Introduction to Approximate Groups

February 12, 2019

Contents

0 Introduction

—Lecture 1—

Example classes: 12 Feb, 5 Mar, 2-3pm.
Venue will be confirmed later.
Examinable material is exactly what is on board (as usual).
No plan for printed notes (as usual).

After 3.5 years of tripos, we finally know what a subgroup is: a *subgroup* $H < G$ is a non-empty set closed under products and inverses.
We would then expect an *approximate subgroup* to be a subset that is only *approximately* closed under products. We'll make this precise soon. Such sets arise naturally in a number of branches of maths, and as such approximate groups have had broad range of applications. In this course we'll 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 CS).[1]

[1] Lecturer actually spent the time to write all this on board, probably implying a slower-paced course than category theory! But every course is slower than category theory.

1 Small Doubling

To start with, we'll look at a preliminary notion of approximate closure called *small doubling*.
In this course, G is always a group, arbitrary unless specified otherwise. Given $A, B \subset G$, we write

$$AB = \{ab : a \in A, b \in B\}$$

called the product set,

$$A^n = \underbrace{A \cdot ... \cdot A}_{n \text{ times}}$$

$$A^{-1} = \{a^{-1} : a \in A\}$$

$$A^{-r} = (A^{-1})^n$$

When G is abelian, we often switch to additive notion, for example $A+B, nA, -A, -nA$ in place of the above (*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 $|A^2|$ 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 we have $|A| \leq |A^2| \leq |A|^2$, and 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.[2] Therefore, we can view sets A satisfying

$$|A^2| = o(|A|^2) \quad (1.1)$$

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

$$|A^2| \leq K|A| \quad (1.2)$$

for some constant $K \geq 1$ fixed a priori.
Obviously, such sets are very far from random, and we can expect (1.2) to impose a significant restriction on A. The main aim of the course is to work out how significant the restriction is.

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

[2] We will see in sheet 1 that random A have quadratic size for $|A^2|$, in the sense of: if A is a set of size n chosen uniformly at random from $\{1, ..., n^{100}\}$ (we'd like to choose from all integers, but there's no uniform measure there), then $\mathbb{E}(|A+A|)$ is *closed to* $\frac{1}{2}|A|^2$. Note that this is the largest we can get, since in an abelian group we have $a + b = b + a$.

Example. • A a finite subgroup;
• $|A| \leq K$;
• $A \subset \mathbb{Z}, A = \{-n, ..., n\}$, $|A + A| \leq 2|A|$.

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

Our main aim will be to prove theorems along the lines of: if A has small doubling, then A has a certain structure. When K is very small, this is quite easy, as follows.

Theorem. (1.1, Freiman[3])
Let $K < \frac{3}{2}$. Suppose $A \subset G$ and $|A^2| \leq K|A|$ (by writing like this, we're obviously assuming A is finite). Then there is a subgroup $H < G$ with $|H| = |A^2|$ $(\leq K|A|)$ such that $A \subset aH = Ha \forall a \in A$.
(i.e., A is a large portion of a coset of a finite subgroup).

Remark. Converse: if $A \subset xH = Hx$ for $x \in G$, $H < G$, $|H| \leq K|A|$, then $|H^2| \leq K|A|$. So this is a complete classsification 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 $> \frac{1}{2}|A|$ pairs of $(x, y) \in A \times A$ s.t. $ax = by$, i.e. $a^{-1}b = xy^{-1}$. This immediately implies $A^{-1}A \subset AA^{-1}$ (as we just needed one such pair (x, y)), and replacing A by A^{-1} we get the other inclusion; so $A^{-1}A = AA^{-1}$ as required.

Since $|A \times A| = |A|^2$, it also implies that $|A^{-1}A| \leq \frac{|A|^2}{\frac{1}{2}|A|} = 2|A|$, 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 exist $> \frac{1}{2}|A|$ pairs $(u, v) \in A \times A$ s.t. $c^{-1}d = uv^{-1}$. This means for at least one pair (x, y) 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 $A^2 = aHa \ \forall a \in A$ (H as before). In particlar, $|H| = |A^2|$.

Proof. First, note that $A \subset aH \cap Ha$ (label this 1.3) by definition of H and $A^{-1}A = AA^{-1}$, so certainly $A^2 \subset aHa$. For the reverse inclusion, let $z \in aHa$. Since H is a subgroup, there exists $|H|$ pairs $(x, y) \in aH \times Ha$ s.t. $z = xy$. Moreover, by (1.3) and the bound $|H| < 2|A|$ from lemma (1.2), more than half of those x and half of those 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. □

[3]The proof presented here is by Tao instead of the original version. Lecturer has no way to tell if the original proof is close to this, especially because the original proof is in Russian.

Proof of theorem (1.1):

Proof. 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, etc.

2 Covering and Higher Sum and Product Sets

—Lecture 2—

Today we will 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, Rusza)
Suppose $A \subseteq \mathbb{F}_p^r$ satisfies $|A+A| \leq K|A|$. Then there is a subgroup $H \subseteq \mathbb{F}_p^r$ with $|H| \leq p^{K^4}K^2|A|$ s.t. $A \subseteq H$.

Remark. It's not ideal that $|A|/|H|$ depends on p. We'll remove this dependency in a few lectures' time.

We'll start by proving the following weaker version:

Proposition. (2.2)
Suppose $A \subseteq \mathbb{F}_p^r$ satisfies $|2A-2A| \leq K|A|$. Then there is a subgroup $H \subseteq \mathbb{F}_p^r$ with $|H| \leq p^K|A-A|$ $(\leq p^K K|A|)$ such that $A \subseteq H$.

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

Lemma. (2.3, Rusza's Covering Lemma)
Suppose $A, B \subseteq G$, and $|AB| \leq K|B|$. Then, $\exists X \subseteq A$ with $|X| \leq K$ s.t. $A \subseteq XBB^{-1}$. Indeed, we may take $X \subseteq A$ maximal such that the sets xB $(x \in X)$ are disjoint.
The term *covering* refers to the conclusion $A \subseteq XBB^{-1}$, which says A *can be covered by few left-translates of* BB^{-1}.

Proof. First, disjointness of $xB \implies |XB| = |X||B|$. Since $X \subseteq A$, $|XB| \leq |AB| \leq K|B|$, so $|X| \leq K$. Maximality implies $\forall a \in A$, $\exists x \in X$ s.t. $aB \cap xB \neq \phi$; and hence $a \in xBB^{-1}$. So $A \subseteq XBB^{-1}$. □

Now we prove some lemmas for proposition 2.2:

Lemma. (2.4)
Suppose $A \subseteq G$ satisfies $|A^{-1}A^2A^{-1}| \leq K|A|$. Then $\exists X \subseteq A^{-1}A^2$, $|X| \leq K$ such that $A^{-1}A^n \subseteq x^{n-1}A^{-1}A \forall n \in \mathbb{N}$.

Proof. Lemma 2.3 gives (as $|A| = |A^{-1}|$) the existence of an $X \subseteq A^{-1}A^2$, $|X| \leq K$ s.t. $A^{-1}A^2 \subseteq XA^{-1}A$. We then have

$$\begin{aligned} A^{-1}A^n &= A^{-1}A^{n-1}A \\ &\subseteq X^{n-2}A^{-1}A^2 \text{ by induction} \\ &\subseteq X^{n-1}A^{-1}A \end{aligned}$$

□

Proof. (of proposition 2.2)
Note that $\mathbb{F}_p^r$ is abelian, so we can apply Lemma 2.4 which gives the existence of X, $|X| \leq K$ s.t. $nA - A \subseteq (n-1)X + A - A \forall n \in \mathbb{N}$. As we are in a finite vector space, this means that $\langle A \rangle \subseteq \langle X \rangle + A - A$, so $|\langle A \rangle| \leq |\langle X \rangle||A - A| \leq p^K K|A|$, as claimed. □

To strenghthen prop 2.2 to theorem 2.1, we use the technique of bounding higher sum/product sets. The key result, at least in the abelian case, is the following:

Theorem. (2.5, Plünecke-Rusza)
Suppose $A \subseteq G$ (abelian) and $|A + A| \leq K|A|$, then $|mA - nA| \leq K^{m+n}|A|$ $\forall m, n \geq 0$.
This was proven intro to discrete analysis last term. We won't redo the whole proof, but we will reprove some parts of it.

Proof. 2.5 gives that $|2A - 2A| \leq K^4|A|$, and $|A - A| \leq K^2|A|$. Then immediately it follows from prop 2.2. □

We'll spend the rest of teh lecture discussing theorem 2.5 and variants of it. We've seen that it is useful, at least in one context. To see more properly why it's useful, let's think about what the genuine closure of subgroups under products and inverse means. One useful feature is that it can be iterated: given $h_1, h_2, ... \in H$, a subgroup, this means that $h_1^{\varepsilon_1}...h_m^{\varepsilon_m} \in H \ \forall \varepsilon_i = \pm 1 \ \forall m \forall h_i \in H$. Then theorem 2.5 allows us to *iterate the approximate closure* of a set of small doubling: $a_1 + ... + a_m - a_1' - ... - a_n'$ may not belong to A, but at least belongs to $mA - nA$ which is (a) not too large, and (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 doesn't hold for non-abelian groups.

Example. (2.6)
Let x generate an infinite cyclic group $\langle x \rangle$, and H be a finite group. Set $G = H * \langle x \rangle$, the free product (just keep in mind that $x^{-1}Hx \neq H$). Set $A = H \cup \{x\}$, then $A^2 = H \cup xH \cup Hx \cup \{x^2\}$, 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 strenghten small doubling slightly, we can recover a form of theorem 2.5. One way is to replace it with *small tripling*, i.e. $|A^3| \leq K|A|$:

Proposition. (2.7)
Suppose $A \subseteq G$, $|A^3| \leq K|A|$. Then $|A^{\varepsilon_1}...A^{\varepsilon_m}| \leq K^{3(m-2)}|A|$, $\forall \varepsilon_i = \pm 1, \forall m \geq 3$.

The key ingredient is the following:

Lemma. (2.8, Rusza's Triangle Inequality)
Given $U, V, W \subseteq G$ all finite, $|U||V^{-1}W| \leq |UV||UW|$.

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 $W(x) \in W$, s.t. $x = V(x)^{-1}W(x)$. Set $\varphi(u, x) = (uV(x), uW(x))$. To see injectivity, first note that $(uV(x))^{-1}(uW(x)) = x$, so x is determined by $\varphi(u, x)$, then $(uV(x))(v(x)^{-1}) = u$, so u is also determined by $\varphi(u, x)$. □

Proof. (of 2.7)
First, we'll do the case $m = 3$.
• $|A^3| = |A^{-3}| \leq K|A|$;
• apply lemma 2.8 with $U = W = A$, $V = A^2$. Get $|A||A^{-2}A| \leq |A^3||A^2| \leq K^2|A|^2$, so $|S^{-2}A| \leq K^2|A|$; • note that $(A^{-2}A)^{-1} = (A^{-1}A^2)$, so $|A^{-1}A^2||A^{-2}A| \leq K^2|A|$;
• replace A by $A-1$ to get $|AA^{-2}| = |A^2A^{-1}| \leq K^2|A|$;
• Finally, apply lemma 2.8 with $U = V = A, W = AA^{-1}$, we get $|A||A^{-1}AA^{-1}| \leq |A^2||A^2A^{-1}| \leq K^3|A|^2$. So, $|A^{-1}AA^{-1}| \leq K^3|A|$. For the last case, swap $A \leftrightarrow A^{-1}$.

For $m \geq 4$, lemma 2.8 gives $|A||A^{\varepsilon_1}...A^{\varepsilon_m}| \leq |AA^{-\varepsilon_2}A^{-\varepsilon_1}||AA^{\varepsilon_3}...A^{\varepsilon_m}| \leq K^3|A|K^{3(m-3)}|A|$ by induction. □

3 Approximate groups

—Lecture 3—

Last time we saw that assuming small tripling instead of small doubling allowed us to control higher product sets of the form $A^{\varepsilon_1}...A^{\varepsilon_m}$. In this lecture, we'll see another possible strengthening of small doubling. We also saw, in the proofs of theorem 2.1 and proposition 2.2, that the advantage of having a *covering* condition in place of a size bound. This motivates the following definition:

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

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, we can write A^m in stead of $A^{\varepsilon_1}...A^{\varepsilon_m}$, and $1 \in A \implies A \subset A^2 \subset A^3 \subset ...$, 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 definition of approximate group, in fact we have $A^m \subset X^{m-1}A$:

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

□

Another advantage is that if $\pi : G \to H$ is a homomorphism and A is a K-approximate group, then $\pi(A)$ is also trivially a K-approximate group (although we'll see that there exists 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 exists $O(K^{12})$-approximate groups 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 part is just lemma 3.1. For the converse, set $\hat{A} = A \cup \{1\} \cup A^{-1}$, and note that $A^2 = \{1\} \cup A \cup A^{-1} \cup A^2 \cup A^{-1}A \cup AA^{-1} \cup A^{-2}$. Each set in this

union has size $\leq K^2|A|$, by prop 2.7, so $|B| \leq 7K^3|A|$, as claimed. Similarly,

$$A^4 = \bigcup_{\varepsilon_i=\pm 1, 0 \leq m \leq 4} A^{\varepsilon_1}...A^{\varepsilon_m}$$

and the sets in this union have size $\leq K^6|A|$. It follows that $|\hat{A}^4| \leq O(K^6)|\hat{A}|$. Lemma 2.4 then tells us that $\exists X \subset G$, $|X| \leq O(K^6)$ s.t. $\hat{A}^n \subset X^{n-2}\hat{A}^2$ for every $n \geq 2$.
In particular, $|X^2| \leq O(K^{12})$, and $\hat{A}^4 = (\hat{A}^2)^2 \subset X^2\hat{A}^2$, so $\hat{A}^2$ is an $O(K^{12})$-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 (but not tripling)? 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 $\exists U \subset A$ with $|U| \geq \frac{1}{K}|A|$ s.t. $|U^m| \leq K^{m-1}|U|$ $\forall m \in \mathbb{N}$.

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

Tao proved a version of theorem 3.3 when he introduced the definition of approximate groups. We'll 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 every finite $C \subset G$, we have $|CUB| \leq R|CU|$.

Proof. It's trivial if $C = \phi$, so we may assume $\exists x \in C$. Defining $C' = C\{x\}$, we may also assume by induction that $|C'UB| \leq R|C'U|$.
Set $W = \{u \in U : xu \in C'U\}$. Then $CU = C'U \cup (xU \setminus xW)$ is a disjoint union, so in particular

$$|CU| = |C'U| + |U| - |W| \quad (3.1)$$

We also have $xWB \subset C'UB$ by definition of W, so $CUB \subset C'UB \cup (xUB \setminus xWB)$, and hence

$$|CUB| \leq |C'UB| + |UB| - |WB| \quad (3.2)$$

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

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

□

Proof. (of 3.3)
Set $U \subset A$ to be non-empty, minimizing $|UA|/|U|$, and write $R = |UA|/|U|$, noting that $R \leq K$ by minimality (since U must have beaten A). Also, U non-empty implies $|UA| \geq |A|$, 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 $|U^m| \leq K^{m-1}|U|$. □

Non-examinable: 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. (3.5, Tao, Petridis)
If $|A^2| \leq K|A|$, and $|AxA| \leq K|A| \ \forall x \in A$, then $|A^m| \leq K^{O(m)}|A| \ \forall m \geq 3$.

Also non-examinable (it's in intro to DA, but lecturer thought it's was showing):

Theorem. (2.5')
After writing a few words, lecturer decided that there's not enough time, and we have written enough this lecture, so wiped everything out from the board.

Check example sheets at *tointon.neocities.org*; they'll be up some time this afternoon.

4 Stability of approximate closure under basic operations

—Lecture 4—

There's a misprint in sheet 1 q7, but it has been corrected last friday.

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 $\pi(H) < \Gamma$;
• 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.

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

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

for all $m \in \mathbb{N}$.
In particular, if $|A^3| \leq K|A|$, then $|\pi(A)^3| \leq K^9|\pi(A)|$ by proposition 2.7.

We'll prove this using an argument of Helfgolt. 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 $|A^{-1}A \cap H| > |A \cap xH|$.

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 lecture will have familiar/trivial analogues for genuine subgroups.

Lemma. (4.3)
Let $H < G$, write $\pi : G \to G/H$ for the quotient map, let $A \subset G$ be finite.
Then $|A^{-1}A \cap H| \geq |A|/|\pi(A)|$.
Note that H is not assumed normal, so G/H is just the *space* of left cosets of xH, not necessarily a group.

Proof. By pigeonhole principle there exists $x \in G$ s.t. $|A \cap xH| \geq |A|/|\pi(A)|$.
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 $\varphi(x)$ s.t. $\pi(\varphi(x)) = x$. Then the cosets of $\varphi(x)H$ for $x \in \pi(A^m)$ are all distinct by definition, so in particular,

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

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

Proof. (of 4.1)
Write $H = \ker \pi$. Lemma (4.4) gives $|\pi(A^m)| \leq \frac{|A^{m+2}|}{|A^2 \cap H|}$. Lemma 4.3 gives $|A^2 \cap H| \geq |A|/|\pi(A)|$ (?). The proposition follows from combining these two 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, containing 1. Let $H \leq G$. Then

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

In particular, by prop 2.7, if $|A^3| \leq K|A|$, then

$$|(A^m \cap H)^3| \leq K^{9m}|A^m \cap H|$$

$\forall m \geq 2$.

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

Proof. Lemma 4.4 gives $|A^m \cap H| \leq |A^{m+1}|/|\pi(A)|$, where $\pi : G \to G/H$ is the quotient map as before).
Lemma 4.3 gives $|A^2 \cap H| \geq |A|/|\pi(A)|$.
The proposition follows from combining 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 $|X| = K^{m-1}$ s.t. $A^m \subset XA$. In particular, $A^m \cap H \subset \cup_{x \in X}(xA \cap H)$.

For each $xA \cap H$ non-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 intro to discrete analysis last term, you saw that when studying small doubling/tripling, there is a more general notion of homomorphism that comes into play – the Freiman homomorphism. To motivate this, consider two sets $A = \{-n, ..., n\} \subset \mathbb{Z}/p\mathbb{Z}$, and $B = \{-n, ..., n\} \subset \mathbb{Z}/q\mathbb{Z}$ for p, q two large primes $\geq 10n$, say. These two sets are intuitively *isomorphic* from perspective of $A+A$, or $B + B$, but there is no way of encoding this with a group homomorphism $\mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/q\mathbb{Z}$.

Definition. (Freiman homomorphism)
Let $m \in \mathbb{N}$, let A, B be subsets of groups. Then a map $\varphi : A \to B$ is a *Freiman m-homomorphism* if for $x_1, ..., x_m, y_1, ..., y_m \in A$ with $x_1...x_m = y_1...y_m$, we have $\varphi(x_1)...\varphi(x_m) = \varphi(y_1)...\varphi(y_m)$.
If $1 \in A$ and $\varphi(1) = 1$, then we say that φ is *centred*. 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*.

Remark. 1) Every map is trivially a 1-homomorphism, so we only care about $m \geq 2$.
2) This notion gets stronger as m increases: we may assume $A \neq \phi$, and then, picking $a \in A$ arbitrarily, if $x_1...x_k = y_1...y_k$ for $k \leq m$, then

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

3) If φ is centred and $a, a^{-1} \in A$, then exercise to check $\varphi(a^{-1}) = \varphi(a)^{-1}$ ($m \geq 2$). When we say φ is a Freiman homomorphism, we mean it is a 2-homomorphism.

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 equality holds.

Proof. Exercise. □

—Lecture 5—

Lemma. (4.8)
Let $A \subset G$ be a K-approximate group, suppose $\varphi : A^3 \to \Gamma$ is a centred ($1 \in A$ and $\varphi(1) = 1$) Freiman 2-homomorphism. Then $\varphi(A)$ is also a K-approximate group.

Proof. $\exists X \subset G$, $|X| \leq K$ s.t. $A^2 \subset XA$. So given $a_1, a_2 \in A$, $\exists x \in X, a_3 \in A$ s.t. $a_1 a_2 = x a_3$. 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

Today we will introduce some non-trivial examples of sets of small doubling in abelian groups.

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

$$P(x; L) = P(x_1, ..., x_r; L_1, ..., L_r) = \{l_1x_1 + ... + l_rx_r : |l_i| \leq L_i \forall i\}$$

is called a *progression of rank* r. If in addition $H \subset G$ 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 $\mathbb{Z}^r$. For example, if $G = \mathbb{Z}$ and $r = 2$:
(diagram)

It's easy to see that such a box B in $\mathbb{Z}^r$ is a 2^r-approximaate group, e.g. in $r = 2$:
(diagram)

Hence $P(x; L)$ is also a 2^r-approximate group, as is $H_P(r; L)$.

Remarkably, these are essentially the only examples:

Theorem. (5.1, Frieman ($G = \mathbb{Z}$); Green-Rusza (arbitrary abelian G))
Suppose $A \subset G$ (abelian) satisfies $|A + A| \leq K|A|$. Then there exists a coset progression $H + P$ of rank $\leq O(K^{O(1)})$ s.t. $A \subset H + P \subset O(K^{O(1)})(A \cup \{0\} \cup -A)$.
In particular, theorem 2.5 gives that $|H + P| \leq \exp(O(K^{O(1)}))|A|$. So A is a large proportion of $H + P$.

A substantial part of this result was proved in intro to discrete analysis, but with a slightly less explicit version of coset progressions:

Definition. (Bohr set)
Let G be a finite abelian group. Let $\Gamma = \{\gamma_1, ..., \gamma_r\} \subset \hat{G} = Hom(G, \mathbb{R}/\mathbb{Z})$, and let $\rho \in [0, \frac{1}{2}]$. Then the set

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

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'll see in sheet 1 that $B(\Gamma, \rho)$ is a 4^r-approximate group, whereas progressions were homomorphic images of boxes, $B(\Gamma, \rho)$ is the pullback(??) of $[-\rho, \rho]^r$ under $(\gamma_1, ..., \gamma_r) = \hat{G}^r$.

It turns out that the notions of coset progression and Bohr set are essentially equivalent. In sheet 2, we'll see that every coset progression is a Freimana 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'll see a special case of that shortly.

Proposition. (5.2, from intro to DA)
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 $|\Gamma| \leq O(K^{O(1)})$, some $\rho \geq \frac{1}{O(K^{O(1)})}$, and a centred Freiman 2-isomorphism $\varphi : B(\Gamma, \rho) \to B$.
($2A-2A$ *contains a large set isomorphic to a Bohr set of bounded rank*).
In intro to DA you saw this in the special case of G torsion-free. The general case is harder, but nonetheless conceptually very similar, so we'll assume this result from now on.

To pass from prop 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, $\rho < \frac{1}{10}$[4]. Then there exists a coset progression $H+P \subset B(\Gamma, \rho)$ with rank r and $|H+P| \geq (\rho/r)^r|G|$.

We'll prove prop 5.3 in the next couple of lectures.

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

Proof. Exercise: if H is a group and $\varphi : H \subset G$ is a centred Freiman 2-homomorphism, then φ is also a group homomorphism. In particular, in this lemma, $\varphi(H)$ is a finite subgroup. Therefore it's sufficient to show that $\varphi(H + P(x; L)) = \varphi(H) + P(\varphi(x_1), ..., \varphi(x_r); L_1, ..., L_r)$.
In fact, we'll show that $\forall h \in H$, $|l_i| \leq L_i$, we have

$$\varphi(h + l_1x_1 + ... + l_rx_r) = \varphi(h) + l_1\varphi(x_1) + ... + l_r\varphi(x_r) \quad (5.1)$$

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

$$\begin{aligned}\varphi(h + l_1x_1 + ... + l_rx_r) &= \varphi(h + l_1x_1 + ... + l_rx_2) + \varphi(0) \\ &= \varphi(h + l_1x_1 + ... + (l_j - 1)x_j + ... + l_rx_r) + \varphi(x_j)\end{aligned}$$

So lemma follows by induction on $\sum_i l_i$. □

Proof. (of theorem 5.1)
By proposition 5.2 and 5.3 and lemma 5.4, $\exists H + P$ coset progression of rank $\leq O(K^{O(1)})$ s.t. $H + P \subset 2A - 2A$ and $|H + P| \geq \exp(-O(K^{O(1)}))|A|$.
We'll now apply a version of Ruzsa's covering lemma due to Chang. Define recursively sets $S_1, S_2, ... \subset A$ s.t. S_i is a maximal subset of size $\leq 2K$ s.t. the

[4]Lecturer thinks he might also get away with a larger choice.

translates $x + S_{i-1} + ... + S_1 + H + P$ are all disjoint. If ever $|S_i| < 2K$, we stop. Now suppose we get as far as $S_1, ..., S_t$. Then

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

So prop 2.5 $\implies |S_t + ... + S_1 + H + P| \leq K^{4+t}|A|$.
On the other hand, disjointness of the translates in the definition of S_i means that

$$|S_t + ... + S_1 + H + P| \geq |S_t|...|S_1||H + P| \geq (2K)^{t-1} \exp(-O(K^{O(1)}))|A|$$

Putting these together, we have $2^{t-1} \leq K^5 \exp(O(K^{O(1)}))$, hence $t \leq O(K^{O(1)})$.
In particular, the process terminates, at S_t say.
But also, since S_t is therefore maximal among all subsets of A s.t. $x + S_{t-1} + ... + S_1 + H + P$ are disjoint for $x \in S_t$, Ruzsa's covering lemma from lecture 2 implies that

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

Enumerating $\bigcup_i S_i$ as $s_1, ..., s_d$, we have $d \leq O(K^{O(1)})$, and

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

as claimed (subject to proving 5.3). □

Exercise: see what bounds you get if you apply Ruzsa's covering lemma directly, instead of Chang's argument.

—Lecture 6—

Proposition. Let G be a finite abelian group. Suppose $\Gamma \subseteq \hat{G}$, $|\Gamma| = r$, and let $\rho < 1/2$. Then $\exists$ coset progression of $H + P \subseteq B(\Gamma, \rho)$ (Bohr set, see previous lecture) of rank r with $|H + P| \geq (\rho/r)^r |H|$.

To prove this, we'll use a field called geometry of numbers, which is concerned with lattices in $\mathbb{R}^d$.
For us, a lattice $\Lambda \subseteq \mathbb{R}^d$ will simply be the additive subgroup (not *subspace*) generated by some basis $x_1, ..., x_d$ for $\mathbb{R}^d$. So $\Lambda = \{l_1 x_1 + ... + l_d x_d | l_i \in i\}$. If $\Gamma \subset \Lambda$ is another lattice, then we'll call it a sublattice, and just write $\Gamma \subset \Lambda$. It is an exercise (in sheet 2) to check that if $\Gamma \subset \Lambda$ with basis $y_1, ..., y_d$, say, then $\det(y_1, ..., y_d)/\det(x_1, ..., x_d) = [\Lambda : \Gamma]$. In particular, two different bases of the same lattice have the same determinant. So we can define this to be $\det \Lambda$.

The relevance of lattices to 5.3 is the following:

6 Geometry of Numbers

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

Proof. Λ is finitely generated as G is finite, and torsion-free as it's 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 $k = d$, $span_{\mathbb{R}}(\Lambda) = \mathbb{R}^d$. So we can take a generating set for Λ of size d, which is then a basis for $\mathbb{R}^d$. The determinant follows by above exercise, as $\det(\mathbb{Z}^d) = 1$. □

We'll investigate the intersection of $[-\rho, \rho]$ with Λ. To do this, we introduce some definitions:

Definition. A set $A \subseteq \mathbb{R}^d$ is *convex* if $\forall x \in \mathbb{R}^d \setminus A^0$, $\exists$ hyperplane h_x with $x \in h_x$ and $h_x \cap A^0 = \phi$ (A^0 is the interior of A).

Definition. A set $B \subseteq \mathbb{R}^d$ is a *convex body* if it is bounded, convex and $B^0 \neq \phi$. It is *symmetric* if $\forall x \in B$, we have $-x \in B$.

Given a symmetric body B and a lattice Λ, define the *successive minima* $\lambda_1 \leq ... \leq \lambda_j$ of B w.r.t. Λ as

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

We may then inductively define $v_1, ..., v_d \in \Lambda$, s.t. $v_1, ..., v_i \in \lambda_i \bar{B}$. We'll call this a *directional basis* for Λ w.r.t. B. Note that this is not unique, and not necessarily a basis for Λ in the earlier sense.

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

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

Proof. Fix a basis $x_1, ..., x_d$ for Λ. Define the *fundamental parallelopiped* P w.r.t. $x_1, ..., x_d$ via $P = \{l_1 x_1 + ... + l_d x_d | l_i \in [0, 1)\}$. Since $x_1, ..., x_d$ is a basis for $\mathbb{R}^d$, $\forall v \in \mathbb{R}^d$, there is a unique $x_v \in \Lambda$ and $p_v \in P$ s.t. $v = x_v + p_v$.
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. So $vol(A) = vol(\varphi(A)) \leq vol(P) = \det(\Lambda)$. □

Proof. (of 6.2)
Let $v_1, ..., v_d$ be a directional basis for Λ w.r.t. B. Set $V_i = span_{\mathbb{R}}(v_1, ..., v_i)$ ($v_0 = \{0\}$), and set $\Lambda_i = \Lambda \cap (V_i \setminus V_{i-1})$, then Λ is a disjoint union $\bigcup_{i=0}^{d} \Lambda_i$.

Claim 1: We have $\lambda_d B^0 \cap (\lambda_d B^0 + \alpha x) = \phi$ whenever $x \in \Lambda$, and $\alpha \geq 2\lambda_d/\lambda_j$.
Proof of claim: given $x \in \Lambda_j$, by definition, $x \notin \lambda_j B^0/$ So by convexity, there exists hyperplane h_x s.t. $x \in h_x$ and $h_x \cap A_j B^0 = \phi$.
By symmetry, we can take $h_{-x} = -hx$. But note that $-hx = hx - 2x$. This means that $\lambda_j B^0$ is contained in the slice of space S_x between the two parallel hyperplanes h_x and $h_x - 2x$. Clearly $S_x \cap (S_x + \alpha_x) = \phi \; \forall \alpha \geq 2$, so in particular $\lambda_j B^0 \cap (\lambda_j B^0 + \alpha x) = \phi$ for all such α as well. Scaling by λ_i/λ_j, then $\lambda_d B^0 \cap (\lambda_d B^0 + \alpha x) = \phi$ whenever $x \geq 2\lambda_d/\lambda_j$.

Claim 2: there exist sets $B_1 \subset ... \subset B_d = \lambda_i B^0$ s.t.
(1) $vol(B_i) = (\lambda_i/\lambda i + 1)^i vol(B_{i+1}) \forall i$;
(2) $B_i \cap (B_i + \alpha x) = \phi$ whenever $x \in \Lambda_i$ and $\alpha \geq 2\max\{\lambda_i/\lambda_j, 1\}$. Proof of claim: Define operations $\gamma_1, ..., \gamma_{d-1}$ on suitable subsets of $\mathbb{R}^d$ as follows: given L bounded and open, define γ_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 $\varphi(z + v) = z + \frac{\lambda_i}{\lambda_{i+1}} v \forall v \in V_i$. Note the following properties:
(i) $vol(\gamma_i(L)) = (\lambda_i/\lambda_{i+1})^i vol(L)$ by Fubini;
(ii) If $L \cap (z + V_i)$ is open and convex $\forall z$, then $\gamma_i(L) \subset L$, because $z \in L$;
(iii) If $L \cap (z + V_i)$ is open and convex, then so is $\gamma_i(L) \cap (z + V_i)$, and indeed so is $\gamma_i(L) \cap (z + V_j \forall j < i)$.
Set $B_d = \lambda_d B^0$, 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) f0r $(i + 1)$, because σ_i scaleds by λ_i/λ_{i+1} in direction x. For $i < j$, it follows from $B_i \subseteq B_{i+1}$.

Now $vol(B_1) = \lambda_1 ... \lambda_d vol(B)$, and by (2), $a - b \notin 2\Lambda \; \forall a, b \in B_1$, so Blichfeldt gives that $vol(B_1) \leq 2^d \det(\Lambda)$. □

—Lecture 7—

Example class Tuesday 1400-1530 MR14. Work to lecturer at pembroke plodge by midday monday.

Some errata for last lecture (correct directly later):
(1)When we had σ_i and $B_1 \subset B_2 \subset ...$ last time, we assume z is the centre of mass of $L \cap (z + V_i)$. To be on safe side, in statement of Minkowski's second theorem, let's assume that B is a polytope (?).
(2) the assumption $h_x \cap A^0 = \phi$ is not sufficient; we also need A^0 is contained in one of the two half spaces into which h_x divides $\mathbb{R}^d$.

Proof. (of 5.3)
Write $\gamma = (\gamma_1, ..., \gamma_r) \in \hat{G}^r$. Define $\Lambda = \gamma(G) + \mathbb{Z}^r$, which is a lattice of determinant $|\ker \gamma|/|G|$ by lemma 6.1. Let $\lambda_1, ..., \lambda_r$ be the successive minima $[-1, 1]^r$ w.r.t. Λ, and $v_1, ..., v_r$ a directional basis. Set $L_i = \lfloor \frac{\rho}{r\lambda_i} \rfloor$ for each i. Then $P(v_1, ..., v_r; L_1, ..., L_r) \subset [-\rho, \rho]^r$ (recall LHS is formula for a coset progression – see start of chapter 5). Pick, for each i, $x_i \in G$ s.t. $\gamma(x_i) = v_i$, and set $H = \ker \gamma$, write $P = P(x_1, ..., x_r; L_1, ..., L_r)$. Then $H + P \subset B(\Gamma, \rho)$.
Now we just need to justify the size constraint. We claim that if $l_1, ..., l_r$ and

$l'_1, ..., l'_r$ satisfy $|l_i|, |l'_i| \leq L_i$, and $l_1x_1 + ... + l_rx_r \in l'_1x_1 + ... + l'_rx_r + H$ (call this (*)), then in fact $l_i = l'_i \forall i$. Indeed, (*) implies that $(l_1 - l'_1)v_1 + ... + (l_r - l'_r)v_r \in \mathbb{Z}^r \in \mathbb{Z}^r \cap [-2\rho, 2\rho]^r$. But $\rho < \frac{1}{2}$, this last intersection is just $\{0\}$. Also v_i are linearly independent, so $l_i = l'_i$ for each i.
Then

$$\begin{aligned} |H + P| &\geq |H|(L_1 + 1)...(L_r + 1) \\ &\geq |H|(\frac{\rho}{r})^r \frac{1}{\lambda_1 ... \lambda_r} \\ &\geq |G|(\rho/r)^r \end{aligned}$$

by Minkowski's second theorem. □

7 Progressions in the Heisenberg group

We know the Heisenberg group is

$$H(\mathbb{Z}) = \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix}$$

where $\mathbb{Z}$ denotes any integer.
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:

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

(just by doing matrix multiplication). 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 group 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 of u_1 and u_2.

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

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

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

(almost the same as the abelian case, but now we care about the order).
Now consider $P = P_{ord}(u_1, u_2; L_1, L_2)$ for $u_1, u_2 \in H(\mathbb{Z})$ as before, and $L_1, L_2 \geq 0$. We have $(u_1^{l_1} u_2^{l_2})(u_1^{l_1'} u_2^{l_2'}) = u_1^{l_1+l_1'} u_2^{l_2+l_2'} u_3^{l_1'l_2}$ (7.2), and it is then easy to check that $|P^2|/|P| \to \infty$ as $L_1, L_2 \to \infty$, essentially because we can change l_1, l_2, l_1', l_2' without changing $l_1 + l_1'$ and $l_2 + l_2'$ but changing $l_1'l_2$, so in some we have an extra degree of freedom in P^2 compared to in 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. (nonabelian progression)
Given $x_1, ..., x_r \in G$, $L_1, ..., L_r \geq 0$, the *nonabelian progression* $P(x; L) = P(...)$ of rank r is defined to consist of all those elements of G expressible as products of $x_1^{\pm 1}, ..., x_r^{\pm r}$, in which each x_i, x_i^{-1} appear at most L_i times between them.

It turns out that $P(u_1, u_2; L_1, L_2)$ *does* have small tripling (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 $G = G_1 > G_2 > ...$ of normal subgroups $G_i \triangleleft G$ 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. (nilpotent group)
A group G is nilpotent if $\exists$ finite central series $G = G_1 > ... > G_{s+1} = \{1\}$. The smallest s which such series exists is called the *step* or *nilpotency class* of G.

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

—Lecture 8—

Example class 1400-1530 MR14 today!

Last time, we said G was *nilpotent* if $\exists$ finite central series $G = H_1 > H_2 > ... > H_{s+1} = \{1\}$. We defined the smallest s for which such a series existed the *step* of G. Today we'll look into more detail at nilpotent groups.

The reasons we focus on this setting are twofolds: there's a clean generalization of Frieman-Green-Ruzsa Theorem to nilpotent groups, and a deep theorem of Brenillard, Green and Tao essentially reduces the general case to the nilpotent case.

Given $x_1, ..., x_k \in G$, we define the simple commutator $[x_1, ..., x_k] (= [x_1, ..., x_k]_k)$ recursively as follows:
$[x_1] = x_1$;
$[x, y] = x^{-1}y^{-1}xy$;
$[x_1, ..., x_k] = [[x_1, ..., x_{k-1}], x_k]$.
Given subgroups $H, N < G$, define $[H, N] = \langle [h, n] : h \in H, n \in N \rangle$, and then given $H_1, ..., H_k < G$, define similarly
$[H_1] = H_1$,
$[H_1, ..., H_k] = [[H_1, ..., H_{k-1}], H_k]$.

Note that

$$[H, N] = [N, H] \quad (8.1)$$

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

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

Proof. Case $k = 1$ is trivial, so consider $k > 1$. If $[s_1, ..., s_k] \in N \forall s_i \in S_i$, then we have $[[s_1, ..., s_{k-1}], s_k] \in N \forall s_i \in S_i$, and hence

$$[s_1, ..., s_{k-1}] \in C_{G/N}(H_k) = \{g \in G : [g, h] \in N \forall h \in H_k\}$$

The centraliser of a normal subgroup is itself normal, so by induction, we have $[H_1, ..., H_{k-1}] \subset C_{G/N}(H_k)$. Hence $[H_1, ..., H_k] \subset N$ as claimed. □

Definition. Given a gruop G, we define the *lower central series* $G = G_1 > G_2 > ...$ of G via $G_k = \langle [g_1, ..., g_k] : g_i \in G \rangle$.

Note that $G_k > G_{k+1}$ because $[g_1, ..., g_{k+1}] = [[g_1, g_2], g_3, ..., g_{k+1}]$. Also, since $[g_1, ..., g_k]^h = [g_1^h, ..., g_k^g]$ (where the notation $\cdot^h$ means conjugating by h), each G_k is normal in G. The fact that this is a *central series* (i.e. G_k/G_{k+1} is central in G/G_{k+1} $\forall K$ follows from the following result:

Proposition. (8.2)
We have $G_{k+1} = [G_k, G]$ $\forall k$. In particular, $G_k = [G, ..., G]_k$.

Proof. First, $G_{k+1} < [G_k, G]$ by definition. The fact that $[G_k, G] < G_{k+1}$ follows from lemma 8.1, since $[g_1, ..., g_{k-1}]$ generate G_k, and G_1, G_k, G_{k+1} are normal. □

Proposition. (8.3)
Let G be a group generated by S. Then $G_k = \langle [s_1, ..., s_k] G_{k+1} : s_i \in S \forall i \rangle$.
(This basically says G_k *is generated mod* G_{k+1} *by simple commutators of generators of* G.)

Proof. Note that $[s_1, ..., s_k]^g \in [s_1, ..., s_k] G_{k+1}$ by definition of G_{k+1}.
So $\langle [s_1, ..., s_k] G_{k+1} : s \in S \rangle$ is normal in G.
Moreover, $[s_1, ..., s_k] \in \langle [t_1, ..., t_k] G_{k+1} : t_i \in S \rangle$ whenever $s_i \in S$ $\forall i$. So lemma 8.1 implies that $[G, ..., G]_k \subset \langle [s_1, ..., s_k] G_{k+1} : s_i \in S \rangle$. Proposition 8.2 gives that $[G, ..., G]_k = G_k$, so we have $G_k \subset \langle [s_1, ..., s_k] G_{k+1} : s_i \in S \rangle$.
The reverse inclusion is immediate. □

Proposition. (8.4)
We have $[G_i, G_j] \subset G_{i+j} \forall i, j$.
For this we'll need the following commutator identity, which we can just check directly by writing it out:

$$[x, y^{-1}, z]^y [y, z^{-1}, x]^z [z, x^{-1}, y]^x = 1 \quad (8.2)$$

Proof. The case $j = 1$ is proposition 8.2. So we can assume $j > 1$, and by induction, that

$$[G_k, G_{j-1}] \subset G_{k+j-1} \forall k \text{ (8.3)}$$

Now note that

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

by proposition 8.2 and formula (8.1) respectively. We also have

$$[[G_{j-1}, G_i], G] = [[G_i, G_{j-1}, G] \subset [G_{i+j-1}, G] = G_{i+j} \text{ (8.5)}$$

by formula (8.1), (8.3), and proposition (8.2) respectively, and

$$[[G_i, G], G_{j-1}] = [G_{i+1}, G_{j-1}] = G_{i+j} \text{ (8.6)}$$

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

$$\begin{aligned}[x, y, z] &= (([y^{-1}, z^{-1}, x]^z [z, x^{-1}, y^{-1}]^x)^{-1})^y \\ &\subset G_{i+j}\end{aligned}$$

by formula (8.2) and formula (8.5) plus (8.6) respectively. The proposition follows from (8.4) and Lemma (8.1). □

Definition. Given a group G, the *upper central series* $\{1\} = Z_0(G) < Z_1(G) < Z_2(G) < ...$ 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.

Proposition. (8.5)
Let $G = H_1 > H_2 > ... > H_{r+1} = \{1\}$ be a finite central series for G (so G is nilpotent). Then we have $H_i \supset G_i \ \forall i = 1, ..., r+1$, and $H_{r+1-i} \subset Z_i(a)$ $\forall i = 0, ..., r$.
This justifies the names *upper* and lower central series: $H_1 > H_2 > ... > H_r > H_{r+1}$ and $Z_{r-i+1}(G) \supset H_i \supset G_i$ for each i.

Corollary. If G is s-step nilpotent, then both the upper and lower central series have length 5.1.

Proof. (of 8.5)
$H_1 \supset G_1$ by definition, so we may assume $i > 1$, and then we have

$$\begin{aligned}H_i &\supset [H_{i-1}, G] \text{ (central series)} \\ &\supset [G_{i-1}, G] \text{ (induction)} \\ &= G_i \text{ (prop 8.2)}\end{aligned}$$

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

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