Sum-Free Sets of Finite Abelian Groups

Notes for a Presentation in Math 159

Robin Truax

March 2021

Contents

1 Introduction

Here, we present a modified (and corrected) proof originally due to Alon and Kleitman in their 1990 paper "Sum-free Subsets" of the following result:

Theorem 1. *Given any set B of nonzero elements from an finite abelian group A, there exists $C \subseteq B$ such that C is sum-free (that is, there do not exist $x, y, z \in C$ such that $x + y = z$) and $|C| > \frac{2}{7}|B|$.*

Since it applies to such a wide class of algebraic structures, this result has interesting consequences for everything from modular arithmetic to chip-firing and sandpiles. To prove this result, we will first use randomness (in the form of the probabilistic method) to show how one can use sum-free subsets of A to generate a sum-free subset of B, and then provide the necessary sum-free subsets of A. We choose this order because the latter process is ultimately less interesting.

2 Even-Spacedness

First, let's discuss why we might expect randomness to be an effective strategy for discovering results about subsets of finite abelian groups. Our hint is an idea I call "even-spacedness", which is a geometric visualization of the algebraic fact that the cosets of a subgroup $A' \leq A$ have identical cardinality and form a partition of A. This fact, often used in the first few days of a group theory class to prove Lagrange's Theorem, has important consequences for randomness in finite abelian groups.

Explicitly, it leads to the following theorem:

Theorem 2. *Suppose $\phi : A \to A'$ is a homomorphism between finite abelian groups, and $B \subseteq A$. Then, uniformly choosing $x \in A$ and considering $\phi(x)$ is the same as uniformly choosing $x' \in \operatorname{im} \phi$.*

Proof. Suppose that $x' \in \operatorname{im}\phi$. Then it is well-known that $\phi^{-1}(x')$ is a coset of $\ker\phi$. For example, the set of elements of A which ϕ maps to $\phi(x)$ is the coset $x + \ker\phi$. In particular, this result is used in the proof of the First Isomorphism Theorem. Now, since all cosets have the same size, a uniform choice of $a \in A$ is implicitly a uniform choice of a coset of $\ker\phi$, which after applying ϕ (via the correspondence between cosets of $\ker\phi$ and elements of $\operatorname{im}\phi$ described above), is a uniform choice of an element in $\operatorname{im}\phi$. □

3 A Simplifying Perspective

Our first step will be to simplify the form of our abelian group A. For this, it is useful to recall the structure theorem for finitely generated abelian groups, which states that

$$A \simeq \mathbb{Z}/n_1\mathbb{Z} \oplus \mathbb{Z}/n_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_k\mathbb{Z}$$

for some integers $n_1, \ldots, n_k$ such that $n_i \mid n_{i+1}$. We call these integers the *invariant factors of* A.

Now, abbreviate n_k by n. Notice that since $n_i \mid n$ for each i, there is a natural injection

$$\phi_i : \mathbb{Z}/n_i\mathbb{Z} \hookrightarrow \mathbb{Z}/n\mathbb{Z}$$

given by $x \bmod n_i \mapsto \frac{xn}{n_i} \bmod n$. Taking the product of these injections gives us a natural injection

$$A \simeq \mathbb{Z}/n_1\mathbb{Z} \oplus \mathbb{Z}/n_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_k\mathbb{Z} \hookrightarrow (\mathbb{Z}/n\mathbb{Z})^k.$$

Note that any element $x \in (\mathbb{Z}/n\mathbb{Z})^k$ can be represented as $(x_1, \ldots, x_k)$ for $x_1, \ldots, x_k \in \mathbb{Z}/n\mathbb{Z}$. This embedding is helpful because any subset $B \subseteq A$ can be now be considered as a subset of $(\mathbb{Z}/n\mathbb{Z})^k$, and finding a large sum-free subset of $B \subseteq (\mathbb{Z}/n\mathbb{Z})^k$ is the same as finding a large sum-free subset of $B \subseteq A$.

4 Defining a Random Variable

Note: For convenience, let $B = \{b_1, \ldots, b_m\}$. Since each element of B is now an element of $(\mathbb{Z}/n\mathbb{Z})^k$, $b_i = (b_{i1}, \ldots, b_{ik})$ for some $b_{ij} \in \mathbb{Z}/n\mathbb{Z}$ for each i.

Now, we want to define a random variable to help us isolate a sum-free subset of B. For this, define

$$\phi_i : (\mathbb{Z}/n\mathbb{Z})^k \to \mathbb{Z}/n\mathbb{Z} \text{ to be given by } (x_1, \ldots, x_k) \mapsto \sum_{j=1}^{k} x_j b_{ij}.$$

Now, suppose that we have a sum-free subset $F \subseteq \mathbb{Z}/n\mathbb{Z}$. Then choose $x \in (\mathbb{Z}/n\mathbb{Z})^k$ uniformly at random, and let X_F be the random variable denoting the number of $b_i \in B$ such that $\phi_i(x) \in F$. In other words, if $C_F \subseteq B$ is the set $\{b_i \in B \mid \phi_i(x) \in F\}$, then $X_F = |C_F|$.

Why is this a useful choice? The answer is that C_F is a sum-free subset. To see why, suppose

$$b_i + b_k = b_l.$$

with $b_i, b_k, b_l \in C_F$. Then, in particular, $b_{ij} + b_{kj} = b_{lj}$ for each j; that is, because the sum is coordinate-wise, the coordinates sum to each other. But then, notice that

$$\phi_i(x) + \phi_k(x) = \sum_{j=1}^{k} x_j b_{ij} + \sum_{j=1}^{k} x_j b_{kj} = \sum_{j=1}^{k} x_j (b_{ij} + b_{kj}) = \sum_{j=1}^{k} x_j (b_{lj}) = \phi_l(x)$$

but this is impossible, since $b_i, b_k, b_l \in C_F$ implies that $\phi_i(x), \phi_k(x), \phi_l(x) \in F$, and F is sum-free. Therefore, by contradiction, C_F must be sum-free.

Note: By our above work, if s is the size of the largest sum-free subset of B, then $s \geq X_F(x)$ for any x.

5 The Expectation of X_F

Notice that $X_F = 0$ with positive probability; for example, given the choice $x = (0, \dots, 0)$, $\phi_i(x) = 0 \notin F$ for every i. Therefore, there must be some outcome (that is, some choice of x) such that $X_F(x) > \mathbb{E}[X_F]$, so in particular if we can show that the expectation of $\mathbb{E}[X_F] \geq \frac{2}{7}m$ then we are immediately done. Therefore, we will compute the expectation of X_F.

To compute $\mathbb{E}[X_F]$, notice that if $d_i = \gcd(b_{i1}, \dots, b_{ik}, n)$, then $\text{im}(\phi_i) = d_i(\mathbb{Z}/n\mathbb{Z})$. Therefore, by even-spacedness, $\phi_i(x)$ is uniformly distributed over $\mathbb{Z}/n\mathbb{Z}$ as x is uniformly distributed over $(\mathbb{Z}/n\mathbb{Z})^k$. Hence,

$$\mathbb{P}(b_i \in C_F) = \mathbb{P}(\phi_i(x) \in F) = \frac{|d_i(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d_i(\mathbb{Z}/n\mathbb{Z})|}$$

Therefore, by the linearity of expectation,

$$\mathbb{E}[X_F] = \mathbb{E}[|C_F|] = \sum_{i=1}^{m} \frac{|d_i(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d_i(\mathbb{Z}/n\mathbb{Z})|}$$

In other words, if we can find a sum-free subset F satsfying

$$\frac{|d(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d(\mathbb{Z}/n\mathbb{Z})|} \geq \frac{2}{7} \tag{1}$$

for each divisor d of n, then we would know that

$$\mathbb{E}[X_F] = \sum_{i=1}^{m} \frac{|d_i(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d_i(\mathbb{Z}/n\mathbb{Z})|} \geq \sum_{i=1}^{m} \frac{2}{7} = \frac{2}{7}m.$$

We call a sum-free subset of $\mathbb{Z}/n\mathbb{Z}$ satisfying (1) a "sufficiently large sum-free subset". By our above work, all we need to do is find a sum-free subset F satisfying (1) for each $\mathbb{Z}/n\mathbb{Z}$.

6 Conclusion

Unfortunately, this is impossible: for various $\mathbb{Z}/n\mathbb{Z}$, no sufficiently large sum-free subsets exist. I verified via computer search that no sufficiently large sum-free subset exists for $\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/12\mathbb{Z}$, and $\mathbb{Z}/16\mathbb{Z}$ (in fact, it is possible that no sufficiently large sum-free subset exists for $\mathbb{Z}/n\mathbb{Z}$ for any $n \equiv 0 \bmod 4$). You can access the (very unoptimized) code I wrote for this search here (clickable link).

Therefore, we'll need to use a different trick. We'll use two different sum-free subsets F and F' of $\mathbb{Z}/n\mathbb{Z}$, which happen to satisfy the following relation

$$\frac{4}{7}\frac{|d(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d(\mathbb{Z}/n\mathbb{Z})|} + \frac{3}{7}\frac{|d(\mathbb{Z}/n\mathbb{Z}) \cap F'|}{|d(\mathbb{Z}/n\mathbb{Z})|} \geq \frac{2}{7} \tag{2}$$

These sum-free subsets are

$$F = \left\{ x \in \mathbb{Z}/n\mathbb{Z} \;\middle|\; \frac{1}{3}n < x \leq \frac{2}{3}n \right\} \qquad F' = \left\{ x \in \mathbb{Z}/n\mathbb{Z} \;\middle|\; \frac{1}{6}n < x \leq \frac{1}{3}n \text{ or } \frac{2}{3}n < x \leq \frac{5}{6}n \right\}.$$

A simple size argument shows that these sets are sum-free, and the proof they satisfy (2) is in the appendix below. However, assuming they satisfy (2), we can complete the result using a small trick. Recall that s is the size of the largest sum-free subset of B, and that $\mathbb{E}[X_F] > s$ for any sum-free subset F of $\mathbb{Z}/n\mathbb{Z}$.

$$s = \frac{4}{7}s + \frac{3}{7}s > \frac{4}{7}\mathbb{E}[X_F] + \frac{3}{7}\mathbb{E}[X_{F'}] = \frac{4}{7}\sum_{i=1}^{m} \frac{|d_i(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d_i(\mathbb{Z}/n\mathbb{Z})|} + \frac{3}{7}\sum_{i=1}^{m} \frac{|d_i(\mathbb{Z}/n\mathbb{Z}) \cap F'|}{|d_i(\mathbb{Z}/n\mathbb{Z})|} =$$
$$= \sum_{i=1}^{m} \left(\frac{4}{7}\frac{|d_i(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d_i(\mathbb{Z}/n\mathbb{Z})|} + \frac{3}{7}\frac{|d_i(\mathbb{Z}/n\mathbb{Z}) \cap F'|}{|d_i(\mathbb{Z}/n\mathbb{Z})|} \right) \geq \sum_{i=1}^{m} \frac{2}{7} = \frac{2}{7}m$$

Hence the result follows assuming F and F' satisfy (2), as desired.

7 Appendix: The Auxiliary Lemma

Lemma 3. *Define*

$$F = \left\{ x \in \mathbb{Z}/n\mathbb{Z} \;\middle|\; \frac{1}{3}n < x \leq \frac{2}{3}n \right\} \qquad F' = \left\{ x \in \mathbb{Z}/n\mathbb{Z} \;\middle|\; \frac{1}{6}n < x \leq \frac{1}{3}n \text{ or } \frac{2}{3}n < x \leq \frac{5}{6}n \right\}.$$

Then, for every $d \mid n$,

$$\frac{4}{7}\frac{|d(\mathbb{Z}/n\mathbb{Z}) \cap F|}{|d(\mathbb{Z}/n\mathbb{Z})|} + \frac{3}{7}\frac{|d(\mathbb{Z}/n\mathbb{Z}) \cap F'|}{|d(\mathbb{Z}/n\mathbb{Z})|} \geq \frac{2}{7}. \tag{2}$$

Proof. First, notice that $|d(\mathbb{Z}/n\mathbb{Z})| = (n/d)$. Then, by inspection, we complete the following table:

$\frac{n}{d}$	$\frac{\lvert d(\mathbb{Z}/n\mathbb{Z}) \cap F\rvert}{\lvert d(\mathbb{Z}/n\mathbb{Z})\rvert}$	$\frac{\lvert d(\mathbb{Z}/n\mathbb{Z}) \cap F'\rvert}{\lvert d(\mathbb{Z}/n\mathbb{Z})\rvert}$
$6k$	$\frac{2k}{6k} = \frac{1}{3}$	$\frac{2k}{6k} = \frac{1}{3}$
$6k+1$	$\frac{2k}{6k+1} \geq \frac{2}{7}$	$\frac{2k}{6k+1} \geq \frac{2}{7}$
$6k+2$	$\frac{2k+1}{6k+2}$	$\frac{2k}{6k+2}$
$6k+3$	$\frac{2k+1}{6k+3} = \frac{1}{3}$	$\frac{2k+1}{6k+3} = \frac{1}{3}$
$6k+4$	$\frac{2k+1}{6k+4}$	$\frac{2k+2}{6k+4}$
$6k+5$	$\frac{2k+2}{6k+5} > \frac{1}{3}$	$\frac{2k+2}{6k+5} > \frac{1}{3}$

In particular, for every case but $\frac{n}{d} = 6k+2$ and $\frac{n}{d} = 6k+4$, it is clear that (2) holds. However, these two cases require slightly more work. The first case, $\frac{n}{d} = 6k+2$, holds because

$$\frac{4}{7} \cdot \frac{2k+1}{6k+2} + \frac{3}{7} \cdot \frac{2k}{6k+2} = \frac{1}{7} \cdot \frac{14k+4}{6k+2} = \frac{1}{7} \cdot \frac{7k+2}{3k+1} \geq \frac{2}{7}.$$

Similarly, the second case, $\frac{n}{d} = 6k+4$, holds because

$$\frac{4}{7} \cdot \frac{2k+1}{6k+4} + \frac{3}{7} \cdot \frac{2k+2}{6k+4} \geq \frac{4}{7} \cdot \frac{1}{4} + \frac{3}{7} \cdot \frac{1}{3} = \frac{1}{7} + \frac{1}{7} = \frac{2}{7}.$$

Hence in these cases (2) also holds, so we are done. □