SUBSET SUMS MODULO A PRIME

HOI H. NGUYEN, ENDRE SZEMERÉDI, AND VAN H. VU

Abstract. Let $\mathbf{Z}_p$ be the finite field of prime order p and A be a subset of $\mathbf{Z}_p$. We prove several sharp results about the following two basic questions:

(1) When can one represent zero as a sum of distinct elements of A ?

(2) When can one represent every element of $\mathbf{Z}_p$ as a sum of distinct elements of A ?

1. Introduction

Let A be an additive group and A be a subset of A. We denote by $\sum(A)$ the collection of subset sums of A:

$$\sum(A) = \{\sum_{x\in B} x | B \subset A, |B| < \infty\}.$$

The following two questions are among the most popular questions in additive combinatorics

Question 1.1. *When* $0 \in \sum(A)$ *?*

Question 1.2. *When* $\sum(A) = G$ *?*

If $\sum(A)$ does not contain the zero element, we say that A is *zero-sum-free*. If $\sum(A) = G$ ($\sum(A) \neq G$), then we say that A is *complete (incomplete)*.

In this paper, we focus on the case $G = \mathbf{Z}_p$, the cyclic group of order p, where p is a large prime. The asymptotic notation will be used under the

2000 Mathematics Subject Classification: 11P99.

Key words and phrases: subset sums, sumsets, incomplete, zero-sum-free.

V. Vu is an A. Sloan Fellow and is supported by an NSF Career Grant.

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.) All logarithms have natural base and $[a, b]$ denotes the set of integers between a and b.

1.3. **A sharp bound on the maximum cardinality of a zero-sum-free set.** How big can a zero-sum-free set be ? This question was raised by Erdős and Heilbronn [4] in 1964. In [8], Szemerédi proved the following.

Theorem 1.4. *There is a positive constant c such that the following holds. If $A \subset \mathbf{Z}_p$ and $|A| \geq cp^{1/2}$, then $0 \in \sum(A)$.*

A result of Olson [6] implies that one can set $c = 2$. More than a quarter of century later, Hamindoune and Zémor [7] showed that one can set $c = \sqrt{2} + o(1)$, which is asymptotically tight.

Theorem 1.5. *If $A \subset \mathbf{Z}_p$ and $|A| \geq (2p)^{1/2} + 5\log p$, then $0 \in \sum(A)$.*

Our first result removes the logarithmic term in Theorem 1.5, giving the best possible bound (for all sufficiently large p). Let $n(p)$ denote the largest integer such that $\sum_{i=1}^{n-1} i < p$.

Theorem 1.6. *There is a constant C such that the following holds for all prime $p \geq C$.*

- *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 $0 \in \sum(A)$.*
- *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 $0 \in \sum(A)$. Furthermore, up to a dilation, the only 0-sum-free set with $n(p)$ elements is $\{-2, 1, 3, 4, \dots, n(p)\}$.*

To see that the bound in the first case is sharp, consider $A = \{1, 2, \dots, n(p) - 1\}$.

1.7. **The structure of zero-sum-free sets with cardinality closed to maximum.** Theorem 1.6 does not provide information about zero-sum-free sets of size slightly smaller than $n(p)$. The archetypical example for a zero-sum-free set is a set whose sum of elements (as positive integers between 1 and $p-1$) is less than p. The general phenomenon we would like to support here is that a zero-sum-free set with sufficiently large cardinality should be close to such a set. In [1], Deshouillers showed the following.

Theorem 1.8. *Let A be a zero-sum-free subset of $\mathbf{Z}_p$ of size at least $p^{1/2}$. Then there is some non-zero element $b \in \mathbf{Z}_p$ such that*

$$\sum_{a \in bA, a < p/2} \|a\| \le p + O(p^{3/4} \log p)$$

and

$$\sum_{a \in bA, a > p/2} \|a\| = O(p^{3/4} \log p).$$

The main issue here is the magnitude of the error term. In the same paper, there is a construction of a zero-sum-free set with $cp^{1/2}$ elements ($c > 1$) where

$$\sum_{a \in bA, a < p/2} \|a\| = p + \Omega(p^{1/2})$$

and

$$\sum_{a \in bA, a > p/2} \|a\| = \Omega(p^{1/2}).$$

It is conjectured [1] that $p^{1/2}$ is the right order of magnitude of the error term. Here we confirm this conjecture, assuming that $|A|$ is sufficiently close to the upper bound.

Theorem 1.9. *Let A be a zero-sum-free subset of $\mathbf{Z}_p$ of size at least $.99n(p)$. Then there is some non-zero element $b \in \mathbf{Z}_p$ such that*

$$\sum_{a \in bA, a < p/2} \|a\| \le p + O(p^{1/2})$$

and

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

The constant .99 is adhoc and can be improved. However, we do not elaborate on this point.

1.10. **Complete sets.** All questions concerning zero-sum-free sets are also natural for incomplete sets. Here is a well-known result of Olson [6].

Theorem 1.11. *Let A be a subset of $\mathbf{Z}_p$ of more than $(4p-3)^{1/2}$ elements, then A is complete.*

Olson's bound is essentially sharp. To see this, observe that if the sum of the norms of the elements of A is less than p, then A is incomplete. Let $m(p)$ be the largest cardinality of a small set. One can easily verify that $m(p) = 2p^{1/2} + O(1)$. We now want to study the structure of incomplete sets of size close to $2p^{1/2}$. Deshouillers and Freiman [3] proved the following.

Theorem 1.12. *Let A be an incomplete subset of $\mathbf{Z}_p$ of size at least $(2p)^{1/2}$. Then there is some non-zero element $b \in \mathbf{Z}_p$ such that*

$$\sum_{a \in bA} \|a\| \leq p + O(p^{3/4} \log p).$$

Similarly to the situation with Theorem 1.8, it is conjectured that the right error term has order $p^{1/2}$ (see [2] for a construction that matches this bound from below). We establish this conjecture for sufficiently large A.

Theorem 1.13. *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*

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

Added in proof. While this paper was written, Deshouillers informed us that he and Prakash have obtained a result similar to Theorem 1.6.

2. Main lemmas

The main tools in our proofs are the following results from [9].

Theorem 2.1. *Let A be a zero-free-sum subset of $\mathbf{Z}_p$. Then we can partition A into two disjoint sets A' and A'' where*

- *A' has negligible cardinality: $|A'| = O(p^{1/2}/\log^2 p)$.*
- *The sum of the elements of (a dilate of) A'' is small: There is a non-zero element $b \in \mathbf{Z}_p$ such that the elements of bA'' belong to the interval $[1, (p-1)/2]$ and their sum is less than p.*

Theorem 2.2. *Let A be an incomplete subset of $\mathbf{Z}_p$. Then we can partition A into two disjoint sets A' and A'' where*

- *A' has negligible cardinality: $|A'| = O(p^{1/2}/\log^2 p)$.*
- *The norm sum of the elements of (a dilate of) A'' is small: There is a non-zero element $b \in \mathbf{Z}_p$ such that the sum of the norms of the elements of bA'' is less than p.*

The above two theorems were proved (without being formally stated) in [**?**]. A stronger version of these theorems will appear in a forth coming paper [5]. We also need the following simple lemmas.

Lemma 2.3. *Let $T' \subset T$ be sets of integers with the following property. There are integers $a \leq b$ such that $[a, b] \subset \sum(T')$ and the non-negative (non-positive) elements of $T \backslash T'$ are less than $b - a$ (greater than $a - b$). Then*

$$[a, b + \sum_{x \in T \backslash T', x \geq 0} x] \subset \sum(T).$$

$$([a + \sum_{x \in T \backslash T', x \leq 0} x, b] \subset \sum(T).)$$

The (almost trivial) proof is left as an exercise.

Lemma 2.4. *Let $K = \{k_1, \dots, k_l\}$ be a subset of $\mathbf{Z}_p$, where the k_i are positive integers and $\sum_{i=1}^{l} k_i \le p$. Then $|\sum(K)| \ge l(l+1)/2$.*

To verify this lemma, notice that the numbers

$$k_1, \dots, k_l, k_1+k_l, k_2+k_l, \dots, k_{l-1}+k_l, k_1+k_{l-1}+k_l, \dots, k_{l-2}+k_{l-1}+k_l, \dots, k_1+\dots+k_l$$

are different and all belong to $\sum(K)$.

3. Proof of Theorem 1.6

Let A be a zero-free-sum subset of $\mathbf{Z}_p$ with size $n(p)$. In fact, as there is no danger for misunderstanding, we will write n instead of $n(p)$. We start with few simple observations.

Consider the partition $A = A' \cup A''$ provided by Theorem 2.1. Without loss of generality, we can assume that the element b equals one. Thus $A'' \subset [1, (p-1)/2]$ and the sum of its elements is less than p. We first show that most of the elements of A'' belong to the set of the first n positive integers $[1, n]$.

Lemma 3.1. $|A'' \cap [1, n]| \ge n - O(n/\log n)$.

Proof By the definition of n and the property of A''

$$\sum_{i=1}^{n} i \ge p > \sum_{a \in A''} a.$$

Assume that A'' has l elements in $[1, n]$ and k elements outside. Then

$$\sum_{a \in A''} a \ge \sum_{i=1}^{l} i + \sum_{j=1}^{k} (n + j).$$

It follows that

$$\sum_{i=1}^{n} i > \sum_{i=1}^{l} i + \sum_{j=1}^{k} (n+j),$$

which, after a routine simplification, yields

$$(l+n+1)(n-l) > (2n+k)k.$$

On the other hand, $n \geq k + l = |A''| \geq n - O(n/\log^2 n)$, thus $n - l = k + O(n/\log^2 n)$ and $n + l + 1 \leq 2n - k + 1$. So there is a constant c such that

$$(2n-k+1)(k+cn/\log^2 n) > (2n+k)k,$$

or equivalently

$$\frac{cn}{k\log^2 n} > \frac{k+1}{2n-k+1}.$$

Since $2n - k + 1 \leq 2n + 1$, a routine consideration shows that $k^2 \log^2 n = O(n^2)$ and thus $k = O(n/\log n)$, completing the proof. ■

The above lemma shows that most of the elements of A'' (and A) belong to $[1, n]$. Let $A_1 = A \cap [1, n]$. It is trivial that

$$|A_1| \geq |A'' \cap [1,n]| = n - O(n/\log n).$$

Let $A_2 = A \backslash A_1$. We have

$$t := |[1,n] \backslash A_1| = |A_2| = |A| - |A_1| = O(n/\log n).$$

Next we show that $\sum(A_1)$ contains a very long interval. Set $I := [2t + 3, (n+1)(\lfloor n/2 \rfloor - t - 1)]$. The length of I is $(1 - o(1))p$; thus I almost covers $\mathbf{Z}_p$.

Lemma 3.2. $I \subset \sum(A_1)$.

Proof We need to show that every element x of in this interval can be written as a sum of distinct elements of A_1. There are two cases:

Case 1. $2t+3 \leq x \leq n$. In this case A_1 contains at least $x-1-t \geq (x+1)/2$ elements in the interval $[1, x-1]$. This guarantees that there are two distinct elements of A_1 adding up to x.

Case 2. $x = k(n+1)+r$ for some $1 \leq k \leq \lfloor n/2 \rfloor - t - 2$ and $0 \leq r \leq n+1$. First, notice that since $|A_1|$ is very close to n (in fact it is enough to have $|A_1|$ slightly larger than $2n/3$ here), one can find three distinct elements $a, b, c \in A_1$ such that $a+b+c = n+1+r$. Consider the set $A_1' = A_1 \backslash \{a, b, c\}$. We will represent $x-(n+1+r) = (k-1)(n+1)$ as a sum of distinct elements of A_1'. Notice that there are exactly $\lfloor n/2 \rfloor$ ways to write $n+1$ as a sum of two different positive integers. We discard a pair if (at least) one of its two elements is not in A_1'. Since $|A_1'| = n-t-3$, we discard at most $t+3$ pairs. So there are at least $\lfloor n/2 \rfloor - t - 3$ different pairs (a_i, b_i) where $a_i, b_i \in A_1'$ and $a_i + b_i = (n+1)$. Thus, $(k-1)(n+1)$ can be written as a sum of distinct pairs. Finally, x can be written as a sum of a, b, c with these pairs. ■

Now we investigate the set $A_2 = A \backslash A_1$. This is the collection of elements of A outside the interval $[1, n]$. Since A is zero sum free, $0 \notin A_2 + I$ thanks to Lemma 3.2. It follows that

$$A_2 \subset \mathbf{Z}_p \backslash ([1, n] \cup (-I) \cup \{0\}) \subset J_1 \cup J_2,$$

where $J_1 := [-2t-2, -1]$ and $J_2 = [(n+1), p - (n+1)(\lfloor n/2 \rfloor - t)] = [(n+1), q]$. We set $B := A_2 \cap J_1$ and $C := A_2 \cap J_2$.

Lemma 3.3. $\sum(B) \subset J_1$.

Proof Assume otherwise. Then there is a subset B' of B such that $\sum_{a\in B'} a \leq -2t-3$ (here the elements of B are viewed as negative integers between -1 and $-2t-3$). Among such B', take one where $\sum_{a\in B'} a$ has the smallest absolute value. For this B', $-4t-4 \leq \sum_{a\in B'} a \leq -2t-3$. On the other hand, by Lemma 3.2, the interval $[2t+3, 4t+4]$ belongs to $\sum(A_1)$. This implies that $0 \in \sum(A_1) + \sum(B') \subset \sum(A)$, a contradiction. ■

Lemma 3.3 implies that $\sum_{a\in B} |a| \leq 2t+2$, which yields

$$|B| \leq 2(t+1)^{1/2}. \tag{1}$$

Set $s := |C|$. We have $s \geq t - 2(t+1)^{1/2}$. Let $c_1 < \cdots < c_s$ be the elements of C and $g_1 < \cdots < g_t$ be the elements of $[1,n]\backslash A_1$.

By the definition of n, $\sum_{i=1}^{n} i > p > \sum_{i=1}^{n-1} i$. Thus, there is an (unique) $h \in [1,n]$ such that

$$p = 1 + \cdots + (h-1) + (h+1) + \cdots + n. \tag{2}$$

A quantity which plays an important role in what follows is

$$d := \sum_{i=1}^{s} c_i - \sum_{j=1}^{t} g_j.$$

Notice that if we replace the g_j by the c_i in (2), we represent $p+d$ as a sum of distinct elements of A

$$p + d = \sum_{a\in X, X\subset A} a. \tag{3}$$

The leading idea now is to try to cancel d by throwing a few elements from the right hand side or adding a few negative elements (of A) or both. If this was always possible, then we would have a representation of p as a sum of distinct elements in A (in other words $0 \in \sum(A)$), a contradiction. To conclude the proof of Theorem 1.6, we are going to show that the only case when it is not possible is when $p = n(n+1)/2 - 1$ and $A = \{-2, 1, 3, 4, \dots, n\}$. We consider two cases:

Case 1. $h \in A_1$**.** Set $A_1' = A_1 \backslash \{h\}$ and apply Lemma 3.2 to A_1', we conclude that $\sum(A_1')$ contains the interval $I' = [2(t+1)+3, (n+1)(\lfloor n/2 \rfloor - t - 2)]$.

Lemma 3.4. $d < 2(t+1)+3$.

Proof Assume $d \geq 2(t+1)+3$. Notice that the largest element in J_2 (and thus in C) is less than the length of I'. So by removing the c_i one by one from d, one can obtain a sum $d' = \sum_{i=1}^{s'} c_i - \sum_{j=1}^{t} g_j$ which belongs to I', for some $s' \leq s$. This implies

$$\sum_{i=1}^{s'} c_i = \sum_{j=1}^{t} g_j + \sum_{a \in X} a$$

for some subset X of A_1'. Since $h \notin A_1'$, the right hand side is a subsum of the right hand side of (2). Let Y be the collection of the missing elements (from the right hand side of (2)). Then $Y \subset A_1$ and $\sum_{i=1}^{s'} c_i + \sum_{a \in Y} a = p$. On the other hand, the left hand side belongs to $\sum(A_1) + \sum(A_2) \subset \sum(A)$. It follows that $0 \in \sum(A)$, a contradiction. ∎

Now we take a close look at the inequality $d < 2(t+1)+3$. First, observe that since A is zero-sum-free, $-\sum(B) \subset \{g_1, \dots, g_t\}$. By Lemma 3.3, $\sum_{a \in B} |a| \leq 2t + 2 < p$. As B has $t - s$ elements, by Lemma 2.4, $\sum(B)$ has at least $(t-s)(t-s+1)/2$ elements, thus $\{g_1, \dots, g_t\}$ contains at least $(t-s)(t-s+1)/2$ elements in $[1, 2t+2]$. It follows that

$$\sum_{i=1}^{t} g_i \leq (2t+2)(t-s)(t-s+1)/2 + \sum_{j=0}^{t-(t-s)(t-s+1)/2-1} (n-j).$$

On the other hand, as all elements of C are larger than n

$$\sum_{i=1}^{s} c_s \geq \sum_{i=1}^{s} (n+i).$$

It follows that d is at least

$$\sum_{i=1}^{s} (n+i) - (2t+2)(t-s)(t-s+1)/2 - \sum_{j=0}^{t-(t-s)(t-s+1)/2-1} (n-j).$$

If $t-s \geq 2$ then $s > t-(t-s)(t-s+1)/2$, we have

$$d \geq n(s-(t-(t-s)(t-s+1)/2)) - (2t+2)(t-s)(t-s+1)/2.$$

Which yields that

$$d \geq (t-s)(t-s-1)(n-3(2t+2))/2.$$

So the last formula has order $\Omega(n) \gg t$, thus $d \gg 2(t+1)+3$, a contradiction. Therefore, $t-s$ is either 0 or 1.

If $t-s=0$, then $d = \sum_{i=1}^{t} c_i - \sum_{i=1}^{t} g_i \geq t^2$. This is larger than $2t+5$ if $t \geq 4$. Thus, we have $t = 0, 1, 2, 3$.

- $t = 0$. In this case $A = [1, n]$ and $0 \in \sum(A)$.
- $t = 1$. In this case $A = [1, n] \backslash \{g_1\} \cup c_1$. If $c_1 - g_1 \neq h$, then we could substitute c_1 for $g_1 + (c_1 - g_1)$ in (2) and have $0 \in \sum(A)$. This

means that $h = c_1 - g_1$. Furthermore, $h < 2t + 5 = 7$ so both c_1 and g_1 are close to n. If $h \geq 3$,

$$p = \sum_{i=1}^{h-1} i + \sum_{j=h+1}^{n} j = \sum_{i=2}^{h-2} i + \sum_{h+1 \leq j \leq n, j \neq g_1} j + c_1.$$

Similarly, if $h = 1$ or 2 then we have

$$p = \sum_{i=1}^{h} i + \sum_{h+2 \leq j \leq n, j \neq g_1} j + c_1.$$

- $t > 1$. Since $d < 2t+5$, $g_1, \dots, g_t$ are all larger than $n - 2t - 4$. As p is sufficiently large, we can assume $n \geq 4t + 10$, which implies that $[1, 2t + 5] \subset A_1$. If $h \neq 1$, then it is easy to see that $[3, 2t + 5] \subset \sum(A_1 \backslash \{h\})$. As $t > 1$, $d \geq t^2 \geq 4$ and can be represented as a sum of elements in $A_1 \backslash \{h\}$. Omitting these elements from (3), we obtain a representation of p as a sum of elements of A. The only case left is $h = 1$ and $d = 4$. But d can equal 4 if and only if $t = 2$, $c_1 = n + 1, c_2 = n + 2, g_1 = n - 1, g_2 = n$. In this case, we have

$$p = \sum_{i=2}^{n} i = 2 + 3 + \sum_{i=5}^{n+2} i.$$

Now we turn to the case $t - s = 1$. In this case B has exactly one element in the interval $[-2t - 2, -1]$ (modulo p) and d is at least $s^2 - (2t + 2) = (t-1)^2 - (2t+2)$. Since $d < 2t+5$, we conclude that t is at most 6. Let $-b$ be the element in B (where b is a positive integer). We have $b \leq 2t+2 \leq 14$. A_1 misses exactly t elements from $[1, n]$; one of them is b and all other are close to n (at least $n - (2t + 4)$). Using this information, we can reduce the bound on b further. Notice that the whole interval $[1, b - 1]$ belongs to A_1. So if $b \geq 3$, then there are two elements x, y of A_1 such that $x + y = b$. Then $x + y + (-b) = 0$, meaning $0 \in \sum(A)$. It thus remain to consider $b = 1$ or 2. Now we consider a few cases depending on the value of d. Notice that $d \geq s^2 - b \geq -2$. In fact, if $s \geq 2$ then $d \geq 2$. Furthermore, if $s = 0$, then $t = 1$ and $d = -g_1 = -b$.

- $d \geq 5$. Since A_1 misses at most one element in $[1, d]$ (the possible missing element is b), there are two elements of A_1 adding up to d. Omitting these elements from (3), we obtain a representation of p as a sum of distinct elements of A.
- $d = 4$. If $b = 1$, write $p = \sum_{a \in X, a \neq 2} a + (-b)$. If $b = 2$, then $p = \sum_{a \in X, a \neq 1,3} a$. (Here and later X is the set in (3).)
- $d = 3$. Write $p = \sum_{a \in X, a \neq 3-b} a + (-b)$.
- $d = 2$. If $b = 1$, then $p = \sum_{a \in X, a \neq 2} a$. If $b = 2$, then $p = \sum_{a \in X} a + (-2)$.
- $d = 1$. If $b = 1$, then $p = \sum_{a \in X} a + (-1)$. If $b = 2$, then $p = \sum_{a \in X, a \neq 1} a$.
- $d = 0$. In this case (3) already provides a representation of p.
- $d = -1$. In this case $s < 2$. But since $h \neq b$, s cannot be 0. If $s = 1$ then $b = 2$ and $c_1 = n + 1$, $g_1 = n$. By (2), we have $p = \sum_{i=1}^{h-1} i + \sum_{j=h+1}^{n} j$ and so

$$p + (h - 1) = \sum_{1 \leq i \leq n+1, i \notin \{2,n\}} i$$

where the right hand side consists of elements of A only. If $h - 1 \in A$ then we simply omit it from the sum. If $h - 1 \notin A$, then $h - 1 = 2$ and $h = 3$. In this case, we can write

$$p = \sum_{1 \leq i \leq n+1, i \notin \{2,n\}} i + (-2).$$

- $d = -2$. This could only occur if $s = 0$ and $b = 2$. In this case $A = \{-2, 1, 3, \ldots, n\}$. If $h = 1$, then $p = \sum_{i=2}^{n} = n(n+1)/2 - 1$ and we end up with the only exceptional set. If $h \geq 3$, then $p + (h - 2) = \sum_{1 \leq i \leq n, i \neq 2} i$. If $h \neq 4$, then we can omit $h - 2$ from the right hand side to obtain a representation of p. If $h = 4$, then we can write

$$p = \sum_{1 \leq i \leq n, i \neq 2} i + (-2).$$

Case 2. $h \notin A$. In this case we can consider A_1 instead of A_1'. The consideration is similar and actually simpler. Since $h \notin A$, we only need to consider $d := \sum_{i=1}^{s} c_i - \sum_{1\leq j\leq t, g_j\neq h} g_j$. Furthermore, as $h \notin A$, if $s = 0$ we should have $h = b$ and this forbid us to have any exceptional structure in the case $d = -2$. The detail is left as an exercise.

4. Proof of Theorem 1.9

We follow the same terminology used in the previous section. Assume that A is zero-sum-free and $|A| = \lambda n = \lambda(2p)^{1/2}$ with some $1 \geq \lambda \geq .99$. Furthermore, assume that the element b in Theorem 2.1 is one. We will use the notation of the previous proof. Let the *core* of A be the collection of $a \in A$ such that $n+1-a \in A$. Theorem 1.9 follows directly from the following two lemmas.

Lemma 4.1. *The core of A has size at least $.6n$.*

Lemma 4.2. *Let A be a zero-sum-free set whose core has size at least $(1/2+\epsilon)n$ (for some positive constant ϵ). Then*

$$\sum_{a\in A, a<p/2} a \leq p + \frac{1}{\epsilon}(n+1)$$

and

$$\sum_{a\in A, a>p/2} \|a\| \leq (\frac{1}{\epsilon}+1)n.$$

Proof (Proof of Lemma 4.1.) Following the proof of Lemma 3.1, with $l = |A'' \cap [1,n]|$ and $k = |A'' \setminus [1,n]|$, we have

$$(l+n+1)(n-l) > (2n+k)k.$$

On the other hand, $n \geq k+l = |A''| = |A| - O(n/\log^2 n)$, thus $n - l = k + n - |A| + O(n/\log^2 n) = (1-\lambda+o(1))n + k$ and $n + l \leq (1+\lambda)n - k$. Putting all these together with the fact that λ is quite close to 1, we can conclude that that $k < .1n$. It follows (rather generously) that $l = \lambda n - k - O(n/\log^2 n) > .8n$.

The above shows that most of the elements of A belong to $[1, n]$, as

$$|A_1| = |A \cap [1, n]| \geq |A'' \cap [1, n]| > .8n.$$

Split A_1 into two sets, A_1' and $A_1'' := A_1 \setminus A_1'$, where A_1' contains all elements a of A_1 such that $n + 1 - a$ also belongs to A_1. Recall that A_1 has at least $\lfloor n/2 \rfloor - t$ pairs (a_i, b_i) satisfying $a_i + b_i = n + 1$. This guarantees that $|A_1'| \geq 2(\lfloor n/2 \rfloor - t) \geq .6n$. On the other hand, A_1' is a subset of the core of A. The proof is complete. ∎

Proof (Proof of Lemma 4.2) Abusing the notation slightly, we use A_1' to denote the core of A. We have $|A_1'| \geq (1/2 + \epsilon)n$.

Lemma 4.3. *Any* $l \in [n(1/\epsilon + 1), n(1/\epsilon + 1) + n]$ *can be written as a sum of* $2(1/\epsilon + 1)$ *distinct elements of* A_1'.

Proof First notice that for any m belongs to $I_\epsilon = [(1 - \epsilon)n, (1 + \epsilon)n]$, the number of pairs $(a, b) \in {A_1'}^2$ satisfying $a < b$ and $a + b = m$ is at least $\epsilon n/2$. Next, observe that any k, $k \in [0, n]$, is a sum of $1/\epsilon + 1$ integers (not necessarily distinct) from $[0, \epsilon n]$. Consider l from $[n(1/\epsilon+1), n(1/\epsilon+1)+n]$; we can represent $l - n(1/\epsilon + 1)$ as a sum $a_1 + \cdots + a_{1/\epsilon+1}$ where $0 \leq a_1, \ldots, a_{1/\epsilon+1} \leq \epsilon n$. Thus l can be written as a sum of $1/\epsilon + 1$ elements (not necessarily distinct) of I_ϵ, as $l = (n + a_1) + \cdots + (n + a_{1/\epsilon+1})$. Now we represent each summand in the above representation of l by two elements of A_1'. By the first observation, the numbers of pairs are much larger than the number of summands, we can manage so that all elements of pairs are different. ∎

Recall that A_1' consists of pairs (a_i', b_i') where $a_i' + b_i' = n + 1$, so

$$\sum_{a' \in A_1'} a' = (n+1)|A_1'|/2.$$

Lemma 4.4. $I' := [n(1/\epsilon+1), \sum_{a' \in A_1'} a' - (n+1)/\epsilon] \subset \sum(A_1')$.

Proof Lemma 4.3 implies that for each $x \in [n(1/\epsilon+1), n(1/\epsilon+1)+n]$ there exist distinct elements $a_1', \dots, a_{2(1/\epsilon+1)}' \in A_1'$ such that $x = \sum_{i=1}^{2(1/\epsilon+1)} a_i'$. We discard all a_i' and $(n+1) - a_i'$ from A_1'. Thus there remain exactly $|A_1'|/2 - 2(1/\epsilon+1)$ different pairs (a_i'', b_i'') where $a_i'' + b_i'' = n+1$. The sums of these pairs represent all numbers of the form $k(n+1)$ for any $0 \le k \le |A_1'|/2 - 2(1/\epsilon+1)$. We thus obtained a representation of $x + k(n+1)$ as a sum of different elements of A_1', in other word $x + k(n+1) \in \sum(A_1')$. As x varies in $[n(1/\epsilon+1), n(1/\epsilon+1)+n]$ and k varies in $[0, |A_1'|/2 - 2(1/\epsilon+1)]$, the proof is completed. ∎

Let $A_2 = A \setminus A_1$ and set $A_2' := A_2 \cap [0, (p-1)/2]$ and $A_2'' = A_2 \backslash A_2'$. We are going to view A_2'' as a subset of $[-(p-1)/2, -1]$.

We will now invoke Lemma 2.3 several times to conclude Lemma 4.2. First, it is trivial that the length of I' is much larger than n, whilst elements of A_1 are positive integers bounded by n. Thus, Lemma 2.3 implies that

$$I'' := [n(1/\epsilon+1), \sum_{a \in A_1} a - (n+1)/\epsilon] \subset \sum(A_1).$$

Note that the length of I'' is greater than $(p-1)/2$. Indeed $n \approx (2p)^{1/2}$ and

$$\begin{aligned} |I''| &= \sum_{a \in A_1} a - (n+1)/\epsilon - n(1/\epsilon+1) \ge \sum_{a \in A_1'} a - O(n) \\ &\ge (1/2+\epsilon)n(n+1)/2 - O(n) > (p-1)/2. \end{aligned}$$

Again, Lemma 2.3 (applied to I'') yields that

$$[n(1/\epsilon+1), \sum_{a\in A_1\cup A_2'} a-(n+1)/\epsilon] \subset \sum(A_1\cup A_2')$$

and

$$[\sum_{a\in A_2''} a+n(1/\epsilon+1), \sum_{a\in A_1} a-(n+1)/\epsilon] \subset \sum(A_1\cup A_2'').$$

The union of these two long intervals belongs to $\sum(A)$

$$[\sum_{a\in A_2''} a+n(1/\epsilon+1), \sum_{a\in A_1\cup A_2'} a-(n+1)/\epsilon] \subset \sum(A).$$

On the other hand, $0\notin\sum(A)$ implies

$$\sum_{a\in A_2''} a+n(1/\epsilon+1)>0$$

and

$$\sum_{a\in A_1\cup A_2'} a-(n+1)/\epsilon<p.$$

The proof of Lemma 4.2 is completed. ■

5. Sketch of the proof of Theorem 1.13

Assume that A is incomplete and $|A|=\lambda p^{1/2}$ with some $2\geq\lambda\geq 1.99$. Furthermore, assume that the element b in Theorem 2.2 is one. We are going to view $\mathbf{Z}_p$ as $[-(p-1)/2,(p-1)/2]$.

To make the proof simple, we made some new invention: $n=\lfloor p^{1/2}\rfloor$, $A_1:=A\cap[-n,n], A_1':=A\cap[0,n], A_1'':=A\cap[-n,-1], A_2':=A\cap[n+1,(p-$

$1)/2], A_2'' := A \cap [-(p-1)/2, -(n+1)], t_1' := |A_1'|, t_1'' := |A_1''|, t_1 := |A_1| = t_1' + t_1''$.

Notice that $|A''|$ (in Theorem 2.2) is sufficiently close to the upper bound. The following holds.

Lemma 5.1. *Most of the elements of A'' belong to $[-n, n]$;*

- *both t_1' and t_1'' are larger than $(1/2+\epsilon)n$,*
- *t_1 is larger than $(2^{1/2}+\epsilon)n$*

with some positive constant ϵ.

As a consequent, both $\sum(A \cap [-n, -1])$ and $\sum(A \cap [1, n])$ contain long intervals thanks to the following Lemma, which is a direct application of Lemma 4.3 and argument provided in Lemma 3.2.

Lemma 5.2. *If X is a subset of $[1, n]$ with size at least $(1/2+\epsilon)n$. Then*

$$[(n+1)(1/\epsilon+1), (n+1)(n/2 - t - c_\epsilon)] \subset \sum(X)$$

where $t = n - |X|$ and c_ϵ depends only on ϵ.

Now we can invoke Lemma 2.3 several times to conclude Theorem 1.13.

Lemma 5.2 implies

$$I' := [(n+1)(1/\epsilon+1), (n+1)(n/2 - t_1' - c_\epsilon)] \subset \sum(A_1').$$

and

$$I'' := [-(n+1)(n/2 - t_1'' - c_\epsilon), -(n+1)(1/\epsilon+1)] \subset \sum(A_1'').$$

Lemma 2.3 (applied to I' and A_1''; I'' and A_1' respectively) yields

$$[\sum_{a_1''\in A_1''} a_1'' + (n+1)(1/\epsilon+1), (n+1)(n/2 - t_1' - c_\epsilon)] \subset \sum(A_1)$$

and

$$[-(n+1)(n/2 - t_1'' - c_\epsilon), \sum_{a_1'\in A_1'} a_1' - (n+1)(1/\epsilon+1)] \subset \sum(A_1).$$

which gives

$$I := [\sum_{a_1''\in A_1''} a_1'' + (n+1)(1/\epsilon+1), \sum_{a_1'\in A_1'} a_1' - (n+1)(1/\epsilon+1)] \subset \sum(A_1).$$

Note that the length of I is greater than $(p-1)/2$. Again, Lemma 2.3 (applied to I and A_2', I and A_2'' respectively) implies

$$[\sum_{a''\in A_1''\cup A_2''} a'' + (n+1)(1/\epsilon+1), \sum_{a_1'\in A_1'} a_1' - (n+1)(1/\epsilon+1)] \subset \sum(A)$$

and

$$[\sum_{a_1''\in A_1''} a_1'' + (n+1)(1/\epsilon+1), \sum_{a'\in A_1'\cup A_2'} a' - (n+1)(1/\epsilon+1)] \subset \sum(A).$$

The union of these two intervals belongs to $\sum(A)$,

$$[\sum_{a''\in A_1''\cup A_2''} a'' + (n+1)(1/\epsilon+1), \sum_{a'\in A_1'\cup A_2'} a' - (n+1)(1/\epsilon+1)] \subset \sum(A).$$

On the other hand, $\sum(A) \neq \mathbf{Z}_p$ implies

$$\sum_{a' \in A_1' \cup A_2'} a' - \sum_{a'' \in A_1'' \cup A_2''} a'' - 2(n+1)(1/\epsilon + 1) < p.$$

In other words

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

References

[1] Jean-Marc Deshouillers, *Quand seule la sous-somme vide est nulle modulo p*, the prodeeding of the Journees Arithmetiques 2005.
[2] Jean-Marc Deshouillers, *Lower bound concerning subset sum wich do not cover all the residues modulo p*, Hardy- Ramanujan Journal, Vol. 28(2005) 30-34.
[3] Jean-Marc Deshouillers and Gregory A. Freiman, *When subset-sums do not cover all the residues modulo p*, Journal of Number Theory 104(2004) 255-262.
[4] Paul Erdős and Heilbronn Hans Arnold, *On the addition of residue classes modulo p*, Acta Arith. 9 (1964) 149–159.
[5] Hoi H. Nguyen, E. Szemerédi and Van H. Vu, Classification theorems for sumsets, submitted.
[6] J. E. Olson, *An addition theorem modulo p*, J. Combinatorial Theory 5(1968), 45-52.
[7] Hamidoune Yahya Ould and Zémor Gilles, *On zero-free subset sums*, Acta Arith. 78 (1996) no. 2, 143–152.
[8] Endre Szemerédi, *On a conjecture of Erdős and Heilbronn*, Acta Arith. 17 (1970) 227-229.
[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.

Department of Mathematics, Rutgers, Piscataway, NJ 08854

Department of Computer Science, Rutgers, Piscataway, NJ 08854

Department of Mathematics, Rutgers, Piscataway, NJ 08854

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