The Classification of Matricies with Non-Negative Integer Coefficients

Milo Moses

July 8, 2023

Abstract

In this article we give a modern, pedagogical, streamlined proof of one of the simplest non-trivial classification theorem in mathematics. Namely, we classify matrices with non-negative integer coefficients whose matrix 2-norm is less than two using A-D-E Dynkin diagrams. We then use this toy model to demonstrate the key features and subtleties of classification theorems.

“One must not be childishly repelled by the examination of the humbler animals, for in all things of nature there is something wonderful”
- Aristotle
1 Introduction

The classification of finite simple groups is widely considered one of the deepest theorems ever proved in mathematics. It states that every finite group with no proper non-trivial normal subgroups will be isomorphic to either an element of one of eighteen infinite families, or to one of twenty seven “sporadic” groups.

Understanding classification theorems is often seen as a daunting task: even the statement of the classification of finite simple groups is quite complicated. The proofs also tend to be extremely long, requiring large amounts of casework. The proof of the classification of finite simple groups consists of tens of thousands of pages across hundreds of journal articles. The proof is so complicated, in fact, that in 1983 Daniel Gorenstein announced the completion of the classification only to discover that he was misinformed about the status of the “quasithin” case. The now accepted announcement of the completion of the classification came over twenty years later, in a 2004 article of Michael Aschbacher [Asc04].

The bulk of the finite simple groups are the so-called finite simple groups of Lie type. They come from finite dimensional simple Lie algebras, objects which themselves admit a classification theorem. The classification of finite dimensional simple Lie algebras is much simpler, but it still requires a large amount of deep theory and case work [GG20].

In this article, we seek to demonstrate the general principles of the classification using the simplest non-trivial example. Namely, we classify matrices \( M \) satisfying the following properties:

1. \( M \) has non-negative integer coefficients,
2. \( M^T = M \), where \( M^T \) denotes the transpose of \( M \),
3. \( M \) has all zeros on its diagonal,
4. All eigenvalues of \( M \) are of absolute value less than 2.

We offer some commentary about these conditions. Property (1) gives the setting of this problem - we are proving a theorem about non-negative integer matrices. Properties (2)-(3) are technical conditions, which are unimportant. In fact, they can be removed by introducing extra structure - we go through this in Appendix A. We include these conditions for now because they make the problem simpler, and all of the key ideas are still demonstrated. Property (4) gives the problem all of its flavour. Without it, the classification is trivial. Namely, matrices are classified by their coefficients and properties (1)-(3) restrict the possibilities for the coefficients in immediately obvious ways. It is property (4) that makes this a linear algebra problem, and makes the solution interesting.

The general outline for the classification is as follows. Just like we only classify groups up to isomorphism, we only classify matrices up to a specified notion of equivalence. We then define a simple way of breaking down matrices into “irreducible” chunks. We then find that irreducible matrices, up to equivalence, are in one to one correspondence with graphs of the following type:
The connection between graphs and matrices is that every graph can be given an adjacency matrix. The rows and columns are labeled by vertices. The $(v,w)$ entry is 1 if there is an edge between $v$ and $w$ and zero otherwise. The conditions (1)-(3) are exactly what is needed to guarantee that some matrix is the adjacency matrix of a “simple” graph.

The structure of this classification is very typical. There are a few (in this case, two) infinite families, and a few (in this case, three) exceptional objects which don’t fit into any infinite family. In turns out that there are a large number of objects across mathematics which are classed in terms of these exact same graphs. These are known as A-D-E classifications. A sample of objects that follow A-D-E classifications are listed below:

1. Platonic solids [VH02],
2. Representations of quivers [Bri08],
3. Special types of singularities of algebraic hypersurfaces [DV34],
4. Coxeter groups generated by reflections [Hum92],
5. Certain 2-dimensional conformal field theories [CZ09],
6. and many more [Sie14].

The number of A-D-E classifications led Vladimir Arnold [Arn76] to pose the following question: why are so many things classified by the same graphs? This was stated in the survey article [HHSV77] as follows:

“The problem is to find the common origin of all the A-D-E classification theorems and to substitute a priori proofs for a posteriori verifications of the parallelism of the classifications”

- Hazewinkel-Hesselink-Siersma-Veldkamp.
Roughly, the classification theorem given in this paper is the answer to the above question. In all A-D-E situations, your data can be somehow represented as matrices with non-negative integer coefficients and small eigenvalues. You will need a bit more data than just a matrix, but this will just correspond to adding a few extra infinite families or exceptional graphs to your classification.

The structure of this paper is as follows:

In Section 2, we give background for the main theorem and state it precisely. This includes discussion of all the components which go into a classification theorem, and the subtleties that can sometimes occur.

In Section 3, we offer a proof of the main theorem using the theory Frobenius-Perron eigenvectors.

In Appendix A, we extend the classification to non-symmetric matrices by introducing bicolorations.

In Appendix B, we compute the size of the largest eigenvalue of the adjacency matrices of A-D-E graphs.

While surely not a difficult exercise, the classification theorem as stated here does not appear in literature. The closest work is the first chapter of Jones et. al’s book [GdLHJ12], which proves the more general theorem we have in Appendix A. Jones et. al’s proof is in turn inspired by the original work of Frobenius on the subject [FFF+12]. While certainly a good reference, Jones et. al’s book leaves many details to references or as exercises to the reader, is unpedagogical in its approach, and does not emphasise key ideas. It is for this reason we see it necessary to give a more modern account which brings to light this wonderful piece of elementary mathematics.

2 The Main Theorem

2.1 Discussing the axioms

Our goal in this section is to state exactly the main theorem of this paper, and give appropriate surrounding discussion. Recall that the main theorem is a classification of matrices $M$ satisfying the following properties:

1. $M$ has non-negative integer coefficients,
2. $M^T = M$, where $M^T$ denotes the transpose of $M$,
3. $M$ has all zeros on its diagonal,
4. All eigenvalues of $M$ are of absolute value less than 2.

Note in particular that the condition (2) implies that $M$ is square matrix, so conditions (3)-(4) are well defined.

The point of conditions (1)-(3) is that they are exactly the conditions needed to make $M$ the adjacency matrix of a simple graph. Here, a graph is a collection of vertices with edges between them. We will restrict our attention to a nice family of graphs we call simple graphs. A simple graph is a graph $\Gamma$ satisfying the following conditions:
1. Γ has no repeating edges between vertices,
2. Γ has undirected edges,
3. Γ has no edges going from vertices to themselves.

Given a simple graph Γ, we define the adjacency matrix of Γ to be the matrix whose rows and columns are labeled by vertices, and whose \((v, w)\) entry is 1 if there is an edge between \(v\) and \(w\) and zero otherwise. We give some examples below:

\[
\begin{align*}
\text{Γ} & = \begin{array}{ccc}
v_0 & v_1 & v_2 \\
\end{array}, \\
\text{adjacency matrix} & = \begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 1 \\
0 & 1 & 0 \\
\end{bmatrix} \\
\text{Γ} & = \begin{array}{ccc}
v_0 & v_1 & v_2 & v_3 & v_4 \\
\end{array}, \\
\text{adjacency matrix} & = \begin{bmatrix}
v_0 & v_1 & v_2 & v_3 & v_4 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 \\
\end{bmatrix}
\end{align*}
\]

The condition (2) on graphs is equivalent to the condition (2) on matrices. This is because edges being undirected means there is an edge between \(v\) and \(w\) if and only if there is an edge between \(w\) and \(v\), so the transpose of the adjacency matrix is equal to itself.

The condition (3) on graphs is equivalent to the condition (3) on matrices. This is because if there is no edge going from a vertex to itself the \((v, v)\) entry is zero for all \(v\). That is, the matrix has zeros on its diagonal.

The condition (1) on graphs, however, is unnecessary to imply condition (1) on matrices. We could have allowed repeating edges, and defined that the \((v, w)\) entry of the adjacency is equal to the number of edges between \(v\) and \(w\). How can we ever expect to recover matrices satisfying conditions (1)-(3) from simple graphs, if we cannot create entries that are bigger than 1? The answer lies in condition (4). It turns out that if any of the entries of the matrix are \(\geq 2\), then the largest eigenvalue of \(M\) will be at least 2 in absolute value. Hence, we know that \(M\) must only have 0s and 1s as entries and adding condition (1) on graphs does not lose us any power. We will prove this in the next section.

Seeing as we will be referring to matrices which satisfy conditions (1)-(4) a lot, we introduce some terminology. We call a matrix satisfying conditions (1)-(3) and whose entries are all 0 or 1 a \textit{simple adjacency matrix}. These are exactly the matrices which appear as the adjacency matrices of simple graphs. The absolute value of the largest eigenvalue of a matrix is known as its \textit{spectral radius}\footnote{The spectrum of a matrix refers to its set of eigenvalues, hence the terminology.}. In this new lingo, we are in the business of classifying simple adjacency matrices of spectral radius less than 2.
2.2 Equivalence and reducibility

The first step of any classification theorem is to define the correct notion of equivalence. It is hopeless to classify groups up to equality: we classify them up to isomorphism. In the case of simple adjacency matrices, we define equivalence as follows. A permutation of a set is a bijective map from that set to itself. Every permutation \( \omega : [1...n] \rightarrow [1...n] \) of the first \( n \) natural numbers has an associated permutation matrix \( \Omega \). This matrix is defined by setting the \((k, \omega(k))\) entry to be 1, and all other entries to be 0. Applying a permutation matrix on the right has the effect of permuting columns, and applying a permutation matrix on the left has the effect of permuting the rows.

We say that two simple adjacency matrices \( M \) and \( N \) are equivalent if there exists a permutation matrix \( \Omega \) such that

\[
M = \Omega N \Omega^{-1}.
\]

In this case, we write \( M \sim N \). Equivalence can also be viewed graph-theoretically:

**Proposition 1.** Let \( M, N \) be simple adjacency matrices with associated graphs \( \Gamma_M, \Gamma_N \). That is, the vertices of \( \Gamma_M \) and \( \Gamma_N \) are labeled by the numbers 1...\( n \) and there is an edge between from \( j \) to \( k \) if and only if the \((j,k)\) entry of the corresponding matrix is 1.

The matrices \( M, N \) are equivalent if and only if there is a graph isomorphism between \( \Gamma_M \) and \( \Gamma_N \). That is, if and only if there is a way of moving vertices of \( \Gamma_M \) onto the vertices of \( \Gamma_N \) so that the edges agree.

**Proof.** Suppose we are given a graph isomorphism from \( \Gamma_M \) to \( \Gamma_N \). This means that every vertex \( k \) of \( \Gamma_M \) is assigned a vertex \( \omega(k) \) of \( \Gamma_N \), and there is an edge from \( j \) to \( k \) in \( \Gamma_M \) if and only if there is an edge from \( \omega(j) \) to \( \omega(k) \) in \( \Gamma_N \). This \( \omega \) is a permutation, and hence defines a permutation matrix \( \Omega \). It is immediate from doing the calculation that \( M = \Omega N \Omega^{-1} \).

Conversely, if \( M = \Omega N \Omega^{-1} \) for some permutation matrix \( \Omega \) then the underlying permutation \( \omega \) defines a graph isomorphism \( \Gamma_M \) to \( \Gamma_N \) in the obvious fashion. Hence, the proof is complete.

We now turn to the second step of any classification theorem: defining a notion irreducibility. It is hopeless to classify groups, even up to isomorphism. We classify simple groups instead. It is here that some classification theorems hide a big subtlety. Even though one might have classified the basic building blocks of some type of object, it is not always easy to build things out of those blocks. For instance, let \( G \) be a group and let \( N \) be a normal subgroup. The basic idea is that \( G \) is built out of the subgroup \( N \) and quotient \( G/N \). However, we do not necessarily have that \( G \cong N \times G/N \). Non-trivial extensions of \( N \) and \( G/N \) are very complicated. Thus, we can say we have classified finite simple groups but we can not say we have classified finite simple groups.

\[2\] Still, there is some sense in which every group’s structure can be broken up into a unique collection of simple groups. This is the Jordan-Holder theorem.
In our case, the notion of reducibility is straightforward. We say that a simple adjacency matrix $M$ is reducible if there exist simple adjacency matrices $M_1, M_2$ such that there is an equivalence

$$M \sim \begin{bmatrix} 0 & M_2 \\ M_1 & 0 \end{bmatrix},$$

where by 0 we indicate a block of zeros of the appropriate size to make $M$ as well defined matrix. We call a matrix irreducible if it is not reducible. Again, we can verify that this is the “correct” notion of irreducibility by checking that is has a nice graph-theoretical analogue:

**Proposition 2.** Let $M$ be a simple adjacency matrix with associated graph $\Gamma_M$. The matrix $M$ is irreducible if and only if $\Gamma_M$ is connected. That is, if and only if one can go from every vertex on $\Gamma_M$ to every other vertex on $\Gamma_M$ by walking along edges.

**Proof.** If a matrix $M$ is reducible into blocks $M_0$ and $M_1$, then it is clear that the vertices corresponding to the rows containing $M_0$ and the vertices corresponding the the rows containing $M_1$ form disconnected components in $\Gamma_M$. Conversely, any disconnected graph can be split into two disconnected components. The adjacency matrices of these disconnected components will clearly induce a reduction of $M$, and thus our proof is complete.

By our definition of reducibility, it is clear that every simple adjacency matrix can be written in the form

$$M \sim \begin{bmatrix} 0 & 0 & \ldots & M_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & M_2 & \ldots & 0 \\ M_1 & 0 & \ldots & 0 \end{bmatrix}$$

where $M_1, \ldots, M_n$ are irreducible. Thus, every simple adjacency matrix is built out of irreducible ones. One still needs to check is that every matrix is uniquely decomposable in terms of irreducible matrices, so that we can indentify matrices by naming their constituent irreducible factors. Mathematically, what we want to say is the following. If there were a second decomposition for $M$ as

$$M \sim \begin{bmatrix} 0 & 0 & \ldots & M'_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & M'_2 & \ldots & 0 \\ M'_1 & 0 & \ldots & 0 \end{bmatrix}$$

with $M'_1, \ldots, M'_n$, irreducible, then

1. $n = n'$,
2. There is a permutation $\omega : [1 \ldots n] \rightarrow [1 \ldots n]$ so that $M_n \sim M'_{\omega(n)}$. 


That is, a block decompositions of simple adjacency matrices are unique up to permuting factors. Graph theoretically, the uniqueness of the decomposition is simply the fact that graphs have well defined connected components. More precisely, let $\Gamma_1...\Gamma_n$ be the connected components of $\Gamma_M$. That is, the subgraphs of $\Gamma_M$ consisting of clusters of vertices that can be reached from one to another by walking along edges. Let $M_1...M_n$ be the corresponding adjacency matrices of $\Gamma_1...\Gamma_n$. We have that

$$
M \sim \begin{bmatrix}
0 & 0 & \ldots & M_n \\
\vdots & \vdots & \ddots & \vdots \\
0 & M_2 & \ldots & 0 \\
M_1 & 0 & \ldots & 0
\end{bmatrix}.
$$

Conversely, every decomposition of $M$ induces a decomposition of $\Gamma_M$ into connected components. Clearly every way of breaking up $\Gamma_M$ into connected components will be the same up to permuting the factors, and hence uniqueness follows.

Up to now we have been ignoring the condition that the spectral radius be less than 2. In theory this could cause some issues. Namely, we might have matrices which decompose into smaller simple adjacency matrices, but do not decompose into smaller adjacency matrices of spectral radius less than 2. Seeing as we will now be dealing a lot with spectral radii, we denote by $\|M\|$ the spectral radius of a matrix $M$. That is, the absolute value of its largest eigenvalue. The following proposition guarantees that we will get no issues around decomposing simple adjacency matrices with spectral radius less than 2:

**Proposition 3.** Let $M$ be a simple adjacency matrix, with a decomposition

$$
M \sim \begin{bmatrix}
0 & 0 & \ldots & M_n \\
\vdots & \vdots & \ddots & \vdots \\
0 & M_2 & \ldots & 0 \\
M_1 & 0 & \ldots & 0
\end{bmatrix}
$$

into (not necessarily irreducible) simple adjacency matrices $M_1...M_n$. It holds that

$$
\|M\| = \max\{\|M_1\|,\ldots,\|M_n\|\}.
$$

In particular, if $\|M\| < 2$ then $M$ decomposes into irreducible blocks which all satisfy $\|M_k\| < 2$.

*Proof.* Every vector $z$ has a unique decomposition $z = \sum_{k=0}^n z_k$, where $z_k$ has non-zero elements only on the columns corresponding to $M_k$. $M$ acts by $M_k$ on each $z_k$. Hence, for $z$ to be an eigenvector for $M$ each $z_k$ has to be 0 or an eigenvector to its corresponding $M_k$. Moreover, all of the eigenvalues of the eigenvectors must be the same. In particular, every eigenvalue for $M$ must be an eigenvalue of one the of the $M_k$s. Conversely, every eigenvalue for one of the
M_ks gives an eigenvalue for M by choosing an M_k-eigenvector and padding it with zeros. Hence, by the definition of the spectral radius as the maximum over eigenvalues, our proof is complete.

We have now proven all of the necessary background results on equivalence and reducibility.

### 2.3 The statement

We are now ready to state the main classification theorem:

**Theorem 1.** Let M be a irreducible simple adjacency matrix of spectral radius less than 2. M is equivalent to one of the graphs A_\ell for \ell \geq 1, D_\ell for \ell \geq 4, or E_\ell for \ell = 6, 7, or 8, where A_\ell, D_\ell, E_\ell are the graphs pictured in the introduction.

In particular, we find that there are infinitely many irreducible simple adjacency matrices. The vast majority of irreducible simple adjacency matrices will be part of one of the two infinite families A_\ell, D_\ell, with E_6, E_7, E_8 giving the three unruly exceptions.

Armed with this classification theorem, it is very easy to prove things about simple adjacency matrices. All one has to do is first prove your result for irreducible simple adjacency matrices, and then show that the statement is preserved as you put matrices together. We demonstrate with an example:

**Theorem 2.** Let M be a simple adjacency adjacency matrix. If \|M\| < 2, then there exists \(n \geq 3\) such that

\[\|M\| = 2 \cos(\pi/n).\]

This is a very surprising theorem. One starts with a seemingly innocuous situation (symmetric non-negative integer matrices with zeros on the diagonal), and we find that if the spectral radius is small enough then it has to fall along some strange discrete sequence of numbers.

Proving this result is very easy once we establish the basics of the theory of Frobenius-Perron eigenvectors, which we do next section. Frobenius-Perron theory allows us to efficiently compute the spectral radius of matrices. In particular, we compute \|M\| for every M in the classification theorem of irreducible simple adjacency matrices. Then, one uses Proposition 3 to find that the spectral radius of any simple adjacency matrix must be the maximum along its irreducible components, and hence the result follows. These computation are performed in Appendix B.

Some form of Theorem 2 dates back to a 1857 paper of Kronecker [Kro57], who proved the key algebraic principle which underlies it. Certainly, his proof does not make use of any sort of classification. Roughly, they key point is the following:

**Theorem 3** (Kronecker). Let \(f(x)\) be a polynomial with integer coefficients. If all of the roots of \(f(x)\) lie on the unit circle (i.e. they have absolute value 1),
then all of the roots of \( f(x) \) are roots of unity (i.e. they are solutions to the equation \( x^n = 1 \) for some \( n \geq 1 \)).

From this, we get following algebraic corollary:

**Corollary 1.** Let \( f(x) \) be a polynomial with integer coefficients. If all of the roots of \( f(x) \) real and lie in the interval \([-2, 2]\), then all of the roots of \( f(x) \) are of the form \( 2 \cos(\pi r) \) for some rational number \( r \).

In particular, the root of \( f(x) \) with largest absolute value will be of the form \( 2 \cos(\pi/n) \) for some \( n \geq 1 \).

**Proof.** Since \(-2 \leq 2 \cos(\theta) \leq 2\) for all angles \( \theta \), we can write all of the roots of \( f(x) \) has \( 2 \cos(\pi \theta_1), 2 \cos(\pi \theta_2), \ldots, 2 \cos(\pi \theta_d) \) where \( d \) is the degree of \( f \). Now, we consider the polynomial \( g(x) = x^n f(x + 1/x) \). It can be expanded as follows:

\[
g(x) = x^n \prod_{k=1}^d (x + 1/x - 2 \cos(\pi \theta_k))
\]

\[
= \prod_{k=1}^d (x^2 - 2 \cos(\pi \theta_k)x + 1)
\]

\[
= \prod_{k=1}^d (x - e^{\pi i \theta_k}) (x + e^{-\pi i \theta_k}).
\]

All of the numbers \( e^{\pm \pi i \theta_k} \) are of absolute value 1, and hence Kronecker’s theorem applies. That is, all of the \( e^{\pm \pi i \theta_k} \) are roots of unity so all of the \( \theta_k \) are rational.

The fact that the largest eigenvalue must be of the form \( 2 \cos(\pi/n) \) for some \( n \geq 1 \) follows from general theory. Namely, if \( f(x) \) has some root \( 2 \cos(\pi k/n) \) with \( k \) and \( d \) relatively prime then it will also have \( 2 \cos(\pi/n) \) as a root because these numbers are Galois conjugates [Was97]. Since \( |2 \cos(k \pi/n)| \leq 2 \cos(\pi/n) \) for all integers \( k \), we conclude our result.

Getting Theorem 2 from Corollary 1 is a simple job. Let \( M \) be a simple adjacency matrix. Let \( f_M(x) \) be its characteristic polynomial. By property (1) it has integer coefficients, by property (2) all of its roots are real, and by property (4) all of its roots lie in the interval \((-2, 2)\). Hence, Kronecker’s theorem applies and we conclude the result.

Theorem 2 is not only a testament to the power of classification theorems, but also gives some intuition for why we must require \( \|M\| < 2 \). When \( \|M\| < 2 \), the eigenvalues are well-behaved - \( 2 \cos(\pi r) \) for rational numbers \( r \) - and presumably will have nice combinatorics coming from this fact. When \( \|M\| = 2 \), these nice combinatorics remain essentially intact. In fact, adding two extra infinite families and a few exceptional graphs the classification can be extended to the case \( \|M\| = 2 \). When \( \|M\| > 2 \) the possible eigenvalues no longer form a discrete set, and they are much harder to control. It is for this reason one has to be very careful when constructing toy models, since tweaking parameters very slightly can drastically change the results.
3 The proof

3.1 Discussion and preliminaries

To begin, we give a general outline of the proof. The first step is to understand eigenvalues and eigenvectors of symmetric non-negative matrices better, mainly through the Frobenius-Perron theorem. With this in hand we will be able to easily compute the spectral radius of matrices. This Frobenius-Perron theory will also allow us to prove a key lemma: if one graph \( \Gamma_2 \) contains another graph \( \Gamma_1 \) as a subgraph, then the spectral radius of \( \Gamma_2 \) (i.e. the spectral radius of its adjacency matrix) will be greater than the spectral radius of its subgraph \( \Gamma_1 \). In particular, if a graph has spectral radius \( < 2 \) then it cannot contain any subgraphs of spectral radius \( \geq 2 \).

The rest of the proof is now a fun game. By constructing more and more graphs with spectral radius \( \geq 2 \), we get tighter and tighter restrictions for what graphs with spectral radius \( < 2 \) can look like, since they cannot contain any of the graphs we constructed as subgraphs. Eventually, these restrictions will be so tight that the only possibilities left are the A-D-E ones.

With this general programme established, we now begin with some necessary results about eigenvectors/eigenvalues. Our first result is a powerful theorem which re-frames the spectral radius in a form which does not use eigenvectors. This will be very useful in our study, since it allows us to get bounds on the spectral radius without needing to compute eigenvectors. We can only prove one half however, since the other half is beyond the scope of this paper.

**Proposition 4.** Let \( M \) be an \( n \) by \( n \) matrix with real coefficients and for which \( M^T = M \). Then,

\[
\| M \| = \max_{z \in \mathbb{C}^n} \frac{\| Mz \|}{\| z \|}
\]

where

\[
\| z \| = \sqrt{\sum_{k=1}^{n} |z_k|^2}
\]

denotes the Euclidean norm of a vector \( z = (z_1...z_k) \).

**Proof.** Let \( z_0 \) be an eigenvector with maximal absolute value for \( M \). That is, a vector for which \( Mz_0 = \lambda z_0 \), \( |\lambda| = \| M \| \). Then,

\[
\| Mz_0 \|/\| z_0 \| = \| \lambda z_0 \|/\| z_0 \| = \| M \|.
\]

Hence, \( \| M \| \leq \max_{z \in \mathbb{C}^n} \frac{\| Mz \|}{\| z \|} \). The inequality the other direction is an immediate consequence of the so-called spectral theorem [Hal63]. \( \square \)

A first consequence of this theorem is a statement which we asserted without proof in Section 2. Namely, if a simple adjacency matrix satisfies \( \| M \| < 2 \) then all of its entries must be 0 or 1. Suppose for contradiction this were not the
case. That is, we had a simple adjacency matrix $M$ with $\|M\| < 2$ and its $(j,k)$ entry was some integer $n \geq 2$. Letting $z$ denote the vector whose $j$ entry is 1 and all of whose other entries is 0, we find that

$$\|Mz\|/\|z\| = n/1 = n.$$  

In particular, Proposition 4 implies that $\|M\| \geq n \geq 2$. This is a contradiction, so we conclude the desired result.

We now move on to the Frobenius-Perron theorem. While seemingly innocuous, it is extremely useful for modeling all sorts phenomena across pure and applied mathematics [PSC05].

**Theorem 4 (Frobenius-Perron).** Let $M$ be an $n \times n$ matrix with non-negative entries. Among the eigenvalues of $M$ with largest absolute value, one of them is real. That is, there is an eigenvector of $M$ with eigenvalue $\|M\|$.

Moreover, there exists a non-negative eigenvector with eigenvalue $\|M\|$. That is, there is an eigenvector $z = (z_1, \ldots, z_n)$ with $z_k \in \mathbb{R}_{\geq 0}$ whose eigenvalue is $\|M\|$.

Moreover, $\|M\|$ is the only eigenvalue for which there exists a non-negative eigenvector. That is, if $z$ is an eigenvector with non-negative components, then its eigenvalue must be $\|M\|$.

We now demonstrate the utility of the Frobenius-Perron theorem by demonstrating the lemma claimed at the beginning of the section:

**Lemma 1.** Let $\Gamma_2$ be a simple graph, and let $\Gamma_1$ be a subgraph. That is, $\Gamma_1$ is a subset of the edges and vertices of $\Gamma_2$. It holds that $\|M_2\| \geq \|M_1\|$, where $M_2, M_1$ are the adjacency matrices of $\Gamma_2, \Gamma_1$.

**Proof.** Adding vertices to a graph does not change its spectral radius, by Proposition 3. Hence, we can restrict to the case that $M_1$ is obtained by removing some edges from $M_2$. Let $z_1$ be a Frobenius-Perron eigenvector for $M_1$, that is, an eigenvector with non-negative entries. By construction, all of the coefficients of $M_2$ are greater than or equal to all of the corresponding coefficients in $M_1$. Hence, all of the entries of $M_2z_1$ are greater than or equal to all of the corresponding entries of $M_1z_1$. Thus,

$$\|M_2\| \geq \|M_2z_1\|/\|z_1\| \geq \|M_1z_1\|/\|z_1\| = \|M_1\|$$  

as desired. 

This completes our linear algebra background. We now have all of the tools at our disposal to prove the theorem.

### 3.2 Body of the proof

We now demonstrate Theorem 1. This will follow by constructing graphs of spectral radius $\geq 2$, concluding that graphs of spectral radius $< 2$ cannot contain them as subgraphs, and repeating until the only possibilities left are the $A_\ell, D_\ell$ and $E_\ell$. 

Before giving any of the constructions, we give some more detail on how the Frobenius-Perron theorem is useful for computing spectral radii. The first step is to observe that given a graph \( \Gamma \), column vectors are constructed by assigning a complex number to each column. Columns are labeled by vertices of \( \Gamma \), and hence specifying a vector amounts to specifying a complex number to each vertex. To compute the spectral radius of graph, all we have to do thus is the following. First, we assign positive real numbers to each vertex. Then, we prove that the vector specified is an eigenvector, and we compute its eigenvalue. The Frobenius-Perron theorem assures us that this computed eigenvalue is the spectral radius.

To begin, we prove a proposition which allows us to easily multiply vectors by adjacency matrices:

**Proposition 5.** Let \( \Gamma \) be a graph, with set \( V \) of vertices and adjacency matrix \( M_\Gamma \). Given any vector \( z = (z_v)_{v \in V} \), we have that

\[
M_\Gamma z = \left( \sum_{w \text{ connected to } v} z_w \right)_{v \in V}.
\]

That is, multiplying by the adjacency matrix has the effect of replacing the value at every vertex by the sum of the values at all of its neighboring vertices.

**Proof.** Given \( v, w \in V \), denote by \( e_{v,w} \) to quantity which is 1 if there is an edge between \( v \) and \( w \) and 0 otherwise. These are exactly the coefficients of the adjacency matrix. Standard multiplication rules tell us that

\[
M_\Gamma z = \left( \sum_{w \in V} e_{v,w} \cdot z_w \right)_{v \in V}.
\]

Thus, the \( v \) entry of \( M_\Gamma z \) is exactly the sum of the \( w \) entries for every \( w \) connected to \( v \), as claimed. \( \square \)

Our first construction demonstrates the procedure very nicely, and shows that graphs of spectral radius < 2 cannot contain any loops:

**Lemma 2.** Define the graphs

\[
A^{(1)}_\ell = \begin{bmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1
\end{bmatrix}
\]

where the 1s on vertices specify non-negative vectors. These vectors are eigenvectors for the \( A^{(1)}_\ell \) with eigenvalue 2. In particular, simple adjacency matrices with spectral radius < 2 cannot contain any loops.
Proof. By Proposition 5, multiplying by the adjacency matrix has the effect of replacing the value at a vertex by the sum of its neighbors. Since every vertex has two neighbors which are both labeled by 1, every vertex will be replaced by 2. Hence, the all 1s vector will be sent to the all 2s vector. Thus, the all 1s vector is an eigenvector with eigenvalue 2 as claimed.

Our second construction demonstrates that graphs of spectral radius \(< 2\) cannot contain points of order four or more (i.e. vertices with \(\geq 4\) edges leaving them), and that they can contain at most one point of order 3:

**Lemma 3.** Define the graphs

\[
D^{(1)}_\ell = \begin{array}{c}
\circ & 2 & 2 & \cdots & 2 & 1 \\
1 & \circ & \cdots & \circ & 1
\end{array} \quad (\ell \geq 4).
\]

The non-negative values at the vertices of \(D^{(1)}_\ell\) specify eigenvectors with eigenvalue 2. In particular, simple adjacency matrices with spectral radius \(< 2\) cannot contain any points of order 4 (since then they would contain a copy of \(D^{(1)}_4\)) and they cannot contain more than one point of order 3 (since then they would contain a copy for \(D^{(1)}_\ell\) for \(\ell \geq 5\)).

Proof. All of the center vertices have two neighbors, both of which are labeled by 2. Hence their original value of 2 gets changed to the new value of 4 after being multiplied by the adjacency matrix by Proposition 5. The vertices on the edge of the central strip have two neighbors labeled by 1 and one neighbor labeled by 2, and hence they get sent to 4 as well. The vertices on the boundary have one neighbor labeled by 2, and hence their original value of 1 gets sent to 2.

All of values at the vertices thus are multiplied by 2, so the specified vector is an eigenvector with eigenvalue 2 as claimed.

We have already drastically reduced our search. Irrducible simple adjacency matrices cannot have loops, cannot have points of order \(\geq 4\), and have at most one point of order 3. That is, they must be of the form

\[
T_{a,b,c} = \begin{array}{c}
\circ & \cdots & \circ \\
\circ & \cdots & \circ & \cdots & \circ \\
\circ & \cdots & \circ & \cdots & \circ \\
\circ & \cdots & \circ & \cdots & \circ
\end{array}
\]

for some \(a, b, c \geq 0\). If any of the values \(a, b, c\) are equal to 0, then \(T_{a,b,c}\) will be isomorphic to a straight line, and hence be isomorphic to \(A_\ell\) for some \(\ell \geq 1\). Thus, we are left with examining the case that \(a, b, c \geq 1\). If two of the values
a, b, c are equal to 1, then we are in the case $D_\ell$ for some $\ell \geq 4$. Thus, we are left with the case that $a, b, c \geq 1$, and at most one of the $a, b, c$ are equal to 1. With one more computation, we are already in a position to conclude that there are finitely many exceptional irreducible simple adjacency matrices:

Lemma 4. Define the graph

$$E_8^{(1)} = \begin{array}{c}
\circ & 2 & 4 & 6 & 5 & 4 & 3 & 2 & 1 \\
\circ & 3 \end{array}$$

The non-negative values at the vertices of $E_8^{(1)}$ specify an eigenvector with eigenvalue 2. In particular, if $a \geq 1$, $b \geq 2$, and $c \geq 2$, then if any one of $a, b$ or $c$ is $\geq 5$ then the graph $T_{a,b,c}$ has spectral radius $\geq 2$. In particular, there are finitely many triples $(a, b, c)$ with $a \geq 1$, $b \geq 2$, and $c \geq 2$ such that $T_{a,b,c}$ has spectral radius $< 2$.

Proof. Using Proposition 5 to compute the action of the adjacency matrix, this follows by a straightforward computation.

We are thus essentially done with our classification. Every irreducible simple adjacency matrix will be of the form $A_\ell$, $D_\ell$, or $T_{a,b,c}$ for one of the finitely many triples $(a, b, c)$ with $a \geq 1$, $b \geq 2$, $c \geq 2$, $a, b, c, \leq 4$. Computing with the given eigenvector, we find that the graph

$$E_6^{(1)} = \begin{array}{c}
\circ & 1 & 2 & 3 & 2 & 1 \\
\circ & 2 \\
\circ & 1 \end{array}$$

has spectral radius 2, and hence $a = 1$. Next,

$$E_7^{(1)} = \begin{array}{c}
\circ & 1 & 2 & 3 & 4 & 3 & 2 & 1 \\
\circ & 2 \end{array}$$

has spectral radius 2 as well, and hence either $b = 2$ or $c = 2$. Thus, the only possibilities are $T_{1,2,2} = E_6$, $T_{1,2,3} = E_7$, or $T_{1,2,4} = E_8$. We are now done with the proof of Theorem 1.

A Extension to non-symmetric matrices

In this section we discuss an extension of this classification to non-symmetric matrices. In fact, with this symmetry condition removed, we can even extend
the classification to non-square matrices. The spectral radius is no longer defined in such generality since non-square matrices do not have eigenvectors. Given an $n$ by $m$ matrix $M$, the quantity

$$\max_{z \in \mathbb{C}^n} \frac{\|Mz\|}{\|z\|}$$

is still defined, however. By Proposition 4 this will agree with the spectral radius in the real symmetric case. Thus, we now define $\|M\| = \max_{z \in \mathbb{C}^n} \frac{\|Mz\|}{\|z\|}$. We will classify matrices with non-negative integer coefficients such that $\|M\| < 2$.

It is important to note that if $M$ is a non-real or non-symmetric the spectral radius of $M$ is still well defined but it may not be equal to $\|M\|$. Hence, this new definition of $\|M\|$ disagrees with our old one in some cases, though this should not cause any confusion because all of our matrices before were real symmetric. We refer to $\|M\|$ as the matrix 2-norm.

The key insight is that given any matrix $M$, one can create a symmetric matrix with zeros along the diagonal by the formula

$$M_{sq} = \begin{bmatrix} 0 & M \\ M^T & 0 \end{bmatrix}.$$ 

If $M$ is an an $n$ by $m$ matrix then $M_{sq}$ is an $n + m$ by $n + m$ matrix. If $M$ has real entries then so will $M_{sq}$. If we know that $\|M\| < 2$, then the below proposition allows us to conclude that $\|M_{sq}\| < 2$. This is very useful, since it will allow us to apply our classification theorem for symmetric matrices with zeros along the diagonal.

**Proposition 6.** Let $M$ be a matrix with real coefficients, and let $M_{sq}$ be as above. It holds that $\|M\| = \|M_{sq}\|$.

**Proof.** By our same argument about block matrices in Proposition 3, we find that $\|M_{sq}\| = \max\{\|M\|, \|M^T\|\}$. Now, it is standard to see that $\|M\| = \|M^T\|$ so our proof is complete. □

The classification theorem for simple adjacency matrices is not enough to deduce the general case yet. The issue is that multiple different matrices $M$ could give equivalent irreducible matrices $M_{sq}$. Hence, we introduce some extra information on the graph of $M_{sq}$ which allows us to distinguish these different origins. This extra information is a bicoloration.

To any matrix $M$ all of whose coefficients are 1 or 0, we associate a bicolored graph $\Gamma_M$ as follows. There are black vertices for every column of $M$ and white vertices for every row of $M$. Given a row index $i$ and a column index $j$, we put an edge connecting the black vertex corresponding to $i$ and the white vertex corresponding to $j$ if the $(i,j)$th entry of $M$ is 1. We give an example below:
It is simple to check that the bicolored graph attached to $M$ and the uncolored graph attached to $M_{sq}$ have the same underlying set of vertices and edges.

As with any classification theorem, we now have to introduce the correct notion of equivalence. Our previous notion will not suffice because, given an $n$ by $m$ matrix $M$ and an $n$ by $n$ permutation matrix $\Omega$, the dimensions don’t line up for the matrix multiplication $\Omega M \Omega^{-1}$ to go through. Instead, we have to choose an $n$ by $n$ permutation matrix for the left and an $m$ by $m$ permutation matrix for the right. Formally, we say that two $n$ by $m$ matrices $M, N$ are pseudo-equivalent if there exists an $n$ by $n$ permutation matrix $\Omega_n$ and an $m$ by $m$ permutation matrix $\Omega_m$ such that

$$M = \Omega_n N \Omega_m.$$ 

We now give a graph theoretic interpretation of pseudo-equivalence:

**Proposition 7.** Let $M, N$ be matrices of the same dimensions with all coefficients $0, 1$. $M$ and $N$ are pseudo-equivalent if and only if the associated graphs $\Gamma_M, \Gamma_N$ are isomorphic as bicolored graphs. That is, there is a way of mapping the black vertices of $\Gamma_M$ onto the black vertices of $\Gamma_N$ and the white vertices of $\Gamma_M$ onto the white vertices of $\Gamma_N$ such that all of edges agree.

**Proof.** Suppose we have a pseudo-equivalence $M = \Omega_n N \Omega_m$. The underlying permutation $\omega_n$ of $\Omega_n$ specifies a way of permuting the black vertices of $M$ onto the black vertices of $N$, and the underlying permutation $\omega_m$ of $\Omega_m$ specifying a way of permuting the white vertices of $M$ onto the white vertices of $N$. The fact that $M = \Omega_n N \Omega_m$ guarantees that edges will line up, and we get a well defined isomorphism of bicolored graphs. Conversely, the permutations of black/white vertices given by an isomorphism of bicolored graphs give exactly the right data to specify a pseudo-isomorphism of matrices so our proof is complete.

We call a matrix with all coefficients $0, 1$ pseudo-irriducible if its associated graph is connected. This again has an interpretation on the level of matrices and block decompositions: we leave the exact statement as an exercise to the reader.

With all of these definitions out of the way, we can state our classification theorem.

**Theorem 5.** Let $M$ be a pseudo-irriducible matrix with non-negative integer coefficients such that $\|M\| < 2$. $M$ is pseudo-equivalent to one of the graphs $A_\ell$ for $\ell \geq 1$, $D_\ell$ for $\ell \geq 4$, or $E_\ell$ for $\ell = 6, 7, 8$, equipped with a bicoloration.

**Proof.** Our conditions imply that $\|M_{sq}\|$ satisfies the conditions for the classification of Theorem 1. Hence, the graph corresponding to $M$ will be isomorphic (as an uncolored graph) to one of the $A_\ell, D_\ell$, or $E_\ell$. This means that the graph corresponding to $M$ will be isomorphic as a bicolored graph to some bicoloration of these, completing the proof.
Some might find the above theorem unsatisfying. It pushes the question back to classifying the possible bicolorations on the $A_\ell, D_\ell, E_\ell$. However, there is a crucial observation to be made: on a given connected graph, there are at most two bicolorations. Upon coloring the first vertex black or white, all of the vertices connecting to it must have opposite color since there cannot be edges between two vertices of the same color. Pushing our way through the graph, this shows that coloring a single vertex forces the colors of all the other vertices.

There is still the issue that a-priori different bicolorations on a graph could happen to be isomorphic as bicolored graphs. For instance, the two colorings

\[
\begin{array}{c}
☺
\
☺
\end{array} , \quad
\begin{array}{c}
☺
\
☺
\end{array}
\]

on $A_2$ are isomorphic as bicolored graphs. In fact, we see that $A_\ell$ will always have a unique bicoloration when $\ell$ is even. When $\ell$ is odd, there are two non-isomorphic bicolorations based on whether the endpoints of the segment are black or white. All in all, we find that there are now five infinite families of pseudo-irrducible matrices with matrix norm less than 2:

1. $A_\ell$ for $\ell \geq 2$ even, equipped with its unique bicoloration,
2. $A_\ell$ for $\ell \geq 1$ odd, bicolored so it’s endpoints are black,
3. $A_\ell$ for $\ell \geq 1$ odd, bicolored so it’s endpoints are white,
4. $D_\ell$ for $\ell \geq 4$, bicolored so it’s two-pronged side is white,
5. $D_\ell$ for $\ell \geq 4$, bicolored so it’s two-pronged side is black.

Each of the $E_6, E_7, E_8$ have two non-isomorphic bicolorations, and hence there are six exceptional pseudo-equivalence classes of matrices with non-negative coefficients and matrix norm less than 2. This completes the explicit description of the classification

\section*{B Computing spectral radii of $A_\ell$, $D_\ell$, and $E_\ell$}

While Theorem 1 as stated has been proved, one key subtlety has been overlooked: we did not prove that $A_\ell, D_\ell$, and $E_\ell$ actually have spectral radius less than 2! Perhaps the classification is even smaller than what we gave. We compute the A-D-E spectral radii now, to show they are all $< 2$. Note that by these computations we will also arrive at a proof of Theorem 2, for which we asserted that the spectral radii of the A-D-E graphs are all of the form $2 \cos(\pi/n)$ for some $n \geq 3$.

\textbf{Proposition 8.} For any $\ell \geq 2$, the quantities
specify a non-negative eigenvector for $A_\ell$. These eigenvectors have eigenvalue $2 \cos(\pi/(\ell + 1))$. Hence, $A_\ell$ has spectral radius $2 \cos(\pi/(\ell + 1))$ for all $\ell \geq 2$.

**Proof.** Using Proposition 5 to compute the action on the adjacency matrix, we find that the entry labeled $\sin\left(\frac{k\pi}{\ell+1}\right)$ will get sent to

$$
\sin\left(\frac{(k-1)\pi}{\ell+1}\right) + \sin\left(\frac{(k+1)\pi}{\ell+1}\right) = 2 \cos\left(\frac{\pi}{\ell+1}\right) \sin\left(\frac{k\pi}{\ell+1}\right),
$$

where the above equality uses the angle addition formula for sine. This is exactly the statement that the specified vector is an eigenvector with eigenvalue $\cos\left(\frac{\pi}{\ell+1}\right)$, so we are done.

Next, we move on to the infinite family $D_\ell$:

**Proposition 9.** For any $\ell \geq 4$, the quantities

specify a non-negative eigenvector for $D_\ell$. These eigenvectors have eigenvalue $2 \cos\left(\frac{\pi}{2\ell}\right)$. Hence, $D_\ell$ has spectral radius $2 \cos\left(\frac{\pi}{2\ell}\right)$ for all $\ell \geq 4$.

**Proof.** This follows from a straightforward use of trigonometric identities.

Finally, we compute the spectral radii of the three exceptional graphs:

**Proposition 10.** The spectral radii of $E_6$ is $2 \cos\left(\frac{\pi}{12}\right)$, the spectral radius of $E_7$ is $2 \cos\left(\frac{\pi}{18}\right)$, and the spectral radius of $E_8$ is $2 \cos\left(\frac{\pi}{30}\right)$.

**Proof.** Frobenius-Perron eigenvalues can be given, but they are very big. It is important to note that for a given graph, computing the spectral radius is not a guess and check process. By factoring the characteristic polynomial one can deterministically find all of the eigenvalues, and then take the maximum. Doing this for the characteristic polynomials of $E_6, E_7, E_8$ we arrive at the spectral radii given.

This completes our calculations.
References


