Lectures on Addi(c)tive Combinatorics

Philippe Michel May 27, 2025

Table of Content

Chapter 1. Basic on sumsets 5
1.1. Examples from number theory 5
1.2. Sumsets in $\mathbb{R}$ 6
1.3. Sumsets in $\mathbb{F}_p$ 7

Chapter 2. Some applications of additive combinatorics 13
2.1. The sum-product phenomenon 13
2.2. Fourier theory for finite abelian groups 15
2.3. Equidistribution and Exponential sums 19

Chapter 3. Growth in groups 25
3.1. Approximate subgroups 25
3.2. The commutative case 29

Chapter 4. Growth vs Energy 33
4.1. Basic properties of the energy 33
4.2. The Balog-Szemeredi-Gowers theorem(s) 34
4.3. The approximate subgroup recognition criterion 37
4.4. Proof of the BSG Theorem (set-theoretic version) 39

Chapter 5. The Sum-Product Theorem 47
5.1. Rough notations and Ruzsa calculus 47
5.2. Warm-up: growth in $\mathbb{F}_p$ under addition and multiplication 50
5.3. The BBSG theorem 52
5.4. Proof of the sum-product Theorem 55

Chapter 6. Growth in $\mathrm{SL}_2(\mathbb{F}_p)$ 57
6.1. A spectral gap property for $\mathrm{SL}_2(\mathbb{F}_p)$ 58
6.2. The Larsen-Pink inequalities 61
6.3. Structure of $\mathrm{SL}_2(k)$ 63
6.4. Special Larsen-Pink inequalities for $\mathrm{SL}_2(\overline{k})$ 65
6.5. Larsen-Pink inequalities for approximate subgroups 72

Chapter 7. Expansion in $\mathrm{SL}_2(\mathbb{F}_p)$ 79
7.1. Basic on graphs 79
7.2. Expander graphs 82
7.3. Expansion in Cayley graphs 86
7.4. The Bourgain-Gamburd expansion machine 90
7.5. Implementing the Bourgain-Gamburg expansion machine 95

Appendix : Harmonic analysis for finite groups 99

7.6. Representations of a finite group 99
7.7. Matrix coefficients 100

Reference 105

CHAPTER 1

Basic on sumsets

Additive combinatorics aims at studying the following kind of very basic

QUESTION. *Given $(G,+)$ a commutative group noted additively and $A,B\subset G$ two non-empty subsets, how big is the subset*

$$A+B=\{a+b,\ a\in A,b\in B\}$$

in terms of $|A|$ and $|B|$?

DEFINITION 1.1. *A subset of the shape*

$$A+B=\{a+b,\ a\in A,b\in B\}$$

for $\emptyset\neq A,B\subset G$ two non-empty subsets is called a sumset of G.

We have the following basic bounds

$$\max(|A|,|B|)\leqslant |A+B|\leqslant |A||B|. \tag{1.1}$$

The upper bound is obvious and the lower bound follow from the fact that for any $a\in A,\ b\in B$

$$A+b,a+B\subset A+B \text{ and } |A+b|=|A|,|a+B|=|B|;$$

we would like to know whether these bounds are sharp and what can be said if $|A+B|$ is very small or very large ("large" or "small" would have to be made more precise if A or B are infinite sets).

The following notations will be useful:

- $(k)A=\{a_1+\cdots+a_k,\ a_1,\cdots,a_k\in A\}\subset G$.
- $k\diamond A=\{k.a=a\cdots+a\ (k \text{ times}),\ a\in A\}\subset (k)A$
- $A+b=A+\{b\}$.
- $A-B=A+(-B),\ -B=\{-b,\ b\in B\}$
- If G is a ring with multiplication noted . and $\xi\in G$ we will write

$$\xi.A=\{\xi.a,\ a\in A\}.$$

1.1. Examples from number theory

For $G=\mathbb{Z}$ examples of sumset problems come from number theory:

Waring's type problems.

THEOREM (Lagrange $4-\square$ Theorem).

$$4.\square(\mathbb{Z})=4\{n^2,\ n\in\mathbb{Z}\}=\mathbb{Z}_{\geqslant 0}.$$

ie. every non-negative integer is the sum of at most four squares.

For $k \geqslant 2$ let $g(k) \in \mathbb{N}$ defined by the property

$$g(k) = \min(g \geqslant 1 \text{ such that } \mathbb{Z}_{\geqslant 0} = g\{n^k,\ n \in \mathbb{N}\}).$$

ie. every non-negative integer is the sum of at most g k-th powers of integers: due to the efforts of many people we we know that $g(k) < \infty$ for any$k \geqslant 2$ and that

$$g(2) = 4,\ g(3) = 9,\ g(4) = 19,\ g(5) = 37.$$

Goldbach's type problems. Let $\mathcal{P}$ be the set of prime numbers

THEOREM (Schnirelman). *There exists $G \geqslant 1$ such that*

$$G(\mathcal{P} \cup \{0\}) = \mathbb{Z}_{\geqslant 0}$$

ie. every non-negative integer is the sum of at most G primes.

THEOREM (Vinogradov-Helfgott). *We have $G \leqslant 4$*

CONJECTURE (Goldbach). *We have $G \leqslant 3$.*

1.2. Sumsets in $\mathbb{R}$

PROPOSITION 1.2. *For $G = \mathbb{R}$; suppose A and B finite. We have*

$$|A + B| \geqslant |A| + |B| - 1.$$

PROOF. This is obvious if $|A|$ or $|B| = 1$

Suppose $|A|, |B| \geqslant 2$ and write $A = \{a_1 < a_2 \cdots < a_m\},\ B = \{b_1 < b_2 \cdots < b_n\}$ then $A + B$ contains

$$a_1 + b_1 < a_2 + b_1 < \cdots < a_m + b_1 < a_m + b_2 < \cdots < a_m + b_n.$$

□

It is no difficult to provide examples for which the upper bound in (1.1) attained : for instance for any integer $N \geqslant 2$ let

$$A = \{1, \cdots, N-1\},\ B = \{N, \cdots, (N-1)N\}$$

then

$$A + B = \{m + nN,\ m, n \in \{1, \cdots, N-1\}\}$$

has

$$|A + B| = (N-1)^2.$$

Nevertheless the lower bound is still sharp and it is possible to characterise the A, B such that $|A + B|$ is as small as possible.

PROPOSITION 1.3. *Suppose that*

$$|A + B| = |A| + |B| - 1$$

then there exists $a, b \in \mathbb{R}$ and $q \in \mathbb{N}_{\geqslant 1}$ such that

$$A = a + q[0, m),\ B = b + q[0, n).$$

We then say that A and B are arithmetic progressions with the same common difference/modulus q.

PROOF. Suppose $|A|, |B| \geqslant 2$ (or this is obvious) we have

$$\begin{aligned} A+B &= \{c_1 < c_2 < \cdots < c_{m+n-1}\} \\ &= \{a_1+b_1 < a_2+b_1 < \cdots < a_m+b_1 < a_m+b_2 < \cdots < a_m+b_n\} \\ &= \{a_1+b_1 < a_1+b_2 < \cdots < a_1+b_n < a_2+b_n < \cdots < a_m+b_n\} \end{aligned}$$

So we have

$$a_2+b_1 = a_1+b_2,\ a_3+b_1 = a_1+b_3,\ \cdots$$

so that

$$a_2-a_1 = b_2-b_1,\ a_3-a_1 = b_3-b_1, \cdots$$

and from there we can conclude that A and B are arithmetic progressions with the same modulus. □

In particular A and B are "intervals" of translates of the subgroup $q\mathbb{Z} \subset \mathbb{R}$.

Observe in general that if $H \subset G$ is a subgroup we have

$$H+H = H.$$

This simple Proposition illustrate the following phenomenon that we will find again in other occasions

A sumset with small size admits some "structure" related to the group law.

1.3. Sumsets in $\mathbb{F}_p$

We now consider the case of G a finite commutative group. We have the following easy

LEMMA 1.4. *Suppose that $|A|+|B| > |G|$ then*

$$A+B = G.$$

PROOF. For any $g \in G$ $A+g$ and $-B$ must intersect so $\exists a \in A,\ b \in B$ such that $a-g = -b$ and

$$g = a+b.$$

□

The simplest finite commutative groups are the cyclic ones $G = \mathbb{Z}/N\mathbb{Z}$. When $N = p$ is a prime we have the

THEOREM 1.5 (Cauchy-Davenport). *Given $A, B \subset \mathbb{Z}/p\mathbb{Z}$. We have*

$$|A+B| \geqslant \min(p, |A|+|B|-1).$$

1.3.1. First proof. Replacing B by the translate $B-b$ for some $b \in B$ we may assume that $0 \in B$. In particular $A \subset A+B$.

If $|B| = 1$ then $A+B = A$ and we are done.

We proceed by induction on $|B|$ and can assume $|B| \geqslant 2$. Moreover by the previous Lemma we may assume that

$$2 \leqslant |A| \leqslant p-2.$$

Suppose that $A = A+B$ then for any $b \in B-\{0\}$ we have $A = A+b$ since $A+b \subset A+B$ and has cardinality $|A| = |A+B|$. In particular b is in the stabilizer of A under the translation action of $\mathbb{Z}/p\mathbb{Z}$ on the set of subsets of $\mathbb{Z}/p\mathbb{Z}$ and since $\mathbb{Z}/p\mathbb{Z}$ has no non-trivial subgroups (p is prime) this stabilizer is $\mathbb{Z}/p\mathbb{Z}$ and $A = \mathbb{Z}/p\mathbb{Z} = A+B$.

Suppose now that $A \neq A+B$: there exists $a_0 \in A$ such that

$$B_0 = \{b \in B,\ a_0+b \notin A\} \neq \emptyset.$$

In particular $A \cap (a_0 + B_0) = \emptyset$ and $0 \notin B_0$.

Let

$$\delta_{a_0}(A) := A \sqcup (a_0 + B_0),\ \delta_{a_0}(B) = B \backslash B_0.$$

We have

$$|\delta_{a_0}(A)| + |\delta_{a_0}(B)| = |A| + |B_0| + |B| - |B_0| = |A| + |B|,$$

$0 \in \delta_{a_0}(B)$ and $|\delta_{a_0}(B)| < |B|$ so by induction

$$|\delta_{a_0}(A) + \delta_{a_0}(B)| \geqslant \min(p, |\delta_{a_0}(A)| + |\delta_{a_0}(B)| - 1) = \min(p, |A| + |B| - 1).$$

Now

$$\delta_{a_0}(A) + \delta_{a_0}(B) = \{a + b',\ a \in A, b' \in B \backslash B_0\} \cup \{a_0 + b + b',\ b \in B_0,\ b' \in B \backslash B_0\} \subset A \cup B.$$

For the first set this is obvious and for the second we have

$$a_0 + b + b' = (a_0 + b') + b$$

and $a_0 + b' \in A$ since $b' \in B \backslash B_0$. □

Remark 1.1. The transformation

$$(A, B) \mapsto (\delta_{a_0}(A), \delta_{a_0}(B))$$

is called the *Dyson transform* at e of (A, B) after Freeman Dyson and is a synthesis of a multitude of previous ad-hoc looking arguments.

1.3.2. Second proof. Here we use explicitly the fact that $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ is a field. This is a example of the so-called *polynomial method*.

We start with the following

Theorem 1.6. *Let k be a field and $P(X_1, \cdots, X_n) \in k[X_1, \cdots, X_n]$. Let $X_1^{d_1} \cdots X_n^{d_n}$ be a monomial of degree*

$$\sum_i d_i = \deg(P) = d$$

having a non zero coefficient in P. For any tuple of subsets $(A_i)_{i \leqslant n}$, $A_i \subset k$ satisfying

$$|A_i| > d_i,\ i = 1, \cdots, n$$

there exists $(a_1, \cdots a_n) \in A_1 \times \cdots A_n$ such that $P(a_1, \cdots a_n) \neq 0$.

Remark 1.2. If $n = 1$ this is the simple fact that a polynomial of degree d has at most d roots in k.

Proof. We proceed by induction on d: if $d = 0$ we are done.

Suppose that $d \geqslant 1$; let $(d_1, \cdots, d_n)$ satisfying the assumption of the Theorem. WLOG-WMA $d_1 \geqslant 1$.

Let $(A_i)_{i \leqslant n}$ satisfying the assumption of the theorem. In particular $|A_1| \geqslant 2$.

Given $a_1 \in A_1$ we have

$$P(X_1, \cdots X_n) = Q(X_1, \cdots X_n)(X_1 - a) + R(X_2, \cdots, X_n)$$

with $\deg Q < d$ and $R(X_2, \cdots, X_n) = P(a_1, X_2, \cdots, X_n)$.

By our assumption $Q(X_1, \cdots X_n)$ has a non-zero coefficient in the monomial $X_1^{d_1 - 1} \cdots X_n^{d_n}$ and $\deg Q = d - 1$.

Since

$$P(a_1, a_2, \cdots a_n) = R(a_2, \cdots a_n)$$

we may assume that for any $(a_2, \cdots a_n) \in A_2 \times \cdots A_n$ we have $R(a_2, \cdots a_n) = 0$ (otherwise we are done by taking $(a_1, a_2, \cdots a_n)$ such that $R(a_2, \cdots a_n) \neq 0$).

Observe that $A_1 - \{a_1\} \neq \emptyset$. By induction on $\deg Q$ there exists

$$(a_1', \cdots a_n) \in (A_1 - \{a_1\}) \times A_2 \times \cdots A_n$$

such that

$$Q(a_1', \cdots a_n) \neq 0$$

and therefore

$$P(a_1', \cdots a_n) = Q(a_1', \cdots a_n)(a_1' - a_1) \neq 0.$$

□

PROOF. (of Cauchy-Davenport) We may assume that $|A| + |B| \leqslant p$.

We proceed by contradiction and assume that

$$|A + B| \leqslant |A| + |B| - 2.$$

Let $m = |A| + |B| - 2 - |A + B| \geqslant 0$ and

$$P(X, Y) = (X + Y)^m \prod_{c \in A+B} (X + Y - c).$$

This is a polynomial of degree $|A| + |B| - 2$ such that for any $(a, b) \in A \times B$

$$P(a, b) = 0.$$

In addition the coefficient of

$$X^{|A|-1} Y^{|B|-1}$$

in P is that of $X^{|A|-1} Y^{|B|-1}$ in $(X + Y)^{|A|+|B|-2}$ and equals the binomial coefficient $\binom{|A|+|B|-2}{|A|-1}$ mod p which is non zero since $|A| + |B| - 2 < p$. This is the contradiction. □

1.3.3. Applications. We have the following immediate extension

THEOREM 1.7 (Cauchy-Davenport). *Given $k \geqslant 2$ and $A_1, \cdots, A_k \subset \mathbb{Z}/p\mathbb{Z}$. We have*

$$|A_1 + \cdots + A_k| \geqslant \min(p, |A_1| + \cdots + |A_k| - k + 1).$$

In particular taking $A_i = A$ we set that

$$|kA| = |A + \cdots + A(k \text{ times})| \geqslant \min(p, k(|A| - 1) + 1)$$

In particular if $k \geqslant \frac{p-1}{|A|-1}$ then $kA = \mathbb{F}_p$.

A more arithmetic application is the following theorem

THEOREM 1.8 (Lagrange). *Given $\alpha, \beta \in \mathbb{F}_p^\times$. For any $x \in \mathbb{F}_p$ there exists $u, v \in \mathbb{F}_p$ such that*

$$x = \alpha u^2 + \beta v^2.$$

PROOF. Suppose $p > 2$. Let $\square(\mathbb{F}_p) = \{u^2,\ u \in \mathbb{F}_p\}$ the set of squares in $\mathbb{F}_p$. We have

$$\square(\mathbb{F}_p) = 1 + \frac{p-1}{2}$$

and

$$|\alpha\square(\mathbb{F}_p) + \beta\square(\mathbb{F}_p)| \geqslant \min(p, p) = p.$$

□

1.3.4. Optimality of CD. Again one may want to compare with the trivial upper bound

$$|A+B| \leqslant \min(|A||B|, p).$$

and again it turn out that most of the time this upper bound is closer to the truth.

THEOREM 1.9. *for any $1 \leqslant m, n \leqslant p-1$ there exists $A, B \subset \mathbb{F}_p^\times$ with $|A| = m,\ |B| = n$ such that*

$$|A+B| \geqslant \frac{1}{2}\min(|A||B|, p-1) = \frac{1}{2}\min(mn, p-1).$$

PROOF. The difficulty in proving that $A+B$ is large comes from the possibility $x \in A+B$ may have a lot of representations in the form $x = a+b$. So we introduce the number of representations

$$r_{A,B}(x) = \sum_{\substack{(a,b)\in A\times B\\ a+b=x}} 1.$$

We have

$$x \in A+B \Longleftrightarrow r_{A,B}(x) \geqslant 1$$

and we have

$$\sum_{x\in\mathbb{F}_p} r_{A,B}(x) = \sum_{x\in\mathbb{F}_p} \sum_{\substack{(a,b)\in A\times B\\ a+b=x}} 1 = |A\times B| = |A||B|.$$

By CS we have (write $r_{A,B}(x) = r_{A,B}(x).1_{A+B}$)

$$\sum_{x\in\mathbb{F}_p} r_{A,B}(x) = \|r_{A,B}.1_{A+B}\|_1 \leqslant \|1_{A+B}\|_2 \|r_{A,B}\|_2 = |A+B|^{1/2}(\sum_{x\in\mathbb{F}_p} r_{A,B}^2(x))^{1/2}.$$

So that

$$|A+B| \geqslant \frac{(|A||B|)^2}{\sum_{x\in\mathbb{F}_p} r_{A,B}^2(x)}$$

The sum

$$\sum_{x\in\mathbb{F}_p} r_{A,B}^2(x) = |\{(a,b,a',b') \in (A\times B)^2,\ a+b = a'+b'\}| = E(A,B)$$

is called the *additive energy* of the pair (A,B) and the smaller additive nrj is the large $|A+B|$.

To find a pair with low additive nrj we will look within the family of "deformations" of $A+B$, namely

$$A+\xi B,\ \xi \in \mathbb{F}_p^\times.$$

For this we evaluate the "expectation" of the random variable

$$\xi \mapsto E(A, \xi B).$$

$$\begin{aligned}\mathbb{E}(E(A,\bullet B)) &= \frac{1}{p-1}\sum_{\xi\in\mathbb{F}_p^\times}|\{(a,b,a',b')\in(A\times B)^2,\ a+b=a'+b'\}|\\ &= \frac{1}{p-1}|\{(a,b,a',b',\xi)\in(A\times B)^2\times\mathbb{F}_p^\times,\ a-a'=\xi(b'-b)\}|\\ &= |A||B|+\frac{(|A|^2-|A|)(|B|^2-|B|)}{p-1}\\ &\leqslant |A||B|(1+\frac{|A||B|}{p-1})\end{aligned}$$

Here the first term comes from the diagonal solutions $a=a',\ b=b'$ and the second from the non-diagonal solutions (since then ξ is determined by (a,b,a',b') Since $E(A,\bullet B)\geqslant 1$ there exists ξ such that

$$E(A,\xi B)\leqslant |A||B|(1+\frac{|A||B|}{p-1})$$

and

$$|A+\xi B|\geqslant \frac{|A||B|}{1+\frac{|A||B|}{p-1}}\geqslant \frac{1}{2}\frac{|A||B|}{\max(1,\frac{|A||B|}{p-1})}\geqslant \frac{1}{2}\min(|A||B|,p-1).$$

□

REMARK 1.3. This proof has introduced two important concepts that we will meet again:

(1) the additive energy $E(A,B)$ of two sets and
(2) the *probabilistic method* which allows to exhibit objects satisfying some generic property without expliciting them but rather because they form a set of of positive measure within a probability space.

An important generalisation of Cauchy-Davenport's theorem is Kneser's (which in fact implies Cauchy-Davenport

THEOREM 1.10 (Kneser). *Given G a commutative group and $A,B\in G$ non-empty subsets and let*

$$H=\mathrm{Stab}_G(A+B)=\{h\in G,\ h+A+B=A+B\}$$

the stabilizer of the set $A+B$ under the translation action of G. We have

$$|A+B|\geqslant |A+H|+|B+H|-|H|.$$

EXERCISE 1.1. Prove that the Cauchy-Davenport is not true for a composite N but extends as follows

THEOREM 1.11 (Cauchy-Davenport). *Suppose that $0\in B$ and for any $b\in B-\{0\}$, $(b,N)=1$. We have*

$$|A+B|\geqslant \min(N,|A|+|B|-1).$$

EXERCISE 1.2. Prove that for $A,B\subset\mathbb{F}_p$

$$|\{a+b,\ (a,b)\in A\times B,\ ab\neq 1\}|\geqslant \min(p,|A|+|B|-3)$$

CHAPTER 2

Some applications of additive combinatorics

2.1. The sum-product phenomenon

Theorem 1.9 gives examples of sumset in $\mathbb{F}_p$ whose size is growing *multiplicatively* with the sizes of the summands. Notice that this construction (which is not explicit) makes use of the existence of the *multiplication* in $\mathbb{F}_p$. The sum-product theorem discovered by Bourgain, Katz and Tao indeed shows that the combitation of addition and multiplication indeed conduct to growth:

THEOREM 2.1 (Sum-Product theorem). *For any $\varepsilon > 0$ there exists C, δ such that for any prime p and any subset $A \subset \mathbb{F}_p^\times$ satisfying*

$$C \leqslant |A| \leqslant p^{1-\varepsilon}$$

one has then

$$|A + A| + |A.A| \geqslant |A|^{1+\delta}.$$

The sum product theorem state that in the commutative ring $\mathbb{F}_p$ a subset exhibit polynomial growth either under addition or multiplication with itself unless (perhaps) it is already quite big ($|A| \geqslant p^{1-\varepsilon}$). Moreover under iterated addition/multiplication A grow to a polynomial size in p after only $O(\log p)$ steps (instead of a linear in p steps with CD).

While the sum-product theorem involves only the commutative ring $\mathbb{F}_p$ this case be interpreted in terms of the presence of growth in a non-commutative (but solvable) finite group: namely the affine group of matrices

$$\mathrm{Aff}_2(\mathbb{F}_p) = \{\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix},\ y \in \mathbb{F}_p^\times,\ x \in \mathbb{F}_p\} = N(\mathbb{F}_p) \ltimes A(\mathbb{F}_p)$$

$$N(\mathbb{F}_p) = \{n(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},\ x \in \mathbb{F}_p\},\ A(\mathbb{F}_p) = \{a(y) = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix},\ y \in \mathbb{F}_p^\times\}.$$

Indeed multiplication in N induces addition while conjugation by A in N induces multiplication:

$$n(x)n(x') = n(x + x),\ a(y).n(x).a(y)^{-1} = n(yx).$$

In fact the sum-product theorem viewed in this light was an important ingredient in the work of Helfgott who exhibited polynomial growth for another (this time highly non-sovable, in fact simple) matrix group:

THEOREM 2.2 (Helfgott). *There exists $k, \delta > 0$ such that for any p and any $A \subset \mathrm{SL}_2(\mathbb{F}_p)$ generating $\mathrm{SL}_2(\mathbb{F}_p)$, one of the following holds*

$$|A^{(3)}| \geqslant |A|^{1+\delta} \text{ or } (A \cup A^{-1} \cup \{\mathrm{Id}_2\})^{(k)} = \mathrm{SL}_2(\mathbb{F}_p).$$

2.1.1. Application to Cayley graphs. Recall that given $(G,.)$ a finite group and $A \subset G$ a non-empty subset, its Cayley graph $\mathrm{Cayley}(G,A)$ is the graph whose vertices are the elements of G and whose edges are the pairs of the shape

$$(g, a.g),\ g \in G.$$

The graph $\mathrm{Cayley}(G,A)$ is connected iff $\langle A\rangle = G$ and has no selfloop iff $e_G \notin A$. Also we say that A is symetric iff

$$A^{-1} = \{a^{-1},\ a \in A\} = A.$$

Recall that a graph $\Gamma = (V,E)$ is equipped with a natural distance on V:

$$d_\Gamma(x,y) := \text{ minimal number of edges necessary to connect } x \text{ and } y$$

and its diameter

$$\mathrm{diam}(\Gamma) = \max_{x,y\in V} d_\Gamma(x,y).$$

Corollary 2.3. *There exists $C \geqslant 1$ such that for $A \subset \mathrm{SL}_2(\mathbb{F}_p)$ a generating subset $\langle A\rangle = \mathrm{SL}_2(\mathbb{F}_p)$ one has*

$$\mathrm{diam}(\mathrm{Cayley}(\mathrm{SL}_2(\mathbb{F}_p), A)) \leqslant C(\log p)^C.$$

Proof. Assume that A is symetric and $\mathrm{Id}_2 \notin A$ (see the exercise for the general case); we have $|A| \geqslant 2$. Observe that for any symetric set B we have

$$(B \cup \{e\})^{(n)} \subset (B \cup \{e\})^{(n+1)}.$$

Apply Helfgott's theorem $j \geqslant 1$ times we have either

$$|A^{(3^j)}| \geqslant |A|^{(1+\delta)^j}$$

or

$$(A^{(3^j)} \cup \{\mathrm{Id}_2\})^{(k)} = \mathrm{SL}_2(\mathbb{F}_p).$$

In particular for

$$j = [\log(\frac{4\log p}{\log 2})/\log(1+\delta)] + 1$$

we have

$$(A^{(3^j)} \cup \{\mathrm{Id}_2\})^{(k)} = \mathrm{SL}_2(\mathbb{F}_p)$$

so that

$$\mathrm{diam}(\mathrm{Cayley}(\mathrm{SL}_2(\mathbb{F}_p), A)) \leqslant 3^j k = k(\frac{4\log p}{\log 2})^{O(1/\delta)}.$$

□

Example 2.1. On can take (Exercise)

$$A_1 = \{\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix}\}$$

since A_1 generate $\mathrm{SL}_2(\mathbb{Z})$. In that case there is another proof using the theory of modular forms to obtain a stronger result.

On the other hand one can also take

$$A_3 = \{\begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & -3\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 3 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ -3 & 1\end{pmatrix}\}$$

since $A_3 \pmod p$ generates $\mathrm{SL}_2(\mathbb{F}_p)$ for any $p > 3$ but in that case the theory of modular forms is not available (as $\langle A_3\rangle \subset \mathrm{SL}_2(\mathbb{Z})$ has infinite index in $\mathrm{SL}_2(\mathbb{Z})$).

2.2. Fourier theory for finite abelian groups

Let $(G,+)$ be a finite commutative group, the group of characters of G, $\widehat{G}$, is the set of group morphisms between G and the multiplicative group $(\mathbb{C}^\times,\times)$:

$$\widehat{G} = \mathrm{Hom}_{Gr}(G,\mathbb{C}^\times) = \{\chi : G \mapsto \mathbb{C}^\times,\ \chi(g+g') = \chi(g)\chi(g')\}.$$

This is a subgroup of the multiplicative group of functions $\mathcal{F}(G;\mathbb{C}^\times)$.

THEOREM 2.4 (Fourier theory). *Let $(G,+)$ be a finite commutative group; its group of characters $\widehat{G} = \mathrm{Hom}_{Grp}(G,\mathbb{C}^\times)$ is finite and has order*

$$|\widehat{G}| = |G|$$

and in fact is isomorphic (non-canonically) to G.

Moreover $\widehat{G}$ is an orthonormal basis of the Hilbert space $\mathcal{F}(G;\mathbb{C})$ when equipped with the hermitian product

$$\langle f_1, f_2\rangle_G = \frac{1}{|G|}\sum_{g\in G} f_1(g)\overline{f}_2(g).$$

In particular we have

$$\forall \chi,\psi\in\widehat{G},\ \langle\chi,\psi\rangle = \frac{1}{|G|}\sum_{g\in G}\chi(g)\overline{\psi}(g) = \delta_{\chi=\psi},\ \langle\chi,1\rangle = \frac{1}{|G|}\sum_{g\in G}\chi(g) = \delta_{\chi=1}.$$

For any $f\in\mathcal{F}(G;\mathbb{C}^\times)$ we have the Fourier decomposition

$$f = \sum_{\chi\in\widehat{G}}\langle f,\chi\rangle\chi.$$

The inner product

$$\langle f,\chi\rangle = \frac{1}{|G|}\sum_{g\in G} f(g)\overline{\chi}(g)$$

is called the χ-th Fourier coefficient of f and the rescaled function

$$\widehat{f} : \chi\in\widehat{G} \to |G|^{1/2}\langle f,\chi\rangle = \frac{1}{|G|^{1/2}}\sum_{g\in G} f(g)\overline{\chi}(g)$$

is called the Fourier transform of f.

The Fourier transform map

$$\widehat{\bullet} : f\in\mathcal{F}(G;\mathbb{C}) \mapsto \widehat{f}\in\mathcal{F}(\widehat{G};\mathbb{C})$$

is an isometry: ie. we have the Plancherel formula

$$\langle f_1,f_2\rangle_G = \frac{1}{|G|}\sum_{g\in G} f_1(g)\overline{f_2}(g) = \frac{1}{|\widehat{G}|}\sum_{\chi\in\widehat{G}}\widehat{f}_1(\chi)\overline{\widehat{f}_2}(\chi) = \langle\widehat{f}_1,\widehat{f}_2\rangle_{\widehat{G}},$$

and after identifying (canonically) the bidual $\widehat{\widehat{G}}$ with G via the evaluation map

$$\mathrm{ev}_\bullet : g\mapsto \mathrm{ev}_g : \chi\mapsto \mathrm{ev}_g(\chi) = \chi(g)$$

the Fourier transform is an anti-involution:

$$\widehat{\widehat{f}}(g) = f(-g).$$

PROOF. The group G act on itself by (right or left) translation and therefore act lineairy on $V := \mathcal{F}(G;\mathbb{C})$: we denote by t_g the corresponding action:

$$t_g f(\bullet) = f(\bullet + g).$$

Moreover by a change over variable the inner product $\langle f_1, f_2\rangle_G$ is translation invariant: $\forall g \in G$

$$\langle t_g f_1, t_g f_2\rangle_G = \langle f_1, f_2\rangle_G.$$

In particular t_g is adjoint ($t_g^* = t_{-g}$) and therefore diagonalisable. Moreover since all the $t_g,\ g \in G$ commute with one another, they are simultaneously diagonalisable. We write the (orthogonal) eigenspace decomposition

$$\mathcal{F}(G;\mathbb{C}) = \bigoplus_{\chi} V_\chi$$

where V_χ is the common eigenspace associated with the system of eigenvalues noted $\chi(g),\ g \in G$:

$$V_\chi = \{f \in V,\ \forall g \in G,\ t_g f = \chi(g) f\}$$

Since $t_{g+g'} = t_g \circ t_{g'}$ we have

$$\chi(g+g') = \chi(g)\chi(g'),\ \chi(e_G) = 1$$

so that the eigenvalue function $\chi : g \to \chi(g)$ is a character of G. Moreover $\chi \in V_\chi$:

$$\forall g \in G,\ t_g \chi = \chi(\bullet + g) = \chi(g)\chi(\bullet) \in V_\chi.$$

conversely, given $f \in V_\chi$ we have

$$t_{g'} f(g) = f(g+g') = (t_g f)(g') = \chi(g) f(g')$$

so that

$$t_{g'} f = \chi(g') f = f(g')\chi \iff f = (f(g')/\chi(g'))\chi.$$

In follows that

$$V_\chi = \mathbb{C}.\chi$$

and that

$$V = \bigoplus_{\chi \in \widehat{G'}} \mathbb{C}.\chi$$

where $\widehat{G'} \subset \widehat{G}$ is a subset of characters. Since any character $\psi \in \widehat{G}$ is an eigenfunction of all t_g (with eigenvalues $\psi(g),\ g \in G$) we conclude that the eigenspace decomposition is made of one-dimensional eigenspaces indexed by all the characters of G

$$V = \bigoplus_{\chi \in \widehat{G}} \mathbb{C}.\chi.$$

From this we conclude that

$$|\widehat{G}| = |G| = \dim V$$

and that the set $\widehat{G}$ form an orthogonal family (since different eigenspaces are mutually orthogonal) and hence an orthogonal basis. Moreover

$$\langle \chi, \chi\rangle = \frac{1}{|G|}\sum_{g \in G} |\chi(g)|^2 = 1$$

since $|\chi(g)| = 1$ (this follows either from the fact that the $\chi(g)$ are eigenvalues of isometries or from Lagrange's theorem $\chi(g)^{|G|} = \chi(g^{|G|}) = \chi(e_G) = 1$).

The Fourier decomposition is a direct consequence of the fact that $\widehat{G}$ form an orthonormal basis of V. The Plancherel formula as well (since $|\widehat{G}| = |G|$)

$$\langle f_1, f_2\rangle_G = \sum_{\chi,\psi\in\widehat{G}}\sum \langle f_1,\chi\rangle\overline{\langle f_2,\psi\rangle}\langle\chi,\psi\rangle_G = \sum_{\chi,\psi\in\widehat{G}}\sum \langle f_1,\chi\rangle\overline{\langle f_2,\psi\rangle}\delta_{\chi=\psi}$$

$$= \sum_{\chi\in\widehat{G}}\langle f_1,\chi\rangle\overline{\langle f_2,\chi\rangle} = \frac{1}{|G|}\sum_{\chi\in\widehat{G}}\widehat{f}_1(\chi)\overline{\widehat{f}_2(\chi)} = \langle\widehat{f}_1,\widehat{f}_2\rangle_{\widehat{G}}.$$

Finally we have for $\psi\in\widehat{G}$

$$\widehat{\psi}(\chi) = |G|^{1/2}\langle\psi,\chi\rangle = |G|^{1/2}\delta_{\psi=\chi}$$

whose Fourier transform is given by

$$|G|^{1/2}\frac{1}{|\widehat{G}|^{1/2}}\sum_{\chi}\delta_{\psi=\chi}\mathrm{ev}_g(\overline{\chi}) = \overline{\psi}(g) = \psi(-g)$$

and we conclude by linearity.

The fact the $\widehat{G}$ is (non-canonically) isomorphic to G follows from the fact that if

$$\varphi : G\simeq G_1\times G_2$$

is isomorphic to a product then we have an isomorphism $\widehat{\varphi}:\widehat{G}_1\times\widehat{G}_2\simeq\widehat{G}$ given by

$$\widehat{\varphi}(\chi_1,\chi_2)(g) = \chi_1(g_1)\chi_2(g_2)$$

where $\varphi(g) = (g_1,g_2)\in G_1\times G_2$.

This reduce the proof to the case of $G = g^{\mathbb{Z}}$ a cyclic group (with generator g). Since a character is completely determined by its values at g, the evaluation map

$$\mathrm{ev}_g : \chi\in\widehat{G}\mapsto\chi(g)\in\mu_{|G|}$$

is an isomorphism and $\mu_{|G|}\simeq\mathbb{Z}/|G|\mathbb{Z}\simeq G$. □

Example 2.2. For the additive group of congruences modulo q, $(\mathbb{Z}/q\mathbb{Z},+)$ we have

$$\widehat{\mathbb{Z}/q\mathbb{Z}} = \{e_q(a\bullet),\ a = 0,\cdots,q-1\}$$

where for $h\,(\mathrm{mod}\,q)\in\mathbb{Z}/q\mathbb{Z}$

$$e_q(ah) = \exp(2\pi i\frac{ah}{q}).$$

We call these the additive characters modulo q.

Notice that $e_q(a\bullet)$ depends only on $a\,(\mathrm{mod}\,q)$ and the isomorphism $\mathbb{Z}/q\mathbb{Z}\simeq\widehat{\mathbb{Z}/q\mathbb{Z}}$ is given by

$$a\,(\mathrm{mod}\,q)\mapsto e_q(a\bullet).$$

To see that we have indeed an isomorphism we observe that this is a group morphism whose kernel is $\{0\,(\mathrm{mod}\,q)\}$ so it is injective and since both groups have the same size it is surjective.

For the multiplicative group of congruences $((\mathbb{Z}/q\mathbb{Z})^\times,\times)$ which has order $\varphi(q)$ the group $(\widehat{\mathbb{Z}/q\mathbb{Z}})^\times$ is the group of Dirichlet characters of modulo q. They are is much less explicit than $\widehat{\mathbb{Z}/q\mathbb{Z}}$.

2.2.1. Convolution. We also recall the convolution operator: for $f_1, f_2 \in \mathcal{F}(G;\mathbb{C})$

$$f_1 \star f_2(g) := \frac{1}{|G|^{1/2}} \sum_{g_1+g_2=g} f_1(g_1) f_2(g_2).$$

We them have

$$\widehat{f_1 \star f_2}(\chi) = \frac{1}{|G|^{1/2}} \sum_{g\in G} f_1 \star f_2(g)\chi(g) = \frac{1}{|G|^{1/2}|G|^{1/2}} \sum_{g_1,g_2\in G} f_1(g_1) f_2(g_2)\chi(g_1+g_2) = \widehat{f_1}(\chi)\widehat{f_2}(\chi).$$

For instance if $f_1 = 1_A$, $f_2 = 1_B$ then

$$|G|^{1/2} 1_A \star 1_B(g) = r_{A,B}(g) = |\{(a,b) \in A \times B,\ a+b=g\}|$$

and

$$\widehat{r_{A,B}}(\chi) = |G|^{1/2} \widehat{1_A}(\chi)\widehat{1_B}(\chi)$$

where

$$\widehat{1_A}(\chi) = \frac{1}{|G|^{1/2}} \sum_{a\in A} \chi(a).$$

By Plancherel formula (or a direct computation) we have also an expression for the additive nrj:

$$E(A,B) = \sum_{g\in G} |r_{A,B}(g)|^2 = \sum_{\chi} |\widehat{r_{A,B}}(\chi)|^2 = |G| \sum_{\chi} |\widehat{1_A}(\chi)|^2 |\widehat{1_B}(\chi)|^2$$

2.2.2. Restriction to a subgroup.

THEOREM 2.5. *Let G be a finite commutative group, $\widehat{G}$ its group of characters, $H \subset G$ a subgroup and let*

$$H^\perp = \{\chi \in \widehat{G},\ \forall h \in H,\ \chi(h) = 1\} \subset \widehat{G}.$$

We have the exact sequence

$$1 \to H^\perp \to \widehat{G} \to \widehat{H} \to 1 \tag{2.1}$$

where the third arrow is the restriction to H. In particular

$$|H^\perp| = |\widehat{G}|/|\widehat{H}| = |G|/|H|.$$

We have

$$\frac{1}{|H|} \sum_{h\in H} \chi(h) = \frac{|G|^{1/2}}{|H|} \widehat{1_H}(\chi) = 1_{H^\perp}(\chi).$$

More generally, for any function $f : G \to \mathbb{C}$ we have

$$\frac{1}{|H|} \sum_{h\in H} f(h) = \sum_{\chi\in H^\perp} \langle f, \chi\rangle = |G|^{1/2} \sum_{\chi\in H^\perp} \widehat{f}(\chi).$$

REMARK 2.1. We have that

$$\widehat{1_H} = \frac{|H|}{|G|^{1/2}} 1_{H^\perp},$$

ie. the Fourier transform of the characteristic function of a sous-group $H \subset G$ is proportional to the characteristic function of a subgroup $H^\perp \subset \widehat{G}$. Moreover the larger H is the smaller the support of the Fourier transform: $H^\perp$ has order the index of H in G.

PROOF. It is clear that the restriction to H map

$$\bullet_{|H} : \chi \in \widehat{G} \to \chi_{|H} \in \widehat{H}.$$

is a group morphism whose kernel is $H^\perp$. Let us prove it is surjective.

Given any $\psi \in \widehat{H}$ let

$$V_\psi := \{f : G \to \mathbb{C},\ \forall h \in H,\ t_h f = \psi(h) f\}.$$

This space is non-zero: the function ψ extended by 0 outside H has this property. Moreover this space is invariant under translation by elements of G: given such an f we have

$$t_h(t_g f) = t_g(t_h f) = \psi(h) t_g f$$

since t_g and t_h commute. In particular there exist in V_ψ a common eigenspace V_χ for some a character $\chi \in \widehat{G}$ and in particular $\chi \in V_\psi$: for all $h \in H$ we have

$$\chi(h+g) = \chi(h)\chi(g) = \psi(h)\chi(g)$$

and therefore $\chi_{|H} = \psi$. From this we obtain the surjectivity of the restriction map hence (2.1).

We have for any $\chi \in \widehat{G}$

$$\frac{1}{|H|} \sum_{h \in H} \chi(h) = \delta_{\chi \in H^\perp}$$

Indeed χ being contained in $H^\perp$ or not is equivalent to $\chi_{|H}$ being a trivial character on H or not.

By Fourier decomposition we have

$$\frac{1}{|H|} \sum_{h \in H} f(h) = \sum_{\chi \in \widehat{G}} \langle f, \chi \rangle \frac{1}{|H|} \sum_{h \in H} \chi(h) = \sum_{\chi \in H^\perp} \langle f, \chi \rangle.$$

□

2.3. Equidistribution and Exponential sums

Consider the "circle" $X = \mathbb{R}/\mathbb{Z}$ which we may identify with the semi-open interval $[0, 1)$.

Suppose given a sequence of finite subsets $(H_n)_{n \geqslant 1}$, $H_n \subset X$ with $|H_n| \to \infty$. We often would like to know how the "image" of $H_n \subset X$ evolves as $n \to \infty$. For instance given some interval $[a, b] \subset \mathbb{R}/\mathbb{Z}$ how many elements of H_n belong to I asymptotically as $n \to \infty$; it is then natural to look on whether there is a limit to the sequence

$$P_n([a, b]) := \frac{|\{h \in H_n,\ h \in [a, b]\}|}{|H_n|},$$

the proportions of elements of H_n contained in $[a, b]$.

DEFINITION 2.6. *A sequence of finite sets $(H_n)_n$ becomes uniformly distributed on X (or equidistributed modulo 1) if for any $0 \leqslant a < b \leqslant 1$*

$$P_n([a, b]) \to b - a = \mu_{Leb}([a, b]).$$

EXAMPLE 2.3. For instance the sequence $H_n = \{h/n,\ 0 \leqslant h \leqslant n-1\}$ become equidistributed modulo 1 as $n \to \infty$.

Notice that if $(H_n)_n$ becomes equidistributed then for any $C \geqslant 1$ any sequence of subset $(C_n)_n$ satisfying

$$C_n \subset H_n,\ |C_n| \leqslant C$$

then $(H_n \backslash C_n)_n$ becomes also equidistributed.

For instance the sets $\{h/n,\ 1 \leqslant h \leqslant n-1\}$ become equidistributed modulo 1.

Notice that the proportion can be rewritten

$$P_n([a,b]) = \frac{1}{|H_n|} \sum_{h \in H_n} 1_{[a,b]}(h) \to b - a = \mu_n(1_{[a,b]})$$

where μ_n denote the probability measure on X given by

$$\mu_n(f) := \frac{1}{|H_n|} \sum_{h \in H_n} f(h)$$

or in other terms the measure

$$\mu_n = \frac{1}{|H_n|} \sum_{h \in H_n} \delta_h.$$

By approximating characteristic functions of intervals by continuous functions and continuous functions by linear combination of characteristic functions of intervals we see (exercise) that uniform distribution is equivalent to showing that for any $f \in \mathcal{C}(X)$ we have

$$\mu_n(f) \to \int_X f(x)dx = \mu_{Leb}(f),\ n \to \infty; \tag{2.2}$$

in other term, this is equivalent to the weak-$\star$ convergence of the sequence of probability measure $(\mu_n)_n$ towards the Lebesgue measure.

THEOREM 2.7 (Weyl equidistribution criterion). *Given*

$$H_n \subset X,\ n \geqslant 1,\ |H_n| \to \infty.$$

TFAE

1. *the sequence $(H_n)_n$ becomes equidistributed.*
2. *For any $a \in \mathbb{Z} - \{0\}$*

$$\frac{1}{|H_n|} \sum_{h \in H_n} \exp(2\pi i a h) \to 0.$$

PROOF. We need only to prove (2) $\Rightarrow$ (1). By approximating continuous functions by smooth functions it is equivalent to prove (2.2) for $f \in \mathcal{C}^\infty(X)$ a smooth function. By Fourier theory we have

$$f(h) = \sum_{a \in \mathbb{Z}} \widehat{f}(a) \exp(2\pi i a h)$$

where (integration par parts)

$$\widehat{f}(a) = \int_{[0,1)} f(x) \exp(-2\pi i a x)dx \ll_f \frac{1}{1 + |a|^{2025}}.$$

In particular

$$\sum_{a \in \mathbb{Z}} |\widehat{f}(a)| < \infty$$

and since for any $a \in \mathbb{Z}$,

$$|\mu_n(\exp(2\pi i a \bullet))| = |\frac{1}{|H_n|} \sum_{h \in H_n} \exp(2\pi i a h)| \leqslant 1$$

we have by Fubini

$$\mu_n(f) = \sum_{a\in\mathbb{Z}} \widehat{f}(a)\frac{1}{|H_n|}\sum_{h\in H_n}\exp(2\pi iah) = \sum_{a\in\mathbb{Z}}\widehat{f}(a)\mu_n(\exp(2\pi ia\bullet)).$$

Given $\varepsilon > 0$ we have

$$\mu_n(f) = \widehat{f}(0) + \sum_{1\leqslant|a|\leqslant 1/\varepsilon}\widehat{f}(a)\mu_n(\exp(2\pi ia\bullet)) + \sum_{|a|>1/\varepsilon}\widehat{f}(a)\mu_n(\exp(2\pi ia\bullet)).$$

The first term is

$$\widehat{f}(0) = \int_{[0,1)} f(x)dx.$$

The third term is bounded by

$$|\sum_{|a|>1/\varepsilon}|\widehat{f}(a)||\mu_n(\exp(2\pi ia\bullet))| \ll \varepsilon^{2024}.$$

while for the second term there is $n(\varepsilon)$ such that for $n \geqslant n(\varepsilon)$ we have

$$\forall 1 \leqslant |a| \leqslant 1/\varepsilon,\ |\mu_n(\exp(2\pi ia\bullet))| \leqslant \varepsilon.$$

It follows that

$$|\sum_{1\leqslant|a|\leqslant 1/\varepsilon}\widehat{f}(a)\mu_n(\exp(2\pi ia\bullet))| \leqslant \varepsilon\sum_a|\widehat{f}(a)| \ll_f \varepsilon.$$

This shows that

$$\mu_n(f) \to \int_{[0,1)} f(x)dx.$$

□

2.3.1. Equidistribution of the multiplicative subgroups of $\mathbb{F}_p^\times$. We now consider the finite field $\mathbb{F}_p$. Any element of $\mathbb{F}_p$ is a congruence class $h \pmod p$ and we can associated to it

$$\frac{h}{p} \pmod 1 \in \mathbb{R}/\mathbb{Z}.$$

We know already that as $p \to \infty$ the sets

$$\{\frac{h}{p} \pmod 1,\ h \in \mathbb{F}_p\}$$

become equidistributed modulo 1, as does the image of the multiplicative group $\mathbb{F}_p^\times$

$$\{\frac{h}{p} \pmod 1,\ h \in \mathbb{F}_p^\times\}.$$

We would like to understand equidistribution modulo 1 but for a (strict) subgroup of the multiplicative group

$$H_p \subset \mathbb{F}_p^\times$$

satisfying $|H_p| \to \infty$.

Example 2.4. Take $H_p = \square(\mathbb{F}_p^\times)$ the subgroup of squares. If p is odd then $|H_p| = \frac{p-1}{2}$ indeed H_p is the image of the morphism $x \mapsto x^2$ whose kernel is $\{\pm 1\}$ (alternatively $H_p^\perp = \{1, (\frac{\bullet}{p})\}$ is the subgroup of order 2 generated by the Legendre symbol); however even if H_p is commensurable with the size of $\mathbb{F}_p^\times$, its equidistribution is not obvious.

Notice that H_p is cyclic (since $\mathbb{F}_p$ is a field): so

$$H_p = \{\xi_p^n,\ n \in \mathbb{Z}\}$$

and we want to study the distribution of

$$\{\frac{\xi_p^n}{p} \,(\mathrm{mod}\, 1),\ n = 1, \cdots, |H_p|\} \subset \mathbb{R}/\mathbb{Z}.$$

but this does not help much since it is hard to anticipate the variations of the function

$$n \mapsto \frac{\xi_p^n}{p} \,(\mathrm{mod}\, 1)$$

(even if ξ_p is say $2 \,(\mathrm{mod}\, p)$).

THEOREM 2.8. *For any $p \geqslant 3$ let $H_p \subset \mathbb{F}_p^\times$ be a multiplicative subgroup of order*

$$|H_p| \geqslant p^{1/2+\varepsilon}$$

then

$$\{\frac{h}{p} \,(\mathrm{mod}\, 1),\ h \in H_p\} \subset \mathbb{R}/\mathbb{Z}$$

is equidistributed as $p \to \infty$.

PROOF. By Weyl equidistribution criterion, it would be sufficient to show that the Weyl's sum converge to 0, forall $a \in \mathbb{Z} - \{0\}$

$$\frac{1}{|H_p|} \sum_{h \in H_p} \exp(2\pi i a \frac{h}{p}) \to 0.$$

Observe that

$$\exp(2\pi i a \frac{h}{p}) = e_p(ah)$$

is an additive character of $\mathbb{F}_p$, in particular it depends only on $a \,(\mathrm{mod}\, p)$. This character maybe trivial but this is the case iff $a \equiv 0 \,(\mathrm{mod}\, p)$ so, as long as $a \neq 0$ and $p > |a|$, $a \not\equiv 0 \,(\mathrm{mod}\, p)$ and $e_p(a\bullet)$ is a non-trivial additive character. We then have

THEOREM 2.9. *Let $H \subset \mathbb{F}_p^\times$ be a subgroup; given $\psi \in \widehat{\mathbb{F}_p}$ an additive character let*

$$\mu_H(\psi) = \frac{1}{H} \sum_{h \in H} \psi(h).$$

If $\psi \neq 1$ we have

$$|\mu_H(\psi)| \leqslant \frac{p^{1/2}}{|H|}.$$

In particular if

$$|H| \geqslant p^{1/2+\varepsilon}$$

for some $\varepsilon > 0$ we have for any $\psi \in \widehat{\mathbb{F}_p} - \{1\}$

$$\mu_H(\psi) \to 0.$$

REMARK 2.2. The sum $\mu_H(\psi)$ is an average of complex numbers of modulus 1 so the fact that $\mu_H(\psi) \to 0$ indicate the presence of oscillations in the values of the additive character along the multiplicative subgroup H.

PROOF. We have

$$\frac{1}{|H|}\sum_{h\in H}\psi(h)=\sum_{\chi\in H^\perp}\langle\psi,\chi\rangle=\sum_{\chi\in H^\perp}\frac{1}{p-1}\sum_{x\in\mathbb{F}_p^\times}\psi(x)\overline{\chi}(x).$$

Set

$$G(\psi,\overline{\chi})=\sum_{x\in\mathbb{F}_p^\times}\psi(x)\overline{\chi}(x).$$

For $\psi\neq 1$ we have

$$G(\psi,1)=-1,\ |G(\psi,\overline{\chi})|=p^{1/2}\text{ for }\chi\neq 1.$$

Indeed if $\chi\neq 1$, $G(\psi,\overline{\chi})$ is a Gauss sum (and a Ramanujan sum if $\chi=1$).

This implies that

$$|\mu_H(\psi)|\leqslant\frac{p^{1/2}}{p-1}|H^\perp|=\frac{p^{1/2}}{p-1}\frac{p-1}{|H|}=\frac{p^{1/2}}{|H|}.$$

□

REMARK 2.3. The character $\psi:\mathbb{F}_p\mapsto\mathbb{C}^\times$ take values in the complex numbers of modulus 1 and if $\psi\neq 1$ satisfies

$$\frac{1}{p}\sum_{x\in\mathbb{F}_p}\psi(x)=0,$$

hence it has a lot of oscillations has x varies in $\mathbb{F}_p$.

The fact that

$$\frac{1}{|H|}\sum_{h\in H}\psi(h)\to 0$$

if $|H|\geqslant^{1/2+\varepsilon}$ shows that some of these oscillation are still present when ψ is restricted to H even if $|H|$ is as small as $p^{1/2+1/2025}$ (since the given bound improves over the trivial bound 1). This is a feature of the additive nature of ψ versus the multiplicative structure of H. For instance, there may exist non-trivial multiplicative characters χ for which $\sum_{h\in H}\chi(h)=|H|$: the non-trivial characters in $H^\perp$!

Using techniques from additive combinatorics Bourgain-Gilibichuk-Konyagin obtained a considerable improvement on the possible size of H and showed that ψ continue to oscillate even along extremely small subgroups.

THEOREM 2.10 (Bourgain-Gilibichuk-Konyagin). *For any $\varepsilon>0$, there exists $\delta>0$ such that for any prime p, any non-trivial additive character ψ and any multiplicative subgroup $H\subset\mathbb{F}_p^\times$ satisfying*

$$|H|\geqslant p^\varepsilon \tag{2.3}$$

we have

$$\mu_H(\psi)\ll p^{-\delta}.$$

COROLLARY 2.11. *Given any $\varepsilon>0$. For any $p\geqslant 3$ let $H_p\subset\mathbb{F}_p^\times$ be a multiplicative subgroup of order*

$$|H_p|\geqslant p^\varepsilon$$

then

$$\{\frac{h}{p}\,(\mathrm{mod}\,1),\ h\in H_p\}\subset\mathbb{R}/\mathbb{Z}$$

is equidistributed as $p\to\infty$.

REMARK 2.4. This is all the more remarkable as the condition

$$|H| \geqslant p^{1/2+\varepsilon}$$

is a recurrent assumption in analytic number theory appearing in many contexts: it is called the *Polyà-Vinogradov range*.

CHAPTER 3

Growth in groups

In this chapter we introduce more sophisticated technique to identify growth for product of subsets $A, B \subset G$ in a finite group (not necessarily commutative).

We have already noted that if $H \subset G$ is a subgroup

$$H.H = H.$$

So H has no growth.

The converse is true:

PROPOSITION 3.1. *If $A \subset G$ satisfies $|A.A| = |A|$ and $e \in A$ then A is a subgroup of G.*

PROOF. Since $e \in A$ we have

$$A.e = A \subset A.A$$

and since $|A.e| = |A.A|$ we have $A = A.A$. In particular

$$A \subset H = \mathrm{Stab}_G(A) = \{g \in G,\ gA = A\}$$

but $He = H \subset A = A$ so that $H = A$. □

More generally we have the following proposition (left as an exercise)

PROPOSITION 3.2. *Let G be a finite group, $A, B \subset G$ non-empty.*
Let

$$H = \mathrm{Stab}_G(B) = \{g \in G,\ gB = B\}$$

If $|A.B| = |B|$ then there exists $g_0 \in G$ such that $A \subset g_0.H$ and $B = H.X$ for $X \subset G$.

We will show that a subset A exhibing a slow growth (like $A^{(2)}$ or $A^{(3)}$ is not much larger than A) then A is "close" to being a subgroup.

3.1. Approximate subgroups

DEFINITION 3.3 (Tao). *Let $K \geqslant 1$ and G a group. A finite set A is a K-approximate subgroup if*

1. $e \in A$,
2. $A = A^{-1}$
3. *There exist $X \subset A$ symmetric $|X| \leqslant K$ such that*

$$A.A \subset X.A.$$

.

REMARK 3.1. A 1-approximate subgroup is a subgroup.

If G is finite any set A is a $|G|$-approximate subgroup. So interesting notions of approximate subgroup occur when $K > 1$ and is small compared to $|G|$.

The following is obvious:

LEMMA 3.4. *Let A be a K approximate subgroup. We have*

$$|A^{(n)}| \leqslant K^{n-1}|A|$$

It is quite remarkable that there is a converse:

THEOREM 3.5. *If $e \in A = A^{-1}$ and*

$$|A^{(3)}| \leqslant K|A|$$

then $A^{(3)}$ is a K'-approximate subgroup with $K' \leqslant 2K^5$.

REMARK 3.2. The important feature is that K' is bounded polynomially in K.

REMARK 3.3. The exponent (3) is optimal and cannot be replaced by (2) excepted when G is commutative (see below).

One general reason explaining why commutativity may play a key role is the following situation: take $A = H$ a subgroup and $B = gH$ a coset; If G is commutative we have

$$HgH = gHH = gH$$

has order $|H|$.

On the other hand if G is not commutative then

$$HgH \neq H$$

excepted when g is in the normalizer of H.

The other extreme is when $H \cap gHg^{-1} = \{e\}$: let us consider the multiplicities of the elements in the double coset HgH. Suppose that we have

$$h_1 g h_1' = h_2 g h_2',$$

we then have

$$h_2{}^{-1} h_1 = g h_2' h_1'{}^{-1} g^{-1} \Longrightarrow \ h_1 = h_2, \ h_1' = h_2'$$

and therefore $|HgH| = |H|^2$.

This argument show more generally that

$$|HgH| = |H|^2/|H \cap gHg^{-1}|.$$

A central tool in the proof of Theorem 3.5 is the notion of

DEFINITION 3.6 (Ruzsa distance). *Let $A, B \subset G$ be nonempty finite sets, their Ruzsa distance is defined as*

$$d(A,B) = \log\left(\frac{|A.B^{-1}|}{\sqrt{|A||B|}}\right)$$

or equivalently

$$|A.B^{-1}| = \sqrt{|A||B|}\exp(d(A,B)).$$

This is not exactly a distance but allmost:

LEMMA 3.7. *We have*

$$d(A,B) \geqslant 0, \ d(A,B) = d(B,A), \ d(A,C) \leqslant d(A,B) + d(B,C).$$

PROOF. Non-negativity: since

$$|A.B^{-1}| \geqslant \max(|A|,|B|) \geqslant \sqrt{|A||B|}$$

so we have

$$d(A,B) \geqslant 0$$

The symmetry follows from the fact that

$$|AB^{-1}| = |(AB^{-1})^{-1}| = |BA^{-1}|.$$

For the triangle inequality, we have

$$d(A,B) + d(B,C) = \log(|A.B^{-1}||B.C^{-1}|/\sqrt{|A||B|^2|C|})$$

so proving that this is $\geqslant d(A,C)$ is equivalent to showing that

$$|A.B^{-1}||B.C^{-1}| \geqslant |B||A.C^{-1}|.$$

To prove this inequality it suffice to construct an injection

$$A.C^{-1} \times B \hookrightarrow A.B^{-1} \times B.C^{-1}$$

For this we choose a section of the surjective map

$$(a,c) \in A \times C \mapsto ac^{-1} \in AC^{-1}$$

ie. a map

$$s : u \in AC^{-1} \mapsto s(u) = (a(u), c(u))$$

such that

$$a(u)c(u)^{-1} = u.$$

We then set

$$\iota(u,b) = (a(u)b^{-1}, bc(u)^{-1}) = (v(u), w(u)).$$

This map is injective since

$$v(u)w(u) = a(u)b^{-1}bc(u)^{-1} = u,\ b = w(u)c(u).$$

□

REMARK 3.4. However the Ruzsa distance is not exactly a distance since it is often the case that

$$d(A,A) = \log(|A.A^{-1}|/|A|) \neq 0.$$

Moreover if $A = gH$ for $g \neq e$ and $B = H$ then $d(A,B) = 0$ although $A \neq B$.

However notice A is symmetric and contain e we have

$$d(A,A) = \log(|A.A|/|A|) = 0 \iff |A.A| = |A| \iff A.A = A$$

so A is a subgroup.

For the proof of the theorem, we need two more results also due to Ruzsa:

LEMMA 3.8 (Controlled Growth Lemma). *Suppose that $e \in A = A^{-1}$ and let*

$$K = |A^{(3)}|/|A|$$

then for any $n \geqslant 3$

$$|A^{(n)}|/|A| \leqslant K^{n-2}.$$

PROOF. By induction: we have $A^{(n+1)} = A^{(n-1)}A^{(2)}$ and hence

$$\frac{|A^{(n+1)}|}{|A|} = \frac{\sqrt{|A^{(n-1)}||A^{(2)}|}}{|A|}\exp(d(A^{(n-1)}, A^{(2)}))$$
$$\leqslant \frac{\sqrt{|A^{(n-1)}||A^{(2)}|}}{|A|}\exp(d(A^{(n-1)}, A) + d(A^{(2)}, A))$$
$$= \frac{\sqrt{|A^{(n-1)}||A^{(2)}|}}{|A|}\frac{|A^{(n)}|}{\sqrt{|A^{(n-1)}||A|}}\frac{|A^{(3)}|}{\sqrt{|A^{(2)}||A|}} \leqslant \frac{|A^{(n)}|}{|A|}K \leqslant K^{n-2}.$$

Here we have used that $A = A^{-1}$ so that

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

□

EXERCISE 3.1. Under the assumptions of the Controlled Growth Lemma, prove that for any sequence $(\varepsilon_i)_{1\leqslant i\leqslant n}$ the product set

$$A^{(\varepsilon_i)_{1\leqslant i\leqslant n}} = A^{(\varepsilon 1)}.\cdots.A^{(\varepsilon n)}$$

satisfies

$$|A^{(\varepsilon_i)_{1\leqslant i\leqslant n}}| \leqslant K^{n-2}|A|.$$

REMARK 3.5. One cannot replace (3) by (2) and $n-2$ by $n-1$ in the above.

Let us take consider H and g such that $g^2 = e$ and

$$H \cap gHg^{-1} = \{e\}$$

so that $|HgH| = H^2$ and suppose $|H|$ can be taken arbitrary large.

Let

$$A = H \cup \{g\};$$

then A is symmetric, contain e and

$$A^{(2)} \supset (Hg \cup gH) \sqcup H$$

has order $3|H| - 1$ (we have $Hg \cap gH = \{g\}$) so that

$$|A^{(2)}|/|A| = (3|H| - 1)/(|H| + 1) \leqslant 3.$$

On the other hand

$$|A^{(3)}|/|A| = |HgH|/(|H| + 1) \geqslant |H|^2/(|H| + 1) \geqslant |H|/2.$$

This would contradict a generalization with (2) replacing (3) as soon as $|H| \geqslant 18$.

EXAMPLE 3.1. For instance, one can take

$$H = \mathrm{Aff}_2(\mathbb{F}_p) \in \mathrm{GL}_2(\mathbb{F}_p),\ g = w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The final ingredient is

LEMMA 3.9 (Ruzsa Covering Lemma). *Suppose that*

$$|AB| \leqslant K|A|$$

there exists $X \subset B$ *such that* $|X| \leqslant K$ *and*

$$B \subset A^{-1}AX$$

PROOF. Let $X \subset B$ be maximal so that the $Ax,\ x \in X$ are disjoint.

We have

$$|A.X| = |A||X| \leqslant |AB| \leqslant K|A|$$

so that $|X| \leqslant K$. In addition by the maximality of X we have that for any $b \in B$

$$Ab \cap AX \neq \emptyset.$$

In other terms for any $b \in B$ there exists $a_1, a_2 \in A$ and $x \in X$ such that

$$a_1 b = a_2 x.$$

Hence

$$b = a_1^{-1} a_2 x \in A^{-1}.A.X.$$

□

PROOF. (of Theorem 3.5) Suppose $e \in A = A^{-1}$. Let $H = A^{(3)}$. We have $e \in H = H^{-1}$ and

$$|AH^{(2)}| = |A^{(7)}| \leqslant K^5|A|.$$

By the covering lemma there exist

$$X \in H^{(2)}$$

with $|X| \leqslant K^5$ such that

$$H^{(2)} \subset A^{-1}.A.X = A^{(2)}.X \subset A^{(3)}X = HX$$

(since $e \in A,\ A^{(2)} \subset A^{(3)}$).

If X is not symmetric we replace it with $X \cup X^{-1}$ whose size is bounded by $2K^5$. □

3.2. The commutative case

THEOREM 3.10 (Pluennecke). *Let $(G, +)$ be a commutative groupe. Suppose that*

$$|A + B| \leqslant K|A|$$

then for any $m, n \geqslant 0$

$$|mB - nB| \leqslant K^{m+n}|A|.$$

In particular if $|2A| \leqslant K|A|$ or $|A - A| \leqslant K|A|$ then $|mA - nA| \leqslant K^{m+n}|A|$.

PROOF. (after Petridis) Let

$$K' := \min_{\emptyset \neq A' \subset A} |A' + B|/|A'| \leqslant K.$$

We claim that for any $A' \subset A$ such that $|A' + B|/|A'| = K'$ and any $C \subset G$ we have

$$|A' + B + C| \leqslant K'|A' + C|.$$

Let us prove the theorem assuming the claim: by induction we have for any $n \geqslant 0$

$$|A' + nB| \leqslant K'^n|A'|.$$

Now

$$|mB - nB| = \sqrt{|mB||nB|}\exp(d(mB, nB))$$
$$\leqslant \sqrt{|mB||nB|}(\exp(d(mB, -A') + d(nB, -A'))$$

We have

$$d(mB, -A') = \log(|A' + mB|/\sqrt{|mB||A'|}) \leqslant \log(K'^m\sqrt{|A'|}/\sqrt{|mB|})$$

and

$$d(nB,-A')=\log(|A'+nB|/\sqrt{|nB||A'|})\leqslant \log(K'^n\sqrt{|A'|}/\sqrt{|nB|})$$

we obtain

$$|mB-nB|\leqslant K'^{m+n}\sqrt{|mB||nB|}\frac{|A'|}{\sqrt{|mB||nB|}}\leqslant K^{m+n}|A|.$$

It remains to prove the claim. We proceed by induction on $|C|$. This is obvious if $|C|=1$ since then

$$|A'+B+C|=|A'+B|\leqslant K'|A'|=K'|A'+C|.$$

Assume $|C|\geqslant 2$ Write $C=C'\sqcup\{c\}$. We have

$$A'+B+C=(A'+B+C')\cup((A'+B+c)\backslash(A'_{B,c}+B+c))$$

where

$$A'_{B,c}=\{a\in A',\ a+B+c\subset A'+B+C'\}.$$

We have by definition of A',

$$|A'_{B,c}+B|\geqslant K'|A'_{B,c}|$$

and by this lower bound an induction

$$\begin{aligned}|A'+B+C|&=|A'+B+C'|+|A'+B+c|-|A'_{B,c}+B+c|\\&=|A'+B+C'|+|A'+B|-|A'_{B,c}+B|\\&\leqslant K'|A'+C'|+K'|A'|-K'|A'_{B,c}|\\&=K'(|A'+C'|+|A'_{B,c}|-|A'_{B,c}|)\end{aligned}$$

We have

$$A'+C=A'+C'\cup((A'+c)\backslash(A'_c+c))$$

where

$$A'_c=\{a\in A',\ a+c\in A'+C'\}\subset A'_{B,c}$$

and

$$\begin{aligned}|A'+C|&=|A'+C'|+|A'+c|-|A'_c+c|\\&=|A'+C'|+|A'|-|A'_c|\\&\geqslant|A'+C'|+|A'|-|A'_{B,c}|.\end{aligned}$$

Hence

$$|A'+B+C|\leqslant K'(|A'+C'|+|A'|-|A'_{B,c}|)\leqslant K'|A'+C|.$$

□

Corollary 3.11. *Suppose that*

$$|A-A|\leqslant K|A|$$

then for any $m,n\geqslant 0$ there exists $X\subset G$ with $|X|\leqslant K^{m+n+1}$ such that

$$mA-nA\subset A-A+X.$$

Proof. We have by Pluennecke formula

$$|A + mA - nA| \leqslant K^{m+n+1}|A|$$

so that by Ruzsa covering Lemma there exists $X \subset mA - nA$ with $|X| \leqslant K^{m+n+1}$ such that

$$mA - nA \subset A - A + X.$$

□

CHAPTER 4

Growth vs Energy

DEFINITION 4.1. *Let G be a group (not necessarily commutative) and $A, B \subset G$ non-empty finite subsets. The (additive/multiplicative) energy*

$$E(A,B) = |\{(a_1,b_1,a_2,b_2) \in A\times B\times A\times B),\ a_1b_1 = a_2b_2\}|.$$

We set $E(A) = E(A,A)$ and the normalized energy is defined as

$$e(A,B) = \frac{E(A,B)}{(|A||B|)^{3/2}}.$$

4.1. Basic properties of the energy

4.1.1. Upper/Lower bounds.

$$E(A,B) \leqslant \min(|A||B|^2, |A|^2|B|) \leqslant \sqrt{|A||B|^2|A|^2|B|} = (|A||B|)^{3/2}.$$

(indeed once we choose 3 amongst (a_1,b_1,a_2,b_2) the fourth is determined). Notice also that by considering the diagonal elements (a,b,a,b) one has

$$E(A,B) \geqslant |A||B|,\ e(A,B) \geqslant 1/\sqrt{|A||B|}.$$

In other terms we have

$$e(A,B) \in [1/\sqrt{|A||B|}, 1].$$

4.1.2. Energy and product sets.

We also recall that

$$E(A,B) = \sum_{g\in G}(r_{A,B}(g))^2,\ r_{A,B}(g) = |\{(a,b)\in A\times B,\ ab = g\}|$$

so that by CS we have

$$(|A||B|)^2 = (\sum_{g\in G} r_{A,B}(g))^2 = (\sum_{g\in A.B} 1_{A.B}(g) r_{A,B}(g))^2 \leqslant \sum_{g\in A.B} 1_{A.B}(g) \sum_{g\in A.B} r^2_{A,B}(g)$$

Or in other terms

$$|A.B| \geqslant \frac{|A|^2|B|^2}{E(A,B)} = \frac{\sqrt{|A||B|}}{e(A,B)}.$$

which we rewrite

$$e(A,B) \geqslant \frac{\sqrt{|A||B|}}{|A.B|} \text{ or equivalently } \frac{|A.B|}{\sqrt{|A||B|}} \geqslant e(A,B)^{-1} \tag{4.1}$$

In particular pairs of subsets with small energy exhibit growth.

EXERCISE 4.1. Prove that if $e(A,B) \geqslant 1/K$, one has

$$K^{-2}|A| \leqslant |B| \leqslant K^2|A|$$

and that if $e(A,A) \leqslant 1/K$ then

$$|A^{(2)}| \geqslant K|A|.$$

4.1.3. Energy and symmetry. We don't have

$$E(A,B) = E(B,A)$$

in general, but since

$$a_1b_1 = a_2b_2 \Longleftrightarrow b_2^{-1}a_2^{-1} = b_1^{-1}a_1^{-1}$$

We have

$$E(A,B) = E(B^{-1},A^{-1}).$$

Also since

$$a_1a_2^{-1} = a_3a_4^{-1} \Longleftrightarrow a_2^{-1}a_4 = a_1^{-1}a_3$$

we have

$$E(A,A^{-1}) = E(A^{-1},A).$$

4.2. The Balog-Szemeredi-Gowers theorem(s)

It says that subsets with big energy are very structured. To get an idea let us look at the maximal case.

PROPOSITION 4.2. *Suppose that $e(A,B) = 1$, there exists a subgroup $H \subset G$ and $a,b \in G$ such that*

$$A = aH,\ B = Hb$$

PROOF. If $e(A,B) = 1$ we have

$$\min(|A||B|^2, |A|^2|B|) = (|A||B|)^{3/2}$$

so $|A| = |B|$. Moreover we also have $E(A,B) = |A||B|^2$. Since $E(A,B)$ count the number of quadruples (a_1,b_1,a_2,b_2) satisfying

$$a_1 = a_2b_2b_1^{-1}$$

we see that any such quadruple is of the shape $(a_2b_2b_1^{-1}, b_1, a_2, b_2)$. In particular for any $(a_2,b_1,b_2) \in A \times B \times B$ we have

$$a_2b_2b_1^{-1} \in A.$$

In other terms

$$B.B^{-1} \subset H = \mathrm{Stab}_{G,r}(A) = \{h \in G,\ Ah = A\}.$$

In particular $|A| = |B| \leqslant |H|$. On the other hand we have $|H| \leqslant |A|$ since $aH \subset A$ and $|A| = |H|$ and $A = aH$. Moreover for any $b \in B$ we have $Bb^{-1} \subset H$ or $B \subset Hb$ which implies $B = Hb$. □

There is an approximate subgroup analog of this result: it was first proven by Balog-Szemeredi and improved by Gowers in the commutative setting and the present non-commutative version is due to Tao:

THEOREM 4.3 (Balog-Szemeredi-Gowers, approximate group version). *There exists $C, d \geqslant 1$ such that if $A, B \subset G$ finite with $e(A,B) \geqslant 1/K$ for some $K \geqslant 1$ then there exist $a, b \in G$, a K'-approximate subgroup H with $K' \leqslant CK^d$, such that*

$$|H| \leqslant K'|A|,\ |A| \leqslant K'|A \cap aH|,\ |B| \leqslant K'|B \cap Hb|.$$

In particular setting

$$A' = A \cap aH,\ B' = B \cap bH$$

we have

$$|A'.B'| \leqslant |H| \leqslant K'|A| \leqslant K'K\sqrt{|A||B|}.$$

The approximate group version will be deduced from the following set-theoretic version:

THEOREM 4.4 (Balog-Szemeredi-Gowers, set-theoretic version). *Given $K \geqslant 1$ and $A, B \subset G$ finite with $e(A,B) \geqslant 1/K$, there exist $A''' \subset A,\ B''' \subset B$ such that*

$$|A'''| \gg |A|/K,\ |B'''| \gg |B|/K$$

and

$$|A'''.B'''| \ll K^8|A|^{1/2}|B|^{1/2}.$$

Here the implicit constant are absolute and explicitable.

To pass from the set-theoretic version to the approximate group version we will use the following

THEOREM 4.5 (Approximate subgroup recognition criterion). *There exists $C, d \geqslant 1$ such that if $A, B \subset G$ satisfy*

$$|A.B| \leqslant K|A|^{1/2}|B|^{1/2}$$

for some $K \geqslant 1$ then there exist $a, b \in G$, a K'-approximate subgroup H with $K' \leqslant CK^d$, such that

$$|H| \leqslant K'|A|,\ |A| \leqslant K'|A \cap aH|,\ |B| \leqslant K'|B \cap Hb|.$$

In particular setting

$$A' = A \cap aH,\ B' = B \cap Hb$$

we have

$$|A'.B'| \leqslant |H| \leqslant K'|A| \leqslant K'K\sqrt{|A||B|}.$$

First we will slightly change the language.

4.2.1. Haar measure notation. We denote by $\mu : \mathcal{F}(G;\mathbb{C}) \to \mathbb{C}$ the couting/uniform/Haar measure (when G is equipped with the discrete topology): the measure which gives mass 1 to any $g \in G$. In other terms

$$\mu(\{g\}) = 1,\ \mu(A) = |A|$$

and

$$\mu(f) = \int_G F(g)dg = \sum_{g \in G} f(g)$$

so that

$$\mu(A) = \mu(1_A).$$

This measure is up to scaling the unique left and right-invariant measure: if we note

$$g.f : h \mapsto f(hg),\ f_{|g} : h \mapsto f(gh)$$

the left and right translation actions, we have by a change of variable

$$\mu(f) = \mu(g.f) = \mu(f_{|g}).$$

It is also invariant under inversion $g \mapsto g^{-1}$:

$$\mu(A^{-1}) = \mu(A),\ \mu(f) = \mu(\tilde{f})$$

where

$$\tilde{f}(g) = f(g^{-1}).$$

This measure induces L^1 and L^2 norms on $\mathcal{F}(G;\mathbb{C})$

$$\|f\|_1 = \mu(|f|),\ \|f\|_2^2 = \mu(|f|^2).$$

We have

$$\|f\|_1,\ \|f\|_2 \leqslant \|f\|_\infty \mu(G).$$

Notice that the L^2-norm comes from the Hermitian product

$$\langle f_1, f_2\rangle := \mu(f_1.\overline{f_2}) = \int_G f_1(g)\overline{f_2}(g)dg = \sum_g f_1(g)\overline{f_2}(g).$$

(non-degenerate because $\mu(g) > 0$ for every $g \in G$) and we have the CS inequality

$$|\mu(f_1.f_2)|^2 = |\langle f_1, \overline{f_2}\rangle|^2 \leqslant \|f_1\|_2^2\|f_2\|_2^2 = \mu(|f_1|^2)|\mu(|f_2|^2).$$

Proof. Easy Exercise. □

The (left-invariant) Ruzsa distance is given by

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

It has the same properties as in the commutative case (same proofs) and satisfies for any $g \in G$

$$d(gA, gB) = d(A,B) = d(Ag, Bg).$$

4.2.2. Multiplicative energy and convolution. Given two functions $f_1, f_2 \in \mathcal{F}(G;\mathbb{C})$ their convolution is defined as

$$f_1 \star f_2(g) = \sum_{g_1 g_2 = g} f_1(g_1)f_2(g_2).$$

Remark 4.1. By comparison with §2.2.1 this definition of the convolution differs by a factor $\mu(G)^{1/2} = |G|^{1/2}$.

Proposition 4.6. *The convolution operation $\star$ is*

1. *Bilinear*
2. *Associative: $(f_1 \star f_2) \star f_3 = f_1 \star (f_2 \star f_3)$*
3. *The Dirac function δ_e is a neutral element:*

$$\delta_e \star f = f \star \delta_e = f.$$

4. *$\star$ is not commutative in general but we have*

$$f_1 \star \tilde{f_2}(e) = \mu(f_1 f_2) = \mu(f_2 f_1) = f_2 \star \tilde{f_1}(e)$$

with

$$\tilde{f} : g \mapsto f(g^{-1}).$$

More generally

$$\widetilde{f_2 \star f_1} = \tilde{f}_1 \star \tilde{f}_2$$

In particular multiplicative convolution give $\mathcal{F}(G;\mathbb{C})$ the structure of noncommutative but associative unital algebra.

We then have as in the commutative case

$$r_{A,B}(g) = 1_A \star 1_B(g)$$

and

$$\mu(1_A \star 1_B) = \mu(r_{A,B}) = \mu(A)\mu(B),\ E(A,B) = \mu(|1_A \star 1_B|^2).$$

Since

$$0 \leqslant r_{A,B}(g) \leqslant \min(\mu(A), \mu(B))$$

we have

$$E(A,B) \leqslant \min(\mu(A),\mu(B))\mu(1_A \star 1_B) \leqslant \mu(A)^{3/2}\mu(B)^{3/2}$$

and we retrieve the normalized energy

$$e(A,B) = \frac{E(A,B)}{\mu(A)^{3/2}\mu(B)^{3/2}} \in [1/\mu(A)^{1/2}\mu(B)^{1/2}, 1].$$

4.3. The approximate subgroup recognition criterion

We will need the following

LEMMA 4.7. *Let $K \geqslant 1$ and $A \subset G$ such that*

$$\mu(AA^{-1}) \leqslant K\mu(A).$$

There exists a symmetric set $e \in S \subset G$ such that

$$\mu(S) \geqslant \mu(A)/2K$$

and for any $n \geqslant 1$

$$\mu(AS^nA^{-1}) \leqslant 2^n K^{2n+1}\mu(A). \tag{4.2}$$

Let us first show how this lemma allows to prove Theorem 4.5.

4.3.1. Proof of Theorem 4.5. The main assumption of Theorem 4.5 states that

$$d(A, B^{-1}) \leqslant \log K.$$

By the triangle inequality we have

$$d(A,A) \leqslant d(A,B^{-1}) + d(B^{-1},A) = 2d(A,B^{-1}) \leqslant \log(K^2).$$

We have therefore

$$\mu(AA^{-1}) \leqslant K^2\mu(A).$$

By Lemma 4.7 there exist a symmetric set S such that

$$\mu(S) \geqslant \mu(A)/2K^2$$

and

$$\mu(AS^3A^{-1}) \leqslant 8(2K^2)^7\mu(A) \ll K^{14}\mu(A).$$

In particular

$$\mu(S), \mu(AS),\ \mu(S^3) \ll K^{14}\mu(A) \ll K^{16}\mu(S).$$

It follows that $H = S^3$ is a $K^{O(1)}$-approximate subgroup containing S. In particular

$$\mu(S.H) \leqslant K^{O(1)}\mu(H) \ll K^{O(1)}\mu(A) \ll K^{O(1)}\mu(S)^{1/2}\mu(H)^{1/2}.$$

This implies that $d(S,H) = O(1+\log K)$ hence $d(A,H) = O(1+\log K)$. By Ruzsa covering Lemma there exists X such that $|X| \ll K^{O(1)}$ such that $A \subset X.H.H$ and X' with $|X'| \ll K^{O(1)}$ with $A \subset X'.H$. We also have

$$d(B^{-1},H) \ll 1+\log K$$

and there exists Y' with $|Y'| \ll K^{O(1)}$ with $B^{-1} \subset Y'.H$. Taking

$$Z = X' \cup Y'^{-1}$$

we have

$$A \subset Z.H,\ B \subset HZ$$

and by the pigeonhole principle we can find $a, b \in Z$ such that

$$A \cap aH,\ B \cap Hb$$

have the required properties.

□

4.3.2. Proof of Lemma 4.7. We will construct S from $s \in G$ such that $A \cap As$ is large.

We have

$$\mu(A \cap As) = \int_G 1_A(g)1_A(gs^{-1})dg = 1_{A^{-1}} \star 1_A(s).$$

In particular

$$\int_G \mu(A \cap Ag)dg = \mu(A)^2.$$

We also have (use (4.1))

$$\int_G \mu(A \cap Ag)^2 dg = \int_G (1_{A^{-1}} \star 1_A)^2(g)dg = E(A, A^{-1}) \geqslant \frac{\mu(A)^4}{\mu(AA^{-1})} \geqslant \mu(A)^3/K. \tag{4.3}$$

Let

$$S := \{s \in G,\ \mu(A \cap Ag) \geqslant \mu(A)/2K\}.$$

The set S is symmetric ($\mu(A \cap Ag^{-1}) = \mu((A \cap Ag)g^{-1}) = \mu(A \cap Ag)$) and contains e.

Moreover we have

$$\int_{G-S} \mu(A \cap Ag)^2 dg \leqslant \frac{\mu(A)}{2K} \int_G \mu(A \cap Ag)dg \leqslant \frac{\mu(A)^3}{2K}$$

and using (4.3) we find

$$\int_S \mu(A \cap Ag)^2 dg \geqslant \frac{\mu(A)^3}{2K}$$

and since $\mu(A \cap Ag) \leqslant \mu(A)$ we have

$$\mu(S) \geqslant \frac{\mu(A)}{2K}.$$

The inequality (4.2) follows from evaluating

$$I_{A,S,n} := \int_{(AA^{-1})^{n+1}} 1_{AS^nA^{-1}}(g_0 \cdots g_n)dg_0 \cdots dg_n$$

We have

$$I_{A,S,n} \leqslant \mu(AA^{-1})^{n+1} \leqslant K^{n+1}\mu(A)^{n+1}.$$

On the other hand changing variable by putting

$$g = g_0 \cdots g_{n-1} g_n,$$

we have

$$I_{A,S,n} = \int_{AS^nA^{-1}} (\int_{(AA^{-1})^n} 1_{AA^{-1}}(g_{n-1}^{-1} \cdots g_0^{-1} g) dg_0 \cdots dg_{n-1}) dg.$$

Given any $g \in AS^nA^{-1}$ we write it as

$$g = a_0 s_1 \cdots s_n a_{n+1}.$$

Making the change of variable $(g_i)_{0 \leqslant i \leqslant n-1} \leftrightarrow (a_i)_{1 \leqslant i \leqslant n}$ with

$$g_0 = a_0 a_1^{-1},\ g_1 = a_1 s_1 a_2^{-1}, \cdots, g_{n-1} = a_{n-1} s_{n-1} a_n^{-1}$$

we have

$$g_{n-1}^{-1} \cdots g_0^{-1} g = a_n s_n a_{n+1}^{-1}$$

and

$$\begin{aligned}
&\int_{(AA^{-1})^n} 1_{AA^{-1}}(g_{n-1}^{-1} \cdots g_0^{-1} g) dg_0 \cdots dg_{n-1} \\
&\qquad = \int_{G^n} 1_{AA^{-1}}(g_{n-1}^{-1} \cdots g_0^{-1} g) \prod_{i=0}^{n-1} 1_{AA^{-1}}(g_i) dg_0 \cdots dg_{n-1} \\
&\qquad\qquad = \int_{G^n} 1_{AA^{-1}}(a_0 a_1^{-1}) \prod_{i=1}^{n} 1_{AA^{-1}}(a_i s_i a_{i+1}^{-1}) da_1 \cdots da_n.
\end{aligned}$$

If $a_1, \cdots, a_n \in A$ and $a_1 s_1, \cdots a_n s_n \in A$ the integrant equals 1 so the integral is lower bounded by

$$\int_{G^n} \prod_{i=1}^{n} 1_A(a_i) 1_A(a_i s_i) da_1 \cdots da_n = \prod_{i=1}^{n} \mu(A \cap As_i) \geqslant (\mu(A)/2K)^n$$

so that

$$K^{n+1}\mu(A)^{n+1} \geqslant I_{A,S,n} \geqslant (\mu(A)/2K)^n \mu(AS^nA^{-1}).$$

Simplify we obtain

$$\mu(AS^nA^{-1}) \leqslant 2^n K^{2n+1} \mu(A)$$

□

4.4. Proof of the BSG Theorem (set-theoretic version)

We now prove a set theoretic version:

4.4.1. First steps. We have by assumption

$$\int_G (1_A \star 1_B(g))^2 dg \geqslant \frac{\mu(A)^{3/2}\mu(B)^{3/2}}{K}.$$

Let

$$E := \{g \in G,\ 1_A \star 1_B(g) \geqslant \frac{\mu(A)^{1/2}\mu(B)^{1/2}}{2K}\}.$$

We have

$$\int_E \frac{\mu(A)^{1/2}\mu(B)^{1/2}}{2K} dg \leqslant \int_E 1_A \star 1_B(g) dg \leqslant \int_G 1_A \star 1_B(g) dg = \mu(A)\mu(B)$$

that is

$$\mu(E) \leqslant 2K\mu(A)^{1/2}\mu(B)^{1/2}. \tag{4.4}$$

On the other hand we have

$$\int_{G-E} (1_A \star 1_B)^2(g) dg \leqslant \frac{\mu(A)^{1/2}\mu(B)^{1/2}}{2K} \int_G (1_A \star 1_B)(g) dg = \frac{\mu(A)^{3/2}\mu(B)^{3/2}}{2K}$$

and therefore

$$\int_E (1_A \star 1_B)^2(g) dg \geqslant \frac{\mu(A)^{3/2}\mu(B)^{3/2}}{K} - \frac{\mu(A)^{3/2}\mu(B)^{3/2}}{2K} = \frac{\mu(A)^{3/2}\mu(B)^{3/2}}{2K}.$$

In particular E is non empty. Moreover since

$$1_A \star 1_B(g) \leqslant \min(\mu(A), \mu(B)) \leqslant \mu(A)^{1/2}\mu(B)^{1/2}$$

we have

$$\mu(A)^{1/2}\mu(B)^{1/2} \int_E 1_A \star 1_B(g) dg \geqslant \int_E (1_A \star 1_B)^2(g) dg \geqslant \frac{\mu(A)^{3/2}\mu(B)^{3/2}}{2K}$$

or

$$\int_E 1_A \star 1_B(g) dg \geqslant \frac{\mu(A)\mu(B)}{2K}$$

We can rewrite this inequality

$$\int_A (\int_B 1_E(ab) db) da \geqslant \int_A (\int_B \frac{1}{2K} db) da$$

Let

$$A' = \{a \in A,\ \int_B 1_E(ab) db \geqslant \mu(B)/4K\}. \tag{4.5}$$

Writing again

$$\int_A \cdots = \int_{A'} \cdots + \int_{A-A'} \cdots$$

and using that

$$\int_{A-A'} (\int_B 1_E(ab) db) da \leqslant \mu(A)\mu(B)/4K$$

we obtain

$$\int_{A'} (\int_B 1_E(ab) db) da \geqslant \mu(A)\mu(B)/4K \tag{4.6}$$

and since $1_E(ab) \leqslant 1$ we have

$$\mu(A')\mu(B) \geqslant \mu(A)\mu(B)/4K$$

or

$$\mu(A') \geqslant \mu(A)/4K.$$

Let

$$R := \mu(A)/\mu(A') \in [1, 4K].$$

From (4.4), we have

$$\mu(E) \leqslant 2KR^{1/2}\mu(A')^{1/2}\mu(B)^{1/2} \tag{4.7}$$

and (4.6) becomes

$$\mu \otimes \mu(\{(a,b) \in A' \times B,\ ab \in E\}) \geqslant \frac{R}{4K}\mu(A')\mu(B).$$

We will prove below the following

LEMMA 4.8 (weak BSG). *Let $K, K' \geqslant 1$ and $A, B, E \subset G$ such that*

$$\mu(E) \leqslant K'\mu(A)^{1/2}\mu(B)^{1/2}$$

and

$$\mu \otimes \mu(\{(a,b) \in A \times B,\ ab \in E\}) \geqslant \frac{1}{K}\mu(A)\mu(B)$$

then for any $\varepsilon \in (0,1)$, there exists $A' \subset A$ and $D \subset A.A^{-1}$ such that

$$\mu(A') \geqslant \frac{\mu(A)}{2K},\ \mu(D) \leqslant \frac{2(KK')^2}{\varepsilon}\mu(A)$$

and

$$\mu \otimes \mu(\{(a,a') \in A' \times A',\ a.a'^{-1} \in D\}) \geqslant (1-\varepsilon)\mu(A')^2.$$

Using this lemma with

$$\varepsilon = 1/32K,\ (A,B,E) \leftrightarrow (A',B,E) \text{ and } (K,K') \leftrightarrow (4K/R, 2KR^{1/2}),$$

we find that there exists $A'' \subset A' \subset A$ and D such that

$$\mu(A'') \gg \mu(A')R/K = \mu(A)/K,$$

$$\mu(D) \ll K^5\mu(A')/R \ll K^6\mu(A'')/R^2 \leqslant K^6\mu(A''). \tag{4.8}$$

and

$$\mu \otimes \mu(\{(a,a') \in A'' \times A'',\ a.a'^{-1} \in D\}) \geqslant (1-1/32K)\mu(A'')^2$$

which we can rewrite this using the complement set as

$$\int_{A''} \mu(a' \in A'',\ a \notin D.a')da \leqslant \mu(A'')^2/32K.$$

Let

$$A''' = \{a \in A'',\ \mu(a' \in A'',\ a \notin D.a') \leqslant \mu(A'')/16K\}.$$

For $a \in A'' - A'''$ we have

$$\mu(a' \in A'',\ a \notin D.a') \geqslant \mu(A'')/16K$$

and therefore We have

$$\frac{\mu(A'')}{16K}\mu(A''-A''') \leqslant \int_{A''-A'''} \mu(a'\in A'',\ a\notin D.a')da$$
$$\leqslant \int_{A''} \mu(a'\in A'',\ a\notin D.a')da \leqslant \frac{\mu(A'')^2}{32K}$$

so that

$$\mu(A''-A''') \leqslant \mu(A'')/2$$

and

$$\mu(A''') \geqslant \mu(A'')/2 \gg \mu(A)/K.$$

Since $A''\subset A'$ we have (see (4.5)) for any $a\in A''$

$$\int_B 1_E(ab)db \geqslant \mu(B)/4K.$$

Hence (by Fubini)

$$\int_B (\int_{A''} 1_E(ab)da)db = \int_{A''}(\int_B 1_E(ab)db)da \geqslant \frac{\mu(A'')\mu(B)}{4K}.$$

Let

$$B''' = \{b\in B,\ \int_{A''} 1_E(ab)da \geqslant \frac{\mu(A'')}{8K}\}.$$

Considering again the complement $B-B'''$ have

$$\int_{B-B'''}(\int_{A''} 1_E(ab)da)db \leqslant \frac{\mu(A'')\mu(B)}{8K}$$

so that

$$\int_{B'''}(\int_{A''} 1_E(ab)da)db \geqslant \frac{\mu(A'')\mu(B)}{8K}.$$

In particular B''' is non empty and since $\int_{A''} 1_E(ab)da \leqslant \mu(A'')$ we have

$$\mu(B''') \geqslant \frac{\mu(B)}{8K}.$$

We have proven that

$$\mu(A''') \gg \mu(A)/K,\ \mu(B''') \gg \frac{\mu(B)}{8K}.$$

It remains to give an upper bound for $\mu(A'''.B''')$ (and to prove Lemma 4.8).

Given $c\in A'''.B'''$ we have $c=a.b$ with $(a,b)\in A'''\times B'''$. By the definition of A''' we have for $a\in A'''$

$$\mu(\{a'\in A'',\ a(a')^{-1}\notin D\}) \leqslant \mu(A'')/16K$$

so that

$$\mu(\{a'\in A'',\ a(a')^{-1}\in D\}) \geqslant \mu(A'')(1-1/16K).$$

By the definition of B''', we have for $b\in B'''$

$$\int_{A''} 1_E(a'b)da' = \mu(\{a'\in A',\ a'b\in E\}) \geqslant \frac{\mu(A'')}{8K}.$$

Hence taking the intersection of the two sets above, we have that for $(a,b)\in A'''\times B'''$

$$\mu(\{a'\in A'',\ a(a')^{-1}\in D,\ a'b\in E\}) \geqslant (\frac{1}{8K}+1-\frac{1}{16K}-1)\mu(A'') = \frac{1}{16K}\mu(A'').$$

Making the change of variable

$$x = a'b,\ cx^{-1} = a(a')^{-1}$$

we have for any $c \in A'''.B'''$

$$\int_G 1_D(cx^{-1})1_E(x)dx = 1_D \star 1_E(c) \geqslant \frac{1}{16K}\mu(A'')$$

and

$$\int_{A'''.B'''} 1_D \star 1_E(c)dc \geqslant \frac{1}{16K}\mu(A'')\mu(A'''.B''').$$

Since

$$\int_{A'''.B'''} 1_D \star 1_E(c)dc \leqslant \int_G 1_D \star 1_E(c)dc = \mu(D)\mu(E)$$

we obtain

$$\frac{1}{16K}\mu(A'')\mu(A'''.B''') \leqslant \mu(D)\mu(E) \ll K^6\mu(A'')K\mu(A)^{1/2}\mu(B)^{1/2}.$$

Hence

$$\mu(A'''.B''') \ll K^8\mu(A)^{1/2}\mu(B)^{1/2}.$$

□

4.4.2. Proof of Lemma 4.8. It remain to prove

LEMMA (weak BSG). *Let $K, K' \geqslant 1$ and $A, B, E \subset G$ such that*

$$\mu(E) \leqslant K'\mu(A)^{1/2}\mu(B)^{1/2}$$

and

$$\mu \otimes \mu(\{(a,b) \in A \times B,\ ab \in E\}) \geqslant \frac{1}{K}\mu(A)\mu(B)$$

then for any $\varepsilon \in (0,1)$, there exists $A' \subset A$ and $D \subset A.A^{-1}$ such that

$$\mu(A') \geqslant \frac{\mu(A)}{2K},\ \mu(D) \leqslant \frac{2(KK')^2}{\varepsilon}\mu(A)$$

and

$$\mu \otimes \mu(\{(a,a') \in A' \times A',\ a.a'^{-1} \in D\}) \geqslant (1-\varepsilon)\mu(A')^2.$$

PROOF. We have

$$\int_B (\int_A 1_E(ab)da)db \geqslant \frac{1}{K}\mu(A)\mu(B).$$

By CS we have

$$(\int_B (\int_A 1_E(ab)da)db)^2 \leqslant \mu(B)\int_B (\int_A 1_E(ab)da)^2 db$$

so that

$$\int_B (\int_A 1_E(ab)da)^2 db \geqslant \frac{1}{K^2}\mu(A)^2\mu(B).$$

By Fubini we have

$$\int_A \int_A (\int_B 1_E(ab)1_E(a'b)db)dada' \geqslant \frac{1}{K^2}\mu(A)^2\mu(B).$$

Let $\Omega_\varepsilon \in A \times A$ be the set of $(a,a') \in A \times A$ such that

$$\int_B 1_E(ab)1_E(a'b)db \leqslant \frac{\varepsilon}{2K^2}\mu(B).$$

Equivalently for $(a,a')\in\Omega_\varepsilon$ we have

$$-\frac{1}{\varepsilon}\int_B 1_E(ab)1_E(a'b)db\geqslant -\frac{1}{2K^2}\mu(B)$$

hence

$$\int_{A\times A}(1-\frac{1}{\varepsilon}1_{\Omega_\varepsilon})(\int_B 1_E(ab)1_E(a'b)db)dada'\geqslant\frac{1}{2K^2}\mu(A)^2\mu(B).$$

By Fubini we have

$$\int_{A\times A}(1-\frac{1}{\varepsilon}1_{\Omega_\varepsilon})(\int_B 1_E(ab)1_E(a'b)db)dada'$$
$$=\int_B(\int_{A\times A}(1-\frac{1}{\varepsilon}1_{\Omega_\varepsilon})1_E(ab)1_E(a'b)dada')db\geqslant\frac{1}{2K^2}\mu(A)^2\mu(B)$$

and there exists $b\in B$ such that

$$\int_{A\times A}(1-\frac{1}{\varepsilon}1_{\Omega_\varepsilon}(a,a'))1_E(ab)1_E(a'b)dada'\geqslant\frac{1}{2K^2}\mu(A)^2.$$

This b being now fixed, the above integral is supported in $A'\times A'$ where

$$A'=A_b=\{a\in A,\ ab\in E\}$$

and the integrant is bounded by 1. We have therefore

$$\mu(A')^2\geqslant\frac{1}{2K^2}\mu(A)^2.$$

Moreover we also have

$$\frac{1}{\varepsilon}\int_{A'\times A'\cap\Omega_\varepsilon}dada'\leqslant\int_{A'\times A'}dada'$$

or

$$\mu(A'\times A'\cap\Omega_\varepsilon)\leqslant\varepsilon\mu(A')^2.$$

Let

$$D=\{a(a')^{-1},\ (a,a')\in A'\times A'\backslash\Omega_\varepsilon\}.$$

We have

$$\mu\otimes\mu(\{(a,a')\in A'\times A',\ a.a'^{-1}\in D\})=\mu\otimes\mu(A'\times A'\backslash\Omega_\varepsilon)\geqslant(1-\varepsilon)\mu(A')^2.$$

It remain to upper bound $\mu(D)$.

Given $d\in D$; write

$$d=a(a')^{-1},\ (a,a')\in A'\times A'\backslash\Omega_\varepsilon.$$

Since $(a,a')\notin\Omega_\varepsilon$ we have

$$\int_B 1_E(ab)1_E(a'b)db\geqslant\frac{\varepsilon}{2K^2}\mu(B).$$

hence writing $e=a'b$ and $ab=a(a')^{-1}e=de$ we have

$$\int_G 1_E(de)1_E(e)de\geqslant\int_B 1_E(ab)1_E(a'b)db\geqslant\frac{\varepsilon}{2K^2}\mu(B)$$

and integrating over $d\in D$ we obtain

$$\int_D\int_G 1_E(de)1_E(e)dedd\geqslant\frac{\varepsilon}{2K^2}\mu(B)\mu(D)$$

and

$$\int_D \int_G 1_E(de)1_E(e)dedd \leqslant \int_G \int_G 1_E(de)1_E(e)dedd \leqslant \mu(E)^2.$$

We have therefore

$$\mu(E)^2 \geqslant \frac{\varepsilon}{2K^2}\mu(B)\mu(D)$$

Which gives

$$\mu(D) \leqslant \frac{2(KK')^2}{\varepsilon}\mu(A).$$

□

The proof of the BSG theorem, approximate group version, follow from the set theoretic version and by applying the Approximate subgroup recognition criterion to A''', B''' which satisfy

$$|A'''.B'''| \ll K^8|A|^{1/2}|B|^{1/2} \ll K^9|A'''|^{1/2}|B'''|^{1/2}.$$

CHAPTER 5

The Sum-Product Theorem

The sum-product theorem discovered by Bourgain, Katz and Tao indeed shows that the combitation of addition and multiplication indeed conduct to growth:

THEOREM 5.1 (Sum-Product theorem). *For any $\delta > 0$ there exists $C, \eta > 0$ such that for any prime p and any subset $A \subset \mathbb{F}_p^\times$ satisfying*

$$C \leqslant |A| \leqslant p^{1-\delta}$$

one has then

$$|A + A| + |A.A| \geqslant |A|^{1+\eta}.$$

REMARK 5.1. The initial version of the Sum-Product theorem due to Bourgain, Katz and Tao included the additional assumption $|A| \geqslant p^\delta$. It was weakened to $|A| \geqslant C$ by Bourgain, Glibichuck and Konyagin. Here we will follow a very compact proof due to Ben Green of the initial version.

REMARK 5.2. Here and after $d(A, B) = d_+(A, B)$ and $e(A, B) = e_+(A, B)$ denote the Ruzsa distance and the additive energy relative to the addition in $\mathbb{F}_p$. Not the multiplication in $\mathbb{F}_p^\times$.

5.1. Rough notations and Ruzsa calculus

Given $K \geqslant 2$ a parameter, X, Y $A, B \subset (R, +)$ finite subsets of an abelian group. We will write

$$X \preceq_K Y,\ Y \succeq_K Y$$

if there exists a (absolute) constant $c > 0$ such that

$$X \leqslant K^c Y.$$

We will write $X \approx_K Y$ if

$$X \preceq_K Y \preceq_K X$$

and we will write

$$A \sim_K B \text{ for } |A - B| \preceq_K (|A||B|)^{1/2}$$

or equivalently

$$d(A, B) = O(\log K).$$

In practice the values of the implicit constants c are allowed to vary from line to line. If the "roughness parameter" K is understood we will omit it from the notation.

With these notations we have the following:

$$A \sim B \Longrightarrow e(A, -B) \approx 1.$$

Indeed if

$$|A - B| \leqslant K\sqrt{|A||B|}$$

we have by CS

$$K^{-1} \leqslant \frac{\sqrt{|A||B|}}{|A-B|} \leqslant e(A,-B) \leqslant 1.$$

Conversely the BSG theorem (set-theoretic version) admits the following immediate Corollary expressed in rough notations:

THEOREM 5.2 (BSG). *Given $K \geqslant 2$. Suppose that*

$$e(A,B) \geqslant 1/K$$

then there exists $A' \subset A,\ B' \subset B$ such that

$$|A'| \approx |A|,\ |B'| \approx |B|,\ A' \sim B'.$$

Here are some general results in additive combinatorics expressions in rough notations:

THEOREM 5.3 (Ruzsa calculus). *Let $A, B, C \subset (R,+)$. Set*

$$\delta(A) = \delta_+(A) := |A+A|/|A|$$

for the doubling constant of A. The following hold:

(1) *Suppose that $A \sim B,\ B \sim C$ then $A \sim C$.*
(2) *Suppose that $A \sim B$ then*

$$A \sim -B, |A| \approx |B|,\ \delta(A) \approx \delta(B) \approx 1. \tag{5.1}$$

(3) *Suppose that $A \sim B \sim C$ then*

$$A \sim B + C. \tag{5.2}$$

(4) *Suppose that $A \sim B,\ \delta(C) \approx 1$ and there exists $x \in R$ such that*

$$|A \cap (x+C)| \approx |A| \approx |B|$$

then

$$A \sim B \sim C.$$

(5) *Suppose that $\delta(A), \delta(B) \approx 1$ and there exists $x \in R$ such that*

$$|A \cap (x+C)| \approx |A| \approx |B|$$

then $A \sim B$.

For the proof the following proposition will be useful:

PROPOSITION 5.4 (Second Ruzsa inequality). *Let $(G,+)$ be a commutative group and $A, B \subset G$. We have*

$$d(A,-B) \leqslant 3d(A,B).$$

PROOF. Recall that

$$\sum_{g \in G} r_{A,B}(g) = \sum_{g \in G} r_{A,-B}(g) = |A||B|$$

and also that

$$a + b = a' + b' \Longleftrightarrow a - b' = a' - b$$

so that

$$E(A,B) = \sum_{g \in G} r_{A,B}(g)^2 = \sum_{g \in G} r_{A,-B}(g)^2 = E(A,-B).$$

By CS we have

$$(|A||B|)^2 \leqslant |A-B| \sum_{g\in A-B} r_{A,-B}(g)^2 = |A-B| \sum_{g\in A+B} r_{A,B}(g)^2.$$

This implies that there exists $g \in A+B$ such that

$$r_{A,B}(g) \geqslant \frac{|A||B|}{|A-B|}.$$

Suppose the contrary: we would have

$$\begin{aligned}(|A||B|)^2 \leqslant |A-B| \sum_{g\in A-B} r_{A,-B}(g)^2 = |A-B| \sum_{g\in A+B} r_{A,B}(g)^2 \\ < |A-B| \frac{|A||B|}{|A-B|} \sum_{g\in A+B} r_{A,B}(g) = (|A||B|)^2.\end{aligned}$$

For g as above, let

$$S_g = \{(a,b) \in A\times B,\ a+b=g\}$$

so that

$$|S_g| \geqslant \frac{|A||B|}{|A-B|}.$$

Choose a section $s=(s_A,s_B)$ (say) of the map

$$(a,b) \in A\times B \mapsto a+b \in A+B$$

and consider the map

$$S_g \times (A+B) \to (A-B)\times(A-B)$$

given by

$$(a,b,x) \to (a-s_B(x), s_A(x)-b).$$

This map is injective: suppose that

$$a-s_B(x) = a'-s_B(x'),\ s_A(x)-b = s_A(x')-b'$$

we have

$$a-s_B(x)-(s_A(x)-b) = a'-s_B(x')-(s_A(x')-b')$$

and since $a+b=a'+b'=g$ we obtain

$$s_A(x)+s_B(x) = x = s_A(x')+s_B(x') = x'$$

and then since $s_A(x)=s_A(x')$, $s_B(x)=s_B(x')$ we have $(a,b)=(a',b')$.

It follows from the injectivity that

$$|S_g||A+B| \leqslant |A-B|^2$$

and therefore

$$\frac{|A||B|}{|A-B|}|A+B| \leqslant |A-B|^2$$

so that

$$\frac{|A+B|}{(|A||B|)^{1/2}} \leqslant \Big(\frac{|A-B|}{(|A||B|)^{1/2}}\Big)^3$$

□

PROOF. (of Thm 5.3) The first property is Ruzsa triangle inequality.

For the second we have

$$|A|, |B| \leqslant |A - B| \preceq (|A||B|)^{1/2}$$

which gives $|A| \approx |B|$. The fact that $A \sim -B$ comes from Ruzsa second inequality. For the third we have

$$d(A, -A) \leqslant d(A, B) + d(B, -A) = O(\log K)$$

which gives

$$1 \leqslant \delta(A) \preceq 1$$

or

$$\delta(A) \approx 1$$

and likewise for B.

We leave the rest as an exercise. □

5.2. Warm-up: growth in $\mathbb{F}_p$ under addition and multiplication

We recall the following Lemma which we have proven in the beginning of this course using the probabilistic method.

LEMMA 5.5. *Given $|A| \subset \mathbb{F}_p$ non-empty; there is $\xi \in \mathbb{F}_p^\times$ such that*

$$|A + \xi A| \geqslant \frac{1}{2} \min(|A|^2, p).$$

THEOREM 5.6 (Glibichuck-Konyagin). *Given $A \subset \mathbb{F}_p$. We have*

$$|3A^{(2)} - 3A^{(2)}| \geqslant \frac{1}{2} \min(|A|^2, p).$$

For the proof and after, the following notation will be very useful. Given $A \subset \mathbb{F}_p$ we set

$$Q[A] := \frac{A - A}{A - A} := \{\frac{a_1 - a_3}{a_2 - a_4},\ a_i \in A,\ a_2 \neq a_4\}.$$

(if A has $\leqslant 1$ element, we set $Q[A] = \emptyset$)

PROOF. Observe that given $\xi \in \mathbb{F}_p^\times$ we have $|A + \xi A| = |A|^2$ unless $\xi \in Q[A]$ (since then $a_1 + \xi a_4 = a_2 + \xi a_3$ with $(a_1, a_4) \neq (a_2, a_3)$).

If $Q[A] \neq \mathbb{F}_p$ there exists $\xi = \frac{a_1 - a_3}{a_2 - a_4} \in Q[A]$ such that $1 + \xi \notin Q[A]$ and therefore

$$|A + (1 + \xi)A| = |A|^2.$$

We have

$$(a_2 - a_4)(A + (1 + \xi)A) \subset (a_2 - a_4)A + (a_1 - a_3 + a_2 - a_4)A \subset 3A^{(2)} - 3A^{(2)}$$

and

$$|3A^{(2)} - 3A^{(2)}| \geqslant |A|^2.$$

If $Q[A] = \mathbb{F}_p$ there is $\xi = \frac{a_1 - a_3}{a_2 - a_4}$ such that

$$|A + \xi A| \geqslant \frac{1}{2} \min(|A|^2, p)$$

and now we have

$$(a_2 - a_4)(A + (1 + \xi)A) \subset 2A^{(2)} - 2A^{(2)} \subset 3A^{(2)} - 3A^{(2)}$$

where the last inclusion follows from the identity

$$b + a - (b' + a) = b - b'$$

applied to $b, b' \in 2A^{(2)}$ and $a \in A$. □

COROLLARY 5.7. *Given $\delta \in (0,1)$ and $|A| \geqslant p^\delta$, there exists $k = k(\delta)$, $l = l(\delta) \geqslant 1$ such that*

$$kA^{(l)} - kA^{(l)} = \mathbb{F}_p.$$

In particular if A is a subgroup of $\mathbb{F}_p^\times$ we have

$$kA - kA = \mathbb{F}_p.$$

PROOF. Iterating the above process k times we find that

$$|4^l.3.A^{(2l)} - 4^l.3.A^{(2l)}| \geqslant \frac{1}{2}\min(|A|^{2l}, p)$$

so for $l \geqslant 1/2\delta$ we have

$$|4^l.3.A^{(2l)} - 4^l.3.A^{(2l)}| \geqslant p/2.$$

Since p is odd we have $|4^l.3.A^{(2l)} - 4^l.3.A^{(2l)}| \geqslant (p+1)/2$. In particular for any $t \in \mathbb{F}_p$ we have

$$t - (4^l.3.A^{(2l)} - 4^l.3.A^{(2l)}) \cap (4^l.3.A^{(2l)} - 4^l.3.A^{(2l)}) \neq \emptyset$$

so that

$$\mathbb{F}_p = (4^l.3.A^{(2l)} - 4^l.3.A^{(2l)}) + (4^l.3.A^{(2l)} - 4^l.3.A^{(2l)}) = 4^l.6.A^{(2l)} - 4^l.6.A^{(2l)}.$$

□

DEFINITION 5.8. *Given $K \geqslant 2$ and $A \subset \mathbb{F}_p$ we define*

$$\mathrm{Alg}_K(A) := \{b \in \mathbb{F}_p^\times,\ |A + bA| \leqslant K|A|\}.$$

PROPOSITION 5.9. *There is an absolute constant $c > 0$ such that*

(1) $b \in \mathrm{Alg}_K(A) \Longrightarrow b^{-1} \in \mathrm{Alg}_K(A)$.
(2) $b \in \mathrm{Alg}_K(A) \Longrightarrow -b \in \mathrm{Alg}_{K^c}(A)$.
(3) $b, b' \in \mathrm{Alg}_K(A),\ b \neq -b' \Longrightarrow b + b' \in \mathrm{Alg}_{K^c}(A)$.
(4) $b, b' \in \mathrm{Alg}_K(A),\ \Longrightarrow b.b' \in \mathrm{Alg}_{K^c}(A)$.

PROOF. Exercise using Ruzsa calculus. □

From this we deduce the following consequence of Corollary 5.7

COROLLARY 5.10. *Given $\delta \in (0, 1/2)$, there exists $\eta = \eta_\delta > 0$ such that for $A, B \subset \mathbb{F}_p$ with*

$$p^\delta \leqslant |A| \leqslant p^{1-\delta},\ |B| \geqslant p^\delta$$

There exists $b \in B$ satisfying

$$|A + bA| \geqslant |A|^{1+\eta}.$$

PROOF. Given $K \geqslant 2$ and assume that

$$\forall b \in B,\ |A + bA| \leqslant K|A|$$

that is $B \subset \mathrm{Alg}_K(A)$.

By Corollary 5.7, there exist $k, l \geqslant 1$ (depending on δ) such that

$$kB^{(l)} - kB^{(l)} = \mathbb{F}_p.$$

By proposition 5.9 applied several times, there exists $c = c(\delta)$ such that

$$(kB^{(l)} - kB^{(l)})\backslash\{0\} \subset \mathrm{Alg}_{K^c}(A)$$

and for any $\xi \in \mathbb{F}_p^\times$ we have

$$|A + \xi A| \leqslant K^c |A|.$$

On the other hand, by Lemma 5.5 there exists $\xi \in \mathbb{F}_p^\times$ such that

$$|A + \xi A| \geqslant \frac{1}{2}\min(|A|^2, p) \geqslant \frac{1}{2}|A|^{\min(2, \frac{1}{1-\delta})} = \frac{1}{2}|A|^{\frac{1}{1-\delta}}.$$

and we obtain

$$K \geqslant \frac{1}{2^{1/c}}|A|^{\frac{1}{c(1-\delta)}}.$$

In particular for $K = \frac{1}{4^{1/c}}|A|^{\frac{1}{c(1-\delta)}}$ there exists $b \in B$ such that

$$|A + bA| \geqslant \frac{1}{4^{1/c}}|A|^{1+\frac{1}{c(1-\delta)}}.$$

□

5.3. The BBSG theorem

The key ingredient is the following version of the BSG Theorem due to Bourgain:

PROPOSITION 5.11 (Multiplicative BBSG). *Given $K \geqslant 2$ and $A \subset \mathbb{F}_p$, $B \subset \mathbb{F}_p^\times$ such that*

$$\forall b \in B,\ e(A, bA) \geqslant 1/K.$$

There exists $A' \subset A$, $b_0 \subset B' \subset B$ satisfying

$$|A'| \approx |A|,\ |B'| \approx |B|$$

and such that

$$\forall b \in B',\ bA' \sim b_0 A'.$$

PROOF. By the BSG, for any $b \in B$ there exists $A'_b, A''_b \subset A$ such that $|A'_b| \approx |A''_b| \approx |A|$ and $A'_b \sim bA''_b$ (where the implicit exponents of K are all independant of b). We would like the A'_b, A''_b to not depend on b up to restricting to the b in a sufficiently large subset of B.

We will use the following

LEMMA 5.12. *Let $\delta \in (0,1)$, S a finite set and $S_n \subset S$, $n \leqslant N$ be subsets of S such that*

$$|S_n| \geqslant \delta |S|.$$

There exists n_0 such that

$$|S_{n_0} \cap S_n| \geqslant \delta^2 |S|$$

for at least $[\delta^2 N/2]$ n's.

PROOF. We have

$$\sum_{n \leqslant N} \sum_{s \in S} 1_{S_n}(s) \geqslant \delta |S| N$$

and by CS

$$\sum_{s \in S} (\sum_{i,j} 1_{S_i}(s) 1_{S_j}(s) \geqslant \frac{\delta^2 |S|^2 N^2}{\sum_{s \in S} 1} = \delta^2 |S| N^2.$$

Hence

$$\sum_n \sum_{n'} |S_n \cap S_{n'}| \geqslant \delta^2 |S| N^2.$$

In particular there exists n_0 such that

$$\sum_n |S_{n_0} \cap S_n| \geqslant \delta^2 |S| N.$$

Consider the set $\mathcal{N}_\delta$ of n's such that

$$|S_{n_0} \cap S_n| \geqslant \frac{\delta^2 |S|}{2}.$$

We have

$$\sum_{n \notin \mathcal{N}_\delta} |S_{n_0} \cap S_n| \leqslant \frac{\delta^2 |S|}{2} N$$

and therefore

$$\sum_{n \in \mathcal{N}_\delta} |S_{n_0} \cap S_n| \geqslant \frac{\delta^2 |S|}{2} N.$$

Since $|S_{n_0} \cap S_n| \leqslant |S|$ we conclude that

$$|\mathcal{N}_\delta| \geqslant \frac{\delta^2}{2} N.$$

□

We apply this Lemma to the family of subsets

$$S_b = A'_b \times A''_b \subset S = A \times A,\ b \in B$$

with $\delta = K^{-c}$ for a suitable absolute constant $c > 0$. There exists $b_0 \in B' \subset B$ with $|B'| \approx |B|$ and such that

$$\forall b \in B', |(A'_b \times A''_b) \cap (A'_{b_0} \times A''_{b_0})| \approx |A \times A|$$

and therefore

$$|A'_b \cap A'_{b_0}| \approx |A''_b \cap A''_{b_0}| \approx |A|.$$

Since for all $b \in B'$ we have

$$A'_b \sim b A''_b,\ A'_{b_0} \sim b_0 A''_{b_0} \tag{5.3}$$

and in particular by (5.1) from Ruzsa calculus we have

$$\delta(A'_{b_0}),\ \delta(A''_{b_0}), \delta(A'_b),\ \delta(A''_b) \approx 1.$$

It then follows from Thm 5.3 (5) that

$$A'_b \sim A'_{b_0},\ A''_b \sim A''_{b_0}.$$

combining this with (5.3) we see that

$$\forall b \in B',\ b_0 A'_{b_0} \sim b_0 A'_b \sim b_0 b A''_b = b b_0 A''_{b_0} \sim b A'_{b_0}$$

and we take $A' = A'_{b_0}$. □

COROLLARY 5.13. *Given $\delta \in (0,1/2)$, there exists $\eta = \eta_\delta > 0$ such that for $A \subset \mathbb{F}_p,\ B \subset \mathbb{F}_p^\times$ with*

$$p^\delta \leqslant |A| \leqslant p^{1-\delta},\ |B| \geqslant p^\delta$$

There exists $b \in B$ satisfying

$$e(A, bA) \leqslant |A|^{-\eta}.$$

PROOF. Let $K = |A|^\eta$ for $\eta > 0$ to be chosen sufficiently small (in terms of δ) and assume that for all $b \in B$ we have

$$e(A, bA) \geqslant 1/K.$$

In particular for any $b \in B$ we have

$$\frac{|A + bA|}{|A|} \leqslant K = |A|^\eta. \tag{5.4}$$

By the BBSG theorem there exists $A' \subset A,\ b_0 \in B' \subset B$ satisfying $|A'| \approx |A|,\ |B'| \approx |B|$ and

$$\forall b \in B',\ b_0 A' \sim bA'.$$

The conditions $|A'| \approx |A|,\ |B'| \approx |B|$ say that there exists an absolute constant $c \geqslant 0$ such that

$$|A'| \geqslant |A|^{1-c\eta},\ |B'| \geqslant |B||A|^{-c\eta}.$$

In particular we have

$$p^{\delta-(1-\delta)c\eta} \leqslant |A|^{1-c\eta} \leqslant |A'| \leqslant |A| \leqslant p^{1-\delta},\ p^{\delta-(1-\delta)c\eta} \leqslant |B'|.$$

Take $\eta \in (0, \frac{\delta}{2c})$ so that

$$p^{\delta/2} \leqslant |A'| \leqslant |A|^{1-\delta/2},\ p^{\delta/2} \leqslant |B'|.$$

By the Corollary 5.10 applied to "(A, B, δ)" $= (A', (b_0)^{-1}B', \delta/2)$, there exists $\eta' = \eta'(\delta) \in (0,1)$ and $b \in B'$ such that

$$|A' + bA'| \geqslant |A'|^{1+\eta'} \geqslant |A|^{(1+\eta')(1-c\eta)}.$$

Suppose that η is chosen so that

$$|A|^{(1+\eta')(1-c\eta)} \geqslant |A|^{1+2\eta} \tag{5.5}$$

We have

$$|A + bA| \geqslant |A' + bA'| \geqslant |A|^{1+2\eta}$$

and we conclude that

$$\frac{|A + bA|}{|A|} \geqslant |A|^{2\eta}$$

which is a contradiction with (5.4).

Let us see that we can pick η with this property: We have

$$\lim_{\eta \to 0}(1+\eta')(1-c\eta) - (1+2\eta) = \eta' > 0$$

so there exists $\eta_0 = \eta_0(c, \eta') = \eta_0(c, \delta) > 0$ such that for $\eta \in (0, \eta_0]$ we have (5.5). We then take

$$\eta = \min(\eta_0, \frac{\delta}{2c}).$$

□

5.4. Proof of the sum-product Theorem

Let $K = |A|^\eta$ for $\eta > 0$ to be chosen sufficiently small depending on δ.

We assume that

$$|A + A|, |A.A| \leqslant K|A|.$$

In particular this means that a lot of the translates

$$bA,\ b \in A$$

have a rather large intersection.

Indeed by Lemma 5.12 (applied to $S = A.A$ and the sets $bA,\ b \in A - \{0\}$) there exists $b_0 \in B \subset A - \{0\}$ such that

$$|B| \gg |A|/K^2$$

and

$$\forall b \in B,\ |b_0 A \cap bA| \gg |A|/K^2.$$

We assume that $\eta > 0$ is chosen sufficiently small (depending only on δ) so that

$$|A|/K^2 = |A|^{1-2\eta} \geqslant p^{\delta/2}.$$

For any $b \in B$ let

$$A_b := A \cap (b/b_0)A.$$

We have

$$|A_b| \approx |A|$$

and since

$$|A_b + A_b| \leqslant |A + A| \leqslant\leqslant K|A| \preceq |A_b|$$

we have

$$\delta(A_b) \approx 1$$

and therefore

$$e(A_b, A_b) \approx 1.$$

Since $A_b \subset A$ we have

$$1 \sim e(A_b, A_b) = \frac{E(A_b, A_b)}{|A_b|^3} \leqslant \frac{E(A, A_b)}{|A_b|^3} \approx \frac{E(A, A_b)}{(|A||A_b|)^{3/2}} = e(A, A_b).$$

In other terms there exists an absolute constant $c \geqslant 0$ such that for all $b \in B$

$$e(A, (b/b_0)A) \geqslant K^{-c} = |A|^{-c\eta}.$$

On the other hand by Corollary 5.13 applied to $(A, (1/b_0)B)$ using that

$$|A|, |B| \geqslant p^{\delta/2}$$

there exists $b \in B$ such that

$$e(A, (b/b_0)A) \leqslant p^{-\eta'}$$

where η' depends on δ. In particular if we choose η such that $c\eta \leqslant \eta'/2$ we obtain a contradiction.

CHAPTER 6

Growth in $\mathrm{SL}_2(\mathbb{F}_p)$

Let

$$\mathrm{SL}_2(\mathbb{F}_p) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{F}_p),\ ad - bc = 1\}$$

be the special linear group. It has order

$$|\mathrm{SL}_2(\mathbb{F}_p)| = p(p^2 - 1) = p^3 - p$$

(more generally for k a finite field $|\mathrm{SL}_2(k)| = |k|(|k|^2 - 1)$).

THEOREM 6.1 (Product Theorem for $\mathrm{SL}_2(\mathbb{F}_p)$; Helfgott version). *There exists $k, \delta > 0$ such that for any p and any subset $A \subset \mathrm{SL}_2(\mathbb{F}_p)$ generating $\mathrm{SL}_2(\mathbb{F}_p)$ as a group, one of the following holds*

$$|A^{(3)}| \geqslant |A|^{1+\delta} \text{ or } (A \cup A^{-1} \cup \{\mathrm{Id}_2\})^{(k)} = \mathrm{SL}_2(\mathbb{F}_p).$$

This will be a consequence of the following:

THEOREM 6.2 (Product Theorem for $\mathrm{SL}_2(\mathbb{F}_p)$; approximate subgroup version). *Let $K \geqslant 2$, there exists an absolute constant $C > 0$ such that given any finite field k and any K-approximate subgroup $A \subset G = \mathrm{SL}_2(k)$ generating G, one has either*

(1) $|A| \leqslant K^C$,
(2) $|A| \geqslant |G|K^{-C}$.

We leave the proof of the implication

$$\text{Thm 6.2} \Longrightarrow \text{Thm 6.1}$$

as an exercise but this still require a "finisher" due to Gowers.

THEOREM 6.3. *There exists an absolute constant $\delta > 0$ such that for any subsets $A, B, C \subset \mathrm{SL}_2(\mathbb{F}_p)$ such that*

$$|A||B||C| \geqslant |\mathrm{SL}_2(\mathbb{F}_p)|^{3-\delta}$$

then

$$A.B.C = \mathrm{SL}_2(\mathbb{F}_p). \tag{6.1}$$

In particular if

$$|A| \geqslant |\mathrm{SL}_2(\mathbb{F}_p)|^{1-\delta/3}$$

then

$$A^{(3)} = \mathrm{SL}_2(\mathbb{F}_p).$$

REMARK 6.1. As we will see, for p large enough one can take any $\delta \in (0, 1/3)$.

6.1. A spectral gap property for $\mathrm{SL}_2(\mathbb{F}_p)$

The proof goes via harmonic analysis on the group $G = \mathrm{SL}_2(\mathbb{F}_p)$ which is recalled in Appendix 7.5.

Consider the convolution

$$f = 1_A \star 1_B \star 1_C : g \mapsto \sum_{abc=g} \sum 1.$$

We need to show that if $|A.B.C|$ is large enough then

$$\forall g \in G,\ f(g) > 0.$$

In fact, up to replacing C by Cg^{-1} (which has the same size as C) it is sufficient to show that

$$f(e_G) > 0.$$

For this we decompose f using non-abelian Fourier theory.

We have the Fourier decomposition (7.18)

$$f(e_G) = \frac{1}{|G|} \sum_{\pi \in \mathrm{Irr}(G)} d_\pi \mathrm{tr}(\pi(f))$$

where

$$\pi(f) : v \in V_\pi \mapsto \sum_{h \in G} f(h)\pi(h)v \in V_\pi.$$

The contribution of the trivial representation $\pi_0 : G \mapsto 1$ is

$$\frac{1}{|G|} \sum_{h \in G} f(h) = \frac{1}{|G|} \sum_{h \in H} \sum_{abc=h} 1 = \frac{|A||B||C|}{|G|}.$$

We now bound the contribution of the non-trivial irreducible representations $\pi \neq \pi_0$. We have

$$f(h) = \sum_{abc=h} 1 = (1_A \star 1_B) \star 1_C(h)$$

and the contribution of π can be bounded using (7.19)

$$\frac{d_\pi}{|G|} |\langle \pi(1_A)\pi(1_B), \pi(\check{1}_C)\rangle_{HS}| \leqslant \frac{d_\pi}{|G|} \|\pi(1_A)\|_{HS} \|\pi(1_B)\|_{HS} \|\pi(1_C)\|_{HS}$$

on using the CS inequality, the sub-multiplicativity of the HS-norm,

$$\|\varphi \circ \psi\|_{HS} \leqslant \|\varphi\|_{HS} \|\psi\|_{HS},$$

and

$$\|\pi(1_C)\|_{HS} = \|\pi(\check{1}_C)\|_{HS}.$$

By Parseval we have

$$|A| = \sum_{h \in G} |1_A(h)|^2 = \frac{1}{|G|} \sum_{\pi' \in \mathrm{Irr}(G)} d_{\pi'} \|\pi'(1_A)\|_{HS}^2 \geqslant d_\pi \|\pi(1_A)\|_{HS}^2$$

so that

$$\|\pi(1_A)\|_{HS} \leqslant (\frac{|A||G|}{d_\pi})^{1/2} \leqslant (\frac{|A||G|}{\min_{\pi' \neq \pi_0} d_\pi})^{1/2}.$$

We have therefore

$$f(e_G) \geqslant \frac{|A||B||C|}{|G|} - \frac{1}{|G|}\Big(\frac{|A||G|}{\min_{\pi'\neq\pi_0} d_\pi}\Big)^{1/2} \sum_{\pi\neq\pi_0} d_\pi \|\pi(1_B)\|_{HS}\|\pi(1_C)\|_{HS}.$$

By CS and Parseval we have

$$\begin{aligned}\sum_{\pi\neq\pi_0} d_\pi \|\pi(1_B)\|_{HS}\|\pi(1_C)\|_{HS} &\leqslant \Big(\sum_{\pi\neq\pi_0} d_\pi \|\pi(1_B)\|_{HS}^2\Big)^{1/2}\Big(\sum_{\pi\neq\pi_0} d_\pi \|\pi(1_C)\|_{HS}^2\Big)^{1/2}\\ &\leqslant (|G||B|)^{1/2}(|G||C|)^{1/2}\end{aligned}$$

and

$$f(e_G) \geqslant \frac{|A||B||C|}{|G|} - \Big(\frac{|A||B||C||G|}{\min_{\pi'\neq\pi_0} d_\pi}\Big)^{1/2}$$

which is > 0 as long as

$$|A||B||C| > \frac{|G|^3}{\min_{\pi'\neq\pi_0} d_\pi}.$$

A remarkable theorem of Frobenius states that $\mathrm{SL}_2(\mathbb{F}_p)$ has no non-trivial representations of small dimension (this is the *spectral gap* in the title of this section):

THEOREM 6.4 (Frobenius). *For $G = \mathrm{SL}_2(\mathbb{F}_p)$ the dimension of any non-trivial irreducible representation π satisfies*

$$d_\pi \geqslant \frac{p-1}{2}.$$

Let us conclude the proof of Theorem 6.3. Since

$$|\mathrm{SL}_2(\mathbb{F}_p)| \sim p^3,\ p\to\infty$$

we see that (6.1) hold when

$$|A||B||C| \geqslant |\mathrm{SL}_2(\mathbb{F}_p)|^{3-\delta}$$

for any $\delta \in (0, 1/3)$ as long as p is large enough. □

PROOF. Let

$$N(\mathbb{F}_p) = \{n(x) := \begin{pmatrix} 1 & x \\ 0 & 1\end{pmatrix},\ x\in\mathbb{F}_p\}$$

The map $x \mapsto n(x)$ is an isomorphism $\mathbb{F}_p \simeq n(\mathbb{F}_p)$. The group $\mathrm{SL}_2(\mathbb{F}_p)$ contains the group of diagonal matrices

$$T(\mathbb{F}_p) = \{t(y) := \begin{pmatrix} y & 0 \\ 0 & 1/y\end{pmatrix},\ y\in\mathbb{F}_p^\times\}$$

and we have

$$t(y)n(x)t(y)^{-1} = n(y^2x)).$$

Consider $n(1)$ (it generated $N(\mathbb{F}_p)$). If $n(1)$ does not act trivially on V_π then it has at least one eigenvalue $\alpha \neq 1$ and since $n(1)^p = \mathrm{Id}_2$ we have

$$\alpha \in \mu_p(\mathbb{C}) - \{1\}$$

(ie. is a non-trivial p-th root of unity). Moreover for any $y \in \mathbb{Z}$ coprime with p, α^{y^2} is also an eigenvalue of $n(1)$ (since $n(y^2) = t(y)n(1)t(y)^{-1}$ is conjugate to $n(1)$, they have the same eigenvalues). It follows that $n(1)$ has at least $\frac{p-1}{2} = |(\mathbb{F}_p^\times)^2|$ distinct eigenvalues; gives the lower bound $d_\pi \geqslant \frac{p-1}{2}$.

Suppose now that $n(1)$ acts trivially on V_π: we have $n(1) \in \ker(\pi)$ and therefore $N(\mathbb{F}_p) = n(1)^{\mathbb{Z}} \subset \ker(\pi)$ and any conjugate of this group is also in the kernel (a kernel is a normal subgroup): in particular the lower triangular unipotent subgroup

$$wN(\mathbb{F}_p)w^{-1} = \left\{ \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix},\ x \in \mathbb{F}_p \right\}$$

is also in $\ker(\pi)$. But we have the following Lemma

LEMMA 6.5. *The set*

$$\left\{ \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}, \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},\ x \in \mathbb{F}_p \right\}$$

generates $\mathrm{SL}_2(\mathbb{F}_p)$.

PROOF. Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$; we want to show this is a product of elements in the above set.

We have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & ax+b \\ b & d+bx \end{pmatrix} \tag{6.2}$$

so if $a \neq 0$ we are reduced to the case $b = 0$ and we have

$$\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} = \begin{pmatrix} a & \\ ax+c & d \end{pmatrix}$$

and we reduced to the case

$$g = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}.$$

We have

$$\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & y \\ x & 1 \end{pmatrix}, \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} = \begin{pmatrix} 1+xy & y \\ x & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} 1+xy & y \\ x & 1 \end{pmatrix} = \begin{pmatrix} a(1+xy) & ay \\ a^{-1}x & a^{-1} \end{pmatrix}$$

so we may force $a = 1$ and reduce to the case $g = \mathrm{Id}_2$.

If $a = 0,\ d \neq 0$ we can replace g by g^{-1} and a and d are exchanged.

If $a = d = 0$ we use (6.2) to make $d \neq 0$ and are reduced to the previous case. □

REMARK 6.2. Recall that $\mathrm{PSL}_2(\mathbb{F}_p) = \mathrm{SL}_2(\mathbb{F}_p)/\{\pm \mathrm{Id}_2\}$ is a simple non-commutative group as long as $p > 3$. Frobenius theorem is in line with this property. In fact Landazuri and Seitz have proven a generalisation of Frobenius's theorem for simple groups of Lie type and bounded rank.

For instance for $n \geqslant 2$ one has

$$\min_{\substack{\pi \in \mathrm{Irr}(\mathrm{SL}_n(\mathbb{F}_p)) \\ \pi \neq \pi_0}} d_\pi \gg_n p^{n-1}.$$

6.2. The Larsen-Pink inequalities

The product theorem is a classification result for the finite *approximate subgroups* of $\mathrm{SL}_2(\mathbb{F}_p)$. A first step is to obtain a classification for the for the finite *subgroups* of $\mathrm{SL}_2(\mathbb{F}_p)$. In fact we will discuss more generally the classification of the finite *subgroups* of $\mathrm{SL}_2(\overline{\mathbb{F}_p})$ where $\overline{\mathbb{F}_p}$ is an algebraic closure of $\mathbb{F}_p$.

Such classification is provided in much greater generality by the fundamental work of Larsen and Pink [?LP]. A key point is the fact that $\mathrm{SL}_2(\overline{\mathbb{F}_p})$ is an *algebraic group*.

6.2.1. Algebraic varieties and algebraic groups. Let $k \subset \overline{k}$ be a field with a choice of algebra closure. The n-dimensional affine space $\overline{k}^n$ is equipped with the *Zariski topology* in which the closed sets are the subsets of the shape

$$V_{\overline{I}}(\overline{k}) = \{\mathbf{x} \in \overline{k}^n,\ P(\mathbf{x}) = 0,\ \forall P \in \overline{I}\}$$

where $\overline{I} \subset \overline{k}[X_1, \cdots, X_n]$ is an ideal and the open sets are the complements.

A closed set is also called an (affine) algebraic variety.

The $\overline{I}$ is finitely generated ($\overline{k}[X_1, \cdots, X_n]$ is noetherian) and if $\overline{I}$ is generated by at most C polynomials of degree $\leqslant C$ then one says that $V_{\overline{I}}$ has complexity $\leqslant C$.

6.2.1.1. *Connected components.* A subvariety is connected if it is connected for the Zariski topology. One can show that a variety decomposes uniquely into a disjoint union of connected subvarieties (the connects components) whose number is $O_C(1)$ where C is a bound on the complexity of V.

6.2.1.2. *Irreducible components and dimension.* A subvariety V is *irreducible* if it is not the union of two proper subvarieties. One can show that any variety can be decomposed (uniquely) into a finite union of irreducible subvarieties (called the irreducible components of V) whose number is $O_C(1)$ where C is a bound on the complexity of V.

The *dimension*, $\dim V$, of an irreducible subvariety V is the maximal length of a strict chain of irreducible subvarieties

$$\emptyset \neq V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_D = V.$$

The dimension, $\dim V$, of a general variety V is the maximal dimension of its irreducible components.

6.2.1.3. *Variety defined over a field.* Suppose that $\overline{I} = \overline{k}.I$ is the ideal generated by some k-ideal $I \subset k[X_1, \cdots, X_n]$. We say that V is defined over k and we define the set of k-points of V as

$$V_I(k) := \{\mathbf{x} \in k^n,\ P(\mathbf{x}) = 0,\ \forall P \in I\} \subset V_I(\overline{k}).$$

We can in fact in the same way $V_I(k')$ the set of k'-point of V_I for any field extension $k \subset k'$ and more generally define the set of K-points $V_I(K)$ for any k-algebra (not necessarily contained in $\overline{k}$).

6.2.2. Linear algebraic groups. Let $M_n(k)$ be the k-algebra of $n \times n$ matrices. Taking elementary matrices as a basis we have the identifications

$$M_n(k) \simeq k^{n^2},\ M_n(\overline{k}) \simeq \overline{k}^{n^2}.$$

The group

$$\mathrm{GL}_n(\overline{k}) = \{g \in M_n(\overline{k})\} = \{g \in M_n(\overline{k}),\ \det(g) \neq 0\}$$

can be seen as the set of $\overline{k}$-points of an affine subvariety of dimension n^2 in a space of dimension n^2+1, namely:

$$\mathrm{GL}_n(\overline{k}) \simeq \{(g,t) \in M_n(\overline{k}) \times \overline{k},\ \det(g)t - 1 = 0\}.$$

via the map

$$(g,t) \mapsto g.$$

In fact GL_n is defined over k. If $n=1$, GL_1 is also noted $\mathbb{G}_m$ and is called the multiplicative group. We have

$$\mathbb{G}_m(k) = k^\times.$$

Definition 6.6. *A linear algebraic group G (defined over k) is an affine subvariety (defined over k) obtained from a system of polynomial equations*

$$P(\mathbf{X},T) = 0,\ P(\mathbf{X},T) \in I(G),\ det(\mathbf{X})T - 1 \in I(G) \subset k[X_{11}, \cdots, X_{nn}, T]$$

and such that $G(\overline{k}) \subset \mathrm{GL}_n(\overline{k})$ is a subgroup $\mathrm{GL}_n(\overline{k})$.

Example 6.1. Given $m \geqslant 1$ let

$$\mu_m(k) = \{\mu \in k^\times,\ \mu^m - 1 = 0\} \subset \mathbb{G}_m(k).$$

This is the group of m-th root of unity defined via the ideal of $k[X,T]$ generated by

$$XT - 1,\ X^m - 1.$$

$\mathrm{SL}_2(\overline{k}) \subset \mathrm{GL}_2(\overline{k})$ is a linear algebraic group defined over k via the ideal generated by

$$\det(g)T - 1 = 0,\ T - 1.$$

The Larsen-Pink inequalities provide a partial classification of the finite subgroups of an algebraic group.

Theorem 6.7 (Larsen-Pink). *Let $\overline{k}$ be algebraically closed and $\mathrm{G}(\overline{k})$ be a connected simple algebraic group.*

For any $D \geqslant 1$ there exists $C = C(D, \dim \mathrm{G}) > 0$ such that the following holds.

For any finite subgroup $A \subset \mathrm{G}(\overline{k})$, either A is contained in a proper algebraic subgroup $\mathrm{H}(\overline{k}) \subset \mathrm{G}(\overline{k})$ such that $[\mathrm{H} : \mathrm{H}^0] \leqslant C$ or for every closed algebraic subvariety $V(\overline{k}) \subset \mathrm{G}(\overline{k})$ of degree $\leqslant D$, one has

$$|A \cap V(\overline{k})| \leqslant C|A|^{\dim V / \dim G}.$$

This Theorem states that unless A is trapped inside some "reasonable" proper subgroup of $\mathrm{G}(\overline{k})$ then A intersects any reasonable algebraic subvariety in at most the right order of point. It admits an almost-subgroup version initiated by Hrushovski:

Theorem 6.8 (Larsen-Pink for almost groups). *Let $\overline{k}$ be algebraically closed and $\mathrm{G}(\overline{k})$ be a connected simple algebraic group.*

For any $D, K \geqslant 1$ there exists $C = C(D, K, \dim \mathrm{G}) > 0$ such that the following holds.

For any K–approximate subgroup $A \subset \mathrm{G}(\overline{k})$, either A is contained in a proper algebraic subgroup $\mathrm{H}(ovk) \subset \mathrm{G}(\overline{k})$ such that $[\mathrm{H} : \mathrm{H}^0] \leqslant C$ or for every closed algebraic subvariety $V \subset \mathrm{G}$ of degree $\leqslant D$, one has

$$|A \cap V(\overline{k})| \leqslant CK^C|A|^{\dim V / \dim G}.$$

The proof of Helfgott theorem, we present here (which is a special case more general results of Breuillard-Green-Tao and Pyber-Szabo) make use of very special cases of the Larsen-Pink inequalities.

6.3. Structure of $\mathrm{SL}_2(k)$

Let k be a field, $\overline{k}$ an algebraic closure. We write

$$G(k) = \mathrm{SL}_2(k), \mathrm{G}(\overline{k}) = \mathrm{SL}_2(\overline{k}).$$

6.3.1. Elements and sous-groups. Let $G = \mathrm{SL}_2$ and let
Given $g \in G$, its eigenvalues are the roots of

$$P_g(X) = X^2 - \mathrm{tr}(g)X + 1.$$

- Semisimple: if $\mathrm{tr}(g) \neq \pm 2$ then g has two distinct eigenvalues $\lambda_1, \lambda_2 \in \overline{k}$ (and even in k if $\mathrm{tr}(g) - 4$ is a square in k).
- In particular g is diagonalizable, ie. conjugate to a diagonal matrix with distinct entries; the conjugating matrix is a priori in $\mathrm{GL}_2(\overline{k})$ but can in fact be taken in $\mathrm{SL}_2(\overline{k})$ (by multiplying any conjugating matrix h by $\det(h)^{-1/2}\mathrm{Id}_2$).
- Central: if $g = \pm\mathrm{Id}_2$ then $\mathrm{tr}(g) = \pm 2$.
- Quasi-unipotent regular: if $\mathrm{tr}(g) = 2\varepsilon,\ \varepsilon = \pm 1$ and $g \neq \varepsilon\mathrm{Id}_2$ then there is one eigenvalue ε of multiplicity 2 and $\varepsilon.g$ is regular unipotent

  $$\varepsilon g \neq \mathrm{Id}_2,\ (\varepsilon g - \mathrm{Id}_2)^2 = 0_2.$$

  Consequently g is conjugate to a matrix of the shape

  $$\varepsilon \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},\ x \in k^\times$$

  and the conjugating matrix can be taken in $\mathrm{SL}_2(\overline{k})$.

6.3.1.1. *Centralizers, normalizers.* Given $g \in \mathrm{G}(\overline{k})$, we note

$$\mathrm{Cent}_g(\overline{k}) = \{h \in \mathrm{G}(\overline{k}),\ hgh^{-1} = g\}$$

its centralizer.

If g is semisimple then $\mathrm{Cent}_g(\overline{k}) =: T_g$ is conjugate to the group of diagonal matrices

$$T = \mathrm{Diag}_2(\overline{k}) = \{\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix},\ t \in \overline{k}^\times\}$$

and is called a maximal torus of $\mathrm{G}(\overline{k})$.

Let

$$\mathrm{Nor}_{T_g} = \{h \in \mathrm{G}(\overline{k}),\ hT_gh^{-1} = T_g\}$$

denote the normalizer of T_g then

$$\mathrm{Nor}_{T_g} = T_g \sqcup w_gT_g$$

where $w_g^2 = \mathrm{Id}_2$. For instance

$$\mathrm{Nor}_T = T \sqcup wT,\ w = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

If g is regular unipotent then

$$\mathrm{Cent}_g(\overline{k}) =: \pm N_g$$

where $g \subset N_g$ is conjugate to the group of unipotent upper triangular matrices

$$N = \{\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},\ x \in \overline{k}\}$$

The group N_g is called a unipotent subgroup.

The normalizer of a unipotent subgroup N_g

$$B_g = \mathrm{Nor}_{U_g} = \{h \in \mathrm{G}(\overline{k}),\ hU_g h^{-1} = U_g\}$$

is called a Borel subgroup is conjugate to the group of unipotent upper triangular matrices

$$B = \{\begin{pmatrix} t & x \\ 0 & t^{-1} \end{pmatrix},\ x \in \overline{k},\ t \in \overline{k}^{\times}\}.$$

It is clear that maximal tori, unipotent subgroups and Borel subgroups are linear algebraic subgroups of $\mathrm{SL}_2(\overline{k})$ of dimensions $1, 1$ and 2.

6.3.1.2. *The action by fractional linear transformations.* Let

$$\mathbb{P}^1(\overline{k}) = \{0 \subset L \subset \overline{k}^2,\ \dim L = 1\}$$

be the set of one dimensional subspaces in $\overline{k}^2$ (the lines passing through the origin).

Any L can be written

$$L = \{\beta y = \alpha x,\ (\alpha, \beta) \neq (0, 0)\}$$

and can be represented by its slope

$$s(L) = [\alpha : \beta] = \begin{cases} \alpha/\beta & \beta \neq 0 \\ \infty & \beta = 0 \end{cases}.$$

The group $\mathrm{SL}_2(\overline{k})$ acts on $\mathbb{P}^1(\overline{k})$ by fractional linear transforrmations: $z \in \mathbb{P}^1(\overline{k})$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \begin{cases} \frac{az+b}{cz+d} & z \neq -d/c \\ \infty & z = -d/c \end{cases}.$$

Its kernel is the center $\{\pm \mathrm{Id}_2\}$. In addition we have

- This action is transitive. The stabilizer of a point $z \in \mathbb{P}^1(\overline{k})$ is a Borel subgroup B_z and conversely a Borel subgroup is the stablizer of a unique L (the common eigenspace of the elements of B). For instance the upper triangulat Borel subgroup is the stabilizer of ∞, B_∞ and the lower triangular Borel subgroup is the stabilizer of 0, ${}^tB = B_0$. In particular we have for any Borel subgroup an isomorphism of $\mathrm{SL}_2(\overline{k})$-spaces
$$\mathbb{P}^1(\overline{k}) \simeq \mathrm{SL}_2(\overline{k})/B.$$
- In fact (check it) the action is 2-transitive: for any $z_1 \neq z_2 \in \mathbb{P}^1(\overline{k})$ there is $g \in \mathrm{SL}_2(\overline{k})$ such that
$$gz_1 = 0,\ gz_2 = \infty.$$
The poinwise stabilizer of a pair of distinct points $z_1 \neq z_2 \in \mathbb{P}^1(\overline{k})$ is the intersection of the two Borel subgroups $B_{z_1} \cap B_{z_2}$ which is in fact a maximal torus T_{z_1,z_2}. To see this transform z_1, z_2 to $0, \infty$ and check that the stabilizer is the diagonal subgroup $T = B_\infty \cap B_0$.
- Finally the pointwise stabilizer of three distinct points z_1, z_2, z_3 (ie. the intersection of three distinct Borel subgroups) is the trivial subgroup $\{\pm \mathrm{Id}_2\}$. To see this we may assume that $z_1 = 0, z_2 = \infty$ and any element stabilizing $0, \infty$ and $z_3 \neq 0, \infty$ would have to be a diagonal matrix $\begin{pmatrix} t & \\ & t^{-1} \end{pmatrix}$ and we would have $t^2 z_3 = z_3$ which implies $t = \pm 1$ (since $z_3 \neq 0$).

6.4. Special Larsen-Pink inequalities for $\mathrm{SL}_2(\overline{k})$

We will now prove some special cases of the LP inequalities for $\mathrm{SL}_2(\overline{k})$ namely for V

- A maximal torus T (of dimension 1),
- a unipotent subgroup U (of dimension 1),
- The conjugacy class $\mathrm{Conj}(g)$ of a non-central element $g \in \mathrm{SL}_2(\overline{)}$.

The proper algebraic subgroup H will be Borel subgroups B.

6.4.1. Intersection with Tori and Unipotent subgroups.

PROPOSITION 6.9 (LP for tori). *There exist constants $C, D > 0$ such that for any finite subgroup $A \subset \mathrm{G}(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds*

- *For any maximal torus T,*
$$|T \cap A| \leqslant C|A|^{1/3}.$$
- *There is a Borel subgroup B such that*
$$|B \cap A| \geqslant C^{-1}|A|.$$

REMARK 6.3. The exponent 3 is the dimension of $\mathrm{SL}(\overline{k})$ as an algebraic subvariety of $M_2(\overline{k})$ and 1 is the dimension of a maximal torus.

PROOF. Suppose that for any Borel subgroup B
$$|A \cap B| \leqslant C^{-1}|A|.$$
Given any $\gamma \in \mathrm{G}(\overline{k})$, consider the intersection $A \cap \gamma B$ and suppose it is non-empty. The group $A \cap B$ acts on $A \cap \gamma B$ by right translations: if $g \in A \cap B$ and $x = a = \gamma b \in A \cap \gamma B$, we have $xg = ag \in A$ and $xg = \gamma bg \in \gamma B$. The action is simple and transitive hence
$$|A \cap \gamma B| = |A \cap B| \leqslant C^{-1}|A|.$$
We claim that there exists
$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A$$
with $abcd \neq 0$. Indeed the set of matrices
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\overline{k}),\ abcd = 0$$
is the union of the cosets of B_∞
$$B_\infty,\ B_0 = wB_\infty w,\ wB_\infty, wB_0$$
(corresponding to $c = 0,\ b = 0,\ a = 0,\ d = 0$). Hence if $C > 4$ and $|A|$ is large enough, we have
$$|A \cap (B_\infty \cup B_0 \cup wB_\infty, wB_0)| < |A|$$
hence g exists.

Up to conjugating A we may assume that
$$T = \{\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix},\ t \in \overline{k}^\times\}$$
is the diagonal torus; let $T_A = A \cap T$. A priori it would be sufficient to construct an injective map
$$\varphi : T_A \times T_A \times T_A \hookrightarrow A$$

for then

$$|T_A|^3 = |\varphi(T_A, T_A, T_A)| \leqslant |A|.$$

In fact we don't need the map to be exactly injective. It will be sufficient that its fibers have uniformly bounded sizes or even that almost all of them have this property. We will do this by producing a map defined via polynomials equations of bounded degrees.

Write

$$T_A = A \cap T = \{\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix},\ t \in H_A\}$$

for some finite subgroup $H_A \subset \overline{k}^\times$.

Given $t_1, t_2, t_2 \in H_A$ we define

$$\begin{aligned}\varphi(t_1, t_2, t_3) &= \begin{pmatrix} t_1 & 0 \\ 0 & t_1^{-1} \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} t_2 & 0 \\ 0 & t_2^{-1} \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} t_3 & 0 \\ 0 & t_3^{-1} \end{pmatrix} \\ &= \begin{pmatrix} t_1 & 0 \\ 0 & t_1^{-1} \end{pmatrix} \begin{pmatrix} a^2 t_2 + bct_2^{-1} & act_2 + bdt_2^{-1} \\ act_2 + cdt_2^{-1} & bct_2 + d^2 t_2^{-1} \end{pmatrix} \begin{pmatrix} t_3 & 0 \\ 0 & t_3^{-1} \end{pmatrix} \in A\end{aligned}$$

Since $abcd \neq 0$ there at most 8-values of t_2 for which one of the entries in the middle matrix is zero. Given any t_2 away from these values, varying $t_1, t_3 \in H_A$ we obtain $|H_A|^2$ distinct elements. In addition the product of the diagonal entries of the triple product does not depend on t_1, t_3 and equals

$$(a^2 t_2 + bct_2^{-1})(bct_2 + d^2 t_2^{-1}).$$

As t_2 varies over H_A this function takes $\geqslant |H_A|/4$ distinct values. In conclusion one obtains at least $\gg |H_A|^3$ distincts elements in A hence

$$|T_A| = |H_A| \ll |A|^{1/3}$$

where the implicit constant is absolute. □

Proposition 6.10. *(LP for unipotent subgroups) There exist constants $C, D > 0$ such that for any finite subgroup $A \subset \mathrm{SL}_2(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds*

- *For any unipotent subgroup N,*

$$|N \cap A| \leqslant C|A|^{1/3}.$$

- *There is a Borel subgroup B such that*

$$|B \cap A| \geqslant C^{-1}|A|.$$

Proof. Exercise. (hint: use a similar method but also involve the inverse of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$) □

6.4.2. Intersection with conjugacy classes. Given $g \in \mathrm{G}(\overline{k})$ we let

$$\mathrm{Conj}(g) = \{hgh^{-1},\ h \in \mathrm{G}(\overline{k})\}$$

the conjugacy class of g. This is an algebraic subvariety of $\mathrm{G}(\overline{k})$.

We have

- If $g = \pm \mathrm{Id}_2$, $\mathrm{Conj}(g) = \{g\}$,
- If g is regular unipotent $\mathrm{Conj}(g)$ is the set of all regular unipotent elements.
- If g is semisimple, $\mathrm{Conj}(g) = \{h \in \mathrm{G}(\overline{k}),\ \mathrm{tr}(h) = \mathrm{tr}(g)\}$.

Here we will use the *upper bounds* from Proposition 6.9 and 6.10 to obtain *lower bounds* on the intersection with conjugacy classes using the orbit-stabilizer theorem:

PROPOSITION 6.11 (LP, large intersection with a conjugacy classes). *There exist constants $C, D > 0$ such that for any finite subgroup $A \subset \mathrm{G}(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds*

- *For any $g \in A$ regular,*
$$|\mathrm{Conj}(g) \cap A| \geqslant C^{-1}|A|^{2/3}.$$
- *There is a Borel subgroup B such that*
$$|B \cap A| \geqslant C^{-1}|A|.$$

PROOF. We treat only the case g regular semisimple and leave the unipotent case as an exercise.

We assume that for any Borel subgroup B, $|A \cap B| \leqslant C^{-1}|A|$ where $C \geqslant 1$ is to be chosen sufficiently large.

Consider the map
$$c_g : h \in A \mapsto hgh^{-1} \in \mathrm{Conj}(g) \cap A =: \mathrm{Conj}(g)_A.$$
Given any $g' \in \mathrm{Conj}(g)_A$, the preimage of g' by c_g is contained in a coset of
$$\mathrm{Cent}_g(\overline{k}) = T_g$$
the unique maximal torus containing g and therefore by Prop. 6.9
$$|c_g^{(-1)}(\{g'\})| \leqslant C|A|^{1/3}$$
but this shows that
$$|c_g(A)| \geqslant C^{-1}|A|^{2/3}.$$
□

Now we prove the (upperbound) LP inequalities for conjugacy classes of regular elements:

PROPOSITION 6.12 (LP, small intersection with conjugacy classes). *There exist constants $C, D > 0$ such that for any finite subgroup $A \subset \mathrm{SL}_2(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds*

- *For any $g \in \mathrm{SL}_2(\overline{k})$, regular (either semisimple or quasi-unipotent)*
$$|\mathrm{Conj}(g) \cap A| \leqslant C|A|^{2/3}.$$
- *There is a Borel subgroup B such that*
$$|B \cap A| \geqslant C^{-1}|A|.$$

PROOF. We prove the case semisimple and leave the quasi-unipotent case as an exercise.

A first basic observation is that for any $h \in \mathrm{SL}_2(\overline{k})$ we have
$$\mathrm{tr}(h) = \mathrm{tr}(h^{-1})$$
(since $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ and therefore if g is regular semisimple we have
$$\mathrm{Conj}(g) = \{h \in \mathrm{SL}_2(\overline{k}),\ \mathrm{tr}(h) = \mathrm{tr}(g)\} = \mathrm{Conj}(g)^{-1} = \mathrm{Conj}(g^{-1}).$$

We assume that for any Borel subgroup B we have $|B \cap A| \leqslant C^{-1}|A|$ where $C \geqslant 1$ to be chosen sufficiently large[1]. We aim to prove that

$$|A \cap \mathrm{Conj}(g)| \leqslant C|A|^{2/3}.$$

For this we (following Larsen-Pink) consider the following map

$$\varphi : (c_1, c_2, c_3) \in \mathrm{Conj}(g)^3 \mapsto (c_1c_2, c_1c_3) \in \mathrm{SL}_2(k)^2$$

or rather its restriction to $\mathrm{Conj}(g)^3_A$ with

$$\mathrm{Conj}(g)_A := A \cap \mathrm{Conj}(g) = \{c \in A,\ \mathrm{tr}(c) = \mathrm{tr}(g)\}$$

(note particular $\mathrm{Conj}(g)_A$ contains only regular semisimple elements).

LEMMA 6.13. *Given $g \in \mathrm{SL}_2(\overline{k})$ a regular element (semisimple or quasi-unipotent) $(a, b) \in \mathrm{SL}_2(k)^2$, the map*

$$(c_1, c_2, c_3) \mapsto c_1 = ac_2^{-1} = bc_2^{-1}$$

a bijection

$$\varphi^{(-1)}(a, b) \simeq \mathrm{Conj}(g) \cap b\mathrm{Conj}(g)^{-1} \cap a\mathrm{Conj}(g)^{-1}.$$

If in addition g is semisimple we have $\mathrm{Conj}(g)^{-1} = \mathrm{Conj}(g)$ so that the bijection can be written

$$\varphi^{(-1)}(a, b) \simeq \mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g).$$

The restriction of φ to $\mathrm{Conj}(g)^3_A$ has image in A^2; we will bound $|\mathrm{Conj}(g)_A|$ by bounding $|\mathrm{Conj}(g)^3_A| = |\mathrm{Conj}(g)_A|^3$ and obtain the later by "slicing" $\mathrm{Conj}(g)^3_A$ into the fibers $\varphi^{(-1)}(a, b) \cap A^3$.

By the previous lemma, we have

$$|\mathrm{Conj}(g)^3_A| = \sum_{a,b \in A} |\mathrm{Conj}(g)_A \cap a\mathrm{Conj}(g)_A \cap b\mathrm{Conj}(g)_A|$$

and we will bound $|\mathrm{Conj}(g)_A \cap a\mathrm{Conj}(g)_A \cap b\mathrm{Conj}(g)_A|$ either pointwise or on average over some suitable families of (a, b). For this we observe that the set

$$\mathrm{Conj}(g)_A \cap a\mathrm{Conj}(g)_A \cap b\mathrm{Conj}(g)_A = \mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g) \cap A$$

is the intersection of A with

$$\mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g)$$

which is an algebraic subvariety of $\mathrm{Conj}(g) \subset \mathrm{SL}_2(\overline{k})$ to which we may want to apply the LP inequalities.

Let us do some rough dimensional analysis: the intersection $\mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g)$ is the intersection of three 2-dimensional varieties in a 3 dimensional space so one would hope that for most values of (a, b) the has dimension 0 ie. is finite absolutely bounded (think of the intersection of three affine planes in a three dimensional space). Hence we expect that for most $(a, b) \in A^2$,

$$|\mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g) \cap A|$$

should be absolutely bounded and the contribution of such generic terms is $\ll |A|^2$.

If $\mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g)$ is one dimensional then a and b are somewhat "related" (for instance we could have $a = b$ or more generally, as we will see, b is contained in a one

[1]As we will see the values of C depends on the "C" in Propositions 6.9 and 6.10

dimensional subvariety associated with a). Applying the LP inequalities to count the b in that subvariety along with the trivial bound

$$|\mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g) \cap A| \leqslant |\mathrm{Conj}(g) \cap A|$$

we obtain that the contribution of such special terms is bounded by

$$\ll \sum_{a \in A} |A|^{1/3} |\mathrm{Conj}(g)_A| \ll |A|^{4/3} |\mathrm{Conj}(g)_A|$$

Finally if $\mathrm{Conj}(g) \cap a\mathrm{Conj}(g) \cap b\mathrm{Conj}(g)$ is two dimensional which basically means that

$$\mathrm{Conj}(g) = a\mathrm{Conj}(g) = b\mathrm{Conj}(g)$$

which can occur only in very rare cases (such as $a = b = \mathrm{Id}$) which are handled easily using that $|\mathrm{Conj}(g)_A| \leqslant |A|$.

Let us implement this strategy (we will proceed a bit differently however) and prove the upper bound

$$|\mathrm{Conj}(g)_A|^3 = |\sum_{a,b \in A} |\mathrm{Conj}(g)_A \cap a\mathrm{Conj}(g)_A \cap b\mathrm{Conj}(g)_A| \ll |A|^2 + |A|^{4/3} |\mathrm{Conj}(g)_A|.$$

If $a = \pm \mathrm{Id}_2$ we use the trivial bound

$$|\mathrm{Conj}(g)_A \cap a\mathrm{Conj}(g)_A \cap b\mathrm{Conj}(g)_A| \leqslant |\mathrm{Conj}(g)_A|$$

so that

$$\sum_{b \in A} |\mathrm{Conj}(g)_A \cap a\mathrm{Conj}(g)_A \cap b\mathrm{Conj}(g)_A| \leqslant |\mathrm{Conj}(g)_A||A| \leqslant |A|^2.$$

Same for $b = \pm \mathrm{Id}_2$ and $a = \pm b$.

We may assume that $\pm \mathrm{Id}_2, \pm a, \pm b$ are distinct and in particular a and b are regular either semisimple or quasi-unipotent.

Suppose that $\mathrm{Id}_2, a^{-1}, b^{-1}$ are k-linearly dependent in the vector space $M_2(k)$. This implies that either, a and b are both regular and either both semisimple (and then b belong to the unique maximal torus containing a) or both quasi-unipotent (and then $\pm b$ belong to the unique unipotent subgroup containing $\pm a$). In particular, by Prop. 6.9 or Prop. 6.10, given any such a, the number of such b's is bounded by $C|A|^{1/3}$ and the sum over such a, b's is bounded by

$$|A| C |A|^{1/3} |\mathrm{Conj}(g)_A| = C |A|^{4/3} |\mathrm{Conj}(g)_A|.$$

It remains to treat the case where $\mathrm{Id}_2, a^{-1}, b^{-1}$ are k-linearly independent; they are in fact $\overline{k}$ linearly independent (put the three elements in a k-basis of $M_2(k)$, their determinant is $\neq 0$ and so the four elements are still linearly independent in $M_2(\overline{k})$).

In this case

$$\mathrm{Conj}(g)_A \cap a\mathrm{Conj}(g)_A \cap b\mathrm{Conj}(g)_A = A \cap L$$

where

$$L = L(g, a, b) := \{\ell \in M_2(\overline{k}),\ \mathrm{tr}(\ell) = \mathrm{tr}(a^{-1}\ell) = \mathrm{tr}(b^{-1}\ell) = \mathrm{tr}(g)\}$$

is an affine line in $M_2(\overline{k})$ (since $\mathrm{Id}_2, a^{-1}, b^{-1}$ are $\overline{k}$-linearly independent the three linear forms $m \mapsto \mathrm{tr}(m),\ \mathrm{tr}(a^{-1}m),\ \mathrm{tr}(b^{-1}m)$ are linearly independent in the dual $M_2(\overline{k})^*$).

If $L \not\subset \mathrm{SL}_2(\overline{k})$ then $L \cap \mathrm{SL}_2(\overline{k})$ consists in at most two points : to see this use the fact that $\mathrm{SL}_2(\overline{k})$ is the algebraic subvariety of $M_2(\overline{k})$ defined by the polynomial equation of degree at most 2, $ad - bc = 1$ and use a parametrisation of the affine line

$$L : x \in \overline{k} \mapsto \ell_0 + \ell' x$$

for $L' \in M_2(\overline{k})$.

Therefore the contribution of such a, b is bounded by

$$2|A|^2.$$

Suppose now that we are in the special case

$$L \subset \mathrm{SL}_2(\overline{k}).$$

Given $\ell_0 \in L$, the shifted line $\ell_0^{-1}.L$ contains Id_2 and admits a parametrization of the shape

$$x \in \overline{k} \mapsto \mathrm{Id}_2 + \ell_0^{-1}\ell' x.$$

We have

$$\begin{aligned}\det(\ell_0^{-1}L(x)) &= \det\begin{pmatrix} 1+a'x & b'x \\ c'x & 1+d'x \end{pmatrix} \\ &= (1+a'x)(1+d'x) - b'c'x^2 \\ &= 1 + \mathrm{tr}(\ell_0^{-1}\ell')x + \det(\ell_0^{-1}\ell')x^2 = 1\end{aligned}$$

and therefore

$$\mathrm{tr}(\ell_0^{-1}\ell') = \det(\ell_0^{-1}\ell') = 0.$$

This shows that $\ell_0^{-1}\ell'$ is conjugate to a nilpotent upper triangular matrix $\begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix}$ and therefore $\ell_0^{-1}L$ is conjugate to the group of unipotent matrices

$$N = \{\begin{pmatrix} 1 & xu \\ 0 & 1 \end{pmatrix},\ x \in \overline{k}\} = \{\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},\ x \in \overline{k}\}$$

(note that $u \neq 0$ since L has more than 1 element).

Therefore there exists $h \in \mathrm{SL}_2(\overline{k})$ such that

$$h\ell_0^{-1}Lh^{-1} = h\ell_0^{-1}h^{-1}hLh^{-1} = N.$$

Setting $\ell_0' = h\ell_0 h^{-1}$ $a' = hah^{-1},\ b' = hbh^{-1}$ we have

$$\begin{aligned}\ell_0' N &= \{hmh^{-1} \in M_2(\overline{k}), \mathrm{tr}(m) = \mathrm{tr}(a^{-1}m) = \mathrm{tr}(b^{-1}m) = \mathrm{tr}(g)\} \\ &= \{\ell_0' n,\ n = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix},\ x \in \overline{k},\ \mathrm{tr}(\ell_0' n) = \mathrm{tr}(a'^{-1}\ell_0' n) = \mathrm{tr}(b'^{-1}\ell_0' n) = \mathrm{tr}(g)\}.\end{aligned}$$

We have for any $x \in \overline{k}$

$$\mathrm{tr}(\begin{pmatrix} s & t \\ u & v \end{pmatrix}\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}) = s + ux + v.$$

Since the traces are independent of x the lower left entries of $\ell_0',\ a',\ b'$ have to be zero and therefore

$$\ell_0',\ a',\ b' \in B_\infty$$

and in particular

$$a, b \in B_h = h^{-1}B_\infty h$$

and

$$\ell_0^{-1}L = h^{-1}Nh = N_h.$$

Moreover N_h is the stabilizer of one of the eigenvectors of a (since N is the stabilizer of the first vector of the canonical basis) so given a, the unipotent subgroup N_h can take at most two values as does B_h (being the normalizer of N_h).

Finally, let $\{t, t^{-1}\}$ be the two eigenvalues of g (these are distincts since g is regular semisimple). The diagonal entries of ℓ_0', $a'^{-1}\ell_0'$, $b'^{-1}\ell_0'$ are equal to either (t, t^{-1}) or (t^{-1}, t) and those of a' and b' are either (t^2, t^{-2}) or (t^{-2}, t^2) or $(1, 1)$; therefore given a', b' is contained in at most three cosets of U_∞ or in other terms, given a, the element b is contained in at most three cosets of U_h. By Proposition 6.10, this implies that given a, there are at most $6C|A|^{1/3}$ possible values of b's so the contribution of (a, b)'s such that Id_2, a^{-1} and b^{-1} are linearly independent is bounded by

$$2|A|^2 + 6C|\mathrm{Conj}(A, g)||A||A|^{1/3}.$$

□

Since the conjugacy class of a regular unipotent element g consists of all regular unipotent elements, and the conjugacy class of the quasi-unipotent $-g$ is $-$ the conjugacy class of g we have

Corollary 6.14. *There exist constants $C, D > 0$ such that for any finite subgroup $A \subset \mathrm{SL}_2(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds*

- *There is a Borel subgroup B such that*

$$|B \cap A| \geqslant C^{-1}|A|.$$

- *A contains at most $2C|A|^{2/3}$ quasi-unipotents elements. In particular A contains at least $(2C)^{-1}|A|^{1/3} - 2$ regular semisimple elements.*

Corollary 6.15 (Large intersection with tori). *There exist constants $C, D > 0$ such that for any finite subgroup $A \subset \mathrm{G}(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds*

- *For any $g \in A$, regular semisimple contained in the maximal torus T_g we have*

$$|T_g \cap A| \geqslant C^{-1}|A|^{1/3}.$$

- *There is a Borel subgroup B such that*

$$|B \cap A| \geqslant C^{-1}|A|.$$

Proof. As usual we assume that we are not in the second situation. We have a map

$$a \in A \mapsto aga^{-1} \in A \cap \mathrm{Conj}(g)$$

with image of size $\leqslant C|A|^{2/3}$ so there is $g' \in A \cap \mathrm{Conj}(g)$ with preimage $\geqslant C^{-1}|A|^{1/3}$ but since all preimages are translates of one another, all have the same size $\geqslant C^{-1}|A|^{1/3}$ and in particular the preimage of g which is $A \cap \mathrm{Cent}_g(\overline{k}) = A \cap T_g$. □

This corollary admits a unipotent version which we leave as an exercise.

Corollary 6.16. *There exist constants $C, D > 0$ such that for any finite subgroup $A \subset \mathrm{G}(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds*

- *For any $g \in A$, regular unipotent contained in the unipotent subgroup U_g we have*

$$|U_g \cap A| \geqslant C^{-1}|A|^{1/3}.$$

- *There is a Borel subgroup B such that*

$$|B \cap A| \geqslant C^{-1}|A|.$$

6.4.3. A rough description of the finite subgroups of $\mathrm{SL}_2(\overline{k})$ over finite fields. From Corollary 6.15 we see that if A is not roughly contained in any Borel subgroup then for any maximal torus T, the following dichotomy holds:

$$\text{either } |A\cap T|\leqslant 2 \text{ or } |A\cap T|\geqslant C^{-1}|A|^{1/3} \tag{6.3}$$

depending on whether $A\cap T$ is contained in $\{\pm \mathrm{Id}_2\}$ or contains at least one regular element.

This dichotomy can be used to prove the following classification result which we will not prove in the live course unless we have time.

THEOREM 6.17 (Rough description of the finite subgroups of $\mathrm{SL}_2(\overline{k})$). *There exist constants $C, D > 0$ such that for $\overline{k} = \overline{\mathbb{F}_p}$ the algebraic closure of a finite field k and any finite subgroup $A \subset \mathrm{SL}_2(\overline{k})$ satisfying $|A| \geqslant D$, one of the following holds:*

- *there is a finite subfield $k \supset \mathbb{F}_p$ satisfying*
$$C^{-1}|A|^{1/3} \leqslant |k| \leqslant C|A|^{1/3}$$
such that A is contained in a conjugate of $\mathrm{SL}_2(k)$ (in particular A has index $\leqslant C$ in that conjugate).
- *There is a Borel subgroup B such that*
$$|B\cap A| \geqslant C^{-1}|A|.$$

PROOF. To be included. □

6.5. Larsen-Pink inequalities for approximate subgroups

We now prepare for the proof Theorem 6.2 by obtaining LP inequality for approximate subgroups.

From now on, we assume that k is a finite field and that $A \subset \mathrm{SL}_d(k)$ is a K-approximate subgroup generating $\mathrm{G}(k) = \mathrm{SL}_d(k)$.

The main tool for the proof will be LP inequalities for A which are proving by using the group theoretic versions from the previous sections.

6.5.1. Approximate subgroup versions of the orbit-stabilizer theorem.

LEMMA 6.18 (Orbit-Stabilizer for approximate subgroups). *Let $G \curvearrowright X$ be a group acting on a set X and $A \subset G$ be a set. We have for $x \in X$*

$$|A.A^{-1}\cap G_x| \geqslant \frac{|A|}{|A.x|} \tag{6.4}$$

and for any $B \subset G$ we have

$$|BA| \geqslant |A\cap G_x||B.x|. \tag{6.5}$$

Here $G_x = \mathrm{Stab}_G(x)$ denote the stablizer of x.

PROOF. We have $a.x = a'.x \iff a'a^{-1} \in G_x$. Hence for any $y = ax \in A.x$ we have

$$r_A(y) = |\{a' \in A,\ a'x = y\}| = |Aa^{-1}\cap G_x| \leqslant |A.A^{-1}\cap G_x|.$$

We then have

$$\sum_{y\in A.x} r_A(y) = \sum_{y\in A.x}\sum_{\substack{a'\in A\\ a'x=y}} 1 = \sum_{a'\in A}\sum_{y\in A.x, a'x=y} 1 = |A|$$

and

$$\sum_{y\in A.x} r_A(y) \leqslant |A.x||A.A^{-1}\cap G_x|.$$

For the second inequality we observe that

$$|BA| \geqslant |B.(A\cap G_x)|$$

and that

$$B.(A\cap G_x).x = B.x.$$

□

LEMMA 6.19. *Let $G \curvearrowright X$ be a group acting on a set X and $A \subset G$ be a symmetric set ($A^{-1} = A$). For any $x \in X$ and $m \geqslant 1$ we have*

$$|A^{(m+1)}| \geqslant \frac{|A^{(m)}\cap G_x|}{|A^{(2)}\cap G_x|}|A|.$$

In particular if $e_G \in A$ (so that $A^{(m)} \subset A^{(m+1)}$) and $|A^{(m+1)}| \leqslant K^m|A|$ we have

$$1 \leqslant \frac{|A^{(m)}\cap G_x|}{|A^{(2)}\cap G_x|} \leqslant K^m.$$

PROOF. We have by (6.5) and (6.4)

$$|A^{(m+1)}| \geqslant |A^{(m)}\cap G_x||A.x| = \frac{|A^{(m)}\cap G_x|}{|A^{(2)}\cap G_x|}|A^{(2)}\cap G_x||A.x| \geqslant \frac{|A^{(m)}\cap G_x|}{|A^{(2)}\cap G_x|}|A|.$$

□

Applying this to the action $G \curvearrowright G/H$ for H a subgroup and $x = e_G.H$ we obtain

COROLLARY 6.20. *Let G be a group and $A \subset G$ be a symmetric set ($A^{-1} = A$). For any subgroup $H \subset G$ and $k \geqslant 1$ we have*

$$|A^{(m+1)}| \geqslant \frac{|A^{(m)}\cap H|}{|A^{(2)}\cap H|}|A|.$$

In particular if $e_G \in A$ (so that $A^{(m)} \subset A^{(m+1)}$) and $|A^{(m+1)}| \leqslant K^m|A|$ we have

$$1 \leqslant \frac{|A^{(m)}\cap H|}{|A^{(2)}\cap H|} \leqslant K^m.$$

6.5.2. Escaping Borel subgroups. In the group theoretic version of the LP inequalities a part of the alternative was that the subgroup A was not roughly contained in any Borel subgroup.

In this section we develop a version of the LP inequalities when A is a generating approximate subgroup. We first show that the Borel subgroup rough containment never occurs (at least of k is large enough).

LEMMA 6.21 (Escape from Borel subgroups). *Let k, A as above. There exists an absolute constant $D \geqslant 1$ such that for any $C \geqslant 1$, one of the following holds*

- *one has $|A| \leqslant K^{DC}$;*
- *for any Borel subgroup $B \subset \mathrm{SL}_2(\overline{k})$*

$$|A^{(2)}\cap B| \leqslant K^{-C}|A|.$$

PROOF. We assume that $|A| > K^{DC}$ (in particular $|k| \gg K^{DC/3}$).

Assume that for some B one has $|A^{(2)} \cap B| > K^{-C}|A|$. Since $A^{(2)}$ is covered by at most K-translates of A there exists $a \in A$ such that

$$|aA \cap B| > K^{-C-1}|A|.$$

We also claim (see below) that for $|k|$ is sufficiently large there exists $g \in \mathrm{SL}_2(k)$ such that $gBg^{-1} \neq B$; since A generates $\mathrm{SL}_2(k)$ this implies that there exists $b \in A$ such that

$$B_b := bBb^{-1} \neq B.$$

In particular since the two Borel subgroups B and B_b are distincts $B \cap B_b = T_b$ is a maximal torus.

We have $bA^{(2)}b^{-1} \cap B_b \subset A^{(4)} \cap B_b$ hence

$$|A^{(4)} \cap B_b| \geqslant |bA^{(2)}b^{-1} \cap bBb^{-1}| = |A^{(2)} \cap B| > K^{-C}|A|.$$

Since $A^{(4)} = A^{(3)}.A$ is covered by at most K^3 successive translates of A there exists $a_3 \in A^{(3)}$ such that

$$|a_3A \cap B_b| > K^{-C-3}|A|.$$

Let

$$A_1 = aA \cap B,\ A_2 = a_3A \cap B_b.$$

We have

$$\sum_{g \in \mathrm{SL}_2(k)} 1_{A_1} \star 1_{A_2^{-1}}(g) = |A_1||A_2| > K^{-2C-4}|A|^2.$$

Since $\mathrm{supp}(1_{A_1} \star 1_{A_2^{-1}}) \subset A^{(6)}$ we have

$$|\mathrm{supp}(1_{A_1} \star 1_{A_2^{-1}})| \leqslant |A^{(6)}| \leqslant K^5|A|,$$

and there exists $a_6 \in A^{(6)}$ such that

$$1_{A_1} \star 1_{A_2^{-1}}(a_6) = |A_1 \cap a_6A_2| \geqslant K^{-2C-9}|A|.$$

Hence the product set

$$(A_1 \cap a_5A_2)^{-1}.(A_1 \cap a_5A_2)$$

which is of size $\geqslant K^{-2C-9}|A|$ is contained in the intersection of

$$A.a^{-1}.a.A = A^{(2)},\ B.B = B \text{ and } B_b.B_b = B_b$$

and therefore

$$|A^{(2)} \cap T_b| = |A^{(2)} \cap B \cap B_b| \geqslant K^{-2C-9}|A|.$$

We claim again that there exist $g \in \mathrm{SL}_2(k)$ such that $gT_bg^{-1} \neq T_b$. Indeed such a g would have to be in the normalizer of T_b which is $T_b \sqcup w_bT_b$ and so would have to be semisimple but if k is large enough $\mathrm{SL}_2(k)$ contains regular non-semisimple elements. Since A generates $\mathrm{SL}_2(k)$ this implies that there is $c \in A$ such that $T_{b,c} = cT_bc^{-1} \neq T_b$ and therefore (since we are intersecting distinct maximal tori)

$$T_{b,c} \cap T_b = \{\pm \mathrm{Id}_2\}.$$

On the other hand the same reasoning as above shows that

$$|A^{(2)} \cap T_b \cap T_{b,c}| \geqslant K^{-2(2C+9)-9}|A|$$

which yields a contradiction if A is large enough. □

It remains to prove the claim which we leave as an exercise.

LEMMA 6.22. *There exists an absolute constant D such that for any finite field k satisfying $|k| \geqslant D$, any Borel subgroup $B \subset \mathrm{SL}_2(\overline{k})$ there is $g \in \mathrm{SL}_2(k)$ such that*

$$gBg^{-1} \neq B.$$

PROOF. (hint) Use the fact that $B = B_z$ for some $z \in \mathbb{P}^1(\overline{k})$ to show that any $g \in B_z$ has its matrix entries satisfying a linear equation (depending on z and with coefficients in $\overline{k}$). Show that if $|k|$ is large enough there is some $g \in \mathrm{SL}_2(k)$ not satisfying this equation (because $|\mathrm{SL}_2(k)| \gg |k|^3$). □

6.5.3. A series of LP inequalities. We can then use this *Escape from the Borels* "wildcard" lemma to obtain approximate subgroup versions of the Larsen-Pink inequalities.

PROPOSITION 6.23 (LP for tori (App. subgroup version)). *There exist a constant $C > 0$ such that for any generating K-approximate subgroup $A \subset \mathrm{SL}_2(k)$ and any maximal torus $T \subset \mathrm{SL}_2(\overline{k})$, we have*

$$|T \cap A^{(2)}| \leqslant K^C |A|^{1/3}.$$

PROOF. We may assume that $|A| \geqslant K^{DC}$ for D the constant in Lemma 6.21 and C to be chosen sufficiently large, for otherwise the trivial bound

$$|T \cap A^{(2)}| \leqslant K|A|$$

will suffice (since $K|A| \leqslant K^C|A|^{1/3}$ if $|A| \leqslant K^{\frac{3}{2}(C-1)}$).

Moreover, by Lemma 6.21 and our assumption $|A| \geqslant K^{DC}$, we have that for any Borel subgroup $B \subset \mathrm{SL}_2(\overline{k})$

$$|A^{(2)} \cap B| \leqslant K^{-C}|A|. \tag{6.6}$$

Up to conjugating T (and A would then generate a group conjugate to $\mathrm{SL}_2(k) \subset \mathrm{SL}_2(\overline{k})$ but this does not change the argument) we may assume that

$$T = \{\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix},\ t \in \overline{k}^\times\}$$

is the diagonal torus. Let

$$T_A = T \cap A^{(2)} = \{\begin{pmatrix} t & \\ & t^{-1} \end{pmatrix},\ t \in H_A\}$$

for $H_A \subset \overline{k}^\times$. If $|H_A| \leqslant K^C$ we have

$$|T_A| \leqslant K^C \leqslant K^{DC/3} \leqslant |A|^{1/3}$$

as long as $D \geqslant 3$.

By the same reasoning as in the proof of Prop. 6.9, using (6.6) there exists $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A^{(2)}$ whose four entries satisfy $abcd \neq 0$ (as long as $K^C > 4$). We then have

$$T_A g T_A g T_A \subset A^{(10)}$$

so for all $t_1, t_2, t_3 \in H_A$

$$\begin{aligned}\varphi(t_1,t_2,t_3) &= \begin{pmatrix} t_1 & 0 \\ 0 & t_1^{-1} \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} t_2 & 0 \\ 0 & t_2^{-1} \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} t_3 & 0 \\ 0 & t_3^{-1} \end{pmatrix} \\ &= \begin{pmatrix} t_1 & 0 \\ 0 & t_1^{-1} \end{pmatrix}\begin{pmatrix} a^2t_2 + bct_2^{-1} & act_2 + bdt_2^{-1} \\ act_2 + cdt_2^{-1} & bct_2 + d^2t_2^{-1} \end{pmatrix}\begin{pmatrix} t_3 & 0 \\ 0 & t_3^{-1} \end{pmatrix} \in A^{(10)}\end{aligned}$$

and by the same reasoning as in Prop. 6.9 we find that

$$|T \cap A^{(2)}|^3 = |H_A|^3 \ll |A^{(10)}| \leqslant K^9|A|$$

where the implicit constant is absolute. □

COROLLARY 6.24. *Assumptions as in Prop. 6.23. For any $m \geqslant 1$, there exists $C_m \geqslant 1$ such that one has*

$$|A^{(m)} \cap T| \leqslant K^{C_m}|A|^{1/3}$$

PROOF. Exercise (use §6.5.1). □

PROPOSITION 6.25. *(LP for unipotent subgroups (App subgroup version)) For any $m \geqslant 1$ exist constants $C_m > 0$ such that for any generating K-approximate subgroup $A \subset \mathrm{SL}_2(k)$ and any unipotent subgroup $N \subset \mathrm{SL}_2(\overline{k})$ one has*

$$|N \cap A^{(m)}| \leqslant K^{C_m}|A|^{1/3}.$$

PROOF. Exercise. □

PROPOSITION 6.26. *(LP, large intersections with conjugacy classes (App subgroup version)) There exist constant $C > 0$ such that for any generating K-approximate subgroup $A \subset \mathrm{SL}_2(k)$ and any $g \in A$ regular, one has*

$$|\mathrm{Conj}(g) \cap A^{(3)}| \geqslant K^{-C}|A|^{2/3}.$$

PROOF. Exercise. □

PROPOSITION 6.27. *(LP, small conjugacy classes (App subgroup version)) There exist constant $C > 0$ such that for any generating K-approximate subgroup $A \subset \mathrm{SL}_2(k)$ and any $g \in \mathrm{SL}_2(\overline{k})$ regular one has*

$$|\mathrm{Conj}(g) \cap A^{(3)}| \leqslant K^{C}|A|^{2/3}.$$

PROOF. Exercise. □

COROLLARY 6.28. *For any $m \geqslant 1$, there exist constant $C_m > 0$ such that for any generating K-approximate subgroup $A \subset \mathrm{SL}_2(k)$ and any $g \in \mathrm{SL}_2(\overline{k})$ regular one has*

$$|\mathrm{Conj}(g) \cap A^{(m)}| \leqslant K^{C_m}|A|^{2/3}.$$

PROOF. Exercise. □

Using that the conjugacy class of a regular unipotent element is the set of all regular unipotent elements one has

COROLLARY 6.29. *(LP, unipotent bound (App subgroup version)) There exist constant $C > 0$ such that for any generating K-approximate subgroup $A \subset \mathrm{SL}_2(k)$, A contains at most $2K^C|A|^{2/3}$ quasi-unipotent elements and at least $(2K^C)^{-1}|A|^{1/3} - 2$ regular semisimple elements.*

PROPOSITION 6.30 (Large intersection with tori (App subgroup version)). *There exist constant $C > 0$ such that for any generating K-approximate subgroup $A \subset \mathrm{SL}_2(k)$ and any $g \in A^{(2)}$ regular semisimple with associated maximal torus T_g we have*

$$|T_g \cap A^{(2)}| \geqslant K^{-C}|A|^{1/3}.$$

PROOF. Given $g \in A^{(2)}$ regular semisimple, the map

$$c_g : a \in A \mapsto aga^{-1} \in A^{(4)} \cap \mathrm{Conj}(g)$$

has image of size

$$|c_g(A)| \leqslant |A^{(4)} \cap \mathrm{Conj}(g)| \leqslant K^C |A|^{2/3}$$

(where $C = C_4$ from Cor. 6.28) so there exists some $g' = a' g a'^{-1} \in c_G(A)$ whose preimage is large:

$$|c_g^{(-1)}(\{g'\})| = |\{a \in A,\ aga^{-1} = g'\}| \geqslant K^{-C}|A|^{1/3}$$

but then any element of $a'^{-1} c_g^{(-1)}(\{g'\}) \subset A^{(2)}$ centralizes g so is contained in T_g; therefore $|A^{(2)} \cap T_g| \geqslant K^{-C}|A|^{1/3}$. □

DEFINITION 6.31. *Let $T \subset \mathrm{SL}_2(\overline{k})$ be a maximal torus. We say that T is involved with A if*

$$T \cap A^{(2)} \not\subset \{\pm \mathrm{Id}_2\}.$$

In this case, from Prop. 6.30, we have

$$|T \cap A^{(2)}| \geqslant K^{-C}|A|^{1/3}.$$

LEMMA 6.32 (Key lemma). *For K and A as above, one of the following holds*

- $|A| \leqslant K^C$,
- *If T is involved with A then for any $a \in A$, aTa^{-1} is also involved.*

PROOF. If T is involved then $|A^{(2)} \cap T| \geqslant K^{-C}|A|^{1/3}$ and conjugating by $a \in A$ we have

$$K^{-C}|A|^{1/3} \leqslant |A^{(2)} \cap T| \leqslant |aA^{(2)}a^{-1} \cap aTa^{-1}| \leqslant |A^{(4)}|$$

and since $A^{(4)}$ is covered by K^3 translates of A there is g such that

$$|gA \cap aTa^{-1}| \geqslant K^{-C-3}|A|^{1/3}$$

but then the product set

$$(Ag^{-1} \cap aTa^{-1}).(gA \cap aTa^{-1}) \subset A^{(2)} \cap aTa^{-1}$$

and has size $\geqslant |gA \cap aTa^{-1}| \geqslant K^{-C-3}|A|^{1/3}$ so that

$$|A^{(2)} \cap aTa^{-1}| \geqslant K^{-C-3}|A|^{1/3}.$$

If we are not in the first case (up adjusting the definition of C) then $K^{-C-3}|A|^{1/3} > 2$ and aTa^{-1} is also involved. □

REMARK 6.4. It is important that the definition of a torus T being "involved" is

$$|T \cap A^{(2)}| > 2$$

instead of $|T \cap A^{(2)}| \geqslant K^{-C}|A|$ (the consequence of Prop. 6.30) because as we have seen in the proof, we obtain upon conjugating by a the lower bound $|A^{(2)} \cap aTa^{-1}| \geqslant K^{-C-3}|A|^{1/3}$ which gets weaker and weaker if we perform more and more conjugations (as we will do below). On the other hand, once we know the much weaker lower bound statement $|aTa^{-a} \cap A^{(2)}| > 2$ we automatically get $|aTa^{-a} \cap A^{(2)}| \geqslant K^{-C}|A|^{1/3}$ (with the same C and not $C+3$).

6.5.4. Proof of Theorem 6.2. (Note: all the implicit constants involved in the symbols $\gg$ or $\approx$ below are absolute).

Assume that $|A| \geqslant K^{DC}$. From Prop. 6.29 A contain $\geqslant K^{-C}|A|^{1/3} - 2 > 0$ semisimple regular elements (for D large enough) and therefore A admits at least one involved torus. Since A generate $\mathrm{SL}_2(k)$, the set of all involved torus is $\mathrm{SL}_2(k)$-invariant under conjugation (reflect on Remark 6.4). Since $|\mathrm{SL}_2(k)| \gg |k|^3$ and for any torus T, $|\mathrm{SL}_2(k) \cap \mathrm{Nor}_T(\overline{k})| \leqslant 2|k|$ (for instance conjugate T to the diagonal torus to see this) there are at least $\gg |k|^3/|k| = |k|^2 \approx |\mathrm{SL}_2(k)|^{2/3}$ involved tori.

For each such torus T we have

$$|A^{(2)} \cap T| \geqslant K^{-C}|A|^{1/3}$$

which implies that (since two distinct tori intersect only in $\pm\mathrm{Id}$)

$$|A^{(2)}| \gg K^{-2C}|\mathrm{SL}_2(k)|^{2/3}|A|^{1/3}$$

and since

$$K|A| \gg |A^{(2)}| \gg K^{-2C}|\mathrm{SL}_2(k)|^{2/3}|A|^{1/3}$$

we have

$$|A| \gg K^{-(2C+1)3/2}|\mathrm{SL}_2(k)|.$$

□

CHAPTER 7

Expansion in $\mathrm{SL}_2(\mathbb{F}_p)$

7.1. Basic on graphs

We recall that a (finite) graph $\mathcal{G} = (V, E)$ where $V = V_{\mathcal{G}}$ is a finite set of "vertices" and $E = E_{\mathcal{G}} \subset V \times V - V^{\Delta}$ is a set[1] of "edges".

- If $(v, w) \in E$ we say that w is directly connected to v or that w is adjacent to v or that there is a path of length 1 between v and w which one write $v \mapsto w$.
- We say that w is connected to v by a path of length $l \geqslant 1$ if there is l edges $e_i = (v_{i-1}, v_i) \in E,\ i \leqslant l$ such that
$$v_0 = v,\ v_l = w,\ \forall i = 1, \cdots, l-1,\ v_i = v_{i+1}.$$
We say that w is connected to v if there is a path of some length between the two.
- If for any $v, w \in V$ one has
$$(v, w) \in E \Longleftrightarrow (w, v) \in E$$
we say that the graph is *undirected* (and otherwise it is a directed graph).
- From now on we assume that graphs are undirected. This implies that the relations being "directly connected" or being "connected" are symmetric and transitive (but not necessarily reflexive).
- The set of w directly connected to v is the set of neighbours of v, $\mathrm{Neig}(v)$. The set on w connected to v is the connected component of v, $\mathrm{Conc}(v)$. A graph is connected if it has only one connected component.
- The set of vertices is equipped with a distance $\mathrm{dist}_{\mathcal{G}} : V \times V \mapsto \mathbb{N} \sqcup \infty$ defined by $\mathrm{dist}_{\mathcal{G}}(v, v) = 0$ and
$$\mathrm{dist}_{\mathcal{G}}(v, w) = \text{minimal length of a path connecting } v \text{ to } w$$
(defined to be ∞ if v is not connected to w).
One define the diameter of $\mathcal{G}$ to be
$$\mathrm{diam}(\mathcal{G}) = \max_{v,w \in V} \mathrm{dist}_{\mathcal{G}}(v, w).$$
- The cardinality $d_v := |\mathrm{Neig}(v)|$ is the degree of $\mathcal{G}$ at the point v. If $d_v = d \geqslant 1$ for every v we say that $\mathcal{G}$ is regular of degree d.

We will write $\mathcal{G}_{n,d}$ to denote a d regular graph with n vertices.

7.1.1. Example: Cayley graphs. Let G be a finite group of order n and $S \subset G$ a symmetric subset of G of order d not containing e_G. We define the (left) Cayley graph $\mathcal{G}_{G,s,l} = (G, E_{S,l})$ where
$$E_{S,l} = \{(g, sg),\ g \in g,\ s \in S\}.$$
The graph is connected iff S generates G.

[1] sometimes E could be a multiset

We will often write a Cayley graph in the form

$$\mathcal{G} = (G, E_{S,l}) = \mathrm{Cayl}(G, S) \text{ or simply } (G, S).$$

REMARK 7.1. Likewise we can define the right Cayley graph $\mathcal{G}_{G,S,r} = (G, E_{S,r})$ where

$$E_{S,r} = \{(g, gs),\ g \in g,\ s \in S\}.$$

7.1.2. The adjacency operator. Suppose that $\mathcal{G}$ is d-regular undirected.
Let $L^2(V)$ be the space of functions on V equipped with the inner product

$$\langle f_1, f_2\rangle = \sum_{v\in V} f_1(v)\overline{f}_2(v)$$

The (normalized) adjacency operator $A : L^2(V) \to L^2(V)$ is the linear map defined by

$$f \mapsto Af : v \mapsto \sum_{(v,w)\in E} f(w).$$

To each vertex v, $Af(v)$ is the sum of f at the immediate neighbours of v.

More generally iterating the above we see that for any $\ell \geqslant 1$

$$A^\ell f(v) = \sum_{\substack{(w_t)_{t\leqslant \ell}\\ w_0=v}} f(w_\ell) \tag{7.1}$$

is the sum of the values of f at the end-points of the paths of length ℓ which start from v.

The *adjacency matrix* (also noted A) is the matrix of A is the basis

$$\{\delta_{v_1},\ \cdots, \delta_{v_n}\}$$

where $n = |V|$ and $\{v_i,\ i \leqslant n\} = V$ is an enumeration of the set of vertices of $\mathcal{G}$.

Writing the matrix $A = (A_{ij})$ then have

$$A_{ij} = \delta_{(v_i,v_j)\in E}$$

and for any $\ell \geqslant 1$

$$(A^\ell)_{ij} = \text{ number of paths of length } \ell \text{ joining } v_i \text{ to } v_j.$$

The adjacency operator is self-adjoint (because the graph is undirected):

$$\langle Af, f\rangle = \sum_v (\sum_{(v,w)\in E} f(w))\overline{f}(v) = \sum_w f(w) \sum_{(w,v)\in E} \overline{f}(v) = \langle f, Af\rangle$$

so is diagonalizable with its (multiset) of eigenvalues being real numbers which we write

$$\mathrm{Spec}(A) = \mathrm{Spec}(\mathcal{G}) = \{\lambda_0 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{n-1} \geqslant \lambda_{n-1}\}.$$

The structure of $\mathrm{Spec}(\mathcal{G})$ contains some of the geometry of $\mathcal{G}$. For instance we have

LEMMA 7.1. *We have $\lambda_0 = d$ and*

$$d = \lambda_0 \geqslant \lambda_2 \cdots \geqslant \lambda_{n-1} \geqslant -d.$$

PROOF. We have

$$A(1) = d.1$$

and for $\varphi \neq 0$ any eigenfunction (using that $|uv| \leqslant \frac{1}{2}(|u|^2 + |v|^2)$) we have

$$|\lambda|\langle\varphi,\varphi\rangle = |\langle A\varphi,\varphi\rangle| = |\sum_{\substack{v,w\\(v,w)\in E}}\sum \varphi(w)\overline{\varphi}(v)| \leqslant \frac{1}{2}\sum_{\substack{v,w\\(v,w)\in E}}\sum |\varphi(w)|^2 + |\varphi(v)|^2$$
$$= \frac{1}{2}(\sum_w |\varphi(w)|^2 d + \sum_v |\varphi(v)|^2 d) = d\langle\varphi,\varphi\rangle.$$

□

In the sequel we denote by $\mathcal{B}$ a ONB of eigenfunction

$$\mathcal{B} = \{\varphi_i,\ A\varphi_i = \lambda_i\varphi_i,\ i = 0,\cdots,n-1,\ \}$$

and take

$$\varphi_0 = 1/n^{1/2}$$

the constant function. We will write

$$\mathcal{B}_0 = \mathcal{B} - \{\varphi_0\} = \{\varphi_i,\ i = 1,\cdots,n-1,\ \}$$

which is an ONB of the subspace of function with mean values 0:

$$L^2(V)_0 = (\mathbb{C}\varphi_0)^\perp = \{f \in L^2(V),\ \langle f,1\rangle = \sum_{v\in V} f(v) = 0\}.$$

We recall the following basic results of graph theory:

PROPOSITION 7.2. *The multiplicity of $\lambda_1 = d$ is the number of connected components of $\mathcal{G}$.*

Suppose $\mathcal{G}$ is connected. We have $\lambda_{n-1} = -d$ iff $\mathcal{G}$ is a bi-partite graph: there exists a decomposition $V = V_+ \sqcup V_-$ with $|V_+| = |V_-|$ and

$$(v,w) \in E \Longleftrightarrow (v,w) \in V_\varepsilon \times V_{-\varepsilon},\ \varepsilon = \pm 1.$$

In that case $\mathrm{Spec}(\mathcal{G})$ *is symmetric relative to the origin:* $\mathrm{Spec}(\mathcal{G}) = -\mathrm{Spec}(\mathcal{G})$ *and the (one dimensional) eigenspace with eigenvalue $-d$ is generated by*

$$1_{V^+} - 1_{V^-}.$$

PROOF. Exercise. □

In particular the graph $\mathcal{G}$ is connected iff $\lambda_0 = d$ has multiplicity 1 or in other terms $\lambda_0 > \lambda_1$.

DEFINITION 7.3. *The graph $\mathcal{G}$ has a (one-sided) spectral gap iff*

$$\lambda_1 < d$$

The graph $\mathcal{G}$ has a two-sided spectral gap iff iff

$$-d < \lambda_{n-1},\cdots,\lambda_1 < d.$$

A graph with a one-sided spectral gap is therefore connected and a graph with a two-sided spectral gap is non bi-partite connected graph.

7.1.3. The averaging operator. It is useful to to normalize A and to consider instead the averaging operator

$$M = \frac{A}{d}$$

so that

$$Mf(v) = \frac{1}{d} \sum_{(v,w)\in E} f(w)$$

is the average value of f along the immediate neighbours of v (M is for "mean" or "moyenne").

We then have

$$\mathrm{Spec}(M) = \{\rho_0 = 1 \geqslant \rho_1 = \lambda_1/d \geqslant \cdots \geqslant \rho_{n-1} = \lambda_{n-1}/d\} \subset [-1, 1].$$

The graph has then a one(resp. two)-sided spectral gap iif

$$\mathrm{Spec}(M) - \{\rho_0\} = \{\rho_1 \geqslant \cdots \geqslant \rho_{n-1}\} \subset [-1, 1), \text{ resp. } \subset (-1, 1).$$

REMARK 7.2. One also define the *Laplace operator*

$$\Delta = \mathrm{Id} - M$$

whose eigenvalues are non negative

$$\mathrm{Spec}(\Delta) = \{0 = 1 - \rho_0 \leqslant 1 - \rho_1 \leqslant \cdots, 1 - \rho_{n-1}\} \subset [0, 1].$$

7.2. Expander graphs

The notion of expander graph is a quantification of the notion of graph with a spectral gap:

DEFINITION 7.4. *Given $\varepsilon \in (0, 1)$, the graph $\mathcal{G}$ is a (one-sided) ε-expander iff*

$$\lambda_1 \leqslant (1 - \varepsilon)d$$

or equivalently

$$\mathrm{Spec}(A) - \{\lambda_0\} \subset [-d, (1 - \varepsilon)d] \text{ or } \mathrm{Spec}(M) - \{\rho_0\} \subset [-1, (1 - \varepsilon)].$$

The graph $\mathcal{G}$ is a two-sided ε-expander iff

$$\forall i = 1, \cdots, n - 1, \ |\lambda_i| \leqslant (1 - \varepsilon)d \text{ or } |\rho_i| \leqslant (1 - \varepsilon)$$

or equivalently

$$\mathrm{Spec}(A) - \{\lambda_0\} \subset [-d(1 - \varepsilon), (1 - \varepsilon)d] \text{ or } \mathrm{Spec}(M) - \{\rho_0\} \subset [-(1 - \varepsilon), (1 - \varepsilon)].$$

REMARK 7.3. Of course any graph with a spectral gap is automatically an ε-expander for some $\varepsilon > 0$. This notion is really interesting as long as ε is not too small compared to either d or n. For instance this notion is interesting when ε does not depend on n (and d remains small compared to n).

DEFINITION 7.5. *Given $\varepsilon \in (0, 1)$ and $d \geqslant 1$ a family of (d, ε) expanders is a sequence $(\mathcal{G}_{n_i,d})_{i\in\mathbb{N}}$ of d regular connected graphs satisfying $n_i \to \infty$ and which are one or two-sided ε-expanders.*

7.2.1. Equidistribution and Mixing for expanders.

LEMMA 7.6. *Given $\mathcal{G}_{n,d}$ be a regular graph and $\mathcal{B} = \{n^{-1/2}\} \sqcup \mathcal{B}_0$ an ONB for M containing the constant function $\varphi_0 = n^{-1/2}$.*

We have for any $f_1, f_2 \in L^2(V)$

$$\langle f_1, M^\ell f_2\rangle = \frac{\langle f_1, 1\rangle\langle 1, f_2\rangle}{n} + \sum_{\varphi\in\mathcal{B}_0} \rho_\varphi^\ell \langle f_1, \varphi\rangle\langle \varphi, f_2\rangle.$$

PROOF. We have (since M is self-adjoint)

$$\begin{aligned}\langle f_1, M^\ell f_2\rangle &= \sum_{\varphi\in\mathcal{B}} \langle f_1, \varphi\rangle\overline{\langle M^\ell f_2, \varphi\rangle} = \sum_{\varphi\in\mathcal{B}} \langle f_1, \varphi\rangle\overline{\langle f_2, M^\ell\varphi\rangle} = \sum_{\varphi\in\mathcal{B}} \rho_\varphi^\ell \langle f_1, \varphi\rangle\overline{\langle f_2, \varphi\rangle}\\ &= \frac{\langle f_1, 1\rangle\overline{\langle f_2, 1\rangle}}{n} + \sum_{\varphi\in\mathcal{B}_0} \rho_\varphi^\ell \langle f_1, \varphi\rangle\overline{\langle f_2, \varphi\rangle}.\end{aligned}$$

□

Since

$$\sum_{\varphi\in\mathcal{B}_0} |\langle f_1, \varphi\rangle\langle \varphi, f_2\rangle| \leqslant \|f_1\|_2\|f_2\|_2$$

we obtain

COROLLARY 7.7. *Let $\mathcal{G}_{n,d}$ be a two sided ε-expander. For any $f_1, f_2 \in L^2(V)$ and any $\ell \geqslant 1$ we have*

$$\left|\langle f_1, M^\ell f_2\rangle - \frac{\langle f_1, 1\rangle\overline{\langle f_2, 1\rangle}}{n}\right| \leqslant (1-\varepsilon)^\ell \|f_1\|_2\|f_2\|_2.$$

where the implicit constant is absolute.

7.2.1.1. *Equidistribution.* Take $f_2 = \delta_{v_0}$ for some $v_0 \in V$ and let $f_1 = 1_W$ be the characteristic function of some subset $W \subset V$ we have from (7.1)

$$\langle 1_W, M^\ell \delta_{v_0}\rangle = \frac{1}{d^\ell} \sum_{\substack{(w_t)_{t\leqslant\ell}\\ w_0=v_0}} 1_W(w_\ell) = \frac{|\{(w_t)_{t\leqslant\ell},\ w_0 = v_0,\ w_\ell \in W\}|}{d^\ell}$$

is the proportion of the paths of length ℓ in the graph that start from v_0 and end up in W. By Corollary 7.7, we obtain

$$\frac{|\{(w_t)_{t\leqslant\ell},\ w_0 = v_0,\ w_\ell \in W\}|}{d^\ell} = \frac{|W|}{n} + O((1-\varepsilon)^\ell |W|^{1/2})$$

REMARK 7.4. Notice that the implicit constant is absolute. In particular it does not depend on v_0.

In other terms as $\ell \to \infty$, the probability that a path of length ℓ starting from v_0 ends in W is asymptotic to $|W|/n$ the measure of W relative to the uniform probability measure on V; moreover, we are sure that some of these paths will end-up in W as soon as

$$\ell \gg \frac{\log(n/|W|^{1/2})}{\log(1/(1-\varepsilon))}. \tag{7.2}$$

Let us generalize this result slightly. Given any probability measure ν_0 on $L^2(V)$: ie. a non-negative linear form $\nu_0 : f \mapsto \nu_0(f)$ ($\nu_0(f) \geqslant 0$ whenever $f \geqslant 0$) such that $\nu_0(1) = 1$ or equivalently a convex linear combination of Dirac masses

$$\nu_0 = \sum_{v \in V} \nu_0(v)\delta_v \text{ s.t. } \nu_0(v) \geqslant 0, \sum_{v \in V} \nu_0(v) = 1.$$

We can then define the sequence of probability measures

$$\nu^{(\ell)} : f \mapsto \nu_0(M^\ell(f)).$$

For instance $\nu_0 = \delta_{v_0}$

$$\nu_{v_0}^{(\ell)}(f) = M^\ell f(v_0) = \frac{1}{d^\ell} \sum_{\substack{(w_t)_{t \leqslant \ell} \\ w_0 = v_0}} f(w_\ell)$$

is the average value of f along the end-points of the paths of length ℓ in $\mathcal{G}$ starting from v_0.

Corollary 7.7 then state that the sequence of measures $\nu^{(\ell)}$ weak-$\star$ converge to the uniform probability measure μ_V on V which assigns mass $1/n$ to any vertex: for any $W \subset V$

$$\mu_V(W) = \frac{|W|}{n}.$$

We can interpret this as a *random walk* along the graph $\mathcal{G}$: $\nu_{v_0}^{(\ell)}$ is the distribution function of the random variable which is the end of the following process:

- Start from some $v_0 \in V$ chosen randomly according to the probability measure ν_0.
- Choose uniformly at random a point v_1 at distance 1 from v_0;
- choose uniformly at random a point v_2 at distance 1 from v_1,
- $\cdots$,
- iterate ℓ time and obtain v_l.

The previous computation shows that this process converge in law to the uniform random variable on V and in $\gg \log(n)$ steps will get very close.

Remark 7.5. This a special case of the convergence of irreducible Markov chains in a space with finitely many states. Recall that for such Markov chains, the key ingredient is the *Perron-Frobenius* theorem. Instead we have used the *spectral theorem* for the self-adjoint operator M.

7.2.2. Geometry of expanders. It turn out that expander graph have nice geometric properties that can be extracted from the properties of their spectrum.

7.2.2.1. *The diameter of an expander.* Let us return to the discussion in §7.2.1.1and take $W = v$ for any vertex $v \in V$, from (7.2) we see that v can be reached by a path of length ℓ starting from v_0 as soon as

$$\ell \gg \frac{\log(n)}{\log(1/(1-\varepsilon))}$$

or in other terms if $\mathcal{G}$ is a two sided ε- expander we have

$$\mathrm{diam}(\mathcal{G}) \ll \frac{\log(n)}{\log(1/(1-\varepsilon))}. \tag{7.3}$$

We will slightly improve this bound:

THEOREM 7.8. *Suppose that $\mathcal{G} = \mathcal{G}_{n,d}$ is a two sided ε-expander. We have*

$$\operatorname{diam}(\mathcal{G}) \ll \frac{\log(2n)}{\log(\frac{1+(1-(1-\varepsilon)^2)^{1/2}}{1-\varepsilon})}.$$

REMARK 7.6. In particular when ε is small

$$\log(\frac{1+(1-(1-\varepsilon)^2)^{1/2}}{1-\varepsilon}) = \varepsilon^{1/2} + O(\varepsilon)$$

and

$$\operatorname{diam}(\mathcal{G}) \ll \frac{\log(2n)}{\varepsilon^{1/2}}.$$

PROOF. Let

$$\mathcal{B} = \{\varphi_0 = 1/n^{1/2},\ \varphi_1, \cdots, \varphi_{n-1}\}$$

be an ONB of $L^2(V)$ made of eigenfunction of A with eigenvalues λ_φ. We have

$$\langle A(\delta_v), \delta_w\rangle = \sum_{\varphi\in\mathcal{B}} \lambda_\varphi \varphi(v)\overline{\varphi}(w)$$

and more generally for any polynomial $P(X) \in \mathbb{C}[X]$

$$\langle P(A)(\delta_v), \delta_w\rangle = \sum_{\varphi\in\mathcal{B}} P(\lambda_\varphi)\varphi(v)\overline{\varphi}(w).$$

Suppose $\operatorname{dist}(v,w) > N$ then for $\deg P \leqslant N$ we have

$$\langle P(A)(\delta_v), \delta_w\rangle = 0$$

because $P(A)$ is a linear combination of A^ℓ for $\ell \leqslant N$ and for each such ℓ $\langle A^\ell\delta_v, \delta_w\rangle = 0$ since v and w are not connected by any path of length ℓ.

We have therefore

$$\begin{aligned}
\frac{P(d)}{n} = P(\lambda_0)\varphi_0(v)\overline{\varphi}_0(w) &= -\sum_{\varphi\neq\varphi_0} P(\lambda_\varphi)\varphi(v)\overline{\varphi}(w)\\
&\leqslant (\sup_{|x|\leqslant(1-\varepsilon)d} |P(x)|)\sum_{\varphi\neq\varphi_0} |\varphi(v)\overline{\varphi}(w)|\\
&\leqslant (\sup_{|x|\leqslant(1-\varepsilon)d} |P(x)|)\frac{1}{2}\sum_{\varphi\neq\varphi_0} |\varphi(v)|^2 + |\varphi(w)|^2\\
&\leqslant \sup_{|x|\leqslant(1-\varepsilon)d} |P(x)|.
\end{aligned}$$

We apply this to

$$P(X) = P_N(X) = T_N(X/(1-\varepsilon)d)$$

where $T_N(X)$ is the N-th Chebycheff polynomial of the first kind

$$T_N(X) = \cos(N\arccos(X)) = \frac{1}{2}((X+\sqrt{X^2-1})^N + (X-\sqrt{X^2-1})^N).$$

We have for $|x| \leqslant (1-\varepsilon)d$

$$|P_N(x)| = |T_N(x/(1-\varepsilon)d)| \leqslant 1$$

so that

$$P_N(d) \leqslant n.$$

On the other hand, since $d/d(1-\varepsilon) > 1$ we have

$$P_N(d) = T_N(d/(1-\varepsilon)d) \geqslant \frac{1}{2}\left((1-\varepsilon)^{-1} + \sqrt{(1-\varepsilon)^{-2}-1}\right)^N.$$

Combining the two inequalities, we have

$$N \leqslant \frac{\log(2n)}{\log(\frac{1+(1-(1-\varepsilon)^2)^{1/2}}{1-\varepsilon})}.$$

□

EXERCISE 7.1. Prove an analogous result when $\mathcal{G}$ is an ε-expander bipartite graph (hint consider A^2).

7.2.2.2. *Independence number.*

DEFINITION 7.9. *An independent set of a graph $\mathcal{G}$ is a subset of $I \subset V$ with no two adjacent vertices. The independence number $i(\mathcal{G})$ of $\mathcal{G}$ is the largest size of an independent set.*

PROPOSITION 7.10. *If $\mathcal{G}_{n,d}$ is a two sided ε-expander one has*

$$i(G) \leqslant (1-\varepsilon)n$$

PROOF. Exercise. □

7.2.2.3. *Chromatic number.*

DEFINITION 7.11. *The chromatic number $\chi(\mathcal{G})$ of a graph $\mathcal{G}$ is the minimum number of colors needed to color V so that in a set of a given color no two elements are adjacent.*

REMARK 7.7. One has $\chi(\mathcal{G}) \leqslant n$ by coloring any vertice with a different color!

PROPOSITION 7.12. *If $\mathcal{G}_{n,d}$ is a two sided ε-expander one has*

$$\chi(G) \geqslant (1-\varepsilon)^{-1}.$$

PROOF. Exercise. □

REMARK 7.8. One can show that

$$\liminf_{n\to\infty} \max_{\mathcal{G}_{n,d}}(|\lambda_1|, |\lambda_{n-1}|) \geqslant 2\sqrt{d-1}.$$

A graph achieving $\max(|\lambda_1|, |\lambda_{n-1}|) = 2\sqrt{d-1}$ is called a Ramanujan graph. Such graphs (which are optimal expanders) exist and are called Ramanujan graphs. The theory of modular form furnished examples of Ramanujan graphs (see [6]).

7.3. Expansion in Cayley graphs

We assume now that $\mathcal{G} = (G, S)$ is a (left) Cayley graph for a generating set $S \subset G$ (in particular the graph is connected); let $n = |G|$ and $d = |S|$ and

$$G = \{g_1 = e_G, \cdots, g_n\}.$$

We can rewrite M as a convolution operator

$$Mf = \mu \star f : g \mapsto \sum_{g_1 g_2 = g} \mu(g_1) f(g_2)$$

where

$$\mu = \mu_S = \frac{1}{S}\sum_{s\in S}\delta_s.$$

Indeed since S is symmetric

$$\mu \star f(g) = \sum_{g_1g_2=g}\frac{1}{|S|}\sum_{s\in S}\delta_{g_1=s}f(g_2) = \frac{1}{|S|}\sum_{s\in S}f(s^{-1}g) = \frac{1}{|S|}\sum_{s\in S}f(sg).$$

Likewise

$$M^\ell f = \mu^{(\ell)} \star f$$

where $\mu^{(\ell)}$ is the ℓ-times self convolution of μ

$$\mu^{(\ell)} = \frac{1}{|S|^\ell}\sum_{(s_1,s_2\cdots,s_\ell)\in S^\ell}\delta_{s_1.s_2\cdots.s_\ell} = \mu\star\cdots\star\mu \ (\ell \text{ times}).$$

Our aim is to discuss Cayley graphs which are expanders; in particular we hope to explain the proof of the following:

THEOREM 7.13 (Bourgain-Gamburd). *Let $S = \{s_1,\cdots,s_d\} \subset \mathrm{SL}_2(\mathbb{Z}) - \{\mathrm{Id}_2\}$ be a finite symmetric set of d elements not containing the identity; for any prime p let*

$$S_p = \{s_1 \,(\mathrm{mod}\, p),\cdots,s_d \,(\mathrm{mod}\, p)\} \subset \mathrm{SL}_2(\mathbb{F}_p)$$

be the set of reductions of elements of S modulo p (for p large enough, S_p has d elements and does not contain Id_2). There exists $\varepsilon = \varepsilon(|S|) > 0$ such that for p sufficiently large and such that S_p generates $\mathrm{SL}_2(\mathbb{F}_p)$, the Cayley graph $(\mathrm{SL}_2(\mathbb{F}_p), S_p)$ is a two sided ε-expander.

EXAMPLE 7.1. The following sets have the property that their reduction modulo p generate $\mathrm{SL}_2(\mathbb{F}_p)$ for p large enough: for $k \geqslant 1$ let

$$S(k) = \{\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}^{-1}, \begin{pmatrix}1 & 0\\ k & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ k & 1\end{pmatrix}^{-1}\}$$

(with $\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}^{-1} = \begin{pmatrix}1 & -k\\ 0 & 1\end{pmatrix}$ and $\begin{pmatrix}1 & 0\\ k & 1\end{pmatrix}^{-1} = \begin{pmatrix}1 & 0\\ -k & 1\end{pmatrix}$). Then $S(k) \,(\mathrm{mod}\, p)$ generates $\mathrm{SL}_2(\mathbb{F}_p)$ for $p > k$ (so that $k \,(\mathrm{mod}\, p)$ is invertible).

Globally the set $S(1)$ generates $\mathrm{SL}_2(\mathbb{Z})$ and $S(2)$ generates the finite index (congruence) subgroup

$$\Gamma(2) = \{\begin{pmatrix}a & b\\ c & d\end{pmatrix} \equiv \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \,(\mathrm{mod}\, 2)\} \subset \mathrm{SL}_2(\mathbb{Z}).$$

For these sets the expansion property can also be deduced from the theory of modular forms.

One the other hand for $k \geqslant 3$ the set $S(k)$ generate a subgroup $L_k \subset \mathrm{SL}_2(\mathbb{Z})$ (the Lubotsky group) of infinite index and which is a free group of rank 2. Moreover for $p \not| k$ $L_k \,(\mathrm{mod}\, p)$ generates $\mathrm{SL}_2(\mathbb{F}_p)$.

REMARK 7.9. A group F_r is said to be free of rank r if F is generated by a symmetric set of $2r$ (distinct) elements

$$S(r) = \{s_1,\cdots,s_r,s_1^{-1},\cdots,s_r^{-1}\}$$

such that any $g \in F_r - \{e\}$ can be written uniquely in a *reduced form* in the elements of $S(r)$:

$$g = w_1.\cdots.w_\ell \text{ for some } \ell \geqslant 1 \text{ and } w_i \in S(r),\ w_{i+1} \neq w_i^{-1}.$$

7.3.1. A expansion criterion for Cayley graphs. To produce expanding Cayley graphs we have to study the spectral properties of the convolution operator

$$f \in L^2(G) \mapsto \mu \star f \in L^2(G).$$

This will depend on the methods discussed in Appendix 7.5 on representations of G Let

$$d_G := \min_{r \in \mathrm{Irr}(G),\ r \neq 1} \dim(r) \tag{7.4}$$

the minimal dimension of a non-trivial irreducible representation of G.

PROPOSITION 7.14 (Bourgain-Gamburd). *For any $\rho \in \mathrm{Spec}(\mu) - \{1\}$ and any $\ell \geqslant 1$ we have*

$$|\rho| \leqslant (\frac{|G|}{d_G} \|\mu_S^{(l)}\|_2^2)^{1/2\ell}.$$

PROOF. Since $\{\delta_{g_i},\ i \leqslant n\}$ is an OBN of $L^2(G)$ we have for any $\ell \geqslant 1$

$$\begin{aligned} \mathrm{tr}(M^\ell) &= \sum_{i=0}^{n-1} \rho_i^\ell = \sum_{i=1}^{n} \langle \mu^{(\ell)} \delta_{g_i}, \delta_{g_i} \rangle \\ &= \sum_{i=1}^{n} \frac{|\text{paths of length } \ell \text{ starting and arriving at } g_i|}{d^\ell} \end{aligned}$$

Observe that by translating by g_i^{-1} on the *right* we have

$$|\{\text{paths of length } \ell \text{ starting and arriving at } g_i\}|$$

is equal to

$$|\{\text{paths of length } \ell \text{ starting and arriving at } e_G|\} = |\{(s_1, \cdots, s_\ell) \in S^\ell,\ s_1 s_2 . \cdots s_\ell = e_G\}|$$

and therefore

$$\mathrm{tr}(M^\ell) = |G| \frac{|\{(s_1, \cdots, s_\ell) \in S^\ell,\ s_1 s_2 . \cdots s_\ell = e_G\}|}{d^\ell}.$$

In particular we have

$$\mathrm{tr}(M^{2\ell}) = |G| \frac{|\{(s_1, \cdots, s_\ell),\ (s_1, \cdots, s_\ell) \in S^\ell,\ s_1 s_2 . \cdots s_\ell = s'_1 s'_2 . \cdots s'_\ell\}|}{d^{2\ell}}$$

which we can rewrite

$$\sum_{i=0}^{n-1} \rho_i^{2\ell} = |G| \langle \mu^{(l)}, \mu^{(l)} \rangle.$$

Since all the terms in the sum on the left are non-negative, we obtain the upper bound

$$|\rho_i| \leqslant (|G| \langle \mu^{(l)}, \mu^{(l)} \rangle)^{1/2\ell}.$$

However one can do better. Given $\rho \in \mathrm{Spec}(M) - \{1\}$, let $L^2(G)_\rho$ be the corresponding eigenspace. Since $M = \mu \star \bullet$ is a convolution operator defined by left multiplication, the eigenspace $L^2(G)_\rho$ is invariant under the right multiplication action of G

$$r_g f : h \mapsto f(hg)$$

and is therefore is a (sub)representation of the right regular representation of G. Moreover this representation does not contain the trivial representation (which is the representation spanned by the constant functions). It follows that

$$\dim L^2(G)_\rho \geqslant d_G.$$

We have therefore

$$d_G.\rho^{2\ell} \leqslant \dim L^2(G)_\rho.\rho^{2\ell} \leqslant \mathrm{tr}(M^{2\ell}) = |G|\langle \mu^{(l)}, \mu^{(l)}\rangle.$$

□

Corollary 7.15 (Bourgain-Gamburd expansion criterion). *Assume that there exists $C \geqslant 1$, and $\alpha, \beta > 0$ satisfying $0 < \beta < \alpha/2 \leqslant 1/2$ and such that the following hold*

(1) $d_G \geqslant |G|^\alpha$,
(2) *There exists $1 \leqslant \ell \leqslant C \log |G|$ such that*

$$\|\mu_S^{(\ell)}\|_2 \leqslant C|G|^{-1/2+\beta}.$$

Then for $|G|$ sufficiently large (depending on C, α, β) there exists $\varepsilon = \varepsilon(C, \alpha, \beta) > 0$ (not depending on $|G|$) such that (G, S) is a two sided ε-expander.

Proof. We have for any $\rho \neq 1$

$$\begin{aligned}|\rho| &\leqslant (\frac{|G|}{d_G}\|\mu_S^{(l)}\|_2^2)^{1/2\ell} \leqslant \exp(\frac{\log|G|}{2\ell}(1-\alpha-1+2\beta+2\frac{\log C}{\log|G|})) \\ &= \exp(\frac{-\log|G|}{2\ell}(\alpha-2\beta-2\frac{\log C}{\log|G|})) \leqslant \exp(-\frac{\alpha-2\beta'}{2C}) := 1-\varepsilon\end{aligned}$$

for (say) $\beta' = \frac{\beta+\alpha/2}{2} < \alpha/2$ and as long as $|G|$ is sufficiently large in term of C and $\alpha/2 - \beta$ □

Remark 7.10. One can see quickly that graph (G, S) is not bipartite: if (G, S) were bipartite the -1-eigenspace $L^2(G)_{-1}$ would be one dimensional but is also a non-trivial representation of G so of dimension $\geqslant d_G > 1$.

Remark 7.11. Note that the function

$$\ell \mapsto \|\mu_S^{(\ell)}\|_2$$

is a decreasing functions. Indeed

$$\|\mu_S^{(\ell+1)}\|_2 = \|\mu \star \mu_S^{(\ell)}\|_2 \leqslant \|\mu \star \bullet\|\|\mu_S^{(\ell)}\|_2$$

where $\|\mu \star \bullet\|$ is the operator norm which is $\leqslant 1$ (since $\mathrm{Spec}(M) \subset [-1, 1]$). Moreover for any probability measure ν we have by CS

$$1 = \sum_g \nu(g) \leqslant \mathrm{Supp}(\nu)^{1/2}\|\nu\|_2$$

or

$$\mathrm{Supp}(\nu) \geqslant 1/\|\nu\|_2^2 \tag{7.5}$$

so the fact the $\|\mu_S^{(\ell)}\|_2$ converge to 0 shows that $\mu_S^{(\ell)}$ is "spreading" through G which is in line with the equidistribution property for expanders.

7.4. The Bourgain-Gamburd expansion machine

The proof of Theorem 7.13 depends on a general expansion criterion called the "Bourgain-Gamburd expansion machine" whose properties can be verified for the Cayley graphs

$$(\mathrm{SL}_2(\mathbb{F}_p), S \,(\mathrm{mod}\, p)).$$

THEOREM 7.16 (Bourgain-Gamburd). *Let G be a finite group and S a symetric set of d generators and let*

$$\mu = \mu_S = \frac{1}{|S|}\sum_{s\in S}\delta_s.$$

Suppose there exists constants $0 < \alpha < 1 < \Lambda$ such that

1. *(Representation gap) $d_G \geqslant |G|^\alpha$.*
2. *(Product Theorem) For any $\delta > 0$ there exists $\delta' > 0$ such that setting $K = |G|^{\delta'}$, any K-approximate subgroup $H \subset G$ satisfying*
$$|G|^\delta \leqslant |H| \leqslant |G|^{1-\delta}$$
generates a proper subgroup of G.
3. *(Non-concentration along proper subgroups) There exists $\ell \leqslant \frac{1}{2}\Lambda \log|G|$ such that for any proper subgroup $H \subsetneq G$*
$$\mu^{(2\ell)}(H) = \frac{|\{(s_1.\cdots s_{2\ell}) \in S^{2\ell},\ s_1.\cdots .s_{2\ell} \in H\}|}{d^{2\ell}} < |G|^{-\alpha}.$$

There exists $\varepsilon = \varepsilon(d,\alpha,\Lambda) > 0$ such that (G,S) is a two sided ε-expander.

REMARK 7.12. For $G = \mathrm{SL}_2(\mathbb{F}_p)$, (1) is Frobenius Theorem 6.4 while (2) is a consequence of Helfgott product Theorem 6.2. It "remains" to discuss the verification of (3) for $\mathrm{SL}_2(\mathbb{F}_p)$ and the proof of the "Expansion Machine" Theorem 7.16.

7.4.1. A weighted BSG Lemma. To be able to exploit Condition (3) we will need a new version of the Balog-Szemeredi-Gowers lemma; here is a direct consequence of the Approximate subgroup version of the BSG lemma Theorem 4.3 (take $A = A^{-1} = B$)

THEOREM 7.17 (Balog-Szemeredi-Gowers, approximate group version). *There exists $C \geqslant 1$ such that for any $K \geqslant 2$, any finite group G and any $A = A^{-1} \subset G$ finite symmetric set whose self-normalized energy satisfies*

$$e(A,A) \geqslant 1/K.$$

There exist $g \in G$, a K^C-approximate subgroup H, such that

$$|H| \leqslant K^C|A|,\ K^{-C}|A| \leqslant |A \cap gH|. \tag{7.6}$$

Remember that the energy can be written in terms of a convolution

$$e(A,A) = \frac{\|1_A \star 1_A\|_2^2}{|A|^3} = \frac{\sum_g |1_A \star 1_A(g)|^2}{|A|^3}.$$

The following generalization (called the *BSG weighted lemma*) recovers Theorem 7.17 when the measure probbility measure ν below is the uniform measure

$$\nu_A = \frac{1_A}{|A|}.$$

LEMMA 7.18 (BSG weighted lemma). *There exists an absolute constant C such that for any $K \geqslant 2$, any group G and $\nu : G \to \mathbb{R}_{\geqslant 0}$ any finitely supported probability measure on G ($\sum_{g\in G} \nu(g) = 1$) which is moreover symmetric ($\forall g \in G,\ \nu(g) = \nu(g^{-1})$) which satisfies*

$$\|\nu \star \nu\|_2 \geqslant \|\nu\|_2/K,$$

then there exists a K^C-approximate subgroup $H \subset G$ and $g \in G$ such that

$$|H| \leqslant K^C/\|\nu\|_2^2,\ \nu(gH) \geqslant K^{-C}.$$

For the proof it will be useful to recall a special case of Young's convolution inequality:

LEMMA 7.19 (Young's convolution inequality (for $(2, 1, 2)$)). *Let G be a group and $\mu, \nu : G \mapsto \mathbb{C}$ be finitely supported functions.*

We have

$$\|\mu \star \nu\|_2 \leqslant \|\mu\|_2\|\nu\|_1$$

where

$$\|\nu\|_1 = \sum_{g\in G} |\nu(g)|.$$

REMARK 7.13. Using Hölder inequality one can obtain a family of Young inequalities: Given $p, q, r \in [1, \infty)$ such that

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$$

one has

$$\|\mu \star \nu\|_r \leqslant \|\mu\|_p\|\nu\|_q$$

where

$$\|\mu\|_p = (\sum_{g\in G} |\mu(g)|^p)^{1/p}$$

and $\|\mu \star \nu\|_r,\ \|\nu\|_q$ are defined similarly.

PROOF. Taking absolute values, we may assume that μ, ν are non-negative.

We have

$$\begin{aligned}
\|\mu \star \nu\|_2^2 &= \sum_{g\in G}(\sum_{g_1g_2=g} \mu(g_1)\nu(g_2))(\mu \star \nu)(g) = \sum_{g_1,g\in G} \mu(g_1)\nu(gg_1^{-1})(\mu \star \nu)(g) \\
&= \sum_{g_1,g\in G} \mu(g_1)\nu(gg_1^{-1})^{1/2}(\mu \star \nu)(g)\nu(gg_1^{-1})^{1/2} \\
&\leqslant (\sum_{g_1,g\in G} \mu(g_1)^2\nu(gg_1^{-1}))^{1/2}(\sum_{g_1,g\in G} (\mu \star \nu)^2(g)\nu(gg_1^{-1}))^{1/2} \\
&= \|\mu\|_2\|\nu\|_1^{1/2}\|\mu \star \nu\|_2\|\nu\|_1^{1/2} = \|\mu\|_2\|\nu\|_1\|\mu \star \nu\|_2
\end{aligned}$$

□

We can now prove the the BSG weighted Lemma.

PROOF. Let

$$W = \text{width}(\nu) := 1/\|\nu\|_2^2$$

(recall that $\text{Supp}(\nu) \geqslant W$ and for $A \subset G$ a finite set and $\nu = \nu_A$ we have $W = |A|$) and let

$$\nu = \nu_{\ll} + \nu_{\asymp} + \nu_{\gg}$$

where

$$\nu_{\ll} = \nu.1_{\nu<1/100K^2W},\ \nu_{\gg} = \nu.1_{\nu>100K/W},\ \nu_{\asymp} = \nu.1_{1/100K^2W\leqslant\nu\leqslant 100K/W}.$$

We have

$$\|\nu_{\ll}\|_2^2 = \sum_{\nu(g)<1/100K^2W} |\nu(g)|^2 \leqslant (100K^2W)^{-1}\sum_g \nu(g) = (100K^2W)^{-1}.$$

By Young's inequality we have

$$\|\nu_{\ll}\star\nu\|_2 \leqslant \|\nu_{\ll}\|_2\|\nu\|_1 \leqslant (10KW)^{-1/2}.$$

By symmetry of ν we also have

$$\|\nu\star\nu_{\ll}\|_2 \leqslant (10KW)^{-1/2}.$$

By CS we have

$$\|\nu_{\gg}\|_1 \leqslant \frac{W}{10K}\|\nu\|_2^2 = 1/10K$$

and therefore we have

$$\|\nu_{\gg}\star\nu\|_2 = \|\nu\star\nu_{\gg}\|_2 \leqslant \|\nu_{\gg}\|_1\|\nu\|_2 \leqslant (10KW^{1/2})^{-1}.$$

By assumption we have $\|\nu\star\nu\|_2 \geqslant 1/KW^{1/2}$ and from this and the previous estimates, we conclude that

$$\|\nu_{\asymp}\star\nu_{\asymp}\|_2 \gg (KW^{1/2})^{-1}.$$

Let

$$A := \{g\in G,\ \nu(g) \geqslant (100K^2W)^{-1}\}.$$

We have by Young's inequality

$$\|1_A\star 1_A\|_2 \leqslant \|1_A\|_2\|1_A\|_1 = |A|^{3/2}$$

and

$$\|1_A\star 1_A\|_2 \geqslant (100K/W)^{-2}\|\nu_{\asymp}\star\nu_{\asymp}\|_2 \gg K^{-5}W^{3/2}$$

and in particular

$$|A| \gg K^{-4}W.$$

On the other hand we have

$$1 \geqslant \nu(A) \geqslant |A|/100K^2W$$

and we conclude that

$$K^{-4}W \ll |A| \leqslant 100K^2W.$$

The upper bound implies that

$$\|1_A\star 1_A\|_2 \gg K^{-8}|A|^{3/2}.$$

By the BSG theorem there exists $C\geqslant 1$ and a K^C-approximate subgroup H and $g\in G$ such that

$$|H| \leqslant K^C|A| \text{ and } K^{-C}|A| \ll |A\cap gH|$$

and therefore

$$|H| \ll K^{C+2}W = K^{C+2}/\|\nu\|_2^2$$

and

$$\nu(gH) \geqslant \nu(gH\cap A) \geqslant (100K^2W)^{-1}|gH\cap A| \gg K^{-C-2}|A|/W \gg K^{-(C+10)}.$$

□

7.4.2. Proof of Theorem 7.16. In this section we assume that the assumptions (1), (2), (3) of Theorem 7.16 hold.

We have to show that there exists $C \geqslant 1$ and $1 \leqslant \ell \leqslant C \log |G|$ such that

$$\|\mu_S^{(\ell)}\|_2 \leqslant |G|^{-1/2+\beta} \tag{7.7}$$

for some $\beta < \alpha/2$.

The next Lemma shows that the non-concentration inequalities (3) a priori valid for a single $\ell = O(\log |G|)$ in fact hold for a whole range of large ℓ's:

PROPOSITION 7.20. *For any $\ell \geqslant \frac{1}{2}\Lambda \log |G|$, any proper subgroup $H \subset G$ and any $g \in G$ we have*

$$\sup_{g \in G} \mu^{(\ell)}(gH) \leqslant |G|^{-\alpha/2}. \tag{7.8}$$

We also have

$$\|\mu^{(\ell)}\|_\infty = \sup_{g \in G} \mu^{(\ell)}(g) \leqslant |G|^{-\alpha/2} \tag{7.9}$$

and

$$\|\mu^{(\ell)}\|_2 \leqslant \|\mu^{(\ell)}\|_\infty \|\mu^{(\ell)}\|_1 \leqslant |G|^{-\alpha/2}. \tag{7.10}$$

PROOF. By Assumption (3), there exists $\ell_0 \leqslant \frac{1}{2}\Lambda \log |G|$ such that for any proper subgroup $H \subset G$ we have

$$\mu^{(2\ell_0)}(H) \leqslant |G|^{-\alpha}.$$

By positivity and symmetry, we have for any $g \in G$,

$$\mu^{(2\ell_0)}(H) = \mu^{(\ell_0)} \star \mu^{(\ell_0)}(Hgg^{-1}H) \geqslant \mu^{(\ell_0)}(Hg)\mu^{(\ell_0)}(g^{-1}H) = \mu^{(\ell_0)}(Hg)^2$$

and therefore, for any proper subgroup H we have

$$\sup_{g \in G} \mu^{(\ell_0)}(Hg) \leqslant |G|^{-\alpha/2}$$

as and

$$\sup_{g \in G} \mu^{(\ell)}(gH) \leqslant |G|^{-\alpha/2}$$

by writing $gH = gHg^{-1}g = H'g$.

Given $\ell \geqslant \frac{1}{2}\Lambda \log |G|$, for any $g \in G$ we have

$$\mu^{(\ell)}(gH) = \mu^{(\ell-\ell_0)} \star \mu^{\ell_0)}(gH) \leqslant |G|^{-\alpha/2}$$

by averaging the previous upper bound over the various products of $\ell - \ell_0$ elements of S: this gives (7.8).

Taking $H = \{e_G\}$ we obtain that

$$\|\mu^{(\ell)}\|_\infty = \sup_{g \in G} \mu^{(\ell)}(g) \leqslant |G|^{-\alpha/2}$$

so that

$$\|\mu^{(\ell)}\|_2 \leqslant \|\mu^{(\ell)}\|_\infty \|\mu^{(\ell)}\|_1 = \|\mu^{(\ell)}\|_\infty \leqslant |G|^{-\alpha/2}.$$

□

LEMMA 7.21 (Flattening lemma). *Let* $\ell \geqslant \frac{1}{2}\Lambda \log |G|$ *such that*

$$\|\mu^{(\ell)}\|_2 \geqslant |G|^{-1/2+\alpha/2}$$

then there exists $\eta = \eta(\alpha) > 0$ *such that*

$$\|\mu^{(2\ell)}\|_2 = \|\mu^{(\ell)} \star \mu^{(\ell)}\|_2 \leqslant |G|^{-\eta}\|\mu^{(\ell)}\|_2.$$

PROOF. Suppose instead that

$$\|\mu^{(\ell)} \star \mu^{(\ell)}\|_2 \geqslant |G|^{-\eta}\|\mu^{(\ell)}\|_2$$

for some $\eta > 0$ (to be chosen later). Set

$$K = |G|^{\eta}.$$

By the weighted BSG lemma, there exists $C \geqslant 1$, a K^C-approximate subgroup $H \subset G$ and $g \in G$ such that

$$|H| \leqslant K^C/\|\mu^{(\ell)}\|_2^2 = |G|^{C\eta}/\|\mu^{(\ell)}\|_2^2 \leqslant |G|^{1-\alpha+C\eta}$$

and

$$\mu^{(\ell)}(gH) \geqslant K^{-C} = |G|^{-C\eta}.$$

By (7.9) we have

$$\|\mu^{(\ell)}\|_\infty \leqslant |G|^{-\alpha/2}$$

so that

$$|H| = |gH| \geqslant |G|^{\alpha/2-C\eta}$$

and

$$|G|^{\alpha/2-C\eta} \leqslant |H| \leqslant |G|^{1-\alpha+C\eta}.$$

Let

$$\delta = \alpha/4$$

and let $\delta' > 0$ be the exponent occurring in Assumption (2).

Let $\eta = \eta(\alpha, C) > 0$ be small enough such that

$$\alpha/4 < \alpha/2 - C\eta < 1 - \alpha + C\eta < 1 - \delta,\ 0 < C\eta \leqslant \min(\delta', \alpha);$$

By the Product Theorem assumption (2), H generates a proper subgroup H' of G and we have

$$\mu^{(\ell)}(gH') \geqslant \mu^{(\ell)}(gH) \geqslant |G|^{-C\eta} \geqslant |G|^{-\alpha}$$

but this contradicts Assumption (3). □

Let us conclude the proof of Theorem 7.16: let

$$\ell_0 = [\frac{1}{2}\Lambda \log |G|] + 1$$

and let

$$\|\mu^{(\ell_0)}\|_2 = |G|^{-\gamma_0};$$

notice that have

$$\gamma_0 \geqslant \alpha/2.$$

If $\gamma_0 \leqslant 1/2 - \alpha/2$, we can apply the Flattening Lemma and obtain

$$\|\mu^{(2\ell_0)}\|_2 \leqslant |G|^{-\gamma_0-\eta}.$$

If $\gamma_0 + \eta \leqslant 1/2 - \alpha/2$ we keep applying the Lemma until we can't that is until $\gamma_0 + k\eta > 1/2 - \alpha/2$. We need to do this at most

$$k \leqslant (1/2 - \alpha/2 - \gamma_0)/\eta = O(1/\eta)$$

times and we then obtain

$$\|\mu^{(2^k \ell_0)}\|_2 \leqslant |G|^{-1/2+\beta}$$

for some $\beta < \alpha/2$. □

7.5. Implementing the Bourgain-Gamburg expansion machine

We want now to describe the proof of Theorem 7.13.

As already explained Assumption (1) and (2) of Theorem 7.16 follow from Frobenius Theorem and Helfgott product Theorem.

It remains to verify Assumption (3).

For this we will use the assumption that the generating set is of the form

$$S_p = S \,(\mathrm{mod}\, p)$$

for

$$S = \{s_1, \cdots, s_d\} \subset \mathrm{SL}_2(\mathbb{Z}) - \{\mathrm{Id}_2\}$$

a fixed set.

One (ie. Bourgain-Gamburd) use this property in the following way: given $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ define

$$\|g\|_\infty = \max(|a|, |b|, |c|, |d|).$$

LEMMA 7.22. *Given $g, g' \in \mathrm{SL}_2(\mathbb{Z})$ such that $\|g\|_\infty, \|g'\|_\infty < p/2$ then*

$$g \equiv g' \,(\mathrm{mod}\, p) \Longleftrightarrow g = g'.$$

PROOF. Indeed if $g \equiv g' \,(\mathrm{mod}\, p)$ we have

$$a - a' \equiv b - b' \equiv c - c' \equiv d - d' \equiv 0 \,(\mathrm{mod}\, p)$$

but since

$$\|g - g'\|_\infty \leqslant \|g\|_\infty + \|g'\|_\infty < p$$

we have

$$|a - a'|,\ |b - b'|,\ |c - c'|,\ |d - d'| < p$$

which implies (since the only integer $< p$ and divisible by p is 0)

$$a - a' = b - b' = c - c' = d - d' = 0.$$

□

Let us pursue this discussion for the special (but representative case) of

$$S(k) = \{\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}^{-1}, \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}^{-1}\}.$$

As was explained before (Exercise) for $k \geqslant 3$, $S(k)$ generates a free group in $\mathrm{SL}_2(\mathbb{Z})$ of rank 2 which implies that for any $(s_1, \cdots, s_\ell) \in S(k)^\ell$

$$s_1 . \cdots . s_\ell = \mathrm{Id}_2 \Longleftrightarrow \ell = 2\ell', \forall i \geqslant 1,\ s_i = s_{2\ell'-i+1}^{-1}.$$

For $g \in \mathrm{SL}_2(\mathbb{Z})$ let

$$\|g\|_o = \max_{\mathbf{x} \neq (0,0)} \frac{\|g.\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$$

be the operator of g acting on $\mathbb{R}^2$. Recall that this norm is semi-multiplicative

$$\|g_1 g_2\|_o \leqslant \|g_1\|_o \|g_2\|_o$$

and satisfies

$$\|g\|_\infty \leqslant \|g\|_0;$$

indeed

$$\max(|a|,|c|) \leqslant (|a|^2+|c|^2)^{1/2} = \|(a,c)\| = \|g.(1,0)\|_2^2 \leqslant \|g\|_0.1$$

and

$$\max(|b|,|d|) \leqslant (|b|^2+|d|^2)^{1/2} = \|(b,d)\| = \|g.(0,1)\|_2^2 \leqslant \|g\|_0.1.$$

This implies the following

LEMMA 7.23. *Let*

$$S_{max} = \max_{s\in S(k)} \|s\|_o.$$

For $\ell < \frac{1}{2}\log(\frac{p}{2S_{max}})$ *the map*

$$(s_1,\cdots,s_\ell) \in S(k)^\ell \mapsto s_1.s_2.\cdots.s_\ell \,(\mathrm{mod}\, p) \in \mathrm{SL}_2(\mathbb{F}_p)$$

is injective.

PROOF. Suppose that

$$s_1.s_2.\cdots.s_\ell = s_1'.s_2'.\cdots.s_\ell' \,(\mathrm{mod}\, p)$$

we have

$$s_1.s_2.\cdots.s_\ell.{s'}_\ell^{-1}.\cdots.{s'}_1^{-1} \equiv \mathrm{Id}_2 \,(\mathrm{mod}\, p)$$

but

$$\|s_1.s_2.\cdots.s_\ell.{s'}_\ell^{-1}.\cdots.{s'}_1^{-1}\|_\infty \leqslant \|s_1.s_2.\cdots.s_\ell.{s'}_\ell^{-1}.\cdots.{s'}_1^{-1}\|_0 \leqslant S_{max}^{2\ell} < p/2$$

From the previous lemma we conclude that

$$s_1.s_2.\cdots.s_\ell.{s'}_\ell^{-1}.\cdots.{s'}_1^{-1} = \mathrm{Id}_2$$

and therefore $s_i = s_i'$, $i \leqslant \ell$ since we are in a free group. □

COROLLARY 7.24. *Suppose that*

$$\ell < \frac{1}{2}\log(\frac{p}{2S_{max}}),$$

we have

$$\|\mu^{(\ell)}\|_2 \leqslant (d^\ell)^{-1/2}$$

PROOF. We have

$$\|\mu^{(\ell)}\|_2^2 = \frac{1}{d^{2\ell}} \sum_{s_1.s_2.\cdots.s_\ell = s_1'.s_2'.\cdots.s_\ell' \,(\mathrm{mod}\, p)} \sum 1 = \frac{1}{d^{2\ell}} \sum_{s_1.s_2.\cdots.s_\ell = s_1'.s_2'.\cdots.s_\ell'} \sum 1 = \frac{1}{d^\ell}$$

since in a free group

$$s_1.s_2.\cdots.s_\ell = s_1'.s_2'.\cdots.s_\ell'$$

implies that

$$s_1 = s_1', \cdots, s_\ell = s_\ell'.$$

□

In particular since $|\mathrm{SL}_2(\mathbb{F}_p)| \asymp p^3$ we see that for any ℓ satisfying

$$\log(|\mathrm{SL}_2(\mathbb{F}_p)|) \ll \ell \ll_{S_{max}} \log(|\mathrm{SL}_2(\mathbb{F}_p)|)$$

we have

$$\|\mu^{(\ell)}\|_\infty^2 \leqslant \|\mu^{(\ell)}\|_2^2 \leqslant |\mathrm{SL}_2(\mathbb{F}_p)|^{-\alpha/2} \tag{7.11}$$

for some absolute $0 < \alpha'$.

In particular we have the non-concentration inequality for the trivial subgroup $H = \{e_G\}$ and more generally for H any subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$ of absolutely bounded size.

For the other subgroups, we have the following classification:

THEOREM 7.25 (Dickson). *For $p > 5$. Given $H \subset \mathrm{SL}_2(\mathbb{F}_p)$ a strict subgroup, one of the following holds*

- *$H/\{\pm \mathrm{Id}_2\} = \mathfrak{A}_4, \mathfrak{S}_4$ or $\mathfrak{A}_5$*
- *H is a subgroup of a dihedral subgroup of order $2\frac{(p\pm 1)}{2}$.*
- *H is a subgroup of a Borel subgroup of order $\frac{p(p-1)}{2}$.*

COROLLARY 7.26. *For $p > 5$. For any $H \subset \mathrm{SL}_2(\mathbb{F}_p)$ a strict subgroup with $|H| > 120$ we have for any $g_1, g_2, g_3, g_4 \in H$ we have*

$$[[g_1, g_2], [g_3, g_4]] = 1.$$

We then have for any $\ell \ll \log(|\mathrm{SL}_2(\mathbb{F}_p)|)$

$$\mu^{(\ell)}(H) = \frac{1}{d^\ell} \sum_{w_\ell \,(\mathrm{mod}\, p) \in H} 1$$

where $w_\ell = s_1.s_2.\cdots.s_\ell$ range over all the words of ℓ letters in the alphabet $S(k)$. We have

$$\mu^{(\ell)}(H)^4 = \frac{1}{d^{4\ell}} \sum_{\substack{i=1,2,3,4 \\ w_{\ell,i} \in H}} 1 \leqslant \frac{1}{d^{4\ell}} \sum_{\substack{w_{\ell,i},\ i=1,2,3,4 \\ [[w_{\ell,1}, w_{\ell,2}], [w_{\ell,3}, w_{\ell,4}]] \equiv \mathrm{Id}_2 \,(\mathrm{mod}\, p)}} 1.$$

The commutator $[[w_{\ell,1}, w_{\ell,2}], [w_{\ell,3}, w_{\ell,4}]]$ is a word in the alphabet $S(k)$ of length 16ℓ so if

$$\ell < \frac{1}{32} \log(\frac{p}{2S_{max}})$$

we have

$$[[w_{\ell,1}, w_{\ell,2}], [w_{\ell,3}, w_{\ell,4}]] \equiv \mathrm{Id}_2 \,(\mathrm{mod}\, p) \Longleftrightarrow [[w_{\ell,1}, w_{\ell,2}], [w_{\ell,3}, w_{\ell,4}]] = \mathrm{Id}_2.$$

We have the following result (see [2, Prop. 8])

PROPOSITION 7.27. *In the free group of rank 2 the number of quadruples $(w_{\ell,1}, w_{\ell,2}, w_{\ell,3}, w_{\ell,4})$ of words of length 2ℓ satisfying*

$$[[w_{\ell,1}, w_{\ell,2}], [w_{\ell,3}, w_{\ell,4}]] = 1$$

is bounded by $\ll \ell^6$.

It follows that for $\ell < \frac{1}{32} \log(\frac{p}{2S_{max}})$ one has

$$\mu^{(\ell)}(H)^4 \ll \frac{\ell^6}{d^{4\ell}}$$

or equivalently

$$\mu^{(\ell)}(H) \ll \frac{\ell^{3/2}}{d^\ell}.$$

This establishes the non-concentration inequality for strict subgroups of $\mathrm{SL}_2(\mathbb{F}_p)$.

Appendix : Harmonic analysis for finite groups

7.6. Representations of a finite group

DEFINITION 7.28. *Let G be a finite group.*

- *A (finite dimensional) representation of G is a group morphism*
$$\pi : G \to \mathrm{GL}(V_\pi)$$
where $V_\pi \neq \{0\}$ is a finite dimensional complex vector space. In other terms a representation is a linear action of G a finite dimensional complex vector space.
The vector space V_π is also called a G-module.
- *A morphism of G-modules (or G-morphism) $\varphi : (\pi, V_\pi) \to (\rho, V_\rho)$ is a linear map $\varphi : V_\pi \to V_\rho$ such that*
$$\varphi \circ \pi = \rho \circ \varphi.$$
We denote by $\mathrm{Hom}_G(\pi, \rho)$ the space of all G-morphisms.
We define injective, surjective, bijective/iso/auto G-morphisms in the evident way.
- *A submodule (sub-representation) $W \subset V_\pi$ is a subspace stable under $\pi(G)$ or equivalently such that the inclusion $\iota : W \subset V_\pi$ is a G-morphism.*
- *A representation (π, V_π) is irreducible iff V_π has no non-trivial G-submodules (no submodules distinct from $\{0\}$ and V_π). Otherwise it is called reducible. We denote by $\mathrm{Irr}(G)$ the set of equivalence classes of irreducible representations.*
- *A representation is unitarizable if there exists an inner product $\langle \bullet, \bullet \rangle_\pi$ on V_π such that $\pi(G) \subset \mathrm{U}(\langle \bullet, \bullet \rangle_\pi)$ (the unitary group of the inner product): for any $g \in G,\ v, v' \in V_\pi$*
$$\langle \rho(v), \rho(v') \rangle_\pi = \langle v, v' \rangle_\pi.$$
Such inner product is called a unitary structure for the representation.

THEOREM 7.29. *Any representation (π, V_π) is unitarizable.*

PROOF. Fix any inner product $\langle \bullet, \bullet \rangle$ and define
$$\langle v, v' \rangle_\pi = \frac{1}{|G|} \sum_{g \in G} \langle g.v, g.v' \rangle,\ g.v := \rho(g)(v)$$

□

EXAMPLE 7.2. The *regular representation* is the $|G|$-dimensional representation of $G \times G$ on $\mathcal{F}(G; \mathbb{C})$ via left and right translations:
$$\mathrm{reg}(g, g')(f) : h \mapsto (g, g').f = f(g'^{-1} h g).$$
It restriction to $G \times \{e_G\}$ (resp. $\{e_G\} \times G$) is called the right (resp. left) regular representation and is noted r_G (resp. l_G).

For the regular representation $(\mathrm{reg}, \mathcal{F}(G;\mathbb{C}))$ the unitary structure is the usual inner product

$$\langle f_1, f_2\rangle_G = \frac{1}{|G|}\sum_{g\in G} f_1(g)\overline{f_2}(g) = \frac{1}{|G|}\int_G |f(g)|^2 dg$$

where dg denote the counting measure.

COROLLARY 7.30. *Any representation is completely reducible: ie. decompose into a direct sum of irreducible representations. Moreover this decomposition is unique up to permutation of the factors.*

PROOF. Let $\langle \bullet, \bullet\rangle_\pi$ be a unitary structure. If $W \subset V_\pi$ is a submodule then

$$W^\perp = \{v \in V_\pi, \forall w \in W,\ \langle v, w\rangle_\pi = 0\}$$

is a G-submodule and we have the decomposition into G-modules

$$V_\pi = W \oplus W^\perp$$

and we iterate. Uniqueness can be obtained from the following "Lemma". □

THEOREM 7.31 (Schur's Lemma). *Let $\pi, \rho \in \mathrm{Rep}(G)$.*

If π is irreducible, any G-map

$$\varphi : V_\pi \to V_\rho$$

is zero or injective.

If ρ is irreducible, any G-map

$$\varphi : V_\pi \to V_\rho$$

is either zero or surjective.

If π and ρ are irreducible, on has

$$\dim_{\mathbb{C}} \mathrm{Hom}_G(\pi, \rho) = \delta_{\pi\sim_G\rho}.$$

PROOF. Suppose π is irreducible and let $\varphi \in \mathrm{Hom}_G(V_\pi, V_\rho)$ non-zero then the kernel is G-invariant so the kernel is either $\{0\}$ or V_π but cannot be V_π: ie. $\ker\varphi = \{0\}$. Same reasonning if $\rho \in \mathrm{Irr}(G)$ with the image.

In particular if $\pi, \rho \in \mathrm{Irr}(G)$ and $\pi \not\simeq \rho$ then

$$\mathrm{Hom}_G(V_\pi, V_\rho) = \{0\}.$$

If $\pi \simeq \rho$ we may assume $\pi = \rho$. Given $\varphi \in \mathrm{Hom}_G(V_\pi, V_\pi) - \{0\}$ and $\lambda \in \mathbb{C}$ an eigenvalue of φ with eigenspace V_λ. Since V_λ is a non-zero G-module we must have $V_\lambda = V_\pi$ and $\varphi = \lambda.\mathrm{Id}_\pi$. □

7.7. Matrix coefficients

DEFINITION 7.32. *Given a unitary representation (π, V_π) and $v, w \in V_\pi$; the (v, w) matrix coefficient of π is the function*

$$\Phi_{\pi,v,w}(\bullet) : g \mapsto \langle g.v, w\rangle_\pi.$$

REMARK 7.14. The term matrix coefficient come from the fact that if v, w are unitary and contained in an ONB of V_π, $\langle g.v, w\rangle_\pi$ is the (v, w) coefficient of the matrix representaing $\pi(g) \in \mathrm{GL}(V_\pi)$ in that basis.

Matrix coefficients allow to construct maps between different representations: given any $w \in V_\pi$ the map

$$\Phi_{\pi,\bullet,w} : v \in V_\pi \mapsto \Phi_{\pi,v,w} \in \mathcal{F}(G;\mathbb{C})$$

is linear and satisfies for any v

$$\Phi_{\pi,g.v,w}(g') = \langle g'.g.v, w\rangle_\pi = \Phi_{\pi,v,w}(g'g) = \mathrm{r}_G(g)(\Phi_{\pi,v,w})(g')$$

and so is a G-map relative to the right regular representation. If particular if $\Phi_{\pi,\bullet,w}$ is injective, the right regular representation will "contain" π as a sub-G module.

More generally given $\pi, \rho \in \mathrm{Rep}(G)$ and $w \in V_\pi,\ w' \in V_\rho$ the map

$$\varphi_{w,w'} : v \in V_\pi \mapsto \int_G \Phi_{\pi,v,w}(g)g^{-1}.w dg = \int_G \langle g.v, w\rangle_\pi g^{-1}.w' dg \in V_\rho \tag{7.12}$$

is G-equivariant: for any $h \in G$

$$\begin{aligned}\varphi_{w,w'}(h.v) = \int_G \langle gh.v, w\rangle_\pi g^{-1}.w' dg &= \int_G \langle g'.v, w\rangle_\pi (hg'^{-1}).w' dg \\ &= h.(\int_G \langle g'.v, w\rangle_\pi g'^{-1}.w' dg) = h.\varphi_{w,w'}(v).\end{aligned}$$

Example 7.3. In particular for $\pi \in \mathrm{Irr}(G)$ and $w \in V_\pi - \{0\}$ we have an injective G-map

$$\Phi_{\pi,\bullet,w} : v \in V_\pi \mapsto \Phi_{\pi,v,w} \in \mathcal{F}(G;\mathbb{C})$$

and we can identify V_π with a subrepresentation of (the right regular) representation $\mathcal{F}(G;\mathbb{C})$. We denote the image by

$$V_{\pi,w} \subset \mathcal{F}(G;\mathbb{C}). \tag{7.13}$$

Theorem 7.33 (Orthogonality of matrix coefficients). *Let $\pi, \rho \in \mathrm{Irr}(G)$ non-isomorphic and of dimension $d_\pi,\ d_\rho$ and $v, w \in V_\pi,\ v', w' \in V_\rho$. Let*

$$\Phi_{\pi,v,w},\ \Phi_{\rho,v',w'} \in \mathcal{F}(G;\mathbb{C})$$

be the corresponding matrix coefficients. We have

$$\langle \Phi_{\pi,v,w},\ \Phi_{\rho,v',w'}\rangle_G = 0 \text{ if } \pi \not\simeq \rho$$

and for $\pi = \rho$

$$\langle \Phi_{\pi,v,w},\ \Phi_{\pi,v',w'}\rangle_G = \frac{\overline{\langle w, w'\rangle}_\pi \langle v, v'\rangle_\pi}{d_\pi}.$$

Proof. We have

$$\begin{aligned}\langle \varphi_{w,w'}(v), v'\rangle_\rho &= \int_G \langle g.v, w\rangle_\pi \langle g^{-1}w', v'\rangle_\rho dg \\ &= \int_G \langle g.v, w\rangle_\pi \langle w', gv'\rangle_\rho dg \\ &= \int_G \langle g.v, w\rangle_\pi \overline{\langle gv', w'\rangle}_\rho dg \\ &= |G| \langle \Phi_{\pi,v,w},\ \Phi_{\rho,v',w'}\rangle_G.\end{aligned}$$

So $\langle \Phi_{\pi,v,v'}, \Phi_{\rho,w,w'}\rangle_G \neq 0$ implies that the G-map $\varphi_{w,w'}$ in (7.12) is non zero and π and ρ are isomorphic.

Assume that $\pi = \rho$ and write V for V_π. We have for $v, v', w, w' \in V$

$$\langle \varphi_{w,w'}(v), v' \rangle_\pi = |G| \langle \Phi_{\pi,v,w},\ \Phi_{\rho,v',w'} \rangle_G.$$

By Schur's lemma we have

$$\varphi_{w,w'} = \lambda(w, w') \mathrm{Id}_2$$

for some $\lambda(w, w') \in \mathbb{C}$ and

$$\langle \varphi_{w,w'}(v), v' \rangle_\pi = \lambda(w, w') \langle v, v' \rangle_\pi.$$

Moreover its trace equals

$$\mathrm{tr}(\varphi_{w,w'}) = \lambda(w, w') d_\pi.$$

Let $\{v_1, \cdots, v_d\}$ be an ONB. We have

$$\begin{aligned} \mathrm{tr}(\varphi_{w,w'}) &= \sum_{i=1}^{d} \langle \varphi_{w,w'}(v_i), v_i \rangle_\pi = \int_G \sum_i \langle g.v_i, w \rangle_\pi \langle g^{-1} w', v_i \rangle_\pi dg \\ &= \int_G \sum_i \langle w', g v_i \rangle_\pi \langle g.v_i, w \rangle_\pi dg = \int_G \langle w', w \rangle_\pi dg = |G| \overline{\langle w, w' \rangle}_\pi \end{aligned}$$

Indeed for any $g \in G$, $\{g.v_i,\ i \leqslant d\}$ is an ONB and for any ONB $\{v'_i,\ i \leqslant d\}$, one has

$$\sum_i \langle w', v'_i \rangle_\pi \langle v'_i, w \rangle_\pi = \langle w', w \rangle_\pi.$$

Hence

$$\lambda(w, w') = \frac{\mathrm{tr}(\varphi_{w,w'})}{d_\pi} = \frac{|G|}{d_\pi} \overline{\langle w, w' \rangle}_\pi$$

and

$$|G| \langle \Phi_{\pi,v,w},\ \Phi_{\rho,v',w'} \rangle_G = |G| \frac{\langle v, v' \rangle_\pi \overline{\langle w, w' \rangle}_\pi}{d_\pi}.$$

□

A direct consequence of Theorem 7.33 is the following.

THEOREM 7.34 (Fourier Theory for finite groups). *For any $\pi \in \mathrm{Irr}(G)$ let*

$$\mathcal{B}_\pi = \{v_1, \cdots, v_{d_\pi}\}$$

be an ONB of V_π and for $j \leqslant d$ let

$$V_{\pi,v_j} := \{\Phi_{\pi,v,v_j},\ v \in V_\pi\} \subset \mathcal{F}(G; \mathbb{C})$$

be the G-space defined in (7.13) and let

$$\mathcal{B}_{\pi,v_j} := \{d_\pi^{1/2} \Phi_{\pi,v_i,v_j}\} \subset \mathcal{F}(G; \mathbb{C}). \tag{7.14}$$

The set $\mathcal{B}_{\pi,v_j}$ is an ONB of the space V_{π,v_j} and we have an orthogonal decomposition into irreducible representations

$$\mathcal{F}(G : \mathbb{C}) = \bigoplus_{\pi \in \mathrm{Irr}(G)} \bigoplus_{j \leqslant d_\pi} V_{\pi,v_j}.$$

In other terms the regular representation contains each irreducible representation V_π, $\pi \in \mathrm{Irr}(G)$ with multiplicity d_π and have

$$|G| = \dim_{\mathbb{C}} \mathcal{F}(G : \mathbb{C}) = \sum_{\pi \in \mathrm{Irr}(G)} d_\pi^2.$$

In addition, the set

$$\mathcal{B}_G := \{d_\pi^{1/2}\Phi_{\pi,v_i,v_j},\ \pi \in \mathrm{Irr}(G),\ v_i, v_j \in \mathcal{B}_\pi\}$$

form an ONB of $\mathcal{F}(G:\mathbb{C})$ *and for any* $f \in \mathcal{F}(G:\mathbb{C})$ *we have the Fourier decomposition*

$$f = \sum_{\pi\in\mathrm{Irr}(G)} d_\pi \sum_{i,j\leqslant d_\pi} c_{\pi,i,j}(f)\Phi_{\pi,v_i,v_j}$$

where

$$c_{\pi,i,j}(f) = \langle f, \Phi_{\pi,v_i,v_j}\rangle_G = \frac{1}{|G|}\sum_{g\in G} f(g)\overline{\Phi_{\pi,v_i,v_j}}(g).$$

We also have the Plancherel-Parseval formula

$$\sum_{g\in G}|f(g)|^2 = |G|\langle f,f\rangle_G = \sum_{\pi\in\mathrm{Irr}(G)} d_\pi^2 \sum_{i,j\leqslant d_\pi}\frac{1}{d_\pi}|\langle f,\Phi_{\pi,v_i,v_j}\rangle_G|^2$$
$$= \frac{1}{|G|}\sum_{\pi\in\mathrm{Irr}(G)} d_\pi \sum_{i,j\leqslant d_\pi}|\sum_{g\in G} f(g)\Phi_{\pi,v_i,v_j}(g)|^2.$$

Proof. The only point which is not "mechanical" is the fact that the orthonormal set $\mathcal{B}_G$ generate $\mathcal{F}(G;\mathbb{C})$. If not the space orthogonal to $\mathbb{C}.\langle\mathcal{B}_\mathrm{G}\rangle$ is G-stable hence a subrepresentation of $\mathcal{F}(G;\mathbb{C})$ which contains an irreducible representation π. But any ONB of it gives matrix coefficient which would have to be contained in the space generated by $\mathcal{B}_\mathrm{G}$. □

Remark 7.15. Let $\pi_0 : G \mapsto 1$ be the trivial 1-dimensional representation. Its only matrix coefficient is the constant function 1 and and the contribution to the Fourier decomposition is

$$\frac{1}{|G|}\sum_{g\in G} f(g)$$

and to the Plancherel-Parseval formula is

$$\frac{1}{|G|}|\sum_{g\in G} f(g)|^2.$$

7.7.1. Interpretation in terms of linear maps. Given $f \in \mathcal{F}(G;\mathbb{C})$ and $\pi \in \mathrm{Rep}(G)$ we define the endomorphism $\pi(f) \in \mathrm{End}(V_\pi)$ by

$$\pi(f) : v \in V_\pi \mapsto \sum_{g\in G} f(g)\pi(g)v.$$

For instance for $g \in G$

$$\pi(g) = \pi(1_{\{g\}}).$$

The coefficient (i,j) of the matrix of $\pi(f)$ in the basis $\mathcal{B}_\pi$ are given by

$$\langle\pi(f)v_i, v_j\rangle_\pi = \sum_{g\in G} f(g)\langle g.v_j, v_j\rangle_\pi = \sum_{g\in G} f(g)\Phi_{\pi,v_i,v_j}(g).$$

From this we see that

$$\sum_{i,j\leqslant d_\pi}|\sum_{g\in G} f(g)\Phi_{\pi,v_i,v_j}(g)|^2 = \|\pi(f)\|_{HS}^2$$

where $\|\bullet\|_{HS}$ denote the Hilbert-Schmidt norm on $\mathrm{End}(V_\pi)$, ie.

$$\|X\|_{HS}^2 = \langle X,X\rangle_{HS} = \mathrm{tr}(X.X^*)$$

where X^* is the dual (whose matrix is the conjugate transpose).

The Parseval formula can then be rewritten in a more compact form

$$\sum_{g\in G}|f(g)|^2=\frac{1}{|G|}\sum_{\pi\in\mathrm{Irr}(G)}d_\pi\|\pi(f)\|^2_{HS} \tag{7.15}$$

and more generally

$$\sum_{g\in G}f_1(g)\overline{f}_2(g)=\frac{1}{|G|}\sum_{\pi\in\mathrm{Irr}(G)}d_\pi\langle\pi(f_1),\pi(f_2)\rangle_{HS} \tag{7.16}$$

where

$$\langle X,Y\rangle_{HS}=\mathrm{tr}(XY^*).$$

Likewise the Fourier decomposition formula can be rewritten

$$f(g)=\frac{1}{|G|}\sum_{\pi\in\mathrm{Irr}(G)}d_\pi\langle\pi(f),\pi(g)\rangle_{HS} \tag{7.17}$$

(directly or by taking $f_1=f,\ f_2=1_{\{g\}}$).

Also since $f(g)=g.f(e_G)$ where

$$g.f:h\mapsto f(hg)$$

we have

$$f(g)=\frac{1}{|G|}\sum_{\pi\in\mathrm{Irr}(G)}d_\pi\langle\pi(g.f),\mathrm{Id}_{V_\pi}\rangle_{HS}=\frac{1}{|G|}\sum_{\pi\in\mathrm{Irr}(G)}d_\pi\mathrm{tr}(\pi(g.f)). \tag{7.18}$$

7.7.2. Convolution. The space $\mathcal{F}(G;\mathbb{C})$ is a non-commutative $\mathbb{C}$-algebra relative to the convolution product

$$f_1\star f_2:g\mapsto\sum_{g_1g_2=g}f_1(g_1)f_2(g_2)$$

with unital element $\delta_{\{e_G\}}$ and the map

$$f\in\mathcal{F}(G;\mathbb{C})\mapsto\pi(f)\in\mathrm{End}(V_\pi)$$

is an algebra morphism:

$$\pi(f_1\star f_2)=\sum_{g_1,g_2\in G}f_1(g_1)f_2(g_2)\pi(g_1g_2)=\pi(f_1)\circ\pi(f_2).$$

This algebra is equipped with an (anti-)involution

$$f\mapsto\check{f}:g\mapsto\overline{f}(g^{-1})$$

and (7.18) can be rewritten

$$f_1\star\check{f}_2(e_G)=\frac{1}{|G|}\sum_{\pi\in\mathrm{Irr}(G)}d_\pi\langle\pi(f_1),\pi(f_2)\rangle_{HS}. \tag{7.19}$$

Reference

[1] N. Alon, M. B. Nathanson, and I. Ruzsa, *The polynomial method and restricted sums of congruence classes*, J. Number Theory **56** (1996), no. 2, 404–417, DOI 10.1006/jnth.1996.0029. MR1373563

[2] J. Bourgain and A. Gamburd, *Uniform expansion bounds for Cayley graphs of* $\mathrm{SL}_2(\mathbb{F}_p)$, Ann. of Math. (2) **167** (2008), no. 2, 625–642, DOI 10.4007/annals.2008.167.625. MR2415383

[3] H. A. Helfgott, *Growth and generation in* $\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$, Ann. of Math. (2) **167** (2008), no. 2, 601–623, DOI 10.4007/annals.2008.167.601. MR2415382

[4] H. A. Helfgott, *Growth in linear algebraic groups and permutation groups: towards a unified perspective*, Groups St Andrews 2017 in Birmingham, London Math. Soc. Lecture Note Ser., vol. 455, Cambridge Univ. Press, Cambridge, 2019, pp. 300–345. MR3931419

[5] E. Kowalski, *An introduction to expander graphs*, Cours Spécialisés [Specialized Courses], vol. 26, Société Mathématique de France, Paris, 2019. MR3931316

[6] P. Sarnak, *Some applications of modular forms*, Cambridge Tracts in Mathematics, vol. 99, Cambridge University Press, Cambridge, 1990. MR1102679

[7] J.-P. Serre, *Linear representations of finite groups*, French edition, Graduate Texts in Mathematics, vol. Vol. 42, Springer-Verlag, New York-Heidelberg, 1977. MR0450380

[8] T. Tao, *Product set estimates for non-commutative groups*, Combinatorica **28** (2008), no. 5, 547–594, DOI 10.1007/s00493-008-2271-7. MR2501249

[9] T. Tao, *Expansion in finite simple groups of Lie type*, Graduate Studies in Mathematics, vol. 164, American Mathematical Society, Providence, RI, 2015. MR3309986