7

Structure of Set Addition

Chapter Highlights

- Freiman's theorem: structure of sets with small doubling
- Inequalities between sizes of sumsets: Ruzsa triangle inequality and Plünnecke's inequality
- Ruzsa covering lemma
- Freiman homomorphisms: preserving partial additive structure
- Ruzsa modeling lemma
- Structure in iterated sumsets: Bogolyubov's lemma
- Geometry of numbers: Minkowski's second theorem
- Polynomial Freiman–Ruzsa conjecture
- Additive energy and the Balog–Szemerédi–Gowers theorem

Let A and B be finite subsets of some ambient abelian group. We define their ***sumset*** to be

$$\boldsymbol{A + B} \coloneqq \{a + b : a \in A, b \in B\}.$$

Note that we view $A + B$ as a set, and do not keep track of the number of ways that each element can be written as $a + b$.

The main goal of this chapter is to understand the following question.

Question 7.0.1 (Sets with small doubling)
What can we say about A if $A + A$ is small?

We will prove Freiman's theorem, which is a deep and foundational result in additive combinatorics. Freiman's theorem tells us that, if $A + A$ is at most a constant factor larger than A, then A must be a large fraction of some generalized arithmetic progression.

Most of this chapter will be devoted toward proving Freiman's theorem. We will see ideas and tools from Fourier analysis, geometry of numbers, and additive combinatorics.

In Section 7.13, we will introduce the **additive energy** of a set, which is another way to measure the additive structure of a set. We will see the Balog–Szemerédi–Gowers theorem, which relates additive energy and doubling. This section can be read independently from the earlier parts of the chapter.

These results on the structure of set addition are not only interesting on their own, but also play a key role in Gowers' proof (2001) of Szemerédi's theorem (although we do not cover it in this book; see Further Reading at the end of the chapter). Gowers' deep and foundational work shows how these topics in additive combinatorics are all highly connected.

Definition 7.0.2 (Sumset notation)
Given a positive integer k, we define the iterated sumset

$$\boldsymbol{kA} := A + \cdots + A \quad (k \text{ times}).$$

This is different from dilating a set, which is denoted by

$$\boldsymbol{\lambda \cdot A} := \{\lambda a : a \in A\}.$$

We also consider the difference set

$$\boldsymbol{A - B} = \{a - b : a \in A, b \in B\}.$$

7.1 Sets of Small Doubling: Freiman's Theorem

How small or large can $A + A$ be, given $|A|$? This is an easy question to answer.

Proposition 7.1.1 (Easy bounds on sumset size)
Let $A \subseteq \mathbb{Z}$ be a finite set. Then

$$2|A| - 1 \le |A + A| \le \binom{|A|+1}{2}.$$

Furthermore, both bounds are best possible as functions of $|A|$.

Proof. Let $n = |A|$. For the lower bound $|A + A| \ge 2n - 1$, note that if the elements of A are $a_1 < a_2 < \cdots < a_n$, then

$$a_1 + a_1 < a_1 + a_2 < \cdots < a_1 + a_n < a_2 + a_n < \cdots < a_n + a_n$$

are $2n - 1$ distinct elements of $A + A$. So $|A + A| \ge 2n - 1$. Equality is attained when A is an arithmetic progression.

The upper bound $|A + A| \le \binom{n+1}{2}$ follows from that there are $\binom{n+1}{2}$ unordered pairs of elements of A. We have equality when there are no nontrivial solutions to $a + b = c + d$ in A, such as when A consists of powers of twos. □

Exercise 7.1.2 (Sumsets in abelian groups). Show that if A is a finite subset of an abelian group, then $|A + A| \ge |A|$, with equality if and only if A is the coset of some subgroup.

What can we say about A if $A + A$ is not too much larger than A?

Definition 7.1.3 (Doubling constant)
The ***doubling constant*** of a finite subset A in an abelian group is the ratio $|A + A|/|A|$.

One of the main results of this chapter, Freiman's theorem, addresses the following question.

Question 7.1.4 (Sets of small doubling)
What is the structure of a set with bounded doubling constant (e.g. $|A + A| \le 100|A|$)?

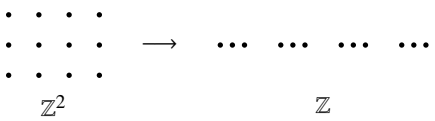
We've already seen an example of such a set in $\mathbb{Z}$, namely arithmetic progressions.

***Example* 7.1.5.** If $A \subseteq \mathbb{Z}$ is a finite arithmetic progression, $|A+A| = 2|A| - 1 \le 2|A|$, so it has doubling constant at most 2.

Moreover if we delete some elements of an arithmetic progression, it should still have small doubling. In fact, if we delete even most of the elements of an arithmetic progression but leave a constant fraction of the progression remaining, we will have small doubling.

***Example* 7.1.6.** If B is a finite arithmetic progression and $A \subseteq B$ has $|A| \ge |B|/K$, then $|A+A| \le |B+B| \le 2|B| \le 2K|A|$, so A has doubling constant at most $2K$.

Now we generalize arithmetic progressions to allow multiple dimensions. Informally, we consider affine images of d-dimensional "grids," as illustrated below.



Definition 7.1.7 (GAP – generalized arithmetic progression)
A ***generalized arithmetic progression (GAP)*** in an abelian group Γ is defined to be an affine map

$$\phi\colon [L_1] \times \cdots \times [L_d] \to \Gamma.$$

That is, for some $a_0, \ldots, a_d \in \Gamma$,

$$\phi(x_1, \ldots, x_d) = a_0 + a_1 x_1 + \cdots + a_d x_d.$$

This GAP has ***dimension*** d and ***volume*** $L_1 \cdots L_d$. We say that this GAP is ***proper*** if ϕ is injective.

We often abuse notation and use the term GAP to refer to the image of ϕ, viewed as a set:

$$a_0 + a_1 \cdot [L_1] + \cdots + a_d \cdot [L_d] = \{a_0 + a_1 x_1 + \cdots + a_d x_d : x_1 \in [L_1], \ldots, x_d \in [L_d]\}.$$

***Example* 7.1.8.** A proper GAP of dimension d has doubling constant $\le 2^d$.

***Example* 7.1.9.** Let P be a proper GAP of dimension d. Let $A \subseteq P$ with $|A| \ge |P|/K$. Then A has doubling constant $\le K2^d$.

While it is often easy to check that certain sets have small doubling, the **inverse problem** is much more difficult. We would like to characterize all sets with small doubling. The following foundational result by Freiman (1973) shows that all sets with bounded doubling must look like Example 7.1.9.

Theorem 7.1.10 (Freiman's theorem)
Let $A \subseteq \mathbb{Z}$ be a finite set satisfying $|A+A| \le K|A|$. Then A is contained in a GAP of dimension at most $d(K)$ and volume at most $f(K)|A|$, where $d(K)$ and $f(K)$ are constants depending only on K.

Freiman's theorem is a deep result. We will spend most the chapter proving it.

***Remark* 7.1.11** (Quantitative bounds)**.** We will present a proof giving $d(K) = \exp(K^{O(1)})$ and $f(K) = \exp(d(K))$, due to Ruzsa (1994). Chang (2002) showed that Freiman's theorem holds with $d(K) = K^{O(1)}$ and $f(K) = \exp(d(K))$ (see Exercise 7.11.2). Schoen (2011) further improved the bounds to $d(K) = K^{1+o(1)}$ and $f(K) = \exp(K^{1+o(1)})$. Sanders (2012, 2013) showed that if we change GAPs to "convex progressions" (see Section 7.12), then an analogous theorem holds with $d(K) = K(\log(2K))^{O(1)}$ and $f(K) = \exp(d(K))$.

It is easy to see that one cannot do better than $d(K) \le K - 1$ and $f(K) = e^{O(K)}$, by considering a set without additive structure.

Also see Section 7.12 on the polynomial Freiman–Ruzsa conjecture for a variant of Freiman's theorem with much better quantitative dependencies.

***Remark* 7.1.12** (Making the GAP proper)**.** The conclusion of Freiman's theorem can be strengthened to force the GAP to be proper, at the cost of potentially increasing $d(K)$ and $f(K)$. For example, it is known that every GAP of dimension d is contained in some proper GAP of dimension $\le d$ with at most $d^{O(d^3)}$ factor increase in the volume; see Tao and Vu (2006, Theorem 3.40).

Remark* 7.1.13** (History)**.** Freiman's original proof (1973) was quite complicated. Ruzsa (1994) later found a simpler proof, which guided much of the subsequent work. We follow Ruzsa's presentation here. Theorem 7.1.10 is sometimes called the ***Freiman–Ruzsa theorem. Freiman's theorem was brought into further prominence due to the role it played in the new proof of Szemerédi's theorem by Gowers (2001).

Remark* 7.1.14** (Freiman's theorem in abelian groups)**.** Green and Ruzsa (2007) proved a generalization of Freiman's theorem in an arbitrary abelian group. A ***coset progression is a set of the form $P + H$ where P is a GAP and H is a subgroup of the ambient abelian group. Define the ***dimension*** of this coset progression to be the dimension of P, and its ***volume*** to be $|H| \operatorname{vol} P$. Green and Ruzsa (2007) proved the following theorem.

Theorem 7.1.15 (Freiman's theorem for general abelian groups)
Let A be a subset of an abelian group satisfying $|A + A| \le K|A|$. Then A is contained in a coset progression of dimension at most $d(K)$ and volume at most $f(k)|A|$, where $d(K)$ and $f(K)$ are constants depending only on K.

7.2 Sumset Calculus I: Ruzsa Triangle Inequality

Here are some basic and useful inequalities relating the sizes of sumsets.

Theorem 7.2.1 (Ruzsa triangle inequality)
If A, B, C are finite subsets of an abelian group, then

$$|A|\,|B - C| \le |A - B|\,|A - C|.$$

Proof. For each $d \in B - C$, define $\mathsf{b}(d) \in B$ and $\mathsf{c}(d) \in C$ such that $d = \mathsf{b}(d) - \mathsf{c}(d)$. In other words, we fix a specific choice of b and c for each element in $B - C$. Define

$$\begin{aligned} \phi : A \times (B - C) &\longrightarrow (A - B) \times (A - C) \\ (a, d) &\longmapsto (a - \mathsf{b}(d), a - \mathsf{c}(d)). \end{aligned}$$

Then ϕ is injective since we can recover (a, d) from $\phi(a, d) = (x, y)$ via $d = y - x$ and then $a = x + \mathsf{b}(d)$. □

***Remark* 7.2.2.** By replacing B with $-B$ and/or C with $-C$, Theorem 7.2.1 implies some additional sumset inequalities:

$$|A|\,|B + C| \le |A + B|\,|A - C|\,;$$
$$|A|\,|B + C| \le |A - B|\,|A + C|\,;$$
$$|A|\,|B - C| \le |A + B|\,|A + C|\,.$$

However, this trick cannot be used to prove the similarly looking inequality

$$|A|\,|B + C| \le |A + B|\,|A + C|\,.$$

This inequality is also true, and we will prove it in the following section.

***Remark* 7.2.3** (Why is it called a triangle inequality?)**.** If we define

$$\rho(A, B) := \log \frac{|A - B|}{\sqrt{|A|\,|B|}}$$

(called a ***Ruzsa distance***), then Theorem 7.2.1 can be rewritten as

$$\rho(B, C) \le \rho(A, B) + \rho(A, C).$$

This is why Theorem 7.2.1 is called a "triangle inequality." However, one should not take the name too seriously. The function ρ is not a metric because $\rho(A, A) \ne 0$ in general.

Exercise 7.2.4 (Iterated sumsets)**.** Let A be a finite subset of an abelian group satisfying

$$|2A - 2A| \le K\,|A|\,.$$

Prove that

$$|mA - mA| \le K^{m-1}\,|A| \quad \text{for every integer } m \ge 2.$$

In the above exercise, we had to start with the assumption that $|2A - 2A| \le K\,|A|$. In the next section, we bound the sizes of iterated sumsets starting with the weaker hypothesis $|A + A| \le K\,|A|$.

7.3 Sumset Calculus II: Plünnecke's Inequality

We prove the following result, which says that having small doubling implies small iterated sumsets, with only a polynomial factor change in the expansion ratios.

Theorem 7.3.1 (Plünnecke's inequality)
Let A be a finite subset of an abelian group satisfying

$$|A + A| \le K\,|A|\,.$$

Then for all integers $m, n \ge 0$,

$$|mA - nA| \le K^{m+n}\,|A|\,.$$

Remark* 7.3.2** (History)**.** Plünnecke (1970) proved a version of the theorem originally using graph theoretic methods. Ruzsa (1989) gave a simpler version of Plünnecke's proof and also extended it from sums to differences. Nevertheless, Ruzsa's proof was still quite long and complex. It sets up a "commutative layered graph," and uses tools from graph theory including Menger's theorem. Theorem 7.3.1 is sometimes called the ***Plünnecke–Ruzsa inequality. See Ruzsa (2009, Chapter 1) or Tao and Vu (2006, Chapter 6) for an account of this proof.

In a surprising breakthrough, Petridis (2012) found a very short proof of the result, which we present here.

We will prove the following more general statement. Theorem 7.3.1 is the special case $A = B$.

Theorem 7.3.3 (Plünnecke's inequality)

Let A and B be finite subsets of an abelian group satisfying

$$|A + B| \le K |A| .$$

Then for all integers $m, n \ge 0$,

$$|mB - nB| \le K^{m+n} |A| .$$

The following lemma plays a key role in the proof.

Lemma 7.3.4 (Expansion ratio bounds)

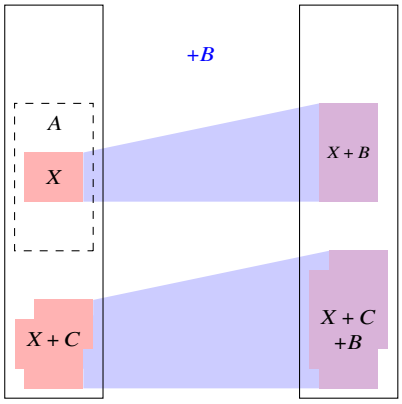
Let X and B be finite subsets of an abelian group, with $|X| > 0$. Suppose

$$\frac{|Y + B|}{|Y|} \ge \frac{|X + B|}{|X|} \quad \text{for all nonempty } Y \subseteq X.$$

Then for any nonempty finite subsets C of the abelian group,

$$\frac{|X + C + B|}{|X + C|} \le \frac{|X + B|}{|X|}.$$

***Remark* 7.3.5** (Interpretation as expansion ratios)**.** We can interpret Lemma 7.3.4 in terms of vertex expansion ratios inside the bipartite graph between two copies of the ambient abelian group, with edges $(x, x + b)$ ranging over all $x \in \Gamma$ and $b \in B$. Every vertex subset X on the left has neighbors $X + B$ on the right and thus has *vertex expansion ratio* $|X + B| / |X|$.



MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

We will apply Lemma 7.3.4 by choosing X among all nonempty subsets of A with the minimum expansion ratio, so that the hypothesis of Lemma 7.3.4 is automatically satisfied. The conclusion of Lemma 7.3.4 then says that a union of translates of X has expansion ratio at most that of X.

Proof of Theorem 7.3.3 given Lemma 7.3.4. Choose X among all nonempty subsets of A with the minimum $|X+B|/|X|$ so that the hypothesis of Lemma 7.3.4 is satisfied. Also we have

$$\frac{|X+B|}{|X|} \le \frac{|A+B|}{|A|} \le K.$$

For every integer $n \ge 0$, applying Lemma 7.3.4 with $C = nB$, we have

$$\frac{|X+(n+1)B|}{|X+nB|} \le \frac{|X+B|}{|X|} \le K.$$

So induction on n yields, for all $n \ge 0$,

$$|X+nB| \le K^n |X|.$$

Finally, applying the Ruzsa triangle inequality (Theorem 7.2.1), for all $m, n \ge 0$.

$$|mB-nB| \le \frac{|X+mB|\,|X+nB|}{|X|} \le K^{m+n}|X| \le K^{m+n}|A|. \qquad \square$$

Proof of Lemma 7.3.4. We will proceed by induction on $|C|$. For the base case $|C| = 1$, note that $X + C$ is a translate of X, so $|X+C+B| = |X+B|$ and $|X+C| = |X|$.

Now for the induction step, assume that for some C,

$$\frac{|X+C+B|}{|X+C|} \le \frac{|X+B|}{|X|}.$$

Now consider $C \cup \{c\}$ for some $c \notin C$. We wish to show that

$$\frac{|X+(C\cup\{c\})+B|}{|X+(C\cup\{c\})|} \le \frac{|X+B|}{|X|}.$$

By comparing the change in the left-hand side fraction, it suffices to show that

$$|(X+c+B)\setminus(X+C+B)| \le \frac{|X+B|}{|X|}\,|(X+c)\setminus(X+C)|. \tag{7.1}$$

Let

$$Y = \{x \in X : x+c+B \subseteq X+C+B\} \subseteq X.$$

Then

$$|(X+c+B)\setminus(X+C+B)| \le |X+B| - |Y+B|.$$

Furthermore, if $x \in X$ satisfies $x+c \in X+C$, then $x+c+B \subseteq X+C+B$ and hence $x \in Y$. So

$$|(X+c)\setminus(X+C)| \ge |X| - |Y|.$$

Thus, to prove (7.1), it suffices to show

$$|X+B| - |Y+B| \le \frac{|X+B|}{|X|}(|X|-|Y|),$$

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

which can be rewritten as

$$|Y+B| \geq \frac{|X+B|}{|X|}\,|Y|\,,$$

which is true due to the hypothesis on X. □

Let us give a quick proof of a variant of the Ruzsa triangle inequality, mentioned in Remark 7.2.2.

Corollary 7.3.6 (Another triangle inequality)
Let A, B, C be finite subsets of an abelian group. Then

$$|A|\,|B+C| \leq |A+B|\,|A+C|\,.$$

Proof. Choose $X \subseteq A$ to minimize $|X+B|/|X|$. Then

$$|B+C| \leq |X+B+C| \overset{\text{Lem. 7.3.4}}{\leq} |X+C|\,\frac{|X+B|}{|X|} \leq |A+C|\,\frac{|A+B|}{|A|}.$$ □

Exercise 7.3.7*. Show that for every sufficiently large K there is some finite set $A \subseteq \mathbb{Z}$ such that

$$|A+A| \leq K\,|A| \qquad \text{and} \qquad |A-A| \geq K^{1.99}\,|A|\,.$$

Exercise 7.3.8* (Loomis–Whitney for sumsets)**.** Show that for every finite subsets A, B, C in an abelian group, one has

$$|A+B+C|^2 \leq |A+B|\,|A+C|\,|B+C|\,.$$

Exercise 7.3.9* (Sumset vs. difference set)**.** Let $A \subseteq \mathbb{Z}$. Prove that

$$|A-A|^{2/3} \leq |A+A| \leq |A-A|^{3/2}\,.$$

7.4 Covering Lemma

Here is a simple and powerful tool in the study of sumsets (Ruzsa 1999).

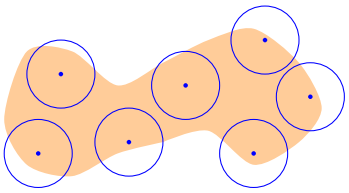
Theorem 7.4.1 (Ruzsa covering lemma)
Let X and B be finite sets in some abelian group. If

$$|X+B| \leq K\,|B|\,,$$

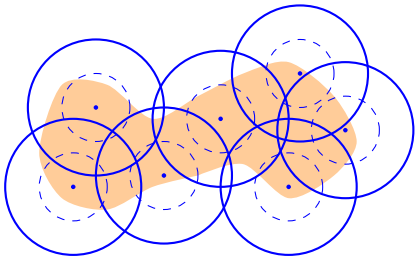
then there exists a subset $T \subseteq X$ with $|T| \leq K$ such that

$$X \subseteq T+B-B.$$

***Remark* 7.4.2** (Geometric intuition)**.** Imagine that B is a unit ball in $\mathbb{R}^n$, and cardinality is replaced by volume. Given some region X (the shaded region below), consider a maximal set $\mathcal{T}$ of disjoint union balls with centers in X (maximal in the sense that one cannot add an additional ball without intersecting some other ball).



Then replacing each ball in $\mathcal{T}$ by a ball of radius 2 with the same center, (i.e., replacing B by $B - B$) the resulting balls must cover the region X (which amounts to the conclusion $X \subseteq T + B - B$), for otherwise at any uncovered point of X we could have added an additional nonoverlapping ball in the previous step.



Similar arguments are important in analysis (e.g., the Vitali covering lemma).

Proof. Let $T \subseteq X$ be a maximal subset such that $t + B$ as t ranges over T are disjoint. Then

$$|T|\,|B| = |T + B| \le |X + B| \le K\,|B|\,.$$

So $|T| \le K$.

By the maximality of T, for all $x \in X$ there exists some $t \in T$ such that $(t+B) \cap (x+B) \neq \emptyset$. In other words, there exist $t \in T$ and $b, b' \in B$ such that $t + b = x + b'$. Hence $x \in T + B - B$ for every $x \in X$. Thus $X \subseteq T + B - B$. □

The following "more efficient" covering lemma can be used to prove a better bound in Freiman's theorem.

Exercise 7.4.3* (Chang's covering lemma)**.** Let A and B be finite sets in an abelian group satisfying

$$|A + A| \le K\,|A| \qquad \text{and} \qquad |A + B| \le K'\,|B|\,.$$

Show that there exists some set X in the abelian group so that

$$A \subseteq \Sigma X + B - B \qquad \text{and} \qquad |X| = O(K \log(KK')),$$

where ΣX denotes the set of all elements that can be written as the sum of a subset of elements of X (including zero as the sum of the empty set).

Hint: Try first finding $2K$ disjoint translates $a + B$.

7.5 Freiman's Theorem in Groups with Bounded Exponent

Let us prove a finite field model analogue of Freiman's theorem. The proof only uses the tools introduced so far, and so it is easier than Freiman's theorem in the integers.

Theorem 7.5.1 (Freiman's theorem in $\mathbb{F}_2^n$)
If $A \subseteq \mathbb{F}_2^n$ has $|A+A| \le K|A|$, then A is contained in a subspace of cardinality at most $f(K)|A|$, where $f(K)$ is a constant depending only on K.

***Remark* 7.5.2** (Quantitative bounds)**.** We will prove Theorem 7.5.1 with $f(K) = 2^{K^4}K^2$. The exact optimal constant $f(K)$ is known for each K (Even-Zohar 2012). Asymptotically, it is $f(K) = \Theta(2^{2K}/K)$.

For a matching lower bound on $f(K)$, let $A = \{0, e_1, \dots, e_n\} \subseteq \mathbb{F}_2^n$, where e_i is the ith standard basis vector. Then $|A+A| \sim n^2/2$, and so $|A+A|/|A| \sim n/2$. However, A is not contained in a subspace of cardinality less than 2^n.

In fact, we prove a more general statement that works for any group with bounded exponent. This result and proof are due to Ruzsa (1999).

Definition 7.5.3 (Exponent of an abelian group)
The ***exponent*** of an abelian group (written additively) is the smallest positive integer r such that $rx = 0$ for all elements x of the group. If no finite r exists, we say that its exponent is infinite (some conventions say that the exponent is zero).

For example, $\mathbb{F}_2^n$ has exponent 2. The cyclic group $\mathbb{Z}/N\mathbb{Z}$ has exponent N. The integers $\mathbb{Z}$ has infinite exponent.

We use $\langle A \rangle$ to refer to the subgroup of a group G generated by some subset A of G. Then the exponent of a group G is $\sup_{x \in G} |\langle x \rangle|$. When the group is a vector space (e.g., $\mathbb{F}_2^n$), $\langle A \rangle$ is the smallest subspace containing A.

Theorem 7.5.4 (Freiman's theorem in groups with bounded exponent)
Let A be a finite set in an abelian group with exponent $r < \infty$. If $|A+A| \le K|A|$, then

$$|\langle A \rangle| \le K^2 r^{K^4} |A|.$$

***Remark* 7.5.5.** This theorem is a converse of the observation that if A is a large fraction of a subgroup, then A has small doubling.

Proof. By Plünnecke's inequality (Theorem 7.3.1), we have

$$|A + (2A - A)| = |3A - A| \le K^4 |A|.$$

By the Ruzsa covering lemma (Theorem 7.4.1 applied with $X = 2A - A$ and $B = A$), there exists some $T \subseteq 2A - A$ with $|T| \le |A + (2A - A)|/|A| \le K^4$ such that

$$2A - A \subseteq T + A - A.$$

Adding A to both sides, we have,

$$3A - A \subseteq T + 2A - A \subseteq 2T + A - A.$$

Iterating, for any positive integer n, we have

$$(n+1)A - A \subseteq nT + A - A \subseteq \langle T \rangle + A - A.$$

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

Since we are in an abelian group with bounded exponent, every element of $\langle A \rangle$ lies in nA for some n. Thus

$$\langle A \rangle \subseteq \bigcup_{n \geq 1} (nA + A - A) \subseteq \langle T \rangle + A - A.$$

Since the exponent of the group is at most $r < \infty$,

$$|\langle T \rangle| \leq r^{|T|} \leq r^{K^4}.$$

By Plünnecke's inequality (Theorem 7.3.1),

$$|A - A| \leq K^2 |A|.$$

Thus we have,

$$|\langle A \rangle| \leq r^{K^4} K^2 |A|. \qquad \square$$

***Remark* 7.5.6.** Note the crucial use of the Ruzsa covering lemma for controlling $nA - A$. Naively bounding nA using Plünnecke's inequality is insufficient.

The above proof for Freiman's theorem over abelian groups of finite exponent does not immediately generalize to the integers. Indeed, in $\mathbb{Z}$, $|\langle T \rangle| = \infty$. We overcome this issue by representing subsets of $\mathbb{Z}$ inside a finite group in a way that partially preserves additive structure.

Exercise 7.5.7. Show that for every real $K \geq 1$ there is some C_K such that for every finite set A of an abelian group with $|A + A| \leq K |A|$, one has $|nA| \leq n^{C_K} |A|$ for every positive integer n.

(If we let $f(n, K)$ denote the smallest real number so that $|A + A| \leq K |A|$ implies $|nK| \leq f(n, K) |A|$, then Plünnecke's inequality gives $f(n, K) \leq K^n$, at most a polynomial in K for a fixed n, whereas the above exercise gives $f(n, K) \leq n^{C_K}$, a polynomial in n for a fixed K. Does this mean that $f(n, K)$ is at most some polynomial in both n and K?)

Exercise 7.5.8* (Ball volume growth in an abelian Cayley graph)**.** Show that there is some absolute constant C so that if S is a finite subset of an abelian group, and k is a positive integer, then

$$|2kS| \leq C^{|S|} |kS|.$$

7.6 Freiman Homomorphisms

Consider two sets of integers, depicted pictorially below as elements on the number line:

$A =$ • • • • • • • • • • • • • • • • • • • • • •

$B =$ •••• ••••• •••••• •••••••

The two sets are very similar from the point of view of additive structure. For example, the obvious bijection between A and B has the nice property that any solution to the equation $w + x = y + z$ in one set is automatically a solution in the other. Sometimes, in additive combinatorics, it is a good idea to treat these two sets as isomorphic. Let us define this

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

notion formally and study what it means for a map between sets to partially preserve additive structure.

Definition 7.6.1 (Freiman homomorphism)
Let A and B be subsets in two possibly different abelian groups. Let $s \geq 2$ be a positive integer. We say that $\phi\colon A \to B$ is a ***Freiman s-homomorphism*** (or ***Freiman homomorphism of order s***), if

$$\phi(a_1) + \cdots + \phi(a_s) = \phi(a_1') + \cdots + \phi(a_s')$$

whenever $a_1, \ldots, a_s, a_1', \ldots, a_s' \in A$ satisfy

$$a_1 + \cdots + a_s = a_1' + \cdots + a_s'.$$

We say that ϕ is a ***Freiman s-isomorphism*** if ϕ is a bijection, and both ϕ and ϕ^{-1} are Freiman s-homomorphisms. We say that A and B are ***Freiman s-isomorphic*** if there exists a Freiman s-isomorphism between them.

***Remark* 7.6.2** (Interpretation)**.** Informally, a Freiman s-homomorphism respects s-fold sums relations. Two sets are Freiman s-isomorphic if there is a bijection between them that respects solutions to the equation $a_1 + \cdots + a_s = a_1' + \cdots + a_s'$.

***Remark* 7.6.3** (Composition)**.** If ϕ_1 and ϕ_2 are both Freiman s-homomorphisms, then their composition $\phi_1 \circ \phi_2$ is also a Freiman s-homomorphism. If ϕ_1 and ϕ_2 are both Freiman s-isomorphisms, then their composition $\phi_1 \circ \phi_2$ is a Freiman s-isomorphism.

***Remark* 7.6.4** (Descension)**.** Every Freiman $(s+1)$-homomorphism is automatically a Freiman s-homomorphism (by setting $a_{s+1} = a_{s+1}'$). Likewise, every Freiman $(s+1)$-isomorphism is automatically a Freiman s-isomorphism.

***Example* 7.6.5** (Freiman homomorphism)**.**

(a) Every abelian group homomorphism is a Freiman homomorphism of every order.
(b) Let S be a set with no nontrivial solutions to $a+b = c+d$ (such a set is called a ***Sidon set***). Then every map from S to an abelian group is a Freiman 2-homomorphism.
(c) The natural embedding $\phi\colon \{0,1\}^n \to (\mathbb{Z}/2\mathbb{Z})^n$ is the restriction of a group homomorphism from $\mathbb{Z}^n$, so it is a Freiman homomorphism of every order. This map ϕ is a bijection. However, the inverse of ϕ does not preserve some additive relations (e.g., $1+1 = 0+0 \pmod 2$). So ϕ is *not* a Freiman 2-isomorphism!
(d) Likewise, the natural embedding $\phi\colon [N] \to \mathbb{Z}/N\mathbb{Z}$ is a Freiman homomorphism of every order but not a Freiman 2-isomorphism. However, when the domain is restricted to all integers less than N/s, then ϕ becomes a Freiman s-isomorphism onto its image (why?).

The last example has the following easy generalization, which we will use later. The ***diameter*** of a set A is defined to be

$$\operatorname{diam} A := \sup_{a,b \in A} |a - b|.$$

Proposition 7.6.6 (Small diameter sets)
If $A \subseteq \mathbb{Z}$ has diameter $< N/s$, then A is Freiman s-isomorphic to its image mod N.

Intuitively, the idea is that there are no wraparound additive relations mod N if A has small diameter.

Proof. The mod N map $\mathbb{Z} \to \mathbb{Z}/N$ is a group homomorphism, and hence automatically a Freiman s-homomorphism. Now, if $a_1, \dots, a_s, a_1', \dots, a_s' \in A$ are such that

$$(a_1 + \cdots + a_s) - (a_1' + \cdots + a_s') \equiv 0 \pmod{N},$$

then the left-hand side, viewed as an integer, has absolute value less than N (since $|a_i - a_i'| < N/s$ for each i). Thus the left-hand side must be 0 in $\mathbb{Z}$. So the inverse of the mod N map is a Freiman s-homomorphism over A, and thus mod N is a Freiman s-isomorphism. □

7.7 Modeling Lemma

The goal of the Ruzsa modeling lemma is to represent a set with bounded doubling inside a small cyclic group in a way that that preserves relevant additive data. This is useful since initially A may contain integers of vastly different magnitudes. On the other hand, if A is a subset of $\mathbb{Z}/N\mathbb{Z}$ with N comparable to A, then we have additional tools such as Fourier analysis (to be discussed in the following section).

As warm-up, let us first prove an easier result in the finite field model.

Proposition 7.7.1 (Modeling lemma in finite field model)
Let $A \subseteq \mathbb{F}_2^n$. Suppose $|sA - sA| \le 2^m$ for some positive integer m. Then A is Freiman s-isomorphic to some subset of $\mathbb{F}_2^m$.

***Remark* 7.7.2.** If $|A + A| \le K|A|$, then Plünnecke's inequality (Theorem 7.3.1) implies $|sA - sA| \le K^{2s}|A|$. By taking m to be the smallest integer with $K^{2s}|A| \le 2^m$, we see that the cardinality of the final vector space $\mathbb{F}_2^m$ is within a constant factor $2K^{2s}$ of $|A|$. In contrast, A initially lived in a space $\mathbb{F}_2^n$ that could potentially be much larger.

Proof. It is easy to check that the following are equivalent for a linear map $\phi\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$:

(1) ϕ is a Freiman s-isomorphism when restricted to A.
(2) ϕ is injective on sA.
(3) $\phi(x) \neq 0$ for all nonzero $x \in sA - sA$.

Then let $\phi\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ be a linear map chosen uniformly at random. Each nonzero $x \in sA - sA$ violates condition (3) with probability 2^{-m}. Since there are $< 2^m$ nonzero elements in $sA - sA$ by hypothesis, (3) is satisfied with with positive probability. Therefore, the desired Freiman s-isomorphism exists. □

Starting with $A \subseteq \mathbb{Z}$ of small doubling, we will find a large fraction of A that can be modeled inside a cyclic group whose size is comparable to $|A|$. It turns out to be enough to model a large subset of A rather than all of A. We will apply the Ruzsa covering lemma later on to recover the structure of the entire set A.

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

Theorem 7.7.3 (Ruzsa modeling lemma)
Let $A \subseteq \mathbb{Z}$. Let $s \geq 2$ and N be positive integers. Suppose $|sA - sA| \leq N$. Then there exists $A' \subseteq A$ with $|A'| \geq |A|/s$ such that A' is Freiman s-isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$.

Proof. Choose any prime $q > \max(sA - sA)$. For every choice of $\lambda \in [q-1]$, we define ϕ_λ as the composition of functions as follows

$$\phi = \phi_\lambda \colon \mathbb{Z} \xrightarrow{\text{mod } q} \mathbb{Z}/q\mathbb{Z} \xrightarrow{\cdot\lambda} \mathbb{Z}/q\mathbb{Z} \xrightarrow{(\text{mod } q)^{-1}} \{0, 1, \dots, q-1\}.$$

The first map is the mod q map. The second map sends x to λx. The last map inverts the mod q map $\mathbb{Z} \to \mathbb{Z}/q\mathbb{Z}$.

If $\lambda \in [q-1]$ is chosen uniformly at random, then each nonzero integer is mapped to a uniformly random element of $[q-1]$ under ϕ_λ, and so is divisible by N with probability $\leq 1/N$. Since there are fewer than N nonzero elements in $sA - sA$, there exists a choice of λ so that

$$N \nmid \phi_\lambda(x) \qquad \text{for any nonzero } x \in sA - sA. \tag{7.2}$$

Let us fix this λ from now on and write $\phi = \phi_\lambda$.

Among the three functions whose composition defines ϕ, the first map (i.e., mod q) and the second map ($\cdot\lambda$ in $\mathbb{Z}/q\mathbb{Z}$) are group homomorphisms, and hence Freiman s-homomorphisms. The last map is not a Freiman s-homomorphism, but it becomes one when restricted to an interval of at most q/s elements (see Proposition 7.6.6). By the pigeonhole principle, we can find an interval I with

$$\operatorname{diam} I < q/s$$

such that

$$A' = \{a \in A \colon \phi(a) \in I\}$$

has $\geq |A|/s$ elements. So ϕ sends A' Freiman s-homomorphically to its image.

We further compose ϕ with the mod N map to obtain

$$\psi \colon \mathbb{Z} \xrightarrow{\phi} \{0, 1, \dots, q-1\} \xrightarrow{\text{mod } N} \mathbb{Z}/N\mathbb{Z}.$$

We claim that ψ maps A' Freiman s-isomorphically to its image. Indeed, we saw that ψ is a Freiman s-homomorphism when restricted to A' (since both $\phi|_{A'}$ and the mod N map are). Now suppose $a_1, \dots, a_s, a_1', \dots, a_s' \in A'$ satisfy

$$\psi(a_1) + \cdots + \psi(a_s) = \psi(a_1') + \cdots + \psi(a_s'),$$

which is the same as saying that N divides

$$y \coloneqq \phi(a_1) + \cdots + \phi(a_s) - \phi(a_1') - \cdots - \phi(a_s') \in \mathbb{Z}.$$

By swapping $(a_1, \dots, a_s)$ with $(a_1', \dots, a_s')$ if needed, we may assume that $y \geq 0$. Since $\phi(A') \subseteq I$, we have $|\phi(a_i) - \phi(a_i')| \leq \operatorname{diam} I < q/s$ for each i, and thus

$$0 \leq y < q.$$

Let

$$x = a_1 + \cdots + a_s - a_1' - \cdots - a_s' \in sA - sA.$$

Since $\phi \bmod q$ is a group homomorphism,

$$\phi(x) \equiv \phi(a_1) + \cdots + \phi(a_s) - \phi(a_1') - \cdots - \phi(a_s') = y \pmod q.$$

Since

$$\phi(x), y \in [0, q) \cap \mathbb{Z} \quad \text{and} \quad \phi(x) \equiv y \pmod q,$$

we have $\phi(x) = y$. Since N divides $y = \phi(x)$, and by (7.2), $N \nmid \phi(x)$ for any nonzero $x \in sA - sA$, we must have $x = 0$. Thus

$$a_1 + \cdots + a_s = a_1' + \cdots + a_s'.$$

Hence A' is a set of size $\geq |A|/s$ that is Freiman s-isomorphic via ψ to its image in $\mathbb{Z}/N\mathbb{Z}$. □

Exercise 7.7.4 (Modeling arbitrary sets of integers)**.** Let $A \subseteq \mathbb{Z}$ with $|A| = n$.

(a) Let p be a prime. Show that there is some integer t relatively prime to p such that $\|at/p\|_{\mathbb{R}/\mathbb{Z}} \leq p^{-1/n}$ for all $a \in A$.

(b) Show that A is Freiman 2-isomorphic to a subset of $[N]$ for some $N = (4 + o(1))^n$.

(c) Show that (b) cannot be improved to $N = 2^{n-2}$.

(You may use the fact that the smallest prime larger than m has size $m + o(m)$.)

Exercise 7.7.5 (Sumset with 3-AP-free set)**.** Let A and B be n-element subsets of the integers. Suppose A is 3-AP free. Prove that $|A + B| \geq n(\log\log n)^{1/100}$ provided that n is sufficiently large.

Hint: Ruzsa triangle inequality, Plünnecke's inequality, Ruzsa model lemma, Roth's theorem

Exercise 7.7.6 (3-AP-free subsets of arbitrary sets of integers)**.** Prove that there is some constant $C > 0$ so that every set of n integers has a 3-AP-free subset of size $\geq ne^{-C\sqrt{\log n}}$.

7.8 Iterated Sumsets: Bogolyubov's Lemma

The goal of this section is to find a large Bohr set inside $2A - 2A$, provided that A is a relatively large subset of $\mathbb{Z}/N\mathbb{Z}$. The idea is due to Bogolyubov (1939).

Let us first explain what happens in the finite field model. Let $A \subseteq \mathbb{F}_2^n$ with $|A| \geq \alpha 2^n$. (Think of α as a constant for now.) Since A is arbitrary, we do not expect it to contain any large subspaces. But perhaps $A + A$ always does.

Question 7.8.1 (Large structure in $A + A$)

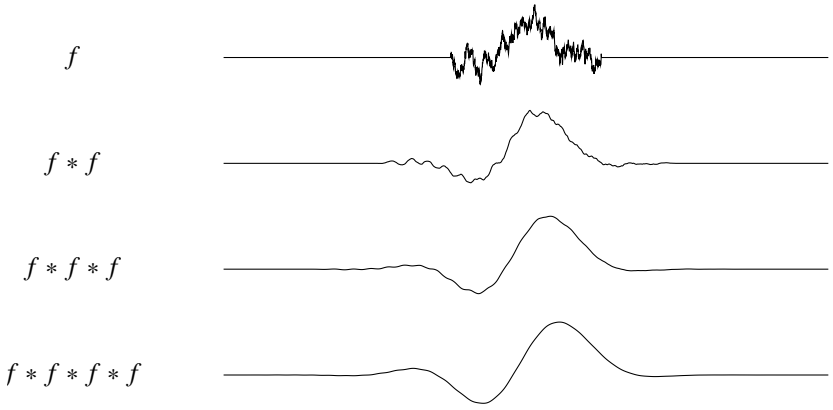
Suppose $A \subseteq \mathbb{F}_2^n$ and $|A| = \alpha 2^n$ where α is a constant independent of n. Must it be the case that $A + A$ contains a large subspace of codimension $O_\alpha(1)$?

The answer to the above question is no, as evidenced by the following example. (Niveau is French for level.)

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

***Example* 7.8.2** (Niveau set)**.** Let A be the set of all points in $\mathbb{F}_2^n$ with Hamming weight (number of 1 entries) at most $(n-c\sqrt{n})/2$. Note by the central limit theorem $|A| = (\alpha+o(1))2^n$ for for some constant $\alpha = \alpha(c) \in (0,1)$. The sumset $A+A$ consists of points in the boolean cube whose Hamming weight is at most $n - c\sqrt{n}$ and thus does not contain any subspace of codimension $< c\sqrt{n}$, by Lemma 6.5.4.

It turns out that the iterated sumset $2A-2A$ (same as $4A$ in $\mathbb{F}_2^n$) always contains a bounded codimensional subspace. The intuition is that taking sumsets "smooths" out the structure of a set, analogous to how convolutions in real analysis make functions more smooth.



Recall some basic properties of the Fourier transform. Given $A \subseteq \mathbb{F}_p^n$ with $|A| = \alpha p^n$, we have

$$\widehat{1_A}(0) = \alpha,$$

and by Parseval's identity

$$\sum_{r \in \mathbb{F}_p^n} |\widehat{1_A}(r)|^2 = \mathbb{E}_{x \in \mathbb{F}_p^n} |1_A(x)|^2 = \alpha.$$

We write $\omega = \exp(2\pi i/p)$ in the proof below.

Theorem 7.8.3 (Bogolyubov's lemma in $\mathbb{F}_p^n$)
If $A \subseteq \mathbb{F}_p^n$ and $|A| = \alpha p^n > 0$, then $2A-2A$ contains a subspace of codimension $< 1/\alpha^2$.

Proof. Let

$$f = 1_A * 1_A * 1_{-A} * 1_{-A},$$

which is supported on $2A-2A$. By the convolution identity (Theorem 6.1.7), noting that $\widehat{1_{-A}}(r) = \overline{\widehat{1_A}(r)}$, we have, for every $r \in \mathbb{F}_p^n$,

$$\widehat{f}(r) = \widehat{1_A}(r)^2 \widehat{1_{-A}}(r)^2 = |\widehat{1_A}(r)|^4.$$

By the Fourier inversion formula (Theorem 6.1.2), we have

$$f(x) = \sum_{r \in \mathbb{F}_p^n} \widehat{f}(r) \omega^{r \cdot x} = \sum_{r \in \mathbb{F}_p^n} |\widehat{1_A}(r)|^4 \omega^{r \cdot x}.$$

It suffices to find a subspace where f is positive since $f(x) > 0$ implies $x \in 2A - 2A$. We will take the subspace defined by large Fourier coefficients. Let

$$R = \left\{ r \in \mathbb{F}_p^n \backslash \{0\} : |\widehat{1_A}(r)| > \alpha^{3/2} \right\}.$$

We can bound the size of R using Parseval's identity:

$$|R|\,\alpha^3 \le \sum_{r \in R} |\widehat{1_A}(r)|^2 < \sum_{r \in \mathbb{F}_p^n} |\widehat{1_A}(r)|^2 = \mathbb{E}_x |1_A(x)|^2 = \alpha.$$

(Skip the preceding step if R is empty.) So

$$|R| < 1/\alpha^2.$$

If $r \notin R \cup \{0\}$, then $|\widehat{1_A}(r)| \le \alpha^{3/2}$. So, applying Parseval's identity again,

$$\begin{aligned} \sum_{r \notin R \cup \{0\}} |\widehat{1_A}(r)|^4 &\le \max_{r \notin R \cup \{0\}} |\widehat{1_A}(r)|^2 \sum_{r \notin R \cup \{0\}} |\widehat{1_A}(r)|^2 \\ &< \alpha^3 \sum_{r \in \mathbb{F}_p^n} |\widehat{1_A}(r)|^2 = \alpha^3 \mathbb{E}_x |1_A(x)|^2 = \alpha^4. \end{aligned}$$

Thus, for all $x \in R^\perp$, so that $x \cdot r = 0$ for all $r \in R$, we have

$$\begin{aligned} f(x) &= \sum_{r \in \mathbb{F}_p^n} |\widehat{1_A}(r)|^4 \operatorname{Re} \omega^{r \cdot x} \\ &\ge |\widehat{1_A}(0)|^4 + \sum_{r \in R} |\widehat{1_A}(r)|^4 - \sum_{r \notin R \cup \{0\}} |\widehat{1_A}(r)|^4 \\ &> \alpha^4 + 0 - \alpha^4 \\ &\ge 0. \end{aligned}$$

Thus $R^\perp \subseteq \operatorname{supp}(f) = 2A - 2A$. Since $|R| < 1/\alpha^2$, we have found a subspace of codimension $< 1/\alpha^2$ contained in $2A - 2A$. □

To formulate an analogous result for a cyclic group $\mathbb{Z}/N\mathbb{Z}$, we need the notion of a Bohr set, which was mentioned earlier in the context of Roth's theorem (Remark 6.4.7).

Definition 7.8.4 (Bohr sets in $\mathbb{Z}/N\mathbb{Z}$)
Let $R \subseteq \mathbb{Z}/N\mathbb{Z}$. Define

$$\operatorname{Bohr}(R, \varepsilon) = \{x \in \mathbb{Z}/N\mathbb{Z} : \|rx/N\|_{\mathbb{R}/\mathbb{Z}} \le \varepsilon, \text{for all } r \in R\}$$

where $\|\cdot\|_{\mathbb{R}/\mathbb{Z}}$ denotes the distance to the nearest integer. Its ***dimension*** is $|R|$ and ***width*** is ε. (Strictly speaking, the definition of a Bohr set includes the data of R and ε and not just the set of elements above.)

Bogolyubov's lemma holds over $\mathbb{Z}/N\mathbb{Z}$ after replacing subspaces by Bohr sets. Note that the dimension of a Bohr set of $\mathbb{Z}/N\mathbb{Z}$ corresponds to the codimension of a subspace in $\mathbb{F}_p^n$.

Theorem 7.8.5 (Bogolyubov's lemma in $\mathbb{Z}/N\mathbb{Z}$)
If $A \subseteq \mathbb{Z}/N\mathbb{Z}$ and $|A| = \alpha N$ then $2A - 2A$ contains some Bohr set $\operatorname{Bohr}(R, 1/4)$ with $|R| < 1/\alpha^2$.

With the right setup, the proof is essentially identical to that of Theorem 7.8.3.

Given $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, we define its ***Fourier transform*** to be the function $\widehat{f} : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ given by

$$\widehat{f}(r) = \mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f(x)\omega^{-rx}$$

where $\omega = \exp(2\pi i/N)$. Fourier inversion, Parseval's identity, and the convolution identity all work the same way.

Proof. Let

$$f = 1_A * 1_A * 1_{-A} * 1_{-A},$$

which is supported on $2A - 2A$. By the convolution identity, for every $r \in \mathbb{Z}/N\mathbb{Z}$,

$$\widehat{f}(r) = \widehat{1_A}^2(r)\widehat{1}^2_{-A}(r) = |\widehat{1_A}(r)|^4.$$

By Fourier inversion, we have (noting that f is real-valued)

$$f(x) = \sum_{r \in \mathbb{Z}/N\mathbb{Z}} \widehat{f}(r)\omega^{rx} = \sum_{r \in \mathbb{Z}/N\mathbb{Z}} |\widehat{1_A}(r)|^4 \omega^{rx}.$$

(Skip the preceding step if R is empty.) Let

$$R = \left\{ r \in \mathbb{Z}/N\mathbb{Z} \backslash \{0\} : |\widehat{1_A}(r)| > \alpha^{3/2} \right\}.$$

As earlier, we can bound the size of R using Parseval's identity:

$$|R|\, \alpha^3 \le \sum_{r \in R} |\widehat{1_A}(r)|^2 < \sum_{r \in \mathbb{F}_p^n} |\widehat{1_A}(r)|^2 = \mathbb{E}_x |1_A(x)|^2 = \alpha.$$

So

$$|R| < 1/\alpha^2.$$

We have

$$\sum_{r \notin R \cup \{0\}} |\widehat{1_A}(r)|^4 \le \alpha^3 \sum_{r \notin R \cup \{0\}} |\widehat{1_A}(r)|^2 < \alpha^4.$$

For all $x \in \mathrm{Bohr}(R, 1/4)$, every $r \in R$ satisfies $\|rx/N\|_{\mathbb{R}/\mathbb{Z}} \le 1/4$, and so $\cos(2\pi rx/N) \ge 0$. Thus every $x \in \mathrm{Bohr}(R, 1/4)$ satisfies

$$\begin{aligned} f(x) &= \sum_{r \in \mathbb{Z}/N\mathbb{Z}} |\widehat{1_A}(r)|^4 \omega^{r \cdot x} \\ &\ge |\widehat{1_A}(0)|^4 + \sum_{r \in R} |\widehat{1_A}(r)|^4 - \sum_{r \notin R \cup \{0\}} |\widehat{1_A}(r)|^4 \\ &> \alpha^4 + 0 - \alpha^4 \ge 0. \end{aligned}$$

Hence $\mathrm{Bohr}(R, 1/4) \subseteq 2A - 2A$. □

***Remark* 7.8.6** (Iterated sumsets and Goldbach conjecture)**.** The above proof hints at why it is easier to understand the iterated sumset kA when $k \ge 3$ than $k = 2$ (roughly speaking, we need two iterations to just apply Parseval, and the extra room is helpful). Exercise 7.8.7 below shows that the three-fold iterated sumset of every large subset of $\mathbb{F}_p^n$ contains a large

affine subspace (we do not always have a large subspace since the origin is not necessarily even in $3A$).

A related phenomenon arises in Goldbach conjecture. Let P denote the set of primes. The still open Goldbach conjecture states that $P + P$ contains all sufficiently large even integers. On the other hand, Vinogradov (1937) showed that $P + P + P$ contains all sufficiently large odd integers (also known as the weak or ternary Goldbach problem).

Our next goal is to find a large GAP in the Bohr set produced by Bogolyubov's lemma. To do this, we need some results from the geometry of numbers.

Exercise 7.8.7 (Bogolyubov with 3-fold sums)**.** Let $A \subseteq \mathbb{F}_p^n$ with $|A| = \alpha p^n$. Prove that $A + A + A$ contains a translate of a subspace of codimension $O(\alpha^{-3})$.

Exercise 7.8.8 (Bogolyubov with better bounds)**.** Let $A \subseteq \mathbb{F}_p^n$ with $|A| = \alpha p^n$.

(a) Show that if $|A + A| < 0.99 \cdot 2^n$, then there is some $r \in \mathbb{F}_p^n \setminus \{0\}$ such that $|\widehat{1_A}(r)| > c\alpha^{3/2}$ for some absolute constant $c > 0$.

(b) By iterating (a), show that $A+A$ contains at least 99% of a subspace of codimension $O(\alpha^{-1/2})$.

(c) Deduce that $2A - 2A$ contains a subspace of codimension $O(\alpha^{-1/2})$.

7.9 Geometry of Numbers

We will need some results concerning lattices and convex bodies belonging to a topic in number theory called the geometry of numbers.

Definition 7.9.1 (Lattice)

A ***lattice*** in $\mathbb{R}^d$ is a set of the form

$$\Lambda = \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_d = \{n_1v_1 + \cdots + n_dv_d : n_1, \ldots, n_d \in \mathbb{Z}\}$$

where $v_1, \ldots, v_d \in \mathbb{R}^d$ are linearly independent vectors.

The ***fundamental parallelepiped*** of a lattice Λ with respect to the basis $v_1, \ldots, v_d$ is

$$\{x_1v_1 + \cdots + x_dv_d : x_1, \ldots, x_d \in [0, 1)\}.$$

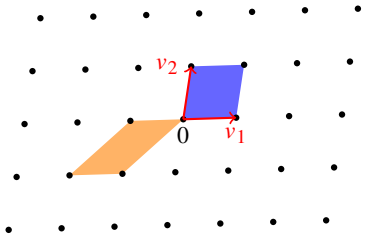
The ***determinant*** of this lattice is defined to be

$$\det \Lambda := \left| \det \begin{pmatrix} | & \cdots & | \\ v_1 & \cdots & v_d \\ | & \cdots & | \end{pmatrix} \right|.$$

This is the absolute value of the determinant of a matrix with $v_1, \ldots, v_d$ as columns.

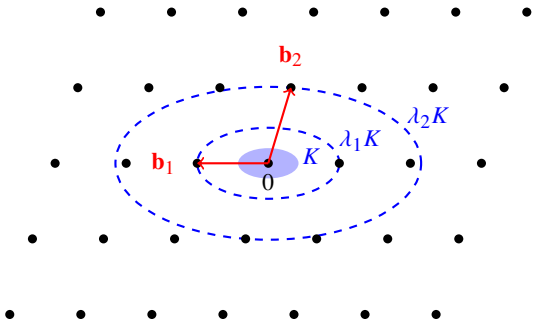
Given a lattice, there are many choices of a basis for the lattice. The determinant of a lattice does not depend on the choice of a basis, and equals the volume of every fundamental parallelepiped. Translations of the fundamental parallelepiped by lattice vectors tiles (i.e., partitions) the space.

An example of a lattice is illustrated below. Two different fundamental parallelepipeds are shaded.



Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice. Let $K \subseteq \mathbb{R}^d$ be a centrally symmetric convex body (here ***centrally symmetric*** means that $-x \in K$ whenever $x \in K$). For each $\lambda \geq 0$, let $\lambda K = \{\lambda x : x \in K\}$ be the dilation of K by a factor λ.

As illustrated below, imagine an animation where at time λ we see λK. This growing convex body initially is just the origin, and at some point it sees its first nonzero lattice point $\mathbf{b}_1$. Let us continue to grow this convex body. Later, at some point, it sees the first lattice point $\mathbf{b}_2$ in a new dimension not seen previously. And we can continue until the convex body grows big enough to contain lattice points that span all directions.



The process of dilating a convex body motivates the next definition.

Definition 7.9.2 (Successive minima)

Let Λ be a lattice in $\mathbb{R}^d$ and $K \subseteq \mathbb{R}^d$ a centrally symmetric convex body. For each $1 \leq i \leq d$, the ***ith successive minimum*** of K with respect to Λ is defined to be

$$\lambda_i = \inf\{\lambda \geq 0 : \dim(\operatorname{span}(\lambda K \cap \Lambda)) \geq i\}.$$

Equivalently, λ_i is the minimum λ such that λK contains i linearly independent lattice vectors from Λ.

A ***directional basis*** of K with respect to Λ is a basis $\mathbf{b}_1, \ldots, \mathbf{b}_d$ of $\mathbb{R}^d$ such that $\mathbf{b}_i \in \lambda_i K \cap \Lambda$ for each $i = 1, \ldots, d$.

Note that there may be more than one possible directional basis.

***Example* 7.9.3** (A directional basis does not necessarily generate the lattice)**.** Let $e_1, \ldots, e_8$ be the standard basis vectors in $\mathbb{R}^8$. Let $v = (e_1 + \cdots + e_8)/2$. Consider the lattice

$$\Lambda = \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_7 \oplus \mathbb{Z}v = \mathbb{Z}^8 + \{0, v\}.$$

Let K be the unit ball in $\mathbb{R}^8$. Note that the directional basis of K with respect to Λ is $e_1, \ldots, e_8$,

as all nonzero lattice points in Λ have length ≥ 1 (in particular, $|v| = \sqrt{2}$). This example shows that the directional basis of a convex body K is not necessarily a $\mathbb{Z}$-basis of Λ.

In the next section, we will apply the following fundamental result from the geometry of numbers (Minkowski 1896).

Theorem 7.9.4 (Minkowski's second theorem)
Let $\Lambda \in \mathbb{R}^d$ be a lattice and $K \subseteq \mathbb{R}^d$ a centrally symmetric convex body. Let $\lambda_1 \leq \cdots \leq \lambda_d$ be the successive minima of K with respect to Λ. Then

$$\lambda_1 \dots \lambda_d \mathrm{vol}(K) \leq 2^d \det(\Lambda).$$

***Example* 7.9.5.** Note that Minkowski's second theorem is tight when

$$K = \left[-\frac{1}{\lambda_1}, \frac{1}{\lambda_1}\right] \times \cdots \times \left[-\frac{1}{\lambda_d}, \frac{1}{\lambda_d}\right]$$

and Λ is the lattice $\mathbb{Z}^d$.

We will prove this theorem in the remainder of the section. The proof, while not long, is rather tricky. Feel free to skip the proof and jump to the next section.

Here is a simple geometric pigeonhole principle (Blichfeldt 1914).

Theorem 7.9.6 (Blichfeldt's theorem)
Let $\Lambda \subseteq \mathbb{R}^d$ be a lattice and $K \subseteq \mathbb{R}^d$ be a measurable set with $\mathrm{vol}(K) > \det(\Lambda)$. Then there are distinct points $x, y \in K$ with $x - y \in \Lambda$.

Proof. Fix a fundamental parallelepiped P. Then $v + P$ tiles $\mathbb{R}^d$ as v ranges over Λ. Partition K by this tiling. For the portion of K lying in $v + P$, translate it by $-v$ to bring it back inside P. Then the parts of K all end up back inside P via translations by lattice vectors. Since $\mathrm{vol}\, K > \mathrm{vol}\, P = \det \Lambda$, some distinct pair of points $x, y \in K$ must end up at the same point of P. This then implies that $x - y \in \Lambda$. □

Here is an easy corollary (though we will not need it).

Theorem 7.9.7 (Minkowski's first theorem)
Let Λ be a lattice in $\mathbb{R}^d$ and $K \subseteq \mathbb{R}^d$ a centrally symmetric convex body. If $\mathrm{vol}(K) > 2^d \det(\Lambda)$, then K contains a nonzero point of Λ.

Proof. We have $\mathrm{vol}(\frac{1}{2}K) = 2^{-d} \mathrm{vol}(K) > \det(\Lambda)$. By Blichfeldt's theorem there exist distinct $x, y \in \frac{1}{2}K$ such that $x - y \in \Lambda$. The point $x - y$ is the midpoint of $2x$ and $-2y$, both of which lie in K (using the face that K is centrally symmetric) and hence $x - y$ lies in K (since K is convex). □

Note that Minkowski's first theorem is tight for $K = [-1, 1]^d$ and $\mathbb{Z}^d$.

Proof of Minkowski's second theorem (Theorem 7.9.4). The idea is to grow K until we hit a point of Λ, and then continue growing, but only in the complementary direction. However rigorously carrying out this procedure is very tricky (and easy to get wrong).

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

In the argument below, K is open (i.e., does not include the boundary). Fix a directional basis $\mathbf{b}_1, \dots, \mathbf{b}_d$. For each $1 \le j \le d$, define map $\phi_j : K \to K$ by sending each point $x \in K$ to the center of mass of the $(j-1)$-dimensional slice of K which contains x and is parallel to $\operatorname{span}_{\mathbb{R}}\{\mathbf{b}_1, \dots, \mathbf{b}_{j-1}\}$. In particular, $\phi_1(x) = x$ for all $x \in K$.

Define a function $\psi : K \to \mathbb{R}^d$ by

$$\psi(x) = \sum_{j=1}^{d} \left(\frac{\lambda_j - \lambda_{j-1}}{2} \right) \phi_j(x),$$

where by convention we let $\lambda_0 = 0$.

For $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_d\mathbf{b}_d \in \mathbb{R}^d$ with $x_1, \dots, x_d \in \mathbb{R}$, we have

$$\phi_j(\mathbf{x}) = \sum_{i<j} c_{j,i}(x_j, \dots, x_d)\mathbf{b}_i + \sum_{i \ge j} x_i \mathbf{b}_i$$

for some continuous functions $c_{j,i}$. By examining the coefficient of each $\mathbf{b}_i$, we find

$$\psi(\mathbf{x}) = \sum_{i=1}^{d} \left(\frac{\lambda_i x_i}{2} + \psi_i(x_{i+1}, \dots, x_d) \right) \mathbf{b}_i$$

for some continuous functions $\psi_i(x_{i+1}, \dots, x_d)$, so its Jacobian $\partial\psi(\mathbf{x})/\partial\mathbf{x}_j$ with respect to the basis $(\mathbf{b}_1, \dots, \mathbf{b}_d)$ is upper triangular with diagonal $(\lambda_1/2, \dots, \lambda_d/2)$. Therefore

$$\operatorname{vol}\psi(K) = \frac{\lambda_1 \cdots \lambda_d}{2^d} \operatorname{vol} K. \tag{7.3}$$

For any distinct points $\mathbf{x} = \sum x_i\mathbf{b}_i$, $\mathbf{y} = \sum y_i\mathbf{b}_i$ in K, let k be the largest index such that $x_k \ne y_k$. Then $\phi_i(\mathbf{x})$ agrees with $\phi_i(\mathbf{y})$ for all $i > k$. So

$$\begin{aligned}
\psi(\mathbf{x}) - \psi(\mathbf{y}) &= \sum_{j=1}^{d} (\lambda_j - \lambda_{j-1}) \left(\frac{\phi_j(\mathbf{x}) - \phi_j(\mathbf{y})}{2} \right) \\
&= \sum_{j=1}^{k} (\lambda_j - \lambda_{j-1}) \left(\frac{\phi_j(\mathbf{x}) - \phi_j(\mathbf{y})}{2} \right) \in \sum_{j=1}^{k} (\lambda_j - \lambda_{j-1}) K = \lambda_k K.
\end{aligned}$$

The $\in$ step is due to K being centrally symmetric and convex. The coefficient of $\mathbf{b}_k$ in $(\psi(\mathbf{x}) - \psi(\mathbf{y}))$ is $\lambda_k(x_k - y_k)/2 \ne 0$. So $\psi(\mathbf{x}) - \psi(\mathbf{y}) \notin \operatorname{span}_{\mathbb{R}}\{\mathbf{b}_1, \mathbf{b}_2, \dots \mathbf{b}_{k-1}\}$. But we just saw that $\psi(\mathbf{x}) - \psi(\mathbf{y}) \in \lambda_k K$ Recall that K is open, and also $\lambda_k K \cap \Lambda$ is contained in $\operatorname{span}_{\mathbb{R}}\{\mathbf{b}_1, \mathbf{b}_2, \dots \mathbf{b}_{k-1}\}$. Thus $\psi(\mathbf{x}) - \psi(\mathbf{y}) \notin \Lambda$.

So $\psi(K)$ contains no two points separated by a nonzero lattice vector. By Blichfeldt's theorem (Theorem 7.9.6), we deduce $\operatorname{vol}\psi(K) \le \det\Lambda$. Combined with (7.3), we deduce

$$\lambda_1 \cdots \lambda_d \operatorname{vol} K \le 2^d \operatorname{vol}\psi(K) \le 2^d \det\Lambda. \qquad \square$$

7.10 Finding a GAP in a Bohr Set

Now we use Minkowski's second theorem to prove that a Bohr set of low dimension contains a large GAP.

Theorem 7.10.1 (Large GAP in a Bohr set)
Let N be a prime. Every Bohr set of dimension d and width $\varepsilon \in (0,1)$ in $\mathbb{Z}/N\mathbb{Z}$ contains a proper GAP with dimension at most d and volume at least $(\varepsilon/d)^d N$.

Proof. Let $R = \{r_1, \dots, r_d\} \subseteq \mathbb{Z}/N\mathbb{Z}$. Recall that

$$\mathrm{Bohr}(R,\varepsilon) = \left\{x \in \mathbb{Z}/N\mathbb{Z} : \|xr/N\|_{\mathbb{R}/\mathbb{Z}} \le \varepsilon \text{ for all } r \in R\right\}.$$

Let

$$v = \left(\frac{r_1}{N}, \dots, \frac{r_d}{N}\right).$$

Thus for each $x = 0, 1, \dots, N-1$, we have $x \in \mathrm{Bohr}(R,\varepsilon)$ if and only if some element of $xv + \mathbb{Z}^d$ lies in $[-\varepsilon,\varepsilon]^d$.

Let

$$\Lambda = \mathbb{Z}^d + \mathbb{Z}v \subseteq \mathbb{R}^d$$

be a lattice consisting of all points in $\mathbb{R}^d$ that are congruent mod 1 to some integer multiple of v. Note $\det(\Lambda) = 1/N$ since there are exactly N points of Λ within each translate of the unit cube. We consider the convex body $K = [-\varepsilon,\varepsilon]^d$. Let $\lambda_1, \dots, \lambda_d$ be the successive minima of K with respect to Λ. Let $\mathbf{b}_1, \dots, \mathbf{b}_d$ be the directional basis. We know

$$\|\mathbf{b}_j\|_\infty \le \lambda_j \varepsilon \text{ for all } j.$$

For each $1 \le j \le d$, let $L_j = \lceil 1/(\lambda_j d)\rceil$. If $0 \le l_j < L_j$ then

$$\|l_j \mathbf{b}_j\|_\infty < \frac{\varepsilon}{d}.$$

If we have integers $l_1, \dots, l_d$ with $0 \le l_i < L_i$ for all i then

$$\|l_1\mathbf{b}_1 + \cdots + l_d\mathbf{b}_d\|_\infty \le \varepsilon.$$

For each $1 \le j \le d$, there is some $0 \le x_j < N$ so that $\mathbf{b}_j \in x_j v + \mathbb{Z}^d$, so its ith coordinate lies in $x_i r_i/N + \mathbb{Z}^d$. The ith coordinate in the above L^∞ bound gives

$$\left\|\frac{(l_1x_1 + \cdots + l_dx_d)r_i}{N}\right\|_{\mathbb{R}/\mathbb{Z}} \le \varepsilon \text{ for all } i.$$

Thus, the GAP

$$l_1x_1 + \cdots + l_dx_d, \qquad 0 \le l_i < L_i \text{ for each } 1 \le i \le d$$

is contained in $\mathrm{Bohr}(R,\varepsilon)$. It remains to show that this GAP is large and proper. It volume is, applying Minkowski's second theorem,

$$L_1 \cdots L_k \ge \frac{1}{\lambda_1 \cdots \lambda_d \cdot d^d} \ge \frac{\mathrm{vol}(K)}{2^d \det(\Lambda) d^d} = \frac{(2\varepsilon)^d}{2^d (1/N) d^d} = \left(\frac{\varepsilon}{d}\right)^d N.$$

Now we check that the GAP is proper. It suffices to show that if

$$l_1x_1 + \cdots + l_dx_d \equiv l_1'x_1 + \cdots + l_d'x_d \pmod N,$$

then we must have $l_i = l_i'$ for all i. Setting

$$\mathbf{b} = (l_1 - l_1')\mathbf{b}_1 + \cdots + (l_d - l_d')\mathbf{b}_d,$$

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

we have $\mathbf{b} \in \mathbb{Z}^d$. Furthermore

$$\|\mathbf{b}\|_\infty \le \sum_{i=1}^{d} \frac{1}{\lambda_i d} \|\mathbf{b}_i\|_\infty \le \varepsilon < 1,$$

so actually $\mathbf{b}$ must be 0. Since $b_1, \dots, b_d$ is a basis we must have $l_i = l_i'$ for all i, as desired. □

7.11 Proof of Freiman's Theorem

We are now ready to prove Freiman's theorem by putting together all the ingredients in this chapter. Let us recall what we have proved.

- **Plünnecke's inequality** (Theorem 7.3.1): $|A+A| \le K|A|$ implies $|mA - nA| \le K^{m+n}|A|$ for all $m, n \ge 0$.
- **Ruzsa covering lemma** (Theorem 7.4.1): if $|X+B| \le K|B|$, then there exist some $T \subseteq X$ with $|T| \le K$ such that $X \subseteq T + B - B$.
- **Ruzsa modeling lemma** (Theorem 7.7.3): if $A \subseteq \mathbb{Z}$ and $|sA - sA| \le N$, then there exists $A' \subseteq A$ with $|A'| \ge |A|/s$ such that A' is Freiman s-isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$.
- **Bogolyubov's lemma** (Theorem 7.8.5): for every $A \subseteq \mathbb{Z}/N\mathbb{Z}$ with $|A| = \alpha N$, $2A - 2A$ contains some Bohr set with dimension $< 1/\alpha^2$ and width $1/4$.
- By a geometry of numbers argument (Theorem 7.10.1), for every prime N, every Bohr set of dimension d and width $\varepsilon \in (0,1)$ contains a proper GAP with dimension $\le d$ and volume $\ge (\varepsilon/d)^d N$.

Now we will prove Freiman's theorem. We restate it below with the bounds that we will prove.

Theorem 7.11.1 (Freiman's theorem)
Let $A \subseteq \mathbb{Z}$ be a finite set satisfying $|A+A| \le K|A|$. Then A is contained in a GAP of dimension at most $d(K)$ and volume at most $f(K)|A|$, where $d(K) \le \exp(K^C)$ and $f(K) \le \exp(\exp(K^C))$ for some absolute constant C.

Proof. By Plünnecke's theorem, we have $|8A - 8A| \le K^{16}|A|$. Let N be a prime with $K^{16}|A| \le N \le 2K^{16}|A|$ (it exists by Bertrand's postulate). By Ruzsa modeling lemma, some $A' \subseteq A$ with $|A'| \ge |A|/8$ is Freiman 8-isomorphic to a subset B of $\mathbb{Z}/N\mathbb{Z}$.

Applying Bogolyubov's lemma on $B \subseteq \mathbb{Z}/N\mathbb{Z}$, with

$$\alpha = \frac{|B|}{N} = \frac{|A'|}{N} \ge \frac{|A|}{8N} \ge \frac{1}{16K^{16}},$$

we deduce that $2B - 2B$ contains a Bohr set with dimension $< 256K^{32}$ and width $1/4$. By Theorem 7.10.1, $2B - 2B$ contains a proper GAP with dimension $d < 256K^{32}$ and volume $\ge (4d)^{-d}N$.

Since B is Freiman 8-isomorphic to A', $2B - 2B$ is Freiman 2-isomorphic to $2A' - 2A'$ (why?). Note GAPs are preserved by Freiman 2-isomorphisms (why?). Hence, the proper GAP in $2B - 2B$ is mapped to a proper GAP $Q \subseteq 2A' - 2A'$ with the same dimension ($\le d$)

and volume ($\geq (4d)^{-d}N$). We have

$$|A| \leq 8\,|A'| \leq 8N \leq 8(4d)^d\,|Q|\,.$$

Since $Q \subseteq 2A' - 2A' \subseteq 2A - 2A$, we have $Q + A \subseteq 3A - 2A$. By Plünnecke's inequality,

$$|Q + A| \leq |3A - 2A| \leq K^5\,|A| \leq 8K^5(4d)^d\,|Q|\,.$$

By the Ruzsa covering lemma, there exists a subset X of A with $|X| \leq 8K^5(4d)^d$ such that $A \subseteq X + Q - Q$. It remains to contain $X + Q - Q$ in a GAP.

By using two elements in each direction, X is contained in a GAP of dimension $|X| - 1$ and volume $\leq 2^{|X|-1}$. Since Q is a proper GAP with dimension $d < 256K^{32}$ and volume $\leq |2A - 2A| \leq K^4\,|A|$, $Q - Q$ is a GAP with dimension d and volume $\leq 2^d K^4\,|A|$. It follows that $A \subseteq X + Q - Q$ is contained in a GAP with

$$\text{dimension} \leq |X| - 1 + d \leq 8(4d)^d K^5 + d - 1 = e^{K^{O(1)}}$$

(recall $d < 256K^{32}$) and

$$\text{volume} \leq 2^{|X|-1+d}K^4\,|A| = e^{e^{K^{O(1)}}}\,|A|\,. \qquad \square$$

The following exercise asks to improve the quantitative bounds on Freiman's theorem.

Exercise 7.11.2 (Improved bounds on Freiman's theorem)**.** Using a more efficient covering lemma from Exercise 7.4.3, prove Freiman's theorem with $d(K) = K^{O(1)}$ and $f(K) = \exp(K^{O(1)})$.

7.12 Polynomial Freiman–Ruzsa Conjecture

Here we explain one of the biggest open problems in additive combinatorics, known as the ***polynomial Freiman–Ruzsa conjecture (PFR)***. As mentioned in Remark 7.1.11, nearly optimal bounds $f(K) = K^{1+o(1)}$ and $d(K) = \exp(K^{1+o(1)})$ are known for Freiman's theorem. However, one can reformulate Freiman's theorem with significantly better quantitative dependencies.

PFR in the Finite Field Model

Let us first explain what happens in the finite field model $\mathbb{F}_2^n$. Theorem 7.5.1 showed that if $A \subseteq \mathbb{F}_2^n$ has $|A + A| \leq K\,|A|$, then A is contained in a subspace of cardinality $\leq f(K)\,|A|$. As mentioned in Remark 7.5.2, the optimal constant is known and satisfies $f(K) = \Theta(2^{2K}/K)$. An example requiring this bound is $A \subseteq \mathbb{F}^{m+n}$ defined by $A = \{e_1, \ldots, e_n\} \times \mathbb{F}_2^m$ (where $e_1, \ldots, e_n$ are the coordinate basis vectors of $\mathbb{F}_2^n$). Here $K = |A + A|\,/|A| \sim n/2$ and $|\langle A\rangle| = (2^n/n)\,|A|$. However, instead of trying to cover A by a single subspace, we can easy cover A by a small number of translates of a subspace with size comparable to A, namely A is covered by $\{e_1\} \times \mathbb{F}_2^m, \ldots, \{e_n\} \times \mathbb{F}_2^m$, which are translates of each other and each has size $\leq |A|$.

The Polynomial Freiman–Ruzsa conjecture in $\mathbb{F}_2^n$ proposes a variant of Freiman's theorem with polynomial bounds, where we are only required to cover a large fraction of A. Ruzsa (1999) attributes the conjecture to Marton.

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

Conjecture 7.12.1 (Polynomial Freiman–Ruzsa in $\mathbb{F}_2^n$)
If $A \subseteq \mathbb{F}_2^n$, and $|A+A| \le K|A|$, then there exists a subspace $V \subseteq \mathbb{F}_2^n$ with $|V| \le |A|$ such that A can be covered by $K^{O(1)}$ cosets of V.

The best current result says that in Conjecture 7.12.1 one can cover A by $\exp((\log K)^{O(1)})$ cosets of V (Sanders 2012). This is called a quasipolynomial bound.

This conjecture has several equivalent forms. Here we give some highlights. For more details, including proofs of equivalence, see the online note accompanying Green (2005c) titled *Notes on the Polynomial Freiman–Ruzsa Conjecture*.

For example, here is a formulation where we just need to use one subspace to cover a large fraction of A.

Conjecture 7.12.2 (Polynomial Freiman–Ruzsa in $\mathbb{F}_2^n$)
If $A \subseteq \mathbb{F}_2^n$, and $|A+A| \le K|A|$, then there exists an affine subspace $V \subseteq \mathbb{F}_2^n$ with $|V| \le |A|$ such that $|V \cap A| \ge K^{-O(1)}|A|$.

Proof of equivalence of Conjecture 7.12.1 and Conjecture 7.12.2. Conjecture 7.12.1 implies Conjecture 7.12.2 since by the pigeonhole principle, at least one of the cosets of V covers $\ge K^{-O(1)}$ fraction of A.

Now assume Conjecture 7.12.2. Let $A \subseteq \mathbb{F}_2^n$ with $|A+A| \le K|A|$. Let V be as in Conjecture 7.12.2. By the Ruzsa covering lemma (Theorem 7.4.1) with $X = A$ and $B = V \cap A$ we find $T \subseteq X$ with $|T| \le |X+B|/|X| \le |A+A|/|A| \le K$ such that $A \subseteq T+B-B \subseteq T+V$. The conclusion of Conjecture 7.12.1 holds. □

Here is another attractive equivalent formulation of the polynomial Freiman–Ruzsa conjecture in $\mathbb{F}_2^n$.

Conjecture 7.12.3 (Polynomial Freiman–Ruzsa in $\mathbb{F}_2^n$)
If $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ satisfies

$$\left|\{f(x,y) - f(x) - f(y) : x, y \in \mathbb{F}_2^n\}\right| \le K,$$

then there exists a linear function $g : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that

$$\left|\{f(x) - g(x) : x \in \mathbb{F}_2^n\}\right| \le K^{O(1)}.$$

In Conjecture 7.12.3, it is straightforward to show a bound of 2^K instead of $K^{O(1)}$, since we can extend f to a linear function based on its values at some basis.

To state our third reformulation, we need the notion of the Gowers uniformity norm. Given a finite abelian group Γ, and $f : \Gamma \to \mathbb{C}$, define the ***$U^3$ uniformity norm*** of f by

$$\|f\|_{U^3} := \Big(\mathbb{E}_{x,y_1,y_2,y_3} f(x)\overline{f(x+y_1)f(x+y_2)f(x+y_3)} \cdot \cdot f(x+y_1+y_2)f(x+y_1+y_3)f(x+y_2+y_3)\overline{f(x+y_1+y_2+y_3)}\Big)^{1/8}.$$

The U^3 norm plays a central role in Gowers' proof of Szemerédi's theorem for 4-APs (the U^3 norm is also discussed in Exercise 6.2.14).

If $f\colon \mathbb{F}_2^n \to \{-1,1\}$ given by $f(x) = (-1)^{q(x)}$ where q is a quadratic polynomial in n variables over $\mathbb{F}_2$ (e.g., $x_1 + x_1x_2 + \cdots$), then it is not hard to check that the expression in the expectation above is identically 1 (it comes from taking three finite differences of q). So $\|f\|_{U^3} = 1$. For proving Szemerédi's theorem for 4-APs, one would like a "1% inverse result" showing that any $f\colon \mathbb{F}_2^n \to [-1,1]$ satisfying $\|f\|_{U^3} \geq \delta$ must correlates with some quadratic polynomial phase function $(-1)^{q(x)}$. Such a result is known but it remains open to find optimal quantitative bounds. The polynomial Freiman–Ruzsa conjecture in $\mathbb{F}_2^n$ is equivalent to a U^3 inverse statement with polynomial bounds (Green and Tao 2010b; Lovett 2012).

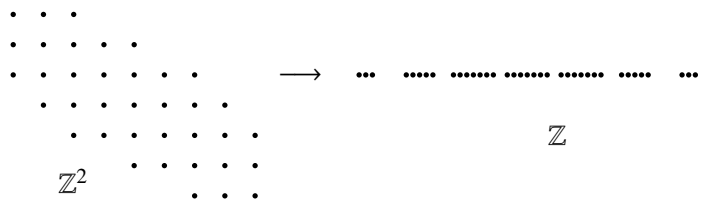
Conjecture 7.12.4 (U^3 inverse with polynomial bounds)
If $f : \mathbb{F}_2^n \to \mathbb{C}$ with $\|f\|_\infty \leq 1$ and $\|f\|_{U_3} \geq 1/K$, then there exists a quadratic polynomial $q(x_1,\dots,x_n)$ over $\mathbb{F}_2$ such that

$$\left|\mathbb{E}_{x\in\mathbb{F}_2^n} f(x)(-1)^{q(x)}\right| \geq K^{-O(1)}.$$

***Remark* 7.12.5** (Quantitative equivalence)**.** It is known that the bounds in each of the above conjectures are equivalent to each other up to a polynomial change. This means that if one statement is true with conclusion $\leq f(K)$ then all the other statements are true with conclusion $\leq Cf(K)^C$ (appropriately interpreted) with some absolute constant C.

PFR in the Integers

Now we formulate the polynomial Freiman–Ruzsa conjecture in $\mathbb{Z}$ instead of $\mathbb{F}_2^n$. It is not enough to use GAPs (Lovett and Regev 2017). Instead, we need to consider convex progressions.



Definition 7.12.6 (Convex progression)
A ***centered convex progression*** in an abelian group Γ is defined to be an affine map

$$\phi\colon \mathbb{Z}^d \cap B \to \Gamma$$

where B is a centrally symmetric convex body. We define its ***dimension*** to be d and its ***volume*** to be $|\mathbb{Z}^d \cap B|$.

Conjecture 7.12.7 (Polynomial Freiman–Ruzsa conjecture in $\mathbb{Z}$)
If $A \subseteq \mathbb{Z}$ satisfies $|A + A| \leq K|A|$, then one can cover A using $K^{O(1)}$ translates of some centered convex progression of dimension $O(\log K)$ and volume at most $|A|$.

More generally, one can formulate the polynomial Freiman–Ruzsa conjecture in an arbitrary abelian group.

Definition 7.12.8 (Centered convex coset progression)
In an abelian group, a ***centered convex coset progression*** is a set of the form $P + H$, where P is a centered convex progression and H is a subgroup. Its ***dimension*** is defined to be the dimension of P, and is ***volume*** is defined to be $|H| \operatorname{vol} P$.

Conjecture 7.12.9 (Polynomial Freiman–Ruzsa conjecture in abelian groups)
If A is a finite subset of an abelian group satisfying $|A + A| \le K|A|$, then one can cover A using $K^{O(1)}$ translates of some centered convex coset progression of dimension $O(\log K)$ and volume at most $|A|$.

For both Conjecture 7.12.7 and Conjecture 7.12.9, the best current result uses $\exp((\log K)^{O(1)})$ translates and dimension bound $(\log K)^{O(1)}$ (Sanders 2012, 2013).

7.13 Additive Energy and the Balog–Szemerédi–Gowers Theorem

We introduce a new way of measuring additive structure by counting the number of solutions to the equation $a + b = c + d$.

Definition 7.13.1 (Additive energy)
Let A be a finite set in an abelian group. Its ***additive energy*** is defined to be

$$E(A) := |\{(a, b, c, d) \in A \times A \times A \times A : a + b = c + d\}|.$$

***Remark* 7.13.2.** The additive energy of A counts 4-cycles in the bipartite Cayley graph with generating set A. It is called an "energy" since we can write it as an L^2 quantity

$$E(A) = \sum_x r_A(x)^2$$

where

$$r_A(x) := |\{(a, b) \in A \times A : a + b = x\}|$$

is the number of ways to write x as the sum of two elements of A.

We have the easy bound

$$2|A|^2 - |A| \le E(A) \le |A|^3.$$

The lower bound is due to trivial solutions $a + b = a + b$ and $a + b = b + a$. The lower bound is tight for sets without nontrivial solutions to $a + b = c + d$. The upper bound is due to d being determined by a, b, c when $a + b = c + d$. It is tight when A is a subgroup.

Here is the main question we explore in this section.

Question 7.13.3
What is the relationship between small doubling and large additive energy? (Both encode some notion of "lots of additive structure.")

One direction is easy.

Proposition 7.13.4 (Small doubling implies large additive energy)
Let A be a finite subset of an abelian group satisfying $|A+A| \le K|A|$. Then

$$E(A) \ge \frac{|A|^3}{K}.$$

Proof. Using $r_A(x)$ from Remark 7.13.2, By the Cauchy–Schwarz inequality

$$E(A) = \sum_{x \in A+A} r_A(x)^2 \ge \frac{1}{|A+A|}\left(\sum_{x \in A+A} r_A(x)\right)^2 = \frac{|A|^4}{|A+A|} \ge \frac{|A|^3}{K}. \qquad \square$$

The next example shows that the converse does not hold.

***Example* 7.13.5** (Large additive energy does not imply small doubling)**.** The set

$$A = [N] \cup \left\{2N+1, 2N+2, \ldots, 2N+2^N\right\}$$

is the union of a set of small doubling and a set without no additive structure. The first component has large additive energy, and so $E(A) = \Theta(N^3)$. On the other hand, the second component gives large doubling $|A+A| = \Theta(N^2)$.

However, we do have a converse if we allow passing to large subsets. Balog and Szemerédi (1994) showed that every set with large additive energy must contain a large subset with small doubling. Their proof used the regularity method, which required tower-type dependencies on the bounds. Gowers (2001) gave a new proof with much better bounds, and this result played a key role in his work on a new proof of Szemerédi's theorem. We will see Gowers' proof here. The presentation stems from Sudakov, Szemerédi, and Vu (2005).

Theorem 7.13.6 (Balog–Szemerédi–Gowers theorem)
Let A be a finite subset of an abelian group satisfying

$$E(A) \ge |A|^3/K.$$

Then there is a subset $A' \subseteq A$ with

$$|A'| \ge K^{-O(1)}|A| \quad \text{and} \quad |A'+A'| \le K^{O(1)}|A'|.$$

We will prove a version of the theorem allowing two different sets. Given two finite sets A and B in an abelian group, define their additive energy to be

$$\boldsymbol{E(A,B)} := |\{(a,b,a',b') \in A \times B \times A \times B : a+b = a'+b'\}|.$$

Then $E(A,A) = E(A)$.

Theorem 7.13.7 (Balog–Szemerédi–Gowers theorem)
Let A and B be finite subsets of the same abelian group. If $|A|, |B| \le n$ and

$$E(A,B) \ge n^3/K,$$

then there exist subsets $A' \subseteq A$ and $B' \subseteq B$ with

$$|A'|, |B'| \ge K^{-O(1)}n \quad \text{and} \quad |A'+B'| \le K^{O(1)}n.$$

Proof that Theorem 7.13.7 implies Theorem 7.13.6. Suppose $E(A) \geq |A|^3/K$. Apply Theorem 7.13.7 with $B = A$ to obtain $A', B' \subseteq A$ with $|A'|, |B'| \geq K^{-O(1)}|A|$ and $|A' + B'| \leq K^{O(1)}|A|$. Then by Corollary 7.3.6, a variant of the Ruzsa triangle inequality, we have

$$|A' + A'| \leq \frac{|A' + B'|^2}{|B'|} \leq K^{O(1)}|A|. \qquad \square$$

We will prove Theorem 7.13.7 by setting up a graph.

Definition 7.13.8 (Restricted sumset)
Let A and B be subsets of an abelian group and let G be a bipartite graph with vertex bipartition $A \cup B$. We define the ***restricted sumset*** $A +_G B$ to be the set of sums along edges of G:

$$A +_G B := \{a + b : (a, b) \in A \times B \text{ is an edge in } G\}.$$

Here is a graphical version of the Balog–Szemerédi–Gowers theorem.

Theorem 7.13.9 (Graph BSG)
Let A and B be finite subsets of an abelian group and let G be a bipartite graph with vertex bipartition $A \cup B$. If $|A|, |B| \leq n$,

$$e(G) \geq \frac{n^2}{K} \quad \text{and} \quad |A +_G B| \leq Kn,$$

then there exist subsets $A' \subseteq A$ and $B' \subseteq B$ with

$$|A'|, |B'| \geq K^{-O(1)}n \quad \text{and} \quad |A' + B'| \leq K^{O(1)}n.$$

Proof that Theorem 7.13.9 implies Theorem 7.13.7. Denote the number of ways to write x as $a + b$ by

$$r_{A,B}(x) := |\{(a, b) \in A \times B : a + b = x\}|.$$

Consider the "popular sums"

$$S = \left\{x \in A + B : r_{A,B}(x) \geq \frac{n}{2K}\right\}.$$

Build a bipartite graph G with bipartition $A \cup B$ such that $(a, b) \in A \times B$ is an edge if and only if $a + b \in S$.

We claim that G has many edges, by showing that "unpopular sums" account for at most half of $E(A, B)$. Note that

$$\frac{n^3}{K} \leq E(A, B) = \sum_{x \in S} r_{A,B}(x)^2 + \sum_{x \notin S} r_{A,B}(x)^2. \tag{7.4}$$

Because $r_{A,B}(x) < n/(2K)$ when $x \notin S$, we can bound the second term as

$$\sum_{x \notin S} r_{A,B}(x)^2 \leq \frac{n}{2K} \sum_{x \notin S} r_{A,B}(x) \leq \frac{n}{2K}|A||B| \leq \frac{n^3}{2K},$$

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

and setting back into (7.4) yields

$$\frac{n^3}{K} \le \sum_{x \in S} r_{A,B}(x)^2 + \frac{n^3}{2K},$$

and so

$$\sum_{x \in S} r_{A,B}(x)^2 \ge \frac{n^3}{2K}.$$

Moreover, because $r_{A,B}(x) \le |A| \le n$ for all x, it follows that

$$e(G) = \sum_{x \in S} r_{A,B}(x) \ge \sum_{x \in S} \frac{r_{A,B}(x)^2}{n} \ge \frac{n^2}{2K}.$$

Furthermore, $A +_G B \subseteq S$,

$$\frac{n}{2K} |A +_G B| \le |A|\,|B| \le n^2,$$

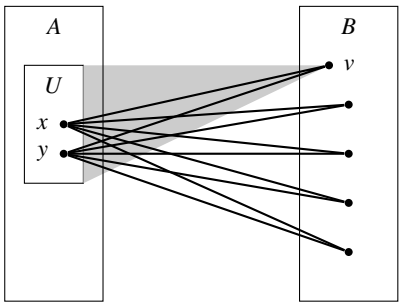
so $|A +_G B| \le 2Kn$. Hence, we can apply Theorem 7.13.9 to find sets $A' \subseteq A$ and $B' \subseteq B$ with the desired properties. □

Proof of Graph BSG

The remainder of this section will focus on proving BSG (Theorem 7.13.9). We begin with a few lemmas.

Lemma 7.13.10 (Path of length 2 lemma)
Let $\delta, \varepsilon > 0$. Let G be a bipartite graph with vertex bipartition $A \cup B$ and at least $\delta |A|\,|B|$ edges. Then there is some $U \subseteq A$ with $|U| \ge \delta |A| / 2$ such that at least $(1 - \varepsilon)$-fraction of the pairs $(x, y) \in U^2$ have at least $\varepsilon \delta^2 |B| / 2$ neighbors common to x and y.

The proof uses the **dependent random choice** technique from Section 1.7. Instead of quoting theorems from that section, let us prove the result from scratch.



Proof. Say that a pair $(x, y) \in A^2$ is "unfriendly" if it has $< \varepsilon \delta^2 |B| / 2$ common neighbors.

Choose $v \in B$ uniformly at random and let $U = N(v)$ be its neighborhood in v. We have

$$\mathbb{E}\,|U| = \mathbb{E}\,|N(v)| = \frac{e(G)}{|B|} \ge \delta |A|.$$

For each fixed pair $(x, y) \in A^2$, we have

$$\mathbb{P}(x, y \in U) = \mathbb{P}(x, y \in N(v)) = \frac{\operatorname{codeg}(x, y)}{|B|}.$$

So if (x, y) is unfriendly, then $\mathbb{P}(x, y \in U) < \varepsilon\delta^2/2$. Let X be the number of unfriendly pairs $(x, y) \in U^2$. Then

$$\mathbb{E}X = \sum_{\substack{(x,y)\in A^2 \\ \text{unfriendly}}} \mathbb{P}(x, y \in U) < \frac{\varepsilon\delta^2}{2} |A|^2 .$$

Hence, we have

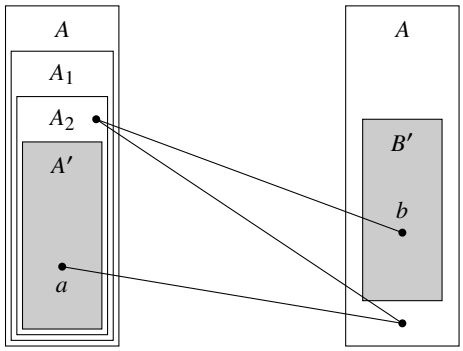
$$\mathbb{E}\left[|U|^2 - \frac{X}{\varepsilon}\right] \geq (\mathbb{E}|U|)^2 - \frac{\mathbb{E}X}{\varepsilon} > \frac{\delta^2}{2} |A|^2 .$$

So for some $v \in B$, $U = N(v)$ satisfies

$$|U|^2 - \frac{X}{\varepsilon} \geq \frac{\delta^2}{2} |A|^2 .$$

Then this $U \subseteq A$ satisfies $|U|^2 \geq \delta^2 |A|^2 /2$, and so $|U| \geq \delta |A| /2$. Moreover, we have $X \leq \varepsilon |U|^2$, so at most ε-fraction of pairs $(x, y) \in U^2$ are unfriendly. □

Lemma 7.13.11 (Path of length 3 lemma)
Let $\delta > 0$. Let G be a bipartite graph with vertex bipartition $A \cup B$ and at least $\delta |A| |B|$ edges. Then there are subsets $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq c\delta^C |A|$ and $|B'| \geq c\delta^C |B|$. such that the number of 3-edge paths joining every pair $(a, b) \in A' \times B'$ is at least $c\delta^C |A| |B|$. Here $c, C > 0$ are absolute constants.



Proof. We repeatedly trim low-degree vertices.

Call a pair of vertices in A "friendly" if they have $\geq \delta^3 |B| /20$ common neighbors. Define

$$A_1 := \left\{a \in A : \deg a \geq \frac{\delta}{2} |B|\right\}.$$

Since each vertex in $A \setminus A_1$ has $< \delta |B| /2$ neighbors, $e(A \setminus A_1, B) \leq \delta |A| |B| /2$. So

$$e(A_1, B) = e(A, B) - e(A \setminus A_1, B) \geq \delta |A| |B| - \frac{\delta}{2} |A| |B| \geq \frac{\delta}{2} |A| |B| .$$

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

Hence $|A_1| \geq \delta |A| /2$.

Construct $A_2 \subseteq A_1$ via the path of length 2 lemma (Lemma 7.13.10) on (A_1, B) with $\varepsilon = \delta/10$. Then, $|A_2| \geq \delta |A_1| /2 \geq \delta^2 |A| /4$ and $\leq (\delta/10)$-fraction pairs of vertices in A_2 are unfriendly.

Set

$$B' = \left\{ b \in B : \deg(b, A_2) \geq \frac{\delta}{4} |A_2| \right\}.$$

Since each vertex in $B \setminus B'$ has $< \delta |A_2| /4$ neighbors in A_2, $e(A_2, B \setminus B') \leq \delta |A_2| |B| /4$. Since every vertex in A_1 has $\geq \delta |B| /2$ neighbors in B, and $A_2 \subseteq A_1$, we have $e(A_2, B) \geq \delta |A_2| |B| /2$. Hence

$$e(A_2, B') = e(A_2, B) - e(A_2, B \setminus B') \geq \frac{\delta}{2} |A_2| |B| - \frac{\delta}{4} |A_2| |B| \geq \frac{\delta}{4} |A_2| |B|.$$

Hence $|B'| \geq \delta |B| /4$.

Let

$$A' = \{a \in A_2 : a \text{ is friendly to } \geq (1 - \delta/5)\text{-fraction of } A_2\}.$$

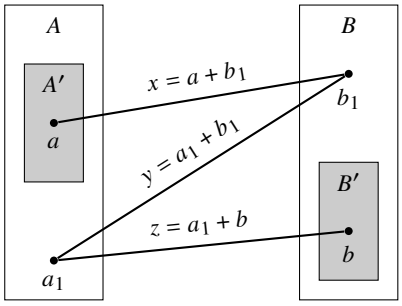
Since $\leq (\delta/10)$-fraction of pairs of vertices in A_2 are unfriendly, we have $|A'| \geq |A_2| /2 \geq \delta^2 |A| /8$.

We claim that that A' and B' satisfy the desired conclusions. Let $(a, b) \in A' \times B'$. Because b is adjacent to $\geq \delta |A_2| /4$ vertices in A_2 and a is friendly to $\geq (1 - \delta/5) |A_2|$ vertices in A_2, there are $\geq \delta |A_2| /20$ vertices in A_2 both friendly to a and adjacent to b. For each such $a_1 \in A_2$, there are $\geq \delta^3 |B| /20$ vertices $b_1 \in B$ for which ab_1a_1b is a path of length 3, so the number of paths of length 3 from a to b is at least

$$\frac{\delta}{20} |A_2| \cdot \frac{\delta^3}{20} |B| \geq \frac{\delta}{20} \cdot \frac{\delta^2}{4} |A| \cdot \frac{\delta^3}{20} |B| \geq \frac{\delta^6}{1600} |A| |B|.$$

Furthermore, recall that $|A'| \geq \delta^2 |A| /8$ and $|B'| \geq \delta |B| /4$. □

We can use the path of length 3 lemma to prove the graph-theoretic analogue of the Balog–Szemerédi–Gowers theorem.



Proof of Theorem 7.13.9 (Graph BSG). Since $e(G) \geq n^2/K$, we have $|A|, |B| \geq n/K$. By the path of length 3 lemma (Lemma 7.13.11), we can find $A' \subseteq A$ and $B' \subseteq B$ each with size $\geq K^{-O(1)}n$ such that for every $(a, b) \in A' \times B'$, there are $\geq K^{-O(1)}n^2$ paths ab_1a_1b in G with $a_1 \in A$ and $b_1 \in B$. Then, with

$$x = a + b_1, \qquad y = a_1 + b_1, \qquad z = a_1 + b,$$

MIT OCW: Graph Theory and Additive Combinatorics --- Yufei Zhao

we have

$$a + b = x - y + z.$$

This shows that every element of $A' + B'$ can be written as $x - y + z$ for some $x, y, z \in A +_G B$ in $\geq K^{-O(1)}n^2$ ways. (For a given $(a, b) \in A' \times B'$, these choices of x, y, z are genuinely distinct; why?) Thus

$$K^{-O(1)}n^2 \, |A' + B'| \leq |A +_G B|^3 \leq K^3 n^3.$$

Therefore $|A' + B'| \leq K^{O(1)}n$. □

Further Reading

See Ruzsa's lecture notes *Sumsets and Structure* (2009) for a comprehensive introduction to many topics related to set addition, including but not limited to Freiman's theorem.

Sanders' article *The Structure of Set Addition Revisited* (2013) provides a modern exposition of Freiman's theorem and his proof of the quasipolynomial Freiman–Ruzsa theorem. Lovett's article *An Exposition of Sanders' Quasi-Polynomial Freiman–Ruzsa Theorem* (2015) gives a gentle exposition of Sanders' proof in $\mathbb{F}_2^n$.

The methods discussed in this chapter play a central role in Gowers' proof of Szemerédi's theorem. The proof for 4-APs is especially worth studying, It contains many beautiful ideas and shows how these the topics in this chapter and the previous chapter are closely linked. See the original paper by Gowers (1998a) on Szemerédi's theorem for 4-APs as well as excellent lecture notes by Gowers (1998b), Green (2009b), and Soundararajan (2007).

Chapter Summary

- **Freiman's theorem.** Every $A \subseteq \mathbb{Z}$ with $|A+A| \le K|A|$ is contained in a generalized arithmetic progression (GAP) of dimension $\le d(K)$ and volume $\le f(K)|A|$.
  - Informally: a set with small doubling is contained in a small GAP.
  - Up to constants, this gives a complete characterization of integer sets with bounded doubling.
- **Rusza triangle inequality.** $|A|\,|B-C| \le |A-B|\,|A-C|$.
- **Plünnecke's inequality.** $|A+A| \le K|A|$ implies $|mA-nA| \le K^{m+n}|A|$.
- **Ruzsa covering lemma.** Idea: take a maximally disjoint set of translates, and their expansions must cover the entire space.
- **Freiman's theorem in groups with bounded exponent.** A set with bounded doubling is contained in a small subgroup.
- **Freiman s-homomorphisms** are maps preserving s-fold sums.
- **Ruzsa modeling lemma.** A set of integers with small doubling can be partially modeled as a large fraction of a cyclic group via a Freiman isomorphism.
- **Bogolyubov's lemma.** If A is large, then $2A-2A$ contains a large subspace (finite field model) or GAP (cyclic group).
- A large **Bohr set** contains a large GAP. Proof uses **Minkowski's second theorem** from the **geometry of numbers**.
- **Polynomial Freiman–Ruzsa conjeture**: a central conjecture in additive combinatorics. The finite field model version has several equivalent and attractive statements, one of which says: if $A \subseteq \mathbb{F}_2^n$, and $|A+A| \le K|A|$, then A can be covered using $K^{O(1)}$ translates of some subspace with cardinality $\le |A|$.
- The **additive energy** $E(A)$ of a set A is the number of solutions to $a+b=c+d$ in A.
- **Balog–Szemerédi–Gowers theorem.** If $E(A) \ge |A|^3/K$, then A has a subset A' with $|A'| \ge K^{-O(1)}|A|$ and $|A'+A'| \le K^{O(1)}|A'|$.
  - Informally: a set with large additive energy contains a large subset with small doubling.

MIT OpenCourseWare
https://ocw.mit.edu

18.225 Graph Theory and Additive Combinatorics

Fall 2023

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.