AN INTRODUCTION TO SPECTRAL GRAPH THEORY

BRIAN PAN
NATHAN NIE
DANIEL WANG

Abstract. Spectral graph theory bridges the world of linear algebra and graph theory by studying how a graph's connectivity and organization are encoded in the eigenvalues and eigenvectors of its associated matrices. We begin with a concise overview of the fundamental concepts of graph theory and the basics of matrices in linear algebra. Then, we introduce the principal matrix representations of a graph: the adjacency matrix and incidence matrix. We highlight their definitions and properties. Next, we introduce the Laplacian operator and its graph analogue, the Laplacian matrix. Finally, we investigate the eigenvalues of the Laplacian matrix and how they reveal deep insights into the connectivity of the graph.

Contents

Date: May 20, 2025.

1. Background

The first recorded discovery of graph theory was by Leonard Euler, a Swiss mathematician, who developed a theorem to solve the famous Seven Bridges of Königsberg problem in 1736 [5]. Euler wanted to find a way to cross all seven bridges in the city of Königsberg, shown in Figure 1, once and exactly once. In the map, the bridges are colored red, while the river is colored blue. Euler eventually found that this was an impossible task, and this discovery laid the foundation for graph theory.

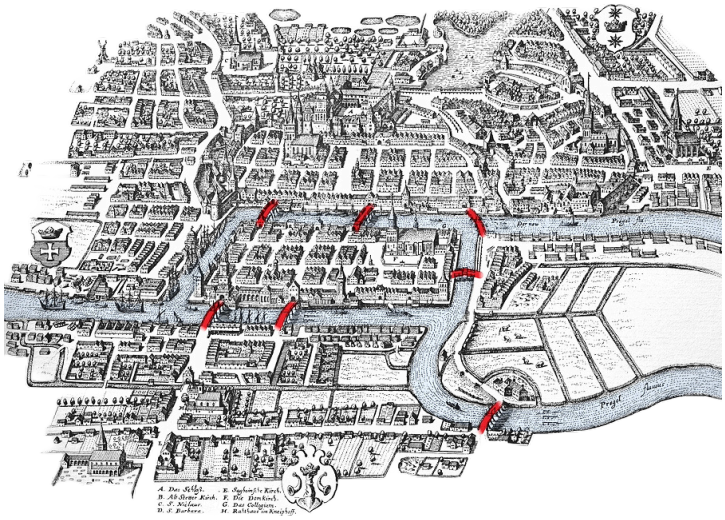


Figure 1. The city and bridges of Königsberg

Since then, the field of graph theory has developed significantly with the work of mathematicians such as Georg Kirchhoff and Arthur Cayley, who have made meaningful contributions to graph theory. We will discuss some of these in this paper, including Kirchhoff's matrix tree theorem [3] and Cayley's tree formula [1].

One unique field in graph theory is spectral graph theory, which combines linear algebra and graph theory. Spectral graph theory originated in the 1950s with the first few pieces of research on structural and spectral properties of graphs, as well as on quantum chemistry. However, this line connecting the two would not be drawn until much later. One of the first books that summarized all of the work in this field was published in 1980 by Cvetković, Doob, and Sachs, which was a monograph called *Spectra of Graphs* [7]. Today, spectral graph theory has developed greatly and is applied to a wide range of fields, including in machine learning, network analysis, autonomous systems, and chemistry.

This paper explores the field of spectral graph theory, which studies the relationships between the structure of a graph and the eigenvalues and eigenvectors of matrices associated with it − particularly the adjacency matrix, incidence matrix, and Laplacian matrix.

We begin with the foundational concepts of graph theory, including the nomenclature and properties of graphs. Shortly after, we introduce linear algebra with the properties of matrices, as well as the eigenvalues and eigenvectors of matrices.

Next, we dive into the relationships between graphs and their matrix representations. More specifically, we look at the properties of the adjacency matrix and the incidence matrix. Then, we introduce the Laplacian operator and its graph analog, the Laplacian matrix, which is a central object in spectral graph theory. Finally, we analyze the key properties of the Laplacian matrix with a focus on its eigenvalues, eigenvectors, and how they relate to the structure of the graph.

1.1. **Introduction to Graphs.** The most basic definition needed in our investigation is that of a graph.

Definition 1.1. [1] A *graph* is an ordered pair $G = (V, E)$ where V is the *vertex set* and E is the *edge set*. An *edge* $e \in E$ is a pair uv of distinct vertices $u, v \in V$.

Hence, there are no loops, meaning that no vertex can have an edge to itself. Moreover, we stipulate that a graph cannot have multiple edges that connect the same pair of vertices. When the edges are given direction, the graph is a *digraph*. Unless stated otherwise, the graphs in this paper are not directed, so we are free to assume $uv = vu$. We also say that $|V|$ is the *order* of the graph, or the number of vertices, of G while $|E|$ is the *size*, or the number of edges, of G.

Two vertices u and v are *adjacent* or *neighbors* if $u, v \in V(G)$ and the edge $e = uv \in E(G)$. The vertex u and the edge e are said to be *incident*.

Consider the following graph in Figure 2. The order $|V|$ is 6, while the size $|E|$ is 6 as well. This graph shows that v_3 is adjacent to v_1 and v_4, while it is incident to $e = v_1v_3$.

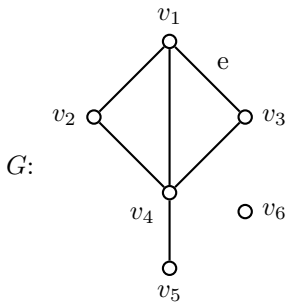


FIGURE 2. A graph

A *subgraph* of a graph $G = (V, E)$ is a smaller graph G' contained within G, formed by taking subsets of V and E. Intuitively, G' is formed by removing vertices and edges from G, but we cannot add vertices or edges.

Definition 1.2. The *degree* $\deg v$ of a vertex v is the number of edges incident with v, equivalent to the number of vertices adjacent to v.

An interesting result involving degrees in graph theory is the *degree sum formula* which states that for any graph, one can relate the number of edges to the degrees of its vertices. This theorem is often referred to as the *First Theorem of Graph Theory* because many who start studying this subject would encounter this result as their first theorem.

Theorem 1.3 (First Theorem of Graph Theory)**.** [1] *If G is a graph of size m, then*

$$\sum_{v \in V(G)} deg\ v = 2m.$$

Proof. Each edge in G represents a degree of two because it is incident with two vertices. Therefore, each edge is counted twice when summing the degrees of the vertices of G. □

As an example, consider the graph in Figure 2. The sum of the degrees of its vertices is $3 + 2 + 2 + 4 + 1 + 0 = 12$, while two times the number of edges is also $2m = 6 \cdot 2 = 12$.

The particular reason why this theorem is also named the *handshaking lemma* by many is because it follows directly from the degree sum formula that there must be an even number of odd vertices in a graph. If there is a party of people who shake hands, the number of people who shake an odd number of other people's hands must be even.

Example 1.4. Consider a graph of order 12 and size 31 in which the degree of each vertex is either 4 or 6. To find the number of vertices of degree 4, let x be the number of vertices of degree 4. Then, there are $12 - x$ vertices of degree 6.

By the degree sum formula, $4x + 6(12 - x) = 2 \cdot 31$, so there are $x = 5$ vertices of degree 4. Consequently, there are $12 - 5 = 7$ vertices of degree 6.

In theory, a graph of order 12 and size 31 with 7 vertices of degree 6 and 5 vertices of degree 4 is tedious to count by hand, so this theorem proves to be extremely beneficial.

Many times, graphs are used to model connections between states in a process, so graph traversals are of great importance. In a graph G, we choose a vertex u to start. Then, we proceed from u to a neighbor of u, then to a neighbor of the neighbor of u, and so on, until we stop at a vertex v. Such a sequence is called a *walk*. The number of edges traversed on a walk is the *length* of that walk.

Definition 1.5. A *$u - v$ walk* W in a graph G is a sequence of vertices starting with u and ending with v where all pairs of consecutive vertices in W are adjacent.

If we continue this walk (until we stop), there is essentially no restriction on what a walk can be. However, there will be instances where we want to place restrictions on walks. If a walk contains no repeated vertices or repeated edges, then it is called a *path*, denoted P_n with n vertices. A path is shown in Figure 3.

Moreover, a *cycle* C_n is a walk with n vertices that starts and ends at the same vertex, but, like paths, cannot have any repeating vertices (except for the first and last) and cannot have any repeating edges. A *k-cycle* is a cycle of length k. A graph is acyclic if it has no cycles. A cycle is shown in Figure 3.

Speaking of graph traversals, it is helpful to know whether we can traverse a graph from any vertex to any other vertex.

Definition 1.6. A graph G is *connected* if G contains a $u - v$ walk for all $u, v \in G$.

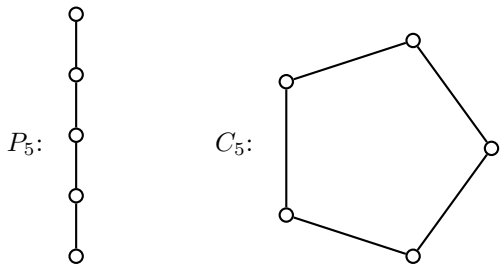


FIGURE 3. A path P_5 and a cycle C_5

In other words, a graph is connected if there exists a walk that connects all pairs of vertices in that graph. If the entire graph is not connected, but rather a specific section, these sections are called *connected components*.

As we continue to study graphs, we will encounter certain graphs with shared characteristics. It is useful to be familiar with some of them.

Definition 1.7. The *complete graph* K_n with n vertices is a graph in which every pair of vertices is adjacent.

Therefore, K_n has the maximum possible size for a graph with n vertices. The complete graph K_4 is shown in Figure 4.

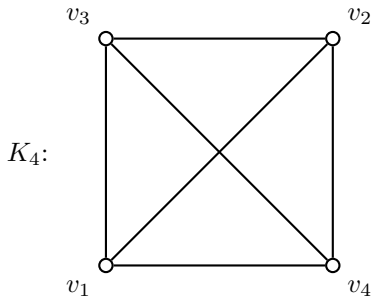


FIGURE 4. A complete graph K_4

Definition 1.8. The *star graph* S_n with n vertices is a graph in which one vertex has degree $n-1$, while every other vertex has a degree of 1.

The star graph takes the shape of a star, hence why it is called a star graph. The star graph S_5 is shown in Figure 5.

1.2. **Introduction to Matrices.** Linear algebra is essential for the study of spectral graph theory, especially in analyzing the matrix representations of graphs, so it is crucial to grasp a solid understanding of the subject.

A *vector* is a tuple of numbers arranged in a vertical fashion. Usually, they are used to represent coordinates in $n-$dimensional space; as such, each entry is well defined according to a single index. A *matrix* is a rectangular array of numbers. An $m \times n$ matrix has m rows and n columns. Each entry in the array is called an *element* of the matrix which can be directly referenced using rows and columns. This is done with the notation A_{ij}, which pinpoints the element in the $i-$th row and the $j-$th column.

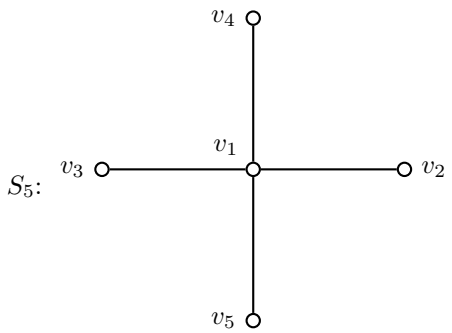


FIGURE 5. A star graph S_5

Definition 1.9. The *transpose* of a matrix A is denoted A^T. The $ij-$th element of A becomes the $ji-$th element of A^T.

Note that the transpose is also well-defined for a vector: instead of being a column of numbers, a transposed vector is a row. We will revisit the matrix transpose in our discussion of the adjacency matrix.

A *square matrix* is a matrix in which the number of rows and columns are the same, say n. The *main diagonal* of a square matrix A is the set of elements A_{ii}, and the *trace* is the sum of the entries on the main diagonal. For example, the identity matrix I, so named because it does not meaningfully transform space, is the square matrix with 1 on the main diagonal and 0 otherwise.

There is a useful way of representing and visualizing transformations in $n-$dimensional space. We do so by multiplying matrices and vectors. When a matrix is applied to a vector, the vector transforms by the matrix A.

Example 1.10. Suppose we multiply the following matrix A and vector v:

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}, \quad v = \begin{bmatrix} 5 \\ 6 \end{bmatrix}.$$

Then, the resulting vector will take the form $[k_1, k_2]^T$. The value of k_1 is the sum of the multiplication of the corresponding elements in the first row of A with v. Thus, we get $k_1 = 2 \cdot 5 + 3 \cdot 6 = 28$.

Similarly, the value of k_2 is the sum of the multiplication of the corresponding elements in the second row of A with v, yielding $k_2 = 4 \cdot 5 + 1 \cdot 6 = 26$.

Therefore, the resulting vector is $Av = [28, 26]^T$.

In summary, to transform space by multiplying a matrix A with a vector v, the number of columns of A must match the number of rows of v. Furthermore, the resulting vector has the number of rows of A, and the individual elements in the resulting vector can be found by multiplying the rows of A with v.

We can also *compose* or multiply matrices with matrices.

Definition 1.11. Let A be a $m \times n$ matrix and B a $n \times p$ matrix. Then

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

In other words, if $A[i,:]$ denotes the $i-$th row of A and $B[:,j]$ is the $j-$th column, then the ij-th entry of AB is exactly $A[i,:]B[j,:]$.

We visualize matrix multiplication with two arbitrary 2×2 matrices shown in Figure 6.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & df+dh \end{bmatrix}.$$

FIGURE 6. The multiplication of two 2×2 matrices

Reversing the order of the two matrices multiplied in Figure 6 produces the result in Figure 7. It does not yield the same result, proving that matrix multiplication is not commutative.

$$\begin{bmatrix} e & f \\ g & h \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ea+fc & eb+fd \\ ga+hc & gb+hd \end{bmatrix}.$$

FIGURE 7. The multiplication of two matrices in reverse order

Next, we examine a crucial part of the matrix: the determinant. The determinant has many uses, including showing how much a matrix scales an area, calculating the inverse of a matrix, and providing useful information on graph structure. We will go back to the discussion of the determinant after the introduction of the eigenvalues of a matrix.

Definition 1.12. The *determinant* of a square matrix A, denoted $\det(A)$, is the product of every eigenvalue of A counted with multiplicities.

Moreover, the *trace* of a square matrix $\mathrm{tr}(A)$, or the sum of the main diagonal, is also equal to the sum of the eigenvalues of A [22].

To visualize an example, we show the determinant for an arbitrary 2×2 and 3×3 matrix in Figure 8. The determinant and trace of a 1×1 matrix is equal to the only element in the matrix.

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = aei + bfh + cdh - ceg - bdi - afh$$

FIGURE 8. The determinant for two arbitrary matrices

1.3. **Eigenvalues and Eigenvectors.** One major backbone of spectral graph theory and linear algebra alike is the use of eigenvalues and eigenvectors. Thus, it is beneficial to introduce the eigenvalues and eigenvectors of matrices. *Eigenvectors* are special vectors such that when a transformation matrix is applied, the resulting vector is simply a scaled version of the original vector. The factor by which it is scaled is the *eigenvalue*.

Definition 1.13. The vector $\vec{v}$ is an *eigenvector* of a matrix A with *eigenvalue* λ if $A\vec{v} = \lambda\vec{v}$.

That is, multiplying the eigenvector $\vec{v}$ by A yields the same result as scaling $\vec{v}$ by the eigenvalue λ.

The eigenvectors and eigenvalues of a matrix can be found using the *characteristic polynomial* of the matrix. The characteristic polynomial is represented by the equation $\det(A - \lambda I) = 0$, which is derived by manipulating the equation $A\vec{v} = \lambda\vec{v}$. The roots of the characteristic polynomial are the eigenvalues. We look at an example to find the eigenvalues and eigenvectors of a matrix for a better understanding.

Example 1.14. Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix}.$$

We start with the equation $A\vec{v} = \lambda\vec{v}$. Here, the left side has a vector $\vec{v}$ that is multiplied by a matrix A, while the right side has a vector $\vec{v}$ being multiplied by a scalar λ. To change the equation where both sides have a vector that is multiplied by a matrix, we multiply the right-hand side by the identity matrix, $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Plugging in A and I yields

$$\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix} \vec{v} = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \vec{v}.$$

Then, subtracting the right side from the left side leads to the equation

$$\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix} \vec{v} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \vec{v} = \vec{0}.$$

From here, we can factor out the eigenvector and multiply the scalar λ into I. We get

$$\left(\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) \vec{v} = \vec{0}.$$

Next, we perform matrix subtraction to get

$$\begin{bmatrix} -6 - \lambda & 3 \\ 4 & 5 - \lambda \end{bmatrix} \vec{v} = \vec{0}.$$

Because we want the eigenvector to be non-zero, the matrix must transform the eigenvector to become the zero vector. In other words, the matrix on the left must squish the eigenvector to the zero vector, and the determinant of the matrix must

be zero. This yields the equation

$$\det \begin{bmatrix} -6-\lambda & 3 \\ 4 & 5-\lambda \end{bmatrix} = 0.$$

Now, we apply the determinant formula for the 2×2 matrix to get:

$$(-6-\lambda)(5-\lambda) - 3 \cdot 4 = 0.$$

Moving forward, this simplifies into the quadratic:

$$\lambda^2 + \lambda - 42 = 0.$$

Therefore, solving for the roots gives $\lambda = -7$, $\lambda = 6$.

Now that we have the eigenvalues, we can solve for the eigenvectors. We start by plugging in the eigenvalue 6 to $A\vec{v} = \lambda\vec{v}$. Since we are working in two dimensions, we can say that $\vec{v} = [x, y]^T$.

$$\begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 6 \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

Multiplying this out and turning it into a system of linear equations yields

$$\begin{cases} -6x + 3y = 6x \\ 4x + 5y = 6y. \end{cases}$$

From either linear equation (both equations will always represent the same line), solving for y leads to the equation $y = 4x$, which means that y is always going to be equal to four times that of x, implying the eigenvectors of A is any non-zero scaler multiplier of $[1, 4]^T$. For example, assuming $x = 2$, we get the eigenvector $\vec{v} = [2, 8]^T$.

Similarly, for $\lambda = -7$, we can follow the process to obtain the following system of equations:

$$\begin{cases} -6x + 3y = -7x \\ 4x + 5y = -7y. \end{cases}$$

For $\lambda = -7$, the eigenvector is $\vec{v} = k[-3, 1]^T$ for some non-zero scalar k.

The graphical representation of the eigenvectors of the matrix in the above example is shown in Figure 9. The thick red vector represents the eigenvector $\vec{v} = [1, 4]^T$, and the thin red vector extending from that represents the result of A being applied, or multiplied, to $\vec{v}$. In addition, we see that the thin red vector is simply the thick red vector scaled by a factor of 6, which is the corresponding eigenvalue of the eigenvector. Similarly, the thick blue vector transforms into the thin blue vector with a transformation of A, which equivalently scales the vector by its eigenvalue -7.

Since the characteristic polynomial was used to find the eigenvectors and eigenvalues of a matrix, it is possible to have no real eigenvalues and eigenvectors, one real eigenvalue and eigenvector, or multiple real eigenvalues and eigenvectors, depending on how many real roots the characteristic polynomial has, which stems from the matrix elements themselves.

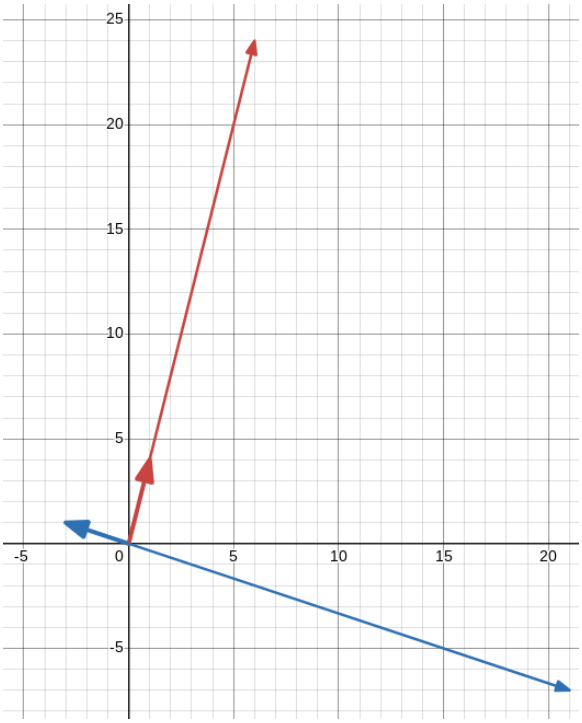


FIGURE 9. Eigenvectors of matrix A

2. Adjacency Matrices

One key matrix representation of graphs is the adjacency matrix. The adjacency matrix is a square matrix with its elements representing whether or not a pair of vertices is adjacent in a graph.

The concept of the adjacency matrix developed back in the early 20th century, and built upon foundational work of mathematicians like Euler. It was in the mid-20th century that published works on the adjacency matrix developed and included the properties of adjacency matrices when they are raised to powers [18], which we look at below.

2.1. Definition of the Adjacency Matrix.

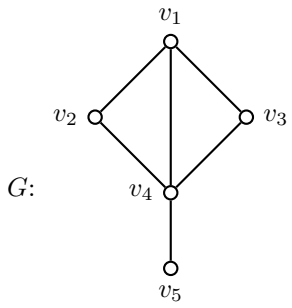
To start our investigation of the adjacency matrix, we first must be familiar with the definition.

Definition 2.1. [1] Let G be a graph of order n and size m, where $V(G) = \{v_1, v_2, v_3, \cdots, v_n\}$ and $E(G) = \{e_1, e_2, e_3, \cdots, e_m\}$. Then, we define the adjacency matrix of G as the $n \times n$ matrix $A = [A_{ij}]$, where

$$A_{ij} = \begin{cases} 1 & \text{if } v_i v_j \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

Thus, an adjacency matrix is symmetric. The graph G and its corresponding adjacency matrix are shown in Figure 10.

For a directed graph, an element $A_{ij} = 1$ if the edge begins at v_i and ends at v_j. Otherwise, it is 0. Therefore, adjacency matrices of directed graphs will not always be symmetric.



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

FIGURE 10. The adjacency matrix of a graph

2.2. **The Power of the Adjacency Matrix.** Another way to view the adjacency matrix is that it provides the number of walks of length 1 from one vertex to another. That, of course, only leads to two possible values: 1 or 0. However, when the adjacency matrix is raised to a higher power, a pattern arises. We investigate the unique transformations the adjacency matrix goes through when it is raised to a power.

Theorem 2.2. [1] *Let G be a graph with vertex set $V(G) = \{v_1, v_2, \cdots, v_n\}$ and adjacency matrix A. Then, the element $(A^n)_{ij}$ is the number of distinct $v_i - v_j$ walks of length n in G.*

Proof. We proceed by induction. First, we set the *base case*. If $n = 1$, then an entry A_{ij} in $A^1 = A$ is 1 if v_i and v_j are adjacent, and 0 otherwise. Therefore, A_{ij} counts the number of walks of length 1 from v_i to v_j.

Next, we perform the *inductive step*. Assume for some $k \geq 1$ that $(A^k)_{ab}$ counts the number of walks of length k from v_a to v_b. Consider a $v_a - v_b - v_c$ walk of length $k+1$. Every such walk must have a $v_a - v_b$ walk of length k, followed by a $v_b - v_c$ walk of length 1. By the definition of matrix multiplication,

$$(A^k A)_{ij} = \sum_n (A^k)_{in} A_{nj}.$$

By the inductive hypothesis, $(A^k)_{ab}$ counts the number of $v_a - v_b$ walks of length k, while A_{bc} is 1 if v_b and v_c are adjacent and 0 otherwise. Hence, an entry $(A_{ij})^{k+1}$ provides the total number of distinct $v_i - v_j$ walks of length $k+1$ in G. □

To visualize this change, raising the adjacency matrix introduced in Figure 10 to the second and third power gives the matrices in Figure 11.

$$A^2 = \begin{bmatrix} 3 & 1 & 1 & 2 & 1 \\ 1 & 2 & 2 & 1 & 1 \\ 1 & 2 & 2 & 1 & 0 \\ 2 & 1 & 1 & 4 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix} \qquad A^3 = \begin{bmatrix} 4 & 5 & 5 & 6 & 2 \\ 5 & 2 & 2 & 6 & 1 \\ 5 & 2 & 2 & 6 & 1 \\ 6 & 6 & 6 & 4 & 4 \\ 2 & 1 & 1 & 4 & 0 \end{bmatrix}$$

FIGURE 11. Powers of an adjacency matrix

Additionally, note that $(A^2)_{ii} = \deg(v_i)$ since the degree represents the number of walks i with length 2. This also implies that $(A^2)_{ii} = \sum_k A_{ik}$.

An interesting property involving the trace also arises when the adjacency matrix is raised to a power.

Proposition 2.3. *The trace of A^2 is twice the number of edges in the graph.*

Proof. Let v_i be some vertex of G. Then, the entry $(A^2)_{ij}$ is equal to the degree of v_i. This means that

$$\operatorname{tr}(A^2) = \sum_{v_i} \deg(v_i) = 2|E(G)|.$$

□

3. Incidence Matrices

Accompanying the adjacency matrix is its sibling, the incidence matrix. The incidence matrix has elements that represent whether or not a vertex and an edge are incident in a graph. The study of the incidence matrix dates back to the 19th century, with one of the first textbooks defining incidence matrices [19].

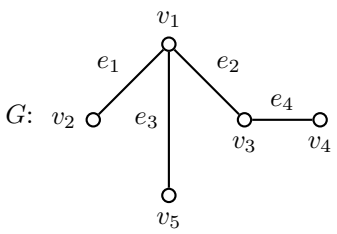
3.1. Definition of the Incidence Matrix.

The incidence matrix, like the adjacency matrix, comes with a simple definition.

Definition 3.1. [1] Let G be a graph of order n and size m, where $V(G) = \{v_1, v_2, v_3, \cdots, v_n\}$ and $E(G) = \{e_1, e_2, e_3, \cdots, e_m\}$. Then, the incidence matrix of G is the $n \times m$ matrix $B = [B_{ij}]$, where

$$B_{ij} = \begin{cases} 1 & \text{if } v_i \text{ is incident with } e_j \\ 0 & \text{otherwise.} \end{cases}$$

A graph G and its corresponding incidence matrix are shown in Figure 12.



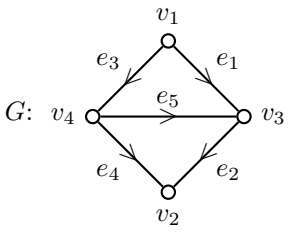
$$B = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Figure 12. The incidence matrix of a graph

The incidence matrix of a directed graph functions in a similar way. On a directed edge, the initial vertex in the incidence matrix is -1, while the ending vertex is 1 [6], with an example shown in Figure 13.

4. The Laplacian Operator

The Laplacian matrix ultimately stems from the Laplacian operator. Therefore, before introducing the Laplacian matrix, we must first become acquainted with the Laplacian operator.



$$B = \begin{bmatrix} -1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix}$$

FIGURE 13. The incidence matrix of a directed graph

The *Laplacian operator*, named after the French mathematician Pierre-Simon Laplace, was first studied in 1786 in which Laplace applied this operator to the study of celestial mechanics. Today, it is encountered all throughout mathematics and physics: in the wave equation, Schrödinger's equation, Poisson's equation, and of course, Laplace's equation.

The Laplacian normally appears in a multivariable calculus or partial differentiation course, so the study of the Laplacian requires a strong calculus background. In simple terms, the Laplacian operator takes the difference between the value of a function at a point and the average of its values at nearby points. However, to gain a more solid understanding of the Laplacian, we first discuss the intuition behind this operator and then introduce the definition.

4.1. **Intuition Behind the Laplacian.** Take the graph of a function $f(x, y)$ that represents the elevation of an area of land. The function receives the specific position of the ground x, y as input and outputs the elevation z of the land at that point.

After graphing $f(x, y)$ on a three-dimensional plane, we may notice that certain x, y points produce relatively large z while other x, y points produce relatively small z. To represent these changes in the graph, the *gradient* of a function $Df(x, y)$ is used [8].

The gradient is a specific type of *vector field* that assigns a vector to every point in space. Every vector field consists of *vectors* with a magnitude and direction. In the case of the gradient, these vectors point to areas with the fastest rate of increase [8]. They represent what and how much change there is between two points. As an example, the vectors in a vector field would point away from spots that are relatively small in value, while the vectors would point toward spots that are relatively large in value. An example of a three-dimensional function gradient is shown in Figure 14.

The behavior of the gradient around a point is the idea of *divergence*, which measures how much the vectors point away from a point. *High* divergence at a point signals that the vectors surrounding that point have a net direction away from it. Similarly, *low* divergence at a point means that the vectors surrounding that point have a net direction towards it [8].

The question arises: How much divergence does the gradient of a function have at a specific point? In other words, exactly how much of a "minimum point" would (x, y) be in terms of $f(x, y)$?

The *Laplacian operator* is used to solve this problem, which gives us a good idea of how functions change across space by describing the divergence of the gradient of a function at a particular point. In other words, it measures how "smooth" a

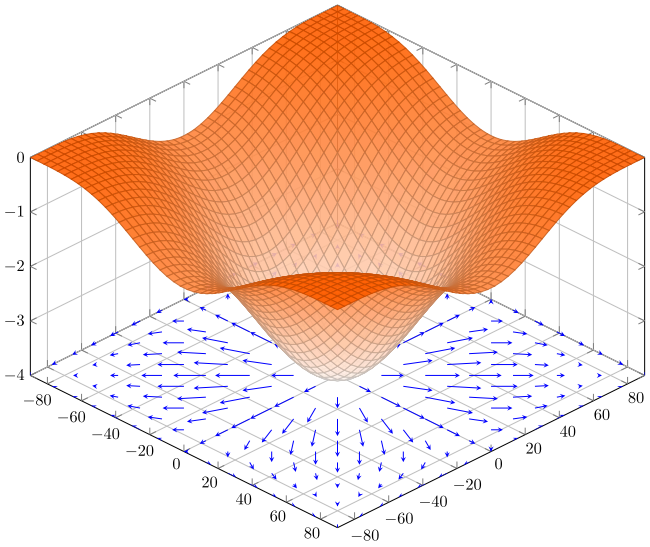


FIGURE 14. The gradient and vector field of $f(x, y) = -(\cos^2 x + \cos^2 y)^2$

function is over its domain. This can be used to help find a local minimum of the function, also known as *minimizing* the function [8].

The Laplacian is especially useful in modern-day machine learning. As stated previously, the Laplacian finds the difference between a point and the points in its vicinity. Now, researchers have applied this to machine learning models, where there are many use cases.

One such use case is with the machine's *cost function*. The cost function describes the difference between the predictions of a model and the actual value of the element it is predicting. Using the Laplacian, the models can measure how "wrong" they are, hence allowing them to improve these models for a better outcome.

4.2. **Laplace's Equation.** Now that we have an intuitive sense of what the Laplacian does, we will briefly define it and discuss some of the more technical details.

Definition 4.1. [9] The *Laplacian* (or *Laplace operator*) of a multivariable function $f(x_1, x_2, \cdots, x_n)$, denoted $\Delta f = \nabla^2 f$, is a differential operator that takes the divergence of the gradient of the function f.

A *differential operator* is an operator that is a function of the differentiation operator. In other words, it uses derivatives. The *derivative* in calculus is the slope of the tangent line of a curve. The *partial derivative* essentially takes the derivative of a multivariable function with respect to one of its variables while keeping all other variables constant. It would be equivalent to finding the "slope" at a point in a multi-dimensional graph. The vector field discussed earlier uses these partial derivatives to map out the directions and magnitudes of each individual arrow at a point in space to create the gradient.

Using the gradient, we've discussed how to find the direction and magnitude of the fastest increase. If we wanted to test whether or not the point we have is the lowest, we would use Laplace's equation.

Definition 4.2. [9] *Laplace's equation* is a second-order partial differential equation, written as $\Delta f = \nabla^2 f = 0$. It sums the second derivatives of each dimension with respect to one dimension, and sets the sum to 0. Laplace's equation in 2 dimensions with respect to u is

$$\nabla^2 u = \frac{\partial^2 u}{\partial^2 x^2} + \frac{\partial^2 u}{\partial^2 y^2} = 0,$$

while a general form with n total dimensions with respect to u is

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = 0.$$

Laplace's equation helps determine whether or not the output of a function is the lowest compared to its surroundings. In other words, it finds a minimum of the function. Referring back to Figure 14, using this equation, we would expect the input $(x, y) = (0, 0)$ to return a true statement because that point is a minimum of the function.

5. Laplacian Matrices

Now, we can dive into the meat and bones of spectral graph theory. The *Laplacian matrix*, also called the *graph Laplacian*, is used greatly across multiple fields of science. One of the first celebrated results with the Laplacian matrix was brought about by one of Kirchhoff's papers on electrical networks in 1847 [16]. Later, the work of Fiedler in 1973 paved the way for more in-depth studies on the Laplacian matrix [16].

Although seemingly arbitrary, this matrix representation of a graph provides us with a powerful way to observe and investigate connectedness in a graph. After discussing the definition and some properties of the Laplacian matrix, we investigate the eigenvalues of the Laplacian matrix and the properties that we can deduce from them, along with some applications of these properties, including spectral partitioning. Finally, we will discuss the famous spanning tree problem.

5.1. Definition of the Laplacian Matrix.

How does the Laplacian translate to graphs? The Laplacian matrix also examines the relationship between one vertex and the others around it, but in a slightly different manner. We will investigate how it does so. First, we define the Laplacian matrix.

Definition 5.1. [4] The Laplacian matrix of a graph G is $L = D - A$, where D is the degree matrix of G and A is the adjacency matrix of G.

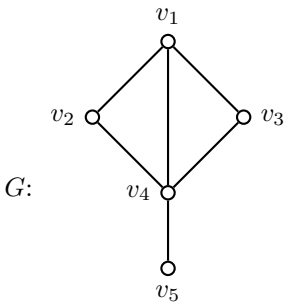
An element D_{ij} of the degree matrix D is $\deg v_i$ if $i = j$. Otherwise, $D_{ij} = 0$. From this definition, we are able to derive another similar definition of the Laplacian matrix by examining the relationships between vertices.

Definition 5.2. [4] Let $G = (V, E)$ be a graph with $V(G) = \{v_1, v_2, \cdots, v_n\}$. Then, the Laplacian matrix $L(G)$ is the $n \times n$ matrix $L = [L_{ij}]$, where

$$L_{ij} = \begin{cases} \deg(v_i) & \text{if } i = j \\ -1 & \text{if } v_i v_j \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

For a directed graph, the Laplacian can be generalized by assigning $L_{ij} = -m$ for m directed edges from v_i to v_j, while L_{ii} is equal to the number of incoming edges to v_i.

We know that the Laplace operator measures the divergence of the gradient of a function f at a point. At the core, it analyzes the values surrounding that point. The Laplacian matrix functions in a similar way to the Laplacian in that each vertex plays the role of a point, or the output of a function, while the incident edges form the surroundings of the point. As an example, a graph and the Laplacian matrix of that graph are shown in Figure 15.



$$L = \begin{bmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 & 0 \\ -1 & 0 & 2 & -1 & 0 \\ -1 & -1 & -1 & 4 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

FIGURE 15. A graph and its Laplacian matrix

In Figure 15, the vertex v_1 may represent the value $f(1)$. Similarly, v_2 represents the value $f(2)$, v_3 represents $f(3)$, and so on. Essentially, the function assigns a value to each vertex of a graph. By replacing the derivatives with differences and swapping the divergence with an average of the incident edges, we arrive at the Laplacian matrix L.

Additionally, a third definition of the Laplacian matrix can be derived from the first two definitions, which may seem random, but is crucial to the further study of its properties.

Lemma 5.3. [11][6] *The Laplacian matrix is also defined as $L = BB^T$ where B is the incidence matrix.*

Proof. For $v_i, v_j \in V(G)$, the corresponding element in BB^T can be expressed as

$$(BB^T)_{ij} = \sum_{e_k \in E(G)} B_{ik} B_{jk}.$$

First, consider the case $i = j$. For each edge incident to v_i, B_{ik} will be either 1 or -1 based on the direction. Then $B_{ik}B_{jk} = (B_{ik})^2 = 1$ for each edge incident to v_i. So $(BB^T)_{ij}$ counts the total number of edges incident to v_i, or $\deg(v_i)$, as claimed.

Next, consider the second case $i \neq j$. If the two vertices v_i and v_j are adjacent, then one of B_{ik} and B_{jk} must be 1, while the other must be -1. This is due to the direction of the edges. Thus, $B_{ik}B_{jk} = -1$. If v_i and v_j are not adjacent, then the result is 0, as claimed. □

5.2. **Eigenvalues of the Laplacian Matrix.** As previously mentioned, the eigenvalues and eigenvectors of a matrix make up the backbone of spectral graph theory, and it does not disappoint for the Laplacian matrix. The eigenvalues of the Laplacian

matrix are extremely interesting because they uncover many properties between a vertex and its neighboring vertices, including information about its connectivity and geometric properties. First, we examine an important property of the Laplacian matrix.

Proposition 5.4. *The Laplacian matrix L is a positive symmetric semi-definite matrix.*

Remark 5.5. [11] The Laplacian matrix is symmetric because both the adjacency matrix A and the degree matrix D are symmetric. The term $x^T Lx$, called the *Dirichlet energy* of a graph, measures the variation between pairs of adjacent vertices where x is some real column vector whose elements represent the vertices in a graph. The expanded form is

$$x^T Lx = \sum_{ij \in E(G)} (x_i - x_j)^2.$$

This means $x^T Lx \geq 0$, so the Laplacian matrix is a positive symmetric semi-definite matrix.

This result directly translates to the eigenvalues being nonnegative as well. Knowing that the Laplacian matrix has nonnegative eigenvalues says a lot about the properties of the matrix. If a matrix has nonnegative eigenvalues, then its main diagonal will be nonnegative, as well as its determinant.

The eigenvalues of the Laplacian matrix is often used to solve abstract problems. For instance, it may not come as a shock that hitting the surface membrane of differently shaped drums will produce different sounds, seemingly at random. However, there is a way to determine the shape of the drum's surface just by listening to it. When the surface of the drum is mapped to a graph, the eigenvalues of the Laplacian matrix of the graph is directly related to the frequency of sound that the drum produces [17]. Thus, the eigenvalues of the Laplacian matrix of a graph of a shape uncovers certain geometric properties of that shape. We will revisit the relation between sound frequencies and the eigenvalues of a matrix later.

Now that we have introduced the eigenvalues of the Laplacian matrix, let us investigate a theorem that bounds the eigenvalues of a matrix to specific areas on a plane.

Theorem 5.6 (Gershgorin Disk Theorem)**.** [15] *Let A be a complex $n \times n$ matrix. In the complex plane, let B_i be a ball with a center at A_{ii} and a radius R_i that is the sum of the absolute values of all other entries in the $i-th$ row for $i = \{1, 2, \ldots, n\}$. Each eigenvalue of A is contained in some ball B_i.*

Example 5.7. To visualize the Gershgorin Disk Theorem, let us find the eigenvalues of the matrix from Example 1.14,

$$A = \begin{bmatrix} -6 & 3 \\ 4 & 5 \end{bmatrix}.$$

We proceed by finding the center and radii of the two circles. The center of the balls are located at $B_1 = (-6, 0i)$ and $B_2 = (5, 0i)$, and they have a radius of $R_1 = 3$ and $R_2 = 4$, respectively. Plotting the circles on the complex plane yields

Figure 16. In this figure, we can see that the eigenvalues $\lambda = 6$ and $\lambda = -7$ are clearly contained in the disks.

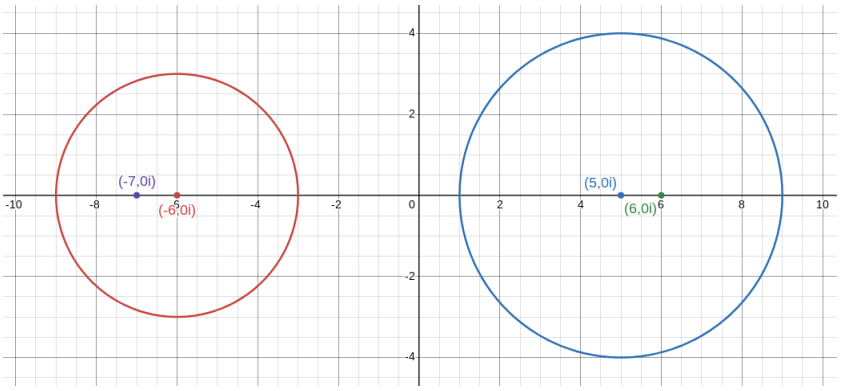
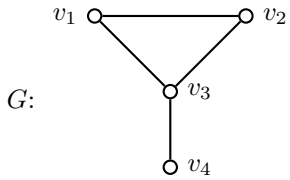


FIGURE 16. Gershgorin Disk Theorem

Thus, the Gershgorin Disk Theorem allows us to set a geometric boundary for what the eigenvalues of a matrix can be. This theorem, combined with the properties of these matrices, also brings a unique bound to the Laplacian matrix. The eigenvalues of the Laplacian matrix of a graph, apart from being bounded by the disks, are also bounded in the interval $[0, 2\Delta]$ where Δ is the maximum degree of G [23].

To visualize this, consider the graph and Laplacian matrix in Figure 17. Its eigenvalues are 0, 1, 3, 4. These eigenvalues are less than two times the maximum degree of G, 6.



G:

$$L = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 3 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

FIGURE 17. A graph and its Laplacian matrix

After examining some general bounds of the eigenvalues of the Laplacian matrix, let us look at some specific eigenvalues of the Laplacian matrix.

First, let $\lambda_i(L)$ be an eigenvalue λ of the Laplacian matrix L such that $\lambda_1(L) \leq \lambda_2(L) \leq \cdots \leq \lambda_n(L)$ where $n = |V(G)|$.

Lemma 5.8. *The smallest eigenvalue $\lambda_1(L)$ is 0.*

Proof. Since L is symmetric and positive semi-definite, we already know that its eigenvalues are real and non-negative. Consider the sum across the $i-$th row of L. This can be expanded as

$$\begin{aligned} \sum_{j=1}^{n} L_{ij} &= \deg(v_i) + \sum_{j \neq i} L_{ij} \\ &= \deg(v_i) + \deg(v_i)(-1) + 0 \\ &= 0. \end{aligned}$$

As such, the vector $\mathbb{1} = (1, \ldots, 1) \in \mathbb{R}^n$ satisfies $L\mathbb{1} = 0$. Thus, zero is an eigenvalue of L. □

Theorem 5.9. *The Laplacian matrix L has an eigenvalue 0 with multiplicity k if and only if the graph has k connected components.*

Proof. Suppose G has k connected components $G_1, G_2, \ldots, G_k$. There always exists an indexing of the vertices such that the first n_1 are associated to G_1 and so on, with the last n_k associated to G_k. Then the Laplacian matrix $L(G)$ precisely takes the form

$$L(G) = \begin{bmatrix} L(G_1) & & \\ & \ddots & \\ & & L(G_k) \end{bmatrix}.$$

This is called a *block diagonal matrix*, where we have square matrices of various sizes along a diagonal with no overlap and otherwise zero.

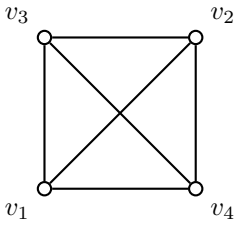
We can take for granted that $\det(L(G)) = \det(L(G_1)) \cdots \det(L(G_k))$. The left size is zero by the lemma. Similarly, each factor of the right side is zero as well. Thus, zero has a multiplicity $\geq k$. That the multiplicity is precisely k follows from constructing the zero-eigenvector for each component and showing that it is unique. □

The special properties of the eigenvalues of the Laplacian matrix also extend to specific types of graphs.

Theorem 5.10. [2] *The Laplacian matrix of a complete graph K_n has eigenvalue 0 with multiplicity 1 and eigenvalue n with multiplicity $n-1$.*

As an example, going back to the complete graph K_4, by Theorem 5.9, the 0 eigenvalue must have a multiplicity of 1 because the graph is connected. Shown in Figure 18 is the graph and the Laplacian matrix L of K_4. Solving for its eigenvalues yields 0, 4, 4, 4. The 4 eigenvalue has a multiplicity of 3, supporting the theorem.

K_4:

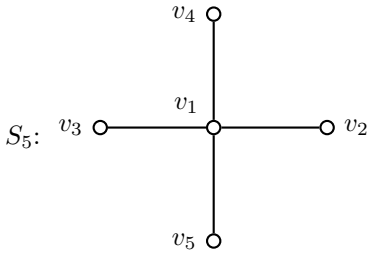


$$L = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}.$$

FIGURE 18. The Laplacian matrix of the complete graph K_4

Theorem 5.11. [2] *The Laplacian matrix of a star graph S_n has eigenvalue 0 with multiplicity 1, eigenvalue 1 with multiplicity $n-2$, and eigenvalue n with multiplicity 1.*

As an example, referring to the star graph S_5 shown in Figure 5, the connected graph must have a 0 eigenvalue with multiplicity 1. The corresponding Laplacian matrix of S_5 is shown in Figure 19. Similarly, solving for the eigenvalues yields 0, 1,



$$L = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

FIGURE 19. The Laplacian matrix of the star graph S_5

1, 1, 5. The eigenvalue 0 has multiplicity 1, the eigenvalue 1 has multiplicity 3, and the eigenvalue 5 has multiplicity 1.

Now that we have gained some familiarity with the eigenvalues of the Laplacian matrix, let us introduce a unique relationship between the eigenvalues of a matrix and the edge count of a graph.

Lemma 5.12. [21] *Let G be a graph. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the eigenvalues of A, the adjacency matrix. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of L, the Laplacian matrix. Show that*

$$2|E(G)| = \sum_i \alpha_i^2 = \sum_i \lambda_i.$$

Proof. First, let us show that $\sum_i \alpha_i^2 = 2|E(G)|$. Let A be the adjacency matrix of G with eigenvalues $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then, the eigenvalues of A^2 are $\alpha_1^2, \alpha_2^2, \ldots, \alpha_n^2$. Since the trace of a square matrix is equal to the sum of its eigenvalues,

$$\operatorname{tr}(A^2) = \sum_i \alpha_i^2.$$

By matrix multiplication, we can rewrite the trace as

$$\operatorname{tr}(A^2) = \sum_{i,j} A_{ij} A_{ji}.$$

Then, since A is a symmetric matrix, we have

$$\operatorname{tr}(A^2) = \sum_{i,j} (A_{ij})^2.$$

Furthermore, any element A_{ij} is either 0 or 1, so we can simplify to get

$$\operatorname{tr}(A^2) = \sum_{i,j} A_{ij}.$$

Finally, since the trace is also the sum of the degrees, by the degree sum formula, we know

$$\operatorname{tr}(A^2) = \sum_i \alpha_i^2 = 2|E(G)|.$$

Now, let us show that $\sum_i \lambda_i^2 = 2|E(G)|$. Since the trace $\operatorname{tr}(L)$ of the Laplacian matrix is also equal to the sum of its eigenvalues,

$$\operatorname{tr}(L) = \sum_i \lambda_i^2.$$

By definition, $L = D - A$, so it follows that

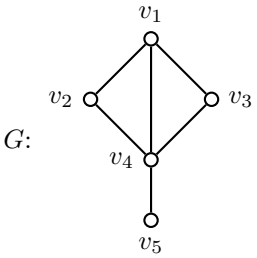
$$\operatorname{tr}(L) = \operatorname{tr}(D) - \operatorname{tr}(A) = \sum_i \deg(v_i) - 0 = 2|E(G)|.$$

Therefore,

$$\operatorname{tr}(L) = \sum_i \lambda_i^2 = 2|E(G)|.$$

Thus, $2|E(G)| = \sum_i \alpha_i^2 = \sum_i \lambda_i$ as desired. □

Example 5.13. As an example, consider the graph and its associated Laplacian matrix and adjacency matrix in Figure 20.



$$G: \quad L = \begin{bmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 & 0 \\ -1 & 0 & 2 & -1 & 0 \\ -1 & -1 & -1 & 4 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

FIGURE 20. The Laplacian and adjacency matrix of a graph

From this graph and its associated matrices, $2|E(G)| = 12$, and we calculate the eigenvalues of the adjacency matrix α_i and the eigenvalues of the Laplacian matrix λ_i to be

$$\{\alpha_i\} \approx \{-1.749, -1.271, 0, 0.335, 2.686\},$$
$$\{\lambda_i\} = \{0, 1, 2, 4, 5\}.$$

Moreover, the sum of all α_i and λ_i are

$$\sum_i \alpha_i^2 = (-1.749)^2 + \cdots + (2.686)^2 \approx 12,$$
$$\sum_i \lambda_i = 0 + 1 + 2 + 4 + 5 = 12.$$

This example demonstrates how $2|E(G)| = \sum_i \alpha_i^2 = \sum_i \lambda_i$.

5.3. **Spectral Partitioning.** The eigenvalues of the Laplacian matrix also reveal how connected a graph is. The *Fiedler Eigenvalue*, named after Miroslav Fiedler, is the second-smallest eigenvalue of the Laplacian matrix of a graph and determines the algebraic connectivity of a graph.

Lemma 5.14. [14] *The Fiedler eigenvalue of a graph G is greater than 0 if and only if G is a connected graph.*

Proof. Since the Laplacian matrix is positive semi-definite, all of its eigenvalues are nonnegative. Moreover, a connected graph must have a 0 eigenvalue with multiplicity 1. Thus, the second-smallest eigenvalue cannot be 0, nor can it be negative, so the second-smallest eigenvalue of the Laplacian matrix of a connected graph must be positive. □

Typically, a graph with a greater Fiedler eigenvalue will have more edges and can therefore be thought of as "more connected" [20]. As an example, the Fiedler eigenvalue of the complete graph K_{100} is 100, while the Fiedler eigenvalue of the cycle graph C_{100} is about 0.004, showing that the cycle graph contains a significantly lower number of edges in contrast to the complete graph, and is "less connected" in contrast to the complete graph.

Likewise, alongside the Fiedler eigenvalue, the *Fiedler eigenvector* is the eigenvector that corresponds to the Fiedler eigenvalue. The Fiedler eigenvector is especially useful in *spectral partitioning*, where the goal is to partition, or group, a graph into two subgraphs in such a way that the subgraphs have nearly the same number of vertices while also minimizing the number of edges between the two subgraphs [20].

The theory proposed by Fiedler for performing spectral partitioning is based on a simple idea.

Definition 5.15. [14][20] Let $G = (V, E)$ be a graph and order the eigenvalues λ_i of $L(G)$ from least to greatest such that $\lambda_1(L) \leq \lambda_2(L) \leq \cdots \leq \lambda_n(L)$ where $n = |V(G)|$. Let the eigenvector corresponding to λ_2 be the Fiedler eigenvector μ_2.

The process of *spectral partitioning* by Fiedler's method is as follows: Each vertex $v_i \in G$ is partitioned into G_1 if the entry in the $i-$th row of μ_2 is positive. Otherwise, it is partitioned into G_2. A *cut edge* is an edge that is incident to a vertex in G_1 and a vertex in G_2.

However, it is not clear why this specific eigenvalue would result in such a partition. Moreover, using the eigenvalues and eigenvectors of the matrix of a graph to partition the graph may be the least obvious thing to do.

To understand the logic behind the Fiedler eigenvalue, we start by visualizing a vibrating string. When the string is plucked to a specific frequency, it will cause the string to appear frozen. This is called a standing wave. As it turns out, the frequency that causes the standing wave is directly related to the eigenvalues and eigenvectors of the matrix that represents the motion of the string at specific points [20]. Coincidentally, the Fiedler eigenvector of the Laplacian matrix of a graph represents the standing wave.

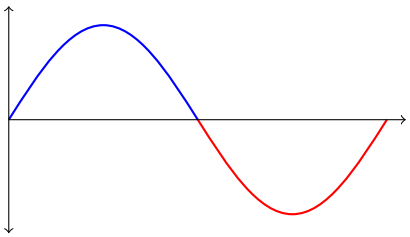


FIGURE 21. A string vibrating at its standing wave frequency

The section of the string that corresponds to the Fiedler eigenvalue and eigenvector is exactly one period, as shown in Figure 21. Each entry of the Fiedler eigenvector equates to one point on the string (usually equally spaced apart), and the heights of these points correspond to the values listed in the Fiedler eigenvector. This results in about half of the entries in the vector being positive (as indicated by the blue

string), while the other half of the entries are negative (as indicated by the red string).

Now that we have gained a solid understanding of Fiedler's method and why it works, let us examine a graph with a bow-tie shape shown in Figure 22 to determine how the partition should be made.

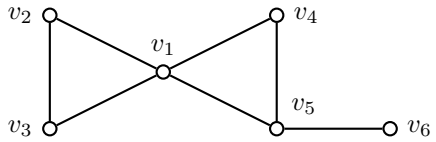


FIGURE 22. A graph shaped like a bow-tie

The bow-tie shape does not necessarily tell us anything concrete, but it does point us to which edges may be chosen to create the partitions. Continuing, calculating the Laplacian matrix of this graph yields

$$L = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 \\ -1 & 0 & 0 & 2 & -1 & 0 \\ -1 & 0 & -1 & 3 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 \end{bmatrix}.$$

The eigenvalues, in order, are

$$\lambda_i = (0, 0.6314, 1.4738, 3, 3.7877, 5.1071).$$

The Fiedler eigenvalue is $\lambda_2 = 0.6314$, and the corresponding Fiedler eigenvector is

$$\mu_2 = \begin{bmatrix} -0.16 \\ -0.44 \\ -0.44 \\ 0.07 \\ 0.26 \\ 0.71 \end{bmatrix}.$$

This tells us that v_1, v_2, v_3 are all partitioned into one group, while v_4, v_5, v_6 are all partitioned into another group, as seen in Figure 23. Note that the sum of the entries in the Fiedler vector is 0. This is true for the Fiedler vector of any graph [20]. As shown, this method of spectral partitioning aims to balance an approximately equal vertex count with the minimum possible number of cut edges.

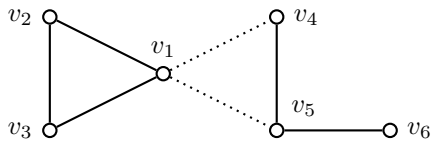


FIGURE 23. A graph partitioned by Fiedler's method

As another example, the partition of another graph by Fiedler's method is shown in Figure 24. Fiedler's method says that $v_1, v_2, v_3, v_4, v_5, v_6$ should be in one group, while v_7, v_8, v_9, v_{10} should be in the other. Note that having a partition with only one cut edge is also possible, but there would be an imbalance of the vertices of each partition.

Do note that another partition has a comparable result to Fiedler's method, with v_1, v_2, v_3, v_4, v_5 in one group and v_6, v_7, v_8, v_9 in the other. This shows how Fiedler's method may not be the only method to successfully partition a graph with these requirements.

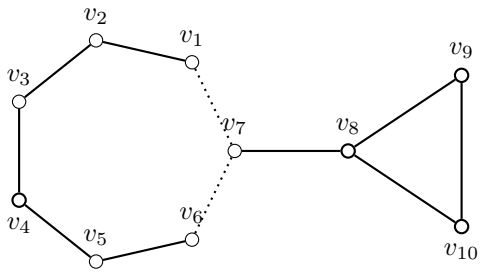


FIGURE 24. A graph and its partition

Another important result Fiedler proved was regarding the connectivity of the two subgraphs formed after spectral partitioning.

Theorem 5.16 (Fiedler's Theorem of Connectivity of Spectral Graph Partitions)**.** [20] *Let G be a connected graph. The two graphs G_1 and G_2 that are formed after the spectral partitioning of G are connected.*

Referring back to Figure 23 and Figure 24, we note that the initial graphs are connected, and so is every partition.

If we wanted to measure how well a graph can be partitioned, we refer to the *Cheeger constant*, also called the *isoperimetric number*.

Intuitively, a large positive Cheeger constant says that the graph is well-connected, and there is no relatively small set of edges that can successfully partition the graph. In other words, every attempt to partition the graph will result in a relatively large number of cut edges. On the other hand, a small positive Cheeger constant points to a graph that has a bottleneck − a relatively small set of edges whose removal separates a relatively large portion of vertices from the graph [27]. The Cheeger constant is positive if and only if the graph is connected [28].

Definition 5.17. [24] Let $G = (V, E)$ where $S \subset V$. Then, the *Cheeger constant* $h(G)$ is

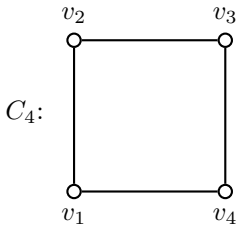
$$h(G) = \min_{0<|S|\leq|V|/2} \frac{|\partial S|}{|S|}$$

where $|\partial S|$ represents the number of edges with one endpoint in S and one endpoint in V that is not in S.

In other words, the Cheeger constant is formed by calculating all possible partitions and taking the minimum of the ratio between the vertices in each partition that form cut edges and those that do not. In summary, for a graph $G = (V, E)$, we

must first identify all subsets S of $V(G)$ such that $0 < |S| \leq \frac{|V|}{2}$. Then, for each subset S, we must calculate the ratio $\frac{|\partial S|}{|S|}$. Finally, we take the minimum ratio over all of these subsets to declare as the Cheeger constant of the graph.

To demonstrate this with a simple example, let us visualize the cycle graph C_4, shown in Figure 25. The complete list of possible subsets are listed in the figure, with its corresponding $|\partial S|$ and $|S|$, as well as its ratio. As a result, the Cheeger constant is the minimum of these ratios, which is 1.



S	$\lvert\partial S\rvert$	$\lvert S\rvert$	$\frac{\lvert\partial S\rvert}{\lvert S\rvert}$
$\{1\}$	2	1	2
$\{2\}$	2	1	2
$\{3\}$	2	1	2
$\{4\}$	2	1	2
$\{1,2\}$	2	2	1
$\{2,3\}$	2	2	1
$\{3,4\}$	2	2	1
$\{1,4\}$	2	2	1

FIGURE 25. The cycle graph C_4 and the calculation of its Cheeger constant

As a more concrete example, consider the bow tie graph from Figure 22. The Cheeger constant of this graph is $\frac{2}{3}$, achieved by the partition $G_1 = \{v_1, v_2, v_3\}$ and $G_2 = \{v_4, v_5, v_6\}$. This tells us that the bow tie graph can be partitioned in a better manner than the cycle graph C_4. Note that while the bow tie graph still has two cut edges, it has more vertices than the cycle graph, causing the Cheeger constant to be lower.

Furthermore, as a more extreme example, a graph with a relatively high Cheeger constant is the complete graph K_8 with $h(K_8) = 4$, while a graph with a relatively low Cheeger constant is the path graph P_8 with $h(P_8) = \frac{1}{4}$. This comparison presents us with a deeper understanding of the Cheeger constant.

The idea of spectral partitioning is useful in many real-life applications. As an example, spectral partitioning can help identify communities based on certain criteria [25]. Each vertex may represent one person, while the edges represent the friendship between two people. Moreover, in biology, spectral partitioning is also helpful in identifying genes that are functionally related within gene regulatory networks [26].

5.4. **Trees in Graphs.** A key type of graph in graph theory is a tree. Trees are an important type of graph because they represent the "least connected" graphs. The nomenclature is rather intuitive, as it represents two main ideas of trees seen in nature; trees are connected, and they are acyclic.

Definition 5.18. A *tree* is an acyclic connected graph.

A *bridge* of a connected graph G is an edge e in G such that $G - e$ is disconnected. Thus, bridges must not lie on any cycle of G, and since trees contain no cycles, all edges of trees are bridges [1].

There are also many properties that define what a tree can, and cannot, be.

Lemma 5.19. [1] *Every nontrivial tree has at least* 2 *end-vertices, or vertices with a degree of* 1.

Proof. Let T be a nontrivial tree containing a $u-v$ path $P = (u = u_0, u_1, \ldots, u_k = v)$ of the greatest possible length. Neither u nor v can be adjacent to any vertex not on P. Otherwise, it would create a new path of greater length. Therefore, $\deg u = \deg v = 1$. □

Lemma 5.20. *Every tree of order n has size $n - 1$.*

Proof. We proceed by induction. First, the *base case*. Every graph with 1 vertex must have 0 edges, so the base case is true.

Next, the *inductive step*. We assume a tree with n vertices has $n - 1$ edges. Now, let G be a tree with $n + 1$ vertices. Form the subgraph H by removing an end-vertex of G, consequently removing the incident edge from G as well. Now, H has $(n + 1) - 1 = n$ vertices. Since H is a tree, we know H has $n - 1$ edges by our previous assumption. Since G has exactly one more edge than H, the edge count of G must be $(n - 1) + 1 = n$. Therefore, G has $n + 1$ vertices and n edges, as claimed. □

Since a tree is considered the "least connected" of all connected graphs, we may come to the conclusion that the size of every connected graph of order n is at least $n - 1$.

Theorem 5.21. *The size of every connected graph of order n is at least $n - 1$.*

Proof. In contrast, assume that there exists a connected graph of order n with size $m = n - 2$. Then $n \geq 4$, so G contains no isolated vertices.

We claim G contains an end-vertex, so assume to the contrary that the degree of every vertex of G is at least 2. Then, the sum of the degrees of the vertices of G is $2m \geq 2n$, so $m \geq n \geq m + 2$, which is impossible. Thus, G must contain an end-vertex. □

The intuitive nomenclature doesn't just stop at trees. A *forest* is a graph that contains trees.

Corollary 5.22. [1] *A forest of order n with k components has size $n - k$.*

Proof. Let a forest F have size m with $G_1, G_2, \ldots, G_k$ be the connected components of F. Let G_i have order n_i and size m_i. Then, n is the sum of all n_i and m is the sum of all m_i. Since each component G_i is a tree, we know $m_i = n_i - 1$. Thus,

$$m = \sum_i m_i = \sum_i (n_i - 1) = n - k.$$

□

Now, we can construct a theorem that classifies graphs into the tree category based on certain conditions.

Theorem 5.23. [1] *Let G be a graph of order n and size m. If G satisfies any two of the properties:*

(1) *G is connected,* (2) *G is acyclic,* (3) $m = n - 1$,

then G is a tree.

Proof. By casework, we first assume G satisfies (1) and (2). Then, by definition, G is a tree. Thus, we now consider the cases where G satisfies (1) and (3) or G satisfies (2) and (3).

Case 1. G satisfies (1) and (3). Since G is connected, we only have to show that G is acyclic. Assume the opposite, that G contains a cycle. Then e is not a bridge of G. Thus, $G - e$ is a connected graph of order n and size $n - 2$, contradicting Lemma 5.20. Therefore, G must be acyclic, and G is a tree.

Case 2. G satisfies (2) and (3). Since G is acyclic, we only have to show that G is connected. Since G satisfies (2) and (3), it follows that G is a forest of order n and size $m = n - 1$. The size of G is $n - k$ for k components in G. Hence $n - 1 = n - k$ so $k = 1$. Therefore, G is connected, and G is a tree. □

A type of subgraph of tress is spanning trees, which takes on a similar definition to the subgraph of a tree, with one key limitation.

Definition 5.24. A *spanning tree* of a graph $G = (V, E)$ is a tree $T = (V, E)$ such that $V(T) = V(G)$ and $E(T) \subseteq E(G)$. In other words, the spanning tree contains the same vertex set and a subset of the edge set.

As an example, there are a total of 3 spanning trees of the complete graph K_3, as shown in Figure 26.

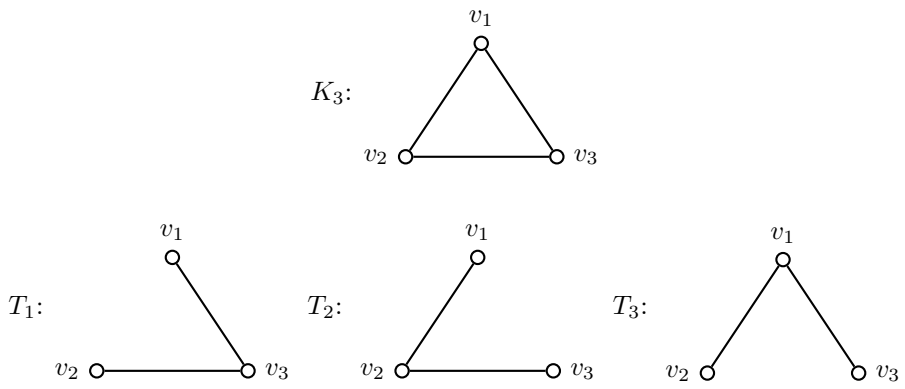


FIGURE 26. The spanning trees of K_3

5.5. **Counting Spanning Trees.** How can we count the total number of spanning trees of a graph? Certainly, if a graph is connected, then it must have at least one spanning tree, but to continue counting them manually would prove to be a laborious and unenlightening task. Thankfully, *Kirchhoff's matrix tree theorem* provides an elegant way to compute the number of spanning trees of a graph.

Kirchhoff's matrix tree theorem provides a formula to find the total number of spanning trees of a graph using the determinant of the Laplacian matrix of a graph. It is a beautiful result in graph theory connecting linear algebra and combinatorics.

Theorem 5.25 (Kirchhoff's Matrix Tree Theorem)**.** *The number of spanning trees $T(G)$ in a graph G is $det(L_0)$ where L_0 is the Laplacian matrix with the $i-$th row and $i-$th column removed for any i.*

The term $\det(L_0)$ is also known as the *cofactor* of an element i of the Laplacian matrix L.

Definition 5.26. The contraction G/e of an edge e from a graph G removes e and merges the two vertices incident to e to become one vertex.

Sketch. [12][6] First, choose any edge e of the graph G. Then, we consider the two cases of the spanning trees $T(G)$ that either contains e as an edge, or does not.

Case 1. If $T(G)$ includes e, then contracting e will not remove or change the tree, so this case can be expressed as $T(G/e)$.

Case 2. If $T(G)$ does not include e, then we can remove e without removing or changing the tree, so this case can be expressed as $T(G-e)$.

Thus, we have the equation:

$$T(G) = T(G-e) + T(G/e).$$

This is the deletion-contraction property. Likewise, $\det(L_0)$ also satisfies this property:

$$\det(L(G)_0) = \det(L(G-e)_0) + \det(L(G/e)_0).$$

By the definition $L = D - A$, each row of L must sum to 0. As a result, the expansion of $\det(L(G)_0)$ will contain only the terms corresponding to edges that are connected and not part of a cycle. Any terms that correspond to cycles on the graph will be canceled out. Thus, we are left with the number of spanning trees. □

Remark 5.27. [13] The theorem presented above only deals with undirected graphs. However, Kirchhoff's theorem also generalizes to directed graphs. Oriented spanning trees of a directed graph would be counted in this case, called spanning arborescences. In this case, $\det L_0$ is equal to the number of arborescences that end at v_i.

We illustrate the Matrix Tree Theorem with a simple example. Figure 27 shows the graph G along with its three spanning trees, as well as its Laplacian matrix. By removing row 3 and column 3 from L to form L_0, we then take the determinant to obtain

$$\begin{aligned}\det\begin{bmatrix}2 & -1 & 0\\ -1 & 2 & 0\\ 0 & 0 & 1\end{bmatrix} &= 0\det\begin{bmatrix}-1 & 0\\ 2 & 0\end{bmatrix} - 0\det\begin{bmatrix}2 & 0\\ -1 & 0\end{bmatrix} + 1\det\begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix}\\ &= 0\cdot 0 - 0\cdot 0 + 1\cdot 3\\ &= 3.\end{aligned}$$

Cayley's tree formula follows from Kirchhoff's theorem as a special case that deals with the complete graph K_n.

Corollary 5.28 (Cayley's Tree Formula)**.** [1] *The number of spanning trees in a complete graph K_n with n vertices is n^{n-2}.*

Cayley's tree formula says that the number of spanning trees of K_3 is $3^{3-2} = 3$, which is confirmed by Figure 26. In general, applying Kirchhoff's matrix tree theorem to the complete graph K_n yields n^{n-2} spanning trees, which matches Cayley's tree formula for the number of spanning trees.

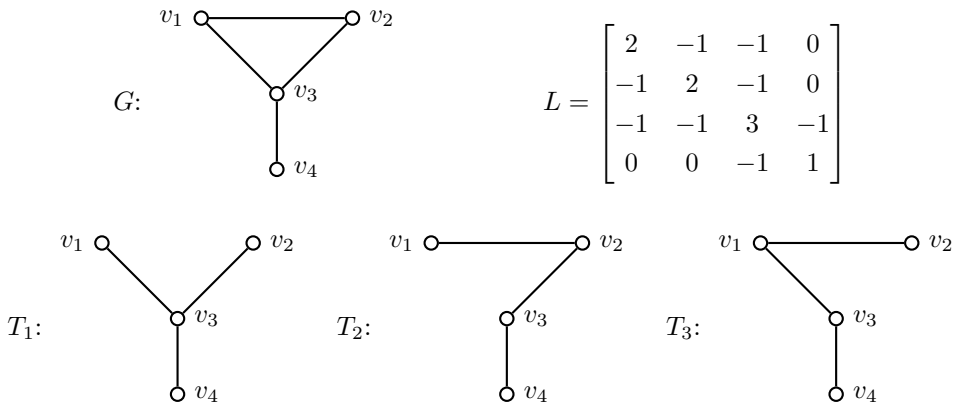


FIGURE 27. A graph, its Laplacian matrix, and its spanning trees

References

[1] G. Chartrand, P. Zhang, *A First Course in Graph Theory*. Dover Publications (2012).

[2] J. Jiang, *An Introduction to Spectral Graph Theory*. University of Chicago (2012).

[3] D. Spielman, *Spectral Graph Theory*. Chapman and Hall (2011).

[4] D. Johnson, *Graph Theory and Linear Algebra*. The University of Utah (2017).

[5] F. Harary, *Independent Discoveries in Graph Theory*. Blackwell Publishing (1979).

[6] R. Stanley, *The Matrix-Tree Theorem*. MIT OpenCourseWare (2014).

[7] D. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs*. Academic Press (1980).

[8] S. Khan, *Laplacian Intuition*. Khan Academy (2016).

[9] S. Khan, *Explicit Laplacian Formula*. Khan Academy (2016).

[10] S. Khan, *Multiplying Matrices*. Khan Academy (2014).

[11] H. Li, *Properties and Applications of Graph Laplacians*. University of Chicago (2022).

[12] O. Dmytruk, *Kirchhoff's Matrix-Tree Theorem*. Nagoya University (2024).

[13] J. Li, *Arborescences: Directed Spanning Trees*. Carnegie Mellon University (2024).

[14] J. Wyss-Gallifent, *Graph Theory*. University of Maryland (2021).

[15] S. Brakken-Thal, *Gershgorin's Theorem for Estimating Eigenvalues*. University of Puget Sound (2007).

[16] X. D. Zhang, *The Laplacian Eigenvalues of Graphs: a Survey*. Shanghai Jiao Tong University (2011).

[17] M. Kac, *Can One Hear The Shape of a Drum?* The Rockefeller University (2012).

[18] C. Berge, *Graphs and Hypergraphs*. University of Paris (1973).

[19] N. Biggs, E. Lloyd, R. Wilson, *Graph Theory 1736-1936*. Clarendon Press (1976).

[20] B. Slininger, *Fiedler's Theory of Spectral Graph Partitioning*. University of California (2013).

[21] P. Rombach, *Spectral Graph Theory*. University of Vermont (2022).

[22] V. Reiner, *Introduction to Combinatorics and Graph Theory*. University of Minnesota (2010).

[23] A. Banerjee, A. Char, B. Mondal, *Spectra of General Hypergraphs*. Indian Institute of Science Education and Research (2017).

[24] K. Munagala, *Graph Laplacian and Cheeger's Inequality*. Duke Trinity College of Arts & Sciences (2017).

[25] M. E. J. Newman, *Spectral Methods for Network Community Detection and Graph Partitioning*. University of Michigan (2017).

[26] S. Roy, S. Feizi, *Regulatory Networks: Inference, Analysis, Application*. Massachusetts Institute of Technology (2012).

[27] P. Rombach, *Cheeger Constant and Conductance*. University of Vermont (2022).

[28] R. Balan, *The Cheeger Constant and the Spectral Gap*. University of Maryland (2021).

MIT, 77 Massachusetts Avenue, Cambridge, MA 02139

Malden High School, 77 Salem Street, Malden, MA 02148
Email address: bpan01217@gmail.com

Boston Latin School, 78 Avenue Louis Pasteur, Boston, MA 02115
Email address: nathannnie395@gmail.com

Wellesley High School, 50 Rice Street, Wellesley, MA 02481
Email address: danielninja222@gmail.com