Part III – Introduction to Approximate Groups (Ongoing course, rough)

Based on lectures by Dr M. Tointon
Notes taken by Bhavik Mehta

Lent 2019

Contents

1 Small doubling

Lecture 1

A subgroup $H < G$ is a non-empty set closed under products and inverses. Roughly, an 'approximate subgroup' is a subset that is only 'approximately closed' under products. (We will make this precise soon). Such sets arise naturally in a number of branches of mathematics, and as such approximate groups have had a broad range of applications. In this course, we will look in detail, for example, at applications to *polynomial growth* (fundamental in geometric group theory) and touch on construction of expander graphs (important in theoretical computer science).

To start with, we will look at a preliminary notion of approximate closure called *small doubling*. In this course, G is always a group, arbitrary unless specified otherwise.

Notation. Given $A, B \subset G$, write

$$\begin{aligned} AB &\coloneqq \{ab \mid a \in A, b \in B\} \quad \text{'Product sets'} \\ A^n &= \underbrace{A \cdot A \cdots A}_{n \text{ times}} \\ A^{-1} &= \{a^{-1} \mid a \in A\} \\ A^{-n} &= (A^{-1})^n \end{aligned}$$

When G is abelian, often switch to additive notation, e.g. $A + B$, nA, $-A$, $-nA$, called 'Sum sets'.

To say A is closed is to say $A^2 = A$. If A is finite, one way to say that A is 'approximately closed' is to say that

$$|A^2| \text{ is 'not much bigger' than } |A|.$$

This is the notion of approximate closure that arises when studying polynomial growth or expansion, for example.

To get a feel for what this should mean, let's look at the possible values of $|A^2|$. Trivially, $|A| \leq |A^2| \leq |A|^2$. Both bounds are attained. However, although the quadratic upper bound on $|A^2|$ in terms of $|A|$ is extremal, in a strict sense, it should not be seen as atypical for the size of A^2. We will see, for example, in Example Sheet 1 that if A is a set of size n chosen uniformly from $\{1, \dots, n^{100}\}$, then $\mathbb{E}(|A + A|)$ is close to $\frac{1}{2}|A|^2$ (about as large as it can be, because abelian). Therefore, we can view sets satisfying

$$|A^2| = o\big(|A|^2\big) \tag{1.1}$$

as being 'exceptional', and so condition (1.1) can already be seen as a form of 'approximate closure'. In this course, we will concentrate on the strongest form of (1.1), where $|A^2|$ is *linear* in $|A|$, in the sense that

$$|A^2| \leq K|A| \tag{1.2}$$

for some $K \geq 1$ fixed a priori.

Since such sets are 'far from random' we can expect (1.2) to impose a significant restriction on A. The main aim of this course is to work out how significant.

Definition. Given $A \subset G$, the ratio $\frac{|A^2|}{|A|}$ is called the **doubling constant** of A. If A satisfies (1.2), we'll say that A has **doubling** at most K, or simply **small doubling**.

Example (Some simple examples)**.**

- (Empty set)

- A a finite subgroup ($K = 1$)
- $|A| \leq K$
- $A \subset \mathbb{Z}$, $A = \{-n, \ldots, n\}$, $|A + A| \leq 2|A|$.

This last example is especially important as it shows the theory does not just reduce to subgroups and 'small' sets. We'll develop these examples later in the course.

One main aim will be to prove theorems along the lines of:

$$A \text{ has small doubling} \Rightarrow A \text{ has a certain structure.}$$

When K is very small, this is quite easy, as follows:

Theorem 1.1 (Freiman; proof due to Tao)**.** Let $K < \frac{3}{2}$. Suppose $A \subset G$ and $|A^2| \leq K|A|$. Then there is a subgroup $H < G$ with $|H| = |A^2|(\leq K|A|)$ such that

$$A \subset aH = Ha \quad \forall a \in A$$

(i.e. A is a large portion of a coset of a finite subgroup).

Remark. Converse: If $A \subset xH = Hx$ for $x \in G$, with $H < G$ and $|H| \leq K|A|$ then $|H^2| \leq K|A|$. So this is a complete classification of sets of very small doubling.

Lemma 1.2 (Identify H)**.** If $|A^2| < \frac{3}{2}|A|$ then $H = A^{-1}A$ is a subgroup. Moreover, $A^{-1}A = AA^{-1}$ and $|H| < 2|A|$.

Proof. Let $a, b \in A$. The hypothesis gives $|aA \cap bA| > \frac{1}{2}|A|$, so there are more than $\frac{1}{2}|A|$ pairs $(x, y) \in A \times A$ such that $ax = by$, i.e. $a^{-1}b = xy^{-1}$. This immediately gives $A^{-1}A \subseteq AA^{-1}$, and replacing A by A^{-1} gives $AA^{-1} \subseteq A^{-1}A$, so $A^{-1}A = AA^{-1}$ as required.

Since $|A \times A| = |A|^2$ it also implies that

$$\left|A^{-1}A\right| \leq \frac{|A|^2}{\frac{1}{2}|A|} = 2|A|,$$

(dividing by number of repetitions), as claimed.

Note also that $A^{-1}A$ is symmetric, so it remains to show that $A^{-1}A$ is closed under products.

Let $c, d \in A$. As above, there are more than $\frac{1}{2}|A|$ pairs $(u, v) \in A \times A$ such that $c^{-1}d = uv^{-1}$. This means that for at least one pair (x, y) as above and one pair (u, v), we have $y = u$. In particular, $a^{-1}bc^{-1}d = xv^{-1} \in AA^{-1} = A^{-1}A$. □

Lemma 1.3 (Size bound)**.** If $|A^2| < \frac{3}{2}|A|$ then $A^2 = aHa \quad \forall a \in A$ (H as before). In particular, $|H| = |A^2|$.

Proof. First, note that

$$A \subset aH \cap Ha \tag{1.3}$$

by definition of H, so certainly $A^2 \subset aHa$. For the reverse inclusion, let $z \in aHa$. Since H is a subgroup, there are $|H|$ pairs $(x, y) \in aH \times Ha$ such that $z = xy$.

Moreover, by (1.3) and the bound $|H| < 2|A|$ from Lemma 1.2, more than half of these x and more than half of these y belong to A. In particular, this means that for at least one pair x, y, both have to belong to A. Hence $z = xy \in A^2$, as required. □

Proof of Theorem 1.1. Given $a \in A$, we have $Aa^{-1} \subset aHa^{-1} \cap H$ so

$$|aHa^{-1} \cap H| \geq |A| > \frac{1}{2}|H|$$

by Lemma 1.2, but the only subgroup of H of size $> \frac{1}{2}|H|$ is H itself. Hence $aHa^{-1} = H$, so indeed $A \subset aH = Ha$ by (1.3). □

Classifying the sets of small doubling is much harder than this in general, and uses a much wider range of techniques, e.g. group theory, harmonic analysis, geometry of numbers...

2 Covering and higher sum and product sets

Lecture 2

We introduce two techniques we'll use repeatedly: *covering* and *bounding higher product sets*. A nice way to do this is by proving the following theorem.

Theorem 2.1 (Ruzsa)**.** Suppose $A \subset \mathbb{F}_p^r$ satisfies $|A+A| \leq K|A|$. Then $\exists H \leq \mathbb{F}_p^r$ with

$$|H| \leq p^{K^4} K^2 |A| \quad \text{such that } A \subset H.$$

So again, like Theorem 1.1, A is a large proportion of a finite subgroup.

Remark. It is not ideal that $\frac{|A|}{|H|}$ depends on p. We will remove this dependency in a few lectures' time.

We'll start by proving the following weaker version:

Proposition 2.2. Suppose $A \subset \mathbb{F}_p^r$ satisfies $|2A-2A| \leq K|A|$. Then $\exists H < \mathbb{F}_p^r$ with

$$|H| \leq p^K |A-A| (\leq p^K K|A|) \quad \text{such that } A \subset H.$$

We'll prove this using 'covering', encapsulated by the following lemma

Lemma 2.3 (Ruzsa's covering lemma)**.** Suppose $A, B \subset G$ and $|AB| \leq K|B|$. Then $\exists X \subset A$ with $|X| \leq K$ such that $A \subset XBB^{-1}$. Indeed, we may take $X \subset A$ maximal such that the sets xB (for $x \in X$) are disjoint.

The term 'covering' refers to the conclusion $A \subset XBB^{-1}$, which says 'A can be covered by a few left-translates of BB^{-1}.

Proof. First, disjointness of xB gives that $|XB| = |X||B|$. Since $X \subset A$,

$$|XB| \leq |AB| \leq K|B|,$$

so $|X| \leq K$. By maximality, for all $a \in A$, there is $x \in X$ such that $aB \cap xB \neq \varnothing$, and hence $a \in xBB^{-1}$. Hence $A \subset XBB^{-1}$, as required. □

Lemma 2.4. Suppose $A \subset G$ satisfies $|A^{-1}A^2A^{-1}| \leq K|A|$. Then $\exists X \subset A^{-1}A^2$, with $|X| \leq K$ such that $A^{-1}A^n \subset X^{n-1}A^{-1}A$ for any $n \in \mathbb{N}$.

Proof. By Lemma 2.3, $\exists X \subset A^{-1}A^2$, $|X| \leq K$ such that

$$A^{-1}A^2 \subset XA^{-1}A. \tag{2.1}$$

We then have

$$\begin{aligned} A^{-1}A^n &= A^{-1}A^{n-1}A \\ &\subset X^{n-2}A^{-1}A^2 \quad \text{by induction} \\ &\subset X^{n-1}A^{-1}A. \quad \text{by (2.1)} \end{aligned}$$

□

Proof of Proposition 2.2. By Lemma 2.4, $\exists X$ with $|X| \leq K$ such that

$$nA - A \subset (n-1)X + A - A \quad \forall n \in \mathbb{N}.$$

This means that $\langle A \rangle \subset \langle X \rangle + A - A$, so

$$|\langle A \rangle| \leq |\langle X \rangle| \, |A-A| \leq p^K |A-A|$$

as claimed. □

To strengthen Proposition 2.2 to Theorem 2.1, we use the second technique of this section, bounding higher sum/product sets. The key result is the following, at least in the abelian case.

Theorem 2.5 (Plünnecke-Ruzsa)**.** Suppose $A \subset G$ for G abelian, and $|A+A| \leq K|A|$. Then for all $m, n \geq 0$,

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

This was proved in Introduction to Discrete Analysis last term. We won't redo the whole proof in lectures, but we will reprove some parts of it. See the Example Sheet for the full result.

Proof of Theorem 2.1. Using Theorem 2.5, $|2A - 2A| \leq K^4|A|$, and $|A - A| \leq K^2|A|$. Then immediate from Proposition 2.2. □

We'll spend the rest of this section discussing Theorem 2.5 and variants of it. We've seen it's useful, at least in one context. To see more philosophically why it's useful, let's think about what the genuine closure of subgroups under products and inverses means. One useful feature is that it can be iterated: given $h_1, h_2, \ldots \in H$, a subgroup, this means that $h_1^{\epsilon_1} \cdots h_m^{\epsilon_m} \in H$ $\forall \epsilon_i = \pm 1$, $\forall n, \forall h_i \in H$. Theorem 2.5 allows us to 'iterate' the 'approximate closure' of a set of small doubling: $a_1 + \cdots + a_m - a_1' - \cdots - a_n'$ may not belong to A, but at least it belongs to $mA - nA$, which is

(a) not too large ($|mA - nA| \leq K^{m+n}|A|$)

(b) itself a set of small doubling ($|2(mA - nA)| \leq K^{2m+2n}|mA - nA|$).

This is an important part of why the theory works so well.

It is therefore unfortunate that Theorem 2.5 does not hold for non-abelian groups:

Example 2.6. Let x generate an infinite cyclic group $\langle x \rangle$, H be a finite group, set $G = H * \langle x \rangle$ (the free product, which has the important property that $x^{-1}Hx \neq H$). Set $A = H \cup \{x\}$. $A^2 = H \cup xH \cup Hx \cup \{x^2\}$, so $|A^2| \leq 3|A|$. But A^3 contains HxH, which has size $|H|^2 \sim |A|^2$.

So as $|H| \to \infty$, Theorem 2.5 cannot hold.

Nonetheless, if we strengthen small doubling slightly, we can recover a form of Theorem 2.5. One way is to replace small doubling with **small tripling**: $|A^3| \leq K|A|$.

Proposition 2.7. Suppose $A \subset G$, $|A^3| \leq K|A|$. Then $|A^{\epsilon_1} \cdots A^{\epsilon_m}| \leq K^{3(m-2)}|A|$ $\forall \epsilon_i = \pm 1$, $\forall m \geq 3$.

The key ingredient is the following:

Lemma 2.8 (Ruzsa's triangle inequality)**.** Given $U, V, W \subset G$, all finite, we have

$$|U||V^{-1}W| \leq |UV||UW|.$$

Proof. We'll define an injection $\varphi : U \times V^{-1}W \to UV \times UW$. First, for $x \in V^{-1}W$, set $v(x) \in V$ and $w(x) \in W$ such that $x = v(x)^{-1}w(x)$. Set $\varphi(u, x) = (uv(x), uw(x))$. To see that φ is injective, first observe

$$(uv(x))^{-1}(uw(x)) = x,$$

so x determined by $\varphi(u, x)$, and then

$$(uv(x))v(x)^{-1} = u,$$

so u is also determined by $\varphi(u, x)$. □

Proof of Proposition 2.7. First we'll do the case $m = 3$.

- $|A^3| = |A^{-3}| \leq K|A|$.
- Apply Ruzsa's triangle inequality with $U = W = A$, $V = A^2$:

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

  so $|A^{-2}A| \leq K^2|A|$.
- Note that $(A^{-2}A)^{-1} = A^{-1}A^2$, so $|A^{-1}A^2| = |A^{-2}A| \leq K^2|A|$.
- Replacing A by A^{-1} we get

$$|AA^{-2}| = |A^2A^{-1}| \leq K^2|A|.$$

- Finally, Ruzsa's triangle inequality with $U = V = A$, $W = AA^{-1}$ gives

$$|A||A^{-1}AA^{-1}| \leq |A^2||A^2A^{-1}| \leq K^3|A|^2.$$

  So $|A^{-1}AA^{-1}| \leq K^3|A|$.
- For the last case, swap A, A^{-1} again.

For $m \geq 4$, Ruzsa's triangle inequality gives

$$\begin{aligned} |A||A^{\epsilon_1} \cdots A^{\epsilon_m}| &\leq |AA^{-\epsilon_2}A^{-\epsilon_1}||AA^{\epsilon_3} \cdots A^{\epsilon_m}| \\ &\leq K^3|A|\ K^{3(m-3)}|A|. \end{aligned}$$

□

3 Approximate Groups

Lecture 3

In the last section, we saw that assuming small tripling instead of small doubling allowed us to control higher product sets of the form $A^{\epsilon_1} \cdots A^{\epsilon_m}$. In this section, we'll see another possible strengthening of small doubling. We also saw, in the proofs of Theorem 2.1 and Proposition 2.2, an advantage of having a 'covering' condition in place of a size bound. This motivates in part the following definition.

Definition. A set $A \subset G$ is called a **K-approximate group** (or K-approximate subgroup) if $1 \in A$, $A^{-1} = A$ and $\exists X \subset G$ with $|X| \leq K$ such that $A^2 \subset XA$.

Note that A need not be finite, although in this course it almost always will be. Also, if A is finite, then $|A^2| \leq K|A|$. The conditions $1 \in A$ and $A^{-1} = A$ are convenient 'notationally': for example, this lets us write A^m instead of $A^{\epsilon_1} \cdots A^{\epsilon_m}$ and $1 \in A$ gives us that $A \subset A^2 \subset A^3 \subset \cdots$, which is also convenient at times. It's the condition $A^2 \subset XA$ that is most important.

For approximate groups, bounding higher product sets is easy:

Lemma 3.1. If A is a finite K-approximate group then $|A^m| \leq K^{m-1}|A|$.

Proof. If X is as in the definition of approximate group, in fact we have $A^m \subset X^{m-1}A$:

$$\begin{aligned} A^m &= A^{m-1}A \\ &\subset X^{m-2}A^2 \quad \text{induction} \\ &\subset X^{m-1}A \quad \text{definition of } X \end{aligned}$$

□

Another advantage of approximate groups is that if $\pi : G \to H$ is a homomorphism and $A \subset G$ is a K-approximate group then $\pi(A)$ is also trivially a K-approximate group (although we'll see that there is a version of this for small tripling).

It turns out that sets of small tripling and approximate groups are essentially equivalent, in the following sense:

Proposition 3.2. Let $A \subset G$ be finite. If A is a K-approximate group then $|A^3| \leq K^2|A|$. Conversely if $|A^3| \leq K|A|$ then there is a $\mathcal{O}(K^{12})$-approximate group B with $A \subset B$ and $|B| \leq 7K^3|A|$ ('A is a large proportion of an approximate group'). In fact, we may take $B = (A \cup \{1\} \cup A^{-1})^2$.

Proof. First bit is just Lemma 3.1. For the converse, set $\hat{A} = A \cup \{1\} \cup A^{-1}$, and note that

$$B = \hat{A}^2 = \{1\} \cup A \cup A^{-1} \cup A^2 \cup A^{-1}A \cup AA^{-1} \cup A^{-2}.$$

Each set in this union has size $\leq K^3|A|$ by Proposition 2.7, so $|B| \leq 7K^3|A|$, as claimed. Similarly,

$$\hat{A}^4 = \bigcup_{\substack{\epsilon_i = \pm 1 \\ 0 \leq m \leq 4}} A^{\epsilon_1} \cdots A^{\epsilon_m},$$

and the sets in this union have size $\leq K^6|A|$. It follows that $|\hat{A}^4| \leq \mathcal{O}(K^6)|\hat{A}|$.

So, Lemma 2.4 implies $\exists X \subset G$ with $|X| \leq \mathcal{O}(K^6)$ such that $\hat{A}^n \subset X^{n-2}\hat{A}^2$ for every $n \geq 2$. In particular, $|X^2| \leq \mathcal{O}(K^{12})$ and $\hat{A}^4 = (\hat{A}^2)^2 \subset X^2\hat{A}^2$, so $\hat{A}^2$ is an $\mathcal{O}(K^{12})$-approximate group, as claimed. □

This is all well and good, but what if we are faced with a set like that from Example 2.6 which only has small doubling? In that specific example, a large proportion of A was a set of small tripling, namely H. Rather helpfully, that turns out to be a general phenomenon:

Theorem 3.3. If $A \subset G$ satisfies $|A^2| \leq K|A|$ then there is $U \subset A$ with $|U| \geq \frac{1}{K}|A|$ such that $|U^m| \leq K^{m-1}|U|$ $\forall m \in \mathbb{N}$.

So small doubling reduces to small tripling, which reduces to approximate groups. In Example Sheet 1, we'll see a direct reduction from small doubling to approximate group.

Tao proved a version of Theorem 3.3 when he introduced the definition of approximate groups. We will use instead a lemma of Petridis, which he proved when proving the Plünnecke-Ruzsa inequalities.

Lemma 3.4 (Petridis)**.** Suppose $A, B \subset G$ are finite, let $U \subset A$ be non-empty, chosen to minimise the ratio $|UB|/|U|$, and write $R = |UB|/|U|$. Then for any finite $C \subset G$, we have

$$|CUB| \leq R|CU|.$$

Proof. Trivial if $C = \varnothing$, so we may assume $\exists x \in C$. Defining $C' = C \setminus \{x\}$, we may also assume by induction that $|C'UB| \leq R|C'U|$. Set $W = \{u \in U \mid xu \in C'U\}$. Then

$$CU = C'U \cup (xU \setminus xW)$$

is a disjoint union, so in particular

$$|CU| = |C'U| + |U| - |W|. \tag{3.1}$$

We also have $xUB \subset C'UB$ by definition of W, so

$$CUB \subset C'UB \cup (xUB \setminus xWB)$$

and hence

$$|CUB| \leq |C'UB| + |UB| - |WB|. \tag{3.2}$$

We have $|C'UB| \leq R|C'U|$ by the induction hypothesis, $|UB| = R|U|$ by definition of R, and $|WB| \geq R|W|$ by minimality in the definition of U. So

$$\begin{aligned} |CUB| &\leq R(|C'U| + |U| - |W|) \quad \text{by (3.2)} \\ &= R|CU| \quad \text{by (3.1).} \end{aligned}$$

□

Proof of Theorem 3.3. Set $U \subset A$ to be non-empty minimising $|UA|/|U|$, and write R for this ratio, noting that $R \leq K$ by minimality. Also, U non-empty, so $|UA| \geq |A|$, so $|U| \geq \frac{|A|}{K}$, as required. Lemma 2.4 also implies that $|U^m A| \leq K|U^m|$ $\forall m$ (taking $C = U^{m-1}$) and since $U \subset A$, this gives $|U^{m+1}| \leq K|U^m|$ $\forall m$, so $|U^m| \leq K^{m-1}|U|$. □

*Bonus content

The reason A in Example 2.6 failed to have small tripling was the existence of $x \in A$ with AxA large. It turns out that this is the only obstruction to small doubling having small tripling:

Theorem (Tao; Petridis)**.** If $|A^2| \leq K|A|$ and $|AxA| \leq K|A|$ $\forall x \in A$ then

$$|A^m| \leq K^{\mathcal{O}(m)}|A| \ \forall m \geq 3.$$

4 Stability of approximate closure under basic operations

Lecture 4

Two familiar properties of genuine subgroups are that they behave well under quotients and intersections: if $H < G$ and $\pi : G \to \Gamma$ is a homomorphism, then $\pi(H) < \Gamma$, and if $H_1, H_2 < G$ then $H_1 \cap H_2 < G$. In this lecture, we'll see versions of these properties for approximate groups and sets of small tripling.

It's trivial that if $A \subset G$ is a K-approximate group then $\pi(A)$ is also a K-approximate group. The following is the corresponding result for sets of small tripling.

Lemma 4.1 (Stability of small tripling under homomorphisms)**.** Let $A \subset G$ be finite, symmetric containing the identity. Suppose $\pi : G \to H$ is a homomorphism. Then

$$\frac{|\pi(A)^m|}{|\pi(A)|} \leq \frac{|A^{m+2}|}{|A|} \quad \forall m \in \mathbb{N}.$$

In particular, if $|A^3| \leq K|A|$ then $|\pi(A)^3| \leq K^9|\pi(A)|$ by Proposition 2.7. We can prove this using an argument of Helfgott. We'll start with a simple observation that we'll use repeatedly in the course.

Lemma 4.2. Let $H < G$, let $A \subset G$ be finite, and let $x \in G$. Then

$$|A^{-1}A \cap H| \geq |A \cap xH|.$$

Proof. We have $(A \cap xH)^{-1}(A \cap xH) \subset A^{-1}A \cap H$. □

Remark. Most of the lemmas and propositions in this section will have familiar/trivial analogues for genuine subgroups. It is a useful exercise to think about what they are.

Lemma 4.3. Let $H < G$, write $\pi : G \to G/H$ for the quotient map, let $A \subset G$ be finite. Then

$$|A^{-1}A \cap H| \geq \frac{|A|}{|\pi(A)|}.$$

Note that H is not assumed normal, so G/H is just the space of left cosets xH, not necessarily a group.

Proof. The pigeonhole principle gives $\exists x \in G$ such that $|A \cap xH| \geq |A|/|\pi(A)|$. Then apply Lemma 4.2. □

Lemma 4.4. Let $H < G$, write $\pi : G \to G/H$ for the quotient map, and let $A \subset G$ be finite. Then

$$|\pi(A^m)||A^n \cap H| \leq |A^{m+n}| \ \forall m, n \geq 0.$$

Proof. Define $\varphi : \pi(A^m) \to A^m$ by picking arbitrarily for each $x \in \pi(A^m)$ some $\varphi(x)$ such that $\pi(\varphi(x)) = x$. Then the cosets $\varphi(x)H$ for $x \in \pi(A^m)$ are all distinct by definition, so

$$|\varphi(\pi(A^m))|\,|(A^n \cap H)| = |\pi(A^m)|\,|A^n \cap H|\,.$$

But also, $\varphi(\pi(A^m))(A^n \cap H) \subset A^{m+n}$. □

Proof of Lemma 4.1. Take $H = \ker \pi$. Lemma 4.4 gives

$$|\pi(A^m)| \leq \frac{|A^{m+2}|}{|A^2 \cap H|}.$$

Lemma 4.3 gives

$$|A^2 \cap H| \geq \frac{|A|}{|\pi(A)|}.$$

The proposition follows by combining these inequalities. □

Now we'll look at intersections.

Proposition 4.5 (Stability of small tripling under intersections with subgroups)**.** Let $A \subset G$ be finite, symmetric, contain 1. Let $H < G$. Then

$$\frac{|A^m \cap H|}{|A^2 \cap H|} \leq \frac{|A^{m+1}|}{|A|}.$$

In particular, by Proposition 2.7, if $|A^3| \leq K|A|$ then $\left|(A^m \cap H)^3\right| \leq K^9 |A^m \cap H|$ for any $m \geq 2$.

Remark. We'll see in Example Sheet 1 that even if A has small tripling, $A \cap H$ need not. So $m \geq 2$ really is important for this last conclusion.

Proof. Take $\pi : G \to G/H$ as before. Lemma 4.4 gives

$$|A^m \cap H| \leq \frac{|A^{m+1}|}{|\pi(A)|}.$$

Lemma 4.3 gives

$$|A^2 \cap H| \geq \frac{|A|}{|\pi(A)|}.$$

Combine these two inequalities. □

Proposition 4.6 (Stability of approximate groups under intersections with subgroups)**.** Let $H < G$, let $A \subset G$ be a K-approximate group. Then $A^m \cap H$ is covered by $\leq K^{m-1}$ left-translates of $A^2 \cap H$. In particular, $A^m \cap H$ is a K^{2m-1}-approximate group (since $A^2 \cap H \subset A^m \cap H$ and $(A^m \cap H)^2 \subset A^{2m} \cap H$).

Proof. By definition, $\exists X \subset G$ with $|X| = K^{m-1}$ such that $A^m \subset XA$. In particular,

$$A^m \cap H \subset \bigcup_{x \in X} (xA \cap H).$$

For each $xA \cap H$ that is not empty, $\exists h = xa \in H$ for some $a \in A$. This means that

$$xA \cap H \subset h(a^{-1}A \cap H) \subset h(A^2 \cap H).$$

Hence each set $xA \cap H$ in this union is contained in a single left translate of $A^2 \cap H$. □

In Introduction to Discrete Analysis, we saw that when studying small doubling or tripling there is a more general notion of homomorphism that comes into play: the Freiman homomorphism. To motivate this, consider two sets $A = \{-n, \dots, n\} \subset \mathbb{Z}/p\mathbb{Z}$, and $B = \{-n, \dots, n\} \subset \mathbb{Z}/q\mathbb{Z}$ for p, q two primes $\geq 10n$, say. These two sets are intuitively 'isomorphic' from the perspective of $A + A$ or $B + B$, but there is no way of encoding this with a group homomorphism $\mathbb{Z}/q\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$. Freiman homomorphisms give us a way to encode this.

Definition. Let $m \in \mathbb{N}$, let A, B be subsets of groups.

- A map $\varphi : A \to B$ is a **Freiman m-homomorphism** if $\forall x_1, \dots, x_m, y_1, \dots, y_m \in A$,
$$x_1 \cdots x_m = y_1 \cdots y_m \implies \varphi(x_1) \cdots \varphi(x_m) = \varphi(y_1) \cdots \varphi(y_m).$$
- If $1 \in A$ and $\varphi(1) = 1$ then we say that φ is **centered**.
- If φ is injective and its inverse $\varphi(A) \to A$ is also a Freiman m-homomorphism then we say $\varphi : A \to \varphi(A)$ is a **Freiman m-isomorphism**.

– We often simply write that φ is a 'Freiman homomorphism' when it is a Freiman 2-homomorphism.

Remark.

(1) Every map is trivially a 1-homomorphism, so we only care about the cases $m \geq 2$.

(2) This definition gets stronger as m increases: we may assume $A \neq \varnothing$, and then, picking $a \in A$ arbitrarily, if $x_1 \cdots x_k = y_1 \cdots y_k$ for $k \leq m$ then

$$x_1 \cdots x_k \underbrace{a \cdots a}_{m-k} = y_1 \cdots y_k \underbrace{a \cdots a}_{m-k}.$$

(3) If φ is a centered m-homomorphism and $a, a^{-1} \in A$ then exercise to check that $\varphi(a^{-1}) = \varphi(a)^{-1}$ (for $m \geq 2$).

Lemma 4.7. Suppose $\varphi : A \to \Gamma$ is a Freiman m-homomorphism. Then

$$|\varphi(A)^m| \leq |A^m|.$$

In particular, if φ is injective then

$$\frac{|\varphi(A)^m|}{|\varphi(A)|} \leq \frac{|A^m|}{|A|},$$

and if φ is a Freiman m-isomorphism then

$$\frac{|\varphi(A)^m|}{|\varphi(A)|} = \frac{|A^m|}{|A|}.$$

Proof. Exercise. □

Lemma 4.8. Let $A \subset G$ be a K-approximate group, and suppose $\varphi : A^3 \to \Gamma$ is a centred Freiman homomorphism. Then $\varphi(A)$ is also a K-approximate group.

Proof. Take X, $|X| \leq K$ such that $A^2 \subset XA$ as in the definition of K-approximate group. So given $a_1, a_2 \in A$ $\exists x \in X, a_3 \in A$ such that $a_1 a_2 = x a_3$. In particular, $x \in A^3$ so $\varphi(x)$ is defined, and

$$\varphi(a_1)\varphi(a_2) = \varphi(x)\varphi(a_3)$$

Hence $\varphi(A)^2 \subset \varphi(X \cap A^3)\varphi(A)$. Also φ is centred so $\varphi(A)$ is symmetric and contains 1. □

5 Coset progressions, Bohr sets and the Freiman-Green-Ruzsa theorem

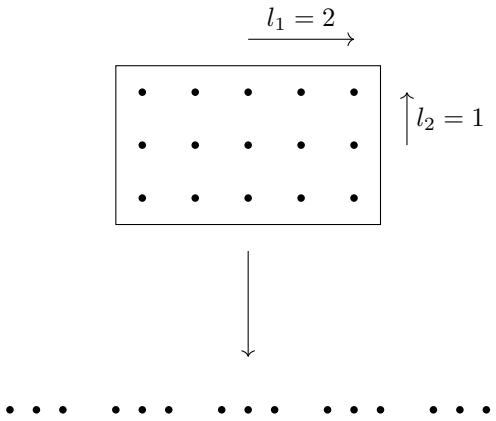
Lecture 5

Definition. Let G be abelian, $x_1, \dots, x_r \in G$, $L_1, \dots, L_r \in \mathbb{N}$. Then the set

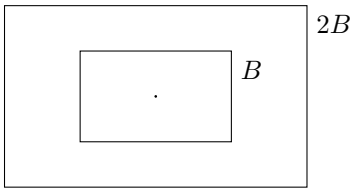
$$P(x; L) = P(x_1, \dots, x_r; L_1, \dots, L_r) = \{l_1 x_1 + \cdots l_r x_r \mid |l_i| \leq L_i \ \forall i\}$$

is called a **progression** of rank r. If in addition $H < G$ is finite, then $H + P(x, L)$ is called a **coset progression** of rank r.

It is useful to think of $P(x; L)$ as a homomorphic image of a 'box' in $\mathbb{Z}^r$, e.g. if $G = \mathbb{Z}$ and $r = 2$:



It is easy to see that such a box B in $\mathbb{Z}^r$ is a 2^r-approximate group, e.g. in $r = 2$



Hence $P(x; L)$ is also a 2^r-approximate group, as is $H + P(x; L)$. Remarkably, these are essentially the only examples:

Theorem 5.1 (Freiman ($G = \mathbb{Z}$); Green-Ruzsa (arbitrary abelian G))**.** Suppose $A \subset G$ (abelian) satisfies $|A + A| \leq K|A|$. Then there is a coset progression $H + P$ of rank $\leq \mathcal{O}(K^{\mathcal{O}(1)})$ such that

$$A \subset H + P \subset \mathcal{O}(K^{\mathcal{O}(1)})(A \cup \{0\} \cup -A).$$

In particular, Plünnecke-Ruzsa gives $|H + P| \leq \exp(\mathcal{O}(K^{\mathcal{O}(1)}))|A|$. So A is a large proportion of $H + P$.

A substantial part of this result was proved in Introduction to Discrete Analysis, but with a slightly less explicit version of coset progressions.

Definition. Let G be a finite abelian group, let

$$\Gamma = \{\gamma_1, \dots, \gamma_r\} \subset \hat{G}$$

(recall $\hat{G} = \mathrm{Hom}(G, \mathbb{R}/\mathbb{Z})$), and let $p \in [0, \frac{1}{2}]$. Then the set

$$B(\Gamma, p) = \{g \in G : \forall i \ \ \|\gamma_i(g)\|_{\mathbb{R}/\mathbb{Z}} \leq p\}$$

is called a **Bohr set** of rank r. Here, given $x \in \mathbb{R}/\mathbb{Z}$ with representative $\hat{x} \in (-\frac{1}{2}, \frac{1}{2}]$, we write $\|x\|_{\mathbb{R}/\mathbb{Z}} = |\hat{x}|$.

We will see on Sheet 1 that $B(\Gamma, p)$ is a 4^r-approximate group. Whereas progressions were homomorphic images of boxes, $B(\Gamma, p)$ is the pullback of $[-p, p]^r$ under $(\gamma_1, \dots, \gamma_r) \in \hat{G}^r$.

It turns out that the notions of coset progression and Bohr set are essentially equivalent. In Sheet 2, we will see that every coset progression is a Freiman image of a Bohr set of the same rank. Moreover, every Freiman image of a Bohr set is a large proportion of some coset progression. We will see a special case of that shortly.

Proposition 5.2. Suppose $A \subset G$ (abelian) with $|A+A| \leq K|A|$. Then $\exists B \subset 2A-2A$, a finite abelian group Z with $|Z| \geq |A|$, a set $\Gamma \subset \hat{Z}$ with $|\Gamma| \leq \mathcal{O}(K^{\mathcal{O}(1)})$, some $p \geq \frac{1}{\mathcal{O}(K^{\mathcal{O}(1)})}$, and a centered Freiman 2-isomorphism $\varphi : B(\Gamma, p) \to B$.

'$2A - 2A$ contains a large set isomorphic to a Bohr set of bounded rank'.

In Introduction to Discrete Analysis, we saw this in the special case of G torsion-free. The general case is harder, but nonetheless conceptually very similar, so we will assume this result from now on.

To pass from Proposition 5.2 to Theorem 5.1, we use the following results.

Proposition 5.3. Suppose G is a finite abelian group, $\Gamma \subset \hat{G}$ is of size r, $p < \frac{1}{10}$. Then there is a coset progression $H + P \subset B(\Gamma, p)$ with rank r and

$$|H + P| \geq \left(\frac{p}{r}\right)^r |G|.$$

Lemma 5.4. Suppose $H + P$ is a coset progression of rank r, and $\varphi : H + P \to G$ (abelian) is a centered Freiman 2-homomorphism. Then $\varphi(H + P)$ is also a coset progression of rank r.

Proof. Exercise: If H a group and $\varphi : H \to G$ is a centered Freiman 2-homomorphism, then φ is also a group homomorphism. In particular, in this lemma $\varphi(H)$ is a finite subgroup. Therefore sufficient to show that

$$\varphi(H + P(x; L)) = \varphi(H) + P(\varphi(x_1), \dots, \varphi(x_i); L_1, \dots, L_r).$$

In fact, we will show that $\forall h \in H$, $|l_i| \leq L_i$ we have

$$\varphi(h + l_1x_1 + \cdots + l_rx_r) = \varphi(h) + l_1\varphi(x_1) + \cdots + l_r\varphi(x_r). \tag{5.1}$$

Since φ centered, $\varphi(-x_i) = -\varphi(x_i)$, so we may assume $l_i \geq 0$ $\forall i$. Also (5.1) is trivial if $l_i = 0$ $\forall i$, so may assume $\exists l_j > 0$. Then

$$\begin{aligned}\varphi(h + l_1x_1 + \cdots + l_rx_r) &= \varphi(h + l_1x_1 + \cdots + l_rx_r) + \varphi(0) \\ &= \varphi(h + l_1x_r + \cdots + (l_j - 1)x_j + \cdots + l_rx_r) + \varphi(x_j)\end{aligned}$$

so the result follows by induction on $\sum_i l_i$. □

Proof of Theorem 5.1. By Proposition 5.2 and Proposition 5.3 and Lemma 5.4, $\exists H + P$ a coset progression of rank $\leq \mathcal{O}(K^{\mathcal{O}(1)})$ such that $H + P \subset 2A - 2A$ and $|H + P| \geq \exp(-\mathcal{O}(K^{\mathcal{O}(1)}))|A|$. We will now apply a version of Ruzsa's covering lemma due to Chang.

Define recursively sets $S_1, S_2, \ldots \subset A$ such that S_i is a maximal subset of size $\leq 2K$ such that the translates

$$x + S_{i-1} + \cdots + S_1 + H + P \quad x \in S_i$$

are all disjoint. If ever $|S_i| < 2K$ we stop. Now suppose we get as far as $S_1, \ldots, S_t$. Then

$$S_t + \ldots + S_1 + H + P \subset 2A - 2A + tA,$$

so Theorem 2.5 gives

$$|S_t + \cdots + S_1 + H + P| \leq K^{4+t}|A|.$$

On the other hand, disjointness of the translates in the definition of S_i means that

$$\begin{aligned} |S_t + \cdots + S_1 + H + P| &\geq |S_t| \cdots |S_1||H + P| \\ &\geq (2K)^{t-1} \exp(-\mathcal{O}(K^{\mathcal{O}(1)}))|A|. \end{aligned}$$

Putting these together, we have $2^{t-1} \leq K^5 \exp(\mathcal{O}(K^{\mathcal{O}(1)}))$, hence $t \leq \mathcal{O}(K^{\mathcal{O}(1)})$. In particular, this process terminates. But also, since S_t is therefore maximal among all subsets of A such that

$$x + S_{t-1} + \cdots + S_1 + H + P$$

are disjoint for $x \in S_t$, Ruzsa's covering lemma implies that

$$A \subset H + 2P + S_1 - S_1 + \cdots + S_{t-1} - S_{t-1} + S_t.$$

Enumerating $\bigcup_i S_i$ as $S_1, \ldots, S_d$, we have $d \leq \mathcal{O}(K^{\mathcal{O}(1)})$ and

$$A \subset H + 2P + P(s_1, \ldots, s_d; 1, \ldots, 1) \subset \mathcal{O}(K^{\mathcal{O}(1)})(A \cup \{0\} \cup -A)$$

as claimed. □

Exercise. See what bounds are given if you apply Ruzsa's covering lemma directly, instead of Chang's argument.

6 Geometry of Numbers

Lecture 6

To prove Proposition 5.3, we'll use a field called the *geometry of numbers* which is concerned with lattices in $\mathbb{R}^d$. For us, a lattice $\Lambda \subset \mathbb{R}^d$ will simply be the additive subgroup (not subspace) generated by some basis $x_1, \dots, x_d$ for $\mathbb{R}^d$. So

$$\Lambda = \{l_1 x_1 + \cdots + \lambda_d x_d \mid l_i \in \mathbb{Z}\}$$

If $\Gamma \subset \Lambda$ is another lattice ew call it a sublattice, and write $\Gamma < \Lambda$. On sheet two, we will see

$$\frac{\det(y_1, \dots, y_d)}{\det(x_1, \dots, x_d)} = [\Lambda : \Gamma] \tag{6.1}$$

where $y_1, \dots, y_d$ are a basis for Λ and $x_1, \dots, x_d$ are a basis for Γ.

In particular, if $x_1, \dots, x_d$ and $x'_1, \dots, x'_d$ are bases for the same lattice Λ then

$$\det(x_1, \dots, x_d) = \det(x'_1, \dots, x'_d).$$

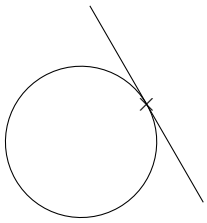
We define this to be $\det(\Lambda)$. The relevance of lattices to Proposition 5.3 is the following:

Lemma 6.1. Let G, Λ be as in Proposition 5.3, and set $\gamma : G \to \mathbb{R}^d/\mathbb{Z}^d$ by enumerating Γ as $\{\gamma_1, \dots, \gamma_d\}$ and setting $\gamma = (\gamma_1, \dots, \gamma_d)$. Then $\Lambda = \gamma(G) + \mathbb{Z}^d$ is a lattice with determinant $\frac{|\ker \gamma|}{|G|}$.

Proof. Λ is finitely generated as G is finite, and torsion-free as in $\mathbb{R}^d$, so isomorphic to $\mathbb{Z}^k$ for some k. Also, Λ has $\mathbb{Z}^d$ as a finite-index subgroup. So $k = d$ and $\operatorname{span}_{\mathbb{R}}(\Lambda) = \mathbb{R}^d$. So we may take a generating set for Λ of size d, which is then a basis for $\mathbb{R}^d$. Determinant follows from (6.1) because $\det(\mathbb{Z}^d) = 1$. □

We'll investigate the interaction of $[-p, p]^d$ with Λ. To do this, we introduce another definition:

Definition. A set $A \subset \mathbb{R}^d$ is convex if $\forall x \in \mathbb{R}^d \setminus A^\circ$ there is a hyperplane h_x with $x \in h_x$ and $h_x \cap A^\circ = \varnothing$.



Definition. A set $B \subset \mathbb{R}^d$ is a convex body if it is bounded and convex and $B^\circ \neq \varnothing$. It is symmetric if $\forall x \in B$ we have $-x \in B$. Given a symmetric convex body B and a lattice Λ, define the successive minima $\lambda_1 \leq \cdots \leq \lambda_d$ of B with respect to Λ via

$$\lambda_i = \inf\{\lambda > 0 \mid \dim \operatorname{span}_{\mathbb{R}}(\lambda \cdot B \cap \Lambda) \geq i\}$$

We may then inductively define linearly independent vectors $v_1, \dots, v_d \in \Lambda^1$ such that $v_1, \dots, v_i \in \lambda_i \bar{B}$. We'll call such a set a directional basis for Λ with respect to B. Note i is not unique, and not necessarily a basis for Λ in the earlier sense (See Ex Sheet 2).

Theorem 6.2 (Minkowski's Second Theorem)**.** Suppose B is a symmetric convex body, Λ a lattice in $\mathbb{R}^d$ and $\lambda_1, \dots, \lambda_d$ are the successive minima. Then $\lambda_1 \cdots \lambda_d \operatorname{vol}(B) \leq 2^d \det(\lambda)$.

Lemma 6.3 (Blichfeldt)**.** Suppose $A \subset \mathbb{R}^d$ is a measurable set, Λ a lattice and $\forall a, b \in A$ distinct we have $a - b \notin \Lambda$. Then $\mathrm{vol}(A) \leq \det(\Lambda)$.

Proof. Fix a basis $x_1, \ldots, x_d$ for Λ and define the **fundamental parallelopiped** P with respect to $x_1, \ldots, x_d$ via

$$P = \{l_1 x_1 + \cdots + l_d x_d \mid l_i \in [0, 1)\}$$

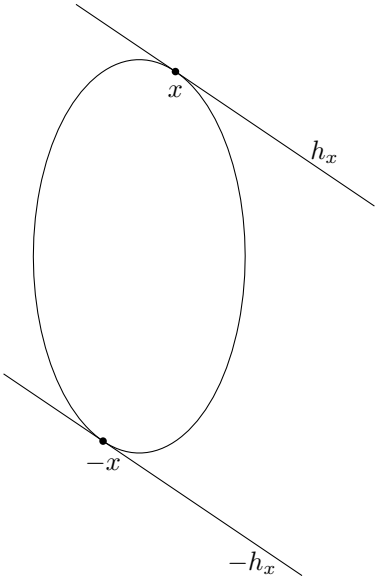
Since $x_1, \ldots, x_d$ is a basis for $\mathbb{R}^d$, $\forall v \in \mathbb{R}^d$ there are unique $x_v \in \Lambda$ and $p_v \in P$ such that $v = x_v + p_v$. Define a map $\varphi : \mathbb{R}^d \to P$ via $\varphi(v) = p_v$. This cuts A into countably many measurable pieces, and translates these pieces into P. It is injective by hypothesis, hence volume preserving, and so

$$\mathrm{vol}(A) = \mathrm{vol}(\varphi(A)) \leq \mathrm{vol}(P) = \det(\Lambda). \qquad \square$$

Proof of Theorem 6.2. Let $v_1, \ldots, v_d$ be a directional basis for Λ with respect to B. Set $V_i = \mathrm{span}_{\mathbb{R}}(v_1, \ldots, v_i)$ (including $V_0 = \{0\}$), and set $\Lambda_i = \Lambda \cap (V_i \setminus V_{i-1})$. Λ is a disjoint union $\bigcup_{i=0}^{d} \Lambda_i$.

Claim 1: We have $\lambda_d B^\circ \cap (\lambda_d B^\circ + \alpha x) = \varnothing$ whenever $x \in \Lambda_j$ and $\alpha \geq \frac{2\lambda_d}{\lambda_j}$.

Proof of claim Given $x \in \Lambda_j$, by definition $x \notin \lambda_j B^\circ$, so by convexity there is a hyperplane h_x such that $x \in h_x$ and $H_x \cap \lambda_j B^\circ = \varnothing$. By symmetry, we may take $h_{-x} = -h_x$.



Note, however, that $-h_x = h_x - 2x$. That means that $x_j B^\circ$ is contained in the slice of space S_x between the parallel hyperplanes h_x and $h_x - 2x$. Clearly $S_x \cap (S_x + \alpha x) = \varnothing$ for all $\alpha \geq 2$, so in particular

$$\lambda_j B^\circ \cap (\lambda_j B^\circ + \alpha x) = \varnothing$$

for all such α as well. Scaling by λ_d / λ_j, we see that $\lambda_d B^\circ \cap (\lambda_d B^\circ + \alpha x) = \varnothing$ whenever $\alpha \geq 2\frac{\lambda_d}{\lambda_j}$, proving the claim. ■

Claim 2: There are sets $B_1 \subset B_2 \subset \cdots \subset B_d = \lambda_d B^\circ$ such that

1. $\mathrm{vol}(B_i) = (\frac{\lambda_i}{\lambda_{i+1}})^i \mathrm{vol}(B_{i+1})$ $\forall i$

2. we have $B_i \cap (B_i + \alpha x) = \varnothing$ whenever $x \in \Lambda_j$ and $\alpha \geq 2\max\{\frac{\lambda_i}{\lambda_j}, 1\}$. $B_i \cap (B_i + \alpha x) = \varnothing$ whenever $x \in \Lambda_j$ and $\alpha \geq 2\max\{\frac{\lambda_i}{\lambda_j}, 1\}$. $B_i \cap (B_i + \alpha x) = \varnothing$ whenever $x \in \Lambda_j$ and $\alpha \geq 2\max\{\frac{\lambda_i}{\lambda_j}, 1\}$. $B_i \cap (B_i + \alpha x) = \varnothing$ whenever $x \in \Lambda_j$ and $\alpha \geq 2\max\{\frac{\lambda_i}{\lambda_j}, 1\}$.

Proof of claim 2: Define operations $\sigma_1, \dots, \sigma_{d-1}$ on suitable subsets of $\mathbb{R}^d$ as follows. Given L bounded and open, define σ_i separately on each affine subspace $z + V_i$ with $z \in L$. For each affine subspace, fix a particular $z \in L$, and define

$$\varphi(z + v) = z + \frac{\lambda_i}{\lambda_{i+1}} v \quad \forall v \in V_i$$

On each slice, σ_i scales L by a factor of $\frac{\lambda_i}{\lambda_{i+1}}$, centred at z parallel to v_i. Note the following properties

(i) $\operatorname{vol}(\sigma_i(L)) = (\frac{\lambda_i}{\lambda_{i+1}}) \operatorname{vol}(L)$ (by Fubini)

(ii) If $L \cap (z + V_i)$ is open and convex for every z, then $\sigma_i(L) \subset L$ because $z \in L$.

(iii) If $L \cap (z + V_i)$ is open and convex then so is $\sigma_i(L) \cap (z + V_i)$, and indeed so is $\sigma_i(L) \cap \bigcap_{j<i}(z + V_j)$

Set $B_d = \lambda_d B^\circ$, and $B_i = \sigma_i(B_{i+1})$ otherwise.

Conclusion 1 is immediate from property (i). Conclusion 2 follows from Claim 1 when $i = d$. For $i < d$, it follows by induction and repeated application of (ii) and (iii). Indeed, (2) for $i \geq j$ follows from (2) for $i + 1$ because σ_i scales by $\frac{\lambda_i}{\lambda_{i+1}}$ in direction x. For $i < j$, it follows from $B_i \subset B_{i+1}$. ■

To prove the theorem, note that $\operatorname{vol}(B_i) = \lambda_i \cdots \lambda_d \operatorname{vol}(B)$ and by 2, $a - b \notin 2\Lambda$ for every $a, b \in B_1$, so Blichfeldt gives $\operatorname{vol}(B_1) \leq 2^d \det(\Lambda)$. □

Lecture 7 *Proof of Proposition 5.3.* Write $\gamma = (\gamma_1, \dots, \gamma_r) \in \hat{G}^r$, define $\Lambda = \gamma(4) + \mathbb{Z}^r$, which is a lattice of determinant $\frac{|\ker \gamma|}{|G|}$ by Lemma 6.1. Let $\lambda_1, \dots, \lambda_r$ be the successive minima of $[-1, 1]^r$ with respect to Λ, and $v_1, \dots, v_r$ a directional basis. Set $L_i = \lfloor \frac{p}{r\lambda_i} \rfloor$ for each i. Then $P(v_1, \dots, v_r; L_1, \dots, L_r) \subset [-p, p]^r$. Pick for each i, pick $x_i \in G$ such that $\gamma(x_i) = v_i$, and set $H = \ker \gamma$. Write $p = P(x_1, \dots, x_r; L_1, \dots, L_r)$. Then $H + P \subset$

Claim: If $l_1, \dots, l_r$ and $l'_1, \dots, l'_r$ satisfy $|l_i|, |L'_i| \leq L_i$, and

$$l_1 x_1 + \cdots + l_r x_r \in l'_1 x_1 + \cdots + l'_r x_r + H \tag{$*$}$$

then in fact $l_i = l'_i$ for all i. Indeed, $(*)$ implies that

$$(l_1 - l'_1)v_1 + \cdots + (l_r - l'_r)v_r \in \mathbb{Z}^r \cap [-2p, 2p]^r.$$

But since $p < \frac{1}{2}$, this last intersection is just $\{0\}$, proving the claim (as v_i are linearly independent in $\mathbb{R}^r$). ■

Then

$$\begin{aligned} |H + P| &\geq |H|(L_1 + 1) \cdots (L_r + 1) \\ &\geq |H| \left(\frac{p}{r}\right)^r \frac{1}{\lambda_1 \cdots \lambda_r} \\ &\geq |G| \left(\frac{p}{r}\right)^r \end{aligned}$$

by Minkowski's second theorem. □

7 Progressions in the Heisenberg group

Let

$$H(\mathbb{Z}) = \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix} = \left\{ \begin{pmatrix} 1 & n_1 & n_2 \\ 0 & 1 & n_1 \\ 0 & 0 & 1 \end{pmatrix} : n_i \in \mathbb{Z} \right\}$$

be the **Heisenberg group**.

Set

$$u_1 = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & 1 \\ & & 1 \end{pmatrix}, u_2 = \begin{pmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}, u_3 = \begin{pmatrix} 1 & 0 & 1 \\ & 1 & 0 \\ & & 1 \end{pmatrix}$$

and note that any element of $H(\mathbb{Z})$ can be expressed in the form

$$\begin{pmatrix} 1 & n_2 & n_3 \\ & 1 & n_1 \\ & & 1 \end{pmatrix} = u_1^{n_1} u_2^{n_2} u_3^{n_3}$$

and we have the following formula for multiplying elements in this form:

$$(u_1^{n_1} u_2^{n_2} u_3^{n_3})(u_1^{n_1'} u_2^{n_2'} u_3^{n_3'}) = u_1^{n_1+n_1'} u_2^{n_2+n_2'} u_3^{n_3+n_3'+n_1' n_2} \tag{7.1}$$

This is easy to verify by multiplying matrices, but there is a more abstract reason for it. To see this, given $x, y \in G$, define the commutator $[x, y] = x^{-1}y^{-1}xy$. In light of the identity $xy = yx[x, y]$, we can view the commutator as being the 'error' or 'cost' incurred when interchanging two elements. For example, the fact that commutators are trivial in abelian groups can be viewed as capturing the notion that elements can be interchanged freely in an abelian group. The $n_1' n_2$ term in (7.1) arises because we swap the order of $n_1' n_2$ pairs of elements u_1 and u_2.

Now let's see one possible generalisation of progression to non-abelian groups.

Definition. Given $x_1, \dots, x_r \in G$, $L_1, \dots, L_r \geq 0$, we define the **ordered progression** of rank r

$$P_{\text{ord}}(x, L) = P_{\text{ord}}(x_1, \dots, x_r; L_1, \dots, L_r) = \{x_1^{l_1} \cdots x_r^{l_r} : |l_i| \leq L_i \ \forall i\}.$$

Now consider $P = P_{\text{ord}}(u_1, u_2', L_1, L_2)$ for $u_1, u_2 \in H(\mathbb{Z})$ as before, and $L_1, L_2 \geq 0$. We have

$$(u_1^{l_1} u_2^{l_2})(u_1^{l_1'} u_2^{l_2'}) = u_1^{l_1+l_1'} u_2^{l_2+l_2'} u_3^{l_1' l_2} \tag{7.2}$$

and it is then easy to check that $\frac{|P^2|}{|P|} \to \infty$ as $L_1, L_2 \to \infty$, essentially because by varying l_1, l_1', l_2, l_2' within their ranges one can change $l_1 l_2'$ without changing $l_1 + l_1'$ or $l_2 + l_2'$. This can be thought of as an extra degree of freedom in P^2 compared to P. Coming back to commutators and recalling that $u_3 = [u_2, u_1]$, we see that this corresponds to the freedom to interchange the order of some of the u_1, u_2 in P^2, as seen in the LHS of (7.2). This is a freedom that the definition of ordered progression explicitly denies us.

It turns out that if we introduce this freedom to P as well then this does force P to have small tripling.

Definition. Given $x_1, \dots, x_r \in G$, $L_1, \dots, L_r \geq 0$, the **nonabelian progression** $P(x; L)$ of rank r is defined to consist of those elements of G expressible as products of $x_1^{\pm 1}, \dots, x_r^{\pm 1}$ in which each x_i, x_i^{-1} appear at most L_i times between them.

It turns out $P(u_1, u_2, L_1, L_2)$ does have small tripling (Example Sheet 2).

A note of caution: nonabelian progressions don't always have small tripling; consider $P(x_1, x_2; L_1, L_2)$ for x_1, x_2 generators of a nonabelian free group. In the case of $H(\mathbb{Z})$, the formula (7.1) is simplified by the fact that $u_3 = [u_2, u_1]$ is central in $H(\mathbb{Z})$. If this were not the case, we'd end up with elements of the form $[[u_2, u_1], u_1]$, for example. This is in fact a specific example of a property called nilpotence.

To define this, first define a normal series for a group G to be a sequence

$$G = G_1 > G_2 > \cdots$$

of normal subgroups $G_i \lhd G$, and a central series to be such a normal series in which each G_i/G_{i+1} is central in G/G_{i+1}.

Definition. A group G is **nilpotent** if there is a finite central series

$$G = G_1 > \cdots > G_{s+1} = \{1\}$$

The smallest s for which such a series exists is called the step or nilpotency class of G.

Exercise. $H(\mathbb{Z})$ is 2-step nilpotent.

8 Nilpotent Groups

Lecture 8

The reasons we focus on nilpotent groups are twofold: there is a clean generalization of Freiman-Green-Ruzsa to nilpotent groups, and a deep theorem of Breuillard, Green and Tao essentially reduces the general case to the nilpotent case.

Given $x_1, \dots, x_k \in G$, define the simple commutator $[x_1, \dots, x_k] = [x_1, \dots, x_k]_t$ recursively as follows:

$$[x_1] = x_1$$
$$[x_1, \dots, x_k] = [[x_1, \dots, x_{k-1}], x_k]$$

(recall $[x, y] = x^{-1}y^{-1}xy$).

Given subgroups $H, N < G$, define $[H, N] = \langle [h, n] \mid h \in H, n \in N \rangle$, and then given $H_1, \dots, H_k < G$, set

$$[H_1] = H_1$$
$$[H_1, \dots, H_k] = [[H_1, \dots, H_{k-1}], H_k].$$

Note that

$$[H, N] = [N, H] \tag{8.1}$$

since $[h, n] = [n, h]^{-1}$.

Lemma 8.1. Let $H_1, \dots, H_k, N \triangleleft G$, let S_i be a generating set of H_i for each i. Suppose $[s_1, \dots, s_k] \in N$ whenever $s_i \in H_i$ for each i. Then $[H_1, \dots, H_k] < N$.

Proof. Case $k = 1$ is trivial, so assume $k > 1$. If $[s_1, \dots, s_k] \in N$ for each $s_i \in S_i$ then we have $[[s_1, \dots, s_{k-1}], s_k] \in N$ for each $s_i \in S_i$, and hence $[s_1, \dots, s_{k-1}] \in C_{G/N}(H_k) = \{g \in G \mid [g, h] \in N \forall h \in H_k\}$. The centraliser of a normal subgroup is itself normal, so by induction we have $[H_1, \dots, H_{k-1}] \subset C_{G/N}(H_k)$, and hence $[H_1, \dots, H_k] \subset N$, as claimed. □

Definition. Given a group G, we define the **lower central series** $G = G_1 > G_2 > \cdots$ of G via

$$G_k := \langle [g_1, \dots, g_k] \mid g_i \in G \rangle.$$

Note that $G_k > G_{k+1}$, because $[g_1, \dots, g_{k+1}] = [[g_1, g_2], g_3, \dots, g_{k+1}]$. Also, since $[g_1, \dots, g_k]^h = [g_1^h, \dots, g_k^h]$, each G_k is normal in G (recall $x^y = y^{-1}xy$). The fact that this is a central series (i.e. G_k/G_{k+1} is central in G/G_{k+1}) follows from the following result.

Proposition 8.2. We have $G_{k+1} = [G_k, G]$ for every k. In particular, $G_k = [G, \dots, G]_k$.

Proof. First, $G_{k+1} < [G_k, G]$ by definition. The fact that $[G_k, G] \subset G_{k+1}$ follows from Lemma 8.1 since $[g_1, \dots, g_{k-1}]$ generate G_k and G_i, G_k, G_{k+1} are normal. □

Proposition 8.3. Let G be a group generated by S, then $G_k = \langle [s_1, \dots, s_k] G_{k+1} \mid s_i \in S \ \forall i \rangle$ 'G_k is generated mod G_{k+1} by simple commutators of generators'.

Proof. Note that $[s_1, \dots, s_k]^g \in [s_1, \dots, s_k] G_{k+1}$ by definition of G_{k+1}, so

$$\langle [s_1, \dots, s_k] G_{k+1} \mid s_i \in S \rangle$$

is normal in G. Moreover,

$$[s_1, \dots, s_k] \in \langle [t_1, \dots, t_k] G_{k+1} \mid t_i \in S \rangle$$

whenever $s_i \in S$ for every i, so Lemma 8.1 implies that $[G, \dots, G]_k \subset \langle [s_1, \dots, s_k] G_{k+1} \mid t_i \in S \rangle$. Proposition 8.2 gives $[G, \dots, G]_k = G_k$, so we have $G_k \subseteq \langle [s_1, \dots, s_k] G_{k+1} \mid s_i \in S \rangle$. The reverse inclusion is immediate. □

Proposition 8.4. We have $[G_i, G_j] \subset G_{i+j}$ for every i, j.

For this we'll need the following commutator identity, which you can check directly:

$$[x, y^{-1}, z]^y [y, z^{-1}, x]^z [z, x^{-1}y]^x = 1. \tag{8.2}$$

Proof of Proposition 8.4. The case $j = 1$ follows from Proposition 8.2. So we can assume $j > 1$ and, by induction,

$$[G_k, G_{j-1}] \subset G_{k+j-1} \forall k. \tag{8.3}$$

Now note that

$$[G_i, G_j] = [G_i, [G_{j-1}, G]] = [[G, G_{j-1}], G_i] \tag{8.4}$$

by Proposition 8.2 and (8.1). We also have

$$[[G_{j-1}, G_i], G] = [[G_i, G_{j-1}], G] \subset [G_{i+j-1}, G] = G_{k+j}. \tag{8.5}$$

by (8.1), (8.3) and Proposition 8.2, and

$$[[G_{i+1}, G], G_{j-1}] = [G_{i+1}, G_{j-1}] \subset G_{i+j}. \tag{8.6}$$

by Proposition 8.2 and (8.3). Given $x \in G$, $y \in G_{j-1}$ and $z \in G_i$, we therefore have

$$\begin{aligned} [x, y, z] &= (([y^{-1}, z^{-1}, x]^z, [z, x^{-1}, y^{-1}]^x)^{-1})^y && \text{by (8.2)} \\ &\subset G_{i+j} && \text{by (8.5) and (8.6).} \end{aligned}$$

The proposition follows from (8.4) and Lemma 8.1. □

Definition. Given a group G, the upper central series

$$\{1\} = Z_0(G) < Z_1(G) < Z_2(G) < \cdots$$

is defined recursively by setting $Z_{i+1}(G)$ so that $Z_{i+1}(G)/Z_i(G)$ is the centre of $G/Z_i(G)$. Note that each $Z_i(G)$ is normal by induction, since the centre of any group is normal.

Theorem 8.5. Let $G = H_1 > H_2 > \cdots > H_{i+1} = \{1\}$ be a finite central series for G (so G is nilpotent). Then we have $H_i > G_i \ \forall i = 1, \ldots, r+1$ and $H_{r+1-i} \subset Z_i(G)$ for every $0 \leq i \leq r$.

This justifies the names upper and lower central series:

$$\begin{array}{ccccccccc} Z_r(G) & & Z_{r-1}(G) & & & & Z_1(G) & & Z_0(G) \\ \vee & & \vee & & & & \vee & & \vee \\ H_1 & > & H_2 & > & \ldots & > & H_r & > & H_{r+1} \\ \vee & & \vee & & & & \vee & & \vee \\ G_1 & & G_2 & & & & G_r & & G_{r+1} \end{array}$$

Corollary 8.6. If G is s-step nilpotent, then both the upper and lower central series have length $s + 1$.

Proof. $H_i > G_1$ by definition, so we may assume $i > 1$, and then we have

$$\begin{aligned} H_i &\supset [H_{i-1}, G] && \text{central series} \\ &\supset [G_{i+1}, G] && \text{by induction} \\ &= G_i && \text{by Proposition 8.2.} \end{aligned}$$

We also have $Z_0(G) > H_{r+1}$ by definition, so we may assume $i > 0$ and, by inudction, that $H_{r+2-i} \subset Z_{i-1}(G)$. But that

$$G/Z_{i-1}(G) = \frac{G/H_{r+2-i}}{Z_i(G)/H_{r+2-i}}$$

because (H_j) is a central series, H_{r+1-i}/H_{r+2-i} is central in G/H_{r+2-i}, so its image in $G/Z_{i-1}(G)$ in the above quotient is central. But the centre of $G/Z_{i-1}(G)$ is $Z_i(G)/Z_{i-1}(G)$, so $H_{r+1-i} \subset Z_i(G)$, as required. □

These results say

- G is nilpotent of step $\leq s \iff G_{s+1} = \{1\} \iff Z_s(G) = G$
- If $G = \langle s \rangle$, we can verify this just by checking that $[t_1, \ldots, t_{s+1}] = 1$ for all $t_i \in S$.
- If G is nilpotent of step $\leq s$, then any commutator like $[[[q_1, q_2], q_3], [q_4, q_5]]$ with more than s entries is trivial.

9 Torsion-free nilpotent approximate groups

Definition. If $\langle x_1, \dots, x_r \rangle$ is s-step nilpotent then the nonabelian progression $P(x; L)$ is called a **nilprogression** of rank r and step s. In this case, we write $P_{\mathrm{nil}}(x; L)$ instead of $P(x; L)$.

Index