University of
Cambridge

Mathematics Tripos

Part III

Introduction to Approximate Groups

Lent, 2019

Lectures by
M. Tointon

Notes by
Qiangru Kuang

Contents

Lecture 1: missed

Lecture 2: Covering and higher sum and product sets

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 0.1 (Ruzsa)**.** *Suppose $A \subseteq \mathbb{F}_p^r$ satisfying $|A+A| \leq K|A|$. Then exists $H \leq \mathbb{F}_p^r$ with $|H| \leq p^{K^4}K^2|A|$ such that $A \subseteq H$.*

So again, like in theorem 1.1., A is a large propotion of a finite subgroup.

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

We'll start by proving the following weaker version:

Proposition 0.2. *Suppose $A \subseteq \mathbb{F}_p^r$ satisfies $|2A-2A| \leq K|A|$. Then exists $H \subseteq \mathbb{F}_p^r$ with $|H| \leq p^K|A-A|$ (so $\leq p^kK|A|$) such that $A \subseteq H$.*

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

Lemma 0.3 (Ruzsa's covering lemma)**.** *Suppose $A, B \subseteq G$ and $|AB| \leq K|B|$. Then there exists $X \subseteq A$ with $|X| \leq K$ such that $A \subseteq XBB^{-1}$. 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 say that A can be covered by a few left translates of BB^{-1}.

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

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

so $|X| \leq K$. Maximality implies that for all $a \in A$ there exists $x \in X$ such that $aB \cap xB \neq \emptyset$, and hence $a \in xBB^{-1}$. Hence $A \subseteq XBB^{-1}$ as required. □

Lemma 0.4. *Suppose $A \subseteq G$ satisfies*

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

Then exists $X \in A^{-1}A^2, |X| \leq K$ such that

$$A^{-1}A^n \subseteq X^{n-1}A^{-1}A$$

for all $n \in \mathbb{N}$.

Proof. By Ruzsa' covering lemma exists $X \subseteq A^{-1}A^2, |X| \leq K|$ such that

$$A^{-1}A^2 \subseteq XA^{-1}A.$$

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. The lemma above implies that exists X with $|X| \leq K$ such that

$$nA - A \subseteq (n-1)X + A - A$$

for all $n \in \mathbb{N}$. Since $\mathbb{F}_p^r$ is a vector space,

$$\langle A \rangle \subseteq \langle X \rangle + A - A$$

so

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

as required. □

To strengthen the proposition to the theorem, we use the second technique: bounding higher sum/product sets. The key result, at least in the abelian case, is the following:

Theorem 0.5 (Plünnecke-Ruzsa)**.** *Suppose $A \subseteq G$ where G is an abelian group and $|A - A| \leq K|A|$. Then*

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

for all $m, n \geq 0$.

This is proved in III Introduction to Discrete Analysis. We won't redo the whole proof, but we will reprove some parts of it.

Proof of Ruzsa's theorem. Plünnecke-Ruzsa implies that $|2A - 2A| \leq K^4|A|$ and $|A - A| \leq K^2|A|$. Then the result follows from prop 2.2. □

We'll spend the rest of the lecture discussing Plünnecke-Ruzsa and variants of it. We've seen it's useful, at least in one context. To see philosophically why it's useful, let's think about what the genuine closure of subgroups under products and inverses mean. One useful feature is that it can be iterated: given $h_1, h_2, \cdots \in H$ a subgroup, this means $h_1^{\varepsilon_1}, \ldots, h_m^{\varepsilon_m}, \cdots \in H$ for all $\varepsilon_i = \pm 1$ for all m, for all $h_i \in H$. The theorem allows us to "iterate" the "approximate closure" of a set of small doubling.

$$a_1 + \ldots a_m - a_1' - \cdots - a_n'$$

may not belong to A but at least it belongs to $mA - nA$, which is "not too large" ($|mA - nA| \leq K^{m+n}|A|$), and is itself a set of small doubing ($2|mA - nA|| \leq K^{2m+2n}|mA - nA|$). This is an important part of why the theory works so well.

It is therfore unfortunate that the theorem does not hold for non-abelian groups.

Example. Let x generates an infinite cyclic group $\langle x \rangle$, H be a finite subgroup. Set $G = H * \langle x \rangle$ (the key point is 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 $|H| \to \infty$, the theorem cannot hold.

Nevertheless, if we strengthen small doubling slightly we can recover a form of the theorem. One way is to replace small doubing with *small tripling*, i.e. $|A^3| \leq K|A|$.

Proposition 0.6 (2.7)**.** *Suppose $A \subseteq G$ and $|A^3| \leq K|A|$. Then*

$$|A^{\varepsilon_1} \cdots A^{\varepsilon_m}| \leq K^{3(m-2)}|A|$$

for all $\varepsilon_i = \pm 1$ for all $m \geq 3$.

The key ingredient is the following:

Lemma 0.7 (Ruzsa's triangle inequality)**.** *Given $U, V, W \subseteq G$, 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 $v(x) \in V$, $w(x) \in W$ such that $x = v(x)^{-1}w(x)$. Set

$$\varphi(u, x) = (uv(x), uw(x)).$$

To see injectivity, first notice that

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

so x is 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 do the case $m = 3$:

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

Apply triangle inequality with $U = W = A, V = A^2$. Get

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

so

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

Next note that $(A^{-1}A)^{-1} = A^{-1}A^2$ so

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

Replace A by A^{-1} we get

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

Finally, use 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 and A^{-1}.

For general m, triangle inequality implies that

$$|A||A^{\varepsilon_1} \cdots A^{\varepsilon_m}| \leq |AA^{-\varepsilon_2}A^{-\varepsilon_1}||AA^{\varepsilon_3} \cdots A^{\varepsilon_m}| \leq K^3|A||K^{3(m-2)}|A|$$

by induction. □

Lecture 3: Approximate groups

Last time we saw that assuming all small tripling instead of small doubling allowed us to control higher product sets of the form $A^{\varepsilon_1} \cdots A^{\varepsilon_m}$. In this lecture we'll see another possible strengening of small doubling. We also saw, in the proof 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 (approximate group)**.** A set $A \subseteq G$ is called a *K-approximate group* or *K-approximate subgroup* if $1 \in A, A^{-1} = A$ and exists $X \subseteq G$ with $|X| \leq K$ such that $A^2 \subseteq XA$.

Remark. Note that A need not to be finite, although in this course it almost always will be. Also if A is finite that $|A^2| \leq K|A|$.

The conditions $1 \in A$ and $A^{-1} = A$ are convenient notationally: for example we can write A^m instead of $A^{\varepsilon_1} \cdots A^{\varepsilon_m}$, and $1 \in A$ implies that $A \subseteq A^2 \subseteq A^3 \subseteq ...$, which is also convenient at times. It is the condition $A^2 \subseteq XA$ that is more important.

For approximate groups, bounding higher product is easy:

Lemma 0.8 (lemma 3.1)**.** *If A is a finite K-approximate group then*

$$A^m| \leq K^{m-1}|A|.$$

Proof. Let X be as in the definition of approximate group. In fact we have $A^m \subseteq X^{m-1}A$:

$$A^m = A^{m-1}A \subseteq X^{m-2}A^2 \subseteq X^{m-1}A$$

by induction. □

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 followin sense:

Proposition 0.9 (proposition 3.2)**.** *Let $A \subseteq G$ be finite. If A is a K-approximate group then $|A^3| \leq K^2|A|$. Conversely if $|A^3| \leq K|A|$ then exists $O(K^{12})$-approximate group B with $A \subseteq B$ and $|B| \leq 7K^3|A|$. In fact, we may take $B = (A \cup \{1\} \cup A^{-1})^2$.*

The interesting direction of the proposition says that A is a large proportion of an approximate group.

Proof. The first part 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

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

and all 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 implies that there exists $X \subseteq G$, $|X| \leq O(K^6)$ such that $\hat{A}^n \subseteq X^{n-1}\hat{A}^2$ for every $n \geq 2$. In particular $|X^2| \leq O(K^{12})$ and $\hat{A}^4 \subseteq X^2\hat{A}^2$, 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? 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 0.10 (theorem 3.3)**.** *If $A \subseteq G$ satisfies $|A^2| \leq K|A|$ then exists $U \subseteq A$ with $|U| \geq \frac{1}{K}|A|$ such that*

$$|U^m| \leq K^{m-1}|U|$$

for all $m \in \mathbb{N}$.

Thus 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 groups.

Tao proved a version of theorem 3.3 when he introduced the definition of apparoximate groups. We'll use instead a lemma of Petridis, which he proved when proving the Plüneccke-Ruzsa inequalities.

Lemma 0.11 (lemma 3.4)**.** *[Petridis] Suppose $A, B \subseteq G$ are finite. Let $U \subseteq A$ be non-empty, chosen to minimise the ratio $|UB|/|U|$ and write $R = |UB|/|U|$. Then for all finite $C \subseteq G$ we have*

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

Proof. Trivial if $C = \emptyset$ so we may assume there exists $x \in C$. Define $C' = C \setminus \{x\}$, we may also assume by induction that $|C'UB| \leq R|C'U|$. We are going to write $CU = C'U \cup xU$ and deal with the overlap. Set

$$W = \{u \in U : xu \in C'U\}.$$

Then

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

is a disjoint union so in particular

$$|CU| = |C'U| + |U| - |W|.$$

We also have $xWB \subseteq C'UB$ by definition of W so

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

and hence

$$|CUB| \leq |C'UB| + |UB| - |WB|.$$

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

$$|CUB| \leq R(|C'U| + |U| - |W|) = R|CU|.$$

□

Proof of theorem 3.3. Set $U \subseteq A$ to be non-empty minimising $|UA|/|U|$ and write $R = |UA|/|A|$. Noting that $R \subseteq K$ by minimality. Also U is non-empty so $|UA| \geq |A|$, so $|U| \geq |A|/K$ as required. Lemma 2.4 also implies that

$$|U^m A| \leq K|U^m|$$

for all m (taking $C = U^{m-1}$) and since $U \subseteq A$, this gives

$$|U^{m+1}| \leq K|U^m|$$

for all m, so $|U^m| \leq K^{m-1}|U|$. □

A bit of non-examinable information:

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 0.12 (theorem 3.5)**.** *[Tao, Petridis] If $|A^2| \leq K|A|$ and $|AxA| \leq K|A|$ for all $x \in A$ then $|A^m| \leq K^{O(m)}|A|$ for all $m \geq 3$.*

Lecture 4: Stability of approximate closure under basic operations

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

It's trivial that if $A \subseteq G$ 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 0.13 (prop 4.1)**.** *[stability of small tripling under homomorphisms] Let $A \subseteq G$ be finite, symmetric, containing the idenity. 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 prop 2.7.

Prove this using an argument of Helfgolt. We's start with a simple observation that we'll use repeatedly in this course.

Lemma 0.14 (lemma 4.2)**.** *Let $H \leq G$. Let $A \subseteq G$ be finite and let $x \in G$. Then*

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

Proof. We have

$$(A \cap xH)^{-1}(A \cap xH) \subseteq A^{-1}A \cap H.$$

□

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

Lemma 0.15 (lemma 4.3)**.** *Let $H \leq G$. Write $\pi : g \to G/H$ for the quotient map. Let $A \subseteq G$ be finite. Then*

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

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

Proof. By pigeonhole principle, there exists $x \in G$ such that

$$|A \subseteq xH| \geq \frac{|A|}{|\pi(A)|}.$$

Then apply lemma 4.2. □

Lemma 0.16 (lemma 4.4)**.** *Let $H \leq G$. Write $\pi : G \to G/H$ for the quotient map and let $A \subseteq G$ be finite. Then*

$$|\pi(A^m)||A^n \cap H| \leq |A^{m+n}|$$

for all $m, n \geq 0$.

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,

$$\varphi(\pi(A^m))(A^n \cap H) \subseteq A^{m+n}.$$

□

Proof of prop 4.1. Write $H = \ker \pi$. By lemma 4.4,

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

The by lemma 4.3

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

The proposition then follows. □

Now we'll look at intersections.

Proposition 0.17 (prop 4.5)**.** *[stability of small tripling uder intersections with subgroups] Let $A \subseteq G$ be finite, symmetric and containing identity. 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

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

for all $m \geq 2$.

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 condition.

Proof. By lemma 4.4

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

where $\pi : G \to G/H$ is the quotient map as before. By lemma 4.3,

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

Just combine these two inequalities. □

Proposition 0.18 (prop 4.6)**.** *[stability of approximate groups under intersections with subgroups] Let $H \leq G$. Let $A \subseteq G$ 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 particlar $A^m \cap H$ is a K^{2m-1}-approximate subgroup (since $A^2 \cap H \leq A^m \cap H$ and $(A^m \cap H)^2 \leq A^{2m} \cap H$).*

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

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

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

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

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

In III Introduction to Discrete Analysis, you saw that when studying small doubling/tripling, there is a more general notion of homomorphism that comes into play: the Freiman homomorphisms. To motivate this, consider two sets

$$A = \{-n, \dots, n\} \subseteq \mathbb{Z}/p\mathbb{Z}$$
$$B = \{-n, \dots, n\} \subseteq \mathbb{Z}/q\mathbb{Z}$$

for p, q two large primes, $\geq 10n$ say. These two sets are intuitively "isomorphic" from the perspective of $A + A$ and $B + B$, 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 all $x_1, \dots, x_m, y_1, \dots, y_m \in A$ with $x_1 \cdots x_m = y_1 \cdots y_m$ we have

$$\varphi(x_1) \cdots \varphi(x_m) = \varphi(y_1) \cdots \varphi(y_m).$$

If $1 \in A$ and $\varphi(1) = 1$ then we say that φ is *centred*. If φ is injective and its inverse $\varphi(A) \to A$ is also a Freiman m-homomorphism we say φ is a *Freiman m-isomorphism*.

Remark.

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

2. This definition gets stronger as m increases: assume $A \neq \emptyset$. Picking $a \in A$. If $x_1 \cdots x_k = y_1 \cdots y_k$ for $k \leq m$ then $x_1 \cdots x_k a \cdots a = y_1 \cdots y_k a \cdots a$.

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

From now on when we say φ is a Freiman homomorphism we mean it is a 2-homomorphism.

Lemma 0.19 (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 this is an equality.

Proof. Exercise. □

Lemma 0.20 (lemma 4.8)**.** *Let $A \subseteq G$ be a K-approximate group. Suppose $\varphi : A^3 \to \Gamma$ is a centred Freiman 2-homomorphism. Then $\varphi(A)$ is also a K-homomorphism.*

Proof. By definition there exists $X \subseteq G, |X| \leq K$ such that $A^2 \subseteq XA$. So given $a_1, a_2 \in A$, there 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 $\varphi(A)^2 \subseteq \varphi(X \cap A^3)\varphi(A)$. Also as φ is centred, $\varphi(A)$ is symmetric and contains 1. □

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

We'll introduce some non-trivial examples of sets of small doubling in abelian groups.

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

$$P(x; L) = P(x_1, \dots, x_r; L_1, \dots, L_r) = \{\ell_1 x_1 + \dots + \ell_r x_r : |\ell_1| \leq L_i \text{ for all} i\}$$

is called a *progression of rank* r. If in addition $H \leq G$ 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 $r = 2$ (picture)

It's easy to see that such a box B in $\mathbb{Z}^r$ is a 2^r-approximate group. For example in $r = 2$ (picture)

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 0.21 (Freiman ($G = \mathbb{Z}$), Green-Ruzsa (arbitrary abelian G))**.** *Suppose $A \subseteq G$ abelian satisfies $|A + A| \leq K|A|$. Then there exists a coset progression $H + P$ of rank $\leq O(K^{O(1)})$ such that*

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

In particular theorem 2.5 (Plünnecke-Ruzsa inequality) implies 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 III Introduction to Discrete Analysis, but with a slightly less explicity version of coset progression.

Definition (Bohr set)**.** Let G be a finite abelian group. Let

$$\Gamma = \{\gamma_1, \dots, \gamma_r\} \subseteq \widehat{G} = \mathrm{Hom}(G, \mathbb{R}/\mathbb{Z})$$

and let $\rho \in [0, \frac{1}{2}]$. Then the set

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

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

$$\|x\|_{\mathbb{R}/\mathbb{Z}} = |\widehat{x}|.$$

We'll see in example sheet 1 that $B(\Gamma, \rho)$ is a 4^r-approximate group. Whereas progression were homomorphic images of boxes, $B(\Gamma, \rho)$ is the pullback of $[-\rho, \rho]^r$ under $(\gamma_1, \dots, \gamma_r) \in \widehat{G}^r$.

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

Proposition 0.22 (from III Introduction to Discrete Analysis)**.** *Suppose $A \subseteq G$ abelian with $|A+A| \leq K|A|$. Then there exists $b \subseteq 2A-2A$, a finite abelian group Z with $|Z| \geq |A|$, a set $\Gamma \subseteq \hat{Z}$ with $|\Gamma| \leq O(K^{O(1)})$, some $\rho \geq \frac{1}{O(K^{O(1)})}$ and a centred Freiman 2-isomorphism $\varphi : B(\Gamma, \rho) \to B$.*

This is saying if A has small doubling then $2A-2A$ contains a large set isomorphic to a Bohr set of bounded rank.

In III Introduction to Discrete Analysis we see this in the special case where G is 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 0.23 (prop 5.3)**.** *Suppose X is a finite abelian group, $\Gamma \subseteq \hat{Z}$ is of size r, $\rho < \frac{1}{10}$. Then there exists a coset progression $H+P \subseteq B(\Gamma, \rho)$ with rank r and $|H+P| \geq (\rho/r)^r |Z|$.*

Lemma 0.24 (lemma 5.4)**.** *Suppose $H+P$ is a coset progression of rank r and $\varphi : H+P \to G$, where G abelian, is a centred Freiman 2-homomorphism. Then $\varphi(H+P)$ is also a coset progression of rank r.*

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

Proof of lemma 5.4. Exercise: if H is a group and $\varphi : H \to G$ is a centred Freiman 2-homomorphism then φ is also a group homomorphism.

In particular in this lemma $\varphi(H)$ is a finite subgropu. Therefore suffices to show that

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

In fact, we'll show that for all $h \in H, |\ell_i| \leq L_i$ we have

$$\varphi(h + \ell_1 x_1 + \cdots + \ell_r x_r) = \varphi(h) + \ell\varphi(x_1) + \cdots + \ell_r \varphi(x_r). \tag{5.1}$$

Since φ is centred, $\varphi(-x_i) = -\varphi(x_i)$ so we may assume that $\ell_i \geq 0$ for all i. Also 5.1 is trivial if $\ell_i = 0$ for all i. So we may assume there there exists $\ell_j > 0$. Then

$$\begin{aligned}\varphi(h + \ell_1 x_1 + \cdots + \ell_r x_r) &= \varphi(h + \ell_1 x_1 + \cdots + \ell_r x_r) + \varphi(0) \\ &= \varphi(h + \ell_1 x_1 + \cdots + (\ell_j - 1)x_j + \cdots + \ell_r x_r) + \varphi(x_j)\end{aligned}$$

so lemma follows by induction on $\sum_i \ell_i$. □

Proof of theorem 5.1. By proposition 5.2 and 5.3 and lemma 5.4, there exists $H + P$ coset progression of rank $\leq O(K^{O(1)})$ such that

$$H + P \subseteq 2A - 2A$$
$$|H + P| \geq \exp(-O(K^{O(1)}))|A|$$

We'll now apply a version of Ruzsa's covering lemma due to Chang. Define recursively sets $S_1, S_2, \cdots \subseteq A$ such that S_i the a maximal subset of size $\leq 2K$ such that the translates $x + S_{i-1} + \cdots + S_1 + H + P$ are all disjoint. If ever $|S_i| < 2K$ we stop. Now suppose we get as far as $S_1, \ldots, S_t$. Then

$$S_t + \cdots + S_1 + H + P \subseteq 2A - 2A + tA$$

so by proposition 2.5

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

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

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

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 this process terminates, at S_t, say.

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 from lecture 2 implies that

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

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

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

as claimed. □

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

Lecture 6: Geometry of numbers

Proposition 0.25 (prop 5.3)**.** *Let G be a finite abelian group. Suppose $\Gamma \subseteq \widehat{G}$ with $|\Gamma| = r$ and let $\rho < \frac{1}{2}$. Then there exists coset progression $H + P \subseteq B(\Gamma, \rho)$ of rank r with*

$$|H + P| \geq (\rho/r)^r |G|.$$

To prove this, we'll use a field called the *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 the subspace) generated by some basis $x_1, \dots, x_d$ for $\mathbb{R}^d$, so

$$\Lambda = \{\sum \ell_i x_i : \ell_i \in \mathbb{Z}\}.$$

If $\Gamma \subseteq \Lambda$ is another lattice then we say it is a *sublattice*, write $\Gamma \leq \Lambda$. It is an exercise (example sheet 2) to check that if $\Gamma \leq \Lambda$ with basis $y_1, \dots, y_d$, say, then

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

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 prop 5.3 is the following:

Lemma 0.26 (lemma 6.1)**.** *Let G, Γ be as in prop 5.3 and set $\gamma : G \to \mathbb{R}^d/\mathbb{Z}^d$ via by enumerating Γ as $\{\gamma_1, \dots, \gamma_d\}$ and set $\gamma = (\gamma_1, \dots, \gamma_d)$. 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 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 $\mathrm{span}_{\mathbb{R}}(\Lambda) = \mathbb{R}^d$. 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 $[-\rho, \rho]^d$ with Λ. To do this we introduce another definition.

Definition (convexity)**.** A set $A \subseteq \mathbb{R}^d$ is *convex* if for all $x \in \mathbb{R}^d \setminus \mathring{A}$, there exists a hyperplane h_x with $x \in h_x$ and $h_x \cap \mathring{A} = \emptyset$.

Definition (convex body)**.** A set $B \subseteq \mathbb{R}^d$ is a *convex body* if it is bounded and convex and $\mathring{B} \neq \emptyset$ and $\mathring{A}$ is contained in one of the two half spaces into which h_x divides $\mathbb{R}^d$. It is *symmetric* if for all $x \in B$, $-x \in B$.

Given a symmetric convex body B and a lattice Λ, define the *successive minima* $\lambda_1 \leq \dots \lambda_d$ of B with respect to Λ via

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

We may then inductively define linear independent vectors $v_1, \dots, v_d \in \Lambda$ such that $v_1, \dots, v_i \in \lambda_i \overline{B}$. Will call such a set a *directional basis* for Λ with respect to B. Note that it is not unique, and not necessarily a basis for Λ in the earlier sense. See example sheet 2.

Theorem 0.27 (theorem 6.2)**.** *[Minkowski's second thereom] Suppose B is a symmetric convex body, Λ a lattice in $\mathbb{R}^d$ and λ_d are successive minima. Then*

$$\lambda_1 \cdots \lambda_d \operatorname{vol}(B) \leq 2^d \det(\Lambda).$$

Lemma 0.28 (lemma 6.3)**.** *[Blichfeldt] Suppose $A \subseteq \mathbb{R}^d$ is a measurable set, Λ a lattice and for all $a, b \in A$ distinct we have $a - b \notin \Lambda$. Then*

$$\operatorname{vol}(A) \leq \det \Lambda.$$

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

$$P = \{\sum \ell_i x_i : \ell_i \in [0, 1)\}.$$

Since $x_1, \dots, x_d$ is a basis for $\mathbb{R}^d$, for all $v \in \mathbb{R}^d$ there exists unique $x_v \in \Lambda, p_v \in P$ such that $v = x_v + p_v$. Define a map

$$\begin{aligned} \varphi : \mathbb{R}^d &\to P \\ v &\mapsto p_v \end{aligned}$$

This cuts A into countably many measurable pieces and translates these pieces to P. It is injective by hypothesis, hence volume preserving, and so

$$\operatorname{vol}(A) = \operatorname{vol}(\varphi(A)) \leq \operatorname{vol}(P) = \det \Lambda.$$

□

Proof of theorem 6.2. Let $v_1, \dots, v_d$ be a directional basis for Λ with respect to B. Set $V_i = \operatorname{span}(v_1, \dots, v_i)$ (with $V_0 = 0$) and set

$$\Lambda_i = \Lambda \cap (V_i \setminus V_{i-i}).$$

Then $\Lambda = \bigcup_{i=0}^{d} \Lambda_i$ as a disjoint union.

Claim 1: we have

$$\lambda_d \mathring{B} \cap (\lambda_d \mathring{B} + \alpha x) = \emptyset$$

whenever $x \in \Lambda_j$ and $\alpha \geq \frac{2\lambda_d}{\lambda_j}$.

Proof. Given $x \in \Lambda_j$, by definition $x \neq \lambda_j \mathring{B}$, so by convexity there exists hyperplane h_x such that $x \in h_x$ and $h_x \cap \lambda_j \mathring{B} = \emptyset$. By symmetry, we may take $h_{-x} = -h_x$. Note, however, that

$$-h_x = h_x - 2x.$$

That means that $\lambda_j \mathring{B}$ is contained in the slice of space S_x between the two parallel hyperplanes h_x and $h_x - 2x$. Clearly

$$S_x \cap (S_x + \alpha x) = \emptyset$$

for all $\alpha \geq 2$, so in particular

$$\lambda_j \mathring{B} \cap (\lambda_j \mathring{B} + \alpha x) = \emptyset$$

for all such α as well. Scaling by λ_d/λ, we see that

$$\lambda_d \mathring{B} \cap (\lambda_d \mathring{B} + \alpha x) = \emptyset$$

whenever $\alpha \geq \frac{2\lambda_d}{\lambda_j}$. □

Claim 2: there exists sets

$$B_1 \subseteq B_2 \subseteq \cdots \subseteq B_d = \lambda_d \mathring{B}$$

such that

1. $\mathrm{vol}(B_i) = \left(\frac{\lambda_i}{\lambda_{i+1}}\right)^i \mathrm{vol}(B_{i+1})$ for all i.
2. We have $B_i \cap (B_i + \alpha x) = \emptyset$ whenever $x \in \Lambda_j$ and $\alpha \geq 2\max\{\frac{\lambda_i}{\lambda_j}, 1\}$.

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

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

for all $v \in V_i$. (on each slice, σ_i scales L b a factor of $\frac{\lambda_i}{\lambda_{i+1}}$ cetred at z parallel to V_i) Note the following properties:

1. $\mathrm{vol}(\sigma_i(L)) = (\lambda_i/\lambda_{i+1})^i \mathrm{vol}(L)$ (by Fubini)
2. If $L \cap (z + V_i)$ is open and convex for all z then $\sigma_i(L) \subseteq L$ because $z \in L$.
3. If $L \cap (z + V_i)$ is open and convex then so is $\sigma_i(L) \cap (z + V_i)$, and indeed so is
$$\sigma_i(L) \cap (z + V_j, j < i).$$

[1] Correction by lecturer: assume z is the centre of mass of $L \cap (z + V_i)$, so that σ's depend continuously on z. Also to be on safe side, in statement of Minkowski, assume that B is a polytope.

Set $B_d = \lambda_d \mathring{B}$ and $B_i \sigma_i(B_{i+1})$. Conclusion 1 is immediate from property 1. Conclusion 2 follows from claim 1 when $n = d$. For $i < d$, it follows by induction and repreated application of 2 and 3. Indeed, 2 for i follows from 2 for $i+1$ because σ_i scales by $\frac{\lambda_i}{\lambda_{i+1}}$ in direction x. For $i < j$, follows from $B_i \subseteq B_{i+1}$. □

To prove the theorem, note that

$$\operatorname{vol}(B_1) = \lambda_1 \cdots \lambda_d \operatorname{vol}(B)$$

and by 2, $a - b \notin 2 \cdot \Lambda$ for all $a, b \in B_1$ so by Blichfeldt,

$$\operatorname{vol}(B_1) \leq 2^d \det \Lambda.$$

□

Proof of prop 5.3. Write $\gamma = (\gamma_1, \dots, \gamma_r) \in \hat{G}^r$. Define $\gamma(H) + \mathbb{Z}^r$, which is a lattice of determinant $|\ker \gamma|/|G|$ 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 $L_i = \left\lfloor \frac{\rho}{r\lambda_i} \right\rfloor$ for each i. Then

$$P(v_1, \dots, v_r; L_1, \dots, L_r) \subseteq [-\rho, \rho]^r.$$

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 $H + P \subseteq B(\Gamma, \rho)$.

Claim that if $\ell_1, \dots, \ell_r$ and $\ell'_1, \dots, \ell'_r$ satisfy $|\ell_i|, |\ell'_i| \leq L_i$, and

$$\rho_1 x_1 + \cdots + \ell_r x_r \in \ell'_1 x_1 + \cdots + \ell'_r x_r + H$$

then in fact $\ell_i = \ell'_i$ for all i. Indeed, the equation implies that

$$(\ell_1 - \ell'_1)v_1 + \cdots + (\ell_r - \ell'_r)v_\ell \in \mathbb{Z}^r \cap [-2\rho, 2\rho]^r$$

but since $\rho < \frac{1}{2}$ this last intersection is just $\{0\}$.

Then

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

by Minkowski's 2nd theorem. □

Progressions in the Heisenberg group

Define *Heisenberg group*

$$H(\mathbb{Z}) = \begin{pmatrix} 1 & \mathbb{Z} & \mathbb{Z} \\ 0 & 1 & \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix} = \left\{ \begin{pmatrix} 1 & n_2 & n_3 \\ 0 & 1 & n_1 \\ 0 & 0 & 1 \end{pmatrix}, n_i \in \mathbb{Z} \right\}$$

Set

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

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

$$\begin{pmatrix} 1 & n_2 & n_3 \\ 0 & 1 & n_1 \\ 0 & 0 & 1 \end{pmatrix} = u_1^{n_1} u_2^{n_2} u_3^{n_3},$$

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

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

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 interchaning 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. The $n_1'n_2$ term 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 (ordered progression)**.** Given $x_1, \dots, x_r \in G, L_1, \dots, L_r \geq 0$. We define the *ordered progression* of rank r

$$P_{\text{ord}}(x; L) = P_{\text{ord}}(x_1, \dots, x_r; L_1, \dots, L_r) = \{x_1^{\ell_1} \cdots x_r^{\ell_r} : |\ell_i| \leq L_1\}$$

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

$$(u_1^{\ell_1} u_2^{\ell_2})(u_1^{\ell_1'} u_2^{\ell_2'}) = u_1^{\ell_1+\ell_1'} u_2^{\ell_2+\ell_2'} u_3^{\ell_1'\ell_2}$$

and it is then easy to check that $|P^2|/|P| \to \infty$ as $L_1, L_2 \to \infty$, essentially because by varying $\ell_1, \ell_1', \ell_2, \ell_2'$ within their ranges one can change $\ell_1'\ell_2$ without changing $\ell_1 + \ell_1'$ or $\ell_2 + \ell_2'$. This can be thought of as an extra 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

$$(u_1^{\ell_1} u_2^{\ell_2})(u_1^{\ell_1'} u_2^{\ell_2'}) = u_1^{\ell_1+\ell_1'} u_2^{\ell_2+\ell_2'} u_3^{\ell_1'\ell_2}$$

This is a freedom that the definition of ordered progression explicitly denies us.

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

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

Note that for abelian groups all three notions coincide.

It turns out that $P(u_1, u_2'; L_1, L_2)$ does have small tripling (see example sheet 2). A note of caution: nonabelian progression 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 gropu. 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 nilpotence, first define a *normal series* for a group G to be a sequence

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

of normal subgroups $G_i \trianglelefteq G$, and a *central series* to be such a normal series in which each G_1/G_{i+1} is central in G/G_{i+1}.

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

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

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.

Lecture 8: Nilpotent groups

Last time, we said G is *nilpotent* if there exists a finite central series

$$G = H_1 > H_2 > ... H_{s+1} = \{1\}.$$

and defined the smallest s for which such a series existed the *step* of G. Today we'll look in more details at nilpotent groups.

The reasons we focus on this setting are two fold: there is a clean generalisation of Freimann-Green-Ruzsa to nilopotnent groups, and a deep theorem of Breuillard, Green and Tao essentially reduces the general case to the nilpotent case.

Given $x_1, ..., x_k \in G$, define the *simple commutator* $[x_1, ..., x_k] = [x_1, ..., x_k]_k$ recursively as follows:

$$[x_1] = x_1$$
$$[x_1, ..., x_k] = [[x_1, ..., x_k], x_k]$$

(Recall that $[x, y] = x^{-1}y^{-1}xy$) Given subgroups $H, N \leq G$, define

$$[H, N] = \langle [h, n] : h \in H, n \in N \rangle$$

and then given $H_1, ..., H_k \leq G$, set

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

Note that

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

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

Lemma 0.29 (lemma 8.1)**.** *Let $H_1, ..., H_k, N \trianglelefteq G$. Let S_i be a generating set for H_i for each i. Suppose $[s_1, ..., s_k] \in N$ whevever $s_i \in S_i$ for all i. Then*

$$[H_1, ..., H_k] \leq N.$$

Proof. Induction on k. $k = 1$ is trivial so assume $k > 1$. If $[s_1, ..., s_k] \in N$ for all $s_i \in S_i$ then we have $[[s_1, ..., s_{k-1}], s_k] \in N$ for all $s_i \in S_i$, hence

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

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

Definition (lower central series)**.** Given a group G, we define the *lower central series*

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

of G via

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

note that $G_k \geq G_{k+1}$ as

$$[g_1, \dots, g_{k+1}] = [[g_1, g_2], g_3, \dots, g_{k+1}].$$

Also since

$$[g_1, \dots, g_k]^h = [g_1^h, \dots, g_k^h]$$

each G_k is normal in G, where $x^y = y^{-1}xy$ for all $x, y \in G$. The fact that this is a *central series* (i.e. G_k/G_{k+1} is central in G/G_{k+1} for all k) follows from the result.

Proposition 0.30 (prop 8.2)**.** *We have $G_{k+1} = [G_k, G]$ for all k. In particular*

$$G_k = [G, \dots, G]_k.$$

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

Proposition 0.31 (prop 8.3)**.** *Let G be a group generated by S. Then*

$$G_k = \langle [s_1, \dots, s_k]G_{k+1} : s_i \in S_i \text{ for all } i \rangle.$$

"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 $\langle [s_1, \dots, s_k]G_{k+1} : s_i \in S \rangle$ is normal in G. Moreover $[s_1, \dots, s_k] \in \langle [t_1, \dots, t_k]G_{k+1} : t_i \in S \rangle$ whenever $s_i \in G$ for all i, so lemma 8.1 implies that

$$[G, \dots, G]_k \subseteq \langle [s_1, \dots, s_k]G_{k+1} s_i \in S \rangle.$$

Proposition 8.2 implies that $[G, \dots, G]_k = G_k$, so we have

$$G_k \subseteq \langle [s_1, \dots, s_k]G_{k+1} : s_i \in S \rangle.$$

The reverse inclusion is immediate. □

Proposition 0.32 (prop 8.4)**.** *We have*

$$[G_i, G_j] \subseteq G_{i+j}$$

for all i, j.

For this we'll use 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. Case $j = 1$ follows from proposition 8.2, so we assume $j > 1$ and, by induction, that for all k

$$[G_k, G_{j=1}] \subseteq G_{k+j-1} \tag{8.3}$$

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

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

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

$$[[G_i, G_i], G_{j-1}] = [G_{i+1}, G_{j-1}] = G_{i+j}$$

by prop 8.2 and (8.3). Given $x \in G, y \in G_j$ and $z \in G_i$, we therefore have

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

by (8.2), which is contained in G_{i+j} by (8.5) and (8.6).

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

Definition. Given a group G, the *upper central series*

$$1 = Z_0(G) \leq Z_1(G) \leq Z_2(G) \leq \cdots$$

is defined recursively 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 0.33 (prop 8.5)**.** *Let $G = H_1 > H_2 > \cdots > H_{r+1} = \{1\}$ be a finite central series for G (so G is nilpotent). Then we have $H_i \supseteq G_i$ for all $i = 1, \ldots, r+1$, and $H_{r+1-i} \subseteq Z_i(G)$ for all $i = 0, \ldots, r$.*

This justifies the name *upper* and *lower* central series:

Corollary 0.34. *If G is s-step nilpotnent then both the upper and lower central series have length $s - 1$.*

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

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

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

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

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 $H_{r+1-i} \subseteq Z_i(G)$ as required. □

These results say

1. G is nilpotent of step $\leq s$ if and only if $G_{s+1} = \{1\}$ if and only if $Z_s(G) = G$.
2. If $G = \langle s \rangle$, we can verify just by checking that $[t, \ldots, t_{s+1}] = 1$ for all $t_i \in S$.
3. If S is nilpotent of step ≤ 5, then any commutator like

$$[[[q_1, q_2], q_3], [q_4, q_5]]$$

with more than 5 entries is trivial.

Lecture 9: Torsion-free nilpotent approximate groups – an overview

Recall from lecture 7 that if $x_1, \dots, x_n \in G$ and $L_1, \dots, L_r \geq 0$ then the *nonabelian progression* $P(x; L)$ consists of those elements of G expressible as products of $x_1^{\pm 1}, \dots, x_r^{\pm 1}$ in which each x_i and x_i^{-1} appear at most ℓ_i times between them.

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

Proposition 0.35 (prop 9.1*)**.** *Given $r, s \in \mathbb{N}$, there exists $\lambda = \lambda_{r,s}$ such that if $x_1, \dots, x_r$ generate an s-step nilpotnent group and $L_2, \dots, L_r \geq \lambda_{r,s}$ then $P_{nil}(x; L)$ is an $O_{r,s}(1)$-approximate group.*

We won't have time to prove this, but we'll do a special case on example sheet 2, where we'll also see that it's necessary to have $L_i \geq \lambda_{r,s}$.

As in the abelian setting, it turns out that these are essentially the only examples of finite nilpotent approximate groups, apart from genunine subgroups.

Theorem 0.36 (theorem 9.2)**.** *Let G be s-step nilpotnent, $A \subseteq G$ a finite K-approximate group. Then there exists $H \trianglelefteq \langle A \rangle$ and a nilprogression P_{nil} of rank $\leq K^{O_s(1)}$ such that*

$$A \subseteq HP_{nil} \subseteq A^{K^{O_s(1)}}.$$

In particular

$$|HP_{nil}| \leq \exp(K^{O_s(1)})|A|.$$

Remark. If $K < 2$ then the theorem is trivial. For $K \geq 2$ we have $O(K^{O(1)}) = K^{O(1)}$, i.e. multiplicative constants can be absorbed into exponents. So we're not cheating when I write $K^{O(1)}$ instead of $K^{O(1)}$.

Unfortunately we won't have time to prove this in full, but in the next few lectures we'll prove some special cases that contain most of the main ideas. We'll start with the case where G is torsion-free, where theorem 9.2 is originally due to Breuillard and Green (although we'll give a different proof).

We shall start with the following weakened version:

Theorem 0.37 (theorem 9.3)**.** *Let G be torsion-free s-step nilpotent, $A \subseteq G$ a finite K-approximate group. Then there exists an ordered progression P_{ord} of rank $\leq K^{O_s(1)}$ such that*

$$A \subseteq P_{ord} \subseteq A^{K^{O_s(1)}}.$$

The basic idea is to write A as a product of approximate groups of step $< s$ and then apply induction to reduce to the step-1 case, aka. the abelian case, and apply the Frieman-Green-Ruzsa theorem (FGR).

The result we use to do this is as follows:

Proposition 0.38 (prop 9.4)**.** *Let G be torsion-free s-step nilpotent, $A \subseteq G$ a finite K-approximate group. Then there exists $k \leq K^{O(1)}$ and $K^{O(1)}$-approximate subgroups $A_1, \dots, A_k \subseteq A^{O(1)}$ such that*

$$A \subseteq A_1 \cdots A_k \subseteq A^{K^{O(1)}},$$

and $\langle A_i \rangle$ is of step $< s$ for all i.

Proof of thm 9.3. An easy induction gievs $K^{O_s(1)}$-approximate groups $B_1, \dots, B_m \subseteq A^{O_s(1)}$ with $m \leq K^{O_s(1)}$. FGR then gives abelian progressions — in particular ordered progression — $P_1, \dots, P_m$, each of rank $\leq K^{O_s(1)}$, such that

$$B_i \subseteq P_i \subseteq B_i^{K^{O_s(1)}}$$

and hence

$$A \subseteq P_1 \cdots P_m \subseteq A^{K^{O_s(1)}}.$$

$P_1 \cdots P_m$ is an order progression of rank $\leq mK^{O_s(1)} \leq K^{O_s(1)}$, so we are done. □

Recall that in proving FGR, we wanted "$A \subseteq$ small progression", but we first proved "$A^c \supseteq$ large progression". We then used Chang's covering argument to get what we wanted. We'll use a similar approach here, starting with the following:

Proposition 0.39 (prop 9.5)**.** *Suppose G is torsion-free nilpotnent and $A \subseteq G$ is a finite K-approximate group. Then exists $r \leq K^{O(1)}$ and $K^{O(1)}$ approximate groups $A_0 A_1 \cdots A_r \subseteq A^{O(1)}$, each generating a group of step $< s$, such that*

$$|A_0 A_1 \cdots A_r| \geq \exp(-K^{O(1)})|A|.$$

Next time we'll see that passing from proposition 9.5 to proposition 9.4 is very similar to Chang covering part of the proof of FGR.

In proving prop 9.5, we actually use the preliminary version of FGR in which A^c contains a large progression. As we noted in that proof, combining prop 5.2 and 5.3 and lemma 5.4 gives the folllowing result:

Theorem 0.40 (theorem 9.6)**.** *[Green-Ruzsa] Let G be abelian and $A \subseteq G$ be a finite K-approximate group. Then exists $H < G$ and $x_1, \dots, x_r \in G$ with $r \leq K^{O(1)}$, and $L_1, \dots, L_r \in \mathbb{N}$ such that $HP(x; L) \subseteq A^4$ and*

$$|HP(x; L)| \geq \exp(-K^{O(1)})|A|.$$

We'll apply this to prop 9.3 via the following result:

Proposition 0.41 (proposition 9.7)**.** *Let G be s-step nilpotnent and $A \subseteq G$ a finite K-approximate gropu. Write $\pi : G \to G/[G,G]$ for the quotient homomorphism. Noting that $G/[G,G]$ is abelian and that $\pi(A)$ is a K-approximate group. Let $H \leq G/[G,G]$ and $x_1, \dots, x_r \in G/[G,G]$ be as*

coming from applying theorem 9.6 to $\pi(A)$. Then

$$\left|(A^{24} \cap \pi^{-1}(H)) \prod_{i=1}^{r} (A^{24} \cap \pi^{-1}(\langle x_i \rangle)\right| \geq \exp(-K^{O(1)})|A|.$$

We'll prove prop 9.7 next time. For now, let's see how this implies prop 9.5. Proposition 4.6 immediately tells us that $A^{24} \cap \pi^{-1}(H)$ and $A^{24} \cap \pi^{-1}(\langle x_1 \rangle)$ are $K^{O(1)}$-approximate groups. It turns out they also generate subgropus of step $< s$, at least when G is torsion-free.

Lemma 0.42 (lemma 9.8)**.** *Let G be s-step nilpotnent, and write $\pi : G \to G/[G,G]$ as before. Then*

1. *for all $x \in G/[G,G]$, $\pi^{-1}(\langle x \rangle)$ is of step $< s$.*
2. *if $H \leq G/[G,G]$ is a finite subgroup and G is torsion-free then $\pi^{-1}(H)$ is of step $< s$.*

Lemma 0.43 (lemma 9.9)**.** *Let G be an arbitrary group. Then the simple commutator map*

$$\begin{aligned} [\cdot, \cdots, \cdot]_k : G^k &\to G_k \\ (x_1, \ldots, x_k) &\mapsto [x_1, \ldots, x_k] \end{aligned}$$

is a homomorphism in each variable mod G_{k+1}. Moreover $[G,G]$ is contained in the kernel of each of these homomorphisms.

Index