GROWTH IN GROUPS

SIMON MACHADO

1. Introduction

The main objective of this course is to give an introduction to recent trends in the study of growth in groups. Our approach will exploit ideas originating in combinatorics - as opposed to earlier ones that focused on the geometric aspect of growth in groups. The two main results of this course are the *structure of approximate groups* (Theorem 3.8) and the *product theorem* (Theorem 3.9). While we will only provide a partial proof of the former, the latter will be fully established.

2. notation

In this course, we call a *group* any set, usually denoted by G, equipped with an associative composition law containing a neutral element and an inverse for each element.

Some examples of groups we will care about:

- *(abelian groups)* the integer lattice $\mathbb{Z}^n$, the mod p group $\mathbb{Z}/p\mathbb{Z}$;
- *(Matrix groups)* the group $SL_n(k)$ of $n \times n$ matrices with coefficients in a field k and determinant 1;
- *(Finite simple groups)* Symmetric groups S_n and the alternating group A_n;
- *(Infinite groups)* $F_2 = \langle a, b \rangle$ the free group on two generators, finitely generated nilpotent groups such as H_3 the group of 3×3 upper triangular matrices with ones on the diagonal and integer coefficients.

Given a group G we will be interested in how its subsets behave (and grow) under the group operation. Precisely, choose two subsets $X, Y \subset G$, we will study the subsets:

- $XY := \{xy \in G : x \in X, y \in Y\}$;
- $X^2 := XX$;
- $X^{-1} := \{x^{-1} \in G : x \in X\}$;
- for an integer $n \geq 0$, $X^n := \{x_1 \cdots x_n \in G : x_1, \ldots, x_n \in X\}$ and $X^{-n} := (X^{-1})^n$.

We will care about the size of these sets and denote by $|X|$ the number of elements in X.

3. Lecture 1: Motivations

Historically, there are three main trends of research regarding growth in groups: *asymptotic*, *quantitative* and *local*. We present them briefly here.

Date: September 2024.

3.1. **Asymptotic growth.** The asymptotic study of growth in groups is perhaps the most established one in mathematics. Fix a group G, and assume it is infinite. Choose now a finite generating subset $X \subset G$ and assume for convenience that it is symmetric (i.e. $e \in X$ and $X^{-1} = X$). That X is generating means that

$$G = \bigcup_{n \geq 0} X^n.$$

The underlying question in the study of growth is how fast $|X^n|$ grows. In the asymptotic setting, this means understanding how $|X^n|$ behaves as n goes to infinity.

Fact 3.1. *(1) $|X^n| \geq n$ for all $n \geq 0$;*
(2) $|X^n| \leq |X|^n$ for all $n \geq 0$.

Proof. Exercise. □

So the growth of $(|X^n|)_{n \geq 0}$ is at least *linear* and at most *exponential*. *Polynomial* growth, a type of growth halfway between linear and exponential, is of particular interest. We say a group of G has *polynomial growth* if there is a finite symmetric generating set $X \subset G$ and $d \geq 2$ such that

$$\forall n \geq 2, |X^n| \leq n^d.$$

As it turns out, the growth type (exponential, polynomial or linear) is independent of the choice of generating set. The groups $\mathbb{Z}^n$ give examples of groups of polynomial growth. We have, in fact a complete characterisation:

Theorem 3.2 (Gromov, 80s)**.** *A group G has polynomial growth if and only if G is virtually nilpotent.*

This result and its proof method were some of the most influential pieces of mathematics of the 80s. Here, we say that G is *virtually nilpotent* if there is a nilpotent subgroup $N \subset G$ such that the quotient G/N is finite. Recall furthermore that a group N is said *nilpotent* if there is positive integer c such that for all $n_1, \ldots, n_c \in N$, we have $[n_1, n_2, \ldots, n_c] = e$. Here $[n_1, n_2]$ is the *commutator* $n_1 n_2 n_1^{-1} n_2^{-1}$ and $[n_1, n_2, \ldots, n_{i+1}]$ is defined inductively as $[[n_1, n_2, \ldots, n_i], n_{i+1}]$ (we will see other characterisations later on).

3.2. **Quantitative growth.** When the ambient group G is finite and X is a symmetric generating set, the study of asymptotic growth is meaningless. On indeed has the obvious bound:

$$\forall n \geq 0, |X^n| \leq |G|.$$

The interesting question, therefore, becomes: how quickly do we reach $|X^n| = G$ (i.e. $X^n = G$)?

Fact 3.3. *If $X^n = G$, then $n \geq \frac{\log |G|}{\log |X|}$.*

Proof. This is a direct consequence of the upper bound $|X^n| \leq |X|^n$. □

For certain groups, this bound should almost be sharp:

Conjecture 3.4 (Babai)**.** *If G is a finite simple group (e.g. $G = A_n$ or $G = SL_n(\mathbb{Z}/p\mathbb{Z})$), then there is $m \leq C_u(\log |G|)^{d_u}$ such that $X^m = G$, where $C_u, d_u \geq 0$ are two universal (indep. of G and X) constants.*

We will establish this conjecture for the groups $G = SL_n(\mathbb{Z}/p\mathbb{Z})$. We can also have a more probabilistic perspective on these problems. Let ν be the uniform probability measure on X and let $Y_1, Y_2, \dots$ be independent random variables drawn with law ν. The question becomes:

Question 3.5. *How fast does the* random walk $S_n := Y_1 \cdots Y_n$ *converge to the uniform probability measure on G?*

When G is a simple group, one hopes for an exponential convergence speed, and this property is called *spectral gap*.

3.3. **Local growth.** The third type of growth might appear more exotic at first:

Lemma 3.6 (Collar lemma)**.** *There is a constant $\epsilon > 0$, often called the* Margulis constant*, such that the following holds. For any hyperbolic surface S, if γ is a closed geodesic of length $l < \epsilon$, then there is a tube $[-\epsilon l^{-1}; \epsilon l^{-1}] \times \gamma \subset S$ around γ.*

In other words, if a hyperbolic surface S is thin somewhere, it has to be thin on a large portion of S. This idea is ubiquitous in the study of surfaces. Note that the above statement is rather non-rigorous, and many terms would require a precise definition. We will not try to make this statement rigorous, but we will mention a growth result that is the crux of the proof of this lemma:

Proposition 3.7. *There is $\epsilon > 0$ such that if $X \subset SL_2(\mathbb{R})$ with $|X^2| \leq 1000|X|$ and all coefficients of matrices in X have modulus $\leq \epsilon$, then there is an abelian subgroup $A \subset SL_2(\mathbb{R})$ such that $X \subset \bigcup_{i=1}^{n} a_i A$ and $n \leq 10000000$.*

3.4. **Goals: small doubling and tripling.** Compared to the above, the modern perspective on growth is more combinatorial than geometric. The goal of the course is the two theorems mentioned in the first paragraph. We will see during the lecture that they can be used to provide swift proofs of all the results mentioned in the previous section.

The first one is the most general:

Theorem 3.8 (Breuillard–Green–Tao, 2013)**.** *For all $K \geq 0$, there is $C > 0$ such that the following holds.*

Let G be a group and $A \subset G$ be such that $|A^2| \leq K|A|$. Then there are a subgroup $N \subset G$ and a subgroup $D \subset N$, normal in N, such that:

1. *N/D is nilpotent of class $\leq C$;*
2. *$D \subset (AA^{-1})^4$;*
3. *A is covered by $\leq C$ cosets of N.*

This theorem can be understood as saying that small doubling (i.e. that $|A^2| \leq K|A|$) comes from nilpotency. Indeed, by (2) and (3), A is 'sandwiched' between D and N, and N/D is nilpotent. We will see a much more precise statement later, which asserts that A looks like some sort of arithmetic progression in N/D. Moreover, while C is a function of K, the dependence is not explicit, and no formula is known to this day.

Theorem 3.8 implies Gromov's theorem as well as the collar lemma. It also indicates that there should not be small tripling in the absence of nilpotent subgroups. The absence of nilpotent subgroups can be observed in particular in finite simple groups. Something much more precise is true.

Theorem 3.9 (Breuillard–Green–Tao; Pyber–Szabo)**.** *Let $n \geq 0$. There is $\delta, \epsilon > 0$ such that for any field k and any symmetric generating subset $A \subset SL_n(k)$ we have*

$$|A^3| \geq |A|^{1+\delta} \text{ or } |A| \geq |G|^{1-\epsilon}.$$

The result says that, unless A is almost all of G, the tripling (i.e. the ratio $\frac{|A^3|}{|A|}$) is always of the same order of magnitude as $|A|$. Hence, there is no *small tripling*. We will see that this result implies a particular case of Babai's conjecture and proves the collar lemma.

3.5. **Two results about (very) small doubling.** We end this first lecture with two results relating the small doubling of a subset to the structural properties of the ambient group. While we will never invoke these results later on, the proof method is a good illustration of the techniques we will use throughout the course.

Fact 3.10. *Let G be a group and $A \subset G$ be a subset. If $|A^2| \leq |A|$, then there is a subgroup $H \subset G$ such that for $a \in A$, $A = Ha = aH$.*

Proof. Exercise. Hint: define $H := \{g \in G : Ag = A\}$. □

Lemma 3.11 (Freiman)**.** *Let G be a group and $A \subset G$ be a subset. If $|A^2| < \frac{3}{2}|A|$, then there is a subgroup $H \subset G$ such that for some (equivalently, every) $a \in A$, $A \subset aH$ and $|H| < \frac{3}{2}|A|$. Moreover, $aH = Ha$.*

The proof is taken from https://terrytao.wordpress.com/2009/11/10/an-elementary-non-commutative-freiman-theorem/.

Proof. We will split the proof into two claims, the first of which is the most important.

Claim 3.12. *$H := AA^{-1}$ is a subgroup, $AA^{-1} = A^{-1}A$ and $|H| < 2|A|$.*

Proof of the claim. The proof relies on the following *fundamental idea*: because of the doubling condition, many products of two elements of A have to be equal (collide). Let us show how this can be exploited.

Take $a, b \in A$ arbitrary. We have

$$|aA \cap bA| \geq 2|A| - |aA \cup bA| \geq 2|A| - |A^2| > \frac{|A|}{2}.$$

Therefore, there are more than $\frac{|A|}{2}$ distinct pairs $(w, x) \in A \times A$ such that $aw = bx$ i.e. $b^{-1}a = xw^{-1}$.

We can readily make two observations:

(1) $b^{-1}a = xw^{-1} \in AA^{-1}$. So $A^{-1}A \subset AA^{-1}$ because a, b were chosen arbitrary. By symmetry, $A^{-1}A = AA^{-1}$.
(2) Because a, b were arbitrary, we know that the map

$$\phi : (w, x) \in A \times A \mapsto xw^{-1} \in AA^{-1} = A^{-1}A$$

is surjective. Moreover, we have established that for every $b^{-1}a \in A^{-1}A$, $|\phi^{-1}(b^{-1}a)| > \frac{|A|}{2}$. Hence, $|A^{-1}A| < \frac{|A \times A|}{\frac{|A|}{2}} < 2|A|$.

It remains only to prove that $H := A^{-1}A$ is a subgroup. Since $A^{-1}A$ is symmetric, it suffices to show that $A^{-1}A$ is closed under product. Let $a, b, c, d \in A$ be four arbitrary elements. As above, we know that there are more than $\frac{|A|}{2}$ pairs $(w, x) \in$

$A \times A$ - respectively $(y, z) \in A \times A$ - such that $b^{-1}a = xw^{-1}$ - respectively $c^{-1}d = yz^{-1}$.

Thus, there are at least one of the pairs (w, x) and one of the pairs (y, z) such that $w = y$ (here, we use the pigeonhole principle and we have used implicitly that if (w, x) and (w', x) are such that $xw^{-1} = b^{-1}a = x(w')^{-1}$, then $w = w'$). So $(b^{-1}a)(c^{-1}d) = (xw^{-1})(yz^{-1}) = xz^{-1} \in AA^{-1} = H$. Which proves that H is closed under product.

□

It remains to prove the right bound on the size of H and the moreover part. Both are a direct consequence of:

Claim 3.13. *For all $a \in A$, $A^2 = aHa$.*

Proof of the claim. Take $a \in A$. Then $A \subset aA^{-1}A = aH$ and $A \subset AA^{-1}a = Ha$. So $A \subset aH \cap Ha$ and indeed $A^2 \subset aHa$.

To prove the reverse inclusion, consider $z \in aHa$ arbitrary. Because H is a subgroup, there are $|H|$ ways to write $z = xy$ with $x \in aH$ and $y \in Ha$. Since $|H| < |A|$ by the first claim, more than half of these ways have $x \in A$ and more than half have $y \in A$. So, at least one has $x \in A$ and $y \in A$. In other words, $z \in A^2$. □

The proof of $aH = Ha$ is left as a (non-trivial) exercise. □

4. Lecture 2: Definition of approximate groups

4.1. **Some useful examples and non-examples.** Here is a list of examples of subsets of interest. Some have small doubling/tripling, and some do not. But each will tell us something about growth under the group operations.

(1) Set $A := \{0; 1\} \subset \mathbb{Z}$. Then $A^2 = \{0; 1; 2\}$ and $|A^2| = \frac{3}{2}|A|$. This shows that Lemma 3.11 is sharp.

(2) *(Intervals and arithmetic progressions)* More generally, for $k \in \mathbb{N}$, set $A := \{-k, \dots, 0, \dots, k\}$. Then $A^2 = \{-2k, \dots, 0, \dots, 2k\}$. So $|A^2| = 2|A| - 1$. Moreover, A^2 is an interval again and has a small doubling.

(3) *(Boxes)* For $k_1, k_2, \dots, k_r \in \mathbb{N}$, the *box* $A := \prod_{i=1}^{r} \{-k_i, \dots, k_i\}$ satisfies
$$|A^2| \leq 2^r |A|.$$

(4) *(Small doubling but large tripling)* Let G be a group and H a finite subgroup. Pick $g \in G$ and define $A := H \cup \{g\}$. Then
$$A^2 = H \cup Hg \cup gH \cup \{g^2\} \text{ and } HgH \subset A.$$
So $|A^2| \leq 3|A| - 2$ and $|A^2| \geq \frac{(|A|-1)^2}{|gHg^{-1} \cap H|}$. If $gHg^{-1} \cap H = \{e\}$ (**exercise**: find such an example), then A has small doubling but large tripling.

(5) *(Chaotic sets do not have small doubling)* For $k \leq n \in \mathbb{N}$, define the random set $A_{k,n} \subset \{0, \dots, n\}$ by drawing a subset of size k uniformly at random among all subsets of size k. Then
$$\limsup_{n \to \infty} \mathbb{E}(|A_{k,n}^2|) \geq \frac{k^2}{3}.$$
This can be interpreted as saying that random sets have little additive structure. **Exercise:** prove this claim.

(6) *(Incompartibility with multiplicative structure)* For $n \in \mathbb{N}$, set $A := \{2^k : 0 \leq k \leq n\}$. Then $|A^2| := \frac{(|A|+1)|A|}{2}$ because of the uniqueness of the base 2 decomposition. This simply illustrates the so-called *sum product-phenomenon*. Namely, a set cannot be nicely behaved for both the multiplicative law and the additive law at the same time.

This list of examples will be helpful throughout the course. Testing the results we will be proving on these examples is always a good safety check.

Another useful habit is to test out the strength of each result on an open-ended question. An interesting one is:

Question 4.1. *Let A be a finite subset of a group G. Suppose that $|A^2| < \alpha|A|$ for some $\alpha < 2$. What can you say about $|A|$?*

A good first step is to find examples of subsets satisfying the assumption. I encourage you to try to answer this question as precisely as possible using the tools we will develop in the following lectures.

4.2. **Doubling? Tripling? Differences?** We have set our sights on studying sets that "do not grow too much" under the group operation. But this is a rather vague notion, and it is *a priori* unclear what the best way to define it is. Is it that $|A^2| \leq K|A|$ for some K small compared to $|A|$? or $|A^3| \leq K|A|$? or $|AA^{-1}| \leq K|A|$ perhaps?

While we have seen in one of the examples above that small doubling does not imply small tripling, we will see that these two notions are still related through an object called an *approximate subgroup*. These will be defined in the next section. We first focus on some technical simplifications.

Lemma 4.2 (Ruzsa's triangle inequality)**.** *Let U, V, W be three finite subsets of a group G. Then*

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

The name *triangle inequality* comes from the following observation. If for $A, B \subset G$ finite one defines

$$d(A, B) := \log \frac{|A^{-1}B|}{|A|^{\frac{1}{2}}|B|^{\frac{1}{2}}}$$

then the lemma is equivalent to the statement

$$d(V, W) \leq d(V, U) + d(U, W).$$

Hence, d satisfies a triangle inequality. This *does not* mean that d is a distance, however, as $d(A, A) \neq 0$ in most cases.

Proof. We will build an injection:

$$\phi : U \times V^{-1}W \to UV \times UW.$$

The existence of this injection implies the lemma immediately.

Choose as we may $v : V^{-1}W \to V$ and $w : V^{-1}W \to W$ two maps such that for $x \in V^{-1}W$ we have $x = v(x)^{-1}w(x)$. Define now $\phi(u, x) := (uv(x), uw(x))$. We claim that ϕ is an injection. Indeed, if $\phi(u, x) := (y, z)$ then $x = y^{-1}z$ and $u = yv(y^{-1}z)^{-1}$ which proves injectivity. □

Although the proof of this lemma is short, it has many valuable consequences.

Lemma 4.3. *Let A be a finite subset of a group G such that $|A^2| \leq K|A|$. Then $|AA^{-1}| \leq K^2|A|$ and $|A^{-1}A| \leq K^2|A|$.*

Proof. Apply the triangle inequality with $U = V = W = A$. Then

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

So $|A^{-1}A| \leq K^2|A|$. Apply now the triangle inequality with $U = V = W = A^{-1}$. Then

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

Since $|X^{-1}| = |X|$ for every subset X, we find $|AA^{-1}| \leq K^2|A|$. □

Under a stronger tripling assumption we can get much more:

Lemma 4.4. *Let A be a finite subset of a group G such that $|A^3| \leq K|A|$. Then for all $m \geq 3$ and $\epsilon_1, \ldots, \epsilon_m \in \{\pm 1\}$,*

$$|A^{\epsilon_1} \cdots A^{\epsilon_m}| \leq K^{3(m-2)}|A|.$$

If moreover $A = A^{-1}$, then

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

Proof. We will proceed by induction on m. Let us start with the case $m = 3$. There is nothing to prove if $A = A^{-1}$. Otherwise, this will follow from the triangle inequality.

- Apply the triangle inequality with $U = V = A$ and $W = A^2$. We get

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

  Since $|A^2| \leq |A^3|$ we find $|A^{-1}A^2| \leq K^2|A|$. By symmetry, $|A^{-2}A| \leq K^2|A|$.
- Apply the triangle inequality with $U = V = A^{-1}$ and $W = A^{-2}$ to get

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

  This yields $|A^2A^{-1}| \leq K^2|A|$.By symmetry, $|AA^{-2}| \leq K^2|A|$.
- Apply now the triangle inequality to $U = A^{-1}$, $V = A^{-1}A$ and $W = A^{-1}$ to get

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

  Using the previous steps we find $|A^{-1}AA^{-1}| \leq K^3|A|$. By symmetry, $|AA^{-1}A| \leq K^3|A|$.

We will now prove by induction the following claim: if for all $\epsilon_1, \epsilon_2, \epsilon_3 \in \{\pm 1\}$ we have $|A^{\epsilon_1}A^{\epsilon_2}A^{\epsilon_3}| \leq k|A|$, then for all $m \geq 3$ and all $\epsilon_1, \ldots, \epsilon_m \in \{\pm 1\}$ we have

$$|A^{\epsilon_1} \cdots A^{\epsilon_m}| \leq k^{m-2}|A|. \tag{1}$$

Suppose we have established (1) for all choices of signs up to some m. Let $\epsilon_1, \ldots, \epsilon_{m+1} \in \{\pm 1\}$ be any choice of signs. Apply now the triangle inequality with $U = A$, $V = A^{-\epsilon_2}A^{-\epsilon_1}$ and $W = A^{\epsilon_3} \cdots A^{\epsilon_{m+1}}$. Then

$$|A||A^{\epsilon_1} \cdots A^{\epsilon_{m+1}}| \leq |AA^{-\epsilon_2}A^{-\epsilon_1}||AA^{\epsilon_3} \cdots A^{\epsilon_{m+1}}|.$$

By the induction hypothesis, we find,

$$|A^{\epsilon_1} \cdots A^{\epsilon_{m+1}}| \leq k^{m-2}|A|.$$

This proves the statement.

Notice finally that the assumption of the claim is satisfied with $k = K^3$ by the first part of the proof. Moreover, when A is symmetric, the assumption is satisfied with $k = K$. This concludes the proof. □

4.3. **approximate subgroups.** We are now ready to talk about the notion of *approximate groups* alluded to above:

Definition 4.5. *A subset A of a group G is called a* K-approximate subgroup *if:*

(1) A is symmetric (i.e. $e \in A$ and $A = A^{-1}$);
(2) there is $F \subset G$ with $|F| \leq K$ such that

$$A^2 \subset FA.$$

Approximate subgroups satisfy all the doubling/tripling one would want them to. If $A^2 \subset FA$ one indeed notes that $A^m \subset F^{m-1}A$. So any K-approximate subgroup A satisfies $|A^m| \leq K^{m-1}|A|$.

Remark 4.6.

- *The definition of approximate subgroups is fairly recent - it was first introduced in 2003 by Terence Tao. It had, however, been studied implicitly for much longer than that.*
- *We have already seen an example of an approximate subgroup without naming it. Indeed, for $k \in \mathbb{N}$ the subset $A := \{-k, \dots, k\}$ satisfies $A^2 = \{-k, k\} + A$. So, it is a 2-approximate subgroup.*
- *More generally, for all $m \geq 1$ and $k_1, \dots, k_m \in \mathbb{N}$, the subset $\prod_{i=1}^m \{-k_i, \dots, k_i\}$ is a 2^m-approximate subgroup.*
- *An approximate subgroup is* ***not required to be finite****. For instance, the unit interval $[-1; 1]$ is an uncountable 2-approximate subgroup. We will see and use more examples later on.*
- *It so happens that intervals are great representatives of approximate subgroups. Indeed, both the celebrated Freiman's theorem and the BGT theorem (Theorem 3.8) assert that all approximate subgroups are built out of intervals and cosets.*

The notion of approximate subgroups is particularly relevant to our study because of the following:

Lemma 4.7. *Let $A \subset G$ be a finite symmetric subset such that $|A^3| \leq K|A|$. Then A^2 is a K^3-approximate subgroup.*

This statement calls for two comments:

- Between the assumption ($|A^3| \leq K|A|$) and the conclusion (A^2 is a K^3-approximate subgroup) the key parameter goes from K to K^3. In most of this course, such a polynomial change in the parameter will be considered harmless. This can be likened to a polynomial-time reduction from one problem to another in complexity theory.
- It asserts that one finds extra structure by taking the square of A. This will be a recurring theme throughout the lecture and is usually a good rule of thumb: the powers of sets are more regular than the sets themselves.

The proof relies on a well-known yet simple covering argument which connects doubling bounds to covering by a few translates:

Lemma 4.8. *(Ruzsa's covering lemma) Let $A, B \subset G$ be finite. If $|AB| \leq K|B|$ then there is $F \subset A$ of size at most K such that $A \subset FBB^{-1}$.*

5. Lecture 3: Finding approximate subgroups

5.1. **Small tripling implies approximate subgroup.** Let us recall the last statement of the previous lecture:

Lemma 5.1. *Let $A \subset G$ be a finite symmetric subset such that $|A^3| \leq K|A|$. Then A^2 is a K^3-approximate subgroup.*

The poof of Lemma 5.1 relies on a covering argument.

Lemma 5.2 (Ruzsa's covering lemma)**.** *Let $X, Y \subset G$ be finite. If $|XY| \leq K|Y|$ then there is $F \subset X$ of size at most K such that $X \subset FYY^{-1}$.*

Proof. Pick $F \subset X$ of maximal size such that the subsets fY for $f \in F$ are pairwise disjoint (such a set F exists by Zorn's lemma). By disjointness

$$|F||Y| = |FY| \leq |XY| \leq K|Y|.$$

So $|F| \leq K$. By maximality, for all $x \in X$, there is $f \in F$ such that $xX \cap fX \neq \emptyset$. So there are $y_1, y_2 \in Y$ such that $xy_1 = fy_2$ i.e.

$$x = fy_2y_1^{-1} \subset FYY^{-1}.$$

□

Proof of Lemma 5.1. Since $|A^3| \leq K|A|$, we have by Lemma 4.4 that $|A^5| \leq K^3|A|$. Apply the covering lemma (Lemma 5.2) to $X = A^4$ and $Y = A$. We get a subset F of size at most K^3 such that $A^4 \subset FA^2$. Since A^2 is symmetric, A^2 is a K^3 approximate subgroup. □

5.2. **Small doubling implies approximate subgroup.** Small doubling is also related to approximate subgroups.

Proposition 5.3. *Let $A \subset G$ be a finite subset of a group. Suppose that*

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

Then there is a $2^{12}K^{36}$-approximate subgroup $S \subset (A^{-1}A)^2$ such that $|S| \leq 16K^{12}|A|$ and $|A \cap Sa| \geq \frac{1}{2K}|A|$ for some $a \in A$.

In other words, a large chunk of A is made of a coset of an approximate subgroup. Many proofs of Proposition 5.3 exist. We mostly follow one due to Terence Tao.

Proof. Define

$$S := \{g \in G : |Ag \cap A| \geq \frac{1}{2K}|A|\}.$$

The letter S stands for 'stabiliser', and the subset S can be interpreted as the set of those elements that do not move A too much. A stabiliser of sorts of A. The subset S is symmetric and contained in $A^{-1}A$. By Lemma 4.3, it has size at most $K^2|A|$.

Claim 5.4. *There is $F \subset A$ with $|F| \leq 2K$ such that $A \subset SF$. In particular, $|S| \geq \frac{1}{2K}$.*

Let us first show how the claim implies Proposition 5.3. We will proceed via a double-counting argument reminiscent of Lemma 3.11. What we will count is the number of quadruples $(a, b, c, d) \in AA^{-1} \times AA^{-1} \times AA^{-1} \times AA^{-1}$ such that $abcd \in AS^3A^{-1}$.

Note first that $|AA^{-1} \times AA^{-1} \times AA^{-1} \times AA^{-1}| = |AA^{-1}|^4 \leq K^8|A|^4$ by Lemma 4.3. Take $x \in AS^3A^{-1}$ and write $x = a_0s_1s_2s_3a_4^{-1}$ with $a_0, a_4 \in A$ and $s_1, s_2, s_3 \in S$. If we choose now

$$a_1 \in As_1^{-1} \cap A, a_2 \in As_2^{-1} \cap A \text{ and } a_3 \in As_3^{-1} \cap A$$

we can rewrite

$$x = (a_0a_1)^{-1}(a_1s_1a_2^{-1})(a_2s_2a_3^{-1})(a_3s_3a_4^{-1}).$$

Because of the choices of a_1, a_2 and a_3, each term in between parenthesis belongs to AA^{-1}. So for each $x \in AS^3A^{-1}$ there are at least

$$|As_1^{-1} \cap A||As_2^{-1} \cap A||As_3^{-1} \cap A| \geq \frac{1}{(2k)^3}|A|^3$$

4-tuples $(a, b, c, d) \in AA^{-1} \times AA^{-1} \times AA^{-1} \times AA^{-1}$ such that $x = abcd$ - where the lower bound follows from the claim. Hence,

$$\frac{1}{(2k)^3}|A|^3|AS^3A^{-1}| \leq K^8|A|^4.$$

So

$$|S^3| \leq |AS^3A^{-1}| \leq 8K^{11}|A| \leq 16K^{12}|S|$$

where the last inequality follows from the claim. So S is a $2^{12}K^{36}$-approximate subgroup.

It remains to prove the claim.

Proof of the claim. Suppose the claim is false. Build now $g_1, \dots, g_{2K+1} \in A$ such that for $i < j$,

$$|Ag_i \cap Ag_j| < \frac{1}{2K}|A|.$$

We construct this sequence inductively as follows: $g_1 \in A$. If $g_1, \dots, g_i$ are built and $i \leq 2K$, notice that $A \nsubseteq \bigcup_{j=1}^{i} Sg_j$ by assumption. So there is $a \in A \setminus \bigcup_{j=1}^{i} Sg_j$. In other words, for all $j \leq i$ we have $ag_i^{-1} \notin S$ i.e.

$$|Aag_i^{-1} \cap A| = |Aa \cap Ag_i| < \frac{1}{2K}|A|.$$

Define $g_{i+1} = a$. This proves that such a sequence exists.

Now, by the inclusion-exclusion principle

$$\begin{aligned} K|A| \geq |A^2| &\geq |\bigcup_{i=1}^{2K+1} Ag_i| \\ &\geq (2K+1)|A| - \sum_{i<j\leq 2K+1} |Ag_i \cap Ag_j| \\ &> (2K+1)|A| - \frac{(2K+1)(2K+2)}{4K}|A| \\ &\geq K|A|. \end{aligned}$$

And we reach a contradiction. □

□

5.3. **Stability of approximate subgroups with intersections and projections.** We wish to show that approximate subgroups are robust with respect to group operations.

Fact 5.5. *Let $A \subset G$ be a K-approximate subgroup. Let $\pi : G \to H$ be a group homomorphism. Then $\pi(A)$ is a K-approximate subgroup.*

Proof. Since A is symmetric, $\pi(A)$ is symmetric. Take $F \subset G$ with $|F| \leq K$ such that $A^2 \subset FA$. Then $\pi(A)^2 \subset \pi(F)\pi(A)$ and $|\pi(F)| \leq |F| \leq K$. □

Intersections of approximate subgroups also behave well. But the proof is a little more involved.

Proposition 5.6. *Let A be a K-approximate subgroup and B be an L-approximate subgroup of a group G. For every $m, n \geq 2$ the set $A^m \cap B^n$ is covered by at most $K^{m-1}L^{n-1}$ left translates of $A^2 \cap B^2$.*

In particular, $A^m \cap B^n$ is a $K^{2m-1}L^{2n-1}$-approximate subgroup.

We say that a $X \subset G$ is covered by at most N left translates of a subset Y if there is $F \subset G$ with $|F| \leq N$ such that $X \subset FY$.

We start with a lemma:

Lemma 5.7. *let $x, y \in G$ and let $A, B \subset G$ be two symmetric subsets such that $xA \cap yB \neq \emptyset$. There is $z \in xA \cap yB$ such that $xA \cap yB \subset z(A^2 \cap B^2)$.*

Proof. Take any $z \in xA \cap yB$. Then $z = xa = yb$ for some $a \in A$ and $b \in B$. So

$$xA \cap yB \subset z(a^{-1}A \cap b^{-1}B) \subset z(A^2 \cap B^2).$$

□

Proof of the Propositon 5.6. Pick $F_1, F_2 \subset G$ with $|F_1| \leq K$ and $|F_2| \leq L$ such that $A^2 \subset F_1 A$ and $B^2 \subset F_2 B$. Then $A^m \subset F_1^{m-1}A$ and $B^n \subset F_2^{n-1}B$. So

$$A^m \cap B^n \subset F_1^{m-1}A \cap F_2^{n-1}B \subset \bigcup_{f_1 \in F_1^{m-1}, f_2 \in F_2^{n-1}} f_1 A \cap f_2 B.$$

According to the previous lemma, for every f_1, f_2 such that $f_1 A \cap f_2 B \neq \emptyset$ there is z such that $f_1 A \cap f_2 B \subset z\left(A^2 \cap B^2\right)$. Hence, there is $Z \subset G$ with $|Z| \leq |F_1^{m-1}||F_2^{n-1}| \leq K^{m-1}L^{n-1}$ such that $A^m \cap B^n \subset Z(A^2 \cap B^2)$.

□

Exercise: Show that there are two (infinite) approximate subgroups A, B of $[-1; 1]$ such $A \cap B$ is *not* an approximate subgroup.

For finite subsets, we have even more.

Lemma 5.8. *Let $m, n \in \mathbb{N}$. Let G be a group, H be a subgroup, and write $\pi : G \to G/H$ the quotient map. Suppose that $A \subset G$ is a finite symmetric. Then*

$$|\pi(A^m)||A^n \cap H| \leq |A^{m+n}| \text{ and } |\pi(A)||A^2 \cap H| \geq |A|.$$

Proof. Define a *section* $\phi : \pi(A^m) \to A^m$. That is, for each $x \in \pi(A^m)$ choose $\phi(x) \in A^m$ such that $\pi(\phi(x)) = x$. On the one hand, we have

$$\phi(\pi(A^m))\,(A^n \cap H) \subset A^{m+n}.$$

On the other hand, since $\phi(\pi(A^m))$ contains at most one element in each coset of H,

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

This proves the first inequality.

For the second inequality, remark that for each $y \in \pi(A)$ we have,

$$|\pi^{-1}(\{y\}) \cap A| \leq |\left(\pi^{-1}(\{y\}) \cap A\right)^{-1}\left(\pi^{-1}(\{y\}) \cap A\right)| \leq |H \cap A^2|.$$

So $|A| \leq |\pi(A)||H \cap A^2|$. □

We will use Lemma 5.8 as follows: given an approximate subgroup A and a subgroup H, we can accurately evaluate the size of A by evaluating independently the size of $\pi(A)$ and $A^2 \cap H$. This is a simple but key idea in the proof of the product theorem (Theorem 3.9).

6. Lecture 4: The shrinking commutator trick

From this point on we will be interested in approximate subgroups of the group $GL_n(k)$ of invertible matrices with entries in a field k. We will mostly be interested in the fields $\mathbb{R}, \mathbb{C}, \mathbb{F}_p$ and their algebraic closure - but the specific properties of the field of definition will rarely matter.

In this lecture, we focus on a topological approach.

6.1. Shrinking commutators and elements with large centre.

When $k = \mathbb{C}$ we can equip $GL_n(\mathbb{C})$ with the norm:

$$|M| = n \sup |m_{ij}| \text{ where } M = (m_{ij})_{1 \leq i,j \leq n}.$$

This norm is *sub-multiplicative* i.e. for all $S, T \in GL_n(\mathbb{C})$, $|ST| \leq |S||T|$. Which implies $|ST - I| \leq |S - I||T| + |T - I|$. Moreover, $|I| = n$ where I denotes the identity matrix.

The crucial *shrinking property* is:

Fact 6.1. *For all $S, T \in GL_n(\mathbb{C})$, write $[S, T] = STS^{-1}T^{-1}$. Then*

$$|[S, T] - I| \leq 2|S^{-1}||T^{-1}||S - I||T - I|.$$

Proof.

$$\begin{aligned} |[S, T] - I| &= |STS^{-1}T^{-1} - I| \\ &\leq |ST - TS||S^{-1}||T^{-1}| \\ &\leq |(S - I)(T - I) - (T - I)(S - I)||S^{-1}||T^{-1}| \\ &\leq 2|S^{-1}||T^{-1}||S - I||T - I|. \end{aligned}$$

□

This simple fact yields:

Lemma 6.2. *Let $A \subset GL_n(\mathbb{C})$ be a finite K-approximate subgroup. Suppose that for all $a \in A$, $|a| \leq C_0$ for some $C_0 > n$. Then there is $\gamma \in A^2$, which commutes with at least $\delta|A|$ elements of A^4 for some δ depending on K and C_0 alone.*

To avoid confusion, we write matrices with lowercase and subsets with uppercase.

Proof. Let A' be the subset of A^2 made of these elements such that $|a - I| \leq C := \frac{C_0^{10}}{8}$. Then for every element $a \in A$ and $b \in A'$ we have:

$$|[a, b] - I| \leq 2C_0^2|a - I||b - I| \leq 2CC_0^2|a - I|.$$

Choose $\gamma \in A^2 \setminus \{I\}$ such that $|\gamma - I|$ is minimal. Then for all $a \in A'$,

$$2CC_0^2|\gamma - I| > |[a, \gamma] - I|$$

. Write $X := \{[a, \gamma] \in GL_n(\mathbb{C}) | a \in A'\}$. We claim that the subsets Ax for x ranging through X are pairwise disjoint. Otherwise, there are $a, b \in A$ and $x, y \in X$ distinct such that $ax = by$. So $yx^{-1} = b^{-1}a \in A^2 \setminus \{I\}$. Hence,

$$\begin{aligned} |\gamma - I| \leq |b^{-1}a - I| = |yx^{-1} - I| &\leq |y - I| + C_0^8|x - I| \\ &\leq 2(1 + C_0^8)CC_0^2|\gamma - I| \\ &< |\gamma - I|. \end{aligned}$$

A contradiction. By disjointness, we have

$$|A||X| = |AX| \leq |A^9| \leq K^8|A|.$$

So $|X| \leq K^8$. This means that the map $a \in A' \mapsto [a, \gamma]$ takes at most K^8 values. So there is $A'' \subset A'$ of size at least $K^{-8}|A'|$ such that for all $a, b \in A''$, $[a, \gamma] = [b, \gamma]$ which is equivalent to $b^{-1}a\gamma = \gamma b^{-1}a$. So $A''^{-1}A'' \subset A^4$ has size at least $K^{-8}|A'|$ and all of its elements commute with γ.

It remains to prove that A' is large enough. Since A is symmetric, $A \subset L := \{a \in GL_n(\mathbb{C}) || a|, |a^{-1}| \leq C_0\}$ which is a compact. Choose $r > 0$, then the ball B_r of radius r centred at I is an open subset containing I. Hence,

$$L \subset \bigcup_{g \in L} gB_r$$

which, by compactness, implies that there is $F \subset L$ finite such that $L \subset FB_r$. So there is $f \in F$ such that $|A \cap fB_r| \geq \frac{|A|}{|F|}$. Hence,

$$\frac{|A|}{|F|} \leq |(A \cap fB_r)^{-1}(A \cap fB_r)| \leq |A^2 \cap B_{2C_0 r}|.$$

Taking $r \leq \frac{C}{2C_0}$ proves the claim (notice that F is chosen independently of A). □

Corollary 6.3. *Let $A \subset SL_2(\mathbb{C})$ be a K-approximate subgroup, all of whose elements have norm at most C_0. Then there is an abelian subgroup $Z \subset GL_2(\mathbb{C})$ such that $A \subset FZ$ for some F of size bounded in terms of K, C_0 alone.*

Proof. Notice first that there is no element of the form λI, $\lambda \neq 1$ in $Sl_2(\mathbb{C})$ such that $|\lambda I - I| \leq \frac{\pi}{4}$. We can adapt the end of the previous proof to choose $A' \subset A$ symmetric such that $A'^4 \subset B_{\frac{\pi}{4}}$ and $|A'| \geq \delta|A|$ for some $\delta > 0$ depending on C_0 alone. Now, A'^2 is a $K^6\delta^{-2}$-approximate subgroup. So by the lemma, there is $\gamma \in A'^4 \setminus \{I\}$ that commutes with at least $\delta'|A|$ elements of A'^8 for some δ' depending on K and δ alone. Write Z the centraliser of γ. Because $\gamma \neq \lambda I$, Z is abelian. So $|A'^4 \cap Z| \geq \delta'|A|$. And $|A(A'^4 \cap Z)| \leq |A|^9 \leq K^8|A| \leq \delta'^{-1}K^8|A'^4 \cap Z|$. By the covering lemma, $A \subset FZ$ for some F of size at most $\delta'^{-1}K^8$. □

6.2. **Gromov's theorem from the theory of approximate subgroups.** The most general theorem for approximate subgroups sounds a lot like what we have proved above:

Theorem 6.4 (Breuillard–Green–Tao)**.** *Let $A \subset G$ be a K-approximate subgroup. There are subgroups $H \subset G$ and $N \subset H$ normal in H such that:*

(1) $A \subset FH$ with F of size bounded above in terms of K alone;

(2) $N \subset A^4$*;*
(3) H/D *is nilpotent. In particular,* H *is virtually nilpotent.*

We will deduce:

Theorem 6.5. *Let* G *be a group. Let* S *be a symmetric generating set such that there is* $C, d > 0$ *with* $|S^n| \leq Cn^d$ *for all* $n \geq 1$*. Then* G *is virtually nilpotent i.e. there is* $H \subset G$ *such that* $|G/H| < \infty$*.*

Proof assuming BGT theorem. For all $n \geq 0$, $|S^{3^n}| \leq C3^{dn}$. So for $n_0 \geq 0$ (to be chosen later),

$$\frac{|S^{3^{n+1}}|}{|S^{3^{n_0}|}} = \prod_{i=n_0}^{n} \frac{|S^{3^{i+1}}|}{|S^{3^i|}} \leq \frac{C3^{d(n+1)}}{S^{3^{n_0}}}.$$

But for n sufficiently large, $\left(\frac{C3^{d(n+1)}}{S^{3^{n_0}}}\right)^{\frac{1}{n-n_0+1}} \leq 4^d$. So there is $i \geq n_0$ such that $\frac{|S^{3^{i+1}}|}{|S^{3^i|}} \leq 4^d$. Hence, by a Lemma 5.1, $S^{2\cdot 3^i}$ is a 4^{3d}-approximate subgroup. According to the BGT theorem there is a virtually nilpotent subgroup $H \subset G$ such that $S^{2\cdot 3^i} \subset FH$ for some F of size bounded in terms of d alone. Suppose we had chosen n_0 such that $2 \cdot 3^i > |F|$. For every $l \leq 2 \cdot 3^i$, $S^l \subset S^{2\cdot 3^i} \subset FH$. So there is $F_l \subset F$ such that $S^l H = F_l H$. We can moreover assume that $F_l \subset F_{l+1}$. Since $2 \cdot 3^i > |F|$, there is l such that $F_{l+1} = F_l$ by the pigeonhole principle. Thus, $S^l H = F_l H = F_{l+1} H = S^{l+1} H$. By an easy induction $FH \supset S^l H = \langle S \rangle H = G$. So H has finite index in G. □

7. Lecture 5: Some algebraic geometry

7.1. **Growth in** $SL_n(k)$**.** We will finally make progress towards the proof of the Product theorem. Let us recall the statement:

Theorem 7.1 (Product theorem, Theorem 3.9)**.** *Let* $n \geq 0$*. There are* $\delta, \epsilon > 0$ *such that for any field* k *and any finite symmetric generating subset* $A \subset SL_n(k)$ *we have*

$$|A^3| \geq |A|^{1+\delta} \text{ or } |A| \geq |G|^{1-\epsilon}.$$

As stated, the theorem is only meaningful when k is finite. We will replace the generating assumption with something more relevant (i.e. Zariski-density) later on. The product theorem can be reformulated in terms of approximate subgroups:

Theorem 7.2. *Let* $n \geq 0$*. There is* $C > 0$ *such that for any field* k *and any finite generating* K*-approximate subgroup* $A \subset SL_n(k)$ *we have*

$$|A| \leq K^C \text{ or } |A| \geq \frac{|G|}{K^C}.$$

The proof will combine the Larsen-Pink inequality with a so-called pivot argument. Most of the work will go towards proving the Larsen-Pink inequality.

Theorem 7.3 (Larsen–Pink inequality)**.** *Let* $M \geq 1$*, let* k *be a field and* $A \subset SL_n(k)$ *be finite and symmetric. Then one of the following is true:*

(1) A *is contained in a subvariety of complexity at most* M *and dimension strictly less than* $\dim SL_n = n^2 - 1$*;*

(2) For every subvariety V of $SL_n(k)$ of complexity at most M, we have

$$|A \cap V| \leq C|A^C|^{\frac{\dim V}{\dim SL_n}}$$

for some $C = O_M(1)$.

There are a lot of terms to define/explain in this statement. We will spend the rest of the lecture doing just that. To start with, we will be using Landau notation much more often. Given some parameters $\underline{x}$ and two non-negative functions f, g we write $f = O_{\underline{x}}(g)$ or, equivalently, $f \ll_{\underline{x}} g$ to mean $f \leq Cg$ where C is a constant depending on $\underline{x}$ alone.

7.2. **Elementary algebraic geometry.** Given a field k, we will denote by $\bar{k}$ its algebraic closure and $k[X_1, \ldots, X_n]$ the ring of polynomials in n variables with coefficients in k.

Definition 7.4. *Given an algebraically closed field $\bar{k}$ and polynomials $P_1, \ldots, P_M \in \bar{k}[X_1, \ldots, X_d]$ of degree at most M, we call the subset*

$$V = V(P_1, \ldots, P_M) := \{(x_1, \ldots, x_d) \in \bar{k}^d | P_1(x_1, \ldots, x_d) = \ldots = P_M(x_1, \ldots, x_d) = 0\}$$

a (sub)variety of $\bar{k}^d$ of complexity at most M.

A few examples and non-examples:

(1) if $M = 1$, $P_1 = 0$, $V(P_1) = \bar{k}^d$;
(2) if $M = 1$, $P_1 = 1$, $V(P_1) = \emptyset$;
(3) $M = 1$ and $P_1 = \det((x_{ij})_{1 \leq i,j \leq d}) - 1$ then
$$V(P_1) = SL_d(\bar{k})$$
where we have implicitly identified $\bar{k}^{d^2}$ and the space of $d \times d$ matrices with entries in $\bar{k}$.
(4) the complex halfspace in $\mathbb{C}$ i.e. $H := \{x + iy | y \geq 0\}$ is *not* a subvariety.

Fact 7.5. *The union (resp. intersection) of subvarieties of complexity M, M' is a subvariety of complexity MM' (resp. $M + M'$).*

Proof. For any $P_1, \ldots, P_M \in \bar{k}[X_1, \ldots, X_d]$ of degree at most M and $Q_1, \ldots, Q_{M'} \in \bar{k}[X_1, \ldots, X_d]$ of degree at most M' we have:

$$V(P_1, \ldots, P_M) \cap V(Q_1, \ldots, Q_{M'}) = V(P_1, \ldots, P_M, Q_1, \ldots, Q_{M'})$$

and

$$V(P_1, \ldots, P_M) \cup V(Q_1, \ldots, Q_{M'}) = V((P_i Q_j)_{i,j}).$$

This proves the fact. □

So given two varieties, we can obtain a third by taking unions.

Definition 7.6. *We say that a subvariety $V \subset \bar{k}^d$ is* irreducible *if there are no two subvarieties $V_1 \neq V \neq V_2$ with $V = V_1 \cup V_2$.*

The type of maps we will be interested in (multiplication, conjugation) will all be polynomial in nature:

Definition 7.7. *A map $P : \bar{k}^{d_1} \to \bar{k}^{d_2}$ is* polynomial *if all its coordinate are polynomials in $\bar{k}$.*

It is tempting to say the image of a subvariety through a polynomial map is a subvariety. While this is false, something close holds.

Proposition 7.8 (Chevalley's theorem)**.** *If $P : \bar{k}^{d_1} \to \bar{k}^{d_2}$ is polynomial and $V \subset \bar{k}^{d_1}$ is a subvariety, then $P(V)$ is a finite union of sets of the form $V_1 \cap \bar{k} \setminus V_2$ where V_1 and V_2 are subvarieties.*

Moreover, if P is defined by polynomials of degree at most M and V has complexity at most M, then the $V_i's$ have complexity $O_{M,d_1,d_2}(1)$.

The map

$$\bar{k}^2 \longrightarrow \bar{k}^2$$
$$(x, y) \longmapsto (x, xy)$$

is an informative example. Its image is the set $\{(x, Y)|x \neq 0\} \cup \{(0, 0)\}$. A proof of Proposition 7.8 can be found in Hartshorne's "algebraic geometry".

7.3. **The Zariski topology.** The subvarieties so happen to be precisely the closed subsets of a topology called the *Zariski topology*. We will prove this fact here, assuming some knowledge of commutative algebra. This will be a good opportunity to give a brief glimpse at a fundamental idea of algebraic geometry.

Namely, to understand the properties of a subvariety $V \subset \bar{k}^d$ we have to understand the properties of its defining equations. But it is not always easy to choose the correct $P_1, \dots, P_M$ such that $V = V(P_1, \dots, P_M)$ for a given purpose, as there are not a unique, or canonical, choice. Instead, one considers

$$I(V) := \{P \in \bar{k}[X_1, \dots, X_d] | \forall (x_1, \dots, x_d) \in V, P(x_1, \dots, x_d) = 0\}.$$

Fact 7.9. *$I(V)$ is an ideal of $\bar{k}[X_1, \dots, X_d]$.*

Fact 7.10 (Noetherianity)**.** *Every ideal of $\bar{k}[X_1, \dots, X_d]$ is finitely generated.*

When $d = 1$ this is a simple consequence of Euclid's division for polynomials. This provides a good first step for an induction, see Hartshorne's "Algebraic geometry". As a consequence:

Lemma 7.11. *If $(V_i)_{i \in \mathbb{N}}$ is a descending ($V_{i+1} \subset V_i$) family of subvarieties of $\bar{k}^d$, then there is i_0 such that $V_{i_0} = V_j$ for all $j \geq i_0$.*

Proof. Write $I_i = I(V_i)$. We will go back and forth between ideals and subvarieties. Since $I_i \subset I_{i+1}$ for all $i \in \mathbb{N}$, $I_\infty = \bigcup_i I_i$ is an ideal. As such, it is finitely generated. Pick $P_1, \dots, P_r$ the generators. Since $P_1, \dots, P_r \subset \bigcup_i I_i$ and the I_i's are an ascending family of ideals, there is i_0 such that $P_1, \dots, P_r \in I_{i_0}$. So $I_{i_0} = I_\infty = I_j$ for all $j \geq i_0$.

We prove that $V(P_1, \dots, P_r) =: V_\infty = \cap_i V_i$. Since $P_1, \dots, P_r$ are in I_j for all $j \geq i_0$, they vanish on V_j. So $V_\infty \supset \cap_i V_i$. Conversely, if $(x_1, \dots, x_d) \in V(P_1, \dots, P_r)$, then $P_1, \dots, P_r$ vanish on $(x_1, \dots, x_d)$. So all the polynomials in the ideal generated by $P_1, \dots, P_r$ vanish on $(x_1, \dots, x_d)$ i.e. all the polynomials in I_i for all i vanish on $(x_1, \dots, x_d)$. In other words, $(x_1, \dots, x_d) \in V_i$ for all $i \in \mathbb{N}$. □

So the subvarieties are indeed the closed subsets of a topology:

Definition 7.12. *The* Zariski topology *is the unique topology for which the subvarieties are the closed subsets. For a subset $X \subset \bar{k}^d$ we write $\overline{X}$ its closure in this topology. This is the* Zariski-closure*.*

When we consider a topological notion with respect to the Zariski-topology, we will use the prefix *Zarsiki-* to avoid any ambiguity.

Another useful consequence of Noetherianity concerns irreducibility.

Fact 7.13. *Any subvariety is a finite union of irreducible subvarieties.*

Hint: Proceed by contradiction and build a descending sequence of Zariski-closed subsets.

Finally, we define the dimension of a subvariety:

Definition 7.14. *A subvariety $V \subset \bar{k}^d$ has* dimension a least D *if there is a sequence:*

$$\emptyset \subsetneq V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_D \subseteq V$$

where the V_i's are irreducible subvarieties. The maximal such D is the dimension *of V.*

8. Lecture 6: Counting in finite fields

Because of the Larsen–Pink inequality, computing the dimension of a given subvariety immediately translates to counting estimates for approximate groups. It is therefore important to be able to compute these dimensions.

Fact 8.1. *(1)* $\dim(\bar{k}^d) = d$*;*

(2) $\dim(M_d(\bar{k})) = d^2$ *where $M_d(\bar{k})$ denotes the space of $d \times d$ matrices with entries in $\bar{k}$;*

(3) $\dim(Diag(d)) = d$ *where $Diag(d) \subset M_d(\bar{k})$ is the subset of diagonal matrices;*

(4) $SL_d(\bar{k}) = d^2 - 1$*;*

(5) $\dim(T_0) = d - 1$ *where T_0 denotes the subset of diagonal matrices of determinant* 1*.*

While the dimension of $\bar{k}^d$ is not so difficult to compute, it requires a certain amount of knowledge about k-algebras that I do not want to cover here, see Proposition 1.9 in Hartshorne's "Algebraic geometry" and references therein. Parts (2) and (3) follow readily from (1). To prove (4) and (5) it is enough to show:

Lemma 8.2. *Let $V \subset \bar{k}^d$ be an irreducible subvariety and $P \subset \bar{k}[X_1, \ldots, X_d]$ that does not vanish everywhere on V. Then*

$$\dim(V \cap V(P)) \leq \dim(V) - 1.$$

Proof. Let D denote the dimension of $V \cap V(P)$. By definition, there are irreducible subvarieties

$$\emptyset \subsetneq V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_D \subset V \cap V(P).$$

Since V is irreducible and $V \cap V(P) \neq V$, we have

$$\emptyset \subsetneq V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_D \subsetneq V.$$

This shows that $\dim(V) \geq D + 1$. □

8.1. **The Schwartz–Zippel lemma.** The first evidence we will collect that dimensions and counting are related is:

Lemma 8.3 (Schwartz–Zippel lemma)**.** *Let k be a finite field and $Q \in k[X_1, \ldots, X_d] \setminus \{0\}$ have degree D. Then:*

$$|\{x \in k^d | Q(x) = 0\}| \leq D|k|^{d-1}.$$

This can be understood as a quantitative version of Lemma 5.2.

Proof. We prove the statement by induction on d. If $d = 1$ then Q is a polynomial in one variable of degree D over k so it has at most D roots i.e. $|\{x \in k | Q(x) = 0\}| \leq D$.

Suppose that we have shown the claim up to $d-1$ and let $Q \in k[X_1, \dots, X_d]$. Then $Q = \sum_{i=1}^{D_1} X_1^i P_i(X_2, \dots, X_d)$ where P_{D_1} is non-zero. Then $D \geq D_i + deg(P_{D_i})$.

Given $(x_2, \dots, x_d) \in k^{d-1}$, either $P_{D_1}(x_2, \dots, x_d) = 0$ or it is not. By the induction hypothesis, there are at most $deg(P_{D_1})|k|^{d-2}$ tuples such that $P_{D_1}(x_2, \dots, x_d) = 0$. In the latter case, $Q(X_1, x_2, \dots, x_d)$ is a polynomial of degree D_1 in one variable. So there are at most D values x_1 such that $Q(x_1, \dots, x_d)$.

This yields

$$\begin{aligned}
|\{x \in k^d | Q(x) = 0\}| &\leq |\{x \in k^d | Q(x) = 0, P_{D_1}(x) = 0\}| + |\{x \in k^d | Q(x) = 0, P_{D_1}(x) \neq 0\}| \\
&\leq |\{x \in k^d | P_{D_1}(x) = 0\}| + D_1 |k|^{d-1} \\
&\leq (deg(P_{D_1}) + D_1) |k|^{d-1} \\
&\leq D|k|^{d-1}.
\end{aligned}$$

□

We can deduce a useful baby case of the Larsen Pink inequality:

Lemma 8.4. *Let V be a* proper *subvarietyof $SL_n(\bar{k})$ of complexity at most M. Let $k \subset \bar{k}$ be a finite subfield. Then*

$$|Sl_n(k) \cap V| \ll_{M,n} |k|^{n^2-2}.$$

Proof. Because V is a proper subvariety of complexity at most M, there is a polynomial P vanishing on V but not everywhere on $SL_n(\bar{k})$. So $V \subset V(P) \subsetneq SL_n(\bar{k})$. We will assume from now on that $V = V(P)$. By the complexity hypothesis, we may also assume that P has degree at most D.

So we want to find an upper bound for $|\{x \in Sl_n(k) | P(x) = 0\}|$. Here, P is understood as a polynomial with variables in the entries of the matrix x. We will find a bound by showing that the matrix x is determined by n^2-1 of its coordinates, which will then allow us to apply Lemma 8.3.

Recall moreover that because $x \in Sl_n(\bar{k})$, $x^{-1} = (C_{ij}(x))$ where $C_{ij}(x)$ is a polynomial of degree $n-1$ in the variables $(x_{kl})_{k \neq i, l \neq j}$ computed as the determinant of a minor of x. Moreover, x^{-1} is non-zero, so at least one of the $C_{ij}(x)$ is non-zero.

In other words,

$$|\{x \in SL_n(k) | P(x) = 0\}| \leq \sum_{1 \leq i,j \leq n} |\{x \in SL_n(k) | P(x) = 0, C_{ij}(x) \neq 0\}|.$$

Write $O_{ij} := \{x \in SL_n(k) | P(x) = 0, C_{ij}(x) \neq 0\}$. We want to bound $|O_{ij}|$. If $x \in O_{11}$, write $x' := (x_{ij})_{(i,j) \neq (1,1)}$. So as x ranges through $SL_n(k)$, x' ranges through k^{n^2-1}. Now, there is a polynomial Q in x' of degree at most n such that $1 = \det(x) = Q(x') + C_{11}(x')x_{11}$. Since $C_{11}(x') \neq 0$ we have $x_{11} = \frac{1-Q(x')}{C_{11}(x')}$. Now, $O_{11} := \{x \in SL_n(k) | P(x) = 0, x_{11} = \frac{1-Q(x')}{C_{11}(x')}, C_{ij}(x) \neq 0\}$. Finally, there are P_1, P_2 of degree $O_{M,n}(1)$ and in the variable $x' = (x_{ij})_{(i,j) \neq (1,1)}$ such that $\frac{P_1(x')}{P_2(x')} = P\left(\frac{1-Q(x')}{C_{11}(x')}, x'\right)$. We find finally,

$|O_{11}| \leq |\{x' \in k^{n^2-1} | P_1(x') = 0\}|$. But P_1 has now $n^2 - 1$ variables and degree $O_{m,n}(1)$. By Lemma 8.3,

$$|\{x' \in k^{n^2-1} | P_1(x') = 0\}| \ll_{M,n} |k|^{n^2-2}.$$

□

Remark 8.5. *Something fishy has happened here. In Lemma 8.3, the polynomial had coefficients in k, but the polynomials we consider in the proof of Lemma 8.4 have coefficients in $\bar{k}$. Why is it not a problem?*

8.2. **Deducing the product theorem.** We are now ready to deduce Theorem 3.9 from the Larsen-Pink inequality. We will consider the two theorems in their approximate subgroup form:

Theorem 8.6. *Let $n \geq 0$. There are $\delta, \epsilon > 0$ such that for any field k and any finite symmetric generating subset $A \subset SL_n(k)$ we have*

$$|A^3| \geq |A|^{1+\delta} \text{ or } |A| \geq |G|^{1-\epsilon}.$$

And:

Theorem 8.7 (Larsen–Pink inequality, approximate subgroup form)**.** *Let $M \geq 1$, let k be a field, $\bar{k}$ be its algebraic closure and $A \subset SL_n(k)$ be finite K-approximate subgroup. Then one of the following is true:*

1. *A is contained in a algebraic subgroup of complexity at most M and dimension strictly less than $\dim SL_n = n^2 - 1$;*
2. *For every subvariety V of $SL_n(\bar{k})$ of complexity at most M and every $m \in \mathbb{N}$, we have*

$$|A^m \cap V| \ll_{M,m,n} K^{O_{M,m,n}(1)} |A|^{\frac{\dim V}{\dim SL_n}}.$$

The Larsen—Pink inequality asserts that computing dimensions is an efficient way to compute intersections with subvarieties. We will use that intuition for the specific varieties we introduce now. They are all related to the notion of tori in one way or another. Our goal will be to prove a dichotomy for the intersection of A with a torus, which will be the starting point of a so-called *pivot argument*.

9. Lecture 7: The proof of the product theorem

All the subvarieties we are interested in have to do with diagonalizability.

1. Write $T_0 \subset SL_n(\bar{k})$ the subgroup of diagonalizable matrices. It has complexity at most n^2 and dimension $n - 1$.
2. We call a *torus* any conjugate of T_0 i.e. a subgroup T of the form gT_0g^{-1} for some $g \in SL_n(\bar{k})$. Any torus also has complexity at most n^2 and dimension $n - 1$.
3. An element $g \in SL_n(\bar{k})$ is *regular* if all of its eigenvalues are distinct. For any $n \times n$ matrix x, we denote its *characteristic polynomial* by χ_x. The coefficients of χ_x are themselves polynomials in the entries of x. Write $Disc(x)$ the discriminant of χ_x, it is a polynomial in the coefficients of χ_x - hence in the entries of x - and it vanishes if and only if two roots of χ_x are equal - that is to say when x is not regular. Thus, the set of non-regular matrices is a proper subvariety of complexity $O_n(1)$. It has dimension at most $n^2 - 2$.

(4) For $g \in SL_n(\bar{k})$ we define its *centralizer* $Z(g) := \{h \in SL_n(\bar{k}) | hg = gh\}$ and its *conjugacy class* $Conj(g) := \{hgh^{-1} \in SL_n(\bar{k}) | h \in SL_n(\bar{k})\}$. The centralizer of g is a Zariski-closed subgroup. Both sets are particularly interesting when g is a regular element.

Claim 9.1. *If g is a regular element, then $Z(g)$ is a torus.*

Proof. Since g is regular, $g = hdh^{-1}$ where $h \in SL(\bar{k})$ and d is a diagonal matrix with pairwise distinct diagonal entries. Notice that if x commutes with g then $h^{-1}xh$ commutes with d. So $h^{-1}Z(g)h \subset Z(d)$ and the reverse inclusion follows in the same way. Therefore, $Z(g) = hZ(d)h^{-1}$. For any matrix y, $yd = dy$ implies $dyd^{-1} = y$. Compute the entries of dyd^{-1} we see that this can only happen if y is diagonal i.e. $Z(d) \subset T_0$. Since two diagonal matrices commute, $Z(d) = T_0$. Hence, $Z(g) = hT_0h^{-1}$. □

The conjugacy class of a regular element also satisfies good properties:

Claim 9.2. *If g is a regular element, then $Conj(g)$ is a subvariety of dimension at most $n^2 - n$.*

Proof. Since g is regular, its characteristic polynomial χ_g splits (i.e. has n simple roots). Moreover, any h that is conjugate to g has the same characteristic polynomial as g. Finally, every h with characteristic polynomial χ_g is conjugate to the diagonal matrix with the roots of χ_g on the diagonal and, hence, every such element is conjugate to g. In other words,

$$Conj(g) := \{h \in SL_n(\bar{k}) | \chi_h = \chi_g\}.$$

But the coefficients of χ_h are polynomial in the entries of h, so $Conj(g)$ is a subvariety.

To prove the dimension bound, consider the subset

$$V = \{(h, hgh^{-1}) \in SL_n(\bar{k})^2 | h \in SL_n(\bar{k})\}.$$

Note first that $(h, x) \in V$ if and only if $h^{-1}xh = g$. So V is a subvariety. The map $\phi : h \in SL_n(\bar{k}) \mapsto (h, hgh^{-1}) \in V$ is polynomial, bijective and its inverse is $\phi^{-1} : (h, x) \in V \mapsto h \in SL_n(\bar{k})$. Since ϕ^{-1} is also polynomial, we have that V and $SL_n(\bar{k})$ have the same dimension i.e. $n^2 - 1$.

Similarly, consider the natural projection $p : (h, x) \in V \mapsto x \in SL_n(\bar{k})$. Then $p(V) = Conj(g)$. Let d denote the dimension of $Conj(g)$ and

$$V_0 \subsetneq \ldots \subsetneq V_d \subset Conj(g)$$

be irreducible subvarieties that witness the dimension of $Conj(g)$. We have that $V_0 = \{x\}$ for some x in $Conj(g)$. Since $x = hgh^{-1}$ for some h, upon considering the irreducible subvarieties $h^{-1}V_ih$ instead of V_i we may also assume that $V_0 = \{g\}$. For all $d \geq i \geq 0$, define $V'_{n-1+i} := p^{-1}(V_i)$. Moreover,

$$p^{-1}(V_0) = \{(h, x) \in V | x = g\} = \{(h, g) \in V | hgh^{-1} = g\} = Z(g) \times \{g\}.$$

Since $Z(g)$ has dimension $n - 1$ we have $V'_0 \subsetneq \ldots \subsetneq V'_{n-1} \subset Z(g) \times \{g\}$ that witness the dimension of $Z(g)$. Hence, $V'_0 \subsetneq \ldots V'_{n-1} \subsetneq V - n - 1 + d \subset V$. So (see Remark 9.3) V has dimension at least $n - 1 + d$. But V has dimension $n^2 - 1$. So $d \leq n^2 - n$. □

Remark 9.3. *The above Claim can be understood intuitively as follows: the conjugacy class is the image of the conjugation map $h \mapsto hgh^{-1}$ and the centralizer $Z(g)$ is in some sense its "kernel". If the conjugation map were a linear map, the rank-nullity theorem would immediately give that the dimension of the image (the rank) plus the dimension of the kernel is equal to the dimension of the source space. Here, in some sense something similar is true for polynomial maps, and the above proof can be generalized.*

Furthermore, to be completely rigorous in the above proof, we would need to prove that the V_i's are irreducible. Which might not be true. This is however not a problem.

Exercise: *Show that with the V_i's defined as above, we can find $V_i'' \subset V_i'$ irreducible such that*

$$V_0'' \subsetneq \ldots \subsetneq V_{n-1+d}'' \subset V.$$

To do so, notice that for $i \leq n-1$,V_i' is already irreducible, so there is nothing to do. Prove then the result by induction on d.

Proof of the product theorem using Larsen-Pink. From the dimension computations above, we obtain using Larsen–Pink:

(1) For all tori T,
$$|A^{10} \cap T| \ll_n K^{O_n(1)}|A|^{\frac{n-1}{n^2-1}} = K^{O_n(1)}|A|^{\frac{1}{n+1}}.$$
(2) For $g \in SL_n(\bar{k})$ regular,
$$|A^{10} \cap Conj(g)| \ll_n K^{O_n(1)}|A|^{\frac{n^2-n}{n^2-1}} = K^{O_n(1)}|A|^{\frac{n}{n+1}}$$
(3) If S denotes the subvariety of non-regular elements and T a torus
$$|A^{10} \cap S| \ll_n K^{O_n(1)}|A|^{\frac{n^2-2}{n^2-1}}$$
and
$$|A^{10} \cap S \cap T| \ll_n K^{O_n(1)}|A|^{\frac{n-2}{n^2-1}}.$$
Where the second inequality is a consequence of the fact the $S \cap T$ is a proper subvariety of T and, hence, has dimension at most $n-2$ (Lemma 8.2).

These inequalities are sufficiently strong to show the following dichotomy:

Claim 9.4. *For any torus T :*

(i) either,
$$K^{O_n(1)}|A|^{\frac{1}{n+1}} \ll_n |A^2 \cap T| \ll_n K^{O_n(1)}|A|^{\frac{1}{n+1}};$$
(ii) or,
$$|A^2 \cap T| \ll_n K^{O_n(1)}|A|^{\frac{n-2}{n^2-1}}.$$

Proof of the Claim. If $T \cap A^2$ does not contain a regular element, then $T \cap A^2 \subset T \cap S \cap A^2$. So
$$|T \cap A^2| \ll_n K^{O_n(1)}|A|^{\frac{n-2}{n^2-1}}$$
according to (3) at the start of the proof. So the inequality (ii) follows from Proposition 5.6.

If $T \cap A^2$ contains a regular element γ, then set $\phi : a \in A \mapsto a\gamma a^{-1} \in Conj(\gamma)$ denote the restriction to A of the conjugation map. We have that $\phi(A) \subset Conj(\gamma) \cap A^4$. Since $|Conj(\gamma) \cap A^4| \ll_n K^{O_n(1)}|A|^{\frac{n}{n+1}}$, there is $a_0 \in A$ such that

$$|\phi^{-1}(\{\phi(a_0)\})| \gg_n K^{O_n(1)}|A|^{1-\frac{n}{n+1}} = K^{O_n(1)}|A|^{\frac{1}{n+1}}.$$

Notice now that if $a, b \in A$ satisfy $\phi(a) = \phi(b)$, then $a\gamma a^{-1} = b\gamma b^{-1}$ so

$$b^{-1}a \in Z(\gamma).$$

So $|Z(\gamma) \cap A^2| \gg_n K^{O_n(1)}|A|^{\frac{1}{n+1}}$. □

□