A CHARACTERIZATION OF INCOMPLETE SEQUENCES IN VECTOR SPACES

HOI H. NGUYEN AND VAN H. VU

Abstract. A sequence A of elements an additive group G is *incomplete* if there exists a group element that *can not* be expressed as a sum of elements from A. The study of incomplete sequences is a popular topic in combinatorial number theory. However, the structure of incomplete sequences is still far from being understood, even in basic groups.

The main goal of this paper is to give a characterization of incomplete sequences in the vector space $\mathbf{F}_p^d$, where d is a fixed integer and p is a large prime. As an application, we give a new proof for a recent result by Gao-Ruzsa-Thangadurai on the Olson's constant of $\mathbf{F}_p^2$ and partially answer their conjecture concerning $\mathbf{F}_p^3$.

1. Introduction

Let G be an additive group and A be a sequence of elements of G. We denote by S_A the collection of subsequence sums of A:

$$S_A = \left\{ \sum_{x \in B} x | B \subset A, 0 < |B| < \infty \right\}.$$

For a positive integer $m \leq |A|$, we denote by m^*A the collection of partial sums over subsequences of cardinality m,

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

If $0 \notin S_A$ (or $0 \notin m^*A$) then we say that A is *zero-sum-free* (or *m-zero-sum-free*). If $S_A \neq G$ (or $m^*A \neq G$) then we say that A is *incomplete* (or *m-incomplete*).

The following questions are among the most popular in classical combinatorial number theory.

The authors are partially supported by research grants DMS-0901216 and AFOSAR-FA-9550-09-1-0167.

Question 1.1. When is A zero-sum-free? When is A incomplete? When is A $m-$zero-sum-free? When is A m-incomplete?

There is a number of results concerning these questions (see for instance [3, 5, 11, 12, 14] for surveys) including classical results such as Olson's theorem and the Erdős-Ginzburg-Ziv theorem. Our goal is to study the above problems for the basic group $\mathbf{F}_p^d$, as d is fixed and p is a large prime. (Here and later $\mathbf{F}_p$ denote the finite field with p elements.)

Our understanding in the case $d = 1$ is more or less satisfying, due to the results from [7, 8] (see also [14] for a survey). However, the proofs of these results do not extend to higher dimensions. The main difficulty in the extension is the existence of non-trivial subgroups (subspaces). In this paper, we develop a new approach that leads to a characterization for incomplete sequences in $\mathbf{F}_p^d$ for $d \geq 2$. This approach makes important use of ideas developed by Alon and Dubiner in [1].

In what follows, it is important to distinguish *sequence* (which means multiple set, where an element may have multiplicity greater than one) and *set* or *subset* (each element appears exactly once).

Let us start by a simple observation. As $S_A = \cup_{1 \leq m \leq |A|} m^* A$, if A is incomplete then it is also m-incomplete for every $m \leq |A|$. So, we are going to consider m-incomplete sequences. It is clear that if A belongs to a translate of a proper subspace of $\mathbf{F}_p^d$, then $m^* A$ belongs to another translate of that proper subspace, and hence A is m-incomplete.

Our leading intuition is that some sort of converse statement must hold. Roughly speaking, we expect that the main reason for a sequence A in $\mathbf{F}_p^d$ to be m-incomplete is that its elements are contained in few translates of a proper subspace. In the special case $d = 1$, the only proper subspace is $\{0\}$. Thus if A is a m-incomplete sequence in $\mathbf{F}_p$, then it consists of few elements of high multiplicities (see [8] for detailed discussion). We are able to quantify this intuition in the following form.

Theorem 1.2 (Characterization for incomplete sequences in $\mathbf{F}_p^d$)**.** *For any positive integer d and positive constants α, β there is a positive constant ϵ such that the following holds for every sequence A in $\mathbf{F}_p^d$, where p is a sufficiently large prime. Either there is an $m \leq \beta p$ such that $m^* A = \mathbf{F}_p^d$ or A can be partitioned into disjoint subsequences $A_0, A_1, \dots, A_l$ such that*

- $|A_0| \leq \alpha p$;
- $|A_i| = \lfloor \epsilon p \rfloor, i \geq 1$;
- *there is a subspace H such that each A_i ($1 \leq i \leq l$) is contained in a translate of H and $m^* A_0$ contains a translate of H for some $1 \leq m \leq |A_0|$.*

The set A_0 is often viewed as the set of exceptional elements. Note that if $H = \{0\}$, then each A_i consists of only one element of multiplicity $\lfloor \epsilon p \rfloor$.

Theorem 1.2 seem to be applicable for various additive problems in vector spaces. In the rest of this section, we discuss a few examples.

Olson's constant. Given an additive group G, let $\mathcal{O}L(G)$, the *Olson constant* of G, be the smallest integer such that no subset of G of cardinality $\mathcal{O}L(G)$ is zero-sum-free. $\mathcal{O}L(G)$ is a parameter of principal interest and its determination has a long history. Erdős, Graham and Heilbronn ([2]) conjectured that $\mathcal{O}L(G) \leq \sqrt{2|G|}$. Szemerédi ([13]) showed that there exists an absolute constant c such that $\mathcal{O}L(G) \leq c\sqrt{|G|}$. Olson ([9],[10]) then improved it to $\mathcal{O}L(G) \leq 2\sqrt{|G|}$. The current record is due to Hamidoune and Zémor ([6]) who showed that $\mathcal{O}L(G) \leq \sqrt{2|G|} + O(|G|^{1/3} \log |G|)$.

Regarding the group $\mathbf{F}_p$, Hamidoune and Zémor proved that $\mathcal{O}L(\mathbf{F}_p) \leq \lceil \sqrt{2p} + 5 \log p \rceil$. They conjectured that the additional log term is not necessary, and this has been recently settled by the current authors with Szemerédi (also due to Deshouillers and Prakash, see [7]). These results give the exact value of $\mathcal{O}L(\mathbf{F}_p)$. Recently, Gao, Ruzsa and Thangadurai [4] showed that $\mathcal{O}L(\mathbf{F}_p^2) = p + \mathcal{O}L(\mathbf{F}_p) - 1$.

Our first application is the following strengthening of Gao-Ruzsa-Thangadurai result.

Theorem 1.3 (Description of optimal zero-sum-free sets in $\mathbf{F}_p^2$)**.** *Suppose that A is a zero-sum-free set of cardinality $p + \mathcal{O}L(\mathbf{F}_p) - 2$ in $\mathbf{F}_p^2$, where p is a sufficiently large prime. Then there is a linear full rank map Φ such that one of the following holds.*

- *$\Phi(A)$ contains $\mathcal{O}L(\mathbf{F}_p) - 1$ points on the line $x = 0$ and $p - 1$ points on the line $x = 1$.*
- *$\Phi(A)$ contains $\mathcal{O}L(\mathbf{F}_p) - 1$ points on the line $x = 0$, $p - 2$ points on the line $x = 1$, and one point on the line $x = 2$.*

Theorem 1.3 not only reproves the bound $p + \mathcal{O}L(\mathbf{F}_p) - 2$, but also characterizes all extremal sets. The proof is short and simple; furthermore, it also classifies zero-sum-free sets of size $\geq \epsilon p$, for any constant $\epsilon > 0$, but we do not elaborate on this point.

It is not hard to see that there exists in $\mathbf{F}_p^d$ a zero-sum-free set of size $p + \mathcal{O}L(\mathbf{F}_p^{d-1}) - 2$. (One can see this by taking the union of a zero-sum-free set of size $\mathcal{O}L(\mathbf{F}_p^{d-1}) - 1$ on the hyperplane $x_d = 0$ and $p - 1$ arbitrary points of the plane $x_d = 1$.) We thus have $\mathcal{O}L(\mathbf{F}_p^d) \geq p + \mathcal{O}L(\mathbf{F}_p^{d-1}) - 1$. Gao, Ruzsa and Thangadurai [4] made the following conjecture.

Conjecture 1.4 (Precise value of Olson's constant for $\mathbf{F}_p^d$)**.** *For any fixed d and sufficiently large p, $\mathcal{O}L(\mathbf{F}_p^d) = p + \mathcal{O}L(\mathbf{F}_p^{d-1}) - 1$.*

The assumption that p is sufficiently large is necessary, see [3].

Since $\mathcal{O}L(\mathbf{F}_p) = O(\sqrt{p}) = o(p)$, this conjecture would imply the following asymptotic version.

Conjecture 1.5. *For any fixed $d \geq 2$ and $\gamma > 0$, the following holds for all sufficiently large p: $\mathcal{O}L(\mathbf{F}_p^d) \leq (d - 1 + \gamma)p$. (Notice that the lower bound $\mathcal{O}L(\mathbf{F}_p^d) \geq (d - 1 - \gamma)p$ is obvious.)*

As far as we know (prior to this paper) there has been no progress on either conjecture. As another application of Theorem 1.2 we settle the $d = 3$ case of Conjecture 1.5.

Theorem 1.6 (Asymptotic value of Olson's constant for $\mathbf{F}_p^3$)**.** *Let $\gamma > 0$ be an arbitrary positive constant, then the following holds for sufficiently large prime p,*

$$\mathcal{O}L(\mathbf{F}_p^3) \leq (2 + \gamma)p.$$

It is possible that one can establish Conjecture 1.5 for arbitrary d using this approach, but the technical details still elude us at this point.

In the rest of this section we introduce our notation. The remaining sections are organized as follows. In Section 2 we provide our main lemmas. The proof of the characterization result is presented in Section 3. The last two sections are devoted to the proofs of Theorem 1.3 and Theorem 1.6, respectively.

Notation

Norm in $\mathbf{F}_p$. For $x \in \mathbf{F}_p$, $\|x\|$ (the norm of x) is the circular distance from x to 0. (For example, the norm of $p - 1$ is 1.)

Dilation of a sequence. For $b \in \mathbf{F}_p$ and a sequence $A \subset \mathbf{F}_p$, $b \cdot A$ is the collection of all ba, where a varies in A.

Projections. Let H and H' be (not necessarily orthogonal) complementary subspaces. For a vector $a \in \mathbf{F}_p^d$ we let $\pi_H(a), \pi_{H'}(a)$ be the unique vectors $a_H, a_{H'}$ satisfying $a = a_H + a_{H'}$ and $a_H \in H, a_{H'} \in H'$ respectively. (Whenever we use this notation, both H and H' will be specified.) For a given sequence A, set $\pi_H(A) := \{\pi_H(a) | a \in A\}$.

Affine basis. A collection of $d+1$ vectors in general position forms an *affine basis* of $\mathbf{F}_p^d$.

2. Technical Lemmas

We are going to use the following results from [1].

Lemma 2.1 (Sumset of affine bases, [1]). *Assume that $s \leq p$. Let $A_1, \dots, A_s$ be s affine bases of $\mathbf{F}_p^d$. Then*

$$|A_1 + \cdots + A_s| \geq (\frac{s}{8d})^d.$$

Lemma 2.2 (Linear independence implies growth in sumset, [1]). *Let W be a number at least one and A be a sequence in $\mathbf{F}_p^d$ such that no hyperplane contains more than $|A|/(4W)$ elements of A. Then for every subset Y of $\mathbf{F}_p^d$ of size at most $p^d/2$, there is an element a of A such that*

$$|(a+Y)\backslash Y| \geq \frac{W}{16p}|Y|.$$

One can immediately derive the following corollary

Corollary 2.3. *Let W be a number at least one and A be a sequence in $\mathbf{F}_p^d$ such that no hyperplane contains more than $|A|/(4W)$ elements of A. Then for every subset Y of $\mathbf{F}_p^d$ of size at most $p^d/2$ and any element $a' \in A$ there is another element $a \in A$ such that*

$$|(a+Y)\backslash(a'+Y)| \geq \frac{W}{16p}|Y|.$$

Proof. (of Corollary 2.3) Apply Lemma 2.2 for the sequence $A' := \{x - a' | x \in A\}$ and Y. If there is a hyperplane H containing $|A'|/(4W) = |A|/(4W)$ elements of A', then the hyperplane $H' := a' + H$ contains $|A|/(4W)$ elements of A. If there is no such H, then there is an element $a - a'$ in A' such that

$$|((a-a')+Y)\backslash Y| \geq \frac{W}{16p}|Y|.$$

Notice that $y \in ((a-a')+Y)\backslash Y$ if and only if $a' + y \in (a+Y)\backslash(a'+Y)$. The claim follows. □

Our proof for the characterization theorem is based on induction on d. We will invoke the following result from our earlier paper [8] as a black box.

Theorem 2.4 (Characterization for incomplete sequences in $\mathbf{F}_p$, [8]). *Assume that A is an incomplete sequence in $\mathbf{F}_p$, with sufficiently large p. Then there is a residue $b \neq 0$ such that we can partition $b \cdot A$ into two disjoint subsequences $b \cdot A = A^\flat \cup A^\sharp$ where*

- $|A^\flat| \leq p^{12/13}$,
- $\sum_{a \in A^\sharp} \|a\| < p$.

We close this section with a trivial, but useful, fact.

Fact 2.5. *Let H_1, H_1' be subspaces such that $H_1 \oplus H_1' = \mathbf{F}_p^d$. Let A_1, A_2 be such sequences in $\mathbf{F}_p^d$ that $m_1^* A_1$ contains a translate of a subspace H_1 of $\mathbf{F}_p^d$ and $m_2^*(\pi_{H_1'}(A_2))$ contains a translate of a subspace H_2 of H_1'. Then $m_1^* A_1 + m_2^* A_2$ contains a translate of the subspace $H_1 + H_2$ in $\mathbf{F}_p^d$.*

Proof. (of Fact 2.5) Since $m_1^* A_1$ contains a translate of H_1, there exists a vector $v_1 \in \mathbf{F}_p^d$ such that $v_1 + H_1 \subset m_1^* A_1$.

Since $m_2^*(\pi_{H_1'}(A_2))$ contains a translate of H_2, there exists a vector $v_2 \in \mathbf{F}_p^d$ such that for any $h_2 \in H_2$ there corresponds a vector $h_1 \in H_1$ satisfying $v_2 + h_1 + h_2 \in m_2^* A_2$. Hence $v_1 + H_1 + v_2 + h_1 + h_2 = v_1 + v_2 + H_1 + h_2 \subset m_1^* A_1 + m_2^* A_2$.

Since the above holds for all $h_2 \in H_2$, we have $v_1 + v_2 + H_1 + H_2 \subset m_1^* A_1 + m_2^* A_2$.

□

3. Proof of Theorem 1.2

3.1. **Existence of rich hyperplanes.**

Lemma 3.2 (Rich hyperplane lemma)**.** *For any positive integer d and positive constants β, δ there is a positive constant ϵ such that the following holds. Let A be a sequence in $\mathbf{F}_p^d$ with at least δp elements so that there is no $1 \leq m \leq \beta p$ such that $m^* A = \mathbf{F}_p^d$. Then, there is a hyperplane H such that*

$$|A \cap H| \geq \epsilon p.$$

Proof. (of Lemma 3.2) Let $c_1 \leq \min\{\beta/4(d+1), \delta/4(d+1)\}$ be a small positive constant. We consider two cases.

Case 1. One cannot find $2c_1 p$ disjoint affine bases in A.

In this case, by the definition of affine basis, there is a hyperplane H containing

$$|A| - 2c_1 p(d+1) \geq \frac{\delta}{2} p$$

elements of A and we are done.

Case 2. One can find $2c_1p$ disjoint affine bases in A.

Set $s := c_1p$ and let $E_1, \dots, E_s, F_1, \dots, F_s$ be the bases. Define

$$\mathcal{E}_0 := E_1 + \cdots + E_s \text{ and } \mathcal{F}_0 := F_1 + \cdots + F_s.$$

By Lemma 2.1,

$$\min\{|\mathcal{E}_0|, |\mathcal{F}_0|\} \geq (c_1p/8d)^d.$$

Partition $A\backslash(\cup_{i=1}^s E_i \cup \cup_{i=1}^s F_i)$ into two sequences $E^\dagger, F^\dagger$ of equal sizes (we can throw one element from A to ensure parity). By choosing c_1 sufficiently small, we can assume that $|E^\dagger| = |F^\dagger| \geq |A|/4$.

Let $W = W(d, \delta, \beta)$ be a large constant and ϵ be a small constant to be determined. Assume, for a contradiction that there is no hyperplane containing $\epsilon|A|$ elements of A.

Let a_0' be an arbitrary element of $E^\dagger$. By the way we set ϵ so that

$$\epsilon|A| \leq (|E^\dagger| - \delta p/8)/(8W). \tag{1}$$

Thus there is no hyperplane containing $|E^\dagger|/(8W)$ elements of $E^\dagger$. By Corollary 2.3, we find $a_0 \in E^\dagger\backslash\{a_0'\}$ such that

$$|(a_0' + \mathcal{E}_0) \cup (a_0 + \mathcal{E}_0)| \geq (1 + \frac{W}{16p})|\mathcal{E}_0|.$$

Define $\mathcal{E}_1 = (a_0' + \mathcal{E}_0) \cup (a_0 + \mathcal{E}_0)$ and $E_1^\dagger := E^\dagger\backslash\{a_0', a_0\}$. Repeating the argument, we find elements $a_1', a_1 \in E_1^\dagger$ such that

$$|(a_1' + \mathcal{E}_1) \cup (a_1 + \mathcal{E}_1)| \geq (1 + \frac{W}{16p})|\mathcal{E}_1|.$$

In general, set $E_i^\dagger := E_{i-1}^\dagger\backslash\{a_{i-1}', a_{i-1}\}$ and $\mathcal{E}_i := (a_{i-1}' + \mathcal{E}_{i-1}) \cup (a_{i-1} + \mathcal{E}_{i-1})$. By induction, we have

$$|\mathcal{E}_i| \geq (1 + \frac{W}{16p})^i|\mathcal{E}_0|,$$

unless $|\mathcal{E}_{i-1}| > p^d/2$. Thus by choosing W sufficiently large (in terms of d, δ and β), there is some $0 \leq k \leq \min\{\beta p/4, \delta p/16\}$ such that

$$|\mathcal{E}_k| > \frac{1}{2}p^d.$$

Notice that in every step, the condition $\epsilon|A| \leq |E_i^\dagger|/(4W)$ is satisfied because of (1) and $|E_i^\dagger| \geq |E^\dagger| - 2k \geq |E^\dagger| - \delta p/8$.

Repeating the argument with $F^\dagger$ and $\mathcal{F}_0$, we have for some $0 \leq l \leq \min\{\beta p/4, \delta p/16\}$ that

$$|\mathcal{F}_l| > \frac{1}{2}p^d.$$

Observe that if X and Y are two subsets of a finite Abelian group G and $|X|, |Y| > |G|/2$, then $X + Y = G$. Thus,

$$\mathcal{E}_k + \mathcal{F}_l = \mathbf{F}_p^d.$$

The left hand side is a subset of m^*A for some small m. Indeed, the elements in $\mathcal{E}_k$ (or $\mathcal{F}_l$) are sums of exactly $c_1p + k$ (or $c_1p + l$) elements of A. Furthermore, by the procedure, the sequence of elements of A involved in $\mathcal{E}_k$ is disjoint from the sequence of elements of A involved in $\mathcal{F}_l$. Finally,

$$m \leq 2c_1p + (k + l) \leq p(2c_1 + \beta/4 + \beta/4) \leq \beta p.$$

This concludes the proof of Lemma 3.2.

□

3.3. **Completing the proof of Theorem 1.2.** Let δ_0 be a small positive constant. There is a positive constant ϵ_0 such that (using Lemma 3.2 iteratively) we can partition $A = B_0 \cup B_1 \cup \cdots \cup B_h$, where $|B_0| \leq \delta_0 p$ and $|B_i| = \lfloor \epsilon_0 p \rfloor$ and each B_i $(1 \leq i \leq h)$ is contained in a hyperplane D_i. Consider two cases:

Case 1. There are some $1 \leq i \leq h$ and $1 \leq m \leq \beta p/2$ such that m^*B_i contains $v + D_i$, a translate of D_i.

We can assume that $i = 1$ and that D_1 is parallel to the subspace H spanned by the basic vectors $e_1, \ldots, e_{d-1}$. Let A' be the projection of $A \backslash B_1$ into H' spanned by e_d.

Consider A'. First, A' is a sequence in $\mathbf{F}_p$ with $|A| - \lfloor \epsilon_0 p \rfloor$ elements. Second, if there is no $1 \leq m \leq \beta p$ such that $m^*A = \mathbf{F}_p^d$, then there is no $1 \leq m \leq \beta p/2$ such that $m^*A' = \mathbf{F}_p$ by Fact 2.5. The classification theorem for $\mathbf{F}_p$, Theorem 2.4, implies that there is a subsequence A'' of A' with at

most $\epsilon_0 p$ elements such that $A'\backslash A''$ contains $M = O_{\epsilon_0}(1)$ different elements $a_1, \dots, a_M$. Indeed, one can set A'' to consist of $b^{-1} \cdot A^\flat$ and those elements $b^{-1} \cdot a$, where $a \in A^\sharp$ and $\|a\| \geq 2/\epsilon_0$.

Consider the system of parallel hyperplanes $a_1 + H, \dots a_M + H$. Set $\epsilon := \alpha/2M$. Partition $A \cap (a_i + H)$ into disjoint subsequences of size exactly $\lfloor \epsilon p \rfloor$ and a remainder sequence of size less than ϵp. Let A_0 be the union of the remainders and B_0, B_1 and A''. We have

$$|A_0| \leq M\epsilon p + |B_0| + |B_1| + |A''| \leq \frac{\alpha}{2} p + \delta_0 p + \epsilon_0 p + \epsilon_0 p \leq \alpha p.$$

Furthermore, there is some $1 \leq m \leq \beta p/2$ such that $m^* B_1 \subset m^* A_0$ contains a hyperplane parallel to H. Finally, $A \backslash A_0$ are partitioned into sequences of size exactly $\lfloor \epsilon p \rfloor$, each of which is contained in a translate of H. This concludes the proof for the first case.

Case 2. There is no $1 \leq i \leq h$ and $1 \leq m \leq \beta p/2$ such that $m^* B_i$ is a translate of D_i.

In this case, we can apply the induction hypothesis to a translate of B_i (which is contained in the subspace parallel with D_i) to obtain a decomposition $B_i = B_{i0} \cup B_{i1} \cup \dots \cup B_{il_i}$, where $B_{ij}, 1 \leq j \leq l_i$ are contained in translates of a subspace H_i and there is an integer m_i such that $m_i^* B_{i0}$ contains a translate of H_i (we choose the parameters to be small enough such that $|B_{i0}| \leq \min(\alpha/(2h), \beta/(2h))$). Without loss of generality, one may assume that all B_{ij} have the same size $\lfloor \epsilon p \rfloor$ with a small positive constant ϵ.

Now take $A_0 := B_0 \cup (\cup_i B_{i0})$ and let A_k $(1 \leq k \leq \sum_{i=1}^h l_i)$ be the B_{ij}. By choosing δ_0 small, we get $|A_0| \leq \alpha p$.

Note that $m_1^* B_{10} + \dots + m_j^* B_{h0}$, and thus $(m_1 + \dots + m_h)^* A_0$, contains a translate of $H = H_1 + \dots + H_h$. Each A_k $(1 \leq k)$ is contained in a translate of H since it is contained in a translate of H_k.

4. Proof of Theorem 1.3

The idea is to "project" the problem into $\mathbf{F}_p$ by using the characterization theorem. Once inside this group, we will be able to invoke Theorem 2.4. In fact, we will also need the following result.

Theorem 4.1 (Erdős-Heilbronn type inequality, [**?**])**.** *Let A be a non-empty subset of $\mathbf{F}_p$. Then $m^* A \geq \min\{p, m|A| - m^2 + 1\}$. In particular, provided that p is large enough, we have*

- *if $|A| \geq 2\sqrt{p} + 1$, then $\lfloor \sqrt{p} \rfloor^* A = \mathbf{F}_p$;*

- *if* $|A| \geq .51p$*, then* $m^*A = \mathbf{F}_p$ *for all* $2 \leq m \leq |A| - 2$*.*

For more general results on $m-$incomplete sequences in $\mathbf{F}_p$, we refer the reader to [8, Theorem 2.8].

Fact 4.2. *Let* B_1, B_2 *be subsets of the lines* $x = b$ *and* $x = p - b$ *of* $\mathbf{F}_p^2 = \{(x, y | x, y \in \mathbf{F}_p)\}$ *respectively. Assume that* $|B_1| + |B_2| > p$*. Then* $B_1 + B_2$ *contains the whole* y*-axis. In particular, the set* $B_1 \cup B_2$ *is not zero-sum-free in* $\mathbf{F}_p^2$*.*

Now we are ready to prove Theorem 1.3. Let α and β be sufficiently small constants. Assume that A is zero-sum-free. By Theorem 1.2, there exists $A_0 \subset A$ of cardinality $|A_0| \leq \alpha p$ such that m^*A_0 contains a line of $\mathbf{F}_p^2$. Without loss of generality, we assume that this line is parallel to the y-axis $\mathcal{L} : x = 0$.

Let $\mathcal{L}'$ be the collection of points on the x-axis, then $\mathcal{L} \oplus \mathcal{L}' = \mathbf{F}_p^2$. Let $B := \pi_{\mathcal{L}'}(A \backslash A_0)$ be the projection of $A \backslash A_0$ into $\mathcal{L}'$, thus B is a sequence in $\mathbf{F}_p$. Since A is zero-sum-free, we have

- $|A \cap \mathcal{L}| < \mathcal{O}L(\mathbf{F}_p) < \sqrt{2p} + 1$ (see discussion prior to Theorem 1.3),
- B is an incomplete sequence in $\mathbf{F}_p$ (by Fact 2.5 and that m^*A_0 contains a translate of $\mathcal{L}$).

Together with Theorem 2.4, the last observation implies that, after an appropriate dilation of A in the direction of the x-axis $\mathcal{L}'$, B contains at least $(1 - 2\alpha)p$ elements of norm 1 (those $b \in B$ with $\|b\| = 1$). (Note that such a dilation does not affect the property of A_0, i.e. m^*A_0 still contains a vertical line.)

Let $A_{-1} := \pi_{\mathcal{L}'}^{-1}(-1) \cap A$ and $A_1 := \pi_{\mathcal{L}}^{-1}(1) \cap A$ be the collections of elements of A whose x-coordinate are -1 and 1 respectively. As these elements correspond to the one of norm 1 in B, we have $|A_{-1}| + |A_1| \geq (1 - 2\alpha)p$.

Without loss of generality, assume that $|A_1| \geq 2\sqrt{p} + 1$. It follows from the first part of Theorem 4.1 that $\lfloor \sqrt{p} \rfloor^* A_1$ contains the whole line $x = \lfloor \sqrt{p} \rfloor$, and hence $|A_{-1}| < \sqrt{p}$ by Fact 4.2. In other words, $|A_1| \geq (1 - 2\alpha)p - \sqrt{p} \geq (1 - 3\alpha)p$.

To make the presentation less technical, we abuse the notation to write $A := A_0 \cup A_1 \cup A_{-1} \cup A'$, where the new set A_0 is the intersection of A with the y-axis, $A_0 := A \cap \mathcal{L}$, and $A' := A \backslash (A_0 \cup A_1 \cup A_{-1})$.

Now comes a crucial observation. Since $|A_1| \geq (1 - 3\alpha)p$, the second part of Theorem 4.1 implies that l^*A_1 covers the whole vertical line $x = l$, for every $2 \leq l \leq |A_1| - 2$. Thus the set $\bigcup_{2 \leq l \leq |A_1|-2} l^*A_1$ covers the whole strip $\{(x, y) \in \mathbf{F}_p^2 : 2 \leq x \leq |A_1| - 2\}$.

The last conclusion immediately implies that $|A_{-1}| \leq 1$, otherwise $2^*A_{-1} + 2^*A_1$ would contain $\mathcal{L}$, and hence the origin, a contradiction.

We next focus on A'. Let $X := \{x_1, \ldots, x_{|A'|}\} = \pi_{\mathcal{L}'}(A')$ be the projection of A' into the x-axis $\mathcal{L}'$. It follows that there does not exist any subset of X whose sum belongs to the "opposite" of $\{2, \ldots, |A_1| - 2\}$, i.e. we must have $S_X \subset \{-1, 0, 1, \ldots, p + 1 - |A_1|\}$ (in $\mathbf{F}_p$).

This partly implies that $x_i \in \{2, \ldots, p + 1 - |A_1|\}$, but more importantly, viewing x_i as real numbers, we must have

$$x_1 + \cdots + x_{|A'|} \leq p - |A_1| + 1. \tag{2}$$

Indeed, if there exists an i such that $x_1 + \cdots + x_i \leq p - |A_1| + 1 < x_1 + \cdots + x_{i+1}$ then $p - |A_1| + 2 \leq x_1 + \cdots + x_i + x_{i+1} = (x_1 + \cdots + x_i) + x_{i+1} \leq p - |A_1| + 1 + (p - |A_1| + 1) \leq 6\alpha p < p/2$. Thus, after taking modulo p, $x_1 + \cdots + x_i + x_{i+1}$ remains in $\{p - |A_1| + 2, \ldots, 6\alpha p\}$, which belongs to the "opposite" of $\{2, \ldots, |A_1| - 2\}$ in $\mathbf{F}_p$, a contradiction.

Since $x_i \geq 2$, (2) implies that $|A'|$ is small, $2|A'| \leq p - |A_1| + 1$. From the definition of A', using $|A_{-1}| \leq 1$, we have $2|A| - p - 3 \leq 2|A_0| + |A_1|$. Inserting the bound of $|A|$, $|A| = p + \mathcal{O}L(\mathbf{F}_p) - 2$, we obtain that $2|A_0| + |A_1| \geq p + 2\mathcal{O}L(\mathbf{F}_p) - 7$.

Note that $|A_0| \leq \mathcal{O}L(\mathbf{F}_p) - 1$, the above result trivially implies that $|A_0| + |A_1| > p$, i.e. $A_0 + A_1$ covers the whole line $x = 1$. Hence A_{-1} must be empty.

Using this new information (instead of $|A_{-1}| \leq 1$) and the upper bounds $|A_0| \leq \mathcal{O}L(\mathbf{F}_p) - 1, |A_1| \leq p - 1$, we easily obtain the lower bounds $|A_1| \geq p - 3$ and $|A_0| \geq \mathcal{O}L(\mathbf{F}_p) - 2$, and hence (2) implies that $x_1 + \cdots + x_{|A'|} \leq 4$. Since $x_i \geq 2$, we must have $|A'| \leq 2$. If $|A'| = 0$, we are done. It remains to consider the following two cases.

Case 1. $|A'| = 2$. We then have $x_1 = x_2 = 2$. Thus, if $|A_1| = p - 2$ and $|A_0| = \mathcal{O}L(\mathbf{F}_p) - 2$, then $(p - 4)^*A_1 + x_1 + x_2$ covers the whole y-axis $x = 0$, contradiction. Furthermore, if $|A_1| = p - 3$ and $|A_0| = \mathcal{O}L(\mathbf{F}_p) - 1$, then $(p - 4)^*A_1 + x_1 + x_2$ covers $p - 3$ elements of the y-axis, and hence $(p - 4)^*A_1 + x_1 + x_2 + A_0$ covers the whole axis, another contradiction.

Case 2. $|A'| = 1$. We then easily eliminate the case $|A_1| = p - 1$ and $|A_0| = \mathcal{O}L(\mathbf{F}_p) - 2$ because in this case, by (2), we must have $x_1 = 2$, and hence $(p - 2)^*A_1 + x_1$ covers the whole y-axis.

Assuming $|A_1| = p - 2$ and $|A_0| = \mathcal{O}L(\mathbf{F}_p) - 1$, it is implied that $x_1 \leq 3$ by (2). But if $x_1 = 3$ then $(p - 3)^*A_1 + x_1$ covers $p - 2$ elements of the y-axis, and hence $(p - 3)^*A_1 + x_1 + A_0$ covers the whole axis, a contradiction. As a

consequence, $x_1 = 2$. The only reason that $A = A_0 \cup A_1 \cup A'$ is zero-sum-free in this case is that the multiset $A_0 \cup \{\sum_{a \in A_1 \cup A'} a\}$ is zero-sum-free in $\mathbf{F}_p$, completing the proof.

5. Proof of Theorem 1.6

To establish Theorem 1.6, we again rely on our characterization theorem to project back to $\mathbf{F}_p^2$. After this step, we are not working with sets anymore, but rather with sequences. For this reason, we need the following statement about the "sequence" counterpart of Olson's constant (which is usually referred to as the Davenport's constant).

Theorem 5.1 (Davenport's constant for $\mathbf{F}_p^d$, [9])**.** *Any collection of $d(p-1)+1$ elements of $\mathbf{F}_p^d$ is not zero-sum-free.*

Now we present the proof of Theorem 1.6. Assume that there exists a set A of size $(2+\gamma)p$ which is zero-sum-free. By Theorem 1.2 (after a bijective linear mapping) we can partition A into disjoint sequences, $A = A_0 \cup A_1 \cdots \cup A_n$ where

- $|A_0| \leq \gamma p/4$;
- $|A_i| = \lfloor \epsilon p \rfloor$;
- $m^* A_0$ contains a translate of a subspace H, for some $m \leq |A_0|$;
- there exist $n \leq n(\gamma)$ vectors $a_1, \dots, a_n \in \mathbf{F}_p^3$ such that $A_i \subset a_i + H$.

Let d' be the dimension of H. We observe that d' can not be either 0 or 3, since the first case would imply that A contains elements of multiplicity $\lfloor \epsilon p \rfloor > 1$ (as ϵ is independent of p), while the second would imply that A is complete. We consider two remaining cases.

Case 1. $d' = 2$. Without loss of generality, we assume that $H = \{z = 0\}$. Consider the projection B of $A \backslash A_0$ onto the z-axis, which can be viewed as a sequence in $\mathbf{F}_p$. Since A is zero-sum-free, $|A \cap H| < \mathcal{O}L(\mathbf{F}_p^2) \leq (1+\gamma/4)p$. Thus there are at least $|A| - |A_0| - (1+\gamma/4)p \geq (1+\gamma/2)p$ elements of B having non-zero norm.

By item 2 of Theorem 2.4, the latter implies that B is complete in $\mathbf{F}_p$ (in fact, it is easy to see that any sequence of p non-zero elements of $\mathbf{F}_p$ is complete). Hence $S_{A \backslash A_0} + m^* A_0 = \mathbf{F}_p^3$, in particular the origin, a contradiction.

Case 2. $d' = 1$. Without loss of generality, we assume that H is the z-axis $\{x = 0, y = 0\}$. By the property of A_i, we may write $A_i \subset v_i + H$, where v_i is the projection of a_i onto the xy-plane $\{z = 0\}$.

Consider the sequence $\{v_1, \dots, v_n\}$, where each v_i has multiplicity $|A_i| - \lfloor 2/\epsilon \rfloor$. Since $\sum_{i=1}^n (|A_i| - 2/\epsilon) \geq (2+\gamma/2)p$, Theorem 5.1 implies that there exist $0 \leq m_i \leq |A_i| - 2/\epsilon$ such that $\sum_{i=1}^n m_i v_i = 0$ (in the xy-plane), where

not all m_i are zero. By multiplying all m_i by $\lfloor 2/\epsilon \rfloor$ if needed, we may assume that at least one of the m_i is greater than $\lfloor 2/\epsilon \rfloor$. Let this be m_1.

We now consider the sumset $\sum_{i=1}^n m_i^* A_i$. By the definition of m_i and v_i, the projection of this set into the xy-plane is the origin, in other words, $\sum_{i=1}^n m_i^* A_i$ belong to the z-axis.

On the other hand, since $|A_1| = \lfloor \epsilon p \rfloor$ and $\lfloor 2/\epsilon \rfloor \leq m_1 \leq |A_1| - 2/\epsilon$, Theorem 4.1 implies that $m_1^* A_1$ contains a translate of H, i. e. , $m_1^* A_1$ contains the whole line which has image $m_1 v_1$ in the xy-plane.

It follows that $\sum_{i=1}^n m_i^* A_i$ covers the whole z-axis, and hence the origin, a contradiction.

Acknowledgements. The authors would like to thank the referees for carefully reading this manuscript and providing very helpful remarks.

References

[1] N. Alon and M. Dubiner, A lattice point problem and additive number theory. Combinatorica, **15** (1995), 301-309.
[2] P. Erdős and R.L. Graham, Old and new problems and results in combinatorial number theory, L' Enseignement Mathematique, Universite de Geneve, Vol. **28**, Geneve, 1980.
[3] W. D. Gao and A. Geroldinger, Zero-sum problems in finite abelian groups : a survey, Expo. Math. **24** (2006), 337-369.
[4] W. D. Gao, I. Z. Ruzsa and R. Thangadurai, Olson's constant for the group $\mathbf{F}_p \oplus \mathbf{F}_p$, J. of Combinatorial Theory, Series A, **107** (2004), 49-67.
[5] A. Geroldinger, Additive group theory and non-unique factorizations, Combinatorial Number Theory and Additive Group Theory, Advanced Courses in Mathematics CRM Barcelona, Birkhäuser, 2008.
[6] Y. O. Hamidoune and G. Zémor, On zero-free subset sums, Acta Arithmetica. **78**(2) (1996), 143-152.
[7] H. H. Nguyen, E. Szemerédi and V. H. Vu, Subset sums modulo a prime, Acta Arithmetica. **131** (2008), 303–316.
[8] H. H. Nguyen and V. H. Vu, Classification theorems for sumsets modulo a prime, J. of Combinatorial Theory, Series A, **116** (2009), 936–959.
[9] J. E. Olson, A combinatorial Problem on finite Abelian groups I, Journal of number theory, **1** (1969), 8-10.
[10] J. E. Olson, Sumset of group elements, Acta Arithmetica, **28** (1975/1976), 147-156.
[11] C. Pomerance and A. Sárközy, Combinatorial number theory, 967-1018, Handbook of Combinatorics, MIT press, edited by R. Graham, M. Grötschel, L. Lovász, 1995.
[12] J. W. Sun, List of publications on restricted sumsets. 2005.
[13] E. Szemerédi, On a conjecture of Erdős and Heilbronn, Acta Arithmetica **17** (1970), 227-229.
[14] V. H. Vu, A structural approach to subset-sum problems, Bulding Bridges (Lovász is 60), Proceeding of the Bolyai Society, 2008.

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