# A BRIEF INTRODUCTION TO SPECTRAL GRAPH THEORY

CATHERINE BABECKI, KEVIN LIU, AND OMID SADEGHI
MATH 563, SPRING 2020

Abstract. There are several matrices that can be associated to a graph. Spectral graph theory is the study of the spectrum, or set of eigenvalues, of these matrices and its relation to properties of the graph. We introduce the primary matrices associated with graphs, and discuss some interesting questions that spectral graph theory can answer. We also discuss a few applications.

## 1. Introduction and Definitions

This work is primarily based on [1]. We expect the reader is familiar with some basic graph theory and linear algebra. We begin with some preliminary definitions.

**Definition 1.** *Let $\Gamma$ be a graph without multiple edges. The **adjacency matrix** of $\Gamma$ is the matrix $A$ indexed by $V(\Gamma)$, where $A_{xy} = 1$ when there is an edge from $x$ to $y$, and $A_{xy} = 0$ otherwise. This can be generalized to multigraphs, where $A_{xy}$ becomes the number of edges from $x$ to $y$.*

**Definition 2.** *Let $\Gamma$ be an undirected graph without loops. The **incidence matrix** of $\Gamma$ is the matrix $M$, with rows indexed by vertices and columns indexed by edges, where $M_{xe} = 1$ whenever vertex $x$ is an endpoint of edge $e$. For a directed graph without loss, the **directed incidence matrix** $N$ is defined by $N_{xe} = -1, 1, 0$ corresponding to when $x$ is the head of $e$, tail of $e$, or not on $e$.*

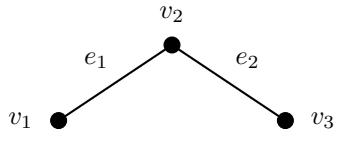
**Definition 3.** *Let $\Gamma$ be an undirected graph without loops. The **Laplace matrix** of $\Gamma$ is the matrix $L$ indexed by $V(G)$ with zero row sums, where $L_{xy} = -A_{xy}$ for $x \neq y$. If $D$ is the diagonal matrix indexed by $V(\Gamma)$ such that $D_{xx}$ is the degree of $x$, then $L = D - A$. The matrix $Q = D + A$ is called the **signless Laplace matrix** of $\Gamma$.*

Both $Q$ and $L$ are positive semidefinite, since $Q = MM^T$ and $L = NN^T$, where a directed incidence matrix $N$ is obtained from an arbitrary orientation of $\Gamma$.

**Definition 4.** *The **(ordinary) spectrum** of a finite graph $\Gamma$ is the spectrum of the adjacency matrix $A$ with multiplicities. The **Laplace spectrum** of a finite undirected graph without loops is the spectrum of the Laplace matrix $L$ with multiplicities.*

Note that the spectrum and Laplace spectrum of a graph $\Gamma$ do not depend on the numbering the vertices chosen, as relabeling the vertices corresponds to permuting the rows and columns of $A$ and $L$.

**Example 1.** *Consider the graph $P_3$, the path on three vertices.*




*Labeling the graph as above, we obtain the matrices*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}, \text{ and } L = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

*The ordinary spectrum is* $\{\sqrt{2}, 0, -\sqrt{2}\}$*, and the Laplace spectrum is* $\{0, 1, 3\}$*.*

We can deduce several preliminary properties about these matrices through linear algebra. First, when $\Gamma$ is an undirected, simple graph with $n$ vertices, both $A$ and $L$ are real, symmetric matrices. Thus, each matrix has all real eigenvalues, and for each eigenvalue, the algebraic and geometric multiplicities coincide. That is, they each have a set of orthogonal eigenvectors which are a basis for $\mathbb{R}^n$.

For general graphs $\Gamma$, notice that $A$ has a zero diagonal. Its trace, and therefore sum of eigenvalues, is thus 0. We noted previously that $L$ is positive semidefinite since $L = NN^T$. Being positive semidefinite is equivalent to having non-negative eigenvalues. The Laplace matrix is also singular, since $L\mathbf{1} = \mathbf{0}$ where $\mathbf{1}$ is the vector of all ones. Hence, the eigenvalues of $L$ can be denoted $0 = \mu_1 \leq \mu_2 \leq ... \leq \mu_n$. Its trace (and therefore sum of eigenvalues) is

$$\sum_{v \in V} \deg(v) = 2|E|.$$

The signless Laplace matrix $Q = MM^T$ also has a real spectrum and nonnegative eigenvalues, but is not necessarily singular. Finally, notice that $\mathrm{Tr}(Q) = \mathrm{Tr}(L)$.

## 2. Spectra and Combinatorial Properties of Graphs

These matrices record a significant amount of combinatorial information of a graph. As one example, we can use the adjacency matrix to count walks.

**Proposition 1.** *Let* $\Gamma$ *be a simple, undirected graph, and let* $h$ *be a non-negative integer. Then* $(A^h)_{ij}$ *is the number of walks of length* $h$ *from vertex* $v_i$ *to* $v_j$*.*

*Proof.* We induct on $h$. Let $\Gamma$ be a graph with $n$ vertices, fix a labeling for its vertices $\{v_i\}_{i=1}^n$, and let $A$ be its adjacency matrix. For $h = 1$, $A_{ij}$ represents the number of edges from $v_i$ to $v_j$, which are equivalent to length 1 walks from $v_i$ to $v_j$.

Now, suppose the statement holds for some $h$. Writing out the matrix multiplication for $A^{h+1} = A^h \cdot A$, we see that

$$(A^{h+1})_{ij} = \sum_{k=1}^{n} (A^h)_{i,k} A_{k,j}.$$

By assumption, $(A^h)_{i,k}$ is the number of walks of length $h$ from $v_i$ to $v_k$, and $A_{k,j}$ is the number of length 1 walks from $v_k$ to $v_j$. Thus, their product is the number of walks of length $h+1$ from $v_i$ to $v_j$ such that the second to last vertex is $v_k$. Summing over all $k$ accounts for all cases. □

**Corollary 1.** *If* $\Gamma$ *is a simple, undirected graph,* $(A^2)_{ii}$ *is the degree of the vertex* $x$*, and* $TrA^2$ *equals the number of edges of* $\Gamma$*. Similarly,* $TrA^3$ *is six times the number of triangles in* $\Gamma$*.*

*Proof.* $A^2_{ii}$ counts the number of closed walks of length 2 from $v_i$ to itself. All such walks must be of the form $v_i \to v_j \to v_i$ for some $v_j$ adjacent to $v_i$. Thus there is one such walk for each edge adjacent to $v$. Now $A^3_{ii}$ counts the number of closed walks of length 3 from $v_i$ to itself. These must be of the form $v_i \to v_j \to v_k \to v_j$, and thus forms a triangle. $\mathrm{Tr}A^3$ then counts each triangle six times when accounting for the base point and the direction of the path. □

**Remark 1.** *The above statement and proof that* $(A^h)_{ij}$ *counts walks of length* $h$ *generalizes to multigraphs and directed graphs, though the corollary involving* $A^2$ *and* $A^3$ *do not.*

Another powerful combinatorial result is known as Kirchhoff's Matrix Tree Theorem. We omit the full proof, but give two outlines.

**Theorem 1** (Kirchhoff)**.** *Let* $\Gamma$ *be an undirected (multi)graph with at least one vertex and Laplace matrix* $L$ *with eigenvalues* $0 = \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$*. Let* $\ell_{xy}$ *be the* $(x, y)$ *cofactor of* $L$*. Let* $\tau(\Gamma)$ *be the number of spanning trees of* $\Gamma$*. Then,*

$$\tau(\Gamma) = \ell_{xy} = \det(L + \frac{1}{n^2} J) = \frac{1}{n} \mu_2 ... \mu_n.$$

For an edge $e$ which is not a loop in $\Gamma$, we can consider the graphs $\Gamma \setminus e$ and $\Gamma/e$, the deletion and contraction of the edge $e$, respectively. It is quick to show that $\tau(\Gamma) = \tau(\Gamma \setminus e) + \tau(\Gamma/e)$. Then, one can induct on on the number of edges.

This result also follows easily using the Cauchy-Binet theorem.

## 3. Spectra, Connectedness, and Bipartite Graphs

We next turn to a handful of results relating the spectra of a graph, connectedness, and bipartite-ness. This is possible largely because the spectra of a graph respects the structure of its connected components.

**Lemma 1.** *Let $\Gamma$ be a graph with connected components $\Gamma_i$. Then the spectrum of $\Gamma$ is the union of the spectra for $\Gamma_i$. The same holds for the Laplace and signless Laplace spectra.*

*Proof.* First, consider a graph $\Gamma$ with $k$ connected components. We show the result for $A$, and the proof follows similarly for $L$ and $Q$.

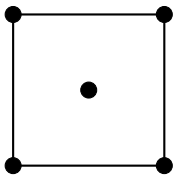
We can label the vertices of $\Gamma$ so that $A$ is block diagonal:

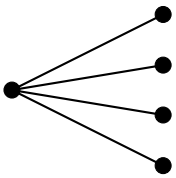
$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{bmatrix},$$

where $A_i$ is the adjacency matrix of the $i$-th connected component. Let $n = |V(\Gamma)|$ and $n_i = |V(\Gamma_i)|$. So we see that $\det(A - \lambda I_n) = 0$ if and only if $\prod_{i=1}^{k} \det(A_i - \lambda I_{n_i}) = 0$. □

Unfortunately, the ordinary spectrum cannot determine if a graph is connected.

**Example 2.** *Both $K_1 \sqcup C_4$ and $K_{1,4}$ have spectrum $\{-2, 0, 0, 0, 2\}$.*





$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

On the other hand, the Laplace spectrum is able to detect connectedness. Moreover, it encodes exactly how many connected components a graph has.

**Proposition 2.** *The multiplicity of 0 as an eigenvalue of the Laplace matrix of an undirected graph $\Gamma$ equals the number of connected components of $\Gamma$.*

*Proof.* Suppose $\Gamma$ is connected. Recall that $L = NN^T$, where $N$ is the incidence matrix of an arbitrary orientation of $\Gamma$. Then, $Lu = 0$ implies $0 = u^T NN^T u = \|N^T u\|^2$. Hence $N^T u = 0$. By definition, row $k$ of $N^T$ has a single entry $N^T_{k,i} = 1$ and $N^T_{k,j} = -1$, with all other components 0.

$$N^T u = \begin{bmatrix} N_{1,1} & \dots & N_{1,i} & \dots & N_{1,j} & \dots & N_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 1 & \dots & -1 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ N_{n,1} & \dots & N_{n,i} & \dots & N_{n,j} & \dots & N_{n,n} \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_j \\ \vdots \\ u_n \end{bmatrix}$$

Thus if there is an edge $ij \in E$, then $u_i = u_j$. Since $\Gamma$ is connected, there is an $ij$ path for each $i, j \in V$. Repeating this argument for the path shows that $u_i = u_j$ for all $i$ and $j$. Thus eigenvectors for 0 must be multiples of $\mathbf{1}$. For general graphs $\Gamma$, each connected component then contributes the 0 eigenvalue exactly once by Lemma 1. □

We can also recover information about whether a graph is bipartite.

**Proposition 3.** *The multiplicity of* 0 *as an eigenvalue of the signless Laplace matrix equals the number of bipartite connected components of* $\Gamma$.

*Proof.* This is done similarly as above. We first assume $\Gamma$ is connected. $Qu = MM^T u = 0$ implies $M^T u = 0$. Row $k$ of $M^T$ has a 1 in two entries $M^T_{k,i}$ and $M^T_{k,j}$, and 0 elsewhere. For the product to be 0, it must be that $u_i = -u_j$. Since $\Gamma$ is connected, this places a condition on all entries of $u$, which may or may not force all entries to be 0. A nonzero vector $u$ corresponds to an assignment of positive and negative signs to vertices such that adjacent vertices are not assigned the same sign. A nontrivial assignment exists if and only if $\Gamma$ is bipartite, so that the eigenspace of 0 has dimension 1 if and only if $\Gamma$ is bipartite. Note that for connected graphs, bipartite partitions are unique, so the eigenspace cannot have dimension more than 1. Now for general graphs, we again apply Lemma 1. □

**Proposition 4.** *A graph is bipartitite if and only if the Laplace spectrum and the signless Laplace spectrum are equal.*

*Proof.* If the spectra are equal, then by the previous two propositions, the number of connected components and bipartite connected components is the same. If $\Gamma$ is bipartite, then then $L$ and $Q$ are similar by a diagonal matrix $D$ with diagonal entries $\pm 1$, i.e. $Q = DLD^{-1}$. Matrix D is diagonal where the diagonal entries corresponding to one partition are all +1 and and the ones corresponding to the other partition are all -1. Then, we have $Q_{ij} = d_{ii} d_{jj}^{-1} L_{ij} = d_{ii} d_{jj} L_{ij}$. So, for $i = j$, we have $Q_{ij} = L_{ij}$ and for $i \neq j$ where $e = (i, j)$ is an edge, $Q_{ij} = -L_{ij}$ holds. □

## 4. Families of Graphs

We next consider the spectra of some common families of graphs.

4.1. **The Complete Graph $\mathbf{K_n}$.** The adjacency matrix of $K_n$ is $J - I$, where $J$ is the $n \times n$ all-ones matrix. There is a one dimensional eigenspace for $Jv = \lambda v$ with $v = \mathbf{1}$ and $\lambda = n$. Any vector $v$ such that $\mathbf{1}^T v = 0$ will be an eigenvector with eigenvalue $\lambda = 0$. This is a $n-1$ dimensional hyperplane, hence 0 is an eigenvalue with multiplicity $n - 1$. Translating, we see that $A = J - I$ then has spectra $(n-1)^1$ and $(-1)^{n-1}$. To consider the Laplacian spectrum, we again translate. We see then that $L = D - A = nI - J$ has 0 as an eigenvalue with multiplicity 1 and eigenvector $v = \mathbf{1}$, and $\lambda = n$ is an eigenvalue with multiplicity $n - 1$ and eigenspace $\mathbf{1}^\perp = \{v : \mathbf{1}^T v = 0\}$.

4.2. **The n-cycle $\mathbf{C_n}$.** We first consider the spectrum of any circulant matrix, as derived in [2]. Since both $A$ and $D - A$ are circulant matrices, the spectra of $C_n$, both directed and undirected, follow.

**Definition 5.** *A matrix $C$ is **circulant** if each row is a cyclic shift by one index of the row above it. That is, it has the form*

$$\begin{bmatrix} c_0 & c_1 & \dots & c_{n-1} \\ c_{n-1} & c_0 & \dots & c_{n-2} \\ \vdots & \ddots & \ddots & \vdots \\ c_1 & \dots & c_{n-1} & c_0 \end{bmatrix}.$$

If one denotes the first row of $C$ as the vector $c$, then a circulant matrix can be characterized by the property $C_{k,j} = c_{(j-k) \mod n}$. Thus, if we seek solutions to $Cv = \lambda v$, we are looking to solve the system of equations

$$\sum_{k=n-m}^{n-1} c_k v_{k-(n-m)} + \sum_{k=0}^{n-1-m} c_k v_{k+m} = \lambda v_m$$

for $m = 0, 1, \dots, n-1$, taking the convention $c_k = 0$ if $k \geq n$ and $v_k = 0$ if $k \leq 0$.

For each $m$, this yields a linear difference equation for $v_m$. Much like for linear differential equations, their continuous counterpart, a standard solution technique for linear difference equations is to make an educated guess and then show that the guess was correct. In the case of constant coefficients, the correct guess is $v_m = r^m$. Making this substitution and factoring out $r^m$, we have

$$\sum_{k=n-m}^{n-1} c_k r^{k-(n-m)} + \sum_{k=0}^{n-1-m} c_k r^{k+m} = \lambda r^m$$

$$r^{-n} \sum_{k=n-m}^{n-1} c_k r^k + \sum_{k=0}^{n-1-m} c_k r^k = \lambda$$

If we choose $r$ such that $r^n = 1$, this reduces to

$$\sum_{k=0}^{n-1} c_k r^k = \lambda.$$

So for each $n$-th root of unity $\zeta_m$, we have an eigenvalue $\lambda_m = \sum_{k=0}^{n-1} c_k \zeta_m^k$ with associated eigenvector $v = (1, \zeta_m, \zeta_m^2, \dots, \zeta_m^{n-1})$.

If we label the vertices of $C_n$ by $0, \dots, n-1$, then the adajency matrix of the undirected cycle is determined by the vector $c$, where

$$c_i = \begin{cases} 1 & i = 1, \text{ or } n-1 \\ 0 & \text{otherwise.} \end{cases}$$

Thus the eigenvalues of $A$ are

$$\begin{aligned} \lambda_m &= e^{2\pi i m/n} + (e^{2\pi i m/n})^{n-1} \\ &= 2\cos\left(\frac{2\pi m}{n}\right), \end{aligned}$$

for $m = 0, \dots, n-1$. Likewise, the eigenvalues of $L = D - A$ are $2 - 2\cos\left(\frac{2\pi m}{n}\right)$ for $m = 0, \dots, n-1$. For the undirected cycle, we see that the ordinary spectrum is the set of $n$-th roots of unity.

## 5. Eigenvalues and Eigenvectors

**Definition 6.** *The largest eigenvalue of a graph is known as its **spectral radius** or **index.***

In order to state some further results, we first develop some linear algebra.

**Definition 7.** *Let $T$ be a real $n \times n$ matrix with nonnegative entries. $T$ is called **primitive** if for some $k$, we have $T^k > 0$; $T$ is called **irreducible** if for all $i, j$, there is a $k$ such that $(T^k)_{ij} > 0$. (Here, $A > 0$ and $A \geq 0$ respectively mean all entries are positive or nonnegative).*

Recall that a directed graph is called strongly connected if there is a directed path from $i$ to $j$ for all vertices $i$ and $j$. Let $\Gamma_T$ be the directed graph with vertices $\{1, ..., n\}$, where $ij$ is an edge whenever $T_{ij} > 0$. Note that a matrix $T$ is irreducible if and only if the directed graph $\Gamma_T$ is strongly connected. This is because $(A^k)_{xy}$ is the number of paths of length $k$ from $x$ to $y$.

**Theorem 2.** *(Perron-Frobenius)*
*Let $T \geq 0$ be irreducible. There is a unique positive real number $\theta_0$ with the following properties:*

(1) *There is a real vector $x_0 > 0$ with $Tx_0 = \theta_0 x_0$.*
(2) *$\theta_0$ has geometric and algebraic multiplicity one.*
(3) *For each eigenvalue $\theta$ of $T$, we have $|\theta| \leq \theta_0$. If $T$ is primitive, then $|\theta| = \theta_0$, implies $\theta = \theta_0$. In general, if $T$ has period $d$, then $T$ has precisely $d$ eigenvalues $\theta$ with $|\theta| = \theta_0$, namely $\theta = \theta_0 e^{2\pi ij/d}$ for $j = 0, 1, ..., d-1$.*
(4) *Any nonnegative left or right eigenvector of $T$ has eigenvalue $\theta_0$. More generally, if $x \geq 0$, $x \neq 0$, and $Tx \leq \theta x$, then $x > 0$ and $\theta \geq \theta_0$; moreover, $\theta = \theta_0$ if and only if $Tx = \theta x$.*
(5) *If $0 \leq S \leq T$ or if $S$ is a principal minor of $T$, and $S$ has eigenvalue $\sigma$, then $|\sigma| \leq \theta_0$; if $|\sigma| = \theta_0$, then $S = T$.*
(6) *Given a complex matrix, let $|S|$ denote the matrix with elements $|S|_{ij} = |S_{ij}|$. If $|S| \leq T$ and $S$ has eigenvalue $\sigma$, then $|\sigma| \leq \theta_0$. If equality holds, then $|S| = T$, and there is a diagonal matrix $E$ with diagonal entires of absolute value 1 and a constant $c$ of absolute value 1 such that $S = cETE^{-1}$.*

Using Perron-Frobenius, we can deduce some results for the spectral radius.

**Proposition 5.** *Each graph $\Gamma$ has a real eigenvalue $\theta_0$ with nonnegative real corresponding eigenvector such that for each eigenvalue $\theta$, we have $|\theta| \leq \theta_0$. The value of $\theta_0(\Gamma)$ does not increase when vertices or edges are removed from $\Gamma$.*
*If $\Gamma$ is strongly connected, then*

(1) *$\theta_0$ has multiplicity 1.*
(2) *If $\Gamma$ is primitive (strongly connected, and such that not all cycles have a length that is a multiple of some integer $d > 1$), then $|\theta| < \theta_0$ for all eigenvalues $\theta$ different from $\theta_0$.*
(3) *The value of $\theta_0(\Gamma)$ decreases when vertices or edges are removed from $\Gamma$.*

**Proposition 6.** *Let $\Gamma$ be a connected, undirected graph with largest eigenvalue $\theta_1$. If $\Gamma$ is regular of valency (degree) $k$, then $\theta_1 = k$. Otherwise, we have $k_{\min} < k < \theta_1 < k_{\max}$, where $k_{\min}, k_{\max}$, and $k$ are the minimum, maximum, and average degree.*

*Proof.* Let $\mathbb{1}$ be the vector of all 1's. Then $A\mathbb{1} \leq k_{\max}\mathbb{1}$, and by Perron-Frobenius, we have $\theta_1 \leq k_{\max}$ with equality if and only if $A\mathbb{1} = \theta_1\mathbb{1}$, that is, if and only if $\Gamma$ is regular of degree $\theta_1$. □

## 6. The Independence Number

One application of these ideas is to the independence number of a graph. Recall that

**Definition 8.** *A **coclique** or **independent set** is a set of pairwise disjoint vertices in a graph $\Gamma$. The **independence number** $\alpha(\Gamma)$ is the maximal size of a coclique in $\Gamma$.*

The eigenvalues associated to a matrix provide bounds on the independence number of a graph. We first develop some more linear algebra, starting with the Rayleigh quotient.

**Definition 9.** *Let $A$ be a real symmetric matrix and $u$ be a nonzero vector. The **Rayleigh quotient** is defined by*

$$\frac{u^T A u}{u^T u}.$$

**Proposition 7.** *Let $\theta_1 \geq \ldots \geq \theta_n$ be the eigenvalues of $A$, a real symmetric matrix, and let $u_1, \ldots, u_n$ be an associated set of orthonormal eigenvectors. Then,*

$$\frac{u^T A u}{u^T u} \geq \theta_i \text{ if } u \in \langle u_1, \ldots, u_i \rangle, \text{ and}$$
$$\frac{u^T A u}{u^T u} \leq \theta_i \text{ if } u \in \langle u_i, \ldots, u_n \rangle.$$

*Proof.* Let $u = \sum a_i u_i$ be the expansion of $u$ in terms of this orthonormal basis. Then, $u^T u = \sum a_i^2$ and $u^T A u = \sum a_i^2 \theta_i$. Hence, if $u \in \langle u_1, \ldots, u_i \rangle$, then

$$\begin{aligned} \frac{u^T A u}{u^T u} &= \frac{\sum_{j=1}^{i} a_j^2 \theta_j}{\sum_{j=1}^{i} a_j^2} \\ &\geq \theta_i \frac{\sum_{j=1}^{i} a_j^2}{\sum_{j=1}^{i} a_j^2} \\ &= \theta_i. \end{aligned}$$

Likewise, if $u \in \langle u_i, \ldots, u_n \rangle$, then

$$\begin{aligned} \frac{u^T A u}{u^T u} &= \frac{\sum_{j=i}^{n} a_j^2 \theta_j}{\sum_{j=i}^{n} a_j^2} \\ &\leq \theta_i \frac{\sum_{j=i}^{n} a_j^2}{\sum_{j=i}^{n} a_j^2} \\ &= \theta_i. \end{aligned}$$

□

We will also need the following definition in our discussion of the independence number.

**Definition 10.** *Given two sequences of real numbers $\theta_1 \geq \ldots \geq \theta_n$ and $\eta_1 \geq \ldots \geq \eta_m$, the second sequence is said to **interlace** the first if*

$$\theta_i \geq \eta_i \geq \theta_{n-m+i} \text{ for } i = 1, \ldots, m.$$

If $m = n - 1$, this yields $\theta_1 \geq \eta_1 \geq \theta_2 \geq \ldots$, hence the name.

**Lemma 2.** *If $B$ is a principal submatrix of $A$, then the eigenvalues of $B$ interlace the eigenvalues of $A$.*

*Proof.* If $B$ is a principal submatrix, then $B = [I_m 0_{n-m}] A [I_m 0_{n-m}]^T$. Let $\eta_1 \geq \ldots \geq \eta_m$ be the eigenvalues of $B$, and $v_i$ their associated eigenvectors, and let $\theta_1 \geq \ldots \geq \theta_m$ be the eigenvalues of $A$, and $u_i$ an associated orthonormal basis of eigenvectors. For each $i$, take a nonzero vector

$$s_i \in \langle v_1, \ldots, v_i \rangle \cap \langle [I_m 0_{n-m}] u_1, \ldots, [I_m 0_{n-m}] u_{i-1} \rangle^{\perp}.$$

Since $[I_m 0_{n-m}][I_m 0_{n-m}]^T = I_m$, we have that

$$[I_m 0_{n-m}]^T s_i \in \langle u_1, \ldots, u_{i-1} \rangle^{\perp}.$$

Hence,

$$\theta_i \geq \frac{([I_m 0_{n-m}] s_i)^T A [I_m 0_{n-m}]^T s_i}{([I_m 0_{n-m}] s_i)^T [I_m 0_{n-m}]^T s_i} = \frac{s_i^T B s_i}{s_i^T s_i} \geq \eta_i.$$

The same argument shows the other inequality.

□

With this background in place, we quickly obtain bounds on the independence number of a graph $\Gamma$.

**Theorem 3.** *$\alpha(\Gamma) \leq |\{i : \theta_i \geq 0\}|$ and $\alpha(\Gamma) \geq |\{i : \theta_i \leq 0\}|$.*

*Proof.* An independent set of size $m$ corresponds to an all 0's principal submatrix of the adjacency matrix $A$ of size $m$. So, by Lemma 2, $\theta_{\alpha(\Gamma)} \geq \eta_{\alpha(\Gamma)} = 0$ and $\theta_{n-\alpha(\Gamma)-1} \leq \eta_{n-\alpha(\Gamma)-1} = 0$

□

## 7. The Chromatic Number

As another application of the eigenvalue bound ideas, we consider graph colorings. An upper bound can be derived by fairly elementary means. We also present a lower bound. Its proof requires more tools, so we omit it.

**Definition 11.** *A **proper vertex coloring** of a graph is an assignment of colors to the vertices so that adjacent vertices get different colors. The **chromatic number** $\chi(\Gamma)$ is the minimum number of colors of a proper vertex coloring of $\Gamma$.*

**Proposition 8.** *Let $\Gamma$ be connected with largest eigenvalue $\theta_1$. Then $\chi(\Gamma) \leq 1 + \theta_1$ with equality if and only if $\Gamma$ is complete or is an odd cycle.*

*Proof.* Let $m = \chi(\Gamma)$, and without loss of generality, we can assume $m > 1$. We claim there exists a submatrix $\Delta$ where vertices have minimum degree $m-1$. Let $\Delta$ be a subgraph such that $\chi(\Delta) = \chi(G)$ and $\Delta$ has the minimum number of vertices, i.e. removing any vertex $v$ from $\Delta$ reduces the chromatic number. Assume some $v \in \Delta$ has degree less than $m-1$. Remove $v$, and color $\Delta \setminus \{v\}$ using $m-1$ colors. Since $v$ is adjacent to at most $m-2$ vertices, we can insert $v$ and color it with one of the $m-1$ colors not used on its neighbors. This gives us a coloring of $\Delta$ with $m-1$ colors, a contradiction. Thus, it must be the minimum degree of $\Delta$ is $m-1$. Now

$$\theta_1(\Gamma) \geq \theta_1(\Delta) \geq d_{\min}(\Delta) \geq m - 1 = \chi(\Gamma) - 1.$$

By Perron-Frobenius, equality holds in the cases $\Gamma = \Delta$, and $\Delta$ must be regular of degree $m-1$. Brook's Theorem states a connected graph in which every vertex has at most $k$ neighbors can be colored with only $k$ colors, except for the two cases of complete graphs and cycle graphs of odd length, which require $k+1$ colors. □

**Theorem 4.** *If $\Gamma$ is not edgeless, then $\chi(\Gamma) \geq 1 - \frac{\theta_1}{\theta_n}$.*

**Remark 2.** *The complete multipartite graph $K_{m\times a}$ has chromatic number $m$ and achieves this bound.*

## 8. Application: Spectral Graph Sparsification

Given a graph $G = (V, E)$, we aim to "approximate" $G$ by a sparse graph $H = (V, E')$. That is, we would like to find a subset $E' \subseteq E$ such that $|E'|$ is as small as possible, say $\mathcal{O}(n)$ or $\mathcal{O}(n\log(n))$ where $n = |V|$, and $H$ is a "good" approximation for $G$. For two matrices $A$ and $B$, we denote $A \succeq B$ or $B \preceq A$ if $A - B$ is a positive semidefinite matrix. We consider the following notion of *spectral* graph approximation.

**Definition 12.** *Consider graphs $G$ and $H$. $H$ **$\epsilon$-spectrally approximates** $G$ for some $\epsilon > 0$ if*

$$(1-\epsilon)L_G \preceq L_H \preceq (1+\epsilon)L_G,$$

*where $L_G$ and $L_H$ are the Laplace matrix of graphs $G$ and $H$ respectively.*

Using the Courant-Fischer theorem, an immediate consequence of the above definition is that all eigenvalues of $L_H$ approximate eigenvalues of $L_G$ up to multiplicative $1 \pm \epsilon$ error.

**Remark 3.** *Spectral sparsification is stronger than cut sparsification. To see this, consider any subset $S \subseteq V$ of the vertices and the corresponding characteristic vector $\nu = \mathbf{1}_S$. Then, we have:*

$$\nu^T L_G \nu = \omega_G(E(S, \bar{S})),$$
$$\nu^T L_H \nu = \omega_H(E(S, \bar{S})),$$

*where $\omega_G(E(S,\bar{S}))$ and $\omega_H(E(S,\bar{S}))$ denote the cumulative weight of the edges in $G$ and $H$ respectively that cross the cut $(S, \bar{S})$ (i.e., the set of edges with one endpoint in $S$ and the other endpoint in $\bar{S}$). Therefore, if*

*$H$ is an $\epsilon$-spectral approximation of $G$, the weight of every cut in $H$ should be within $1 \pm \epsilon$ of the weight of the corresponding cut in $G$.*

**Theorem 5.** [3] *For every $\epsilon > 0$, the following holds:*
*For every unweighted, connected graph $G = (V, E)$, there exists a weighted graph $H = (V, E')$ such that $E' \subseteq E$, $|E'| \leq \mathcal{O}(\frac{n\log(n)}{\epsilon^2})$ and $H$ $\epsilon$-spectrally approximates $G$.*

*Proof.* First, to see why the approximation graph $H$ needs to be weighted, consider the case where $G$ is a complete graph. Since the weight of every cut in $H$ should be close to that of the corresponding cut in $G$ and considering that $|E(S, \bar{S})|=|S||\bar{S}|$, we need to put large weights on the edges of the sparse graph $H$.
Let $p$ be a probability distribution on the edges of the graph $G$, i.e., we have some probabilities $p_e \geq 0$ such that $\sum_{e\in E} p_e = 1$. The random sampling algorithm is provided in Algorithm 1.

**Algorithm 1** Graph Sparsification Algorithm

**Input:** Graph $G = (V, E)$, probabilities $p_e$ over the edges $e \in E$, $k$: number of edges of the sparse graph (i.e., $|E'| = k$ where $H = (V, E')$ is the $\epsilon$-spectral approximation of $G$).
**Output:** $\epsilon$-spectral approximation $H = (V, E')$ of $G = (V, E)$ where $E' \subseteq E$.
Initialize edge weights $\omega_e := 0$ for every $e \in E$.
**for** $i = 1$ **to** $k$ **do**
  Sample an edge $e \in E$ i.i.d. according to probabilities $\{p_e\}_{e\in E}$.
  Update $\omega_e = \omega_e + \frac{1}{kp_e}$.
**end for**

If we denote the edge sampled at round $i \in [k]$ by $e^{(i)}$, then we have $L_H = \sum_{i=1}^{k} X_i$ where $X_i = \frac{1}{kp_{e^{(i)}}} L_{e^{(i)}}$. Therefore, we can write:

$$\mathbb{E}[L_H] = \sum_{i=1}^{k} \mathbb{E}[X_i] = \sum_{i=1}^{k} \sum_{e\in E} p_e \frac{1}{kp_e} L_e = \frac{1}{k} \sum_{i=1}^{k} \underbrace{\sum_{e\in E} L_e}_{=L_G} = L_G.$$

So $L_H$ is an unbiased estimator of $L_G$. However, in order for $H$ to be an $\epsilon$-spectral approximation of $G$ with a small value of $k$, we need the sum of i.i.d. random matrices $L_H = \sum_{i=1}^{k} X_i$ to be concentrated around its mean. We use the following concentration inequality for a sum of i.i.d. random matrices:
For any $\alpha \geq 1$, if $X_i \preceq \alpha\mathbb{E}[X_i] = \frac{\alpha}{k} L_G$ holds with probability one for all $i \in [k]$, then for any $\epsilon \in (0, 1)$, we have:

$$\mathbb{P}\big[(1-\epsilon)L_G \preceq L_H = \sum_{i=1}^{k} X_i \preceq (1+\epsilon)L_G\big] \geq 1 - 2n\exp(-\frac{\epsilon^2 k}{4\alpha}).$$

Therefore, if we choose $k = \mathcal{O}(\frac{\alpha}{\epsilon^2}\log(n))$, the desired result holds with probability at least $1 - \frac{1}{n}$. So if we choose the sampling probabilities $p_e$ such that $X_i \preceq \alpha\mathbb{E}[X_i] = \frac{\alpha}{k} L_G$ holds with probability one for all $i \in [k]$ and $\alpha \leq \mathcal{O}(n)$, the proof is complete. A natural choice for the sampling probabilities is uniform sampling where $p_e = \frac{1}{|E|}$ $\forall e \in E$. However, this approach fails in a general graph. For instance, Let G be the *Barbell* graph, i.e., union of two $K_{\frac{n}{2}}$ connected by an edge $e$. Since the weight of the cut $e$ in G and its corresponding cut in $H$ need to be approximately equal, we have to include the edge $e$ in the sparse graph $H$ and using uniform random sampling, we need to sample $\Theta(n^2)$ edges for edge $e$ to be sampled with high probability. Thus we need an alternative choice of sampling probabilities $p_e$ $\forall e \in E$.

**Definition 13.** *For every edge $e \in E$, we define the **effective resistance** of the edge $e$ as follows:*

$$R_e := \mathrm{Tr}(L_e L_G^{\dagger}),$$

*where $L_G^{\dagger}$ is the pseudo-inverse of $L_G$. That is, if $L_G = \sum_{i=2}^{n} \lambda_i u_i u_i^T$ is the eigenvalue decomposition of the Laplace matrix $L_G$, we have $L_G^{\dagger} = \sum_{i=2}^{n} \frac{1}{\lambda_i} u_i u_i^T$.*

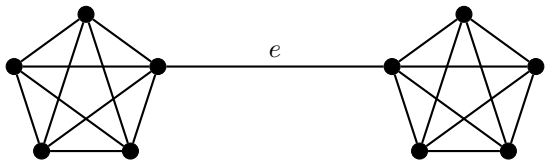



FIGURE 1. Barbell graph

Using the above definition, we have $\sum_{e\in E} L_e L_G^\dagger = L_G L_G^\dagger = I_{\mathrm{Im}(L_G)}$, where $\mathrm{Im}(L_G)$ is the span of non-zero eigenvectors of $L_G$ (i.e., the $(n-1)$-dimensional vector space orthogonal to $\mathbf{1}$).

**Remark 4.** *We can obtain the above formula using the definition of effective resistance in electrical networks as well. To see that, we first remind the reader that $L_G = NN^T$, where $N$ is the directed incidence matrix of the graph. Effective resistance between two endpoints $u, w$ of an edge $e$ is defined as the potential difference between them when a unit current is injected at one and extracted at the other. Using Kirchoff's current law, if we denote the current injected in vertices and the current in edges by $i_{inj}$ and $i_f$ respectively, we have $Ni_f = i_{inj}$. We can use Ohm's law to write $i_f = N^T v$, where $v$ denotes the potential induced at the vertices. Combining these two equalities, we obtain:*

$$i_{inj} = NN^T v = L_G v.$$

*If $i_{inj}$ is orthogonal to the kernel of the Laplace matrix $L_G$ (i.e., $i_{inj} \perp \mathbf{1}$), we can write $v = L_G^\dagger i_{inj}$. To obtain the effective resistance across the edge $e = (u, w)$, let $i_{inj} = 1_w - 1_u$. First, note that $i_{inj} \perp \mathbf{1}$ holds. Therefore, we have:*

$$R_e = v(w) - v(u) = (1_w - 1_u)^T v = (1_w - 1_u)^T L_G^\dagger (1_w - 1_u) = \mathrm{Tr}\big(\underbrace{(1_w - 1_u)(1_w - 1_u)^T}_{=L_e} L_G^\dagger\big) = \mathrm{Tr}(L_e L_G^\dagger).$$

We define our sampling probabilities to be proportional to the effective resistance of the edges. To be precise, for edge $e \in E$, we set $p_e = \frac{R_e}{n-1}$. To see why this is indeed a probability distribution over the edges, we can write:

$$\sum_{e\in E} R_e = \sum_{e\in E} \mathrm{Tr}(L_e L_G^\dagger) = \mathrm{Tr}\left(\sum_{e\in E} L_e L_G^\dagger\right) = \mathrm{Tr}(L_G L_G^\dagger) = \mathrm{Tr}(I_{\mathrm{Im}(L_G)}) = n - 1.$$

Using the aforementioned sampling probabilities, for all $i \in [k]$, we have:

$$X_i \preceq \alpha \mathbb{E}[X_i] = \frac{\alpha}{k} L_G \iff \frac{n-1}{k R_{e^{(i)}}} L_{e^{(i)}} \preceq \frac{\alpha}{k} L_G \iff L_{e^{(i)}} \preceq \alpha \frac{R_{e^{(i)}}}{n-1} L_G.$$

Therefore, if we set $\alpha = n-1$, we need to verify $L_e \preceq R_e L_G$ for all $e \in E$ to complete the proof. Equivalently, we have to show that $\nu^T L_e \nu \leq R_e \nu^T L_G \nu$ holds for every $\nu \in \mathbb{R}^n$. Considering that the vector of all ones $\mathbf{1}$ is in the kernel of both matrices $L_e$ and $L_G$, it suffices to only consider $\nu$ such that $\nu \perp \mathbf{1}$. Moreover, denoting the square root of $L_G^\dagger = \sum_{i=2}^n \frac{1}{\lambda_i} u_i u_i^T$ with $L_G^{\dagger/2} = \sum_{i=2}^n \frac{1}{\sqrt{\lambda_i}} u_i u_i^T$, since $\mathrm{Im}(L_G^{\dagger/2})$ contains every vector orthogonal to $\mathbf{1}$, it suffices to prove the inequality for $\nu = L_G^{\dagger/2}\omega$ for any $\omega \in \mathbb{R}^n$ such that $\omega \perp \mathbf{1}$. So we need to show the following:

$$(L_G^{\dagger/2}\omega)^T L_e (L_G^{\dagger/2}\omega) \leq R_e (L_G^{\dagger/2}\omega)^T L_G (L_G^{\dagger/2}\omega) = R_e \omega_T L_G^{\dagger/2} L_G L_G^{\dagger/2} \omega = R_e \omega^T \omega.$$

The above inequality holds because:

$$\omega^T L_G^{\dagger/2} L_e L_G^{\dagger/2} \omega \leq \|L_G^{\dagger/2} L_e L_G^{\dagger/2}\| \omega^T \omega \leq \mathrm{Tr}(L_G^{\dagger/2} L_e L_G^{\dagger/2}) \omega^T \omega = \mathrm{Tr}(L_e L_G^\dagger) \omega^T \omega = R_e \omega^T \omega,$$

where the second inequality is due to positive semidefiniteness of $L_G^{\dagger/2} L_e L_G^{\dagger/2}$. □

## References


[1] Andries E. Brouwer and Willem H. Haemers. *Spectra of Graphs*. Springer, 2012.

[2] Robert M. Gray. *Toeplitz and Circulant Matrices: A Review*. Now Publishers, 2006.

[3] Daniel A. Spielman and Nikhil Srivastava. Graph sparsification by effective resistances. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, page 563–568, New York, NY, USA, 2008. Association for Computing Machinery.