Expander graphs and group theory

Ben-Gurion University

Izhar Oppenheim

June 26, 2019

Disclaimer: these lecture notes are written on a week-by-week basis in order to assist both myself (the lecturer) and the students. Due to my nature, these notes are almost bound to have typos, grammatical mistakes and mathematical mistakes (hopefully the last ones will be minor).

Preface

The aim of this course is to expose the students to the theory of expander graphs and its deep connection to expansion properties of groups. Since these topics are very broad, a lot of material was left out. The choices on which subjects to include or exclude depended on personal taste and on the ability to explain the subjects in a relatively elementary level. The course outline is the following:

1. Expander graphs
   (a) Preliminaries: Linear algebra, some terminology regarding graphs.
   (b) Basic spectral theory of graphs.
   (c) Different definitions of expansion: spectral and combinatorial. The connections between the different definitions (Buser-Cheeger inequality).
   (d) The expander mixing lemma.
   (e) Alon-Boppana Theorem.
   (f) Existence of expander graphs - non-constructive proof.
   (g) Metric distortion of expander graphs.
   (h) Application - error-correcting codes.
   (i) Gabber-Galil-Margulis expanders.
2. Expansion in groups
   (a) Needed basic concepts from group theory: finite and infinite countable groups, group actions, normal subgroups and quotients, Cayley and Schreier graphs.
   (b) Basic representation theory of finite groups.
   (c) Connection between representation of finite groups and expansion.
   (d) The "mother group approach" to expander construction.

(e) Hilbert spaces - quick overview.

(f) Kazhdan Property (T) - definition and relations to expansion.

(g) Elementary matrices groups and property (T): strategy for construction of expanders using matrix groups over $\mathbb{F}_p[t]$.

(h) Criterion for property (T) for a group generated by finite subgroups using angle considerations.

(i) Proof that elementary matrix groups over $\mathbb{F}_p[t]$ have property (T) given that p is sufficiently large.

Part I: Expander graphs

1 Preliminaries

1.0.1 Linear algebra - Quick reminder

Given a vector space U over $\mathbb{C}$, an *inner-product* is a function $\langle .,. \rangle : U \times U \to \mathbb{C}$ such that:

1. Positive definite: for every $x \in U$, $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = \vec{0}$.
2. (Bi)linear: for every $x, y, z \in U$ and every $\alpha \in \mathbb{C}$,
$$\langle \alpha x + y, z \rangle = \alpha \langle x, z \rangle + \langle y, z \rangle.$$
3. Conjugate symmetric: $x, y \in U$, $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

An inner-product induces a norm $\|x\|^2 = \langle x, x \rangle$ and the norm induces a topology, i.e., a notion of convergence: we define $x_n \to x_0$ by $\|x_n - x_0\| \to 0$.

Also, two vectors $x, y \in U$ are called orthogonal if $\langle x, y \rangle = 0$.

Proposition 1.1 (Cauchy-Schwarz inequality - without proof)**.** *Let U be an inner-product space, then for every $x, y \in U$, $|\langle x, y \rangle| \leq \|x\| \|y\|$.*

Exercise 1.2 (Informal)**.** *Use Cauchy-Schwarz inequality to prove that the inner-product is continuous, i.e., that if $x_n \to x_0, y_n \to y_0$, then $\langle x_n, y_n \rangle \to \langle x_0, y_0 \rangle$.*

Proposition 1.3 (Parallelogram equality)**.** *Let U be an inner-product space, then for every $x, y \in U$,*

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Proof. Expand the left side of the equation. □

A *linear operator* $T : U \to U$ is an operator satisfying $T(\alpha x + y) = \alpha T x + T y$. The *operator norm* of T is defined as

$$\|T\| = \sup_{x \in U, \|x\| = 1} \|Tx\| = \sup_{x \in U, x \neq \vec{0}} \frac{\|Tx\|}{\|x\|}.$$

We recall that T is continuous if and only if $\|T\| < \infty$ and in particular, if U is finite dimensional, then T is always continuous. A linear product over an inner-product space U is called *self-adjoint* if for every $x, y \in U$, $\langle Tx, y \rangle = \langle x, Ty \rangle$.

Proposition 1.4. *Let U be an inner-product space and $T : U \to U$ be a self-adjoint linear operator. Then the following holds:*

1. *For every x, $\langle Tx, x \rangle$ is real.*
2. *Every eigenvalue of T is real.*
3. *Two eigenvectors corresponding to different eigenvalues are orthogonal.*
4. *If $W \subseteq U$ is a linear subspace such that $TW \subseteq W$, then $TW^{\perp} \subseteq W^{\perp}$, where*
$$W^{\perp} = \{y \in U : \forall x \in W, \langle y, x \rangle = 0\}.$$

Proof. 1. $\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$.

2. If $Tx = \lambda x$, then $\langle Tx, x\rangle = \lambda\|x\|^2$, i.e., $\lambda = \frac{\langle Tx,x\rangle}{\|x\|^2}$ and by 1., both the numerator and denominator are real.

3. If $Tx = \lambda x, Ty = \mu y$ and $\lambda \neq \mu$, then

$$\lambda\langle x, y\rangle = \langle Tx, y\rangle = \langle x, Ty\rangle = \mu\langle x, y\rangle,$$

and thus $\langle x, y\rangle = 0$.

4. Let $y \in W^\perp$, then for every $x \in W$, $Tx \in W$ and therefore

$$\langle x, Ty\rangle = \langle Tx, y\rangle = 0,$$

i.e., $Ty \in W^\perp$.

□

Proposition 1.5. *Let U be an inner-product space and $T : U \to U$ be a self-adjoint linear operator. Then*

$$\|T\| = \sup_{x\in U, \|x\|=1} |\langle Tx, x\rangle| = \sup_{x\in U, x\neq\vec{0}} \frac{|\langle Tx, x\rangle|}{\|x\|^2}.$$

($\frac{|\langle Tx,x\rangle|}{\|x\|^2}$ is called Rayleigh quotient of T at x).

Proof. If $Tx = 0$ for every x there is nothing to prove. Assume that $T \neq 0$. Denote $M = \sup_{x\in U,\|x\|=1}|\langle Tx, x\rangle|$. First, by Cauchy-Schwarz, for every $x \in U, \|x\| = 1$,

$$|\langle Tx, x\rangle| \leq \|Tx\|\|x\| \leq \|T\|,$$

thus $M \leq \|T\|$.

In the other direction, we note that for every $x, y \in U$,

$$\begin{aligned}
&\langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle \\
&= \langle Tx, x\rangle + \langle Ty, x\rangle + \langle Tx, y\rangle + \langle Ty, y\rangle - (\langle Tx, x\rangle - \langle Ty, x\rangle - \langle Tx, y\rangle + \langle Ty, y\rangle) \\
&= 2(\langle Ty, x\rangle + \langle Tx, y\rangle) \\
&= 2(\langle y, Tx\rangle + \overline{\langle Tx, y\rangle}) \\
&= 2(\langle Tx, y\rangle + \overline{\langle Tx, y\rangle}) \\
&= 4Re(\langle Tx, y\rangle).
\end{aligned}$$

Therefore,

$$4|Re(\langle Tx, y\rangle)| \leq |\langle T(x+y), x+y\rangle| + |\langle T(x-y), x-y\rangle| \leq M(\|x+y\|^2 + \|x-y\|^2) = 2M(\|x\|^2 + \|y\|^2).$$

Therefore, for every x with $\|x\| = 1$ and $Tx \neq 0$, we can take $y = \frac{Tx}{\|Tx\|}$ and get

$$4\|Tx\| \leq 2M + 2M = 4M,$$

as needed. □

Remark 1.6. *We know that is T is self-adjoint, then for every x, $\langle Tx, x\rangle$ is real and therefore by the previous proposition*

$$\|T\| = \max\{\sup_{x,\|x\|=1} \langle Tx, x\rangle, -\inf_{x,\|x\|=1} \langle Tx, x\rangle\}.$$

Proposition 1.7. *Let U be an inner-product space and $T : U \to U$ be a self-adjoint linear operator. If U is finite dimensional, then*

$$\lambda = \sup_{x,\|x\|=1} \langle Tx, x\rangle,$$

$$\mu = \inf_{x,\|x\|=1} \langle Tx, x\rangle,$$

are eigenvalues of T.

Proof. It is enough to prove that λ is an eigenvalue (because then, by the same argument, $-\mu$ is an eigenvalue of $-T$).

Let $\{x_n\}_{n=1}^{\infty}$ by a sequence such that $\|x_n\|= 1$ for every n and $\langle Tx_n, x_n\rangle \to \lambda$. Since U is finite dimensional, its unit ball is compact and therefore x_n has a converging subsequence. Therefore, up to passing to a subsequence, we can assume that $x_n \to x_0$ (and $\|x_0\|= 1$ by continuity of the norm).

Define $T' = T - \mu I$. Then T' is self-adjoint and

$$\|T'\|= \max\{\sup_{x,\|x\|=1} \langle T'x, x\rangle, -\inf_{x,\|x\|=1} \langle T'x, x\rangle\} = \max\{\lambda - \mu, 0\} = \lambda - \mu.$$

Using this,

$$\begin{aligned}&\|Tx_n - \lambda x_n\|^2 = \|Tx_n - \mu x_n - (\lambda - \mu)x_n\|^2 = \\ &\|(T - \mu I)x_n\|^2 + \|(\mu - \lambda)x_n\|^2 - 2(\lambda - \mu)\langle Tx_n - \mu x_n, x_n\rangle \leq \\ &2(\lambda - \mu)^2 - 2(\lambda - \mu)(\langle Tx_n, x_n\rangle - \mu) \to 0.\end{aligned}$$

By continuity, $\|Tx_0 - \lambda x_0\|^2= \lim_n \|Tx_n - \lambda x_n\|^2= 0$ and therefore $Tx_0 = \lambda x_0$ as needed. □

Corollary 1.8. *Let U be an inner-product space and $T : U \to U$ be a self-adjoint linear operator. Then there is an orthonormal basis of U that is composed of eigenvectors of T.*

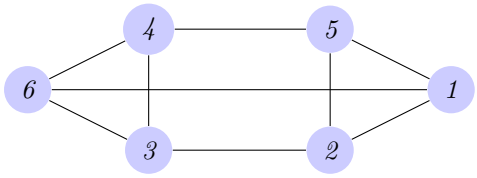
Proof. The proof is by induction on the dimension of U. If U is of dimension 1, then there is nothing to prove. Let U of dimension n.

By the above proposition, there is an eigenvector x with an eigenvalue $\lambda = \sup_{x,\|x\|=1}\langle Tx, x\rangle$. Let $U' = (span\{x\})^{\perp}$. By a previous proposition, $T : U' \to U'$ is well-defined (since $TU' \subseteq U'$) and the dimension of U' is $n-1$ and therefore U' and an orthonormal basis $\{y_1, ..., y_{n-1}\}$ composed of eigenvectors of T. Since $U' \perp x$, we deduce that $\{y_1, ..., y_{n-1}, x\}$ is an orthonormal basis of U and we are done. □

1.1 Graphs - some terminology

A simple graph is a pair (V, E), where V is a set called the *vertices* of the graph and E is a set of pairs of the form $\{u, v\}$, where $u, v \in V$, called the *edges* of the graph and u, v are called neighbors. Throughout, we will assume that all our graphs have no isolated vertices, i.e., that for every $v \in V$ there is at least one $u \in V$ such that $\{v, u\} \in E$.

Example 1.9. $V = \{1, 2, 3, 4, 5, 6\}$,
$E = \{\{1, 2\}, \{1, 6\}, \{4, 6\}, \{3, 6\}, \{4, 5\}, \{1, 5\}, \{2, 5\}, \{2, 3\}, \{3, 4\}\}$.



Some basic terminology regarding graphs:

1. A graph is said to have no isolated vertices if for every $v \in V$ there is at least one $u \in V$ such that $\{v, u\} \in E$. **Throughout, we will always assume that our graphs have no isolated vertices**.

2. A *loop* in a graph (V, E) is an edge of the form $\{u, u\}$. **Unless stated otherwise, we will always assume that our graphs have no loops.**

3. A graph (V, E) is called *finite*, if $|V| < \infty$ (e.g., the graph in Example 1.9 is finite).

4. Given a graph (V, E), a *path* in the graph is a sequence of vertices $v_0, ..., v_n \in V$ such that $\{v_i, v_{i+1}\} \in E$ for every $0 \leq i \leq n-1$. The number n is called the length of the path.

5. Two vertices $u, v \in V$ are said to be connected by a path if there is a path $v_0, ..., v_n \in V$ such that $v_0 = u, v_n = v$. The (graph) distance between two vertices, is the length of the shortest path connecting them (if there is no such path, we say that the distance is ∞).

6. A graph (V, E) is called *connected*, if every two vertices are connected by a path (e.g., the graph in Example 1.9 is connected).

7. For a graph (V, E) a *connected component* of the graph is a (sub)graph (V', E') such that $\emptyset \neq V' \subseteq V, E' \subseteq E$ and:

   - If $u, v \in V'$ and $\{u, v\} \in E$, then $\{u, v\} \in E'$.
   - Every two vertices $u, v \in V'$ are connected by a path in (V, E) (and therefore in (V', E')).
   - For every $u \in V', v \in V \setminus V'$, there is no path connecting u and v.

   We note that a graph is connected if and only if it has only one connected component.

8. For $v \in V$, the *degree (or valency)* of v is the number $m(v) = |\{u \in V : \{u, v\} \in E\}|$.

9. For $d \in \mathbb{N}$, a graph (V, E) is called *d-regular*, if for every $v \in V$, $m(v) = d$ (e.g., the graph in Example 1.9 is 3-regular). A graph is called *regular*, if it is d-regular for some d.

Below, for the rest of part I, we will assume that all the graphs are finite.

2 Introduction to spectral graph theory

The basic idea behind spectral graph theory is that given a (finite) graph, one can associate a operator/matrix with it and the eigenvalues of this operator reflect some properties of the graph (terminology - the spectrum of a symmetric matrix is the set of its eigenvalues). There are several standard operators/matrices associated with a given graph (the following list is not comprehensive):

1. The adjacency matrix.
2. The graph non-normalized Laplacian.
3. The Markov kernel (normalized adjacency matrix).
4. The normalized graph Laplacian.

These matrices are all equivalent in some sense, when working with regular graphs, but the normalized versions behave different the the non-normalized versions in non-regular graphs. For our purposes, we will always work with the normalized operators (the non-normalized adjacency matrix will be discussed in the homework exercises).

Given a finite graph $G = (V, E)$, define $\ell^2(V)$ to be the space of functions $\phi : V \to \mathbb{C}$, with the inner-product

$$\langle \phi, \psi \rangle = \sum_{v \in V} m(v)\phi(v)\overline{\psi(v)},$$

(this is just a weighted version of the standard inner-product of $\mathbb{C}^{|V|}$). We denote the norm induced by the inner-product by $\|.\|$, i.e.,

$$\|\phi\|^2 = \sum_{v \in V} m(v)|\phi(v)|^2.$$

Definition 2.1 (The Markov kernel)**.** *Given a graph $G = (V, E)$, the Markov kernel (or simple random walk operator) of the graph is the $|V|\times|V|$ matrix indexed by V, defined as*

$$M(v,u) = \begin{cases} \frac{1}{m(v)} & \{v,u\} \in E \\ 0 & \{v,u\} \notin E \end{cases}.$$

We note that as an operator $M : \ell^2(V) \to \ell^2(V)$ acts as follows - given $\phi \in \ell^2(V)$,

$$(M\phi)(v) = \sum_{u \in V} M(v,u)\phi(u) = \frac{1}{m(v)} \sum_{u \in V, \{v,u\} \in E} \phi(u).$$

In other words, $M\phi(v)$ is the average of the values of ϕ taken over the neighbors of v.

Proposition 2.2. *With respect to the inner-product of $\ell^2(V)$, M is a self-adjoint operator, i.e., for every $\phi, \psi \in \ell^2(V)$, $\langle M\phi, \psi \rangle = \langle \phi, M\psi \rangle$.*

Proof.

$$\begin{aligned}
\langle M\phi, \psi\rangle &= \sum_{v\in V} m(v)\left(\frac{1}{m(v)}\sum_{u\in V,\{v,u\}\in E}\phi(u)\right)\overline{\psi(v)}\\
&= \sum_{v\in V}\sum_{u\in V,\{v,u\}\in E}\phi(u)\overline{\psi(v)}\\
&= \sum_{u\in V}\phi(u)\left(\sum_{v\in V,\{v,u\}\in E}\overline{\psi(v)}\right)\\
&= \sum_{u\in V} m(u)\phi(u)\overline{\left(\frac{1}{m(u)}\sum_{v\in V,\{v,u\}\in E}\psi(v)\right)}\\
&= \langle \phi, M\psi\rangle.
\end{aligned}$$

□

Corollary 2.3. *All the eigenvalues of M are real, eigenfunctions of different eigenvalues are orthogonal and $\ell^2(V)$ as an orthogonal basis of eigenfunctions of M.*

Remark 2.4. *We can always choose the basis of the eigenfunction above to be composed of only real valued functions: we note that for every eigenfunction ϕ, $\phi = \phi_1 + i\phi_2$, where ϕ_1, ϕ_2 are real valued. By the definition of M, $M\phi_1, M\phi_2$ are also real valued and therefore*

$$\lambda\phi_1 + i\lambda\phi_2 = M\phi = M\phi_1 + iM\phi_2,$$

implies that $\lambda\phi_1 = M\phi_1, \lambda\phi_2 = M\phi_2$. Thus, the induction proving that there is an orthonormal basis of eigenfunctions can be done by choosing a real-valued eigenfunction at each step.

We observe that 1 is always an eigenvalue of M that corresponds to the constant function $\mathbb{1}$:

$$(M\mathbb{1})(v) = \frac{1}{m(v)}\sum_{u\in V,\{v,u\}\in E}\mathbb{1}(u) = 1 = \mathbb{1}(v).$$

Denote

$$\ell_0^2(V) = \{\phi \in \ell^2(V) : \phi \perp \mathbb{1}\}.$$

By a direct computation,

$$\langle \phi, \mathbb{1}\rangle = \sum_{v\in V} m(v)\phi(v),$$

and therefore

$$\ell_0^2(V) = \{\phi \in \ell^2(V) : \sum_{v\in V} m(v)\phi(v) = 0\}.$$

The *non-trivial spectrum of M* is the spectrum of $M|_{\ell_0^2(V)}$: $\ell_0^2(V) \to \ell_0^2(V)$.

Proposition 2.5. *Let $G = (V, E)$ be a finite graph. The graph is connected if and only if 1 is always an eigenvalue of M with multiplicity 1. Moreover, 1 is an eigenvalue of M with multiplicity k if and only if G has k connected components.*

Proof. Assume that the graph is not connected. Let (V', E') be a connected component of $G = (V, E)$ and define $\chi_{V'} \in \ell^2(V)$ to be the indicator function of V'. Then for every $v \in V'$, all the neighbors of v are also in V' and therefore $(M\chi_{V'})(v) = 1 = \chi_{V'}(v)$. Also, for every $v \notin V'$ all the neighbors of v are also not in V' and therefore $(M\chi_{V'})(v) = 0 = \chi_{V'}(v)$. Thus, $M\chi_{V'} = \chi_{V'}$. Since this is true for every connected component, we get that if there are two different connected components $(V', E'), (V'', E'')$, then $\chi_{V'}, \chi_{V''}$ are eigenfunctions with eigenvalue 1 and since $V' \cap V'' = \emptyset$, $\langle \chi_{V'}, \chi_{V''} \rangle = 0$. It follows that the number of connected components is greater or equal to the multiplicity of 1.

We will show that if ϕ is an eigenfunction of M with the eigenvalue 1, then it has to be constant on every connected component and thus ϕ is a linear combination of $\{\chi_{V'} : (V', E') \text{ is a connected component}\}$. This is done using the maximum principle. Without loss of generality, it is enough to prove this for real valued eigenfunctions. Let ϕ be a real valued function such that $M\phi = \phi$ and let (V', E') be a connected component. Let $C = \max_{v \in V'} \phi(v)$ (recall that V' is a finite set). We will call a vertex $v_0 \in V'$ maximal in (V', E'), if $\phi(v_0) = C$. Let v_0 be a maximal

$$C = \phi(v_0) = (M\phi)(v_0) = \frac{1}{m(v_0)} \sum_{u, \{v_0, u\} \in E'} \phi(u) \leq \frac{1}{m(v_0)} \sum_{u, \{v_0, u\} \in E'} C = C.$$

This yields that the inequality above is in fact an equality, i.e.,

$$\frac{1}{m(v_0)} \sum_{u, \{v_0, u\} \in E'} \phi(u) = C$$

and it follows that $\phi(u) = C$ for every u such that $\{v_0, u\} \in E'$. In other words, if $\phi(v_0) = C$, then for every neighbor u of v_0, $\phi(u) = C$. By induction, this argument yields that for every u, if u is connected v_0 by a path, then $\phi(u) = C$, and therefore $\forall v \in V', \phi(v) = C$. □

Proposition 2.6. *For every finite graph $G = (V, E)$ all the eigenvalues of M are in the interval $[-1, 1]$, i.e.,* $\mathrm{Spec}(M) \subseteq [-1, 1]$.

Proof. The proof is again by the maximum principle. Let ϕ be an eigenfunction of M with eigenvalue λ. There is some v_0, such that for every $v \in V$, $|\phi(v)| \leq |\phi(v_0)|$. Then

$$|\lambda||\phi(v_0)| = |(M\phi)(v_0)| \leq \left| \frac{1}{m(v_0)} \sum_{u, \{v_0, u\} \in E} \phi(u) \right| \leq \frac{1}{m(v_0)} \sum_{u, \{v_0, u\} \in E} |\phi(u)| \leq |\phi(v_0)|,$$

and after dividing by $|\phi(v_0)|$, we get that $|\lambda| \leq 1$. □

Definition 2.7. *A graph $G = (V, E)$ is called bipartite if there are $S_0, S_1 \subset V$, called the sides of the graph, such that for every $\{u, v\} \in E$, $|\{u, v\} \cap S_0| = |\{u, v\} \cap S_1| = 1$. In other words, a graph is bipartite if every edge connects a vertex in one side to a vertex in the other side.*

Proposition 2.8. *Let $G = (V, E)$ be a finite connected graph. This graph is bipartite if and only if -1 is an eigenvalue of M.*

Proof. Assume that the graph is bipartite with sides S_0, S_1. Define

$$\phi(u) = \begin{cases} 1 & u \in S_0 \\ -1 & u \in S_1 \end{cases}.$$

Let $v \in V$. If $v \in S_0$ all the neighbors of v are is S_1 and therefore $(M\phi)(v) = -1 = -\phi(v)$. If $v \in S_1$ all the neighbors of v are is S_0 and therefore $(M\phi)(v) = 1 = -\phi(v)$.

Conversely, assume that -1 is an eigenvalue of M and let ϕ be a real eigenfunction with $M\phi = -\phi$. Again, we will apply the maximum principle. Without loss of generality, we can assume that $\max_{v\in V}|\phi(v)| = 1$ and that the is $v_0 \in V$ such that $\phi(v_0) = 1$. Then applying the same considerations as above, we deduce that all the neighbors u of v_0, $\phi(u) = -1$. Again, by the same argument, if $\phi(u) = -1$, then for every neighbor v of u, $\phi(v) = 1$. Since the graph is connected, we deduce that ϕ has only the values ± 1 and assigns different values for every two neighboring vertices. Thus, $S_0 = \{v : \phi(v) = 1\}, S_1 = \{v : \phi(v) = -1\}$ are the two sides of G. □

Exercise 2.9. *Let $G = (V, E)$ a bipartite graph with sides S_0, S_1. Show that the spectrum of M for this graph is symmetric in the following sense - if λ is an eigenvalue of M, then $-\lambda$ is also an eigenvalue of M with the same multiplicity (hint: if ϕ is an eigenfunction of M with eigenvalue λ, show that $\psi(v) = \begin{cases} \phi(v) & v \in S_0 \\ -\phi(v) & v \in S_1 \end{cases}$ is an eigenfunction of M with eigenvalue $-\lambda$).*

Definition 2.10 (The normalized Laplacian)**.** *Let $G = (V, E)$ be a finite graph. Define $L : \ell^2(V) \to \ell^2(V)$ as $L = I - M$. In other words, for every $\phi \in \ell^2(V)$,*

$$(L\phi)(v) = \phi(v) - \frac{1}{m(v)} \sum_{u,\{u,v\}\in E} \phi(u).$$

From the above discussion regarding M, the following are immediate (left as an informal exercise):

1. $\mathrm{Spec}(L) \subseteq [0, 2]$.
2. The space $\ell^2(V)$ has an orthogonal basis of real-valued eigenfunction of L.
3. The constant function $\mathbb{1}$ is an eigenfunction of L with an eigenvalue 0.
4. The graph $G = (V, E)$ is connected if and only if 0 is an eigenvalue of multiplicity 1 of L.
5. For a connected graph (V, E), the graph is bipartite if and only if $2 \in \mathrm{Spec}(L)$.
6. For a bipartite graph, if λ is an eigenvalue of L, then $2-\lambda$ is an eigenvalue of L with the same multiplicity.

Proposition 2.11. *Let $G = (V, E)$ be a graph. For every $\phi \in \ell^2(V)$,*

$$\langle L\phi, \phi \rangle = \sum_{\{u,v\} \in E} |\phi(u) - \phi(v)|^2.$$

Proof. Fix $\phi \in \ell^2(V)$, then

$$\begin{aligned}
\langle L\phi, \phi \rangle &= \sum_{v \in V} m(v)\phi(v) \overline{\left(\phi(v) - \frac{1}{m(v)} \sum_{u, \{u,v\} \in E} \phi(u) \right)} \\
&= \sum_{v \in V} m(v)|\phi(v)|^2 - \sum_{v \in V} \sum_{u, \{u,v\} \in E} \phi(v)\overline{\phi(u)} \\
&= \sum_{v \in V} m(v)|\phi(v)|^2 - \sum_{\{u,v\} \in E} 2\phi(v)\overline{\phi(u)} \\
&=
\end{aligned}$$

We note that

$$\sum_{v \in V} m(v)|\phi(v)|^2 = \sum_{v \in V} \sum_{u, \{u,v\} \in E} |\phi(v)|^2 = \sum_{\{u,v\} \in E} |\phi(v)|^2 + |\phi(u)|^2.$$

Combining this with the above computation yields that

$$\begin{aligned}
\langle L\phi, \phi \rangle &= \sum_{\{u,v\} \in E} |\phi(v)|^2 + |\phi(u)|^2 - 2\phi(v)\overline{\phi(u)} \\
&= \sum_{\{u,v\} \in E} |\phi(u) - \phi(v)|^2.
\end{aligned}$$

□

2.1 Exercises

Definition 2.12 (Adjacency matrix)**.** *Given a graph $G = (V, E)$, its adjacency matrix A is a $|V| \times |V|$ matrix indexed by V, defined as*

$$A(v, u) = \begin{cases} 1 & \{v, u\} \in E \\ 0 & \{v, u\} \notin E \end{cases}.$$

1. Prove that for a d-regular graph (V, E), λ is an eigenvalue of M if and only if $d\lambda$ is an eigenvalue of A.
2. Prove that for every n, $A^n(u, v)$ is the number of paths of length n between u and v.
3. Calculate the spectrum of M for the following graphs:
   (a) The complete graph on n vertices: for $n \geq 2$, $V = \{1, ..., n\}$ and $E = \{\{i, j\} : 1 \leq i < j \leq n\}$.
   (b) The complete bipartite graph on $n_0 + n_1$ vertices: for $n_0, n_1 \in \mathbb{N}$, $S_0 = \{v_1, ..., v_{n_0}\}, S_1 = \{u_1, ..., u_{n_1}\}$, $V = S_0 \cup S_1$, $E = \{\{v_i, u_j\} : 1 \leq i \leq n_0, 1 \leq j \leq n_1\}$.

4. Given a graph (V, E), we say that the random walk on the graph mixes, if there is a constant $0 \leq \alpha < 1$, such that for every $\phi \in \ell^2(V)$ and every $k \in \mathbb{N}$, $\|M^k\phi - \phi^0\| \leq \alpha^k\|\phi - \phi^0\|$, where ϕ^0 is the orthogonal projection of ϕ on the subspace of constant functions (i.e., on the space $\{\psi : V \to \mathbb{C} : \psi \equiv C\} \subseteq \ell^2(V)\}$).

   (a) Show that if (V, E) is bipartite or if (V, E) is not connected, then the random walk does not mix.

   (b) Recall that if (V, E) is connected not bipatite, then there is $0 \leq \lambda < 1$, such that $\mathrm{Spec}(M) \subseteq [-\lambda, \lambda] \cup \{1\}$. Show that in that case the random walk mixes and find the connection between α and λ.

   (c) Define the lazy random walk as $M' = \frac{1}{2}M + \frac{1}{2}I$. Show that if (V, E) is connected, then M' mixes: there is a constant $0 \leq \alpha < 1$, such that for every $\phi \in \ell^2(V)$ and every $k \in \mathbb{N}$, $\|(M')^k\phi - \phi^0\| \leq \alpha^k\|\phi - \phi^0\|$, where ϕ^0 is as above.

3 Expander graphs

There are several definitions of expander graphs and more importantly expander family of graphs. The two most prominent definitions are:

- Spectral expansion
- Edge expansion

The basic idea behind is to "measure" how connected the graph is. These two definitions are equivalent "qualitatively" but not "quantitatively" (this statement will be explained below).

3.1 Spectral expansion

Notation: for a connected graph $G = (V, E)$, denote by λ_G the second largest eigenvalue of M.

Definition 3.1 (Spectral expansion)**.** *For a constant $\lambda < 1$, a graph $G = (V, E)$ is called a (one-sided) λ-spectral expander, if it is connected and if $\lambda_G \leq \lambda$.*

A graph (V, E) is called a two-sided λ-spectral expander, if it is connected and if $\mathrm{Spec}(M) \subseteq [-\lambda, \lambda] \cup \{1\}$.

Remark 3.2. *When referring to spectral expansion, we will mean the one-sided expansion, but the reader should note that some people/books refer to spectral expansion as the two-sided spectral expansion.*

Obtaining spectral expansion is very easy: for instance, if G is the complete graph on n vertices, then it is a two-sided $\frac{1}{n-1}$-spectral expander and if G is a complete bipartite graph, then it is always a one-sided 0-spectral expander. However, these examples are not good, since the maximal degree grows with n. The idea is to find examples where the maximal degree is uniformly bounded. More precisely:

Definition 3.3. *A family of finite graphs $\{G_n = (V_n, E_n)\}_{n=1}^{\infty}$ is called an (one sided spectral) expander family if:*

1. $\lim_n |V_n| = \infty$.
2. $\sup_n (\max_{v \in V_n} m(v)) < \infty$.
3. *There is a constant* $\lambda < 1$ *such that every* G_n *is a one-sided* λ*-spectral expander.*

Remark 3.4. *In some books, condition* 2. *above is replaced with the stronger condition that there is a constant* $d \in \mathbb{N}$ *such that for every* n, G_n *is* d*-regular.*

3.2 Edge expansion

Before defining edge expansion, we will need the following notations: for a graph $G = (V, E)$ denote:

1. For a subset $\emptyset \neq U \subseteq V$, denote $m(U) = \sum_{u \in U} m(u)$. Note that if (V, E) is d-regular, then $m(U) = d|U|$.
2. For subsets $\emptyset \neq U, U' \subseteq V$, denote $E(U, U') = \{\{u, u'\} : u \in U, u' \in U'\}$.
3. For a subset $\emptyset \neq U \subsetneqq V$, denote $\partial U = E(U, V \setminus U)$.

Definition 3.5 (Cheeger constant)**.** *Given a finite graph* $G = (V, E)$*, the Cheeger constant* h_G *is the constant defined as follows. First, for* $\emptyset \neq U \subsetneqq V$*, define*

$$h_G(U) = \frac{|\partial U|}{\min\{m(U), m(V \setminus U)\}}.$$

Second, define

$$h_G = \min_{\emptyset \neq U \subsetneqq V} h_G(U) = \min_{\emptyset \neq U \subsetneqq V, m(U) \leq \frac{1}{2} m(V)} \frac{|\partial U|}{m(U)}.$$

Terminology: given a finite graph $G = (V, E)$, a set $\emptyset \neq U \subsetneqq V, m(U) \leq \frac{1}{2} m(V)$ is called a *cut* of G. The Cheeger constant compares the number of edges going out of U and the "size" of U.

Observations:

- For every cut U, $|\partial U| \leq m(U)$ and thus $h_G \leq 1$.
- G is connected if and only if $h_G > 0$.
- Large h_G means no "bottleneck" (network example).

Again, it is not hard to get a large h_G for a single graph (again, consider the complete graph or the complete bipartite graph), but we want $|E|$ is linear with respect to $|V|$.

Definition 3.6. *A family of finite graphs* $\{G_n = (V_n, E_n)\}_{n=1}^{\infty}$ *is called an (edge-) expander family if:*

1. $\lim_n |V_n| = \infty$.
2. $\sup_n (\max_{v \in V_n} m(v)) < \infty$.
3. *There is a constant* $\alpha > 0$ *such that every* G_n, $h_{G_n} \geq \alpha$.

Remark 3.7. *In some books, condition 2. above is replaced with the stronger condition that there is a constant $d \in \mathbb{N}$ such that for every n, G_n is d-regular.*

Some non-examples:

1. $G_n = (V_n, E_n)$ where $V_n = \{0, ..., 2n+1\}$, $E_n = \{\{i, i+1\} : 0 \leq i \leq 2n-1\}$. The cut $U = \{0, ..., n\}$ has $m(V \setminus U) = m(U) = 1 + 2(n-1)$ and $|\partial U| = 1$, thus $h_{G_n} = \frac{1}{1+2(n-1)}$ which tends to 0 as $n \to \infty$.

2. Fix $k \in \mathbb{N}$ and define $G_n = (V_n, E_n)$ as

$$V_n = \{(a_1, ..., a_k) : a_1, ..., a_k \in \mathbb{Z}, 0 \leq a_1, ..., a_k \leq 2n+1\},$$

$$E_n = \{\{(a_1, ..., a_k), (b_1, ..., b_k)\} : \exists i_0, a_{i_0} + 1 = b_{i_0}, \forall i \neq i_0, a_i = b_i\}.$$

   This is a non-example, because for $U = \{(a_1, ..., a_k) : a_1, ..., a_k \in \mathbb{Z}, 0 \leq a_1, ..., a_k \leq n\}$, $m(U) \geq n^k$ and $|\partial U| \leq 2n^{k-1}$ (this is basically due to the isoperimetric inequality in $\mathbb{R}^k$).

3.3 Buser-Cheeger inequality

The Buser-Cheeger inequality (sometimes known only as the Cheeger inequality) connects the two notion of expansion defined above. Namely:

Theorem 3.8 (Buser-Cheeger inequality)**.** *For every connected finite graph $G = (V, E)$, $2h_G \geq 1 - \lambda_G \geq \frac{h_G^2}{2}$.*

Remark 3.9. *We note that this inequality do not give a "quantitative" equivalence between the two notions of expansion:*

- *Even if $\lambda_G = 0$, we only derive from the inequality that $h_G \geq \frac{1}{2}$ (and not that $h_G = 1$).*
- *Even if $h_G = 1$, we only derive from the inequality that $\lambda_G \leq \frac{1}{2}$ (and not that $\lambda_G = 0$).*

What can be derived is a "qualitative" equivalence: by the Buser-Cheeger inequality, a family of graphs G_n is an expander family according to the spectral definition if and only if it is an expander family according to the edge-expansion definition.

We will divide the proof into two parts: proving $2h_G \geq 1 - \lambda_G$, which is simple and proving $1 - \lambda_G \geq \frac{h_G^2}{2}$, which is much more involved.

Proof of $2h_G \geq 1 - \lambda_G$:

Proof. Let $G = (V, E)$ be a finite connected graph. By the definition of the Laplacian, $\mathrm{Spec}(L) \subseteq [1 - \lambda_G, 2] \cup \{0\}$ and therefore, for every $\phi \in \ell_0^2(V)$, $\langle L\phi, \phi \rangle \geq (1 - \lambda_G)\|\phi\|^2$. Let $\emptyset \neq U \subsetneq V$ be a cut (i.e., $m(U) \leq \frac{1}{2}m(V)$). Define a function

$$\phi(v) = \begin{cases} -\frac{1}{m(U)} & v \in U \\ \frac{1}{m(V \setminus U)} & v \notin U \end{cases}.$$

Then

$$\begin{aligned}\sum_{v\in V} m(v)\phi(v) &= \sum_{v\in U} m(v)\phi(v) + \sum_{v\in V\setminus U} m(v)\phi(v)\\ &= -\sum_{v\in U} m(v)\frac{1}{m(U)} + \sum_{v\in V\setminus U} m(v)\frac{1}{m(V\setminus U)}\\ &= -1+1\\ &= 0.\end{aligned}$$

In other words, $\phi \in \ell_0^2(V)$ and therefore $\langle L\phi,\phi\rangle \geq (1-\lambda_G)\|\phi\|^2$. Be a previous proposition,

$$\begin{aligned}\langle L\phi,\phi\rangle &= \sum_{\{u,v\}\in E} |\phi(u)-\phi(v)|^2\\ &= \sum_{\{u,v\}\in E, u\in U, v\in V\setminus U} \left(\frac{1}{m(U)}+\frac{1}{m(V\setminus U)}\right)^2\\ &= |E(U,V\setminus U)|\left(\frac{1}{m(U)}+\frac{1}{m(V\setminus U)}\right)^2.\end{aligned}$$

Also,

$$\begin{aligned}\|\phi\|^2 = \sum_{v\in V} m(v)\phi(v)^2 &= \sum_{v\in U} m(v)\phi(v)^2\\ &+ \sum_{v\in V\setminus U} m(v)\phi(v)^2 = \sum_{v\in U} m(v)\frac{1}{m(U)^2}\\ &+ \sum_{v\in V\setminus U} m(v)\frac{1}{m(V\setminus U)^2} = \frac{1}{m(U)}+\frac{1}{m(V\setminus U)}.\end{aligned}$$

From the inequality, $\langle L\phi,\phi\rangle \geq (1-\lambda_G)\|\phi\|^2$, it follows that

$$|\partial U| \geq (1-\lambda_G)\frac{1}{\frac{1}{m(U)}+\frac{1}{m(V\setminus U)}} = (1-\lambda_G)\frac{m(U)m(V\setminus U)}{m(V)}.$$

Therefore

$$h_G(U) = \frac{|\partial U|}{m(U)} \geq (1-\lambda_G)\frac{m(V\setminus U)}{m(V)} \geq \frac{1-\lambda_G}{2}.$$

Since this is true for every cut U, the needed inequality follows. □

Proof of $1-\lambda_G \geq \frac{h_G^2}{2}$:

Proof. Let $\phi \in \ell_0^2(V)$ be a real-valued eigenfunction of M such that $M\phi = \lambda_G\phi$. Then $L\phi = (1-\lambda_G)\phi$. It follows that

$$(1-\lambda_G) = \frac{\langle L\phi,\phi\rangle}{\|\phi\|^2}.$$

Label the vertices of V according to the value of ϕ, i.e., denote $V=\{v_1,...,v_n\}$ such that $\phi(v_1)\geq\phi(v_n)\geq...\geq\phi(v_n)$. We consider $U_i=\{v_1,...,v_i\}$ and denote $\alpha_G = \min_i h_G(U_i)$. Obviously, $\alpha_G \geq h_G$ and we will show that $1-\lambda_G \geq \frac{\alpha_G^2}{2}$.

Define $1\leq k\leq n$ as follows to be the maximal index such that $m(U_k)\leq \frac{1}{2}m(V)$. In other words, $1\leq k\leq n$ is the unique index such that $m(U_k)\leq$

$\frac{1}{2}m(V)$ and $m(U_{k+1}) > \frac{1}{2}m(V)$. Denote $c = \phi(v_k)$ (note that c is the roughly the median value of ϕ when accounting to the degrees) and define $\psi(v_i) = \phi(v_i) - c$. Then $\|\psi\|^2 = \|\phi\|^2 + c^2 \geq \|\phi\|^2$ and

$$\langle L\phi, \phi \rangle = \sum_{\{u,v\} \in E} |\phi(u) - \phi(v)|^2 = \sum_{\{u,v\} \in E} |(\phi(u) - c)(\phi(v) - c)|^2 = \langle L\psi, \psi \rangle.$$

It follows that

$$1 - \lambda \geq \frac{\langle L\psi, \psi \rangle}{\|\psi\|^2}.$$

Next, the positive and negative parts of ψ:

$$\psi^+(v_i) = \begin{cases} \psi(v_i) & i \leq k \\ 0 & i > k \end{cases},$$

$$\psi^-(v_i) = \begin{cases} -\psi(v_i) & i < k \\ 0 & i \leq k \end{cases}.$$

Note that $\psi = \psi^+ - \psi^-$ and that ψ^+, ψ^- are both non-negative functions. It follows that

$$1 - \lambda_G \geq \frac{\sum_{\{u,v\} \in E} |\psi(u) - \psi(v)|^2}{\sum_{v \in V} m(v)(\psi(v))^2}.$$

Regrading the numerator, we note that for every $\{u, v\} \in E$ there are 2 options:

1. The values $\psi(u)$ and $\psi(v)$ have the same sign. In that case, either $\psi^+(u) = \psi^+(v) = 0$ or $\psi^-(u) = \psi^-(v) = 0$. It follows that

$$|\psi(u) - \psi(v)|^2 = |\psi^+(u) - \psi^+(v)|^2 + |\psi^-(u) - \psi^-(v)|^2.$$

2. The values $\psi(u)$ and $\psi(v)$ have the opposite signs. In that case

$$|\psi(u) - \psi(v)| = |\psi(u)| + |\psi(v)|,$$

and therefore

$$|\psi(u) - \psi(v)|^2 \geq |\psi(u)|^2 + |\psi(v)|^2 = |\psi^+(u) - \psi^+(v)|^2 + |\psi^-(u) - \psi^-(v)|^2.$$

It follows that

$$\sum_{\{u,v\} \in E} |\psi(u) - \psi(v)|^2 \geq \sum_{\{u,v\} \in E} |\psi^+(u) - \psi^+(v)|^2 + \sum_{\{u,v\} \in E} |\psi^-(u) - \psi^-(v)|^2.$$

Also, we note that

$$\sum_{v \in V} m(v)(\psi(v))^2 = \sum_{v \in V} m(v)(\psi^+(v))^2 + \sum_{v \in V} m(v)(\psi^-(v))^2.$$

Therefore

$$1 - \lambda_G \geq \frac{\sum_{\{u,v\} \in E} |\psi^+(u) - \psi^+(v)|^2 + \sum_{\{u,v\} \in E} |\psi^-(u) - \psi^-(v)|^2}{\sum_{v \in V} m(v)(\psi^+(v))^2 + \sum_{v \in V} m(v)(\psi^-(v))^2}.$$

For every four positive numbers a, b, c, d it holds that $\frac{a+b}{c+d} \geq \min\{\frac{a}{c}, \frac{b}{d}\}$, and therefore

$$1 - \lambda_G \geq \min\left\{ \frac{\sum_{\{u,v\}\in E}|\psi^+(u) - \psi^+(v)|^2}{\sum_{v\in V} m(v)(\psi^+(v))^2}, \frac{\sum_{\{u,v\}\in E}|\psi^-(u) - \psi^-(v)|^2}{\sum_{v\in V} m(v)(\psi^-(v))^2} \right\}.$$

Without loss of generality, we will assume that

$$1 - \lambda_G \geq \frac{\sum_{\{u,v\}\in E}|\psi^+(u) - \psi^+(v)|^2}{\sum_{v\in V} m(v)(\psi^+(v))^2},$$

and finish the proof by showing that

$$\frac{\sum_{\{u,v\}\in E}|\psi^+(u) - \psi^+(v)|^2}{\sum_{v\in V} m(v)(\psi^+(v))^2} \geq \frac{\alpha_G^2}{2}.$$

$$\frac{\sum_{\{u,v\}\in E}|\psi^+(u) - \psi^+(v)|^2}{\sum_{v\in V} m(v)(\psi^+(v))^2} =$$
$$\frac{\left(\sum_{\{u,v\}\in E}|\psi^+(u) - \psi^+(v)|^2\right)\left(\sum_{\{u,v\}\in E}|\psi^+(u) + \psi^+(v)|^2\right)}{\left(\sum_{v\in V} m(v)(\psi^+(v))^2\right)\left(\sum_{\{u,v\}\in E}|\psi^+(u) + \psi^+(v)|^2\right)} \geq^{CS}$$
$$\frac{\left(\sum_{\{u,v\}\in E}|\psi^+(u) - \psi^+(v)||\psi^+(u) + \psi^+(v)|\right)^2}{\left(\sum_{v\in V} m(v)(\psi^+(v))^2\right)\left(\sum_{\{u,v\}\in E}|\psi^+(u) + \psi^+(v)|^2\right)} \geq$$
$$\frac{\left(\sum_{\{u,v\}\in E}|(\psi^+(u))^2 - (\psi^+(v))^2|\right)^2}{\left(\sum_{v\in V} m(v)(\psi^+(v))^2\right)\left(\sum_{\{u,v\}\in E} 2(\psi^+(u))^2 + 2(\psi^+(v))^2\right)} =$$
$$\frac{\left(\sum_{\{u,v\}\in E}|(\psi^+(u))^2 - (\psi^+(v))^2|\right)^2}{2\|\psi^+\|^4}.$$

Thus, in order to finish the proof, we need to prove that

$$\sum_{\{u,v\}\in E}|(\psi^+(u))^2 - (\psi^+(v))^2| \geq \alpha_G\|\psi^+\|^2.$$

By our indexing,

$$\sum_{\{u,v\}\in E}|(\psi^+(u))^2 - (\psi^+(v))^2| = \sum_{\{v_t,v_j\}\in E, t>j}(\psi^+(v_j))^2 - (\psi^+(v_t))^2.$$

We note that for every $\{v_t, v_j\} \in E, t > j$,

$$(\psi^+(v_j))^2 - (\psi^+(v_t))^2 = \sum_{i=j}^{t-1}(\psi^+(v_i))^2 - (\psi^+(v_{i+1}))^2.$$

We recall that $\psi^+(v_k) = ... = \psi^+(v_n) = 0$ and therefore if $t > j \geq k$,

$$(\psi^+(v_j))^2 - (\psi^+(v_t))^2 = 0,$$

and if $t \geq k > j$,

$$(\psi^+(v_j))^2 - (\psi^+(v_t))^2 = (\psi^+(v_{k-1}))^2 + \sum_{i=j}^{k-2} (\psi^+(v_i))^2 - (\psi^+(v_{i+1}))^2.$$

Therefore,

$$\begin{aligned}
&\sum_{\{v_t, v_j\} \in E, t>j} (\psi^+(v_j))^2 - (\psi^+(v_t))^2 \\
&= |\partial U_{k-1}|(\psi^+(v_{k-1}))^2 + \sum_{i=1}^{k-2} |\partial U_i| \left((\psi^+(v_i))^2 - (\psi^+(v_{i+1}))^2 \right) \\
&\geq \alpha_G m(U_{k-1})(\psi^+(v_{k-1}))^2 + \sum_{i=1}^{k-2} \alpha_G m(U_i) \left((\psi^+(v_i))^2 - (\psi^+(v_{i+1}))^2 \right) \\
&= \alpha_G \left(m(U_1)(\psi^+(v_1))^2 + \sum_{i=2}^{k-1} (\psi^+(v_i))^2 (m(U_i) - m(U_{i-1})) \right) \\
&= \alpha_G \left(\sum_{i=1}^{k-1} m(v_i)(\psi^+(v_i))^2 \right) \\
&= \alpha_G \|\psi^+\|^2.
\end{aligned}$$

□

3.4 Exercises

1. Given a finite connected graph $G = (V, E)$, prove that for every $\alpha > 0$, $h_G \geq \alpha$ if and only if

$$\max_{\emptyset \neq U \subseteq V, m(U) \leq \frac{1}{2} m(V)} \frac{|E(U, U)|}{m(U)} \leq \frac{1 - \alpha}{2}$$

   (This fact is called "Alon-Chung Lemma").

2. For a finite connected graph $G = (V, E)$, the *diameter* of G is the longest path-distance between two vertices, i.e.,

$$\operatorname{diam}(G) = \max_{u,v \in V} \operatorname{dist}(u, v).$$

   The next exercise is aimed to prove the fact that in an expander family, the diameter grows at most logarithmicly with respect to the number of vertices:

   Proposition 3.10. *Let $\{G_n = (V_n, E_n)\}_{n \in \mathbb{N}}$ be an expander family. Then there is a constant C, such that* $\operatorname{diam}(G_n) \leq C \log(|V_n|)$.

   More specifically, for a finite graph $G = (V, E)$, if $d = \max_{v \in V} m(v)$, then $\operatorname{diam}(G) \leq \frac{2}{\log(\frac{h_G}{d} + 1)} \log(|V|)$.

A reader looking for a challenge can try to prove this fact without following guidance.

Guidance: for $v \in V$, $r \in \mathbb{N} \cup \{0\}$, denote

$$B_r(v) = \{u \in V : \text{dist}(v,u) \leq r\},$$

$$S_r(v) = \{u \in V : \text{dist}(v,u) = r\}.$$

(a) For every $v \in V$, denote $r(v)$ to be the minimal non-negative integer such that $m(B_{r(v)}(v)) > \frac{1}{2}m(V_n)$. Note that $m(B_0(v)) \leq \frac{1}{2}m(V)$ and therefore $r(v) \geq 1$.

(b) Show that for every $v \in V$ and every $0 \leq r \leq r(v)-1$, $d|S_{r+1}(v)| \geq h_G|B_r(v)|$.

(c) Deduce that for every $0 \leq r \leq r(v)-1$, $|B_{r+1}(v)| \geq (1+\frac{h_G}{d})|B_r(v)|$ and as a result $|B_{r(v)}(v)| \geq (1+\frac{h_G}{d})^{r(v)}$.

(d) Note that $|V| \geq |B_{r(v)}(v)|$ and thus by the previous inequality,

$$r(v) \leq \log_{1+\frac{h_G}{d}}(|V|) = \frac{\log(|V|)}{\log(1+\frac{h_G}{d})}.$$

(e) Note that for every two vertices $u, v \in V$, $m(B_{r(v)}(v)) + m(B_{r(u)}(u)) > m(V)$ and thus $B_{r(v)}(v) \cap B_{r(u)}(u) \neq \emptyset$. Deduce that for every two vertices

$$\text{dist}(u,v) \leq r(u) + r(v) \leq \frac{2}{\log(1+\frac{h_G}{d})}\log(|V|),$$

as needed.

3. The converse of the previous fact does not hold: give an example that logarithmic growth of diameter does **not** imply expansion, i.e., give an example of a family of finite graphs $\{G_n = (V_n, E_n)\}_{n=1}^{\infty}$ such that

(a) $\lim_n |V_n| = \infty$.

(b) $\sup_n (\max_{v \in V_n} m(v)) < \infty$.

(c) There is a constant C such that every G_n, $\text{diam}(G_n) \leq C\log(|V_n|)$.

(d) The Cheeger constant tends to 0: $\lim_n h_{G_n} = 0$.

4 The expander mixing lemma

Here we will show that assuming two-sided spectral gap (or one sided spectral gap + bipartite) leads to a stronger notion of expansion:

Notation: given a graph $G = (V, E)$ and non-empty subsets $U, W \subseteq V$, denote $e(U,W) = |\{(u,v) \in U \times W : \{u,v\} \in E\}|$.

Note that there is a technicality here: if U, W are disjoint, then $e(U,W) = |E(U,W)|$, but if U, W are not disjoint and $u, v \in U \cap W, \{u,w\} \in E$, then the edge $\{u,w\}$ is counted twice, e.g., $e(U,U) = 2|E(U,U)|$.

Theorem 4.1 (The expander mixing lemma - two-sided spectral gap)**.** *Let $G = (V, E)$ be a finite connected graph such that* $\mathrm{Spec}(M) \subseteq [-\lambda, \lambda] \cup \{1\}$*, then for any two non-empty sets $U, W \subseteq V$,*

$$\left| e(U,W) - \frac{m(U)m(W)}{m(V)} \right| \leq \lambda \sqrt{m(U)(1 - \frac{m(U)}{m(V)})m(W)(1 - \frac{m(W)}{m(V)})}.$$

Proof. Let $U, W \subseteq V$ be any two non-empty sets. Take

$$\phi_U = \chi_U - \frac{\langle \chi_U, \mathbb{1} \rangle}{\langle \mathbb{1}, \mathbb{1} \rangle} \mathbb{1},$$

$$\phi_W = \chi_W - \frac{\langle \chi_W, \mathbb{1} \rangle}{\langle \mathbb{1}, \mathbb{1} \rangle} \mathbb{1}.$$

Note that by definition, $\phi_U \perp \mathbb{1}, \phi_W \perp \mathbb{1}$ and therefore $\|M\phi_U\| \leq \lambda \|\phi_U\|$. Thus,

$$|\langle M\phi_U, \phi_W \rangle| \leq \|M\phi_U\| \|\phi_W\|.$$

By calculation, $\langle \chi_U, \mathbb{1} \rangle = m(U), \langle \mathbb{1}, \mathbb{1} \rangle = m(V)$ and therefore

$$\phi_U = \chi_U - \frac{m(U)}{m(V)} \mathbb{1}.$$

It follows that

$$\|\phi_U\|^2 = \|\chi_U\|^2 - 2\frac{m(U)}{m(V)} \langle \chi_U, \mathbb{1} \rangle + \left(\frac{m(U)}{m(V)} \right)^2 \langle \mathbb{1}, \mathbb{1} \rangle = m(U) - \frac{m(U)^2}{m(V)} = m(U)(1 - \frac{m(U)}{m(V)}),$$

and similarly,

$$\|\phi_W\|^2 = m(W)(1 - \frac{m(W)}{m(V)}).$$

Therefore, we get that

$$|\langle M\phi_U, \phi_W \rangle| \leq \lambda \sqrt{m(U)(1 - \frac{m(U)}{m(V)})m(W)(1 - \frac{m(W)}{m(V)})}.$$

Next, we turn to compute the left-hand side of the inequality. Note that $M\phi_U = M\chi_U - \frac{m(U)}{m(V)} \mathbb{1}$ and that

$$(M\chi_U)(v) = \frac{|E(v,U)|}{m(v)}.$$

Therefore

$$\begin{aligned}
\langle M\phi_U, \phi_W \rangle &= \langle M\chi_U, \phi_W \rangle - \frac{m(U)}{m(V)} \langle \mathbb{1}, \phi_W \rangle \\
&=^{\phi_W \perp \mathbb{1}} \langle M\chi_U, \phi_W \rangle \\
&= \langle M\chi_U, \chi_W \rangle - \frac{m(W)}{m(V)} \langle M\chi_U, \mathbb{1} \rangle \\
&=^{M \text{ is self-adjoint}} \langle M\chi_U, \chi_W \rangle - \frac{m(W)}{m(V)} \langle \chi_U, M\mathbb{1} \rangle \\
&=^{M\mathbb{1}=\mathbb{1}} \left(\sum_{v \in W} m(v) \frac{|E(v,U)|}{m(v)} \right) - \frac{m(U)m(W)}{m(V)} \\
&= e(U,W) - \frac{m(U)m(W)}{m(V)}.
\end{aligned}$$

Combining this with the previous inequality yields that

$$\left|e(U,W) - \frac{m(U)m(W)}{m(V)}\right| \leq \lambda\sqrt{m(U)(1-\frac{m(U)}{m(V)})m(W)(1-\frac{m(W)}{m(V)})}.$$

□

Remark 4.2. *In some sources, when referring to the expander mixing lemma, a slightly weaker inequality is stated:*

$$\left|e(U,W) - \frac{m(U)m(W)}{m(V)}\right| \leq \lambda\sqrt{m(U)m(W)}.$$

Remark 4.3. *If the graph G is d-regular, then the inequality of expander mixing lemma reads as*

$$\left|e(U,W) - \frac{d|U||W|}{|V|}\right| \leq d\lambda\sqrt{|U|(1-\frac{|U|}{|V|})|W|(1-\frac{|W|}{|V|})}.$$

Remark 4.4. *Bilu and Linial also proved the converse of the expander mixing lemma: they proved that for a d-regular graph $G = (V, E)$, if there is a constant α such that for every two sets U, W,*

$$\left|e(U,W) - \frac{d|U||W|}{|V|}\right| \leq d\alpha\sqrt{|U||W|},$$

then all the non-trivial eigenvalues of M are bounded in absolute value by $O(\alpha(1+\log(\frac{d}{\alpha}))$.

Note the the expander mixing lemma also gives another proof of the easy side of the Cheeger inequality, under stricter conditions:

Proof. Let G be a connected graph such that $\mathrm{Spec}(M) \subseteq [-\lambda, \lambda] \cup \{1\}$. Let $\emptyset \neq U \subseteq V$ such that $m(U) \leq \frac{1}{2}m(V)$. Then $|\partial U| = e(U, V \setminus U)$ and therefore using $W = V \setminus U$ in the expander mixing lemma, yields

$$\left||\partial U| - \frac{m(U)(m(V)-m(U))}{m(V)}\right| \leq \lambda\sqrt{m(U)(1-\frac{m(U)}{m(V)})(m(V)-m(U))(1-\frac{m(V)-m(U)}{m(V)})}.$$

Thus,

$$\left||\partial U| - \frac{m(U)(m(V)-m(U))}{m(V)}\right| \leq \lambda\frac{m(U)(m(V)-m(U))}{m(V)}.$$

In particular

$$\frac{|\partial U|}{m(U)} \geq (1-\lambda)\frac{m(V)-m(U)}{m(V)} \geq \frac{1-\lambda}{2}.$$

□

A variation of the expander mixing lemma is the bipartite expander mixing lemma in which that assumption of two-sided spectral gap is replaced by the assumption of bipartiteness.

Theorem 4.5 (Bipartite expander mixing lemma)**.** *Let $G = (V, E)$ be a connected bipartite graph with sided S_0, S_1. Then for every two non-empty sets $U_0 \subseteq S_0, U_1 \subseteq S_1$,*

$$\left| \frac{|E(U_0, U_1)|}{|E|} - \frac{m(U_0)}{m(S_0)} \frac{m(U_1)}{m(S_1)} \right| \leq \lambda_G \sqrt{\frac{m(U_0)}{m(S_0)} \frac{m(U_1)}{m(S_1)} (1 - \frac{m(U_0)}{m(S_0)})(1 - \frac{m(U_1)}{m(S_1)})}.$$

Proof. Note that by a previous exercise, -1 is an eigenvalue of G with multiplicity 1 and eigenfunction $\chi_{S_0} - \chi_{S_1}$ and that the spectrum of M is symmetric and therefore $\mathrm{Spec}(M) = \{\pm 1\} \cup [-\lambda_G, \lambda_G]$. Therefore, if $\phi \in \ell^2(V)$ and $\phi \perp \mathrm{span}\{\mathbb{1}, \chi_{S_0} - \chi_{S_1}\}$, then $\|M\phi\| \leq \lambda_G \|\phi\|$. Note that $\mathrm{span}\{\mathbb{1}, \chi_{S_0} - \chi_{S_1}\} = \mathrm{span}\{\chi_{S_0}, \chi_{S_1}\}$. Thus $\phi \perp \mathrm{span}\{\mathbb{1}, \chi_{S_0} - \chi_{S_1}\}$ if and only if $\langle \phi, \chi_{S_0} \rangle = \langle \phi, \chi_{S_1} \rangle = 0$. Define

$$\phi_{U_0} = \chi_{U_0} - \frac{\langle \chi_{U_0}, \chi_{S_0} \rangle}{\langle \chi_{S_0}, \chi_{S_0} \rangle} \chi_{S_0},$$

$$\phi_{U_1} = \chi_{U_1} - \frac{\langle \chi_{U_1}, \chi_{S_1} \rangle}{\langle \chi_{S_1}, \chi_{S_1} \rangle} \chi_{S_1}.$$

Note that by definition, $\phi_{U_0} \perp \chi_{S_0}$ and since ϕ_{U_0} is supported on S_0, then also $\phi_{U_0} \perp \chi_{S_1}$. Similarly, $\phi_{U_1} \perp \mathrm{span}\{\chi_{S_0}, \chi_{S_1}\}$. Thus,

$$|\langle M\phi_{U_0}, \phi_{U_1} \rangle| \leq \lambda \|\phi_{U_0}\| \|\phi_{U_1}\|.$$

By calculations similar to those in the proof of the expander mixing lemma,

$$\|\phi_{U_0}\|^2 = m(U_0)(1 - \frac{m(U_0)}{m(S_0)}),$$

$$\|\phi_{U_1}\|^2 = m(U_1)(1 - \frac{m(U_1)}{m(S_1)}),$$

and

$$\begin{aligned} \langle M\phi_{U_0}, \phi_{U_1} \rangle &= \langle M\chi_{U_0}, \phi_{U_1} \rangle \\ &= \langle M\chi_{U_0}, \chi_{U_1} \rangle - \frac{m(U_1)}{m(S_1)} \langle \chi_{U_0}, \chi_{S_0} \rangle \\ &= |E(U_0, U_1)| - \frac{m(U_0)m(U_1)}{m(S_1)}. \end{aligned}$$

Thus,

$$\left| |E(U_0, U_1)| - \frac{m(U_0)m(U_1)}{m(S_1)} \right| \leq \lambda \sqrt{m(U_0)(1 - \frac{m(U_0)}{m(S_0)}) m(U_1)(1 - \frac{m(U_1)}{m(S_1)})},$$

and dividing the inequality by $m(S_1) = m(S_0) = |E|$ yields the needed result. □

5 Alon-Boppana bound

A natural question is how good of an expander can a graph be? This question is ill-defined as stated because we saw that without bounding the degree, one can take the complete graph on n vertices and for such a graph both the spectrum

and the Cheeger constant are optimal. Thus, the more precise question is how good of an expander can a d-regular graph be?

The Alon-Boppana bound address this question with respect to spectral expansion.

Theorem 5.1 (Alon-Boppana bound)**.** *Let $G = (V, E)$ be a d-regular connected graph with $|V| = n$. If the diameter of G is larger than 5, then*

$$\lambda_G \geq \frac{2\sqrt{d-1}}{d} - \frac{2\sqrt{d-1}-1}{d(\lfloor \frac{\mathrm{diam}(G)}{2} \rfloor - 1)}.$$

where $\mathrm{diam}(G)$ *is the diameter of G with respect to the path-metric.*

Remark 5.2. *The Alon-Boppana bound is not the one theorem stated above, but several different theorems showing that for a d-regular graph $G = (V, E)$, λ_G is bounded from below by $\left(\frac{2\sqrt{d-1}}{d} - \text{error term}\right)$, where the error term tends to 0 as the diameter (or the number of vertices) tends to infinity.*

Proof. Let A be the adjacency matrix of G, i.e., A is an $n \times n$ matrix indexed by V such that

$$A(u, v) = \begin{cases} 1 & \{u, v\} \in E \\ 0 & \{u, v\} \notin E \end{cases}.$$

Since G is d-regular, $M = \frac{1}{d}A$ and we showed in the homework that for every $u, v \in V$, $k \in \mathbb{N}$, $A^k(u, v)$ is the number of path of length k-from u to v. Fix $a, b \in V$ such that $\mathrm{diam}(G) = \mathrm{dist}(a, b)$ and denote $k = \lfloor \frac{\mathrm{diam}(G)}{2} - 2 \rfloor$ (note that $\mathrm{diam}(G) \geq 5$ and therefore $k \geq 0$).

Fix some a' such that $\{a, a'\} \in E$ and define

$$\mathrm{dist}(v, \{a, a'\}) = \min\{\mathrm{dist}(v, a), \mathrm{dist}(v, a')\}.$$

Denote

$$S_t(\{a, a'\}) = \{v \in V : \mathrm{dist}(v, \{a, a'\}) = t\}.$$

Define a function

$$\phi(v) = \begin{cases} \frac{1}{(\sqrt{d-1})^t} & v \in S_t(\{a, a'\}), t \leq k \\ 0 & \text{otherwise} \end{cases}.$$

We want to show that

$$\langle M\phi, \phi \rangle \geq \|\phi\|^2 \left(\frac{2\sqrt{d-1}}{d} - \frac{2\sqrt{d-1}-1}{d(k+1)} \right).$$

This is equivalent to showing that

$$\langle L\phi, \phi \rangle = \|\phi\|^2 - \langle M\phi, \phi \rangle \leq \|\phi\|^2 \left(1 - \frac{2\sqrt{d-1}}{d} + \frac{2\sqrt{d-1}-1}{d(k+1)} \right).$$

For a vertex $v \in V$, denote

$$O_v = \{u : \{u, v\} \in E, \mathrm{dist}(\{a, a'\}, u) = \mathrm{dist}(\{a, a'\}, v) + 1\}.$$

Then

$$\begin{aligned}
\langle L\phi,\phi\rangle &= \sum_{\{u,v\}\in E}|\phi(v)-\phi(u)|^2\\
&=\sum_{t=0}^{k-1}\sum_{v\in S_t(\{a,a'\})}\sum_{u\in O_v}|\phi(v)-\phi(u)|^2+\sum_{v\in S_k(\{a,a'\})}|O_v||\phi(v)|^2\\
&=\sum_{t=0}^{k-1}\sum_{v\in S_t(\{a,a'\})}|O_v||\phi(v)-\frac{\phi(v)}{\sqrt{d-1}}|^2+\sum_{v\in S_k(\{a,a'\})}|O_v||\phi(v)|^2\\
&=\sum_{t=0}^{k-1}\sum_{v\in S_t(\{a,a'\})}|O_v||\phi(v)|^2\frac{d-2\sqrt{d-1}}{d-1}+\sum_{v\in S_k(\{a,a'\})}|O_v||\phi(v)|^2\\
&\leq^{|O_v|\leq d-1}\sum_{t=0}^{k-1}\sum_{v\in S_t(\{a,a'\})}|\phi(v)|^2(d-2\sqrt{d-1})+\sum_{v\in S_k(\{a,a'\})}(d-1)|\phi(v)|^2\\
&=\sum_{t=0}^{k}\sum_{v\in S_t(\{a,a'\})}d|\phi(v)|^2(1-\frac{2\sqrt{d-1}}{d})+\sum_{v\in S_k(\{a,a'\})}(2\sqrt{d-1}-1)|\phi(v)|^2\\
&=\|\phi\|^2(1-\frac{2\sqrt{d-1}}{d})+\frac{2\sqrt{d-1}-1}{d}\sum_{v\in S_k(\{a,a'\})}d|\phi(v)|^2.
\end{aligned}$$

To finish the proof of the inequality, it remains to prove that

$$\sum_{v\in S_k(\{a,a'\})}d|\phi(v)|^2\leq\frac{1}{k+1}\|\phi\|^2.$$

We note that

$$\sum_{v\ \in S_k(\{a,a'\})}|\phi(v)|^2=|S_k|\frac{1}{(d-1)^k}.$$

Since $|S_k|\leq(d-1)|S_{k-1}|\leq\ldots\leq(d-1)^k|S_0|$, it follows that

$$|S_k|\frac{1}{(d-1)^k}\leq\frac{1}{k+1}\sum_{t=0}^{k}|S_t|\frac{1}{(d-1)^t}=\frac{1}{d(k+1)}\|\phi\|^2.$$

Define ψ similarly:

$$\psi(v)=\begin{cases}\frac{1}{(\sqrt{d-1})^t} & v\in S_t(\{b,b'\}),t\leq k\\ 0 & \text{otherwise}\end{cases}.$$

Then, by the same considerations,

$$\langle M\psi,\psi\rangle\geq\|\psi\|^2\left(\frac{2\sqrt{d-1}}{d}-\frac{2\sqrt{d-1}-1}{d(k+1)}\right).$$

Note that ϕ is supported on the ball

$$\{v:\operatorname{dist}(a,v)\leq\frac{\operatorname{diam}(G)}{2}-1\},$$

and ψ is supported on the ball

$$\{v : \text{dist}(b, v) \leq \frac{\text{diam}(G)}{2} - 1\}.$$

Since, $\text{dist}(a, b) = \text{diam}(G)$, it follows that the supports of ϕ and ψ are disjoint and therefore $\langle \phi, \psi \rangle = 0$. Also, $M\phi$ is supported on the ball

$$\{v : \text{dist}(a, v) \leq \frac{\text{diam}(G)}{2}\},$$

and therefore $\langle M\phi, \psi \rangle = 0$ (and since M is self-adjoint, $\langle M\psi, \phi \rangle = 0$). We note that ϕ, ψ are orthogonal and therefore span a subspace of dimension 2. It follows that there are non-zero α, β such that $(\alpha\phi + \beta\psi) \perp \mathbb{1}$. Without loss of generality, we can choose such α, β such that $\|\alpha\phi + \beta\psi\|^2 = 1$, and we note that by the orthogonality of ϕ, ψ,

$$\|\alpha\phi + \beta\psi\|^2 = |\alpha|^2\|\phi\|^2 + |\beta|^2\|\psi\|^2.$$

For such α, β, the following holds:

$$\begin{aligned}
\lambda_G &\geq \langle M(\alpha\phi + \beta\psi), \alpha\phi + \beta\psi \rangle \\
&= |\alpha|^2\langle M\phi, \phi \rangle + |\beta|^2\langle M\psi, \psi \rangle \\
&\geq \left(|\alpha|^2\|\phi\|^2 + |\beta|^2\|\psi\|^2\right)\left(\frac{2\sqrt{d-1}}{d} - \frac{2\sqrt{d-1}-1}{k+1}\right) \\
&= \frac{2\sqrt{d-1}}{d} - \frac{2\sqrt{d-1}-1}{k+1},
\end{aligned}$$

as needed.

□

By the following general proposition, we can replace the diameter in the above theorem, with $\log_d(|V|)$:

Proposition 5.3. *If $G = (V, E)$ is a connected finite graph such that the maximal degree is bounded from above by $d \geq 2$, then for every $k \in \mathbb{N} \cup \{0\}$ and every $v \in V$, $|B_k(v)| \leq d^{k+1}$. As a result,* $\text{diam}(G) + 1 \geq \log_d(|V|)$.

Proof. Denote

$$B_k(v) = \{u \in V : \text{dist}(v, u) \leq k\},$$
$$S_k(v) = \{u \in V : \text{dist}(v, u) = k\}.$$

The maximal degree of G is bounded by d, thus for every $v \in V$, $|S_0(v)| = 1, |S_1(v)| = d$ and for every $k > 1$,

$$|S_k(v)| \leq (d-1)|S_{k-1}(v)| \leq ... \leq (d-1)^{k-1}d.$$

Therefore for every $k > 1$

$$|B_k(v)| = \sum_{i=0}^{k} |S_k(v)| \leq$$

$$1 + d + d(d-1) + ... + d(d-1)^{k-1} \leq 1 + d + ... + d^k = \frac{d^{k+1} - 1}{d-1} \leq d^{k+1}.$$

Note that for every vertex v, $B_{\text{diam}(G)}(v) = V$ and therefore $|V| \leq d^{\text{diam}(G)+1}$, which implies that $\log_d(|V|) \leq \text{diam}(G) + 1$. □

Corollary 5.4. *For every d-regular connected finite graph, $G = (V, E)$,*

$$\lambda_G \geq \frac{2\sqrt{d-1}}{d} - \frac{2\sqrt{d-1}-1}{\lfloor \frac{\log_d(|V|)-1}{2} \rfloor - 1}.$$

Proof. Combining the above proposition with the Alon-Boppana bound, yields that for every d-regular graph $G = (V, E)$,

$$\lambda_G \geq \frac{2\sqrt{d-1}}{d} - \frac{2\sqrt{d-1}-1}{\lfloor \frac{\log_d(|V|)-1}{2} \rfloor - 1}.$$

□

Corollary 5.5. *For every $\varepsilon > 0$, there are only finitely many d-regular graphs G for which $\lambda_G \leq \frac{2\sqrt{d-1}}{d} - \varepsilon$.*

Also, if $\{G_n = (V_n, E_n)\}_{n\in\mathbb{N}}$ is a family of d-regular λ-spectral expanders, then $\lambda \geq \frac{2\sqrt{d-1}}{d}$.

Proof. By the above corollary, for every $\varepsilon > 0$, there is a constant N such that if $|V| > N$, then

$$\lambda_G > \frac{2\sqrt{d-1}}{d} - \varepsilon.$$

Since there are only finitely many different d-regular graphs on $\leq N$ vertices, the assertion follows.

Regarding the second statement - one should note that by definition of an expander family $|V_n|$ tends to infinity and in particular there are infinitely many non-isomorphic graphs in the family. □

The Alon-Boppana bound (and the above corollary) raises the question whether there are such optimal family of expanders. To be precise, define:

Definition 5.6 (Ramanujan graphs)**.** *A d-regular graph G is Ramanujan if it is a $\frac{2\sqrt{d-1}}{d}$-two-sided spectral expander.*

A d-regular bipartite G graph is bipartite Ramanujan if $\lambda_{G_n} \leq \frac{2\sqrt{d-1}}{d}$.

A family of d-regular graphs $G_n = (V_n, E_n)$, where $|V_n| \to \infty$ is called Ramanujan if for every n, G_n is a Ramanujan.

A family of d-regular bipartite graphs $G_n = (V_n, E_n)$, where $|V_n| \to \infty$ is called bipartite Ramanujan if for every n, G_n is bipartite Ramanujan.

A brief history of Ramanujan graphs:

- In 1988, Lubotzky, Phillips, Sarnak first defined Ramanujan graphs and gave the first construction of a family of such graphs for $d = p + 1$, where p is prime and $p \equiv 1 \mod 4$. Their construction relied on the Ramanujan conjecture (and hence the name).
- In 1994, Morgenstern extended the above result and gave a construction for a family of such graphs for $d = p + 1$, where p is any prime power.
- In 2002, Friedman proved that random d-regular graphs are almost Ramanujan, i.e., that for every $\varepsilon > 0$, there is a positive function $f(n)$ tending to 1 as $n \to \infty$ such that the probability that a randomly chosen d-regular graph on n vertices will be a $(\frac{2\sqrt{d-1}}{d} + \varepsilon)$-two-sided spectral expander is greater or equal to $f(n)$.

- In 2013, Marcus, Spielman and Srivastava proved that bipartite Ramanujan graphs exists for every $d \geq 3$ and later in 2015 they proved that bipartite Ramanujan graphs exists for every $d \geq 3$ and any number of vertices.

- Open conjectures (at least at the time these notes where written):

  1. (non-bipartite) Ramanujan graphs exists for every $d \geq 3$.
  2. There is some positive $\alpha > 0$ such that for every (large enough) n, the probability that a random d-regular graph on n vertices is Ramanujan is $\geq \alpha$.

6 Existence of expander families - a non-constructive proof

Friedman's result mentioned above states that a randomly chosen d-regular graph (with $d \geq 3$) on n vertices ($n >> d$) will have almost the best expansion one can hope for both in a spectral sense. Thus fixing some d and randomly choosing d-regular graphs on n vertices, when n tends to infinity will give us an expander family with positive probability (and thus such a family exists).

We will not prove this fact in its full strength, but only show that a randomly chosen d-regular bipartite graph G_n on $2n$ vertices, will have with positive probability $h_{G_n} \geq \frac{1}{4d}$. We will actually use a variant of the Cheeger constant known as vertex expansion (and not edge expansion):

Definition 6.1 (Vertex expansion)**.** *Let $G = (V, E)$ be a finite connected graph. For a non-empty set $U \subseteq V$, we define $\partial_{out}U$ as the set of vertices in $V \setminus U$ that have at least one neighbor in U and define*

$$h_{out}(G) = \min_{\emptyset \neq U \subseteq V, m(U) \leq \frac{1}{2}m(V)} \frac{|\partial_{out}U|}{m(U)}.$$

We also define $\partial_{in}U$ as the set of vertices in U that have a neighbor in $V \setminus U$ and define

$$h_{in}(G) = \min_{\emptyset \neq U \subseteq V, m(U) \leq \frac{1}{2}m(V)} \frac{|\partial_{in}U|}{m(U)}.$$

Observe that for every graph $G = (V, E)$ and every non-empty set $U \subseteq V$, $|\partial_{out}U| \leq |\partial U|$ and $|\partial_{in}U| \leq |\partial U|$ and therefore $h_{out}(G) \leq h_G$ and $h_{in}(G) \leq h_G$. Also, if G the maximal degree of G is bounded by d, then $dh_{out}(G) \geq h_G$, $dh_{in}(G) \geq h_G$.

In order to avoid technical difficulties, we will generalize our definition of a graph and allow multiple edges between two vertices. In other words, between two vertices there can be $k \in \mathbb{N} \cup \{0\}$ edges. The degree of a vertex is still the number of edges connected to it. This is not important to this example, but all our previous results regarding the random walk operator, the Laplacian pass and the Cheeger inequality pass verbatim to this setting, when one takes $M(v, u) = \frac{m(v,u)}{m(v)}$, where $m(v, u)$ is the number of edges connecting u and v.

Fix $d \geq 8$ and let $n \in \mathbb{N}$ (assume that $n >> d$). Recall that a permutation on $\{1, ..., n\}$ is a bijective map $\pi : \{1, ..., n\} \to \{1, ..., n\}$ and there are $n!$ such maps.

Given d such permutations $\pi_1, ..., \pi_d$ define a bipartite graph $G_n(\pi_1, ..., \pi_d)$ on $2n$ vertices as follows:

- The vertices of the graph are $V_n = S_0 \cup S_1$, where $S_0 = \{v_1, ..., v_n\}, S_1 = \{u_1, ..., u_n\}$.
- Given $1 \leq i, j \leq n$, the number of edges between v_i and u_j is $|\{1 \leq k \leq d : \pi_k(i) = j\}$. In other words, the edges are the **multiset** $\{\{v_i, u_{\pi_k(i)}\} : 1 \leq i \leq n, 1 \leq k \leq d\}$.

For example, if $d = 8$ and $\pi_1, ..., \pi_8$ are all equal to the trivial permutation (mapping i to i for all i), then $G(\pi_1, ..., \pi_8)$ will be the graph where v_i and u_j are not connected if $i \neq j$ and there are 8 edges between v_i and u_i for every i.

Note that for every choice of permutations $\pi_1, ..., \pi_d$, $G_n(\pi_1, ..., \pi_d)$ is a d-regular bipartite graph.

Theorem 6.2. *Let $d > 7$ be fixed. Then*

$$\lim_n \frac{|\{\pi_1, ..., \pi_d : h_{out}(G_n(\pi_1, ..., \pi_d)) \geq \frac{1}{4d}\}|}{(n!)^d} = 1.$$

In other words, for a randomly chosen $\pi_1, ..., \pi_d$, the probability that $h_{out}(G_n(\pi_1, ..., \pi_d)) \geq \frac{1}{2}$ tends to 1 as n tends to infinity.

Proof. We will bound from above the number choices of $\pi_1, ..., \pi_d$ for which $h_{out}(G_n(\pi_1, ..., \pi_d)) < \frac{1}{4d}$.

We note that we are talking about d-regular graphs and therefore $m(U) = d|U|$ and $m(U) \leq \frac{1}{2}m(V)$ is equivalent to $|U| \leq n$. For a set $U \subseteq V$, denote $U_0 = U \cap S_0, U_1 = U \cap S_1$. Then

$$\partial_{out} U = ((\partial_{out} U_0) \setminus U_1) \cup ((\partial_{out} U_1) \setminus U_0).$$

Therefore, if $h_{out}(G) < \alpha$, then for there is a set U such that

$$\alpha > \frac{|\partial_{out} U|}{d|U|} \geq \begin{cases} \frac{|(\partial_{out} U_0) \setminus U_1|}{d|U|} & |U_0| \geq |U_1| \\ \frac{|(\partial_{out} U_1) \setminus U_0|}{d|U|} & |U_1| > |U_0| \end{cases}.$$

Note that if $|U_0| - 2 \geq |U_1|$, we get a set U' for which $|U'_0| \geq |U'_1|$ and

$$\frac{|(\partial_{out} U_0) \setminus U_1|}{d|U|} \geq \frac{|(\partial_{out} U'_0) \setminus U'_1|}{d|U'|},$$

by transferring one vertex from U_0 to ∂U_0 (and similarly if $|U_1| - 2 \geq |U_0|$). Thus, up to a rounding error, we can assume that $|U_0| = |U_1| \leq \frac{n}{2}$ and get that there is $i = 0, 1$, and $\emptyset \neq U_i \subseteq S_i$ such that $|U_i| \leq \frac{n}{2}$ and

$$\alpha > \frac{|(\partial_{out} U_i) \setminus U_{i+1}|}{d|U|} = \frac{|\partial_{out} U_i| - |U_i|}{2d|U_i|}.$$

As a result, for $\alpha = \frac{1}{4d}$, there is such $U_i \subseteq S_i$ such that $\frac{3}{2}|U_i| > |\partial_{out} U_i|$. Given d, n, we will call a choice $\pi_1, ..., \pi_d$ bad if there is such U_0 exists.

Let $\emptyset \neq A \subseteq S_0$, with $|A| \leq \frac{n}{2}$ and let $B \subseteq S_1$ with $|B| = \frac{3}{2}|A|$. For how many choices of $\pi_1, ..., \pi_d$ does $\partial_{out} A \subseteq B$?

For a single permutation π, there are $\binom{|B|}{|A|}|A|!\,(n-|A|)!$ possibilities and thus for d permutations there are $\left(\binom{|B|}{|A|}|A|!\,(n-|A|)!\right)^d$ possibilities. Thus, running over all the possible choices of A and B yields that the total number of bad choices is bounded from above by

$$\begin{aligned}
&\sum_{t=1}^{\frac{n}{2}} \binom{n}{t}\binom{n}{\lceil\frac{3}{2}t\rceil}\left(\binom{\lceil\frac{3}{2}t\rceil}{t}t!\,(n-t)!\right)^d \\
&= \sum_{t=1}^{\frac{n}{2}} \binom{n}{t}\binom{n}{t}\frac{(n-t)!}{(n-\lceil\frac{3}{2}t\rceil)!\,(\lceil\frac{3}{2}t\rceil-t)!}\left(\binom{\lceil\frac{3}{2}t\rceil}{t}\right)^{d-1}(t!\,(n-t)!\,)^d \\
&= \sum_{t=1}^{\frac{n}{2}} (n!\,)^2\frac{(n-t)!}{(n-\lceil\frac{3}{2}t\rceil)!\,(\lceil\frac{3}{2}t\rceil-t)!}\left(\binom{\lceil\frac{3}{2}t\rceil}{t}\right)^{d-1}(t!\,(n-t)!\,)^{d-2}.
\end{aligned}$$

Denote

$$a_t = (n!\,)^2\frac{(n-t)!}{(n-\lceil\frac{3}{2}t\rceil)!\,(\lceil\frac{3}{2}t\rceil-t)!}\left(\binom{\lceil\frac{3}{2}t\rceil}{t}\right)^{d-1}(t!\,(n-t)!\,)^{d-2}.$$

We want to understand at what values does a_t attains its maximum. We consider $\frac{a_{t+2}}{a_t}$ in order to determined when is the sequence increasing and when it is decreasing (we take $t+2$ and not $t+1$ to avoid rounding issues).

$$\begin{aligned}
\frac{a_{t+2}}{a_t} &= \frac{\frac{(n-t-2)!}{(n-\lceil\frac{3}{2}t\rceil-3)!(\lceil\frac{3}{2}t\rceil+1-t)!}}{\frac{(n-t)!}{(n-\lceil\frac{3}{2}t\rceil)!(\lceil\frac{3}{2}t\rceil-t)!}}\left(\frac{\binom{\lceil\frac{3}{2}t\rceil+3}{t+2}}{\binom{\lceil\frac{3}{2}t\rceil}{t}}\right)^{d-1}\left(\frac{(t+2)!\,(n-(t+2))!}{t!\,(n-t)!}\right)^{d-2} \\
&= \left(\frac{(\lceil\frac{3}{2}t\rceil+3)(\lceil\frac{3}{2}t\rceil+2)(\lceil\frac{3}{2}t\rceil+1)}{(t+2)(t+1)(\lceil\frac{3}{2}t\rceil+1-t)}\right)^{d-1}\cdot \\
&\left(\frac{(n-\lceil\frac{3}{2}t\rceil)(n-\lceil\frac{3}{2}t\rceil-1)(n-\lceil\frac{3}{2}t\rceil-2)}{(n-t)(n-t-1)(\lceil\frac{3}{2}t\rceil+1-t)}\right)\left(\frac{(t+2)(t+1)}{(n-t)(n-t-1)}\right)^{d-2}.
\end{aligned}$$

Note that:

$$\left(\frac{(\lceil\frac{3}{2}t\rceil+3)(\lceil\frac{3}{2}t\rceil+2)(\lceil\frac{3}{2}t\rceil+1)}{(t+2)(t+1)(\lceil\frac{3}{2}t\rceil+1-t)}\right)^{d-1} \to^{t\to\infty} \left(\frac{27}{4}\right)^{d-1},$$

and this sequence is bounded by some constant M (independent of n). Also note that assuming that for $n >> t$, $\left(\frac{(n-\lceil\frac{3}{2}t\rceil)(n-\lceil\frac{3}{2}t\rceil-1)(n-\lceil\frac{3}{2}t\rceil-2)}{(n-t)(n-t-1)(\lceil\frac{3}{2}t\rceil+1-t)}\right)\left(\frac{(t+2)(t+1)}{(n-t)(n-t-1)}\right)^{d-2}$ is a positive number that is very close to 0. Thus for small values of t (assuming that n is sufficiently large), $\left(\frac{(n-\lceil\frac{3}{2}t\rceil)(n-\lceil\frac{3}{2}t\rceil-1)(n-\lceil\frac{3}{2}t\rceil-2)}{(n-t)(n-t-1)(\lceil\frac{3}{2}t\rceil+1-t)}\right)\left(\frac{(t+2)(t+1)}{(n-t)(n-t-1)}\right)^{d-2} << \frac{1}{M}$ and thus $\frac{a_{t+2}}{a_t} < 1$. For large values of t, $t = \beta n$,

$$\left(\frac{(t+2)(t+1)}{(n-t)(n-t-1)}\right)^{d-2} \approx \left(\frac{\beta^2}{(1-\beta)^2}\right)^{d-2},$$

and

$$\left(\frac{(n-\lceil\frac{3}{2}t\rceil)(n-\lceil\frac{3}{2}t\rceil-1)(n-\lceil\frac{3}{2}t\rceil-2)}{(n-t)(n-t-1)(\lceil\frac{3}{2}t\rceil+1-t)}\right) \approx \frac{(1-\frac{3}{2}\beta)^3}{(1-\beta)^2\frac{1}{2}\beta} \geq^{0\leq\beta\leq\frac{1}{2}} \frac{2}{4^3}.$$

Thus when $t \geq \beta n$ and $\frac{2}{4^3}\left(\frac{\beta^2}{(1-\beta)^2}\right)^{d-2} \geq \frac{1}{\left(\frac{27}{4}\right)^{d-1}}$ the sequence is increasing. This argument is cheating by a little, since we replaced all the expression with their limits, but it can be shown that this does not change much: the sequence first decrease rapidly, then stagnate a little and then increase (when the proportion of the increasing ratio is also increasing). Thus for even t's the sequence gets its maximum at $t = 2$ and $t = \frac{n}{2}$ and for odd t's the sequence gets its maximum at $t = 1$ and $t = \frac{n}{2}$. It is easy to check that $a_1 > a_2$ and therefore the maximum of a_t is achieved either in $t = 1$ or in $t = \frac{n}{2}$. Thus, the number of bad choices is bounded by $\frac{n}{2}(a_1 + a_{\frac{n}{2}})$. To finish the proof, we will show that

$$\frac{na_1}{(n!)^d} \to 0, \text{ and } \frac{na_{\frac{n}{2}}}{(n!)^d} \to 0.$$

First,

$$\frac{na_1}{(n!)^d} = \frac{n(n!)^2(n-1)(n-2)2^{d-1}((n-1)!)^{d-2}}{2(n!)^d} = 2^{d-1}\frac{(n-1)(n-2)}{n^{d-3}} \to 0.$$

Second,

$$\frac{na_{\frac{n}{2}}}{(n!)^d} = n\frac{\frac{n}{2}!}{\frac{n}{4}!\,\frac{n}{4}!}\left(\binom{\frac{3}{4}n}{\frac{n}{2}}\right)^{d-1}\left(\frac{((\frac{n}{2})!)^2}{(n!)}\right)^{d-2} \leq n2^{\frac{3}{4}n(d-1)+\frac{n}{2}}\left(\frac{((\frac{n}{2})!)^2}{(n!)}\right)^{d-2}.$$

Recall that by Stirling approximation

$$\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n} \leq n! \leq en^{n+\frac{1}{2}}e^{-n}.$$

Thus

$$\frac{((\frac{n}{2})!)^2}{(n!)} \leq \frac{e^2(\frac{n}{2})^{n+1}}{\sqrt{2\pi}n^{n+\frac{1}{2}}} \leq C\sqrt{n}2^{-n}.$$

Therefore

$$\frac{na_{\frac{n}{2}}}{(n!)^d} \leq Cn^{\frac{3}{2}}2^{n(\frac{3}{4}(d-1)+\frac{1}{2}-(d-2))},$$

and if $d > 7$, then $\frac{3}{4}(d-1) + \frac{1}{2} - (d-2) < 0$ and this expression tends to 0 as $n \to \infty$. □

7 Metric distortion of expander graphs

The following result shows that spectral expansion is "non-euclidean", specifically, the graph metric of a spectral expander is distorted when mapped into a euclidean space. Given a graph G, define

$$c_2(G) = \inf_{k\in\mathbb{N}}\inf_{\Phi:V\to\mathbb{R}^k}\left(\sup_{u,v\in V}\frac{\|\Phi(u)-\Phi(v)\|_{\mathbb{R}^k}}{\text{dist}(u,v)}\right)\left(\sup_{u,v\in V}\frac{\text{dist}(u,v)}{\|\Phi(u)-\Phi(v)\|_{\mathbb{R}^k}}\right),$$

where $\|.\|_{\mathbb{R}^k}$ denotes the euclidean norm.

Note that by definition, $c_2(G) \geq 1$ and it measures the distortion of the metric of G when it is be mapped into some $\mathbb{R}^k$. Examples:

1. $G_n = (V_n, E_n)$ where $V_n = \{1, ..., n\}$, $E_n = \{\{i, i+1\} : 0 \leq i \leq 2n-1\}$. In this case $c_2(G_n) = 1$ for every n, isometrically embedding the graph in $\mathbb{R}$.

2. Fix $k \in \mathbb{N}$ and define $G_n = (V_n, E_n)$ as

$$V_n = \{(a_1, ..., a_k) : a_1, ..., a_k \in \mathbb{Z}, 1 \leq a_1, ..., a_k \leq n\},$$

$$E_n = \{\{(a_1, ..., a_k), (b_1, ..., b_k)\} : \exists i_0, a_{i_0} + 1 = b_{i_0}, \forall i \neq i_0, a_i = b_i\}.$$

Here there is also a very natural embedding of the graph into $\mathbb{R}^k$. We leave it to the reader to check that using that embedding, $c_2(G) \leq \sqrt{k}$ for every n.

In the above examples, the distortion $c_2(G_n)$ is fixed and independent of the size of G_n. The reader may recall that these examples were used before as examples of families of graphs that are clearly not expanders.

Theorem 7.1. *Let $0 \leq \lambda < 1$ be a constant and $G = (V, E)$ be a finite connected λ-spectral expander. Let $d \in \mathbb{N}, d \geq 2$ such that $d \geq \max_{v \in V} m(v)$. Then for any $k \in \mathbb{N}$ and any map $\Phi : V \to \mathbb{R}^k$, if $\|\Phi(v) - \Phi(u)\| \leq \beta$ for every $\{u, v\} \in E$, then there are vertices $v, u \in V$ such that $\mathrm{dist}(u, v) \geq \log_d(|V|) - 3$ and $\|\Phi(u) - \Phi(v)\|_{\mathbb{R}^k} \leq 2\frac{\beta}{\sqrt{1-\lambda}}$.*

Proof. We note that $\Phi = (\phi_1, ..., \phi_k)$, when $\phi_i \in \ell^2(V)$. For every $1 \leq i \leq k$, we denote $\phi_i^0 \in \ell^2(V)$ to be the projection of ϕ_i on the space of constant maps in $\ell^2(V)$ and define $\Psi : V \to \mathbb{R}^k$ as $\Psi = (\phi_1 - \phi_1^0, ..., \phi_k - \phi_k^0) = (\psi_1, ..., \psi_k)$. We note that since $\phi_i^0 \in \ell^2(V)$ are constant maps for every $1 \leq i \leq k$, we have that for every two vertices $u, v \in V$,

$$\|\Psi(u) - \Psi(v)\|_{\mathbb{R}^k}^2 = \sum_{i=1}^{k} |(\phi_i(u) - \phi_i^0(u)) - (\phi_i(v) - \phi_i^0(v))|^2 =$$

$$\sum_{i=1}^{k} |\phi_i(u) - \phi_i(v)|^2 = \|\Phi(u) - \Phi(v)\|_{\mathbb{R}^k}^2.$$

Therefore, it is enough to prove the theorem for Ψ instead of Φ. Using a previous proposition regarding the Laplacian, we note that for every $1 \leq i \leq k$,

$$\langle L\phi_i, \phi_i \rangle = \sum_{\{u,v\} \in E} |\phi_i(u) - \phi_i(v)|^2.$$

By the assumption of λ-spectral expansion, it follows that

$$\sum_{\{u,v\} \in E} |\phi_i(u) - \phi_i(v)|^2 = \langle L\phi_i, \phi_i \rangle \geq (1 - \lambda) \sum_{v \in V} m(v) |\phi_i - \phi_i^0|^2.$$

Summing this inequality over all $1 \leq i \leq k$ yields

$$(1 - \lambda) \sum_{v \in V} m(v) \|\Psi(v)\|_{\mathbb{R}^k}^2 \leq \sum_{\{u,v\} \in E} \|\Phi(u) - \Phi(v)\|_{\mathbb{R}^k}^2.$$

Let $\mathrm{Med}\,(\{\|\Psi(v)\|_{\mathbb{R}^k} : v \in V\})$ be the median that the of the set $\{\|\Psi(v)\|_{\mathbb{R}^k} : v \in V\}$ when accounting to the degree defined as follows: order V such that $\|\Psi(v_1)\|_{\mathbb{R}^k} \leq \|\Psi(v_2)\|_{\mathbb{R}^k} \leq ...$, denote i_{Med} to be the maximal index such that $m(\{v_1, ..., v_{i_{\mathrm{Med}}}\}) \leq \frac{1}{2} m(V)$ and set $\mathrm{Med}\,(\{\|\Psi(v)\|_{\mathbb{R}^k} : v \in V\}) = \|\Psi(v_{\mathrm{Med}})\|_{\mathbb{R}^k}$.

By this definition,

$$\sum_{v\in V} m(v)\|\Psi(v)\|^2_{\mathbb{R}^k} \geq \frac{m(V)}{2}\left(\operatorname{Med}\left(\{\|\Psi(v)\|_{\mathbb{R}^k}\colon v\in V\}\right)\right)^2.$$

Thus,

$$(1-\lambda)\frac{m(V)}{2}\left(\operatorname{Med}\left(\{\|\Psi(v)\|_{\mathbb{R}^k} : v\in V\}\right)\right)^2 \leq \sum_{\{u,v\}\in E}\|\Phi(u)-\Phi(v)\|^2_{\mathbb{R}^k} \leq |E|\beta^2 \leq \frac{m(V)}{2}\beta^2.$$

Therefore

$$\operatorname{Med}\left(\{\|\Psi(v)\|_{\mathbb{R}^k}\colon v\in V\}\right) \leq \frac{\beta}{\sqrt{1-\lambda}}.$$

Let $U\subseteq V$ such that

$$U=\{v\in V : \|\Psi(v)\|_{\mathbb{R}^k} \leq \operatorname{Med}\left(\{\|\Psi(v)\|_{\mathbb{R}^k}\colon v\in V\}\right)\}.$$

By definition, $m(U)+d\geq\frac{m(V)}{2}$ and therefore

$$\frac{|V|}{2}-d\leq\frac{m(V)}{2}-d\leq m(U)\leq d|U|.$$

It follows that $|U|\geq\frac{|V|}{2d}-1$. Recall that we saw that $|B_r(v)|\leq d^{r+1}$ and note

$$|U|\geq\frac{|V|}{2d}-1>\frac{\frac{|V|}{2d}}{d}=d^{\log_d(\frac{|V|}{2d})-1}.$$

Therefore it follows that for any fixed $v\in U$,

$$|U|> d^{\log_d(\frac{|V|}{2d})-1}\geq|B_{\log_d(\frac{|V|}{2d})-2}(v)|.$$

In other words, there is some $u\in U$ such that $u\notin B_{\log_d(\frac{|V|}{2d})-2}(v)$ and thus

$$\operatorname{dist}(u,v)\geq\log_d(\frac{|V|}{2d})-1=\log_d(|V|)-\log_d(d)-\log_d(2)-1\geq\log_d(|V|)-3.$$

On there other hand, since $u,v\in U$, it follows that

$$\|\Psi(u)-\Psi(v)\|\leq\|\Psi(u)\|_{\mathbb{R}^k}+\|\Psi(v)\|_{\mathbb{R}^k}\leq 2\operatorname{Med}\left(\{\|\Psi(v)\|_{\mathbb{R}^k}\colon v\in V\}\right)\leq 2\frac{\beta}{\sqrt{1-\lambda}}.$$

□

Corollary 7.2. *Let $0\leq\lambda<1$ be a constant and $G=(V,E)$ be a finite connected λ-spectral expander. Let $d\in\mathbb{N}, d\geq 2$ such that $d\geq\max_{v\in V}m(v)$. Then $c_2(G)\geq\frac{\sqrt{1-\lambda}}{2}(\log_d(|V|)-3)$.*

In particular, if $\{G_n=(V_n,E_n)\}_{n\in\mathbb{N}}$ is an expander family, then $\lim_n c_2(G_n)=\infty$.

8 Application of expanders to error-correcting codes

Definition 8.1. *A binary code of block length n is a subset $C \subseteq \{0,1\}^n$. A binary code C is called linear if it is a linear subspace of $\{0,1\}^n$ over the field $\mathbb{F}_2$, i.e., if for every $\overline{v}, \overline{u} \in C$, it follows that $\overline{v} + \overline{u} \in C$ where addition is performed* mod 2*.*

Some terminology: let $C \subseteq \{0,1\}^n$ be a code.

- Every $\overline{v} \in C$ is called a *codeword*.
- The *rate* of the code is $\text{Rate}(C) = \frac{\log_2(|C|))}{n}$. If C is a linear code, then we note that the rate is equal to $\frac{\dim(C)}{n}$.
- Denote $\|.\|_1$ to be the Hamming norm (i.e., the ℓ^1 norm) in $\{0,1\}^n$ and define the *relative distance* of C to be

$$\delta(C) = \min_{\overline{v},\overline{u} \in C, \overline{v} \neq \overline{u}} \frac{1}{n} \|\overline{v} - \overline{u}\|_1,$$

(the word "relative" in the definition refers to normalizing by $\frac{1}{n}$). We note that if C is linear then the relative distance of C is equal to

$$w(C) = \min_{\overline{v} \in C, \overline{v} \neq \overline{0}} \frac{1}{n} \|\overline{v}\|_1,$$

which is called the *weight* of C.

Motivation: The idea behind error correcting codes is sending a massage over a noisy channel such that the person receiving this massage can correct it after it has been (slightly) corrupted by the noise. Think about Phonetic Alphabet in Ham radio: people using Ham radio do not say "Bee" and "Cee" on it, since these words are "too close" to each other and when there is noise in the channel, the person on the other side may think you said "Bee" when you actually said "Cee". Instead, Phonetic Alphabet is used: "Bravo" and "Charlie" are used in lieu of "Bee" and "Cee". Since "Bravo" and "Charlie" are "far apart", the listener will not confuse one for the other.

In our case, the codewords are words that we allow to transmit and the noise is changing some of the bits in our codewords. There is a trade-off between the rate of the code and the distance of the code. For instance, we can consider the following extremes:

1. If we just want to send a yes/no answer, then we can take $C = \{\overline{0}, \overline{1}\} \subseteq \{0,1\}^n$. It is easy to check that $\text{Rate}(C) = \frac{1}{n}$ (which is the smallest possible) and $\delta(C) = 1$ (which is the best possible).
2. If we want the rate of C to be 1, we have to take $C = \{0,1\}^n$. For this C, $\delta(C) = \frac{1}{n}$ (which is the smallest possible).

Proposition 8.2 (Hamming bound)**.** *Given a code $C \subseteq \{0,1\}^n$, if $\delta(C) = \delta$, then*

$$\text{Rate}(C) \leq 1 - \frac{1}{n} \log_2\Big(\sum_{j=0}^{\lfloor \frac{n\delta - 1}{2} \rfloor} \binom{n}{j}\Big).$$

Proof. Denote for s,

$$B_s(\overline{x}) = \{\overline{y} \in \{0,1\}^n : \|\overline{y} - \overline{x}\|_1 \leq s\}.$$

Note that for $s_0 = \lfloor \frac{n\delta - 1}{2} \rfloor$, we have for every $\overline{x}, \overline{x}' \in C, \overline{x} \neq \overline{x}'$,

$$B_{s_0}(\overline{x}) \cap B_{s_0}(\overline{x}') = \emptyset.$$

Thus

$$2^n \geq \sum_{\overline{x} \in C} |B_{s_0}(\overline{x})|.$$

Also, note that for every $\overline{x}$,

$$|B_{s_0}(\overline{x})| = \sum_{j=0}^{\lfloor \frac{n\delta - 1}{2} \rfloor} \binom{n}{j}.$$

Therefore

$$2^n \geq |C| \left(\sum_{j=0}^{\lfloor \frac{n\delta - 1}{2} \rfloor} \binom{n}{j} \right),$$

and the proposition follows. □

As in the case of expander graphs, we will be less interested in a single code and more in a family of "good codes":

Definition 8.3. *A family of codes $\{C_i \subseteq \{0,1\}^{n_i}\}_{i \in \mathbb{N}}$ will be called asymptotically good codes if*

1. *The block length grows to infinity:* $\lim_i n_i = \infty$.
2. *The rates is uniformly bounded from below:* $\inf_i \mathrm{Rate}(C_i) > 0$.
3. *The relative distances are uniformly bounded from below:* $\inf_i \delta(C_i) > 0$.

Moreover, we will be interested in good linear codes with an addition extra-feature defined below.

It is not hard to see that a linear code $C \subseteq \{0,1\}^n$ is determined by a basis of C. The following terminology connects a linear codes to the following matrices.

1. A *generator matrix* for a linear code C of rate $\frac{k}{n}$ is an $n \times k$ matrix that A whose columns form a basis of C. By definition, A is a generator matrix if and only if for every $\overline{v} \in \{0,1\}^k$, $A\overline{v} \in C$.
2. The space $C^\perp$ is define as

   $$C^\perp = \{\overline{u} \in \{0,1\}^n : \forall \overline{v} \in C, \overline{u}^t \overline{v} = 0\}.$$

   Equivalently, if A is the generator matrix of C, then $C^\perp$ is the kernel of the map $\overline{u} \to \overline{u}^t A$. Thus, from linear algebra, $C^\perp$ is of rank $n - k$.
3. A *parity check* matrix of C, is a generator matrix of $C^\perp$. In other words, H is an $n \times (n-k)$ matrix whose columns form a basis for $C^\perp$. We note that H is a parity check matrix of C if and only if for every $\overline{v} \in C$, $H^t \overline{v} = \overline{0}$.

4. (non-standard name) We will call a parity check matrix H, D-sparse if the number of 1's in every row and every column is less then D.

Definition 8.4. *A family of codes $\{C_i \subseteq \{0,1\}^{n_i}\}_{i\in\mathbb{N}}$ will be called LDPC (low density parity check) codes if $n_i \to \infty$ and there is a fixed D such that for every i, the parity check matrix of C_i is D-sparse.*

In 1995, Sipser and Spielman showed how to use expander graphs in order to construct a family of asymptotically good LDPC codes which we shall now describe.

Let $G = (V, E)$ be a d-regular connected bipartite graph with sides S_0, S_1. Fix a linear code $C_0 \subseteq \{0,1\}^d$ with rate $R_0 > \frac{1}{2}$ and $\delta_0 = \delta(C_0)$. Define a code $C(G, C_0)$ with block length $|E|$ as follows

1. For every vertex $v \in V$, denote $E_v = \{e \in E : v \in E\}$ and choose a map $\pi_v : E_v \to \{1, ..., d\}$.
2. Index the space $\{0,1\}^{|E|}$ as $\{0,1\}^E$.
3. For a word $(x_e)_{e\in E} = \overline{x} \in \{0,1\}^E$, $\overline{x} \in C(G, C_0)$ if and only if for every $v \in V$, $(x_{\pi_v(e)})_{e\in E_v} \in C_0$.

Theorem 8.5. *If $\lambda_G \geq \lambda$, then $C = C(G, C_0)$ is a d-sparse code with rate $\geq 2R_0 - 1$ and relative distance $\geq \delta_0(\delta_0 - \lambda)$.*

Proof. First, consider the parity matrix of C_0: this is a $d \times (d - dR_0)$ matrix. Thus, C is defined by $|V|(d - dR_0) = 2|E|(1 - R_0)$ equations (which may be linearly dependent) and thus it is a space of dimension $\geq$

$$|E| - 2|E|(1 - R_0) = |E|(2R_0 - 1),$$

and thus the rate is $\geq 2R_0 - 1$ as needed.

Second, the parity check matrix of C_0 is of size less than $d \times (d(1 - R_0))$, thus every equation defining the parity check matrix of H is of length $\leq d$ and every edge is in at most $2(d(1 - R_0)) \leq d$ equations ($d(1 - R_0)$ for each of its vertices).

Last, let $\overline{0} \neq \overline{x} \in C$. We need to prove that $\|\overline{x}\|_1 \geq |E|\delta_0(\delta_0 - \lambda)$. Denote

$$U = \{v \in V : \exists e \in E_v, x_e = 1\},$$

and $U_0 = U \cap S_0, U_1 = U \cap S_1$. Note that for every $v \in U$,

$$|\{e \in E_v : x_e = 1\}| \geq d\delta_0.$$

Therefore,

$$\|\overline{x}\|_1 \geq d\delta_0|U_0|,$$

and

$$\|\overline{x}\|_1 \geq d\delta_0|U_1|.$$

Thus,

$$\|\overline{x}\|_1 \geq d\delta_0\sqrt{|U_0||U_1|}.$$

Also note that

$$\{e \in E : x_e = 1\} \subseteq E(U_0, U_1),$$

and thus

$$d\delta_0\sqrt{|U_0||U_1|} \leq |E(U_0,U_1)|. \tag{1}$$

Combining the previous inequalities yields that

By the expander mixing lemma (for bipartite graphs)

$$\frac{|E(U_0,U_1)|}{|E|} \leq \frac{|U_0||U_1|}{|S_0||S_1|}+\lambda\sqrt{\frac{|U_0||U_1|}{|S_0||S_1|}} =^{|E|=d|S_0|=d|S_1|} d^2\frac{|U_0||U_1|}{|E|^2}+d\lambda\frac{1}{|E|}\sqrt{|U_0||U_1|}.$$

Therefore, after multiplying by $\frac{|E|^2}{d}$ and using the inequality (1), we get

$$\delta_0|E|\sqrt{|U_0||U_1|} \leq d|U_0||U_1|+\lambda|E|\sqrt{|U_0||U_1|}.$$

Thus,

$$\delta_0(\delta_0-\lambda)|E| \leq d\delta_0\sqrt{|U_0||U_1|} \leq \|\overline{x}\|_1,$$

as needed.

□

Corollary 8.6. *Fix C_0 to be a code with block length d, rate $R_0 > \frac{1}{2}$ and $\delta_0 = \delta(C_0)$. For any family of d-regular bipartite λ-expanders $\{G_i\}_{i\in\mathbb{N}}$ with $\lambda < \delta_0$, the family of codes $C(G_i, C_0)$ is a family of asymptotically good LDPC codes.*

A decoding algorithm of an error-correcting codes is an algorithm from finding the closest codeword in C for a given (corrupt) word $\overline{y} \in \{0,1\}^n$. A brute force algorithm for decoding always exists: for a given word $\overline{y} \in \{0,1\}^n$, calculate all the distances $\|\overline{y}-\overline{x}\|_1$ over all $\overline{x} \in C$ and find $\overline{x} \in C$ with the minimal distance. Since calculating each distance requires n step, the brute force algorithm decodes a corrupt word in $n|C| = n2^{n\,\mathrm{Rate}(C)}$ steps. For the Sipser-Spielman codes defined above, Zemor gave an efficient decoding algorithm that can decode a corrupt word given that the number of corrupt bits is not too large.

Zemor decoding algorithm: given a code $C(G, C_0)$ first decode the edges using brute force by the constraints induced by the vertices of S_0 (note that every edge is connected to only one vertex in S_0, therefore there is no conflict here). Second, decode decode the edges using brute force by the constraints induced by the vertices of S_1.

Theorem 8.7. *Assume the following:*

- *The are fixed constants $\delta_0, \lambda, \varepsilon$ such that $0 < \varepsilon < 1$ and $\lambda < \frac{\delta_0}{2}\frac{\varepsilon}{2-\varepsilon}$.*
- *The graph $G = (V, E)$ is a connected d-regular bipartite finite graph with $\lambda_G \leq \lambda$.*
- *Fix C_0 to be a linear code with block length d, relative distance δ_0 and rate $R_0 > \frac{1}{2}$.*
- *Fix $C = C(G, C_0)$ as above.*

Then, for any word $\overline{y} \in \{0,1\}^{|E|}$, if there is a word $\overline{x} \in C$ such that

$$\frac{\|\overline{x}-\overline{y}\|_1}{|E|} \leq (1-\varepsilon)\frac{\delta_0}{2}(\frac{\delta_0}{2}-\lambda),$$

then applying Zemor decoding algorithm on $\overline{y}$ converges to $\overline{x}$ after applying the algorithm $O(\log|E|)$ times and thus it decodes $\overline{y}$ in $O(|E|\log|E|)$ steps (each application of the algorithm takes $O(|E|)$ steps).

Proof. For convenience assume that $\overline{y}$ is a corrupt word such that the closest code word to $\overline{y}$ is $\overline{0}$. This makes no difference in the proof, but psychologically it allows us to think of error edges as edges with the value 1 and correct edges as edges with the value 0.

Also, for convenience, let us think about S_0 as "left" and S_1 as "right" and refer to the first decoding in Zemor's algorithm as "left decoding" and to the second decoding in Zemor's algorithm as "right decoding". For $i \in \mathbb{N}$ denote by A_i the false decoded vertices in S_0 after the i-th left decoding step and by B_i the false decoded vertices after the i-th right decoding step. Repeat.

Note that for every $v \in A_1$, the number of errors in E_v was $\geq d\frac{\delta_0}{2}$ before the decoding. Thus,

$$|A_1| \leq \frac{\|\overline{x}-\overline{y}\|_1}{d\frac{\delta_0}{2}} \leq \frac{|E|(1-\varepsilon)\frac{\delta_0}{2}(\frac{\delta_0}{2}-\lambda)}{d\frac{\delta_0}{2}} = |E|\frac{(1-\varepsilon)(\frac{\delta_0}{2}-\lambda)}{d}.$$

Observe that every vertex in B_1 have at least $d\frac{\delta_0}{2}$ error edges before the right decoding step and each of these edges have a vertex in A_1. Thus,

$$d\frac{\delta_0}{2}|B_1| \leq |E(A_1,B_1)|.$$

By the bipartite expander mixing lemma,

$$|E(A_1,B_1)| \leq |E|\left(d^2\frac{|A_1||B_1|}{|E|^2}+\lambda\sqrt{d^2\frac{|A_1||B_1|}{|E|^2}}\right) = d^2\frac{|A_1||B_1|}{|E|}+\lambda d\sqrt{|A_1||B_1|}.$$

Therefore

$$d\frac{\delta_0}{2}|B_1| \leq d^2\frac{|A_1||B_1|}{|E|}+\lambda d\sqrt{|A_1||B_1|} \leq d(1-\varepsilon)(\frac{\delta_0}{2}-\lambda)|B_1|+\lambda d\sqrt{|A_1||B_1|}.$$

This yields that

$$|B_1|\frac{\varepsilon\frac{\delta_0}{2}-(1-\varepsilon)\lambda}{\lambda} \leq \sqrt{|A_1||B_1|}.$$

After squaring both sides and applying obvious algebraic manipulations yields

$$|B_1| \leq \left(\frac{\lambda}{\varepsilon\frac{\delta_0}{2}-(1-\varepsilon)\lambda}\right)^2|A_1|.$$

Denote $\alpha = \left(\frac{\lambda}{\varepsilon\frac{\delta_0}{2}-(1-\varepsilon)\lambda}\right)^2$. The assumption that $\lambda < \frac{\delta_0}{2}\frac{\varepsilon}{2-\varepsilon}$ implies that $\alpha < 1$ and repeating the same argument yields that

$$|A_2| \leq \alpha^2|A_1|.$$

By induction,

$$|A_i| \leq (\alpha^2)^{i-1}|A_1| < (\alpha^2)^{i-1}|E|.$$

Since $|A_i|$ is always an integer, it follows that for $i_0 = -\lceil \log_{\alpha^2}(|E|)\rceil + 1$, $|A_{i_0}| < 1$ and therefore $|A_{i_0}| = 0$ and there are no corrupt edges. □

9 Gabber-Galil-Margulis expanders

Here we present an explicit construction of an expander family, given by Gabber and Galil who simplified the construction of Margulis.

Define the graph $G_n = (V_n, E_n)$ (in which we allow loops and multiple edges) as follows: Denote $\mathbb{Z}_n = \mathbb{Z}/(n\mathbb{Z})$ and define

$$V_n = \mathbb{Z}_n \times \mathbb{Z}_n,$$

and each vertex (x, y) has 8 edges that contain it and is connected to

$$(x, y+1), (x, y-1), (x+1, y), (x-1, y), (x, y+x), (x, y-x), (x+y, y), (x-y, y).$$

where all operations are $\mod n$.

Theorem 9.1. *For every $n \geq 4$, $\lambda_{G_n} \leq 1 - \frac{1}{1000}$.*

Remark 9.2. *The bound in the Theorem above is far from optimal - it was chosen to keep the proof (relatively) simple.*

Proof. To avoid cumbersome notation, fix some n and denote $G_n = G, V_n = V, E_n = E$. Assume that the second largest eigenvalue of G is $1 - \varepsilon$ or equivalently, that there is a function $\phi \perp \mathbb{1}$, such that

$$\langle L\phi, \phi \rangle = \varepsilon \|\phi\|^2 = \varepsilon \left(\sum_{v \in V} 8|\phi(v)|^2 \right).$$

Below, we will show that $\varepsilon \geq \frac{1}{1000}$.

We think of G as composed of 3 graphs on V:

1. The 4-regular graph G^1 in which (x, y) is connected to

$$(x, y+1), (x, y-1), (x+1, y), (x-1, y).$$

2. The 2-regular graph G^2 in which (x, y) is connected to

$$(x, y+x), (x, y-x).$$

3. The 2-regular graph G^3 in which (x, y) is connected to

$$(x+y, y), (x-y, y).$$

We denote $L_i = I - M_{G^i}$ and note that

$$\begin{aligned}
\langle L\phi, \phi\rangle = & \\
&\sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n} \big(|\phi((x,y)) - \phi((x+1,y))|^2 + |\phi((x,y)) - \phi((x,y+1))|^2 \\
&+ |\phi((x,y)) - \phi((x,y+x))|^2 + |\phi((x,y)) - \phi((x+y,y))|^2\big) \\
&= \sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n} \big(|\phi((x,y)) - \phi((x+1,y))|^2 + |\phi((x,y)) - \phi((x,y+1))|^2\big) + \\
&\sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n} |\phi((x,y)) - \phi((x,y+x))|^2 + \\
&\sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n} |\phi((x,y)) - \phi((x+y,y))|^2 \\
&= \langle L_1\phi, \phi\rangle + \langle L_2\phi, \phi\rangle + \langle L_3\phi, \phi\rangle.
\end{aligned}$$

Thus by our assumption, for $\phi \perp \mathbb{1}$, and for $i = 1, 2, 3$,

$$\langle L_i\phi, \phi\rangle \le \varepsilon\|\phi\|^2.$$

To avoid confusion, throughout $\|.\|$ will denote the norm in the graph G and not in the subgraphs, i.e.,

$$\|\phi\|^2 = \sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n} 8|\phi(x,y)|^2.$$

For the graph G^1, we have an explicit basis of orthogonal eigenfuctions:

$$\phi_{k,j}(x,y) = e^{\frac{2\pi ikx}{n}} e^{\frac{2\pi ijy}{n}}.$$

(we will discuss how did we know this, when we get to group representations). Indeed,

$$\begin{aligned}
(M_{G^1}\phi_{k,j})(x,y) = \frac{1}{4}\Big(&e^{\frac{2\pi ik(x+1)}{n}} e^{\frac{2\pi ijy}{n}} + e^{\frac{2\pi ik(x-1)}{n}} e^{\frac{2\pi ijy}{n}} + e^{\frac{2\pi ikx}{n}} e^{\frac{2\pi ij(y+1)}{n}} \\
&+ e^{\frac{2\pi ikx}{n}} e^{\frac{2\pi ij(y-1)}{n}}\Big) \\
= \phi_{k,j}(x,y)&\left(\frac{1}{2}\cos(\frac{2\pi k}{n}) + \frac{1}{2}\cos(\frac{2\pi j}{n})\right).
\end{aligned}$$

(checking orthogonality is straight-forward and left to the reader).

Note that

$$\|\phi_{k,j}\|^2 = \sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n} 8|\phi_{k,j}(x,y)|^2 = 8n^2,$$

and that $\phi_{0,0} = \mathbb{1}$. Thus if $\phi \perp \mathbb{1}$, then there are constants $a_{k,j} \in \mathbb{C}$ such that

$$\phi = \sum_{(k,j)\in\mathbb{Z}_n\times\mathbb{Z}_n\setminus\{(0,0)\}} a_{k,j}\phi_{k,j},$$

and

$$\|\phi\|^2 = 8n^2 \sum_{(k,j)\in\mathbb{Z}_n\times\mathbb{Z}_n\setminus\{(0,0)\}} |a_{k,j}|^2.$$

Without loss of generality, we can assume that $\sum_{(k,j)\in\mathbb{Z}_n\times\mathbb{Z}_n\setminus\{(0,0)\}}|a_{k,j}|^2=1$.

Our proof strategy: We define a measure μ on the space $X=\mathbb{Z}_n\times\mathbb{Z}_n\setminus\{(0,0)\}$ as follows: for every $A\subseteq X$, define

$$\mu(A)=\sum_{(k,j)\in A}|a_{k,j}|^2.$$

By our normalization, μ is a probability measure. Then, we partition X into 5 disjoint subsets $A_1,...,A_5$ and show that for each such subset

$$\mu(A_i)\leq \text{Bound that tends to 0 as } \varepsilon\to 0.$$

Thus, we will get that

$$1=\mu(X)=\sum_{i=1}^{5}\sum_{(k,j)\in A_i}\mu(A_i)\leq \text{sum of our bounds},$$

and deduce that ε cannot be too small.

Step 1 - actions on sets in X: Observe that for every $(k,j)\in\mathbb{Z}_n\times\mathbb{Z}_n$ and for every (x,y),

$$\phi_{k,j}(x,y+x)=e^{\frac{2\pi ikx}{n}}e^{\frac{2\pi j(y+x)}{n}}=e^{\frac{2\pi i(k+j)x}{n}}e^{\frac{2\pi jy}{n}}=\phi_{k+j,j}(x,y)$$

By this

$$\begin{aligned}
\varepsilon(8n^2)&=\varepsilon\|\phi\|^2\\
&\geq\langle L_2\phi,\phi\rangle\\
&=\sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n}|\phi(x,y)-\phi(x,y+x)|^2\\
&=\sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n}|\sum_{(k,j)\in X}a_{k,j}\phi_{k,j}(x,y)-\sum_{(k,j)\in X}a_{k,j}\phi_{k,j}(x,y+x)|^2\\
&=\sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n}|\sum_{(k,j)\in X}a_{k,j}\phi_{k,j}(x,y)-\sum_{(k,j)\in X}a_{k,j}\phi_{k+j,j}(x,y)|^2\\
&=\sum_{(x,y)\in\mathbb{Z}_n\times\mathbb{Z}_n}|\sum_{(k,j)\in X}(a_{k,j}-a_{k-j,j})\phi_{k,j}(x,y)|^2\\
&=\frac{1}{8}\|\sum_{(k,j)\in X}(a_{k,j}-a_{k-j,j})\phi_{k,j}(x,y)\|^2\\
&=\frac{1}{8}8n^2\sum_{(k,j)\in X}|a_{k,j}-a_{k-j,j}|^2.
\end{aligned}$$

Thus,

$$2\sqrt{2\varepsilon}\geq\sqrt{\sum_{(k,j)\in X}|a_{k,j}-a_{k-j,j}|^2},$$

or equivalently,

$$2\sqrt{2\varepsilon}\geq\sqrt{\sum_{(k,j)\in X}|a_{k,j}-a_{k+j,j}|^2}.$$

Define $T_1 : X \to X$ by $T_1((k,j)) = (k+j, j)$. Then for every $A \subseteq X$,

$$\begin{aligned}\sqrt{\mu(A)} &= \sqrt{\sum_{(k,j)\in A} |a_{k,j}|^2} \\ &\leq^{\text{triangle ineq.}} \sqrt{\sum_{(k,j)\in A} |a_{k+j,j}|^2} + \sqrt{\sum_{(k,j)\in A} |a_{k,j} - a_{k+j,j}|^2} \\ &\leq \sqrt{\mu(T_1(A))} + 2\sqrt{2\varepsilon}.\end{aligned}$$

After squaring both sides,

$$\mu(A) \leq \mu(T_1(A)) + 4\sqrt{2\varepsilon} + 8\varepsilon.$$

By similar considerations,

$$\mu(A) \leq \mu(T_1^{-1}(A)) + 4\sqrt{2\varepsilon} + 8\varepsilon.$$

Using the inequality regarding L_3 and the fact that

$$\phi_{k,j}(x+y, y) = \phi_{k,j+k}(x,y),$$

we deduce a similar result: define $T_2 : X \to X$:

$$T_2((k,j)) = (k, j+k),$$

then for every $A \subseteq X$,

$$\mu(A) \leq \mu(T_2(A)) + 4\sqrt{2\varepsilon} + 8\varepsilon,$$

$$\mu(A) \leq \mu(T_2^{-1}(A)) + 4\sqrt{2\varepsilon} + 8\varepsilon.$$

Step 2 - bound on the set A_1: By our computation above, for every $(k,j) \in \mathbb{Z}_n \times \mathbb{Z}_n$,

$$L_1\phi_{k,j} = \left(\frac{1}{2}(1-\cos(\frac{2\pi k}{n})) + \frac{1}{2}(1-\cos(\frac{2\pi j}{n}))\right)\phi_{k,j}.$$

Denote

$$\kappa_{k,j} = \left(\frac{1}{2}(1-\cos(\frac{2\pi k}{n})) + \frac{1}{2}(1-\cos(\frac{2\pi j}{n}))\right),$$

and by this notation, $L_1\phi_{k,j} = \kappa_{k,j}\phi_{k,j}$. Denote

$$A_1 = \{(k,j) \in \mathbb{Z}_n \times \mathbb{Z}_n : \frac{1}{4}n \leq k \leq \frac{3}{4}n \text{ or } \frac{1}{4}n \leq j \leq \frac{3}{4}n\},$$

and observe that for every $(k,j) \in A_1$, $\kappa_{k,j} \geq \frac{1}{2}$.

Thus, by our assumptions

$$\varepsilon\left(8n^2\right) = \varepsilon\|\phi\|^2 \geq \langle L_1\phi, \phi\rangle = \sum_{(k,j)\in\mathbb{Z}_n\times\mathbb{Z}_n\setminus\{(0,0)\}} \kappa_{k,j}|a_{k,j}|^2\|\phi_{k,j}\|^2 =$$

$$8n^2 \sum_{(k,j)\in\mathbb{Z}_n\times\mathbb{Z}_n\setminus\{(0,0)\}} \kappa_{k,j}|a_{k,j}|^2 \geq 8n^2 \sum_{(k,j)\in A_1} \frac{1}{2}|a_{k,j}|^2.$$

As a result,

$$2\varepsilon \geq \sum_{(k,j)\in A_1} |a_{k,j}|^2 = \mu(A_1).$$

Step 3 - bounds on the sets $A_2, ..., A_5$: In order to define the sets $A_2, ..., A_5$, we will think about $\mathbb{Z}_n$ as numbers between $-\frac{n}{2}$ and $\frac{n}{2}$ (mod n) with the usual absolute value. With this notation,

$$A_1 = \{(k,j) \in X : \frac{n}{4} \leq |k| \text{ or } \frac{n}{4} \leq |j|\}.$$

Define

$$A_2 = \{(k,j) \in X : |j| \leq |k| < \frac{n}{4} \text{ and } kj > 0\},$$

$$A_3 = \{(k,j) \in X : |k| < |j| < \frac{n}{4} \text{ and } kj \geq 0\},$$

$$A_4 = \{(k,j) \in X : |k| \leq |j| < \frac{n}{4} \text{ and } kj < 0\},$$

$$A_5 = \{(k,j) \in X : |j| < |k| < \frac{n}{4} \text{ and } kj \leq 0\}.$$

Observe that

$$T_1(A_2 \cup A_3) \subseteq A_1 \cup A_2 :$$

by the assumption that $kj \geq 0$ in both A_2, A_3, we have that

$$|k + j| = |k| + |j|.$$

Therefore for every $(k,j) \in A_2 \cup A_3$, $T_1(k,j) = (k+j, j)$ and $|k+j| \geq |j|$. Therefore if $|k+j| < \frac{n}{4}$, then $T_1(k,j) \in A_2$ and if $|k+j| \geq \frac{n}{4}$, then $T_1(k,j) \in A_1$. Note that by the previous steps:

$$\begin{aligned}
\mu(A_2) + \mu(A_3) &= \mu(A_2 \cup A_3) \\
&\leq \mu(T_1(A_2 \cup A_3)) + 4\sqrt{2\varepsilon} + 8\varepsilon \\
&\leq \mu(A_1) + \mu(A_2) + 4\sqrt{2\varepsilon} + 8\varepsilon \\
&\leq \mu(A_2) + 4\sqrt{2\varepsilon} + 10\varepsilon
\end{aligned}$$

As a result,

$$\mu(A_3) \leq 4\sqrt{2\varepsilon} + 10\varepsilon.$$

By similar reasoning:

$$T_2(A_2 \cup A_3) \subseteq A_1 \cup A_3,$$

$$T_1^{-1}(A_4 \cup A_5) \subseteq A_1 \cup A_4,$$

$$T_2^{-1}(A_4 \cup A_5) \subseteq A_1 \cup A_5.$$

Therefore, for every $i = 2, ..., 5$,

$$\mu(A_i) \leq 4\sqrt{2\varepsilon} + 10\varepsilon.$$

After summing,

$$1 = \mu(A_1) + ... + \mu(A_5) \leq 42\varepsilon + 16\sqrt{2\varepsilon},$$

Thus, $\varepsilon \geq \frac{1}{1000}$ as needed. □

10 Exercises

1. (Power of a graph) Let $G = (V, E)$ be a connected d-regular graph. Define $G^k = (V, E_k)$ to be the graph in which we allow multiple edges and loops such that the number of edges between $u, v \in V$ is the number of paths of length k between u and v. Prove that:

   (a) The graph G^k is d^k-regular.

   (b) If $\lambda_G \leq \lambda$ then for every odd k, $\lambda_{G_k} \leq \lambda^k$.

   (c) If G is bipartite, then for every even k, G^k has two connected components and the spectrum of the random walk in each of these components is contained in $[0, \lambda^k] \cup \{1\}$.

   (d) If λ is an eigenvalue of G with multiplicity n, then λ^k is an eigenvalue of G^k with multiplicity n.

2. (Bipartite cover of a graph) Let $G = (V, E)$ be a graph. Define the bipratite cover G' of G as follows:

   - The vertices of G' are two copies of V, denoted S_0, S_1. Formally, $|S_0| = |S_1| = |V|$ and there are bijection maps $\pi_i : S_i \to V$).
   - The edges of G' are $\{\{u, v\} : u \in S_0, v \in S_1, \{\pi_0(u), \pi_1(v)\} \in E\}$, i.e., $u \in S_0, v \in S_1$ are connected by an edge if and only if u and v are connected in G.

   (a) Prove that $\lambda \neq 0$ is an eigenvalue of M_G with multiplicity k if and only if $\pm\lambda$ is an eigenvalue of $M_{G'}$ with multiplicity k. Prove that 0 is an eigenvalue of M_G with multiplicity k if and only if 0 is an eigenvalue of $M_{G'}$ with multiplicity $2k$.

   (b) Show that using the bipartite cover, one can deduce the expander mixing lemma in the non-bipartite case from the bipartite case.

3. (Tensor product of graphs) Let $G_1 = (V_1, E_1), G_2 = (V_2, E_1)$ be finite connected graphs. Define $G_1 \otimes G_2$ as follows: $V_{G_1 \otimes G_2} = V_1 \times V_2$, $E_{G_1 \otimes G_2} = \{\{(v_1, v_2), (u_1, u_2)\} : \{v_1, u_1\} \in E_1, \{v_2, u_2\} \in E_2\}$. Prove that if λ_i is an eigenvalue of M_{G_i}, then $\lambda_1 \lambda_2$ is an eigenvalue of $M_{G_1 \otimes G_2}$ and all the eigenvalues of $M_{G_1 \otimes G_2}$ are of this form (Hint: for $\phi_i \in \ell^2(V_i)$ eigenfunction of M_{G_i} with eigenvalue λ_i, consider $\phi((v_1, v_2)) = \phi_1(v_1)\phi_2(v_2)$).

4. (Kroncker sum of graphs) Let $G_1 = (V_1, E_1), G_2 = (V_2, E_1)$ be finite connected graphs such that G_i is d_i-regular for $i = 1, 2$. Define $G_1 \oplus G_2$ as follows: $V_{G_1 \oplus G_2} = V_1 \times V_2$, $E_{G_1 \oplus G_2} = \{\{(v_1, v_2), (u_1, u_2)\} : v_1 = u_1 \text{ and } \{v_2, u_2\} \in E_2 \text{ or } v_2 = u_2 \text{ and } \{v_1, u_1\} \in E_1\}$. Prove that if λ_i is an eigenvalue of M_{G_i} for $i = 1, 2$, then $\frac{d_1}{d_1+d_2}\lambda_1 + \frac{d_2}{d_1+d_2}\lambda_2$ is an eigenvalue of $M_{G_1 \oplus G_2}$ and all the eigenvalues of $M_{G_1 \oplus G_2}$ are of this form (Hint: use the hint regarding the Tensor product).

5. (Line graph of a graph) Let $G = (V, E)$ be a finite connected d-regular graph, where $d > 2$. Define the Line graph of G, denoted Line(G) as follows: the vertices of Line(G) are the edges of G or formally, $V_{\text{Line}} = \{u_e : e \in E\}$ and two vertices of Line(G) are connected if the edges in G share a vertex, or formally, $E_{\text{Line}} = \{\{u_e, u_{e'}\} : e, e' \in E, e \cap e' \neq \emptyset\}$.

The aim of the following exercise is to prove that the eigenvalues of $M_{\text{Line}(G)}$ can be computed if one knows the eigenvalues of M_G.

Define the incidence matrix of G, denoted by H, to be $|V|\times|E|$ matrix indexed by $V \times E$ as follows:

$$H(v,e) = \begin{cases} 1 & v \in e \\ 0 & v \notin e \end{cases}.$$

(a) Prove that $HH^t = dI_{|V|} + A_G$, where $I_{|V|}$ is the $|V|\times|V|$ identity matrix. Also prove that $H^tH = 2I_{|E|} + A_{\text{Line}(G)}$.

(b) Prove that for any real matrix H and any $\lambda \neq 0$, λ is an eigenvalue of H^tH of multiplicity k if and only if λ is an eigenvalue of HH^t of multiplicity k (this is a general claim regarding matrices).

(c) Observe that $\text{Line}(G)$ is $(2d-2)$-regular (and recall that G is d-regular) and therefore $\frac{1}{d}HH^t = I_{|V|} + M_G$ and $\frac{1}{2d-2}H^tH = \frac{2}{2d-2}I_{|E|} + M_{\text{Line}(G)}$.

(d) Find the connection between the spectrum of M_G and $M_{\text{Line}(G)}$.

(e) Show that $M_{\text{Line}(G)}$ always have the eigenvalue $-\frac{1}{d-1}$ and calculate the multiplicity of this eigenvalue as a function of $|V|$ (remark: note that you have to distinguish between the case where G is bipartite and the case where G is not bipartite).

6. (The incidence graph of a finite projective plane) Fix $p \in \mathbb{N}, p \geq 2$ to be a prime and denote $\mathbb{F}_p = \{0, ..., p-1\}$ to be the field with p elements (both addition and multiplication are taken $\mod p$). We use the following conventions in $\mathbb{F}_p^3$: a vector $\underline{x} \in \mathbb{F}_p^3$ is considered to be a column vector. For two vectors $\underline{x}, \underline{y} \in \mathbb{F}_p^3$, $\underline{y}^t\underline{x}$ is the usual inner-product taken modulo p.

Define the following bipartite graph G with sides S_0, S_1:

- The vertices in S_0 are the linear spans of non-zero column vectors in $\mathbb{F}_p^3$: for $\underline{x} \in \mathbb{F}_p^3 \setminus \{(0,0,0)\}$ we take

$$[\underline{x}] = \{k\underline{x} : k \in \mathbb{F}_p\}$$

and define

$$S_0 = \{[\underline{x}] : \underline{x} \in \mathbb{F}_p^3 \setminus \{(0,0,0)\}\},$$

for convenience we will call this vertices "point vertices" (or simply "points").

- The vertices in S_1 are the linear spans of non-zero column vectors in $\mathbb{F}_p^3$:

$$S_1 = \{[\underline{y}^t] : \underline{y} \in \mathbb{F}_p^3 \setminus \{(0,0,0)\}\}.$$

for convenience we will call this vertices "line vertices" (or simply "lines").

- A point vertex $[\underline{x}]$ and a line vertex $[\underline{y}^t]$ are connected by an edge if

$$\underline{y}^t\underline{x} = 0.$$

(The motivation behind all these definitions is the point-line geometry where S_0 are the points, S_1 are the lines and a line passes through a point if the corresponding vertices are connected by an edge. By the exercise below, this geometry do not have parallel (i.e., not intersecting) lines and thus it is a geometry of a projective plane).

(a) Show that the graph is well-defined, i.e., that for $[\underline{x_1}] = [\underline{x_2}]$ and $[\underline{y_1^t}] = [\underline{y_2^t}]$, $\underline{y_1}^t \underline{x_1} = 0$ if and only if $\underline{y_2}^t \underline{x_2} = 0$.

(b) Show that $|S_0| = |S_1| = p^2 + p + 1$.

(c) Show that G is $(p+1)$-regular.

(d) Show that for every two (different) point vertices there is a unique line vertex that is connected by an edge to both of them. Similarly, show that for every two (different) line vertices there is a unique point vertex that is connected by an edge to both of them.

(e) Calculate the spectrum of M_G (Hint: consider G^2 and use the relevant result regarding power of a bipartite graph).

Part II: Expansion in groups

11 Locally finite graphs

A graph G is called *locally finite* if every vertex has a finite degree. From now on, when we talk about graphs, we mean locally finite connected graphs that can be infinite.

12 Groups - basic concepts

This section is aimed to review some basic concepts regarding groups.

12.1 Basic definitions

Definition 12.1 (Generating set, Finite generation)**.** *For a group Γ a set $S \subseteq \Gamma$ is a generating set if $\Gamma = \bigcup_{k\in\mathbb{N}} S^k$, where $S^k = \{s_1...s_k \in \Gamma : s_1, ..., s_k \in S\}$. In other words, every element in Γ can be written as a finite product of elements in S.*

A generating set S is called symmetric if that $S = S^{-1}$.

A group Γ is called finitely generate if it has a finite generating set.

Throughout, Γ will denote a countable (i.e., either finite or infinite denumerable) finitely generated group.

Remark 12.2. *Not every countable group is finitely generated, e.g., $\bigoplus_{n\in\mathbb{N}} \mathbb{Z}$ is countable but not finitely generated.*

Definition 12.3 (Cosets)**.** *Given a group Γ and a subgroup $H < \Gamma$, a left-coset is a set of the form*

$$gH = \{gh : h \in H\},$$

where $g \in \Gamma$. Similarly, a right-coset is a set of the form

$$Hg = \{hg : h \in H\},$$

where $g \in \Gamma$.

The index of H, denoted $[G : H]$, is the number of left (or right) cosets.

Remark 12.4. *Note that there is a natural bijection between left and right cosets: $gH \to Hg^{-1}$ and $g_1H = g_2H$ if and only if $Hg_1^{-1} = Hg_2^{-1}$.*

We recall that a subgroup $N < \Gamma$ is called normal and we denote $N \triangleleft \Gamma$ if for every $g \in \Gamma$, $gN = Ng$ or equivalently, if for every $g \in \Gamma$, $g^{-1}Ng = N$.

We also recall that for a normal subgroup $N \triangleleft \Gamma$, the cosets of N form a group with the multiplication defined as

$$(g_1N) \cdot (g_2N) = g_1g_2N,$$

and this group is denoted by Γ/N and called the quotient of Γ by N.

12.2 Group action on sets

Definition 12.5. *Given a group Γ and a set X, an action of Γ on X is a group homomorphism $\rho : \Gamma \to \mathrm{Aut}(X)$, where $\mathrm{Aut}(X)$ is a group of bijections from X to itself. In other words, for every $g \in \Gamma$, $\rho(g) : X \to X$ is a bijection and for every $g_1, g_2 \in \Gamma$, $\rho(g_1 g_2) = \rho(g_1)\rho(g_2)$.*

Sometimes, ρ is implicit: we denote $\Gamma \curvearrowright X$ and for $x \in X$, $g \in \Gamma$, $g.x$ is the image of x under (the implicit) $\rho(g)$.

Examples:

1. $\Gamma = \mathbb{Z}$ acts on $X = \mathbb{Z}$ by translation: $a.b = a + b$.
2. More generally, every Γ acts on itself $X = \Gamma$ in three natural ways:
   (a) Left-action: $g.g' = gg'$.
   (b) Right-action: $g.g' = g'g^{-1}$. Note that the inverse is necessary -
   $$(g_1 g_2).g' = g'(g_1 g_2)^{-1} = g' g_2^{-1} g_1^{-1} = g_1.(g' g_2^{-1}) = g_1.(g_2.g').$$
   (c) Conjugation: $g.g' = gg'g^{-1}$.
3. $\Gamma = \mathbb{Z}$ acts on $X = \mathbb{R}$ by translation: $a.b = a + b$.
4. More generally, for every Γ if $\Gamma < \Omega$ then it acts on it in the three ways defined above.
5. $\Gamma = \mathbb{Z}$ acts on $X = \mathbb{Z}/N\mathbb{Z}$ by translation (with addition $\mod N$).
6. More generally, for every Γ and every $N \triangleleft \Gamma$, Γ acts on Γ/N from the left.

Terminology: let $\Gamma \curvearrowright X$,

1. For $x_0 \in X$, the *stabilizer of* x_0, denoted $\mathrm{Stab}(x_0)$, is
   $$\mathrm{Stab}(x_0) = \{g \in \Gamma : g.x_0 = x_0\}.$$
   Note that this is a subgroup of Γ.
2. The action of Γ on X is called *free* if for every $x \in X$, $\mathrm{Stab}(x) = \{e\}$.
3. For $x_0 \in X$, the *orbit of* x_0 is the set
   $$\mathrm{Orbit}(x_0) = \{g.x_0 : g \in \Gamma\}.$$
4. Note that for every $x, y \in X$, either $\mathrm{Orbit}(x) = \mathrm{Orbit}(y)$ or $\mathrm{Orbit}(x) \cap \mathrm{Orbit}(y) = \emptyset$. Thus, we define an equivalence relation $x \sim_\Gamma y$ if $\mathrm{Orbit}(x) = \mathrm{Orbit}(y)$. Define $\Gamma\backslash X$ to be the equivalence classes of $\sim_\Gamma$.
5. A *fundamental domain* of the action is a subset $D \subseteq X$ such that $X = \bigcup_{x \in D} \mathrm{Orbit}(x)$ and for every $x, y \in D$, if $x \neq y$, then $\mathrm{Orbit}(x) \cap \mathrm{Orbit}(y) = \emptyset$.
6. The action of Γ on X is called *transitive* if for every $x \in X$, $\mathrm{Orbit}(x) = X$. In other words, the action is transitive if every $\{x\} \subseteq X$ is a fundamental domain.

Exercise 12.6. *Prove that $\sim_\Gamma$ is an equivalence relation.*

Usually, one studies an action of a group on X that has some structure such that the action preserves this structure. We will be interested in two types of group actions - action on a graph and action on a vector space (action on a vector space is called representation).

12.3 Group action on a graph

Definition 12.7 (Group action on a graph)**.** *Given a group Γ and a graph $G = (V, E)$, an action of Γ on G is an action of Γ on the set V such that for every $g \in \Gamma$ and every $\{u, v\} \in E$, we have that $\{g.u, g.v\} \in E$. In other words, if we denote $\mathrm{Aut}(G)$ to be the group of symmetries of G, the action of Γ on G is a homomorphism of Γ into $\mathrm{Aut}(G)$.*

Examples:

1. A finite group Γ, with $|\Gamma| = n$ always acts on the complete graph on n vertices: label the vertices by elements of Γ and act by the left (or right) group action on itself. This action is transitive and free. This is just another way of stating the known fact that a finite group with n elements is isomorphic to a a subgroup of the permutation group.

2. The ring graph on n vertices is the graph on $\mathbb{Z}/n\mathbb{Z}$, where $E = \{\{i, i+1\} : i \in \mathbb{Z}/n\mathbb{Z}\}$ and addition is $\mod n$. The group $\Gamma = \mathbb{Z}$ acts on the ring graph of n vertices $a.i = a + i$ (this action is transitive, but not free). The group of $\Gamma = \mathbb{Z}/2\mathbb{Z}$ on the ring graph of n vertices by reflection: $1.i = n - i \mod n$ (for $n > 1$ this action is neither free nor transitive).

Proposition 12.8. *Let Γ be a group acting transitively on a locally finite, connected graph G. Fix $v_0 \in V$ and let $\{u_1, ..., u_d\} \in V$ be the neighbors of v_0. For every $1 \le i \le d$, choose $s_i \in \Gamma$ such that $s_i.v_0 = u_i$ and denote $S_1 = \{s_1, ..., s_d\}$ and $H = \mathrm{Stab}(v_0) < \Gamma$. Then the set $S = S_1 \cup H$ is a generating set of Γ.*

Proof. Let $g \in \Gamma$ and denote $k = \mathrm{dist}(v_0, g.v_0)$. We will show by induction on k, that $g \in S_1^k H$.

If $k = 0$, then $g.v_0 = v_0$ and by definition $g \in \mathrm{Stab}(v_0) = H$.

Next, we assume that the assertion is true for k and prove it for $k + 1$. Let $v_0, v_1, ..., v_{k+1} = g.v_0$ be a path between v_0 and $g.v_0$. Note that v_1 is a neighbor of v_0 and therefore there is some $s \in S_1$ such that $s.v_0 = v_1$. Then $s^{-1}.v_1 = v_0$ and $s^{-1}.v_1, ..., (s^{-1}g).v_0$ is a path of length k from v_0 to $(s^{-1}g).v_0$. By the induction assumption, $s^{-1}g \in S_1^k H$ and therefore $g \in sS_1^k H \subseteq S_1^{k+1} H$ as needed. □

Exercise 12.9. *Let Γ be a group acting on a locally finite, connected graph G. Show that:*

1. *If $\Gamma\backslash G$ is finite, then we can choose a fundamental domain of the action that is a connected subgraph of G.*

2. *If the action is transitive, then graph G is regular.*

3. *Assume that the action is transitive and free. Then for every $v_0 \in V$, if $\{u_1, ..., u_d\} \in V$ are the neighbors of v_0 then there is a symmetric set $S_1 = \{s_1, ..., s_d\} \in \Gamma$ such that $s_i.v_0 = u_i$.*

Exercise 12.10. *Show that Proposition 12.8 can be generalized as follows: let Γ be a group acting on a locally finite, connected graph G such that $\Gamma\backslash G$ is finite. Choose a fundamental domain D and recall that $\partial_{out}D$ as the set of vertices in $V \setminus D$ that have at least one neighbor in D. For every $u \in \partial_{out}D$ choose $s_u \in \Gamma$ such that $u \in s_u.D$ and denote $S_1 = \{s_u : u \in \partial_{out}D\}$. Prove that for every $v \in D$, $S = S_1 \cup \mathrm{Stab}(v)$ is a generating set for Γ.*

Definition 12.11. *Let Γ be a group acting a locally finite, connected graph G. The quotient graph $\Gamma\backslash G$ is a graph with the vertex set $\Gamma\backslash V$ and the edge set $\Gamma\backslash E$.*

Remark 12.12. *Note that $\Gamma\backslash G$ may have loops and multiple edges even if G did not.*

12.4 Cayley and Schreier graphs

Definition 12.13 (Cayley graph)**.** *Let Γ be a finitely generated group with a finite, symmetric generating set S. The Cayley graph $\mathrm{Cay}(\Gamma; S)$ is defined as follows: $V = \{g : g \in \Gamma\}$, $E = \{\{g, gs\} : g \in \Gamma, s \in S\}$.*

There is a natural action of Γ on $\mathrm{Cay}(\Gamma; S)$ is defined by the left-action of Γ on itself, i.e., $g'.g = g'g$. We note that this is indeed an action on the graph, since for every edge $\{g, gs\}$, $g'.\{g, gs\} = \{g'g, g'gs\}$ is also an edge.

Examples:

1. For every finite group Γ, we can take $S = \Gamma\setminus\{e\}$ and it that case $\mathrm{Cay}(\Gamma; S)$ is the complete graph on $|\Gamma|$ vertices.
2. For $\Gamma = \mathbb{Z}/n\mathbb{Z}$ and $S = \{\pm 1\}$, $\mathrm{Cay}(\Gamma; S)$ is the ring graph on n vertices.
3. For $\Gamma = \mathbb{Z}$ and $S = \{\pm 1\}$, $\mathrm{Cay}(\Gamma; S)$ is the infinite 2-regular graph.

Definition 12.14 (Schreier graph)**.** *Let Γ be a finitely generated group with a finite, symmetric generating set S, and $H < \Gamma$. The Schreier graph, $\mathrm{Sch}(\Gamma, H; S)$ is $H\backslash\mathrm{Cay}(\Gamma; S)$ or more explicitly, the graph with vertices $V = \{Hg : g \in \Gamma\}$ and edges $E = \{\{Hg, Hgs\} : g \in \Gamma, s \in S\}$.*

Remark 12.15. *Note that there is generally no action of Γ on $\mathrm{Sch}(\Gamma, H; S)$ and $\mathrm{Sch}(\Gamma, H; S)$ does not have to be symmetric. Indeed, a Theorem of Gross states that every d-regular finite connected graph with d even is a Schreier graph of a permutation group.*

Observation 12.16. *If $N \triangleleft \Gamma$, then $\mathrm{Sch}(\Gamma, N; S)$ is (isomorphic to) $\mathrm{Cay}(N\backslash\Gamma; N\backslash S)$.*

13 Representations of finite groups - a crash course

Throughout U is a finite dimensional vector space over $\mathbb{C}$. Denote $\mathrm{GL}(U)$ to be the linear invertible maps $T : U \to U$.

Definition 13.1 (Representation of a group)**.** *Let Γ be a finite group. A representation of (π, U) of Γ is a homomorphism $\pi : \Gamma \to \mathrm{GL}(U)$. The dimension of U is called the degree of the representation.*

Example 13.2. *1. For every group Γ and every space U, there is the trivial representation $\pi(g) = I$ for every $g \in \Gamma$, where $I \in \mathrm{GL}(U)$ is the identity operator.*

2. For $\Gamma = \mathbb{Z}/n\mathbb{Z}$, define the representation $\pi_j \in \mathrm{GL}(\mathbb{C})$ for $0 \leq j \leq n-1$ by

$$\pi_j(m) = (e^{\frac{2\pi ijm}{n}}),$$

where $(e^{\frac{2\pi ijm}{n}})$ should be thought of as a matrix, i.e.,

$$(e^{\frac{2\pi ijm}{n}}).x = e^{\frac{2\pi ijm}{n}}x.$$

(verify that for every j this is indeed a representation).

Definition 13.3 (Equivalence of representations)**.** *Let Γ be a finite group and $(\pi_1, U_1), (\pi_2, U_2)$ be representations of Γ. Two representations $(\pi_1, U_1), (\pi_2, U_2)$ are called equivalent if there is an invertible linear map $T : U_1 \to U_2$ such that $\pi_2 \circ T = T \circ \pi_1$, i.e., for every $g \in \Gamma$,*

$$\pi_2(g)T = T\pi_1(g).$$

Example 13.4. *Let $\Gamma = \mathbb{Z}/n\mathbb{Z}$ and $U = \mathbb{C}^2$. Define the following representations:*

$$\pi_1(m) = \begin{pmatrix} \cos(\frac{2\pi m}{n}) & -\sin(\frac{2\pi m}{n}) \\ \sin(\frac{2\pi m}{n}) & \cos(\frac{2\pi m}{n}) \end{pmatrix}$$

and

$$\pi_2(m) = \begin{pmatrix} e^{\frac{2\pi im}{n}} & 0 \\ 0 & e^{-\frac{2\pi im}{n}} \end{pmatrix}.$$

(verify that those are in fact representations). Then π_1 and π_2 are equivalent: take

$$A = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}.$$

Then

$$A^{-1} = \frac{1}{2i}\begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix},$$

and

$$\begin{aligned} A^{-1}\pi_1(m)A &= \frac{1}{2i}\begin{pmatrix} 1 & i \\ -1 & i \end{pmatrix}\begin{pmatrix} \cos(\frac{2\pi m}{n}) & -\sin(\frac{2\pi m}{n}) \\ \sin(\frac{2\pi m}{n}) & \cos(\frac{2\pi m}{n}) \end{pmatrix}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{2i}\begin{pmatrix} e^{\frac{2\pi im}{n}} & ie^{\frac{2\pi im}{n}} \\ e^{-\frac{2\pi im}{n}} & -ie^{-\frac{2\pi im}{n}} \end{pmatrix}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{\frac{2\pi im}{n}} & 0 \\ 0 & e^{-\frac{2\pi im}{n}} \end{pmatrix} \\ &= \pi_2(m). \end{aligned}$$

Definition 13.5 (Direct sum of representations)**.** *Let Γ be a finite group and let $(\pi_1, U_1), (\pi_2, U_2)$ be representations of Γ. We denote $\pi_1 \oplus \pi_2 : \Gamma \to \mathrm{GL}(U_1 \oplus U_2)$ to be the direct sum of the representations defined as*

$$(\pi_1 \oplus \pi_2)(g).(x_1, x_2) = (\pi_1(g).x_1, \pi_2(g).x_2).$$

Example 13.6. *For* $\Gamma = \mathbb{Z}/n\mathbb{Z}$, $U_1 = U_2 = \mathbb{C}$,

$$\pi_1(m) = (e^{\frac{2\pi i m}{n}}),$$

$$\pi_2(m) = (e^{-\frac{2\pi i m}{n}}),$$

$\pi_1 \oplus \pi_2$ *is a representation on* $\mathbb{C}^2$ *by*

$$(\pi_1 \oplus \pi_2)(m) = \begin{pmatrix} e^{\frac{2\pi i m}{n}} & 0 \\ 0 & e^{-\frac{2\pi i m}{n}} \end{pmatrix}.$$

Definition 13.7 (Γ-invariant subspace)**.** *Let* Γ *be a finite group and* (π, U) *be a representation of* Γ*. Subspace* $U' \subseteq U$ *is called* Γ*-invariant, if for every* $x \in U'$ *and every* $g \in \Gamma$, $\pi(g).x \in U'$.

Note that if U' *is a* Γ*-invariant space, then we can define a representation* $(\pi|_{U'}, U')$ *of* Γ *by restricting* π *to* U' *and such a representation is called a sub-representation* π.

Example 13.8. *For every* $(\pi_1 \oplus \pi_2, U_1 \oplus U_2)$, $U_1 \oplus \{0\}$ *and* $\{0\} \oplus U_2$ *are both* Γ*-invariant subspaces.*

Definition 13.9 (Irreducible representation)**.** *Let* Γ *be a finite group and* (π, U) *be a representation of* Γ*. The representation* π *is called irreducible if the only* Γ*-invariant subspaces of* U *are* $\{0\}$ *and* U.

Example 13.10. *Every* 1*-dimensional representation is irreducible.*

Definition 13.11 (Decomposable representation)**.** *Let* Γ *be a finite group and* (π, U) *be a representation of* Γ*. The representation* π *is called decomposable if there are two non-zero* Γ*-invariant subspaces* $U_1, U_2 \subseteq U$ *such that* $U = U_1 \oplus U_2$.

Definition 13.12 (Completely reducible representation)**.** *Let* Γ *be a finite group and* (π, U) *be a representation of* Γ*. The representation* π *is called Completely reducible if there are non-zero subspaces* $U_1, ..., U_r \subseteq U$ *such that* $\pi|_{U_j}$ *is irreducible and* $U = U_1 \oplus ... \oplus U_r$.

Lemma 13.13. *Let* Γ *be a finite group and* $(\pi_1, U_1), (\pi_2, U_1)$ *be equivalent representation of* Γ*. Then* π_1 *is irreducible/decomposable/completely reducible if and only if* π_1 *is irreducible/decomposable/completely reducible.*

Proof. We will only prove the lemma for decomposable representations - the other two proofs are similar.

Assume that π_1 is decomposable and let $T : U_1 \to U_2$ invertible such that $\pi_2 T = T\pi_1$. By our assumption, there are non-zero Γ-invariant $U_1', U_1'' \subseteq U_1$ such that $U_1 = U_1' \oplus U_1''$. Define $U_2' = TU_1', U_2'' = TU_1''$. We claim that U_2', U_2'' are non-zero Γ-invariant subspaces and that $U_2 = U_2' \oplus U_2''$. The fact that if $U_1 = U_1' \oplus U_1''$, U_1', U_1'' are non-zero and $T : U_1 \to U_2$ is invertible, then $U_2 = U_2' \oplus U_2''$ and U_2', U_2'' are non-zero is an elementary exercise in linear algebra and is left for the reader.

To finish the proof, we will show that U_2' is Γ-invariant. Let $y \in U_2'$ and $g \in \Gamma$ arbitrary. By definition, there is $x \in U_1'$ such that $Tx = y$ and therefore

$$\pi_2(g).y = \pi_2(g)T.x = T(\pi_1(g).x).$$

Note that $x \in U_1'$ implies that $\pi_1(g).x \in U_1'$ and therefore $T(\pi_1(g).x) \in U_2'$ as needed. □

Fact 13.14 (Unitary operators - reminder)**.** *For a finite dimensional inner-product space U, a linear operator $T : U \to U$ is called unitary if $T^*T = I$. The following are equivalent:*

1. *The operator T is unitary.*
2. *The operator T preserves the inner-product, i.e., for every $x, y \in U$, $\langle Tx, Ty\rangle = \langle x, y\rangle$.*
3. *The operator T preserves the norm, i.e., for every $x \in U$, $\|Tx\| = \|x\|$.*

Denote $\mathcal{U}(U)$ to be all the unitary operators from U to U and note that this is a group.

Definition 13.15 (Unitary representation)**.** *Let $(U, \langle ., .\rangle)$ be an inner-product space and Γ a finite group. A representation $(\pi, (U, \langle ., .\rangle))$ is called unitary if for every $g \in \Gamma$, $\pi(g)$ is unitary.*

Definition 13.16. *Let $(U_1, \langle ., .\rangle_1), (U_1, \langle ., .\rangle_2)$ be inner-product spaces and Γ a finite group. Two unitary representation $(\pi_1, U_1), (\pi_2, U_2)$ are called unitary equivalent if there is an invertible linear operator such that for every $g \in \Gamma$, $T\pi_1(g) = \pi_2(g)T$ and such that for every $x, y \in U_1$, $\langle x, y\rangle_1 = \langle Tx, Ty\rangle_2$.*

Proposition 13.17. *For a finite group Γ, every representation on an inner-product space is equivalent to a unitary representation.*

Proof. Let $(U, \langle ., .\rangle)$ be an inner-product space $(\pi, (U, \langle ., .\rangle))$ a representation. Define an inner-product $\langle ., .\rangle'$ on U by

$$\langle x, y\rangle' = \sum_{g' \in \Gamma} \langle \pi(g').x, \pi(g').y\rangle.$$

Then for every $x, y \in U$ and every $g \in \Gamma$,

$$\langle \pi(g).x, \pi(g).y\rangle' = \sum_{g' \in \Gamma} \langle \pi(g)\pi(g').x, \pi(g)\pi(g').y\rangle$$

□

Proposition 13.18. *Let Γ be a finite group. Every unitary representation of Γ is either irreducible or decomposable. Moreover, if (π, U) is a decomposable unitary representation, then there are non-zero $U_1, U_2 \subseteq U$ that are Γ-invariant, $U_1 \perp U_2$ and $U = U_1 \oplus U_2$.*

Proof. Let (π, U) be a unitary representation of Γ. If π is irreducible, we are done. Assume that there is a Γ-invariant subspace $U_1 \subseteq U$ such that $U_1 \neq \{0\}, U_1 \neq U$. Denote

$$U_2 = U_1^{\perp} = \{y \in U : \forall x \in U_1, \langle x, y\rangle = 0\}.$$

By definition $U_1 \perp U_2$ and by linear algebra $U = U_1 \oplus U_2$ and $U_2 \neq \{0\}, U_2 \neq U$. Thus, we are left to prove that U_2 is Γ-invariant. We need to prove that for every $y \in U_2$ and every $g \in \Gamma$, $\pi(g).y \in U_2 = U_1^{\perp}$. In other words, we need to

show that every $y \in U_2$, every $x \in U_1$ and every $g \in \Gamma$, $\langle \pi(g).y, x\rangle = 0$. Fix $y \in U_2$, $x \in U_1$ and $g \in \Gamma$.

$$\begin{aligned}\langle \pi(g).y, x\rangle &=^{\pi(g^{-1}) \text{ is unitary}} \langle \pi(g^{-1})\pi(g).y, \pi(g^{-1}).x\rangle \\ &= \langle y, \pi(g^{-1}).x\rangle \\ &=^{\pi(g^{-1}).x \in U_1} 0,\end{aligned}$$

as needed. □

Corollary 13.19. *Let Γ be a finite group. Every representation of Γ is either irreducible or decomposable.*

Proof. Follows from the fact that every representation is equivalent to a unitary representation. □

Theorem 13.20 (Maschke's Theorem for vector spaces over $\mathbb{C}$)**.** *Let Γ be a finite group. Every unitary representation of Γ is completely reducible and moreover the Γ-invariant subspace can be taken to be orthogonal.*

Proof. Let (π, U) be a unitary representation of Γ. We prove the theorem by induction on the dimension of π.

If π is 1-dimensional, we are done since every 1-dimensional representation is irreducible. Assume that π is n-dimensional with $n > 1$. If π is irreducible, we are done. Otherwise, π is decomposable into orthogonal Γ-invariant non-zero subspaces U', U''. The dimension of U', U'' is $< n$ and thus by the induction they are completely reducible and we are done. □

Corollary 13.21. *Let Γ be a finite group. Every representation of Γ is completely reducible.*

Definition 13.22 (Morphism of representations)**.** *Let Γ be a finite group and $(\pi_1, U_1), (\pi_2, U_2)$ be representations of Γ. A morphism $T \in Hom(\pi_1, \pi_2)$ is a linear map $T : U_1 \to U_2$ such that $\pi_2 \circ T = T \circ \pi_1$, i.e., for every $g \in \Gamma$,*

$$\pi_2(g)T = T\pi_1(g).$$

Recall that two representations $(\pi_1, U_1), (\pi_2, U_2)$ are called equivalent if there is $T \in \mathrm{Hom}(\pi_1, \pi_2)$ such that T is invertible (i.e., bijective) and in that case, $T^{-1} \circ \pi_2 \circ T = \pi_1$.

Proposition 13.23. *For every $(\pi_1, U_1), (\pi_2, U_2)$, $\mathrm{Hom}(\pi_1, \pi_2)$ is a linear space over $\mathbb{C}$.*

Proof. Let $T, S \in \mathrm{Hom}(\pi_1, \pi_2)$ and let $\alpha \in \mathbb{C}$, then for every $g \in \Gamma$,

$$\begin{aligned}\pi_2(g)(T + \alpha S) &= \pi_2(g)T + \alpha\pi_2(g)S \\ &= T\pi_1(g) + \alpha S\pi_1(g) \\ &= (T + \alpha S)\pi_1(g).\end{aligned}$$

□

Proposition 13.24. *For every $(\pi_1, U_1), (\pi_2, U_2)$, and every $T \in \mathrm{Hom}(\pi_1, \pi_2)$, $\ker(T) \subseteq U_1$ and $\mathrm{Im}(T) \subseteq U_2$ are Γ-invariant subspaces.*

Proof. Let $x \in \ker(T)$ and $g \in \Gamma$, then

$$T(\pi_1(g).x) = \pi_2(g).(Tx) = \pi_2(g).0 = 0,$$

and therefore $\pi_1(g).x \in \ker(T)$ as needed.

Let $y \in \mathrm{Im}(T)$ and $g \in \Gamma$, then there is $x \in U_1$ such that $Tx = y$ and

$$\pi_2(g).y = \pi_2(g).(Tx) = T(\pi_1(g).x) \in \mathrm{Im}(T).$$

□

Lemma 13.25 (Schur's Lemma)**.** *Let Γ be a finite group and let $(\pi_1, U_1), (\pi_2, U_2)$ two irreducible representations of Γ. Then every $T \in \mathrm{Hom}(\pi_1, \pi_2)$ is either $T = 0$ or invertible. Consequently:*

1. *If $\pi_1 \nsim \pi_2$, then $\mathrm{Hom}(\pi_1, \pi_2) = \{0\}$.*
2. *If (π, U) is irreducible, then $\mathrm{Hom}(\pi, \pi) = \{\lambda I : \lambda \in \mathbb{C}\}$.*

Proof. Let $T \in \mathrm{Hom}(\pi_1, \pi_2)$. If $T = 0$, we are done. Assume $T \neq 0$, then $\ker(T) \neq U_1$. By Proposition 13.24, $\ker(T)$ is Γ-invariant and by the assumption that π_1 is irreducible it follows that $\ker(T) = \{0\}$ (note that we established that $\ker(T) \neq U_1$), i.e., T is injective. It follows that $\mathrm{Im}(T) \neq \{0\}$. By Proposition 13.24, $\mathrm{Im}(T)$ is Γ-invariant and since π_2 is irreducible, it follows that $\mathrm{Im}(T) = U_2$, i.e., T is also surjective as needed.

It follows immediately that if $\pi_1 \nsim \pi_2$, then $\mathrm{Hom}(\pi_1, \pi_2) = \{0\}$.

To prove the last assertion, we note that for every $\lambda \in \mathbb{C}$, λI commutes with every matrix and therefore $\pi(g)(\lambda I) = (\lambda I)\pi(g)$, i.e., $\{\lambda I : \lambda \in \mathbb{C}\} \subseteq \mathrm{Hom}(\pi, \pi)$. In the other direction, let $T \in \mathrm{Hom}(\pi, \pi)$. Since we are working over $\mathbb{C}$, $T : U \to U$ has some an eigenvalue which we done $\lambda \in \mathbb{C}$. By definition, $T - \lambda I$ is not invertible and by the above proposition, $(T - \lambda I) \in \mathrm{Hom}(\pi, \pi)$. It follows from the first part of the Lemma that $T - \lambda I = 0$ or $T = \lambda I$ as needed. □

Corollary 13.26. *Let Γ be a finite Abelian group. All the irreducible representations of Γ are 1-dimensional.*

Proof. Let (π, U) be an irreducible representation of Γ. Fix $x \in U$ be some non-zero vector. We will show that $\mathrm{span}\{x\}$ is Γ-invariant and therefore U is 1-dimensional.

Let $g_0 \in \Gamma$ and denote $\pi(g_0) = T$. Then for every $g \in \Gamma$,

$$T\pi(g) = \pi(g_0)\pi(g) = \pi(g_0 g) = \pi(g g_0) = \pi(g)\pi(g_0) = \pi(g)T,$$

i.e., $\pi(g_0) = T \in \mathrm{Hom}(\pi, \pi)$ and by Schur's Lemma there is $\lambda \in \mathbb{C}$ such that $\pi(g_0) = \lambda_{g_0} I$. Then $\pi(g_0).x = \lambda x \in \mathrm{span}\{x\}$. Since this is true for every $g_0 \in \Gamma$, it follows that $\mathrm{span}\{x\}$ is Γ-invariant as needed. □

Definition 13.27 (Regular representations)**.** *Let Γ be a finite group. Denote $\ell^2(\Gamma)$ to be the vector space of functions $\phi : \Gamma \to \mathbb{C}$ with the inner-product*

$$\langle \phi, \psi \rangle = \sum_{g \in \Gamma} \phi(g)\overline{\psi(g)}.$$

The left-regular representation of Γ *on* $\ell^2(\Gamma)$ *is defined by*

$$(\lambda(g).\phi)(g') = \phi(g^{-1}g').$$

The right-regular representation of Γ *on* $\ell^2(\Gamma)$ *is defined by*

$$(\rho(g).\phi)(g') = \phi(g'g).$$

Proposition 13.28. *The left and right regular representations are indeed group representation and moreover, they are both unitary representation.*

Proof. We will prove the proposition only for the right-regular representation and leave the (symmetric) proof for the left-regular representation for the reader.

First, we verify that ρ is in fact a representation. For every $g \in \Gamma$, $\rho(g)$ is obviously linear and therefore we are left to show that $\rho(g_1g_2) = \rho(g_1)\rho(g_2)$. Let $\phi \in \ell^2(\Gamma)$, $g_1, g_2 \in \Gamma$, then for every $g' \in \Gamma$,

$$\begin{aligned}(\rho(g_1)(\rho(g_2)\phi))(g') &= (\rho(g_2)\phi)(g'g_1)\\ &= \phi(g'g_1g_2)\\ &= (\rho(g_1g_2)\phi)(g').\end{aligned}$$

Second, we show that ρ is unitary: let $\phi, \psi \in \ell^2(\Gamma)$, and $g \in \Gamma$, then

$$\begin{aligned}\langle \rho(g)\phi, \rho(g)\psi\rangle &= \sum_{g'\in\Gamma} (\rho(g)\phi)(g')\overline{(\rho(g)\psi)(g')}\\ &= \sum_{g'\in\Gamma} \phi(g'g)\overline{\psi(g'g)}\\ &=^{g''=g'g} \sum_{g''\in\Gamma} \phi(g'')\overline{\psi(g'')}\\ &= \langle \phi, \psi\rangle.\end{aligned}$$

□

Theorem 13.29 (Peter-Weyl Theorem - without proof)**.** *Let* Γ *be a finite group. Let* Λ *be a set of representatives for the equivalence classes of irreducible unitary representations of* Γ*, i.e.,*

$$\Lambda = \{(\pi, E_\pi) : \pi \text{ is an irreducible unitary repsentation of } \Gamma\},$$

such that no two different $\pi_1, \pi_2 \in \Lambda$ *are equivalent and for every irreducible unitary representation* π' *of* Γ *there is* $\pi \in \Lambda$ *such that* π' *and* π *are equivalent. Then the right (or left) regular representation decomposes orthogonally as*

$$\rho = \bigoplus_{\pi\in\Lambda} \pi^{\deg(\pi)}.$$

In other words, there are orthogonal subspaces of $\ell^2(\Gamma)$*,* $\{U_j^\pi : 1 \le j \le \deg(\pi) : \pi \in \Lambda\}$ *that are* Γ*-invariant such that*

$$\ell^2(\Gamma) = \bigoplus_{\pi\in\Lambda} \bigoplus_{1\le j\le \deg(\pi)} U_j^\pi,$$

and such that for every π *and every* j*,* $\rho|_{U_j^\pi}$ *is unitary equivalent to* π*.*

Corollary 13.30. *For every finite group Γ, we have that*

$$|\Gamma| = \sum_{\pi \in \Lambda} \deg(\pi)^2.$$

In Peter-Weyl Theorem the decomposition is given more explicitly than stated above: using the language of matrix coefficients (which we will not define), one can describe the subspaces U_j^π. We will not do so, but only observe the following: the subspace of constant functions in $\ell^2(\Gamma)$ is Γ-invariant with respect to the left and right regular representations and the representation ρ restricted to this subspace is the trivial representation. Thus, we define

$$\ell_0^2(\Gamma) = \{\phi \in \ell^2(\Gamma) : \sum_{g \in \Gamma} \phi(g) = 0\},$$

and note that by Peter-Weyl Theorem,

$$\rho|_{\ell_0^2(\Gamma)} = \bigoplus_{\pi \in \Lambda, \pi \not\sim \pi_{trivial}} \pi^{\deg(\pi)}.$$

13.1 Exercises

1. Prove that for every finite group Γ, the left-regular representation and the right-regular representation are equivalent.

2. Recall that for $\Gamma = \mathbb{Z}/n\mathbb{Z}$, we saw that

   $$\pi_j(m) = (e^{\frac{2\pi i j m}{n}}),$$

   are representations for every $j = 0, ..., n-1$.

   (a) Prove that $\Lambda = \{\pi_j : j = 0, ..., n-1\}$ is a set of representatives for the equivalence classes of irreducible unitary representations of Γ.

   (b) Let $\Gamma = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Show that all the irreducible unitary representations (up to equivalence) of Γ are given by $\pi_{k,j}((m_1, m_2)) = \pi_k(m_1)\pi_j(m_2)$, where $k, j = 0, ..., n-1$.

3. Let Γ be a finite group and $N \triangleleft \Gamma$ be a normal subgroup. Show that if (π, U) is a representation of Γ/N, then $\pi' : \Gamma \to \mathrm{GL}(U)$ defined as $\pi'(g) = \pi(gN)$ is a representation of Γ.

4. Let Γ be a finite group acting on a finite graph (V, E) and define $\ell^2(V)$ as above. Prove that $(\pi, \ell^2(V))$ defined as

   $$\pi(g).\phi(v) = \phi(g^{-1}.v),$$

   is a unitary representation of Γ and that the constant functions and $\ell_0^2(V)$ are both Γ-invariant subspaces of this representation.

14 Representation theory of finite groups and expansion of graphs

Let Γ be a finite group with a symmetric generating set S. Let $\mathrm{Cay}(\Gamma; S) = (V, E)$ be the Cayley graph of Γ. We observe the following: the space $\ell^2(V)$ and the space $\ell^2(\Gamma)$ are isomorphic as inner-product spaces, where the inner-product differs by a factor of $|S|$. In other words, both $\ell^2(V)$ and $\ell^2(\Gamma)$ are spaces of functions $\phi : \Gamma \to \mathbb{C}$, with the following (slight) difference in the inner-product:

$$\langle \phi, \psi \rangle_{\ell^2(V)} = \sum_{g \in \Gamma} |S| \phi(g) \overline{\psi(g)},$$

$$\langle \phi, \psi \rangle_{\ell^2(\Gamma)} = \sum_{g \in \Gamma} \phi(g) \overline{\psi(g)}.$$

Thus, $\langle \phi, \psi \rangle_{\ell^2(V)} = |S| \langle \phi, \psi \rangle_{\ell^2(\Gamma)}$ and we identify between $\ell^2(V)$ and $\ell^2(\Gamma)$ (the difference in the inner-product will not matter when considering expansion). We also identify between $\ell^2_0(V)$ and $\ell^2_0(\Gamma)$.

Proposition 14.1. *Let Γ be a finite group with a generating set S. The graph $\mathrm{Cay}(\Gamma; S) = (V, E)$ is a λ-spectral expander if and only if for every $\phi \in \ell^2_0(\Gamma)$,*

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) . \phi, \phi \rangle_{\ell^2(\Gamma)} \leq \lambda \|\phi\|^2_{\ell^2(\Gamma)}.$$

Proof. Let $\phi \in \ell^2(V)$ and $g \in V = \Gamma$,

$$(M\phi)(g) = \frac{1}{|S|} \sum_{s \in S} \phi(gs) = \frac{1}{|S|} \sum_{s \in S} \rho(s).\phi(g) = (\left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) . \phi)(g),$$

where ρ is the right-regular representation.

Recall that $\mathrm{Cay}(\Gamma; S) = (V, E)$ is a λ-expander if and only if for every $\phi \in \ell^2_0(V)$,

$$\langle M\phi, \phi \rangle_{\ell^2(V)} \leq \lambda \|\phi\|^2_{\ell^2(V)}.$$

Passing to $\ell^2(\Gamma)$, this reads as follows - for every $\phi \in \ell^2_0(\Gamma)$,

$$|S| \langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) . \phi, \phi \rangle_{\ell^2(\Gamma)} \leq \lambda |S| \|\phi\|^2_{\ell^2(\Gamma)},$$

and dividing by $|S|$ yields the needed inequality. □

Proposition 14.2 (Expansion of the Cayley graph of a finite group can be deduced using the unitary representations)**.** *Let Γ be a finite group with a generating set S. The graph $\mathrm{Cay}(\Gamma; S) = (V, E)$ is a λ-spectral expander if and only if for every non-trivial irreducible unitary representation (π, U), for every $x \in U$*

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) . x, x \rangle_U \leq \lambda \|x\|^2_U.$$

Proof. Assume first that for every non-trivial irreducible unitary representation (π, U), for every $x \in U$

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) .x, x \rangle_U \leq \lambda \|x\|_U^2.$$

Recall that from Peter-Weyl Theorem,

$$\ell^2(\Gamma) = \bigoplus_{\pi \in \Lambda} \bigoplus_{1 \leq j \leq \deg(\pi)} U_j^\pi,$$

where Λ are representatives of equivalence class of irreducible representations, U_j^π are Γ-invariant subspace and for every π and every j, $\rho|_{U_j^\pi}$ is unitary equivalent to π. Thus $\phi \in \ell_0^2(\Gamma)$ decomposes orthogonally to

$$\phi = \sum_{\pi \in \Lambda, \pi \nsim \pi_{trivial}} \sum_{1 \leq j \leq \deg(\pi)} \phi_j^\pi,$$

where $\phi_j^\pi \in U_j^\pi$.

The composition above is orthogonal, thus

$$\|\phi\|^2 = \sum_{\pi \in \Lambda, \pi \nsim \pi_{trivial}} \sum_{1 \leq j \leq \deg(\pi)} \|\phi_j^\pi\|^2.$$

Note that for every $\phi_j^\pi \in U_j^\pi$, $(\left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) .\phi_j^\pi) \in U_j^\pi$. Therefore,

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) .\phi, \phi \rangle = \sum_{\pi \in \Lambda, \pi \nsim \pi_{trivial}} \sum_{1 \leq j \leq \deg(\pi)} \langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) .\phi_j^\pi, \phi_j^\pi \rangle.$$

By our assumption, for every π, j,

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) .\phi_j^\pi, \phi_j^\pi \rangle \leq \lambda \|\phi_j^\pi\|^2,$$

and thus by summation,

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) .\phi, \phi \rangle = \sum_{\pi \in \Lambda, \pi \nsim \pi_{trivial}} \sum_{1 \leq j \leq \deg(\pi)} \langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) .\phi_j^\pi, \phi_j^\pi \rangle \leq$$
$$\sum_{\pi \in \Lambda, \pi \nsim \pi_{trivial}} \sum_{1 \leq j \leq \deg(\pi)} \lambda \|\phi_j^\pi\|^2 = \lambda \|\phi\|^2.$$

In the other direction, assume towards contradiction that $\mathrm{Cay}(\Gamma; S)$ is a λ-expander and that there is some non-trivial irreducible unitary representation (π, U), for every $x \in U$

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) .x, x \rangle_U > \lambda \|x\|_U^2.$$

From Peter-Weyl, that there is $\phi = \phi_1^\pi \in \ell_0^2(\Gamma)$ such that

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \rho(s) \right) .\phi, \phi \rangle > \lambda \|\phi\|^2,$$

and the contradiction follows. □

Definition 14.3. *Let Γ be a group and (π, U) be a representation. The representation π is said to have a non-trivial invariant vector if there is $x \in U, x \neq 0$ such that for every $g \in \Gamma$, $\pi(g).x = x$.*

Observation 14.4. *A representation π has a non-trivial invariant vector if and only if the trivial representation is a sub-representation of π.*

Corollary 14.5. *Let Γ be a finite group with a generating set S and $\lambda < 1$ some constant. If $\mathrm{Cay}(\Gamma; S)$ is a λ-expander, if and only if for every unitary representation (π, U) that does not have a non-trivial invariant vector*

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) .x, x \rangle_U \leq \lambda \|x\|_U^2.$$

Proof. Assume that $\mathrm{Cay}(\Gamma; S)$ is a λ-expander, then for every non-trivial irreducible unitary representation the assertion holds. By Maschke's Theorem every unitary representation is an orthogonal sum of unitary irreducible representations and by our assumption non of them is the trivial representation and we finish by a computation similar to that in the proof of proposition above.

In the other direction, observe that $\rho|_{\ell_0^2(\Gamma)}$ is a representation that does not have a non-trivial invariant vector. □

Exercise 14.6. *Let Γ be a finite group with a generating set S and $(\pi, \mathbb{C})$ a one dimensional unitary representation of Γ. Show $\phi_\pi \in \ell^2(\Gamma)$ defined as $\phi(g) = \pi(g).1$ is an eigenfunction of $M : \ell^2(\Gamma) \to \ell^2(\Gamma)$, where M is the random walk operator on $\mathrm{Cay}(\Gamma; S)$. Deduce that if Γ is an Abelian group, we can find explicitly all the eigenfunction of M on the Cayley graph and note that this is exactly what we did for $\mathbb{Z}_n \times \mathbb{Z}_n$ in the beginning of the proof of Gabber-Galil-Margulis expanders.*

15 The mother group approach to expanders

Proposition 15.1. *Let Γ be a finite group with a generating set S and $\lambda < 1$ some constant. Assume that for every (π, U) unitary representation of Γ that does not have a non-trivial invariant vector, it holds that:*

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) .x, x \rangle_U \leq \lambda \|x\|_U^2.$$

Then for every subgroup $H < \Gamma$, $\mathrm{Sch}(\Gamma, H; S)$ is a λ-expander.

Proof. Recall that the vertex set of $\mathrm{Sch}(\Gamma, H; S)$ are cosets $V = \{Hg : g \in \Gamma\}$ and Hg, Hg' are connected by an edge if and only if $g^{-1}g' \in S$ (or in other words, there is $s \in S$ such that $g' = gs$). Define a representation π as follows:

$$U = \ell_0^2(V) = \{\phi : V \to \mathbb{C} : \sum_{v \in V} \phi(v) = 0\},$$

and for every $g \in \Gamma$ define $(\pi(g).\phi)(Hg') = \phi(Hg'g)$ (verify that this is indeed a representation). Note that this representation has no non-trivial invariant vector, because if ϕ is invariant, then for every $g \in \Gamma$,

$$\phi(Hg) = (\pi(g).\phi)(He) = \phi(He),$$

i.e., ϕ is the constant function, but the only constant function in $\ell_0^2(V)$ is the 0-function.

Thus, by the assumption of the proposition, for every $\phi \in \ell_0^2(V)$,

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) .\phi, \phi \rangle_{\ell^2(V)} \leq \lambda \|\phi\|^2_{\ell^2(V)}.$$

We finish by noting that for every $Hg \in V$,

$$(\left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) .\phi)(Hg) = \frac{1}{|S|} \sum_{s \in S} \phi(Hgs) = (M\phi)(Hg),$$

and therefore we showed that for every $\phi \in \ell_0^2(V)$,

$$\langle M\phi, \phi \rangle \leq \lambda \|\phi\|^2,$$

as needed. □

The above proposition hints towards a way to construct expanders that is called the "Mother group approach". Start with a finitely generated group Γ which is expanding in the following sense: there is a finite generating set S and a constant $\lambda < 1$ such that if (π, U) is a unitary representation that does not have a non-trivial invariant vector, then

$$\langle \left(\frac{1}{|S|} \sum_{s \in S} \pi(s) \right) .x, x \rangle_U \leq \lambda \|x\|^2_U.$$

For every $H < \Gamma$, $\mathrm{Sch}(\Gamma, H; S)$ will be a λ-expander and thus if Γ has "many" subgroups, then we get many graphs. However, if Γ is a finite group, this approach cannot yield and infinite family of graphs, since for every $H < \Gamma$, $\mathrm{Sch}(\Gamma, H; S)$ has at most $|\Gamma|$ vertices.

To make this approach work, we need to pass to infinite groups. We note that Proposition 15.1, did not use the fact that Γ was a finite group, but only the fact that $H\backslash\Gamma$ was finite and therefore $\mathrm{Sch}(\Gamma, H; S)$ was finite.

We will show below that the expansion property of Γ stated above is in fact a slightly "watered-down" version of what is known as Kazhdan's property (T). To define Kazhdan's property (T), we will first need to give a (very) brief overview on Hilbert spaces.

16 Hilbert spaces - a crash course

A Hilbert space is meant to generalize the finite-dimensional inner-product spaces to the setting of an infinite dimension. We will no go into the theory of Hilbert spaces - just give the basic definitions and state a few facts.

16.1 Normed spaces

Definition 16.1 (Norm)**.** *Given a vector space U over $\mathbb{C}$, a norm on U is a function $\|.\| \to \mathbb{R}_{\geq 0}$ such that:*

1. $\|x\| = 0$ *if and only if* $x = 0$.

2. *For every* $\alpha \in \mathbb{C}$ *and every* $x \in U$, $\|\alpha x\| = |\alpha|\|x\|$.

3. *For every* $x, y \in U$, $\|x + y\| \leq \|x\| + \|y\|$.

The couple $(U, \|.\|)$ *is called a normed space.*

The example to keep in mind the norm induced by an inner-product.

Definition 16.2 (Convergence)**.** *Let* $(U, \|.\|)$ *be a normed space. A sequence* $\{x_n\}_{n\in\mathbb{N}} \subseteq U$ *is said to converge to* $x_0 \in U$ *if for every* $\varepsilon > 0$*, there is* N *such that for every* $n > N$, $\|x_n - x_0\| < \varepsilon$.

Definition 16.3 (Operator terminology)**.** *Let* $(U, \|.\|)$ *be a normed space and* $T : U \to U$ *be a linear operator.*

1. *The operator* T *is called continuous if for every convergent sequence* $\{x_n\}_{n\in\mathbb{N}} \subseteq U$ *such that* $x_n \to x_0$*, we have that* $Tx_n \to Tx_0$.

2. *The operator norm of* T *is*

$$\|T\| = \sup_{x\in U, \|x\|=1} \|Tx\| = \sup_{x\in U, x\neq 0} \frac{\|Tx\|}{\|x\|}.$$

3. *The operator* T *is called bounded if* $\|T\| < \infty$*. We denote* $B(U)$ *to be the space of all bounded linear operators.*

Fact 16.4. *Let* $(U, \|.\|)$ *be a normed space and* $T : U \to U$ *be a linear operator. The operator* T *is continuous if and only if it is bounded.*

Definition 16.5 (Cauchy sequence)**.** *Let* $(U, \|.\|)$ *be a normed space. A sequence* $\{x_n\}_{n\in\mathbb{N}} \subseteq U$ *is called a Cauchy sequence if for every* $\varepsilon > 0$ *there is* N *such that for every* $m, n > N$, $\|x_n - x_m\| < \varepsilon$.

Remark 16.6. *It is an easy exercise to show that every convergent sequence is Cauchy.*

Definition 16.7 (Banach spaces)**.** *A normed space* $(U, \|.\|)$ *is called complete if every Cauchy sequence is convergent. A complete normed space is called a Banach space.*

16.2 Hilbert spaces

Definition 16.8 (Hilbert Space)**.** *An inner-product space* $\mathcal{H}$ *is called a Hilbert space if it is complete with respect to the norm induced by the inner-product.*

Example: $\mathbb{C}^n$ is a complete space with respect to the norm and therefore a Hilbert space.

The most classical example of a Hilbert space of infinite dimension is the space ℓ^2:

Example 16.9 (ℓ^2)**.** *Let Ω be a countable set. We denote $\ell^2(\Omega)$ to be the following space*

$$\ell^2(\Omega) = \{\phi : \Omega \to \mathbb{C} : \sum_{v\in\Omega} |\phi(v)|^2 < \infty\}.$$

We claim that $\ell^2(\Omega)$ is a Hilbert space with respect to the inner-product defined as

$$\langle \phi, \psi \rangle = \sum_{v\in\Omega} \phi(v)\overline{\psi(v)},$$

and the induced norm is

$$\|\phi\| = \sqrt{\sum_{v\in\Omega} |\phi(v)|^2}.$$

We omit the proof that $\ell^2(\Omega)$ is a Hilbert space, and we note that there are few things we need to check here:

1. *ℓ^2 is a vector space.*
2. *$\langle ., . \rangle$ defined above is well defined and an inner-product.*
3. *ℓ^2 is complete with respect to the norm induced by the inner-product.*

16.3 Unitary operators and orthogonal projections

Definition 16.10 (Adjoint operator)**.** *Let $\mathcal{H}$ be a Hilbert space and $T \in B(\mathcal{H})$. The adjoint of T is an operator denoted $T^* : \mathcal{H} \to \mathcal{H}$ and defined such that for every $x, y \in \mathcal{H}$,*

$$\langle Tx, y \rangle = \langle x, T^* y \rangle.$$

Fact 16.11. *For every $T \in B(\mathcal{H})$ there is $T^* \in B(\mathcal{H})$ (i.e., T^* exists, T^* is unique and T^* is bounded).*

Example 16.12. *If $\mathcal{H} = \mathbb{C}^n$ with the usual inner-product and T is given by a matrix A, then T^* is given by the matrix A^*.*

Definition 16.13 (Unitary operator)**.** *Let $\mathcal{H}$ be a Hilbert space. An operator $T \in B(\mathcal{H})$ is called unitary if $T^*T = TT^* = I$ (I is the identity operator). We denote $\mathcal{U}(\mathcal{H})$ to be the group of all unitary operators of $\mathcal{H}$.*

Fact 16.14. *The following are equivalent:*

1. *The operator T is unitary.*
2. *The operator T is invertible and preserves the inner-product, i.e., for every $x, y \in \mathcal{H}$, $\langle Tx, Ty \rangle = \langle x, y \rangle$.*
3. *The operator T is invertible and preserves the norm, i.e., for every $x \in \mathcal{H}$, $\|Tx\| = \|x\|$.*

Remark 16.15. *Note that since $T \in \mathcal{U}(\mathcal{H})$ preserves the norm, it follows that $\|T\| = 1$.*

Definition 16.16 (Orthogonal projection)**.** *Let $\mathcal{H}$ be a Hilbert space and $T \in B(\mathcal{H})$. The operator T is called an orthogonal projection if $T = T^*$ and $T^2 = T$ (T^2 denotes $T \circ T$). An orthogonal projection P is called in orthogonal projection on $\mathcal{H}'$ if $\mathrm{Im}(P) = \mathcal{H}'$.*

Exercise 16.17. *If P is an orthogonal projection on $\mathcal{H}' \subseteq \mathcal{H}$, then $I - P$ is an orthogonal projection on*

$$(\mathcal{H}')^{\perp} = \{x \in \mathcal{H} : \forall y \in \mathcal{H}', \langle x, y\rangle = 0\}.$$

Fact 16.18. *An operator $T \in B(\mathcal{H})$ is an orthogonal projection if and only if $T^2 = T$ and $\|T\| \leq 1$.*

Definition 16.19 (Closed subspace)**.** *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{H}' \subseteq \mathcal{H}$ be a linear subspace of $\mathcal{H}$. The subspace $\mathcal{H}'$ is called closed if for any convergent sequence $x_n \to x_0$, if $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathcal{H}'$, then $x_0 \in \mathcal{H}'$.*

Fact 16.20. *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{H}' \subseteq \mathcal{H}$ be a closed subspace of $\mathcal{H}$. There is a unique orthogonal projection $P_{\mathcal{H}'} \in B(\mathcal{H})$ such that $P_{\mathcal{H}'}$ is an orthogonal projection on $\mathcal{H}'$ and for this $P_{\mathcal{H}'}$ the following holds: for every $x \in \mathcal{H}$,*

$$\|x - P_{\mathcal{H}'}x\| = \inf_{y\in\mathcal{H}'} \|y - x\|.$$

In other words, $P_{\mathcal{H}'}$ sends each $x \in \mathcal{H}$ to the point closest to it in $\mathcal{H}'$.

17 Unitary representations of infinite groups and definition of Kazhdan property (T)

Definition 17.1. *Let Γ be a countable (finite or infinite) group and $\mathcal{H}$ be a Hilbert space (of finite or infinite dimension). A unitary representation $(\pi, \mathcal{H})$ of Γ is a group homomorphism $\pi : \Gamma \to \mathcal{U}(\mathcal{H})$.*

This definition generalizes our definition for finite dimensional representations for finite groups. To see an infinite dimensional example, we consider the regular representation:

Example 17.2. *Let Γ be an infinite countable group. Define $\ell^2(\Gamma)$ as:*

$$\{\phi : \Gamma \to \mathbb{C} : \sum_{g\in\Gamma} |\phi|^2 < \infty\},$$

with the inner-product

$$\langle \phi, \psi\rangle = \sum_{g\in\Gamma} \phi(g)\overline{\psi(g)}.$$

(As noted above, it is a Hilbert space, but this is not a trivial fact). Define $\rho : \Gamma \to \mathcal{U}(\ell^2(\Gamma))$ to be the right-regular representation: $(\rho(g).\phi)(g') = \phi(g'g)$. One can verify that this representation is unitary.

Note that if Γ is infinite, then ρ does not have non-trivial invariant vector, since every invariant vector must be a constant function and if Γ is infinite, the only constant function in $\ell^2(\Gamma)$ is the zero function.

We recall that our original motivation was to define an expansion notion of a group regarding unitary representations. This notion is called "Kazhdan Property (T)". We will start by giving the standard definition of property (T) and then show that it coincides with expansion:

Definition 17.3 (Kazhdan Property (T))**.** *Let Γ be a countable group. The group Γ is said to have property (T) if there is a finite set $S \subseteq \Gamma$ and a constant $\varepsilon > 0$ such that for every unitary representation $(\pi, \mathcal{H})$ that does not have a non-trivial invariant vector it follows that for every $x \in \mathcal{H}$,*

$$\max_{s \in S} \|x - \pi(s).x\| \geq \varepsilon \|x\|.$$

We quantify this definition as follows: let Γ and $S \subseteq \Gamma$ as above. For a unitary representation $(\pi, \mathcal{H})$ without an invariant vector, we denote $\kappa_\pi(\Gamma, S)$ to be:

$$\kappa_\pi(\Gamma, S) = \max\{\varepsilon \geq 0 : \forall x \in \mathcal{H}, \max_{s \in S} \|x - \pi(s).x\| \geq \varepsilon \|x\|\}.$$

We further denote, $\kappa(\Gamma, S) = \inf_\pi \kappa_\pi(\Gamma, S)$, where the infimum is taken over all unitary representations that do not have a non-trivial invariant vector. The constant $\kappa(\Gamma, S)$ is called Kazhdan's cosntant of Γ with respect to S and Γ has property (T) if and only if there is a finite set S such that $\kappa(\Gamma, S) > 0$.

Note that S in the definition above does not have to be a generating set. In fact, originally, Kazhdan defined property (T) as a tool to prove finite generation (he showed that every group with property (T) is finitely generated). However, if we know that Γ is finitely generated, then S can be taken to be a finite generating set as this exercise implies:

Exercise 17.4. *Let Γ be a finitely generated group and $S' \subseteq \Gamma$ be a finite set such that $\kappa(\Gamma, S') > 0$. Prove that for every finite generating set S, $\kappa(\Gamma, S) > 0$.*

In the literature the following are standard variations on the definition:

1. Normalizing: we note that for every $x \neq 0$,

   $$\max_{s \in S} \|x - \pi(s).x\| \geq \varepsilon \|x\|,$$

   is equivalent to

   $$\max_{s \in S} \| \frac{x}{\|x\|} - \pi(s). \frac{x}{\|x\|} \| \geq \varepsilon,$$

   and thus the definition is sometimes written as follows:

   The group Γ is said to have property (T) if there is a finite set $S \subseteq \Gamma$ and a constant $\varepsilon > 0$ such that for every unitary representation $(\pi, \mathcal{H})$ that does not have a non-trivial invariant vector it follows that for every **unit vector** $x \in \mathcal{H}$,

   $$\max_{s \in S} \|x - \pi(s).x\| \geq \varepsilon.$$

2. Almost invariant vectors imply invariant vectors (contra-positive): for a group Γ, a finite set $S \subseteq \Gamma$, a constant $\varepsilon > 0$ and a representation $(\pi, \mathcal{H})$, a unit vector $x \in \mathcal{H}$ is called ε-almost invariant, if

   $$\max_{s \in S} \|x - \pi(s).x\| < \varepsilon.$$

   By passing to the contra-positive, one can give the definition of property (T) in the following form:

   The group Γ is said to have property (T) if there is a finite set $S \subseteq \Gamma$ and a constant $\varepsilon > 0$ such that for every unitary representation $(\pi, \mathcal{H})$ if there is a ε-almost invariant unit vector $x \in \mathcal{H}$, then it has a non-trivial invariant vector.

At this point, the reader might be confused since earlier when discussing the "mother group" approach to expanders we have put forward a different notion of group expansion, namely, we asked that for Γ there is a finite symmetric set S and a constant $\lambda < 1$, such that for every unitary representation $(\pi, \mathcal{H})$ that does not have a non-trivial invariant vector, it follows for every $x \in \mathcal{H}$ that

$$\langle \frac{1}{|S|} \sum_{s \in S} \pi(s).x, x \rangle \leq \lambda \|x\|^2.$$

The following proposition states that this condition is in fact equivalent to property (T):

Proposition 17.5. *Let Γ be a group with a finite generating set S. Then Γ has property (T) if and only if there is $\lambda < 1$, such that for every unitary representation $(\pi, \mathcal{H})$ that does not have a non-trivial invariant vector, it follows for every $x \in \mathcal{H}$ that*

$$\langle \frac{1}{|S|} \sum_{s \in S} \pi(s).x, x \rangle \leq \lambda \|x\|^2.$$

Moreover, $\kappa(\Gamma, S) \geq \sqrt{2(1-\lambda)}$ and $\lambda \leq (1 - \frac{\kappa(\Gamma,S)^2}{2|S|})$.

This proposition will easily follow from the following lemma:

Lemma 17.6. *Let Γ be a group with a finite generating set S and $(\pi, \mathcal{H})$ be any unitary representation. Then,*

$$\langle (I - \frac{1}{|S|} \sum_{s \in S} \pi(s)).x, x \rangle = \frac{1}{2|S|} \sum_{s \in S} \|x - \pi(s).x\|^2.$$

Proof of the proposition based on the lemma. If $\kappa(\Gamma, S) \geq \varepsilon > 0$, then for every unitary representation $(\pi, \mathcal{H})$ that does not have a non-trivial invariant vector, it follows that

$$\|x\|^2 - \langle \frac{1}{|S|} \sum_{s \in S} \pi(s)).x, x \rangle =$$

$$\langle (I - \frac{1}{|S|} \sum_{s \in S} \pi(s)).x, x \rangle = \frac{1}{2|S|} \sum_{s \in S} \|x - \pi(s).x\|^2 \geq \frac{\varepsilon^2}{2|S|} \|x\|^2,$$

and thus,

$$\langle \frac{1}{|S|} \sum_{s \in S} \pi(s)).x, x \rangle \leq (1 - \frac{\varepsilon^2}{2|S|}) \|x\|^2.$$

In the other direction, assume that there is $\lambda < 1$ such that for every unitary representation $(\pi, \mathcal{H})$ that does not have a non-trivial invariant vector,

$$\langle \frac{1}{|S|} \sum_{s \in S} \pi(s).x, x \rangle \leq \lambda \|x\|^2.$$

Then

$$(1-\lambda)\|x\|^2 \leq \langle (I - \frac{1}{|S|} \sum_{s \in S} \pi(s).x), x \rangle =$$

$$\frac{1}{2|S|} \sum_{s \in S} \|x - \pi(s).x\|^2 \leq \frac{1}{2} \left(\max_{s \in S} \|x - \pi(s).x\| \right)^2,$$

and it follows that $\kappa(\Gamma, S) \geq \sqrt{2(1-\lambda)} > 0$. □

Proof of the lemma. Let Γ, S as above and let $(\pi, \mathcal{H})$ some representation. Then for every $x \in \mathcal{H}$,

$$\sum_{s \in S} \|x - \pi(s).x\|^2 = \sum_{s \in S} \langle (I - \pi(s)).x, (I - \pi(s)).x \rangle =$$
$$\sum_{s \in S} \langle (I - \pi(s))^*(I - \pi(s)).x, x \rangle =$$
$$\sum_{s \in S} \langle (I - \pi(s^{-1}))(I - \pi(s)).x, x \rangle = \sum_{s \in S} \langle (2I - \pi(s) - \pi(s^{-1})).x, x \rangle =$$
$$\sum_{s \in S} \langle (I - \pi(s)).x, x \rangle + \sum_{s \in S} \langle (I - \pi(s^{-1})).x, x \rangle =$$
$$2 \sum_{s \in S} \langle (I - \pi(s)).x, x \rangle =$$
$$2 \langle (|S| I - \sum_{s \in S} \pi(s)).x, x \rangle =$$
$$2|S| \langle (I - \frac{1}{|S|} \sum_{s \in S} \pi(s)).x, x \rangle.$$

□

Example 17.7. *Every finite group Γ has property (T): take $S = \Gamma$ and let $(\pi, \mathcal{H})$ be some unitary representation without a non-trivial invariant vector. Let $x \in \mathcal{H}$. Note that $x' = \frac{1}{|S|} \sum_{s \in \Gamma} \pi(s).x$ is an invariant vector: for every $g \in \Gamma$,*

$$\pi(g).x' = \frac{1}{|S|} \sum_{s \in \Gamma} \pi(g)\pi(s).x = \frac{1}{|S|} \sum_{s \in \Gamma} \pi(gs).x =^{s'=gs} \frac{1}{|S|} \sum_{s' \in \Gamma} \pi(s').x = x'.$$

Thus, for every x, $\frac{1}{|S|} \sum_{s \in \Gamma} \pi(s).x = 0$ and it follows that

$$\langle \frac{1}{|S|} \sum_{s \in \Gamma} \pi(s).x, x \rangle = 0 = 0 \|x\|^2,$$

i.e., $\lambda \leq 0$ and thus $\kappa(\Gamma, S) \geq \sqrt{2}$.

Example 17.8. *The group $\Gamma = \mathbb{Z}$ does not have property (T). Consider $\ell^2(\mathbb{Z})$ with the regular representation $(\rho(n).\phi)(a) = \phi(n + a)$. This representation is unitary, since it preserves the norm and does not have a non trivial invariant vector (any invariant vector is a constant function, and the only constant function in $\ell^2(\mathbb{Z})$ is the zero function). Fix $S = \{\pm 1\}$ (recall that Γ has property (T) if and only if it has it with respect to any finite generating set). Denote $\phi_n = \frac{1}{\sqrt{n}} \chi_{\{1,\ldots,n\}}$. Then, $\|\phi_n\| = 1$, $\rho(1).\phi_n = \frac{1}{\sqrt{n}} \chi_{\{0,\ldots,n-1\}}$ and*

$$\|\rho(-1).\phi_n - \phi_n\| = \|\rho(1).\phi_n - \phi_n\| = \frac{\sqrt{2}}{\sqrt{n}}.$$

Thus, $\kappa_\rho(\Gamma, S) = 0$.

One can give an entire one semester course on Kazhdan's property (T) alone and not cover half of the essentials regarding this property. Of course, we will not do this here, but mainly use property (T) in order to construct expanders. A reader who wants to learn more about this property is referred to the book titled "Kazhdan's property (T)" by Bekka, de la Harpe and Valette.

17.1 Exercises

Throughout, Γ is a countable group and $S \subseteq \Gamma$ is a finite symmetric generating set.

1. Let $N \triangleleft \Gamma$ be a normal subgroup. Show that if Γ has property (T), then $N\backslash\Gamma$ has property (T). Moreover, show that if $S \subseteq \Gamma$ is a finite set, then $\kappa(N\backslash\Gamma, N\backslash S) \geq \kappa(\Gamma, S)$.

2. Let $(\pi, \mathcal{H})$ be a unitary representation of Γ. We denote $\mathcal{H}^{\pi(\Gamma)}$ to be the subspace of invariant vectors, i.e.,

$$\mathcal{H}^{\pi(\Gamma)} = \{x \in \mathcal{H} : \forall g \in \Gamma, \pi(g).x = x\}.$$

   Denote $P_{\mathcal{H}^{\pi(\Gamma)}}$ to be the orthogonal projection on $\mathcal{H}^{\pi(\Gamma)}$.

   (a) Show that the image of $I - P_{\mathcal{H}^{\pi(\Gamma)}}$ is a Γ-invariant subspace of $\mathcal{H}$ with no non-trivial invariant vectors.

   (b) Note that for every $x \in \mathcal{H}$, we can write $x = P_{\mathcal{H}^{\pi(\Gamma)}}x + (I - P_{\mathcal{H}^{\pi(\Gamma)}})x$. Using this decomposition, show that that for every $g \in \Gamma$ and every $x \in \mathcal{H}$,

$$x - \pi(g).x = (I - P_{\mathcal{H}^{\pi(\Gamma)}})x - \pi(g).(I - P_{\mathcal{H}^{\pi(\Gamma)}})x.$$

   (c) Deduce that for every $x \in \mathcal{H}$,

$$\max_{s \in S} \|x - \pi(s).x\| \geq \kappa(\Gamma, S)\|(I - P_{\mathcal{H}^{\pi(\Gamma)}})x\|.$$

   (d) Let $0 < \delta \leq 1$ be a constant. Show that for $x \in \mathcal{H}$, if

$$\max_{s \in S} \|x - \pi(s).x\| < \delta\kappa(\Gamma, S)\|x\|,$$

   then

$$\|(I - P_{\mathcal{H}^{\pi(\Gamma)}})x\| < \delta\|x\|,$$

   and by triangle inequality,

$$\|P_{\mathcal{H}^{\pi(\Gamma)}}x\| > (1 - \delta)\|x\|.$$

   (one can actually deduce a better inequality using the Pythagorean theorem, which states that $\|P_{\mathcal{H}^{\pi(\Gamma)}}x\|^2 + \|(I - P_{\mathcal{H}^{\pi(\Gamma)}})x\|^2 = \|x\|^2$).

3. A countable group Γ is called *amenable* if there is a sequence of finite sets $\{F_n \subseteq \Gamma\}_n$ such that for every $g \in \Gamma$, $\lim_n \frac{|F_n \triangle g.F_n|}{|F_n|} = 0$ (for sets A, B, $A \triangle B$ denotes the symmetric difference: $A \triangle B = (A \setminus B) \cup (B \setminus A)$).

   (a) Show that every finite group is amenable.

   (b) Show that if Γ is infinite and amenable, then it does not have property (T). (Hint: consider the indicator functions $\chi_{F_n} \in \ell^2(\Gamma)$ and show that for every $s \in S$, $\frac{\|\chi_{F_n} - \rho(s)\chi_{F_n}\|}{\|\chi_{F_n}\|}$ tends to 0).

   (c) Show for a finitely generated group Γ with a finite generating set S, Γ is amenable if and only if there is a sequence of finite sets $\{F_n \subseteq \Gamma\}_n$ such that for every $s \in S$, $\lim_n \frac{|F_n \triangle s.F_n|}{|F_n|} = 0$.

(d) Show that for every k, $\mathbb{Z}^k$ is amenable.

4. Let Γ be a group Γ with a finite generating set S. For $n \in \mathbb{N}$ denote $B_n(e)$ to be the ball of radius $\leq n$ in $\mathrm{Cay}(\Gamma; S)$ centred at e (e is the identity of Γ). A group Γ is said to have exponential growth if there is a constant $a > 1$ such that $|B_n(e)| \geq a^n$ for every n. Show that if Γ is infinite and has property (T), then it has exponential growth (Hint: you can either show this directly, by taking the definition of Kazhdan constant and applying it on indicator functions of balls or show more and prove that if a group does not have exponential growth it is amenable and then use the previous exercise).

18 Elementary matrix groups over polynomial rings

Our aim for the rest of the course is to show that elementary matrix groups over $\mathbb{F}_p[t]$ defined below have property (T) (when p is large enough) and can be used to explicitly construct expanders.

Notation: for a prime p, $\mathbb{F}_p$ denotes the field with p elements, i.e., the set $\{0, ..., p-1\}$ with addition and multiplication modulo p. Also, $\mathbb{F}_p[t]$ denotes polynomials with coefficients in $\mathbb{F}_p$, i.e., expressions of the form $a_0 + a_1 t + ... + a_n t^n$, where $a_0, ..., a_n \in \mathbb{F}_p$.

We will be interested in the group (generated by) 3×3 elementary matrices over $\mathbb{F}_p[t]$. Given $1 \leq i, j \leq 3$, $i \neq j$, and $r \in \mathbb{F}_p[t]$, denote $e_{i,j}(r)$ to be the 3×3 matrix with 1's along the main diagonal, r in the i-th row and j-th column and 0 in all the other entries. For example,

$$e_{1,2}(2 + t^2) = \begin{pmatrix} 1 & 2 + t^2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Matrix of the form $e_{i,j}(r)$ is called an *elementary matrix*.

Define $\Gamma = \mathrm{EL}_3(\mathbb{F}_p[t])$ to be the group generated by all elementary matrices with entries in $\mathbb{F}_p[t]$, i.e., $\mathrm{EL}_3(\mathbb{F}_p[t])$ is the group generated by the matrices of the form

$$\{e_{i,j}(r) : 1 \leq i, j \leq 3, i \neq j, r \in \mathbb{F}_p[t]\}.$$

We note that this is in fact a group: matrix multiplication is associative, the identity matrix I is the trivial element and for every i, j, r as above, $e_{i,j}(r)e_{i,j}(-r) = I$ and thus any product of elementary matrices has an inverse.

We note that Γ is infinite (since $\mathbb{F}_p[t]$ is infinite) and our goal is to show that we can apply the "mother group approach" to expanders on Γ and use it to produce an infinite family of expander graphs. Namely, we need to show two things:

Theorem 18.1. *Let $\Gamma = \mathrm{EL}_3(\mathbb{F}_p[t])$. Then the following holds:*

1. *There is an infinite sequence $N_n \lhd \Gamma$ such that $|N_n \backslash \Gamma| < \infty$ and $\lim_n |N_n \backslash \Gamma| = \infty$.*

2. *If $p > 49$, then the group Γ has property (T).*

Consequently: For any generating set S of Γ, the family $\mathrm{Cay}(N_n\backslash\Gamma; N_n\backslash S)$ is an expander family.

The easy part is showing that 1. above is satisfied:

Proof that the of elementary matrices has "many" finite index normal subgroups. For $n \in \mathbb{N}$, denote $\mathbb{F}_p[t]/\langle t^n\rangle$ to be polynomials in $\mathbb{F}_p[t]$ with degree $< n$ such that multiplication is taken modulo t^n, i.e., by eliminating all the summands of degree $\geq n$. For example, let $1+t+t^2 \in \mathbb{F}_{11}[t]/\langle t^3\rangle$, then

$$(1+t+t^2)(1+t+t^2) = 1+2t+2t^2.$$

The group $\Gamma_n = \mathrm{EL}_3(\mathbb{F}_p[t]/\langle t^n\rangle)$ of matrices generated by

$$\{e_{i,j}(r) : 1 \leq i,j \leq 3, i \neq j, r \in \mathbb{F}_p[t]/\langle t^n\rangle\},$$

is a finite group and there is an obvious homomorphism $\Phi_n : \mathrm{EL}_3(\mathbb{F}_p[t]) \to \mathrm{EL}_3(\mathbb{F}_p[t]/\langle t^n\rangle)$, where every entry of a matrix in $\mathrm{EL}_3(\mathbb{F}_p[t])$ is taken modulo t^n. Thus, if we denote $\ker(\Phi_n) = N_n$ we get a a finite index subgroup of $\mathrm{EL}_3(\mathbb{F}_p[t])$ and $|N_n\backslash\mathrm{EL}_3(\mathbb{F}_p[t])| = |\mathrm{EL}_3(\mathbb{F}_p[t]/\langle t^n\rangle)|$ grows to infinity with n (since $|\mathbb{F}_p[t]/\langle t^n\rangle|$ grows to infinity with n).

Therefore if we show that Γ has property (T), then for a fixed generating set S of Γ, $\mathrm{Cay}(\Gamma_n; N_n\backslash S)$ will be an expander family. □

We note that above we actually showed more than needed as explained below:

Definition 18.2. *A group Γ is called residually finite if there is a sequence of finite index normal subgroups $\{N_n \lhd \Gamma\}_{n\in\mathbb{N}}$ such that $N_{n+1} \lhd N_n$ and $\bigcap_n N_n = \{e\}$.*

Observation 18.3. *Above we showed that $\mathrm{EL}(\mathbb{F}_p[t])$ is residually finite.*

Proposition 18.4. *Let Γ be a countable group with a finite generating set S. Show that if Γ is residually finite, then for every $R \in \mathbb{N}$, there is a finite group Γ_R and a homomorphism $\Phi : \Gamma \to \Gamma_R$ that induces an isomorphism of graphs between the ball of radius R centred at e in $\mathrm{Cay}(\Gamma, S)$ and the ball of radius R center at e in $\mathrm{Cay}(\Gamma_R, \Phi(S))$.*

Proof. Fix some $R \in \mathbb{N}$. Denote dist_Γ to be the distance in $\mathrm{Cay}(\Gamma; S)$. By the definition of residual finiteness, there is a finite index normal subgroup $N \lhd \Gamma$ such that for every $g \in \Gamma, g \neq e$ if $\mathrm{dist}_\Gamma(e,g) \leq 2R$, then $g \notin N$. Fix such a normal subgroup and take $\Gamma_R = N\backslash\Gamma$. Define $\Phi : \Gamma \to \Gamma_R$ to be $\Phi(g) = Ng$.

First, we will show that Φ is injective when restricted to the ball of radius R in $\mathrm{Cay}(\Gamma; S)$. Let $g_1, g_2 \in \Gamma$ such that $\mathrm{dist}_\Gamma(e, g_1) \leq R$, $\mathrm{dist}_\Gamma(e, g_2) \leq R$ and $\Phi(g_1) = \Phi(g_2)$. Then $Ng_1 = Ng_2$, i.e., there is $g \in N$ such that $gg_1 = g_2$. We will show that in that case $g = e$. Assume toward contradiction that $g \neq e$, then $\mathrm{dist}_\Gamma(e, g) > 2R+1$, and by triangle inequality,

$$\begin{aligned} 2R+1 &\leq \mathrm{dist}_\Gamma(e,g) \\ &= \mathrm{dist}_\Gamma(e, gg_1) + \mathrm{dist}_\Gamma(gg_1, g) \\ &= \mathrm{dist}_\Gamma(e, gg_1) + \mathrm{dist}_\Gamma(g_1, e) \\ &\leq \mathrm{dist}_\Gamma(e, gg_1) + R \end{aligned}$$

This implies that $R \geq \text{dist}_\Gamma(e, g_2) = \text{dist}_\Gamma(e, gg_1) \geq R+1$ which is a contradiction.

Second, we note that for every g_1, g_2 in the ball of radius R in $\text{Cay}(\Gamma; S)$, g_1 and g_2 are connected by an edge if and only if $g_2 = g_1 s$ and by the injectivity we showed above this holds if and only if $Ng_2 = Ng_1 s$, i.e., if and only if $\Phi(g_1)$ and $\Phi(g_2)$ are connected by an edge in $\text{Cay}(\Gamma_R, \Phi(S))$ □

The proof that $\text{EL}_3(\mathbb{F}_p[t])$ has property (T) will require much more work, namely we will show the following:

1. The group $\text{EL}_3(\mathbb{F}_p[t])$ is generated by 4 finite subgroups.
2. There is a method to prove property (T) for groups generated by finite subgroups, using the notion of the angle projections.
3. Applying the above method for $\text{EL}_3(\mathbb{F}_p[t])$ show that it has property (T) given that p is large enough.

19 The Steinberg relations and generation by finite subgroups for the elementary matrix group

The group $\text{EL}_3(\mathbb{F}_p[t])$ has the following relations, called the Steinberg relations:

1. For every $1 \leq i, j \leq 3$, $i \neq j$ and every $r_1, r_2 \in \mathbb{F}_p[t]$, $e_{i,j}(r_1)e_{i,j}(r_2) = e_{i,j}(r_1 + r_2)$.
2. For every $1 \leq i, j \leq 3$, $i \neq j, j \neq k, k \neq i$ and every $r_1, r_2 \in \mathbb{F}_p[t]$, $[e_{i,j}(r_1), e_{j,k}(r_2)] = e_{i,k}(r_1 r_2)$ ($[.,.]$ is the commutator: $[g,h] = ghg^{-1}h^{-1}$).

Next, we will note an important fact:

Proposition 19.1. $\Gamma = \text{EL}_3(\mathbb{F}_p[t])$ *is generated by 4 finite subgroups: denote*

$$K_1 = \{e_{1,2}(at) : a \in \mathbb{F}_p\},$$

$$K_2 = \{e_{1,2}(a) : a \in \mathbb{F}_p\},$$

$$K_3 = \{e_{2,3}(a) : a \in \mathbb{F}_p\},$$

$$K_4 = \{e_{3,1}(a) : a \in \mathbb{F}_p\}.$$

Observe that these are all subgroups of size p. We claim that $K_1 \cup ... \cup K_4$ *is a generating set of* Γ.

Proof. First, by Steinberg relations, the subgroup generated by K_2, K_3, K_4 contains all matrices of the form $e_{i,j}(a)$, $a \in \mathbb{F}_p$. For example: $e_{1,3}(a) = [e_{1,2}(a), e_{2,3}(1)]$.

Second, using the first step, the subgroup generated by K_1, K_2, K_3, K_4 contains all matrices of the form $e_{i,j}(a + bt)$. Indeed,

$$e_{1,3}(bt) = [e_{1,2}(bt), e_{2,3}(1)],$$

$$e_{3,2}(bt) = [e_{3,1}(1), e_{1,2}(bt)],$$

$$e_{2,3}(bt) = [e_{2,1}(1), e_{1,3}(bt)],$$
$$e_{3,1}(bt) = [e_{3,2}(bt), e_{2,1}(1)],$$
$$e_{2,1}(bt) = [e_{2,3}(bt), e_{3,1}(1)].$$

Last, we will show that for every $n \in \mathbb{N}$, the group generated by $K_1, ..., K_4$ contains $e_{i,j}(a_0 + ... + a_n t^n)$ for every $i \neq j$ and $a_0, ..., a_n \in \mathbb{F}_p$. The proof is by induction: we proved $n = 1$ above. Assume that the assertion is true for n. It is enough to show that for every $e_{i,j}(bt^{n+1})$ for every $i \neq j$ and every $b \in \mathbb{F}_p$. For every $i \neq j$, we pick $k \neq i, k \neq j$ and by the Steinberg relations:

$$e_{i,j}(bt^{n+1}) = [e_{i,k}(bt), e_{k,j}(t^n)],$$

where $e_{i,k}(bt), e_{k,j}(t^n)$ are in the group generated by $K_1, ..., K_4$ by the induction assumption. □

20 Criterion for property (T) for a group generated by finite subgroups

Proposition 20.1. *Let K be a finite group and let $(\pi, \mathcal{H})$ be a unitary representation of K. Denote $\mathcal{H}^{\pi(K)}$ to be the space of K-invariant vectors, i.e.,*

$$\mathcal{H}^{\pi(K)} = \{x \in \mathcal{H} : \forall g \in K, \pi(g).x = x\}.$$

Define $P^\pi = \frac{1}{|K|} \sum_{g \in K} \pi(g)$, then P^π is the orthogonal projection on $\mathcal{H}^{\pi(K)}$.

Proof. To avoid cumbersome notation, we denote $P^\pi = P$.

Note that for every $g' \in K$,

$$\pi(g')P = \frac{1}{|K|} \sum_{g \in K} \pi(g'g) = \frac{1}{|K|} \sum_{g \in K} \pi(g) = P.$$

This implies two things, first,

$$P^2 = \frac{1}{|K|} \sum_{g' \in K} \pi(g')P = P,$$

and second, for every $x \in \mathcal{H}$ and every $g' \in K$,

$$\pi(g')Px = Px,$$

i.e., $\mathrm{Im}(P) \subseteq \mathcal{H}^{\pi(K)}$.

Also note that for every $x \in \mathcal{H}^{\pi(K)}$,

$$Px = \frac{1}{|K|} \sum_{g \in K} \pi(g).x = \frac{1}{|K|} \sum_{g \in K} x = x,$$

i.e., $\mathrm{Im}(P) = \mathcal{H}^{\pi(K)}$.

Last, we note that π is a unitary representation and therefore $\pi(g)^* = \pi(g^{-1})$. This implies that

$$P^* = \frac{1}{|K|} \sum_{g \in K} \pi(g)^* = \frac{1}{|K|} \sum_{g \in K} \pi(g^{-1}) = P,$$

as needed. □

Let Γ be a countable group and let $K_1, ..., K_n$ be finite subgroups such that $K_1 \cup ... \cup K_n$ is a generating set of Γ. Let $(\pi, \mathcal{H})$ to be a unitary representation. Define $P_i^\pi = \frac{1}{|K_i|} \sum_{g \in K_i} \pi(g)$. By the above proposition, P_i^π in the orthogonal projection on $\mathcal{H}^{\pi(K_i)}$.

Observation 20.2. *In the setting above, $\bigcap_i \mathcal{H}^{\pi(K_i)}$ is the subspace of Γ-invariant vectors. In particular, if $(\pi, \mathcal{H})$ is a unitary representation without a non-trivial invariant vector, it follows that $\bigcap_i \mathcal{H}^{\pi(K_i)} = \{0\}$.*

Indeed, let $x \in \bigcap_i \mathcal{H}^{\pi(K_i)}$. For every $g \in K_1 \cup ... \cup K_n$, $\pi(g).x = x$. By our assumption, $K_1 \cup ... \cup K_n$ is a generating set and therefore for every $g \in \Gamma$, $\pi(g).x = x$.

Proposition 20.3. *Let $\Gamma, K_1, ..., K_n$ as above. In order to prove that Γ has property (T), it is sufficient to prove that there is a constant $k \in \mathbb{N}$ such that for every unitary representation $(\pi, \mathcal{H})$ without a non-trivial invariant vector, it holds that*

$$\left\| \left(\frac{P_1^\pi + ... + P_n^\pi}{n} \right)^k \right\| \leq \frac{1}{2}.$$

Proof. Assume that the conditions of the claim holds, then for every unitary representation $(\pi, \mathcal{H})$ without a non-trivial invariant vector it holds for every $x \in \mathcal{H}$. Denote $T = \frac{P_1^\pi + ... + P_n^\pi}{n}$. By the assumption,

$$\begin{aligned} \left\| \left(I - T^k\right) x \right\| &\geq \|x\| - \left\| T^k x \right\| \\ &\geq (1 - \frac{1}{2}) \|x\|. \end{aligned}$$

Thus,

$$\begin{aligned} &\frac{1}{2}\|x\| \leq \left\| \left(I - T^k\right) x \right\| \leq \left\| \left(I + T + ... + T^{k-1}\right) (I - T) x \right\| \leq \\ &(\|I\| + \|T\| + ... + \|T\|^{k-1}) \|(I - T)x\|. \end{aligned}$$

Observe that $\|T\| \leq \frac{1}{n} \sum_i \|P_i^\pi\| \leq 1$ (an orthogonal projection has norm ≤ 1). Thus,

$$\frac{1}{2(k+1)} \|x\| \leq \|(I - T)x\|.$$

Therefore, the definition of T and the triangle inequality yields

$$\begin{aligned} &\frac{1}{2(k+1)} \|x\| \leq \frac{1}{n} \sum_{i=1}^n \|(I - P_i^\pi)x\| = \\ &\frac{1}{n} \sum_{i=1}^n \| \frac{1}{|K_i|} \sum_{g \in K_i} (I - \pi(g)).x \| \leq \\ &\frac{1}{n} \sum_{i=1}^n \frac{1}{|K_i|} \sum_{g \in K_i} \|x - \pi(g).x\| \leq \max_{g \in K_1 \cup ... \cup K_n} \|x - \pi(g).x\|. \end{aligned}$$

Thus, $\kappa(\Gamma, K_1 \cup ... \cup K_n) \geq \frac{1}{2(k+1)} > 0$ as needed. □

This leads us to a general question in functional analysis: Let $P_1, ..., P_n$ be orthogonal projections such that $\mathrm{Im}(P_1) \cap ... \cap \mathrm{Im}(P_n) = \{0\}$. How can we bound the norm of $\| \left(\frac{P_1 + ... + P_n}{n} \right)^k \|$?

It turns out that this can be done using the notion of angles between projections:

Definition 20.4 (Angle between projections)**.** *Let $\mathcal{H}$ be a Hilbert space and P_1, P_2 be orthogonal projections. Denote $P_{1,2}$ to be the orthogonal projection on $\mathrm{Im}(P_1) \cap \mathrm{Im}(P_2)$ and define*

$$\cos(\angle(P_1, P_2)) = \|P_1 P_2 - P_{1,2}\|.$$

Remark 20.5. *Recall that $\|T^*\| = \|T\|$. Then*

$$\|P_1 P_2 - P_{1,2}\| = \|(P_1 P_2 - P_{1,2})^*\| = \|P_2^* P_1^* - P_{1,2}^*\| = \|P_2 P_1 - P_{1,2}\|.$$

Thus, $\cos(\angle(P_1, P_2)) = \cos(\angle(P_2, P_1))$.

Using this notion, one can prove the following theorem:

Theorem 20.6. *Let $\mathcal{H}$ be a Hilbert space and let $P_1, ..., P_n$ be orthogonal projections such that there is some $0 \leq \varepsilon < \frac{1}{2n-1}$ such that for every $1 \leq i < j \leq n$, $\cos(\angle(P_i, P_j)) \leq \varepsilon$, then there are constants $C > 0$, $r < 1$ that depend only on ε such that for every k, $\| \left(\frac{P_1 + ... + P_n}{n} \right)^k \| \leq C r^k$.*

In particular, there is some k that depends only on ε such that $\| \left(\frac{P_1 + ... + P_n}{n} \right)^k \| \leq \frac{1}{2}$.

Remark 20.7. *This theorem is not sharp - one can actually demand $\varepsilon < \frac{1}{n-1}$ and get the same result, but the proof becomes more complicated. In the sake of completeness, we only use the non-sharp version in these notes.*

We delay the proof of this Theorem (it is given in an appendix) for later and for now see how we apply it.

Combining Proposition 20.3 and Theorem 20.6 we get the following criterion for property (T):

Theorem 20.8. *Let Γ be a group generated by finite subgroups $K_1, ..., K_n$. Given a representation $(\pi, \mathcal{H})$, denote as above: $P_i^\pi = \frac{1}{|K_i|} \sum_{g \in K_i} \pi(g)$.*

If there is $\varepsilon < \frac{1}{2n-1}$ such that for every unitary representation $(\pi, \mathcal{H})$ without a non-trivial invariant vector, it holds for every i, j that

$$\cos(\angle(P_i^\pi, P_j^\pi)) \leq \varepsilon,$$

then Γ has property (T).

21 Property (T) for $\mathrm{EL}_3(\mathbb{F}_p[t])$

In order to apply this criterion, we need a way to bound $\cos(\angle(P_i^\pi, P_j^\pi))$ for all unitary representations. In the example that interests us, this will be done analysing representation of a specific finite group.

Recall that in the example of $\Gamma = \mathrm{EL}_3(\mathbb{F}_p[t])$, the generating subgroups were

$$K_1 = \{e_{1,2}(at) : a \in \mathbb{F}_p\},$$

$$K_2 = \{e_{1,2}(a) : a \in \mathbb{F}_p\},$$
$$K_3 = \{e_{2,3}(a) : a \in \mathbb{F}_p\},$$
$$K_4 = \{e_{3,1}(a) : a \in \mathbb{F}_p\}.$$

We note that in this example not only $K_1, ..., K_4$ are finite, but each subgroup generated by 2 of these subgroups are finite. Denote $K_{i,j}$ to be the subgroup generated by K_i and K_j and note that by a previous observation,

$$\mathcal{H}^{\pi(K_{i,j})} = \mathcal{H}^{\pi(K_i)} \cap \mathcal{H}^{\pi(K_j)}.$$

Thus $P^\pi_{i,j}$ is given by

$$\frac{1}{|K_{i,j}|} \sum_{g \in K_{i,j}} \pi(g).$$

Observation 21.1. *In the example of* $\Gamma = \mathrm{EL}_3(\mathbb{F}_p[t])$ *with the subgroups above,* $\cos(\angle(P^\pi_1, P^\pi_2)) = 0$ *for every unitary representation. Indeed,*

$$K_{1,2} = \{e_{1,2}(a + bt) : a, b \in \mathbb{F}_p\},$$

and

$$P^\pi_2 P^\pi_1 = \frac{1}{p^2} \sum_{a,b \in \mathbb{F}_p} \pi(e_{1,2}(a + bt)) = P^\pi_{1,2}.$$

Another way to prove the same result, is noting that K_1 *and* K_2 *commute and therefore* $P^\pi_2 P^\pi_1 = P^\pi_1 P^\pi_2$*. Observe that* $\mathrm{Im}(P^\pi_2 P^\pi_1) \subseteq \mathrm{Im}(P^\pi_2)$ *and* $\mathrm{Im}(P^\pi_1 P^\pi_2) \subseteq \mathrm{Im}(P^\pi_1)$*, but since the two are equal, we get that*

$$\mathrm{Im}(P^\pi_1 P^\pi_2) \subseteq \mathrm{Im}(P^\pi_1) \cap \mathrm{Im}(P^\pi_2) = \mathrm{Im}(P^\pi_{1,2}).$$

Proposition 21.2. *In the example of* $\Gamma = \mathrm{EL}_3(\mathbb{F}_p[t])$ *with the subgroups above, for every* $2 \leq i < j \leq 4$ *and every unitary representation* $(\pi, \mathcal{H})$*,*

$$\cos(\angle(P^\pi_i, P^\pi_j)) \leq \frac{1}{\sqrt{p}}.$$

Proof. We will give the proof for $i = 2, j = 3$, the proofs in the other cases are similar. Let $(\pi, \mathcal{H})$ be some unitary representation. We need to show that for every unit vector $x \in \mathcal{H}$,

$$\|(P^\pi_2 P^\pi_3 - P^\pi_{2,3})x\|^2 \leq \frac{1}{p}.$$

Fix $x \in \mathcal{H}$ to be some unit vector, then

$$\begin{aligned}\|(P^\pi_2 P^\pi_3 - P^\pi_{2,3})x\|^2 &= \langle (P^\pi_2 P^\pi_3 - P^\pi_{2,3})x, (P^\pi_2 P^\pi_3) - P^\pi_{2,3})x\rangle = \\ &\langle (P^\pi_3 P^\pi_2 P^\pi_3 - P^\pi_{2,3})x, (I - P^\pi_{2,3})x\rangle \leq \\ &\|(P^\pi_3 P^\pi_2 P^\pi_3 - P^\pi_{2,3})x\|\|(I - P^\pi_{2,3})x\| \leq \|P^\pi_3 P^\pi_2 P^\pi_3 - P^\pi_{2,3}\|.\end{aligned}$$

Thus, it is sufficient to prove that

$$\|(P^\pi_3 P^\pi_2 P^\pi_3 - P^\pi_{2,3})\| \leq \frac{1}{p}.$$

This can be done via direct computation:

$$P_{2,3}^{\pi} = \frac{1}{p^3} \sum_{a,b,c \in \mathbb{F}_p} \pi \left(\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \right).$$

$$\begin{aligned} P_3^{\pi} P_2^{\pi} P_3^{\pi} &= \frac{1}{p^3} \sum_{a,b,b' \in \mathbb{F}_p} \pi \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \right) \\ &= \frac{1}{p^3} \sum_{a,b,b' \in \mathbb{F}_p} \pi \left(\begin{pmatrix} 1 & a & ab \\ 0 & 1 & b+b' \\ 0 & 0 & 1 \end{pmatrix} \right) \\ &= \frac{1}{p^3} \sum_{a,b,c \in \mathbb{F}_p, a \neq 0} \pi \left(\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \right) + \frac{1}{p^2} \sum_{b \in \mathbb{F}_p} \pi \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \right). \end{aligned}$$

Thus,

$$\begin{aligned} P_3^{\pi} P_2^{\pi} P_3^{\pi} - P_{2,3}^{\pi} &= \frac{1}{p} \left(\frac{1}{p} \sum_{b \in \mathbb{F}_p} \pi \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \right) \right) \left(I \right. \\ &\qquad \left. - \frac{1}{p} \sum_{c \in \mathbb{F}_p} \pi \left(\begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \right) \\ &= \frac{1}{p} P_3^{\pi} (I - \frac{1}{p} \sum_{c \in \mathbb{F}_p} \pi(e_{1,3}(c))). \end{aligned}$$

Note that P_3^{π} is an orthogonal projection and therefore $\|P_3^{\pi}\| \leq 1$. Also note that $\frac{1}{p} \sum_{c \in \mathbb{F}_p} \pi(e_{1,3}(c))$ is an orthogonal projection and therefore $I - \frac{1}{p} \sum_{c \in \mathbb{F}_p} \pi(e_{1,3}(c))$ is an orthogonal projection and thus of norm ≤ 1. It follows that

$$\|P_3^{\pi} P_2^{\pi} P_3^{\pi} - P_{2,3}^{\pi}\| \leq \frac{1}{p},$$

as needed. □

Corollary 21.3. *Let $p > 49$ be a prime number and let $\Gamma = \mathrm{EL}_3(\mathbb{F}_p[t])$, then Γ has property (T).*

Proof. Let $K_1, ..., K_4$, be the subgroups of Γ defined above. Then for every unitary representation $(\pi, \mathcal{H})$ of Γ it follows that for every $1 \leq i < j \leq 4$,

$$\cos(\angle(P_i^{\pi}, P_j^{\pi})) \leq \frac{1}{\sqrt{p}} < \frac{1}{7},$$

and therefore by Theorem 20.8, Γ has property (T). □

Remark 21.4. *In fact, the corollary above can be improved such that it is enough to ask that $p \geq 5$. This is done in the following way: First, using a sharp version of Theorem 20.6, that only requires that the cosines as $< \frac{1}{n-1}$. Second, defining*

$$K_1' = \{e_{1,2}(a + bt) : a, b \in \mathbb{F}_p\},$$

and using K_1', K_3, K_4. Then $n = 3$ and one can prove (using representation theory of the subgroups $K_{1,3}', K_{1,4}'$ that the cosines are $\leq \frac{1}{\sqrt{p}}$. Then we need that $\frac{1}{\sqrt{p}} < \frac{1}{2}$ and it is enough to take $p \geq 5$.

A Complex numbers

The aim of this appendix is to cover very basic facts regarding complex numbers. A complex number is a number of the form $a+ib$, where $a, b \in \mathbb{R}$ and $i^2 = -1$. Thus, we can add and multiply two complex numbers:

$$(a+ib)+(c+id) = (a+c)+i(b+d),$$

$$(a+ib)(c+id) = ac-bd+i(bc+ad).$$

We denote by $\mathbb{C}$ the set of the complex number and note that under the definitions of addition and multiplication defined above, this is a field.

The norm on $\mathbb{C}$ is the function $|a+ib| = \sqrt{a^2+b^2}$ and with this norm we can think of $\mathbb{C}$ as a norm space isometric to $\mathbb{R}^2$ with the euclidean norm. Thus, we think about $(a+ib)$ as a vector written in Cartesian coordinates on two axes - the real axis and the imaginary axis.

As in the case of polar coordinates in $\mathbb{R}^2$, every complex number can be written as

$$(a+ib) = r(\cos(\theta)+i\sin(\theta)),$$

where $0 \leq r < \infty$ equals to the norm of $a+ib$ and θ is the angle between $(a+ib)$ and the real axis. We note that for every $m \in \mathbb{Z}$,

$$r(\cos(\theta+2\pi m)+i\sin(\theta+2\pi m)) = r(\cos(\theta)+i\sin(\theta)).$$

We also norm that when multiplying, we get the following:

$$\begin{aligned}&(r_1(\cos(\theta_1)+i\sin(\theta_1)))(r_2(\cos(\theta_2)+i\sin(\theta_2)))\\ &\quad = r_1r_2((\cos(\theta_1)\cos(\theta_2)-\sin(\theta_1)\sin(\theta_2))+i(\sin(\theta_1)\cos(\theta_2)+\sin(\theta_2)\cos(\theta_1)))\\ &\quad = r_1r_2(\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2)).\end{aligned}$$

We define $e^{i\theta} = \cos(\theta)+i\sin(\theta)$ and note that by the computation above, $e^{i\theta_1}e^{i\theta_2} = e^{i(\theta_1+\theta_2)}$. Moreover, for every $l \in \mathbb{R}$, we define $e^{l+i\theta} = e^l e^{i\theta}$ and note that for every $r > 0$,

$$r(\cos(\theta)+i\sin(\theta)) = e^{\ln(r)+i\theta}.$$

The reason that we work with $\mathbb{C}$ when studying group representations is that $\mathbb{C}$ is an algebraically closed field:

Theorem A.1 (Fundamental Theorem of Algebra)**.** *Let $p(z) = a_nz^n + ... + a_1z + a_0$ be a polynomial, where $n \in \mathbb{N}$ and $a_n, ..., a_0 \in \mathbb{C}, a_n \neq 0$, then $p(z)$ has a root $\alpha \in \mathbb{C}$, i.e., there is $\alpha \in \mathbb{C}$ such that $p(\alpha) = 0$. Consequently, there are $\alpha_1, ..., \alpha_n \in \mathbb{C}$ (not necessarily distinct) such that*

$$p(z) = a_n(z-\alpha_1)(z-\alpha_2)...(z-\alpha_n).$$

Corollary A.2. *Every $n \times n$ matrix A with complex entries has at least one complex eigenvalue.*

Proof. The characteristic polynomial $\det(A-zI)$ has at least one complex root. □

Remark A.3. *This corollary is false for $\mathbb{R}$, e.g.,*

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

has no real eigenvalue.

Last, by the Fundamental Theorem of Algebra, the equation $z^n = 1$ has n solutions (when counting with multiplicity) and these solutions are called the *n-roots of unity*. These solutions can be written explicitly as $\{e^{\frac{2\pi ij}{n}} : 0 \leq j < n\}$ (note that we see that the multiplicity of every root is in fact 1).

B Angle criterion for convergence of averaged projections

Definition B.1 (Angle between projections)**.** *Let $\mathcal{H}$ be a Hilbert space and P_1, P_2 be orthogonal projections. Denote $P_{1,2}$ to be the orthogonal projection on $\mathrm{Im}(P_1) \cap \mathrm{Im}(P_2)$ and define*

$$\cos(\angle(P_1, P_2)) = \|P_1 P_2 - P_{1,2}\|.$$

Remark B.2. *Recall that $\|T^*\| = \|T\|$. Then*

$$\|P_1 P_2 - P_{1,2}\| = \|(P_1 P_2 - P_{1,2})^*\| = \|P_2^* P_1^* - P_{1,2}^*\| = \|P_2 P_1 - P_{1,2}\|.$$

Thus, $\cos(\angle(P_1, P_2)) = \cos(\angle(P_2, P_1))$.

We will show that for n orthogonal projections $P_1, ..., P_n$, if $\cos(\angle(P_i, P_j))$ is small enough for every i, j, then $\|\left(\frac{P_1+...+P_n}{n}\right)^k\|$ decays to 0 exponentially fast. In order to do that we will need some preliminary results.

A basic fact that we will use below (without proof) is the following: given a Hilbert space $\mathcal{H}$ with closed subspaces $\mathcal{H}'' \subseteq \mathcal{H}'$, then

$$P_{\mathcal{H}''} P_{\mathcal{H}'} = P_{\mathcal{H}''} = P_{\mathcal{H}'} P_{\mathcal{H}''}.$$

Thus, in the notations above, for any two projections P_1, P_2, $P_1 P_{1,2} = P_{1,2} P_1 = P_{1,2}$ and similarly, $P_2 P_{1,2} = P_{1,2} P_2 = P_{1,2}$. It follows that $P_1 - P_{1,2}$ and $P_2 - P_{1,2}$ are both orthogonal projections:

$$(P_1 - P_{1,2})^* = P_1^* - P_{1,2}^* = P_1 - P_{1,2},$$

$$(P_1 - P_{1,2})^2 = P_1^2 - P_1 P_{1,2} - P_{1,2} P_1 + P_{1,2}^2 = P_1 - 2P_{1,2} + P_{1,2} = P_1 - P_{1,2}.$$

It also follows that

$$P_1 P_2 - P_{1,2} = P_1 P_2 - P_1 P_{1,2} = P_1(P_2 - P_{1,2}) = P_1(P_2 - P_{1,2})^2 = (P_1 P_2 - P_{1,2})(P_2 - P_{1,2}).$$

and similarly,

$$P_2 P_1 - P_{1,2} = (P_2 P_1 - P_{1,2})(P_1 - P_{1,2}).$$

Lemma B.3. *Let $\mathcal{H}$ be a Hilbert space, P_1, P_2 be orthogonal projections. Then for every $x \in \mathcal{H}$,*

$$\|(P_1 P_2 - P_2 P_1)x\| \leq \frac{\cos(\angle(P_1, P_2))}{1 - \cos(\angle(P_1, P_2))} \left(\|(I - P_1)x\| + \|(I - P_2)x\|\right).$$

Proof. We start by observing that

$$\begin{aligned}&\|(P_1P_2 - P_2P_1)x\| \leq \\ &\quad \|(P_1P_2 - P_{1,2})x + (P_{1,2} - P_2P_1)x\| \leq \\ &\quad \|(P_1P_2 - P_{1,2})x\| + \|(P_2P_1 - P_{1,2})x\| = \\ &\quad \|(P_1P_2 - P_{1,2})(P_2 - P_{1,2})x\| + \|(P_2P_1 - P_{1,2})(P_1 - P_{1,2})x\| \leq \\ &\quad \cos(\angle(P_1, P_2))(\|(P_2 - P_{1,2})x\| + \|(P_1 - P_{1,2})x\|).\end{aligned}$$

Thus, we are left to show that

$$\frac{\|(P_1 - P_{1,2})x\| + \|(P_2 - P_{1,2})x\|}{1 - \cos(\angle(P_1, P_2))} \leq \|(I - P_1)x\| + \|(I - P_2)x\|.$$

We note that

$$\begin{aligned}&\|(P_2 - P_{1,2})x\| \leq \\ &\quad \|(P_2 - P_{1,2})P_1x\| + \|(P_2 - P_{1,2})(I - P_1)x\| \leq \\ &\quad \|(P_2P_1 - P_{1,2})x\| + \|P_2 - P_{1,2}\|\|(I - P_1)x\| \leq \\ &\quad \|(P_2P_1 - P_{1,2})(P_1 - P_{1,2})x\| + \|(I - P_1)x\| \leq \\ &\quad \cos(\angle(P_1, P_2))\|(P_1 - P_{1,2})x\| + \|(I - P_1)x\|.\end{aligned}$$

Similarly,

$$\begin{aligned}&\|(P_1 - P_{1,2})x\| \leq \\ &\quad \cos(\angle(P_1, P_2))\|(P_2 - P_{1,2})x\| + \|(I - P_1)x\|.\end{aligned}$$

Adding these two inequalities yields

$$\begin{aligned}&\|(P_1 - P_{1,2})x\| + \|(P_2 - P_{1,2})x\| \leq \\ &\quad \cos(\angle(P_1, P_2))(\|(P_1 - P_{1,2})x\| + \|(P_2 - P_{1,2})x\|) + \|(I - P_1)x\| + \|(I - P_2)x\|.\end{aligned}$$

Thus,

$$\frac{\|(P_1 - P_{1,2})x\| + \|(P_2 - P_{1,2})x\|}{1 - \cos(\angle(P_1, P_2))} \leq \|(I - P_1)x\| + \|(I - P_2)x\|,$$

as needed. □

Let $\mathcal{H}$ be a Hilbert space and let $P_1, ..., P_n$ be orthogonal projections. Define the following energy function on $\mathcal{H}$:

$$E(x) = \frac{1}{n}\sum_{i=1}^{n}\|(I - P_i)x\|.$$

Note that by definition, $E(x) \geq 0$ for every x. Also note that since $I - P_1, ..., I - P_n$ are orthogonal projections (and as such of norm ≤ 1), it follows that for every x, $E(x) \leq \|x\|$.

Proposition B.4. *If* $\mathrm{Im}(P_1) \cap ... \cap \mathrm{Im}(P_n) = \{0\}$*, then* $E(x) = 0$ *implies that* $x = 0$.

Proof. Assume that $E(x) = 0$, then for every i, $(I - P_i)x = 0$ or equivalently $x = P_ix$. This means that $x \in \mathrm{Im}(P_i)$ for every i and thus

$$x \in \mathrm{Im}(P_1) \cap ... \cap \mathrm{Im}(P_n) = \{0\},$$

as needed. □

Lemma B.5. *Let $\mathcal{H}$ be a Hilbert space and let $P_1, ..., P_n$ be orthogonal projections such that there is some $0 \leq \varepsilon < \frac{1}{2n-1}$ such that for every $1 \leq i < j \leq n$, $\cos(\angle(P_i, P_j)) \leq \varepsilon$. Denote $T = \frac{P_1+...+P_n}{n}$. Then there is a constant $r < 1$ that depend only on ε such that for every $x \in \mathcal{H}$, $E(Tx) \leq rE(x)$.*

Consequently, for every x and every $k \in \mathbb{N}$, $E(T^k x) \leq r^k E(x) \leq r^k \|x\|$.

Proof. Fix some $x \in \mathcal{H}$. We note that

$$\|(I-P_1)Tx\| = \|(I-P_1)\frac{P_1+...+P_n}{n}x\| \leq$$
$$\frac{1}{n}\sum_{i=2}^{n}\|(I-P_1)P_i x\| = \frac{1}{n}\sum_{i=2}^{n}\|(P_i - P_iP_1)x + (P_iP_1 - P_1P_i)x\| \leq$$
$$\frac{1}{n}\sum_{i=2}^{n}\|P_i\|\|(I-P_1)x\| + \|(P_iP_1 - P_1P_i)x\| \leq^{\text{by previous lemma}}$$
$$\frac{1}{n}\sum_{i=2}^{n}\|(I-P_1)x\| + \frac{\cos(\angle(P_1,P_i))}{1-\cos(\angle(P_1,P_i))}(\|(I-P_1)x\| + \|(I-P_i)x\|) \leq^{\text{bound on cosine}}$$
$$\frac{1}{n}\sum_{i=2}^{n}\|(I-P_1)x\| + \frac{\varepsilon}{1-\varepsilon}(\|(I-P_1)x\| + \|(I-P_i)x\|) =$$
$$\frac{n-1}{n}(1+\frac{\varepsilon}{1-\varepsilon})\|(I-P_1)x\| + \sum_{i=2}^{n}\frac{\varepsilon}{1-\varepsilon}\|(I-P_i)x\|.$$

By a similar computation, for every j,

$$\|(I-P_j)Tx\| \leq \frac{n-1}{n}(1+\frac{\varepsilon}{1-\varepsilon})\|(I-P_j)x\| + \sum_{i=1, i\neq j}^{n}\frac{\varepsilon}{1-\varepsilon}\|(I-P_i)x\|.$$

Thus, by summing over all the j's and dividing by n, we get

$$E(Tx) \leq \frac{1}{n}\sum_{j=1}^{n}\frac{n-1}{n}(1+2\frac{\varepsilon}{1-\varepsilon})\|(I-P_j)x\| = \frac{n-1}{n}(1+2\frac{\varepsilon}{1-\varepsilon})E(x).$$

Denote $r = \frac{n-1}{n}(1+2\frac{\varepsilon}{1-\varepsilon})$. The assumption $\varepsilon < \frac{1}{2n-1}$ implies that $r < 1$ and thus $E(Tx) \leq rE(x)$.

Since this is true for every x, we can apply induction and get $E(T^k x) \leq r^k E(x) \leq r^k \|x\|$. □

After this, we are ready to state the theorem regarding convergence of $\|\left(\frac{P_1+...+P_n}{n}\right)^k\|$:

Theorem B.6. *Let $\mathcal{H}$ be a Hilbert space and let $P_1, ..., P_n$ be orthogonal projections such that there is some $0 \leq \varepsilon < \frac{1}{2n-1}$ such that for every $1 \leq i < j \leq n$, $\cos(\angle(P_i, P_j)) \leq \varepsilon$, then there are constants $C > 0$, $r < 1$ that depend only on ε such that for every k, $\|\left(\frac{P_1+...+P_n}{n}\right)^k\| \leq Cr^k$.*

In particular, there is some k that depends only on ε such that $\|\left(\frac{P_1+...+P_n}{n}\right)^k\| \leq \frac{1}{2}$.

Proof. Denote $T = \frac{P_1+...+P_n}{n}$. Fix some $k \in \mathbb{N}$. Then for every $x \in \mathcal{H}$,

$$\|(T^{k+1} - T^k)x\| = \|(I - \frac{P_1 + ... + P_n}{n})T^k x\| \leq \frac{1}{n}\sum_{i=1}^{n}\|(I - P_i)T^k x\| =$$
$$E(T^k x) \leq r^k \|x\|,$$

where $r < 1$ is as in the previous lemma. Therefore, for every k, j and every x,

$$\|(T^{k+j} - T^k)x\| \leq$$
$$\|(T^{k+j} - T^{k+j-1})x\| + \|(T^{k+j-1} - T^{k+j-2})x\| + ... + \|(T^{k+1} - T^k)x\| \leq$$
$$(r^{k+j-1} + r^{k+j-2} + ...r^k)\|x\| \leq \frac{r^k}{1-r}\|x\|.$$

Thus, for every x , $\{T^k x\}_{k\in\mathbb{N}}$ is a Cauchy sequence and therefore convergent. Let $x \in \mathcal{H}$ be arbitrary and denote $\lim_k T^k x = x_0$. Then $Tx_0 = x_0$, but by the previous lemma, $E(x_0) = E(Tx_0) \leq rE(x_0)$, which implies that $E(x_0) = 0$ and thus $x_0 = 0$. It follows that for every x, $\lim_k T^k x = 0$.

Fix $x \in \mathcal{H}$ to be some unit vector, then by the previous computation, for every k, j,

$$\|(T^{k+j} - T^k)x\| \leq \frac{r^k}{1-r},$$

and taking $j \to \infty$, yields that $\|T^k x\| \leq \frac{r^k}{1-r}$ as needed (in the notations of the theorem $C = \frac{1}{1-r}$). □