# Expander Graphs and their Applications*
# Lecture 2: Graph Expansion & Eigenvalues

Notes taken by Danny Harnik

March 2002

**Summary:** After defining families of expander graphs and giving some examples of such families, we discuss some algebraic properties of graphs. Mainly, we discuss the connection between the expansion property of a graph to the eigenvalues of the graph's adjacency matrix. We also see an application of expander graphs for error amplification with a small amount of random bits.

## 1 Definitions

We begin with some notes and notations:

- Throughout this lecture (and course) we discuss $d$-**regular** graphs (graphs in which all vertices have the same degree $d$). denote a graph by $G = (V, E)$ and $|V| = n$. We allow self loops and multiple edges in the graph.
- Unlike the previous lecture, we discuss general graphs and not only bipartite graphs.
- For $S, T \subset V$ denote the set of all edges between $S$ and $T$ by $E(S, T) = \{(u, v) | u \in S, v \in T, (u, v) \in E\}$.

**Definition 1.**

1. The **Edge Boundary** of a set $S$, denoted $\partial S$, is $\partial S = E(S, \overline{S})$. This is actually the set of outgoing edges from $S$.
2. The **Expansion Parameter** of $G$, denoted $h(G)$, is defined as:

$$h(G) = \min_{\{S \mid |S| \leq \frac{n}{2}\}} \frac{|\partial S|}{|S|}$$

We note that there are other notions of expansion that can be studied. The most popular is counting the number of neighboring vertices of any small set $S$, rather than the number of outgoing edges.

**Definition 2. Family Of Expander Graphs**
A family of Expander graphs $\{G_i\}$ where $i \in \mathbf{N}$ is a collection of graphs with the following properties:

- The graph $G_i$ is a $d$-regular graph of size $n_i$ ($d$ is the same constant for the whole family). $\{n_i\}$ is a monotone growing series that doesn't grow too fast (e.g. $n_{i+1} \leq n_i^2$).
- For all $i$, $h(G_i) \geq \epsilon > 0$.

When discussing a family of expander graphs one should also consider the time required to construct such a graph. There are two natural versions for the requirement on the constructibility of graphs:

**Definition 3.**

* Lecture Notes for a course given by Nati Linial and Avi Wigderson at the Hebrew University, Israel.

1. A family of expander graphs is called **Mildly Explicit** if there is a Polynomial-Time algorithm that given $1^i$ creates $G_i$.
2. A family of expander graphs is called **Very Explicit** if there is a Polynomial-Time algorithm that given $(i, v, k)$ (where $i \in \mathbf{N}, v \in V$ and $k \in \{1, \cdots, d\}$) computes the $k$th neighbor of vertex $v$ in the graph $G_i$.

The second definition is useful for very large graphs, where one cannot construct the whole graph, but rather works locally on a small part of the graph.

# 2 Examples of Expander Graphs

1. This family of graphs $G_m$ lies on a grid: $V_m = \mathbf{Z}_m \times \mathbf{Z}_m$.
   The degree is $d = 4$ and the edges are described as follows:
   Vertex $(x, y)$ has edges to $(x+y, y), (x-y, y), (x, y+x)$ and $(x, x-y)$ (all operations are done modulo $m$).
   Margulis (73) showed that this is an expander family.
   Gaber & Galil (80) showed that this is an $\epsilon$-expander family ( for a specific $\epsilon$ ).
2. This family has graphs of size $p$ (for all prime $p$). Here $V_p = \mathbf{Z}_p$ and $d = 3$. Each vertex $x$ is connected to its neighbors and its inverse (i.e. $x+1, x-1$ and $x^{-1}$).
   This was shown to be an $\epsilon$-expander family by Lubotsky, Philips and Sarnak (88).

# 3 The Spectrum of a Graph

The **Adjacency Matrix** of a graph $G$, denoted $A(G)$, is an $n \times n$ matrix that for each $(u, v)$ contains the number of edges in $G$ between vertex $u$ and vertex $v$. Since the graph is $d$-regular, the sum of each row and column in $A(G)$ is $d$.

By definition the matrix $A(G)$ is symmetric and therefore has an orthonormal base $v_0, \cdots, v_{n-1}$, with eigenvalues $\mu_0, \mu_1, \cdots, \mu_{n-1}$ such that for all $i$ we have $Av_i = \mu_i v_i$. Without loss of generality we assume the eigenvalues are sorted in descending order $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_{n-1}$. The eigenvalues of $A(G)$ are called the **Spectrum** of the graph $G$.

The spectrum of a graph contains a lot of information regarding the graph. Here are some examples of observations that demonstrate this connection between the spectrum of a $d$-regular graph and its properties:

- $\mu_0 = d$
- The graph is connected iff $\mu_0 > \mu_1$
- The graph is bipartite iff $\mu_0 = -\mu_{n-1}$

In the rest of the lecture we will discuss the connection between the expansion of a graph and its spectrum. In particular, the graphs second eigenvalue is related to the expansion parameter of the graph.

**Theorem 4.**

$$\frac{d - \mu_1}{2} \leq h(G) \leq \sqrt{2d(d - \mu_1)}$$

This Theorem is due to Cheeger & Buser in the continuous case, and to Tamner, Alon & Milman in the discrete case.

The theorem actually proves that $d - \mu_1$, also known as the **Spectral Gap**, can give a good estimate on the expansion of a graph. Moreover, the graph is an expander ($h(G) > \epsilon$) if and only if the spectral gap is bounded ($d - \mu_1 > \epsilon'$). We do not prove this theorem at this stage (will be proved later in the course). Instead we show a Lemma that allows us to find connections between the expansion property and the second eigenvalue.

## 4 The Expander Mixing Lemma and Applications

Denote $\lambda = \max(|\mu_1|, |\mu_{n-1}|)$.
Since the eigenvalues are already sorted, this means that $\lambda$ is larger than the absolute value of all eigenvalues (except $\mu_0 = d$).

**Lemma 5. Expander Mixing Lemma** *for all* $S, T \subseteq V$*:*

$$||E(S,T)| - \frac{d|S||T|}{n}| \leq \lambda\sqrt{|S||T|}$$

This lemma can be viewed as relating the second eigenvalue to the question of how "random" the graph is. The left hand side compares the expected number of edges between $S$ and $T$ in a random graph ($\frac{d|S||T|}{n}$) and the actual number of edges between the two sets ($|E(S,T)|$). This difference is small when $\lambda$ is small. So a small $\lambda$ (or large spectral graph) means a graph with allot of "randomness".

*Proof.* Denote by $\chi_S$ and $\chi_T$ the characteristic vectors of $S$ and $T$ ($\chi_S$ is a vector with ones for all $v \in S$ and zeros in all other places). Let $\chi_S = \Sigma_i \alpha_i v_i$ and $\chi_T = \Sigma_j \beta_j v_j$ be their representation as linear combinations of the orthonormal base $v_0, \cdots, v_{n-1}$, where $v_0 = \bar{1}/\sqrt{n}$. We have:

$$\begin{aligned} |E(S,T)| &= \chi_S A \chi_T \\ &= (\Sigma_i \alpha_i v_i) A (\Sigma_j \beta_j v_j) \end{aligned}$$

and since the $v_i$'s are eigenvectors and orthonormal:

$$\begin{aligned} |E(S,T)| &= (\Sigma_i \alpha_i v_i)(\Sigma_j \beta_j A v_j) \\ &= (\Sigma_i \alpha_i v_i)(\Sigma_j \beta_j \mu_j v_j) \\ &= \Sigma_i \mu_i \alpha_i \beta_i \end{aligned}$$

Since $\alpha_0 = \langle \chi_S, \frac{\bar{1}}{\sqrt{n}} \rangle = \frac{|S|}{\sqrt{n}}$ and $\beta_0 = \frac{|T|}{\sqrt{n}}$:

$$\begin{aligned} |E(S,T)| &= \mu_0 \frac{|S||T|}{n} + \Sigma_{i=1}^{n-1} \mu_i \alpha_i \beta_i \\ &= d\frac{|S||T|}{n} + \Sigma_{i=1}^{n-1} \mu_i \alpha_i \beta_i \end{aligned}$$

Due to the triangle inequality and the definition of $\lambda$:

$$\begin{aligned} ||E(S,T)| - d\frac{|S||T|}{n}| &= |\Sigma_{i=1}^{n-1} \mu_i \alpha_i \beta_i| \\ &\leq \Sigma_{i=1}^{n-1} |\mu_i \alpha_i \beta_i| \\ &\leq \lambda \Sigma_{i=1}^{n-1} |\alpha_i \beta_i| \end{aligned}$$

And by the Cauchy-Schwartz inequality:

$$\begin{aligned} ||E(S,T)| - d\frac{|S||T|}{n}| &\leq \lambda ||\alpha||_2 ||\beta||_2 \\ &= \lambda ||\chi_S||_2 ||\chi_T||_2 \\ &= \lambda\sqrt{|S||T|} \end{aligned}$$

$\square$

Following is an example of an application of the lemma above:

## 4.1 Deterministic Error Amplification for BPP

In this example we are given a function in $BPP$. This means a function $f : \{0,1\}^n \to \{0,1\}$ and a probabilistic polynomial time algorithm $\mathcal{A}$ that approximates $f$ in the sense that for random $r \in \{0,1\}^m$ we have:

$$Pr_r[\mathcal{A}(x,r) \neq f(x)] \leq \frac{1}{4}$$

Our goal is to reduce the probability of error. This can be achieved by simply repeating $\mathcal{A}$ with different random coins $r$, and taking a majority vote. Using $t$ calls to $\mathcal{A}$ one can reduce the error to $c^{-t}$ for some constant $c$. However this requires a large number of coin tosses ($t \cdot m$ in this case). The question is: can we make the error smaller with a small number of random coins?

We introduce an algorithm $\mathcal{B}$ that uses just $m$ random coins: $\mathcal{B}$ uses a $d$-regular expander graph $G_{2^m}$ (of size $N = 2^m$). The algorithm chooses a random vertex $v \in G_{2^m}$ and takes a majority vote on the output of $\mathcal{A}$ on each of the neighbors of $v$ (denote by $\Gamma(v)$ the set of the neighbors of $v$ in $G_{2^m}$).

$$\mathcal{B}(x,v) = Majority_{u \in \Gamma(v)} \mathcal{A}(x,u)$$

**Claim 6.**

$$Pr_v[\mathcal{B}(x,v) \neq f(x)] \leq 4(\frac{\lambda}{d})^2$$

*Proof.* Let $S$ be the set of vertices that algorithm $\mathcal{B}$ makes errors on, and $T$ be the set of vertices that algorithm $\mathcal{A}$ makes errors on ($S = \{v | \mathcal{B}(x,v) \neq f(x)\}$ and $T = \{u | \mathcal{A}(x,u) \neq f(x)\}$). By definition, every $v \in S$ has at least $\frac{d}{2}$ neighbors in $T$. So: $|E(S,T)| \geq |S|\frac{d}{2}$ On the other hand, due to the Expander Mixing Lemma and since $|T| \leq \frac{N}{4}$:

$$\begin{aligned} |E(S,T)| &\leq \frac{d|S||T|}{N} + \lambda\sqrt{|S||T|} \\ &\leq \frac{d|S|}{4} + \lambda\sqrt{|S|\frac{N}{4}} \end{aligned}$$

Combined together we get:

$$\frac{d|S|}{4} \leq \lambda\sqrt{|S|\frac{N}{4}}$$

And finally:

$$\begin{aligned} Pr[\mathcal{B}(x,v) \neq f(x)] &= \frac{|S|}{N} \\ &\leq 4(\frac{\lambda}{d})^2 \end{aligned}$$

□

The above claim shows that error amplification can be done, but gives a rather poor amplification rate. There are a few ways to improve this:

- One can take instead of $G_{2^m}$, with adjacency matrix $A$, the graph with matrix $A^k$. In this new graph (containing an edge for any path of length $k$ in the original graph) we have degree $d^k$, but also second eigenvalue $\lambda^k$ giving a reduced error of $(\frac{\lambda}{d})^{2k}$.
- It is worth noting that in a good expander we can achieve $\lambda \approx \frac{1}{\sqrt{d}}$ giving an error of about $\frac{1}{d}$. So to get an error smaller than a given $\epsilon$ one should use graphs of degree $\frac{1}{\epsilon}$.
- We will see in the next lecture constructions with better amplification using random walks on expander graphs.

## 5 How Big Can the Spectral Gap be?

We conclude with the question of how small can $\lambda$ be.

As an example we can check the most connected graph - The Clique.
The $n$-clique $K_n$ is the graph were every vertex is connected to all its neighbors and has degree $d = n - 1$. The spectrum of the Clique can be easily calculated by viewing the Clique's matrix as $J - I$ were $J$ is the all ones matrix and $I$ is the identity matrix. The spectrum of $K_n$ is $[n-1, -1, -1, \cdots, -1]$. hence $\lambda = 1$.

But this is for $d \approx n$ and we are interested in the behavior of $\lambda$ when $d$ is much smaller (usually a constant). This case is discussed in the following Theorem due to Alon-Boppana:

**Theorem 7.** *for every $d$-regular graph:*

$$\lambda \geq 2\sqrt{d-1} - o_n(1)$$

We will not prove this theorem here, but instead show a weaker statement:

**Claim 8.** *for every $d$-regular graph:*

$$\lambda \geq \sqrt{d}(1 - o_n(1))$$

*Proof.*
note: In this discussion we don't allow multiple edges (the Adjacency matrix contains only zeros and ones).

Given a $d$-regular graph $G$ with adjacency matrix $A$, we look at the trace of $A^2$ (trace is the sum of the values in the diagonal) . On one hand, since the matrix is symmetric, and the sum of each row/column is $d$ we have $(A^2)_{ii} = d$ for all $i$:

$$Trace(A^2) = n \cdot d.$$

On the other hand:

$$\begin{aligned} Trace(A^2) &= \sum_i \mu_i^2 \\ &\leq d^2 + (n-1)\lambda^2 \end{aligned}$$

together we get

$$\lambda^2 \geq d\frac{n-d}{n-1}$$

and:

$$\lambda \geq \sqrt{d}(1 - o_n(1))$$

$\square$