Graph Theory and Additive Combinatorics

Lecturer: Prof. Yufei Zhao

7
Structure of set addition

7.1 Structure of sets with small doubling

One of the main goals of additive combinatorics can be roughly described as understanding the behavior of sets under addition. In order to discuss this more precisely, we will begin with a few definitions.

Definition 7.1. Let A and B be finite subsets of an abelian group. Their ***sumset*** is defined as $A + B = \{a + b | a \in A, b \in B\}$. We can further define $A - B = \{a - b | a \in A, b \in B\}$ and $kA = \underbrace{A + A + \cdots + A}_{k \text{ times}}$ where k is a positive integer. Note that this is different from multiplying every element in A by k, which we denote the ***dilation*** $k \cdot A = \{kA | a \in A\}$.

Given a finite set of integers A, we want to understand how its size changes under these operations, giving rise to the following natural question:

Question 7.2. How large or small can $|A + A|$ be for a given value of $|A|$ where $A \subset \mathbb{Z}$?

It turns out that this is not a hard question. In $\mathbb{Z}$, we have precise bounds on the size of the sumset given the size of the set.

Proposition 7.3. *If A is a finite subset of $\mathbb{Z}$, then*

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

Proof. The right inequality follows from the fact that there are only $\binom{|A|+1}{2}$ unordered pairs of elements of A.

If the elements of A are $a_1 < a_2 < \cdots < a_{|A|}$, then note that $a_1 + a_1 < a_1 + a_2 < \cdots < a_1 + a_{|A|} < a_2 + a_{|A|} < \cdots < a_{|A|} + a_{|A|}$ is an increasing sequence of $2|A| - 1$ elements of $|A + A|$, so the left inequality follows. □

The upper bound is tight when there are no nontrivial collisions in $A + A$, that is, there are no nontrivial solutions to $a_1 + a_2 = a_1' + a_2'$ for $a_1, a_2, a_1', a_2' \in A$.

Example 7.4. If $A = \{1, a, a^2, \ldots, a^{n-1}\} \subset \mathbb{Z}$ for $a > 1$, then $|A + A| = \binom{n+1}{2}$.

The lower bound is tight when A is an arithmetic progression. Even if we instead consider arbitrary abelian groups, the problem is similarly easy. In a general abelian group G, we only have the trivial inequality $|A + A| \geq |A|$, and equality holds if A is a coset of some finite subgroup of G. The reason we have a stronger bound in $\mathbb{Z}$ is that there are no nontrivial finite subgroups of $\mathbb{Z}$.

A more interesting question that we can ask is what can we say about sets where $|A + A|$ is small. More precisely:

Definition 7.5. The ***doubling constant*** of a finite subset A of an abelian group is the ratio $|A + A|/|A|$.

Question 7.6. What is the structure of a set with bounded doubling constant (e.g. $|A + A| \leq 100|A|$)?

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

Example 7.7. If $A \subset \mathbb{Z}$ is a finite arithmetic progression, $|A + A| = 2|A| - 1 \leq 2|A|$, 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 7.8. If B is a finite arithmetic progression and $A \subseteq B$ has $|A| \geq C|B|$, then $|A + A| \leq |B + B| \leq 2|B| \leq 2C^{-1}|A|$, so A has doubling constant at most $2/C$.

A more substantial generalization of this is a d-dimensional arithmetic progression.

Definition 7.9. A ***generalized arithmetic progression (GAP)*** of dimension d is a set of the form

$$\{x_0 + \ell_1 x_1 + \cdots + \ell_d x_d \mid 0 \leq \ell_1 < L_d, \ldots, 0 \leq \ell_d < L_d, \ell_1, \ldots, \ell_d \in \mathbb{Z}\}$$

where $x_0, x_1, \ldots, x_d \in \mathbb{Z}$ and $L_1, \ldots, L_d \in \mathbb{N}$. The ***size*** of a GAP is defined as $L_1 L_2 \cdots L_d$. If there are no nontrivial coincidences among the elements of the GAP, it is called ***proper***.

Remark 7.10. Note that if a GAP is not proper, the size is not equal to the number of distinct elements, i.e. its cardinality.

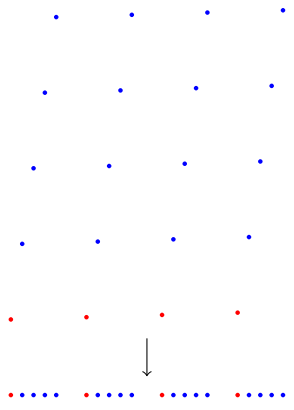

Figure 7.1: Picture of a 2-dimensional arithmetic progression as a projection of a lattice in $\mathbb{Z}^2$ into $\mathbb{Z}$.

It is not too hard to see that a proper GAP of dimension d has doubling constant at most 2^d. Furthermore, we have the same property that deleting a constant fraction of the elements of a GAP will still leave a set of small doubling constant. We have enumerated several examples of sets of small doubling constant, so it is natural to ask whether we can give an exact classification of such sets. We have an "inverse problem" to Question 7.6, asking whether every set with bounded doubling constant must be one of these examples.

This is not such an easy problem. Fortunately, a central result in additive combinatorics gives us a positive answer to this question.

Theorem 7.11 (Freiman's theorem). *If $A \subset \mathbb{Z}$ is a finite set and $|A+A| \leq K|A|$, then A is contained in a GAP of dimension at most $d(K)$ and size at most $f(K)|A|$, where $d(K)$ and $f(K)$ are constants depending only on K.*

Freiman (1973)

Remark 7.12. The conclusion of the theorem can be made to force the GAP to be proper, at the cost of increasing $d(K)$ and $f(K)$, using the fact below, whose proof we omit but can be found as Theorem 3.40 in the textbook by Tao and Vu.

Tao and Vu (2006)

Theorem 7.13. *If P is a GAP of dimension d, then P is contained in a proper GAP Q of dimension at most d and size at most $d^{C_0 d^3}|P|$ for some absolute constant $C_0 > 0$.*

Freiman's theorem gives us significant insight into the structure of sets of small doubling. We will see the proof of Freiman's theorem in the course of this chapter. Its proof combines ideas from Fourier analysis, the geometry of numbers, and classical additive combinatorics.

Freiman's original proof was difficult to read and did not originally get the recognition it deserved. Later on Ruzsa found a simpler proof, whose presentation we will mostly folllow. The theorem is sometimes called the Freiman–Ruzsa theorem. Freiman's theorem was brought into prominence as it and its ideas play central roles in Gowers' new proof of Szemerédi's theorem.

Ruzsa (1994)

If we consider again Example 7.4, then we have $K = \frac{|A|+1}{2} = \Theta(|A|)$. There isn't really a good way to embed this into a GAP. If we let the elements of A be $a_1 < a_2 < \cdots < a_{|A|}$, we can see that it is contained in a GAP of dimension $|A|-1$ and size $2^{|A|-1}$, by simply letting $x_0 = a_1$, $x_i = a_{i+1} - a_1$, and $L_i = 2$ for $1 \leq i \leq |A|-1$. Then this indicates that the best result we can hope for is showing $d(K) = O(K)$ and $f(K) = 2^{O(K)}$. This problem is still open.

Open problem 7.14. Is Theorem 7.11 true with $d(K) = O(K)$ and $f(K) = 2^{O(K)}$?

The best known result is due to Sanders, who also has the best known bound for Roth's Theorem (Theorem 6.12).

Theorem 7.15 (Sanders). *Theorem 7.11 is true with* $d(K) = K(\log K)^{O(1)}$, $f(K) = e^{K(\log K)^{O(1)}}$.

Sanders (2012)

In the asymptotic notation we assume that K is sufficiently large, say $K \geq 3$, so that $\log K$ is not too small.

Similar to how we discussed Roth's theorem, we will begin by analyzing a finite field model of the problem. In $\mathbb{F}_2^n$, if $|A+A| \leq K|A|$, then what would A look like? If A is a subspace, then it has doubling constant 1. A natural analogue of our inverse problem is to ask if all such A are contained in a subspace that is not much larger than A.

Theorem 7.16 ($\mathbb{F}_2^n$-analogue of Freiman). *If $A \subset \mathbb{F}_2^n$ has $|A+A| \leq K|A|$, 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 7.17. If we let A be a linearly independent set (i.e. a basis), then $K = \Theta(|A|)$ and the smallest subspace containing A will have cardinality $2^{|A|}$. Thus $f(K)$ must be exponential in K at least. We'll prove Theorem 7.16 in Section 7.3.

7.2 *Plünnecke–Ruzsa inequality*

Before we can prove Freiman's theorem (Theorem 7.11) or its finite field version (Theorem 7.16), we will need a few tools. We begin with one of many results named after Ruzsa.

Theorem 7.18 (Ruzsa triangle inequality). *If A, B, C are finite subsets of an abelian group, then*

$$|A||B-C| \leq |A-B||A-C|.$$

Proof. We will construct an injection

$$\phi : A \times (B-C) \hookrightarrow (A-B) \times (A-C).$$

For each $d \in B-C$, we can choose $b(d) \in B, c(d) \in C$ such that $d = b(d) - c(d)$. Then define $\phi(a,d) = (a-b(d), a-c(d))$. This is injective because if $\phi(a,d) = (x,y)$, then we can recover (a,d) from (x,y) because $d = y-x$ and $a = x + b(y-x)$. □

Remark 7.19. By replacing B with $-B$ and/or C with $-C$, we can change some of the plus signs into minus signs in this inequality. Unfortunately, this trick cannot be used to prove the similar inequality $|A||B+C| \leq |A+B||A+C|$. Nevertheless, we will soon see that this inequality is still true.

Remark 7.20. Where's the triangle? If we define $\rho(A,B) = \log \frac{|A-B|}{\sqrt{|A||B|}}$, then Theorem 7.18 states that $\rho(B,C) \leq \rho(A,B) + \rho(A,C)$. This looks like the triangle inequality, but unfortunately ρ is not actually a metric because $\rho(A,A) \neq 0$ in general. If we restrict to only looking at subgroups, however, then ρ is a bona fide metric.

The way that we use Theorem 7.18 is to control further doublings of a set of small doubling. Its usefulness is demonstrated by the following example.

Example 7.21. Suppose A is a finite subset of an abelian group with $|2A - 2A| \leq K|A|$. If we set $B = C = 2A - A$ in Theorem 7.18, then we get

$$|3A - 3A| \leq \frac{|2A - 2A|^2}{|A|} \leq K^2|A|.$$

We can repeat this with $B = C = 3A - 2A$ to get

$$|5A - 5A| \leq \frac{|3A - 3A|^2}{|A|} \leq K^4|A|$$

and so on, so for all m we have that $|mA - mA|$ is bounded by a constant multiple of $|A|$.

The condition $|2A - 2A| \leq K|A|$ is stronger than the condition $|A + A| \leq K|A|$. If we want to bound iterated doublings given just the condition $|A + A| \leq K|A|$, we need the following theorem.

Theorem 7.22 (Plünnecke–Ruzsa inequality)**.** *If A is a finite subset of an abelian group and $|A + A| \leq K|A|$, then $|mA - nA| \leq K^{m+n}|A|$.*

Plünnecke (1970)

Ruzsa (1989)

We think of polynomial changes in K as essentially irrelevant, so this theorem just says that if a set has small doubling then any iteration of the set is also small.

Remark 7.23. Plünnecke's original proof of the theorem did not receive much attention. Ruzsa later gave a simpler proof of Plünnecke's theorem. Their proofs involved the study of an object called a commutative layered graph, and involved Menger's theorem for flows and the tensor power trick. Recently Petridis gave a significantly simpler proof which uses some of the earlier ideas, which we will show here.

Petridis (2012)

In proving this theorem, we will generalize to the following theorem.

Set $B = A$ to recover Theorem 7.22

Theorem 7.24. *If A and B are finite subsets of an abelian group and $|A + B| \leq K|A|$, then $|mB - nB| \leq K^{m+n}|A|$.*

Petridis' proof relies on the following key lemma.

Lemma 7.25. *Suppose A and B are finite subsets of an abelian group. If $X \subseteq A$ is a nonempty subset which minimizes $\frac{|X+B|}{|X|}$, and $K' = \frac{|X+B|}{|X|}$, then $|X + B + C| \leq K'|X + C|$ for all finite sets C.*

Remark 7.26. We can think of this lemma in terms of a bipartite graph. If we consider the bipartite graph on vertex set $G_1 \sqcup G_2$, where G_1, G_2 are copies of the ambient abelian group G, with edges from g to $g + b$ for any $g \in G_1$, $g + b \in G_2$ where $b \in B$. Then if $N(S)$ denotes the neighborhood of a set of vertices S, then the lemma is considering

the ***expansion ratio*** $\frac{|N(A)|}{|A|} = \frac{|A+B|}{|A|}$. The lemma states that if X is a set whose expansion ratio K' is less than or equal to the expansion ratio of any of its subsets, then for any set C, $X + C$ also has expansion ratio at most K'.

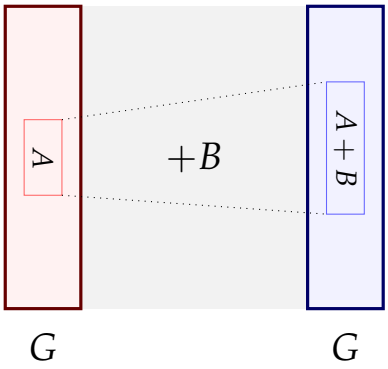

Figure 7.2: Bipartite graph where edges correspond to addition by an element of B.

Proof of Theorem 7.24 assuming Lemma 7.25. Assuming the key lemma, let us prove the theorem. Let X be a nonempty subset of A minimizing $\frac{|X+B|}{|X|}$, and let $K' = \frac{|X+B|}{|X|}$. Note that $K' \leq K$ by minimality. Applying the lemma with $C = rB$ where $r \geq 1$, we have $|X + (r+1)B| \leq K'|X + rB| \leq K|X + rB|$, so by induction $|X + rB| \leq K^r|X|$ for all $r \geq 0$. Applying Theorem 7.18 we have $|mB - nB| \leq \frac{|X+mB||X+nB|}{|X|} \leq K^{m+n}|X| \leq K^{m+n}|A|$. □

Proof of Lemma 7.25. We will proceed by induction on $|C|$. The base case of $|C| = 1$ is clear because for any finite set S, $S + C$ is a translation of S so $|S + C| = |S|$, thus $|X + B + C| = |X + B| = K'|X| = K'|X + C|$.

For the inductive step, assume $|C| > 1$, let $\gamma \in C$ and $C' = C \backslash \{\gamma\}$. Then

$$X + B + C = (X + B + C') \cup \big((X + B + \gamma) \backslash (Z + B + \gamma)\big)$$

where

$$Z = \{x \in X | x + B + \gamma \subseteq X + B + C'\}.$$

$Z \subseteq X$ so by minimality $|Z + B| \geq K'|Z|$. We have

$$\begin{aligned} |X + B + C| &\leq |X + B + C'| + |(X + B + \gamma) \backslash (Z + B + \gamma)| \\ &= |X + B + C'| + |X + B| - |Z + B| \\ &\leq K'|X + C'| + K'|X| - K'|Z| \\ &= K'(|X + C'| + |X| - |Z|). \end{aligned}$$

Now we want to understand the right hand side $X + C$. Note that

$$X + C = (X + C') \sqcup \big((X + \gamma) \backslash (W + \gamma)\big)$$

where

$$W = \{x \in X | x + \gamma \in X + C'\}.$$

In particular this is a disjoint union, so

$$|X + C| = |X + C'| + |X| - |W|.$$

We also have $W \subseteq Z$ because $x + \gamma \in X + C'$ implies $x + B + \gamma \subseteq X + B + C'$. Thus $|W| \leq |Z|$, so

$$|X + C| \geq |X + C'| + |X| - |Z|,$$

which, when combined with the above inequality, completes the induction. □

The key lemma also allows us to replace all the minus signs by pluses in Theorem 7.18 as promised.

Corollary 7.27. *If A, B, C are finite subsets of an abelian group, then $|A||B+C| \leq |A+B||A+C|$.*

Proof. Let $X \subseteq A$ be nonempty such that $\frac{|X+B|}{|X|}$ is minimal. Let $K = \frac{|A+B|}{|A|}$, $K' = \frac{|X+B|}{|X|} \leq K$. Then

$$\begin{aligned} |B+C| &\leq |X+B+C| \\ &\leq K'|X+C| && \text{(Lemma 7.25)} \\ &\leq K'|A+C| \\ &\leq K|A+C| \\ &= \frac{|A+B||A+C|}{|A|} \end{aligned}$$

□

7.3 *Freiman's theorem over finite fields*

We have one final lemma to establish before we can prove the finite field analogue of Frieman's theorem (Theorem 7.16).

Theorem 7.28 (Ruzsa covering lemma)**.** *Let X and B be subsets of an abelian group. If $|X+B| \leq K|B|$, then there exists a subset $T \subset X$ with $|T| \leq K$ such that $X \subset T + B - B$.*

Ruzsa (1999)
In essence, this theorem says that if it looks like $X + B$ is coverable by K translates of the set B (based off only size data), then X is in fact coverable by K translates of the slightly larger set $B - B$.

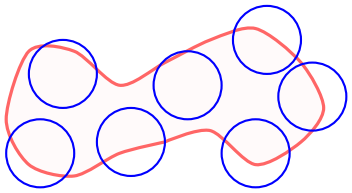

Figure 7.3: A maximal packing of a region with half balls

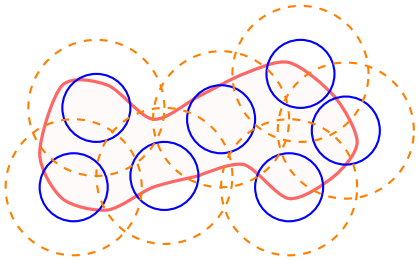

Figure 7.4: The maximal packing leads to a proper covering

The covering analogy provides the intuition for our proof. We treat the covering sets as balls in a metric space. Now, if we have a maximal packing of half-sized balls, expanding each to become a unit ball should produce a covering of the region. Note that maximal here means no more balls can be placed, not that the maximum possible number of balls have been placed. We formalize this to prove the Ruzsa covering lemma.

Proof. Let $T \subset X$ be a maximal subset such that $t + B$ is disjoint for all $t \in T$. Therefore, $|T||B| = |T+B| \leq |X+B| \leq K|B|$. So, $|T| \leq K$.

Now, as T is maximal, for all $x \in X$ there exists some $t \in T$ such that $(t+B) \cap (x+B) \neq \emptyset$. In other words, there exists $b, b' \in B$ such that $t + b = x + b'$. Hence $x \in t + B - B$ for some $t \in T$. Since this applies to all $x \in X$, we have $X \subset T + B - B$. □

The Ruzsa covering lemma is our final tool required for the proof of Freiman's theorem over finite fields (Theorem 7.16). The finite field model is simpler than working over $\mathbb{Z}$, and so it can be done with fewer tools compared to the original Freiman's theorem (Theorem 7.11).

Now, we will prove Freiman's theorem in groups with bounded exponent. This setting is slightly more general than finite fields.

Definition 7.29. The ***exponent*** of an abelian group (written additively) is the smallest positive integer r (if it exists) such that $rx = 0$ for all elements x of the group.

We also use $\langle A \rangle$ to refer to the subgroup of a group G generated by some subset A of G. By this notation, the exponent of a group G is $\max_{x \in G} |\langle x \rangle|$. With that notation, we can finally prove Ruzsa's analogue of Freiman's theorem over finite exponent abelian groups.

Theorem 7.30 (Ruzsa)**.** *Let A be a finite set in an abelian group with exponent $r < \infty$. If $|A + A| \leq K|A|$, then*

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

Ruzsa (1999)
This theorem is, in a sense, the converse of our earlier observation that if A is a large enough subset of some subgroup H, then A has small doubling

Proof. By the Plünnecke–Ruzsa inequality (Theorem 7.22), we have

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

Now, from the Ruzsa Covering Lemma (with $X = 2A - A$, $B = A$), there exists some $T \subset 2A - A$ with $|T| \leq K^4$ such that

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

Adding A to both sides, we have,

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

Iterating this, we have for any positive integer n,

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

Using the Ruzsa Covering Lemma allowed us to control the expression $nA - A$ nicely. If we had only used the Plünnecke–Ruzsa inequality (Theorem 7.22), the argument would have failed as the exponent of K would've blown up.

For sufficiently large n, we have $nA = \langle A \rangle$. Thus we can say,

$$\langle A \rangle \subset \langle T \rangle + A - A.$$

Due to the bounded exponent, we have,

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

And by the Plünnecke–Ruzsa inequality (Theorem 7.22),

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

Thus we have,

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

□

Example 7.31. In $\mathbb{F}_2^n$, if A is an independent subset (e.g. the basis of some subgroup), then A has doubling constant $K \approx |A|/2$, and $|\langle A \rangle| = 2^{|A|} \approx 2^{2K}|A|$. Thus the bound on $|\langle A \rangle|$ must be at least exponential in K.

It has recently been determined very precisely the maximum possible value of $|\langle A \rangle| / |A|$ over all $A \subset \mathbb{F}_2^\infty$ with $|A+A|/|A| \leq K$. Asymptotically, it is $\Theta\left(2^{2K}/K\right)$. Even-Zohar (2012)

For general r, we expect a similar phenomenon to happen. Ruzsa conjectured that $|\langle A \rangle| \leq r^{CK}|A|$. This result is proven for some r such as the primes. Ruzsa (1999) Even-Zohar and Lovett (2014)

Our proof for Freiman's theorem over abelian groups of finite exponent (Theorem 7.30) does not generalize to the integers. Indeed, in our proof above, $|\langle T \rangle|$ if we were working in $\mathbb{Z}$. The workaround is to model subsets of $\mathbb{Z}$ inside a finite group in a way that partially preserves additive structure.

7.4 *Freiman homomorphisms*

To understand any object, you should understand maps between them and the properties preserved by those maps. This is one of the fundamental principles of mathematics. For example, when studying groups we are not concerned with what the labels of the elements are, but the the relations between them according to the group operation. With manifolds, we do not focus on embeddings in space but instead maps (e.g. diffeomorphisms) which preserve various fundamental properties.

In additive combinatorics, our object of study is set addition. So we must understand maps between sets which preserve, or at least partially preserve, additive structure. Such maps are referred to as ***Freiman homomorphisms***.

Definition 7.32. Let A, B be subsets in (possibly different) abelian groups. We say that $\phi\colon A \to B$ is a ***Freiman s-homomorphism*** (or a ***Freiman homomorphism of order s***), if

Freiman s-homomorphism partially remembers additive structure, up to s-fold sums.

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

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

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

Definition 7.33. If $\phi\colon A \to B$ is a bijection, and both ϕ and ϕ^{-1} are Freiman s-homomorphisms, then ϕ is said to be a ***Freiman s-isomorphism***.

Let us look at some examples:

Example 7.34. Every group homomorphism is a Freiman homomorphism for any order.

Example 7.35. 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. And if ϕ_1 and ϕ_2 are both Freiman s-isomorphisms, then their composition $\phi_1 \circ \phi_2$ is a Freiman s-isomorphism.

Example 7.36. Suppose S has no additive structure (e.g. $\{1, 10, 10^2, 10^3\}$). Then an arbitrary map $\phi\colon S \to \mathbb{Z}$ is a Freiman 2-homomorphism.

Example 7.37. Suppose S_1 and S_2 are both sets without additive structure. Then any bijection $\phi\colon S_1 \to S_2$ is a Freiman 2-isomorphism.

Note that Freiman isomorphism and group homomorphisms have subtle differences!

Example 7.38. The natural embedding $\phi\colon \{0,1\}^n \to (\mathbb{Z}/2\mathbb{Z})^n$ is a group homomorphism, so it is a Freiman homomorphism of every order. It is also a bijection. But its inverse map does not preserve some additive relations, thus it is not a Freiman 2-isomorphism!

In general, the mod N map $\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ is a group homomorphism, but not a Freiman isomorphism. This holds even if we restrict the map to $[N]$ rather than $\mathbb{Z}$. However, we can find Freiman isomorphisms by restricting to subsets of small diameter.

Proposition 7.39. *If $A \subset \mathbb{Z}$ has diameter smaller than N/s, then (mod N) maps A Freiman s-isomorphically to its image.*

If A is restricted to a small interval, then it does not have its additive relations wrap around mod N. Thus it becomes a Freiman isomorphism.

Proof. If $a_1, \dots, a_s, a'_1, \dots, a'_s \in A$ are such that

$$\sum_{i=1}^{s} a_i - \sum_{i=1}^{s} a'_i \equiv 0 \pmod{N},$$

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. □

7.5 *Modeling lemma*

When trying to prove Freiman's theorem over the integers, our main difficulty is that a subset A with small doubling might be spread out over $\mathbb{Z}$. But we can use a Freiman isomorphism to model A inside a smaller space, preserving relative additive stucture. In this smaller space, we have better tools such as Fourier Analysis. To set up this model, we prove a modeling lemma. To warm up, let us prove this in the finite field model.

Theorem 7.40 (Modeling lemma in finite field model)**.** *Let $A \subset \mathbb{F}_2^n$ with $2^m \geq |sA - sA|$ for some positive integer m. Then A is Freiman s-isomorphic to some subset of $\mathbb{F}_2^m$.*

$\mathbb{F}_2^n$ could potentially be very large. But we can model the additive structure of A entirely within $\mathbb{F}_2^m$, which has bounded size.

Remark 7.41. If $|A + A| \leq K|A|$, then by the Plünnecke–Ruzsa inequality (Theorem 7.22) we have $|sA - sA| \leq K^{2s}|A|$, so the hypothesis if the theorem would be satisfied for some $m = O(s \log K + \log |A|)$.

Proof. The following are equivalent for linear maps $\phi\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$:

1. ϕ is Freiman s-isomorphic 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 the uniform random linear map. Each $x \in sA - sA$ violates condition (3) with probability 2^{-m}. Thus if $2^m \geq |sA - sA|$, then the probability that condition (3) is satisfied is nonzero. This implies the existence of a Freiman s-isomorphism. □

This proof does not work directly in $\mathbb{Z}$ as you cannot just choose a random linear maps. In fact, the model lemma over $\mathbb{Z}$ shows that, in fact, if $A \subset \mathbb{Z}$ has small doubling, then a large fraction of A can be modeled inside a small cyclic group whose size is comparable to $|A|$. It turns out to be enough to model a large subset of A, and we will use the Ruzsa covering lemma later on to recover the structure of the entire set A.

Theorem 7.42 (Ruzsa modeling lemma)**.** *Let $A \subset \mathbb{Z}$, $s \geq 2$, and N be a positive integer such that $N \geq |sA - sA|$. Then there exists $A' \subset A$ with $|A'| \geq |A|/s$ such that A' is Freiman s-isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$.*

Ruzsa (1992)

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

We just want to take q large enough to not have to worry about any pesky details. Its actual size does not really matter.

$$\phi\colon \mathbb{Z} \to \mathbb{Z}/q\mathbb{Z} \xrightarrow{\times\lambda} \mathbb{Z}/q\mathbb{Z} \to [q].$$

Any unspecified maps refer to the natural embeddings to and from mod q. The first two maps are group homomorphisms, so they must be Freiman s-homomorphisms. The last map is not a group homomorphism over the whole domain, but it is over small intervals. In fact, by the pigeonhole principle, for all λ there exists an interval $I_\lambda \subset [q]$ of length less than q/s such that $A_\lambda = \{a \in A\colon \phi(a) \in I_\lambda\}$ has more than $|A|/s$ elements. Thus ϕ, when restricted to A_λ, is a Freiman s-homomorphism.

Now, we take this map and send it to a cyclic group, while preserving Freiman s-homomorphism. We define,

$$\psi\colon \mathbb{Z} \xrightarrow{\phi} [q] \to \mathbb{Z}/N\mathbb{Z}.$$

Claim 7.43. If ψ does not map A_λ Freiman s-isomorphically to its image, then there exists some nonzero $d = d_\lambda \in sA - sA$ such that $\phi(d) \equiv 0 \pmod N$.

Proof. Suppose ψ does not map A_λ Freiman isomorphically to its image. Thus, there exists $a_1, \dots, a_s, a'_1, \dots, a'_s \in A_\lambda$ such that

$$a_1 + \cdots + a_s \neq a'_1 + \cdots + a'_s,$$

but

$$\phi(a_1) + \cdots + \phi(a_s) \equiv \phi(a'_1) + \cdots + \phi(a'_s) \pmod N.$$

Since $\phi(A_\lambda) \subset I_\lambda$, which is an interval of length less than q/s, we have,

$$\left|\phi(a_1) + \cdots + \phi(a_s) - \phi(a'_1) - \cdots - \phi(a'_s)\right| \in (-q, q).$$

By swapping $(a_1, \dots, a_s)$ with $(a'_1, \dots, a'_s)$ if necessary, we assume that the LHS above is nonnegative, i.e., lies in the interval $[0, q)$.

We set $d = a_1 + \cdots + a_s - a'_1 - \cdots - a'_s$. Thus $d \in (sA - sA) \setminus \{0\}$. Now, as all the functions composed to form ϕ are group homomorphisms mod q, we have

$$\phi(d) \equiv \phi(a_1) + \cdots + \phi(a_s) - \phi(a'_1) - \cdots - \phi(a'_s) \pmod q,$$

and $\phi(d)$ lies in $[0, q)$ by the definition of ϕ. Thus the two expressions above are equal. As a result,

$$\phi(d) \equiv 0 \pmod N.$$

□

Now, for each $d \in (sA - sA)\setminus\{0\}$, the number of λ such that $\phi(d) \equiv 0 \pmod N$ equals the number of elements of $[q-1]$ divisible by N. This number is at most $(q-1)/N$.

Note that we are fixing d, but ϕ is determined by λ.

Therefore, the total number of λ such that there exists $d \in (sA - sA)\setminus\{0\}$ with $\phi(d) \equiv 0 \pmod N$ is at most $(|sA - sA| - 1)(q - 1)/N < q - 1$. So there exists some λ such that ψ maps A_λ Freiman s-isomorphically onto its image. Taking $A' = A_\lambda$, our proof is complete. □

By summing up everything we know so far, we establish a result that will help us in the proof of Freiman's theorem.

Corollary 7.44. *If $A \subset \mathbb{Z}$ with $|A + A| \leq K|A|$, then there exists a prime $N \leq 2K^{16}|A|$ and some $A' \subset A$ with $|A'| \geq |A|/8$ such that A' is Freiman 8-isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$.*

Proof. By the Plünnecke–Ruzsa inequality (Theorem 7.22), $|8A - 8A| \leq K^{16}|A|$. We choose a prime $K^{16} \leq N < 2K^{16}$ by Bertrand's postulate. Then we apply the modeling lemma with $s = 8$ and $N \geq |8A - 8A|$. Thus there exists a subset $A' \subset A$ with $|A'| \geq |A|/8$ which is Freiman 8-isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$. □

7.6 Bogolyubov's lemma

In the Ruzsa modeling lemma (Theorem 7.42) we proved that for any set A of integers with small doubling constant, a large fraction of A is Freiman isomorphic to a subset of $\mathbb{Z}/N\mathbb{Z}$ with N not much larger than the size of A. To prove Freiman's Theorem, we need to prove that we can cover A with GAPs. This leads to the natural question of how to cover large subsets of $\mathbb{Z}/N\mathbb{Z}$ with GAPs. In this section, we first show how to find additive structure within subsets of $\mathbb{Z}/N\mathbb{Z}$. Later on, we will show how to use this additive structure to obtain a covering. It will be easier to first consider the analogous question in the finite field $\mathbb{F}_2^n$. Note a subset of $\mathbb{F}_2^n$ of size $\alpha 2^n$ does not necessarily contain any large structure such as a subspace. However, the key intuition for this section is the following: given a set A, the sumset $A + A$ smooths out the structure of A. With this intuition, we arrive at the following natural question:

Question 7.45. Suppose $A \subset \mathbb{F}_2^n$ 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.

Example 7.46. Let A_n 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

$$|A_n| \sim k2^n$$

where $k > 0$ is a constant depending only on c. However, $A_n + A_n$ 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 dimension $> n - c\sqrt{n}$. The proof of this claim is left as an exercise to the reader. (The same fact was also used in the proof of (6.3).)

Returning to the key intuition that the sumset $A + A$ smooths out the structure of A, it is natural to consider sums of more copies of A. It turns out that if we replace $A + A$ with $2A - 2A$ in Question 7.45 then the answer is affirmative.

Theorem 7.47 (Bogolyubov's lemma). *If $A \subset \mathbb{F}_2^n$ and $|A| = \alpha 2^n$ where α is a constant independent of n then $2A - 2A$ contains a subspace of codimension at most $1/\alpha^2$.* Bogolyubov (1939)

Proof. Let $f = 1_A * 1_A * 1_{-A} * 1_{-A}$. Note that f is supported on $2A - 2A$. Next, by the convolution property in Proposition 6.4,

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

By Fourier inversion, we have

$$f(x) = \sum_{r \in \mathbb{F}_2^n} \widehat{f}(r)(-1)^{r\cdot x} = \sum_{r \in \mathbb{F}_2^n} |\widehat{1}_A(r)|^4 (-1)^{r\cdot x}.$$

Note that it suffices to find a subspace where f is positive since $f(x) > 0$ would imply $x \in 2A - 2A$. We will choose this subspace by looking at the size of the Fourier coefficients. Let

$$R = \{r \in \mathbb{F}_2^n \backslash \{0\} : |\widehat{1}_A(r)| > \alpha^{3/2}\}.$$

By Parseval's identity, $|R| < 1/\alpha^2$. Next note

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

If x is in $R^\perp$, the orthogonal complement of R, then

$$\begin{aligned} f(x) &= \sum_{r \in \mathbb{F}_2^n} |\widehat{1}_A(r)|^4 (-1)^{r\cdot x} \\ &\geq |\widehat{1}_A(0)|^4 + \sum_{r \in R} |\widehat{1}_A(r)|^4 (-1)^{r\cdot x} - \sum_{r \notin R \cup \{0\}} |\widehat{1}_A(r)|^4 \\ &> \alpha^4 + \sum_{r \in R} |\widehat{1}_A(r)|^4 - \alpha^4 \\ &\geq 0. \end{aligned}$$

Thus $R^\perp \subset \operatorname{supp}(f) = 2A - 2A$ and since $|R| < 1/\alpha^2$, we have found a subspace with the desired codimension contained in $2A - 2A$. □

Our goal is now to formulate an analogous result for a cyclic group $\mathbb{Z}/N\mathbb{Z}$. The first step is to formulate an analog of subspaces for the cyclic group $\mathbb{Z}/N\mathbb{Z}$. Note we encountered a similar issue in transferring the proof of Roth's theorem from finite fields to the integers (see Theorem 6.2 and Theorem 6.12). It turns out that the correct analog is given by a Bohr set. Recall the definition of a Bohr set:

Definition 7.48. Suppose $R \subset \mathbb{Z}/N\mathbb{Z}$. Define

$$\operatorname{Bohr}(R, \epsilon) = \{x \in \mathbb{Z}/N\mathbb{Z} : \left\| \frac{rx}{N} \right\| \leq \epsilon, \text{for all } r \in R\}$$

where $\|\cdot\|$ denotes the distance to the nearest integer. We call $|R|$ the dimension of the Bohr set and ϵ the width.

It turns out that Bogolyubov's lemma holds over $\mathbb{Z}/N\mathbb{Z}$ after replacing subspaces by Bohr sets of the appropriate dimension. Note that the dimension of a Bohr set of $\mathbb{Z}/N\mathbb{Z}$ corresponds to the codimension of a subspace of $\mathbb{F}_2^n$.

Theorem 7.49 (Bogolyubov's lemma in $\mathbb{Z}/N\mathbb{Z}$)**.** *If $A \subset \mathbb{Z}/N\mathbb{Z}$ and $|A| = \alpha N$ then $2A - 2A$ contains some Bohr set $\operatorname{Bohr}(R, 1/4)$ with $|R| < 1/\alpha^2$.* Bogolyubov (1939)

Recall the definition of the Fourier Transform over $\mathbb{Z}/N\mathbb{Z}$.

Definition 7.50. Fourier transform of $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ is the function $\hat{f} : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ given by

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

where $\omega = e^{(2\pi i)/N}$.

We leave it as an exercise to the reader to verify the Fourier inversion formula, Parseval's identity, Plancherel's identity and the other basic properties of the Fourier transform. Now we will prove Theorem 7.49. It follows the same outline as the proof of Theorem 7.47 except for a few minor details.

Proof of Theorem 7.49. Let $f = 1_A * 1_A * 1_{-A} * 1_{-A}$. Note that f is supported on $2A - 2A$. Next, by the convolution property in Proposition Proposition 6.4,

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

By Fourier inversion, we have

$$f(x) = \sum_{r \in \mathbb{Z}/N\mathbb{Z}} \widehat{f}(r)\omega^{rx} = \sum_{r \in \mathbb{Z}/N\mathbb{Z}} |\widehat{1}_A(r)|^4 \cos\left(\frac{2\pi rx}{N}\right).$$

Let

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

By Parseval's identity, $|R| < 1/\alpha^2$. Next note

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

Now note the condition $x \in \mathrm{Bohr}(R, 1/4)$ is precisely equivalent to

$$\cos\left(\frac{2\pi rx}{N}\right) > 0 \text{ for all } r \in R.$$

For $x \in \mathrm{Bohr}(R, 1/4)$, we have

$$\begin{aligned} f(x) &= \sum_{r \in \mathbb{Z}/N\mathbb{Z}} |\widehat{1}_A(r)|^4 \cos\left(\frac{2\pi rx}{N}\right) \\ &\geq |\widehat{1}_A(0)|^4 + \sum_{r \notin R \cup \{0\}} |\widehat{1}_A(r)|^4 \cos\left(\frac{2\pi rx}{N}\right) \\ &> 0. \end{aligned}$$

□

We have now shown that for a set A that contains a large fraction of $\mathbb{Z}/N\mathbb{Z}$, the set $2A - 2A$ must contain a Bohr set of dimension less than $1/\alpha^2$. In the next section we will analyze additive structure within Bohr sets. In particular, we will show that Bohr sets of low dimension contain large GAPs.

7.7 *Geometry of numbers*

Before we can prove the main result of this section, we first introduce some machinery from the geometry of numbers. The geometry of numbers involves the study of lattices and convex bodies and has important applications in number theory.

Definition 7.51. A ***lattice*** in $\mathbb{R}^d$ is a set given by $\Lambda = \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_d$ where $v_1, \ldots, v_d \in \mathbb{R}^d$ are linearly independent vectors.

Definition 7.52. The ***determinant*** $\det(\Lambda)$ of a lattice $\Lambda = \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_d$ is the absolute value of the determinant of a matrix with $v_1, \ldots, v_d$ as columns.

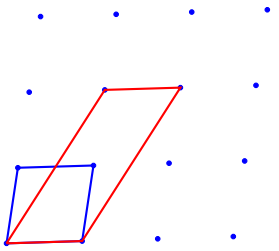

Figure 7.5: A lattice in $\mathbb{R}^2$, the blue shape is a fundamental parallelepiped while the red is not.

Remark 7.53. Note the determinant of a lattice is also equal to the volume of the fundamental parallelepiped.

Example 7.54. $\mathbb{Z} + \mathbb{Z}\omega$ where $\omega = e^{(2\pi i)/3}$ is a lattice. Its determinant is $\sqrt{3}/2$.

Example 7.55. $\mathbb{Z} + \mathbb{Z}\sqrt{2} \subset \mathbb{R}$ is **not** a lattice because 1 and $\sqrt{2}$ are not linearly independent.

We now introduce the important concept of successive minima of a convex body K with respect to a lattice Λ.

Definition 7.56. Given a centrally symmetric convex body $K \subset \mathbb{R}^d$ (by centrally symmetric we mean $x \in K$ if and only if $-x \in K$), define the ***i^{th} successive minimum*** of K with respect to a lattice Λ as

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

for $1 \leq i \leq d$. Equivalently, λ_i is the minimum λ 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 $\mathbf{b}_i \in \lambda_i K$ for each $i = 1, \ldots, d$. (Note that there may be more than one possible directional basis.)

Example 7.57. 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

$$\Lambda = \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_7 \oplus \mathbb{Z}v.$$

Let K be the unit ball in $\mathbb{R}^8$. Note that the directional basis of K with respect to Λ is $e_1, \ldots, e_8$. This example shows that the directional basis of a convex body K is not necessarily a $\mathbb{Z}$-basis of Λ.

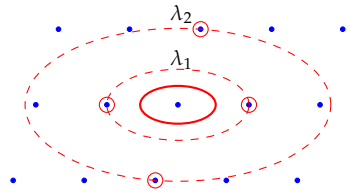

Figure 7.6: A diagram showing the successive minima of the body outlined by the solid red line with respect to the lattice of blue points.

Minkowski's second theorem gives us an inequality to control the product of the successive minima in terms of the volume of K and the determinant of the lattice Λ.

Theorem 7.58 (Minkowski's second theorem)**.** *Let $\Lambda \in \mathbb{R}^d$ be a lattice and K a centrally symmetric body. Let $\lambda_1 \leq \cdots \leq \lambda_d$ be the successive minima of K with respect to Λ. Then*

Minkowski (1896)

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

Example 7.59. 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 $\mathbb{Z}^d$.

The proof of Minkowski's second theorem is omitted. We will now use Minkowski's second theorem to prove that a Bohr set of low dimension contains a large GAP.

Theorem 7.60. *Let N be a prime. Every Bohr set of dimension d and width $\epsilon \in (0,1)$ in $\mathbb{Z}/N\mathbb{Z}$ contains a proper GAP with dimension at most d and size at least $(\epsilon/d)^d N$.*

Proof. Let $R = \{r_1, \dots, r_d\}$. Let

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

Let $\Lambda \subset \mathbb{R}^d$ 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 $K = [-\epsilon, \epsilon]^d$. 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

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

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

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

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

$$\|l_1 \mathbf{b}_1 + \cdots + l_d \mathbf{b}_d\|_\infty \leq \epsilon. \tag{7.1}$$

Each $\mathbf{b}_j$ is equal to $x_j v$ plus a vector with integer coordinates for some $0 \leq x_j < N$. The bound for the i^{th} coordinate in (7.1) implies

$$\left\| \frac{(l_1 x_1 + \cdots + l_d x_d) r_i}{N} \right\|_{\mathbb{R}\setminus\mathbb{Z}} \leq \epsilon \text{ for all } i.$$

Thus, the GAP

$$\{l_1 x_1 + \cdots + l_d x_d : 0 \leq l_i < L_i \text{ for all } i\}$$

is contained in $\text{Bohr}(R, \epsilon)$. It remains to show that this GAP is large and that it is proper. First we show that it is large. Using Minkowski's second theorem, its size is

$$\begin{aligned} L_1 \cdots L_k &\geq \frac{1}{\lambda_1 \cdots \lambda_d \cdot d^d} \\ &\geq \frac{\text{vol}(K)}{2^d \det(\Lambda) d^d} \\ &= \frac{(2\epsilon)^d}{2^d \frac{1}{N} d^d} \\ &= \left(\frac{\epsilon}{d}\right)^d N. \end{aligned}$$

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

$$l_1 x_1 + \cdots + l_d x_d \equiv l_1' x_1 + \cdots + l_d' x_d \pmod{N},$$

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

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

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

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

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

7.8 *Proof of Freiman's theorem*

So far in this chapter, we have demonstrated a number of useful methods and theorems in additive combinatorics on our quest to prove Freiman's theorem (Theorem 7.11). Now, we finally put these tools together to form a complete proof.

The proof method will be as follows. Starting with a set A with small doubling constant, we first map A to a subset, B, of $\mathbb{Z}/N\mathbb{Z}$ using the corollary of the Ruzsa modeling lemma (Theorem 7.42). We then find a large GAP within $2B - 2B$ using Bogolyubov's lemma (Theorem 7.47) and results on the geometry of numbers. This in turn gives us a large GAP in $2A - 2A$. Finally, we apply the Ruzsa covering lemma (Theorem 7.28) to create a GAP that contains A from this GAP contained in $2A - 2A$. Recall the statement of Freiman's theorem (Theorem 7.11):

> If $A \subset \mathbb{Z}$ is a finite set and $|A + A| \leq K|A|$, then A is contained in a GAP of dimension at most $d(K)$ and size at most $f(K)|A|$.

Proof. Because $|A + A| \leq K|A|$, by the corollary to Ruzsa modeling lemma (Corollary 7.44), there exists a prime $N \leq 2K^{16}|A|$ and some $A' \subset A$ with $|A'| \geq |A|/8$ such that A' is Freiman 8-isomorphic to a subset B of $\mathbb{Z}/N\mathbb{Z}$.

Applying Bogolyubov's lemma (Theorem 7.47) on B with

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

gives that $2B - 2B$ contains some Bohr set, $\text{Bohr}(R, 1/4)$, where $|R| < 256K^{32}$. Thus, by Theorem 7.60, $2B - 2B$ contains a proper GAP with dimension $d < 256K^{32}$ and size at least $(4d)^{-d}N$.

As B is Freiman 8-isomorphic to A', we have $2B - 2B$ is Freiman 2-isomorphic to $2A' - 2A'$. This follows from the definition of Freiman s-isomorphism and by noting that every element in $2B - 2B$ is the sum and difference of four elements in B with a similar statement for $2A' - 2A'$. Note that arithmetic progressions are preserved by Freiman 2-isomorphisms as the difference between any two elements in $2B - 2B$ is preserved. Hence, the proper GAP in $2B - 2B$ is mapped to a proper GAP, Q, in $2A' - 2A'$ with the same dimension and size.

Next we will use the Ruzsa covering lemma to cover the entire set A with translates of Q. Because $Q \subset 2A - 2A$, we have $Q + A \subset 3A - 2A$. By the Plünnecke-Ruzsa inequality (Theorem 7.22), we have

$$|Q + A| \leq |3A - 2A| \leq K^5|A|.$$

As $A' \subset \mathbb{Z}/N\mathbb{Z}$, we have $N \geq |A'| \geq |A|/8$. Because $|Q| \geq (4d)^{-d}N$, we have $K^5|A| \leq K'|Q|$ where $K' = 8(4d)^d K^5 = e^{K^{O(1)}}$. In particular, the above inequality becomes $|Q + A| \leq K'|Q|$. Hence, by the Ruzsa covering lemma (Theorem 7.42), there exists a subset X of A with $|X| \leq K'$ such that $A \subset X + Q - Q$.

All that remains is to show that $X + Q - Q$ is contained in a GAP with the desired bounds on dimension and size. Note that X is trivially contained in a GAP of dimension $|X|$ with length 2 in every direction. Furthermore, because every element in $Q - Q$ lies on some arithmetic progression contained in Q translated to the origin, we have the dimension of $Q - Q$ is d. Hence, by the bounds outlined above, $X + Q - Q$ is contained in a GAP P with dimension

$$\dim(P) \leq |X| + d \leq K' + d = 8(4d)^d K^5 + d = e^{K^{O(1)}}.$$

Because Q is a proper GAP with dimension d and the doubling constant of an arithmetic progression is 2, we have that $Q - Q$ has size at most $2^d|Q|$. The GAP containing X has size $2^{|X|}$. Hence, applying the Plünnecke-Ruzsa inequality, we have that the size of P is

$$\text{size}(P) \leq 2^{|X|}2^d|Q| \leq 2^{K'+d}|2A - 2A| \leq 2^{K'+d}K^4|A| = e^{e^{K^{O(1)}}}|A|.$$

Taking $d(K) = e^{K^{O(1)}}$ and $f(K) = e^{e^{K^{O(1)}}}$ completes the proof of Freiman's theorem. □

Remark 7.61. By considering $A = \{1, 10, 10^2, 10^3, \dots, 10^{|A|-1}\}$ we see that Freiman's theorem is false for $d(K) < \Theta(K)$ and $f(K) < 2^{\Theta(K)}$. It is also conjectured that Freiman's holds for $d(K) = \Theta(K)$ and $f(K) = 2^{\Theta(K)}$.

While the bounds given in the above proof of Freiman's theorems are quite far off this (exponential rather than linear), Chang showed that Ruzsa's arguments can be made to give polynomial bounds ($d(K) = K^{O(1)}$ and $f(k) = \exp(K^{O(1)})$). When we apply Ruzsa's covering lemma, we are somewhat wasteful. Rather than cover A all at once, a better method is to cover A bit by bit. In particular starting with Q we cover parts of A with $Q - Q$. We then repeat the proof on what remains of A to find Q_1 with smaller dimension. We then cover the rest of A with $Q_1 - Q_1$. This method significantly reduces the amount we lose in this step and gives the desired polynomial bounds.

Chang (2002)

As noted before, the best known bound (Theorem 7.15) is given by $d(K) = K(\log K)^{O(1)}$ and $f(K) = e^{K(\log K)^{O(1)}}$, whose proof is substantially more involved.

7.9 *Freiman's theorem for general abelian groups*

We have proved Freiman's theorem for finite fields and for integers, so one might wonder whether Freiman's theorem holds for general abelian groups. This is indeed the case, but first we must understand what such a Freiman's theorem might state.

For $\mathbb{F}_p^n$ for fixed primes p, Freiman's theorem gives that any set with small doubling constant exists in a not too much larger subgroup, while for integers, Freiman's theorem gives the same but for a not too much larger GAP. Because finitely generated abelian groups can always be represented as the direct sum of cyclic groups of prime power orders and copies of $\mathbb{Z}$, to find a generalization of GAPs and subgroups, one might try taking the direct sum of these two types of structures.

Definition 7.62. Define a ***coset progression*** as the direct sum $P + H$ where P is a proper GAP and H is a subgroup. The ***dimension*** of a coset progression is defined as the dimension of P and the ***size*** of a coset progression is defined as the cardinality of the whole set.

By a ***direct sum*** $P + H$ we mean that if $p + h = p' + h'$ for some $p, p' \in P$ and $h, h' \in H$ then $p = p'$ and $h = h'$.

Theorem 7.63 (Freiman's theorem for general abelian groups)**.** *If A is a subset of a arbitrary abelian group and $|A + A| \leq K|A|$, then A is*

Green and Ruzsa (2007)

contained in a coset progression of dimension at most $d(K)$ and size at most $f(k)|A|$, where $d(K)$ and $f(K)$ are constants depending only on K.

Remark 7.64. The proof of this theorem follows a similar method to the given proof of Freiman's theorem but with some modifications to the Ruzsa modeling lemma. The best known bounds for are again given by Sanders and are $d(K) = K(\log K)^{O(1)}$ and $f(K) = e^{K(\log K)^{O(1)}}$. It should be noted that these functions depend only on K, so they remain the same regardless of what abelian group A is a subset of.

Sanders (2013)

7.10 The Freiman problem in nonabelian groups

We may ask a similar question for nonabelian groups: what is the structure of subsets of a nonabelian group that have small doubling? Subgroups still have small doubling just as in the abelian case. Also, we can take a GAP formed by any set of commuting elements. However, it turns out that there are other examples of sets of small doubling, which are not directly derived from either of these examples from abelian groups.

Example 7.65. The ***discrete Heisenberg group*** $H_3(\mathbb{Z})$ is the set of upper triangular matrices with integer entries and only ones on the main diagonal. Multiplication in this group is as follows:

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+x & c+z+ay \\ 0 & 1 & b+y \\ 0 & 0 & 1 \end{pmatrix}.$$

Now, let S be the following set of generators of H.

$$S = \left\{ \begin{pmatrix} 1 & \pm 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \pm 1 \\ 0 & 0 & 1 \end{pmatrix} \right\}.$$

Consider the set S^r, which is taken by all products of r sequences of elements from S. By the multiplication rule, the elements of S^r are all of the form

$$\begin{pmatrix} 1 & O(r) & O(r^2) \\ 0 & 1 & O(r) \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, $|S^r| \leq O(r^4)$, since there are at most $O(r^4)$ possibilities for such a matrix. It can also be shown that $|S^r| = \Omega(r^4)$, and thus $|S^r| = \Theta(r^4)$. Thus, the doubling of S^r is $|S^{2r}|/|S^r| \approx 16$, so S^r has bounded doubling.

It turns out that this is an example of a more general type of construction in a group which is "almost abelian." This is captured by the notion of a nilpotent group.

Definition 7.66. A ***nilpotent group*** G is one whose lower central series terminates. In other words,

$$[\dots[[G,G],G]\dots,G] = \{e\},$$

for some finite number of repetitions. (The commutator subgroup $[H,K]$ is defined as $\{hkh^{-1}k^{-1} : h \in H, k \in K\}$.)

All nilpotent groups have polynomial growth similarly to Example 7.65, defined in general as follows.

Definition 7.67. Let G be a finitely generated group generated by a set S. The group G is said to have ***polynomial growth*** if there are constants $C, d > 0$ such that $|S^r| \leq Cr^d$ for all r. (This definition does not depend on S since for any other set of generators S', there exists r_0 such that $S' \subset S^{r_0}$.)

Gromov's theorem is a deep result in geometric group theory that provides a complete characterization of groups of polynomial growth.

Theorem 7.68 (Gromov's theorem)**.** *A finitely generated group has polynomial growth if and only if it is virtually nilpotent, i.e., has a nilpotent subgroup of finite index.*

Gromov (1981)

The techniques used by Gromov relate to Hilbert's fifth problem, which concerns characterization of Lie groups. A more elementary proof of Gromov's theorem was later given by Kleiner in 2010.

Kleiner (2010)

Now, we have a construction of a set with small doubling in any virtually nilpotent group G: the "nilpotent ball" S^r, where S generates G. It is then natural to ask the following question.

Question 7.69. Must every set of small doubling (or equivalently, sets known as *approximate groups*) behave like some combination of subgroups and nilpotent balls?

Lots of work has been done on this problem. In 2012, Hrushovski, using model theoretic techniques, showed a weak version of Freiman's theorem for nonabelian groups. Later, Breuillard, Green, and Tao, building on Hrushovski's methods, proved a structure theorem for approximate groups, generalizing Freiman's theorem to nonabelian groups. However, these methods provide no explicit bounds due to their use of ultrafilters.

Hrushovski (2012)

Breuillard, Green, and Tao (2012)

7.11 *Polynomial Freiman–Ruzsa conjecture*

In $\mathbb{F}_2^n$, if A is an independent set of size n, its doubling constant is $K = |A+A|/|A| \approx n/2$, and the size of any subgroup that contains A must be at least $2^{\Theta(K)}|A|$.

Another example, extending the previous one, is to let A be a subset of $\mathbb{F}^{m+n}$ defined by $A = \mathbb{F}_2^m \times \{e_1, \dots, e_n\}$ (where $e_1, \dots, e_n$ are generators of $\mathbb{F}_2^n$). This construction has the same bounds as the previous one, but with arbitrarily large $|A|$. This forms an example showing that the bound in the abelian group version of Freiman's theorem cannot be better than exponential.

However, note that in this example, A must contain the very large (affine) subspace $\mathbb{F}_2^m \times \{e_1\}$, which has size comparable to A. We may thus ask whether we could get better bounds in Freiman's theorem if we only needed to cover a large subset of A. In this vein, the Polynomial Freiman–Ruzsa conjecture in $\mathbb{F}_2^n$ asks the following. Green (2004)

Conjecture 7.70 (Polynomial Freiman–Ruzsa conjecture in $\mathbb{F}_2^n$)**.** *If $A \subset \mathbb{F}_2^n$, and $|A+A| \le K|A|$, then there exists an affine subspace $V \subseteq \mathbb{F}_2^n$ with $|V| \le |A|$ such that $|V \cap A| \ge K^{-O(1)}|A|$.*

This conjecture has several equivalent forms. For example, the following three are equivalent to Conjecture 7.70:

Conjecture 7.71. *If $A \subset \mathbb{F}_2^n$, 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.*

Proof of equivalence of Conjecture 7.70 and Conjecture 7.71. Clearly Conjecture 7.71 implies Conjecture 7.70.

Now suppose the statement of Conjecture 7.70 is true, and suppose we have $A \subset \mathbb{F}_2^n$ satisfying $|A+A| \le K|A|$. Then by Conjecture 7.70, there exists some affine subspace V with size at most $|A|$ such that $|V \cap A| \ge K^{-O(1)}|A|$ Applying the Ruzsa covering lemma (Theorem 7.28) with $X = A, B = V \cap A$ gives a set X of size $K^{O(1)}$ such that $A \subseteq V - V + X$. The conclusion of Conjecture 7.71 follows immediately, where the cosets are the shifts of the vector space $V - V$ by each of the elements of X. □

Conjecture 7.72. *If $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ satisfies*

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

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

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

(In this version, 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.)

Conjecture 7.73. *If $f : \mathbb{F}_2^n \to \mathbb{C}$ with $\|f\|_\infty \leq 1$ and $\|f\|_{U_3} \geq \delta$ (where $\|f\|_{U_3}$ is the Gowers U_3 norm, and relates to 4-AP counts), then there exists a quadratic polynomial $q(x_1, \dots, x_n)$ over $\mathbb{F}_2$ such that*

$$|\mathbb{E}_{x \in \mathbb{F}_2^n}[f(x)(-1)^{q(x)}]| \geq \delta^{O(1)}.$$

It turns out that these versions of the conjectures are all equivalent up to polynomial changes in the bounds (or equivalently, linear relations between the $O(1)$ terms). The best bound to date is due to Sanders and achieves a quasipolynomial bound of $e^{(\log K)^{O(1)}}$. The polynomial Freiman–Ruzsa conjecture would be implied by the following strengthening of Bogolyubov's lemma: Sanders (2012)

Conjecture 7.74 (Polynomial Bogolyubov-Ruzsa conjecture in $\mathbb{F}_2^n$)**.** *If $A \subset \mathbb{F}_2^n$ with $|A| = \alpha 2^n$, then $2A - 2A$ contains a subspace of codimension $O(\log(1/\alpha))$.*

The standard form of Bogolyubov's lemma (Theorem 7.47) shows a bound of $O(\alpha^{-2})$. The best result on this conjecture is also due to Sanders, who obtained a quasipolynomial bound of $(\log(1/\alpha))^{O(1)}$. Sanders (2012)

One may similarly make a version of the polynomial Freiman–Ruzsa conjecture in $\mathbb{Z}$ instead of $\mathbb{F}_2^n$. First, we must define a centered convex progression, the analog of a subspace.

Definition 7.75. A ***centered convex progression*** is a set of the form

$$P = \{x_0 + \ell_1 x_1 + \cdots + \ell_d x_d : (\ell_1, \dots, \ell_d) \in \mathbb{Z}^d \cap B\},$$

where B is some convex centrally symmetric body in $\mathbb{R}^d$. In other words, it is a shift of the image of $\mathbb{Z}^d \cap B$ under some homomorphism $\mathbb{Z}^d \to \mathbb{Z}$. Its ***dimension*** is d and its ***size*** is $|\mathbb{Z}^d \cap B|$.

Then, the polynomial Freiman–Ruzsa conjecture in $\mathbb{Z}$ states the following.

Conjecture 7.76 (Polynomial Freiman–Ruzsa conjecture in $\mathbb{Z}$)**.** *If $A \subset \mathbb{Z}$ with $|A + A| \leq K|A|$, then there exists a centered convex progression of dimension $O(\log K)$ and size at most $|A|$ whose intersection with A has size at least $K^{-O(1)}|A|$.*

More generally, the Polynomial Freiman–Ruzsa conjecture in abelian groups uses *centered convex coset progressions*, which are defined as a direct sum $P + H$, where P is the image of some $\mathbb{Z}^d \cap B$ under a homomorphism from $\mathbb{Z}^d$ to the group, and H is some coset of a subgroup.

The best bound on this conjecture (in both the $\mathbb{Z}$ and the abelian group cases) is once again quasipolynomial due to Sanders, who derived it from a quasipolynomial bound for the polynomial Bogolyubov-Ruzsa conjecture: Sanders (2012)

Conjecture 7.77 (Polynomial Bogolyubov-Ruzsa conjecture in $\mathbb{Z}$)**.** *If $A \subset \mathbb{Z}/N\mathbb{Z}$ with N prime, then $2A - 2A$ contains a proper centered convex progression of dimension $O(\log(1/\alpha))$ and size at least $\alpha^{O(1)}N$.*

Again, the version for general abelian groups can be obtained by instead using proper centered convex coset progressions instead.

7.12 *Additive energy and the Balog–Szémeredi–Gowers theorem*

So far, we have measured the amount of additive structure in a set using the doubling constant. Here we introduce ***additive energy***, a new measurement of additive structure in a set; where previously we were interested in sets of high doubling, we are now interested in sets with high additive energy.

Definition 7.78. Let A and B be finite subsets of an abelian group. Their ***additive energy*** is defined to be

$$E(A,B) = |\{(a_1,a_2,b_1,b_2) \in A \times A \times B \times B : a_1 + a_2 = b_1 + b_2\}|.$$

We set the additive energy of a single subset A to be $E(A) := E(A,A)$.

Remark 7.79. We can think of the additive energy as counting 4-cycles in an appropriate Cayley graph. Just as counting 4-cycles turned out to be fundamental in graph theory, we will see that additive energy is fundamental in additive combinatorics.

Definition 7.80. For two finite subsets A and B of an abelian group, define $r_{A,B}(x) := |\{(a,b) \in A \times B : x = a + b\}|$ to count the number of ways x is expressible as a sum in $A + B$.

Remark 7.81. We can compute additive energy as

$$E(A,B) = \sum_x r_{A,B}(x)^2.$$

For additive energy, we have the following analogue of Proposition 7.3.

Proposition 7.82. *If A is a finite subset of $\mathbb{Z}$ then $|A|^2 \leq E(A) \leq |A|^3$.*

Proof. The lower bound comes from the fact that all 4-tuples of the form $(a_1,a_2,a_1,a_2) \in A^4$ are counted by the additive energy $E(A)$. The upper bound is because for any triple $(a_1,a_2,a_3) \in A^3$, we have that $E(A)$ counts at most one 4-tuple with those first three coordinates, with fourth coordinate $a_1 + a_2 - a_3$. □

Remark 7.83. Proposition 7.82 is tight. The lower bound holds when A has no additive structure, while the upper bound holds asymptotically when $A = [n]$.

Thus far, we have likened sets of small doubling and large additive energy. In fact, the former implies the latter.

Proposition 7.84. *If $|A+A| \leq K|A|$ then $E(A) \geq |A|^3/K$.*

Proof. We use Remark 7.81 and the Cauchy-Schwarz inequality to show

$$\begin{aligned} E(A) &= \sum_{x \in A+A} r_{A,A}(x)^2 \geq \frac{1}{|A+A|}\left(\sum_{x \in A+A} r_{A,A}(x)\right)^2 \\ &= \frac{|A|^4}{|A+A|} \geq \frac{|A|^3}{|K|}. \end{aligned}$$

□

It is natural to ask whether the converse of Proposition 7.84 holds. In fact, a set with large additive energy may also have high doubling, as described in Example 7.85 below.

Example 7.85. Consider the set $A = [N/2] \cup \left\{-2, -4, -8, \ldots, -2^{N/2}\right\}$. Note that A is the union of a set of small doubling and a set with no additive structure. The first component forces the additive energy to be $E(A) = \Theta(N^3)$, while the second forces a large doubling $|A+A| = \Theta(N^2)$.

However, Balog and Szemerédi showed that every set with large additive energy must have a highly structured subset with small doubling, even if the set has relatively little additive structure overall. Their proof was later refined by Gowers, who proved polynomial bounds on the constants, and this is the version we will present here.

Theorem 7.86 (Balog–Szemerédi–Gowers theorem)**.** *Let A be a finite subset of an abelian group. If $E(A) \geq |A|^3/K$ then there is a subset $A' \subset A$ with $|A'| \geq K^{-O(1)}|A|$ and $|A'+A'| \leq K^{O(1)}|A'|$.*

Balog and Szemerédi (1994)
Gowers (1998)

We present a stronger version of the theorem, which considers the additive structure between two different sets.

Theorem 7.87. *Let A and B be finite subsets of the same abelian group. If $|A|, |B| \leq n$ and $E(A,B) \geq n^3/K$ then there exist subsets $A' \subset A$ and $B' \subset B$ with $|A'|, |B'| \geq K^{-O(1)}n$ and $|A'+B'| \leq K^{O(1)}n$.*

Proof that Theorem 7.87 implies Theorem 7.86. Suppose $E(A) \geq |A|^3/K$. Apply Theorem 7.87 with $B = A$ to obtain $A', B' \subset A$ with $|A'|, |B'| \geq K^{-O(1)}n$ and $|A'+B'| \leq K^{O(1)}n$. Then by Corollary 7.27, a variant of the Ruzsa triangle inequality, we have

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

□

To prove Theorem 7.87, we once again reduce from additive combinatorics to graph theory. The proof of Theorem 7.87 relies on the following graph analogue.

Definition 7.88. Let A and B be subsets of an abelian group and let G be a bipartite graph with vertex bipartition $A \cup B$. Then we define the ***restricted sumset*** $A +_G B$ to be the set of sums along edges of G:

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

Theorem 7.89. *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 $|A|, |B| \leq n$ and G has at least n^2/K edges and $|A +_G B| \leq Kn$ then there exist subsets $A' \subset A$ and $B' \subset B$ with $|A'|, |B'| \geq K^{-O(1)}n$ and $|A' + B'| \leq K^{O(1)}n$.*

Proof that Theorem 7.89 implies Theorem 7.87. Define $r_{A,B}$ as in Definition 7.80. Let $S = \{x \in A + B : r_{A,B}(x) \geq n/2K\}$ be the set of "popular sums." 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

$$\frac{n^3}{K} \leq E(A, B) = \sum_{x \in S} r_{A,B}(x)^2 + \sum_{x \notin S} r_{A,B}(x)^2. \tag{7.2}$$

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

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

and setting back into (7.2) yields

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

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

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

Hence, we can apply Theorem 7.89 to find sets $A' \subset A$ and $B' \subset B$ with the desired properties. □

The remainder of this section will focus on proving Theorem 7.89. We begin with a few lemmas.

Lemma 7.90 (Path of length 2 lemma)**.** *Fix $\delta, \epsilon > 0$. Let G be a bipartite graph with bipartition $A \cup B$ and at least $\delta|A||B|$ edges. Then there is some $U \subset A$ with $|U| \geq \delta|A|/2$ such that at least $(1 - \epsilon)$-fraction of the pairs $(x, y) \in U^2$ have at least $\epsilon\delta^2|B|/2$ neighbors common to x and y.*

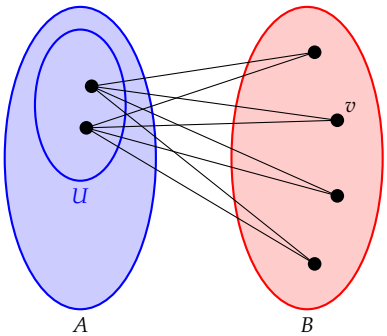

Figure 7.7: Paths of length 2 between two points in U.

Proof. We use the *dependent random choice* method from Section 2.9. Choose $v \in B$ uniformly at random, and let $U = N(v) \subset A$. We have $\mathbb{E}[|U|] \geq \delta|A|$.

We note that pairs with few common neighbors are unlikely to be contained in U. Indeed, if $x, y \in A$ share fewer than $\epsilon\delta^2|B|/2$ common neighbors then $\Pr[\{x, y\} \subset U] < \epsilon\delta^2/2$.

Say two points are *friendly* if they share at least $\epsilon\delta^2|B|/2$ common neighbors. Let X be the number of unfriendly pairs $(x, y) \in U^2$. Then

$$\mathbb{E}[X] = \sum_{\substack{(x,y)\in A^2 \\ \text{unfriendly}}} \Pr[\{x, y\} \subset U] < \frac{\epsilon\delta^2}{2}|A|^2.$$

Hence, we have

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

so there is a choice of U with $|U|^2 - X/\epsilon \geq \delta^2|A|^2/2$. For this choice of U, we have $|U|^2 \geq \delta^2|A|^2/2$, so $|U| \geq \delta|A|/2$. Moreover, we have $X \leq \epsilon|U|^2$, so at most ϵ-fraction of pairs $(x, y) \in U^2$ have fewer than $\epsilon\delta^2|B|/2$ common neighbors. □

Lemma 7.91 (Path of length 3 lemma)**.** *There are constants $c, C > 0$ such that the following holds. Fix any $\epsilon, \delta > 0$ and let G be any bipartite graph with bipartition $A \cup B$ and at least $\delta|A||B|$ edges. Then there are subsets $A' \subset A$ and $B' \subset B$ such that every pair $(a, b) \in A' \times B'$ is joined by at least $\eta|A||B|$ paths of length 3, where $\eta = c\delta^C$.*

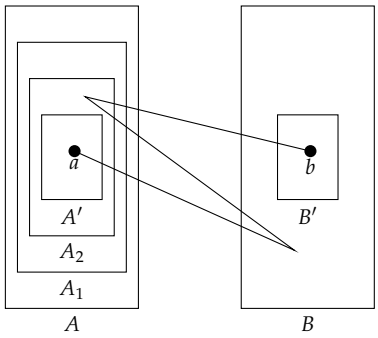

Figure 7.8: The construction for a path of length 3.

Proof. Call vertices a pair of vertices in A *friendly* if they have at least $\frac{\delta^3|B|}{20}$ common neighbors.

Define

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

Restricting A to A_1 maintains an edge density of at least δ between A_1 and B and removes fewer than $\delta|A||B|/2$ edges from G. Because we are left with at least $\delta|A||B|/2$ edges and the max degree of $a \in A_1$ is $|B|$, we have $|A_1| \geq \delta|A|/2$.

Construct $A_2 \subset A_1$ via the path of length 2 lemma (Lemma 7.90) on (A_1, B) with $\epsilon = \delta/10$. Then, $|A_2| \geq \delta|A_1|/2 \geq \delta^2|A|/4$ and at most ϵ-fraction pairs of vertices in A_2 are unfriendly.

Set

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

Restricting from (A_2, B) to (A_2, B') removes at most $\delta|A_2||B|/4$ edges. Because the minimum degree in A_2 is at least $\delta/2$, there are at least $\delta|A_2||B|/2$ edges between A_2 and B. Hence, there are at least

$\delta|A_2||B|/4$ edges between A_2 and B' and because the maximum degree of $b \in B'$ is $|A_2|$, we have $|B'| \geq \delta|B|/4$.

Define

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

Then $|A'| \geq |A_2|/2 \geq \delta^2|A|/8$.

We now fix $(a, b) \in A' \times B'$ and lower-bound the number of length-3 paths between them. Because b is adjacent to at least $\delta|A_2|/4$ vertices in A_2 and a is friendly to at least $(1 - \delta/5)|A_2|$ vertices in A_2, there are at least $\delta|A_2|/20$ vertices in A_2 both friendly to a and adjacent to b. For each such $a_1 \in A_2$, there are at least $\delta^3|B|/20$ points $b_1 \in B$ 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

$$\frac{\delta}{20}|A_2| \cdot \frac{\delta^3}{20}|B| \geq \frac{\delta}{20} \cdot \frac{\delta^2}{4}|A| \cdot \frac{\delta^3}{20}|B| = \frac{\delta^6}{20 \cdot 4 \cdot 80}|A||B|.$$

Taking η equal to the above coefficient, we note that $|A'| \geq \delta^2|A|/8 \geq \eta|A|$ and $|B'| \geq \delta|B|/4 \geq \eta|B|$. □

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

Proof of Theorem 7.89. Note that we have $|A|, |B| \geq \frac{n}{K}$. By the path of length 3 lemma (Lemma 7.91), we can find $A' \subset A$ and $B' \subset B$ of sizes $|A'|, |B'| \geq K^{-O(1)}n$ such that for every $(a, b) \in A' \times B'$, there are at least $K^{-O(1)}n^2$ paths ab_1a_1b with $(a_1, b_1) \in A \times B$. Hence, for each $(a, b) \in A' \times B'$, there are at least $K^{-O(1)}n^2$ solutions $x, y, z \in A +_G B$ to the equation $x - y + z = a + b$, as $(x, y, z) = (a + b_1, a_1 + b_1, a_1 + b)$ is a solution along each path ab_1a_1b. It follows that

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

so $|A' + B'| \leq K^{O(1)}n$. □

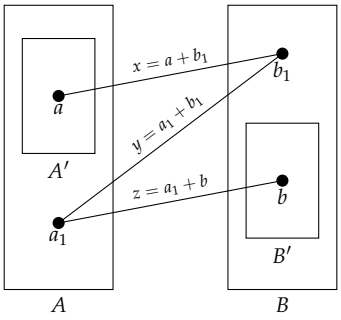

Figure 7.9: Using the path of length 3 lemma to prove the Balog–Szemerédi–Gowers theorem

18.217 Graph Theory and Additive Combinatorics

Fall 2019