A New Lower Bound on the Chromatic Number of a Graph

Eunsuk Jung

Abstract

In this MTH 501 project, we present the results of the article "Proof of a conjectured lower bound on the chromatic number of a graph" by Tsuyoshi Ando and Minghua Lin [1]. Their work states and proves a conjectured lower bound for the chromatic number of a graph G based on the eigenvalues of the adjacency matrix of G and the properties of the Frobenius norm. The proof techniques reveal the close interplay between graph theory, linear algebra, and matrix analysis.

A Math 501 Project
Under the direction of

Dr. John S. Caughman
Second Reader
Dr. Jong-Sung Kim

Submitted in partial fulfillment of the requirements for the degree of

Master of Science in Mathematics

Portland State University
February 28, 2020

Contents

1 Introduction

We let G be a simple graph with $n \geq 1$ vertices and A_G be its adjacency matrix. The chromatic number $\chi(G)$ is the minimum number of colors needed to label the vertices so that adjacent vertices receive different colors. Since each color class must be an independent set, $\chi(G) \leq r$ if and only if G is r-partite (Recall that a graph is said to be *r-partite* whenever the vertex set can be expressed as the union of r independent sets).

It follows that if $\chi(G) = r$, we can partition the adjacency matrix A_G into $r \times r$ blocks such that $A_G{=}[A_{ij}]_{i,j=1}^{r}$ with

$$A_{ii} = 0 \quad \text{for all} \quad 1 \leq i \leq r.$$

Note that the diagonal blocks A_{ii} for all $1 \leq i \leq r$ are not only zero matrices, but they are also square matrices by construction. Since the adjacency matrix A_G is real and symmetric, it must have real eigenvalues by the spectral theorem. Accordingly, we let $\mu_1, \mu_2, \ldots, \mu_n$ be the eigenvalues of A_G, sorted in non-increasing order.

The Hoffman lower bound [2] on the chromatic number is one of the best known results in spectral graph theory. It states that

$$\chi(G) \geq 1 + \frac{\mu_1}{-\mu_n}, \tag{1}$$

where μ_1, μ_2 are the maximum and minimum eigenvalues, respectively.

Over the years, many studies have found lower bounds for $\chi(G)$. In 2013, Wocjan and Elphick [5] found a significant improvement over the existing lower bounds for the chromatic number $\chi(G)$. Specifically, they proved that

$$\chi(G) \geq 1 + \max_{m} \frac{\sum_{i=1}^{m} \mu_i}{-\sum_{i=1}^{m} \mu_{n-i+1}} \quad \text{for} \quad m = 1, \ldots, n-1. \tag{2}$$

Their results generalize the Hoffman lower bound (2) given above, which occurs as the special case of (2) when $m{=}1$. They also conjectured the lower bound

$$\chi(G) \geq 1 + \frac{s^+}{s^-} \tag{3}$$

where s^+ and s^- are the sums of the squares of positive and negative eigenvalues, respectively. To corroborate this conjecture, they looked for counterexamples, but none were found. They left this as an open question, but the result was proved in [1], which is the subject of this 501 project. Specifically, this project will prove the conjecture, stated as Theorem 1.1 below, and will showcase the close interplay between graph theory, linear algebra, and matrix analysis.

Theorem 1 *Let A_G denote the adjacency matrix of a graph G, and let π, ν and δ denote the numbers of positive, negative and zero eigenvalues of A_G, respectively. Let*

$$s^+ = \mu_1^2 + \mu_2^2 + \cdots + \mu_\pi^2, \quad \textit{and} \quad s^- = \mu_{n-\nu+1}^2 + \mu_{n-\nu+2}^2 + \cdots + \mu_n^2. \tag{4}$$

Then

$$\chi(G) \geq 1 + \frac{s^+}{s^-}. \tag{5}$$

This spectral approach to studying the structure of graphs is somewhat limited by the fact that there exist non-isomorphic graphs that are co-spectral, meaning that the eigenvalues of the adjacency matrices of the graphs are the same. This shows that the spectrum of a graph can only provide partial information when we study the structure of the graphs. Nevertheless, spectral methods can be very effective in cases when graphs have special properties, such as symmetries. Spectral graph theory has found applications in chemistry, network design, coding theory, and computer science. It also was useful to Larry Page of Google, whose patented PageRank algorithm uses the Perron-Frobenius eigenvector of the graph of the world-wide web.

2 Preliminaries

We begin with some basic definitions from graph theory and linear algebra, then prove a number of inequalities involving the Frobenius norm. This norm is central to our main result giving the new lower bound for the chromatic number. Finally, we will apply these results to the adjacency matrix to obtain the theorem.

2.1 Graphs

We begin with a few basic terms regarding graphs.

A **graph** G consists of a **vertex set** $V(G)$, an **edge set** $E(G)$, and a relation that associates with each edge two vertices called its **endpoints**. A graph is **simple** if it has no loops or multiple edges. We specify a simple graph by its vertex set and its edge set, and we will be concerned only with finite undirected graphs. When u and v are the endpoints of an edge, they are said to be **adjacent**, written $u \sim v$. In this case, we refer to the shared edge as uv. For more detail on these concepts, see [4].

2.2 Chromatic number

We now discuss the chromatic number of a graph.

An ***r*-coloring** of a graph G is a labeling $f : V(G) \to S$, where $|S| = r$. We refer to the labels as colors, and the vertices of the same color form a **color class**. In a **proper** r-coloring, adjacent vertices must have different colors, so the color classes must be independent sets. If a graph requires at least r colors to be properly colored, we refer to r as the **chromatic number** $\chi(G)$. In other words, the chromatic number is the minimum number of colors needed to label the vertices so that adjacent vertices receive different colors. A graph G is ***r*-partite** if $V(G)$ can be expressed as the union of r independent sets.

The **clique number** of a graph G, written as $\omega(G)$, is the maximum size of a set of pairwise adjacent vertices in G. The **independence number**, written as $\alpha(G)$, is the maximum size of a set of pairwise non-adjacent vertices in G. For every graph G, we have $\chi(G) \geq \omega(G)$ and $\chi(G) \geq \frac{n(G)}{\alpha(G)}$. The first bound holds because vertices of a clique require distinct colors. The second bound holds because each color class is an independent set and has at most $\alpha(G)$ vertices.

As an example, **the Petersen graph** is an undirected graph with 10 vertices and 15 edges (see Figure 1). The Petersen graph has chromatic number 3, meaning that its vertices can be properly colored with three colors — but not with two.

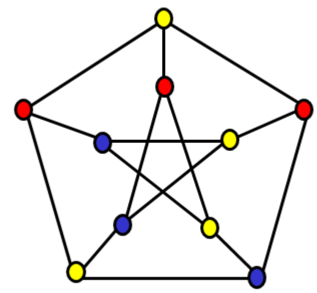


Figure 1: Proper 3-coloring of the Petersen graph

2.3 Adjacency matrix and conformal partitions

When we specify a graph, we typically list the vertices and edges. Let G be a loopless graph with vertex set $V(G)=\{v_1,\ldots,v_n\}$ and edge set $E(G)=\{e_1,\ldots,e_m\}$. The **adjacency matrix** of G, written as $A(G)$, is the $n \times n$ matrix in which entry $a_{i,j}$ is 0 or 1, depending on whether there is an edge in G with endpoints $\{v_i, v_j\}$. In other words, adjacency matrix A is defined by

$$A_{ij} = \begin{cases} 1 & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases} .$$

Note that an adjacency matrix depends upon the given vertex ordering to index the rows and columns. Every adjacency matrix of a simple graph is **symmetric** (i.e. $a_{ij}=a_{ji}$) and has entries 0 or 1, with only zeros on the diagonal. The **degree** of a vertex v is the sum of the entries in the row for v in $A(G)$.

Returning to our example, let A denote the adjacency matrix of the Petersen graph. Then if we order the vertices clockwise, first around the outside, then around the inside, we obtain the following adjacency matrix:

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

We close this section with one more definition. We say two block matrices A and B are **conformal** for multiplication when their block sizes are suitable for block multiplication. In this case we find that the product of an $m \times n$ block matrix with an $n \times p$ block matrix results in an $m \times p$ block matrix, so that

$$A_{m\times n} \times B_{n\times p} = C_{m\times p},$$

where $C_{i,j} = \sum_{k=1}^{n} A_{i,k} \times B_{k,j}$.

For example, if A is 2×3 matrix and B is 3×4 matrix, then they are conformal for multiplication and we can get 2×4 matrix C.

2.4 Eigenvalues and Eigenvectors

The **eigenvalues** of a matrix A are the numbers μ such that

$$Ax = \mu x$$

has a nonzero solution vector x. The vector x is called an **eigenvector** associated with the eigenvalue μ. The **eigenvalues** of a graph are the eigenvalues of its adjacency matrix A. These are roots $\mu_1, \ldots, \mu_n$ of the **characteristic polynomial**

$$\phi(G;\mu) = \det(\mu I - A) = \prod_{i=1}^{n} (\mu - \mu_i). \tag{6}$$

The **spectrum** is the list of distinct eigenvalues with their multiplicities $m_1, \ldots, m_t$. We can write

$$\mathrm{Spec}(G) = \begin{pmatrix} \mu_1 & \ldots & \mu_t \\ m_1 & \ldots & m_t \end{pmatrix}.$$

For example, since the characteristic polynomial of the Petersen graph P is

$$(x-3)(x-1)^5(x+2)^4,$$

the spectrum of P is

$$\mathrm{Spec}(P) = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 5 & 4 \end{pmatrix}.$$

The **trace** of a matrix is the sum of the diagonal elements or the sum of the eigenvalues. So the trace of a simple graph is 0. Indeed, the trace of Petersen graph above is 0 since the sum of the diagonal entries (and the sum of eigenvalues) equals 0.

2.5 Spectral Theorem

Relating the eigenvalues to other graph parameters will require several results from linear algebra, including the spectral theorem for real symmetric matrices.

Theorem 2 *Let A be any real symmetric $n \times n$ matrix. Then*

1. *A has n real eigenvalues $\mu_1, \ldots, \mu_n$ (not necessarily distinct).*
2. *There is a set of n corresponding eigenvectors $v_1, \ldots, v_n$ that constitute an orthonormal basis of $\mathbb{R}^n$, that is, $v_i^T v_j = \delta_{ij}$ for all i, j.*

Proof. 1. Assume A is a real symmetric $n \times n$ matrix. Since the characteristic polynomial has degree n, we know A has n complex eigenvalues, counting multiplicity. To see that these are real, suppose $Ax = \mu x$ with $x \neq 0$ and $\mu \in \mathbb{C}$. Then

$$\begin{aligned} \mu \bar{x}^T x &= \bar{x}^T(\mu x) \\ &= \bar{x}^T(Ax) \\ &= (\bar{x}^T \bar{A}^T)x \\ &= \overline{(Ax)}^T x \\ &= \bar{\mu}\bar{x}^T x. \end{aligned}$$

Because $x \neq 0$, we know that $\bar{x}^T x \neq 0$. It follows that $\mu = \bar{\mu}$ and $\mu \in \mathbb{R}$.

2. This fundamental result is found in many introductory linear algebra texts. For example, see [3, p.104]. □

The next corollary recasts the spectral theorem in terms of a matrix decomposition.

Corollary 1 *Let A be a real symmetric $n \times n$ matrix with eigenvalues $\mu_1, \ldots, \mu_n$, Let M denote the diagonal matrix of eigenvalues and let V denote the matrix whose columns are the corresponding orthonormal eigenvectors $v_1, \ldots, v_n$. Then*

1. *The matrix A can be written as $A = VMV^T$, so that*

$$A = \begin{pmatrix} v_1 & v_2 & \ldots & v_n \end{pmatrix} \begin{pmatrix} \mu_1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & \mu_n \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{pmatrix}$$

2. *The matrix A satisfies*

$$\begin{aligned} A &= \mu_1 v_1 v_1^T + \mu_2 v_2 v_2^T + \cdots + \mu_n v_n v_n^T \\ &= \sum_{i=1}^{n} \mu_i v_i v_i^T. \end{aligned}$$

Proof. 1. Since the v_i's are orthonormal, we have V is invertible with $V^{-1} = V^T$ and $VV^T = I$. For any i, if e_i denotes the standard basis vector, then

$$V^T AVe_i = V^T Av_i = V^T \mu_i v_i = \mu_i V^T v_i = \mu_i e_i = Me_i.$$

Thus $M = V^T AV$, which implies $A = VMV^T$.

2. For any j,

$$\left(\sum_{i=1}^{n} \mu_i v_i v_i^T \right) v_j = \mu_j v_j = Av_j.$$

Since the v_i form a basis, it follows that $A = \sum_{i=1}^{n} \mu_i v_i v_i^T$. □

2.6 Frobenius Norm

For an $m \times n$ matrix A, the **Frobenius norm** $\|A\|_F$ is a matrix norm defined as the square root of the sum of the absolute squares of its elements, so that

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}. \tag{7}$$

The **conjugate transpose** of an $m \times n$ matrix is obtained by interchanging the rows and columns and then taking the complex conjugate of each entry. The Frobenius norm is also equal to the square root of the matrix trace of $\overline{A^T}A$ where $\overline{A^T}$ is the conjugate transpose, i.e.,

$$\|A\|_F = \sqrt{\mathrm{tr}(\overline{\mathrm{A}^{\mathrm{T}}}\mathrm{A})}. \tag{8}$$

If $\mathbb{F} = \mathbb{R}$, then the conjugate transpose of a matrix is the same as its transpose, which is the matrix obtained by interchanging the rows and columns.

Example. The conjugate transpose of the matrix

$$\begin{pmatrix} 2 & 3+4i & 7 \\ 6 & 5 & 8i \end{pmatrix}$$

is the matrix

$$\begin{pmatrix} 2 & 6 \\ 3-4i & 5 \\ 7 & -8i \end{pmatrix}.$$

A **Hermitian matrix** is a complex square matrix that is equal to its own conjugate transpose. In other words, the element in the i-th row and j-th column is equal to the complex conjugate of the element in the j-th row and i-th column, for all indices i and j. So if A_{ij} denotes the ij-entry of a matrix A, then

$$A \text{ is Hermitian} \quad \Leftrightarrow \quad A_{ij} = \overline{A_{ji}}.$$

Writing this in matrix form gives us:

$$A \text{ is Hermitian} \quad \Leftrightarrow \quad A = \overline{A^T}.$$

Proposition 1 *The entries on the main diagonal of any Hermitian matrix are real.*

Proof. Let A be a Hermitian matrix. Then by definition,

$$A_{ij} = \overline{A_{ji}}.$$

When i=j, the result follows. □

Only the main diagonal entries are necessarily real. Hermitian matrices can have arbitrary complex-valued entries in their off-diagonal elements, as long as diagonally-opposite entries are complex conjugates.

Proposition 2 *A matrix that has only real entries is Hermitian if and only if it is symmetric.*

Proof. Assume that A has only real entries. By definition, A is Hermitian if and only if

$$A_{ij} = \overline{A_{ji}}$$

for all i and j. Since A_{ij} is real, $A_{ij}=\overline{A_{ij}}$. So A is Hermitian if and only if it is symmetric. $\square$

We now explore some properties of Hermitian matrices and the Frobenius norm.

Proposition 3 *The sum of any two Hermitian matrices is Hermitian.*

Proof. Assume that A and B are Hermitian. Then

$$(A+B)_{ij} = A_{ij} + B_{ij} = \overline{A_{ji}} + \overline{B_{ji}} = \overline{(A+B)_{ji}}.$$

It follows that the matrix $A+B$ is also Hermitian. $\square$

Theorem 3 *A matrix norm is a function $\|\cdot\|$ from the set of all real (or complex) matrices of finite size into $\mathbb{R}^{\geq 0}$ that satisfies*

1. $\|A\| \geq 0$ *and* $\|A\| = 0$ *iff* $A = O$ *(a matrix of all zeros).*
2. $\|\alpha A\| = \|\alpha\|\|A\|$ *for all* $\alpha \in \mathbb{R}$.
3. $\|A+B\| \leq \|A\| + \|B\|$

   Properties (4)-(6) are additional properties of the Frobenius norm.

4. $\|A\| = \|A^T\|$
5. $\|AA^T\| = \|A\|^2 = \|A^T A\|$
6. $\|AB\| \leq \|A\|\|B\|$

Proof. These are standard derivations that can be found in most textbooks on linear algebra. For example, see [3]. $\square$

2.7 Positive Semidefinite Matrices

A real symmetric $n \times n$ matrix A is said to be **positive semidefinite** whenever

$$v^T A v \geq 0$$

for all $v \in \mathbb{R}^n$. Positive semidefinite matrices are also easily characterized in terms of eigenvalues.

Theorem 4 *Let A be a real symmetric $n \times n$ matrix. Then A is positive semidefinite iff all its eigenvalues $\mu_i \geq 0$ (i.e. all eigenvalues are non-negative).*

Proof. ($\Rightarrow$) Suppose A is positive semidefinite. Let $\mu_1, \ldots, \mu_n$ be the eigenvalues of A with corresponding basis of orthonormal eigenvectors $v_1, \ldots, v_n$. Then, by definition, $v^T Av \geq 0$ for all $v \in \mathbb{R}^n$. So for all i, it must be the case that $\mu_i = v_i^T Av_i \geq 0$.

($\Leftarrow$) Suppose that all of the eigenvalues of A satisfy $\mu_i \geq 0$. Then for any $v \in \mathbb{R}^n$, we have

$$v^T Av = v^T \left(\sum_{i=1}^{n} \mu_i v_i v_i^T \right) v = \sum_{i=1}^{n} \mu_i \left(v \cdot v\right)^2 \geq 0.$$

So A is positive semidefinite. $\square$

Example 1 *The $n \times n$ identity matrix is positive semidefinite. All the eigenvalues are 1 and every vector is an eigenvector. It is the only 3×3 symmetric matrix with all eigenvalues 1.*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Example 2 *The following matrix is positive semidefinite:*

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

To see why, note that if $x = (x_1\, x_2)^T$ is any vector, then $x^T Mx = (x_1 - x_2)^2 \geq 0$.

Note that, although the sum of two positive semidefinite matrices is positive semidefinite, the product of two positive semidefinite matrices need not be semidefinite. The next result considers the direct sum.

Proposition 4 *The direct sum matrix $A \oplus B$, where*

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

is positive semidefinite iff A and B both are positive semidefinite.

Proof. To see why, first note that $A \oplus B$ is symmetric iff A and B are symmetrix. Now, consider the quadratic for $x^T(A \oplus B)x$. The vector x can be divided into x_1 and x_2, and

$$x^T(A \oplus B)x = x_1^T Ax_1 + x_2^T Bx_2.$$

The result follows. $\square$

We conclude this introduction with a few key properties about positive definite matrices that will be used in our proof of the main result. Their derivations can be found in many textbooks on linear algebra, for example [3].

The first of these tells us that when a matrix is positive semidefinite, its diagonal should dominate the non-diagonal elements.

Proposition 5 *If A is positive semidefinite, then each off-diagonal element of A is less than or equal to the diagonal element in its row or in its column.*

Proof. Assume A is positive semidefinite. The quadratic form of A is

$$x^T A x = \sum_{i,j} A_{i,j} x_i x_j, \tag{9}$$

where x_i denotes the respective components of x. Since A is positive semidefinite then (9) is non-negative for every x. By choosing x to be a standard basis vector e_i, we get $A_{ii} \geq 0$. Let x have only two nonzero entries, say $x = e_i + e_j$. Then,

$$A_{i,j} \leq \frac{A_{ii} + A_{jj}}{2}.$$

This shows that any off-diagonal element is less than or equal to the diagonal element in its row or in its column. $\square$

For a **block matrix**, the diagonal blocks are going to be assumed to be square in this paper. For example, consider a block matrix of the form

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where A, B, C and D are themselves matrices. In this example, let

$$A = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix}, \; C = \begin{pmatrix} 4 & 4 \\ 4 & 4 \\ 4 & 4 \end{pmatrix} \text{ and } D = \begin{pmatrix} 5 & 0 & 5 \\ 0 & 5 & 0 \\ 5 & 0 & 5 \end{pmatrix}.$$

Therefore, the matrix is

$$M = \left(\begin{array}{cc|ccc} 0 & 2 & 3 & 3 & 3 \\ 2 & 0 & 3 & 3 & 3 \\ \hline 4 & 4 & 5 & 0 & 5 \\ 4 & 4 & 0 & 5 & 0 \\ 4 & 4 & 5 & 0 & 5 \end{array}\right).$$

Theorem 5 *(Schur's complement) Let M be any symmetric 2×2 block matrix:*

$$M = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}.$$

Then M is positive semidefinite iff D and the Schur complement $A - BD^{-1}B^T$ are both positive semidefinite.

Proof. This is a standard derivation that can be found in most textbooks on linear algebra. For example, see [3]. $\square$

Finally, we include the following important inequality regarding the Frobenius norm of a 2×2 block matrix.

Theorem 6 *[3, p.209] Let M be any symmetric 2×2 block matrix:*

$$M = \begin{pmatrix} A & B \\ B^T & D \end{pmatrix}.$$

If M is positive semidefinite, then the Frobenius norm of the blocks of M satisfy

$$\|B\|^2 \leq \|A\|\|D\|.$$

Proof. The proof of this important result can be found in [3, p.209]. $\square$

3 Main Result

In this section we present the proof of the main result of the paper by Tsuyoshi Ando and Minghua Lin [1].

For any simple graph G, the adjacency matrix A_G is real symmetric, so the eigenvlaues of A_G are real numbers, which we denote by $\mu_1, \ldots, \mu_n$ such that $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ (in non-increasing order). Let the π, ν, and δ be the numbers of positive, negative and zero eigenvalues of A_G respectively. Let

$$s^+ = \mu_1^2 + \cdots + \mu_\pi^2, \tag{10}$$

and

$$s^- = \mu_{n-\nu+1}^2 + \cdots + \mu_n^2. \tag{11}$$

We aim to prove the following theorem.

Theorem 1 *Let s^+, s^- be defined as above. Then,*

$$\chi(G) \geq 1 + \frac{s^+}{s^-}. \tag{12}$$

To begin, we recall the definition of the Frobenius norm. For any matrix X, we have

$$\|X\|_F = \sqrt{\mathrm{tr(X^*X)}} \tag{13}$$

If X is Hermitian, so that X=X^* (where X^* is a conjugate transpose of X), then $\|X\|_F^2$ is equal to the sum of the squares of all the eigenvalues of X.

Example 3 *Let X denote the following Hermitian matrix:*

$$X = \begin{pmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{pmatrix}.$$

Then, the trace of X^*X is the sum of squares of all diagonal entries. In particular,

$$\|X\|_F = \sqrt{1 + (1-i)^2 + 4 + (1+i)^2 + 9 + i^2 + 4 + (-i)^2 + 0} = \sqrt{16} = 4.$$

Since the sum of the n eigenvalues of X is the same as the trace of X, the sum of eigenvalues is also 4. This is confirmed, since $1 + 3 + 0 = 4$.

Our next result relates the norm of such a matrix to the norms of its blocks.

Lemma 1 *Let* $X{=}[X_{ij}]_{i,j=1}^r$ *be a positive semidefinite matrix that is given in* $r \times r$ *block form. Then*

$$\|X\|_F^2 \leq r \sum_{i=1}^{r} \|X_{ii}\|_F^2. \tag{14}$$

Note that we will use $\|X\|^2$ instead of $\|X\|_F^2$

Proof. Since $X = [X_{ij}]_{i,j=1}^r$ is positive semidefinite, so is any principle submatrix. Therefore, the 2×2 block matrix $\begin{pmatrix} X_{ii} & X_{ij} \\ X_{ji} & X_{jj} \end{pmatrix}$ must be positive semidefinite. We claim that

$$\|X_{ij}\|^2 \leq \|X_{ii}\|\|X_{jj}\| \leq \frac{\|X_{ii}\|^2 + \|X_{jj}\|^2}{2} \tag{15}$$

First inequality comes from Theorem 6. To derive the second inequality, note that

$$(\|X_{ii}\| - \|X_{jj}\|)^2 \geq 0.$$

This is equivalent to

$$\|X_{ii}\|^2 - 2\|X_{ii}\|\|X_{jj}\| + \|X_{jj}\|^2 \geq 0.$$

Rearranging, we obtain

$$\|X_{ii}\|\|X_{jj}\| \leq \frac{\|X_{ii}\|^2 + \|X_{jj}\|^2}{2}$$

Summing over all $i \neq j$ and applying (15), we have

$$\sum_{i \neq j} \|X_{ij}\|^2 \leq \sum_{i \neq j} \frac{\|X_{ii}\|^2 + \|X_{jj}\|^2}{2}$$

But

$$\begin{aligned}
\sum_{i \neq j} \frac{\|X_{ii}\|^2 + \|X_{jj}\|^2}{2} &= \frac{\|X_{11}\|^2 + \|X_{22}\|^2}{2} + \cdots + \frac{\|X_{11}\|^2 + \|X_{rr}\|^2}{2} \\
&\quad + \frac{\|X_{22}\|^2 + \|X_{11}\|^2}{2} + \cdots + \frac{\|X_{22}\|^2 + \|X_{rr}\|^2}{2} + \ldots \\
&\quad + \frac{\|X_{rr}\| + \|X_{11}\|^2}{2} + \cdots + \frac{\|X_{rr}\|^2 + \|X_{r-1r-1}\|^2}{2} \\
&= (r-1)\|X_{11}\|^2 + (r-1)\|X_{22}\|^2 + \cdots + (r-1)\|X_{rr}\|^2 \\
&= (r-1) \sum_{i=1}^{r} \|X_{ii}\|^2.
\end{aligned}$$

Therefore,

$$\sum_{i \neq j} \|X_{ij}\|^2 \leq (r-1) \sum_{i=1}^{r} \|X_{ii}\|^2.$$

Hence,

$$\begin{aligned}
\|X\|^2 &= \sum_{i=1}^{r} \|X_{ii}\|^2 + \sum_{i \neq j} \|X_{ij}\|^2 \\
&\leq \sum_{i=1}^{r} \|X_{ii}\|^2 + (r-1) \sum_{i=1}^{r} \|X_{ii}\|^2 \\
&= r \sum_{i=1}^{r} \|X_{ii}\|^2.
\end{aligned}$$

The result follows. □

Lemma 2 *Let X and Y be real symmetric matrices such that $XY = 0$. Then*

$$XY^* = X^*Y = Y^*X = YX^* = 0.$$

Proof. Since X and Y are real and symmetric, $X^* = X$ and $Y^* = Y$. Because $XY = 0$, it follows immediately that

$$X^*Y = XY^* = 0.$$

But since $0^* = (XY)^* = Y^*X^*$, we also have

$$Y^*X = YX^* = 0,$$

and the result is proved. □

Lemma 3 *Let X and Y be matrices, and assume that $XY = 0$. Then*

$$\|X + Y\|^2 = \|X - Y\|^2.$$

Proof. Let X and Y be as given. Then we have

$$\begin{aligned}
\|X + Y\|^2 &= \operatorname{tr}(X+Y)^*(X+Y) \\
&= \operatorname{tr}(X^* + Y^*)(X+Y) \\
&= \operatorname{tr}(X^*X + Y^*X + X^*Y + Y^*Y) \\
&= \operatorname{tr}(X^*X + 0 + 0 + Y^*Y) \\
&= \operatorname{tr}(X^*X) + \operatorname{tr}(Y^*Y) \\
&= \|X\|^2 + \|Y\|^2
\end{aligned}$$

Similarly, we have that

$$\begin{aligned}
\|X - Y\|^2 &= \operatorname{tr}(X-Y)^*(X-Y) \\
&= \operatorname{tr}(X^* - Y^*)(X-Y) \\
&= \operatorname{tr}(X^*X - Y^*X - X^*Y + Y^*Y) \\
&= \operatorname{tr}(X^*X - 0 - 0 + Y^*Y) \\
&= \operatorname{tr}(X^*X) + \operatorname{tr}(Y^*Y) \\
&= \|X\|^2 + \|Y\|^2
\end{aligned}$$

The result now follows. □

Lemma 4 *Let $X = [X_{ij}]_{i,j=1}^r$ and $Y = [Y_{ij}]_{i,j=1}^r$ be positive semidefinite matrices that are conformally partitioned into blocks, and assume that $XY = 0$. Further, assume that the diagonal blocks of X and Y coincide, so that $X_{ii} = Y_{ii}$ for all i $(1 \leq i \leq r)$. Then*

$$4 \sum_{i=1}^{r} \|X_{ii}\|^2 + \sum_{i \neq j} \|X_{ij} + Y_{ij}\|^2 = \sum_{i \neq j} \|X_{ij} - Y_{ij}\|^2. \tag{16}$$

Proof. First we consider $\|X - Y\|^2$. The $(1,1)$-block of $(X+Y)^*(X+Y)$ is:

$$\underline{(X_{11}^* + Y_{11}^*)(X_{11} + Y_{11})} + (X_{21}^* + Y_{21}^*)(X_{21} + Y_{21}) + \cdots + (X_{r1}^* + Y_{r1}^*)(X_{r1} + Y_{r1}).$$

Likewise, the $(2,2)$-block of $(X+Y)^*(X+Y)$ is:

$$(X_{12}^* + Y_{12}^*)(X_{12} + Y_{12}) + \underline{(X_{22}^* + Y_{22}^*)(X_{22} + Y_{22})} + \cdots + (X_{r2}^* + Y_{r2}^*)(X_{r2} + Y_{r2}).$$

Continuing in this way, we find that the (r,r)-block of $(X+Y)^*(X+Y)$ is:

$$(X_{1r}^* + Y_{1r}^*)(X_{1r} + Y_{1r}) + (X_{2r}^* + Y_{2r}^*)(X_{2r} + Y_{2r}) + \cdots + \underline{(X_{rr}^* + Y_{rr}^*)(X_{rr} + Y_{rr})}.$$

So, when we consider the sums of the underlined terms above,

$$\begin{aligned} \operatorname{tr}(X_{ii}^* + Y_{ii}^*)(X_{ii} + Y_{ii}) &= \operatorname{tr}(X_{ii}^* + X_{ii}^*)(X_{ii} + X_{ii}) \\ &= \operatorname{tr}(2X_{ii}^*)(2X_{ii}) \\ &= 4 \cdot \operatorname{tr}(X_{ii}^* X_{ii}) \\ &= 4 \cdot \|X_{ii}\|^2. \end{aligned}$$

Therefore,

$$\|X + Y\|^2 = 4 \cdot \sum_{i=1}^{r} \|X_{ii}\|^2 + \sum_{i \neq j} \|X_{ij} + Y_{ij}\|^2. \tag{17}$$

On the other hand, we consider $\|X - Y\|^2$. The $(1,1)$-block of $(X-Y)^*(X-Y)$ is:

$$\underline{(X_{11}^* - Y_{11}^*)(X_{11} - Y_{11})} + (X_{21}^* - Y_{21}^*)(X_{21} - Y_{21}) + \cdots + (X_{r1}^* - Y_{r1}^*)(X_{r1} - Y_{r1}).$$

Likewise, the $(2,2)$-block of $(X-Y)^*(X-Y)$ is:

$$(X_{12}^* - Y_{12}^*)(X_{12} - Y_{12}) + \underline{(X_{22}^* - Y_{22}^*)(X_{22} - Y_{22})} + \cdots + (X_{r2}^* - Y_{r2}^*)(X_{r2} - Y_{r2}).$$

Continuing in this way, we find that the (r,r)-block of $(X-Y)^*(X-Y)$ is:

$$(X_{1r}^* - Y_{1r}^*)(X_{1r} - Y_{1r}) + (X_{2r}^* - Y_{2r}^*)(X_{2r} - Y_{2r}) + \cdots + \underline{(X_{rr}^* - Y_{rr}^*)(X_{rr} - Y_{rr})}.$$

Since the diagonal blocks of X and Y coincide, the underlined terms are all zero. It follows that

$$\|X - Y\|^2 = \sum_{i \neq j} \|X_{ij} - Y_{ij}\|^2. \tag{18}$$

Thus, by Lemma 3, since $\|X + Y\|^2 = \|X - Y\|^2$, lines (17), (18) imply the result. $\square$

Lemma 5 *Let $X = [X_{ij}]_{i,j=1}^r$ and $Y = [Y_{ij}]_{i,j=1}^r$ be positive semidefinite matrices that are conformally partitioned into blocks, and assume that $XY = 0$. Further, assume that the diagonal blocks of X and Y coincide, so that $X_{ii} = Y_{ii}$ for all i $(1 \leq i \leq r)$. Then*

$$\sum_{i=1}^{r} \|X_{ii}\|^2 = -\sum_{i \neq j} \Re\left(trX_{ij}^* Y_{ij}\right).$$

Proof. By Lemma 4,

$$\begin{aligned}
4 \cdot \sum_{i=1}^{r} \|X_{ii}\|^2 &= \sum_{i \neq j} \|X_{ij} - Y_{ij}\|^2 - \sum_{i \neq j} \|X_{ij} + Y_{ij}\|^2 \\
&= \sum_{i \neq j} \left(\mathrm{tr}\left(X_{ij} - Y_{ij}\right)^* \left(X_{ij} - Y_{ij}\right) - \mathrm{tr}\left(X_{ij} + Y_{ij}\right)^* \left(X_{ij} + Y_{ij}\right)\right) \\
&= \sum_{i \neq j} \mathrm{tr}\left(\left(X_{ij} - Y_{ij}\right)^* \left(X_{ij} - Y_{ij}\right) - \left(X_{ij} + Y_{ij}\right)^* \left(X_{ij} + Y_{ij}\right)\right) \\
&= \sum_{i \neq j} \mathrm{tr}\left(-2X_{ij}^* Y_{ij} - 2Y_{ij}^* X_{ij}\right) \\
&= -2\sum_{i \neq j} \mathrm{tr}\left(X_{ij}^* Y_{ij} + Y_{ij}^* X_{ij}\right) \\
&= -2\sum_{i \neq j} \left(\mathrm{tr}\left(X_{ij}^* Y_{ij}\right) + \mathrm{tr}\left(Y_{ij}^* X_{ij}\right)\right) \\
&= -2\sum_{i \neq j} \left(\mathrm{tr}\left(X_{ij}^* Y_{ij}\right) + \mathrm{tr}\overline{\left(X_{ij}^* Y_{ij}\right)}\right) \\
&= -2\sum_{i \neq j} \left(\mathrm{tr}\left(\left(X_{ij}^* Y_{ij}\right) + \overline{\left(X_{ij}^* Y_{ij}\right)}\right)\right) \\
&= -2\sum_{i \neq j} \mathrm{tr}\left(2 \cdot \Re\left(X_{ij}^* Y_{ij}\right)\right) \\
&= -4\sum_{i \neq j} \mathrm{tr}\left(\Re\left(X_{ij}^* Y_{ij}\right)\right),
\end{aligned}$$

and the result follows. $\square$

If you take $r = 3$, $X = \frac{2}{3}J_3$ and $Y = I_3 - \frac{1}{3}J_3$, then from *Lemma 1* to *Lemma 5* can be verified.

Now we present the following theorem, then we apply it to prove the main result.

Theorem 7 *Let $X = [X_{ij}]_{i,j=1}^r$ and $Y = [Y_{ij}]_{i,j=1}^r$ be positive semidefinite matrices that are conformally partitioned into blocks, and assume that $XY = 0$. Further, assume that the diagonal blocks of X and Y coincide, so that $X_{ii} = Y_{ii}$ for all i $(1 \leq i \leq r)$. Then*

$$\|X\|^2 \leq (r-1)\|Y\|^2.$$

Proof. By Lemma 3,

$$\|X + Y\|^2 = \|X - Y\|^2.$$

By Lemma 4,

$$4\sum_{i=1}^{r}\|X_{ii}\|^2+\sum_{i\neq j}\|X_{ij}+Y_{ij}\|^2=\sum_{i\neq j}\|X_{ij}-Y_{ij}\|^2.$$

So by Lemma 5,

$$\sum_{i=1}^{r}\|X_{ii}\|^2=-\sum_{i\neq j}\Re(tr(X_{ij}^*Y_{ij}).$$

So the Cauchy-Schwarz inequality implies, with $d:=\sum_{i=1}^{r}\|X_{ii}\|^2$, that

$$d\leq\sum_{i\neq j}\|X_{ij}\|\|Y_{ij}\|. \tag{19}$$

Again applying the Cauchy-Schwarz inequality, we now have

$$d\leq\left(\sum_{i\neq j}\|X_{ij}\|^2\right)^{\frac{1}{2}}\left(\sum_{i\neq j}\|Y_{ij}\|^2\right)^{\frac{1}{2}}.$$

The right side above can be written as

$$\left(\|X\|^2-d\right)^{\frac{1}{2}}\left(\|Y\|^2-d\right)^{\frac{1}{2}},$$

so it follows that

$$d^2\ \leq\ \|X\|^2\|Y\|^2-(\|X\|^2+\|Y\|^2)d+d^2.$$

Rearranging, we obtain

$$(\|X\|^2+\|Y\|^2)\,d\ \leq\ \|X\|^2\|Y\|^2$$

By Lemma 1,

$$\|X\|^2\leq r\sum_{i=1}^{r}\|X_{ii}\|^2=rd.$$

Combining the two inequalities above with $d>0$ yields

$$\|X\|^2+\|Y\|^2\leq r\|Y\|^2,$$

as desired. □

Now we are ready to present the proof of **Theorem 1**.

Proof of Theorem 1. Let A_G denote the adjacency matrix of a graph G, and let π,ν and δ denote the numbers of positive, negative and zero eigenvalues of A_G respectively. Consider the spectral decomposition of A_G:

$$A_G=\sum_{i=1}^{n}\mu_i\upsilon_i\upsilon_i^*. \tag{20}$$

Write the adjacency matrix as $A_G=B-C$, where

$$B=\sum_{i=1}^{\pi}\mu_i\upsilon_i\upsilon_i^*\quad\text{and}\quad C=\sum_{i=n-\nu+1}^{n}\left(-\mu\right)_i\upsilon_i\upsilon_i^*, \tag{21}$$

.

If we let r denote the chromatic number $\chi(G)$, then the color classes of a proper r-coloring partition the matrix A_G into $r \times r$ blocks. This partition is inherited conformally by the matrices B, C. Since A_G has zero blocks on the diagonal, B and C must have identical diagonal blocks.

Since B and C both have all positive eigenvalues, they are both positive semidefinite. By the orthonormality of the basis of eigenvectors, they also satisfy the condition that $BC = 0$.

Theorem 8 gives us

$$\|B\|^2 \leq (r-1)\,\|C\|^2.$$

Since $\|C\|^2$ is not zero, this implies

$$1 + \frac{\|B\|^2}{\|C\|^2} \leq r$$

Taking into account that

$$s^+ = \mu_1^2 + \mu_2^2 + \cdots + \mu_\pi^2 = \|B\|^2,$$

and

$$s^- = \mu_{n-\nu+1}^2 + \mu_{n-\nu+2}^2 + \cdots + \mu_n^2 = \|C\|^2,$$

we obtain the desired result. □

Interchanging the roles of B and C if necessary, we have the following.

Corollary 2 *Let G, s^+, s^- be as previously defined. Then*

$$\chi(G) \geq 1 + \max\{\frac{s^-}{s^+}, \frac{s^+}{s^-}\}. \tag{22}$$

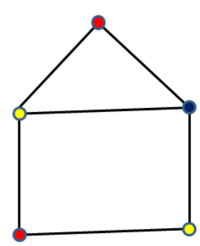
Theorem 8 *(Hoffman [2, p.79-91]) The chromatic number is bounded by*

$$\chi(G) \geq 1 + \frac{\mu_1}{-\mu_n}. \tag{23}$$

4 Applications

We conclude with a number of examples to illustrate the power of the results and compare the Hoffman lower bound.

Example 4 *Let H be the House graph, shown below.*



Proper 3-coloring of the House graph

Then the corresponding adjacency matrix A_H is given by

$$A_H = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

So we can verify the lower bound for the House graph as follows. Since the eigenvalues are

$$\mu_1 \approx 2.481,\ \mu_2 \approx 0.689,\ \mu_3 = 0\ \mu_4 \approx -1.170,\ \mu_5 = -2,$$

we have that

$$s^+ = \mu_1^2 + \mu_2^2 \approx 6.63$$

and

$$s^- = \mu_4^2 + \mu_5^2 + \approx 5.369.$$

Therefore,

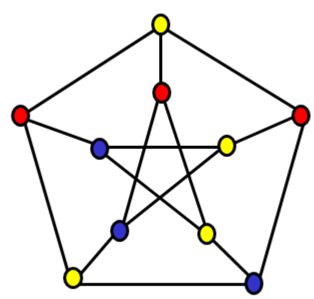
$$1 + \frac{s^+}{s^-} \approx 2.235,$$

which is lower bound for the chromatic number of the House graph. As we must use at least 3 colors to properly color H, we have $\chi(H)$ is 3. In this case, the Hoffman bound is

$$1 + \frac{\mu_1}{-\mu_5} \approx 2.241,$$

so the Hoffman bound is better than the new bound.

Example 5 *Let P be the Petersen graph, shown below.*



Proper 3-coloring of the Petersen graph

The spectrum of eigenvalues are below:

$$Spec(P) = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 5 & 4 \end{pmatrix}$$

So $s^+ = 14$ and $s^- = 16$. Therefore,

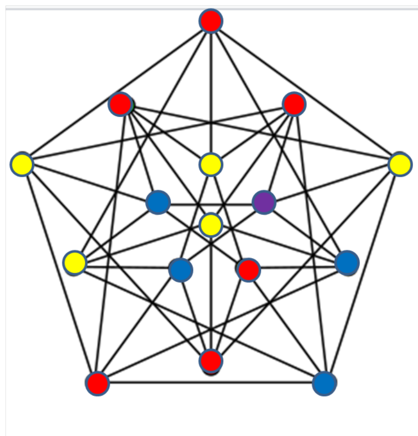
$$1 + \frac{s^-}{s^+} \approx 2.143,$$

which is lower bound for the chromatic number of the Petersen graph. The actual value of $\chi(P)$ is 3. The Hoffman bound of the Petersen graph is

$$1 + \frac{\mu_1}{-\mu_{10}} = 2.5,$$

so the Hoffman bound is better than the new bound.

Example 6 *Let C be the Clebsch graph as shown below.*



Clebsch graph

The Clebsch graph consists of a 5-regular graph with 40 edges. Since the characteristic polynomial is

$$(\mu + 3)^5 (\mu - 1)^{10} (\mu - 5),$$

the spectrum of eigenvalues are below:

$$Spec(C) = \begin{pmatrix} -3 & 1 & 5 \\ 5 & 10 & 1 \end{pmatrix}$$

Since $s^+ = 35$ and $s^- = 45$,

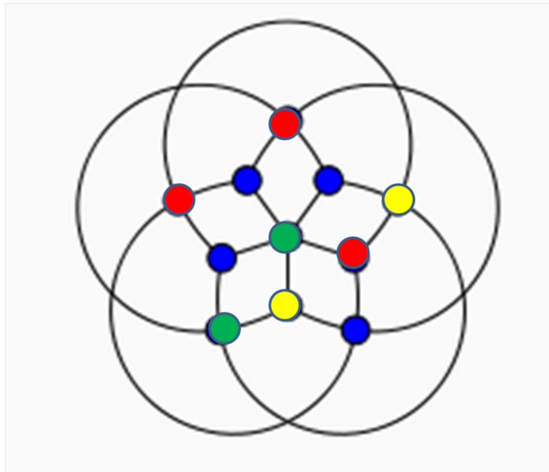
$$1 + \frac{s^-}{s^+} \approx 2.286,$$

which is a lower bound for the chromatic number of the Clebsch graph. The chromatic number $\chi(C)$ is known to be 4. The Hoffman bound of the Clebsch graph is

$$1 + \frac{\mu_1}{-\mu_{16}} \approx 2.667,$$

so the Hoffman bound is better than the new bound.

Example 7 *Let G be the Grötzsch graph, shown below.*



Grötzsch graph

The Grötzsch graph is a triangle-free graph with 11 vertices and 20 edges. Since the characteristic polynomial of the Grötzsch graph is

$$(\mu-1)^5\left(\mu^2-\mu-10\right)\left(\mu^2+3\mu+1\right)^2,$$

we find that

$$s^+ \approx 18.702, \quad \text{and} \quad s^- \approx 21.298.$$

Therefore,

$$1+\frac{s^-}{s^+}\approx 2.139,$$

which is a lower bound for the chromatic number of the Grötzsch graph. It is known that $\chi(G)=4$*. The Hoffman bound of the Grötzsch graph is*

$$1+\frac{\mu_1}{-\mu_{11}}\approx 2.370,$$

so the Hoffman bound is better than the new bound.

5 Conclusion

Spectral bounds in graph theory typically give graph theoretic information in terms of some function of the eigenvalues of the adjacency matrix. In this paper, we have investigated a new lower bound for the chromatic number which involve all of the eigenvalues of the adjacency matrix. By proving the conjecture of Wocjan and Elphick, this work provides an unexpected relationship between the signs of the eigenvalues of the adjacency matrix and the chromatic number of the corresponding graph. Also, we verified the power of the new lower bound for some examples comparing the Hoffman bound.

References

[1] T. Ando and M. Lin. Proof of a conjectured lower bound on the chromatic number of a graph. *Linear Algebra and its Applications*, 485:480–484, 2015.

[2] A.J. Hoffman. *On eigenvalues and colourings of graphs in: Graph Theory and its Applications.* Academic Press, 1970.

[3] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, 1991.

[4] D. B. West. *Introduction to graph theory.* Pearson Modern Classic, 2018.

[5] P. Wocjan and C. Elphick. New spectral bounds on the chromatic number encompassing all eigenvalues of the adjacency matrix. *Electron. J. Combin.*, 20(3):P39, 2013.