

RANDOM WALK ON FINITE GROUPS

ALIREZA SALEHI GOLSEFIDY

CONTENTS

1. Lecture 1: Random-walk via a group action. Spectral gap, and groups with property (T).	1
2. Lecture 2: Groups with property(T), Kazhdan constant, contraction, and spectral gap	3
3. Lecture 3: Expanders, mixing inequality, and discrete Laplacian	7
4. Lecture 4: Cheeger's inequality for regular graphs and explicit construction of expanders.	11
5. Lecture 5: Fourier analysis on finite groups, I.	14
6. Lecture 6: Fourier analysis on finite groups, II.	17
7. Lecture 7: Fourier analysis on finite groups, III. Margulis-Gabber-Galil expanders, I.	19
8. Lecture 8: Margulis-Gabber-Galil expanders, II. Gowers's quasi-randomness	23
9. Lecture 9: Quasi-randomness and product of large sets. Entropies and the Bourgain-Gamburd method.	26
10. Lecture 10: Quasi-random groups. Random-walk on free groups, I.	29
11. Lecture 11: Random-walk on free groups, II. Bourgain-Gamburd method, I.	32
12. Lecture 12: Bourgain-Gamburd method, II. Multiplicative energy. Approximate subgroups.	35
13. Lecture 13: A detour: the sum-product phenomenon over finite fields.	39
14. Lecture 14: Product theorem, I.	41
15. Lecture 15: Prodcut theorem, II.	42
16. Lecture 16: Approximate subgroup, III. Ping-pong argument and escaping algebraic subgroups.	45
17. Lecture 17: Ping-Pong argument and escaping proper subgroups, II.	48
18. Lecture 18: Random-walk on group extensions, I	51
19. Additional readings	54
20. Problems	54

1. LECTURE 1: RANDOM-WALK VIA A GROUP ACTION. SPECTRAL GAP, AND GROUPS WITH PROPERTY (T).

Suppose G is a group and it acts on a set X . Suppose S is a finite subset of G . Consider the following stochastic process:

- (1) We start with a random point y_0 in X with the probability law given by the measure ν_0 .
- (2) At step i , we choose an element g_i randomly from the set S and let it act on the random point y_{i-1} to obtain the random point y_i .

This process after ℓ steps gives us a random-variable Y_ℓ with values in X .

We would like to describe the probability law of Y_ℓ . This means finding the probability measure ν_ℓ such that for every measurable set A ,

$$\nu_\ell(A) := \mathbb{P}(Y_\ell \in A).$$

Note that

$$(1) \quad \nu_{\ell+1}(A) = \mathbb{P}(Y_{\ell+1} \in A) = \frac{1}{|S|} \sum_{s \in S} \mathbb{P}(s \cdot Y_\ell \in A) = \frac{1}{|S|} \sum_{s \in S} \nu_\ell(s^{-1} \cdot A).$$

For a measure ν on X and a measure μ with finite support $S \subseteq G$, we define

$$(\mu * \nu)(A) := \sum_{s \in S} \mu(s) \nu(s^{-1}A).$$

Hence, by (1), it follows that

$$\nu_{\ell+1} = \mu_S * \nu_\ell = \mu_S * (\mu_S * (\cdots * (\mu_S * \nu_0) \cdots)),$$

where ν_S is the probability counting measure on S . Note that if μ_1 and μ_2 are two probability measures with finite support on G and ν is a probability measure on X , then

$$(\mu_1 * \mu_2) * \nu = \mu_1 * (\mu_2 * \nu),$$

where $(\mu_1 * \mu_2)(g) = \sum_{g_1 g_2 = g} \mu_1(g_1) \mu_2(g_2)$. Hence, $\nu_\ell = \mu_S^{(\ell)} * \nu_0$, where

$$\mu_S^{(\ell)} := \underbrace{\mu_S * \cdots * \mu_S}_{\ell\text{-times}}.$$

So, to understand ν_ℓ , we may begin by studying $\mu_S^{(\ell)}$. This means analyzing the ℓ -step random-walk X_ℓ on G with respect to the measure μ_S . Hence, we take an *observer* $f \in C(G)$ and investigate the expected value of $f(X_\ell)$. Note

$$\mathbb{E}[f(X_\ell)] = \sum_{g \in G} \mu_S^{(\ell)}(g) f(g).$$

Assuming S is symmetric and the group generated by S is dense in G , the Kawada-Itô theorem implies that

$$\lim_{\ell \rightarrow \infty} \mathbb{E}[f(X_\ell)] = \frac{1}{m_G(G)} \int_G f(x) dm_G(x),$$

where m_G is a Haar measure of G . Recall that every compact second countable group G has a regular Borel measure which is G -invariant.

Applying the same result to $(\rho_x f)(g) = f(gx)$, we deduce that

$$\lim_{\ell \rightarrow \infty} \sum_{g \in G} \mu_S^{(\ell)}(g) f(gx) = \frac{\langle f, \mathbb{1}_G \rangle}{\langle \mathbb{1}_G, \mathbb{1}_G \rangle}$$

Since S is symmetric,

$$\sum_{g \in G} \mu_S^{(\ell)}(g) f(gx) = \sum_{g \in G} \mu_S^{(\ell)}(g^{-1}) f(gx) = (\mu_S^{(\ell)} * f)(x).$$

Therefore, for all $x \in G$

$$\lim_{\ell \rightarrow \infty} (\mu_S^{(\ell)} * f)(x) = \frac{\langle f, \mathbb{1}_G \rangle}{\langle \mathbb{1}_G, \mathbb{1}_G \rangle},$$

which means

$$\lim_{\ell \rightarrow \infty} \mu_S^{(\ell)} * f = \frac{\langle f, \mathbb{1}_G \rangle}{\langle \mathbb{1}_G, \mathbb{1}_G \rangle} \mathbb{1}_G,$$

pointwise. We would like to understand the error term of this convergence. Instead of considering the pointwise convergence, we focus on the L^2 -convergence. Let

$$T_S: L^2(G, m_G) \rightarrow L^2(G, m_G), \quad T_S(f) = \mu_S * f,$$

where

$$(\mu_S * f)(x) = \frac{1}{|S|} \sum_{s \in S} f(s^{-1}x).$$

Then, $\mu_S^{(\ell)} * f = T_S^\ell(f)$. Note that the left-translation gives us a unitary representation of G ,

$$\lambda_G: G \rightarrow \mathcal{U}(L^2(G)), \quad (\lambda_G(x)(f))(g) = f(x^{-1}g),$$

and we have

$$T_S(f) = \frac{1}{|S|} \sum_{s \in S} (\lambda_G(s))(f).$$

Since T_S is an averaging operator, for all $f \in L^2(G)$,

$$\|T_S(f)\|_2 \leq \frac{1}{|S|} \|(\lambda_G(s))(f)\|_2 = \|f\|_2.$$

Hence, $\|T_S\|_{\text{op}} \leq 1$. We also observe that $T_S(\mathbb{1}_G) = \mathbb{1}_G$. Therefore,

$$T_S^\ell(f) - \frac{\langle f, \mathbb{1}_G \rangle}{\langle \mathbb{1}_G, \mathbb{1}_G \rangle} \mathbb{1}_G = T_S^\ell \left(f - \frac{\langle f, \mathbb{1}_G \rangle}{\langle \mathbb{1}_G, \mathbb{1}_G \rangle} \mathbb{1}_G \right).$$

Note that $f - \frac{\langle f, \mathbb{1}_G \rangle}{\langle \mathbb{1}_G, \mathbb{1}_G \rangle} \mathbb{1}_G$ is the orthogonal projection to the space $L^2(G)^\circ$ of constant functions. Thus

$$\|T_S^\ell(f) - \frac{\langle f, \mathbb{1}_G \rangle}{\langle \mathbb{1}_G, \mathbb{1}_G \rangle} \mathbb{1}_G\|_2 \leq \|T_S|_{L^2(G)^\circ}\|_{\text{op}}^\ell \|f\|_2.$$

Therefore, we obtain an exponentially small error if

$$(2) \quad \|T_S|_{L^2(G)^\circ}\|_{\text{op}} < 1.$$

It is useful to define a similar averaging operator for other unitary representations. A group homomorphism $\pi: G \rightarrow \mathcal{U}(\mathcal{H})$ to the group of unitary operators on a Hilbert space which is strongly continuous is called a unitary representation of G – π is called strongly continuous if for every $v \in \mathcal{H}$,

$$G \rightarrow \mathcal{H}, \quad x \mapsto \pi(x)v$$

is continuous. For a measure μ with finite support, we let

$$\pi(\mu) := \sum_{s \in G} \mu(s) \pi(s).$$

Note that again $\|\pi(\mu)\|_{\text{op}} \leq 1$. In the next lecture, we will define groups that have property (T) and use them to provide examples where (2) holds.

2. LECTURE 2: GROUPS WITH PROPERTY(T), KAZHDAN CONSTANT, CONTRACTION, AND SPECTRAL GAP

For a measure μ with finite support on a compact group G and a unitary representation π of G , we defined

$$\pi(\mu) := \sum_{x \in G} \mu(x) \pi(x).$$

We mentioned that

$$\|\mu_S^{(\ell)} * f - \left(\frac{1}{|G|} \int_G f(x) dx \right) \mathbb{1}_G\|_2 \leq \|(\lambda_G(\mu_S))|_{L^2(G)^\circ}\|_{\text{op}}^\ell \|f\|_2.$$

So, we get a very good rate of L^2 -convergence of $\mu_S^{(\ell)} * f$ to the constant function given by the average of f if $\pi(\mu_S)$ is a contraction where $\pi := \lambda_G|_{L^2(G)^\circ}$.

We set

$$\lambda(\mu; \pi) := \|\pi(\mu)\|_{\text{op}} \quad \text{and} \quad \mathcal{L}(\mu; \pi) := -\log \lambda(\mu; \pi).$$

Note that, for every $v, w \in \mathcal{H}$

$$\begin{aligned} \langle \pi(\mu)v, w \rangle &= \sum_s \mu(s) \langle \pi(s)v, w \rangle \\ &= \sum_s \mu(s) \langle v, \pi(s^{-1})w \rangle = \langle v, \pi(\check{\mu})w \rangle, \end{aligned}$$

where $\check{\mu}(s) = \mu(s^{-1})$. This means $\pi(\mu)^* = \pi(\check{\mu})$. In particular, if μ is symmetric, then $\pi(\mu)$ is self-adjoint. Hence, it has an orthonormal eigenbasis with real eigenvalues at least when $\mathcal{H}$ is of finite dimension. A similar statement holds for *compact self-adjoint operators*. Suppose $\{v_i\}_{i \in I}$ is an orthonormal eigenbasis of $\mathcal{H}$, and $\pi(\mu)v_i = \lambda_i v_i$. The set $\{\lambda_i\}_i$ is called the *spectrum* of $\pi(\mu)$. Then for every unit vector v ,

$$\begin{aligned} v &= \sum_i (v \cdot v_i) v_i, \quad 1 = \|v\|^2 = \sum_i |v \cdot v_i|^2, \\ \pi(\mu)v &= \sum_i \lambda_i (v \cdot v_i) v_i, \quad \|\pi(\mu)v\|^2 = \sum_i |v \cdot v_i|^2 \lambda_i^2 \leq \sup_i |\lambda_i|^2. \end{aligned}$$

Hence, when μ is symmetric and either $\pi(\mu)$ is a compact operator or π is a finite-dimensional representation, then $\lambda(\mu; \pi)$ is the maximum of eigenvalues of $\pi(\mu)$. That is why we say μ has the *spectral gap property* for π if $\lambda(\mu; \pi) < 1$. When G acts on a measure space X , for a measure μ with finite support on G , we say $\mu \curvearrowright X$ has the *spectral gap property* if $\lambda(\mu; L^2(X)^\circ) < 1$.

Note that if $\pi(s)v$ is *almost* v for every s in the support of μ , then $\pi(\mu)v$ is *almost* v . This means we cannot expect to obtain a contraction. Let's make this idea precise.

For a finite subset S of G , a unitary representation π of G , and $v \in \mathcal{H} \setminus \{0\}$, we let

$$\delta_S(v) := \max_{s \in S} \frac{\|\pi(s)v - v\|}{\|v\|}.$$

We define

$$\kappa(S; \pi) := \inf_{v \in \mathcal{H}} \delta_S(v).$$

For a measure μ with finite support S , we let

$$\delta_\mu(v) := \delta_S(v) \quad \text{and} \quad \kappa(\mu; \pi) := \kappa(S; \pi).$$

The following proposition gives us a connection between $\kappa(\mu; \pi)$ and $\|\pi(\mu)\|_{\text{op}}$.

Proposition 1 (Displacement-Contraction Equivalence). *Suppose π is a unitary representation of a group G and μ is a measure with finite support on G . Then the following inequalities hold:*

$$1 - \|\pi(\mu)\|_{\text{op}} \leq \kappa(\mu; \pi) \quad \text{and} \quad \|\pi(\mu)\|_{\text{op}} \leq \sqrt{1 - p_0^2 \kappa(\mu; \pi)^2},$$

where $p_0 := \min_{s \in S} \mu(s)$.

Proof. Let S be the support of μ . For every unit vector $v \in \mathcal{H}$,

$$\|v - \pi(\mu)v\| \geq \|v\| - \|\pi(\mu)v\| \geq 1 - \|\pi(\mu)\|_{\text{op}},$$

and

$$\|v - \pi(\mu)v\| = \left\| \sum_{s \in S} \mu(s)(v - \pi(s)v) \right\| \leq \sum_{s \in S} \|v - \pi(s)v\| \leq \max_{s \in S} \|\pi(s)v - v\|.$$

Hence, $1 - \|\pi(\mu)\|_{\text{op}} \leq \delta_\mu(\pi)$.

To prove the second inequality, write $S = \{s_1, \dots, s_n\}$, and, for a unit vector v , set $v_i := \pi(s_i)v$. Note that

$$\|v_i\| = 1 \quad \text{for all } i, \quad \text{and} \quad \max_{i,j} \|v_i - v_j\| \geq \delta_\mu(v),$$

and we must show

$$\left\| \sum_{i=1}^n \mu(s_i) v_i \right\| \leq \sqrt{1 - p_0^2 \delta_\mu(v)^2}.$$

Hence, the second inequality quantitatively expresses the *strict convexity* of the unit ball in a Hilbert space: a weighted average of points on the unit sphere that are not very close to one another lies strictly inside the ball. Using dot product, it follows that

$$\|v_i - v_j\|^2 = 2 - 2\text{Re}(v_i \cdot v_j),$$

and

$$\left\| \sum_{i=1}^n \mu(s_i) v_i \right\|^2 = \sum_{i=1}^n \mu(s_i)^2 + 2 \sum_{i < j} \mu(s_i) \mu(s_j) \text{Re}(v_i \cdot v_j).$$

Hence

$$\begin{aligned} \left\| \sum_{i=1}^n \mu(s_i) v_i \right\|^2 &= \sum_{i=1}^n \mu(s_i)^2 + \sum_{i < j} \mu(s_i) \mu(s_j) (2 - \|v_i - v_j\|^2) \\ &= \left(\sum_{i=1}^n \mu(s_i) \right)^2 - \sum_{i < j} \mu(s_i) \mu(s_j) \|v_i - v_j\|^2 \leq 1 - p_0^2 \max_{i < j} \|v_i - v_j\|^2. \end{aligned}$$

Therefore for every unit vector v ,

$$\|\pi(\mu)v\|^2 \leq 1 - p_0^2 \delta_\mu(v)^2 \leq 1 - p_0^2 \kappa(\mu; \pi)^2,$$

which implies $\|\pi(\mu)\|_{\text{op}} \leq \sqrt{1 - p_0^2 \kappa(\mu; \pi)^2}$. This finishes the proof. $\square$

We say a group Γ has Kazhdan's property(T) if there exists a finite subset S of Γ and a positive number ε such that for every unitary representation $\pi: \Gamma \rightarrow \mathcal{U}(\mathcal{H})$ with no non-zero fixed vector,

$$(3) \quad \kappa(S; \pi) \geq \varepsilon.$$

Let $\kappa(S) := \inf\{\kappa(S; \pi) \mid \pi \text{ is a unitary representation with no non-zero fixed vector}\}$. So, (3) is equivalent to $\kappa(S) \geq \varepsilon$. The constant $\kappa(S)$ is called the *Kazhdan constant of Γ with respect to S* .

Lemma 2. *Suppose $\Gamma = \langle S \rangle$ and S is a finite set. If Γ has property(T), then $\kappa(S) > 0$. Moreover, if S' is another generating set and $S' \subseteq \prod_{\ell_0} S$, then $\kappa(S') \geq \frac{\kappa(S)}{\ell_0}$.*

Proof. Since Γ has property(T), there exists a finite subset S_0 such that $\kappa(S_0) > 0$. Since S_0 is a finite set and S is a generating set of Γ , there exists a natural number ℓ_0 such that

$$(4) \quad S_0 \subseteq \prod_{\ell_0} (S \cup S^{-1}).$$

For every unitary representation π with no non-zero Γ -fixed point and a unit vector $v \in M_\pi$, there exists $s_0 \in S_0$ such that

$$\|\pi(s_0)v - v\| \geq \kappa(S_0).$$

By (4), there exist $s_1, \dots, s_k \in S \cup S^{-1}$ for some $k \leq \ell_0$ such that

$$s_0 = s_1 s_2 \cdots s_k.$$

Hence,

$$\begin{aligned}\kappa(S_0) &\leq \|\pi(s_1 \cdots s_k)v - v\| = \left\| \sum_{i=1}^k (\pi(s_1 \cdots s_i)v - \pi(s_1 \cdots s_{i-1})v) \right\| \\ &\leq \sum_{i=1}^k \|\pi(s_1 \cdots s_{i-1})(\pi(s_i)v - v)\| \leq \sum_{i=1}^k \|\pi(s_i)v - v\| \\ &\leq k\delta_S(v) \leq \ell_0\delta_S(v).\end{aligned}$$

Therefore for every unit v , $\delta_S(v) \geq \kappa(S_0)/\ell_0$, which implies that $\kappa(S; \pi) \geq \kappa(S_0)/\ell_0$. Thus

$$\kappa(S) \geq \kappa(S_0)/\ell_0 > 0.$$

The same argument implies the other claim. $\square$

Lemma 2 implies that a finitely generated group $\Gamma = \langle S \rangle$ has property(T) if and only if $\kappa(S) > 0$. It is worth pointing out that Gelfand and Zuk proved that there are many finitely generated groups with property(T) such that

$$\inf\{\kappa(S) \mid S \text{ a generating set of } \Gamma\} = 0.$$

Corollary 3. *Let S be a finite subset of a compact group G . Suppose $\Gamma := \langle S \rangle$ is a dense subgroup of G and Γ has property(T). Then $\mu_S \curvearrowright G$ has the spectral gap property. Moreover*

$$\lambda(\mu_S; L^2(G)^\circ) \leq \sqrt{1 - \frac{\kappa(S)^2}{|S|^2}}.$$

Proof. Since $\lambda_G: G \rightarrow \mathcal{U}(L^2(G)^\circ)$ is strongly continuous and Γ is dense in G , a Γ -fixed function is G -invariant, which means it is constant. Thus the only constant function of $L^2(G)^\circ$ is 0. Therefore, there is no non-zero Γ -fixed point in $L^2(G)^\circ$. Hence, by Proposition 1,

$$\lambda(\mu_S; L^2(G)^\circ) \leq \sqrt{1 - \frac{\kappa(\mu_S; L^2(G)^\circ)^2}{|S|^2}} \leq \sqrt{1 - \frac{\kappa(\mu_S)^2}{|S|^2}}.$$

This finishes the proof. $\square$

Corollary 4. *Suppose $\Gamma = \langle S \rangle$ has property(T) where S is a finite set. Then for every normal subgroup N of finite index*

$$\lambda(\mu_S; L^2(G/N)^\circ) \leq \sqrt{1 - \frac{\kappa(\mu_S)^2}{|S|^2}}.$$

This means all these finite-dimensional representations have a uniform spectral gap.

By Corollary 4, we deduce that random-walks on finite groups Γ/N have a uniform rate of convergence to the uniform distribution. In the next lecture, we will define expander graphs and see why Corollary 4 implies that the family of Cayley graphs

$$\{\text{Cay}(\Gamma/N; S) \mid N \text{ normal subgroup of finite-index}\}$$

is a family of expanders.

Kazhdan proved that if G is an almost simple group of real rank at least 2, then every lattice Γ of G has property(T). We say Γ is a lattice in G if

- (1) Γ is discrete.
- (2) G/Γ has a finite G -invariant Radon measure.

A Lie group G is called almost simple if its only normal closed subgroups are finite and central. The real rank of an almost simple Lie group is the dimension of the largest closed subgroup of G which is simultaneously diagonalizable (over $\mathbb{R}$).

Using reduction theory, Minkowski proved that $\mathrm{SL}_n(\mathbb{Z})$ is a lattice in $\mathrm{SL}_n(\mathbb{R})$. Later Siegel gave a more modern treatment of this result. Borel and Harish-Chandra, and independently Mostow and Tamagawa proved that

$$G(\mathcal{O}_k) \subseteq G(k \otimes_{\mathbb{Q}} \mathbb{R})$$

is a lattice where G is a semisimple algebraic group defined over k and k is a number field, i.e. it is a finite extension of $\mathbb{Q}$. For instance, let

$$q(x_1, \dots, x_m, y_1, \dots, y_n) := x_1^2 + \dots + x_m^2 - \sqrt{2}y_1^2 - \dots - \sqrt{2}y_n^2.$$

The corresponding special orthogonal group

$$\mathrm{SO}(q)(A) := \{g \in \mathrm{SL}_{m+n}(A) \mid q(g\mathbf{x}) = q(\mathbf{x})\},$$

for every $\mathbb{Z}[\sqrt{2}]$ -algebra A . By the aforementioned result of Borel and Harish-Chandra, we deduce that

$$\mathrm{SO}(q)(\mathbb{Z}[\sqrt{2}]) \subseteq \mathrm{SO}(q)(\mathbb{Q}[\sqrt{2}] \otimes_{\mathbb{Q}} \mathbb{R}) \simeq \mathrm{SO}(q)(\mathbb{R} \oplus \mathbb{R}) \simeq \mathrm{SO}(m, n) \times \mathrm{SO}(m + n)$$

is a lattice. Moreover, the projection of this lattice to the compact group $\mathrm{SO}(m + n)$ is dense. The real rank of $\mathrm{SO}(m, n)$ is $\min\{m, n\}$, and $\mathrm{SO}(q)$ is an almost simple group unless $(m, n) = (2, 2)$. Hence, we obtain a finite subset S such that $\mu_S \curvearrowright \mathrm{SO}(k)$ has the spectral gap property if $k \geq 5$. This implies that $\mu_S \curvearrowright S^{k-1}$ has the spectral gap property if $k \geq 5$. Independently, Margulis and Sullivan proved that the same statement holds for $k \geq 3$, and they used this to give an affirmative answer to the Banach-Ruziewicz problem: a finitely additive, rotation-invariant measure on the sphere S^k of dimension at least 2 is the Lebesgue measure.

3. LECTURE 3: EXPANDERS, MIXING INEQUALITY, AND DISCRETE LAPLACIAN

The Cheeger constant of a finite graph $G = (V, E)$ is defined as

$$h_G := \min_{A \subseteq V} \frac{|E(A, A^c)|}{\min\{|A|, |A^c|\}},$$

where $A^c := V \setminus A$ and $E(A, A^c)$ is the set of edges connecting a vertex in A to a vertex in A^c . Cheeger defined a similar notion for Riemannian manifolds: given a finite volume manifold M of dimension n , its Cheeger constant is defined as

$$\inf_S \frac{\mathrm{vol}_{n-1}(S)}{\min\{\mathrm{vol}_n(A), \mathrm{vol}_n(B)\}},$$

where S is a smooth submanifold of dimension $n - 1$ that cuts M into two manifolds A and B .

The Cheeger constant provides a quantitative measure of how well the graph G is connected. For example, consider the road network of a city spread across two sides of a river with only a few bridges. If those bridges were to be closed for repair, one half of the city would become disconnected from the other.

A family $\{G_i\}_{i=1}^{\infty}$ of k -regular graphs is called a family of *expanders* if $\inf_i h_{G_i} > 0$ and $|V(G_i)| \rightarrow \infty$. Note that such graphs are extremely *sparse*: the number of edges of G_i is $\frac{k|V(G_i)|}{2}$, so many pairs of vertices are not adjacent. Yet these graphs are nevertheless highly connected, as there is a uniform positive lower bound on their Cheeger constant.

We will show that the connectivity of a finite graph G can be understood by examining how quickly a random-walk on G becomes equidistributed. Starting with a distribution given by the

function $f \in L^2(V)$, the new distribution after one step random walk on the graph is given by

$$(Af)(v) = \sum_{w, v \in E} f(w) \mathbb{P}(w \rightarrow v) = \frac{1}{k} \sum_{w \in N(v)} f(w),$$

where $N(w)$ is the set of neighbors of v . Note that $L^2(V)$ is a Hilbert space with

$$\langle f_1, f_2 \rangle := \sum_{v \in V} f_1(v) \overline{f_2(v)},$$

and $\{\delta_v\}_{v \in V}$ is an orthonormal basis of $L^2(V)$. Hence,

$$f(v) = \langle f, \delta_v \rangle \quad \text{and} \quad f = \sum_{v \in V} \langle f, \delta_v \rangle \delta_v$$

for all $f \in L^2(V)$. Note that

$$\langle A\delta_v, \delta_{v'} \rangle = (A\delta_v)(v') = \frac{1}{k} \sum_{w \in N(v')} \delta_v(w) = \frac{1}{k} [v \in N(v')] = \frac{1}{k} [v \sim v'] = \langle \delta_v, A\delta_{v'} \rangle.$$

Therefore, A is self-adjoint. Hence, it has an orthonormal eigenbasis $\{\phi_1, \dots, \phi_n\}$ that are real-valued functions. Suppose

$$A\phi_i = \lambda_i \phi_i \quad \text{and} \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n.$$

Note that

$$A\mathbb{1} = \mathbb{1} \quad \text{and} \quad \|Af\|_2 \leq \|f\|_2.$$

Hence $1 = \lambda_1 \geq \dots \geq \lambda_n \geq -1$. Similar to the case of random-walk on groups, the rate of convergence to the equidistribution is controlled by

$$\|A\|_{L^2(V)^\circ} \|_{\text{op}}.$$

Note that $L^2(V)^\circ = \bigoplus_{i=2}^n \mathbb{C}\phi_i$. Therefore,

$$\lambda_G := \|A\|_{L^2(V)^\circ} \|_{\text{op}} = \max\{|\lambda_i| \mid i = 2..n\} = \max\{|\lambda_2|, |\lambda_n|\}.$$

Lemma 5. *In the above setting:*

- (1) G is connected if and only if $\lambda_2 < 1$.
- (2) Assuming G is connected, G is bipartite if and only if $\lambda_n = -1$.

Proof. Both parts follow from a *maximal principle* argument. Changing ϕ_i to $-\phi_i$, if necessary, we may assume that

$$\max_v \phi_i(v) = \|\phi_i\|_\infty.$$

Let $V_{i,\max} := \{v \in V \mid \phi_i(v) = \|\phi_i\|_\infty\}$. Suppose $\lambda_2 = 1$. Then for every $v \in V_{2,\max}$,

$$\phi_2(v) = A\phi_2(v) = \frac{1}{k} \sum_{w \in N(v)} \phi_2(w) \leq \|\phi_2\|_\infty = \phi_2(v).$$

Hence, $v \in V_{2,\max}$ implies that $N(v) \subseteq V_{2,\max}$. Therefore, the connected component of G which contains v is a subset of $V_{2,\max}$ and on this connected component ϕ_2 is always $\|\phi_2\|_\infty$. Since ϕ_2 is not a constant function, we deduce that G is not connected.

Conversely, if H is a connected component of G , then $A\mathbb{1}_{V_H} = \mathbb{1}_{V_H}$. Hence, if G is not connected, then $\lambda_2 = 1$.

Suppose $\lambda_n = -1$ and $v \in V_{n,\max}$. Then

$$-\phi_n(v) = A\phi_n(v) = \frac{1}{k} \sum_{w \in N(v)} \phi_n(w) \geq -\|\phi_n\|_\infty = -\phi_n(v).$$

Hence, $\phi_n(v) = \|\phi_n\|_\infty$ implies that $N(v) \subseteq \{w \in V \mid \phi_n(w) = -\|\phi_n\|_\infty\}$. By a similar argument, $\phi_n(v) = -\|\phi_n\|_\infty$ implies that $N(v) \subseteq V_{n,\max}$. Since G is connected, we deduce that G is a bipartite graph where the partition of vertices is given by

$$V_{n,\max} \sqcup \{w \in V \mid \phi_n(w) = -\|\phi_n\|_\infty\}.$$

Conversely, if G is bipartite and the partition $A \sqcup B$ gives us the side vertices, then

$$A(\mathbb{1}_A - \mathbb{1}_B) = -(\mathbb{1}_A - \mathbb{1}_B),$$

and so $\lambda_n = -1$. □

Remark 6. The above argument implies that

$$\dim \ker(A - \text{id}) = \# \text{connected components of } G$$

$$\dim \ker(A + \text{id}) = \# \text{bipartite connected components of } G.$$

For every function $f \in L^2(V)$, the orthogonal projection of f onto the space $L^2(V)^\circ$ is given by $f - \frac{\langle f, \mathbb{1}_V \rangle}{\langle \mathbb{1}_V, \mathbb{1}_V \rangle} \mathbb{1}_V$. Hence

$$(5) \quad \|A^\ell f - \text{Ave}(f) \mathbb{1}_V\|_2 = \left\| A^\ell \left(f - \frac{\langle f, \mathbb{1}_V \rangle}{\langle \mathbb{1}_V, \mathbb{1}_V \rangle} \mathbb{1}_V \right) \right\|_2 \leq \lambda_G^\ell \|f\|_2.$$

By (5) and the Cauchy-Schwarz inequality,

$$(6) \quad |\langle A^\ell f, g \rangle - |V| \text{Ave}(f) \text{Ave}(g)| = |\langle A^\ell (f - \text{Ave}(f) \mathbb{1}_V), g \rangle| \leq \lambda_G^\ell \|f\|_2 \|g\|_2.$$

We refer to (6) as a *mixing inequality*. If we start with a point randomly chosen from a subset X of vertices, the distribution after ℓ -step random-walk on the vertices of the graph is given by

$$A^\ell \mu_X.$$

Hence, the probability of ending up in another subset Y of vertices is

$$\sum_{v \in Y} (A^\ell \mu_X)(v) = \langle A^\ell \mu_X, \mathbb{1}_Y \rangle.$$

Therefore, by (6)

$$(7) \quad |\mathbb{P}(Z^{(\ell)} \in Y) - \frac{|Y|}{|V|}| \leq \lambda_G^\ell \sqrt{\frac{|Y|}{|X|}}.$$

In particular, starting with a vertex v_0 ,

$$|\mathbb{P}(Z^{(\ell)} = v) - \frac{1}{|V|}| \leq \lambda_G^\ell,$$

which implies that $\mathbb{P}(Z^{(\ell)} = v) > 0$ if $\lambda_G^\ell < \frac{1}{|V|}$. This is equivalent to

$$\ell > \frac{1}{\mathcal{L}_G} \log |V|, \quad \text{where} \quad \mathcal{L}_G := -\log \lambda_G.$$

Therefore

$$\text{diam}(G) \leq \frac{1}{\mathcal{L}_G} \log |V| + 1.$$

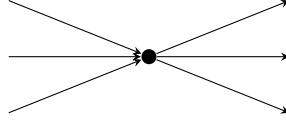
We fix an orientation for the edges of G . For an edge e , the initial point of e is denoted by e^- and its terminal point is denoted by e^+ . We let

$$d: L^2(V) \rightarrow L^2(E), \quad df(e) := f(e^+) - f(e^-).$$

We can think about this as either a type of *directional derivative* or the differential map from functions on 1-cells and to functions on 0-cells. Then for $f \in L^2(V)$ and $g \in L^2(E)$, we have

$$\begin{aligned}\langle df, g \rangle_E &= \sum_{e \in E} df(e) \overline{g(e)} = \sum_{e \in E} (f(e^+) - f(e^-)) \overline{g(e)} \\ &= \sum_{v \in V} f(v) \left(\sum_{e^+=v} \overline{g(e)} - \sum_{e^-=v} \overline{g(e)} \right).\end{aligned}$$

Hence, $d^*g(v) = \sum_{e^+=v} g(e) - \sum_{e^-=v} g(e)$. Thinking about $g(e)$ as the amount of a *liquid* going through the edge e , then $d^*g(v)$ measures how much source or sink the vertex v is.



Then

$$d^*df(v) = \sum_{e^+=v} df(e) - \sum_{e^-=v} df(e) = \sum_{e^+=v} (f(v) - f(e^-)) - \sum_{e^-=v} (f(e^+) - f(v)) = kf(v) - \sum_{w \sim v} f(w).$$

Hence $\Delta := \frac{1}{k}d^*d = \text{id} - A$. This is called the *discrete Laplacian*. Note that Δ and A share the same orthonormal eigenbasis $\phi_1, \dots, \phi_n$, and

$$\Delta\phi_i = \mu_i\phi_i, \quad \text{where} \quad \mu_i = 1 - \lambda_i, \quad \text{so} \quad 0 = \mu_1 \leq \mu_2 \leq \dots \leq \mu_n \leq 2.$$

Lemma 7. *In the above setting, $\min\{\|df\|_2/\|f\|_2 \mid 0 \neq f \in L^2(V)^\circ\} = \sqrt{k\mu_2}$.*

Proof. Since $f \in L^2(V)^\circ$, $f = \sum_{i=2}^n c_i\phi_i$ for some $c_i \in \mathbb{C}$. As $\|f\|_2^2 = 1$, we deduce that $\sum_{i=2}^n |c_i|^2 = 1$. Note that

$$\|df\|_2^2 = \langle df, df \rangle_E = \langle f, d^*df \rangle_V = k\langle f, \Delta f \rangle_V = k\left\langle \sum_{i=2}^n c_i\phi_i, \sum_{i=2}^n c_i\mu_i\phi_i \right\rangle = k\sum_{i=2}^n |c_i|^2\mu_i \geq k\mu_2.$$

We also deduce that $\|d\phi_2\|_2^2 = k\mu_2$. This finishes the proof. $\square$

To perform an L^1 -optimization, we will look for good *cuts*. In a social network, *communities* can be recongnized based on certain *features*. A mathematical model of a social network is a graph $G = (V, E)$; a feature is a real-valued function $f: V \rightarrow \mathbb{R}$ on the vertices, and a cut is defined according on whether $f(v) \geq t$ or not for some thershold t . A threshold t is a good choice if the number of edges cut is relatively *small*; that is, if

$$\frac{|E(V_{f,t}^+, V_{f,t}^-)|}{\min\{|V_{f,t}^+|, |V_{f,t}^-|\}}$$

is *small*, where

$$\begin{aligned}V_{f,t}^+ &:= \{v \in V \mid f(v) \geq t\} \quad \text{and,} \\ V_{f,t}^- &:= \{v \in V \mid f(v) < t\}.\end{aligned}$$

We will study this ratio using a probabilistic argument.

4. LECTURE 4: CHEEGER'S INEQUALITY FOR REGULAR GRAPHS AND EXPLICIT CONSTRUCTION OF EXPANDERS.

Suppose $G = (V, E)$ is a finite graph, $f: V \rightarrow \mathbb{R}$ is a function, and μ is a Borel measure on $\mathbb{R}$ such that $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}$. For $t \in \mathbb{R}$, let $V_t := \{v \in V \mid f(v) \geq t\}$. For $e \in E$, let I_e denote the interval whose endpoints are $f(e^-)$ and $f(e^+)$. Let m_0 be the median of f , and define I_v to be the interval with the end points m_0 and $f(v)$. Then

$$\begin{aligned}
 \int_t |E(V_t, V \setminus V_t)| d\mu(t) &= \sum_{e \in E} [t \in I_e] d\mu(t) = \sum_{e \in E} \mu(I_e), \\
 \int_t \min\{|V_t|, |V \setminus V_t|\} d\mu(t) &= \int_{t \leq m_0} |V \setminus V_t| d\mu(t) + \int_{t \geq m_0} |V_t| d\mu(t) \\
 (8) \quad &= \sum_{v \in V} \left(\int_{t \leq m_0} [f(v) < t] d\mu + \int_{t \geq m_0} [f(v) \geq t] d\mu(t) \right) \\
 &= \sum_{v \in V} \mu(I_v).
 \end{aligned}$$

Note that $h_G(\min\{|V_t|, |V \setminus V_t|\}) \leq |E(V_t, V \setminus V)|$ for all t . Therefore, by (8), it follows that

$$(9) \quad h_G \leq \frac{\sum_{e \in E} \mu(I_e)}{\sum_{v \in V} \mu(I_v)}.$$

Let's apply (9) for μ equal to the Lebesgue measure. Then

$$\begin{aligned}
 \mu(I_e) &= |f(e^+) - f(e^-)| = |df(e)| \quad \text{so} \quad \sum_{e \in E} \mu(I_e) = \|df\|_1, \\
 (10) \quad \mu(I_v) &= |f(v) - m_0| \quad \text{so} \quad \sum_{v \in V} \mu(I_v) = \|f - m_0 \mathbb{1}_V\|_1.
 \end{aligned}$$

By (9) and (10), we deduce

$$h_G \leq \frac{\|df\|_1}{\|f - m_0 \mathbb{1}_V\|_1}.$$

Lemma 8. *In the above setting, $\min\{\|df\|_1/\|f\|_1 \mid f \in L^1(V), \text{median}(f) = 0\} = h_G$.*

Proof. We have already proved that the left hand side is more than or equal to the right hand side. Now, suppose $h_G = \frac{|E(A, B)|}{|A|}$, $|A| \leq |B|$, and $B = V \setminus A$. Let $f := \mathbb{1}_A$. Then $|df(e)| = [e \in E(A, B)]$, so $\|df\|_1 = |E(A, B)|$.

If $|A| < |B|$, then $\text{median}(f) = 0$. Hence, $h_G = \|df\|_1/\|f\|_1$ and $\text{median}(f) = 0$.

If $|A| = |B|$, then $\text{median}(f) = 1/2$. Hence, $|f(v) - 1/2| = 1/2$ for all $v \in V$. Therefore, $\|f - \text{median}(f)\|_1 = |V|/2 = |A|$. Thus, $h_G = \|df\|_1/\|f - \text{median}(f)\|_1$. Let $\bar{f} := f - \text{median}(f)$. Note that $df = d\bar{f}$ and $\text{median}(\bar{f}) = 0$. Hence, $h_G = \|d\bar{f}\|_1/\|\bar{f}\|_1$ and $\text{median}(\bar{f}) = 0$. This finishes the proof. $\square$

Lemma 9. $h_G/k \geq 0.5\mu_2$.

Proof. Suppose $h_G = |E(A, B)|/|A|$, $|A| \leq |B|$, and $B := V \setminus A$. Let $f := \mathbb{1}_A$. Then by Lemma 7,

$$k\mu_2 \leq \frac{\|df\|_2^2}{\|f - \text{Ave}(f)\mathbb{1}_V\|_2^2}.$$

Note that $|df(e)| = [e \in E(A, B)]$, so $\|df\|_2^2 = |E(A, B)|$. We also have

$$\|f - \text{Ave}(f)\mathbb{1}_V\|_2^2 = |A| \left(1 - \frac{|A|}{|V|}\right)^2 + (|V| - |A|) \left(\frac{|A|}{|V|}\right)^2 = \frac{|A||B|}{|V|} \geq \frac{|A|}{2}.$$

Hence,

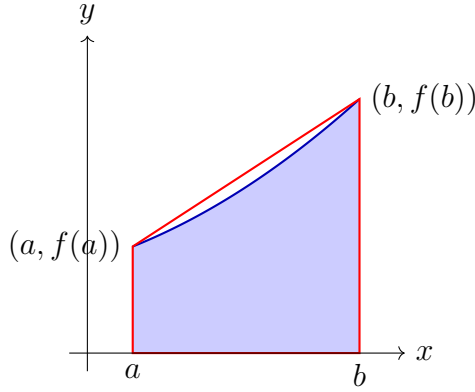
$$k\mu_2 \leq \frac{|E(A, B)|}{|A|/2} = 2h_G.$$

This finishes the proof. $\square$

To obtain a lower bound for μ_2 in terms of h_G , we use (9) with a measure inspired by torque. Suppose $\text{median}(f) = 0$ and let $d\mu(t) = |t|dt$. In this setting, we have

$$\mu([a, b]) = \begin{cases} \frac{b^2 - a^2}{2} & \text{if } a \geq 0 \\ \frac{b^2 + a^2}{2} & \text{if } a < 0 \leq b \\ \frac{a^2 - b^2}{2} & \text{if } b < 0. \end{cases} = \begin{cases} \frac{|a^2 - b^2|}{2} & \text{if } ab \geq 0 \\ \frac{a^2 + b^2}{2} & \text{if } ab \leq 0. \end{cases}$$

Therefore $\mu(I_v) = |f(v)|^2/2$, and $\mu([a, b]) \leq |b - a| \frac{|a| + |b|}{2}$ (One can see this inequality based on the convexity of the function $t \mapsto |t|$).



Thus

$$\sum_{v \in V} \mu(I_v) = \|f\|_2^2/2 \quad \text{and} \quad \sum_{e \in E} \mu(I_e) \leq \sum_{e \in E} |df(e)| \left(\frac{|f(e^+)| + |f(e^-)|}{2} \right).$$

Hence, by (9), it follows

$$h_G \leq \frac{\sum_{e \in E} |df(e)| \left(\frac{|f(e^+)| + |f(e^-)|}{2} \right)}{\|f\|_2^2/2}.$$

By the Cauchy-Schwarz inequality, it follows that

$$\sum_{e \in E} |df(e)| \left(\frac{|f(e^+)| + |f(e^-)|}{2} \right) \leq \|df\|_2 \sqrt{\sum_{e \in E} \left(\frac{|f(e^+)| + |f(e^-)|}{2} \right)^2}$$

Since $\left(\frac{|f(e^+)| + |f(e^-)|}{2} \right)^2 \leq \frac{|f(e^+)|^2 + |f(e^-)|^2}{2}$, it follows that

$$\sum_{e \in E} \left(\frac{|f(e^+)| + |f(e^-)|}{2} \right)^2 \leq \sum_{e \in E} \frac{|f(e^+)|^2 + |f(e^-)|^2}{2} = \frac{k}{2} \|f\|_2^2.$$

Altogether, we deduce

$$h_G \leq \frac{\|df\|_2 \sqrt{\frac{k}{2}} \|f\|_2}{\|f\|_2^2/2} = \sqrt{2k} \frac{\|df\|_2}{\|f\|_2} \leq \sqrt{2k} \frac{\|df\|_2}{\|f - \text{Ave}(f)\mathbb{1}_V\|_2}$$

if $\text{median}(f) = 0$. The last inequality implies that

$$h_G \leq \sqrt{2k} \frac{\|df\|_2}{\|f\|_2}$$

for every $f \in L^2(V)^\circ$. Hence, by Lemma 7, we deduce that

$$h_G \leq \sqrt{2k} \sqrt{k\mu_2} = k\sqrt{2\mu_2}.$$

Theorem 10. $\frac{1}{2}\mu_2 \leq \frac{h_G}{k} \leq \sqrt{2\mu_2}$.

In particular, we obtain the following corollary:

Corollary 11. *A family $\{G_i\}_{i=1}^\infty$ of k -regular graphs is a family of expanders if $\lim_{i \rightarrow \infty} |V(G_i)| = \infty$ and $\sup_i \lambda_{G_i} < 1$.*

Proof. Let $\lambda_0 := \sup_i \lambda_{G_i}$. Recall that $\lambda_{G_i} := \max\{|\lambda_2|, |\lambda_n|\}$. So, $\mu_2(G_i) \geq 1 - \lambda_0$. Therefore, by Theorem 10,

$$\inf_i h_{G_i} \geq \frac{k}{2} \mu_2(G_i) > \frac{k}{2} (1 - \lambda_0) > 0.$$

This finishes the proof. $\square$

Theorem 12 (1st explicit construction due to Margulis). *Suppose Γ has property (T) and is generated by a finite symmetric set S . Assume further that Γ has normal subgroups of arbitrarily large finite index. Then*

$$\{\text{Cay}(\Gamma/N; S) \mid N \triangleleft \Gamma, |\Gamma/N| < \infty\}$$

is a family of expanders, where $\text{Cay}(\Gamma/N; S)$ is the Cayley graph of Γ/N with respect to the generating set given by the cosets of S .

Proof. Note that random-walk on the Cayley graph $\text{Cay}(\Gamma/N; S)$ is the same as a random-walk on Γ/N with respect to the probability counting measure on the set $\bar{S} := \{sN \mid s \in S\}$. By Corollary 3,

$$\lambda(\mu_{\bar{S}}; L^2(\Gamma/N)) \leq \sqrt{1 - \frac{\kappa(\mu_{\bar{S}})^2}{|\bar{S}|^2}} \leq \sqrt{1 - \frac{\kappa(\mu_S)^2}{|S|^2}}.$$

Hence,

$$\sup\{\lambda_{\text{Cay}(\Gamma/N; S)} \mid N \triangleleft \Gamma, |\Gamma/N| < \infty\} \leq \sqrt{1 - \frac{\kappa(\mu_S)^2}{|S|^2}} < 1.$$

This finishes the proof. $\square$

By a result of Kazhdan $\text{SL}_n(\mathbb{Z})$ has property (T) if $n \geq 3$. More generally every lattice in an almost simple group of rank at least 2 has property (T). Therefore, we deduce the following corollary:

Corollary 13. *Let S be a symmetric generating set of $\text{SL}_n(\mathbb{Z})$ where $n \geq 3$. Then the following is a family of expanders*

$$\{\text{Cay}(\text{SL}_n(\mathbb{Z}/m\mathbb{Z}), \pi_m(S)) \mid m \in \mathbb{N}\}$$

where $\pi_m(S)$ is the set of S modulo m .

Prior to the mentioned work of Kazhdan, Selberg had studied the spectrum of the Laplacian of congruence covers of the modular surface, and as a result deduced that Corollary 13 holds for $n = 2$ as well. Due to works of many mathematician, to name a few Burger, Sarnak, Jacquet, Langlands, Clozel, Ullmo, and Oh, on automorphic representations a similar result holds for all S -arithmetic semisimple groups. This takes us to **Lubotzky's 1-2-3 problem**:

Is the following a family of expanders

$$\{\text{Cay}(\text{SL}_2(\mathbb{Z}/p\mathbb{Z}), \pi_p(S_i))\} \mid p \text{ prime more than } 3\},$$

$$\text{where } S_i := \left\{ \begin{pmatrix} 1 & \pm i \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \pm i & 1 \end{pmatrix} \right\}?$$

Due to the aforementioned Selberg's result, the answer is affirmative for $i = 1$ or 2 , but for the group generated by S_3 is not an arithmetic group. Hence, automorphic representation theory is not suitable for answering this question. Bourgain and Gamburd gave an affirmative answer to this question. Their method has been instrumental in many results in the past 20 years. Learning their method is an important goal of this course.

5. LECTURE 5: FOURIER ANALYSIS ON FINITE GROUPS, I.

In this lecture, we give a quick review of the Fourier analysis on finite groups. Almost all the results that we cover today are part of the Peter-Weyl theorem for compact groups.

Definition 1. (1) For a finite group G , we say (ρ, V) is a representation of G if V is a vector space over $\mathbb{C}$ and $\rho: G \rightarrow \text{GL}(V)$ is a group homomorphism. The dimension of V is called the degree of this representation and it is denoted by $\deg \rho$.
 (2) For two finite degree representations (ρ_1, V_1) and (ρ_2, V_2) , $T: V_1 \rightarrow V_2$ is called an *intertwining operator* if it is a linear isomorphism, and, for every $g \in G$, the following is a commuting diagram:

$$\begin{array}{ccc} V_1 & \xrightarrow{\rho_1(g)} & V_1 \\ T \downarrow & & \downarrow T \\ V_2 & \xrightarrow{\rho_2(g)} & V_2 \end{array}$$

- (3) We say two representations (ρ_1, V_1) and (ρ_2, V_2) are equivalent if there exists an intertwining operator $T: V_1 \rightarrow V_2$ which is an isomorphism. In this case, we write $\rho_1 \simeq \rho_2$.
 (4) We say (ρ_1, V_1) is a subrepresentation of (ρ_2, V_2) if there exists an injective intertwining operator $T: V_1 \rightarrow V_2$.
 (5) A representation (ρ, V) is called *irreducible* if it has no non-trivial subrepresentation.
 (6) Suppose $\{(\rho_i, V_i)\}_{i \in I}$ is a family of representations of G . Then $\rho := \oplus_{i \in I} \rho_i$ denotes the representation

$$\rho: G \rightarrow \text{GL}\left(\bigoplus_{i \in I} V_i\right), \quad \rho(g)((v_i)_i) := (\rho_i(g)(v_i))_i.$$

- (7) A representation (ρ, V) is called *completely reducible* if $\rho \simeq \oplus_i \rho_i$ and ρ_i is irreducible.

Note that there exists a natural bijection between representations (ρ, V) of a group G and modules of the group ring $\mathbb{C}G$. One can translate the above concepts in the language of modules:

Representation theory of G	$\mathbb{C}G$ -modules
intertwining operator	module homomorphism
subrepresentation	submodule
irreducible	simple
equivalence	isomorphism

Lemma 14. Suppose G is a finite group. Then every finite degree representation is isomorphic to a unitary representation.

Proof. Suppose $\rho: G \rightarrow \text{GL}(V)$ is a finite degree representation. For $v, w \in V$, let

$$(v, w) := \frac{1}{|G|} \langle \rho(g)(v), \rho(g)(w) \rangle.$$

Then $(,): V \times V \rightarrow \mathbb{C}$ is semilinear, $(v, v) = \frac{1}{|G|} \|\rho(g)(v)\|^2 \geq 0$, and equality holds precisely when $v = 0$. Hence $(,)$ is an inner product on V . Moreover, for all $v, w \in V$ and $g \in G$,

$$(\rho(g)(v), \rho(g)(w)) = (v, w).$$

Let $\{u_1, \dots, u_n\}$ be a basis of V which is orthonormal with respect to $(,)$; this means $(u_i, u_j) = [i = j]$ for all i, j . Let $T: V \rightarrow V$ be the linear map which sends e_i to u_i . Note that T is a linear isomorphism and for every $v, w \in V$,

$$\langle v, w \rangle = (T(v), T(w)).$$

Define

$$\pi: G \rightarrow \text{GL}(V), \quad \pi(g) := T^{-1} \circ \rho(g) \circ T.$$

Then, for every $v, w \in V$,

$$\begin{aligned} \langle \pi(g)(v), \pi(g)(w) \rangle &= (T(\pi(g)(v)), T(\pi(g)(w))) \\ &= (\rho(g)(T(v)), \rho(g)(T(w))) \\ &= (T(v), T(w)) = \langle v, w \rangle. \end{aligned}$$

Hence $\rho \simeq \pi$ and π is a unitary representation. □

Lemma 15. *Every finite degree representation of a finite group is completely reducible.*

Proof. We proceed by induction on the degree of the representation. By Lemma 14, we may assume that ρ is a unitary representation. If it is irreducible, there is nothing to show. Otherwise, there exists a proper non-zero G -invariant subspace W . Let

$$W^\perp := \{v \in V \mid \forall w \in W, \langle v, w \rangle = 0\}.$$

Then, for all $g \in G$, $v \in W^\perp$, and $w \in W$,

$$\langle \rho(g)(v), w \rangle = \langle v, \rho(g^{-1})(w) \rangle = 0$$

we $\rho(g^{-1})(w) \in W$. Hence $W^\perp$ is G -invariant. Note that $V = W \oplus W^\perp$. We can apply the induction hypothesis to W and $W^\perp$ and write them as direct sums of simple $\mathbb{C}G$ -modules. Therefore, V can be written a direct sum of simple $\mathbb{C}G$ -modules. This finishes the proof. □

Lemma 16 (Schur's lemma). *Suppose ρ_1 and ρ_2 are two irreducible representations of G . Then $\text{Hom}_G(\rho_1, \rho_2) = 0$ if $\rho_1 \not\simeq \rho_2$, and $\dim_{\mathbb{C}} \text{Hom}_G(\rho_1, \rho_2) = 1$ if $\rho_1 \simeq \rho_2$.*

Proof. Suppose $T: V_1 \rightarrow V_2$ is a $\mathbb{C}G$ -module homomorphism. Then $\ker T$ and $\text{Im}(T)$ are $\mathbb{C}G$ -modules. Since V_i is a simple $\mathbb{C}G$ -module, either $\ker T = 0$ and $\text{Im}(T) = V_2$, or $\ker T = V_1$ and $\text{Im}(T) = 0$. Hence, every non-zero element of $\text{Hom}_{\mathbb{C}G}(V_1, V_2)$ is an isomorphism. Therefore, $\rho_1 \not\simeq \rho_2$ if and only if $\text{Hom}_{\mathbb{C}G}(V_1, V_2) \neq 0$.

Suppose $T_1, T_2 \in \text{Hom}_{\mathbb{C}G}(V_1, V_2) \setminus \{0\}$. Then $T := T_2^{-1} \circ T_1 \in \text{End}_{\mathbb{C}G}(V_1)$. By the above argument, $\text{End}_{\mathbb{C}G}(V_1)$ is a division ring. If $\lambda \in \mathbb{C}$ is an eigenvalue of T , then $T - \lambda 1_{V_1} \in \text{End}_{\mathbb{C}G}(V_1)$ and it is not invertible. Therefore, $T - \lambda 1_{V_1} = 0$. Thus,

$$T_1 = \lambda T_2.$$

Hence, $\dim_{\mathbb{C}} \text{Hom}_{\mathbb{C}G}(V_1, V_2) = 1$. This finishes the proof. □

For a unitary representation $(\pi, \mathcal{H})$ of a finite group G , $v, w \in \mathcal{H}$, let

$$f_{w,v}^{(\pi)}: G \rightarrow \mathbb{C}, \quad f_{w,v}^{(\pi)}(g) := \langle w, \pi(g)(v) \rangle.$$

Define

$$M_\pi := \text{span}\{f_{w,v}^{(\pi)} \mid v, w \in \mathcal{H}\} \subseteq L^2(G).$$

Lemma 17. *Suppose $(\pi, \mathcal{H})$ is an irreducible unitary representation of a finite group G . Then the following statements hold:*

- (1) *The space of matrix coefficients M_π is invariant under the action of $G \times G$ on $L^2(G)$.*
- (2) *For every $w \in \mathcal{H}$, $v \mapsto f_{v,w}$ is an embedding of π into the left-regular representation of G .*

Proof. Note that

$$\begin{aligned} f_{\pi(g_1)(w), \pi(g_2)(v)}^{(\pi)}(g) &= \langle \pi(g_1)(w), \pi(g)(\pi(g_2)(v)) \rangle \\ &= \langle w, \pi(g_1^{-1}gg_2)(v) \rangle = f_{w,v}^{(\pi)}(g_1^{-1}gg_2) \\ &= ((g_1, g_2) \cdot f_{w,v}^{(\pi)})(g). \end{aligned}$$

Note that $v \mapsto f_{v,w}$ is a linear map, and if for every v , $f_{v,w} = 0$, then $w = 0$. Hence, it is an embedding. By the above computation, $f_{\pi(g)v, w} = \lambda(g)(f_{v,w})$. This finishes the proof. $\square$

Lemma 18. *Suppose π_1 and π_2 are two non-equivalence irreducible unitary representations of a finite group G . Then $M_{\pi_1} \perp M_{\pi_2}$ as subspaces of $L^2(G)$.*

Proof. Note that the following gives us a linear action of G the space $\text{Hom}_{\mathbb{C}}(\mathcal{H}_1, \mathcal{H}_2)$ of linear maps from $\mathcal{H}_1$ to $\mathcal{H}_2$:

$$g \cdot T := \pi_2(g) \circ T \circ \pi_1(g)^{-1}.$$

Note that T is in $\text{Hom}_{\mathbb{C}G}(\mathcal{H}_1, \mathcal{H}_2)$ if and only if $g \cdot T = T$ for all $g \in G$. Therefore for all $T \in \text{Hom}_{\mathbb{C}G}(\mathcal{H}_1, \mathcal{H}_2)$,

$$\text{Ave}_G(T) := \frac{1}{|G|} \sum_{g \in G} g \cdot T \in \text{Hom}_{\mathbb{C}G}(\mathcal{H}_1, \mathcal{H}_2).$$

Since $\pi_1 \not\cong \pi_2$, by Schur's lemma, we deduce that

$$\text{Ave}_G(T) = 0, \quad \forall T \in \text{Hom}_{\mathbb{C}}(\mathcal{H}_1, \mathcal{H}_2).$$

Suppose $\{e_i\}_{i=1}^{n_1}$ and $\{f_i\}_{i=1}^{n_2}$ are orthonormal basis of $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Let

$$T_{i,j}: \mathcal{H}_1 \rightarrow \mathcal{H}_2, \quad T_{i,j}(e_k) = [i = k]f_j.$$

Then

$$\begin{aligned} 0 &= \langle (\text{Ave}_G(T_{i,j}))(e_k), f_\ell \rangle = \frac{1}{|G|} \sum_{g \in G} \langle T_{i,j}(\pi_1(g)(e_k)), \pi_2(g)(f_\ell) \rangle \\ &= \frac{1}{|G|} \sum_{g \in G} \langle T_{i,j}(\sum_r \langle \pi_1(g)(e_k), e_r \rangle e_r), \pi_2(g)(f_\ell) \rangle \\ &= \frac{1}{|G|} \sum_{g \in G} \langle \langle \pi_1(g)(e_k), e_i \rangle f_j, \sum_s \langle \pi_2(g)(f_\ell), f_s \rangle f_s \rangle \\ &= \frac{1}{|G|} \sum_{g \in G} \langle \pi_1(g)(e_k), e_i \rangle \overline{\langle \pi_2(g)(f_\ell), f_j \rangle} \\ &= \frac{1}{|G|} \sum_{g \in G} \overline{f_{e_i, e_k}^{(\pi_1)}(g)} f_{f_\ell, f_j}^{(\pi_2)}(g). \end{aligned}$$

This finishes the proof. $\square$

6. LECTURE 6: FOURIER ANALYSIS ON FINITE GROUPS, II.

We continue with our study of representations of finite groups.

Lemma 19. *Suppose $(\pi, \mathcal{H})$ is an irreducible representation of a finite G . Suppose $(e_1, \dots, e_n)$ is an orthonormal basis of $\mathcal{H}$. Then*

$$\{\sqrt{n}f_{e_i, e_j} \mid 1 \leq i, j \leq n\}$$

is an orthonormal basis of $M_\pi \subseteq L^2(G)$, where the inner product on $L^2(G)$ is defined using the probability counting measure.

Proof. Following proof of Lemma 18, we consider the linear action of G on $\text{End}_{\mathbb{C}}(\mathcal{H})$ given by

$$g \cdot T := \pi(g) \circ T \circ \pi(g)^{-1},$$

and let

$$\text{Ave}_G(T) := \frac{1}{|G|} \sum_{g \in G} g \cdot T.$$

Since $g \cdot \text{Ave}_G(T) = \text{Ave}_G(T)$, $\text{Ave}_G(T)$ is an intertwining operator. Therefore, by Schur's lemma, there exists $c(T) \in \mathbb{C}$ such that

$$(11) \quad \text{Ave}_G(T) = c(T) \text{id}.$$

Taking traces of both sides of (11), we deduce that $nc(T) = \text{Tr}(T)$. Hence,

$$(12) \quad \text{Ave}_G(T) = \frac{\text{Tr}(T)}{n} \text{id}.$$

For every i, j , let $T_{i,j}: \mathcal{H} \rightarrow \mathcal{H}$ be the linear map such that $T_{i,j}(e_k) = [i = k]e_j$. Then, since $\text{Tr}(T_{i,j}) = [i = j]$, by (12), it follows that

$$(13) \quad \text{Ave}_G(T_{i,j}) = \frac{[i = j]}{n} \text{id}.$$

On the other hand, we have:

$$\begin{aligned} \langle (\text{Ave}_G(T_{i,j}))(e_k), e_\ell \rangle &= \frac{1}{|G|} \sum_{g \in G} \langle T_{i,j}(\pi(g)(e_k)), \pi(g)(e_\ell) \rangle \\ &= \frac{1}{|G|} \sum_{g \in G} \overline{\langle f_{e_k, e_i}^{(\pi)}(g) e_j, \pi(g)(e_\ell) \rangle} \\ &= \frac{1}{|G|} \sum_{g \in G} \overline{f_{e_k, e_i}^{(\pi)}(g)} f_{e_\ell, e_j}^{(\pi)}(g). \end{aligned}$$

Therefore, by (13),

$$[i = j][k = \ell] = n \langle f_{e_k, e_i}^{(\pi)}, f_{e_\ell, e_j}^{(\pi)} \rangle,$$

with respect to the probability counting measure on G . By the definition of M_π , f_{e_i, e_j} 's span M_π . This finishes the proof. $\square$

Corollary 20. *If $(\pi, \mathcal{H}_\pi)$ is an irreducible unitary representation of a finite group G , then*

$$M_\pi \simeq \mathcal{H}_\pi^{\deg \pi}$$

as a G -module, and moreover,

$$M_\pi \simeq \text{End}_{\mathbb{C}}(\mathcal{H}_\pi)$$

as a $G \times G$ -module.

Proof. Let

$$\theta: \text{End}_{\mathbb{C}}(\mathcal{H}_{\pi}) \rightarrow M_{\pi}, \quad \theta(T) := f_T, \quad \text{where } f_T(x) := \text{Tr}(\pi(x)^* \circ T).$$

Note that, if $(e_1, \dots, e_n)$ is an orthonormal basis of $\mathcal{H}_{\pi}$, then

$$(14) \quad f_T(x) = \sum_i \langle \pi(x)^* \circ T(e_i), e_i \rangle = \sum_i f_{T(e_i), e_i}^{(\pi)}(x),$$

and so $f_T \in M_{\pi}$.

Next, we show that θ is surjective. Let $T_{i,j} \in \text{End}_{\mathbb{C}}(\mathcal{H}_{\pi})$ be the map given by $T_{i,j}(e_k) = [i = k]e_j$. Then, by (14),

$$\theta(T_{i,j}) = f_{e_j, e_i}^{(\pi)},$$

so θ is surjective.

Now, the first claim follows from Part (2) of Lemma 17 and Lemma 19.

Clearly, θ is a linear map. For every $(g_1, g_2) \in G \times G$,

$$\begin{aligned} ((g_1, g_2) \cdot f_T)(x) &= f_T(g_1^{-1}xg_2) = \text{Tr}(\pi(g_2^{-1}x^{-1}g_1) \circ T) \\ &= \text{Tr}(\pi(x)^*(\pi(g_1) \circ T \circ \pi(g_2)^{-1})) = f_{(g_1, g_2) \cdot T}(x). \end{aligned}$$

Therefore, $(g_1, g_2) \cdot \theta(T) = \theta((g_1, g_2) \cdot T)$. This finishes the proof. $\square$

Lemma 21. *Suppose G is a finite group and $\mathcal{H}_{\pi} \subseteq L^2(G)$ is an irreducible subrepresentation π of the left-regular representation. Then $\mathcal{H}_{\pi} \subseteq M_{\pi}$.*

Proof. Let $(h_1, \dots, h_n)$ be an orthonormal basis of $\mathcal{H}$. Then for every $g \in G$,

$$\pi(g)(h_i) = \sum_j \overline{f_{h_i, h_j}^{(\pi)}}(g) h_j.$$

Hence, for every $g, x \in G$,

$$h_i(g^{-1}x) = \sum_j \overline{f_{h_i, h_j}^{(\pi)}}(g) h_j(x) = \sum_j f_{h_j, h_i}^{(\pi)}(g^{-1}) h_j(x)$$

Applying the above formula for $x = 1$ and g^{-1} , we deduce

$$h_i(g) = \sum_j f_{h_j, h_i}^{(\pi)}(g).$$

Therefore $h_i \in M_{\pi}$. This finishes the proof. $\square$

Definition 2. (1) Suppose G is a finite group. The class of equivalence classes of unitary irreducible representations of G is called the *unitary dual* of G and it is denoted by $\widehat{G}$.

(2) For a finite group G , the Fourier of a function $f \in L^1(G)$ is a function which sends a unitary representation $(\pi, \mathcal{H}_{\pi})$ of G to an operator of the space $\mathcal{H}_{\pi}$ given by

$$\widehat{f}(\pi) := \sum_{x \in G} f(x) \pi(x) \in \text{End}_{\mathbb{C}}(\mathcal{H}_{\pi}).$$

Lemma 22. *Suppose $f, g \in L^1(G)$ and π is an irreducible unitary representation of G . Then*

$$\widehat{f * g}(\pi) = \widehat{f}(\pi) \widehat{g}(\pi).$$

Proof. We have

$$\begin{aligned} \widehat{f * g}(\pi) &= \sum_{x \in G} \widehat{f * g}(x) \pi(x) = \sum_{x_1, x_2 \in G} f(x_1) g(x_2) \pi(x_1 x_2) \\ &= \left(\sum_{x_1} f(x_1) \pi(x_1) \right) \left(\sum_{x_2} g(x_2) \pi(x_2) \right) = \widehat{f}(\pi) \widehat{g}(\pi). \end{aligned}$$

Theorem 23 (Fourier analysis on finite groups). *Suppose G is a finite group. Then the following statements hold:*

- (1) *Every irreducible unitary representation of G is a subrepresentation of the left-regular representation. In particular, the unitary dual of G is finite.*
- (2) *Suppose $\widehat{G} := \{[\pi_1], \dots, [\pi_m]\}$ where π_i 's are irreducible subrepresentations of $L^2(G)$. Then*

$$L^2(G) = \bigoplus_{i=1}^m M_{\pi_i} \simeq \bigoplus_{i=1}^m \mathcal{H}_{\pi_i}^{\deg \pi_i}.$$

- (3) *Let $\mathcal{E}(G) := \bigoplus_{i=1}^m \text{End}_{\mathbb{C}}(\mathcal{H}_{\pi_i})$, and define*

$$\mathcal{F}: L^1(G) \rightarrow \mathcal{E}(G), \quad \mathcal{F}(f) := (\widehat{f}(\pi_i))_{i=1}^m.$$

Then $\mathcal{F}$ is a $\mathbb{C}$ -algebra isomorphism where the multiplication on $L^1(G)$ is given by the convolution operator.

- (4) *For $(x_i), (y_i) \in \mathcal{E}(G)$, define*

$$\langle (x_i), (y_i) \rangle := \frac{1}{|G|} \sum_i \deg \pi_i \text{Tr}(x_i y_i^*).$$

Then for every $f, g \in L^2(G)$,

$$\langle \mathcal{F}(f), \mathcal{F}(g) \rangle = \langle f, g \rangle.$$

In particular, we have the following:

$$\|f\|_2^2 = \frac{1}{|G|} \sum_{i=1}^m \deg \pi_i \|\widehat{f}(\pi_i)\|_{\text{HS}}^2 \quad (\text{Parseval's formula}).$$

- (5) *For every $f \in L^2(G)$,*

$$f(x) = \frac{1}{|G|} \sum_i \deg \pi_i \text{Tr}(\widehat{f}(\pi_i) \pi_i(x)^*).$$

7. LECTURE 7: FOURIER ANALYSIS ON FINITE GROUPS, III. MARGULIS-GABBER-GALIL EXPANDERS, I.

We start by proving Theorem 23.

Proof of Theorem 23. Part (1) follows from Corollary 20.

Part (2) is a consequence of Lemma 15, Lemma 21, and Corollary 20.

Note that by Lemma 21, $\mathcal{F}$ is a ring homomorphism, and by Part (2),

$$\dim_{\mathbb{C}} L^1(G) = \dim_{\mathbb{C}} \mathcal{E}(G).$$

To show that $\mathcal{F}$ is an isomorphism, it is sufficient to prove Part (4). To prove a linear map preserves inner product, it is necessary and sufficient to show that it is an isometry.

By Lemma 18, Lemma 19, and Part (2),

$$\left\{ \sqrt{\frac{\deg \pi_k}{|G|}} f_{e_i^{(k)}, e_j^{(k)}}^{(\pi_k)} \mid 1 \leq k \leq m, 1 \leq i, j \leq \deg \pi_k \right\}$$

is an orthonormal basis of $L^2(G)$, where

$$\{e_1^{(k)}, \dots, e_{\deg \pi_k}^{(k)}\}$$

is an orthonormal basis of $\mathcal{H}_{\pi_k}$. Let

$$g_{i,j}^{(k)} := \sqrt{\frac{\deg \pi_k}{|G|}} f_{e_i^{(k)}, e_j^{(k)}}^{(\pi_k)}$$

Hence, for every $f \in L^2(G)$,

$$f = \sum_{k,i,j} \langle f, g_{i,j}^{(k)} \rangle g_{i,j}^{(k)} \quad \text{and} \quad \|f\|_2^2 = \sum_{k,i,j} |\langle f, g_{i,j}^{(k)} \rangle|^2.$$

Note that

$$\begin{aligned} \langle f, g_{i,j}^{(k)} \rangle &= \sum_{x \in G} f(x) \overline{g_{i,j}^{(k)}(x)} = \sqrt{\frac{\deg \pi_k}{|G|}} \sum_{x \in G} f(x) \langle \pi_k(x)(e_j^{(k)}), e_i^{(k)} \rangle \\ &= \sqrt{\frac{\deg \pi_k}{|G|}} \langle \widehat{f}(\pi_k)(e_j^{(k)}), e_i^{(k)} \rangle = \sqrt{\frac{\deg \pi_k}{|G|}} [\widehat{f}(\pi_k)]_{ij}, \end{aligned}$$

where $[\widehat{f}(\pi_k)]_{ij}$ is the ij -entry of the matrix representation of $\widehat{f}(\pi_k)$ written in the basis $(e_i^{(k)})_i$. Therefore,

$$\begin{aligned} f &= \frac{1}{|G|} \sum_{k,i,j} \deg \pi_k [\widehat{f}(\pi_k)]_{ij} \overline{[\pi_k(x)]_{ij}} = \frac{1}{|G|} \sum_k \deg \pi_k \text{Tr}(\widehat{f}(\pi_k) \pi_k(x)^*), \quad \text{and} \\ \|f\|_2^2 &= \frac{1}{|G|} \sum_{k,i,j} \deg \pi_k |[\widehat{f}(\pi_k)]_{ij}|^2 = \frac{1}{|G|} \sum_k \deg \pi_k \|\widehat{f}(\pi_k)\|_{\text{HS}}^2. \end{aligned}$$

This finishes the proof. $\square$

Corollary 24. For $\pi \in \widehat{G}$, let $\chi_\pi: G \rightarrow \mathbb{C}$, $\chi_\pi(x) := \text{Tr}(\pi(x))$. Then the following statements hold:

- (1) The center $Z(L^1(G))$ of $L^1(G)$ consists of all the class functions; that means functions which satisfy $f(x) = f(yxy^{-1})$ for all $x, y \in G$.
- (2) $\dim_{\mathbb{C}} Z(L^1(G)) = \kappa(G) = |\widehat{G}|$, where $\kappa(G)$ is the number of conjugacy classes of G .
- (3) $\{\frac{1}{\sqrt{|G|}} \chi_\pi \mid \pi \in \widehat{G}\}$ is an orthonormal basis of $Z(L^1(G)) \subseteq L^2(G)$.

Proof. Exercise. $\square$

Theorem 25 (Fourier analysis on finite abelian groups). Suppose G is an abelian group. Then the following holds:

- (1) For every $\pi \in \widehat{G}$, $\deg \pi = 1$, so $\widehat{G} = \text{Hom}(G, S^1)$.
- (2) $\widehat{G}$ is an abelian group under pointwise multiplication.
- (3) $(L^1(G), +, *) \rightarrow (L^1(\widehat{G}), +, \cdot)$, $f \mapsto \widehat{f}$ is a $\mathbb{C}$ -algebra isomorphism, where

$$\widehat{f}(\pi) = \sum_{x \in G} f(x) \pi(x) \in \mathbb{C}.$$

- (4) $L^2(G) \rightarrow L^2(\widehat{G})$, $f \mapsto \widehat{f}$ is an isometry of Hilbert spaces; that means

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

(5) For every $f \in L^2(G)$,

$$f(x) = \frac{1}{|G|} \sum_{\pi \in \widehat{G}} \widehat{f}(\pi) \overline{\pi(x)}.$$

(6) The following is a non-degenerate bilinear pairing:

$$\widehat{G} \times G \rightarrow S^1, \quad (\pi, x) \mapsto \pi(x).$$

So $\widehat{\widehat{G}}$ is naturally isomorphic to G .

(7) Sending $\theta \in \text{Hom}(G, H)$ to $\widehat{\theta} \in \text{Hom}(\widehat{H}, \widehat{G})$ given by

$$\widehat{\theta}(\pi) := \pi \circ \theta$$

defines an exact contravariant functor from the category of finite abelian groups to itself.

(8) There exists a natural isomorphism $\widehat{G_1 \oplus G_2} \simeq \widehat{G_1} \oplus \widehat{G_2}$.

(9) There exists an isomorphism $\widehat{G} \simeq G$.

Proof. By Theorem 23, $|G| = \sum_{\pi \in \widehat{G}} (\deg \pi)^2$ and by Corollary 24, $|\widehat{G}| = \kappa(G) = |G|$ as G is abelian. Therefore, $\deg \pi = 1$ for all $\pi \in \widehat{G}$.

Parts (3), (4), and (5) are special cases of Theorem 23.

To prove Part (6), let δ_{x_0} be the dirac function at x_0 . Then $\widehat{\delta_{x_0}}(\pi) = \pi(x_0)$. Hence, if $(\pi, x_0) = 1$ for all $\pi \in \widehat{G}$, then by Part (5), $\delta_{x_0} = \delta_1$, so $x_0 = 1$.

Parts (7) and (8) are left as an exercise.

For Part (9), use the classification of finite abelian groups, Part (8), and the following observation: for every $y + \mathbb{Z} \in (\frac{1}{m}\mathbb{Z})/\mathbb{Z}$, the following defines an element of $\widehat{\mathbb{Z}/m\mathbb{Z}}$

$$\mathbb{Z}/m\mathbb{Z} \rightarrow S^1, \quad x + m\mathbb{Z} \mapsto e^{2\pi i xy}.$$

(Note that $\square$)

Theorem 26 (Margulis-Gabber-Galil expanders). *Let G_m be a graph with m^2 vertices that are labelled by elements of $(\mathbb{Z}/m\mathbb{Z})^2$. The vertex (x, y) is connected to the following 8 vertices:*

$$(x \pm y, y), (x, y \pm x), (x \pm 1, y), \quad \text{and} \quad (x, y \pm 1).$$

Then $\{G_m\}_{m=1}^\infty$ is a family of expanders.

Proof. Note that G_m is a Schreier graph of the special affine group

$$\Gamma := \mathbb{Z}^2 \rtimes \text{SL}_2(\mathbb{Z}) \curvearrowright (\mathbb{Z}/m\mathbb{Z})^2,$$

using the following generating set

$$S := \{(0, u_1^{\pm 1}), (0, u_2^{\pm 1}), (\pm e_1, I), (\pm e_2, I)\},$$

where $u_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $u_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Hence, to show the claim, we have to find a uniform lower bound for $\kappa(\mu_S; \pi_m)$ where $\pi_m: \Gamma \rightarrow \mathcal{U}(L^2((\mathbb{Z}/m\mathbb{Z})^2)^\circ)$: by Proposition 1

$$\lambda_2(G_m) = \|\pi_m(\mu_S)\|_{\text{op}} \leq \sqrt{1 - \frac{\kappa(\mu_S; \pi_m)^2}{|S|^2}},$$

where

$$\kappa(\mu_S; \pi_m) := \max_{s \in S, \|f\|_2=1} \|\pi_m(s)(f) - f\|_2.$$

We use Fourier analysis on the finite abelian group $A_m := (\mathbb{Z}/m\mathbb{Z})^2$, and write what it means to have $\kappa(\mu_S; \pi_m) \leq \varepsilon$.

Lemma 27. Suppose A is a finite abelian group, $a \in A$, and $f \in L^2(A)$. Then, for every $\pi \in \widehat{A}$,

$$\widehat{\lambda(a)(f)}(\pi) = \pi(a)\widehat{f}(\pi).$$

Proof. Note that $\widehat{\lambda(a)(f)}(\pi) = \sum_{x \in A} f(a^{-1}x)\pi(x) = \sum_{x' \in X} f(x')\pi(ax') = \pi(a)\widehat{f}(\pi)$. $\square$

Let $\overline{A}_m := ((\frac{1}{m}\mathbb{Z})/\mathbb{Z})^2 \subseteq (\mathbb{Q}/\mathbb{Z})^2 \subseteq (\mathbb{R}/\mathbb{Z})^2$. For $\mathbf{v} \in \overline{A}_m$, let

$$\pi_{\mathbf{v}}: A_m \rightarrow S^1, \quad \pi_{\mathbf{v}}(\mathbf{a}) := e(\langle \mathbf{v} | \mathbf{a} \rangle),$$

where $e: \mathbb{R}/\mathbb{Z} \rightarrow S^1$, $e(x) := e^{2\pi i x}$ and $\langle \mathbf{v} | \mathbf{a} \rangle := v_1 a_1 + v_2 a_2 \in \mathbb{R}/\mathbb{Z}$. Recall that

$$(\frac{1}{m}\mathbb{Z})/\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}, \quad (y + \mathbb{Z}, x + m\mathbb{Z}) \mapsto xy + \mathbb{Z}$$

is well-defined. Then,

$$\overline{A}_m \rightarrow \widehat{A}_m, \quad \mathbf{v} \mapsto \pi_{\mathbf{v}}$$

is an isomorphism, and using this isomorphism, we identify these two groups. Therefore, for $f \in L^2(A_m)$ with $\|f\|_2 = 1$, we may view $\widehat{f}$ as a function on $\overline{A}_m$, and by the Fourier analysis on finite abelian groups,

$$1 = \|f\|_2^2 = \sum_{\mathbf{v} \in \overline{A}_m} \frac{|\widehat{f}(\mathbf{v})|^2}{|\overline{A}_m|}.$$

Hence, we get a probability measure ν_f on the torus $(\mathbb{R}/\mathbb{Z})^2$ whose support is a subset of $\overline{A}_m$ and it is given by the distribution function

$$\mathbf{v} \mapsto \frac{|\widehat{f}(\mathbf{v})|^2}{|\overline{A}_m|}.$$

Assuming $\kappa(\mu_S; \pi_m) \leq \varepsilon$, there exists $f \in L^2(A_m)^\circ$ such that $\|f\|_2 = 1$ and for all $s \in S$,

$$\|\pi_m(s)(f) - f\|_2 \leq \varepsilon.$$

Note that for every finite abelian group A ,

$$f \in L^2(A)^\circ \quad \text{if and only if} \quad \widehat{f}(1) = 0,$$

where $1 \in \widehat{A}$ is the trivial representation.

By Lemma 27 and $\pm e_i \in S$, we deduce that

$$\frac{1}{|\overline{A}_m|} \sum_{\mathbf{v} \in \overline{A}_m} |\pi_{\mathbf{v}}(\pm e_i) - 1|^2 |\widehat{f}(\mathbf{v})|^2 \leq \varepsilon^2.$$

Hence,

$$\int_{(\mathbb{R}/\mathbb{Z})^2} (2 - 2\cos(2\pi t_i)) d\nu_f(t) \leq \varepsilon^2.$$

(We will continue in the next lecture.) $\square$

8. LECTURE 8: MARGULIS-GABBER-GALIL EXPANDERS, II. GOWERS'S QUASI-RANDOMNESS

Let's continue the proof of *Margulis-Gabber-Galil expanders*.

Proof of Theorem 26 (continue). Let's recall that $A_m := (\mathbb{Z}/m\mathbb{Z})^2$, $\bar{A}_m := ((\frac{1}{m}\mathbb{Z})/\mathbb{Z})^2 \subseteq (\mathbb{R}/\mathbb{Z})^2$, and

$$\bar{A}_m \rightarrow \widehat{A}_m, \quad \mathbf{v} \mapsto \pi_{\mathbf{v}}, \quad \text{where} \quad \pi_{\mathbf{v}}(\mathbf{x}) := e(\langle \mathbf{v} | \mathbf{x} \rangle)$$

is an isomorphism, where $e(x) := e^{2\pi i x}$ and $\langle v | \mathbf{x} \rangle = v_1 x_1 + v_2 x_2$. For $f \in L^2(A_m)$ and $\mathbf{v} \in \bar{A}_m$, we let $\widehat{f}(\mathbf{v}) := \widehat{f}(\pi_{\mathbf{v}})$. By Lemma 27, for every $\mathbf{a} \in A_m$,

$$(15) \quad \widehat{(\mathbf{a} \cdot f)}(\mathbf{v}) = e(\langle \mathbf{v} | \mathbf{a} \rangle) \widehat{f}(\mathbf{v}).$$

For $s \in \text{SL}_2(\mathbb{Z})$ and $f \in L^2(A_m)$,

$$(16) \quad \begin{aligned} \widehat{s \cdot f}(\mathbf{v}) &= \sum_{\mathbf{x} \in A_m} (s \cdot f)(\mathbf{x}) e(\langle \mathbf{v} | \mathbf{x} \rangle) = \sum_{\mathbf{x} \in A_m} f(s^{-1} \mathbf{x}) e(\langle \mathbf{v} | \mathbf{x} \rangle) \\ &= \sum_{\mathbf{x}' \in A_m} f(\mathbf{x}') e(\langle \mathbf{v} | s \mathbf{x}' \rangle) = \widehat{f}(s^t \mathbf{v}). \end{aligned}$$

Equation (16) can be also written as

$$\widehat{s \cdot f} = (s^t)^{-1} \cdot \widehat{f}.$$

Assuming that $\kappa(\mu_S; \pi_m) \leq \varepsilon$, there exists $f \in L^2(A_m)^\circ$ such that

$$(17) \quad \|f\|_2 = 1, \quad \|e_i \cdot f - f\|_2 \leq \varepsilon, \quad \text{and} \quad \|u_i^{\pm 1} \cdot f - f\|_2 \leq \varepsilon.$$

Hence, by Fourier analysis on finite abelian groups,

$$(18) \quad \begin{aligned} \widehat{f}(0) &= 0, \\ \frac{1}{|A_m|} \sum_{\mathbf{v} \in \bar{A}_m} |\widehat{f}(\mathbf{v})|^2 &= 1, \\ \frac{1}{|A_m|} \sum_{\mathbf{v} \in \bar{A}_m} |e(\langle \mathbf{v} | e_i \rangle) - 1|^2 |\widehat{f}(\mathbf{v})|^2 &\leq \varepsilon^2, \\ \frac{1}{|A_m|} \sum_{\mathbf{v} \in \bar{A}_m} |\widehat{f}((u_i^t)^{\pm 1} \mathbf{v}) - \widehat{f}(\mathbf{v})|^2 &\leq \varepsilon^2. \end{aligned}$$

Let $\nu_f(\mathbf{t}) := \frac{|\widehat{f}(\mathbf{t})|^2}{|A_m|}$, so

$$\int_T g(\mathbf{t}) d\nu_f := \sum_{\mathbf{t} \in \bar{A}_m} g(\mathbf{t}) \nu_f(\mathbf{t})$$

defines a probability measure on T . We have $\nu_f(0) = 0$ as $f \perp 1_{A_m}$, and

$$(19) \quad \begin{aligned} \int_T (2 - 2 \cos(2\pi t_i)) d\nu_f &\leq \varepsilon^2, \\ \sum_{\mathbf{t} \in \bar{A}_m} |\sqrt{\nu_f(s|\mathbf{t})} - \sqrt{\nu_f(\mathbf{t})}|^2 &\leq \varepsilon^2, \end{aligned}$$

where $s = (u_i^t)^{\pm 1}$. We take $[-1/2, 1/2]^2$ as a set of coset representatives for $\mathbb{R}^2/\mathbb{Z}^2$, and note that if $1/4 \leq |t_i| \leq 1/2$, then $2 - 2 \cos(2\pi t_i) \geq 2 - \sqrt{2}$. Hence,

$$(20) \quad \nu_f(T \setminus [-1/4, 1/4]^2) \leq \frac{2}{(2 - \sqrt{2})^2} \varepsilon^2 = (3 + 2\sqrt{2}) \varepsilon^2.$$

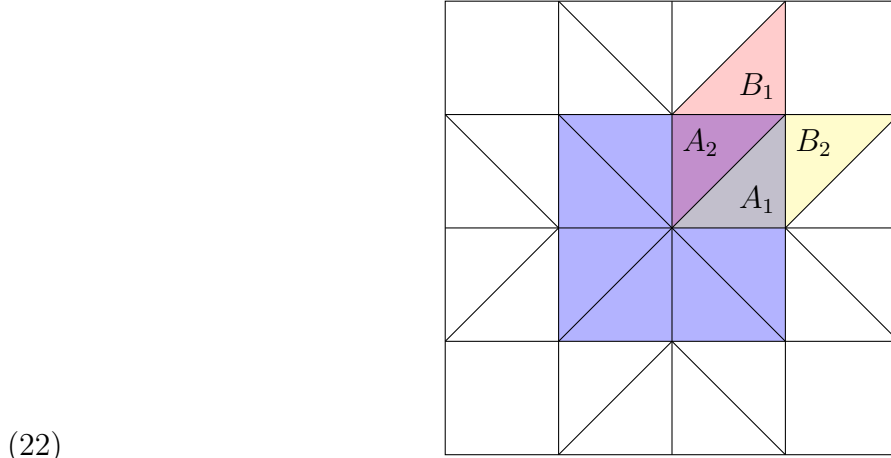
For every function g on T and $s = (u_i^t)^{\pm 1}$, we have

$$\begin{aligned}
\left| \sum_{\mathbf{t} \in \bar{A}_m} g(\mathbf{t}) v_f(s|\mathbf{t}) - \sum_{\mathbf{t} \in \bar{A}_m} g(\mathbf{t}) v_f(\mathbf{t}) \right| &\leq \sum_{\mathbf{t} \in \bar{A}_m} |g(\mathbf{t})| |v_f(s|\mathbf{t}) - v_f(\mathbf{t})| \\
&\leq \sum_{\mathbf{t} \in \bar{A}_m} |g(\mathbf{t})| \left| \sqrt{v_f(s|\mathbf{t})} - \sqrt{v_f(\mathbf{t})} \right| \left| \sqrt{v_f(s|\mathbf{t})} + \sqrt{v_f(\mathbf{t})} \right| \\
&\leq \varepsilon \left(\sum_{\mathbf{t} \in \bar{A}_m} |g(\mathbf{t})|^2 \left| \sqrt{v_f(s|\mathbf{t})} + \sqrt{v_f(\mathbf{t})} \right|^2 \right)^{1/2} \\
&\leq \sqrt{2\varepsilon} \left(\sum_{\mathbf{t} \in \bar{A}_m} |g(\mathbf{t})|^2 v_f(s|\mathbf{t}) + \sum_{\mathbf{t} \in \bar{A}_m} |g(\mathbf{t})|^2 v_f(\mathbf{t}) \right)^{1/2}.
\end{aligned}$$

Applying the above inequality for g equals to the characteristic function 1_B of a subset B , we obtain

$$(21) \quad |\nu_f(s^{-1}B) - \nu_f(B)| \leq \sqrt{2\varepsilon} \sqrt{\nu_f(s^{-1}B) + \nu_f(B)} \leq 2\varepsilon,$$

for $s = (u_i^t)^{\pm 1}$. By (21), we deduce that in the figure given in (22),



we have

$$|\nu_f(A_1 \cup A_2) - \nu_f(A_1 \cup B_2)| \leq 2\varepsilon \quad \text{and} \quad |\nu_f(A_1 \cup A_2) - \nu_f(A_2 \cup B_1)| \leq 2\varepsilon.$$

Therefore,

$$\nu_f(A_1 \cup A_2) \leq 4\varepsilon + \nu_f(B_1 \cup B_2)$$

Hence, by symmetry and (20), we deduce

$$1 \leq 16\varepsilon + (3 + 2\sqrt{2})\varepsilon^2.$$

Therefore, we get a contradiction if $\varepsilon < 0.06$. □

The above Figure is an approach by Martin Kassabov based on earlier works by Marc Burger and Yehuda Shalom. One can be much less generous and find a much better bound for $\kappa(\mu_S; \pi_m)$. In fact, Marc Burger combined the above approach with the spectral theory of self-adjoint operators and proved the same result for every unitary representation of $\mathbb{Z}^2 \rtimes \text{SL}_2(\mathbb{Z})$ that does not have a non-zero $\mathbb{Z}^2$ -invariant vector. This property is known as *relative property (T)* and its qualitative form was proved by Kazhdan and used by Margulis to construct these expanders.

More conceptually, the small square represents *low frequencies* in the Fourier space and the outside the small square represents *high frequencies*.

How can we think about low and high frequencies in a non-abelian setting? In the non-abelian setting, we use the degree of a representation to quantify how noisy the corresponding matrix coefficients are: let T be a threshold parameter, and define

$$H(f; T) := \frac{1}{|G|} \sum_{\pi \in \widehat{G}, \deg \pi \geq T} \deg \pi \|\widehat{f}(\pi)\|_{\text{HS}}^2.$$

Lemma 28 (Mixing in high frequencies). *For every $f, f_1, f_2 \in L^2(G)$ and $T \in \mathbb{R}^+$, the following statements hold:*

- (1) $H(f; T) \leq \|f\|_2^2$.
- (2) $H(f_1 * f_2; T) \leq \frac{|G|}{T} H(f_1; T) H(f_2; T)$.

Proof. The first part follows from the Parseval formula for finite groups. For the second part, we have

$$\begin{aligned} H(f_1 * f_2; T) &= \frac{1}{|G|} \sum_{\pi \in \widehat{G}, \deg \pi \geq T} \deg \pi \|\widehat{f_1 * f_2}(\pi)\|_{\text{HS}}^2 = \frac{1}{|G|} \sum_{\pi \in \widehat{G}, \deg \pi \geq T} \deg \pi \|\widehat{f_1}(\pi) \widehat{f_2}(\pi)\|_{\text{HS}}^2 \\ &\leq \frac{1}{|G|} \sum_{\pi \in \widehat{G}, \deg \pi \geq T} \deg \pi \|\widehat{f_1}(\pi)\|_{\text{HS}}^2 \|\widehat{f_2}(\pi)\|_{\text{HS}}^2 \\ &\leq \frac{|G|}{T} \left(\frac{1}{|G|} \sum_{\pi \in \widehat{G}, \deg \pi \geq T} \deg \pi \|\widehat{f_1}(\pi)\|_{\text{HS}}^2 \right) \left(\frac{1}{|G|} \sum_{\pi \in \widehat{G}, \deg \pi \geq T} \deg \pi \|\widehat{f_2}(\pi)\|_{\text{HS}}^2 \right) \\ &= \frac{|G|}{T} H(f_1; T) H(f_2; T). \end{aligned}$$

This finishes the proof. $\square$

This takes us to the notion of *quasi-random groups* introduced by Gowers. For $c > 0$, we say a finite group G is c -quasi-random if

$$\deg \pi \geq |G|^c$$

for all $\pi \in \widehat{G} \setminus \{1\}$. We obtain the following corollary of Lemma 28.

Proposition 29. *Suppose G is a finite c -quasi-random group. For $f \in L^2(G)$, let*

$$T_f: L^2(G) \rightarrow L^2(G), \quad T_f(g) := f * g.$$

Then $T_f(L^2(G)^\circ) \subseteq L^2(G)^\circ$, and

$$\|T_f|_{L^2(G)^\circ}\|_{\text{op}} \leq |G|^{1/2-c/2} \|f\|_2.$$

Proof. Note that $\sum_{x \in G} (f * g)(x) = \sum_{x_1, x_2 \in G} f(x_1)g(x_2) = (\sum_x f(x))(\sum_x g(x))$. Therefore

$$T_f(L^2(G)^\circ) \subseteq L^2(G)^\circ.$$

Hence, for $g \in L^2(G)^\circ$, we have

$$\widehat{g}(1) = \widehat{f * g}(1) = 0,$$

so by the definition of c -quasi-random groups,

$$\|g\|_2^2 = H(g; |G|^c) \quad \text{and} \quad \|f * g\|_2^2 = H(f * g; |G|^c).$$

Therefore, by Lemma 28, we deduce

$$\|T_f(g)\|_2^2 \leq \frac{|G|}{|G|^c} H(f; |G|^c) H(g; |G|^c) \leq (|G|^{1-c} \|f\|_2^2) \|g\|_2^2,$$

for every $g \in L^2(G)^\circ$. The claim follows. $\square$

9. LECTURE 9: QUASI-RANDOMNESS AND PRODUCT OF LARGE SETS. ENTROPIES AND THE BOURGAIN-GAMBURD METHOD.

In the previous lecture, we proved that if $\deg \pi \geq |G|^c$ for all non-trivial $\pi \in \widehat{G}$, then

$$\|T_f|_{L^2(G)^\circ}\|_{\text{op}} \leq |G|^{1/2-c/2} \|f\|_2.$$

Before we formulate some of the consequences of the above inequality, we define *entropy* of a measure (or a random-variable). Suppose μ is a probability measure with finite support. Then the α -entropy of μ is

$$H_\alpha(\mu) := \begin{cases} \frac{\alpha}{1-\alpha} \log \|\mu\|_\alpha & \text{if } \alpha \neq 0, 1, \infty \\ -\sum_x \mu(x) \log \mu(x) & \text{if } \alpha = 1 \\ \log |\text{supp}(\mu)| & \text{if } \alpha = 0 \\ -\log \|\mu\|_\infty & \text{if } \alpha = \infty. \end{cases}$$

The 1-entropy is known as the Shannon entropy. We refer to the 2-entropy as the Rényi entropy of μ . Here are some of its basic properties:

- Lemma 30.** (1) For a fixed μ , $\alpha \mapsto H_\alpha(\mu)$ is continuous and decreasing.
 (2) Suppose $|\text{Supp}(\mu)| = n$. Then $H_\alpha(\mu) \leq \log n$ and for $\alpha > 0$ equality holds if and only if μ is the probability counting measure on $\text{Supp}(\mu)$.
 (3) $H_\alpha(\mu_1 * \mu_2) \geq \max\{H_\alpha(\mu_1), H_\alpha(\mu_2)\}$ for every pair (μ_1, μ_2) of probability measures with finite support on a group G .

Proof. Note that $\frac{d}{d\alpha} \log(\sum_x \mu(x)^\alpha) = \sum_x \nu_\alpha(x) \log \mu(x)$, where

$$\nu_\alpha(x) := \frac{\mu(x)^\alpha}{\sum_y \mu(y)^\alpha}.$$

By the l'hôpital rule

$$\lim_{\alpha \rightarrow 1} H_\alpha(\mu) = -\lim_{\alpha \rightarrow 1} \sum_x \nu_\alpha(x) \log \mu(x) = H_1(\mu).$$

Note that

$$\lim_{\alpha \rightarrow 0} H_\alpha(\mu) = \lim_{\alpha \rightarrow 0} \log \sum_x \mu(x)^\alpha = \log |\text{Supp}(\mu)| = H_0(\mu).$$

By another application of the l'hôpital rule,

$$\lim_{\alpha \rightarrow \infty} H_\alpha(\mu) = -\lim_{\alpha \rightarrow \infty} \sum_x \nu_\alpha(x) \log \mu(x).$$

Note that $\lim_{\alpha \rightarrow \infty} \nu_\alpha(x) = 0$ if $\mu(x) < \|\mu\|_\infty$ and 1 otherwise. Therefore,

$$\lim_{\alpha \rightarrow \infty} H_\alpha(\mu) = H_\infty(\mu).$$

Observe that

$$\begin{aligned} \frac{dH_\alpha}{d\alpha} &= \frac{1}{(1-\alpha)^2} \left(\left(\sum_x \nu_\alpha(x) \log \mu(x) \right) (1-\alpha) + \log \sum_x \mu(x)^\alpha \right) \\ &= \frac{-1}{(1-\alpha)^2} \left(\sum_x \nu_\alpha(x) \log \frac{\mu(x)^{\alpha-1}}{\sum_y \mu(y)^\alpha} \right) \\ &= \frac{1}{(1-\alpha)^2} \left(\sum_x \nu_\alpha(x) \log \frac{\mu(x)}{\nu_\alpha(x)} \right) \end{aligned}$$

By the Jensen inequality,

$$\sum_x \nu_\alpha(x) \log \frac{\mu(x)}{\nu_\alpha(x)} \leq \log \sum_x \mu(x) = 0$$

Therefore, H_α is decreasing on α .

By Part (1), $H_\alpha(\mu) \leq H_0(\mu) = \log |\text{Supp}(\mu)|$. Suppose $H_\alpha(\mu) = \log n$. Then for every $\alpha' \in [0, \alpha]$, $\log n = H_{\alpha'}(\mu)$. Deduce that $\alpha' \mapsto \sum_{i=1}^n (n\mu(x_i))^{\alpha'}$ is a constant function. Looking at its derivative, deduce that $n\mu(x_i) = 1$ for all i .

For $\alpha > 1$, $\|\mu_1 * \mu_2\|_\alpha = \|\sum_x \mu_1(x)\lambda(x)(\mu_2)\|_\alpha \leq \sum_x \mu_1(x)\|\lambda(x)(\mu_2)\|_\alpha = \|\mu_2\|_\alpha$. Therefore,

$$H_\alpha(\mu_1 * \mu_2) = \frac{\alpha}{1-\alpha} \log \|\mu_1 * \mu_2\|_\alpha \geq \frac{\alpha}{1-\alpha} \log \|\mu_2\|_\alpha = H_\alpha(\mu_2).$$

By symmetry, we also have $H_\alpha(\mu_1 * \mu_2) \geq H_\alpha(\mu_1)$. For $\alpha < 1$, using concativity of $x \mapsto x^\alpha$, we deduce

$$(\mu_1 * \mu_2)(y)^\alpha \geq \sum_x \mu_1(x)\mu_2(x^{-1}y)^\alpha.$$

Therefore, for $0 < \alpha < 1$,

$$\sum_y (\mu_1 * \mu_2)(y)^\alpha \geq \sum_{y,x} \mu_1(x)\mu_2(x^{-1}y)^\alpha = \sum_y \mu_2(y)^\alpha.$$

This implies that $H_\alpha(\mu_1 * \mu_2) \geq H_\alpha(\mu_2)$, which finishes the proof. $\square$

By Lemma 30, we see that entropy gets larger as we do random-walk on a group, and $H_\alpha(\mu) \leq \log n$ if the support of μ has n elements.

Proposition 31. *Let G be a c -quasi-random group. Suppose X_1, X_2 , and X_3 are three independent random-variables with values in G . Suppose*

$$\frac{H_0(X_1) + H_0(X_2) + H_0(X_3)}{3} > \left(1 - \frac{c}{3}\right) \log |G|.$$

Then $H_0(X_1X_2X_3) = \log |G|$.

This proposition is often written in the following form:

Proposition 32 (due to Gowers). *Suppose G is a c -quasi-random group, $A_1, A_2, A_3 \subseteq G$, and*

$$|A_1||A_2||A_3| > |G|^{3-c}.$$

Then $G = A_1A_2A_3$.

Proof. Let μ_i be the probability counting measure on A_i . Then, for every $x \in G$, by the Cauchy-Scharwz inequality

$$\begin{aligned} \left| \mu_1 * \mu_2 * \mu_3(x) - \frac{1}{|G|} \right| &= \left| \mu_1 * \mu_2 * \mu_3(x) - m_G(x) \right| \\ &= \left| \mu_1 * (\mu_2 * (\mu_3 - m_G)) \right| \\ &\leq \|\mu_1\|_2 \|\mu_2 * (\mu_3 - m_G)\|_2. \end{aligned}$$

Hence, by Proposition 29, we deduce

$$\begin{aligned} \left| \mu_1 * \mu_2 * \mu_3(x) - \frac{1}{|G|} \right| &\leq |G|^{1/2-c/2} \|\mu_1\|_2 \|\mu_2\|_2 \|\mu_3\|_2 \\ &= |G|^{1/2-c/2} |A_1|^{-1/2} |A_2|^{-1/2} |A_3|^{-1/2} \\ &< |G|^{1/2-c/2} |G|^{-3/2-c/2} = \frac{1}{|G|}. \end{aligned}$$

Hence, $\mu_1 * \mu_2 * \mu_3(x) > 0$ for all $x \in G$. This means that $x \in A_1A_2A_3$, so $A_1A_2A_3 = G$. $\square$

A consequence of Proposition 32 is that if $A \subseteq G$ and $|A| > |G|^{1-c/3}$, then A is not *product free*: that means $A \cdot A \cap A \neq \emptyset$.

Note that by Lemma 30, for a random-walk $\{X^{(\ell)}\}_\ell$ on a finite group G , the following is an increasing sequence

$$\ell \mapsto \frac{H_\alpha(X^{(\ell)})}{\log |G|}$$

with values in $(0, 1]$. Proposition 31 implies that

$$\{\ell \in \mathbb{N} \mid \frac{H_0(X^{(\ell)})}{\log |G|} \in (1 - c/3, 1)\}$$

is a finite set. In fact, Proposition 31 implies that

$$|\{\ell \in \mathbb{N} \mid \frac{H_0(X^{(\ell)})}{\log |G|} \in (1 - c/3, 1)\}| \leq 2 \min\{\ell \in \mathbb{N} \mid \frac{H_0(X^{(\ell)})}{\log |G|} \in (1 - c/3, 1)\}.$$

Another important consequence of Proposition 29 is the following bound on the Lyapunov exponent of a symmetric random-variable.

Proposition 33. *Suppose G is a c -quasi-random group and X is a symmetric random-variable with values in G . Suppose for $\ell < C \log |G|$,*

$$H_2(X^{(\ell)}) > \left(1 - \frac{c}{2}\right) \log |G|$$

Then

$$\mathcal{L}(X) > \frac{c}{4C}.$$

Proof. Let μ be the probability law of X . Since X is symmetric, $\lambda(X^{(\ell)}) = \lambda(\mu)^\ell$. So,

$$\mathcal{L}(X^{(\ell)}) = \ell \mathcal{L}(X).$$

By Proposition 29,

$$\|T_{\mu^{(\ell)}}|_{L^2(G)^\circ}\|_{\text{op}} \leq |G|^{1/2-c/2} \|\mu^{(\ell)}\|_2.$$

Hence,

$$\begin{aligned} \mathcal{L}(X^{(\ell)}) &= -\log \|T_{\mu^{(\ell)}}|_{L^2(G)^\circ}\|_{\text{op}} \\ &\geq \left(\frac{c}{2} - \frac{1}{2}\right) - \log \|\mu^{(\ell)}\|_2 \\ &= \left(\frac{c}{2} - \frac{1}{2}\right) \log |G| - \frac{1}{2} H_2(X^{(\ell)}) \\ &> \frac{c}{4} \log |G|. \end{aligned}$$

Therefore,

$$\mathcal{L}(X) > \frac{c \log |G|}{4\ell} > \frac{c}{4C}.$$

□

Proposition 33 gives us a scheme for obtaining a lower bound for the Lyapunov exponent:

Showing that the sequence $\frac{H_2(X^{(\ell)})}{\log |G|}$ reaches to the interval $(c/2, 1]$ in $O(\log |G|)$ -steps.

We proved the mixing inequality using Fourier analysis on finite groups and then used it to obtain the above scheme. Gowers's original proof relied on the *high multiplicity* of the largest eigenvalue. Sarnak and Xue used the same idea in their study automorphic representations. As we will explain later, Bourgain and Gamburd placed the study of the sequence $\frac{H_2(X^{(\ell)})}{\log |G|}$ as the center of their approach to the Lyapunov exponent $\mathcal{L}(X)$. Sarnak, being a chess player, referred to this part as the *endgame* of the whole scheme.

We have proved results on c -quasi-random groups. Next, we provide examples of such groups and mention some of their basic properties.

Proposition 34. *Suppose π is a non-trivial irreducible representation of $\mathrm{SL}_2(\mathbb{F}_p)$. Then*

$$\deg \pi \geq \frac{p-1}{2}.$$

Proof. To study a representation π of $\mathrm{SL}_2(\mathbb{F}_p)$, we often start with its restriction to a *Borel subgroup*: the subgroup B of upper-triangular matrices. Note that $B = DU$ where

$$U := \{u(a) \mid a \in \mathbb{F}_p\} = \langle u(1) \rangle \quad \text{and} \quad D := \{d(b) \mid b \in \mathbb{F}_p^\times\},$$

where $u(a) := \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ and $d(b) := \mathrm{diag}(b, b^{-1})$. Moreover, $U \trianglelefteq B$ with

$$d(b)u(a)d(b)^{-1} = u(b^2a).$$

Since $u(1)^p = 1$, $\pi(u(1))$ is similar to

$$\mathrm{diag}(\zeta_1, \dots, \zeta_{\deg \pi})$$

for some $\zeta_i \in \mathbb{C}$ such that $\zeta_i^p = 1$. For every $b \in \mathbb{F}_p^\times$, $\pi(d(b)u(1)d(b)^{-1}) = \pi(u(b^2)) = \pi(u(1))^{b^2}$ is similar to $\pi(u(1))$. Hence,

$$(23) \quad \mathrm{diag}(\zeta_1, \dots, \zeta_{\deg \pi}) \sim \mathrm{diag}(\zeta_1^{b^2}, \dots, \zeta_{\deg \pi}^{b^2}).$$

(To be continued). □

10. LECTURE 10: QUASI-RANDOM GROUPS. RANDOM-WALK ON FREE GROUPS, I.

Proof of Proposition 34 (continue). By (23), it follows that, for every i ,

$$\{\zeta_i^{b^2} \mid b \in \mathbb{F}_p^\times\} \subseteq \{\zeta_1, \dots, \zeta_{\deg \pi}\}.$$

Note that if $\zeta_i \neq 1$, then it is of order p . So, $\zeta_i^b \neq \zeta_i^{b'}$ if $b \neq b'$ are in $\mathbb{F}_p^\times$. Hence, if $\zeta_i \neq 1$ for some i , then

$$\deg \pi \geq \frac{p-1}{2}.$$

Using the reduced row and column process, one can show that $\mathrm{SL}_2(\mathbb{F}_p)$ is generated by $u(1)$ and its transpose $u(1)^t$. Also note that

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

Hence, $\pi(u(1)) \neq 1$, which implies that $\zeta_i \neq 1$ for some i . This finishes the proof. □

Proposition 35. *For every d , there exists $c := c(d)$ such that for every prime p , $\mathrm{SL}_d(\mathbb{F}_p)$ is c -quasi-random.*

Proof. Note that for every $i < j$, we may embed $\mathrm{SL}_2(\mathbb{F}_p)$ into $\mathrm{SL}_d(\mathbb{F}_p)$ by placing the entries of a 2-by-2 matrix in the (i, i) , (i, j) , (j, i) , and (j, j) entries and filling out the remaining entries as in the identity matrix. Moreover, these copies of $\mathrm{SL}_2(\mathbb{F}_p)$ generate $\mathrm{SL}_d(\mathbb{F}_p)$. Hence, the restriction of a non-trivial representation π to at least one of these subgroups is non-trivial. Therefore, $\deg \pi \geq \frac{p-1}{2}$. Since

$$|\mathrm{SL}_d(\mathbb{F}_p)| = p^{d^2-1} \prod_{i=2}^d (1 - p^{-i})^{-1},$$

we deduce that, for $p \geq 7$,

$$\deg \pi \geq p^{1/2} > |\mathrm{SL}_d(\mathbb{F}_p)|^{(2(d^2-1))^{-1}}.$$

□

Note that the above argument works for every group generated by copies of $\mathrm{SL}_2(\mathbb{F}_p)$ and has order bounded by a fixed power of p :

Proposition 36. *For every prim p , let G_p be a finite group generated by copies of $\mathrm{SL}_2(\mathbb{F}_p)$. Suppose $|G_p| \leq p^C$. Then G_p is $(2C)^{-1}$ -quasi-random for every $p \geq 7$.*

So, one obtains a similar result for a bounded product of all almost simple groups of Lie type over $\mathbb{F}_p$: for a given n_1, n_2, n_3, m_1, m_2 , and m_3 ,

$$\mathrm{SL}_{n_1}(\mathbb{F}_p)^{m_1} \times \mathrm{Sp}_{2n_2}(\mathbb{F}_p)^{m_2} \times \mathrm{Spin}_{n_3}(\mathbb{F}_p)^{m_3}$$

is a c -quasi-random group for some $c := c(n_1, n_2, n_3, m_1, m_2, m_3)$. One can also get an almost inverse of this result:

Proposition 37. *Suppose G is a c -quasi-random group. Then the following statements hold:*

- (1) G is a perfect group.
- (2) If G has a non-trivial action on a finite set X , then $|X| \geq |G|^c$.
- (3) G has at most c^{-1} non-abelian composition factors.

Proof. By the Fourier analysis on finite abelian groups, $\widehat{G/[G, G]}$ consists of group homomorphisms from $G/[G, G]$ to S^1 , and it has the same number of elements as $G/[G, G]$. So, if G is not perfect, then there exists a non-trivial group homomorphism $\chi: G/[G, G] \rightarrow S^1$. Therefore, the composition $\chi \circ \pi$ of χ and the natural quotient map $\pi: G \rightarrow G/[G, G]$ gives us a non-trivial unitary representation of G that is of degree 1. This contradicts the assumption that degree of every non-trivial irreducible representation is at least $|G|^c$.

Also note that if G has a non-trivial action on a finite set X , then we obtain a non-trivial group homomorphism $\rho: G \rightarrow \mathcal{U}(L^2(X))$. Since every representation is completely reducible, ρ is non-trivial, and every irreducible representation is of degree at least $|G|^c$, we deduce that $|X| \geq |G|^c$.

Suppose $G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_m = 1$ is a composition series. For every i , G acts on G_i/G_{i+1} by conjugation. Therefore, by Part (2), either G acts trivially on G_i/G_{i+1} or $|G_i/G_{i+1}| \geq |G|^c$. Note that if G acts trivially on G_i/G_{i+1} , then G_i/G_{i+1} is abelian. Hence,

$$|G| \geq \prod_{G_i/G_{i+1} \text{ not abelian}} |G_i/G_{i+1}| \geq \prod_{G_i/G_{i+1} \text{ not abelian}} |G|^c.$$

So,

$$|\{i \mid G_i/G_{i+1} \text{ not abelian}\}| \leq c^{-1}.$$

□

Bourgain and Gamburd considered the sequence

$$\ell \mapsto \frac{H_2(X^{(\ell)})}{\log |G|}$$

to study $\mathcal{L}(X)$. They did so, by dividing the study of this dynamical system into three regimes: initial (opening), intermediate (middle-game), and final (endgame).

This type of approach is common in dynamical systems. Typically, the system initially grows in a wild and unstructured manner. In the intermediate regime, a new form of structure emerges, which eventually leads to saturation in the final regime.

As we have explained earlier (Proposition 33), Bourgain and Gamburd's goal was to show that $\frac{H_2(X^{(\ell)})}{\log |G|}$ reaches to the threshold $(1 - c/2, 1]$ in logarithmic number of steps. Having Lubotzky's 1-2-3 problem in mind, they noted that $u_-(3)$ and $u_+(3)$ freely generate a subgroup of $\mathrm{SL}_2(\mathbb{Z})$, so

$$H_2(X^{(\ell)} \pmod{p}) = H_2(X^{(\ell)})$$

for $\ell < \log_5 p$. To see this claim, note that for a word $w(a, b)$ of length ℓ in the free group generated by a and b , every entry of $w(u_-(3), u_+(3))$ has absolute value at most 4^ℓ . Hence, for a prime $p > 4^\ell + 1$, $w(u_-(3), u_+(3)) \equiv I \pmod{p}$ if and only if $w(a, b) = 1$. Therefore, for every prime $p > 2^{2\ell+1}$ and distinct words w_1, w_2 of length at most ℓ in the free group generated by a and b ,

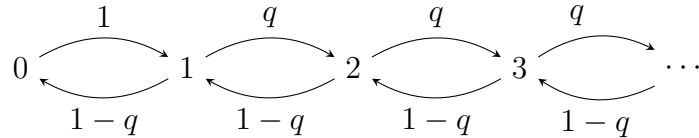
$$w_1(u_-(3), u_+(3)) \not\equiv w_2(u_-(3), u_+(3)) \pmod{p}.$$

This implies that for every $x \in \mathrm{SL}_2(\mathbb{Z})$,

$$\mathbb{P}(X^{(\ell)} = x) = \mathbb{P}(X^{(\ell)} \equiv x \pmod{p})$$

for every prime p and $\ell < 0.5 \log_2 p - 0.5$. Thus, in this case, analyzing the initial regime reduces to studying a random-walk on a free group.

Note that the Cayley graph of a free group on k generators is a $2k$ -regular graph. For vertices v and w in a regular tree, and for a non-negative integer ℓ , let $p^{(\ell)}(v, w)$ denote the probability of starting at v and reaching w in ℓ steps. Note that $p^{(\ell)}(v, w)$ depends only on the distance $d(v, w)$. So, we consider a new Markov process based on the distance from a given fixed vertex v : we label the states with the non-negative integers. The transition probability from i to $i + 1$ is $q := \frac{2k-1}{2k}$ if $i > 0$ and 1 if $i = 0$, and the transition probability from i to $i - 1$ is $1 - q$.



Let $p(\ell) := p^{(\ell)}(v, v)$. Then $p(\ell)$ is the probability of returning to 0 after ℓ steps in the above Markov process. Let $p^*(\ell)$ be the probability of return to zero for the first time after ℓ -steps. Define the generating functions

$$F(z) := \sum_{\ell=0}^{\infty} p(\ell) z^\ell, \quad \text{and} \quad F^*(z) := \sum_{\ell=0}^{\infty} p^*(\ell) z^\ell.$$

Note that

$$p(\ell) = \sum_{\ell=\ell_1+\dots+\ell_k} p^*(\ell_1) \cdots p^*(\ell_k),$$

so

$$F(z) = 1 + F^*(z) + F^*(z)^2 + \cdots = \frac{1}{1 - F^*(z)}.$$

Take a path $c_0 = 0, c_1, \dots, c_\ell = 0$ of length ℓ that starts from 0 and returns to 0 for the first time at the end. Hence,

$$c_{i+1} = c_i \pm 1, \quad c_i > 0 \quad \text{if} \quad 0 < i < \ell, \quad \text{and} \quad 2|\ell.$$

Every such path has $\ell/2$ moves to the left with exactly one such move starting from 0, and $\ell/2$ moves to the right. Hence, the probability of taking such a path is

$$q^{\ell/2-1}(1-q)^{\ell/2}.$$

The number of such paths equals the number of paths from $(1, 1)$ to $(\ell-1, 1)$ using steps $(1, 1)$ or $(1, -1)$, without ever going below the line $y = 1$. This number is the $(\ell/2 - 1)$ -th *Catalan number*. So

$$F^*(z) = 1 + \sum_{\ell'=1}^{\infty} q^{\ell'-1}(1-q)^{\ell'} C_{\ell'-1} z^{2\ell'}.$$

(To be continued!)

11. LECTURE 11: RANDOM-WALK ON FREE GROUPS, II. BOURGAIN-GAMBURD METHOD, I.

We continue our investigation of a random-walk on a free group generated by k elements. Recall that

$$F^*(z) = 1 + (1-q)z^2 \sum_{\ell=0}^{\infty} C_\ell (q(1-q)z^2)^\ell,$$

where C_ℓ is the ℓ -th Catalan number. Let $G(z)$ be the generating function of the Catalan sequence: $G(z) := \sum_{\ell=0}^{\infty} C_\ell z^\ell$. As before, let C_ℓ^* denote the number of Catalan paths that return to the x -axis for the first time after 2ℓ steps, and $G^*(z)$ denote the corresponding generating function. Then

$$C_\ell = \sum_{\ell_1 + \dots + \ell_k = \ell} \prod_{i=1}^k C_{\ell_i}^* \quad \text{and} \quad C_\ell^* = C_{\ell-1}.$$

Hence,

$$G(z) = \frac{1}{1 - G^*(z)} \quad \text{and} \quad G^*(z) = zG(z).$$

Therefore,

$$zG(z)^2 - G(z) + 1 = 0 \quad \text{which implies} \quad G(z) = \frac{-1 - \sqrt{1 - 4z}}{2z}.$$

Note that the other solution is not acceptable because of sign of its coefficients. This implies that

$$F^*(z) = 1 + (1-q)z^2 \left(\frac{-1 - \sqrt{1 - 4q(1-q)z^2}}{2q(1-q)z^2} \right) = 1 + \frac{-1 - \sqrt{1 - 4q(1-q)z^2}}{2q},$$

and so the radius of convergence of $F^*(z)$ and $F(z)$ is

$$\limsup_{\ell \rightarrow \infty} p(\ell)^{1/\ell} = 2\sqrt{q(1-q)} = \frac{\sqrt{2k-1}}{k}.$$

Note that $p(\ell_1 + \ell_2) \geq p(\ell_1)p(\ell_2)$, so $\ell \mapsto -\log p(\ell)$ is a subadditive function.

Lemma 38 (Fekete's subadditive lemma). *Suppose $\{a_n\}_{n=0}^{\infty}$ is a subadditive sequence of non-negative numbers; that is*

$$a_{n+m} \leq a_n + a_m.$$

Then $\lim_{n \rightarrow \infty} \frac{a_n}{n}$ exists and equals $\inf_n \frac{a_n}{n}$.

Proof. Let $L := \limsup_{n \rightarrow \infty} \frac{a_n}{n}$. So there exist $n_1 < n_2 < \dots$ such that $\lim_{i \rightarrow \infty} \frac{a_{n_i}}{n_i} = L$. For every $m \in \mathbb{N}$, after passing to a subsequence of n_i 's, if needed, we may assume that $n_i = mq_i + r$. Then, for all i ,

$$\frac{a_{n_i}}{n_i} \leq \frac{q_i a_m + a_r}{q_i m + r},$$

which implies

$$L \leq \frac{a_m}{m}$$

for all m . Thus $L \leq \inf_m \frac{a_m}{m}$. This finishes the proof. $\square$

By Lemma 38, we deduce that

$$(24) \quad \lim_{\ell \rightarrow \infty} p(\ell)^{1/\ell} = \sup_{\ell} p(\ell)^{1/\ell} = \frac{\sqrt{2k-1}}{k}; \quad \text{in particular} \quad p(\ell) \leq \left(\frac{\sqrt{2k-1}}{k} \right)^\ell.$$

Note that

$$p(2\ell) = p^{(2\ell)}(v, v) = \sum_w p^{(\ell)}(v, w)^2, \quad \text{so} \quad \sum_w p^{(\ell)}(v, w)^2 \leq \left(\frac{\sqrt{2k-1}}{k} \right)^{2\ell},$$

which implies

$$p^{(\ell)}(v, w) \leq \left(\frac{\sqrt{2k-1}}{k} \right)^\ell$$

for all v, w and ℓ . Hence,

$$\frac{H_\infty(X_\ell)}{\ell} \geq \log\left(\frac{k}{\sqrt{2k-1}}\right).$$

Before we continue with the study of the Bourgain-Gamburd method, let's point out the following corollary of (24):

Theorem 39 (due to Alon and Boppana). *Suppose $\{G_i\}_{i=1}^\infty$ is a family of $2k$ -regular expander graphs. Then*

$$\sup_i \lambda(G_i) \geq \frac{\sqrt{2k-1}}{k}.$$

Proof. Note that for every i , there exists a covering map $\pi_i: T_{2k} \rightarrow G_i$ where T_{2k} is a $2k$ -regular tree. A random-walk on G_i may be viewed as the *shadow* of a random-walk on T_{2k} . Let X_ℓ be the random-variable denoting an ℓ -step random-walk on T_{2k} which starts at the vertex v_0 , and $\bar{X}_\ell := \pi_i(X_\ell)$. Let $n_i := |V(G_i)|$. Then, by (7),

$$\frac{1}{n_i} + \lambda(G_i)^\ell \geq \mathbb{P}(\bar{X}_\ell = \pi_i(v_0)) \geq p(\ell).$$

Therefore,

$$p(\ell)^{1/\ell} \leq (\lambda(G_i)^\ell + 1/n_i)^{1/\ell} \leq \lambda(G_i) + (1/n_i)^{1/\ell}$$

Let $\ell_i := \lfloor \sqrt{\log n_i} \rfloor$. Then

$$\log(1/n_i)^{1/\ell_i} = \frac{-\log n_i}{\lfloor \sqrt{\log n_i} \rfloor} \rightarrow -\infty \quad \text{as} \quad i \rightarrow \infty.$$

Hence

$$\frac{\sqrt{2k-1}}{k} = \sup_i p(\ell_i)^{1/\ell_i} \leq \sup_i \lambda(G_i).$$

This completes the proof. $\square$

A k -regular finite graph G is called *Ramanujan* if

$$\lambda(G) \leq \frac{2\sqrt{k-1}}{k}.$$

Lubotzky-Philip-Sarnak and, independently, Margulis proved the existence of $(p+1)$ -regular Ramanujan graphs, where p is prime. Morgenstern proved the existence of (p^n+1) -regular Ramanujan graphs. Existence of k -regular Ramanujan graphs for arbitrary k is still open. One can define *bi-partite Ramanujan graphs* and Marcus, Spielman, and Srivastava showed the existence of such graphs.

Let's go back to the Bourgain-Gamburd method, and focus on Lubotzky's 1-2-3 problem: analysis of random-walk on $\text{Cay}(\text{SL}_2(\mathbb{F}_p), S)$, where

$$S := \left\{ \begin{pmatrix} 1 & \pm 3 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \pm 3 & 1 \end{pmatrix} \right\}.$$

So far, we have proved that for $\ell_0 = c \log p$,

$$\frac{H_2(\overline{X}_{\ell_0})}{\ell_0} \geq \frac{H_\infty(\overline{X}_{\ell_0})}{\ell_0} \geq \log(2/\sqrt{3}).$$

Hence, we have gained an *initial entropy*, and we are done with the *opening* and the *endgame*: recall we have proved that it suffices to show the existence of ℓ with the following properties:

$$\ell \leq C \log |\text{SL}_2(\mathbb{F}_p)| \quad \text{and} \quad H_2(\overline{X}_\ell) \geq (1 - \frac{c_0}{2}) \log |\text{SL}_2(\mathbb{F}_p)|,$$

where $c_0 := 1/4$ is the quasi-randomness parameter of $\text{SL}_2(\mathbb{F}_p)$. Consider the sequence

$$n \mapsto h_n := \frac{H_2(\overline{X}_{2^n \ell_0})}{\log |G_p|}.$$

We have proved that there exists $c_1 > 0$ such that

$$\begin{aligned} \text{(Initial entropy)} \quad & c_1 \leq h_0, \quad \text{and it suffices to show that} \\ \text{(Reaching threshold)} \quad & 1 - \frac{c_0}{2} \leq h_n \end{aligned}$$

for some n which is at most a number independent of p . If we show that there exists $c_3 > 0$ (independent of p) such that

$$h_{k+1} - h_k > c_3 \quad \text{unless} \quad h_k > 1 - \frac{c_0}{2},$$

then we obtain the desired result. Bourgain and Gamburd give us the precise condition under which one does not gain enough entropy after a random-walk.

Theorem 40 (Bourgain-Gamburd). *Suppose X and Y are two independent random-variables with values in a group G and finite range. If $K \geq 2$ and*

$$(25) \quad H_2(XY) \leq \frac{H_2(X) + H_2(Y)}{2} + \log K,$$

then there exists $A \subseteq G$ with the following properties:

- (1) (Approximate structure) A is a K^C -approximate subgroup.
- (2) (Controlling size) $|\log |A| - H_2(X)| \leq C \log K$.
- (3) (Equidistribution) $\mathbb{P}(XX' \in A) \geq \frac{1}{K^C}$, where X' is a random-variable independent of X with the probability law given by X^{-1} .

12. LECTURE 12: BOURGAIN-GAMBURD METHOD, II. MULTIPLICATIVE ENERGY.
APPROXIMATE SUBGROUPS.

Theorem 40 was initially studied for random-variables that were uniformly distributed on two subsets B and C of G . Note that if X is uniformly distributed on B and Y is uniformly distributed on C , then

$$\begin{aligned}\mathbb{P}(XY = g)^2 &= \left(\frac{|\{(b, c) \in B \times C \mid bc = g\}|}{|B||C|} \right)^2 \\ &= \frac{|\{(b_1, c_1, b_2, c_2) \in B \times C \times B \times C \mid b_1c_1 = b_2c_2 = g\}|}{|B|^2|C|^2}.\end{aligned}$$

Therefore,

$$(26) \quad \begin{aligned}H_2(XY) &= -\log \left(\frac{E(B, C)}{|B|^2|C|^2} \right), \quad \text{where} \\ E(B, C) &:= |\{(b_1, c_1, b_2, c_2) \in B \times C \times B \times C \mid b_1c_1 = b_2c_2\}|.\end{aligned}$$

The quantity $E(B, C)$ is called the *multiplicative energy*: think of it as a kind of bond energy between atoms. So, more energy indicates more structure. Inequality 25 may be reformulated as

$$E(B, C) \geq K^{-1}|B|^{3/2}|C|^{3/2}.$$

The multiplicative energy was first studied in the context of abelian groups by Balog and Szemerédi using a density theorem from graph theory. Later, Gowers proved a much stronger quantitative version of their result and, as a consequence, obtained a description of when two subsets of comparable size can have large energy. These results were often formulated in terms of the doubling quantity

$$D_2[A] := \frac{|A \cdot A|}{|A|}.$$

Tao developed this theory in the setting of non-abelian groups. Following Helfgott's work, he knew that the triple product expansion is the right quantity to consider in this context:

$$D_3[A] := \frac{|A \cdot A \cdot A|}{|A|}.$$

He made the theory more elegant by introducing *K-approximate subgroups*: $A \subseteq G$ is called an K -approximate subgroup if $1 \in A$, $A^{-1} = A$, and there exists $B \subseteq G$ such that

$$A \cdot A \subseteq A \cdot B \cap B \cdot A, \quad \text{and} \quad |B| \leq K.$$

Theorem 41 (Tao). *Suppose B and C are two finite subsets of a group G . If $K \geq 2$ and*

$$E(B, C) \geq K^{-1}|B|^{3/2}|C|^{3/2},$$

then there exists $A \subseteq G$ with the following properties:

- (1) [Approximate structure] A is a K^C -approximate subgroup.
- (2) [Controlling size] $|\log |A| - \log |B|| \leq C \log K$.
- (3) [Equidistribution] *There exist $x, y \in G$ such that*

$$\frac{|B \cap xA|}{|B|} \geq \frac{1}{K^C} \quad \text{and} \quad \frac{|C \cap Ay|}{|C|} \geq \frac{1}{K^C}.$$

Bourgain-Gamburd's theorem may be viewed as a weighted version of Tao's theorem.

Proof of Theorem 40. The idea is to *approximate* the probability laws $\mu(g) := \mathbb{P}(X = g)$ and $\nu(g) := \mathbb{P}(Y = g)$ by some weighted characteristic functions, and use Tao's theorem.

Comparable entropies. First observe that, by Lemma 30,

$$H_2(XY) \geq \max\{H_2(X), H_2(Y)\}.$$

Hence,

$$(27) \quad |H_2(X) - H_2(Y)| \leq 2 \log K, \quad \text{so} \quad H_2(XY) \leq \min\{H_2(X), H_2(Y)\} + 2 \log K.$$

Note that

$$\|\mu\|_2^2 = \sum_{x \in G} \mu(x)\mu(x) = \int_G \mu(x) d\mu(x),$$

so $\|\mu\|_2^2$ is the average of μ with respect to the probability measure given by μ . This inspires us to define the *tails* of μ as follows:

$$\begin{aligned} \mathcal{C}(\mu; >) &:= \{x \in G \mid \mu(x) > K^{10} \|\mu\|_2^2\}, \\ \mathcal{C}(\mu; <) &:= \{x \in G \mid \mu(x) < K^{-10} \|\mu\|_2^2\}, \end{aligned}$$

and the middle by

$$\mathcal{C}(\mu; \sim) := \{x \in G \mid K^{-10} \|\mu\|_2^2 < \mu(x) < K^{10} \|\mu\|_2^2\}.$$

Let $\mu_{>} := \mu \mathbb{1}_{\mathcal{C}(\mu; >)}$, $\mu_{<} := \mu \mathbb{1}_{\mathcal{C}(\mu; <)}$, and $\mu_{\sim} := \mu \mathbb{1}_{\mathcal{C}(\mu; \sim)}$.

Controlling tails. $\|\mu_{>}\|_1 < K^{-10}$ and $\|\mu_{<}\|_2 < K^5 \|\mu\|_2$.

Proof. The first inequality follows from the Markoff inequality:

$$\|\mu\|_2^2 = \int_G \mu(x) d\mu(x) > \mu(\mathcal{C}(\mu, >)) K^{10} \|\mu\|_2^2,$$

which implies that

$$\|\mu_{>}\|_1 = \mu(\mathcal{C}(\mu, >)) < K^{-10}.$$

The second inequality is rather an observation:

$$\|\mu_{<}\|_2^2 = \int_{\mathcal{C}(\mu; <)} \mu(x) d\mu(x) < \mu(\mathcal{C}(\mu; <)) K^{-10} \|\mu\|_2^2 \leq K^{-10} \|\mu\|_2^2,$$

which implies that

$$\|\mu_{<}\|_2 < K^{-5} \|\mu\|_2.$$

□

Cutting tails. $\|\mu_{\sim} * \nu_{\sim}\|_2 \geq (2K)^{-1} \|\mu\|_2^{1/2} \|\nu\|_2^{1/2}$.

Proof. Let us point out that $H_2(XY) \geq H_2(X)$ implies that

$$\|f * g\|_2 \leq \|f\|_1 \|g\|_2$$

for any two complex valued functions on G . Hence, by the previous steps,

$$\begin{aligned} \|\mu_{>} * \nu\|_2 &\leq \|\mu_{>}\|_1 \|\nu\|_2 < K^{-10} \|\nu\|_2 \leq K^{-9} (\|\mu\|_2 \|\nu\|_2)^{1/2}, \\ \|\mu_{<} * \nu\|_2 &\leq \|\nu\|_1 \|\mu_{<}\|_2 < K^{-5} \|\mu\|_2 \leq K^{-4} (\|\mu\|_2 \|\nu\|_2)^{1/2}, \\ \|\mu_{\sim} * \nu_{>}\|_2 &\leq \|\mu_{\sim}\|_2 \|\nu_{>}\|_1 < K^{-10} \|\mu\|_2 \leq K^{-9} (\|\mu\|_2 \|\nu\|_2)^{1/2}, \\ \|\mu_{\sim} * \nu_{<}\|_2 &\leq \|\mu_{\sim}\|_1 \|\nu_{<}\|_2 < K^{-5} \|\nu\|_2 \leq K^{-4} (\|\mu\|_2 \|\nu\|_2)^{1/2}. \end{aligned}$$

Since $\|\mu * \nu\|_2 > K^{-1} (\|\mu\|_2 \|\nu\|_2)^{1/2}$, we deduce that

$$\|\mu_{\sim} * \nu_{\sim}\|_2 \geq K^{-1} (1 - 2K^{-3} - 2K^{-8}) (\|\mu\|_2 \|\nu\|_2)^{1/2} \geq (2K)^{-1} (\|\mu\|_2 \|\nu\|_2)^{1/2}.$$

□

Size of the middle set. $4^{-1}K^{-24}\|\mu\|_2^{-2} \leq |\mathcal{C}(\mu; \sim)| \leq K^{20}\|\mu\|_2^{-2}$.

Proof. The upper bound is a consequence of the Markoff inequality:

$$\|\mu\|_2^2 \geq |\mathcal{C}(\mu; \sim)|(K^{-10}\|\mu\|_2^2)^2 \quad \text{which implies} \quad |\mathcal{C}(\mu; \sim)| \leq K^{20}\|\mu\|_2^{-2}.$$

To prove the lower bound, we begin with the following consequence of the previous step:

$$\|\mu_\sim\|_2 \geq \|\mu_\sim * \nu_\sim\|_2 \geq (2K)^{-1}(\|\mu\|_2\|\nu\|_2)^{1/2} \geq (2^{-1}K^{-2})\|\mu\|_2.$$

Note that

$$\|\mu_\sim\|_2^2 \leq |\mathcal{C}(\mu; \sim)|(K^{10}\|\mu\|_2^2)^2.$$

Therefore,

$$((2^{-1}K^{-2})\|\mu\|_2)^2 \leq |\mathcal{C}(\mu; \sim)|(K^{10}\|\mu\|_2^2)^2 \quad \text{which implies} \quad (4^{-1}K^{-24})\|\mu\|_2^{-2} \leq |\mathcal{C}(\mu; \sim)|.$$

□

High energy. $E(\mathcal{C}(\mu; \sim), \mathcal{C}(\nu; \sim)) \geq K^{-C}|\mathcal{C}(\mu; \sim)|^{3/2}|\mathcal{C}(\nu; \sim)|^{3/2}$ for some fixed $C > 0$.

Proof. Note that $\mu_\sim \leq (K^{10}\|\mu\|_2^2)\mathbb{1}_{\mathcal{C}(\mu; \sim)}$, $\nu_\sim \leq (K^{10}\|\nu\|_2^2)\mathbb{1}_{\mathcal{C}(\nu; \sim)}$, and

$$E(\mathcal{C}(\mu; \sim), \mathcal{C}(\nu; \sim)) = \|\mathbb{1}_{\mathcal{C}(\mu; \sim)} * \mathbb{1}_{\mathcal{C}(\nu; \sim)}\|_2^2.$$

Therefore, by the *cutting tails* step, we obtain

$$K^{40}(\|\mu\|_2^2\|\nu\|_2^2)^2 E(\mathcal{C}(\mu; \sim), \mathcal{C}(\nu; \sim)) \geq \|\mu_\sim * \nu_\sim\|_2^2 \geq (2K)^{-2}(\|\mu\|_2\|\nu\|_2),$$

which implies that

$$E(\mathcal{C}(\mu; \sim), \mathcal{C}(\nu; \sim)) \geq 4^{-1}K^{-42}(\|\mu\|_2^{-2}\|\nu\|_2^{-2})^{3/2} \geq K^{-C}|\mathcal{C}(\mu; \sim)|^{3/2}|\mathcal{C}(\nu; \sim)|^{3/2}.$$

□

Final step. By Theorem 41, there exists $A \subset G$ with the following properties:

- (1) (Approximate structure) A is $K^{C'}$ -approximate subgroup.
- (2) (Controlling size) $|\log |A| - \log |\mathcal{C}(\mu; \sim)|| \leq C' \log K$.
- (3) (Equidistribution) There exist $x, y \in G$ such that

$$\frac{|\mathcal{C}(\mu; \sim) \cap xA|}{|\mathcal{C}(\mu; \sim)|} \geq K^{-C'} \quad \text{and} \quad \frac{|\mathcal{C}(\nu; \sim) \cap Ay|}{|\mathcal{C}(\nu; \sim)|} \geq K^{-C'}.$$

Using our bounds on the **size of the middle sets**, we deduce that

$$|H_2(X) - \log |A|| = |\log \|\mu\|_2^{-2} - \log |A|| < 26 \log K + |\log |A| - \log |\mathcal{C}(\mu; \sim)|| \leq C'' \log K.$$

We also have

$$\begin{aligned} \mu(xA) &\geq \mu(\mathcal{C}(\mu; \sim) \cap xA) \\ &\geq K^{-10}\|\mu\|_2^2 |\mathcal{C}(\mu; \sim) \cap xA| \\ &\geq K^{-10-C'}\|\mu\|_2^2 |\mathcal{C}(\mu; \sim)| \geq K^{-36-C'}. \end{aligned}$$

Similarly, we obtain

$$\mu(Ay) \geq K^{-36-C'}.$$

The claim follows. □

Let's review what results we have so far: for a pair of integers n, ℓ_0 , we define

$$h_{n, \ell_0} := \frac{H_2(\overline{X}_{2^n \ell_0})}{\log |G_p|}.$$

Proved that there exists $\ell_0 \ll \log |G_p|$ and $c_1 > 0$ such that

$$H_\infty(\overline{X}_{\ell_0}) \geq c_1 \log |G_p| \quad \text{in particular,} \quad h_{0, \ell_0} \geq c_1.$$

It suffices to show that for some $n \ll 1$, $h_{n, \ell_0} \geq 1 - \frac{c_0}{2}$. To that end, we will prove that for sufficiently small $\delta > 0$, for every n either $h_{n, \ell_0} > 1 - \frac{c_0}{2}$ or $h_{n+1, \ell_0} - h_{n, \ell_0} > \delta$. Suppose to the contrary for some n and small enough (to be determined later) c_3 , $h_{n+1, \ell_0} - h_{n, \ell_0} < \delta$. This implies that

$$H_2(\overline{X}_{2^{n+1} \ell_0}) < H_2(\overline{X}_{2^n \ell_0}) + \delta \log |G_p|.$$

So by the Bourgain-Gamburd theorem, there exists $A \subseteq G_p$ with the following properties:

- (1) A is an $|G_p|^{C\delta}$ -approximate subgroup.
- (2) $|\log |A| - h_{n, \ell_0} \log |G_p|| \leq C\delta \log |G_p|$.
- (3) $\mathbb{P}(X_{2^{n+1} \ell_0} \in A \cdot A) > |G_p|^{-C\delta}$.

Note that $c_1 \log |G_p| < h_{n, \ell_0} < (1 - \frac{c_0}{2}) \log |G_p|$ implies that

$$(28) \quad |G_p|^{\frac{c_1}{2}} < |G_p|^{c_1 - C\delta} < |A| < |G_p|^{1 - \frac{c_0}{2} + C\delta} < |G_p|^{1 - \frac{c_0}{4}}$$

assuming that $\delta \ll \min\{c_0, c_1\}$. In order to reach to a contradiction, we must investigate approximate subgroups A with *positive Hausdorff dimension*, by this I mean having a lower bound on $\frac{\log |A|}{\log |G_p|}$, and show that we can *escape* them.

Helfgott proved the following result, which motivated Bourgain and Gamburd to revisit Lubotzky's 1-2-3 problem.

Theorem 42 (Helfgott). *For every $c > 0$, there exist $e(c) > 0$ and $c'(c) > 0$ such that if A is a generating set of $\text{SL}_2(\mathbb{F}_p)$ and $|A| < |\text{SL}_2(\mathbb{F}_p)|^{1-c}$, then $|A \cdot A \cdot A| > c'|A|^{1+e}$.*

Note that if A is a $|G_p|^{C\delta}$ -approximate subgroup which generates $\text{SL}_2(\mathbb{F}_p)$, and

$$|G_p|^{\frac{c_1}{2}} < |A| < |G_p|^{1 - \frac{c_0}{4}},$$

then by Theorem 42, there exist e and c' depeneding only on c_0 such that

$$|G_p|^{2C\delta} > \frac{|A \cdot A \cdot A|}{|A|} > c'|A|^e > c'|G_p|^{\frac{c_1}{2}e} > |G_p|^{\frac{c_1 e}{3}}$$

where the last inequality holds for sufficiently large p . We obtain a contradiction if $\delta < \frac{c_1 e}{6C}$. This would imply the approximate subgroup obtained by the Bourgain-Gamburd theorem cannot generate $\text{SL}_2(\mathbb{F}_p)$, which means it is a subset of a proper subgroup. Later, we will argue why we can *escape* proper subgroups.

Helfgott's result is extended to all finite almost simple groups of Lie type by Breuillard, Green and Tao, and independently, by Pyber and Szabo.

Theorem 43 (Breuillard-Green-Tao). *Suppose F is a finite field and $\underline{G}$ is an absolutely almost simple F -algebraic group. Then there exists $C > 0$ depending only on the rank of $\underline{G}$ such that the following holds. Suppose A is a generating set of $\underline{G}(F)$ and A is a K -approximate subgroup. Then either $|A| < K^C$ or $|A| > K^{-C}|\underline{G}(F)|$.*

13. LECTURE 13: A DETOUR: THE SUM-PRODUCT PHENOMENON OVER FINITE FIELDS.

The following result of Bourgain, Katz, and Tao motivated Helfgott to think about growth in $\mathrm{SL}_2(\mathbb{F}_p)$, and it was a crucial ingredient in his original proof.

Theorem 44 (Bourgain-Katz-Tao). *For every $c > 0$, there exist positive numbers $e(c)$ and $c'(c)$ such that for every $A \subseteq \mathbb{F}_p$ with $p^c < |A| < p^{1-c}$,*

$$\max\{|A + A|, |A \cdot A|\} > c'(c)|A|^{1+e(c)}.$$

We present a *bounded generation* result which is the backbone of Theorem 44. For $A \subseteq \mathbb{F}_q$ with $0, 1 \in A$ and a positive integer ℓ , let

$$\langle A \rangle_\ell := \sum_\ell \prod_\ell A - \sum_\ell \prod_\ell A.$$

Note that $\bigcup_\ell \langle A \rangle_\ell$ is the subring generated by A . So the parameter ℓ is measuring how many steps are needed to generate this subring.

Theorem 45 (Bourgain-Katz-Tao). *For every $c > 0$, there exists a positive integer $C(c)$ such that the following holds. Suppose $\mathbb{F}_q$ is a finite field order $q = p^m$ and $A \subseteq \mathbb{F}_q$ such that $0, 1 \in A$ and $|A| > q^c$. Then $\langle A \rangle_{C(c)}$ is a subfield of $\mathbb{F}_q$.*

We start with the following sum-scalar product lemma.

Lemma 46. *Suppose $A, B \subseteq \mathbb{F}_q$. Then there exists $c \in \mathbb{F}_q^\times$ such that*

$$|A + cB| > \frac{1}{2} \min\{|A||B|, q - 1\}.$$

Proof. Let X and Y be independent random-variables with uniform distributions on A and B , respectively. Then for every $c \in \mathbb{F}_q^\times$, the range of $X + cY$ is $A + cB$, and so

$$(29) \quad \log |A + cB| \geq H_2(X + cY).$$

We will compute the average of $H_2(X + cY)$ as c ranges in $\mathbb{F}_q^\times$. This average may be viewed as a *conditional Rényi entropy* $H_2(X + cY | U)$ where X, Y, U mutually independent and U is uniform on $\mathbb{F}_q^\times$.

$$(30) \quad \begin{aligned} \frac{1}{|\mathbb{F}_q^\times|} \sum_{c \in \mathbb{F}_q^\times} H_2(X + cY) &= \frac{1}{|\mathbb{F}_q^\times|} \sum_{c \in \mathbb{F}_q^\times} -\log \|\mu_A * \mu_{cB}\|_2^2 \\ &\geq -\log \left(\frac{1}{|\mathbb{F}_q^\times|} \sum_{c \in \mathbb{F}_q^\times} \|\mu_A * \mu_{cB}\|_2^2 \right), \end{aligned}$$

as $-\log(x)$ is a convex function. Recall that by (26),

$$\|\mu_A * \mu_{cB}\|_2^2 = \frac{|\{(a_1, b_1, a_2, b_2) \in A \times B \times A \times B \mid a_1 + cb_1 = a_2 + cb_2\}|}{|A|^2 |B|^2}.$$

Hence,

$$\frac{1}{|\mathbb{F}_q^\times|} \sum_{c \in \mathbb{F}_q^\times} \|\mu_A * \mu_{cB}\|_2^2 = \frac{|\{(a_1, b_1, a_2, b_2, c) \in A \times B \times A \times B \times \mathbb{F}_q^\times \mid a_1 + cb_1 = a_2 + cb_2\}|}{(q-1)|A|^2 |B|^2}.$$

Note that $a_1 + cb_1 = a_2 + cb_2$ holds if and only if

$$\text{either } a_1 \neq a_2, b_1 \neq b_2, c = \frac{a_1 - a_2}{b_2 - b_1} \quad \text{or} \quad a_1 = a_2, b_1 = b_2.$$

Therefore,

$$\begin{aligned} \frac{1}{|\mathbb{F}_q^\times|} \sum_{c \in \mathbb{F}_q^\times} \|\mu_A * \mu_{cB}\|_2^2 &= \frac{|A|(|A|-1)|B|(|B|-1) + |A||B|(q-1)}{(q-1)|A|^2|B|^2} \\ &= \frac{(|A|-1)(|B|-1)}{(q-1)|A||B|} + \frac{1}{|A||B|} \\ &< 2 \max\left\{\frac{1}{q-1}, \frac{1}{|A||B|}\right\} \end{aligned}$$

Claim follows from the above upper bound, (29), and (30). $\square$

We repeatedly apply the sum-scalar product expansion till we reach to the threshold $\frac{q-1}{2}$. Then after adding one more element, we may apply the following observation:

Lemma 47. *Suppose G is a finite group, $A \subseteq G$, and $|A| > \frac{|G|}{2}$. Then $A \cdot A = G$.*

Proof. For every $x \in G$, $A \cap xA^{-1} \neq \emptyset$ as $|A| + |xA^{-1}| > |G|$. Suppose $a \in A \cap xA^{-1}$. Then there exists $a' \in A$ such that $a = xa'^{-1}$, so $x = aa' \in A \cdot A$. $\square$

Proposition 48. *For every $\varepsilon > 0$, there exists a positive integer $C := C(\varepsilon)$ such that the following holds. For every power of a prime q with $q^\varepsilon > 4$ and a subset A of a finite field $\mathbb{F}_q$ with $|A| > q^\varepsilon$, there exist $c_1, \dots, c_C \in \mathbb{F}_q^\times$ such that*

$$c_1A + \dots + c_CA = \mathbb{F}_q.$$

Proof. We prove by induction on k that there exist $c_1, \dots, c_k \in \mathbb{F}_q^\times$ such that

$$|c_1A + \dots + c_kA| > \min\{q^{(1+.5k)\varepsilon}, (q-1)/2\}.$$

The base of induction is clear. Applying Lemma 46 to the pair of sets $(c_1A + \dots + c_kA, A)$, we find $c_{k+1} \in \mathbb{F}_q^\times$ such that

$$\begin{aligned} |c_1A + \dots + c_kA + c_{k+1}A| &> \min\left\{\frac{|c_1A + \dots + c_kA||A|}{2}, \frac{q-1}{2}\right\} \\ &> \min\left\{\frac{q^{(1+.5k)\varepsilon}q^\varepsilon}{2}, \frac{q-1}{2}\right\} \\ &> \min\left\{q^{(1+.5(k+1))\varepsilon}, \frac{q-1}{2}\right\} \end{aligned}$$

Hence, for $k \leq 2\varepsilon^{-1}$, we have

$$|c_1A + \dots + c_kA| > \frac{q-1}{2}.$$

Adding one more element and applying Lemma 47, we deduce that $C(\varepsilon) := 4\varepsilon^{-1} + 2$ satisfies the desired property. $\square$

Lemma 49. *Suppose $\mathbb{F}_q = c_1A + \dots + c_kA$ and $1 \in A$. Then one of the following statements hold:*

- (1) **(Reduction)** *There exists i_0 such that $\mathbb{F}_q = c_1\langle A \rangle_4 + \dots + c_{i_0-1}\langle A \rangle_4 + c_{i_0+1}\langle A \rangle_4 + \dots + c_k\langle A \rangle_4$.*
- (2) **(Injectivity)** *The function $\ell_c: A^k \rightarrow \mathbb{F}_q$, $\ell_c(a_1, \dots, a_k) := \sum_{i=1}^k c_i a_i$ is injective.*

Proof. Suppose ℓ_c is not injective. Then, there exist $\mathbf{a} \neq \mathbf{a}' \in A^k$ such that $\ell_c(\mathbf{a}) = \ell_c(\mathbf{a}')$. We may assume that $a_k \neq a'_k$. Hence,

$$c_k(a'_k - a_k) = - \sum_{i=1}^{k-1} c_i(a_i - a'_i) \quad \text{and} \quad \mathbb{F}_q = c_1(a'_k - a_k)A + \dots + c_k(a'_k - a_k)A.$$

Therefore,

$$\begin{aligned}\mathbb{F}_q &= c_1((a'_k - a_k)A + (a_1 - a'_1)A) + \cdots + c_{k-1}((a'_k - a_k)A + (a_{k-1} - a'_{k-1})A) \\ &\subseteq c_1\langle A \rangle_4 + \cdots + c_{k-1}\langle A \rangle_4\end{aligned}$$

This finishes the proof. $\square$

Lemma 50. Suppose $\mathbb{F}_q = c_1A + \cdots + c_kA$ and $0, 1 \in A$. If

$$\ell_{\mathbf{c}}: \langle A \rangle_2^k \rightarrow \mathbb{F}_q, \quad \ell_{\mathbf{c}}(x_1, \dots, x_k) := \sum_{i=1}^k c_i x_i$$

is injective, then A is a subfield of $\mathbb{F}_q$.

Proof. Since $\mathbb{F}_q$ is a finite field, it suffices to prove that $A \cdot A = A$ and $A + A = A$. For every $b_1, b_2 \in A$, there exists $\mathbf{a}, \mathbf{a}' \in A^k$ such that

$$\ell_{\mathbf{c}}(\mathbf{a}) = c_1(b_1 - b_2) = \ell_{\mathbf{c}}(b_1 - b_2, 0, \dots, 0) \quad \text{and} \quad \ell_{\mathbf{c}}(\mathbf{a}') = c_1(b_1 b_2) = \ell_{\mathbf{c}}(b_1 b_2, 0, \dots, 0).$$

Hence, by the injectivity of $\ell_{\mathbf{c}}$ and $A \subseteq \langle A \rangle_2$, we deduce that

$$b_1 - b_2 = a_1 \in A \quad \text{and} \quad b_1 b_2 = a'_1 \in A.$$

This finishes the proof. $\square$

Proof of Theorem 45. By Proposition 48, there exists $c_1, \dots, c_C \in \mathbb{F}_q^\times$ such that

$$(31) \quad c_1A + \cdots + c_CA = \mathbb{F}_q, \quad \text{so} \quad c_1\langle A \rangle_2 + \cdots + c_C\langle A \rangle_2 = \mathbb{F}_q.$$

Hence, by Lemma 49, either (31) holds for $\langle A \rangle_8$ with one less term or by Lemma 50 A is a subfield. Thus, after at most $C(\varepsilon)$ -many **reductions**, we obtain a subfield. Hence, $\langle A \rangle_{8^C}$ is a subfield of $\mathbb{F}_q$. This finishes the proof. $\square$

14. LECTURE 14: PRODUCT THEOREM, I.

Our plan is to prove Helfgott's product theorem (theorem 42). We present a proof which contains the main ideas of the general case.

Let's recall the functor of common zeros of a family of polynomials. For an ideal $\mathfrak{a}$ of the ring of polynomials $F[x_1, \dots, x_n]$,

$$V_{\mathfrak{a}}: F\text{-}\underline{\text{Alg}} \rightarrow \underline{\text{Set}}, \quad V_{\mathfrak{a}}(A) := \{\mathbf{a} \in A^n \mid \forall f \in \mathfrak{a}, f(\mathbf{a}) = 0\}.$$

This functor is represented by the F -algebra $F[x_1, \dots, x_n]/\mathfrak{a}$. For an algebraic closure $\overline{F}$ of F , there exists a bijection between $V_{\mathfrak{a}}(\overline{F})$ and the set of maximal ideals of $F[\mathbf{x}]/\mathfrak{a}$. So it can be viewed as the subset of $\text{Spec}(F[\mathbf{x}]/\mathfrak{a})$. We take the induced Zariski-topology on $V_{\mathfrak{a}}(\overline{F})$; this means closed sets are of the form $V_{\mathfrak{b}}(\overline{F})$ where $\mathfrak{b}|\mathfrak{a}$. The dimension of $V_{\mathfrak{a}}$ is defined to be the Krull dimension of $F[\mathbf{x}]/\mathfrak{a}$.

For $\mathbf{p} \in V_{\mathfrak{a}}(A)$, the A -points of the tangent space of $V_{\mathfrak{a}}$ at $\mathbf{p}$ is the fiber of the map

$$V_{\mathfrak{a}}(A[t]/\langle t \rangle) \rightarrow V_{\mathfrak{a}}(A)$$

over $\mathbf{p}$. This takes us to the following definition

$$T_{\mathbf{p}}V_{\mathfrak{a}}(A) := \{\mathbf{v} \in A^n \mid \forall f \in \mathfrak{a}, df_{\mathbf{p}}(\mathbf{v}) = 0\}.$$

Following Breuillard, Green, and Tao, we say the complexity of $V_{\mathfrak{a}}(\overline{F})$ is at most M if $n \leq M$ and $\mathfrak{a}$ can be generated by M polynomials of degree at most M . For two ideals $\mathfrak{a}$ and $\mathfrak{b}$ of $F[\mathbf{x}]$, we let

$$V_{\mathfrak{a}, \mathfrak{b}}(A) := \{\mathbf{p} \in V_{\mathfrak{a}}(A) \mid \langle e_{\mathbf{p}}(\mathfrak{b}) \rangle = A\},$$

where $e_{\mathbf{p}}: F[\mathbf{x}] \rightarrow A$ is the evaluation map at $\mathbf{p}$. Note that for a field extension E/F ,

$$V_{\mathbf{a};\mathbf{b}}(E) = V_{\mathbf{a}}(E) \setminus V_{\mathbf{b}}(E).$$

We say $V_{\mathbf{a};\mathbf{b}}$ is of complexity at most M if $n \leq M$ and $\mathbf{a}$ and $\mathbf{b}$ can be generated by at most M polynomials of degree at most M .

We say $\underline{G}$ is an algebraic subgroup of $(\underline{\mathrm{GL}}_n)_F$ if the following statements hold:

- (1) For some ideal $\mathbf{a}$ of $F[x_{ij}]$, $\underline{G} = V_{\mathbf{a};\langle \det[x_{ij}] \rangle}$; that means for every F -algebra A the set $\underline{G}(A)$ is the common solutions of $f \in \mathbf{a}$ in $\underline{\mathrm{GL}}_n(A)$.
- (2) For every F -algebra A , $\underline{G}(A)$ is a subgroup of $\underline{\mathrm{GL}}_n(A)$.
- (3) $\dim \underline{G} = \dim \mathrm{Lie}(\underline{G})(\overline{F})$, where $\mathrm{Lie}(\underline{G}) = T_I \underline{G}$.

The following is at the center of proofs of the product theorems.

Theorem 51 (Larsen-Pink inequality). *Suppose $\underline{G} \subseteq (\underline{\mathrm{GL}}_n)_{\mathbb{F}}$ is an absolutely almost simple group of complexity at most M where $\mathbb{F}$ is a finite field. Suppose $V_{\mathbf{a}} \subseteq \underline{G}$ is of complexity at most M . Let A be a symmetric generating set of $\underline{G}(\mathbb{F})$.*

Proof of the Larsen-Pink Inequality for dimension 1 varieties should be typed.

15. LECTURE 15: PRODCUT THEOREM, II.

In the previous lecture, we mentioned the *Larsen-Pink inequality* (due to Breuillard-Green-Tao): if A is a symmetric generating set of $\mathrm{SL}_n(\mathbb{F}_q)$ and $V_{\mathbf{a}} \subseteq (\underline{\mathrm{SL}}_n)_{\mathbb{F}_q}$ is of complexity at most M , then there exist $m := m(M)$ and $c := c(M)$ such that

$$|A \cap V_{\mathbf{a}}(\mathbb{F}_q)| \leq c |\prod_m A|^{\frac{\dim V_{\mathbf{a}}}{\dim(\underline{\mathrm{SL}}_n)_{\mathbb{F}_q}}}.$$

We proved this statement when $\dim V_{\mathbf{a}} = 1$. We mentioned that the general case is proved by induction on $\dim V_{\mathbf{a}}$. We will prove a special case of the Larsen-Pink inequality needed to study approximate subgroups of $\mathrm{SL}_2(\mathbb{F}_q)$. To that end, we want to show the intersection of a conjugacy class and a bounded folded product of A is *small*. A conjugacy class of $\mathrm{SL}_2(\mathbb{F}_q)$ is more or less determined by its trace. For $t \in \mathbb{F}_q$,

$$V_t(A) := \{a \in M_2(A) \mid \det(a) = 1, \mathrm{Tr}(a) = t\}.$$

Clearly, the complexity of V_t is 2. We also observe that $\dim V_t = 2$. Hence, the following is a special case of the Larsen-Pink inequality.

Proposition 52 (Escaping conjugacy classes: non-weighted version). *There exist positive integer m and positive number c such that for every symmetric generating set A of $\mathrm{SL}_2(\mathbb{F}_q)$ and $t \in \mathbb{F}_q$,*

$$|A \cap V_t(\mathbb{F}_q)| \leq c |\prod_m A|^{\frac{2}{3}}.$$

Proof. Following the case of 1 subvarieties, we consider the map:

$$f: V_t(\overline{\mathbb{F}}_q) \times V_t(\overline{\mathbb{F}}_q) \rightarrow \mathrm{SL}_2(\overline{\mathbb{F}}_q), \quad f(x_1, x_2) := x_1 x_2^{-1}.$$

If $f|_{(A \cap V_t(\mathbb{F}_q))^2 \setminus \Delta(A \cap V_t(\mathbb{F}_q))}$ is injective, then

$$|A \cap V_t(\mathbb{F}_q)|^2 \leq 2 |\prod_2 A| \quad \text{which implies that} \quad |A \cap V_t(\mathbb{F}_q)| < 2 |\prod_2 A|^{1/2}.$$

The claim follows. So, we may assume that f is not injective. Note that this is not surprising as a typical fiber of f is of dimension

$$\dim(V_t \times V_t) - \dim(\mathrm{SL}_2)_{\mathbb{F}_q} = 1.$$

For $w \in \mathrm{SL}_2(\mathbb{F}_q)$, the fiber of f over w is given by

$$\{(wx, x) \in \mathrm{SL}_2(\overline{\mathbb{F}}_q) \mid \mathrm{Tr}(x) = \mathrm{Tr}(wx) = t\}.$$

Let $V_w := \{x \in \mathrm{SL}_2(\overline{\mathbb{F}}_q) \mid \mathrm{Tr}(x) = \mathrm{Tr}(wx) = t\}$. Note that $\dim V_w = 1$ for $w \neq 1$. Hence, by the Larsen-Pink inequality for dimension 1 closed subsets, there exist $m_1 \in \mathbb{N}$ and $c_1 \in \mathbb{R}^+$ such that

$$|A \cap V_w(\mathbb{F}_q)| \leq c_1 |\prod_{m_1} A|^{1/3},$$

for every $w \in \mathrm{SL}_2(\mathbb{F}_q) \setminus \{1\}$. Therefore,

$$\begin{aligned} |A \cap V_t(\mathbb{F}_q)|^2 &\leq |A \cap V_t(\mathbb{F}_q)| + \sum_{w \in A \cdot A \setminus \{1\}} |A \cap V_w(\mathbb{F}_q)| \\ &\leq |A| + c_1 |A \cdot A| |\prod_{m_1} A|^{1/3} \\ &\leq (c_1 + 1) |\prod_{m_1} A|^{4/3}, \end{aligned}$$

which implies that

$$|A \cap V_t(\mathbb{F}_q)| \leq (c_1 + 1)^{1/2} |\prod_{m_1} A|^{2/3}.$$

This finishes the proof. $\square$

We need a bit more connection between conjugation and a symmetric subset A of a group G .

Lemma 53 (Orbit-Stabilizer: subsets). *Suppose G acts on a set X and $A \subseteq G$ is a finite symmetric subset. Then, for every $x \in X$,*

$$|A| \leq \underbrace{|A \cdot x|}_{\text{orbit}} \underbrace{|AA \cap G_x|}_{\text{stabilizer}} \leq |AAA|.$$

This finishes the proof.

Proof. By the orbit-stabilizer theorem, we know that $f: G/G_x \rightarrow G \cdot x$, $f(gG_x) := g \cdot x$ is a bijection. So,

$$|A \cdot x| = |\{aG_x \mid a \in A\}|.$$

Note that $a_1G_x = a_2G_x$ if and only if $a_1^{-1}a_2 \in G_x$. Hence, each fiber of $A \rightarrow G/G_x$ has at most $|AA \cap G_x|$ -many elements. Therefore,

$$|A \cdot x| \geq \frac{|A|}{|AA \cap G_x|}.$$

To prove the other inequality, note that for every $B \subseteq G$, each fiber of

$$h: AB \rightarrow A \cdot x, \quad h(g) := g \cdot x$$

has at least $|B \cap G_x|$ -many elements as $h(a(B \cap G_x)) = a \cdot x$. Therefore,

$$|AB| \geq |B \cap G_x| |A \cdot x|.$$

In particular,

$$|AAA| \geq |AA \cap G_x| |A \cdot x|.$$

$\square$

Corollary 54 (Conjugacy class: subsets). *Suppose $A \subseteq G$ is symmetric. Then, for every $x \in A$,*

$$|A| \leq |\{axa^{-1} \mid a \in A\}| |AA \cap C_G(x)| \leq |AAA|.$$

Proof. Just consider the action of G on itself by conjugation and apply the Orbit-Stabilizer for subsets. $\square$

Corollary 55 (Rich centralizers). *There exist $m \in \mathbb{N}$ and $c \in \mathbb{R}^+$ such that the following holds. Suppose A is a symmetric generating set of $\mathrm{SL}_2(\mathbb{F}_q)$. Then for every $x \in A \setminus \{\pm 1\}$,*

$$|AA \cap C_G(x)| \geq c^{-1} \frac{|A|}{|\prod_m A|^{2/3}}.$$

In particular, if A is a K -approximate subgroup, then

$$|AA \cap C_G(x)| \geq c^{-1} K^{-2m/3} |A|^{1/3}.$$

Proof. Note that $|\{axa^{-1} \mid a \in A\}| \leq |A \cap V_{\mathrm{Tr}(x)}(\mathbb{F}_q)|$. Hence, by Proposition 52,

$$|\{axa^{-1} \mid a \in A\}| \leq c |\prod_m A|^{2/3}.$$

Therefore, the first inequality follows from Corollary 54. The second inequality follows from the observation that

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

as A is a K -approximate subgroup. □

Note that if $t \neq \pm 2$, then every element of $V_t(\overline{\mathbb{F}}_q)$ has distinct eigenvalues. So, they are diagonalizable. We also observe that

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}^{-1} = \begin{pmatrix} x_{11} & a^2 x_{12} \\ a^{-2} x_{21} & x_{22} \end{pmatrix}.$$

In particular, the centralizer of a diagonal matrix which is not ± 1 is the set of diagonal matrices.

Corollary 56. *Suppose $t \neq \pm 2$. Then for every $x \in V_t(\mathbb{F}_q)$, $C_{\mathrm{SL}_2(\overline{\mathbb{F}}_q)}(x)$ is abelian and it is diagonalizable over $\overline{\mathbb{F}}_q$. In particular, $\dim C_{(\mathrm{SL}_2)_{\mathbb{F}_q}}(x) = 1$.*

Theorem 57 (Product theorem due to Helfgott). *There exist $m \in \mathbb{N}$ such that the following holds. Suppose A is a generating set of $\mathrm{SL}_2(\mathbb{F}_q)$ and it is a K -approximate subgroup where $K > 2$. Then*

$$\text{either } |A| < K^m \quad \text{or} \quad |A| > K^{-m} |\mathrm{SL}_2(\mathbb{F}_q)|.$$

Proof. There exists $x_0 \in \prod_{m_0} A$ such that $\mathrm{Tr}(x_0) \neq \pm 2$. Let $\underline{T} := C_{(\mathrm{SL}_2)_{\mathbb{F}_q}}(x_0)$. Then by Corollary 56, $\underline{T}$ is diagonalizable. We say $x \in \mathrm{SL}_2(\mathbb{F}_q)$ is a *pivot* if fibers of

$$\phi_x: A \times \underline{T}(\mathbb{F}_q) \rightarrow \mathrm{SL}_2(\mathbb{F}_q), \quad \phi_x(a, t) := axtx^{-1}$$

modulo $\pm 1^2$ have one element. Let $T := \underline{T}(\mathbb{F}_q)$.

Case 1. There exists a pivot x_1 in A .

Proof of Case 1. In this case, $\phi_{x_1}(A \times (AA \cap T)) \subseteq \prod_4 A$ and by Corollary 55

$$|\phi_{x_1}(A \times (AA \cap T))| \gg |A| |AA \cap T| \gg K^{-2m/3} |A|^{4/3}.$$

Therefore,

$$K^{-2m/3} |A|^{4/3} \ll \left| \prod_4 A \right| \leq K^3 |A|.$$

So

$$|A| \ll K^{O(1)},$$

which means A is controlled by $\{1\}$. □

Case 2. There is no pivot in $\mathrm{SL}_2(\mathbb{F}_q)$.

Proof of Case 2. Note that x is not a pivot if and only if there exists

$$\pm 1 \neq y \in AA \cap xTx^{-1}.$$

Hence, by Corollary 55 and Corollary 56, we deduce that for every $x \in \mathrm{SL}_2(\mathbb{F}_q)$,

$$|AA \cap xTx^{-1}| \gg K^{-2m/3}|A|^{4/3}.$$

We also note that Corollary 56 implies that

$$xTx^{-1} \cap x'Tx'^{-1} = \pm 1 \quad \text{if} \quad xTx^{-1} \neq x'Tx'^{-1}.$$

Therefore,

$$\begin{aligned} |AA| &\gg \sum_{x \in \mathrm{SL}_2(\mathbb{F}_q)} (|AA \cap xTx^{-1}| - 2) \\ &\gg [\mathrm{SL}_2(\mathbb{F}_q) : N_{\mathrm{SL}_2(\mathbb{F}_q)}(T)] K^{-2m/3} |A|^{1/3} \\ &\gg K^{-2m/3} q^2 |A|^{1/3}. \end{aligned}$$

Hence, using $|AA| \leq K|A|$, we obtain

$$|A| \ll K^{O(1)} q^3 \ll K^{O(1)} |\mathrm{SL}_2(\mathbb{F}_q)|,$$

which means A is controlled by $\mathrm{SL}_2(\mathbb{F}_q)$. □

(To be continued). □

16. LECTURE 16: APPROXIMATE SUBGROUP, III. PING-PONG ARGUMENT AND ESCAPING ALGEBRAIC SUBGROUPS.

Proof of Theorem 57. Case 3. There are some pivots but none in A .

Proof of Case 3. Let x' be a pivot with smallest distance from the identity in the Cayley graph of $\mathrm{SL}_2(\mathbb{F}_q)$ with respect to the generating set A . Hence, $x' = a_0x$ where $a_0 \in A$ and x is not a pivot.

Since x is not a pivot, by a similar argument as in Case 2,

$$|AA \cap xTx^{-1}| \gg K^{-2m/3}|A|^{1/3}.$$

Since a_0x is a pivot, the map

$$A \times xTx^{-1} \rightarrow \mathrm{SL}_2(\mathbb{F}_q), \quad (a, xtx^{-1}) \mapsto aa_0xtx^{-1}a_0^{-1}$$

is injective modulo ± 1 . Therefore,

$$|A||AA \cap xTx^{-1}| \ll |Aa_0AAa_0^{-1}| \leq |AAAAA| \leq K^4|A|.$$

Hence,

$$K^{-2m/3}|A|^{4/3} \ll K^4|A|.$$

Thus

$$|A| \ll K^{O(1)}.$$

This completes the proof. □

□

Let's revisit the Bourgain-Gamburd method:

Theorem 58 (The Bourgain-Gamburd method). *Let the following hypotheses hold.*

(H1) G is a c_0 -quasi-random group.

(H2) If A is a K -approximate subgroup of G and $\langle A \rangle = G$, then either

$$\text{either } |A| \leq c_1 K^{c_1} \quad \text{or} \quad |A| \geq c_1^{-1} K^{-c_1} |G|.$$

(H3) X is a symmetric random-variable with values in G .

(H4) (Initial entropy) For some $\ell_0 \leq c_3 \log |G|$, $H_\infty(X_{\ell_0}) \geq c_4 \log |G|$.

Then, for $\gamma := \gamma(c_0, c_1, c_3, c_4) > 0$ which is sufficiently small and for every non-negative integer k , one of the following statements hold:

(C0) (Exceptional) $|G|^\gamma \leq 3$.

(C1) (Enough entropy) $H_2(X_{2^k \ell_0}) > (1 - \frac{c_0}{2}) \log |G|$.

(C2) (Gaining entropy) $H_2(X_{2^{k+1} \ell_0}) \geq H_2(X_{2^k \ell_0}) + \gamma \log |G|$.

(C3) (Trapped in a subgroup) There exists a proper subgroup H such that $\log |H| > \frac{c_4}{2} \log |G|$ and $\mathbb{P}(X_{2^{k+1} \ell_0} \in H) > |G|^{-C\gamma}$, where C is a fixed universal constant.

Proof. For a non-negative integer k and a sufficiently small γ (to be determined later), we assume that (C0), (C1) and (C2) do not hold: $|G|^\gamma > 3$, and we do *not have enough entropy* but do *not gain enough entropy*, either. Then by Theorem 40 there exists $A \subset G$ and a universal constant C such that the following statements hold:

(1) A is a $|G|^{C\gamma}$ -approximate subgroup.

(2) $|\log |A| - H_2(X_{2^k \ell_0})| \leq C\gamma \log |G|$.

(3) $\mathbb{P}(X_{2^{k+1} \ell_0} \in AA) \geq |G|^{-C\gamma}$.

By the second property and not having enough entropy, we deduce that

$$(32) \quad \log |A| \leq H_2(X_{2^k \ell_0}) + C\gamma \log |G| \leq \left(1 - \frac{c_0}{4}\right) \log |G|$$

assuming $\gamma < \frac{c_0}{4C}$.

Claim 1. $\log |A| > \frac{c_4}{2} \log |G|$.

Proof of Claim 1. Note that $H_\infty(X_{2^{k+1} \ell_0}) \geq H_\infty(X_{\ell_0})$ (see Lemma 30). Hence, by (H4) and the third property, it follows that

$$\begin{aligned} |G|^{-C\gamma} &\leq \mathbb{P}(X_{2^{k+1} \ell_0} \in AA) \\ &= |AA| \max_g \mathbb{P}(X_{2^{k+1} \ell_0} = g) \\ &= |G|^{C\gamma} |A| |G|^{-c_4}. \end{aligned}$$

Hence,

$$|A| > |G|^{c_4 - 2C\gamma} > |G|^{c_4/2}$$

assuming $\gamma < \frac{c_4}{4C}$. This finishes proof of the Claim 1. $\square$

Claim 2. $H := \langle A \rangle$ is a proper subgroup.

Proof of Claim 2. Suppose to the contrary $H = G$. Then, by (H2),

$$\text{either } \log |A| \leq c_1 \gamma \log |G| + \log c_1 \quad \text{or} \quad \log |A| \geq (1 - c_1 \gamma) \log |G| - \log c_1.$$

Since we assumed (C0) does not hold, we may assume that

$$\gamma \log |G| > 1.$$

Therefore, we have

$$\text{either } \log |A| \leq 2c_1 \gamma \log |G| \quad \text{or} \quad \log |A| \geq (1 - 2c_1 \gamma) \log |G|.$$

Assuming $2c_1 \gamma < \frac{c_4}{2}$ the first inequality cannot hold because of Claim 1, and assuming $2c_1 \gamma < \frac{c_0}{4}$, the second inequality cannot hold because of (32). Thus, we get a contradiction. $\square$

By the third property of A , Claim 1, and Claim 2, we see that H satisfies (C3). This finishes the proof. $\square$

Applying the Bourgain-Gamburd method for a random-walk on $\mathrm{SL}_2(\mathbb{F}_p)$, we have proved that (H1) holds for $c_0 < 0.25$, (H2) holds by Helfgott's theorem, and (H4) can be obtained using Kesten's result on random-walks on a free group. So, in order to reach to the desired entropy threshold in $O(\log |G|)$ -steps, it suffices to *escape* proper subgroups that are *not too small*. To that end, we recall Dickson's classification of subgroups of $\mathrm{SL}_2(\mathbb{F}_p)$.

Theorem 59. *Suppose H is a subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$. Then either $|H| \ll 1$ or a conjugate of H is contained in*

$$D := \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{F}_p^\times \right\} \cup \left\{ \begin{pmatrix} 0 & a \\ -a^{-1} & 0 \end{pmatrix} \mid a \in \mathbb{F}_p^\times \right\},$$

or a conjugate of H is contained in

$$B := \left\{ \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix} \mid a \in \mathbb{F}_p^\times, b \in \mathbb{F}_p \right\}.$$

Note that D and B may be described in terms of the action of $\mathrm{SL}_2(\mathbb{F}_p)$ on the projective line $\mathbb{P}^1(\mathbb{F}_p)$. The subgroup D is the stabilizer of the

We consider the action of $\mathrm{SL}_2(\mathbb{C})$ on the projective line $\mathbb{P}^1(\mathbb{C})$, and we identify $\mathbb{P}^1(\mathbb{C})$ with $\mathbb{C} \cup \{\infty\}$. Note that if $a \in \mathrm{SL}_2(\mathbb{R})$ and $\mathrm{Tr}(a) > 2$, then it has two positive real eigenvalues $\alpha > 1$ and $\alpha^{-1} < 1$. Let's assume $a = \mathrm{diag}(\alpha, \alpha^{-1})$. Then a has precisely two fixed points in $\mathbb{C} \cup \{\infty\}$: 0 and ∞ . The point 0 is an *unstable* fixed point of a and ∞ is a *stable* fixed point of a . We also note that for every neighborhoods V_0 and V_∞ of 0 and ∞ , respectively. There exists n_0 such that for every $n \geq n_0$,

$$a^n(\mathbb{P}^1(\mathbb{C}) \setminus V_0) \subseteq V_\infty.$$

Since Γ is Zariski-dense in $(\mathrm{SL}_2)_{\mathbb{Q}}$, there exists $b \in \Gamma$ such that

$$b(\mathrm{Fix}(a)) \cap \mathrm{Fix}(a) = \emptyset.$$

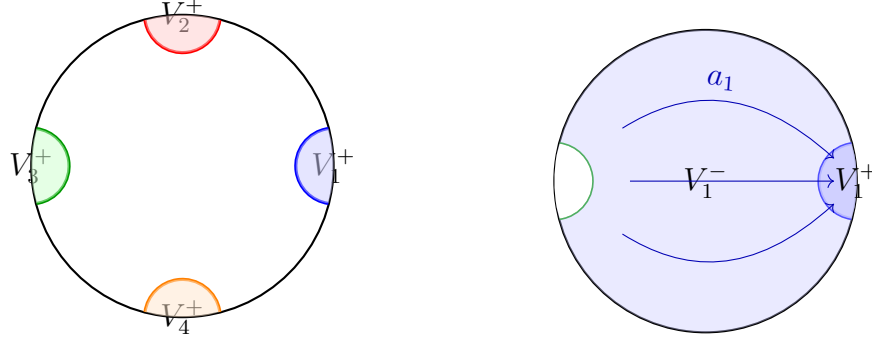
Hence, $\mathrm{Fix}(bab^{-1}) \cup \mathrm{Fix}(a)$ consists of 4 points, and we can consider disjoint neighborhoods V_i^+ of them. Let $\bar{a}_1 := a$, $\bar{a}_2 := bab^{-1}$, $\bar{a}_3 := a^{-1}$, and $\bar{a}_4 := ba^{-1}b^{-1}$. We set the indexing in a way that V_i^+ is a neighborhood of the stable fixed point of a_i . Let V_i^- be the complement of a closed neighborhood of the unstable fixed point of a_i . So, there exists n_0 such that for all $n \geq n_0$,

$$\bar{a}_i^n(V_i^-) \subseteq V_i^+.$$

Moreover, we may make sure that $\bigcup_i V_i^- = \mathbb{P}^1(\mathbb{C})$. Let's summarize what we have proved:

Lemma 60 (Ping-pong players). *Suppose $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ is Zariski-dense. Then, there exist open subsets $V_i^\pm$ of $\mathbb{P}^1(\mathbb{C})$ and $a_1, a_2, a_3 := a_1^{-1}, a_4 := a_2^{-1} \in \Gamma$ with the following properties:*

- (1) V_i^+ 's are disjoint.
- (2) $V_j^+ \subseteq V_i^-$ if $|i - j| \neq 2$.
- (3) $\bigcup_i V_i^- = \mathbb{P}^1(\mathbb{C})$; moreover for every $p \in \mathbb{P}^1(\mathbb{R})$, there exists $i_1 \neq i_2$ such that $p \in V_{i_1}^- \cap V_{i_2}^-$.
- (4) $a_i(V_i^-) \subseteq V_i^+$.



Let's remark that when working with subgroups of $\mathrm{SL}_d(\mathbb{Z})$ with $d > 2$, V_i^- 's are complement of closed neighborhoods of certain subspaces called *repelling spaces* and V_i^+ 's are neighborhoods of contracting subspaces that are of minimum possible dimension.

Proposition 61 (Due to Varjú). *Suppose a_i 's are as in Lemma 60. Let μ be the probability counting measure on $\{a_1, \dots, a_4\}$ and X_ℓ be an ℓ -step random-walk with respect to μ . Then, there exists $C > 1$ such that for every $p \in \mathbb{P}^1(\mathbb{C})$,*

$$\mathbb{P}(X_\ell \cdot p_0 = p_0) < C e^{-C^{-1}\ell}.$$

Note that, by the ping-pong lemma, a_1 and a_2 freely generate a free subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

Lemma 62. *In the above setting, $\langle a_1, a_2 \rangle$ is a free group of rank 2.*

Proof. Suppose $w(x_1, x_2)$ is a non-trivial reduced word in the free group freely generated by x_1 and x_2 . Let $x_3 := x_1^{-1}$ and $x_4 := x_2^{-1}$. Then w may be viewed as a word $x_{i_k} \cdots x_{i_1}$ in the free semigroup freely generated by x_i 's. Pick a point p_0 from $V_{i_1}^- \setminus V_{i_k}^+$. Let $p_{j+1} := a_{i_j} \cdot p_j$. Then, by induction, it follows that $p_j \in V_j^+$ for every $j \geq 1$. In particular, $p_k \in V_{i_k}^+$. Hence, $p_k \neq p_0$, so $w(a_1, a_2, a_3, a_4) \neq 1$. This finishes the proof. $\square$

In the above argument, you see the importance of having a point on the *side* of the next *ping-pong player* in order to predict where the *ball* lands next. Next we show that for every point p_0 most of the words do not fix p_0 .

Lemma 63. *In the above setting, for every point $p_0 \in \mathbb{P}^1(\mathbb{C})$,*

$$\{w \in B_\ell \mid w \cdot p_0 = p_0\} < 2^{\ell+1},$$

where B_ℓ is the set of reduced words of length ℓ in the free group generated by a_1 and a_2 .

(We started the proof, but it is typed as part of the next lecture.)

17. LECTURE 17: PING-PONG ARGUMENT AND ESCAPING PROPER SUBGROUPS, II.

Proof of Lemma 63. Suppose $w = a_{i_\ell} \cdots a_{i_1}$ is in B_ℓ and $w \cdot p_0 = p_0$. Let $p_j := a_{i_j} \cdot p_{j-1}$. Suppose k is the smallest non-negative integer such that p_k is on *side* of the i_{k+1} -player; that means $p_k \in V_{i_{k+1}}^-$ and, for all $j < k$, $p_j \notin V_{i_{j+1}}^-$. By induction, it follows that

$$p_j \in V_{a_{i_j}}^+$$

for every $j \geq k+1$. In particular, $p_0 = p_\ell \in V_{a_{i_\ell}}^+$. Since $V_{a_j}^+$'s are pairwise disjoint and $p_0 \in V_{a_{i_\ell}}^+$, the index i_ℓ is uniquely determined by p_0 . Therefore,

$$p_{\ell-1} = a_{i_\ell}^{-1} \cdot p_0$$

is uniquely determined by $p_\ell = p_0$. By a similar argument and induction, we deduce that all the indexes

$$i_\ell, i_{\ell-1}, \dots, i_{k+1}$$

are uniquely determined by p_0 .

For every $0 \leq j \leq k-1$, p_{i_j} is not in $V_{a_{i_{j+1}}}^-$, which implies that i_{j+1} at most $4-2 = 2$ possibilities. Altogether, we see that such a word w has at most 2^k possibilities. The integer k ranges from 0 to ℓ . Hence,

$$|\{w \in B_\ell \mid w \cdot p_0 = p_0\}| \leq \sum_{k=0}^{\ell} 2^k = 2^{\ell+1} - 1.$$

This finishes the proof. $\square$

Proof of Proposition 61. Note that

$$\mathbb{P}(X_\ell \cdot p_0 = p_0) \leq |\{w \in B_\ell \mid w \cdot p_0 = p_0\}| \max_w \mathbb{P}(X_\ell) \leq |\{w \in B_\ell \mid w \cdot p_0 = p_0\}| \mathbb{P}(X_{2\ell} = 1)^{1/2}.$$

By Lemma 62, (24), and Lemma 63,

$$\mathbb{P}(X_\ell \cdot p_0 = p_0) \leq 2^{\ell+1} \left(\frac{\sqrt{3}}{4}\right)^\ell = 2 \left(\frac{\sqrt{3}}{2}\right)^\ell$$

This finishes the proof. $\square$

To prove that (C3) does not hold in Theorem 58. We must prove a connection between fixing points on $\mathbb{P}^1(\mathbb{F}_p)$ to fixing points on $\mathbb{P}^1(\mathbb{C})$. To this end, we consider *small lifts* of elements of $\mathrm{SL}_2(\mathbb{F}_p)$ that fix $p_0 \in \mathbb{P}^1(\mathbb{F}_p)$. For $\delta > 0$ and $p_0 \in \mathbb{P}^1(\mathbb{F}_p)$, let

$$L_\delta(p_0) := \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \|\gamma\|_\infty < p^\delta, \pi_p(\gamma) \cdot p_0 = p_0\},$$

where $\pi_p: \mathrm{SL}_2(\mathbb{Z}) \rightarrow \mathrm{SL}_2(\mathbb{F}_p)$ is the group homomorphism induced by the residue map modulo p .

Proposition 64. *Then there exists a positive number δ such that, for every sufficiently large prime p , the following statement hold. For every $p_0 \in \mathbb{P}^1(\mathbb{F}_p)$, there exists $q \in \mathbb{P}^1(\mathbb{C})$ such that*

$$\langle L_\delta(p_0) \rangle \cdot q = q.$$

Proof. Note that for every field F , $[x : y] \in \mathbb{P}^1(F)$ is fixed by $\gamma := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ if and only if

$$\frac{ax + by}{cx + dy} = \frac{x}{y} \quad \text{if and only if} \quad by^2 + (a - d)xy - cx^2 = 0.$$

Let $p_\gamma(x, y) := by^2 + (a - d)xy - cx^2$. We further observe that only elements of $F^2 \setminus \{(0, 0)\}$ can represent a point of $\mathbb{P}^1(F)$. A point $(x, y) \in F^2$ is not zero if and only if there exists $(x, y, x', y') \in F^4$ such that $f_0(x, y, x', y') := xx' + yy' = 1$. Hence, the contrary assumption implies that the family of polynomials

$$\{p_\gamma(x, y) \mid \gamma \in L_\delta(p_0)\} \cup \{f_0(x, y, x', y')\}$$

do not have a common zero in $\mathbb{C}$. Then by Nullstellensatz, there exists $q_\gamma, q_0 \in \mathbb{C}[x, y, x', y']$ such that

$$q_0 f_0 + \sum_{\gamma \in L_\delta(p_0)} q_\gamma p_\gamma = 1.$$

It is clear that we may choose polynomials q_γ and q_0 with rational coefficients. There is an effective version of Nullstellensatz which controls the degree and the *height* of these polynomials. The height of a rational number $\frac{m}{n}$ is defined to be $\max\{|m|, |n|\}$, and the height of a polynomial

with rational coefficients is the maximum of the heights of its coefficients. The effective versions of Nullstellensatz imply that we may further assume

$$\deg q_\gamma, \deg q_0 \ll 1 \quad \text{and} \quad H(q_\gamma), H(q_0) \ll \max\{\|\gamma\|_\infty\}^{O(1)}$$

where the bound on the degree is obtained based on the bound on the degrees of p_γ and p_0 , and the number of variables. Clearing the denominators, we deduce that there exist $\tilde{q}_\gamma, \tilde{q}_0 \in \mathbb{Z}[x, y, x', y']$ and $d_0 \in \mathbb{Z}$ such that

$$(33) \quad \tilde{q}_0 f_0 + \sum_{\gamma \in L_\delta(p_0)} \tilde{q}_\gamma p_\gamma = d_0 \quad \text{and} \quad |d_0| \ll p^{C'\delta},$$

where C' is a fixed universal constant. Assuming $p^\delta \gg 1$, we deduce that $|d_0| < p$. Hence looking at (33) modulo p , we deduce that f_0 and p_γ 's do not have a common zero in $\mathbb{F}_p$. This contradicts the assumption that p_0 is a common fixed of elements of $L_\delta(p_0)$ modulo p . $\square$

Proposition 65. *Suppose a_i 's are as in Lemma 60. Then, there exist $\delta' > 0$ such that for every prime p with $p^{\delta'} \gg 1$, $p_0 \in \mathbb{P}^1(\mathbb{F}_p)$, and $\ell > \delta' \log p$,*

$$\mathbb{P}(\pi_p(X_\ell) \cdot p_0) < p^{-\delta'},$$

where π_p is the group homomorphism induced by the residue map modulo p .

Proof. Suppose $\ell_0 := \lfloor c\delta \log p \rfloor$ where c is a sufficiently small fixed constant and δ is the positive number given by Proposition 64. Let $p_\ell := \mathbb{P}(\pi_p(X_\ell) \cdot p_0)$. Then as $\ell_0 \leq \ell$,

$$\begin{aligned} p_\ell &= \sum_g \mathbb{P}(\pi_p(X_{\ell-\ell_0}) = g) \mathbb{P}(g^{-1} \pi_p(X_{\ell_0}) \cdot p_0 = p_0) \\ &\leq \max_g \mathbb{P}(g^{-1} \pi_p(X_{\ell_0}) \cdot p_0 = p_0) \\ &\leq \mathbb{P}(\pi_p(X_{2\ell_0}) \cdot p_0)^{1/2}. \end{aligned}$$

Choosing c small enough, we obtain

$$\mathbb{P}(X_{2\ell_0} = \gamma) = \mathbb{P}(\pi_p(X_{2\ell_0}) = \pi_p(\gamma))$$

for all $\gamma \in \text{SL}_2(\mathbb{Z})$, and we further deduce that

$$\mathbb{P}(\pi_p(X_{2\ell_0}) \cdot p_0) \leq \mathbb{P}(X_{2\ell_0} \in L_\delta(p_0)).$$

By Proposition 64 and Proposition 61, it follows that

$$\mathbb{P}(X_{2\ell_0} \in L_\delta(p_0)) < C e^{-2C^{-1}\ell_0} = C e^{-2C^{-1}\lfloor c\delta \log p \rfloor} < p^{-C^{-1}c\delta},$$

where the last inequality holds when p^δ is sufficiently large. Altogether, we deduce that

$$p_\ell < p^{-(2C)^{-1}c\delta}.$$

Hence, $\delta' := (2C)^{-1}c\delta$ satisfies the desired property. This finishes the proof. $\square$

Now, we can put all the pieces together and finish the proof of Bourgain-Gamburd's result.

Theorem 66 (Bourgain-Gamburd). *Suppose S is a finite symmetric subset of $\text{SL}_2(\mathbb{Z})$ and the group Γ generated by S is Zariski-dense in $(\text{SL}_2)_\mathbb{Q}$. Then the family of Cayley graphs*

$$\{\text{Cay}(\pi_p(\Gamma), \pi_p(S))\}$$

as p ranges over primes is a family of expanders.

Proof. Let $a_1, a_2, a_3, a_4 \in \Gamma$ be the *ping-pong players* as in Lemma 60. By Proposition 65 and Dickson's classification of proper subgroups of $\mathrm{SL}_2(\mathbb{F}_p)$, for sufficiently large p , $\pi_p(a_i)$'s generate $\mathrm{SL}_2(\mathbb{F}_p)$. Let X_ℓ be an ℓ -step random-walk with respect to the probability counting measure on a_i 's. By Lemma 62, it follows that

$$(\text{Initial entropy}) \quad H_\infty(\pi_p(X_{\ell_0})) > c_1 \log |\mathrm{SL}_2(\mathbb{F}_p)|$$

where c_1 is a universal fixed positive number and ℓ_0 is a positive integer less than $c_2 \log |\mathrm{SL}_2(\mathbb{F}_p)|$. Let γ be the positive number given by Theorem 58. Assuming $p^\gamma \gg 1$, (C0) (exceptional cases) cannot hold. We may assume that $C\gamma < \delta'$ where δ' is the positive number given by Proposition 65. Then, by Proposition 65, (C3) (being trapped in a proper subgroup) cannot hold. Hence, we gain $\gamma \log |\mathrm{SL}_2(\mathbb{F}_p)|$ entropy and after at most γ^{-1} doubling of the number of steps, we gain *enough entropy*; that means (C1) holds. Therefore by Proposition 33, $\mathcal{L}(\pi_p(X)) > c' > 0$ for some fixed universal constant c' . Thus, by Proposition 1,

$$\kappa(\pi_p(X); L^2(\mathrm{SL}_2(\mathbb{F}_p))^\circ) \geq 1 - e^{-c'}$$

for sufficiently large p . There exists k_0 such that

$$\{a_1, a_2\} \subseteq \prod_{k_0} S.$$

By the same argument as in the proof of Lemma 2, for every $f \in L^2(\mathrm{SL}_2(\mathbb{F}_p))^\circ$,

$$\|a_i \cdot f - f\|_2 \leq k_0 \delta_S(f).$$

Hence,

$$\kappa(S; L^2(\mathrm{SL}_2(\mathbb{F}_p))^\circ) \geq \kappa(\pi_p(X); L^2(\mathrm{SL}_2(\mathbb{F}_p))^\circ) / k_0 > (1 - e^{-c'}) / k_0.$$

This finishes the proof. $\square$

Our next goal is to investigate random-walks on group extensions. Lindenstrauss and Varjú proved that if S is a symmetric generating set of $\mathrm{SL}_n(\mathbb{F}_p) \ltimes \mathbb{F}_p^n$, then one can bound the Lyapunov exponent $\mathcal{L}(S)$ in terms of bounds for $|S|$ and the Lyapunov exponent of the quotient $\mathrm{pr}_{\mathrm{SL}_n(\mathbb{F}_p)}(S)$ to the simple part. They used Fourier analysis on the finite abelian group $\mathbb{F}_p^n$ to prove this result. They further conjectured that an analogous statement should hold for $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$: bounds for the Lyapunov exponents of the marginals, together with a bound on the size of the generating set, should give us a bound on the Lyapunov exponent. Srinivas and I proved this conjecture. In the remaining part of the course, I will present the proofs of these results.

18. LECTURE 18: RANDOM-WALK ON GROUP EXTENSIONS, I

Our goal is to prove the following theorems.

Theorem 67. *Suppose A is a simple $\mathbb{F}_p[H]$ -module, where $H = \mathrm{SL}_2(\mathbb{F}_p)$. Let $X = (Y, Z)$ be a symmetric random-variable with values in $G := H \ltimes A$. Suppose $\mathcal{L}(Y) > \varepsilon_0 > 0$ and $\mathbb{P}(X = g) > e_0$ for every g in the range of X . Then there exists $\varepsilon > 0$ depending only on ε_0 and e_0 such that $\mathcal{L}(X) > \varepsilon$.*

Theorem 68. *Let $X = (X_L, X_R)$ be a random-variable with values in $G := H \times H$, where $H := \mathrm{SL}_2(\mathbb{F}_p)$. Suppose $\mathcal{L}(X_L), \mathcal{L}(X_R) > \varepsilon_0 > 0$ and $\mathbb{P}(X = g) > e_0$ for every g in the range of X . Then there exists $\varepsilon > 0$ depending only on ε_0 and e_0 such that $\mathcal{L}(X) > \varepsilon$.*

Lindenstrauss and Varjú proved theorem 67 for $A = \mathbb{F}_p^2$.

We start by exploring what we can gain using the Bourgain-Gamburd method.

Is G a c_0 -quasi-random group where c_0 is a universal fix number?

Lemma 69. *The group G is perfect.*

Proof. Note that $[d(a), u^+(b)] = u^+((a^2 - 1)b)$ for every $a \in \mathbb{F}_p^\times$ and $b \in \mathbb{F}_p$, where

$$d(a) := \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \quad \text{and} \quad u^+(b) := \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}.$$

Since the abelian subgroup of $\mathbb{F}_p$ generated by $(a^2 - 1)b$ as (a, b) ranges in $\mathbb{F}_p^\times \times \mathbb{F}_p$ is $\mathbb{F}_p$, we deduce that the upper triangular unipotent matrices are in the derived subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$. Similarly, the lower triangular matrices are in that group. Since $\mathrm{SL}_2(\mathbb{F}_p)$ is generated by upper and lower unipotent subgroups, we deduce that $H := \mathrm{SL}_2(\mathbb{F}_p)$ is perfect. Hence, $H \times H$ is also perfect.

Let $N := [G, G]$ where $G := H \ltimes V$. Then $H \subseteq N$ as H is perfect. For every $v \in V$ and $h \in H$,

$$(34) \quad (h, 0)(1, v)(h, 0)^{-1}(1, -v) = (1, h \cdot v - v) \in N.$$

Hence, the $\mathbb{F}_p$ -span W of $h \cdot v - v$'s is in N . Note that for every $h' \in H$ and $w \in W$,

$$\begin{aligned} h' \cdot w &= h' \cdot \left(\sum_{h \in H} c_h (h - 1) \cdot v \right) \\ &= \sum_{h \in H} c_h ((h'h - 1) - (h' - 1)) \cdot v \\ &= \sum_{h'' \in H} c_{h'^{-1}h''} (h'' - 1) \cdot v - \left(\sum_{h \in H} c_h \right) (h' - 1) \cdot v \in W. \end{aligned}$$

Hence, W is a non-zero submodule of V , which implies $W = V$. Therefore $W \subseteq N$. Therefore $N = G$, so G is perfect. $\square$

Proposition 70. *Suppose N is a proper normal subgroup of $H := \mathrm{SL}_2(\mathbb{F}_p)$. Then $N \subseteq \{\pm 1\}$.*

Proof. Let $\omega_0 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and B denote the subgroup of upper-triangular matrices of H . Then

$$(\text{Bruhat decomposition}) \quad H = B \sqcup B\omega_0 B.$$

To see this, consider the action of $\mathrm{SL}_2(\mathbb{F}_p)$ on the projective line $\mathbb{P}^1(\mathbb{F}_p)$. Note that this action is transitive and the stabilizer subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$ associated with $[e_1]$ is B . Also note that $\mathbb{P}^1(\mathbb{F}_p) \setminus \{[e_1]\}$ may be identified with the affine line $\mathbb{A}^1(\mathbb{F}_p)$ and $U^+ \subseteq B$ acts simply transitively on this set. So, for every $h \in \mathrm{SL}_2(\mathbb{F}_p)$, if $h[e_1] = [e_1]$, then $h \in B$. If not, then there exists u^+ such that

$$h[e_1] = u^+[e_2] = u^+\omega_0[e_1],$$

which implies that

$$h \in u^+\omega_0 B \subseteq B\omega_0 B.$$

Since double cosets of B form a partition of H , we deduce that this special case of the Bruhat decomposition.

By the Bruhat decomposition, it is clear that B is a maximal subgroup of H : if $B \subsetneq K$ and K is a subgroup of H , then $\omega_0 \in K$, which implies that $K = H$.

Note that

$$\omega_0 B \omega_0^{-1} \cap B = \{\mathrm{diag}(a, a^{-1}) \mid a \in \mathbb{F}_p^{-1}\},$$

so $\omega_0 B \omega_0^{-1} \cap B$ fixes the set $\{[e_1], [e_2]\}$ pointwise. Therefore, $u^+(\omega_0 B \omega_0^{-1} \cap B)u^{+^{-1}}$ fixes the set $\{u^+[e_1], u^+[e_2]\}$ pointwise. Therefore, $x \in \bigcap_{h \in H} h B h^{-1}$ fixes all the points of $\mathbb{P}^1(\mathbb{F}_p)$. Hence, x is a scalar matrix. Thus

$$(35) \quad \bigcap_{h \in H} h B h^{-1} = \{\pm 1\}.$$

Suppose N is a normal subgroup of H . If $N \subseteq B$, then by (35), $N \subseteq \{\pm 1\}$. So, we may assume that $N \not\subseteq B$. Therefore $BN = H$ as B is a maximal subgroup of H and BN is a subgroup which properly contains B . We deduce that

$$H/N = BN/N.$$

Note that U^+ is a normal subgroup of B , so U^+N/N is a normal subgroup of H/N . Therefore,

$$U^+N = \omega_0 U^+ N \omega_0^{-1} = U^-N.$$

Since $H = \langle U^+, U^- \rangle$, it follows that

$$H/N = U^+N/N \simeq U^+/(U^+ \cap N)$$

is an abelian group. Since H is perfect, $N = H$. This finishes the proof. $\square$

Lemma 71. *Suppose N is a proper normal subgroup of $G := H \rtimes V$ in the above setting. Then either $V \subseteq N \subseteq \{\pm 1\} \times V$ or $N \subseteq \{\pm 1\}$.*

Proof. Note that V is normal subgroup of G . So, NV/V is a normal subgroup of $G/V \simeq H$. Hence, either $NV = G$ or by Proposition 70 $N \subseteq \{\pm 1\} \times V$. Suppose $NV = G$. Then, $V \not\subseteq N$, so $N \cap V$ is a proper $\mathbb{F}_p[H]$ -submodule of V . Therefore, $N \cap V$ is trivial. For every $h \in H$, there exists $v(h) \in V$ such that $n(h) := (h, v(h)) \in N$. Then, for every $v' \in V$,

$$(36) \quad [n(h), (1, v')] \in N \cap V \quad \text{which implies that} \\ (h, 1)(1, v(h))(1, v')(1, -v(h))(h^{-1}, 1)(1, -v') = (1, (h-1) \cdot v') \in N \cap V.$$

So, we may assume that $N \subseteq \{\pm 1\} \times V$. Note that $V \cap N$ is a submodule of V . Hence, either $V \subseteq N$ or $V \cap N$ is trivial. In the latter, $|N| \leq 2$, so $N \subseteq \{\pm 1\}$. This finishes the proof. $\square$

Lemma 72. *The groups $H \times H$ and $H \rtimes V$ are c_0 -quasi-random where c_0 depends only on $\dim_{\mathbb{F}_p} V$.*

Proof. Suppose π is a non-trivial representation of G . Let N be the kernel of π . Then $N \cap H \times \{1\} \subseteq \{\pm 1\}$ or $N \cap \{1\} \times H \subseteq \{\pm 1\}$ or $N \cap H \subseteq \{\pm 1\}$. Therefore, the restriction of π to a copy of H is non-trivial. Thus, $\dim \pi \geq \frac{p-1}{2}$. Since $|G| \leq p^{d_0}$ where $d_0 = 6$ or $d_0 = 3 + \dim_{\mathbb{F}_p} V$. This finishes the proof. $\square$

By (7), we obtain that, for $\ell_0 := \lceil \frac{4}{\varepsilon_0} \rceil \log |H|$ and every $h \in H$,

$$(37) \quad \left| \mathbb{P}(Y_{\ell_0} = h) - \frac{1}{|H|} \right| \leq \frac{1}{|H|^2},$$

where Y is X_L or X_R in the setting of Theorem 68. In particular,

$$(38) \quad (\text{Initial entropy}) \quad H_\infty(X_{\ell_0}) > \log |H| - \log 2 > c_1 \log |G|,$$

where c_1 depends only on $\dim_{\mathbb{F}_p} V$ and p is sufficiently large.

Following the Bourgain-Gamburd method, for a fixed small enough γ , one of the following statements hold:

- (1) (Enough entropy) $H_2(X_{2^k \ell_0}) > (1 - \frac{c_0}{2}) \log |G|$.
- (2) (Gaining entropy) $H_2(X_{2^{k+1} \ell_0}) > H_2(X_{2^k \ell_0}) + \gamma \log |G|$.
- (3) (Trapped in an approximate subgroup) There exists $A \subseteq G$ such that
 - (a) A is $|G|^{C\gamma}$ -approximate subgroup.
 - (b) $|\log |A| - H_2(X_{2^k \ell_0})| < C\gamma \log |G|$.
 - (c) $\mathbb{P}(X_{2^{k+1} \ell_0} \in AA) > |G|^{-C\gamma}$.

In the next lecture, we will focus on the third possibility, and interpret it as *fixing a point* with high probability under a well-chosen group action.

19. ADDITIONAL READINGS

Measure spaces. A measure space is a triple $(X, \mathcal{B}, \mu)$, where X is a non-empty set, $\mathcal{B} \subseteq P(X)$ which is a σ -algebra, meaning that:

- (1) $X \in \mathcal{B}$,
- (2) for every $A \in \mathcal{B}$, $X \setminus A \in \mathcal{B}$,
- (3) for every countable subset $\{A_i \mid i \in \mathbb{N}\}$ of $\mathcal{B}$, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$,

and $\mu: \mathcal{B} \rightarrow [0, \infty]$ is a measure, meaning that:

- (1) $\mu(\emptyset) = 0$, $\mu(X) \neq 0$,
- (2) if $A_i \in \mathcal{B}$ are pairwise disjoint, then $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

A subset of X is called μ -measurable if it is an element of $\mathcal{B}$.

When X is a topological space, the σ -algebra generated by open sets is called the Borel σ -algebra. A measure $(X, \mathcal{B}, \mu)$ on a topological space X is called *Borel-regular outer measure* if for every $A \in \mathcal{B}$ there exists a Borel set A' such that $A \subseteq A'$ and $\mu(A) = \mu(A')$. Assume that every closed set of X is a countable intersection of open sets. For instance this property holds for a metric space. Then the following important regularity holds:

Theorem 73. *Suppose μ is a Borel-regular outer measure on a topological space as above. Suppose there exists open sets $\{\mathcal{O}_i\}_{i=1}^{\infty}$ such that $X = \bigcup_{i=1}^{\infty} \mathcal{O}_i$ and $\mu(\mathcal{O}_i) < \infty$. Then for every μ -measurable set A ,*

$$\begin{aligned}\mu(A) &= \sup\{\mu(C) \mid C \subseteq A, C \text{ is closed}\} \\ \mu(A) &= \inf\{\mu(\mathcal{O}) \mid A \subseteq \mathcal{O}, \mathcal{O} \text{ is open}\}.\end{aligned}$$

Exercise 1. Prove that there is no measure μ defined on the entire power set $P(\mathbb{R})$ with the following properties:

- (1) for all $x \in \mathbb{R}$ and $A \subseteq \mathbb{R}$, $\mu(A + x) = \mu(A)$, and
- (2) for all $a < b$, $\mu([a, b]) := b - a$.

(**Hint.** Take a set Y of coset representatives of $\mathbb{R}/\mathbb{Q}$ and cover $[0, 1]$ using shifts Y by $\mathbb{Q}$.)

Self-adjoint compact operators.

20. PROBLEMS

20.1. Property(T) and finite abelianization.

- (1) Suppose Γ is a finitely generated group that has property(T). Prove that every quotient of Γ has property(T).

(**Hint.** Note that if π is a unitary representation of Γ/N with no non-zero fixed points, then $\pi \circ \pi_N$ is a unitary representation of Γ with no non-zero vectors.)

- (2) Prove that $\mathbb{Z}$ does not have property(T).

(**Hint.** Let $\pi_n: \mathbb{Z} \rightarrow S^1$, $\pi_n(k) := e^{2\pi i k/n}$ be a one-dimensional unitary representation of $\mathbb{Z}$. Note that $\delta_{\{1\}}(1) = |e^{2\pi i/n} - 1|$ and deduce that $\lim_{n \rightarrow \infty} \kappa(\{1\}; \pi_n) = 0$.)

- (3) Prove that if Γ is a finitely generated group that has property(T), then $\Gamma/[\Gamma, \Gamma]$ is finite.

20.2. GNS Construction (Gelfand, Naimark, and Segal). Suppose X is a topological space. A continuous function $\Phi: X \times X \rightarrow \mathbb{C}$ is called a *kernel of positive type* if $\Phi(x_1, x_2) = \overline{\Phi(x_2, x_1)}$,

for every $c_1, \dots, c_n \in \mathbb{C}$ and for every $x_1, \dots, x_n \in X$,

$$(c_1 \ \cdots \ c_n) \begin{pmatrix} \Phi(x_1, x_1) & \cdots & \Phi(x_1, x_n) \\ \vdots & \ddots & \vdots \\ \Phi(x_n, x_1) & \cdots & \Phi(x_n, x_n) \end{pmatrix} \begin{pmatrix} \bar{c}_1 \\ \vdots \\ \bar{c}_n \end{pmatrix} = \sum_{i,j} c_i \bar{c}_j \Phi(x_i, x_j) \geq 0.$$

A continuous function $\Psi: X \times X \rightarrow \mathbb{R}$ is called a *kernel of conditionally of negative type* if $\Psi(x, x) = 0$, $\Psi(x_1, x_2) = \overline{\Psi(x_2, x_1)}$, and for every $c_1, \dots, c_n \in \mathbb{R}$ with $\sum_{i=1}^n c_i = 0$ and for every $x_1, \dots, x_n \in X$,

$$\sum_{i,j} c_i c_j \Psi(x_i, x_j) \leq 0.$$

Note that for conditionally of negative type coefficients should be real.

- (1) Suppose $f: X \rightarrow \mathcal{H}$ is a continuous function where $\mathcal{H}$ is a Hilbert space. Prove that

$$\Phi(x_1, x_2) := f(x_1) \cdot f(x_2)$$

is a kernel of positive type.

(**Hint.** Note that $\sum_{i,j} c_i \bar{c}_j \Phi(x_i, x_j) = \|\sum_i c_i f(x_i)\|^2$.)

- (2) Suppose Φ is a kernel of positive type. For each $x \in X$, let $\phi_x: X \rightarrow \mathbb{C}$, $\phi_x(x') := \Phi(x, x')$. Let V be the space spanned by ϕ_x 's. Define

$$\left\langle \sum_{i=1}^n a_i \phi_{x_i}, \sum_{j=1}^n b_j \phi_{x_j} \right\rangle := \sum_{i,j} a_i \bar{b}_j \Phi(x_i, x_j).$$

Prove that $\langle, \rangle$ is a well-defined inner product.

(**Hint.** Let $f := \sum_{i=1}^n a_i \phi_{x_i}$ and $g := \sum_{i=1}^n b_i \phi_{x_i}$. Show that

$$\langle f, g \rangle = \sum_i a_i \overline{g(x_i)} = \sum_j \bar{b}_j f(x_j),$$

and deduce that $\langle, \rangle$ is well-defined. Clearly, $\langle, \rangle$ is semilinear and $\langle f, f \rangle \geq 0$ for all f .

Note that for all $f \in V$, $\langle f, \phi_x \rangle = f(x)$, and so by the Cauchy-Schwarz inequality

$$(39) \quad |f(x)|^2 \leq \Phi(x, x) \langle f, f \rangle.$$

Deduce that if $\langle f, f \rangle = 0$, then $f = 0$.)

- (3) (Existence) Suppose Φ is a kernel of positive type. Then there exist a Hilbert space $\mathcal{H}$ and a continuous function $f: X \rightarrow \mathcal{H}$ such that the space spanned by $\{f(x)\}_{x \in X}$ is dense in $\mathcal{H}$ and $\Phi(x_1, x_2) = f(x_1) \cdot f(x_2)$.

(**Hint.** Let $(V, \langle, \rangle)$ be as in Problem(2). If $\{\psi_n\}_n$ is a Cauchy sequence with respect to the norm given by the inner product, then by (39), $\{\psi_n(\gamma)\}_n$ is a Cauchy sequence. Hence, the completion $\mathcal{H}$ of V can be viewed as a subset of functions from X to $\mathbb{C}$. Let

$$f: X \rightarrow \mathcal{H}, \quad f(x) := \phi_x.$$

Then

$$\langle f(x_1), f(x_2) \rangle = \langle \phi_{x_1}, \phi_{x_2} \rangle = \Phi(x_1, x_2).$$

To show that f is continuous, note that

$$\|f(x_1) - f(x_2)\|^2 = \Phi(x_1, x_1) - 2\operatorname{Re}\Phi(x_1, x_2) + \Phi(x_2, x_2),$$

and Φ is continuous.

- (4) (Uniqueness) Suppose $\mathcal{H}_i$ is a Hilbert space, $f_i: X \rightarrow \mathcal{H}_i$ is a continuous function, and the span of $\{f_i(x) \mid x \in X\}$ is dense on $\mathcal{H}_i$. Suppose for every $x, y \in X$,

$$\langle f_1(x), f_1(y) \rangle_{\mathcal{H}_1} = \langle f_2(x), f_2(y) \rangle_{\mathcal{H}_2}.$$

Prove that there exists an isomorphism $U: \mathcal{H}_1 \rightarrow \mathcal{H}_2$ such that $f_2 = U \circ f_1$. This means the following is a commuting diagram and the Hilbert space constructed in Problem (3) is unique up to an isomorphism.

$$\begin{array}{ccc} & X & \\ f_1 \swarrow & & \searrow f_2 \\ \mathcal{H}_1 & \xrightarrow{U} & \mathcal{H}_2 \end{array}$$

(**Hint.** Let V_i be the span of $\{f_i(x) \mid x \in X\}$. Argue why $\|\sum_j c_j f_1(x_j)\|^2 = \|\sum_j c_j f_2(x_j)\|^2$, and deduce that $\sum_j c_j f_1(x_j) = 0$ if and only if $\sum_j c_j f_2(x_j) = 0$. This implies that the following is a well-defined isomorphism

$$U: V_1 \rightarrow V_2, \quad U\left(\sum_j c_j f_1(x_j)\right) := \sum_j c_j f_2(x_j),$$

which preserves the norm. Since $\mathcal{H}_2$ is a complete space and V_i is dense in $\mathcal{H}_i$, we obtain an extension $U: \mathcal{H}_1 \rightarrow \mathcal{H}_2$ which is an isomorphism of Hilbert spaces such that $U \circ f_1 = f_2$.)

- (5) Suppose Φ_1 and Φ_2 are of positive type. Then $\Phi_1 \Phi_2$ is of positive type.

(**Hint.** Note that H is a semi-positive definite matrix if and only if $H = AA^*$. Hence there are vectors $v_1, \dots, v_n$ and $w_1, \dots, w_n$ such that

$$\Phi_1(x_i, x_j) = v_i \cdot v_j \quad \text{and} \quad \Phi_2(x_i, x_j) = w_i \cdot w_j.$$

Then

$$\begin{aligned} \sum_{i,j} c_i \bar{c}_j \Phi_1(x_i, x_j) \Phi_2(x_i, x_j) &= \sum_{i,j} c_i \bar{c}_j (v_i \cdot v_j) (w_i \cdot w_j) \\ &= \sum_{i,j} c_i \bar{c}_j (v_i \otimes w_i) \cdot (v_j \otimes w_j) \\ &= \left\| \sum_i c_i (v_i \otimes w_i) \right\|^2 \geq 0. \end{aligned}$$

Here is an alternative approach: for $i = 1, 2$, there exists $f_i: X \rightarrow \mathcal{H}_i$ such that

$$\Phi_i(x, x') = f_i(x) \cdot f_i(x').$$

Let $f: X \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$, $f(x) := f_1(x) \otimes f_2(x)$, and observe that

$$\Phi_1(x, x') \Phi_2(x, x') = f(x) \cdot f(x').$$

)

- (6) Suppose $f: X \rightarrow \mathcal{H}$ is a continuous function where $\mathcal{H}$ is a real Hilbert space. Prove that $\Psi(x_1, x_2) := \|f(x_1) - f(x_2)\|^2$ is conditionally of negative type.

(**Hint.** Show that $\sum_{i,j} c_i c_j \|f(x_i) - f(x_j)\|^2 = -2 \left\| \sum_i c_i f(x_i) \right\|^2$ if $\sum_i c_i = 0$.)

- (7) Suppose $\Psi: X \rightarrow \mathbb{R}$ is a kernel of conditionally of negative type. Let

$$\widehat{V} := \{\psi: X \rightarrow \mathbb{C} \mid |\text{supp}(\psi)| < \infty, \sum_{x \in X} \psi(x) = 0\}.$$

Note that

$$\langle \psi_1, \psi_2 \rangle := - \sum_{x, x'} \psi_1(x) \psi_2(x') \Psi(x, x')$$

is a well-defined semi-inner product. Let

$$K := \{\psi \in \widehat{V} \mid \langle \psi, \psi \rangle = 0\}.$$

Prove that K is a subspace, $\langle [\psi_1], [\psi_2] \rangle := \langle \psi_1, \psi_2 \rangle$ is a well-defined inner product on $V := \widehat{V}/K$.

(Hint. For every $x \in X$, let $\delta_x: X \rightarrow \mathbb{C}$ be the characteristic function of $\{x\}$. Then $\widehat{V}$ is the span of $\{\delta_x - \delta_1\}_{x \in \Gamma}$. By the Cauchy-Schwarz inequality

$$|\langle \psi, \delta_x - \delta_1 \rangle|^2 \leq \langle \psi, \psi \rangle \langle \delta_x - \delta_1, \delta_x - \delta_1 \rangle.$$

Deduce that K is a subspace and $\langle \widehat{V}, K \rangle = 0$.)

- (8) Suppose $\Psi: X \rightarrow \mathbb{R}$ is a kernel of conditionally of negative type and $(V, \langle, \rangle)$ is the space given in Problem 5. Let $\mathcal{H}$ be the real Hilbert space obtain by the completion of V . Define

$$f: X \rightarrow \mathcal{H}, \quad f(x) := [\delta_x - \delta_1].$$

Prove that f is continuous and

$$\Psi(x, x') = \frac{1}{2} \|f(x) - f(x')\|^2.$$

(Hint. $\|f(x) - f(x')\|^2 = \langle \delta_x - \delta_{x'}, \delta_x - \delta_{x'} \rangle = 2\Psi(x, x')$.)

- (9) Suppose $\Psi: X \rightarrow \mathbb{R}$ is conditionally of negative type. Let

$$\Phi(x, x') := \Psi(x, 1) + \Psi(x', 1) - \Psi(x, x').$$

Prove that Φ is of positive type.

(Hint. There exists $f: X \rightarrow \mathcal{H}$ such that $\Psi(x, x') = \|f(x) - f(x')\|^2$. Show that

$$\sum_{i,j} c_i c_j \Phi(x_i, x_j) = 2 \left\| \sum_i c_i f(x_i) \right\|^2 \geq 0$$

for all $c_i \in \mathbb{R}$.)

- (10) (Schoenberg) Suppose $\Psi: X \rightarrow \mathbb{R}$ is a kernel of conditionally of negative type. Prove that for every $t \geq 0$, $\exp(-t\Psi)$ is of positive type.

(Hint. First note that it is sufficient to prove the claim for $t = 1$. Let Φ be the function defined in Problem 8. Then by Problem 5, $\exp(\Phi)$ is of positive type. Let $a_i := e^{\Psi(x_i, 1)}$. Then

$$\exp(-\Psi(x_i, x_j)) = a_i^{-1} a_j^{-1} \exp(\Phi(x_i, x_j)).$$

Hence $\sum_{i,j} c_i c_j \exp(-\Psi(x_i, x_j)) = \sum_{i,j} (c_i a_i^{-1}) (c_j a_j^{-1}) \exp(\Phi(x_i, x_j)) \geq 0$.)

- (11) (GNS Construction) Suppose G is a topological group. We say $\phi: G \rightarrow \mathbb{C}$ is of positive type if $\Phi(x, x') := \phi(x'^{-1}x)$ is a kernel of positive type. Then there exist a unitary representation $\pi: G \rightarrow \mathcal{U}(\mathcal{H})$ and $v \in \mathcal{H}$ such that

$$\phi(g) = \langle \pi(g)(v), v \rangle,$$

$g \mapsto \pi(g)(v)$ is continuous for all $v \in \mathcal{H}$, and the span of $\pi(G)(v)$ is dense in $\mathcal{H}$. Moreover $(\mathcal{H}, \pi, v)$ is unique up to an isomorphism.

(Hint. Since $\Phi(x, x')$ is of positive type, there exists a continuous function $f: G \rightarrow \mathcal{H}$ such that

$$\phi(x'^{-1}x) = \Phi(x, x') = \langle f(x), f(x') \rangle,$$

and the span of $\{f(x)\}_x$ is dense in $\mathcal{H}$. Note that $\Phi(xx_1, xx_2) = \Phi(x_1, x_2)$, and so $\Phi(x_1, x_2) = \langle f(xx_1), f(xx_2) \rangle$. Therefore, the continuous function $\lambda(x^{-1})(f): G \rightarrow \mathcal{H}$ also satisfies the same properties as f . By the uniqueness, there exists an isomorphism $u_x: \mathcal{H} \rightarrow \mathcal{H}$ such that $u_x \circ f = \lambda(x^{-1})(f)$. Observe that

$$u_x \circ u_{x'}(f(x'')) = u_x(f(x'x'')) = f(xx'x'') = u_{xx'}(f(x''')).$$

Since the span of $\{f(x) \mid x \in G\}$ is dense in $\mathcal{H}$, deduce that $u_x \circ u_{x'} = u_{xx'}$. Hence, $\pi: G \rightarrow \mathcal{U}(\mathcal{H}), \pi(x) := u_x$ is a well-defined group homomorphism. Note that

$$G \rightarrow \mathcal{H}, x \mapsto \pi(x)(f(x')) = f(xx')$$

is continuous for all $x' \in G$. Deduce that π is strongly continuous.)

(Remark. We say a unitary representation $\pi: G \rightarrow \mathcal{U}(\mathcal{H})$ is *cyclic* if there exists $v \in \mathcal{H}$ such that the span of $\pi(G)(v)$ is dense in $\mathcal{H}$. So there exists a bijection between functions of positive type of G and cyclic unitary representations above representation is cyclic.)

- (12) (1-parameter unitary cyclic representations) Suppose $\psi: G \rightarrow \mathbb{R}$ is a function such that $\Psi(g, g') := \psi(g'^{-1}g)$ is conditionally of negative type. Then for every $t \geq 0$, there exist $(\mathcal{H}_t, \pi_t, v_t)$ such that
- (a) $\mathcal{H}_t$ is a Hilbert space, $\pi_t: G \rightarrow \mathcal{U}(\mathcal{H}_t)$ is a strongly continuous unitary representation, $v_t \in \mathcal{H}_t$,
 - (b) The span of $\pi_t(G)(v_t)$ is dense in $\mathcal{H}_t$.
 - (c) For all $x \in G$, $\langle \pi_t(x)(v_t), v_t \rangle = \exp(-t\psi(x))$.

(Hint. Immediate consequence of Schoenberg's theorem and the GNS construction.)

20.3. Property(FH). Suppose Γ is a finitely generated group. An affine isometric action of Γ on a real Hilbert space $\mathcal{H}$ is an action such that

$$(40) \quad \gamma \cdot x = \pi(\gamma)(x) + a(\gamma)$$

where $\pi: \Gamma \rightarrow O(\mathcal{H})$ is a group homomorphism and $a(\gamma) \in \mathcal{H}$. We say Γ has property(FH) if every affine isometric action of Γ on a real Hilbert space has a fixed point. Note that for a representation π and a function a , (40) defines an action if and only if for every $\gamma_1, \gamma_2 \in \Gamma$,

$$(41) \quad (1\text{-Cocycle relation}) \quad a(\gamma_1\gamma_2) = \pi(\gamma_1)(a(\gamma_2)) + a(\gamma_1).$$

A function $a: \Gamma \rightarrow \mathcal{H}$ which satisfies (41) is called an 1-cocycle. The set of all 1-cocycles is denoted by $Z^1(\Gamma, \pi)$. Note that $Z^1(\Gamma, \pi)$ is a group under addition.

We also observe that an action given by (40) fixes x_0 if and only if for all $\gamma \in \Gamma$,

$$(42) \quad (1\text{-coboundary}) \quad a(\gamma) = x_0 - \pi(\gamma)(x_0).$$

For $x_0 \in \mathcal{H}$, let $b_{x_0}: \Gamma \rightarrow \mathcal{H}$, $b_{x_0}(\gamma) := x_0 - \pi(\gamma)(x_0)$. For every x_0 , the function b_{x_0} is called a 1-coboundary. The set of all the 1-coboundaries is denoted by $B^1(\Gamma, \pi)$. It is easy to see that $B^1(\Gamma, \pi)$ is a subgroup of $Z^1(\Gamma, \pi)$. We let

$$(\text{First cohomology group}) \quad H^1(\Gamma, \pi) := \frac{Z^1(\Gamma, \pi)}{B^1(\Gamma, \pi)}.$$

The goal of this sequence of problems is to investigate property(FH). In particular, we want to show that property(T) implies property(FH). In fact, these properties are equivalent.

- (1) Prove that Γ has property(FH) if and only if $H^1(\Gamma, \pi) = 0$ for every orthogonal representation $\pi: \Gamma \rightarrow O(\mathcal{H})$.
- (2) Suppose $C \subseteq \mathcal{H}$ is a bounded subset. Let

$$r_C := \inf\{r \mid \exists x \in \mathcal{H}, B(x, r) \supseteq C\}.$$

Prove that there exists a unique $x(C) \in \mathcal{H}$ such that $C \subseteq B(x(C), r(C))$, where $B(x, r)$ is the closed ball of radius r about x .

(Hint. Suppose $C \subseteq B(x, r(C)) \cap B(x', r(C))$. Use

$$\begin{aligned} \|v + w\|^2 + \|v - w\|^2 &= 2(\|v\|^2 + \|w\|^2) \quad \text{to deduce} \\ \|(y - x) + (y - x')\|^2 + \|x - x'\|^2 &= 2(\|y - x\|^2 + \|y - x'\|^2). \end{aligned}$$

Therefore

$$\frac{\|y - x\|^2 + \|y - x'\|^2}{2} = \left\|y - \left(\frac{x + x'}{2}\right)\right\|^2 + \left\|\frac{x - x'}{2}\right\|^2.$$

Show that $C \subseteq B(\frac{x+x'}{2}, r(C) - \|\frac{x-x'}{2}\|^2)$ and get a contradiction.)

(Remark. This argument works for CAT(0) spaces: a metric space X such that the following properties hold:

- (a) For every two points $p, p' \in X$, there exists a midpoint m ; it means

$$d(p, m) = d(p', m) = d(p, p')/2.$$

- (b) For every three points $p_1, p_2, p_3 \in X$,

$$d(p_3, m)^2 \leq \frac{d(p_3, p_1)^2 + d(p_3, p_2)^2}{2} - \frac{d(p_1, p_2)^2}{4},$$

where m is a midpoint of $p_1 p_2$.)

- (3) Prove that an affine isometric action of Γ has a fixed point if and only if it has a bounded orbit if and only if all Γ -orbits are bounded.

(Hint. Let C be a bounded orbit of Γ . Argue why $r(C)$ is a fixed point of Γ .)

(Remark. The same argument shows that if a group G acts by isometries on a CAT(0) space X , then G has a fixed point if and only if it has a bounded orbit. This is used by Cartan to show that all the maximal compact subgroups of a semisimple Lie group are conjugate of each other. Cartan proved that $X := G/K$ is a Riemannian manifold and it is a CAT(0) space, where K is a maximal compact subgroup of G . Similarly, this result was employed by Bruhat and Tits to investigate compact subgroups of semisimple p -adic analytic groups. They defined simplicial spaces which are known as Bruhat-Tits buildings, and they proved that these spaces are CAT(0).)

- (4) Suppose Γ is a finitely generated group, has property(T) and it does not have property(FH). Then there exists a group homomorphism $\rho: \Gamma \rightarrow \mathcal{H} \rtimes O(\mathcal{H})$, $\rho(\gamma) = (a(\gamma), \pi(\gamma))$ such that $\{a(\gamma) \mid \gamma \in \Gamma\}$ is unbounded. Let $G := \mathcal{H}_{\pi} \rtimes \Gamma$. Let

$$\psi: G \rightarrow \mathbb{R}, \psi(v, \gamma) = \|v + a(\gamma)\|^2.$$

Note that

$$\begin{aligned} \Psi((v, \gamma), (v', \gamma')) &:= \psi((v', \gamma')^{-1}(v, \gamma)) = \psi(-\pi(\gamma'^{-1})(v') + \pi(\gamma'^{-1})(v), \gamma'^{-1}\gamma) \\ &= \|\pi(\gamma'^{-1})(v - v') + a(\gamma'^{-1}\gamma)\|^2 \\ &= \|\pi(\gamma'^{-1})(v - v') + \pi(\gamma'^{-1})(a(\gamma)) + a(\gamma'^{-1})\|^2 \\ &= \|\pi(\gamma'^{-1})(v - v' + a(\gamma) - a(\gamma'))\|^2 \\ &= \|(v + a(\gamma)) - (v' + a(\gamma'))\|^2. \end{aligned}$$

Hence ψ is a function of conditional negative type. Therefore, using Schoenberg's theorem and the GNS construction, we obtain a one-parameter family of cyclic unitary representations $(\mathcal{H}_t, \Pi_t, v_t)$ of G such that $\langle \Pi_t(v, \gamma)(v_t), v_t \rangle = \exp(-t\|v + a(\gamma)\|^2)$. Let $\pi_t: \Gamma \rightarrow \mathcal{U}(\mathcal{H}_t)$ be the restriction of Π_t to Γ , and $f_t: \mathcal{H} \rightarrow \mathcal{H}_t$, $f_t(v) := \Pi_t(v, 1)(v_t)$. Note that f_t is continuous.

- (a) Prove that $\langle f_t(v), f_t(v') \rangle = e^{-t\|v-v'\|^2}$ for all $v, v' \in \mathcal{H}$.

(Hint. $\langle f_t(v), f_t(v') \rangle = \langle \Pi_t(v - v', 1)(v_t), v_t \rangle$.)

- (b) Prove that $\pi_t(\gamma)(v_t) = f_t(a(\gamma))$ for all $\gamma \in \Gamma$.

(Hint. Show that $\|v_t\| = 1$ using $\langle \Pi_t(0, 1)(v_t), v_t \rangle = 1$. Next observe

$$\langle \Pi_t(-a(\gamma), \gamma)(v_t), v_t \rangle = 1,$$

and then deduce $\Pi_t(-a(\gamma), \gamma)(v_t) = v_t$.)

- (c) Prove that $f_t(\gamma \cdot v) = \pi_t(\gamma)(f_t(v))$, where $\gamma \cdot v := a(\gamma) + \pi(\gamma)(v)$ is the corresponding affine isometric action.

(Hint. Note $(0, \gamma^{-1})(\pi(\gamma)(v), 1)(0, \gamma) = (v, 1)$ and $\Pi_t(0, \gamma)(v_t) = \Pi_t(a(\gamma), 1)(v_t)$.)

- (d) Prove that the linear span of $f_t(\mathcal{H})$ is dense in $\mathcal{H}_t$.

(Hint. Use Part(c) to show that $\Pi_t(G)(v_t)$ is in the image of $f_t(\mathcal{H})$.)

- (e) Prove that if $\{w_n\}_n \subseteq \mathcal{H}$ and $w_n \rightarrow \infty$, then for every $u \in \mathcal{H}_t$, $\lim_{n \rightarrow \infty} \langle f_t(w_n), u \rangle = 0$.

(Hint. By Part (a), $\|f_t(w_n)\| = 1$ and

$$\lim_{n \rightarrow \infty} \langle f_t(w_n), f_t(w) \rangle = \lim_{n \rightarrow \infty} e^{-t\|w_n - w\|^2} = 0.$$

Use Part (d).)

- (f) Prove that $\mathcal{H}_t^\Gamma = 0$.

(Hint. Suppose $u \in \mathcal{H}_t^\Gamma$. For every $w \in \mathcal{H}$, the Γ -orbit $\{\gamma \cdot w \mid \gamma \in \Gamma\}$ is unbounded. So there exists $\{\gamma_n\}_n \subseteq \Gamma$ such that $\gamma_n \cdot w \rightarrow \infty$. So, $\lim_{n \rightarrow \infty} \langle f_t(\gamma_n \cdot w), u \rangle = 0$. Use Part (c) to deduce

$$0 = \lim_{n \rightarrow \infty} \langle \pi_t(\gamma_n)(f_t(w)), u \rangle = \lim_{n \rightarrow \infty} \langle f_t(w), \pi_t(\gamma_n^{-1})(u) \rangle = \lim_{n \rightarrow \infty} \langle f_t(w), u \rangle$$

Thus $u \perp f_t(\mathcal{H})$. Use Part (d) to obtain $u = 0$.)

- (g) Prove that, for every $s \in \Gamma$, $\lim_{t \rightarrow 0^+} \|\pi_t(s)(v_t) - v_t\| = 0$, and argue why this contradicts the assumption that Γ has property(T).

20.4. **Property(FA).**

20.5. **Characters of finite groups.**

20.6. **Clifford's theorem.** MATHEMATICS DEPARTMENT, UNIVERSITY OF CALIFORNIA, SAN DIEGO, CA 92093-0112, USA

Email address: golsefidy@ucsd.edu