LOWER BOUNDS FOR SUMSETS OF MULTISETS IN $\mathbb{Z}_P^2$

Greg Martin
Dept. of Mathematics, University of British Columbia, Vancouver, BC, Canada
gerg@math.ubc.ca

Alexis Peilloux
Université Pierre-et-Marie Curie, aculté de Mathématiques,Paris, France
alexis.peilloux@etu.upmc.fr

Erick B. Wong
Dept. of Mathematics, University of British Columbia, Vancouver, BC, Canada
erick.wong@alumni.ubc.ca

Received: 8/31/12, Revised: 10/7/13, Accepted: 10/31/13, Published: 11/11/13

Abstract

The classical Cauchy–Davenport theorem implies the lower bound $n+1$ for the number of distinct subsums that can be formed from a sequence of n elements of the cyclic group $\mathbb{Z}_p$ (when p is prime and $n < p$). We generalize this theorem to a conjecture for the minimum number of distinct subsums that can be formed from elements of a multiset in $\mathbb{Z}_p^m$; the conjecture is expected to be valid for multisets that are not "wasteful" by having too many elements in nontrivial subgroups. We prove this conjecture in $\mathbb{Z}_p^2$ for multisets of size $p+k$, when k is not too large in terms of p.

1. Introduction

Determining the number of elements in a particular abelian group that can be written as sums of given sets of elements is a topic that goes back at least two centuries. The most famous result of this type, involving the cyclic group $\mathbb{Z}_p$ of prime order p, was established by Cauchy in 1813 [1] and rediscovered by Davenport in 1935 [2, 3] (here $\#A$ denotes the cardinality of A):

Lemma 1.1 (Cauchy–Davenport Theorem). *Let A and B be subsets of $\mathbb{Z}_p$, and define $A+B$ to be the set of all elements of the form $a+b$ with $a \in A$ and $b \in B$. Then $\#(A+B) \geq \min\{p, \#A + \#B - 1\}$.*

The lower bound is easily seen to be best possible by taking A and B to be intervals, for example. It is also easy to see that the lower bound of $\#A+\#B-1$ does not hold for general abelian groups $\mathbf{G}$ (take A and B to be the same nontrivial subgroup of

G). There is, however, a well-known generalization obtained by Kneser in 1953 [4], which we state in a slightly simplified form that will be quite useful for our purposes (see [8, Theorem 4.1] for an elementary proof):

Lemma 1.2 (Kneser's Theorem). *Let A and B be subsets of a finite abelian group $\mathbf{G}$, and let m be the largest cardinality of a proper subgroup of $\mathbf{G}$. Then $\#(A+B) \geq \min\{\#\mathbf{G}, \#A + \#B - m\}$.*

Given a sequence $A = (a_1, \ldots, a_k)$ of (not necessarily distinct) elements of an abelian group $\mathbf{G}$, a related result involves its *sumset* ΣA, which is the set of all sums of any number of elements chosen from A (not to be confused with $A + A$, which it contains but usually properly):

$$\Sigma A = \left\{ \sum_{j \in J} a_j \colon J \subseteq \{1, \ldots, k\} \right\}.$$

(Note that we allow J to be empty, so that the group's identity element is always an element of ΣA, even when A itself is empty) When $\mathbf{G} = \mathbb{Z}_p$, one can prove the following result by writing $\Sigma A = \{0, a_1\} + \cdots + \{0, a_k\}$ and applying the Cauchy–Davenport theorem inductively:

Lemma 1.3. *Let $A = (a_1, \ldots, a_k)$ be a sequence of nonzero elements of $\mathbb{Z}_p$. Then $\#\Sigma A \geq \min\{p, k+1\}$.*

This result can also be proved directly by induction on k, and in fact such a proof will discover why the order p of the cyclic group must be prime (intuitively, the sequence A could lie completely within a nontrivial subgroup). For a formal proof, see [6, Lemma 2]. Again the lower bound is easily seen to be best possible, by taking $a_1 = \cdots = a_k$.

It is a bit misleading to phrase such results in terms of sequences, since the actual order of the elements in the sequence is irrelevant (given that we are considering only abelian groups). We prefer to use *multisets*, which are simply sets that are allowed to contain their elements with multiplicity. If we let m_x denote the multiplicity with which the element x occurs in the multiset A, then the definition of ΣA can be written in the form

$$\Sigma A = \left\{ \sum_{x \in \mathbf{G}} \delta_x x \colon 0 \leq \delta_x \leq m_x \right\},$$

where $\delta_x x$ denotes the group element $x + \cdots + x$ obtained by adding δ_x summands all equal to x.

When using multisets, we should choose our notation with care: the hypotheses of such results tend to involve the total number of elements of the multiset A counting multiplicity, while the conclusions involve the number of distinct elements of ΣA. Consequently, throughout this paper, we use the following notational conventions:

- $|S|$ denotes the total number of elements of the multiset S, counted with multiplicity;
- $\#S$ denotes the number of distinct elements of the multiset S, or equivalently the cardinality of S considered as a (mere) set.

In this notation, Lemma 1.3 can be restated as:

Lemma 1.4. *Let A be a multiset contained in $\mathbb{Z}_p$ such that $0 \notin A$. Then $\#\Sigma A \geq \min\{p, |A|+1\}$.*

The purpose of this paper is to improve, as far as possible, this lower bound for multisets contained in the larger abelian group $\mathbb{Z}_p^2$. We cannot make any progress without some restriction upon our multisets: if a multiset is contained within a nontrivial subgroup of $\mathbb{Z}_p^2$ (of cardinality p), then so is its sumset, in which case the lower bound $\min\{p, |A|+1\}$ from Lemma 1.4 is the best we can do. Therefore we restrict to the following class of multisets. We use the symbol $\mathbf{0} = (0,0)$ to denote the identity element of $\mathbb{Z}_p^2$.

Definition 1.5. A multiset A contained in $\mathbb{Z}_p^2$ is called *valid* if:

- $\mathbf{0} \notin A$; and
- every nontrivial subgroup contains fewer than p points of A, counting multiplicity.

The exact number p in the second condition has been carefully chosen: any nontrivial subgroup of $\mathbb{Z}_p^2$ is isomorphic to $\mathbb{Z}_p$, and so Lemma 1.4 applies to these nontrivial subgroups. In particular, any multiset A containing $p-1$ nonzero elements of a nontrivial subgroup will automatically have that entire subgroup contained in its sumset ΣA, so allowing p nonzero elements in a nontrivial subgroup would always be "wasteful".

We believe that the following lower bound should hold for sumsets of valid multisets:

Conjecture 1.6. Let A be a valid multiset contained in $\mathbb{Z}_p^2$ such that $p \leq |A| \leq 2p-3$. Then $\#\Sigma A \geq (|A|+2-p)p$. In other words, if $|A| = p+k$ with $0 \leq k \leq p-3$, then $\#\Sigma A \geq (k+2)p$.

It is easy to see that this conjectured lower bound would be best possible: if A is the multiset that contains the point $(1,0)$ with multiplicity $p-1$ and the point $(0,1)$ with multiplicity $k+1$, then the set ΣA is precisely $\big\{(s,t)\colon s \in \mathbb{Z}_p,\, 0 \leq t \leq k+1\big\}$, which has $(k+2)p$ distinct elements. Conjecture 1.6 is actually part of a larger assertion (Conjecture 4.3) concerning lower bounds for sumsets in $\mathbb{Z}_p^m$. We note that Conjecture 1.6 is vacuous for $p = 2$, and we shall see in Section 4 that the

conditions of the more general conjecture render it similarly unremarkable when $p = 2$.

One of our results completely resolves the first two cases $k = 0$ and $k = 1$ of this conjecture:

Theorem 1.7. *Let p be a prime.*

a. *If A is any valid multiset contained in $\mathbb{Z}_p^2$ with $|A| = p$, then $\#\Sigma A \geq 2p$.*

b. *Suppose that $p \geq 5$. If A is any valid multiset contained in $\mathbb{Z}_p^2$ with $|A| = p+1$, then $\#\Sigma A \geq 3p$.*

It turns out that proving part (b) of the theorem requires a certain amount of computation for a finite number of primes (see the remarks following the proof of the theorem in Section 3). Extending the conjecture to larger values of k would require, by our methods, more and more computation to take care of small primes p as k grows. However, we are able to establish the conjecture when p is large enough with respect to k, or equivalently when k is small enough with respect to p:

Theorem 1.8. *Let p be a prime, and let $2 \leq k \leq \sqrt{p/(2\log p + 1)} - 1$ be an integer. If A is any valid multiset contained in $\mathbb{Z}_p^2$ with $|A| = p + k$, then $\#\Sigma A \geq (k + 2)p$.*

A contrapositive version of Theorem 1.8 is also enlightening:

Corollary 1.9. *Let p be a prime, and let $2 \leq k \leq \sqrt{p/(2\log p + 1)} - 1$ be an integer. Let A be a multiset contained in $\mathbb{Z}_p^2 \setminus \{\mathbf{0}\}$ with $|A| = p + k$. If $\#\Sigma A < (k + 2)p$, then there exists a nontrivial subgroup of $\mathbb{Z}_p^2$ that contains at least p points of A, counting multiplicity.*

Our methods of proof stem from two main ideas. First, we will obviously exploit the structure of $\mathbb{Z}_p^2$ as a direct sum of cyclic groups of prime order, within which we can apply the known Lemma 1.4 after using projections. Section 2 contains several elementary lemmas in this vein (see in particular Lemma 2.8). It is important for us to utilize the flexibility coming from the fact that $\mathbb{Z}_p^2$ can be decomposed as the direct sum of two subgroups in many different ways. Second, our methods work best when there exists a single subgroup that contains many elements of the given multiset; however, by selectively replacing pairs of elements with their sums, we can increase the number of elements in a subgroup in a way that improves our lower bounds upon the sumset (see Lemma 3.2). These methods, which appear in Section 3, combine to provide the proofs of Theorems 1.7 and 1.8. Finally, Section 4 contains a generalization of Conjecture 1.6 to higher-dimensional direct sums of $\mathbb{Z}_p$, together with examples demonstrating that the conjecture would be best possible.

Subsequent to the completion of this paper, K. Matomäsi [7] has given an elegant argument which reduces our Conjecture 4.3 to the single subcase (c), and thereby establishes Conjecture 1.6 in full generality, by a result of Peng [9].

2. Sumsets in Abelian Groups and Direct Products

All of the results in this section are valid for general finite abelian groups and have correspondingly elementary proofs, although the last two lemmas seem rather less standard than the first few. In this section, $\mathbf{G}$, $\mathbf{H}$, and $\mathbf{K}$ denote finite abelian groups, and e denotes a group's identity element.

Lemma 2.1. *Let $B_0, B_1, B_2, \ldots, B_j$ be multisets in $\mathbf{G}$, and set $A = B_0 \cup B_1 \cup \cdots \cup B_j$. For each $1 \le i \le j$, specify an element $x_i \in \Sigma B_i$, and set $C = B_0 \cup \{x_1, \ldots, x_j\}$. Then $\Sigma C \subseteq \Sigma A$.*

Proof. For each $1 \le i \le j$, choose a submultiset $D_i \subseteq B_i$ such that the sum of the elements of D_i equals x_i. By definition, every element y of ΣC equals the sum of the elements of some subset E of B_0, plus $\sum_{i \in I} x_i$ for some $I \subseteq \{1, \ldots, j\}$. But then y equals the sum of the elements of $E \cup \bigcup_{i \in I} D_i$, which is an element of ΣA since $E \cup \bigcup_{i \in I} D_i \subseteq B_0 \cup \bigcup_{1 \le i \le j} B_i = A$. □

Lemma 2.2. *Let $A_1, A_2, \ldots, A_j$ be multisets in $\mathbf{G}$, and set $A = A_1 \cup \cdots \cup A_j$. If m is the largest cardinality of a proper subgroup of $\mathbf{G}$, then either $\Sigma A = \mathbf{G}$ or $\#\Sigma A \ge (\sum_{i=1}^{j} \#\Sigma A_i) - (j-1)m$.*

Proof. Since $\Sigma A = \Sigma A_1 + \Sigma A_2 + \cdots + \Sigma A_j$ (viewed as ordinary sets), this follows immediately by inductive application of Kneser's theorem. □

For the remainder of this section, we will be dealing with groups that can be decomposed into a direct sum.

Definition 2.3. A subgroup $\mathbf{H}$ of $\mathbf{G}$ is called an *internal direct summand* if there exists a subgroup $\mathbf{K}$ of $\mathbf{G}$ such that $\mathbf{G}$ is the internal direct sum of $\mathbf{H}$ and $\mathbf{K}$, or in other words, such that $\mathbf{H} \cap \mathbf{K} = \{e\}$ and $\mathbf{H} + \mathbf{K} = \mathbf{G}$. Equivalently, $\mathbf{H}$ is an internal direct summand of $\mathbf{G}$ if there exists a *projection homomorphism* $\pi_{\mathbf{H}} \colon \mathbf{G} \to \mathbf{H}$ that is the identity on $\mathbf{H}$. Note that this projection homorphism does depend on the choice of $\mathbf{K}$ but is uniquely determined by $\pi_{\mathbf{H}}^{-1}(e) = \mathbf{K}$.

Lemma 2.4. *For any homomorphism $f \colon \mathbf{G} \to \mathbf{H}$, and any subset X of $\mathbf{G}$, we have $f(\Sigma X) = \Sigma(f(X))$. In particular, if $\mathbf{H}$ is an internal direct summand of $\mathbf{G}$, then $\pi_{\mathbf{H}}(\Sigma X) = \Sigma(\pi_{\mathbf{H}}(X))$ for any subset X of $\mathbf{G}$.*

Proof. Given $y \in f(\Sigma X)$, there exists $x \in \Sigma X$ such that $f(x) = y$. Hence we can find $x_1, \ldots, x_j \in X$ such that $x_1 + \cdots + x_j = x$, and so $f(x_1 + \cdots + x_j) = y$. But f is a homomorphism, and so $f(x_1) + \cdots + f(x_j) = y$, so that $y \in \Sigma(f(X))$. This shows that $f(\Sigma X) \subseteq \Sigma(f(X))$; the proof of the reverse inclusion is similar. □

Lemma 2.5. *Let $\mathbf{G} = \mathbf{H} \oplus \mathbf{K}$, and let D and E be multisets contained in $\mathbf{H}$ and $\mathbf{K}$, respectively. For any $z \in \mathbf{G}$,*

$$z \in \Sigma(D \cup E) \quad \text{if and only if} \quad \pi_{\mathbf{H}}(z) \in \Sigma D \text{ and } \pi_{\mathbf{K}}(z) \in \Sigma E.$$

Proof. Since $z = \pi_{\mathbf{H}}(z) + \pi_{\mathbf{K}}(z)$, the "if" direction is obvious. For the converse, note that

$$\pi_{\mathbf{H}}(z) \in \pi_{\mathbf{H}}\big(\Sigma(D \cup E)\big) = \Sigma\big(\pi_{\mathbf{H}}(D \cup E)\big)$$

by Lemma 2.4. On the other hand, $\pi_{\mathbf{H}}(D) = D$ and $\pi_{\mathbf{H}}(E) = \{e\}$, and so

$$\pi_{\mathbf{H}}(z) \in \Sigma\big(\pi_{\mathbf{H}}(D) \cup \pi_{\mathbf{H}}(E)\big) = \Sigma\big(D \cup \{e\}\big) = \Sigma D$$

(since the sumset is not affected by whether e is an allowed summand). A similar argument shows that $\pi_{\mathbf{K}}(z) \in \Sigma E$, which completes the proof of the lemma. □

Lemma 2.6. *Let $\mathbf{H}$ and $\mathbf{K}$ be subgroups of $\mathbf{G}$ satisfying $\mathbf{H} \cap \mathbf{K} = \{e\}$. Let D and E be multisets contained in $\mathbf{H}$ and $\mathbf{K}$, respectively. Then $\#\Sigma(D \cup E) = \#\Sigma D \cdot \#\Sigma E$.*

Proof. Notice that every element of $\Sigma(D \cup E)$ is contained in $\mathbf{H} + \mathbf{K}$; therefore we may assume without loss of generality that $\mathbf{G} = \mathbf{H} \oplus \mathbf{K}$. In particular, we may assume that $\mathbf{H}$ and $\mathbf{K}$ are internal direct summands of $\mathbf{G}$, so that the projection maps $\pi_{\mathbf{H}}$ and $\pi_{\mathbf{K}}$ exist and every element $z \in \mathbf{G}$ has a unique representation $z = x + y$ where $x \in \mathbf{H}$ and $y \in \mathbf{K}$; note that $x = \pi_{\mathbf{H}}(z)$ and $y = \pi_{\mathbf{K}}(z)$ in this representation.

To establish the lemma, it therefore suffices to show that $z = \pi_{\mathbf{H}}(z) + \pi_{\mathbf{K}}(z) \in \Sigma(D \cup E)$ if and only if $\pi_{\mathbf{H}}(z) \in \Sigma D$ and $\pi_{\mathbf{K}}(z) \in \Sigma E$; but this is exactly the statement of Lemma 2.5. □

The next lemma is a bit less standard yet still straightforward: in a direct product of two abelian groups, it characterizes the elements of a sumset that lie in a given coset of one of the direct summands.

Lemma 2.7. *Let $\mathbf{H}$ and $\mathbf{K}$ be subgroups of $\mathbf{G}$ satisfying $\mathbf{H} \cap \mathbf{K} = \{e\}$. Let D and E be multisets contained in $\mathbf{H}$ and $\mathbf{K}$, respectively. For any $y \in \mathbf{K}$:*

a. if $y \in \Sigma E$, then $(\mathbf{H} + \{y\}) \cap \Sigma(D \cup E) = \Sigma D + \{y\}$;

b. if $y \notin \Sigma E$, then $(\mathbf{H} + \{y\}) \cap \Sigma(D \cup E) = \emptyset$.

Proof. As in the proof of Lemma 2.6, we may assume without loss of generality that $\mathbf{G} = \mathbf{H} \oplus \mathbf{K}$. Suppose that z is an element of $(\mathbf{H} + \{y\}) \cap \Sigma(D \cup E)$. Since $z \in \mathbf{H} + \{y\}$, we may write $z = x + y$ for some $x \in \mathbf{H}$, whence $\pi_{\mathbf{K}}(z) = \pi_{\mathbf{K}}(x) + \pi_{\mathbf{K}}(y) = e + y = y$. On the other hand, since $z \in \Sigma(D \cup E)$, we see that $y \in \Sigma E$ by Lemma 2.5. In other words, the presence of any element $z \in (\mathbf{H} + \{y\}) \cap \Sigma(D \cup E)$ forces $y \in \Sigma E$, which establishes part (b) of the lemma.

We continue under the assumption $y \in \Sigma E$ to prove part (a). The inclusions $\Sigma D + \{y\} \subseteq \mathbf{H} + \{y\}$ and $\Sigma D + \{y\} \subseteq \Sigma(D \cup E)$ are both obvious, and so $\Sigma D + \{y\} \subseteq (\mathbf{H} + \{y\}) \cap \Sigma(D \cup E)$. As for the reverse inclusion, let $z \in (\mathbf{H} + \{y\}) \cap \Sigma(D \cup E)$ as above; then $\pi_{\mathbf{H}}(z) \in \Sigma D$ by Lemma 2.5, whence $z = \pi_{\mathbf{H}}(z) + \pi_{\mathbf{K}}(z) = \pi_{\mathbf{H}}(z) + y \in \Sigma D + \{y\}$ as required. □

Finally we can establish the lemma that we will make the most use of when we return to the setting $\mathbf{G} = \mathbb{Z}_p^2$ in the next section.

Lemma 2.8. *Let $\mathbf{G} = \mathbf{H} \oplus \mathbf{K}$, and let C be a multiset contained in $\mathbf{G}$. Let $D = C \cap \mathbf{H}$, let $F = C \setminus D$, and let $E = \pi_{\mathbf{K}}(F)$. Then $\#\Sigma C \geq \#\Sigma D \cdot \#\Sigma E$.*

Proof. Lemma 2.6 tells us that $\#\Sigma(D \cup E) = \#\Sigma D \cdot \#\Sigma E$, and so it suffices to show that $\#\Sigma C \geq \#\Sigma(D \cup E)$. We accomplish this by showing that

$$\#\big((\mathbf{H} + \{y\}) \cap \Sigma C\big) \geq \#\big((\mathbf{H} + \{y\}) \cap \Sigma(D \cup E)\big) \tag{1}$$

for all $y \in \mathbf{K}$.

For any $y \in \mathbf{K} \setminus \Sigma E$, Lemma 2.7 tells us that $(\mathbf{H} + \{y\}) \cap \Sigma(D \cup E) = \emptyset$, in which case the inequality (1) holds trivially. For any $y \in \Sigma E$, Lemma 2.7 tells us that $(\mathbf{H} + \{y\}) \cap \Sigma(D \cup E) = \Sigma D + \{y\}$, and so the right-hand side of the inequality (1) equals $\#\Sigma D$.

On the other hand, since $\Sigma E = \Sigma(\pi_{\mathbf{K}}(F)) = \pi_{\mathbf{K}}(\Sigma F)$ by Lemma 2.4, there exists at least one element $z \in \Sigma F$ satisfying $\pi_{\mathbf{K}}(z) = y$; as $\mathbf{G} = \mathbf{H} \oplus \mathbf{K}$, this is equivalent to saying that $z \in \mathbf{H} + \{y\}$. Since $\Sigma D \subseteq \mathbf{H}$, we have $\Sigma D + \{z\} \subseteq \mathbf{H} + \{y\}$ as well. But the inclusion $\Sigma D + \{z\} \subseteq \Sigma D + \Sigma F = \Sigma C$ is trivial, and therefore $\Sigma D + \{z\} \subseteq (\mathbf{H} + \{y\}) \cap \Sigma C$; in particular, the left-hand side of the inequality (1) is at least $\#\Sigma D$. Combined with the observation that the right-hand side equals $\#\Sigma D$, this lower bound establishes the inequality (1) and hence the lemma. □

These lemmas might be valuable for studying sumsets in more general abelian groups. They will prove to be particularly useful for studying sumsets in $\mathbb{Z}_p^2$, however, essentially because there are many ways of writing $\mathbb{Z}_p^2$ as an internal direct sum of two subgroups (which are simply lines through $\mathbf{0}$).

3. Lower Bounds for Sumsets

In this section we establish Theorems 1.7 and 1.8; the proofs employ two combinatorial propositions which we defer to the next section. It would be possible to prove these two theorems at the same time, at the expense of a bit of clarity; however, we find it illuminating to give complete proofs of Theorem 1.7 (the cases $|A| = p$ and $|A| = p + 1$) first, as the proofs will illustrate the methods used to prove the more

general Theorem 1.8. Seeing the limitations of the proof of Theorem 1.7 will also motivate the formulation of our main technical tool, Lemma 3.2.

Throughout this section, A will denote a valid multiset contained in $\mathbb{Z}_p^2$. For any $x \in \mathbb{Z}_p^2$, we let $\langle x \rangle$ denotes the subgroup of $\mathbb{Z}_p^2$ generated by x (that is, the line passing through both the origin $\mathbf{0}$ and x), and we let m_x denote the multiplicity with which x appears in A, so that $|A| = \sum_{x \in \mathbb{Z}_p^2} m_x$. The fact that A is valid means that $m_{\mathbf{0}} = 0$ and $\sum_{t \in \langle x \rangle} m_t < p$ for every $x \in \mathbb{Z}_p^2 \setminus \{\mathbf{0}\}$.

Our first lemma quantifies the notion that we can establish sufficiently good lower bounds for the cardinality of ΣA if we know that there are enough elements of A lying in one subgroup of $\mathbb{Z}_p^2$. Naturally, the method of proof is to partition A into the elements lying in that subgroup and all remaining elements, project the remaining elements onto a complementary subgroup, and then use Lemma 1.4 in each subgroup separately.

Lemma 3.1. *Let A be any valid multiset contained in $\mathbb{Z}_p^2$. Suppose that for some $x \in \mathbb{Z}_p^2 \setminus \{\mathbf{0}\}$,*

$$\sum_{y \in \langle x \rangle} m_y \geq |A| - (p-1). \tag{2}$$

Then $\#\Sigma A \geq (|A| + 2 - p)p$.

Remark. The conclusion is trivial if $|A| < p-1$; also, the fact that A is valid means that the left-hand side of equation (2) is at most $p-1$, and so the lemma is vacuous if $|A| > 2p-2$. Therefore in practice the lemma will be applied only to multisets A satisfying $p-1 \leq |A| \leq 2p-2$.

Proof. Let $D = A \cap \langle x \rangle$; note that $|D| \leq p-1$ since A is a valid multiset, and note also that $|D| = \sum_{y \in \langle x \rangle} m_y \geq |A| - (p-1)$ by assumption. Set $F = A \setminus D$. Choose any nontrivial subgroup $\mathbf{K}$ of $\mathbb{Z}_p^2$ other than $\langle x \rangle$, and set $E = \pi_{\mathbf{K}}(F)$. Then by Lemma 2.8, we know that $\#\Sigma A \geq \#\Sigma D \cdot \#\Sigma E$. By Lemma 1.4 and the fact that $\mathbf{0} \notin D \cup E$, we obtain

$$\begin{aligned} \#\Sigma A &\geq \min\big\{p, 1+|D|\big\} \cdot \min\big\{p, 1+|E|\big\} \\ &= \min\big\{p, 1+|D|\big\} \cdot \min\big\{p, 1+|A|-|D|\big\}, \end{aligned} \tag{3}$$

since $|E| = |F| = |A| - |D|$. The inequalities $|D| \leq p-1$ and $|A| - |D| \leq p-1$ ensure that p is the larger element in both minima, and so we have simply

$$\#\Sigma A \geq (1+|D|)(1+|A|-|D|) = \tfrac{1}{4}|A|^2 + |A| + 1 - \big(|D| - \tfrac{1}{2}|A|\big)^2.$$

The pair of inequalities $|D| \leq p-1$ and $|A| - |D| \leq p-1$ is equivalent to the inequality $\big||D| - \frac{1}{2}|A|\big| \leq p - 1 - \frac{1}{2}|A|$; therefore

$$|\Sigma A| \geq \tfrac{1}{4}|A|^2 + |A| + 1 - \big(p - 1 - \tfrac{1}{2}|A|\big)^2 = (|A| + 2 - p)p,$$

as claimed. □

This lemma alone is sufficient to establish Theorem 1.7.

Proof of Theorem 1.7(a). When $|A| = p$, the right-hand side of the inequality (2) equals 1, and so the inequality holds for any $x \in A$. Therefore Lemma 3.1 automatically applies, yielding $\#\Sigma A \geq (|A|+2-p)p = 2p$ as desired. (In fact essentially the same proof gives the more general statement: if A is a multiset contained in $\mathbb{Z}_p^2 \setminus \{\mathbf{0}\}$ but not contained in any proper subgroup, and $|A| \geq p$, then $\#\Sigma A \geq 2|A|$.) □

Proof of Theorem 1.7(b). We are assuming that $|A| = p+1$. Suppose first that there exists a nontrivial subgroup of $\mathbb{Z}_p^2$ that contains at least two points of A (including possibly two copies of the same point). Choosing any nonzero element x in that subgroup, we see that the inequality (2) is satisfied, and so Lemma 3.1 yields $\#\Sigma A \geq (|A|+2-p)p = 3p$ as desired.

From now on we may assume that there does not exist a nontrivial subgroup of $\mathbb{Z}_p^2$ that contains at least two points of A. Since there are only $p+1$ nontrivial subgroups of $\mathbb{Z}_p^2$, it must be the case that A consists of exactly one point from each of these $p+1$ subgroups; in particular, the elements of A are distinct. We can verify the assertion for $p \leq 11$ by exhaustive computation (see the remarks after the end of this proof), so from now on we may assume that $p \geq 13$.

Suppose first that all sums of pairs of distinct elements from A are distinct. All these sums are elements of ΣA, and thus $\#\Sigma A \geq \binom{p+1}{2} > 3p$ since $p \geq 13$.

The only remaining case is when two pairs of distinct elements from A sum to the same point of $\mathbb{Z}_p^2$. Specifically, suppose that there exist $x_1, y_1, x_2, y_2 \in A$ such that $x_1 + y_1 = x_2 + y_2$. Partition $A = B_0 \cup B_1 \cup B_2$ where $B_1 = \{x_1, y_1\}$ and $B_2 = \{x_2, y_2\}$ and hence $B_0 = A \setminus \{x_1, y_1, x_2, y_2\}$; note that this really is a partition of A, as the fact that $x_1 + y_1 = x_2 + y_2$ forces all four elements to be distinct. Moreover, if we define $z = x_1 + y_1 = x_2 + y_2$, then we know that $z \neq \mathbf{0}$ since x_1 and y_1 are in different subgroups.

Define C to be the multiset $B_0 \cup \{z, z\}$; by Lemma 2.1, we know that $\#\Sigma A \geq \#\Sigma C$. Define $D = C \cap \langle z \rangle$; we claim that $|D| = 3$. To see this, note that A has exactly one point in every nontrivial subgroup, and in particular A has exactly one point in $\langle z \rangle$. Furthermore, that point cannot be x_1 for example, since then $y_1 = z - x_1$ would also be in that subgroup; similarly that point cannot be x_2, y_1, or y_2. We conclude that B_0 has exactly one point in $\langle z \rangle$, whence C has exactly three points in $\langle z \rangle$.

Now define $F = C \setminus D$, so that $|F| = |C| - |D| = (|B_0|+2)-3 = (|A|-4+2)-3 = p-4$. Let $\mathbf{K}$ be any nontrivial subgroup other than $\langle z \rangle$, and set $E = \pi_{\mathbf{K}}(F)$. The lower bounds $\#\Sigma D \geq 4$ and $\#\Sigma E \geq p-3$ then follow from Lemma 1.4. By Lemma 2.8, we conclude that $\#\Sigma C \geq \#\Sigma D \cdot \#\Sigma E = 4(p-3) > 3p$ since $p \geq 13$. □

Remark. The computation that verifies Theorem 1.7(b) for $p \leq 11$ should be done a little bit intelligently, since there are 10^{12} subsets A of $\mathbb{Z}_{11}^2$ (for example) consisting

of exactly one nonzero element from each nontrivial subgroup. We describe the computation in the hardest case $p = 11$. Let us write the elements of $\mathbb{Z}_{11}^2$ as ordered pairs (s, t) with s and t considered modulo 11. By separately dilating the two coordinates of $\mathbb{Z}_{11}^2$ (which does not alter the cardinality of ΣA), we may assume without loss of generality that A contains both $(1, 0)$ and $(0, 1)$. We also know every such A contains a subset of the form $\{(i, i), (j, 2j), (k, 3k), (\ell, 4\ell)\}$ for some integers $1 \le i, j, k, \ell \le 10$. Therefore the cardinality of every such ΣA is at least as large as the cardinality of one of the subsumsets $\Sigma\big(\{(1, 0), (0, 1), (i, i), (j, 2j), (k, 3k), (\ell, 4\ell)\}\big)$.

There are 10^4 such subsumsets, and direct computation shows that all of them have more than 33 distinct elements except for the cases $\Sigma\big(\{(1, 0), (0, 1), \pm(1, 1), \pm(1, 2), \pm(1, 3), \pm(1, 4)\}\big)$, which each contain 32 distinct elements. It is then easily checked that any subsumset of the form $\Sigma\big(\{(1, 0), (0, 1), \pm(1, 1), \pm(1, 2), \pm(1, 3), \pm(1, 4), (m, 5m)\}\big)$ with $1 \le m \le 10$ contains more than 33 distinct elements. This concludes the verification of Theorem 1.7(b) for $p = 11$, and the cases $p \le 7$ are verified even more quickly.

We now foreshadow the proof of Theorem 1.8 by reviewing the structure of the proof of Theorem 1.7(b). In that proof, we quickly showed that the desired lower bound held if there were enough elements of A in the same subgroup. Also, the desired lower bound certainly held if there were enough distinct sums of pairs of elements of A. If however no subgroup contained enough elements of A and there were only a few distinct sums of pairs of elements of A, then we showed that we could find multiple pairs of elements summing to the same point in $\mathbb{Z}_p^2$. Replacing those elements in A with multiple copies of their joint sum, we found that the corresponding subgroup now contained enough elements to carry the argument through.

The following lemma quantifies the final part of this strategy, where we replace j pairs of elements of A with their joint sum and then use our earlier ideas to bound the cardinality of the sumset from below.

Lemma 3.2. *Let A be any valid multiset contained in $\mathbb{Z}_p^2$, and let $z \in \mathbb{Z}_p^2 \setminus \{\mathbf{0}\}$. For any integer j satisfying*

$$0 \le j \le \tfrac{1}{2} \sum_{t \in \mathbb{Z}_p^2 \setminus \langle z \rangle} \min\{m_t, m_{z-t}\}, \tag{4}$$

we have

$$\#\Sigma A \ge \min\left\{p, 1 + j + \sum_{y \in \langle z \rangle} m_y\right\} \min\left\{p, 1 + |A| - 2j - \sum_{y \in \langle z \rangle} m_y\right\}.$$

Remark. This can be seen as a generalization of Lemma 3.1, as equation (3) is the special case $j = 0$ of this lemma.

Proof. Partition $A = B_0 \cup B_1 \cup \cdots \cup B_j$, where for each $1 \le i \le j$, the multiset B_i has exactly two elements, neither contained in $\langle z \rangle$, that sum to z (the complementary submultiset B_0 is unrestricted). The upper bound (4) for j is exactly what is required for such a partition to be possible; the factor of $\frac{1}{2}$ arises because the sum on the right-hand side of (4) double-counts the pairs $(t, z-t)$ and $(z-t, t)$. Then set C equal to B_0 with j additional copies of z inserted. By Lemma 2.1, we know that $\#\Sigma A \ge \#\Sigma C$.

Now let D be the intersection of C with the subgroup $\langle z \rangle$, and let $F = C \setminus D$. Let $\mathbf{K}$ be any nontrivial subgroup other than $\langle z \rangle$, and set $E = \pi_{\mathbf{K}}(F)$. By Lemma 2.8, we know that $\#\Sigma C \ge \#\Sigma D \cdot \#\Sigma E$. However, the number of elements of D (counting multiplicity) is $j + |B_0 \cap \langle z \rangle|$; this is the same as $j + |A \cap \langle z \rangle|$ (since no elements of $B_1, \ldots, B_j$ lie in $\langle z \rangle$), or in other words $j + \sum_{y \in \langle z \rangle} m_y$. Similarly, the number of elements of E (equivalently, of F) is equal to the number of elements of $B_0 \setminus \langle z \rangle$; this is the same as $|A \setminus \langle z \rangle| - 2j$, or in other words $|A| - 2j - \sum_{y \in \langle z \rangle} m_y$. The lower bounds $\#\Sigma D \ge \min\left\{p, 1 + j + \sum_{y \in \langle z \rangle} m_y\right\}$ and $\#\Sigma E \ge \min\left\{p, 1 + |A| - 2j - \sum_{y \in \langle z \rangle} m_y\right\}$ then follow from Lemma 1.4; the chain of inequalities $\#\Sigma A \ge \#\Sigma C \ge \#\Sigma D \cdot \#\Sigma E$ establishes the lemma. □

We are now ready to use Lemma 3.2 to establish Conjecture 1.6 when $|A| = p + k$, for all but finitely many primes p depending on k. Let $H_k = 1 + \frac{1}{2} + \cdots + \frac{1}{k}$ denote the kth harmonic number.

Theorem 3.3. *Let $k \ge 2$ be any integer, and let A be any valid multiset contained in $\mathbb{Z}_p^2$ such that $|A| = p + k$. If $p \ge 4(k+1)^2 H_k - 2k$, then $\#\Sigma A \ge (k+2)p$.*

Remark. Using the elementary bound $H_k \le \gamma + \log(k+1)$, where γ denotes the Euler–Mascheroni constant, we see that Theorem 3.3 holds as long as $p \ge 4(k+1)^2(\gamma + \log(k+1))$. Theorem 1.8 can thus be readily deduced from Theorem 3.3 as follows: If $k + 1 \le \sqrt{p/(2\log p + 1)}$ then $p \ge 4(k+1)^2(\frac{1}{4} + \frac{1}{2}\log p)$. In this case we have $p \ge (1 + 2\log 2)(k+1)^2$, whence $\log p \ge \frac{4}{5} + 2\log(k+1)$ and $\frac{1}{4} + \frac{1}{2}\log p \ge \gamma + \log(k+1)$.

Proof. If there are $k+1$ elements of A in some nontrivial subgroup, then we are done by Lemma 3.1. Therefore we may assume that there are at most k points in each subgroup; in particular, $m_x \le k$ for all $x \in \mathbb{Z}_p^2$. We now argue that if ΣA is small, then there must be lots of pairs of elements of A that add to the same element of $\mathbb{Z}_p^2$, at which point we will be able to invoke Lemma 3.2. We may assume that $\Sigma A \ne \mathbb{Z}_p^2$, for otherwise we are done immediately.

For each $1 \le i \le k$, we define the level set $A_i = \{x \in \mathbb{Z}_p^2 : m_x \ge i\}$. Notice that A can be written precisely as the multiset union $A_1 \cup A_2 \cup \cdots \cup A_k$, and so $\sum_{i=1}^{k} \#A_i = |A| = p + k$. Let B_i be the multiset formed by the sums of pairs of elements of A_i not in the same subgroup:

$$B_i = \left\{x + y \colon x, y \in A_i,\ \langle x \rangle \ne \langle y \rangle\right\}.$$

Note that $\mathbf{0} \notin B_i$ (the restriction $\langle x\rangle \neq \langle y\rangle$ ensures that $x \neq -y$) and that every element of B_i occurs with even multiplicity (the restriction $\langle x\rangle \neq \langle y\rangle$ ensures that $x \neq y$). It is not hard to estimate the relative sizes of $\#A_i$ and $|B_i|$: for each $x \in A_i$ there are at most k elements of A lying in the subgroup $\langle x\rangle$. Since each such x occurs with multiplicity at least i in A, there are at most k/i distinct values of y excluded by the condition $\langle x\rangle \neq \langle y\rangle$. Hence $|B_i| \geq \#A_i(\#A_i - k/i)$, which implies that

$$\#A_i \leq \frac{k}{i} + \sqrt{|B_i|}. \tag{5}$$

Since $\sum_{i=1}^{k} \#A_i$ is fixed, this shows that $|B_i|$ cannot be very small on average. At the same time, $\#B_i$ cannot get very large: if $\sum_{i=1}^{k} \#B_i \geq (2k+1)p$, then (under our assumption that $\Sigma A \neq \mathbb{Z}_p^2$) Lemma 2.2 already yields

$$\#\Sigma A \geq \sum_{i=1}^{k} \#\Sigma A_i - (k-1)p > \sum_{i=1}^{k} \#B_i - (k-1)p \geq (k+2)p.$$

where the middle inequality holds because $B_i \subseteq \Sigma A_i$. We may therefore assume henceforth that

$$\sum_{i=1}^{k} \#B_i < (2k+1)p. \tag{6}$$

Let us now introduce the weighted height parameter

$$\eta = \max_{1 \leq i \leq k} \left\{ \frac{i|B_i|}{2\#B_i} : \#B_i > 0 \right\}. \tag{7}$$

We shall show shortly that $\eta > k+1$. Assuming so, then for some $1 \leq i \leq k$, we have

$$\frac{|B_i|}{2\#B_i} > \frac{k+1}{i},$$

so by the pigeonhole principle, there exists some $z \in B_i$ (in particular $z \neq \mathbf{0}$) occurring with multiplicity greater than $2(k+1)/i$; since this multiplicity is an even integer, it must be at least $2(k+2)/i$. For each solution $x+y=z$ corresponding to an occurrence of z in B_i, we have by construction that $x, y \notin \langle z\rangle$ and $m_x, m_y \geq i$, so for this particular choice of z,

$$\tfrac{1}{2} \sum_{t \in \mathbb{Z}_p^2 \setminus \langle z\rangle} \min\{m_t, m_{z-t}\} \geq k+2.$$

Furthermore, $\sum_{y \in \langle z\rangle} m_y \leq k$ by assumption. Therefore we are free to apply Lemma 3.2 with $j = (k+2) - \sum_{y \in \langle z\rangle} m_y$, which gives the lower bound

$$\#\Sigma A \geq \min\{p, k+3\} \min\left\{p, p-k-3+\sum_{y \in \langle z\rangle} m_y\right\} \geq (k+3)(p-k-3) \geq (k+2)p$$

(the last step used the inequality $p \geq (k+3)^2$, which holds under the hypotheses of the theorem).

It remains only to verify that $\eta > k+1$. Summing the inequality (5) over all $1 \leq i \leq k$ yields

$$p + k = \sum_{i=1}^{k} \#A_i \leq kH_k + \sum_{i=1}^{k} \sqrt{|B_i|} \leq kH_k + \sqrt{2\eta} \sum_{i=1}^{k} \sqrt{\frac{\#B_i}{i}},$$

using the definition (7) of η. We estimate the rightmost sum using Cauchy–Schwarz together with the inequality (6):

$$\sum_{i=1}^{k} \sqrt{\frac{\#B_i}{i}} \leq \left(\sum_{i=1}^{k} \#B_i \right)^{1/2} \left(\sum_{i=1}^{k} \frac{1}{i} \right)^{1/2} < \sqrt{(2k+1)pH_k}.$$

Combining the previous two inequalities gives $p + k - kH_k < \sqrt{\eta(4k+2)pH_k}$, so that

$$\eta > \frac{(p+k-kH_k)^2}{(4k+2)pH_k} > \frac{p(p+2(k-kH_k))}{(4k+2)pH_k} = \frac{(p+2k)-2kH_k}{(4k+2)H_k} \geq \frac{4(k+1)^2H_k - 2kH_k}{(4k+2)H_k}$$

by the hypothesis on the size of p. In other words,

$$\eta > \frac{2(k+1)^2 - k}{2k+1} = k + 1 + \frac{1}{2k+1},$$

which completes the proof of the theorem. □

4. A Wider Conjecture

As mentioned earlier, Conjecture 1.6 is just one part of a more far-reaching conjecture concerning sumsets of multisets in $\mathbb{Z}_p^m$. Before formulating that wider conjecture, we must expand the definition of a valid multiset to $\mathbb{Z}_p^m$.

Definition 4.1. Let p be an odd prime, and let m be a positive integer. A multiset A contained in $\mathbb{Z}_p^m$ is *valid* if:

- $\mathbf{0} \notin A$; and
- for each $1 \leq d \leq m$, every subgroup of $\mathbb{Z}_p^m$ that is isomorphic to $\mathbb{Z}_p^d$ contains fewer than dp points of A, counting multiplicity.

When $m = 1$, a multiset contained in $\mathbb{Z}_p$ is valid precisely when it does not contain 0; when $m = 2$ and $|A| < 2p$, this definition of valid agrees with Definition 1.5 for multisets contained in $\mathbb{Z}_p^2$. Note that in particular, Definition 4.1(b) implies

that every valid multiset contained in $\mathbb{Z}_p^m$ has at most $mp-1$ elements, counting multiplicity. We now give an example showing that this upper bound $mp-1$ can in fact be achieved. Throughout this section, let $\{x_1,\ldots,x_m\}$ denote a generating set for $\mathbb{Z}_p^m$, and let $\mathbf{K}_d = \langle x_1,\ldots,x_d\rangle$ denote the subgroup of $\mathbb{Z}_p^m$ generated by $\{x_1,\ldots,x_d\}$, so that $\mathbf{K}_d \cong \mathbb{Z}_p^d$.

Example 4.2. Let A_1 be the multiset consisting of $p-1$ copies of x_1; for $2 \le j \le m$ let $A_j = \{x_j + ax_1 \colon 0 \le a \le p-1\}$; and define $B_m = \bigcup_{j=1}^m A_j$. Then $|B_m| = (p-1)+(m-1)p = mp-1$ and $\mathbf{0} \notin B_m$. To verify that B_m is a valid subset of $\mathbb{Z}_p^m$, let $\mathbf{H}$ be any subgroup of $\mathbb{Z}_p^m$ that is isomorphic to $\mathbb{Z}_p^d$; we need to show that B_m contains fewer than dp points of $\mathbf{H}$.

First suppose that $x_1 \notin \mathbf{H}$, which implies that $bx_1 \notin \mathbf{H}$ for every nonzero multiple bx_1 of x_1. Then for each $2 \le j \le m$, at most one of the elements of A_j can be in $\mathbf{H}$, since the difference of any two such elements is a nonzero multiple of x_1. Therefore $|B_m \cap \mathbf{H}| = \ell$ for some $1 \le \ell \le m-1$, and in fact all ℓ of these elements are of the form $x_j + ax_1$ for ℓ distinct values of j. Since no such element is in the subgroup spanned by the others, we conclude that $d \ge \ell$, and so the necessary inequality $|B_m \cap \mathbf{H}| = \ell \le d < dp$ is amply satisfied.

Now suppose that $x_1 \in \mathbf{H}$. Then for each $2 \le j \le m$, either all or none of the elements of A_j are in $\mathbf{H}$. By reindexing the x_i, we may choose an integer $1 \le \ell \le m$ such that $\mathbf{H}$ contains $A_1 \cup \cdots \cup A_\ell$ and is disjoint from $A_{\ell+1} \cup \cdots \cup A_m$. In particular, $|B_m \cap \mathbf{H}| = (p-1)+(\ell-1)p = \ell p - 1$. But $\mathbf{H}$ contains $\{x_1,\ldots,x_\ell\}$ and hence $d \ge \ell$, so that $\ell p - 1 \le dp - 1$ as required.

We may now state our wider conjecture; Conjecture 1.6 is the special case $q = 1$ of part (a) of this conjecture.

Conjecture 4.3. Let p be an odd prime. Let m be a positive integer, and let A be a valid multiset of $\mathbb{Z}_p^m$ with $|A| \ge p$. Write $|A| = qp + k$ with $0 \le k \le p-1$.

a. If $0 \le k \le p-3$, then $\#\Sigma A \ge (k+2)p^q$.

b. If $k = p-2$, then $\#\Sigma A \ge p^{q+1} - 1$.

c. If $k = p-1$, then $\#\Sigma A \ge p^{q+1}$.

In particular, if $|A| = mp-1$ then $\Sigma A = \mathbb{Z}_p^m$.

We remark that the quantity dp in Definition 4.1, bounding the number of elements in a valid multiset that can lie in a rank-d subgroup, has been carefully chosen in light of this conjecture: by Conjecture 4.3(c), any valid multiset A with at least $dp-1$ elements counting multiplicity must satisfy $\#\Sigma A \ge p^d$. In particular, if A is a valid multiset contained in a subgroup $\mathbf{H} < \mathbb{Z}_p^m$ that is isomorphic to $\mathbb{Z}_p^d$, then $|A| \ge dp-1$ implies that $\Sigma A = \mathbf{H}$. Therefore allowing dp elements

in such a subgroup would always be "wasteful". Of course, the validity of Definition 4.1 for rank-d subgroups depends crucially upon the truth of Conjecture 4.3(c) for $q = d - 1$.

The conjecture is restricted to multisets A with $|A| \geq p$ because we already know the truth for smaller multisets, for which the definition of "valid" is simply the condition that $\mathbf{0} \notin A$: when $|A| \leq p - 1$, the best possible lower bound is $\#\Sigma A \geq |A| + 1$ as in Lemma 1.4. We remark that Peng [9, Theorem 2] has proved Conjecture 4.3(c) in the case $m = 2$ and $q = 1$, even under a slightly weaker hypothesis; in other words, he has shown that if A is a valid multiset contained in $\mathbb{Z}_p^2$ with $|A| = 2p - 1$, then $\Sigma A = \mathbb{Z}_p^2$. (Mann and Wou [5] have proved in the case that A is actually a set—that is, a multiset with distinct elements—that $\#A = 2p - 2$ suffices to force $\Sigma A = \mathbb{Z}_p^2$.) Peng considers the higher-rank groups $\mathbb{Z}_p^m$ as well, but the multisets he allows (see [10, Theorem 1]) form a much wider class than our valid multisets for $q \geq 2$, and so $|A|$ must be much larger than required by Conjecture 4.3(c) in order to imply $\Sigma A = \mathbb{Z}_p^m$ when $m \geq 3$. Finally, we mention that we have completely verified Conjecture 4.3 by exhaustive computation for the groups $\mathbb{Z}_p^2$ with $p \leq 7$ and also for the group $\mathbb{Z}_3^3$.

It is easy to see that all of the lower bounds in Conjecture 4.3(a), if true, would be best possible. Given $q \geq 1$ and $0 \leq k \leq p - 3$, let A' be any valid multiset contained in $\mathbf{K}_q$ with $|A'| = qp - 1$ (such as the one given in Example 4.2 with $m = q$), and let A be the union of A' with $k + 1$ copies of x_{q+1}. Then $\Sigma A = \{y + ax_{q+1} \colon y \in \Sigma A', 0 \leq a \leq k + 1\}$ and thus $\#\Sigma A' = (k + 2)\#\Sigma A \leq (k + 2)p^q$ since ΣA is contained in $\mathbf{K}_q$. Similarly, the fact that there exists a valid multiset contained in $\mathbf{K}_{q+1}$ with $qp + (q - 1) = (q + 1)p - 1$ elements (such as the one given in Example 4.2 with $m = q + 1$) shows that the lower bound in Conjecture 4.3(c) would be best possible, since the sumset of this multiset would still be contained in $\mathbf{K}_{q+1}$ and thus would have at most p^{q+1} distinct elements.

The lower bound in Conjecture 4.3(b) might seem counterintuitive, especially in comparison with the pattern established in Conjecture 4.3(a). However, we can give an explicit example showing that the lower bound $p^{q+1} - 1$ for $\#\Sigma A$ cannot be increased:

Example 4.4. When p is an odd prime, define B'_m to be the set B_m from Example 4.2 with one copy of x_1 removed, so that B'_m contains x_1 with multiplicity only $p - 2$. Since B_m is a valid multiset contained in $\mathbb{Z}_p^m$, so is B'_m. We have $|B'_m| = |B_m| - 1 = (mp - 1) - 1 = (m - 1)p + (m - 2)$, and we claim that $-x_1 \notin \Sigma B'_m$; this will imply that $\#\Sigma B'_m \leq p^m - 1$, and so the lower bound for $\#\Sigma A$ in Conjecture 4.3(b) cannot be increased. (In fact it is not hard to show that every other element of $\mathbb{Z}_p^m$ is in $\Sigma B'_m$, and so $\#\Sigma B'_m$ is exactly equal to $p^m - 1$.)

Suppose for the sake of contradiction that $-x_1 \in \Sigma B'_m$, and let C be a submultiset of B'_m such that $-x_1 = \sum_{y \in C} y$. For each $2 \leq j \leq m$, define $\ell_j = |C \cap A_j| =$

$\#\big(C \cap \{x_j + ax_1 : 0 \le a \le p-1\}\big)$. Then

$$-x_1 = \sum_{y \in C} y = tx_1 + \ell_2 x_2 + \ell_3 x_3 + \cdots + \ell_m x_m$$

for some integer t. It follows from this equation that each ℓ_j must equal either 0 or p. However, if $\ell_j = p$ then

$$\sum_{y \in C \cap A_j} y = \sum_{0 \le a \le p-1} (x_j + ax_1) = px_j + \frac{p(p-1)}{2} x_1 = \mathbf{0}$$

(since p is odd). So in either case, if $s = |C \cap A_1|$ is the multiplicity with which x_1 appears in C, then

$$-x_1 = \sum_{y \in C} y = sx_1 + \sum_{j=2}^{m} \sum_{y \in C \cap A_j} y = sx_1 + \mathbf{0} + \cdots + \mathbf{0}.$$

This is a contradiction, however, since s must lie between 0 and $p-2$. Therefore $-x_1$ is indeed not an element of $\Sigma B'_m$, as claimed.

A naive application of Lemma 2.8 seems to relatively ineffective in extending our main theorem to partial results toward Conjecture 4.3 for the higher-rank groups $\mathbb{Z}_p^m$, say using the decomposition $\mathbb{Z}_p^m = \mathbb{Z}_p^{m-1} \oplus \mathbb{Z}_p$ inductively. If a multiset C is very well-distributed across subgroups, then we expect the cardinality of $D = C \cap \mathbb{Z}_p^{m-1}$ to be about p times smaller than that of C. Without exploiting the structure of C as in the proof of Theorem 1.8, we would thus require $|C| \gg p^{m-1}$ elements to obtain $\Sigma C = \mathbb{Z}_p^m$. This is comparable to the aforementioned results of Peng, rather than the linear growth in m which is predicted by Conjecture 4.3.

The line of questioning in this section turns out to be uninteresting when $p = 2$: when the multiset A does not contain $\mathbf{0}$, the condition that no rank-1 subgroup of $\mathbb{Z}_2^m$ contain 2 points of A is simply equivalent to A not containing any element with multiplicity greater than 1. It is easy to check that if A consists of any q points in $\mathbb{Z}_2^m$ that do not lie in any subgroup isomorphic to $\mathbb{Z}_2^{q-1}$, then ΣA fills out the entire rank-q subgroup generated by A. In other words, the analogous definition of "valid" for multisets in $\mathbb{Z}_2^m$ would simply be a set of q points that generate a rank-q subgroup of $\mathbb{Z}_2^m$, and we would always have $\#\Sigma A = 2^{|A|} = 2^{\#A}$ for valid (multi)sets in $\mathbb{Z}_2^m$.

Acknowledgments The collaboration leading to this paper was made possible thanks to Jean–Jacques Risler, Richard Kenyon, and especially Ivar Ekeland; the authors also thank the University of British Columbia and the Institut d'Études Politiques de Paris for their undergraduate exchange program. The first author thanks Andrew Granville for conversations that explored this topic and eventually led to the formulation of the conjectures herein.

References

[1] A. Cauchy, *Recherches sur les nombres*, Jour. Ecole Polytechn. **9** (1813), 99–116.

[2] H. Davenport, *On the addition of residue classes*, J. London Math. Soc. **10** (1935), 30–32.

[3] H. Davenport, *A historical note*, J. London Math. Soc. **22** (1947), 100–101.

[4] M. Kneser, *Abschätzungen der asymptotischen Dichte von Summenmengen*, Math. Z. **58** (1953), 459–484.

[5] H.B. Mann and Y.F. Wou, *An addition theorem for the elementary abelian group of type* (p, p), Monatsh. Math. **102** (1986), 273–308.

[6] G. Martin, *Dense Egyptian fractions*, Trans. Amer. Math. Soc. **351** (1999), no. 9, 3641–3657.

[7] K. Matomäki, *On sumsets of multisets in* $\mathbb{Z}_p^m$, Electr. J. Combin. **20** (2013), P30.

[8] M.B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, vol. 165 of Graduate Texts in Mathematics, Springer–Verlag, New York, 1996.

[9] C. Peng, *Addition theorems in elementary abelian groups, I*, J. Number Th. **27** (1987), 46–57.

[10] C. Peng, *Addition theorems in elementary abelian groups, II*, J. Number Th. **27** (1987), 58–62.