0 Sumsets

Let G be an abelian group, let $A, B \subseteq G$ and let $g \in G$. We define $A + B = \{a + b : a \in A \text{ and } b \in B\}$ and we call any such set a *sumset.* Continuing with this theme, we define $-A = \{-a : a \in A\}$ and $A + g = A + \{g\}$. Any set of the form $A + g$ is called a *shift* of A, and we also call the operation of replacing A by $A + g$ the act of *shifting* A. Our main initial focus will be on small sumsets, namely, we are interested in the following two problems.

1. How small can $|A + B|$ be? (say in terms of $|A|$, $|B|$)

2. if $|A + B|$ is small, what can be said about the structure of A, B?

One familiar type of set which gives rise to small sumsets is an arithmetic progression. We define a set $A \subseteq G$ to be an *arithmetic progression* with *difference* g if there exist a positive integer n and $a \in A$ so that $A = \{a + ig : 1 \leq i \leq n\}$. If A, B are arithmetic progressions with difference g and respective sizes m, n, then $A + B$ will be an arithmetic progression with difference g and size $\leq m + n - 1$ (strict inequality can be achieved if say $A = B$ is a finite subgroup of G generated by g).

1 Sumsets in $\mathbb{Z}$

Our goal here will be to provide answers to questions 1 and 2 from the previous section in the special case when the group is $\mathbb{Z}$. We begin with an easy observation which resolves the first question in this case.

Observation 1.1 *If A, B are nonempty finite subsets of $\mathbb{Z}$, then $|A + B| \geq |A| + |B| - 1$*

Proof: Shifting either A or B only shifts the sumset $A + B$, it has no effect on the sizes of our sets. Thus, we are free to shift A and B, and therefore may assume that 0 is the maximum element in A and 0 is the minimum element in B. Then $A \cap B = \{0\}$ and $A \cup B \subseteq A + B$ so we have $|A + B| \geq |A| + |B| - 1$ as desired. $\square$

Our next theorem gives a characterization of those pairs $A, B \subseteq \mathbb{Z}$ which satisfy the bound given in the previous theorem with equality.

Observation 1.2 *Let $A, B \subseteq \mathbb{Z}$ be nonempty finite subsets of $\mathbb{Z}$. If $|A+B| = |A|+|B|-1$, then one of the following holds:*

- $|A| = 1$, *or* $|B| = 1$.
- *A, B are arithmetic progressions with a common difference.*

Proof: Let $A = \{a_1, a_2, \ldots, a_m\}$ and $B = \{b_1, b_2, \ldots, b_n\}$ with $a_1 < a_2 \ldots < a_m$ and $b_1 < b_2 \ldots < b_n$. If $m = 1$ or $n = 1$ then we have nothing to prove, so we may assume that $m, n \geq 2$. Now consider the integer lattice $\mathbb{Z} \times \mathbb{Z}$. We call a sequence of points $(q_1, r_1), (q_2, r_2), \ldots, (q_\ell, r_\ell)$ from this lattice a *North/East walk* if $(q_{j+1}, r_{j+1}) - (q_j, r_j) \in \{(1,0),(0,1)\}$ for every $1 \leq j \leq n-1$. Observe that if $(q_1, r_1), \ldots, (q_{m+n-1}, r_{m+n-1})$ is a North/East walk from $(1,1)$ to (m,n), then we have $a_{q_1}+b_{r_1} < a_{q_2}+b_{r_2} < \ldots a_{q_{m+n-1}}+b_{r_{m+n-1}}$, so these are $m+n-1$ distinct points in the sumset $A+B$ and therefore this list contains the entire sumset. For every $1 \leq i \leq m-1$ and $1 \leq j \leq n-1$ there is a North/East walk from $(1,1)$ to (m,n) whose $(i+j-1)^{st}$ entry is $(i+1,j)$ and one whose $(i+j-1)^{st}$ entry is $(i, j+1)$ so we must have $a_i + b_{j+1} = a_{i+1} + b_j$ (as otherwise our sumset would have size $\geq m+n$). Equivalently, $a_{i+1} - a_i = b_{j+1} - b_j$. It follows immediately from this that A and B are arithmetic progressions with a common difference. $\square$

2 Sumsets in $\mathbb{Z}_p$

For every positive integer n, we let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. Throughout this section we shall assume that p is a prime. Next we have a classical result which gives a natural lower bound on the size of a sumset in $\mathbb{Z}_p$.

Theorem 2.1 (Cauchy-Davenport) *If $A, B \subseteq \mathbb{Z}_p$ are nonempty, then $|A+B| \geq \min\{p, |A|+|B|-1\}$.*

Proof: We proceed by induction on $|A|$. The result holds trivially if $|A| = 1$ or if $|B| = p$, so we may assume that $|A| > 1$ and $|B| < p$. By shifting A, we may assume that $\{0, g\} \subseteq A$ for some $g \neq 0$. Since $\emptyset \neq B \neq \mathbb{Z}_p$, there must exist an integer n so that $ng \in B$ and $(n+1)g \notin B$, so by shifting B we may assume that $0 \in B$ and $g \notin B$. Now consider the sets $A \cap B$ and $A \cup B$ and note that $(A \cap B) + (A \cup B) \subseteq A + B$. Since $0 \in A \cap B$ and $g \notin A \cap B$ we have that

$A \cap B$ is a proper nonempty subset of A. Thus, by applying induction to the pair $A \cap B$, $A \cup B$ we find $|A+B| \geq |(A\cap B)+(A\cup B)| \geq \min\{p, |A\cap B|+|A\cup B|-1\} = \min\{p, |A|+|B|-1\}$ as desired. □

Motivated by Observation 1.1 and Theorem 2.1, we define a pair of nonempty finite subsets (A, B) of an abelian group G to be *critical* if $|A + B| < |A| + |B|$. We will be interested in understanding the structure of critical pairs. Let us begin by observing a rather trivial type of critical pair. If G is finite and $A, B \subseteq G$ satisfy $|A| + |B| > |G|$, then for every $g \in G$ we have $B \cap (g - A) \neq \emptyset$ and it follows that $A + B = G$. So every pair (A, B) with $|A| + |B| > |G|$ is critical, but has sumset equal to the entire group. We shall call every such pair *trivial*.

Let $(A, B) \subseteq G \times G$ be nontrivial and critical, and define $C = G \setminus -(A + B)$. Two key properties of this triple are indicated below.

$$0 \notin A + B + C$$
$$|A| + |B| + |C| = |A| + |B| + |G| - |A + B| > |G|$$

It follows from the first equation above that $-A$ is disjoint from $B+C$ so $|B+C| \leq |G|-|A|$. Combining this with the second equation we get $|B + C| \leq |G| - |A| < |B| + |C|$, so we deduce that (B, C) is critical. Similarly (C, A) is critical. Thus, every critical pair is actually part of a triple of subsets, any two of which form a critical pair. Next we will establish a bit of terminology for these objects.

We define a triple of sets (A, B, C) of a finite group G to be a *trio* if $0 \notin A + B + C$. We call (A, B, C) *nontrivial* if $A, B, C \neq \emptyset$ and we call it *critical* if $|A| + |B| + |C| > |G|$. It follows from the previous discussion that every nontrivial critical pair extends to a nontrivial critical trio, and every pair of sets from a nontrivial critical trio forms a nontrivial critical pair. With these definitions in place, we now have the following corollary of Theorem 2.1.

Corollary 2.2

1. *If (A, B) is a nontrivial critical pair in $\mathbb{Z}_p$, then $|A + B| = |A| + |B| - 1$*
2. *If (A, B, C) is a nontrivial critical trio in $\mathbb{Z}_p$, then $|A| + |B| + |C| = p + 1$*

Proof: Part 1 follows immediately from Theorem 2.1. For part 2, observe that since $-C$ is disjoint from $A+B$ it must have size $\leq p-|A+B| = p-|A|+|B|+1$. Thus $|A|+|B|+|C| \leq p+1$ and to be critical, we must have equality. □

Our next goal is to prove a theorem of Vosper which characterizes the critical trios (and thus the critical pairs) in $\mathbb{Z}_p$. Before proving this, we will require a couple of relatively simple lemmas.

Lemma 2.3 *If (A, B, C) is a nontrivial critical trio in $\mathbb{Z}_p$, and A is a nontrivial arithmetic progression with difference g, then B and C are arithmetic progressions with difference g.*

Proof: Consider the sets $A' = A \cap (A+g)$ and $B' = B \cup (B-g)$. By constuction $A' + B' = (A \cap (A+g)) + (B \cup (B-g)) \subseteq A + B$. Thus, by Cauchy-Davenport, we have

$$\begin{aligned} |A| + |B| &= |A+B| + 1 \\ &\geq |A' + B'| + 1 \\ &\geq |A'| + |B'| \\ &= (|A| - 1) + |B \cup (B-g)| \end{aligned}$$

So $|B \cup (B-g)| \leq |B| + 1$. It follows immediately from this that B is an arithmetic progression with difference g as desired. A similar argument shows that C has the same property. □

A subset A of an abelian group G is called a *unique difference set* if the only solutions to the equation $a - a' = b - b'$ with $a, a', b, b' \in A$ are those for which $a = b$ and $a' = b'$.

Lemma 2.4 *Let A, B be finite subsets of an abelian group G and assume that $k \leq |A| \leq |B|$. If B is a unique difference set, then $|A+B| \geq k|B| - k(k-1)/2$.*

Proof: Choose distinct elements $a_1, a_2, \ldots, a_k \in A$, let $1 \leq i \leq k$ and consider the set $B_i = B + a_i \setminus B + \{a_1, a_2, \ldots, a_{i-1}\}$. Since B is a unique difference set $|(B+a_i) \cap (B+a_j)| \leq 1$

for ever $1 \leq j < i$ so $|B_i| \geq |B| - (i-1)$. Now we have

$$\begin{aligned} |A+B| &\geq |\{a_1, a_2, \ldots, a_k\} + B| \\ &\geq \sum_{i=1}^{k} |B_i| \\ &\geq k|B| - \sum_{i=1}^{k}(i-1) \\ &= k|B| - k(k-1)/2 \end{aligned}$$

This completes the proof. □

We are now ready to characterizes the critical trios in $\mathbb{Z}_p$.

Theorem 2.5 (Vosper) *If p is prime and (A, B, C) is a nontrivial crtitical trio in $\mathbb{Z}_p$, then one of the following holds:*

- $|A| = 1$, $|B| = 1$, *or* $|C| = 1$.
- A, B, C *are arithmetic progressions with a common difference.*

Proof: We proceed by induction on $|A|$ and for fixed $|A|$ by induction on $|B|$. If one of A, B, C has size 1, then we are finished. Similarly, if one of A, B, C has size 2, then the result follows from Lemma 2.3. Since $\mathbb{Z}_p$ is commutative, we may now assume that $3 \leq |A| \leq |B| \leq |C|$.

If B is a unique difference set, then by applying Lemma 2.4 with $k = 3$ we find $|A+B| \geq 3|B| - 3 \geq |B| + |A|$, a contradiction. Thus B is not a unique difference set, so we may choose $g \in G$ so that $B' = B \cap (B+g)$ has size ≥ 2. Set $C' = C \cup (C-g)$, $B'' = B \cup (B+g)$ and $C'' = C \cap (C-g)$. By construction, $B', C', B'' \neq \emptyset$. If $C'' = C \cap (C-g) = \emptyset$ then we may choose $b_1, b_2 \in B$ with $b_1 + g = b_2$ (since $B \cap (B+g) \neq \emptyset$) and we find $|B+C| \geq |\{b_1, b_2\} + C| \geq 2|C| \geq |B| + |C|$, a contradiction. Therefore, C'' is also nonempty. By construction, $B' + C' \subseteq B + C$ and $B'' + C'' \subseteq B + C$ so we find that (A, B', C') and (A, B'', C'') are both nontrivial trios. Furthermore, $(|A| + |B'| + |C'|) + (|A| + |B''| + |C''|) = 2|A| + 2|B| + 2|C| = 2p + 2$ so both (A, B', C') and (A, B'', C'') are critical trios. Since B' is a proper subset of B with size ≥ 2, by applying the theorem inductively to (A, B', C') we deduce that A is a nontrivial arithmetic progression. It now follows from Lemma 2.3 that A, B, C are arithmetic progressions with a common difference, as required. □