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 matricies 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 extremly 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].

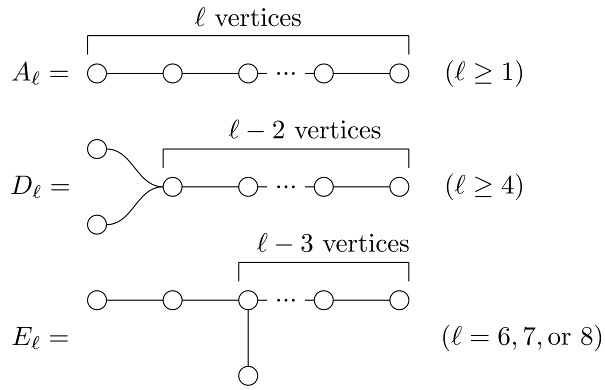
The bulk of the finite simple groups are the so-called finite simple groups of Lie type. They come from finite dimensional simple Lie algebras, objects which themselves admit a classification theorem. The classification of finite dimensional simple Lie algebras is much simpler, but it still requires 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 matricies 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 probelm - we are proving a theorem about non-negative integer matricies. 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, matricies 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 isomorphim, we only classify matricies up to a specified notion of equivalence. We then define a simple way of breaking down matricies into "irriducible" chunks. We then find that irriducible matricies, up to equivalence, are in one to one correspondance with graphs of the following type:



The connection between graphs and matricies is that every graph can be given an *adjacency matrix*. The rows and collumns are labeled by verticies. 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 classied 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 matricies 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 matricies by introducing bicolorations.

In Appendix B, we compute the size of the largest eigenvalue of the adjacency matricies 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 neccecary 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 matricies 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.

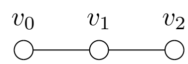
Note in particular that the condition (2) implies that M is square matrix, so conditions (3)-(4) are well defined.

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 verticies with edges between them. We will restrict our attention a nice family of graphs we call *simple graphs*. A simple graph is a graph Γ satisfying the following conditions:

1. Γ has no repeating edges between verticies,
2. Γ has undirected edges,
3. Γ has no edges going from verticies to themselves.

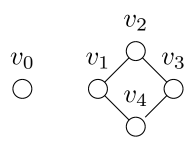
Given a simple graph Γ, we define the adjacency matrix of Γ to be the matrix whose rows and collumns are labeled by verticies, and whose (v, w) entry is 1 if there is an edge between v and w and zero otherwise. We give some examples below:

$\Gamma =$



$$, \quad \begin{matrix}\text{adjacency}\\ \text{matrix}\end{matrix} = \begin{array}{cc} & \begin{matrix} v_0 & v_1 & v_2 \end{matrix} \\ \begin{matrix} v_0 \\ v_1 \\ v_2 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

$\Gamma =$



$$, \quad \begin{matrix}\text{adjacency}\\ \text{matrix}\end{matrix} = \begin{array}{cc} & \begin{matrix} v_0 & v_1 & v_2 & v_3 & v_4 \end{matrix} \\ \begin{matrix} v_0 \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{matrix} & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

The condition (2) on graphs is equivalent to the condition (2) on matricies. 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 matricies. 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 unnececary to imply condition (1) on matricies. 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 matricies 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 ≥ 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 refering to matricies 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 *simple adjacency matrix*. These are exactly the matricies which appear as the adjacency matricies of simple graphs. The absolute value of the largest eigenvalue of a matrix is known as its *spectral radius*[1]. In this new lingo, we are in the buisness of classifying simple adjacency matricies of spectral radius less than 2.

[1] The spectrum of a matrix refers to its set of eigenvalues, hence the terminology.

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 matricies, we define equivalence as follows. A *permutation* of a set is a bijective map from that set to itself. Every permutation $\omega : [1...n] \to [1...n]$ of the first n natural numbers has an associated permutation matrix Ω. 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 collumns, and applying a permutation matrix on the left has the effect of permuting the rows.

We say that two simple adjacency matricies M and N are equivalent if there exists a permutation matrix Ω 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 matricies with associated graphs Γ_M, Γ_N. That is, the verticies of Γ_M and Γ_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 matricies M, N are equivalent if and only if there is a graph isomorphism between Γ_M and Γ_N. That is, if and only if there is a way of moving verticies of Γ_M onto the verticies of Γ_N so that the edges agree.

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

Conversely, if $M = \Omega N \Omega^{-1}$ for some permutation matrix Ω then the underlying permutation ω defines a graph isomorphism Γ_M to Γ_N in the obvious fashion. Hence, the proof is complete. □

We now turn to the second step of any classification theorem: defining a notion irriducibility. 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 subtelty. 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 neccecarily have that $G \cong N \times G/N$. Non-trivial extensions of of N and G/N are very complicated. Thus, we can say we have classifited finite simple groups but we can not say we have classified finite simple groups.[2]

[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 *reducibe* if there exist simple adjacency matricies 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 irriducible if it is not reducible. Again, we can verify that this is the "correct" notion of irriducibility by checking that is has a nice graph-theoretical analogue:

Proposition 2. *Let M be a simple adjacency matrix with associated graph Γ_M. The matrix M is irreducible if and only if Γ_M is connected. That is, if and only if one can go from every vertex on Γ_M to every other vertex on Γ_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 verticies corresponding to the rows containing M_0 and the verticies corresponing the the rows containing M_1 form disconencted components in Γ_M. Conversely, any disconnected graph can be split into two disconnected components. The adjacency matricies 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

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

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

$$M \sim \begin{bmatrix} 0 & 0 & ... & M'_{n'} \\ ... & ... & ... & ... \\ 0 & M'_2 & ... & 0 \\ M'_1 & 0 & ... & 0 \end{bmatrix}$$

with $M'_1 ... M'_{n'}$ irriducible, then

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

That is, a block decompositions of simple adjacency matricies 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 Γ_M. That is, the subgraphs of Γ_M consisting of clusters of verticies that can be reached from one to another by walking along edges. Let $M_1...M_n$ be the corresponding adjacency matricies of $\Gamma_1...\Gamma_n$. We have that

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

Conversely, every decomposition of M induces a decomposition of Γ_M into connected components. Clearly every way of breaking up Γ_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 matricies which decompose into smaller simple adjacency matricies, but do not decompose into smaller adjacency matricies of spectral 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 matricies with spectral radius less than 2:

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

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

into (not neccecarily irriducible) simple adjacency matricies $M_1...M_n$. It holds that

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

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

Proof. Every vector $\mathbf{z}$ has a unique decomposition $\mathbf{z} = \sum_{k=0}^{n} \mathbf{z}_k$, where $\mathbf{z}_k$ has non-zero elements only on the collumns corresponding to M_k. M acts by M_k on each $\mathbf{z}_k$. Hence, for $\mathbf{z}$ to be an eigenvector for M each $\mathbf{z}_k$ 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_ks. 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 neccecary 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 irriducible simple adjacency matrix of spectral radius less than 2. M is equivalent to one of the graphs A_ℓ for $\ell \geq 1$, D_ℓ for $\ell \geq 4$, or E_ℓ for $\ell = 6, 7,$ or 8, where A_ℓ, D_ℓ, E_ℓ are the graphs pictured in the introduction.*

In particular, we find that there are infinitely many irriducible simple adjacency matricies. The vast majority of irriducible simple adjacency matricies will be part of one of the two infinite families A_ℓ, D_ℓ, 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 matricies. All one has to do is first prove your result for irriducible simple adjacency matricies, and then show that the statement is preserved as you put matricies together. We demonstrate with an example:

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

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

This is a very surprising theorem. One starts with a seemingly innocuous situation (symmetric non-negative integer matricies 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 matricies. In particular, we compute $\|M\|$ for every M in the classification theorem of irriducible simple adjacency matricies. Then, one uses Proposition 3 to find that the spectral radius of any simple adjacency matrix must be the maximum along its irriducible 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 θ, we can write all of the roots of $f(x)$ has $2\cos(\pi\theta_1), 2\cos(\pi\theta_2)...2\cos(\pi\theta_d)$ 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:

$$\begin{aligned} g(x) &= x^n \prod_{k=1}^{d} (x + 1/x - 2\cos(\pi\theta_k)) \\ &= \prod_{k=1}^{d} (x^2 - 2\cos(\pi\theta_k)x + 1) \\ &= \prod_{k=1}^{d} \left(x - e^{\pi i\theta_k}\right)\left(x + e^{-\pi i\theta_k}\right). \end{aligned}$$

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 θ_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 $2\cos(k\pi/n)$ 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. $\square$

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 propety (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 $\|M\| >$ 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 paramaters 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 matricies better, mainly through the Frobenius-Perron theorem. With this in hand we will be able to easily compute the spectral radius of matricies. This Frobenius-Perron theory will also allow us to prove a key lemma: if one graph Γ_2 contains another graph Γ_1 as a subgraph, then the spectral radius of Γ_2 (i.e. the spectral radius of its adjacency matrix) will be greater than the spectral radius of its subgraph Γ_1. In particular, if a graph has spectral radius < 2 then it cannot contain any subgraphs of spectral radius ≥ 2.

The rest of the proof is now a fun game. By constructing more and more graphs with spectral radius ≥ 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 neccecary 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,*

$$\|M\| = \max_{\mathbf{z}\in\mathbb{C}^n} \|M\mathbf{z}\|/\|\mathbf{z}\|$$

where

$$\|\mathbf{z}\| = \sqrt{\sum_{k=1}^{n} |z_k|^2}$$

denotes the Euclidean norm of a vector $\mathbf{z} = (z_1 ... z_k)$.

Proof. Let $\mathbf{z}_0$ be an eigenvector with maximal absolute value for M. That is, a vector for which $M\mathbf{z}_0 = \lambda\mathbf{z}_0$, $|\lambda| = \|M\|$. Then,

$$\|M\mathbf{z}_0\|/\|\mathbf{z}_0\| = \|\lambda\mathbf{z}_0\|/\|\mathbf{z}_0\| = \|M\|.$$

Hence, $\|M\| \leq \max_{\mathbf{z}\in\mathbb{C}^n} \|M\mathbf{z}\|/\|\mathbf{z}\|$. The innequality the other direction is an immediate consequence of the so-called spectral theorem [Hal63]. □

A first consequence of this theorem is a statement which we asserted without proof is 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 $\mathbf{z}$ denote the vector whose j entry is 1 and all of whose other entries is 0, we find that

$$\|M\mathbf{z}\|/\|\mathbf{z}\| = n/1 = n.$$

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 phenomina accross pure and applied mathematics [PSC05].

Theorem 4 (Frobenius-Perron)**.** *Let M be an n by 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 $\mathbf{z} = (z_1...z_n)$ 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 $\mathbf{z}$ 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 Γ_2 be a simple graph, and let Γ_1 be a subgraph. That is, Γ_1 is a subset of the edges and verticies of Γ_2. It holds that $\|M_2\| \geq \|M_1\|$, where M_2, M_1 are the adjacency matricies of Γ_2, Γ_1.*

Proof. Adding verticies 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 $\mathbf{z}_1$ 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 $M_2\mathbf{z}_1$ are greater than or equal to all of the corresponding entries of $M_1\mathbf{z}_1$. Thus,

$$\|M_2\| \geq \|M_2\mathbf{z}_1\|/\|\mathbf{z}_1\| \geq \|M_1\mathbf{z}_1\|/\|\mathbf{z}_1\| = \|M_1\|$$

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 ≥ 2, concluding that graphs of spectral radius < 2 cannot contain them as subgraphs, and repeating until the only possibilities left are the A_ℓ, D_ℓ and E_ℓ.

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 Γ, collumn vectors are constructed by assigning a complex number to each collumn. Collumns are labeled by verticies of Γ, 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 matricies:

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

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

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 verticies.

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

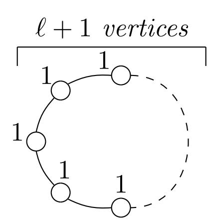
$$M_\Gamma \mathbf{z} = \left(\sum_{w \in V} e_{v,w} \cdot z_w \right)_{v \in V}.$$

Thus, the v entry of $M_\Gamma \mathbf{z}$ is exactly the sum of the w entries for every w connected to v, as claimed. □

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*

$A_\ell^{(1)} =$



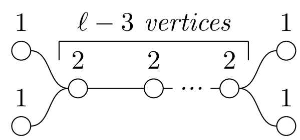
$(\ell \geq 2)$

where the 1s on verticies specify non-negative vectors. These vectors are eigenvectors for the $A_\ell^{(1)}$ with eigenvalue 2. In particular, simple adjacency matricies 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 labled 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. $\square$

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

Lemma 3. *Define the graphs*

$D_\ell^{(1)} =$ 

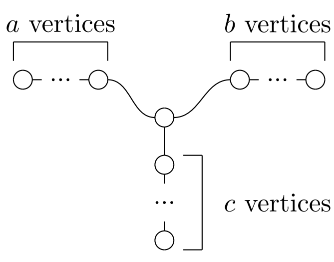
$(\ell \geq 4)$.

The non-negative values at the verticies of $D_\ell^{(1)}$ specify eigenvectors with eigenvalue 2. In particular, simple adjacency matricies with spectral radius < 2 cannot contain any points of order 4 (since then they would contain a copy of $D_4^{(1)}$) and they cannot contain more than one point of order 3 (since then they would contain a copy for $D_\ell^{(1)}$ for $\ell \geq 5$).

Proof. All of the center verticies 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 verticies on the edge of the central strip have two neighbors labeled by 1 and one neighbor labled by 2, and hence they get sent to 4 as well. The verticies on the boundary have one neighbor labled by 2, and hence their original value of 1 gets sent to 2.

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

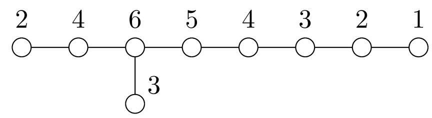
We have already drastically reduced our search. Irriducible simple adjacency matricies cannot have loops, cannot have points of order ≥ 4, and have at most one point of order 3. That is, they must be of the form

$T_{a,b,c} =$ 

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_ℓ 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_ℓ 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 irriducible simple adjacency matricies:

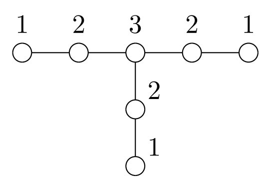
Lemma 4. *Define the graph*

$E_8^{(1)} =$ 

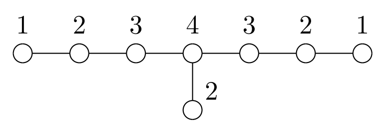
The non-negative values at the verticies of $E_8^{(1)}$ specify an eigenvector with eigenvalue 2. In particular, if $a \geq 1$, $b \geq 2$, and $c \geq 2$, then if any one of a, b or c is ≥ 5 then the graph $T_{a,b,c}$ has spectral radius ≥ 2. In particular, there are finitely many triples (a, b, c) 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 irriducible simple adjacency matrix will be of the form A_ℓ, D_ℓ, 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

$E_6^{(1)} =$ 

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

$E_7^{(1)} =$ 

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 matricies

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

$$\max_{\mathbf{z}\in\mathbb{C}^n} \|M\mathbf{z}\|/\|\mathbf{z}\|$$

is still defined, however. By Proposition 4 this will agree with the spectral radius in the real symmetric case. Thus, we now *define* $\|M\| = \max_{\mathbf{z}\in\mathbb{C}^n} \|M\mathbf{z}\|/\|\mathbf{z}\|$. We will classify matricies 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 matricies 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

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

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 $\|M_{\mathrm{sq}}\| < 2$. This is very useful, since it will allow us to apply our classification theorem for symmetric matricies with zeros along the diagonal.

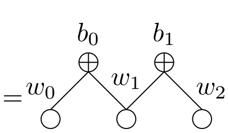
Proposition 6. *Let M be a matrix with real coefficients, and let M_{sq} be as above. It holds that $\|M\| = \|M_{sq}\|$.*

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

The classification theorem for simple adjacency matricies is not enough to deduce the general case yet. The issue is that multiple different matricies M could give equivalent irriducible matricies 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 Γ_M as follows. There are black verticies for every collumn of M and white verticies for every row of M. Given a row index i and a collumn 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:

$$M = \begin{array}{c c} & \begin{array}{cc} b_0 & b_1 \end{array} \\ \begin{array}{c} w_0 \\ w_1 \\ w_2 \end{array} & \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \end{array} \quad , \quad \text{associated graph} =$$



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 verticies 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 Ω, 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 matricies M, N are pseudo-equivalent if there exists an n by n permutation matrix Ω_n and an m by m permutation matrix Ω_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 matricies of the same dimensions with all coefficients $0, 1$. M and N are pseudo-equivalent if and only if the associated graphs Γ_M, Γ_N are isomorphic as bicolored graphs. That is, there is a way of mapping the black verticies of Γ_M onto the black verticies of Γ_N and the white verticies of Γ_M onto the white verticies of Γ_N such that all of edges agree.*

Proof. Suppose we have a pseudo-equivalence $M = \Omega_n N \Omega_m$. The underlying permutation ω_n of Ω_n specifies a way of permuting the black verticies of M onto the black verticies of N, and the underlying permutation ω_m of Ω_m specifying a way of permuting the white verticies of M onto the white verticies 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 verticies given by an isomorphism of bicolored graphs give exactly the right data to specify a pseuo-isomorphism of matricies 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 matricies and block decompostions: 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_ℓ for $\ell \geq 1$, D_ℓ for $\ell \geq 4$, or E_ℓ for $\ell = 6, 7, 8$, equipped with a bicoloration.*

Proof. Our conditions imply that $\|M_{\text{sq}}\|$ 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_ℓ, D_ℓ, or E_ℓ. 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_ℓ, D_ℓ, E_ℓ. However, there is a crucial observation to be made: on a given connected graph, there are at most two bicolations. Upon coloring the first vertex black or white, all of the verticies connecting to it must have opposite color since there cannot be edges between two verticies of the same color. Pushing our way through the graph, this shows that coloring a single vertex forces the colors of all the other verticies.

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

$$\oplus\!\!-\!\!\!-\!\!\!-\!\!\bigcirc \quad , \qquad \bigcirc\!\!-\!\!\!-\!\!\!-\!\!\oplus$$

on A_2 are isomorphic as bicolored graphs. In fact, we see that A_ℓ will always have a unique bicoloration when ℓ is even. When ℓ 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-irriducible matricies with matrix norm less than 2:

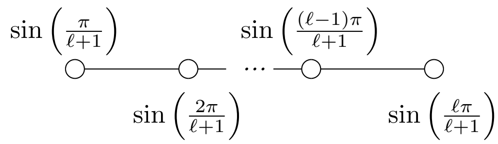
1. A_ℓ for $\ell \geq 2$ even, equipped with its unique bicoloration,
2. A_ℓ for $\ell \geq 1$ odd, bicolored so it's endpoints are black,
3. A_ℓ for $\ell \geq 1$ odd, bicolored so it's endpoints are white,
4. D_ℓ for $\ell \geq 4$, bicolored so it's two-pronged side is white,
5. D_ℓ for $\ell \geq 4$, 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 calsses of matricies with non-negative coefficients and matrix norm less than 2. This completes the explicit description of the classification

B Computing spectral radii of A_ℓ, D_ℓ, and E_ℓ

While Theorem 1 as stated has been proved, one key subtlety has been overlooked: we did not prove that A_ℓ, D_ℓ, and E_ℓ actually have spectral radius less than 2! Perhpas the classification is even smalller 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 $n \geq 3$.

Proposition 8. *For any $\ell \geq 2$, the quantites*



specify a non-negative eigenvector for A_ℓ. These eigenvectors have eigenvalue $2\cos(\pi/(\ell+1))$. Hence, A_ℓ has spectral radius $2\cos(\pi/(\ell+1))$ for all $\ell \geq 2$.

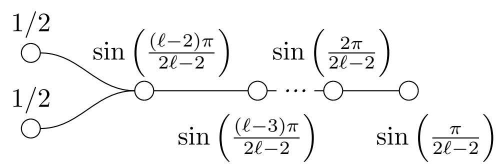
Proof. Using Proposition 5 to compute the action on the adjacency matrix, we find that the entry labled $\sin\left(\frac{k\pi}{\ell+1}\right)$ will get sent to

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

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_ℓ:

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



specify a non-negative eigenvector for D_ℓ. These eigenvectors have eigenvalue $2\cos(\pi/(2\ell))$. Hence, D_ℓ has spectral radius $2\cos(\pi/(2\ell))$ 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 $2\cos(\pi/12)$, the spectral radius of E_7 is $2\cos(\pi/18)$, and the spectral radius of E_8 is $2\cos(\pi/30)$.*

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 charactaristic 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

[Arn76] VI Arnold. Problems of present day mathematics, xvii (dynamical systems and differential equations). In *Proc. Symp. Pure Math*, volume 28, 1976.

[Asc04] Michael Aschbacher. The status of the classification of the finite simple groups. *Notices of the American Mathematical Society*, 51(7):736–740, 2004.

[Bri08] Michel Brion. Representations of quivers. 2008.

[CZ09] Andrea Cappelli and Jean-Bernard Zuber. Ade classification of conformal field theories. *arXiv preprint arXiv:0911.3242*, 2009.

[DV34] Patrick Du Val. On isolated singularities of surfaces which do not affect the conditions of adjunction (part i.). In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 30, pages 453–459. Cambridge University Press, 1934.

[FFF+12] Georg Frobenius, Ferdinand Georg Frobenius, Ferdinand Georg Frobenius, Ferdinand Georg Frobenius, and Germany Mathematician. *Über Matrizen aus nicht negativen Elementen.* Königliche Akademie der Wissenschaften Berlin, 1912.

[GdLHJ12] Frederick M Goodman, Pierre de La Harpe, and Vaughan FR Jones. *Coxeter graphs and towers of algebras*, volume 14. Springer Science & Business Media, 2012.

[GG20] Morikuni Goto and Frank D Grosshans. *Semisimple lie algebras.* CRC Press, 2020.

[Hal63] Paul R Halmos. What does the spectral theorem say? *The American Mathematical Monthly*, 70(3):241–247, 1963.

[HHSV77] M Hazewinkel, W Hesselink, D Siersma, and FD Veldkamp. The ubiquity of coxeter-dynkin diagrams. *Nieuw Arch. Wisk*, 25:257–307, 1977.

[Hum92] James E Humphreys. *Reflection groups and Coxeter groups.* Number 29. Cambridge university press, 1992.

[Kro57] Leopold Kronecker. Zwei sätze über gleichungen mit ganzzahligen coefficienten. 1857.

[PSC05] S Unnikrishna Pillai, Torsten Suel, and Seunghun Cha. The perron-frobenius theorem: some of its applications. *IEEE Signal Processing Magazine*, 22(2):62–75, 2005.

[Sie14] Kyler Siegel. The ubiquity of ade classifications in nature. 2014.

[VH02] J Van Hoboken. *Platonic solids, binary polyhedral groups, Kleinian singularities and Lie algebras of type A, D, E*. PhD thesis, University of Amsterdam Amsterdam, The Netherlands, 2002.

[Was97] Lawrence C Washington. *Introduction to cyclotomic fields*, volume 83. Springer Science & Business Media, 1997.