such that $1 \leq k \leq m$ and all the l-sums of A_3^m are contained in an interval of length less than p.

However, the extremal example for l-incomplete sequences, in general, is not unique (for instance if $l = m = p$ then any sequence $\{-1^{[n]}, 0^{[p]}, 1^{[p-2-n]}\}$ is $l-$incomplete and of maximum cardinality). Nevertheless, Theorem 2.8 still allows us to conclude that any $l-$incomplete sequence of size close to $|A_3^m|$ can be dilated and translated into one of the extremal examples, as in the spirit of Theorems 3.1 and 4.1.

Let us discuss in detail the special case $l = p$. This is motivated by the classical theorem of P. Erdős, A. Ginzburg and A. Ziv [3], one of the starting points of combinatorial number theory.

Theorem 6.1. *(Erdős-Ginzburg-Ziv) For any sequence $A \in Z_p$ of cardinality $2p-1$ there is a subsequence $A' \subset A$ of cardinality p such that $\sum_{a \in A'} a = 0$.*

In fact, P. Erdős, A. Ginzburg and A. Ziv proved the statement for any finite abelian group G, by reducing it to the case $G = \mathbf{Z}_p$ above.

In the context of this paper, Theorem 6.1 stated that any sequence of cardinality $2p-1$ in $\mathbf{Z}_p$ is not p-zero-sum-free. The bound $2p-1$ is sharp as shown by the example $A = \{a^{[p-1]}, b^{[p-1]}\}$, for any two different elements $a, b \in \mathbf{Z}_p$. Using Theorem 2.8, we prove that if A is p-zero-sum-free and $|A| - p \gg f(p,p) = \lfloor p^{12/13} \log^2 p \rfloor$, then A has two elements of high multiplicities.

Theorem 6.2. *There is a positive constant C such that the following holds for all primes $p > C$. Assume that A is a p-zero-sum-free sequence and $p + c_3 f(p,p) \leq |A| \leq 2p-2$. Then $\{a^{[m_a]}, b^{[m_b]}\} \subset A$, where a, b are two different elements of $\mathbf{Z}_p$ and $m_a + m_b \geq 2(|A| - p - (c_3+3)f(p,p))$.*

Notice that A must have at least p elements so that the notion of p-zero-sum-free makes sense. Our theorem already yields a non-trivial conclusion when A has slightly more than p elements. A similar statement was proved in [4], but under the stronger assumption that $|A| \geq \frac{3}{2}p$.

As a corollary, one obtains the following refinement of Theorem 6.1.

Corollary 6.3 ([2])**.** *The following holds for all sufficiently large primes p. Let A be a p-zero-sum-free sequence of cardinality $2p-2$ in $\mathbf{Z}_p$. Then $A = \{a^{[p-1]}, b^{[p-1]}\}$, where a, b are two different elements of $\mathbf{Z}_p$.*

Proof (Proof of Corollary 6.3) By Theorem 6.2, we may assume that

$$A = \{0^{[p-k_1]}, 1^{[p-k_2]}, a_1, \ldots, a_l\}$$

where $1 \leq k_1 = o(p), 1 \leq k_2 = o(p), l = k_1 + k_2 - 2$ and a_i are (not necessarily distinct) integers in $[-p/2, p/2] \backslash \{0, 1\}$. If $l = 0$ then we are done. Assume that $l \geq 1$. We are going to construct a subsequence of A of length p whose elements sum up to zero modulo p.

Case 1: There is some a_i with absolute value at least $p/6$.

Assume that $p/2 > a_1 \geq p/6$. The subsequence $\{0^{[a_1-1]}, 1^{[p-a_1]}, a_1\}$ has cardinality p and sums up to zero modulo p. In the case $-p/2 < a_1 \leq -p/6$, consider the subsequence $\{0^{[p-|a_1|-1]}, 1^{[|a_1|]}, a_1\}$.

Case 2: All a_i have absolute value less than $p/6$ and there are at least $\max\{1, k_1 - 1\}$ negatives among them.

By a greedy algorithm, one can find a non-empty sequence (say, $a_1, \ldots, a_{l_1}$) of negative elements such that $l_1 + |a_1 + \cdots + a_{l_1}| \geq k_1$. Then the subsequence

$$\{0^{[p-l_1-|a_1+\ldots a_{l_1}|]}, 1^{[|a_1+\ldots a_{l_1}|]}, a_1, \ldots, a_{l_1}\}$$

sums up to zero modulo p.

Case 3: All a_i have absolute value less than $p/6$ and there are at least $\min\{l, k_2\}$ positives among them.

As each positive element is at least 2 and at most $p/6$, there is a subsequence of (say, l_2) positive elements whose sum is at least k_2 and at most $p/3$. Assume that $a_1, \dots, a_{l_2}$ are these elements. Then the subsequence

$$\{0^{[(a_1+\dots a_{l_1})-l_2]}, 1^{[p-(a_1+\dots a_{l_2})]}, a_1, \dots, a_{l_2}\}$$

sums up to zero modulo p. ■

We conclude this section by sketching the proof of Theorem 6.2.

Proof (Sketch of proof of Theorem 6.2) Since A is p-zero-sum-free in $\mathbf{Z}_p$, A is also p-incomplete. By Theorem 2.8, after a linear transform, we can find a subsequence A' of A such that

$$\max\{l_1^* A'\} - \min\{l_1^* A'\} < p, \tag{4}$$

where $l_1 \geq p - 2f(p,p)$ and $|A'| \geq |A| - c_3 f(p,p)$ and where c_3 is a positive constant. (Recall that $\max(X)$ (respectively, $\min(X)$) refers to the maximum (respectively, minimum) element in X.)

Let $A' = \{a_1, \dots, a_q\}$, where $a_i \leq a_{i+1}$ for $1 \leq i \leq q-1 = |A'|-1$ and rewrite (4) as

$$\sum_{i=1}^{l_1} a_{q-l_1+i} - \sum_{i=1}^{l_1} a_i = \sum_{i=1}^{k} a_{q-k+i} - \sum_{i=1}^{k} a_i < p, \tag{5}$$

where $k = \min(l_1, q-l_1)$. Note that

$$\sum_{i=1}^{k} a_{q-k+i} - \sum_{i=1}^{k} a_i \geq \sum_{i=i_0}^{j_0} a_{i+p} - \sum_{i=i_0}^{j_0} a_i = \sum_{i=i_0}^{j_0} (a_{i+p} - a_i), \tag{6}$$

where $i_0 = \max(1, q-l_1-p+1)$ and $j_0 = \min(l_1, q-p)$.

Since A has maximum multiplicity less than p, we have, for any i, that $a_{i+p} - a_i \geq 1$. Thus by (6) we obtain that

$$j_0 - i_0 \leq \sum_{i=i_0}^{j_0} (a_{i+p} - a_i) < p,$$

and we infer that the number of $i \in [i_0, j_0]$ such that $a_{i+p} - a_i = 1$ is at least $2(j_0 - i_0) - p + 3$. Next let i_1 and j_1 be the smallest and largest index i in $[i_0, j_0]$ such that $a_{i+p} - a_i = 1$. Thus $a_{i_1+p} - a_{i_1} = a_{j_1+p} - a_{j_1} = 1$ and

$$2(j_0 - i_0) - p + 2 \le j_1 - i_1 \le j_0 - i_0 < p. \tag{7}$$

In what follows, a_{i_1} plays a special role, so we denote it by a to distinguish it from the other a_i. Let $B = \{a_{i_1}, \ldots, a_{j_1+p}\}$. Obviously $|B| = j_1 - i_1 + p + 1$ and $a_{j_1+p} - a_{i_1} \le 2$.

Set $\gamma := j_0 - i_0$. Then $0 \le \gamma \le l_1 - 1$. We consider two cases.

Case 1: $a_{j_1} = a$**.** In this case $a_{j_1+p} = 1$ and $B = \{x^{[m_0]}, (x+1)^{[m_1]}\}$ where

$$m_0 + m_1 = j_1 - i_1 + p + 1 \ge 2(j_0 - i_0) - p + 2 + p + 1 = 2\gamma + 3. \tag{8}$$

Case 2: $a_{j_1} = a+1$**.** Recall that the number of pairs (a_i, a_{i+p}) such that $a_{i+p} - a_i = 1$ is at least $2(j_0 - i_0) - p + 2 = 2\gamma - p + 2$. Furthermore if $a_{i+p} - a_i = 1$ then either a_i or a_{i+p} must be $a+1$. By this observation, none of the elements in $\{a_{j_1+1}, \ldots, a_{p+i_1-1}\}$ belongs to any pair (a_i, a_{i+p}) with $a_{i+p} - a_i = 1$. Furthermore, we have $a_i = a + 1$ for $j_1 + 1 \le i \le p + i_1 - 1$. As a consequence, the multiplicity m_1 of $a + 1$ in B is at least

$$m_1 \ge 2\gamma - p + 2 + (p + i_1 - j_1 - 1) = 2\gamma - (j_1 - i_1) + 1. \tag{9}$$

It is convenient to write $B = \{a^{[m_0]}, (a+1)^{[m_1]}, (a+2)^{[m_2]}\}$. Clearly we have $\min(p^*B) = \min(p - m_0, m_1) + 2(p - m_0 - \min(p - m_0, m_1))$ and $\max(p^*B) = 2m_2 + \min(p - m_2, m_1)$.

Besides, it is not hard to show that

$$p^*B = [\min(p^*B), \max(p^*B)]. \tag{10}$$

The p-zero-sum-free assumption implies that $\max(p^*B) < p$. It follows that

$$2m_2 + \min(p - m_2, m_1) < p. \tag{11}$$

Consequently,

$$2m_2 + m_1 < p. \tag{12}$$

From (9) and (12) we deduce that $m_2 \le (p - 2\gamma + (j_1 - i_1) - 2)/2$. On the other hand, $m_0 + m_1 + m_2 = |B| = j_1 - i_1 + p + 1$. Thus

$$m_0 + m_1 \geq j_1 - i_1 + p + 1 - (p - 2\gamma + (j_1 - i_1) - 2)/2 \geq \gamma + 2 + (j_1 - i_1 + p)/2.$$

The latter inequality, together with (7), yields

$$m_0 + m_1 \geq 2\gamma + 3. \tag{13}$$

To summarize, in both cases ((8) and (13)) we have $m_0 + m_1 \geq 2\gamma + 3$. Combining this with the estimates $l_1 \geq p - 2f(p,p)$ and $q \geq |A| - c_3 f(p,p)$ we get

$$\begin{aligned} m_0 + m_1 &\geq 2(\min(l_1, q-p) - \max(1, q - l_1 - p + 1)) + 3 \\ &\geq 2(|A| - p) - (2c_3 + 6)f(p,p). \end{aligned}$$

■

7. The key lemmas.

The key lemmas we use in proofs are the following results from [10].

Theorem 7.1. *For any fixed positive integer d there exist positive $C = C(d)$ and $c = c(d)$ depending on d such that the following holds. If A is a subset of $[n]$ and l is a positive integer such that $l^d|A| \geq C(d)n$ and $l \leq |A|/2$. Then l^*A contains an arithmetic progression of length $c(d)l|A|^{1/d}$.*

Theorem 7.2. *For any fixed positive integer d there exist positive $C = C(d)$ and $c = c(d)$ depending on d such that the following holds. If A is a subset of $\mathbf{Z}_p$, $|A| \geq 2$ and l is a positive integer such that $l^{d+1}|A| \geq C(d)p$, then l^*A contains all residue classes modulo p or contains an arithmetic progression of length $c(d)l|A|^{1/d}$.*

Theorem 7.3. *For any fixed positive integer d there exist positive $C = C(d)$ and $c = c(d)$ depending on d such that the following holds. Let $A_1, \dots, A_l$ be subsets of cardinality $|A|$ of $\mathbf{Z}_p$ where l and $|A|$ satisfy $l^{d+1}|A| \geq C(d)p$. Then $A_1 + \dots + A_l$ contains all residue classes modulo p or an arithmetic progression of length $c(d)l|A|^{1/d}$.*

In our proofs we will be mainly interested in the case $d = 1$ and $d = 2$. We will also use the following lemmas. The proofs are left as exercises.

Lemma 7.4. [7] *There are positive constants C_0 and c_0 such that the following holds. Let A be a set of $\mathbf{Z}_p$ satisfying $|A| \leq C_0 p^{1/2}$. Then*

$$|l^*A| \geq c_0|A|^2$$

where $l = \lfloor |A|/2 \rfloor$.

Lemma 7.5. *Let D be a positive integer and X be a sequence of cardinality D in $\mathbf{Z}_D$. Then S_X contains the zero element. Furthermore, if the elements of X are co-prime with D, then $S_X = \mathbf{Z}_D$.*

Lemma 7.6. [9] *Let $d_1, \dots, d_n$ be distinct positive integers and $D = lcm(d_1, \dots, d_n)$. Then for any $0 \le r \le D-1$ there exist $0 \le a_i \le d_i - 1$ such that $\sum_{i=1}^n a_i/d_i = r/D(\mathrm{mod} 1)$.*

Lemma 7.7. *(a consequence of Chinese remainder theorem) Let $d_1, \dots, d_n, D$ be distinct positive integers and $\gcd(d_1, \dots, d_n, D) = 1$. Then for any $0 \le r \le D-1$ there exist $0 \le a_i \le D$ such that $\sum_{i=1}^n a_i \le D$ and $\sum_{i=1}^n a_i d_i/D = r/D(\mathrm{mod} 1)$.*

We will mainly focus on the proof of Theorem 2.8, which is the most difficult among the three theorems in Section 2. Theorem 2.5 can be proved by invoking the same technique in a simpler manner and we will sketch its proof. Theorem 2.2 can be deduced from Theorem 2.5 by several applications of Lemma 7.1.

8. Proof of Theorem 2.8

Our plan consists of four main steps

- We first obtain a long arithmetic progression (say P) by using the subset sums of a small subsequence of A.
- Next we show that (after a linear transform) one can find a reasonably short interval (say A_0) around 0 which contains many elements of A.
- Since A is l-incomplete, the sum of the subset sums of the remaining part $A \backslash (A_0 \cup P)$ with A_0 and P does not cover $\mathbf{Z}_p$. Thus the main part of A concentrates around a few points which are evenly distributed in $\mathbf{Z}_p$.
- Finally we use this structural information to deduce the statement of the theorem.

8.1. **Creating a long arithmetic progression.** Assume that A is an l-incomplete sequence with maximal multiplicity less than m. Recall that

$$f(p, m) = \lfloor (pm)^{6/13} \log^2 p \rfloor.$$

In what follows, we think of m and p as fixed and use shorthand f for $f(p, m)$. By setting c_3 large, we can assume that $|A|/f$ is large, whenever needed. If there is an element a such that $m_a(A) \ge |A| - f$ then the theorem is trivial, as we can take $A^\flat = \{b \in A, b \ne a\}$. Thus we can assume that $m(A) < |A| - f$.

Let λ be a sufficiently large constant. We execute the first step of the plan by showing the following.

Lemma 8.2. *There is a subsequence $A^\flat \subset A$ of cardinality at most f whose $l^\flat$-sums, for some integer $l^\flat \le f$, contain an arithmetic progression of length $\lambda(pm)^{12/13}/m$.*

Here we abuse the notation $A^\flat$ slightly. The current $A^\flat$ is not necessarily the $A^\flat$ in Theorem 2.8. However, as the reader will see, the latter will be the union of the current $A^\flat$ with a very small sequence of A.

Proof (Proof of Lemma 8.2) We consider three cases.

Case 1: $m > (pm)^{6/13}$.

Since $m(A) \leq |A| - f$ by assumption, we can find in A f disjoint sets $A_1, \dots, A_f$, each has exactly two different elements. Let $A' = A \setminus \cup_{i=1}^f A_i$. By the assumption $m > (pm)^{6/13}$, it follows that for each $i = 1, \dots, f$,

$$f^2|A_i| = 2f^2 > (pm)^{12/13} \gg p.$$

Thus we can apply Theorem 7.3 to the f sets $A_1, \dots, A_f$ and conclude that their sum $A_1 + \dots + A_f$ contains an arithmetic progression P of length $|P| \geq c(1)f|A_i| > c(1)(pm)^{6/13}\log^2 p$, for some positive constant $c(1)$.

On the other hand, the assumption $m > (pm)^{6/13}$ yields that $(pm)^{6/13} \geq (pm)^{12/13}/m$. Thus

$$|P| \geq \lambda(pm)^{12/13}/m$$

for any fixed constant λ. We complete by letting $A^\flat = \bigcup_{i=1} A_i$ and $l^\flat = f$.

Case 2: $p^{1/5} < m \leq (pm)^{6/13}$.

Let $A^\flat$ be an arbitrary subsequence of cardinality f in A. Since $m(A^\flat) \leq m(A) \leq m$, we can find in $A^\flat$ disjoint sets $A_1, \dots, A_m$ each of which has cardinality

$$\lfloor |A_i| = |A^\flat|/m \rfloor = \lfloor f/m \rfloor.$$

Let $k = \lfloor |A_1|/2 \rfloor$. Since $|A_i| \ll p^{1/2}$, by Lemma 7.4 we have

$$|k^* A_i| \geq c_0 |A_i|^2.$$

Next choose a set B_i of cardinality $|B_i| = c_0|A_i|^2$ from $k^* A_i$ for all i. Since

$$m^2|B_i| \geq m^2 c_0 (\frac{f}{m} - 1)^2 > c_0 m^2 \frac{f^2}{4m^2} > (pm)^{12/13} > p^{12/13+2/13} \gg p,$$

we can apply Theorem 7.3 to the m sets $B_1, \dots, B_m$ to conclude that the sumset $B_1 + \cdots + B_m$ contains an arithmetic progression P of length

$$|P| = c(1)m|B_i| = c(1)c_0 m|A_i|^2 > \frac{c(1)c_0}{4} m \frac{f}{m}^2 > \frac{\lambda(pm)^{12/13}}{m},$$

for any fixed λ, thanks to the definition of $f = f(p, m)$.

Let $l^\flat = mk$. Note that the arithmetic progression P is contained in $k^*A_1 + \cdots + k^*A_m$. But the latter sumset is a subset of $(l^\flat)^*A^\flat$. Thus the set $(l^\flat)^*A^\flat$ contains an arithmetic progression P of length $|P| \geq \lambda(pm)^{12/13}/m$.

Case 3: $m \leq p^{1/6}$.

Again let $A^\flat$ be an arbitrary subsequence of cardinality f of A. For each element a, let m_a be its multiplicity in $A^\flat$. We partition $A^\flat$ according the magnitudes of these multiplicities. For $0 \leq i \leq \log m - 1$, let n_i be the number of element a of $A^\flat$ such that $2^i \leq m_a < 2^{i+1}$. It is easy to see that $f = |A^\flat| \leq \sum_{i=0}^{\log m - 1} n_i 2^{i+1}$ (here the log has base 2), which implies that there exists an index $0 \leq i_0 \leq \log m - 1$ satisfying

$$n_{i_0} 2^{i_0+1} \geq \frac{f}{\log m}. \tag{14}$$

Let $a_1, \dots, a_{n_{i_0}}$ be elements of $A^\flat$ whose multiplicity belongs to $[2^{i_0}, 2^{i_0+1})$. Set $B_1 := \cdots = B_{2^{i_0}} := \{a_1, \dots, a_{n_{i_0}}\}$. Then the union of the B_j is a subsequence of $A^\flat$. Furthermore,

$$|B_1| = n_{i_0} \geq \frac{f}{2^{i_0+1} \log m} > \frac{(pm)^{6/13}}{m} \tag{15}$$

because $2^{i_0} \leq m \leq p$. Let $l_1 = \lfloor |B_1|/2 \rfloor$. By the assumption $m \leq p^{1/6}$ we have

$$l_1^2 |B_1| > (pm)^{18/13}/(8m^3) \gg p.$$

Theorem 7.2 applied to B_1 with $d = 1$, yields an arithmetic progression $P_1 \subset l_1^* B_1$ of length

$$|P_1| \geq c(1) l_1 |B_1| > c(1)|B_1|^2/4.$$

Since each B_i is a duplicate of B_1, we obtain 2^{i_0} duplicates $P_1, P_2, \dots, P_{2^{i_0}}$ of P_1 in $l_1^* B_1, \dots, l_1^* B_{2^{i_0}}$ respectively. Now consider $P = P_1 + \cdots + P_{2^{i_0}}$. Notice that

$$|P| = 2^{i_0}|P_1| - (2^{i_0} - 1) \geq 2^{i_0}|P_1|/2.$$

By (14) and (15), we have

$$|P| \geq 2^{i_0} c(1)|B_1|^2/8 = c(1)2^{i_0} n_{i_0}|B_1|/8 \geq$$

$$\geq (c(1)/8)(f/(2\log m))((pm)^{6/13}/m) > \lambda(pm)^{12/13}/m$$

for any fixed λ. Now observe that

$$P \subset l_1^* B_1 + \cdots + l_1^* B_{2^{i_0}} \subset (2^{i_0} l_1)^* A^\flat.$$

Thus by setting $l^\flat = 2^{i_0} l_1$ we conclude that the collection of $l^\flat$-sums of $A^\flat$ contains an arithmetic progression of length $\lambda(pm)^{12/13}/m$.

■

By a dilation of A with some nonzero $b' \in \mathbf{Z}_p$, we can assume that the arithmetic progression P obtained by Lemma 8.2 is an interval, $P = [p_0, p_0 + L]$ for some residue p_0 and $L \geq \lambda(pm)^{12/13}/m$.

8.3. **Dense subsequence around zero.** Let $Q = \lfloor (pm)^{3/13} \rfloor$ and $A' = A \backslash A^\flat$.

Lemma 8.4. *There exists a residue $c' \in \mathbf{Z}_p$ such that $(A'+c') \cap [-p/(2Q^2), p/(2Q^2)]$ contains a subsequence of cardinality $3Q$.*

Proof (Proof of Lemma 8.4) Call a pair (x, y) of $\mathbf{Z}_p \times \mathbf{Z}_p$ *nice* if

$$p/Q^2 < \|y - x\| < L.$$

Note that if (x, y) is a nice pair then $x + P \cap y + P \neq \emptyset$ and $x + P \cup y + P$ is an interval of length

$$|x + P \cup y + P| \geq \min(|P| + p/Q^2, p). \tag{16}$$

Assume that $B = \{x_1, y_1, \ldots, x_r, y_r\}$ is a (maximal) sequence of nice pairs in A' (this means that there is no more nice pair left in $A' \backslash B$). We are going to show that $r < Q^2$. Assume otherwise. By (16),

$$P' = \bigcup_{z_i \in \{x_i, y_i\}, 1 \le i \le Q^2} z_1 + \cdots + z_{Q^2} + P = \mathbf{Z}_p.$$

On the other hand, by the assumption of the Theorem,

$$\left| A' \backslash \bigcup_{i=1}^{Q^2} \{x_i, y_i\} \right| = |A| - |A^\flat| - 2Q^2 \ge |A| - 2f \ge l.$$

So we are able to choose a subsequence C in $A' \backslash B$ of cardinality $l - l^\flat - Q^2$.

But then

$$\mathbf{Z}_p = P' + \sum_{c \in C} c \subset l^* A,$$

which means that A is l-complete, impossible. Thus $r < Q^2$.

We define a new $A^\flat$ by taking the union of the existing one with B. The bound on $|B|$ shows that the new $A^\flat$ is still of cardinality $O((pm)^{6/13} \log^2 p)$. We keep using the notation A' for $A \backslash A^\flat$, but the reader should keep in mind that the new A' has no nice pair as we have discarded B. This implies that there are intervals $A_0, \ldots, A_n$ of $\mathbf{Z}_p$ such that $|A_i| \le p/Q^2$ and $\min\{\|x - y\| \Big| x \in A_i, y \in A_j\} \ge L$ for any $i \ne j$ and the union $\cup_{i=1}^n A_i$ contains A'. It then follows that

$$n + 1 \le p/L.$$

But by pigeon-hole principle there is an interval, say A_0, which contains at least $|A'|/(n+1)$ elements of A'. Recall that the length of A_0 is less than p/Q^2 and

$$|A'|/(n+1) \ge |A'|L/p > (pm)^{6/13 + 12/13}/(pm) = (pm)^{5/13} > 3Q.$$

■

We infer from Lemma 8.4 that, by an appropriate translation, one can find a reasonably short interval around 0 which contains many elements of A. (Notice that the translation shifts P to another interval of the same length). We will work with this translated image of A.

8.5. **Distribution of the elements of A.** Let I_0 and J_0 be two disjoint subsequences of $A' \cap [-p/(2Q^2), p/(2Q^2)]$ of cardinality Q and $2Q$ respectively.

Let $A'' = A' \backslash (I_0 \cup J_0)$. We show that almost all elements of A''(and thus almost all elements of A) concentrate around a few points which are regularly distributed in $\mathbf{Z}_p$.

Lemma 8.6. *There is a subsequence $A''' \subset A''$ and an integer D such that*

- $$|A'''| \leq 2(pm)^{6/13},$$
- $$D \leq (pm)^{1/13},$$
- *for any $a \in A'' \backslash A'''$ there is an integer $0 \leq h \leq D-1$ satisfying*

$$|a - \frac{hp}{D}| \leq \frac{p}{Q}.$$

We postpone the proof of Lemma 8.6 until Proposition 8.6.2.

Let a be any element of A''. Then by Dirichlet's theorem, there is a pair of positive integers i and d satisfying $1 \leq d \leq Q$ and $\gcd(i,d) = 1$ such that

$$|a - \frac{ip}{d}| \leq \frac{p}{dQ}.$$

Next let

$$X_d = \{a \in A'' : |a - \frac{ip}{d}| \leq \frac{p}{dQ}, 1 \leq i \leq d, 1 \leq d \leq Q, \gcd(i,d) = 1\}.$$

Call the index d *rich* if $|X_d| \geq 2d$. Let us denote the *rich* indices by

$$d_1 < d_2 < \cdots < d_s.$$

We will collect some facts about the *rich* indices.

Proposition 8.6.1.

$$d_j \leq (pm)^{1/13}.$$

Proof (Proof of Proposition 8.6.1) Let $X'_{d_j} = \{a_1, \ldots, a_{d_j}\}$ be any subsequence of d_j elements of X_{d_j}. By Lemma 7.5, for $0 \leq i \leq d_j - 1$ there exists $A^i_{d_j} \subset X'_{d_j}$ such that

$$|\sum_{a\in A^i_{d_j}} a - \frac{ip}{d_j}| \le \frac{p}{Q}.$$

Choose a sequence $B^i_{d_j} \subset I_0$ such that $|B^i_{d_j}| = d_j - |A^i_{d_j}|$. By the definition of I_0 we have

$$\sum_{b\in B^i_{d_j}} |b| \le |B^i_{d_j}|p/(2Q^2) \le d_j p/(2Q^2) \le p/2Q.$$

Thus

$$|\sum_{a\in A^i_{d_j}} a + \sum_{b\in B^i_{d_j}} b - \frac{ip}{d_j}| \le 2p/Q. \tag{17}$$

By definition, $\sum_{a\in A^i_{d_j}} a + \sum_{b\in B^i_{d_j}} b \subset d^*_j(X_{d_j}\cup I_0)$. Thus the inequality (17) implies that $d^*_j(X'_{d_j}\cup I_0)$ forms a K-net of $\mathbf{Z}_p$ with $K \le p/d_j + 4p/Q$.

Now we claim that $K > L$. Seeking a contradiction, suppose that $K \le L$. Then

$$d^*_j(X'_{d_j}\cup I_0) + P = \mathbf{Z}_p. \tag{18}$$

Because the cardinality of $A''\backslash X'_{d_j}$ is larger than l,

$$|A''\backslash X'_{d_j}| = |A'| - |I_0| - |J_0| - |X'_{d_j}| \ge |A| - |A^\flat| - 4Q \ge l,$$

we can choose $C \subset A''\backslash X'_{d_j}$ of cardinality $|C| = l - d_j - l^\flat$. Next, by (18) we have

$$\mathbf{Z}_p = d^*_j(X'_{d_j}\cup I_0) + P = d^*_j(X'_{d_j}\cup I_0) + P + \sum_{c\in C} c \subset l^* A.$$

Thus A is l-complete, a contradiction.

Observe that beside the inequality $K > L$ we also have

$$L \gg p/Q \text{ and } L \ge \lambda(pm)^{12/13}/m \ge 2(pm)^{12/13}/m.$$

Thus

$$d_j \leq 2p/L \leq (pm)^{1/13}.$$

■

Proposition 8.6.1, in particular, implies that the number of *rich* indices is also small,

$$s \leq (pm)^{1/13}.$$

In the following, we prove a stronger fact.

Proposition 8.6.2. *Let $D = lcm(d_1, \dots, d_s)$. Then we have*

$$D \leq (pm)^{1/13}.$$

Proof (Proof of Proposition 8.6.2) For each $1 \leq i \leq s$ let X'_{d_i} be a subsequence of cardinality d_i in X_{d_i}. We claim that $(\sum_{i=1}^s d_i)^*(\bigcup_{i=1}^s X'_{d_i} \bigcup I_0)$ is a K-net in $\mathbf{Z}_p$ with

$$K \leq p/D + 4sp/Q.$$

To prove the claim, first let r be any integer between 0 and $D-1$. By Lemma 7.6 there exist $0 \leq a_i \leq d_i - 1$ such that $\sum_{i=1}^s a_i p/d_i = rp/D$.

Next choose $A^r_{d_i} \subset X'_{d_i}$ such that $|\sum_{a \in A^r_{d_i}} a - a_i p/d_i| \leq p/Q$. Summing these inequalities over $1 \leq i \leq s$ we obtain

$$|\sum_{a \in \bigcup_{i=1}^s A^r_{d_i}} a - rp/D| \leq sp/Q. \tag{19}$$

In addition, because

$$\sum_{i=1}^s d_i \leq \lfloor s(pm)^{1/9} \rfloor \leq \lfloor (pm)^{2/9} \rfloor = Q = |I_0|,$$

there are disjoint subsequences $B^r_{d_1}, \dots, B^r_{d_s}$ of I_0 such that $|B^r_{d_i}| = d_j - |A^r_{d_j}|$. And by the definotion of I_0 we have

$$\sum_{b\in\bigcup_{i=1}^{s} B^r_{d_i}} |b| \leq (\sum_{i=1}^{s} d_i)p/(2Q^2) \leq Qp/(2Q^2) = p/2Q. \tag{20}$$

Putting the estimates (19),(20) together to obtain

$$|\sum_{a\in\cup A^r_{d_i}} a + \sum_{b\in\cup B^r_{d_i}} b - rp/D| \leq sp/Q + p/2Q \leq 2sp/Q. \tag{21}$$

Notice that $\sum_{i=1}^{s}(|A^r_{d_i}| + |B^r_{d_i}|) = \sum_{i=1}^{s} d_i$. Point (21) concludes the claim.

We now claim that $K > L$. Assume otherwise. Then

$$(\sum_{i=1}^{s} d_i)^*(\bigcup_{i=1}^{s} X'_{d_i} \bigcup I_0) + P = \mathbf{Z}_p. \tag{22}$$

But

$$|A''\backslash \bigcup_{i=1}^{s} X'_{d_i}| = |A'| - |I_0| - |J_0| - \sum_{j=1}^{s} d_j \geq |A| - |A^\flat| - 4Q \geq l,$$

there exists a subsequence C in $A''\backslash \bigcup_{i=1}^{s} X'_{d_i}$ of cardinality $|C| = l - \sum_{j=1}^{s} d_j - l^\flat$.

Adding elements of C to (22) we achieve

$$\mathbf{Z}_p = d^*_j(X'_{d_j} \cup I_0) + P = d^*_j(X'_{d_j} \cup I_0) + P + \sum_{c\in C} c.$$

The last sum of the equality above is a subset of l^*A. Thus A is l-complete, a contradiction.

In conclusion we have just proved that $(\sum_{i=1}^{s} d_i)^*(\bigcup_{i=1}^{s} X'_{d_i} \bigcup I_0)$ is a K-net in $\mathbf{Z}_p$ with

$$L \leq K \leq p/D + 4sp/Q.$$

In particular,

$$L \leq p/D + 4sp/Q,$$

$$\lambda(pm)^{12/13}/m - 4p(pm)^{1/13}/(pm)^{3/13} \leq p/D.$$

Hence (because $\lambda \geq 2$)

$$D \leq (pm)^{1/13}.$$

■

For brevity set $t := \sum_{i=1}^{s} d_i$, $H := \bigcup_{i=1}^{s} X'_{d_i} \cup I_0$ and

$$T := t^*H = (\sum_{i=1}^{s} d_i)^*(\bigcup_{i=1}^{s} X'_{d_i} \bigcup I_0).$$

Recall that T is a K-net with $K \leq p/D + 4sp/Q$. We remove H from A'' and record the set T for latter use. Let us now prove Lemma 8.6 by putting everything together.

Proof (Proof of Lemma 8.6) Call an element a of A'' *single* if $a \notin \bigcup_{j=1}^{s} X_{d_j}$. By Dirichlet's theorem, any single point is an element of some X_d where d is not rich. But $|X_d| < 2d$ if d is not rich. Thus by double counting, the number of single points, denoted by A''', is bounded by

$$|A'''| \leq \sum_{d \leq Q} (2d - 1) < 2Q^2 = 2(pm)^{6/13}.$$

Let a be any element of $A''\backslash A'''$, then $a \in X_{d_j}$ for some rich d_j. Put $h = iD/d_j$. Then by definition

$$|a - \frac{hp}{D}| = |a - \frac{ip}{d_j}| \leq \frac{p}{d_j} \leq \frac{p}{Q}.$$

Furthermore, by Proposition 8.6.1,

$$D \leq (pm)^{1/13}.$$

■

Add A''' to $A^\flat$, the cardinality of $A^\flat$ is still $O((pm)^{6/13} \log^2 p)$. For $1 \leq h \leq D$ we let

$$J_h = \{a | a \in A'', \frac{hp}{D} - \frac{p}{Q} \leq a \leq \frac{hp}{D} + \frac{p}{Q}\}$$

and

$$R_h = \{a - \frac{hp}{D} | a \in J_h\}.$$

By throwing away a small number ($\leq sD \leq (pm)^{2/13}$) of elements to $A^\flat$, we can assume that the cardinalities of R_h, $1 \leq h \leq s$, are divisible by D. Note that the sum of any D elements of R_h is an integer. We denote by R the sequence of all reduced elements,

$$R = \bigcup_{h=1}^{s} R_h.$$

Hence for any $r \in R$ we have $|r| \leq p/Q$.

Let us summarize what we attain up to this step. Up to a proper dilation (with b') and translation (with c'), there is a partition of A, $A = A^\flat \cup J_0 \cup H \cup A''$ such that

- $|A^\flat| = O((pm)^{6/13} \log^2 p)$ and ${l^\flat}^* A^\flat$ contains an interval $P = [a, a + L]$ of length $L = \lambda(pm)^{12/13}/m$ with some $l^\flat \leq (pm)^{6/13} \log^2 p$.
- $|H| \leq 2(pm)^{3/13}$ and $t^* H$ contains a $p/D + 4sp/Q$-net (named T).
- $|J_0| = 2Q$ and $J_0 \subset [-p/(2Q^2), p/(2Q^2)]$.

8.7. **Completing the proof of Theorem 2.8.** Set

$$l_0 := l - l^\flat - t.$$

Since the elements of R are small, the set $l_0^* R (\subset \mathbf{Q})$ is dense in the interval in which it is contained. We show that $l_0^* R \cap \mathbf{Z}$ is also dense in this interval. Suppose for the moment that this interval is longer than $p/D + 4sp/Q$. Then $(l_0^* R \cap \mathbf{Z}) + P$ contains another interval of length $p/D + 4sp/Q$ (in $\mathbf{Z}$, as P is viewed as an interval of $\mathbf{Z}$). We then infer that $l_0^* A'' \cap \mathbf{Z} + P$ contains an interval of that same length in $\mathbf{Z}_p$. So

$$Z_p = l_0^* A'' + P + T \subset l^* A.$$

Which is impossible. We conclude that $l_0^* R$ must be supported by a short interval of $\mathbf{Q}$. In the following we explain the argument in detail.

Set

$$l_2 := l_0 - D^2 \text{ and } l_1 := l_2 - Q = l_2 - \lfloor (pm)^{1/3} \rfloor.$$

Then

$$l_2 > l_1 \geq l - 2(pm)^{6/13} \log^2 p.$$

Viewing R as a subsequence of $\mathbf{Q}$ in $[-p/Q, p/Q]$, our goal is to establish the following.

Lemma 8.8. *Let* $m_1 = \min_{l_1 \leq l' \leq l_2} (\min l'^* R)$ *and* $m_2 = \max_{l_1 \leq l' \leq l_2} (\max l'^* R)$. *Then we have*

$$m_2 - m_1 < p/D.$$

Proof (Sketch of proof of Lemma 8.8) Add several (at most D^2) elements of R to the representations of m_1 and m_2 respectively to make the number of summands from each class R_h divisible by D. We obtain m_1', m_2' with the following properties.

- $m_i' \in l_i'^* R$, where $l_1 \leq l_i' \leq l_2 + D^2$.
- $|m_i' - m_i| \leq D^2 p/Q$. (Because to create m_i' we added at most D^2 elements from R, whose element is bounded by p/Q.)
- $m_1', m_2' \in \mathbf{Z}$. (As the sum of any D elements of R_h is an integer.)

By the properties above, we are done with the Lemma if $m_2' - m_1' \leq p/D - 2D^2 p/Q$. Seeking for contradiction, suppose that

$$m_2' - m_1' > p/D - 2D^2 p/Q. \tag{23}$$

Let $U_1, U_2 \subset R$ be sequences of cardinality l_1', l_2' respectively such that

$$\sum_{u \in U_i} u = m_i'.$$

The reader should find it straightforward to construct sequences $V_1, V_2, \ldots, V_n$ in R such that all the following properties hold.

- $$V_1 = U_1, V_n = U_2.$$

- $$\min\{l_1', l_2'\} \le |V_i| \le \max\{l_1', l_2'\} \text{ for } 1 \le i \le n.$$

- $$|V_{i+1} \backslash V_i| \le D. \tag{24}$$

- For any $1 \le h \le s$ the cardinality of $V_i \cap R_h$ is divisible by D, i.e.,

$$D||V_i \cap R_h| \text{ for } 1 \le h \le s. \tag{25}$$

Notice that condition (25) guarantees that $\sum_{v \in V_i} v$ is an integer, and (24) implies that

$$|\sum_{v \in V_{i+1}} v - \sum_{v \in V_i} v| \le Dp/Q \text{ for } 1 \le i \le n.$$

Thus the set $\{\sum_{v \in V_i} v | i = 1, \dots, n\}$ is a pD/Q-net (of $\mathbf{Z}$) in the interval $[m_1', m_2']$. Recall that

$$|J_0| = 2Q > Q + D^2 = l_0 - l_1 \ge l_0 - |V_i|,$$

i.e. for each $1 \le i \le n$ one can choose a sequence W_i of J_0 of cardinality $l_0 - |V_i|$ (W_i's are not necessarily disjoint). Denote $V_i \cup W_i$ by X_i. Then we have $|X_i| = l_0$ and

$$|\sum_{x \in X_i} x - \sum_{v \in V_i} v| \le (l_0 - |V_i|)p/Q^2 \le (l_0 - l_1)p/Q^2 \le p/Q. \tag{26}$$

Because $\{\sum_{v \in V_i} v | i = 1, \dots, n\}$ is a Dp/Q-net in $[m_1', m_2']$, we have

$$[m_1', m_2'] \subset \{\sum_{v \in V_i} v | i = 1, \dots, n\} + [0, Dp/Q] (\mathrm{mod} p);$$

and it follows from (26) that

$$[m_1' + p/Q, m_2' - p/Q] \subset \{\sum_{x \in X_i} x | i = 1, \dots, n\} + [0, 2Dp/Q]. \tag{27}$$

We proceed by claiming the following.

Claim 8.8.1. *Suppose that (23) holds. Then the set*

$$\{\sum_{x\in X_i} x + T | i = 1, \dots, n\}$$

is a $8D^2p/Q$-net of $\mathbf{Z}_p$.

Proof (Proof of Claim 8.8.1) Obtain from (27) that

$$[m_1' + p/Q, m_2' + 7D^2p/Q] \subset \{\sum_{x\in X_i} x | i = 1, \dots, n\} + [0, 8D^2p/Q].$$

Consequently,

$$[m_1' + p/Q, m_2' + 7D^2p/Q] + T \subset \{\sum_{x\in X_i} x | i = 1, \dots, n\} + [0, 8D^2p/Q] + T. \tag{28}$$

Notice that because T is a $p/D + 4sp/Q$-net of $\mathbf{Z}_p$, and by (23) that

$$m_2' + 7D^2p/Q - m_1' - p/Q \geq p/D + 4D^2p/Q > p/D + 4sp/Q,$$

we have

$$\mathbf{Z}_p = [m_1' + p/Q, m_2' + 7D^2p/Q] + T.$$

Together with (28) this gives

$$(\{\sum_{x\in X_i} x | i = 1, \dots, n\} + T) + [0, 8D^2p/Q] = \mathbf{Z}_p.$$

■

To finish the proof of Lemma 8.8 one observes that

$$L \geq \lambda p/(pm)^{1/13} \geq 8p/(pm)^{1/13} \geq 8D^2p/Q.$$

Thus Claim 8.8.1 would give

$$\{\sum_{x\in X_i} x + T | i = 1, \dots, n\} + P = \mathbf{Z}_p.$$

However, $\{\sum_{x\in X_i} x + T | i = 1, \dots, n\} + P \subset l^*A$. Hence A is $l-$complete, a contradiction. As a consequence, (23) can not hold.

■

Now we close the proof of Theorem 2.8. Dilate the whole set A with D. By viewing $D \cdot A''$ as a sequence of $\mathbf{Z}$ in $[-Dp/Q, Dp/Q]$, one sees that

$$\max_{l_1 \le l' \le l_2} \max(l'^*(D \cdot A'')) - \min_{l_1 \le l' \le l_2} \min(l'^*(D \cdot A'')) = Dm_2 - Dm_1 < p.$$

Thus if Φ denotes the linear map $b' \cdot X + c'$ then the statement of Theorem 2.8 holds for $A^\flat($ of the statement $) := \Phi^{-1}(A^\flat \cup J_0 \cup H)$ and $b := Db', c := Dc'$.

9. Sketch of proof of Theorem 2.5

Theorem 2.5 can be verified by following the proof of Theorem 2.8 above. In fact, the situation here is somewhat simpler. Since the subset sums in Theorem 2.5 do not need to have a fixed number of summands, we do not have to consider I_0 and J_0.

Keep the same notation as in the proof of Theorem 2.8. As an analogue of Lemma 8.8, we can establish the following lemma.

Lemma 9.1. *Let $m_1 = \min(S_R)$ and $m_2 = \max(S_R)$. Then we have*

$$m_2 - m_1 < p/D.$$

Then by dilating the whole set A with D, one obtains Theorem 2.5.

10. Proof of Theorem 2.2

By Theorem 2.5 there exists a non-zero residue b and a small set $A^\flat \subset A$ of cardinality at most $c_2 f(p, m)$ such that

$$\sum_{a \in b\cdot(A \backslash A^\flat)} \|a\| < p. \tag{29}$$

Consider the sequence of positive and negative elements of $b \cdot (A \backslash A^\flat)$,

$$A^+ := b \cdot (A \backslash A^\flat) \cap [1, (p-1)/2] \text{ and } A^- := b \cdot (A \backslash A^\flat) \cap [-(p-1)/2, -1].$$

We shall prove the following.

Lemma 10.1. *There exists an absolute constant β such that either $|A^+| \le \beta f(p,m)$ or $|A^-| \le \beta f(p,m)$.*

Assume for the moment, and without loss of generality, that $|A^-| \le \beta f(p,m)$. Then Theorem 2.2 holds for $A^\flat$ (of the statement) $:= A^\flat \cup b^{-1} \cdot A^-$ and $c_1 := c_2 + \beta$.

Proof (Proof of Lemma 10.1) Assume otherwise that

$$|A^+|, |A^-| \ge \beta f(p,m) \text{ for large positive constant } \beta. \tag{30}$$

Note that from (29) we have

$$\sum_{a \in A^+} a < p, \text{ and } \sum_{a \in A^-} |a| < p. \tag{31}$$

Set $q := \lfloor p/f(p,m) \rfloor$. Let $B^+ := A^+ \cap [1,q]$ and $B^- := A^- \cap [-1,-q]$ respectively.

We infer from (31) that

$$|B^+| \ge (\beta - 1) f(p,m) \text{ and } |B^-| \ge (\beta - 1) f(p,m).$$

Viewing B^+ and B^- as sequence of integers in $[-q,q]$, we then reach a contradiction with the zero-sum-freeness property of A by showing that there exist some elements of B^+ and B^- whose sum is 0.

Consider the following two cases.

Case 1: $m \ge p^{4/9}$.

By pigeon-hole principle there are two elements $a^+ \in B^+, a^- \in B^-$ whose multiplicities (denoted by m_{a^+}, m_{a^-} respectively) are large.

$$m_{a^+} \ge |B^+|/q \ge (\beta - 1) f(p,m)/(p/f(p,m)) > (pm)^{12/13}/p > p/(pm)^{6/13} \ge q,$$

and similarly

$$m_{a^-} > q.$$

Note that $0 \le |a^-|, a^+ \le q$. Thus $|a^-| < m_{a^-}$ and $a^+ < m_{a^+}$, which yield

$$0 = |a^-|a^+ + a^+a^- \in S_{B^+} + S_{B^-} \subset S_A, \text{contradiction.}$$

Case 2: $1 \le m < p^{4/9}$.

Without loss of generality assume that

$$|\sum_{a\in B^-} a| \ge \sum_{a\in B^+} a. \tag{32}$$

Fix any subset X of B^+ of cardinality $|X| = f(p,m)/\max(\log p, m)$.

First, one sees that

$$(f(p,m)/\log p)^2 \gg p/f(pm) = q,$$

and

$$(f(p,m)/m)^2 \gg p/f(p,m) = q.$$

Thus, Theorem 7.1 applied to X (with $l = \lfloor |X|/2 \rfloor$ and $d = 1$) yields an arithmetic progression $P = \{a, a+d, \dots, a+Ld\}$ of length $L \ge c(1)|X|^2/2$.

Note that $P \subset S_X \subset [1, |X|q]$, thus the difference d of P is bounded, i.e.,

$$d \le |X|q/L \le 2q/(c(1)|X|) \ll (pm)^{1/13}/\log p. \tag{33}$$

Next, view $(B^+\backslash X) \cup B^-$ as a sequence of residues modulo d. We throw away residues of multiplicity less than d. Let W be the sequence of thrown elements. So obviously,

$$|W| \le d^2 \le (pm)^{2/9}/\log^2 p.$$

We consider two subcases.

Subcase 2.1: There exists a nontrivial divisor d_1 of d which divides all the remaining residues.

Set

$$B_1^+ := \{\frac{b}{d_1} | b \in B^+ \backslash (X \cup W)\} \text{ and } B_1^- := \{\frac{b}{d_1} | b \in B^- \backslash W\}.$$

Observe that

$$|B_1^+|, |B_1^-| \geq (\beta - 1) f(pm) - 2f(p,m)/\log p.$$

Also, $B_1^+, B_1^- \subset [-q_1, q_1] := [-\lfloor q/d_1 \rfloor, \lfloor q/d_1 \rfloor]$.

Viewing B_1^+ and B_1^- as B^+ and B^-, we reconsider **Case 1** and **Case 2**. Thus either a contradiction is obtained or we get B_2^+ and B_2^- whose elements are divisible by some integer $d_2 \geq 2$. Repeat the process until we get a contradiction thanks to **Case 1** or **Subcase 2.2** as follows. (Notice that the process stops after at most $\log p$ steps because q_i decreases by a factor of at least 2 with each step, while $|B_i^+|, |B_i^-| \geq (\beta - 2) f(p,m)$ always.)

Subcase 2.2: There does not exist such divisor of d. Thus the residues are mutually co-prime with d.

By Lemma 7.7 there exist $x_1, \dots, x_u \in X \backslash X, y_1, \dots, y_v \in B^-$ with $u + v \leq d$ and

$$a = -\sum_{i=1}^{u} x_i - \sum_{j=1}^{v} y_j (\mathrm{mod} d). \tag{34}$$

Note that

$$\sum_{i=1}^{u} x_i + |\sum_{j=1}^{v} y_j| \leq dq \ll dL. \tag{35}$$

We consider the following two possibilities.

Subcase 2.2.1: $|\sum_{j=1}^{v} y_i| - \sum_{i=1}^{u} x_i \geq a$.

Then by (34) and (35) we get

$$|\sum_{j=1}^{v} y_i| - \sum_{i=1}^{u} x_i \in P.$$

Thus

$$|\sum_{j=1}^{v} y_i| \in \sum_{i=1}^{u} x_i + S_X,$$

and so

$$0 \in \sum_{j=1}^{v} y_i + \sum_{i=1}^{u} x_i + S_X \subset S_{X\cup B^-} \subset S_A, \text{ contracdition .}$$

Subcase 2.2.2: $|\sum_{j=1}^{v} y_i| - \sum_{i=1}^{u} x_i < a$.

Then let $Y_0 =: \{y_1, \dots, y_v\}$. By Lemma 7.5 one can find $Y_0' \subset B^-\backslash Y_0$ such that $|Y_0'| \le d$ and $d|\sum_{y\in Y_0'} y$.

Set $Y_1 := Y_0 \cup Y_0'$. If $|\sum_{y\in Y_1} y| - \sum_{i=1}^{u} x_i$ is still less than a then we again use Lemma 7.5 to find $Y_1' \subset B^-\backslash Y_1$ such that Y_1' has the same property as Y_0'. We next increase Y_1 by $Y_2 := Y_1 \cup Y_1'$. Repeat the process until we get $Y_N \subset B^-$ such that

$$|\sum_{y\in Y_{N-1}} y| - \sum_{i=1}^{u} x_i < a \text{ and } |\sum_{y\in Y_N} y| - \sum_{i=1}^{u} x_i \ge a.$$

Notice that by (31) we have

$$\sum_{i=1}^{u} x_i + a + Ld \le \sum_{a\in B^+} a \le |\sum_{y\in B^-} y|.$$

In addition, since $q \ll L$,

$$\sum_{i=1}^{u} x_i + a \le |\sum_{y\in Y} y| \tag{36}$$

for any $Y \subset B^-$ with cardinality $|Y| \ge |B^-| - d$. Lemma 7.5 and (36) thus ensure the existence of N above.

In sum,

$$0 \le |\sum_{y\in Y_N} y| - \sum_{i=1}^{u} x_i - a \le Ld$$

and

$$d|| \sum_{y \in Y_N} y| - \sum_{i=1}^{u} x_i - a.$$

It then follows that

$$| \sum_{y \in Y_N} y| - \sum_{i=1}^{u} x_i \in P.$$

$$0 \in \sum_{y \in Y_N} y + \sum_{i=1}^{u} x_i + S_{X'} \subset S_{X \cup Y} \subset S_A, \text{ contradiction .}$$

■

References

[1] G. E. Andrews, *The theory of partitions.* Cambridge university press, 1998.

[2] A. Bialostocki and P. Dierker, *On Erdős-GinzburgiZiv theorem and the Ramsey number for stars and matchings.* Discrete Mathematics, 110, (1992), 1-8.

[3] P. Erdős, A. Ginzburg and A. Ziv, *Theorem in the additive number theory.* Bull. Res. Council Israel 10F (1961), 41-43.

[4] W. D. Gao, A. Panigrahi and R. Thangadurai, *On the structure of p-zero-sum-free sequences and its application to a variant of Erdős-Ginzburg-Ziv theorem.* Proc. Indian Acad. Sci. Vol. 115, No. 1 (2005), 67-77.

[5] H. H. Nguyen, E. Szemerédi and V. H. Vu, *Subset sums modulo a prime*, to appear, Acta Arithmetica.

[6] J. E. Olson, *An addition theorem modulo p*, J. Combinatorial Theory 5(1968), 45-52.

[7] J. E. Olson, *Sums of sets of group elements.* Acta Arithmetica, 28 (1975), 147-156.

[8] J. W. Sun, *List of publications on restricted sumsets.* 2005.

[9] Endre Szemerédi and Van H. Vu , *Long arithmetic progression in sumsets and the number of x-free sets.* Proceeding of London Math Society, 90(2005) 273-296.

[10] E. Szemerédi and V. H. Vu , *Long arithmetic progressions in sumsets: Thresholds and Bounds.* Journal of the A.M.S, 19 (2006), no 1, 119-169.

Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA

E-mail address: hoi@math.rutgers.edu

Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA

E-mail address: vanvu@math.rutgers.edu