

Spectral graph theory: Three common spectra*

Steve Butler

September 2006

Abstract

In this first talk we will introduce three of the most commonly used types of matrices in spectral graph theory. They are the adjacency matrix, the combinatorial Laplacian, and the normalized Laplacian. We also will give some simple examples of how the spectrum can be used for each of these types.

1 Introduction

Using the spectrum of light scientists have had great success in being able to indirectly determine the compounds of chemicals that could not be directly measured. For instance, by examining the light given off by distant stars we can determine their chemical composition, even though we could never directly gather any material from those stars.

In an analogous way we can use the spectra of various matrices (i.e., the eigenvalues of the matrices) to get information about a graph that would otherwise be difficult to obtain. In this series of talks we will give some introductory comments about connections between the eigenvalues of matrices and the properties of graphs. The study of the relations between these two objects is spectral graph theory.

Thus a student of spectral graph theory not only needs to be familiar with graph theory but also must understand the basic tools of linear algebra. Eigenvalues, eigenvectors, determinants, Courant-Fischer Theorem, Perron-Frobenius, and so on are the tools of the trade. A student interested in studying spectral graph theory should spend some time learning how to use these tools, and searching for connections to graphs.

*First of three talks given at the *Center for Combinatorics*, Nankai University, Tianjin

In this first talk we will present the three most common matrices associated with graphs (namely the adjacency matrix, the combinatorial Laplacian, and the normalized Laplacian), and give some simple examples for each about how the eigenvalues can be used to give some information about the graph. In the second talk we will use the Courant-Fischer Theorem to derive some various results about the normalized Laplacian, including some interlacing inequalities. In the final talk we will look at edge expansion and in particular a Cheeger inequality and discrepancy result.

Throughout this talk we will make use of the graph in Figure 1 as a basic example, and will give the spectrum of this graph for each of the three matrices. In this talk we will be dealing with simple undirected graphs (i.e., graphs without loops or multiple edges).

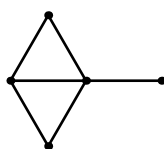


Figure 1: A simple example of a graph.

2 The adjacency matrix

Given a graph G we can form a matrix A , called the adjacency matrix, by letting the vertices index the columns and rows, and then letting

$$A_{i,j} = \begin{cases} 1 & \text{if } i \text{ is adjacent to } j; \\ 0 & \text{if } i \text{ is not adjacent to } j. \end{cases}$$

For example one adjacency matrix for the graph in Figure 1 is

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

In some sense the adjacency matrix is not unique, this is because we can relabel the vertices of the graph which would cause a simultaneous permutation of the rows and columns. So for example we could also have gotten the following matrix as an

adjacency matrix of the graph,

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

It should be noted however that the eigenvalues of the two matrices will always be the same for any relabeling. And if we tie eigenfunctions to the vertices it will also follow that eigenfunctions are independent of the choice of labeling. In particular, we have the following.

Lemma 1. *If the eigenvalues of two graphs do not match, then the graphs are not isomorphic.*

The converse is not true. That is there exists graphs which are not isomorphic but do still have the same eigenvalues (in general showing two graphs are not isomorphic is a nontrivial problem), one such example is shown in Figure 2.



Figure 2: Two nonisomorphic graphs whose adjacency matrices have eigenvalues $-2, 0, 0, 0, 2$.

For the graph in Figure 1 a calculation (for which computers are especially good at) shows that the eigenvalues for the graph are:

$$2.68554393\dots, 0.33490398\dots, 0, -1.27133037\dots, -1.74911754\dots \quad (1)$$

Comment. When dealing with the spectrum of a graph we are dealing with the set of eigenvalues. Most frequently, it is more correct to say that we are dealing with a multi-set, i.e., we allow for repetitions. So for example in the graphs given in Figure 1 we have 0 listed as an eigenvalue three times.

Comment. Note that all the eigenvalues given in (1) are real. This follows from the fact that the adjacency matrix is symmetric which in turn follows from the fact that the graph is undirected. In addition, since the adjacency matrix is symmetric the matrix has a full set of eigenvectors which are mutually orthogonal. This latter fact comes in useful in some applications. These are the main reasons that the vast majority of results in spectral graph theory deal with undirected graphs. In the third talk we will give an example of what can be done when dealing with directed graphs.

Comment. In the adjacency matrix we have used 0s and 1s. This is useful for emphasizing the discrete nature of the graph. However, more generally we can allow for other entries besides 0s and 1s. The differing entries can then be thought of as weights on the edges, and many results relating to the spectra can then be generalized to their “weighted” versions.

Estimating the number of walks of length k

We can use the eigenvalues of the adjacency matrix to count the number of walks of length k . To do this, first notice that by the rules of matrix multiplication we have

$$(A^k)_{i,j} = \sum_{i_1, i_2, \dots, i_{k-1}} A_{i, i_1} A_{i_1, i_2} \cdots A_{i_{k-2}, i_{k-1}} A_{i_{k-1}, j}.$$

Looking at the term in the summand we see that $a_{i, i_1} a_{i_1, i_2} \cdots a_{i_{k-2}, i_{k-1}} a_{i_{k-1}, j}$ will be 1 if and only if vertex i is adjacent to i_1 which is adjacent to i_2 and so on until we get to j . In other words, the term in the summand is 1 if and only if it corresponds to a walk of length k starting at vertex i and ending at vertex j . So we have the following.

Lemma 2. *Let A be the adjacency matrix of G . Then $(A^k)_{i,j}$ is the number of walks of length k starting at vertex i and ending at vertex j .*

Let us now use the lemma with the eigenvalues of the adjacency matrix. We will make use of the following two facts. First, the trace of the determinant is the sum of the eigenvalues of the matrix. Secondly, the eigenvalues of A^k are the eigenvalues of A raised to the k th power. In what follows we will let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of A .

- $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 0$. This follows by noting that the sum of the eigenvalues is the trace of the adjacency matrix which is 0 since A is 0 on the diagonal.
- $\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2 = 2E(G)$, where $E(G)$ is the number of edges of the graph. This follows by noting that the sum of the square of the eigenvalues is the same as the trace of A^2 . The diagonal entries of A^2 counts the number of closed walks of length 2 (a closed walk is a walk that starts and ends at the same vertex, since we are on the diagonal the starting and ending vertices are the same), for which each edge is counted exactly twice.
- $\lambda_1^3 + \lambda_2^3 + \cdots + \lambda_n^3 = 6T(G)$, where $T(G)$ is the number of triangles of the graph. This follows by noting that the sum of the eigenvalues cubed is the same as the trace of A^3 , i.e., the same as the number of closed walks of length 3. Each triangle will be counted exactly six times (i.e., a choice of 3 initial vertices and 2 directions for each triangle).

We leave as an exercise to the interested reader to determine what $\lambda_1^4 + \dots + \lambda_n^4$ counts.

Adding up the eigenvalues, the eigenvalues squared, and the eigenvalues cubed as given in (1) gives respectively, 0, 12, and 6 as expected from the above and the structure of the graph in Figure 1.

Next suppose that we wanted to count the total number of walks of length k . This can be achieved by considering $\mathbf{1}^T A^k \mathbf{1}$, where $\mathbf{1}$ is an all 1s vector of length n and we use T in an exponent to denote the transpose of a matrix. Since A is symmetric then, as noted above, we have a full set of (real) orthonormal eigenvectors. So let ϕ_i be the eigenvector associated with λ_i . Then for some appropriate constants a_i we have that $\mathbf{1} = \sum_i a_i \phi_i$. Putting this in for $\mathbf{1}$ we have that the total number of walks of length k is

$$\left(\sum_i a_i \phi_i^T \right) A^k \left(\sum_i a_i \phi_i \right) = \left(\sum_i a_i \phi_i^T \right) \left(\sum_i a_i \lambda_i^k \phi_i \right) = \sum_i a_i^2 \lambda_i^k.$$

Lemma 3. *Given a connected, non-bipartite graph G , the number of walks of length k (for k very large) is $\approx a_n^2 \lambda_n^k$.*

Without loss of generality we can assume that the graph has an edge, from which it easily follows that $\lambda_n > 0$. Now we have

$$\lim_{k \rightarrow \infty} \frac{\mathbf{1}^T A^k \mathbf{1}}{\lambda_n^k} = \lim_{k \rightarrow \infty} \sum_i a_i^2 \frac{\lambda_i^k}{\lambda_n^k} = a_n^2.$$

In the last step we used that $|\lambda_i| < \lambda_n$ for $i \neq n$. The latter statement is an easy consequence of the Perron-Frobenius Theorem which we will discuss in the third lecture. This shows that λ_n tells us the growth rate for the number of walks of length k in non-bipartite graphs (which is sometimes referred to as “capacity”).

Comment. Note in the above derivation that the only eigenvalue which was important was the largest eigenvalue. In spectral graph theory almost all of the focus and energy has been put into only the few largest and few lowest eigenvalues, the middle range of the spectra being usually neglected. This perhaps is an indication of how little we know how to use the spectrum to examine the graph.

3 The combinatorial Laplacian

The second type of matrix that we will consider is the combinatorial Laplacian matrix, denoted as L . In some settings this is referred to as the Laplacian. Here we have added the term combinatorial to help distinguish the two types of Laplacian matrices

that we will consider. Beyond this we will see that the combinatorial Laplacian is involved in an interesting enumeration problem (some partial justification for the term combinatorial).

Again we let the vertices index the columns and rows and define L entrywise as follows:

$$L_{i,j} = \begin{cases} d_i & \text{if } i = j; \\ -1 & \text{if } i \text{ is adjacent to } j; \\ 0 & \text{otherwise,} \end{cases}$$

where d_i is the degree of the i th vertex. This is closely related to the adjacency matrix and is sometimes written as $L = D - A$, where D is the diagonal matrix with the degree on the diagonals and A is the adjacency matrix.

For the graph in Figure 1 the combinatorial Laplacian will be

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

with eigenvalues

$$5, 4, 2, 1, 0. \tag{2}$$

Comment. The fact that all the eigenvalues are integers is coincidence and does not hold in general. However there is one special eigenvalue, namely 0. The fact that 0 is always an eigenvalue is easy to see by noting that all of the row sums are 0, i.e., $\mathbf{1}$ is an eigenvector for the eigenvalue 0. All the other eigenvalues are nonnegative, so in other words the combinatorial Laplacian is positive semi-definite. This follows from, for example, the Gersgorin Disc Theorem. We will also give another proof below using the incidence matrix below.

That $\mathbf{1}$ is an eigenvector turns out to be very useful in some applications. We will make use of a similar idea for the normalized Laplacian in the third lecture. Note for the adjacency matrix that $\mathbf{1}$ is an eigenvector if and only if the graph is regular. This is one reason why spectral results are usually first proved for regular graphs (i.e., for regular graphs we have good control on the eigenvectors).

The combinatorial Laplacian is associated with the incidence matrix. The incidence matrix, which we will denote C , has rows indexed by the vertices and columns indexed by the edges. To define the entries, given an edge $e = \{i, j\}$ then in the column corresponding to e we put in 1 and -1 for the vertices to which the edge is

incident and 0s otherwise, i.e.,

$$C = \begin{matrix} & & e \\ & & \vdots \\ i & \left(\begin{array}{ccc} \dots\dots\dots & 1 & \dots\dots\dots \\ & \vdots & \\ & \vdots & \\ j & \dots\dots\dots & -1 & \dots\dots\dots \\ & & \vdots & \end{array} \right) & \end{matrix},$$

which entry is negative turns out to be unimportant for our purposes.

The relation between C and L is that $L = CC^T$. To see this we note that $(CC^T)_{i,j}$ can be found by taking the inner product of the i th and j th row of C . It is easy to check that for $i \neq j$ that this inner product is 0 if there is no edge and -1 otherwise, while for $i = j$ we add 1 for each edge incident to i , i.e., we get d_i .

Comment. If instead of using ± 1 in C we had only used 1, the resulting matrix would be the unsigned Laplacian, also known as the quasi-Laplacian, which differs from the combinatorial Laplacian in that the off-diagonal entries are positive. This matrix has not been as extensively studied.

We can use this representation of L to now show that the all of the eigenvalues are nonnegative. Suppose that σ is an eigenvalue with a (real) normal eigenvector ϕ . Then

$$\sigma = \phi^T(\sigma\phi) = \phi^T L\phi = \phi^T CC^T\phi = (C^T\phi)^T C^T\phi = \|C^T\phi\|^2 \geq 0.$$

The Matrix Tree Theorem

One of the more interesting results related to the combinatorial Laplacian is in counting the number of spanning trees for connected graphs. A spanning tree of a graph G is, as its name implies, a subgraph of G which is a tree and is incident to all the vertices.

Theorem 4 (Matrix Tree Theorem). *Let G be a graph, and $0 = \sigma_0 \leq \sigma_1 \leq \dots \leq \sigma_{n-1}$ be the eigenvalues of the combinatorial Laplacian of G . Then the number of spanning trees of G is given by*

$$\frac{\sigma_1\sigma_2 \cdots \sigma_{n-1}}{n}.$$

As an example, taking the eigenvalues given in (2) for the graph in Figure 1 we would expect $5 \cdot 4 \cdot 2 \cdot 1/5 = 8$ spanning trees. It is easy to check that this is in the case and we show them below.

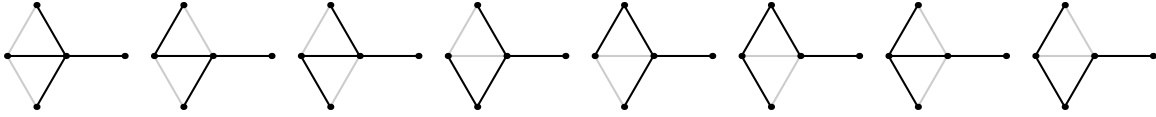


Figure 3: The eight spanning trees of the graph in Figure 1.

We will not include a proof of the Matrix Tree Theorem here but give a sketch of the ideas. The first thing we note is that the Matrix Tree Theorem is usually stated as saying that any cofactor of the matrix is the number of spanning trees. Knowing this it can be shown that the coefficient of x in the characteristic polynomial is n times the number of spanning trees (since this coefficient is found by adding the n cofactors from the diagonal). On the other hand the coefficient of x is also found by summing all possible products of $n - 1$ of the n eigenvalues. Since 0 is an eigenvalue only one of these will be nonzero and the result follows.

So now let us compute the cofactor for the i th diagonal term. This involves taking the determinant of L_0 where we have removed the i th row and column of L . It is easy to check that $L_0 = C_0 C_0^T$ where C_0 is C with the i th row removed. Using the Cauchy-Binet formula we have

$$\det(L_0) = \det(C_0 C_0^T) = \sum_{\substack{X \subseteq E \\ |X|=n-1}} \det(C_X C_X^T) = \sum_{\substack{X \subseteq E \\ |X|=n-1}} (\det(C_X))^2,$$

where C_X is the matrix formed by the columns in B_0 corresponding to X . The remainder of the proof is to show that

$$\det(C_X) = \begin{cases} 0 & \text{if the edges corresponding to } X \text{ have a cycle;} \\ \pm 1 & \text{if the edges corresponding to } X \text{ form a tree.} \end{cases}$$

From which the proof then follows. The first statement easily follows by finding a linear dependence in the columns of B_X . The second one takes more work and we will not attempt it here but can be found in several textbooks.

4 The normalized Laplacian

The final type of matrix that we will consider is the normalized Laplacian matrix, denoted \mathcal{L} . As the name suggests this is closely related to the combinatorial Laplacian that we have just looked at. For graphs with no isolated vertices the relationship is given by $\mathcal{L} = D^{-1/2} L D^{-1/2} = D^{-1/2} (D - A) D^{-1/2} = I - D^{-1/2} A D^{-1/2}$. (Throughout the rest of this section we will assume no isolated vertices since they contribute little

more than technicalities to the arguments.) Entrywise we have,

$$\mathcal{L}_{i,j} = \begin{cases} 1 & \text{if } i = j; \\ \frac{-1}{\sqrt{d_i d_j}} & \text{if } i \text{ is adjacent to } j; \\ 0 & \text{otherwise.} \end{cases}$$

For the graph in Figure 1 the normalized Laplacian will be

$$\mathcal{L} = \begin{pmatrix} 1 & \frac{-1}{2} & 0 & 0 & 0 \\ \frac{-1}{2} & 1 & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{12}} & \frac{-1}{\sqrt{8}} \\ 0 & \frac{-1}{\sqrt{8}} & 1 & \frac{-1}{\sqrt{6}} & 0 \\ 0 & \frac{-1}{\sqrt{12}} & \frac{-1}{\sqrt{6}} & 1 & \frac{-1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{8}} & 0 & \frac{-1}{\sqrt{6}} & 1 \end{pmatrix},$$

with eigenvalues

$$1.72871355\dots, 1.5, 1, 0.77128644\dots, 0. \quad (3)$$

Comment. As before we have that 0 is an eigenvalue (now with eigenvector $D^{1/2}\mathbf{1}$) and the remaining eigenvalues are nonnegative. A major difference between the two spectrums though is that while for the combinatorial Laplacian the eigenvalues can be essentially as large as desired, the normalized Laplacian has eigenvalues always lying in the range between 0 and 2 inclusive (due in great part to the normalization), we will prove this statement in the next lecture.

One advantage to this is that it makes it easier to compare the distribution of the eigenvalues for two different graphs, especially if there is a large difference in the “size” of the graphs.

Comment. Comparing the eigenvalues in (1), (2) and (3) we see that they can be quite different. In general it will make a big difference as to which spectrum is used, and some results which might hold for one spectrum will not hold for another. Knowing which spectrum to use is a combination of experience as well as trial and error.

The normalized Laplacian has connections with many interesting properties of graphs. We will see this more in the third lecture. For now we will consider the problem of random walks.

Random walks—A rambling introduction

A random walk on a graph G can be thought of as a walk where we start at a vertex on the graph and at each step of time pick randomly (in our case uniformly) one of the edges incident to the current vertex and go along that edge to the next vertex. Repeating as often as desired.

As an example, consider the problem of shuffling cards. In this setting the graph is all possible ways to arrange a deck (a large graph!) and the edges represent shuffles, i.e., starting with a deck of cards which orderings can be reached using one shuffle. In this case a random walk corresponds to doing a random sequence of shuffles.

One problem of interest for people shuffling cards is how many times do we need to shuffle until the cards are sufficiently “random”. In this setting “random” can be taken to mean as saying that knowing the initial configuration of cards before starting the shuffling will not give you any significant information about the current placement of cards (i.e., all of the initial information has been lost). The study of random walks on graphs can help answer such questions.

To work this problem we can keep track of the probability distribution of the various destinations after k steps. In other words, what is the probability that we are at some particular vertex in the graph after k steps.

Pictorially, imagine that we have a cup of full water at our initial vertex, and empty cups everywhere else (see Figure 4). At each step we will simultaneously redistribute the water at each vertex to the vertices neighbors, and thus overtime the water should diffuse throughout the graph. In our problem the water represents the probability distribution and the fractional amount of water at a vertex at k steps is the probability that we are at that vertex in k steps.

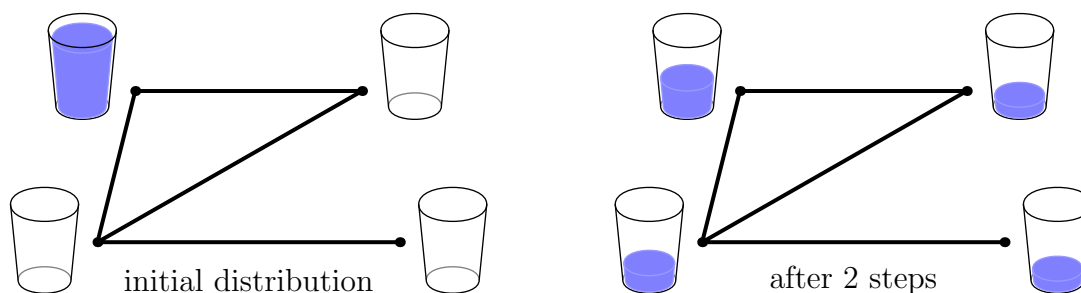


Figure 4: A “watered” down approach to random walks.

The question that we want to ask is how many steps does it take before we are sufficiently random. As a first step we should decide what is meant by random. We will say that a probability distribution is random if the probability of being at any vertex is proportional to its degree. More explicitly, we are random if the probability

of being at vertex i is equal to $d_i / \sum_{\ell \in G} d_\ell$.

Convergence of random walks—The technical approach

In what follows we will change our convention slightly in that all of the vectors will be treated as *row* vectors (so now we will be focusing on multiplication on the left hand side).

Putting the discussion of the previous section in terms of matrices, if A is the adjacency matrix, then $D^{-1}A$ is the “probability transition matrix”. That is,

$$(D^{-1}A)_{i,j} = \begin{cases} 0 & \text{if } i \text{ is not adjacent to } j; \\ \frac{1}{d_i} & \text{if } i \text{ is adjacent to } j; \end{cases}$$

is the probability that given you are at vertex i you move to vertex j . An initial probability distribution will be a vector f , and the probability distribution after k steps will be $f(D^{-1}A)^k$.

Comment. The two requirements to be a probability distribution is that all of the entries are nonnegative and sum to 1. To see that the matrix $D^{-1}A$ takes one probability distribution to another it suffices to check that these two properties are always maintained. Nonnegativity follows easily since all of the terms are nonnegative. To show that the sum of the entries of a vector g is 1 it suffices to show that $g\mathbf{1}^T = 1$, in our case we have that $fD^{-1}A\mathbf{1}^T = fD^{-1}D\mathbf{1}^T = f\mathbf{1}^T = 1$. The last step following since we started with a probability distribution.

To see the connection between the probability transition matrix and the normalized Laplacian, note that

$$D^{-1/2}(I - \mathcal{L})D^{1/2} = D^{-1/2}(D^{-1/2}AD^{-1/2})D^{1/2} = D^{-1}A,$$

showing that $D^{-1}A$ and $I - \mathcal{L}$ are similar. As a consequence if λ is an eigenvalue of \mathcal{L} then $1 - \lambda$ is an eigenvalue of \mathcal{L} . In particular, 1 is always an eigenvalue of $D^{-1}A$ (since 0 is always an eigenvalue of \mathcal{L}) and its *left* eigenvector is easily shown to be $\mathbf{1}D$. If we normalize the vector to a probability distribution the resulting vector is

$$\frac{\mathbf{1}D}{\sum_{\ell} d_{\ell}},$$

which we refer to as the stationary distribution. This is also the distribution which, from the above discussion, we want our random walk to converge to.

Now let ϕ_i be an orthonormal set of eigenvectors associated with λ_i for \mathcal{L} (here we will follow the convention that $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1}$). By the above ϕ_i is also

associated with $1 - \lambda_i$ for $D^{-1/2}AD^{-1/2}$. It is easy to check that $\phi_0 = \mathbf{1}D^{1/2}/\sqrt{\sum_\ell d_\ell}$. Finally, since we have the full set of eigenvalues and orthonormal eigenvectors, we can use the idea of projections onto eigenspaces to write

$$D^{-1/2}AD^{-1/2} = \sum_i (1 - \lambda_i) \phi_i^T \phi_i.$$

To check how close our random walk is after k steps to the stationary distribution we can use several types of measurements. Here we will choose the L^2 -norm, i.e., $\|g\| = \sqrt{\sum_i |g_i|^2}$ (different types of measurements will give different types of bounds for the rate of convergence). In particular, the L^2 -norm distance between the random walk after k steps and the stationary distribution is

$$\begin{aligned} \left\| f(D^{-1}A)^k - \frac{\mathbf{1}D}{\sum_\ell d_\ell} \right\| &= \left\| fD^{-1/2}(D^{-1/2}AD^{-1/2})^k D^{1/2} - \frac{\mathbf{1}D}{\sum_\ell d_\ell} \right\| \\ &= \left\| fD^{-1/2} \left(\sum_i (1 - \lambda_i) \phi_i^T \phi_i \right)^k D^{1/2} - \frac{\mathbf{1}D}{\sum_\ell d_\ell} \right\| \\ &= \left\| fD^{-1/2} \left(\sum_i (1 - \lambda_i)^k \phi_i^T \phi_i \right) D^{1/2} - \frac{\mathbf{1}D}{\sum_\ell d_\ell} \right\| \\ &= \left\| fD^{-1/2} \left(\sum_{i \neq 0} (1 - \lambda_i)^k \phi_i^T \phi_i \right) D^{1/2} \right\| \\ &\leq \max_{i \neq 0} |1 - \lambda_i|^k \frac{\max_i \sqrt{d_i}}{\min_j \sqrt{d_j}}. \end{aligned}$$

Going from the third to the fourth line is an easy calculation to show that the $i = 0$ term in the summand will cancel out with the other expression. In the last step we used matrix norms. In particular, we used that $\|fB\| \leq \|f\|\|B\|$ three times, where $\|f\|$ is the L^2 -norm and $\|B\|$ is the operator norm. For the case of symmetric matrices (which all three are) $\|A\|$ is the maximal absolute value of an eigenvalue, with the matrices as above it is easy to find these values and then get the bound (of course $\|f\| \leq 1$ since it is a probability distribution).

In particular, this tells us that we can use eigenvalues to get an estimate on the rate of convergence of a random walk. The more closely the eigenvalues are gathered around 1 for the normalized Laplacian the faster we should expect to converge to the stationary distribution. This allows us to give an estimate on the number of steps needed to produce random-like results.

As an example, in the graph of Figure 1 if we start at any particular vertex and take 10 steps the probability distribution to where we end up will be at most 0.08444976... away from the stationary distribution in the L^2 -norm (though we

are most likely much closer). If we wanted to be within 0.000001 of the stationary distribution then using the above results it is easy to check that it will take at most 46 steps.

Comment. Note that since the eigenvalues of the normalized Laplacian are between 0 and 2, we have $\max_{i \neq 0} |1 - \lambda_i| \leq 1$. When can this equal 1? The first possibility is that we have 0 as an eigenvalue multiple times. In the next lecture we will see that this means that the graph is not connected. The second possibility is that we have 2 as an eigenvalue. Again in the next lecture we will see that this means that the graph has a bipartite component.

Therefore, for a random walk to converge to the stationary distribution it suffices to be on a graph which is connected and not bipartite. In terms of ergodic theory (which can be thought of as the study of mixing systems) the first condition that the graph be connected is equivalent to the requirement that starting at any point in the system any other point can be reached (i.e., transitive) while the second condition that the graph not be bipartite is equivalent to the requirement that the system be aperiodic.