1 Small Sumsets

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).

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 $|A + B| \geq |A| + |B| - 1$. Furthermore, 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 $b_1 < b_2 \ldots < b_n$. To see that $|A + B| \geq |A| + |B| - 1$, note that $A + B$ contains the following (distinct) elements

$$a_1 + b_1 < a_2 + b_1 \ldots < a_m + b_1 < a_m + b_2 \ldots < a_m + b_n.$$

Now suppose that $|A + B| = |A| + |B| - 1$ and $|A|, |B| \geq 2$. In this case, we can consider forming more general sequences of elements of $A+B$ by starting from a_1+b_1 and stepping to a_m+b_n by incrementing one of the two indices at each step. Since $|A+B| = |A|+|B|-1$ we must get the same sequence of elements each time. It is possible to move from the element $a_1 + b_i$ to either $a_2 + b_i$ or $a_1 + b_{i+1}$, so it must be that $b_{i+1} - b_i = a_2 - a_1$ holds for every $1 \leq i \leq n - 1$. A similar argument shows that $a_{i+1} - a_i = b_2 - b_1$ for every $1 \leq i \leq m - 1$ and this gives us the second outcome. $\square$

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 1.2 (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. $\square$

Subsequence Sums

In this section we will establish the following delightful result.

Theorem 1.3 (Erdös-Ginzburg-Ziv) *If $a_1, \ldots, a_{2n-1}$ is a sequence of integers, there is a subsequence of length n which sums to a multiple of n.*

The proof of this result will call upon an application of Cauchy-Davenport, which we distill as follows.

Lemma 1.4 *Let p be prime and let $a_1, \ldots, a_{2p-1} \in \mathbb{Z}_p$. Then either one element appears at least p times in this sequence, or every $g \in \mathbb{Z}_p$ is the sum of some p-term subsequence.*

Proof: For simplicity, we will identify $\mathbb{Z}_p$ with the representatives $0, 1, \ldots, p-1$. We may then assume that our sequence is ordered so that $a_1 \leq a_2 \ldots \leq a_{2p-1}$. Now consider the following sumset.

$$\{a_1, a_p\} + \{a_2, a_{p+1}\} + \ldots + \{a_{p-1}, a_{2p-2}\} + \{a_{2p-1}\}$$

If each set of the form $\{a_i, a_{i+p-1}\}$ has size two, then the Cauchy-Davenport Theorem implies that the above expression is equal to $\mathbb{Z}_p$ and we are finished. On the other hand if $\{a_i, a_{i+p-1}\}$ has size one, then $a_i, a_{i+1}, \ldots, a_{i+p-1}$ are the same element. $\square$

Next we prove a theorem which immediately implies Theorem 1.3.

Theorem 1.5 (Erdös-Ginzburg-Ziv) *Let G be an abelian group of order n and let $a_0, \ldots, a_{2n-2}$ be a sequence of elements from G. Then there exists a subsequence of length n whose terms sum to 0.*

Proof: We proceed by induction on n. As a base case, observe that the result holds trivially when $n = 1$. So, we may suppose that $n = hp$ where p is prime and choose a subgroup $H < G$ with index p (and size h). Now consider the following sequence of elements from the quotient group G/H

$$a_0 + H, a_1 + H, \ldots, a_{2hp-2} + H.$$

By applying the previous lemma to this sequence, we may assume that the first p terms sum to H (the identity of G/H). If $h = 1$ then we are finished. Otherwise, after this first application, the remainder of our sequence $a_p + H, a_{p+1} + H, \ldots, a_{2hp-2} + H$ still has length $\geq 2p - 1$, so we may again apply the previous lemma to arrange $(a_p + H) + \ldots + (a_{2p-1} + H) = H$. By repeating this argument, we may assume that $(a_{jp} + H) + \ldots + (a_{(j+1)p-1} + H) = H$ for

$0 \leq j \leq 2h-2$. Now for every $0 \leq j \leq 2h-2$ define $b_j = a_{jp} + \ldots + a_{(j+1)p-1}$ and note that $b_j \in H$. By applying our theorem inductively to the sequence $b_0, \ldots, b_{2h-2}$ we may choose a subsequence of length h which sums to 0. Replacing each b_i term in this sequence with the corresponding block of terms from the a_i sequence yields the desired result. $\square$

Very Small Sumsets in $\mathbb{Z}_p$

We now turn our attention to the question of determining which pairs $A, B \subseteq \mathbb{Z}_p$ meet the lower bound from the Cauchy-Davenport Theorem with equality. More generally, let us call the pair (A, B) *deficient* if $|A+B| < |A| + |B|$. We will be interested in understanding the structure of deficient pairs. Let us begin by observing a rather trivial type of deficient 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 deficient, 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 deficient, 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 deficient. Similarly (C, A) is deficient. Thus, every deficient pair is actually part of a triple of subsets, any two of which form a deficient 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 *deficient* if $|A| + |B| + |C| > |G|$. It follows from the previous discussion that every nontrivial deficient pair extends to a nontrivial deficient trio, and every pair of sets from a nontrivial deficient trio forms a nontrivial deficient pair. With these definitions in place, we now have the following corollary of Theorem 1.2.

Corollary 1.6

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

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

Proof: Part 1 follows immediately from Theorem 1.2. For part 2, observe that since $-C$ is disjoint from $A+B$ we have $|C| \leq p-|A+B| = p-|A|-|B|+1$. Thus $|A|+|B|+|C| \leq p+1$ and to be deficient, we must have equality. $\square$

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

Lemma 1.7 *If (A, B, C) is a nontrivial deficient 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. $\square$

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 1.8 *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 $a_1, a_2, \ldots, a_k \in A$, and for every $1 \leq i \leq k$ 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

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

This completes the proof. $\square$

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

Theorem 1.9 (Vosper) *If p is prime and (A, B, C) is a nontrivial deficient 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 1.7. By our inductive hypothesis (and the fact that $\mathbb{Z}_p$ 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 1.8 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 deficient 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 1.7 that A, B, C are arithmetic progressions with a common difference, as required. $\square$