Lectures on
Spectral Graph Theory

Fan R. K. Chung

Author address:

University of Pennsylvania, Philadelphia, Pennsylvania 19104

E-mail address: chung@math.upenn.edu

Contents

CHAPTER 1

Eigenvalues and the Laplacian of a graph

1.1. Introduction

Spectral graph theory has a long history. In the early days, matrix theory and linear algebra were used to analyze adjacency matrices of graphs. Algebraic methods are especially effective in treating graphs which are regular and symmetric. Sometimes, certain eigenvalues have been referred to as the "algebraic connectivity" of a graph [**126**]. There is a large literature on algebraic aspects of spectral graph theory, well documented in several surveys and books, such as Biggs [**25**], Cvetković, Doob and Sachs [**90, 91**], and Seidel [**224**].

In the past ten years, many developments in spectral graph theory have often had a geometric flavor. For example, the explicit constructions of expander graphs, due to Lubotzky-Phillips-Sarnak [**193**] and Margulis [**195**], are based on eigenvalues and isoperimetric properties of graphs. The discrete analogue of the Cheeger inequality has been heavily utilized in the study of random walks and rapidly mixing Markov chains [**224**]. New spectral techniques have emerged and they are powerful and well-suited for dealing with general graphs. In a way, spectral graph theory has entered a new era.

Just as astronomers study stellar spectra to determine the make-up of distant stars, one of the main goals in graph theory is to deduce the principal properties and structure of a graph from its graph spectrum (or from a short list of easily computable invariants). The spectral approach for general graphs is a step in this direction. We will see that eigenvalues are closely related to almost all major invariants of a graph, linking one extremal property to another. There is no question that eigenvalues play a central role in our fundamental understanding of graphs.

The study of graph eigenvalues realizes increasingly rich connections with many other areas of mathematics. A particularly important development is the interaction between spectral graph theory and differential geometry. There is an interesting analogy between spectral Riemannian geometry and spectral graph theory. The concepts and methods of spectral geometry bring useful tools and crucial insights to the study of graph eigenvalues, which in turn lead to new directions and results in spectral geometry. Algebraic spectral methods are also very useful, especially for extremal examples and constructions. In this book, we take a broad approach with emphasis on the geometric aspects of graph eigenvalues, while including the algebraic aspects as well. The reader is not required to have special background in geometry, since this book is almost entirely graph-theoretic.

From the start, spectral graph theory has had applications to chemistry [**27**]. Eigenvalues were associated with the stability of molecules. Also, graph spectra arise naturally in various problems of theoretical physics and quantum mechanics, for example, in minimizing energies of Hamiltonian systems. The recent progress on expander graphs and eigenvalues was initiated by problems in communication networks. The development of rapidly mixing Markov chains has intertwined with advances in randomized approximation algorithms. Applications of graph eigenvalues occur in numerous areas and in different guises. However, the underlying mathematics of spectral graph theory through all its connections to the pure and applied, the continuous and discrete, can be viewed as a single unified subject. It is this aspect that we intend to cover in this book.

1.2. The Laplacian and eigenvalues

Before we start to define eigenvalues, some explanations are in order. The eigenvalues we consider throughout this book are not exactly the same as those in Biggs [**25**] or Cvetković, Doob and Sachs [**90**]. Basically, the eigenvalues are defined here in a general and "normalized" form. Although this might look a little complicated at first, our eigenvalues relate well to other graph invariants for general graphs in a way that other definitions (such as the eigenvalues of adjacency matrices) often fail to do. The advantages of this definition are perhaps due to the fact that it is consistent with the eigenvalues in spectral geometry and in stochastic processes. Many results which were only known for regular graphs can be generalized to all graphs. Consequently, this provides a coherent treatment for a general graph. For definitions and standard graph-theoretic terminology, the reader is referred to [**31**].

In a graph G, let d_v denote the degree of the vertex v. We first define the Laplacian for graphs without loops and multiple edges (the general weighted case with loops will be treated in Section 1.4). To begin, we consider the matrix L, defined as follows:

$$L(u,v) = \begin{cases} d_v & \text{if } u = v, \\ -1 & \text{if } u \text{ and } v \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

Let T denote the diagonal matrix with the (v,v)-th entry having value d_v. The *Laplacian* of G is defined to be the matrix

$$\mathcal{L}(u,v) = \begin{cases} 1 & \text{if } u = v \text{ and } d_v \neq 0, \\ -\dfrac{1}{\sqrt{d_u d_v}} & \text{if } u \text{ and } v \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

We can write

$$\mathcal{L} = T^{-1/2} L T^{-1/2}$$

with the convention $T^{-1}(v,v) = 0$ for $d_v = 0$. We say v is an isolated vertex if $d_v = 0$. A graph is said to be nontrivial if it contains at least one edge.

$\mathcal{L}$ can be viewed as an operator on the space of functions $g : V(G) \to \mathbb{R}$ which satisfies

$$\mathcal{L}g(u) = \frac{1}{\sqrt{d_u}} \sum_{\substack{v \\ u \sim v}} \left(\frac{g(u)}{\sqrt{d_u}} - \frac{g(v)}{\sqrt{d_v}} \right)$$

When G is k-regular, it is easy to see that

$$\mathcal{L} = I - \frac{1}{k} A,$$

where A is the adjacency matrix of G,(i. e., $A(x, y) = 1$ if x is adjacent to y, and 0 otherwise,) and I is an identity matrix. All matrices here are $n \times n$ where n is the number of vertices in G.

For a general graph, we have

$$\begin{aligned} \mathcal{L} &= T^{-1/2} L T^{-1/2} \\ &= I - T^{-1/2} A T^{-1/2}. \end{aligned}$$

We note that $\mathcal{L}$ can be written as

$$\mathcal{L} = S\ S^*,$$

where S is the matrix whose rows are indexed by the vertices and whose columns are indexed by the edges of G such that each column corresponding to an edge $e = \{u, v\}$ has an entry $1/\sqrt{d_u}$ in the row corresponding to u, an entry $-1/\sqrt{d_v}$ in the row corresponding to v, and has zero entries elsewhere. (As it turns out, the choice of signs can be arbitrary as long as one is positive and the other is negative.) Also, S^* denotes the transpose of S.

For readers who are familiar with terminology in homology theory, we remark that S can be viewed as a "boundary operator" mapping "1-chains" defined on edges (denoted by C_1) of a graph to "0-chains" defined on vertices (denoted by C_0). Then, S^* is the corresponding "coboundary operator" and we have

$$C_1 \underset{S^*}{\overset{S}{\rightleftarrows}} C_0$$

Since $\mathcal{L}$ is symmetric, its eigenvalues are all real and non-negative. We can use the variational characterizations of those eigenvalues in terms of the Rayleigh quotient of $\mathcal{L}$ (see, e.g. [**162**]). Let g denote an arbitrary function which assigns to each vertex v of G a real value $g(v)$. We can view g as a column vector. Then

$$\begin{aligned} \frac{\langle g, \mathcal{L}g \rangle}{\langle g, g \rangle} &= \frac{\langle g, T^{-1/2} L T^{-1/2} g \rangle}{\langle g, g \rangle} \\ &= \frac{\langle f, Lf \rangle}{\langle T^{1/2} f, T^{1/2} f \rangle} \\ &= \frac{\sum_{u \sim v} (f(u) - f(v))^2}{\sum_v f(v)^2 d_v} \end{aligned} \tag{1.1}$$

where $g = T^{1/2}f$ and $\sum_{u\sim v}$ denotes the sum over all unordered pairs $\{u, v\}$ for which u and v are adjacent. Here $\langle f, g\rangle = \sum_x f(x)g(x)$ denotes the standard inner product in $\mathbb{R}^n$. The sum $\sum_{u\sim v}(f(u) - f(v))^2$ is sometimes called the *Dirichlet sum* of G and the ratio on the left-hand side of (1.1) is often called the *Rayleigh quotient*. (We note that we can also use the inner product $\langle f, g\rangle = \sum \overline{f(x)}g(x)$ for complex-valued functions.)

From equation (1.1), we see that all eigenvalues are non-negative. In fact, we can easily deduce from equation (1.1) that 0 is an eigenvalue of $\mathcal{L}$. We denote the eigenvalues of $\mathcal{L}$ by $0 = \lambda_0 \le \lambda_1 \le \cdots \le \lambda_{n-1}$. The set of the λ_i's is usually called the *spectrum* of $\mathcal{L}$ (or the spectrum of the associated graph G.) Let $\mathbf{1}$ denote the constant function which assumes the value 1 on each vertex. Then $T^{1/2}\mathbf{1}$ is an eigenfunction of $\mathcal{L}$ with eigenvalue 0. Furthermore,

$$\lambda_G = \lambda_1 = \inf_{f\perp T\mathbf{1}} \frac{\sum_{u\sim v}(f(u) - f(v))^2}{\sum_v f(v)^2 d_v}. \tag{1.2}$$

The corresponding eigenfunction is $g = T^{1/2}f$ as in (1.1). It is sometimes convenient to consider the nontrivial function f achieving (1.2), in which case we call f a *harmonic eigenfunction* of $\mathcal{L}$.

The above formulation for λ_G corresponds in a natural way to the eigenvalues of the Laplace-Beltrami operator for Riemannian manifolds:

$$\lambda_M = \inf \frac{\int_M |\nabla f|^2}{\int_M |f|^2},$$

where f ranges over functions satisfying

$$\int_M f = 0.$$

We remark that the corresponding measure here for each edge is 1 although in the general case for weighted graphs the measure for an edge is associated with the edge weight (see Section 1.4.) The measure for each vertex is the degree of the vertex. A more general notion of vertex weights will be considered in Section 2.5.

We note that (1.2) has several different formulations:

$$\lambda_1 = \inf_f \sup_t \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_v (f(v)-t)^2 d_v} \tag{1.3}$$

$$= \inf_f \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_v (f(v)-\bar{f})^2 d_v}, \tag{1.4}$$

where

$$\bar{f} = \frac{\sum_v f(v)d_v}{\text{vol } G},$$

and vol G denotes the volume of the graph G, given by

$$\text{vol } G = \sum_v d_v.$$

By substituting for $\bar{f}$ and using the fact that $N\sum_{i=1}^N (a_i - a)^2 = \sum_{i<j}(a_i - a_j)^2$ for $a = \sum_{i=1}^N a_i/N$, we have the following expression (which generalizes the one in [**126**]):

$$\lambda_1 = \text{vol } G \ \inf_f \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_{u,v}(f(u)-f(v))^2 d_u d_v}, \tag{1.5}$$

where $\sum_{u,v}$ denotes the sum over all unordered pairs of vertices u, v in G. We can characterize the other eigenvalues of $\mathcal{L}$ in terms of the Rayleigh quotient. The largest eigenvalue satisfies:

$$\lambda_{n-1} = \sup_f \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_v f^2(v)d_v}. \tag{1.6}$$

For a general k, we have

$$\lambda_k = \inf_f \sup_{g\in P_{k-1}} \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_v (f(v)-g(v))^2 d_v} \tag{1.7}$$

$$= \inf_{f\ \perp TP_{k-1}} \frac{\sum_{u\sim v}(f(u)-f(v))^2}{\sum_v f(v)^2 d_v} \tag{1.8}$$

where P_i is the subspace generated by the harmonic eigenfunctions corresponding to λ_i, for $i \leq k-1$.

The different formulations for eigenvalues given above are useful in different settings and they will be used in later chapters. Here are some examples of special graphs and their eigenvalues.

EXAMPLE 1.1. For the complete graph K_n on n vertices, the eigenvalues are 0 and $n/(n-1)$ (with multiplicity $n-1$).

EXAMPLE 1.2. For the complete bipartite graph $K_{m,n}$ on $m+n$ vertices, the eigenvalues are 0, 1 (with multiplicity $m+n-2$), and 2.

EXAMPLE 1.3. For the star S_n on n vertices, the eigenvalues are $0, 1$ (with multiplicity $n-2$), and 2.

EXAMPLE 1.4. For the path P_n on n vertices, the eigenvalues are $1-\cos\frac{\pi k}{n-1}$ for $k = 0, 1, \cdots, n-1$.

EXAMPLE 1.5. For the cycle C_n on n vertices, the eigenvalues are $1-\cos\frac{2\pi k}{n}$ for $k = 0, \cdots, n-1$.

EXAMPLE 1.6. For the n-cube Q_n on 2^n vertices, the eigenvalues are $\frac{2k}{n}$ (with multiplicity $\binom{n}{k}$) for $k = 0, \cdots, n$.

More examples can be found in Chapter 6 on explicit constructions.

1.3. Basic facts about the spectrum of a graph

Roughly speaking, half of the main problems of spectral theory lie in deriving bounds on the distributions of eigenvalues. The other half concern the impact and consequences of the eigenvalue bounds as well as their applications. In this section, we start with a few basic facts about eigenvalues. Some simple upper bounds and lower bounds are stated. For example, we will see that the eigenvalues of any graph lie between 0 and 2. The problem of narrowing the range of the eigenvalues for special classes of graphs offers an open-ended challenge. Numerous questions can be asked either in terms of other graph invariants or under further assumptions imposed on the graphs. Some of these will be discussed in subsequent chapters.

LEMMA 1.7. *For a graph G on n vertices, we have*

(i):

$$\sum_i \lambda_i \leq n$$

with equality holding if and only if G has no isolated vertices.

(ii): *For $n \geq 2$,*

$$\lambda_1 \leq \frac{n}{n-1}$$

with equality holding if and only if G is the complete graph on n vertices. Also, for a graph G without isolated vertices, we have

$$\lambda_{n-1} \geq \frac{n}{n-1}.$$

(iii): *For a graph which is not a complete graph, we have* $\lambda_1 \leq 1$.
(iv): *If* G *is connected, then* $\lambda_1 > 0$. *If* $\lambda_i = 0$ *and* $\lambda_{i+1} \neq 0$ *, then* G *has exactly* $i+1$ *connected components.*
(v): *For all* $i \leq n-1$, *we have*

$$\lambda_i \leq 2.$$

with $\lambda_{n-1} = 2$ *if and only if a connected component of* G *is bipartite and nontrivial.*
(vi): *The spectrum of a graph is the union of the spectra of its connected components.*

PROOF. (i) follows from considering the trace of $\mathcal{L}$.

The inequalities in (ii) follow from (i) and $\lambda_0 = 0$.

Suppose G contains two nonadjacent vertices a and b, and consider

$$f_1(v) = \begin{cases} d_b & \text{if } v = a, \\ -d_a & \text{if } v = b, \\ 0 & \text{if } v \neq a, b. \end{cases}$$

(iii) then follows from (1.2).

If G is connected, the eigenvalue 0 has multiplicity 1 since any harmonic eigenfunction with eigenvalue 0 assumes the same value at each vertex. Thus, (iv) follows from the fact that the union of two disjoint graphs has as its spectrum the union of the spectra of the original graphs.

(v) follows from equation (1.6) and the fact that

$$(f(x) - f(y))^2 \leq 2(f^2(x) + f^2(y)).$$

Therefore

$$\lambda_i \leq \sup_f \frac{\sum_{x \sim y} (f(x) - f(y))^2}{\sum_x f^2(x) d_x} \leq 2.$$

Equality holds for $i = n-1$ when $f(x) = -f(y)$ for every edge $\{x, y\}$ in G. Therefore, since $f \neq 0$, G has a bipartite connected component. On the other hand, if G has a connected component which is bipartite, we can choose the function f so as to make $\lambda_{n-1} = 2$.

(vi) follows from the definition. □

For bipartite graphs, the following slightly stronger result holds:

LEMMA 1.8. *The following statements are equivalent:*

(i): G *is bipartite.*
(ii): G *has* $i+1$ *connected components and* $\lambda_{n-j} = 2$ *for* $1 \leq j \leq i$.
(iii): *For each* λ_i, *the value* $2 - \lambda_i$ *is also an eigenvalue of* G.

PROOF. It suffices to consider a connected graph. Suppose G is bipartite graph with vertex set consisting of two parts A and B. For any harmonic eigenfunction f with eigenvalue λ, we consider the function g

$$g(x) = \begin{cases} f(x) & \text{if } x \in A, \\ -f(x) & \text{if } x \in B. \end{cases}$$

It is easy to check that g is a harmonic eigenfunction with eigenvalue $2 - \lambda$. □

For a connected graph, we can immediately improve the lower bound of λ_1 in Lemma 1.7. For two vertices u and v, the *distance* between u and v is the number of edges in a shortest path joining u and v. The *diameter* of a graph is the maximum distance between any two vertices of G. Here we will give a simple eigenvalue lower bound in terms of the diameter of a graph. More discussion on the relationship between eigenvalues and diameter will be given in Chapter 3.

LEMMA 1.9. *For a connected graph G with diameter D, we have*

$$\lambda_1 \geq \frac{1}{D \text{ vol } G}$$

PROOF. Suppose f is a harmonic eigenfunction achieving λ_1 in (1.2). Let v_0 denote a vertex with $|f(v_0)| = \max_v |f(v)|$. Since $\sum_v f(v) = 0$, there exists a vertex u_0 satisfying $f(u_0)f(v_0) < 0$. Let P denote a shortest path in G joining u_0 and v_0. Then by (1.2) we have

$$\begin{aligned} \lambda_1 &= \frac{\sum_{x \sim y} (f(x) - f(y))^2}{\sum_x f^2(x) d_x} \\ &\geq \frac{\sum_{\{x,y\} \in P} (f(x) - f(y))^2}{\text{vol } G \ f^2(v_0)} \\ &\geq \frac{\frac{1}{D} (f(v_0) - f(u_0))^2}{\text{vol } G \ f^2(v_0)} \\ &\geq \frac{1}{D \text{ vol } G} \end{aligned}$$

by using the Cauchy-Schwarz inequality. □

LEMMA 1.10. *Let f denote a harmonic eigenfunction achieving λ_G in (1.2). Then, for any vertex $x \in V$, we have*

$$\frac{1}{d_x} \sum_{\substack{y \\ y \sim x}} (f(x) - f(y)) = \lambda_G f(x).$$

PROOF. We use a variational argument. For a fixed $x_0 \in V$, we consider f_ϵ such that

$$f_\epsilon(y) = \begin{cases} f(x_0) + \dfrac{\epsilon}{d_{x_0}} & \text{if } y = x_0, \\ f(y) - \dfrac{\epsilon}{\text{vol } G - d_{x_0}} & \text{otherwise.} \end{cases}$$

We have

$$\begin{aligned} & \frac{\displaystyle\sum_{\substack{x,y\in V \\ x\sim y}} (f_\epsilon(x) - f_\epsilon(y))^2}{\displaystyle\sum_{x\in V} f_\epsilon^2(x) d_x} \\ = & \frac{\displaystyle\sum_{\substack{x,y\in V \\ x\sim y}} (f(x)-f(y))^2 + \sum_{\substack{y \\ y\sim x_0}} \frac{2\epsilon(f(x_0)-f(y))}{d_{x_0}} - \sum_{\substack{y \\ y\neq x_0}} \sum_{\substack{y' \\ y\sim y'}} \frac{2\epsilon(f(y)-f(y'))}{\text{vol } G - d_{x_0}}}{\displaystyle\sum_{x\in V} f^2(x) d_x + 2\epsilon f(x_0) - \frac{2\epsilon}{\text{vol } G - d_{x_0}} \sum_{y\neq x_0} f(y) d_y} \\ & + O(\epsilon^2) \\ = & \frac{\displaystyle\sum_{\substack{x,y\in V \\ x\sim y}} (f(x)-f(y))^2 + \frac{2\epsilon \displaystyle\sum_{\substack{y \\ y\sim x_0}} (f(x_0)-f(y))}{d_{x_0}} + \frac{2\epsilon \displaystyle\sum_{\substack{y \\ y\sim x_0}} (f(x_0)-f(y))}{\text{vol } G - d_{x_0}}}{\displaystyle\sum_{x\in V} f^2(x) d_x + 2\epsilon f(x_0) + \frac{2\epsilon f(x_0) d_{x_0}}{\text{vol } G - d_{x_0}}} \\ & + O(\epsilon^2) \end{aligned}$$

since $\displaystyle\sum_{x\in V} f(x) d_x = 0$, and $\displaystyle\sum_y \sum_{y'} (f(y) - f(y')) = 0$. The definition in (1.2) implies that

$$\frac{\displaystyle\sum_{\substack{x,y\in V \\ x\sim y}} (f_\epsilon(x) - f_\epsilon(y))^2}{\displaystyle\sum_{x\in V} f_\epsilon^2(x) d_x} \geq \frac{\displaystyle\sum_{\substack{x,y\in V \\ x\sim y}} (f(x) - f(y))^2}{\displaystyle\sum_{x\in V} f^2(x) d_x}$$

If we consider what happens to the Rayleigh quotient for f_ϵ as $\epsilon \to 0$ from above, or from below, we can conclude that

$$\frac{1}{d_{x_0}} \sum_{\substack{y \\ y\sim x_0}} (f(x_0) - f(y)) = \lambda_G f(x_0).$$

and the Lemma is proved. □

One can also prove the statement in Lemma 1.10 by recalling that $f = T^{-1/2}g$, where $\mathcal{L}g = \lambda_G g$. Then

$$T^{-1}Lf = T^{-1}(T^{1/2}\mathcal{L}T^{1/2})(T^{-1/2}g) = T^{-1/2}\lambda_G g = \lambda_G f,$$

and examining the entries gives the desired result.

With a little linear algebra, we can improve the bounds on eigenvalues in terms of the degrees of the vertices.

We consider the trace of $(I - \mathcal{L})^2$. We have

$$\begin{aligned} Tr(I-\mathcal{L})^2 &= \sum_i (1-\lambda_i)^2 \\ &\leq 1 + (n-1)\bar{\lambda}^2, \end{aligned} \tag{1.9}$$

where

$$\bar{\lambda} = \max_{i \neq 0} |1 - \lambda_i|.$$

On the other hand,

$$\begin{aligned} Tr(I-\mathcal{L})^2 &= Tr(T^{-1/2}AT^{-1}AT^{-1/2}) \\ &= \sum_{x,y} \frac{1}{\sqrt{d_x}} A(x,y) \frac{1}{d_y} A(y,x) \frac{1}{\sqrt{d_x}} \\ &= \sum_x \frac{1}{d_x} - \sum_{x \sim y} (\frac{1}{d_x} - \frac{1}{d_y})^2, \end{aligned} \tag{1.10}$$

where A is the adjacency matrix. From this, we immediately deduce

LEMMA 1.11. *For a k-regular graph G on n vertices, we have*

$$\max_{i \neq 0} |1 - \lambda_i| \geq \sqrt{\frac{n-k}{(n-1)k}} \tag{1.11}$$

This follows from the fact that

$$\max_{i \neq 0} |1 - \lambda_i|^2 \geq \frac{1}{n-1}(tr(I-\mathcal{L})^2 - 1).$$

Let d_H denote the harmonic mean of the d_v's, i.e.,

$$\frac{1}{d_H} = \frac{1}{n} \sum_v \frac{1}{d_v}.$$

It is tempting to consider generalizing (1.11) with k replaced by d_H. This, however, is not true as shown by the following example due to Elizabeth Wilmer.

EXAMPLE 1.12. Consider the m-petal graph on $n = 2m+1$ vertices, $v_0, v_1, \cdots, v_{2m}$ with edges $\{v_0, v_i\}$ and $\{v_{2i-1}, v_{2i}\}$, for $i \geq 1$. This graph has eigenvalues $0, 1/2$ (with multiplicity $m-1$), and $3/2$ (with multiplicity $m+1$). So we have $\max_{i \neq 0} |1 - \lambda_i| = 1/2$. However,

$$\sqrt{\frac{n - d_H}{(n-1)d_H}} = \sqrt{\frac{m - 1/2}{2m}} \to \frac{1}{\sqrt{2}}$$

as $m \to \infty$.

Still, for a general graph, we can use the fact that

$$\frac{\sum_{x\sim y}(\frac{1}{d_x}-\frac{1}{d_y})^2}{\sum_{x\in V}(\frac{1}{d_x}-\frac{1}{d_H})^2 d_x} \;\le\; \lambda_{n-1} \;\le\; 1+\bar{\lambda}. \tag{1.12}$$

Combining (1.9), (1.10) and (1.12), we obtain the following:

LEMMA 1.13. *For a graph G on n vertices, $\bar{\lambda} = \max_{i\neq 0}|1-\lambda_i|$ satisfies*

$$1+(n-1)\bar{\lambda}^2 \geq \frac{n}{d_H}(1-(1+\bar{\lambda})(\frac{k}{d_H}-1)),$$

where k denotes the average degree of G.

There are relatively easy ways to improve the upper bound for λ_1. From the characterization in the preceding section, we can choose any function $f : V(G) \to \mathbb{R}$, and its Rayleigh quotient will serve as an upper bound for λ_1. Here we describe an upper bound for λ_1 (see [**204**]).

LEMMA 1.14. *Let G be a graph with diameter $D \geq 4$, and let k denote the maximum degree of G. Then*

$$\lambda_1 \leq 1 - 2\frac{\sqrt{k-1}}{k}\left(1-\frac{2}{D}\right)+\frac{2}{D}.$$

One way to bound eigenvalues from above is to consider "contracting" the graph G into a weighted graph H (which will be defined in the next section). Then the eigenvalues of G can be upper-bounded by the eigenvalues of H or by various upper bounds on them, which might be easier to obtain. We remark that the proof of Lemma 1.14 proceeds by basically contracting the graph into a weighted path. We will prove Lemma 1.14 in the next section.

We note that Lemma 1.14 gives a proof (see [**5**]) that for any fixed k and for any infinite family of regular graphs with degree k,

$$\limsup \lambda_1 \leq 1 - 2\frac{\sqrt{k-1}}{k}.$$

This bound is the best possible since it is sharp for the Ramanujan graphs (which will be discussed in Chapter **??**). We note that the cleaner version of $\lambda_1 \leq 1-2\sqrt{k-1}/k$ is not true for certain graphs (e.g., 4-cycles or complete bipartite graphs). This example also illustrates that the assumption in Lemma 1.14 concerning $D \geq 4$ is essential.

1.4. Eigenvalues of weighted graphs

Before defining weighted graphs, we will say a few words about two different approaches for giving definitions. We could have started from the very beginning with weighted graphs, from which simple graphs arise as a special case in which the weights are 0 or 1. However, the unique characteristics and special strength of

graph theory is its ability to deal with the $\{0,1\}$-problems arising in many natural situations. The clean formulation of a simple graph has conceptual advantages. Furthermore, as we shall see, all definitions and subsequent theorems for simple graphs can usually be easily carried out for weighted graphs. A weighted undirected graph G (possibly with loops) has associated with it a weight function $w : V \times V \to \mathbb{R}$ satisfying

$$w(u,v) = w(v,u)$$

and

$$w(u,v) \geq 0.$$

We note that if $\{u,v\} \notin E(G)$, then $w(u,v) = 0$. Unweighted graphs are just the special case where all the weights are 0 or 1.

In the present context, the degree d_v of a vertex v is defined to be:

$$d_v = \sum_u w(u,v),$$

$$\text{vol } G = \sum_v d_v.$$

We generalize the definitions of previous sections, so that

$$L(u,v) = \begin{cases} d_v - w(v,v) & \text{if } u = v, \\ -w(u,v) & \text{if } u \text{ and } v \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

In particular, for a function $f : V \to \mathbb{R}$, we have

$$Lf(x) = \sum_{\substack{y \\ x \sim y}} (f(x) - f(y))w(x,y).$$

Let T denote the diagonal matrix with the (v,v)-th entry having value d_v. The Laplacian of G is defined to be

$$\mathcal{L} = T^{-1/2} L T^{-1/2}.$$

In other words, we have

$$\mathcal{L}(u,v) = \begin{cases} 1 - \dfrac{w(v,v)}{d_v} & \text{if } u = v, \text{ and } d_v \neq 0, \\ -\dfrac{w(u,v)}{\sqrt{d_u d_v}} & \text{if } u \text{ and } v \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

We can still use the same characterizations for the eigenvalues of the generalized versions of $\mathcal{L}$. For example,

$$\begin{aligned} \lambda_G := \lambda_1 &= \inf_{g \perp T^{1/2}\mathbf{1}} \frac{\langle g,\ \mathcal{L}g\rangle}{\langle g,\ g\rangle} \\ &= \inf_{\sum f(x)d_x = 0} \frac{\sum_{x\in V} f(x)Lf(x)}{\sum_{x\in V} f^2(x)d_x} \\ &= \inf_{\sum f(x)d_x = 0} \frac{\sum_{x\sim y}(f(x)-f(y))^2 w(x,y)}{\sum_{x\in V} f^2(x)d_x}. \end{aligned} \tag{1.13}$$

A contraction of a graph G is formed by identifying two distinct vertices, say u and v, into a single vertex v^*. The weights of edges incident to v^* are defined as follows:

$$\begin{aligned} w(x,v^*) &= w(x,u) + w(x,v), \\ w(v^*,v^*) &= w(u,u) + w(v,v) + 2w(u,v). \end{aligned}$$

LEMMA 1.15. *If H is formed by contractions from a graph G, then*

$$\lambda_G \le \lambda_H$$

The proof follows from the fact that an eigenfunction which achieves λ_H for H can be lifted to a function defined on $V(G)$ such that all vertices in G that contract to the same vertex in H share the same value.

Now we return to Lemma 1.14.

SKETCHED PROOF OF LEMMA 1.14:
Let u and v denote two vertices that are at distance $D \ge 2t+2$ in G. We contract G into a path H with $2t+2$ edges, with vertices $x_0, x_1, \ldots x_t, z, y_t, \ldots, y_2, y_1, y_0$ such that vertices at distance i from u, $0 \le i \le t$, are contracted to x_i, and vertices at distance j from v, $0 \le j \le t$, are contracted to y_j. The remaining vertices are contracted to z. To establish an upper bound for λ_1, it is enough to choose a suitable function f, defined as follows:

$$\begin{aligned} f(x_i) &= a(k-1)^{-i/2}, \\ f(y_j) &= b(k-1)^{-j/2}, \\ f(z) &= 0, \end{aligned}$$

where the constants a and b are chosen so that

$$\sum_x f(x)d_x = 0.$$

It can be checked that the Rayleigh quotient satisfies

$$\frac{\sum_{u\sim v}(f(u)-f(v))^2 w(u,v)}{\sum_v f(v)^2 d_v} \leq 1 - \frac{2\sqrt{k-1}}{k}\left(1-\frac{1}{t+1}\right)+\frac{1}{t+1},$$

since the ratio is maximized when $w(x_i, x_{i+1}) = k(k-1)^{i-1} = w(y_i, y_{i+1})$. This completes the proof of the lemma. □

1.5. Eigenvalues and random walks

In a graph G, a walk is just a sequence of vertices $(v_0, v_1, \cdots, v_s)$ with $\{v_{i-1}, v_i\} \in E(G)$ for all $1 \leq i \leq s$. A random walk is determined by the transition probabilities $P(u,v) = Prob(x_{i+1} = v | x_i = u)$, which are independent of i. Clearly, for each vertex u,

$$\sum_v P(u,v) = 1.$$

For any initial distribution $f : V \to \mathbb{R}$ with $\sum_v f(v) = 1$, the distribution after k steps is just fP^k (i.e., a matrix multiplication with f viewed as a row vector where P is the matrix of transition probabilities). The random walk is said to be *ergodic* if there is a unique *stationary distribution* $\pi(v)$ satisfying

$$\lim_{s\to\infty} fP^s(v) = \pi(v).$$

It is easy to see that necessary conditions for the ergodicity of P are (i) *irreducibility*, *i.e.*, for any $u, v \in V$, there exists some s such that $P^s(u,v) > 0$ (ii) *aperiodicity*, *i.e.*, g.c.d. $\{s : P^s(u,v) > 0\} = 1$. As it turns out, these are also sufficient conditions. A major problem of interest is to determine the number of steps s required for P^s to be *close* to its stationary distribution, given an arbitrary initial distribution.

We say a random walk is *reversible* if

$$\pi(u)P(u,v) = \pi(v)P(v,u).$$

An alternative description for a reversible random walk can be given by considering a weighted connected graph with edge weights satisfying

$$w(u,v) = w(v,u) = \pi(v)P(v,u)/c$$

where c can be any constant chosen for the purpose of simplifying the values. (For example, we can take c to be the average of $\pi(v)P(v,u)$ over all (v,u) with $P(v,u) \neq 0$, so that the values for $w(v,u)$ are either 0 or 1 for a simple graph.) The random walk on a weighted graph has as its transition probabilities

$$P(u,v) = \frac{w(u,v)}{d_u},$$

where $d_u = \sum_z w(u,z)$ is the (weighted) degree of u. The two conditions for ergodicity are equivalent to the conditions that the graph be (i) connected and (ii) non-bipartite. From Lemma 1.7, we see that (i) is equivalent to $\lambda_1 > 0$ and

(ii) implies $\lambda_{n-1} < 2$. As we will see later in (1.15), together (i) and (ii) deduce ergodicity.

We remind the reader that an unweighted graph has $w(u,v)$ equal to either 0 or 1. The usual random walk on an unweighted graph has transition probability $1/d_v$ of moving from a vertex v to any one of its neighbors. The transition matrix P then satisfies

$$P(u,v) = \begin{cases} 1/d_u & \text{if } u \text{ and } v \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

In other words,

$$fP(v) = \sum_{\substack{u \\ u \sim v}} \frac{1}{d_u} f(u)$$

for any $f : V(G) \to \mathbb{R}$.

It is easy to check that

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

where A is the adjacency matrix.

In a random walk with an associated weighted connected graph G, the transition matrix P satisfies

$$\mathbf{1}TP = \mathbf{1}T$$

where $\mathbf{1}$ is the vector with all coordinates 1. Therefore the stationary distribution is exactly $\pi = \mathbf{1}T/\mathrm{vol}\ G$, We want to show that when k is large enough, for any initial distribution $f : V \to \mathbb{R}$, fP^k converges to the stationary distribution.

First we consider convergence in the L_2 (or Euclidean) norm. Suppose we write

$$fT^{-1/2} = \sum_i a_i \phi_i,$$

where ϕ_i denotes the orthonormal eigenfunction associated with λ_i.

Recall that $\phi_0 = \mathbf{1}T^{1/2}/\sqrt{\mathrm{vol}\ G}$ and $\|\cdot\|$ denotes the L_2-norm, so

$$a_0 = \frac{\langle fT^{-1/2}, \mathbf{1}T^{1/2}\rangle}{\|\mathbf{1}T^{1/2}\|} = \frac{1}{\sqrt{\mathrm{vol}\ G}}$$

since $\langle f, \mathbf{1}\rangle = 1$. We then have

$$\begin{aligned} \|fP^s - \pi\| &= \|fP^s - \mathbf{1}T/\mathrm{vol}\ G\| \\ &= \|fP^s - a_0\phi_0 T^{1/2}\| \\ &= \|fT^{-1/2}(I-\mathcal{L})^s T^{1/2} - a_0\phi_0 T^{1/2}\| \\ &= \|\sum_{i\neq 0}(1-\lambda_i)^s a_i \phi_i T^{1/2}\| \\ &\le (1-\lambda')^s \frac{\max_x \sqrt{d_x}}{\min_y \sqrt{d_y}} \\ &\le e^{-s\lambda'} \frac{\max_x \sqrt{d_x}}{\min_y \sqrt{d_y}} \end{aligned} \tag{1.14}$$

where

$$\lambda' = \begin{cases} \lambda_1 & \text{if } 1-\lambda_1 \geq \lambda_{n-1}-1 \\ 2-\lambda_{n-1} & \text{otherwise.} \end{cases}$$

So, after $s \geq 1/\lambda' \log(\max_x \sqrt{d_x}/\epsilon \min_y \sqrt{d_y})$ steps, the L_2 distance between fP^s and its stationary distribution is at most ϵ.

Although λ' occurs in the above upper bound for the distance between the stationary distribution and the s-step distribution, in fact, only λ_1 is crucial in the following sense. Note that λ' is either λ_1 or $2-\lambda_{n-1}$. Suppose the latter holds, *i.e.,* $\lambda_{n-1}-1 \geq 1-\lambda_1$. We can consider a modified random walk, called the lazy walk, on the graph G' formed by adding a loop of weight d_v to each vertex v. The new graph has Laplacian eigenvalues $\tilde{\lambda}_k = \lambda_k/2 \leq 1$, which follows from equation (1.13). Therefore,

$$1-\tilde{\lambda}_1 \geq 1-\tilde{\lambda}_{n-1} \geq 0,$$

and the convergence bound in L_2 distance in (1.14) for the modified random walk becomes

$$2/\lambda_1 \log(\frac{\max_x \sqrt{d_x}}{\epsilon \min_y \sqrt{d_y}}).$$

In general, suppose a weighted graph with edge weights $w(u,v)$ has eigenvalues λ_i with $\lambda_{n-1}-1 \geq 1-\lambda_1$. We can then modify the weights by choosing, for some constant c,

$$w'(u,v) = \begin{cases} w(v,v)+cd_v & \text{if } u=v \\ w(u,v) & \text{otherwise.} \end{cases} \tag{1.15}$$

The resulting weighted graph has eigenvalues

$$\lambda'_k = \frac{\lambda_k}{1+c} = \frac{2\lambda_k}{\lambda_{n-1}+\lambda_k}$$

where

$$c = \frac{\lambda_1+\lambda_{n-1}}{2} - 1 \leq \frac{1}{2}.$$

Then we have

$$1-\lambda'_1 = \lambda'_{n-1} - 1 = \frac{\lambda_{n-1}-\lambda_1}{\lambda_{n-1}+\lambda_1}.$$

Since $c \leq 1/2$ and we have $\lambda'_k \geq 2\lambda_k/(2+\lambda_k) \geq 2\lambda_k/3$ for $\lambda_k \leq 1$. In particular we set

$$\lambda = \lambda'_1 = \frac{2\lambda_1}{\lambda_{n-1}+\lambda_1} \geq \frac{2}{3}\lambda_1.$$

Therefore the modified random walk corresponding to the weight function w' has an improved bound for the convergence rate in L_2 distance:

$$\frac{1}{\lambda} \log \frac{\max_x \sqrt{d_x}}{\epsilon \min_y \sqrt{d_y}}.$$

We remark that for many applications in sampling, the convergence in L_2 distance seems to be too weak since it does not require convergence at each vertex. There are several stronger notions of distance several of which we will mention.

A strong notion of convergence that is often used is measured by the relative pointwise distance (see [**224**]): After s steps, the *relative pointwise distance* (r.p.d.) of P to the stationary distribution $\pi(x)$ is given by

$$\Delta(s) = \max_{x,y} \frac{|P^s(y,x) - \pi(x)|}{\pi(x)}.$$

Let ψ_x denote the characteristic function of x defined by:

$$\psi_x(y) = \begin{cases} 1 & \text{if } y = x, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose

$$\psi_x T^{1/2} = \sum_i \alpha_i \phi_i,$$

$$\psi_y T^{-1/2} = \sum_i \beta_i \phi_i.$$

where ϕ_i's denote the eigenfunction of the Laplacian $\mathcal{L}$ of the weighted graph associated with the random walk. In particular,

$$\alpha_0 = \frac{d_x}{\sqrt{\text{vol } G}},$$

$$\beta_0 = \frac{1}{\sqrt{\text{vol } G}}.$$

Let A^* denote the transpose of A. We have

$$\begin{aligned}
\Delta(t) &= \max_{x,y} \frac{|\psi_y \, P^t \, \psi_x^* - \pi(x)|}{\pi(x)} \\
&= \max_{x,y} \frac{|\psi_y \, T^{-1/2} \, (I - \mathcal{L})^t \, T^{1/2} \, \psi_x^* - \pi(x)|}{\pi(x)} \\
&\leq \max_{x,y} \frac{\sum_{i \neq 0} |(1-\lambda_i)^t \alpha_i \beta_i|}{d_x/\text{vol } G} \\
&\leq \bar{\lambda}^t \max_{x,y} \frac{\sum_{i \neq 0} |\alpha_i \beta_i|}{d_x/\text{vol } G} \\
&= \bar{\lambda}^t \max_{x,y} \frac{\|\psi_x T^{1/2}\| \; \|\psi_y T^{-1/2}\|}{d_x/\text{vol } G} \\
&\leq \bar{\lambda}^t \frac{\text{vol } G}{\min_{x,y} \sqrt{d_x d_y}} \\
&\leq e^{-t(1-\bar{\lambda})} \frac{\text{vol } G}{\min_x d_x}
\end{aligned}$$

where $\bar{\lambda} = \max_{i \neq 0} |1 - \lambda_i|$. So if we choose t such that

$$t \geq \frac{1}{1-\bar{\lambda}} \log \frac{\text{vol } G}{\epsilon \min_x d_x},$$

then, after t steps, we have $\Delta(t) \leq \epsilon$.

When $1 - \lambda_1 \neq \bar{\lambda}$, we can improve the above bound by using a lazy walk as described in (1.15). The proof is almost identical to the above calculation except for using the Laplacian of the modified weighted graph associated with the lazy walk. This can be summarized by the following theorem:

THEOREM 1.16. *For a weighted graph G, we can choose a modified random walk P so that the relative pairwise distance $\Delta(t)$ is bounded above by:*

$$\Delta(t) \leq e^{-t\lambda} \frac{\text{vol } G}{\min_x d_x} \leq \exp^{-2t\lambda_1/(2+\lambda_1)} \frac{\text{vol } G}{\min_x d_x}.$$

where $\lambda = \lambda_1$ if $2 \geq \lambda_{n-1} + \lambda_1$ and $\lambda = 2\lambda_1/(\lambda_{n-1} + \lambda_1)$ otherwise.

COROLLARY 1.17. *For a weighted graph G, we can choose a modified random walk P so that we have*

$$\Delta(t) \leq e^{-c}$$

if

$$t \geq \frac{1}{\lambda} \log \frac{\text{vol } G}{\min_x d_x}$$

where $\lambda = \lambda_1$ if $2 \geq \lambda_{n-1} + \lambda_1$ and $\lambda = 2\lambda_1/(\lambda_{n-1} + \lambda_1)$ otherwise.

We remark that for any initial distribution $f : V \to \mathbb{R}$ with $\langle f, \mathbf{1} \rangle = 1$ and $f(x) \geq 0$, we have, for any x,

$$\begin{aligned} \frac{|fP^s(x) - \pi(x)|}{\pi(x)} &\leq \sum_y f(y) \frac{|\, P^s(y,x) - \pi(x) \,|}{\pi(x)} \\ &\leq \sum_y f(y) \Delta(s) \\ &\leq \Delta(s). \end{aligned}$$

Another notion of distance for measuring convergence is the so-called *total variation distance*, which is just half of the L_1 distance:

$$\begin{aligned} \Delta_{TV}(s) &= \max_{A \subset V(G)} \max_{y \in V(G)} |\, \sum_{x \in A} (P^s(y,x) - \pi(x)) \,| \\ &= \frac{1}{2} \max_{y \in V(G)} \sum_{x \in V(G)} |\, P^s(y,x) - \pi(x) \,|\,. \end{aligned}$$

The total variation distance is bounded above by the relative pointwise distance, since

$$\begin{aligned}\Delta_{TV}(s) &= \max_{\substack{A\subset V(G)\\ \mathrm{vol}A\leq\frac{\mathrm{vol}G}{2}}} \max_{y\in V(G)} \mid \sum_{x\in A}(P^s(y,x)-\pi(x)) \mid \\ &\leq \max_{\substack{A\subset V(G)\\ \mathrm{vol}A\leq\frac{\mathrm{vol}G}{2}}} \sum_{x\in A}\pi(x)\Delta(s) \\ &\leq \frac{1}{2}\Delta(s).\end{aligned}$$

Therefore, any convergence bound using relative pointwise distance implies the same convergence bound using total variation distance. There is yet another notion of distance, sometimes called *χ-squared distance*, denoted by $\Delta'(s)$ and defined by:

$$\begin{aligned}\Delta'(s) &= \max_{y\in V(G)}\left(\sum_{x\in V(G)}\frac{(P^s(y,x)-\pi(x))^2}{\pi(x)}\right)^{1/2} \\ &\geq \max_{y\in V(G)}\sum_{x\in V(G)} \mid P^s(y,x)-\pi(x) \mid \\ &= 2\Delta_{TV}(s),\end{aligned}$$

using the Cauchy-Schwarz inequality. $\Delta'(s)$ is also dominated by the relative pointwise distance (which we will mainly use in this book).

$$\begin{aligned}\Delta'(s) &= \max_{x\in V(G)}\left(\sum_{y\in V(G)}\frac{(P^s(x,y)-\pi(y))^2}{\pi(y)}\right)^{1/2} \\ &\leq \max_{x\in V(G)}(\sum_{y\in V(G)}(\Delta(s))^2\cdot\pi(y))^{\frac{1}{2}} \\ &\leq \Delta(s).\end{aligned}$$

We note that

$$\begin{aligned}\Delta'(s)^2 &\geq \sum_x \pi(x)\sum_y\frac{(P^s(x,y)-\pi(y))^2}{\pi(y)} \\ &= \sum_x \psi_x T^{1/2}(P^s-I_0)T^{-1}(P^s-I_0)T^{1/2}\psi_x^* \\ &= \sum_x \psi_x((I-\mathcal{L})^{2s}-I_0)\psi_x^*,\end{aligned}$$

where I_0 denotes the projection onto the eigenfunction ϕ_0, ϕ_i denotes the i-th orthonormal eigenfunction of $\mathcal{L}$ and ψ_x denotes the characteristic function of x. Since

$$\psi_x = \sum_i \phi_i(x)\phi_i,$$

we have

$$\begin{aligned}(1.16)\qquad \Delta'(s)^2 &\geq \sum_x \psi_x((I-\mathcal{L})^{2s}-I_0)\psi_x^* \\ &= \sum_x (\sum_i \phi_i(x)\phi_i)((I-\mathcal{L})^{2s}-I_0)(\sum_i \phi_i(x)\phi_i)^* \\ &= \sum_x \sum_{i\neq 0} \phi_i^2(x)(1-\lambda_i)^{2s} \\ &= \sum_{i\neq 0}\sum_x \phi_i^2(x)(1-\lambda_i)^{2s} \\ &= \sum_{i\neq 0}(1-\lambda_i)^{2s}.\end{aligned}$$

Equality in (1.16) holds if, for example, G is vertex-transitive, i.e., there is an automorphism mapping u to v for any two vertices in G, (for more discussions, see Chapter 7 on symmetrical graphs). Therefore, we conclude

THEOREM 1.18. *Suppose G is a vertex transitive graph. Then a random walk after s steps converges to the uniform distribution under total variation distance or χ-squared distance in a number of steps bounded by the sum of $(1-\lambda_i)^{2s}$, where λ_i ranges over the non-trivial eigenvalues of the Laplacian:*

$$(1.17)\qquad \Delta_{TV}(s) \leq \frac{1}{2}\Delta'(s) = \frac{1}{2}(\sum_{i\neq 0}(1-\lambda_i)^{2s})^{1/2}.$$

The above theorem is often derived from the *Plancherel formula*. Here we have employed a direct proof. We remark that for some graphs which are not vertex-transitive, a somewhat weaker version of (1.17) can still be used with additional work (see [**81**] and the remarks in Section 4.6). Here we will use Theorem 1.18 to consider random walks on an n-cube.

EXAMPLE 1.19. For the n-cube Q_n, our (lazy) random walk (as defined in (1.15)) converges to the uniform distribution under the total variation distance, as estimated as follows: From Example (1.6), the eigenvalues of the Q_n are $2k/n$ of multiplicity $\binom{n}{k}$ for $k = 0, \cdots, n$. The adjusted eigenvalues for the weighted graph corresponding to the lazy walk are $\lambda'_k = 2\lambda_k/(\lambda_{n-1}+\lambda_1) = \lambda_k n/(n+1)$. By using Theorem 1.18 (also see [**104**]), we have

$$\begin{aligned}\Delta_{TV}(s) \leq \frac{1}{2}\Delta'(s) &\leq \frac{1}{2}(\sum_{k=1}^{n}\binom{n}{k}(1-\frac{2k}{n+1})^{2s})^{1/2} \\ &\leq \frac{1}{2}(\sum_{k=1}^{n} e^{k\log n - \frac{4ks}{n+1}})^{1/2} \\ &\leq e^{-c}\end{aligned}$$

if $s \geq \frac{1}{4}n\log n + cn$.

We can also compute the rate of convergence of the lazy walk under the relative pointwise distance. Suppose we denote vertices of Q_n by subsets of an n-set

$\{1, 2, \cdots, n\}$. The orthonormal eigenfunctions are ϕ_S for $S \subset \{1, 2, \cdots, n\}$ where

$$\phi_S(X) = \frac{(-1)^{|S \cap X|}}{2^{n/2}}$$

for any $X \subset \{1, 2, \cdots, n\}$. For a vertex indexed by the subset S, the characteristic function is denoted by

$$\psi_S(X) = \begin{cases} 1 & \text{if } X = S, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly,

$$\psi_X = \sum_S \frac{(-1)^{|S \cap X|}}{2^{n/2}} \phi_S.$$

Therefore,

$$\begin{aligned} \frac{|P^s(X,Y) - \pi(Y)|}{\pi(Y)} &= |2^n \psi_X P^s \psi_Y^* - 1| \\ &\leq |2^n \psi_X P^s \psi_X^* - 1| \\ &= \sum_{S \neq \emptyset} (1 - \frac{2|S|}{n+1})^s \\ &= \sum_{k=1}^{n} \binom{n}{k} (1 - \frac{2k}{n+1})^s \end{aligned}$$

This implies

$$\begin{aligned} \Delta(s) &= \sum_{k=1}^{n} \binom{n}{k} (1 - \frac{2k}{n+1})^s \\ &\leq \sum_{k=1}^{n} e^{k \log n - \frac{2ks}{n+1}} \\ &\leq e^{-c} \end{aligned}$$

if

$$s \geq \frac{n \log n}{2} + cn.$$

So, the rate of convergence under relative pointwise distance is about twice that under the total variation distance for Q_n.

In general, $\Delta_{TV}(s), \Delta'(s)$ and $\Delta(s)$ can be quite different [**81**]. Nevertheless, a convergence lower bound for any of these notions of distance (and the L_2-norm) is λ^{-1}. This we will leave as an exercise. We remark that Aldous [**4**] has shown that if $\Delta_{TV}(s) \leq \epsilon$, then $P^s(y,x) \geq c_\epsilon \pi(x)$ for all vertices x, where c_ϵ depends only on ϵ.

Notes

For an induced subgraph of a graph, we can define the Laplacian with boundary conditions. We will leave the definitions for eigenvalues with Neumann boundary conditions and Dirichlet boundary conditions for Chapter **??**.

The Laplacian for a directed graph is also very interesting. The Laplacian for a hypergraph has very rich structures. However, in this book we mainly focus on the Laplacian of a graph since the theory on these generalizations and extensions is still being developed.

In some cases, the factor $\log \frac{\mathrm{vol}\ G}{\min_x d_x}$ in the upper bound for $\Delta(t)$ can be further reduced. Recently, P. Diaconis and L. Saloff-Coste [**100**] introduced a discrete version of the logarithmic Sobolev inequalities which can reduce this factor further for certain graphs (for $\Delta'(t)$). In Chapter 12, we will discuss some advanced techniques for further bounding the convergence rate under the relative pointwise distance.