High Dimensional Expanders

Lecture 5: Cayley Graphs and Expanders

Instructor: Irit Dinur *Scribe: Limor Friedman*

In this lecture we introduce Cayley graphs and examine some of their spectral properties. We then discuss ϵ-biased distributions, which we use to construct Cayley graphs which are good expanders. At the end we present the notion of error correcting codes which will be further studied in the next lecture.

1 Introduction - groups and Cayley graphs

Definition 1.1 (Group)**.** *A group is a set G together with a binary operation $* : G \times G \to G$ satisfying:*

1. *For every $a, b, c \in G$, $a * (b * c) = (a * b) * c$.*
2. *There exists an element $e \in G$ s.t. $e * a = a * e = a$ for every $a \in G$.*
3. *For every $a \in G$ there exists an element $a^{-1} \in G$ s.t. $a * a^{-1} = a^{-1} * a = e$.*

Examples:

1. The additive cyclic group $(\mathbb{Z}_n, +)$ (integers modulo n).
2. The additive group $(\mathbb{Z}_2^n, +)$ (n-dimensional vector space over the field $\mathbb{Z}_2$).
3. $(SL_2(p), \cdot)$, where $SL_2(p) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ : \ a, b, c, d \in \mathbb{Z}, \ \det A = 1 \ (\mod p) \right\}$.
4. Given a set of symbols S, denote by $S^{-1} = \left\{ s^{-1} \ : \ s \in S \right\}$. A word in S is a written product of elements of $S \cup S^{-1}$, i.e. $w_1 w_2 \dots w_k$ where $k \in \mathbb{N}$ and $w_i \in S \cup S^{-1}$. We say a word $w_1 w_2 \dots w_k$ in S is reduced if for every $1 \leq i \leq k-1$, $w_{i+1} \neq w_i^{-1}$.
   The free group F_S is defined to be the group of all reduced words in S with concatenation of words as the group operation (followed by reduction if necessary).

Remark: A group G is called abelian if for every $a, b \in G$, $a * b = b * a$. In the previous examples (1) and (2) are abelian groups while (3) is not. (4) is abelian only in case of one gnerator.

Definition 1.2 (Generating set)**.** *A subset S of a group G is called a generating set if for every $g \in G$, $g = w_1 w_2 \dots w_\ell$ where $w_i \in S$ or $w_i^{-1} \in S$ for every $i \in [\ell]$.*
For example the set $\{1\}$ generates $\mathbb{Z}_n$, the set S generates the free group F_S.

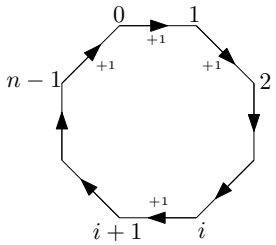
Definition 1.3 (Cayley graph)**.** *Let G be a group and let $S \subseteq G$ be a subset, the Cayley graph $Cay(G, S)$ is a graph with vertex set G and edge set $\{(g, gs) \ : \ s \in S, g \in G\}$.*

Remarks:

1. G as defined above is a directed graph. Assuming S is symmetric, i.e. $S = S^{-1}$, we can think of the graph as an undirected graph.
2. The graph $\mathrm{Cay}(G, S)$ is connected iff S generates G.

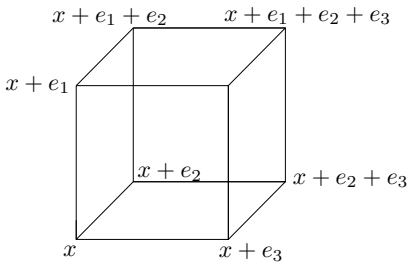
Examples:

- $G = \mathbb{Z}_n$, $S = \{-1, 1\}$



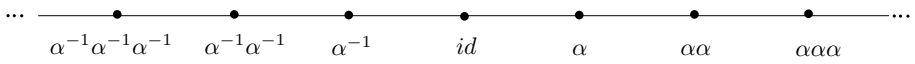
$Cay(\mathbb{Z}_n, \{\pm 1\})$

- $G = \mathbb{Z}_2^n$, $S = \{e_i\}_{i=1}^n$



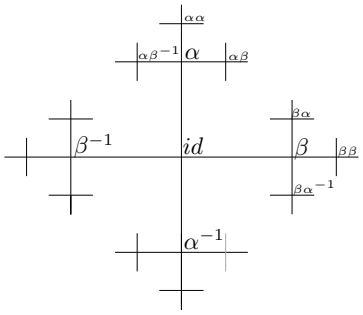
$Cay(\mathbb{Z}_2^n, \{\pm e_i\}_{i=1}^n)$

- Free group on 1 element α



$Cay(F_1, \{\alpha\})$

- Free group on 2 elements α, β



$Cay(F_2, \{\alpha, \beta\})$

2 Spectral properties of Cayley graphs

Definition 2.1 (Character)**.** *Let G be a group. A character of G is a function $\chi : G \to \mathbb{C}\setminus\{0\}$ s.t. for every $g, h \in G$, $\chi(gh) = \chi(g)\chi(h)$.*

Examples:

1. The group $\mathbb{Z}_2$ has two characters - the trivial character $I_G \equiv 1$ and χ defined by $\chi(g) = (-1)^g$ for every $g \in G$.
2. The group $\mathbb{Z}_2^n$ has 2^n characters - for every $T \subseteq [n]$ define χ_T by $\chi_T(z) = (-1)^{\sum_{i\in T} z_i}$ for every $z \in \mathbb{Z}_2^n$. Note that χ_T is indeed a character of G as for every $z, z' \in \mathbb{Z}_2^n$

$$\chi_T(z+z') = (-1)^{\sum_{i\in T}(z+z')_i} = (-1)^{\sum_{i\in T} z_i + \sum_{i\in T} z'_i} = (-1)^{\sum_{i\in T} z_i}(-1)^{\sum_{i\in T} z'_i} = \chi_T(z)\cdot\chi_T(z').$$

Remark: The characters of G are vectors in the linear space of all complex functions on G, $\mathbb{C}^G$. This space has an inner product $\langle f, g\rangle = \mathop{\mathbb{E}}_{z\in G} f(z)\overline{g(z)}$.

Lemma 2.2. *For every $T \subseteq [n]$, χ_T is an eigenvector of $Cay(\mathbb{Z}_2^n, \{e_i\}_{i=1}^n)$ with eigenvalue $\lambda_T = 1 - \frac{2|T|}{n}$.*

Proof. Let A be the transition matrix of $Cay(\mathbb{Z}_2^n, \{e_i\}_{i=1}^n)$, then

$$A\chi_T(z) = \mathop{\mathbb{E}}_{y\sim z}\chi_T(y) = \mathop{\mathbb{E}}_{i\in[n]}\chi_T(z+e_i) = \mathop{\mathbb{E}}_{i\in[n]}\chi_T(z)\chi_T(e_i) = \chi_T(z)\underbrace{\mathop{\mathbb{E}}_{i\in[n]}\chi_T(e_i)}_{\lambda_T} = \lambda_T\chi_T(z)$$

where

$$\lambda_T = \mathop{\mathbb{E}}_{i\in[n]}\chi_T(e_i) = \mathop{\mathbb{E}}_{i\in[n]}(-1)^{\sum_{j\in T}(e_i)_j} = \frac{|T^c| - |T|}{n} = \frac{n - |T| - |T|}{n} = 1 - \frac{2|T|}{n}.$$

□

Lemma 2.3. *For every $T, S \subseteq [n]$, $\langle\chi_T, \chi_S\rangle = \begin{cases} 0 & T \neq S \\ 1 & T = S \end{cases}$.*

Proof. For every $T \subseteq [n]$

$$\langle\chi_T, \chi_T\rangle = \mathop{\mathbb{E}}_z\chi_T(z)\chi_T(z) = \mathop{\mathbb{E}}_z 1 = 1.$$

For $T \neq S \subseteq [n]$, single out $j \in T \triangle S$, we have

$$\begin{aligned}\langle\chi_T, \chi_S\rangle &= \mathop{\mathbb{E}}_z\chi_T(z)\chi_S(z) = \mathop{\mathbb{E}}_z(-1)^{\sum_{i\in T} z_i}(-1)^{\sum_{i\in S} z_i} \\ &= \mathop{\mathbb{E}}_{\substack{z\\ z_j=0}}(-1)^{\sum_{i\in T\setminus\{j\}} z_i}(-1)^{\sum_{i\in S\setminus\{j\}} z_i} - \mathop{\mathbb{E}}_{\substack{z\\ z_j=1}}(-1)^{\sum_{i\in T\setminus\{j\}} z_i}(-1)^{\sum_{i\in S\setminus\{j\}} z_i} = 0.\end{aligned}$$

Corollary 2.4. *The characters $\{\chi_T\}_{T\subseteq[n]}$ form a basis of orthogonal eigenvectors for the Hamming cube. In particular, every function $f : \mathbb{Z}_2^n \to \mathbb{R}$ can be written as $f = \sum_T \alpha_T\chi_T$, where $\{\alpha_T\}_T$ are the "Fourier coefficients" usually denoted by $\alpha_T = \hat{f}(T)$.*

Remark: Let $B \subseteq \mathbb{Z}_2^n$, $f = I_B : \mathbb{Z}_2^n \to \{0, 1\}$ the indicator of B. Let A be the transition matrix of the Hamming cube graph, then

$$\langle Af, f\rangle = \left\langle \sum_T \alpha_T\lambda_T\chi_T, \sum_T \alpha_T\chi_T \right\rangle = \sum_T \alpha_T^2\lambda_T = \sum \alpha_T^2\left(1 - \frac{2|T|}{n}\right)$$

on the other hand

$$\langle Af, f\rangle = \mathop{\mathbb{E}}_z Af(z)f(z) = \mathop{\mathbb{E}}_z\left(\mathop{\mathbb{E}}_i f(z+e_i)f(z)\right) = \text{fraction of the edges inside } B$$

We see that B has more edges in it if its Fourier weight is on high degrees.

3 ϵ-biased distributions and expander graphs

Definition 3.1 (ϵ-biased bit)**.** *Let $\mathcal{D}$ be a distribution over $\{0,1\}^n$, a bit i is ϵ-biased if for all $\phi \neq T \subseteq [n]$*

$$\left| \Pr_{X\sim\mathcal{D}} [x_i = 0] - \Pr_{X\sim\mathcal{D}} [x_i = 1] \right| \leq \epsilon.$$

Definition 3.2 (ϵ-biased distribution)**.** *Let $\mathcal{D}$ be a distribution over $\{0,1\}^n$. We say $\mathcal{D}$ is ϵ-biased if for all $\phi \neq T \subseteq [n]$*

$$bias_T(\mathcal{D}) = \left| \Pr_{X\sim\mathcal{D}} \left[\sum_{i\in T} x_i = 0 \right] - \Pr_{X\sim\mathcal{D}} \left[\sum_{i\in T} x_i = 1 \right] \right| \leq \epsilon$$

For example, U_n, the uniform distribution on $\{0,1\}^n$ is 0-biased.

Lemma 3.3. *Suppose $\mathcal{D}$ is an ϵ-biased distribution that is uniform over a set $S \subseteq \mathbb{Z}_2^n$, then $\lambda(Cay(\mathbb{Z}_2^n, S)) \leq \epsilon$.*

Remark: This means that given such an ϵ-biased distribution we can construct an expander.

Proof. Let A be the transition matrix of this graph. For every $T \subseteq [n]$

$$A\chi_T(z) = \mathop{\mathbb{E}}_{s\in S} \chi_T(z+s) = \mathop{\mathbb{E}}_{s\in S} \chi_T(z)\chi_T(s) = \chi_T(z) \underbrace{\mathop{\mathbb{E}}_{s\in S} \chi_T(s)}_{=\lambda_T}$$

meaning $\{\chi_T\}_T$ are the eigenvectors of A with eigenvalues $\lambda_T = \mathop{\mathbb{E}}_{s\in S} \chi_T(s)$ (since $\{\chi_T\}_T$ are orthogonal these are all the eigenvalues).
For every $\phi \neq T \subseteq [n]$,

$$\lambda_T = \mathop{\mathbb{E}}_{s\in S} \chi_T(s) = \sum_{x\in\mathbb{Z}_2^n} \mathcal{D}(x)\chi_T(x) = \sum_{x\in\mathbb{Z}_2^n} \mathcal{D}(x)(-1)^{\sum_{i\in T} x_i} = \Pr_{x\sim\mathcal{D}} \left[\sum_{i\in T} x_i = 0 \right] - \Pr_{x\sim\mathcal{D}} \left[\sum_{i\in T} x_i = 1 \right]$$

and since $\mathcal{D}$ is ϵ-biased we get $|\lambda_T| \leq \epsilon$ and therefore $\lambda(Cay(\mathbb{Z}_2^n, S)) \leq \epsilon$. □

Corollary 3.4. *Let S be as in Lemma 3.3, then the graph $Cay(\mathbb{Z}_2^n, S)$ is connected, in particular S generates $\mathbb{Z}_2^n$ implying $|S| \geq n$.*

How small can S be s.t. $\mathcal{D}$, the uniform distribution over S is ϵ-biased? It turns out that $|S| = \frac{n}{\epsilon^2}$ is enough (this is logarithmic in the size of the space $\mathbb{Z}_2^n$).

4 Error correcting codes

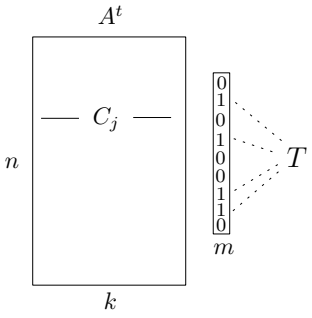
A linear error correcting code is an $\mathbb{F}_2$-linear subspace $C \subseteq \{0,1\}^n$.
The dimension of the code measures how much information we can send.
The distance of a code is the minimum distance between distinct codewords, i.e.

$$\text{dist}(C) = \min_{x\neq y\in C} \text{dist}(x,y) = \min_{x\neq y\in C} w(x-y) = \min_{0\neq x\in C} w(x)$$

where $w(x)$ (the weight of x) is the number of non-zero coordinates in x.
The distance measures how many errors can be tolerated.

The subspace C can be described as $Im(A^t)$, where A is a $k \times n$ matrix whose rows are a basis of C. We call A a generating matrix for the code C.

Let A be a generating matrix of a code C and let $Col\,(A)$ be the uniform ditribution over the columns of A.



Lemma 4.1. *If $Col\,(A)$ is an ϵ-biased k-bit distribution then $dist\,(C) \geq \frac{1}{2} - \frac{\epsilon}{2}$.*

Proof. For every $0 \neq m \in \{0,1\}^k$, we will show that $w\,(A^t m) \geq \left(\frac{1}{2} - \frac{\epsilon}{2}\right) n$.
Denote thecolumns of A by $\{C_i\}_{i=1}^n$ and let $T = \{i \in [k] \ : \ m_i = 1\}$.
Notice that for every $j \in [n]$

$$\left(A^t m\right)_j = \langle C_j, m \rangle = \sum_i \left(C_j\right)_i m_i = \sum_{i \in T} \left(C_j\right)_i$$

Since $Col\,(A)$ is ϵ-biased

$$\left| \underbrace{\left| j \in [n] \ : \ \sum_{i \in T} \left(C_j\right)_i = 0 \right|}_{\#\text{ 0's in } A^T m} - \underbrace{\left| j \in [n] \ : \ \sum_{i \in T} \left(C_j\right)_i = 1 \right|}_{\#\text{ 1's in } A^T m} \right| \leq \epsilon n$$

therefore $|n - 2w\,(A^t m)| \leq \epsilon n$, implying $w\,(A^t m) \geq \left(\frac{1}{2} - \frac{\epsilon}{2}\right) n$. □