EIGENVALUES OF THE LAPLACIAN AND THEIR RELATIONSHIP TO THE CONNECTEDNESS OF A GRAPH

ANNE MARSDEN

Abstract. This paper develops the necessary tools to understand the relationship between eigenvalues of the Laplacian matrix of a graph and the connectedness of the graph. First we prove that a graph has k connected components if and only if the algebraic multiplicity of eigenvalue 0 for the graph's Laplacian matrix is k. We then prove Cheeger's inequality (for d-regular graphs) which bounds the number of edges between the two subgraphs of G that are the least connected to one another using the second smallest eigenvalue of the Laplacian of G.

Contents

1. Introduction

We can learn much about a graph by creating an adjacency matrix for it and then computing the eigenvalues of the Laplacian of the adjacency matrix. In section three this paper shows that the multiplicity of the second smallest eigenvalue indicates how many connected components exist in the graph. Recall that there are no edges between any two connected components of a graph. Often, however, a graph has subgraphs that are almost connected components but for a few edges between them. In this case the machinery used to identify connected components in section three no longer works. This more subtle problem is addressed by Cheeger's Inequality. In section four we develop the idea of the expansion of a graph. The expansion of a graph is, roughly speaking, a real number which indicates how close G is to having two connected components (the expansion approaches zero as G gets closer to having connected components). Cheeger's Inequality puts lower and upper bounds on the expansion of the graph, which is useful because the expansion of a graph can be difficult to compute. This paper proves Cheeger's Inequality for only two connected components. However, recent research has developed Cheeger's Inequality for k connected components [1]. These results are very important in the

Date: August 31, 2013.

analysis of Markov chains, mixing time of random walks, graph partitioning, image segmentation and more [2].

2. Spectral Theorem for Real Matrices and Rayleigh Quotients

Let $M_n(\mathbb{R})$ denote the ring of n by n matrices with entries in $\mathbb{R}$. For this paper, unless otherwise noted, all vectors are assumed to be column vectors.

Definition 2.1. A matrix $A \in M_n(\mathbb{R})$ is symmetric if $A = A^T$

Lemma 2.2. *If $A \in M_n(\mathbb{R})$ is symmetric then A has a real eigenvalue.*

The proof is omitted.

Definition 2.3. A matrix C is orthogonal if $CC^T = C^TC = I$

Note that for an orthogonal matrix C, $C^T = C^{-1}$. Note also that an $n \times n$ matrix whose columns form an orthonormal basis of R^n is an orthogonal matrix.

Theorem 2.4 (Spectral Theorem)**.** *For every symmetric matrix $A \in M_n(\mathbb{R})$ there exists an orthogonal matrix C whose columns are eigenvectors of A which form an orthonormal basis of $\mathbb{R}^n$ so that $C^{-1}AC$ is a diagonal matrix*

$$\begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$$

where $\lambda_1, \dots, \lambda_n$ are the corresponding eigenvalues of A.

Proof. We prove by induction on k. The $k = 1$ case is clear. Now suppose it holds for any matrix in $M_k(\mathbb{R})$. Then consider $A \in M_{k+1}(\mathbb{R})$. By Lemma 2.2, A has at least one real eigenvalue λ_1. Take the corresponding eigenvector v_1 and form an orthonormal basis

$$\{v_1, \dots, v_{k+1}\}$$

using Gram-Schmidt method starting with v_1. Let U be a matrix whose columns are the vectors in this orthonormal basis. That is,

$$U = \begin{bmatrix} | & \dots & | \\ v_1 & \dots & v_{k+1} \\ | & \dots & | \end{bmatrix}$$

One can compute that

$$U^TAU = \begin{bmatrix} \lambda_1 & x_{1\times k} \\ 0_{k\times 1} & A_2 \end{bmatrix} \tag{2.5}$$

where $x_{1\times k}$ is some row vector, $0_{k\times 1}$ is a column vector of all zeros, and $A_2 \in M_k(\mathbb{R})$. But U^TAU is symmetric since

$$(U^TAU)^T = (AU)^T(U^T)^T = U^TA^TU = U^TAU \tag{2.6}$$

Therefore $x_{1\times k} = 0_{1\times k}$ and so now we have,

$$U^T A U = \begin{bmatrix} \lambda_1 & 0_{1\times k} \\ 0_{k\times 1} & A_2 \end{bmatrix} \tag{2.7}$$

$A_2 \in M_k(\mathbb{R})$ so by inductive hypothesis there exists an orthogonal matrix C whose

columns form an orthonormal basis of eigenvectors of A_2. Denote these column vectors as $w_1, \ldots, w_k$. Then we claim that

$$\left\{ v_1, U\begin{pmatrix} 0 \\ w_1 \end{pmatrix}, \ldots, U\begin{pmatrix} 0 \\ w_k \end{pmatrix} \right\}$$

forms an orthonormal basis of eigenvectors of A. We must show for all i,

$$AU\begin{pmatrix} 0 \\ w_i \end{pmatrix} = \lambda_i U \begin{pmatrix} 0 \\ w_i \end{pmatrix} \tag{2.8}$$

where λ_i is the corresponding eigenvalue for w_i. But this is true if and only if,

$$U^T A U\begin{pmatrix} 0 \\ w_i \end{pmatrix} = \lambda_i \begin{pmatrix} 0 \\ w_i \end{pmatrix} \tag{2.9}$$

One can confirm (2.9) by using the equality given by (2.7). That is,

$$\begin{bmatrix} \lambda_1 & 0_{1\times k} \\ 0_{k\times 1} & A_2 \end{bmatrix} \begin{bmatrix} 0 \\ w_i \end{bmatrix} = \lambda_i \begin{bmatrix} 0 \\ w_i \end{bmatrix}$$

So now it remains to show that

$$\left\{ v_1, U\begin{pmatrix} 0 \\ w_1 \end{pmatrix}, \ldots, U\begin{pmatrix} 0 \\ w_k \end{pmatrix} \right\}$$

forms an orthonormal basis. It suffices to show that the vectors are orthogonal to each other since this implies that they are linearly independent. Furthermore, if there are $k+1$ linearly independent vectors in $\mathbb{R}^{k+1}$ then they must be a spanning set and so they must form a basis. If we write $\begin{bmatrix} 0 \\ w_i \end{bmatrix}$ as $\begin{bmatrix} 0 \\ w_{i1} \\ w_{i2} \\ \vdots \\ w_{ik} \end{bmatrix}$ we can compute that for all i,

$$U \begin{bmatrix} 0 \\ w_{i1} \\ w_{i2} \\ \vdots \\ w_{ik} \end{bmatrix} = \sum_j w_{ij} v_{j+1} \tag{2.10}$$

where the v_{j+1} are the vectors in the orthonomal basis for A created at the beginning of the proof. Then for all i we have,

$$\langle v_1, U\begin{pmatrix} 0 \\ w_i \end{pmatrix}\rangle = \langle v_1, \sum_j w_{ij} v_{j+1}\rangle = \sum_j w_{ij}\langle v_1, v_{j+1}\rangle = 0\rangle \tag{2.11}$$

and for any $i \neq k$ we have,

$$\langle U(0, w_i), U(0, w_k)\rangle = \langle \sum_j w_{ij} v_{j+1}, \sum_k w_{ik} v_{k+1}\rangle = \langle w_i, w_k\rangle = 0 \tag{2.12}$$

□

Corollary 2.13. *If an eigenvalue λ of a symmetric matrix A has algebraic multiplicity k then there are k linearly independent eigenvectors of A with corresponding eigenvalue λ*

Proof. The characteristic polynomials of A and $C^{-1}AC$ are the same since

$$det(xI - C^{-1}AC) = det(C^{-1}(xI - A)C) = det(C^{-1})det(xI - A)det(C) = det(xI - A)$$

Then the algebraic multiplicities of the eigenvalues of A and $C^{-1}AC$ must also be the same. This means that if λ_i has algebraic multiplicity k it appears as a diagonal entry k times in $C^{-1}AC$. In the proof of Spectral Theorem we showed that for each diagonal entry of $C^{-1}AC$ we can find a corresponding eigenvector of A that is linearly independent of the rest. Then if λ_i appears k times in $C^{-1}AC$ we can find k linearly independent eigenvectors of A with eigenvalue λ_i. □

Definition 2.14. Let A be an $n \times n$ matrix. The Rayleigh Quotient of a vector $x \in \mathbb{R}^n$ with respect to this matrix A is defined to be $\frac{x^T Ax}{x^T x}$. It is sometimes written as $\mathbb{R}_A(x)$ [5].

Note that Theorem 2.4 implies that all the eigenvalues of a real symmetric matrix are real, so it makes sense to order them.

Theorem 2.15. *For any symmetric matrix $A \in M_n(\mathbb{R})$ with eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, we have $\lambda_1 = \underset{x\in\mathbb{R}^n}{min}\mathbb{R}_A(x)$*

Proof. By Spectral Theorem, there exists an orthonormal basis $\{v_1, \ldots, v_n\}$ of eigenvectors of A (where v_i is the eigenvector with eigenvalue λ_i). Then for any $x \in \mathbb{R}^n, x = \sum\limits_i \beta_i v_i$. Thus $x^T Ax = \left(\sum\limits_i \beta_i v_i\right)^T A \left(\sum\limits_i \beta_i v_i\right) = \sum\limits_i \beta_i^2 \lambda_i$ and $x^T x = \sum\limits_i \beta_i^2$. By definition of λ_1, for any i $\lambda_1 \leq \lambda_i$. Then

$$\sum_i \beta_i^2 \lambda_i \geq \sum_i \beta_i^2 \lambda_1 = \lambda_1 \sum_i \beta_i^2 \tag{2.16}$$

Therefore for all $x \in \mathbb{R}^n$ we have,

$$\frac{x^T A x}{x^t x} \geq \lambda_1 \frac{\sum_i \beta_i^2}{\sum_i \beta_i^2} = \lambda_1 \tag{2.17}$$

It remains to show that an x such that $\frac{x^T A x}{x^t x} = \lambda_1$ indeed exists. Pick $x = v_1$, then

$v_1^T A v_1 = v_1^T \lambda_1 v_1 = \lambda_1$ and $v_1^T v_1 = 1$ so $\frac{v_1^T A v_1}{v_1^T v_1} = \lambda_1$. □

3. The Laplacian and the Connected Components of a Graph

For a graph G we let E denote the set of all edges in G and V denote the set of vertices in G. An edge that connects vertex i to vertex j is denoted as $[i, j]$. We do not allow multiple edges to be between the same pair of vertices. We assume G is an undirected graph, which means an edge that connects vertex i to vertex j also connects vertex j to vertex i so that $[i, j] = [j, i]$.

Definition 3.1. Let $I = \{k \in V | [i, k] \in E\}$. The degree of vertex i is equal to the cardinality of set I.

Definition 3.2. G is a d-regular graph if and only if the degree of each vertex in G is d.

Definition 3.3. Let G be a graph with n vertices. An adjacency matrix of G is an $n \times n$ matrix (a_{ij}) such that

$$a_{ij} = \left\{ \begin{array}{c} 1, \text{if } [i, j] \in E \\ 0, \text{otherwise} \end{array} \right\}$$

Definition 3.4. The Laplacian matrix L is defined to be $D - A$ where $D = \begin{bmatrix} deg(1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & deg(n) \end{bmatrix}$ and A is the adjacency matrix of the graph.

Note that for an undirected graph the adjacency matrix and the Laplacian matrix are symmetric.

Lemma 3.5. *Let G be a d-regular graph. Then 0 is an eigenvalue for the Laplacian matrix of G.*

Proof. If G is a d-regular graph the sum of the entries in row i gives the degree of vertex i so we have,

$$A \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} d \\ d \\ \vdots \\ d \end{pmatrix}$$

Then one can see that

$$(D - A) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} d - d \\ d - d \\ \vdots \\ d - d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

□

Lemma 3.6. *Let $x \in \mathbb{R}^n$ where $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Let L denote the Laplacian matrix of graph G and E denote the set of edges in G. Then $x^T Lx = \sum_{[i,j]\in E} (x_i - x_j)^2$*

Proof. By definition of the Laplacian we have,

$$x^T Lx = x^T (D - A)x = x^T Dx - x^T Ax$$

Expanding this out we get

$$\begin{pmatrix} x_1 \dots x_n \end{pmatrix} \begin{bmatrix} deg(1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & deg(n) \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} x_1 \dots x_n \end{pmatrix} (a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} =$$

$$\sum_i deg(i) x_i^2 - \begin{pmatrix} x_1 \dots x_n \end{pmatrix} (a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Recall that $a_{ij} = 1$ if there is an edge between vertex i and j and $a_{ij} = 0$ otherwise, so $\begin{pmatrix} x_1 \dots x_n \end{pmatrix} (a_{ij}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{[i,j]\in E} a_{ij} x_i x_j = \sum_{[i,j]\in E} 2xixj$. Thus $x^T Lx = \sum_i deg(i) x_i^2 - \sum_{[i,j]\in E} a_{ij} xixj = \sum_{[i,j]\in E} (x_i^2 - 2xixj + x_j^2) = \sum_{[i,j]\in E} (x_i - x_j)^2$ □

Lemma 3.7. *Let B be a block diagonal matrix with block matrices $A_1, A_2 \dots, A_n$. Then $detB = detA_1 detA_2 \dots detA_n$.*

The proof is omitted because the lemma is rather intuitive and the proof is technical.

Definition 3.8. A connected component of an undirected graph is a connected subgraph such that there are no edges between vertices of the subgraph and vertices of the rest of the graph.

Theorem 3.9. *Let G be a d-regular graph. The algebraic multiplicity of eigenvalue 0 for the Laplacian matrix is exactly 1 iff G is connected.*

Proof. We prove the first direction by contrapositive. Suppose G is not connected. Then it is possible to write G's adjacency matrix so that it is block diagonal. To do this, label vertices contained in the same connected component of G consecutively so that vertices $\{1, 2, \dots, a - 1, a\}$ are part of one connected component; vertices $\{a + 1, a + 2, \dots, a + b - 1, a + b\}$ are part of another connected component and so on. Let A_1 denote the adjacency matrix for the connected component with vertices $\{1, 2, \dots, a - 1, a\}$, A_2 denotes the adjacency matrix for the connected component with vertices $\{a + 1, a + 2, \dots, a + b - 1, a + b\}$, and so on. Then our adjacency matrix for G can be written as

$$\begin{bmatrix} A_1 & \dots & & & \\ \vdots & A_2 & \dots & & \\ & \vdots & \ddots & & \\ & & & \ddots & \vdots \\ & & & \dots & A_r \end{bmatrix}$$

Since there are no edges between these connected components, all values outside the A_is will be 0, so the matrix is block diagonal. Note that the Laplacian is also block diagonal. Let the blocks be denoted $L_1, \dots, L_r$. By Lemma 3.5 each block matrix in the Laplacian has eigenvalue 0. That is, for any i, $det(xI - L_i)$ has a $(x - 0)$ term. By Lemma 3.7 $det(xI - L) = det(xI - L_1) \dots det(xI - L_r)$, so $det(xI - L)$ has more than one $(x - 0)$ term so the algebraic multiplicity is greater than 1. Now we prove the second direction by contradiction. Suppose G is connected but the algebraic multiplicity of eigenvalue 0 is greater than 1. Then by Corollary 2.13, there exist at least two linearly independent eigenvectors v_1, v_2 with eigenvalue 0. Then $v_1^T L v_1 = 0$ and $v_2^T L v_2 = 0$ so by Lemma 3.6,

$$\sum_{[i,j] \in E} (v_{i1} - v_{j1})^2 = 0$$

and

$$\sum_{[i,j] \in E} (v_{i2} - v_{j2})^2 = 0$$

Since all entries in the sum are positive, this can only be zero when $\forall [i,j] \in E$, $v_{i1} = v_{j1}$ and $v_{i2} = v_{j2}$. This also implies that if there is a path between two vertices m and n, then $v_{m1} = v_{n1}$ and $v_{m2} = v_{n2}$ even if $[m,n] \notin E$. But if G is connected there is a path between any pair of vertices. Therefore v_1 and v_2 are constant vectors (every entry is the same). Thus v_1 and v_2 are scalar multiples of each other. But v_1 and v_2 are linearly independent which is a contradiction. □

Theorem 3.10. *A graph G has k connected components iff the algebraic multiplicity of 0 in the Laplacian is k*

Proof. G has k connected components iff its adjacency matrix can be block diagonal with k blocks iff its Laplacian matrix has k blocks. Let the blocks be denoted $L_1, \dots, L_k$. By Lemma 3.5 each block matrix in the Laplacian has eigenvalue 0. By Theorem 3.9 each block matrix has eigenvalue 0 with multiplicity exactly one. By Lemma 3.7 $det(xI - L) = det(xI - L_1) \dots det(xI - L_k)$, so the algebraic multiplicity of 0 for the entire Laplacian matrix is the sum of the algebraic multiplicity of 0 of each L_i. If G has k connected components then this sum is $1 + \dots + 1$ k times, so k. For the other direction, if the algebraic multiplicity of the Laplacian is k then the sum is k so there must be k blocks in the Laplacian, and so G must have k connected components. □

4. Cheeger's Inequality

For this section the following notation will be used. Let V be the set of vertices of a graph G. We will assume $|V| = n$. Let S denote any subset of V where $|S| \leq \frac{n}{2}$ and S^C be its complement so that $S \cup S^C = V$ and $S \cap S^C = \varnothing$. Let $E(S, S^C) =$

$\{[i,j] \in E | i \in S, j \in S^C \text{ or } j \in S, i \in S^C\}$ so $|E(S,S^C)|$ gives the number of edges between S and S^C

Definition 4.1. The expansion of set $S \subseteq V$ is $\frac{|E(S,S^C)|}{\sum_i deg(i)}$. We will often write the expansion of set S as ϕ_S.

We will let ϕ_G be defined as $\min_{S \subseteq G} \frac{|E(S,S^C)|}{\sum_i deg(i)}$

I now present Cheeger's Inequality. I present the theorem for a d-regular graph for simplicity. However, there is a version of the theorem for graphs that are not d regular [1]

Theorem 4.2. *(Cheeger's Inequality) Let λ_2 be the second smallest eigenvalue of the Laplacian matrix for a graph G (i.e., $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$). Then $\frac{\lambda_2}{2} \leq \phi_G \leq 2\sqrt{\frac{\lambda_2}{d}}$*

Lemma 4.3. *Let λ_2 be the second smallest eigenvalue for some real symmetric matrix L (like the Laplacian), let K be a two dimensional vector subspace of $\mathbb{R}^n$. Then $\lambda_2 = \underset{dim K=2}{min} \, \underset{x \in K}{max} \frac{x^T L x}{x^T x}$*

Proof. L is real symmetric so by Spectral Theorem we can find an orthonormal basis of eigenvectors of L. Denote this basis $u_1, u_2, \ldots, u_n$. Consider the vector subspace $K = \text{span}(u_1, u_2)$. Since $\forall x \in K$, $x = \alpha_1 u_1 + \alpha_2 u_2$ we have that

$$\frac{x^T L x}{x^T x} = \frac{(\alpha_1 u_1 + \alpha_2 u_2)^T L (\alpha_1 u_1 + \alpha_2 u_2)}{(\alpha_1 u_1 + \alpha_2 u_2)^T (\alpha_1 u_1 + \alpha_2 u_2)} = \frac{\alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2}{\alpha_1^2 + \alpha_2^2} \leq \frac{\alpha_1^2 \lambda_2 + \alpha_2^2 \lambda_2}{\alpha_1^2 + \alpha_2^2} = \lambda_2$$

This shows that there exists a two dimensional vector space K such that

$$\lambda_2 = \underset{x \in K}{max} \frac{x^T L x}{x^T x}$$

and so it follows that

$$\lambda_2 \geq \min_{\dim K = 2} \max_{x \in K} \frac{x^T L x}{x^T x}$$

Now we show that $\lambda_2 \leq \min_{\dim K = 2} \max_{x \in K} \frac{x^T L x}{x^T x}$. First note that for any K such that $\dim K = 2$ we have,

$$K \cap \text{span}(u_2, u_3, \ldots, u_n) \neq \varnothing$$

If this were not true then $K = \text{span}(u_1)$ and so it would no longer be two dimensional. Then since

$$K \cap \text{span}(u_2, \ldots, u_n) \subseteq K$$

it follows that,

$$\max_{x \in K} \frac{x^T L x}{x^T x} \geq \max_{x \in K \cap \text{span}(u_2, \ldots, u_n)} \frac{x^T L x}{x^T x}$$

Now for any $x \in \text{span}(u_2, \ldots, u_n)$ one can see that

$$\frac{x^T L x}{x^T x} \geq \lambda_2$$

so it follows that for any $x \in K \cap \text{span}(u_2, \ldots, u_n)$ we have,

$$\frac{x^T L x}{x^T x} \geq \lambda_2$$

Thus for any K, $\max_{x \in K} \frac{x^T L x}{x^T x} \geq \lambda_2$ and so it follows that

$$\min_{\dim K = 2} \max_{x \in K} \frac{x^T L x}{x^T x} \geq \lambda_2$$

Then since

$$\lambda_2 \geq \min_{\dim K = 2} \max_{x \in K} \frac{x^T L x}{x^T x}$$

and

$$\min_{\dim K = 2} \max_{x \in K} \frac{x^T L x}{x^T x} \geq \lambda_2$$

we have our result. $\square$

Now we prove the first part of Cheeger's Inequality, that is $\frac{\lambda_2}{2} \leq \phi_G$. We prove this for for d-regular graphs.

Proof. Choose some S such that $\phi_S = \phi_G$. We will create two vectors in $\mathbb{R}^n$ using this S. Define x_1 to be a column vector where $x_{i1} = 1$ if $i \in S$ and $x_{i1} = 0$ otherwise. Define x_2 to be a column vector where $x_{i2} = 1$ if $i \in S^C$ and $x_{i2} = 0$ otherwise. Consider $\frac{x_1^T L x_1}{x_1^T x_1}$; using Lemma 3.6, this is equal to

$$\frac{\sum_{[i,j] \in E} (x_{i1} - x_{j1})^2}{\sum_i x_{i1}^2}$$

Note that $(x_{i1} - x_{j1})^2$ will be 0 if $i, j \in S$ since $x_{i1} = x_{j1} = 1$ and it will be 1 if i, j is an edge between S and S^C. Thus the numerator is equal to the number of edges between S and S^C, which we will denote as $|E(S, S^C)|$. One can see that the denominator is equal to $|S|$. Thus we have shown so far that,

$$\frac{x_1^T L x_1}{x_1^T x_1} = \frac{\sum_{[i,j] \in E} (x_{i1} - x_{j1})^2}{\sum_i x_{i1}^2} = \frac{|E(S, S^C)|}{|S|}$$

Since $\phi_G = \phi_S = \frac{|E(S, S^C)|}{d|S|}$ it follows that

$$\frac{1}{d} \frac{x_1^T L x_1}{x_1^T x_1} = \phi_S = \phi_G \tag{4.4}$$

Now consider $\operatorname{span}(x_1, x_2)$. We claim that for any $x \in \operatorname{span}(x_1, x_2)$,

$$\max_{x \in \operatorname{span}(x_1, x_2)} \frac{x^T L x}{x^T x} \leq 2\phi_G$$

Note that $\{x_1, x_2\}$ is a linearly indepedent set so dim(span(x_1, x_2))=2. Then by Lemma 4.3 this would prove that $\lambda_2 \leq 2\phi_G$. For $x \in \operatorname{span}(x_1, x_2)$, we write

$$x = \alpha_1 x_1 + \alpha_2 x_2$$

so

$$\frac{x^TLx}{x^Tx} = \frac{\sum\limits_{[i,j]\in E}((\alpha_1 x_{i1}+\alpha 2x_{i2})-(\alpha_1 x_{j1}+\alpha 2x_{j2}))^2}{\sum\limits_{i}(\alpha_1 x_{i1}+\alpha 2x_{i2})^2} = \frac{\alpha_1^2(x_{i1}-x_{j1})^2+2\alpha_1\alpha_2(x_{i1}-x_{j1})(x_{i2}-x_{j2})+\alpha_2^2(x_{i2}-x_{j2})^2}{\alpha_1^2\sum\limits_i x_{i1}^2+\alpha_2^2\sum\limits_i x_{i2}^2+\alpha_1\alpha_2\sum x_{i1}x_{i2}}$$

One can see that

$$\sum_{[i,j]\in E}\alpha_1\alpha_2(x_{i1}-x_{j1})(x_{i2}-x_{j2}) = -|E(S,S^C)|$$

This is true because $\alpha_1\alpha_2(x_{i1}-x_{j1})(x_{i2}-x_{j2}) = 0$ if edge $[i,j]$ is either in S or S^C and $\alpha_1\alpha_2(x_{i1}-x_{j1})(x_{i2}-x_{j2}) = 1$ if edge $[i,j]$ is between S and S^C. One can also see that

$$\sum_{[i,j]\in E}(x_{i1}-x_{j1})^2 = |E(S,S^C)| \text{ and } \sum_{[i,j]\in E}(x_{x2}-x_{j2})^2 = |E(S,S^C)|$$

So the numerator becomes

$$|E(S,S^C)|(\alpha_1^2-2\alpha_1\alpha_2+\alpha_2^2) = |E(S,S^C)|(\alpha_1-\alpha 2)^2$$

For the denominator note that $\alpha_1\alpha_2\sum x_{i1}x_{i2} = 0$ since $x_{i1} = 1$ iff $i\in S$ iff $i\notin S^C$ iff $x_{i2} = 0$. The same reasoning shows that $x_{i1} = 0$ iff $x_{i2} = 1$. So the denominator is

$$\alpha_1^2|S| + \alpha_2^2|S^C|$$

Therefore

$$\frac{x^TLx}{x^Tx} = \frac{E(S,S^C)|(\alpha_1-\alpha 2)^2}{\alpha_1^2|S|+\alpha_2^2|S^C|}$$

We want to show that this is less than or equal to $2\phi_G$. By Eq. 4.4 we can do this by showing

$$\frac{|E(S,S^C)|(\alpha_1-\alpha 2)^2}{\alpha_1^2|S|+\alpha_2^2|S^C|} \leq 2\frac{|E(S,S^C)|}{|S|} = 2\phi_G$$

For simplicity sake we will prove this with different notation. Let $a = |E(S,S^C)|$, $b = |S|$, $c = |S^C|$. We aim to show that

$$\frac{a(\alpha_1-\alpha 2)^2}{\alpha_1^2 b+\alpha_2^2 c} \leq 2\frac{a}{b}$$

To prove this first recall that $|S| \leq \frac{n}{2}$ so $b \leq \frac{n}{2}$. Then $3b \leq 3\frac{n}{2}$ and since $3\frac{n}{2} \leq 2n$ we know $3b \leq 2n$ so $b \leq 2n-2b$ so $1 \leq \frac{2(n-b)}{b}$. Since $c = |S^C|$ and $|S^C|+|S| = n$ we have $c = (n-b)$. Then $1 \leq \frac{2c}{b}$ so $0 \leq \frac{2c}{b}-1$. Since we know

$$0 \leq (\alpha_1+\alpha_2)^2 = \alpha_1^2+2\alpha_1\alpha_2+\alpha_2^2$$

then since $0 \leq \frac{2c}{b}-1$ we have,

$$0 \leq \alpha_1^2+2\alpha_1\alpha_2+(\tfrac{2c}{b}-1)\alpha_2^2$$

Adding $\alpha_1^2-2\alpha_1\alpha 2+\alpha_2^2$ to both sides we get,

$$\alpha_1^2-2\alpha_1\alpha 2+\alpha_2^2 \leq 2\alpha_1^2+\tfrac{2c}{b}\alpha_2^2$$

We can rewrite this as,

$$(\alpha_1-\alpha_2)^2 \leq \tfrac{2}{b}(\alpha_1^2 b+\alpha_2^2 c)$$

Since $0 < \alpha_1^2 b + \alpha_2^2 c$ we can divide to get

$$\frac{(\alpha_1-\alpha_2)^2}{\alpha_1^2 b+\alpha_2^2 c} \leq \frac{2}{b}$$

Since $0 \leq a$ we can multiply each side by a to get

$$\frac{a(\alpha_1-\alpha 2)^2}{\alpha_1^2 b+\alpha_2^2 c} \leq 2\frac{a}{b}$$

Which is what we wanted to show. This shows that for any $x \in \text{span}(x_1, x_2)$, $\max\limits_{x\in\text{span}(x_1,x_2)} \frac{x^T Lx}{x^T x} \leq 2\phi_G$. As previously mentioned, by Lemma 4.3 this proves that $\lambda_2 \leq 2\phi_G$. □

Now we present a few preliminary definitions and lemmas required to prove the right side of Cheeger's Inequality for d-regular graphs.

Definition 4.5. Let X be a discrete random variable which can take values $x_1, x_2, \ldots, x_n$ with corresponding probabilities $p_1, p_2, \ldots, p_n$. The expectation of X is denoted as $\mathbb{E}[X]$ and is defined to be $\sum\limits_i x_i p_i$

Lemma 4.6. *Suppose X_1 can take on values $\{x_{11}, x_{12}, \ldots, x_{1k}\}$ where each $x_{1i} \in \mathbb{R}^+$ and X_2 can take on values $\{x_{21}, x_{22}, \ldots, x_{2l}\}$ where each $x_{2i} \in \mathbb{R}^+$. Let $\beta \in \mathbb{R}$. If $\frac{\mathbb{E}[X_1]}{\mathbb{E}[X_2]} \leq \beta$ then there exists some x_{1i}, x_{2j} such that $\frac{x_{1i}}{x_{2j}} \leq \beta$.*

Proof. We prove by contradiction. Suppose for any $x_{1i} \in \{x_{11}, x_{12}, \ldots, x_{1k}\}$ and for any $x_{2j} \in \{x_{21}, x_{22}, \ldots, x_{2l}\}$ we have

$$\frac{x_{1i}}{x_{2j}} > \beta \text{ so } x_{1i} > \beta x_{2j}$$

Let $x_{1r} = \min\{x_{11}, x_{12}, \ldots, x_{1k}\}$ and $x_{2s} = \max\{x_{21}, x_{22}, \ldots, x_{2k}\}$. We have $x_{1r} > \beta x_{2s}$. Let p_{1i} denote the probability that $X_1 = x_{1i}$ and p_{2j} denote the probability that $X_2 = x_{2j}$. Then

$$\sum_i p_{1i} = \sum_j p_{2j} = 1$$

so now

$$x_{1r}\sum_i p_{1i} > \beta x_{2s}\sum_j p_{2j}$$

Since $x_{1r} = \min\{x_{11}, \ldots, x_{1k}\}$ and $x_{2s} = \max\{x_{21}, \ldots, x_{2k}\}$ we have that

$$\sum_i x_{1i}p_{1i} \geq x_{1r}\sum_i p_{1i} > \beta x_{2s}\sum_j p_{2j} \geq \beta\sum_j x_{2j}p_{2j}$$

This implies that

$$\sum_i x_{1i}p_{1i} > \beta\sum_j x_{2j}p_{2j}$$

By definition of expectation this is the same as

$$\mathbb{E}[X_1] > \beta\mathbb{E}[X_2]$$

Which implies that

$$\frac{\mathbb{E}[X_1]}{\mathbb{E}[X_2]} > \beta$$

This is a contradiction since $\frac{\mathbb{E}[X_1]}{\mathbb{E}[X_2]} \leq \beta$. □

Lemma 4.7. *Let* $\mathbf{1}$ *be the column vector in* $\mathbb{R}^n$ *defined by* $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. *Let* x *be a column vector in* $\mathbb{R}^n$ *such that* $< x, \mathbf{1} >= 0$. *Let* $\alpha \in \mathbb{R}$ *and let* L *be some Laplacian matrix. Then* $\frac{(x+\alpha\mathbf{1})^T L(x+\alpha\mathbf{1})}{(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1})} \leq \frac{x^T Lx}{x^T x}$.

Proof. By Lemma 3.6,

$$\frac{(x+\alpha\mathbf{1})^T L(x+\alpha\mathbf{1})}{(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1})} = \frac{\sum\limits_{[i,j]\in E}[(x_i+\alpha)-(x_j+\alpha)]^2}{\sum\limits_{i}(x_i+\alpha)^2}$$

Note that

$$\sum_{[i,j]\in E}[(x_i+\alpha)-(x_j+\alpha)]^2 = \sum_{[i,j]\in E}[x_i - x_j]^2 = x^T Lx$$

So the numerator of $\frac{(x+\alpha\mathbf{1})^T L(x+\alpha\mathbf{1})}{(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1})}$ is equal to the numerator of $\frac{x^T Lx}{x^T x}$. Then we must show that the denominator of $\frac{(x+\alpha\mathbf{1})^T L(x+\alpha\mathbf{1})}{(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1})}$ is greater than the denominator of $\frac{x^T Lx}{x^T x}$, i.e. $(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1}) \geq x^T x$. One can see that

$(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1}) = \sum\limits_i (x_i+\alpha)^2 = \sum\limits_i x_i^2 + 2\sum\limits_i x_i + \sum \alpha^2 = x^T x + 2\langle x, \mathbf{1}\rangle + \alpha^2\langle \mathbf{1}, \mathbf{1}\rangle$

Since $\langle x, \mathbf{1}\rangle = 0$ we get that,

$$(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1}) = x^T x + \alpha^2\langle \mathbf{1}, \mathbf{1}\rangle$$

And

$$\alpha^2\langle \mathbf{1}, \mathbf{1}\rangle \geq 0 \text{ so } x^T x + \alpha^2\langle \mathbf{1}, \mathbf{1}\rangle \geq x^T x$$

This implies that

$$(x+\alpha\mathbf{1})^T(x+\alpha\mathbf{1}) \geq x^T x$$

□

Note that this lemma explains why we take the second smallest eigenvalue. By Lemma 3.5 $x = \mathbf{1}$ is an eigenvector with eigenvalue 0, which is the smallest eigenvalue for the Laplacian. By Spectral Theorem we have a basis of eigenvectors that are orthogonal to each other so the eigenvector that corresponds to λ_2 must be orthogonal to $\mathbf{1}$. So the eigenvector corresponding to λ_2 satisfies our lemma above, which is necessary to prove Cheeger's Inequality.

Lemma 4.8. *Let* $R_L(x)$ *denote* $\frac{x^T Lx}{x^T x}$. *For any* $\alpha \in \mathbb{R}^n \backslash \{0\}$ *it is true that* $R_L(\alpha x) = R_L(x)$.

Proof. By Lemma 3.6,

$$R_L(\alpha x) = \frac{\sum\limits_{[i,j]\in E}(\alpha x_i - \alpha x_j)^2}{\sum\limits_i(\alpha x_i)^2} = \frac{\alpha^2 \sum\limits_{[i,j]\in E}(x_i - x_j)^2}{\alpha^2 \sum\limits_i x_i^2} = \frac{\sum\limits_{[i,j]\in E}(x_i - x_j)^2}{\sum\limits_i x_i^2} = R_L(x)$$

[4] □

Lemma 4.9. *For $a, b \in \mathbb{R}$, it is true that $a^2 + b^2 \geq 2ab$*

Now we prove that for a d-regular graph G, $\phi_G \leq 2\sqrt{\frac{\lambda_2}{d}}$. This proof assumes we know the eigenvector that corresponds to λ_2 and aims to show that there exists a subset of vertices S of the graph G such that $\phi_S \leq 2\sqrt{\frac{\lambda_2}{d}}$ which would imply that $\phi_G \leq 2\sqrt{\frac{\lambda_2}{d}}$ since $\phi_G \leq \phi_S$.

Proof. Suppose we are given an eigenvector

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

that corresponds to λ_2, that is $Lx = \lambda_2 x$. From discussion above we know that this x must satisfy $< x, \mathbf{1} >= 0$ and note that $\frac{x^T Lx}{x^T x} = \lambda_2$. Pick $\alpha \in \mathbb{R}$ such that $\{(x_1 + \alpha, x_2 + \alpha, \ldots, x_n + \alpha\}$ has median value 0. By Lemma 4.7 it follows that

$$\frac{(x + \alpha\mathbf{1})^T L(x + \alpha\mathbf{1})}{(x + \alpha\mathbf{1})^T (x + \alpha\mathbf{1})} \leq \frac{x^T Lx}{x^T x} \leq \lambda_2 \tag{4.10}$$

Although it may not be apparent at present, we shift our eigenvector to have median

value 0 to ultimately ensure that the subset of vertices S has less than $\frac{n}{2}$ vertices. Let

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

denote $x + \alpha 1$. Then by Eq. 4.10 we know $R_L(y) \leq \lambda_2$.

Define $y^+ = (y_i^+)$ where $y_i^+ = y_i$ if $y_i \geq 0$ and $y_i^+ = 0$ otherwise. Define $y^- = (y_i^-)$ where $y_i^- = y_i$ if $y_i \leq 0$ and $y_i^- = 0$ otherwise. We claim that either $R_L(y^+) \leq 2R_L(y)$ or $R_L(y^-) \leq 2R_L(y)$. By Lemma 3.6,

$$R_L(y^+) = \frac{\sum_{[i,j]\in E} (y_i^+ - y_j^+)}{\sum_i (y_i^+)^2}$$

and

$$R_L(y^-) = \frac{\sum_{[i,j]\in E} (y_i^- - y_j^-)}{\sum_i (y_i^-)^2}$$

In this form one can see that

$$\sum_{[i,j]\in E} (y_i^+ - y_j^+) \leq \sum_i (y_i - y_j) \text{ so } \mathrm{num} R_L(y^+) \leq \mathrm{num} R_L(y)$$

Similarly,

$$\sum_{[i,j]\in E}(y_i^- - y_j^-) \leq \sum_i (y_i - y_j)$$

so num$R_L(y^-) \leq$ num$R_L(y)$. Now we must consider the denominators. Note that denom$R_L(y^+)$ denotes the number of $y_i \geq 0$, denom$R_L(y^-)$ denotes the number of $y_i \leq 0$, and denom$R_L(y)$ denotes the total number of y_i. Thus

$$\text{denom}R_L(y^+) + \text{ denom}R_L(y^-) = \text{denom}R_L(y)$$

Suppose

$$\text{denom}R_L(y^+) \geq \text{denom}R_L(y^-)$$

then

$$2\text{denom}R_L(y^+) \geq \text{denom}R_L(y)$$

Then it follows that

$$R_L(y^+) \leq 2R_L(y)$$

The same argument shows that if

$$\text{denom}R_L(y^-) \geq \text{ denom}R_L(y^+)$$

then

$$R_L(y^-) \leq 2R_L(y)$$

For the rest of the proof suppose without loss of generality that denom$R_L(y^+) \geq$ denom$R_L(y^-)$ so that $R_L(y^+) \leq 2R_L(y)$. Recall that $R_L(y) \leq \lambda_2$. Thus $R_L(y^+) \leq 2\lambda_2$. Now we scale y^+ by some $\beta \in \mathbb{R}$ so that

$$\{\beta y_1^+, \beta y_2^+, \ldots, \beta y_n^+\} \subseteq [0,1]$$

By Lemma 4.8,

$$R_L(\beta y^+) = R_L(y^+)$$

Therefore,

$$R_L(\beta y^+) \leq 2\lambda_2$$

We construct S as follows: Pick $t \in [0,1]$ at random and let $S = \{i, \text{s.t. } (\beta y_i^+)^2 \geq t\}$ (Recall that i denotes a vertex in our graph G). We will now show that with this construction of S it is true that

$$\frac{\mathbb{E}[|(S,S^C)|]}{d\mathbb{E}[|S|]} \leq \sqrt{\frac{2}{d}R_L(y^+)}$$

so that

$$\frac{\mathbb{E}[|(S,S^C)|]}{d\mathbb{E}[|S|]} \leq 2\sqrt{\frac{\lambda_2}{d}}$$

Let $Y_{[i,j]} = 1$ if $i \in S$ and $j \in S^C$ or if $j \in S$ and $i \in S^C$, and $Y_{[i,j]} = 0$ otherwise. Then if $[i,j]$ is an edge between S and S^C we get that $Y_{[i,j]} = 1$ (but note that the converse is not necessarily true). Then

$$\mathbb{E}[|E(S,S^C)|] \leq \mathbb{E}[\sum_{[i,j]\in E} Y_{[i,j]}]$$

By linearity of expectation this is equal to $\sum\limits_{[i,j]\in E} \mathbb{E}[Y_{[i,j]}]$. Note that $Y_{[i,j]} = 1$ iff $(y_i^+)^2 \le t \le (y_j^+)^2$ since this implies that $i \in S$ and $j \in S^C$. So now we have that

$$\sum_{[i,j]\in E} \mathbb{E}[Y_{[i,j]}] = \sum_{[i,j]\in E} P((y_i^+)^2 \le t \le (y_j^+)^2)$$

To get this probability, recognize that each $(y_i^+)^2 \in [0,1]$, so the probability that our random t is between $(y_i^+)^2$ and $(y_j^+)^2$ is the distance between them which is $|(y_i^+)^2 - (y_j^+)^2|$. Now,

$$\mathbb{E}[|E(S,S^C)|] \le \sum_{[i,j]\in E} |(y_i^+)^2 - (y_j^+)^2| = \sum_{[i,j]\in E} |y_i^+ - y_j^+||y_i^+ + y_j^+|$$

By the Cauchy-Schwarz Inequality this is less than or equal to

$$\sqrt{\sum_{[i,j]\in E} (y_i^+ - y_j^+)^2}\sqrt{\sum_{[i,j]\in E} (y_i^+ + y_j^+)^2}$$

By Lemma 4.9 this is less than or equal to

$$\sqrt{\sum_{[i,j]\in E} (y_i^+ - y_j^+)^2}\sqrt{\sum_{[i,j]\in E} 2(y_i^+)^2 + (y_j^+)^2}$$

Note that this sum is over all edges $[i,j]$ of G and so in $\sum\limits_{[i,j]\in E} (y_i^+)^2 + (y_j^+)^2$ each $(y_i^+)^2$ term will appear $d(i)$ times (where $d(i)$ denotes the degree of vertex i). For a d-regular graph we get

$$2\sum_{[i,j]\in E} (y_i^+)^2 + (y_j^+)^2 = \underset{[i,j]\in E}{2} \sum d(y_i^+)^2 = 2d\sum_{[i,j]\in E} (y_i^+)^2$$

So we've shown that

$$\mathbb{E}[|E(S,S^C)|] \le \sqrt{\sum_{[i,j]\in E} (y_i^+ - y_j^+)^2}\sqrt{2d\sum_{[i,j]\in E} (y_i^+)^2}$$

This is a great step towards showing

$$\frac{\mathbb{E}[|(S,S^C)|]}{d\mathbb{E}[|S|]} \le \sqrt{\frac{2}{d}R_L(y^+)}$$

but now we must consider the denominator. Let $Z_i = 1$ if $i \in S$ and $Z_I = 0$ otherwise. With this construction we have $\mathbb{E}[|S|] = \mathbb{E}[\sum\limits_i Z_i]$. By lineary of expectation this is equal to $\sum\limits_i \mathbb{E}[Z_i]$ which is just $\sum\limits_i P(t \le (y_i^+)^2)$. To get this probability, recognize that each $(y_i^+)^2 \in [0,1]$, so the probability that our random t is less than or equal to $(y_i^+)^2$ is given by $(y_i^+)^2$ itself. So

$$\mathbb{E}[|S|] = \sum_i P(t \le (y_i^+)^2) = \sum_i (y_i^+)^2$$

Thus

$$\frac{\mathbb{E}[|(S,S^C)|]}{d\mathbb{E}[|S|]} \le \frac{\sqrt{\sum\limits_{[i,j]\in E}(y_i^+ - y_j^+)^2}\sqrt{2d\sum\limits_{[i,j]\in E}(y_i^+)^2}}{d\sum\limits_i (y_i^+)^2} = \sqrt{\frac{2}{d}}\sqrt{\frac{\sum\limits_{[i,j]\in E}(y_i^+ - y_j^+)^2}{\sum\limits_i (y_i^+)^2}} = \sqrt{\frac{2}{d}R_L(y^+)} \le \sqrt{\frac{2}{d}2\lambda_2} = 2\sqrt{\frac{\lambda_2}{d}}$$

By Lemma 4.6 this implies there exists some S such that $\frac{|(S,S^C)|}{d|S|} \leq 2\sqrt{\frac{\lambda_2}{d}}$. Therefore $\phi_G \leq 2\sqrt{\frac{\lambda_2}{d}}$.

[4] $\square$

Acknowledgments. It is a pleasure to thank my mentor, Clark Butler, for his help, patience, and advice. I would also like to thank Professor Madhur Tulsiani and Professor Laszlo Babai. My primary guide in writing this paper has been their lectures. Furthermore they showed me the power and excitement of algebraic graph theory, which has had a major influence on my life. Finally, I would like to thank Peter May for organizing this REU. It is a wonderful program and I am grateful to have been given the chance to take part.

References

[1] James Lee, Shayan Oveis Gharan, Luca Trevisan. Muti-way spectral partitioning and higher-order Cheeger inequalities. Jan. 26, 2013.
[2] Luca Trevisan. Cheeger's Inequality and Other Approximations of Edge Expansion. April 21, 2008. http://lucatrevisan.wordpress.com/2008/04/21/cheegers-inequality-and-other-approximations-of-edge-expansion/
[3] Luca Trevisan. Expanders Lectures 1-6. http://www.youtube.com/channel/UCRhz0iV4GG90gX9PeBPyveA?feature=watch
[4] Madhur Tulsiani. Gram-Schmidt orthogonalization, Schur's theorem, spectral theorem for normal matrices, eigenvalues and eigenbasis for Hermitian matrices, adjacency matrices of graphs. (July 8, 2013 Lecture Notes) http://ttic.uchicago.edu/ madhurt/courses/reu2013/lecture7.pdf
[5] Madhur Tulsiani. Eigenvalues and eigenvectors of graphs, Rayleigh quotients. (July 9, 2013 Lecture Notes) http://ttic.uchicago.edu/ madhurt/courses/reu2013/lecture8.pdf