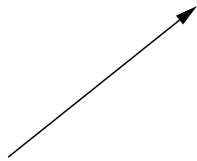
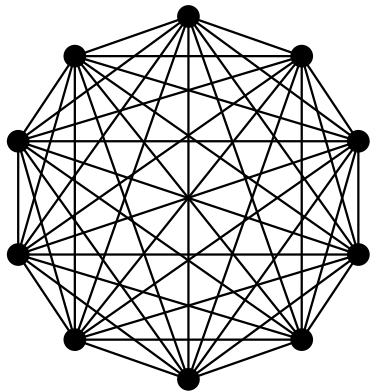


Expanding Graphs

9th February 2006



Joris Borgdorff

Contents

1	Preface	2
2	Introduction to graphs	4
2.1	Graphs	4
2.2	Cayley graphs	7
3	Expanders	8
3.1	Introduction to expanders	8
3.1.1	Calculating the Cheeger constant	12
3.2	Adjacency Matrix	12
3.3	Eigenvalues of the adjacency matrix	13
3.4	The Laplacian	14
3.4.1	Linear algebra	14
3.4.2	Properties of the Laplacian	15
4	Constructing an expander	19
4.1	Margulis construction	19
4.2	Morgenstern construction	22
5	An expander as randomizer	26
5.1	Random number generators	26
5.2	Primality test	27
5.3	Random walk on an expander	28
5.3.1	Some notation	28
5.3.2	Random walk	28
5.3.3	Resource analysis	30
6	Afterword	31

1 Preface

This thesis is about so-called expanding graphs. These graphs have the property that it is easy to get from one point to any other in the graph while it does not need many connections. In real life this expander graph could be viewed as a country with many cities where you have to create an infrastructure, reducing most of the traffic jams while not having to construct too many roads.

We assume the reader has some knowledge of mathematical notation and linear algebra while no knowledge on graph theory is presumed. Some group

theory is needed for section 2.2 and section 4 but is not necessary for the rest of the thesis.

Graph theory will be discussed in section 2. Most of the basics about graphs will be mentioned, as well as some properties that are of special interest for expanding graphs.

In section 3 we will go into what an expanding graph exactly is, as well as mentioning some of the properties of expanding graphs. For this we will notably use the so-called Laplacian. Some knowledge of linear algebra becomes important.

Section 4 will show us two methods of construction for expanding graphs. While proof and background of these constructions are difficult, we can apply the constructions.

In section 5 we will discuss random walks on expanding graphs, which finally yields some (theoretic) practical use. Random walks can help reduce the number of random bits (or coin flips) some random algorithm need.

Realise that this is just an introduction to expanding graphs. Most literature about these graphs use highly advanced mathematics so to comprehend them one would need a full study on its own. This said, it is interesting to see what we can learn about them in a shorter timespan.

Literature I have used intensively to study expanding graphs are:

- [Ser80] which gives a lot of basic graph theory and backgrounds on graphs. It is a difficult book which requires quite a lot of group and ring theoretic knowledge in some parts, particularly the end.
- chapters 1 and 4.2 from [Lub94]. These chapters focus on expanders as graphs. The rest of the book was too complex for me to understand as a 3rd year student. It brings together different sorts of mathematics using expanders.
- [DSV03] is a recent and very comprehensible book. Especially chapter 1 was of great use to me, other chapters were less relevant for my thesis.
- [LWd03] are lecture notes of a course on expanding graphs. These notes are very well understandable for undergraduate students and very useful as they give some applications of expanding graphs, which are hard to find in the literature, especially as clear as they are given here.

The most recent copy of this thesis will be available at <http://mooi.mine.nu/bachelorthesis/>. The Java program described in Subsection 3.1.1 will also be located there, along with the source.

This thesis is formatted using L^AT_EX 2_ε and uses a graph drawing package available at <http://www.cs.umu.se/~drewes/graphs/>.

2 Introduction to graphs

2.1 Graphs

As we will use graphs extensively we will start by defining what a directed graph is exactly. In this thesis we will not be using undirected graphs. Most of the following definitions are slightly modified versions of those in chapter 1.2.1 of [Ser80].

Definition 2.1. A directed **graph** X is an ordered pair $X = X(V, E)$ where V and E are finite sets. There are two maps

$$E \rightarrow V \times V \quad e \mapsto (o(e), t(e))$$

and

$$E \rightarrow E \quad e \mapsto \bar{e}$$

which satisfy the following conditions: for each $e \in E$ we have $\bar{\bar{e}} = e$, $\bar{e} \neq e$ and $o(e) = t(\bar{e})$.

As an interpretation V is called the set of vertices of X and is also denoted as $V(X)$. E is called the set of edges between these vertices and also denoted as $E(X)$. So an element $v \in V(X)$ is called a **vertex** of X ; an element $e \in E$ is called an **(oriented) edge** and \bar{e} is called the *inverse* edge. The vertex $o(e)$ is called the *origin* of e and the vertex $t(e)$ is called the *terminus* of e . Two vertices $p, q \in V$ are called **adjacent** if there is an edge $y \in E$ so that $o(y) = p$ and $t(y) = q$ or the other way around.

To get some feeling for graphs we will often use *diagrams*. In such diagrams vertex is drawn with a dot and every edge with a line or a vector. These are usually drawn in the plane but can also be drawn in 3-dimensional space. Lines may cross but we will not assign many importance to that fact in this thesis but mention that it is a point to worry about for a more formal approach.

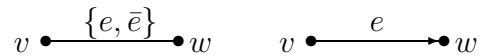


Figure 1: Types of edges

In Figure 1 we see that edge $o(e) = v$ and $t(e) = w$. This means v and w are adjacent. As we see in the first part there is a line and in the second a vector. This is to make the difference between edges *with* an inverse (a line) and edges *without* an inverse (a vector) clear in diagrams.

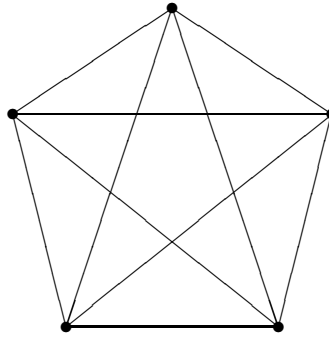


Figure 2: Diagram of K_5

An often used graph is K_n , the complete graph. K_n is a graph where any vertex is adjacent to any other and has n vertices. In Figure 2 we see for example the graph K_5 .

In this thesis we will not allow graphs with *loops* or *multiple edges*. A loop is an edge e so that $o(e) = t(e)$ and double edges are edges $e_i, \dots, e_j \in E(X), e_k \neq e_l$ so that $o(e_k) = o(e_l)$ and $t(e_k) = t(e_l)$ for $0 \leq i \leq k, l \leq j \leq n - 1$.

To move on: here is a list of definitions we will need.

Definition 2.2. A **subgraph** $X' := (V', E')$ of a graph X is defined by $V' \subseteq V(X)$ and $E' \subseteq E(X)$ so that for $e \in E'$ $o(e) \in V'$ and $t(e) \in V'$.

Definition 2.3. Two graphs X and X' are called **isomorphic** if there are bijective functions $\phi : V(X) \mapsto V(X')$ and $\psi : E(X) \mapsto E(X')$ so that if $o(e) = v$ and $t(e) = w$ then $o(\psi(e)) = \phi(v)$ and $t(\psi(e)) = \phi(w)$ for $e \in E(X)$ and $v, w \in V(X)$.

Definition 2.4. The **order** of a graph X , denoted by $|X|$ is the number of vertices it contains. We call X **k -regular** or of degree k if every vertex has k adjacent vertices. $|E(X)|$ is the number of edges, not counting inverses.

Definition 2.5. An **orientation** $X^+(V, E^+)$ on a graph $X(V, E)$ is a subgraph of X with E^+ an arbitrary subset of E so that for every $e \in E$ either $e \in E^+$ or $\bar{e} \in E^+$ but not both.

Notation 2.6. Let X be a graph and an orientation X^+ . Given $A, B \subseteq V(X)$ two disjoint subsets the number of edges $e \in E(X^+)$ so that $o(e) \in A$ and $t(e) \in B$ or $o(e) \in B$ and $t(e) \in A$ is denoted by $E(A, B)$.

Two other terms we will need are **path** and **circuit**.

Definition 2.7. Path_n is an oriented graph with $n + 1$ vertices and the following properties:

$$E(X) = \{p_1, \dots, p_n\}$$

so that $t(p_l) = o(p_{l+1})$ for $1 \leq l \leq n - 1$ and $p_i \neq p_j$, $p_i \neq \bar{p}_j$, $t(p_i) \neq t(p_j)$, $t(p_i) \neq o(p_1)$ for $1 \leq i, j \leq n$, $i \neq j$.

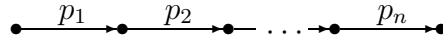


Figure 3: Diagram of Path_n

Definition 2.8. Circ_n is an oriented graph with n vertices and the following properties:

$$E(X) = \{p_1, \dots, p_n\}$$

so that $t(p_l) = o(p_{l+1})$ for $1 \leq l \leq n - 1$ and $p_i \neq p_j$, $p_i \neq \bar{p}_j$, $t(p_i) \neq t(p_j)$ for $1 \leq i, j \leq n$, $i \neq j$. Also $t(p_n) = o(p_1)$

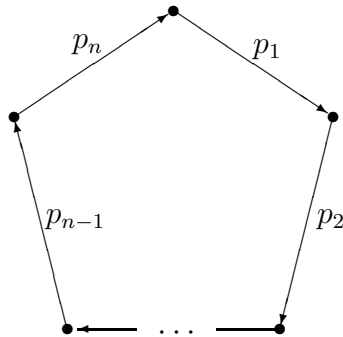


Figure 4: Diagram of Circ_n

Definition 2.9. A subgraph of a graph is called a **path** (without backtracking) of length n if it is isomorphic to Path_n . Similarly, a subgraph of a graph is called a **circuit** (without backtracking) of size n if it is isomorphic to Circ_n .

Definition 2.10. A graph is called **connected** if there is a path from any vertex to any other.

Definition 2.11. The **distance function** $d : V(X) \times V(X) \mapsto \mathbb{Z}_{\geq 0}$ on a graph X where $d(A, B)$ is the *length of the shortest path* from any vertex in A to any vertex in B , where $A, B \subseteq V$. If there is no such path $d(A, B) = \infty$.

Definition 2.12. A graph $X(V, E)$ is **bipartite** if there are two disjoint sets $I, O \subseteq V$ with $V = I \cup O$ so that either $[o(e) \in I \text{ and } t(e) \in O]$ or $[o(e) \in O \text{ and } t(e) \in I]$ for each $e \in E$. Graph X is also written as $X(I, O, E)$.

An interpretation often given to a bipartite graph is that I is the set of *inputs* of X and O is the set of *outputs* of X . Inputs could be sensors and outputs could be different processing units.

A notion which can help us later is that of a tree.

Definition 2.13. A **tree** is a connected graph $X(V, E)$ where for every $v \in V$ there is no more than one $e \in E$ so that $t(e) = v$.

An example of a tree is Path_n .

2.2 Cayley graphs

Definition 2.14. A graph $X(G, S)$ of a group G with $S \subseteq G$ is called a **Cayley graph** and has G as its set of vertices, and S determines the edges by the following property: $a \in G$ is connected to $as \in G$ for every $s \in S$.

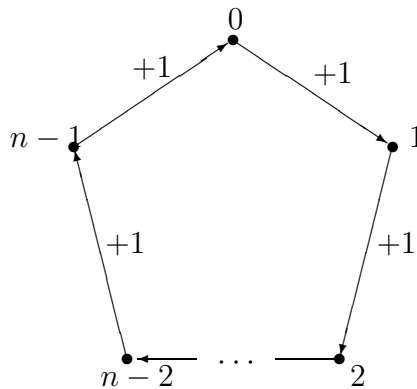


Figure 5: Cayley graph $X((\mathbb{Z}_n, +), \{1\})$

A simple example of a Cayley graph $X(G, S)$ is one where $G = (\mathbb{Z}_n, +)$ (the group under normal addition modulo n) and $S = \{1\}$. With one look at Figure 5 we see it is isomorphic to Circ_n . Would we take $S = \{1, -1\}$ then we would have an unoriented graph instead of an oriented graph. This fact

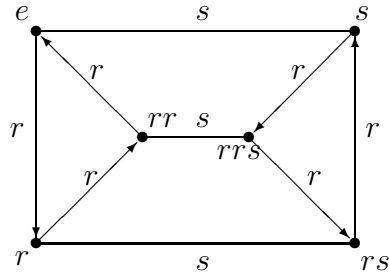


Figure 6: Cayley graph $X(D_3, \{r, s\})$

is easy to see intuitively with the argument that when $a, a^{-1} \in S$ all edges a are unoriented (has an inverse \bar{a}). In Figure 6 is the diagram of the Cayley graph $X(D_3, \{r, s\})$ which is a bit more complex. Since $s = s^{-1}$ in D_3 , all edges s are unoriented.

3 Expanders

We are particularly interested in a special type of graphs called *expanding graphs* or in short **expanders**. Expanders are used to construct sparse k -regular graphs which are highly connected. These expanders are of great interest for a number of reasons:

- they help create efficient computer networks with use of so-called *superconcentrators*,
- they have helped computer science to create error correcting codes,
- expanders can also help reducing the need for a true randomizer.

Section 5 deals with this last item. The two other items will not be discussed further. For reference please look into [LWd03]. Expanders can also help in some mathematical issues which go beyond the scope of this thesis. One might read [Lub94] for some further information, although this requires quite some prior knowledge.

3.1 Introduction to expanders

Definition 3.1. A finite connected graph $X(V, E)$ where V has n vertices and X is of degree k (so E contains $\frac{kn}{2}$ edges) is called an (n, k, c) -**expander** if for every non-empty non-equal subset A of V ,

$$|\delta A| \geq c \left(1 - \frac{|A|}{n}\right) |A| \tag{1}$$

where $\delta A = \{y \in V \mid d(y, A) = 1\}$ is the *boundary* of A and d is the distance function on X .

Why does equation 1 give a good definition for an expander? What it roughly says is that whatever subset we have, the number of neighbours is always larger than the size of our subset times some constant. This is given some nuance as we multiply this with the chance that a vertex is not in our subset. If we would not do this, for instance this would lead to very unreasonable results when we take A to be all vertices but one.

There is also a different definition for a special kind of expander, the bounded concentrator. A separate definition is useful because bounded concentrators are bipartite graphs, a special case.

Definition 3.2. Let a graph $X(I, O, E)$ be bipartite and k -regular. Then X is called a (n, θ, k, α) -**bounded concentrator** if

- $|I| = n$,
- $|O| = \theta n$ for $0 < \theta < 1$,
- $|\delta I| \geq |I|$ if $|I| \leq \alpha n$ for $0 \leq \alpha \leq \theta$,

Particularly the last property shows why such a bounded concentrator has a good expansion.

Now that we know what an expander is it might be useful to define a measure of expanders, something to compare the quality of different expanders. For this reason I introduce the a **Cheeger constant**¹ of a graph X .

Definition 3.3. Let $X(V, E)$ be a finite graph. Define the Cheeger constant of X , denoted $h(X)$, by:

$$h(X) = \min_{\emptyset \subsetneq A \subsetneq V} \frac{|E(A, V \setminus A)|}{\min(|A|, |V \setminus A|)}. \tag{2}$$

The higher this Cheeger constant of a graph X is, the better we call the expanding quality of X . This is easy to grasp: the Cheeger constant is the lower bound on the number of edges outside your set compared to the vertices

¹The Cheeger constant was initially used with Riemannian manifolds and imported to graph theory later. As Riemannian manifolds go beyond the extent of this thesis I will not discuss the original meaning of the Cheeger constant.

in your set. So a higher Cheeger constant means a graph expands quicker. To see this take for example K_n and Circ_n .

With K_n the number of edges between A and $V \setminus A$ is always $|A||V \setminus A|$. We can suppose $|A| \leq |V \setminus A|$. In this case $|V \setminus A| \geq \lceil \frac{n}{2} \rceil$ and

$$\begin{aligned} h(X) &= \min_{0 \leq |A| \leq \frac{n}{2}} \frac{E(A, V \setminus A)}{\min(|A|, |V \setminus A|)} \\ &= \min_{0 \leq |A| \leq \frac{n}{2}} \frac{|A||V \setminus A|}{|A|} \\ &= \lceil \frac{n}{2} \rceil \end{aligned}$$

With Circ_n take A as a half-circle. Then $E(A, V \setminus A) = 2$ so

$$h(X) \leq \frac{2}{\lceil \frac{n}{2} \rceil} \leq \frac{4}{n-1}.$$

We see the Cheeger constant is a lot smaller for Circ_n than for K_n , especially as n grows large. This makes sense: K_n is clearly much more connected than Circ_n . While K_n is highly connected there is one problem: the number of edges grows quadratically as n grows. This often not wanted and that is why expander graphs are defined to be k -regular, this means the number of edges grows linearly.

For a small graph, c is often quite large. For large n the c often gets small. We are interested in a **family of expanding graphs**, where c and k are fixed while n gets large. With a family we mean graphs constructed with the same method. As we shall see in section 4, to create such a family is not an easy task.

Now we can formulate the following proposition. This proposition shows there is quite a strong relation between the Cheeger constant and c of an (n, k, c) -expander.

Proposition 3.4. *Let X be a k -regular graph with n vertices. Then*

- i. If X is an (n, k, c) -expander then $h(X) \geq \frac{c}{2}$*
- ii. X is an $(n, k, \frac{h(X)}{k})$ -expander.*

Proof. i. It is easy to see that $|E(A, V \setminus A)| \geq |\delta A|$ and $|E(A, V \setminus A)| \geq |\delta(V \setminus A)|$: for every edge only one vertex on either side is used. For each vertex there may be multiple edges going to the other side. From this

and Definition 3.1 and 3.3 follows

$$\begin{aligned} h(X) &\geq \min_{\emptyset \subsetneq A \subsetneq V} \frac{|\delta A|}{\min(|A|, |V \setminus A|)} \\ &\geq \min_{\emptyset \subsetneq A \subsetneq V} \frac{c(1 - \frac{|A|}{n})|A|}{\min(|A|, |V \setminus A|)}. \end{aligned}$$

This given we can easily write $|A|$ as m and $|V \setminus A|$ as $n - m$. Define $h_1(n)$ as

$$\begin{aligned} h_1(n) &= \min_{0 \leq m \leq \lfloor \frac{n}{2} \rfloor} \frac{c(1 - \frac{m}{n})m}{m} \\ &= \min_{0 \leq m \leq \lfloor \frac{n}{2} \rfloor} c(1 - \frac{m}{n}) \end{aligned}$$

and similarly define $h_2(n)$ as

$$\begin{aligned} h_2(n) &= \min_{\lceil \frac{n}{2} \rceil \leq m \leq n} \frac{c(1 - \frac{m}{n})m}{n - m} \\ &= \min_{\lceil \frac{n}{2} \rceil \leq m \leq n} c \frac{m}{n}. \end{aligned}$$

Clearly these two functions are symmetrical and their minimum lies in $m = \lfloor \frac{n}{2} \rfloor$ and $m = \lceil \frac{n}{2} \rceil$ respectively. So now

$$h(X) \geq \min(h_1(|X|), h_2(|X|)) = c \frac{\lceil \frac{n}{2} \rceil}{n} \geq \frac{c}{2}$$

- ii. First assess that $|\delta A| \geq \frac{E(A, V \setminus A)}{k}$: every vertex in X has k edges so for every $v \in V \setminus A$ there can only be k edges so that $e \in E(X)$ $o(e) \in A$ and $t(e) = v$. This means $k|\delta A| \geq E(A, V \setminus A)$. This leads to the following:

$$\begin{aligned} |\delta A| &\geq |\delta A| \left(1 - \frac{|A|}{n}\right) \\ &\geq \frac{E(A, V \setminus A)}{k} \left(1 - \frac{|A|}{n}\right) \\ &= \frac{E(A, V \setminus A)}{k|A|} \left(1 - \frac{|A|}{n}\right) |A| \\ &\geq \frac{h(X)}{k} \left(1 - \frac{|A|}{n}\right) |A| \end{aligned}$$

which proves X is an $(n, k, \frac{h(X)}{k})$ -expander. □

3.1.1 Calculating the Cheeger constant

Calculating $h(X)$ by using its definition would take exponential time in the number of vertices; for a modern computer it might take 24 hours to calculate the Cheeger constant for a graph X with $|X| = 15$. What is possible though is making a fast estimate. For this purpose I have adapted a Java program written by Jeffrey Lijffijt. This program is available on <http://mooi.mine.nu/bachelorthesis/> along with the source. For a given unoriented graph X this program makes a few assumptions which do not always lead to the actual Cheeger constant, but to an estimate.

The algorithm first picks a vertex. This is put in a set A . The algorithm then picks one of the adjacent vertices of A with the least edges not going to A and adds this to A . Every time a vertex is added a Cheeger estimate is taken. This repeats until all but one vertices are in A . The program repeats the algorithm until every vertex is picked first once, then the minimum of all Cheeger estimates is taken.

Let the algorithm use an unoriented graph $X(V, E)$ and let $n = |V|$, $m = |E|$. Then the program has a performance of $O(n^2m \log n)$. This is due to the algorithm being $O(nm \log n)$, which is repeated n times for each vertex being picked first once. The vertex with the least number of edges not going to A is picked by using a priority queue, which uses $O(\log n)$ time. In the algorithm $n - 1$ vertices are added to A . When a vertex v is added to A , the algorithm checks every edge of v to calculate $E(A, V \setminus A)$ and informs the adjacent vertices that v has been added to A . A performance of $O(n^2m \log n)$ is good, much better than exponential running time.

The reason why this program does not always find the actual Cheeger constant is that the set A for which the minimum is found in Eq. (2) does not have to be connected. The algorithm assumes that A is connected, as this will most often be the case and lead to a good estimate. Furthermore the algorithm is greedy, which means it searches for a local minimum. However, local minima do not always lead to a global minimum. Apart from getting an estimate the algorithm also gives an upper bound on the Cheeger constant. I am not able to give further guarantees about the program like a lower bound or how correct the Cheeger constant estimate is.

3.2 Adjacency Matrix

To make the analysis of graphs somewhat easier we can also write a graph X as a matrix.

Definition 3.5. The **adjacency matrix** A of a graph X is a matrix so that

$$A_{ij} = \begin{cases} 0 & \text{if not } \exists e \in E(X) \text{ so } o(e) = v_i, t(e) = v_j \\ |E_{ij}| & \text{if } \exists e \in E(X) \text{ so } o(e) = v_i, t(e) = v_j \end{cases}$$

where $\{v_1, \dots, v_n\} = V(X)$ is an arbitrary ordering and E_{ij} is the set of edges from v_i to v_j .

Because the ordering is arbitrary another used notation is to use vertices instead of indices for adjacency matrix A . So notation becomes A_{vw} , $v, w \in V(X)$. Note that $|E_{ij}|$ is 1 because we assumed that graphs do not contain multiple edges.

Take for instance adjacency matrix A of K_5 and adjacency matrix B of Circ_5 :

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Some characteristics of these matrices immediately draw attention.

Proposition 3.6. *Let A be the adjacency matrix of a graph X . Then*

- i. $A_{vv} = 0$,*
- ii. if X is unoriented then A is a symmetric matrix.*

Proof. i. On page 5 we restricted ourselves to graphs without loops. Now if $A_{vv} = 1$ then there is an edge $e \in E(X)$ so that $o(e) = t(e)$ which is exactly a loop.

- ii. If X is unoriented then for every $e \in E(X)$ there is an inverse \bar{e} of e . This means that if $o(e) = v$ and $t(e) = w$ then also $o(\bar{e}) = w$ and $t(\bar{e}) = v$. This makes $A_{vw} = A_{wv}$ for every $v, w \in V(X)$.

□

3.3 Eigenvalues of the adjacency matrix

Now that we have defined this adjacency matrix we can analyse its eigenvalues. This will lead some interesting results.

Definition 3.7. Let $\lambda_0, \dots, \lambda_{n-1}$ be the eigenvalues of the adjacency matrix A of a k -regular graph X so that

$$\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{n-1}.$$

We call these ordered eigenvalues the **spectrum** of A or informally the spectrum of X .

Proposition 3.8. Let $\lambda_0, \dots, \lambda_{n-1}$ be the spectrum of a k -regular graph X . Then

i. $\lambda_i \leq k$ for all $0 \leq i \leq n - 1$,

ii. $\lambda_0 = k$.

Proof. i. Let \mathbf{p} be an eigenvector of A and let p_j be the largest component of \mathbf{p} . Then by Definition 3.5

$$(A\mathbf{p})_j = \sum_{i=0}^{n-1} A_{ji}p_i \leq \sum_{i=0}^{n-1} A_{ji}p_j = kp_j.$$

This means that for every eigenvector \mathbf{p}

$$A\mathbf{p} \leq k\mathbf{p}.$$

ii. This is quite easy: take the vector $\mathbf{u} = (1, 1, \dots, 1)$. Because every vertex v has k adjacent vertices every row obviously contains k times the value 1, which results in $A\mathbf{u} = k\mathbf{u}$.

□

3.4 The Laplacian

The **Laplacian** of a graph X is an averaging operator which we will use to create an upper and lower bound of the Cheeger constant of X based on the eigenvalues of the adjacency matrix of X . First we will freshen up our linear algebra a bit.

3.4.1 Linear algebra

As we saw in Proposition 3.6 the adjacency matrix of an unoriented graph X is symmetric. We will need this, and in particular need the following theorem.

Theorem 3.9 (Fundamental Theorem of Real Symmetric Matrices). *Every $n \times n$ real symmetric matrix has n real eigenvalues counted with their algebraic multiplicity, and is diagonalizable by a real orthogonal matrix.*

I will not state the proof here but refer to p. 480 of [FB95] or any other introductory book on linear algebra.

Proposition 3.10. *Let M be a real $m \times n$ -matrix and let $\mathbf{a} \in \mathbb{C}^n$ and $\mathbf{b} \in \mathbb{C}^m$. Then*

$$\langle M\mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{a}, M^T\mathbf{b} \rangle$$

Proof.

$$\begin{aligned} \langle M\mathbf{a}, \mathbf{b} \rangle &= (M\mathbf{a})^T\mathbf{b} \\ &= \mathbf{a}^T M^T\mathbf{b} \\ &= \langle \mathbf{a}, M^T\mathbf{b} \rangle \end{aligned}$$

□

For vectors we will define an ordering.

Definition 3.11. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$. Then

- $\mathbf{a} > \mathbf{b}$ if $a_i > b_i$, $0 \leq i \leq n-1$,
- $\mathbf{a} \geq \mathbf{b}$ if $a_i \geq b_i$, $0 \leq i \leq n-1$,
- $\mathbf{a} \leq \mathbf{b}$ if $a_i \leq b_i$, $0 \leq i \leq n-1$,
- $\mathbf{a} < \mathbf{b}$ if $a_i < b_i$, $0 \leq i \leq n-1$.

3.4.2 Properties of the Laplacian

Let $X^+(V, E^+)$ be an orientation on a graph X and $n = |X|$ and $m = |E^+|$. We then define the linear operator $D : \mathbb{C}^n \mapsto \mathbb{C}^m$ with $e \in E^+$ and $v \in V$.

$$D_{e,v} = \begin{cases} 1 & \text{if } t(e) = v, \\ -1 & \text{if } o(e) = v, \\ 0 & \text{otherwise.} \end{cases}$$

This operator makes the following transformation for $\mathbf{r} \in \mathbb{C}^n$:

$$(D\mathbf{r})_e = r_{t(e)} - r_{o(e)},$$

while for $D^T : \mathbb{C}^m \mapsto \mathbb{C}^n$ and $\mathbf{s} \in \mathbb{C}^m$:

$$(D^T\mathbf{s})_v = \sum_{e \in E^+ | t(e)=v} s_e - \sum_{e \in E^+ | o(e)=v} s_e$$

Definition 3.12. $\Delta := D^T D$ is the Laplacian on graph X .

Let us take K_4 as an example. K_4 is already oriented so we do not have to choose an orientation. For K_4

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

then the Laplacian Δ on K_4 becomes

$$\Delta = D^T D = \begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix}.$$

Proposition 3.13. Let Δ be the Laplacian on k -regular graph X , A the adjacency matrix of X and $\mathbf{r}, \mathbf{s} \in \mathbb{C}^n$. Then

- i. $\langle D\mathbf{r}, D\mathbf{s} \rangle = \langle \mathbf{r}, \Delta\mathbf{s} \rangle$,
- ii. $\Delta = kI - A$,
- iii. The eigenvalues of Δ are $\mu_0, \mu_1, \dots, \mu_{n-1}$ where $\mu_0 \leq \mu_1 \leq \dots \leq \mu_{n-1}$ and $\mu_i = k - \lambda_i$ for all $0 \leq i \leq n-1$ where the λ_i are the eigenvalues of the adjacency matrix of X ,
- iv. μ_1 is the smallest eigenvalue with an eigenvector \mathbf{p} so that

$$\langle \mathbf{p}, (1, 1, \dots, 1) \rangle = 0.$$

- v. For any $\mathbf{p} \in \mathbb{C}^n$ so that

$$\langle \mathbf{p}, (1, 1, \dots, 1) \rangle = 0$$

is

$$\Delta\mathbf{p} \geq \mu_1\mathbf{p}$$

Proof. i. Following Proposition 3.10 we have

$$\langle D\mathbf{r}, D\mathbf{s} \rangle = \langle \mathbf{r}, D^T D\mathbf{s} \rangle = \langle \mathbf{r}, \Delta\mathbf{s} \rangle.$$

ii.

$$\begin{aligned}
(\Delta \mathbf{r})_v &= (D^T D \mathbf{r})_v \\
&= \sum_{e \in E^+ | t(e)=v} (r_{t(e)} - r_{o(e)}) - \sum_{e \in E^+ | o(e)=v} (r_{t(e)} - r_{o(e)}) \\
&= \sum_{e \in E^+ | t(e)=v} (r_v - r_{o(e)}) - \sum_{e \in E^+ | o(e)=v} (r_{t(e)} - r_v) \\
&= kr_v - \sum_{e \in E^+ | t(e)=v} r_{o(e)} - \sum_{e \in E^+ | o(e)=v} r_{t(e)}
\end{aligned}$$

As we see for each vertex v the value of the vertex itself is multiplied by k , while the value of its neighbours is subtracted. This linear transformation is exactly $kI - A$.

iii. The eigenvalues of A are values λ_i so that

$$\det(A - \lambda_i I) = 0.$$

We have due to item ii that the eigenvalues of Δ are values μ_i so that

$$\det(\Delta - \mu_i I) = \det(kI - A - \mu_i I) = \det(-A + (k - \mu_i)I) = 0.$$

From this obviously follows that $\mu_i = k - \lambda_i$. We already put an ordering on the eigenvalues λ_i of A so naturally the eigenvalues μ_i of Δ have the inverse ordering.

iv. Because of the ordering on the eigenvalues μ_i it is clear that μ_0 is the smallest eigenvalue. However in the proof of Proposition 3.8 we saw that the eigenvalue λ_0 of A is associated with the eigenvector $(1, 1, \dots, 1)$ and therefore so is μ_0 . However, all eigenvectors of a symmetric matrix are perpendicular to one another. This means that the second smallest eigenvalue μ_1 of Δ is the smallest eigenvalue so that its eigenvector \mathbf{p} is orthogonal to $(1, 1, \dots, 1)$.

v. By Theorem 3.9 there is an orthogonal basis of eigenvectors for Δ . Let us call them $\mathbf{v}_0, \dots, \mathbf{v}_{n-1}$, associated to μ_0, \dots, μ_{n-1} . As \mathbf{p} is orthogonal to \mathbf{v}_0 , \mathbf{p} is spanned by the other eigenvectors. So we can write

$$\mathbf{p} = \sigma_1 \mathbf{v}_1 + \dots + \sigma_{n-1} \mathbf{v}_{n-1}$$

for some constants $\sigma_1, \dots, \sigma_{n-1}$. From Definition 3.11 and iv follows

$$\begin{aligned}
\Delta \mathbf{p} &= \Delta(\sigma_1 \mathbf{v}_1 + \dots + \sigma_{n-1} \mathbf{v}_{n-1}) \\
&= \mu_1 \sigma_1 \mathbf{v}_1 + \dots + \mu_{n-1} \sigma_{n-1} \mathbf{v}_{n-1} \\
&\geq \mu_1 \sigma_1 \mathbf{v}_1 + \dots + \mu_1 \sigma_{n-1} \mathbf{v}_{n-1} \\
&= \mu_1 \mathbf{p}
\end{aligned}$$

□

A very important notion is the **spectral gap** of a k -regular graph X . This spectral gap is $\lambda_0 - \lambda_1$ or $k - \lambda_1$ where $\lambda_0, \dots, \lambda_{n-1}$ is the spectrum of X . While it might not seem obvious why this is important, quite some theorems depend on it, like the following.

Theorem 3.14.

$$h(X) \geq \frac{k - \lambda_1}{2}$$

Proof. We begin by taking $F \subsetneq V$. We will use the vector $\mathbf{p} \in \mathbb{R}^n$ which is constructed as followed:

$$p_v = \begin{cases} -|V \setminus F| & \text{if } v \in F \\ |F| & \text{if } v \in V \setminus F \end{cases}$$

where we can assume that $|F| \leq \frac{|V|}{2} = \frac{n}{2}$. In this case

$$\langle \mathbf{p}, (1, 1, \dots, 1) \rangle = -|V \setminus F||F| + |F||V \setminus F| = 0.$$

We have then with Proposition 3.13.v

$$\|D\mathbf{p}\|^2 = \langle \mathbf{p}, \Delta\mathbf{p} \rangle \geq \langle \mathbf{p}, \mu_1\mathbf{p} \rangle = \mu_1\|\mathbf{p}\|^2 \quad (3)$$

and

$$\langle \mathbf{p}, \mathbf{p} \rangle = |V \setminus F|^2|F| + |F|^2|V \setminus F| = |F||V \setminus F|(|F| + |V \setminus F|) = |F||V \setminus F||V|.$$

\mathbf{p} is defined so that

$$D\mathbf{p}_e = \begin{cases} 0 & \text{if } o(e) \in F, t(e) \in F \text{ or } o(e) \in V \setminus F, t(e) \in V \setminus F, \\ \pm|V| & \text{otherwise.} \end{cases}$$

Hence and due to Eq. (3),

$$\|D\mathbf{p}\|^2 = |V|^2|E(F, V \setminus F)| \geq \mu_1|F||V \setminus F||V| \quad (4)$$

Then by Definition 3.3 and Eq. (4) and with our assumption that $|F| \leq \frac{n}{2}$

$$h(X) \geq \frac{E(F, V \setminus F)}{|F|} \geq \mu_1 \frac{|V \setminus F|}{|V|} \geq \frac{k - \lambda_1}{2}.$$

□

Theorem 3.14 gives us a very important result namely that how higher spectral gap of a graph X is, the better is the expanding quality of X . This transforms the problem of finding expanders from graph analysis to finding matrices with certain properties with linear algebra, which is much easier to deal with.

We can for instance look at the spectral gap for K_5 and Circ_5 . One can do this by first calculating their adjacency matrix and then calculating the eigenvalues of their adjacency matrix. $\lambda_0^{K_5} = 4$, $\lambda_1^{K_5} = -1$ which is a spectral gap of 5, while $\lambda_0^{\text{Circ}_5} = 2$ and $\lambda_1^{\text{Circ}_5} = (\sqrt{5} - 1)/2 \sim 0.62$ which gives a spectral gap of $(5 - \sqrt{5})/2 \sim 1.38$. This is a lot smaller. We already saw intuitively that K_5 was a better expander than Circ_5 so this makes sense.

4 Constructing an expander

The hardest part about expanders is really the construction. How do we make a family of graphs with a good expansion? There are roughly three different constructions:

1. the Margulis construction [Mar73], which was improved multiple times, notably by Gabber-Galil [GG81]. This construction is easy to perform but it does not give a really good expansion.
2. the Lubotzky, Phillips, and Sarnak construction [LPS86], using the Ramanujan conjecture.
3. The construction of Moshe Morgenstern [Mor91] which relies heavily on ring theory and results in a bounded concentrator.

It would be too elaborate to construct all those expanders and give proofs. Instead, we will construct one expander like Margulis' and calculate its second smallest eigenvalue. We will also shed some light on how the Morgenstern construction works.

4.1 Margulis construction

The construction is based on $\mathbb{Z}_m \times \mathbb{Z}_m$ and consequently the resulting expander $X(V, E)$ has m^2 vertices. Furthermore, the expander is 8-regular. We assumed in section 2 that graphs contain no loops and multiple edges and that we were dealing with directed graphs. The Margulis construction does need loops, multiple edges and is undirected to guarantee that the graph is 8-regular. This seemingly creates some difficulties as we assumed different

throughout section 2. However, if we look closely we will see that this assumption is only used in Proposition 3.6 and this proposition is not referred to elsewhere in the section. The difference between the undirected and directed graphs is easily solved: substitute every undirected edge with a directed edge and its inverse.

The edges of X are constructed so that for every vertex $v = (x, y) \in V$ the neighbours of v are given by the permutations π_0, \dots, π_3 :

$$\begin{aligned}\pi_0(v) &= (x + 2y, y), \\ \pi_1(v) &= (x + 2y + 1, y), \\ \pi_2(v) &= (x, 2x + y), \\ \pi_3(v) &= (x, 2x + y + 1)\end{aligned}$$

where addition is modulo m . These create 4 edges, one also has to find their inverses which creates a total of 8 edges per vertex. Actually this happens automatically if one calculates all the edges from other vertices. Note that if $\pi_i(v) = v$ for some $0 \leq i \leq 3$, $v \in V$ then two loops on v are created: one for π_i and one for π_i^{-1} .

$\pi_0(0, 0) = (0, 0)$	$\pi_0(0, 1) = (2, 1)$	$\pi_0(0, 2) = (1, 2)$
$\pi_1(0, 0) = (1, 0)$	$\pi_1(0, 1) = (0, 1)$	$\pi_1(0, 2) = (2, 2)$
$\pi_2(0, 0) = (0, 0)$	$\pi_2(0, 1) = (0, 1)$	$\pi_2(0, 2) = (0, 2)$
$\pi_3(0, 0) = (0, 1)$	$\pi_3(0, 1) = (0, 2)$	$\pi_3(0, 2) = (0, 0)$
$\pi_0(1, 0) = (1, 0)$	$\pi_0(1, 1) = (0, 1)$	$\pi_0(1, 2) = (2, 2)$
$\pi_1(1, 0) = (2, 0)$	$\pi_1(1, 1) = (1, 1)$	$\pi_1(1, 2) = (0, 2)$
$\pi_2(1, 0) = (1, 2)$	$\pi_2(1, 1) = (1, 0)$	$\pi_2(1, 2) = (1, 1)$
$\pi_3(1, 0) = (1, 0)$	$\pi_3(1, 1) = (1, 1)$	$\pi_3(1, 2) = (1, 2)$
$\pi_0(2, 0) = (2, 0)$	$\pi_0(2, 1) = (1, 1)$	$\pi_0(2, 2) = (0, 2)$
$\pi_1(2, 0) = (0, 0)$	$\pi_1(2, 1) = (2, 1)$	$\pi_1(2, 2) = (1, 2)$
$\pi_2(2, 0) = (2, 1)$	$\pi_2(2, 1) = (2, 2)$	$\pi_2(2, 2) = (2, 0)$
$\pi_3(2, 0) = (2, 2)$	$\pi_3(2, 1) = (2, 0)$	$\pi_3(2, 2) = (2, 1)$

Table 1: Transformations of the Margulis construction

Let us perform such a construction of such a graph X based on $\mathbb{Z}_3 \times \mathbb{Z}_3$. For all the transformations refer to Table 1. With these transformations we can also make a diagram. This is displayed in Figure 7. The adjacency

matrix A of X becomes

$$A = \begin{pmatrix} 4 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 2 & 0 & 0 & 2 \\ 1 & 0 & 0 & 4 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 4 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 & 2 & 0 & 0 & 2 \\ 1 & 0 & 0 & 1 & 0 & 0 & 2 & 2 & 2 \\ 0 & 1 & 0 & 0 & 1 & 0 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 & 0 & 2 & 2 & 2 & 0 \end{pmatrix},$$

when we order the vertices as follows: $(0, 0), (1, 0), (2, 0), (0, 1), \dots, (2, 2)$. One can check that the second greatest eigenvalue is $1 + \sqrt{13} \sim 4.60555$, while the greatest eigenvalue is 8, ofcourse. Using Proposition 3.14

$$h(X) \geq \frac{8 - (1 + \sqrt{13})}{2} = \frac{7 - \sqrt{13}}{2} \sim 1.69722.$$

When we use the algorithm from Subsection 3.1.1 we find that

$$h(X) \leq 2.$$

This combined gives us a good estimate of $h(X)$.

Notice that the diagram is symmetrical, which is caused by the transformation being symmetrical.

Theorem 4.1. *Let an 8-regular graph X be a Margulis construction on $\mathbb{Z}_m \times \mathbb{Z}_m$, $m \in \mathbb{N}$, and $\lambda_0, \dots, \lambda_{n-1}$ its spectrum. Then*

- i. $\lambda_1 \leq 5\sqrt{2} \sim 7.071$,
- ii. X is an $(m^2, 8, 1 - \frac{5}{4\sqrt{2}})$ -expander.

Proof. i. A comparibly easy proof using direct Fourier transformation is given by [LWd03].

- ii. Due to Proposition 3.4.ii and i, X is an $(m^2, 8, \frac{h(X)}{8})$ -expander, where due to Proposition 3.14

$$\frac{h(X)}{8} \geq \frac{8 - \lambda_1}{16} \geq 1 - \frac{5}{4\sqrt{2}} \sim 0.116.$$

□

Defining what a good (n, k, c) -expander is, is a bit arbitrary, one has to pick a constant c above which one says the expander is good. The value $c = 1 - \frac{5}{4\sqrt{2}}$ is not generally considered very good, however for quite some time this was the best expander available.

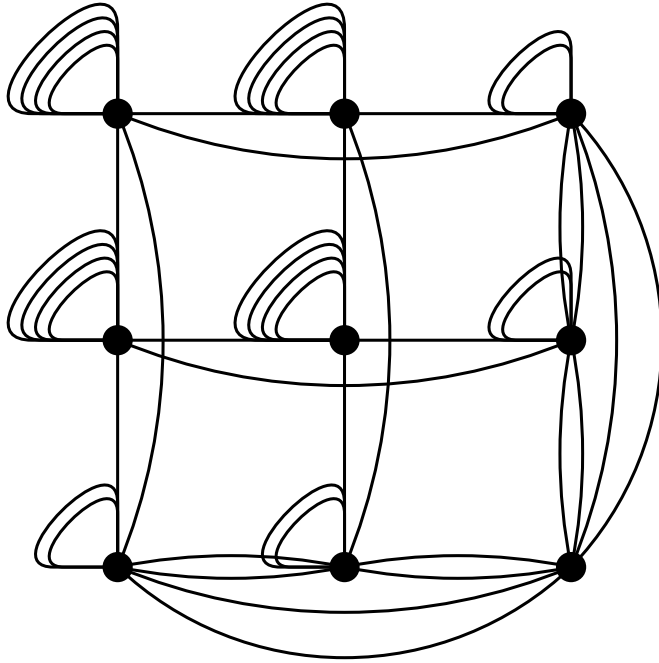


Figure 7: Diagram of Margulis construction

4.2 Morgenstern construction

For the remainder of this section some background in ring theory is necessary. One might also read chapter 2.1 of [Ser80] which [Mor91] is strongly based on as this shows more background than we will now.

In a Morgenstern construction the resulting graph is $q + 1$ -regular, where we can choose $q \geq 5$ prime. If we follow a later improvement by [Sch98] we can choose $q \geq 3$ prime. While this proof is even more difficult than that of Morgenstern we will gladly take the opportunity to take $q = 3$ as this will save us a lot of work. Once we have chosen q , we can define some fields we will be working on, most importantly we will work with \mathbb{F}_q (the field of q elements), in our case \mathbb{F}_3 . We also have to choose a polynomial $g(x) \in \mathbb{F}_3[x]$, the polynomials in one variable x , so that $\text{degree}(g) \geq 2$. We have chosen to use [Sch98] in which case it is necessary to have $\text{degree}(g) = 2$. Analysis is simplified if we choose an irreducible polynomial over \mathbb{F}_3 ; we will choose $g(x) = x^2 + 1$.

The construction is fairly simple, while the background of this construction is not. Sadly there was not enough time to shed more light on the background.

Recall that $GL_2(R)$ (General Linear group) are the invertible 2×2 -matrices over some field R . An addition to this is $PGL_2(R)$: the Projective General Linear group. These are the invertible 2×2 -matrices divided by their centre. Recall that the centre of $GL_2(R)$ is composed as followed:

$$Z(GL_2(R)) = \{A \in GL_2(R) \mid AM = MA, \forall M \in GL_2(R)\} = \{rI \mid \forall r \in R^*\} \quad (5)$$

Let us recall how quotients on matrix groups work again:

$$M \cong MrI \cong rM, \quad M \in PGL_2(R) = GL_2(R)/Z(GL_2(R)), \quad r \in R^*.$$

Let $R = \mathbb{F}_3[x]/g(x)\mathbb{F}_3[x]$ and

$$H = PGL_2(R).$$

Here $\mathbb{F}_3[x]/g(x)\mathbb{F}_3[x]$ are the polynomials over \mathbb{F}_3 modulo $g(x)$. This is a field with 9 elements, which are the following:

$$R = \mathbb{F}_3[x]/g(x)\mathbb{F}_3[x] = \{\bar{0}, \bar{1}, \bar{2}, \bar{x}, \overline{x+1}, \overline{x+2}, \overline{2x}, \overline{2x+1}, \overline{2x+2}\}. \quad (6)$$

Now let

$$A = \left\{ \begin{pmatrix} a & b+cx \\ 0 & 1 \end{pmatrix} \mid a \in \mathbb{F}_3^*; b, c \in \mathbb{F}_3 \right\}$$

and $B = PGL_2(\mathbb{F}_3)$. We will create a bipartite graph $X(I, O, E)$ with $I = H/A$ and $O = H/B$. We write aA for a vertex in I , with $a \in H$. Similarly we can write bB for a vertex in O , with $b \in H$. We now calculate $|I|$ and $|O|$ by calculating the orders of some of the associated groups. First of all

$$|R| = 9,$$

as we see in Eq. (6). Then

$$|GL_2(R)| = (9^2 - 1)(9^2 - 9) = 9(9^2 - 1)(9 - 1),$$

which we can see as the number of different combinations we can make in the first column (which is all but $(0, 0)^T$), times the number of combinations can make linearly independent in the second column.

The order of the centre follows from Eq. (5):

$$|Z(GL_2(R))| = 9 - 1$$

which makes

$$|H| = |PGL_2(R)| = |GL_2(R)|/|Z(GL_2(R))| = 9(9^2 - 1) = 9 \cdot 80.$$

Fairly straightforward is

$$|A| = |\mathbb{F}_3| |\mathbb{F}_3|^2 = 3^2(3-1) = 9 \cdot 2,$$

and

$$|B| = |GL_2(\mathbb{F}_3)| / |Z(GL_2(\mathbb{F}_3))| = (3^2 - 1)(3^2 - 3) / (3 - 1) = 3 \cdot 8.$$

This means

$$|I| = |H| / |A| = \frac{9 \cdot 80}{9 \cdot 2} = 40$$

and

$$|O| = |H| / |B| = \frac{9 \cdot 80}{3 \cdot 8} = 30.$$

So we know we have to draw 40 vertices at one side and 30 vertices at the other. Now the edges are a different story. We have to connect vertices aA and bB if $aA \cap bB \neq \emptyset$, or otherwise put, if there is a $h \in H$ so that

$$h \cong aA \quad \text{and} \quad h \cong bB.$$

One can imagine that finding all edges in this way is a frightful task to do by hand. Luckily Gunther Cornelissen had already done this based on a computer program to generate the adjacency matrix of the Morgenstern graph written by the group of E.-U. Gekeler at Saarbrücken, see Figure 8 for a diagram.

Theorem 4.2. *For $q \geq 5$ the Morgenstern construction yields an $(|H/A|, \frac{q}{q+1}, q+1, \frac{q-4}{q-3})$ -bounded concentrator. With the improvement in [Sch98], for $q \geq 3$ the construction yields a $(|H/A|, \frac{q}{q+1}, q+1, \frac{q-2}{q-1})$ -bounded concentrator.*

Proof. We refer to [Mor91] and [Sch98] for the proof. \square

We have used a polynomial g of the smallest possible degree. If we repeat the arguments from above with g of degree d for $d \geq 2$ we will see that

$$|H| = |GL_2(R)| / |Z(GL_2(R))| = \frac{(q^{2d} - 1)(q^{2d} - q^d)}{q^d - 1} = q^d(q^{2d} - 1)$$

while A and B are independent of d . This means that the size of the expander is exponential in d and one will not likely choose a large d . What one can do is choose different g_1, g_2, \dots, g_l with $\text{degree}(g_i) = d$. This will result in different expanders, although they will all be the same size. We see that with

$$|H/A| = \frac{q^d(q^{2d} - 1)}{q^2(q - 1)} = q^{d-2}(1 + q + \dots + q^{2d-1}),$$

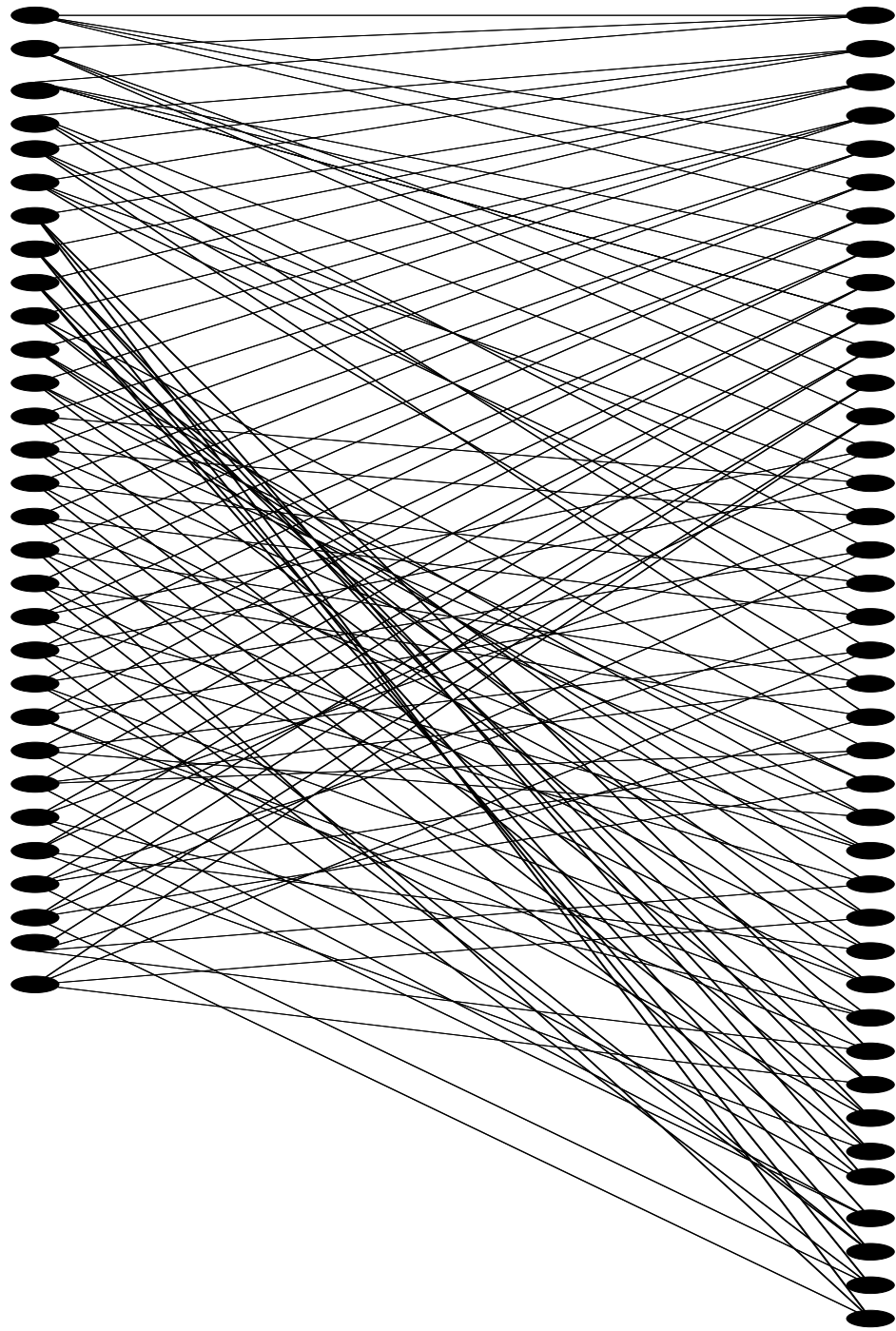


Figure 8: Diagram of Morgenstern construction

the size of I is polynomial in q . So instead of raising d we could also raise q . However this creates a problem: the family of expanders no longer is k -regular, as $k = q + 1$. As the number of vertices get larger, the proportion of edges also rises. This goes against our principle of not having to use too many edges.

So while the Morgenstern construction leads to far better expanding constants than Margulis', usability remains limited if we need a graph of size n . However, if the exact size is not of much importance while the expanding quality is, this construction is a very good contestant. Note that one can make different constructions with the same size by varying g but keeping it the same degree.

5 An expander as randomizer

As we mentioned in Section 3 an expander can be used to reduce the need of a true randomizer. Let us first explain how random numbers in computers work and give an example of an algorithm that requires random bits.

5.1 Random number generators

While a random algorithm requires random numbers, to measure the randomness needed we will count the bits it outputs (where a bit is 0 or 1). In a computer, bits seeming random are generated by all kinds of system events, these can be the mouse moving or the disk doing some work. All these events are deterministic, though, so while these bits may look random and act random (based on all previous bits you have seen, the chance that you guess the next bit right is about $1/2$) they really are not, which is why we call these bits pseudorandom bits. There are a lot of different **pseudorandom number generators (PRNG)**, all using a deterministic algorithm to create a next pseudorandom bit. Some PRNG's are less random or slow, while others nearly fulfill the job.

There are also hardware **random number generators (RNG)**, these are based on quantumphysics which is non-deterministic in certain aspects and because of this feature can create truly random bits.

For both the PRNG and the RNG it takes time to gather random bits, this is why computer scientists want algorithms that use random bits sparingly if they need them.

5.2 Primality test

Lets say we want to test if an odd number N is a prime, say, for cryptographic purposes. There are quite some primality tests from number theory which can give an correct answer to this question *with a certain probability*. One has to repeat the test with different inputs to be more sure of the correctness of the answer. Since a few years there also exists a polynomial primality test which does not require random numbers by [AKS02], but follow me through for the sake of the argument.

Let $f : R \mapsto \{0, 1\}$, where R is a set, be a function where $f(x)$ determines whether $x \in \text{LANG}$. $\text{LANG} \subset R$ can be any set, in our case R is the set of integers and LANG is the set of all primes, so $f : \mathbb{Z} \mapsto \{0, 1\}$. There is a complexity class of functions called **Random Polynomial (RP)**. Functions in RP are functions $g : R \times \{0, 1\}^m \mapsto \{0, 1\}$ that have the property that if $f(x) = 0$ then $g(x, r) = 0$ and if $f(x) = 1$ then $\mathbb{P}[g(x, r) = 1] = c$ where $0 < c < 1$. Likewise there is a class called Co-RP where functions $h : R \times \{0, 1\}^m \mapsto \{0, 1\}$ have the property that if $f(x) = 0$ then $\mathbb{P}[h(x, r) = 0] = d$ where $0 < d < 1$ and if $f(x) = 1$ then $h(x, r) = 1$.

One of the primality tests is the **Rabin test**, which is taken from and is further explained in chapter [Beu99].

Theorem 5.1 (Rabin). *Let N be an odd integer and $a \in \mathbb{Z}$ not a multiple of N . Suppose $N - 1 = 2^k \cdot m$ with m odd and $k \in \mathbb{N}$. If*

$$a^m \not\equiv 1 \pmod{N} \quad \text{and} \quad \forall 0 \leq j \leq k - 1 : a^{2^j \cdot m} \not\equiv -1 \pmod{N}$$

then N is composite. Furthermore if N is composite then the equation is true for at least $3/4$ of choices of $a \in \{1, 2, \dots, N - 1\}$.

If the Rabin test is true for an a , we call a a witness for N being composite. The problem with this theorem is that if a is a bad witness, we can not be sure that N is prime until we have tested more than $1/4$ of the set $\{1, 2, \dots, N - 1\}$, and this will take too long for large N . We could try some numbers deterministically but there might be a system in the numbers for which the test is false. We should better choose random numbers in the set $\{1, 2, \dots, N - 1\}$. In this case the probability that N is composite if the test is false for one number is $1/4$. The probability that N is composite after 20 tests which turned false with different numbers is $(1/4)^{20} \sim 10^{-12}$. This means the probability that N is prime is about $1 - 10^{-12} = 0.999999999999$.

Theorem 5.1 also shows that if used with random numbers the Rabin test is in RP for testing whether N is composite and is in Co-RP for testing whether N is prime: let $f(x)$ determine whether x is composite and let $g(x, r)$ be the Rabin test, then $g(x, r) = 0$ if $f(x) = 0$ (x is not composite) and if

$f(x) = 1$ then $\mathbb{P}[g(x, r) = 1] = 3/4$. Each Rabin test needs $\lceil \log_2 N \rceil$ random bits, for the choice of a . This does not depend on the a that actually gets chosen. If a is picked at random and is small then first a lot of zero-bits were randomly chosen and then some one-bits. If a is large then first some one-bits were randomly chosen and then a lot more bits.

5.3 Random walk on an expander

Now we can use the ‘randomness’ of an expander. This randomness lies in the fact that when we perform a random walk of length s on a k -regular expander X it seems in some ways that we have randomly chosen s independent vertices of the graph. We need less random bits for a random walk though: we need about $s^2 \log k$, instead of about $s^2 \log |X|$.

5.3.1 Some notation

To do some analysis on random walks we will introduce some notions. First of all we will use a second norm on vectors $\mathbf{v} \in \mathbb{R}^n$:

$$\|\mathbf{v}\|_1 := \sum_{i=1}^n |v_i|,$$

while we will write $\|\mathbf{v}\|_2$ for the Euclidean norm.

Definition 5.2. Let $\mathbf{v} \in \mathbb{R}^n$ so that $1 \leq i \leq n$, $v_i \geq 0$. Then \mathbf{v} is a probability vector if $\|\mathbf{v}\|_1 = 1$ and \mathbf{v} is distribution vector if $\|\mathbf{v}\|_1 \leq 1$.

By $\mathbf{u} \in \mathbb{R}^n$ we denote the uniform distribution, that is, $\mathbf{u} = \frac{1}{n}(1, \dots, 1)$.

We make the adjacency matrix A of X better compatible for working with probabilities by defining $\hat{A} = \frac{1}{k}A$. This way all columns and rows of \hat{A} add up to 1. Of course, all eigenvalues of \hat{A} are those of A divided by k .

From Theorem 3.14 we know that the spectral gap is a measurement for how good an expander is. Taking that in mind we take $\alpha = \max\{|\hat{\lambda}_1|, |\hat{\lambda}_{n-1}|\}$.

5.3.2 Random walk

Let now our expander X have $N - 1$ vertices: one for every a we can choose for our Rabin test. If N is composite we have a collection of vertices B which produces bad witnesses (these say N could be prime while it is not). From Theorem 5.1 we know that $|B| \leq \frac{|X|}{4}$. We call (B, s) the event that a random walk of length $s \in \mathbb{Z}_{\geq 0}$, which starts at a random vertex, only visits B . We need this to have a small probability, because if we visit one or more vertices

outside B , witnesses, we will know at once that N is composite. This leads to the main proposition of this section.

Proposition 5.3.

$$\mathbb{P}[(B, s)] \leq (1/4 + \alpha)^s.$$

From this we have again a bound on the spectral gap of X . If $\hat{\lambda}_1$ is close to $3/4$ we have a very large probability that the random walk visits only B . In other words, we want the spectral gap to be as large as possible, it should at least be larger than $1/4$.

We need two lemmas to prove Proposition 5.3. For this we need to define a projection matrix P which sends a distribution vector to a distribution inside B :

$$P_{ij} = \begin{cases} 1 & \text{if } i = j \in B \\ 0 & \text{otherwise} \end{cases}.$$

Lemma 5.4.

$$\mathbb{P}[(B, s)] = \|(P\hat{A})^s P\mathbf{u}\|_1.$$

Proof. We start with the uniform distribution \mathbf{u} , as we take a random vertex of X , where each vertex has the same probability of being chosen. We are interested in whether that first vertex was in B which is obviously $\|P\mathbf{u}\|_1$, so for $s = 0$ the equality holds. The effect of \hat{A} on a distribution vector is walking one step forward in every direction, so the effect of $P\hat{A}$ is walking one step but just preserving the chances that this step will end up in B . This means $(P\hat{A})^s P\mathbf{u}$ results in the distribution vector of walking s steps only through B . To extract the probability of this event is merely taking $\|(P\hat{A})^s P\mathbf{u}\|_1$. \square

Lemma 5.5. *For any distribution vector \mathbf{v} :*

$$\|P\hat{A}P\mathbf{v}\|_2 \leq (1/4 + \alpha) \cdot \|\mathbf{v}\|_2$$

Proof. First we decompose $P\mathbf{v}$ into two parts:

$$P\mathbf{v} = (P\mathbf{v})_c + (P\mathbf{v})_s$$

where $(P\mathbf{v})_c = b \cdot \mathbf{u}$ for some constant b and $\langle (P\mathbf{v})_c, (P\mathbf{v})_s \rangle = 0$. From the triangle inequality we have

$$\|P\hat{A}P\mathbf{v}\|_2 = \|P\hat{A}((P\mathbf{v})_c + (P\mathbf{v})_s)\|_2 \leq \|P\hat{A}(P\mathbf{v})_c\|_2 + \|P\hat{A}(P\mathbf{v})_s\|_2$$

First we will look at how we can simplify $(P\mathbf{v})_c$. As \mathbf{u} is an eigenvector of \widehat{A} with eigenvalue 1 and $(P\mathbf{v})_c$ is a multiple of \mathbf{u} we have

$$\|P\widehat{A}(P\mathbf{v})_c\|_2 = \|P(P\mathbf{v})_c\|_2.$$

And because P reduces $3/4$ of a vector to 0 and $(P\mathbf{v})_c$ is a multiple of \mathbf{u} we have

$$\|P(P\mathbf{v})_c\|_2 = \sqrt{1/4}\|(P\mathbf{v})_c\|_2 \leq 1/4\|\mathbf{v}\|_2.$$

The last inequality is due to the fact that $(P\mathbf{v})_c$ is maximized if \mathbf{v} is a multiple of \mathbf{u} .

Now we will look at how we can simplify $(P\mathbf{v})_s$. Remember that we took $\alpha = \max\{|\hat{\lambda}_1|, |\hat{\lambda}_{n-1}|\}$ and use Proposition 3.13.v. This leads to

$$\|P\widehat{A}(P\mathbf{v})_s\|_2 \leq \|\widehat{A}(P\mathbf{v})_s\|_2 \leq \alpha\|(P\mathbf{v})_s\|_2 \leq \alpha\|\mathbf{v}\|_2.$$

Bringing this together again we get

$$\|P\widehat{A}P\mathbf{v}\|_2 \leq \|P\widehat{A}(P\mathbf{v})_c\|_2 + \|P\widehat{A}(P\mathbf{v})_s\|_2 \leq (1/4 + \alpha)\|\mathbf{v}\|_2.$$

□

Bringing these two lemmas together creates a fairly straightforward proof of Proposition 5.3.

Proof of Proposition 5.3.

$$\begin{aligned} \mathbb{P}[(B, s)] &= \|(P\widehat{A})^s P\mathbf{u}\|_1 \\ &\leq \sqrt{n} \cdot \|(P\widehat{A})^s P\mathbf{u}\|_2 \\ &= \sqrt{n} \cdot \|(P\widehat{A}P)^s \mathbf{u}\|_2 \\ &\leq \sqrt{n} \cdot (1/4 + \alpha)^s \|\mathbf{u}\|_2 \\ &= (1/4 + \alpha)^s \end{aligned}$$

□

5.3.3 Resource analysis

We now have Proposition 5.3, does this help us? For the Rabin test using the random walk we need $\lceil \log_2(N-1) \rceil$ random bits for a first vertex and $\lceil \log_2 k \rceil$ random bits for every next vertex. This yields a smaller amount of random bits needed than picking a random vertex every time. It costs more computational power, though, as it also needs more steps to get the

same odds: while picking s random vertices give an error probability of 4^{-s} , performing a random walk of s steps gives an error probability of $(1/4 + \alpha)^s$ (the error in the Rabin test is that a composite number is recognised as prime). Another important point we have not discussed is how to get all the neighbours of a vertex, given a vertex in a good k -regular expander, where each vertex knows which number it represents. One solution would be to construct an expander of $N - 1$ vertices and assigning numbers to each of the vertices. For a very large N this is considered too cumbersome, we might just as well run the Rabin test $1/4N$ times. We could do a $\mathbb{Z}_{N-1} \times \mathbb{Z}_{N-1}$ Margulis construction and use the x component of each vertex as a for the Rabin test. However, the bound obtained from Theorem 4.1 yields $\alpha \geq \frac{5\sqrt{2}}{8} > 3/4$ which makes it useless. With the Morgenstern construction it is not possible to exactly specify the needed vertices which also does not make it a good candidate. However, at the time, other solutions for random walks do not seem to be possible in the literature. This keeps the random walks on expander graphs largely in the theoretic domain.

We could use a random walk on an expander to produce a PNRG. This will not always give the desired result though, as the random bits produced are not totally random. Suppose we use a random walk on a 5-regular expander as PRNG. If this expander was constructed in a deterministic manner and the m bits of one vertex are output then there are only 5 different options for the next m bits, instead of 2^m . For algorithms in RP (like primality tests) this is good enough (as we have assessed a good bound with a random walk), for a lot of real life applications it is not. In cryptography, for example, it is a *lot* easier to guess a part of an encrypted passwords out of 5 different choices than to guess it out of 2^m options. If the expander is built randomly for each use, we will need a lot more random bits than if we don't use an expander at all, or, if we use this random expander multiple times, a pattern will get visible.

6 Afterword

I very much enjoyed studying expanding graphs as it became obvious to me that this was quite a multidisciplinary subject, both within mathematics (from ring theory to harmonic functions) and between mathematics and computational sciences (from construction to application). It also became obvious that as a bachelor student a lot of the theory goes beyond my reach, combined with the fact that a I could not treat a lot of the applications of expanders.

Something also worth noting is that the applications I did come across

could not yet be implemented with the current knowledge about expanders, so there is quite some work to be done in this field yet.

I would like to thank Gunther Cornelissen for assisting me with my thesis and his lessons on [Ser80] (I hope I'll get it some day...)

References

- [AKS02] M. Agrawal, N. Kayal, and N. Saxena. Primes is in p , 2002.
- [Beu99] Frits Beukers. *Getaltheorie voor beginners*. Epsilon Uitgaven, 1999.
- [DSV03] Guiliana Davidoff, Peter Sarnak, and Alain Valette. *Elementary Number Theory, Group Theory, and Ramanujan Graphs*. Number 55 in London Mathematical Society Student Texts. Cambridge University Press, 2003.
- [FB95] Fraleigh and Beauregard. *Linear Algebra*. Addison-Wesley, third edition, 1995.
- [GG81] Ofer Gabber and Zvi Galil. Explicit constructions of linear-sized superconcentrators. *Journal of Computer and System Sciences*, 22:307–420, 1981.
- [LPS86] A. Lubotzky, R. Phillips, and P. Sarnak. Explicit expanders and the ramanujan conjectures. In *Proceedings of the Eighteenth Annual Symposium on the Theory of Computing*. ACM, 1986.
- [Lub94] Alexander Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures*. Number 125 in Progress in Mathematics. Birkhäuser Verlag, 1994.
- [LWd03] Nati Linial and Avi Wigderson. Expander graphs and their applications, Jan 2003. Lecture notes from a course on expanding graphs at the Hebrew University, Israel.
- [Mar73] G. A. Margulis. Explicit construction of concentrators. *Problemy Peredači Informacii*, 9(4):71–80, 1973. English translation in Problems Inform. Transmission (1975).
- [Mor91] Moshe Morgenstern. Explicit construction of natural bounded concentrators. In *32nd Annual Symposium on Foundations of Computer Science*, pages 392–404. IEEE Computer Society Press, Oct 1991.

- [Sch98] Ortwin Scheja. On zeta functions of arithmetically defined graphs. *Finite Fields and Their Applications*, 5:314–343, 1998.
- [Ser80] Jean-Pierre Serre. *Trees*. Springer-Verlag, 1980.

Index

- (n, k, c) -expander, 8
- (n, θ, k, α) -bounded concentrator, 9
- adjacency matrix, 13
- adjacent, 4
- bipartite, 7
- Cayley graph, 7
- Cheeger constant, 9
- Circ_n , 6
- circuit, 5, 6
- connected, 6
- distance function, 7
- $E(A, B)$, 5
- edge, 4
- expanders, 8
- family of expanding graphs, 10
- graph, 4
- isomorphic, 5
- k -regular, 5
- Laplacian, 14
- order, 5
- orientation, 5
- Path_n , 6
- path, 5, 6
- PRNG, 26
- pseudorandom number generators,
26
- Rabin test, 27
- random number generators, 26
- Random Polynomial, 27
- RNG, 26
- RP, 27
- spectral gap, 18
- spectrum, 14
- subgraph, 5
- tree, 7
- vertex, 4