ELEMENTARY NUMBER THEORY, GROUP THEORY, AND RAMANUJAN GRAPHS

GIULIANA DAVIDOFF
Mount Holyoke College

PETER SARNAK
Princeton University & NYU

ALAIN VALETTE
Universite de Neuchatel

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 2003

Printed in the United States of America

Typeface Times Roman 10.25/13 pt. *System* LaTeX 2$_\varepsilon$ [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Davidoff, Giuliana P.

Elementary number theory, group theory, and Ramanujan graphs / Guiliana Davidoff,
Peter Sarnak, Alain Valette.

p. cm. – (London Mathematical Society student texts; 55)

Includes bibliographical references and index.

ISBN 0-521-82426-5 – ISBN 0-521-53143-8 (pb.)

1. Graph theory. 2. Number theory. 3. Group theory. I. Sarnak, Peter.
II. Valette, Alain. III. Title. IV. Series.

QA166 .D35 2003

511′.5 – dc21 2002074057

ISBN 0 521 82426 5 hardback
ISBN 0 521 53143 8 paperback

Contents

Chapter 1

Graph Theory

1.1. The Adjacency Matrix and Its Spectrum

We shall be concerned with graphs $X = (V, E)$, where V is the set of vertices and E is the set of edges. As stated in the Overview, we always assume our graphs to be undirected, and most often we will deal with finite graphs.

We let $V = \{v_1, v_2, \ldots\}$ be the set of vertices of X. Then the *adjacency matrix* of the graph X is the matrix A indexed by pairs of vertices $v_i, v_j \in V$. That is, $A = (A_{ij})$, where

$$A_{ij} = \text{number of edges joining } v_i \text{ to } v_j.$$

We say that X is *simple* if there is at most one edge joining adjacent vertices; hence, X is simple if and only if $A_{ij} \in \{0, 1\}$ for every $v_i, v_j \in V$.

Note that A completely determines X and that A is symmetric because X is undirected. Furthermore, X has no loops if and only if $A_{ii} = 0$ for every $v_i \in V$.

1.1.1. Definition. Let $k \geq 2$ be an integer. We say that the graph X is *k-regular* if for every $v_i \in V : \sum_{v_j \in V} A_{ij} = k$.

If X has no loop, this amounts to saying that each vertex has exactly k neighbors.

Assume that X is a finite graph on n vertices. Then A is an n-by-n symmetric matrix; hence, it has n real eigenvalues, counting multiplicities, that we may list in decreasing order:

$$\mu_0 \geq \mu_1 \geq \cdots \geq \mu_{n-1}.$$

The *spectrum* of X is the set of eigenvalues of A. Note that μ_0 is a simple eigenvalue, or has multiplicity 1, if and only if $\mu_0 > \mu_1$.

For an arbitrary graph $X = (V, E)$, consider functions $f : V \to \mathbb{C}$ from the set of vertices of X to the complex numbers, and define

$$\ell^2(V) = \{f : V \to \mathbb{C} : \sum_{v\in V} |f(v)|^2 < +\infty\}.$$

The space $\ell^2(E)$ is defined analogously.

Clearly, if V is finite, say $|V| = n$, then every function $f : V \to \mathbb{C}$ is in $\ell^2(V)$. We can think of each such function as a vector in $\mathbb{C}^n$ on which the adjacency matrix acts in the usual way:

$$Af = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ \vdots & \vdots & & \vdots \\ A_{i1} & A_{i2} & \dots & A_{in} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} f(v_1) \\ f(v_2) \\ \vdots \\ f(v_n) \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}\ f(v_1) + A_{12}\ f(v_2) + \dots + A_{1n}\ f(v_n) \\ \vdots \\ A_{i1}\ f(v_1) + A_{i2}\ f(v_2) + \dots + A_{in}\ f(v_n) \\ \vdots \\ A_{n1}\ f(v_1) + A_{n2}\ f(v_2) + \dots + A_{nn}\ f(v_n) \end{pmatrix}.$$

Hence, $(Af)(v_i) = \sum_{j=1}^{n} A_{ij}\ f(v_j)$. It is very convenient, both notationally and conceptually, to forget about the numbering of vertices and to index matrix entries of A directly by pairs of vertices. So we shall represent A by a matrix $(A_{xy})_{x,y\in V}$, and the previous formula becomes $(Af)(x) = \sum_{y\in V} A_{xy}\ f(y)$, for every $x \in V$.

1.1.2. Proposition. Let X be a finite k-regular graph with n vertices. Then

(a) $\mu_0 = k$;
(b) $|\mu_i| \le k$ for $1 \le i \le n-1$;
(c) μ_0 has multiplicity 1, if and only if X is connected.

Proof. We prove (a) and (b) simultaneously by noticing first that the constant function $f \equiv 1$ on V is an eigenfunction of A associated with the eigenvalue k. Next, we prove that, if μ is any eigenvalue, then $|\mu| \le k$. Indeed, let f be a real-valued eigenfunction associated with μ. Let $x \in V$ be such that

$$|f(x)| = \max_{y \in V} |f(y)|.$$

Replacing f by $-f$ if necessary, we may assume $f(x) > 0$. Then

$$f(x)\,|\mu| = |f(x)\,\mu| = \left|\sum_{y \in V} A_{xy}\; f(y)\right| \leq \sum_{y \in V} A_{xy}\,|f(y)|$$

$$\leq f(x) \sum_{y \in V} A_{xy} = f(x)\,k.$$

Cancelling out $f(x)$ gives the result.

To prove (c), assume first that X is connected. Let f be a real-valued eigenfunction associated with the eigenvalue k. We have to prove that f is constant. As before, let $x \in V$ be a vertex such that $|f(x)| = \max_{y \in V} |f(y)|$. As $f(x) = \frac{(Af)(x)}{k} = \sum_{y \in V} \frac{A_{xy}}{k} f(y)$, we see that $f(x)$ is a convex combination of real numbers which are, in modulus, less than $|f(x)|$. This implies that $f(y) = f(x)$ for every $y \in V$, such that $A_{xy} \neq 0$, that is, for every y adjacent to x. Then, by a similar argument, f has the same value $f(x)$ on every vertex adjacent to such a y, and so on. Since X is connected, f must be constant.

We leave the proof of the converse as an exercise. $\square$

Proposition 1.1.2(c) shows a first connection between spectral properties of the adjacency matrix and combinatorial properties of the graph. This is one of the themes of this chapter.

1.1.3. Definition. A graph $X = (V, E)$ is *bipartite*, or *bicolorable*, if there exists a partition of the vertices $V = V_+ \cup V_-$, such that, for any two vertices x, y with $A_{xy} \neq 0$, if $x \in V_+$ (resp. V_-), then $y \in V_-$ (resp. V_+).

In other words, it is possible to paint the vertices with two colors in such a way that no two adjacent vertices have the same color. Bipartite graphs have very nice spectral properties characterized by the following:

1.1.4. Proposition. Let X be a connected, k-regular graph on n vertices. The following are equivalent:

(i) X is bipartite;
(ii) the spectrum of X is symmetric about 0;
(iii) $\mu_{n-1} = -k$.

Proof.

(i) $\Rightarrow$ (ii) Assume that $V = V_+ \cup V_-$ is a bipartition of X. To show symmetry of the spectrum, we assume that f is an eigenfunction of A with associated eigenvalue μ. Define

$$g(x) = \begin{cases} f(x) & \text{if } x \in V_+ \\ -f(x) & \text{if } x \in V_- \end{cases}.$$

It is then straightforward to show that $(Ag)(x) = -\mu\, g(x)$ for every $x \in V$.

(ii) $\Rightarrow$ (iii) This is clear from Proposition 1.1.2.

(iii) $\Rightarrow$ (i) Let f be a real-valued eigenfunction of A with eigenvalue $-k$.

Let $x \in V$ be such that $|f(x)| = \max_{y \in V} |f(y)|$. Replacing f by $-f$ if necessary, we may assume $f(x) > 0$. Now

$$f(x) = -\frac{(Af)(x)}{k} = -\sum_{y \in V} \frac{A_{xy}}{k} f(y) = \sum_{y \in V} \frac{A_{xy}}{k} (-f(y)).$$

So $f(x)$ is a convex combination of the $-f(y)$'s which are, in modulus, less than $|f(x)|$. Therefore, $-f(y) = f(x)$ for every $y \in V$, such that $A_{xy} \neq 0$, that is, for every y adjacent to x. Similarly, if z is a vertex adjacent to any such y, then $f(z) = -f(y) = f(x)$. Define $V_+ = \{y \in V : f(y) > 0\}$, $V_- = \{y \in V : f(y) < 0\}$; because X is connected, this defines a bipartition of X. □

Thus, every finite, connected, k-regular graph X has largest positive eigenvalue $\mu_0 = k$; if, in addition, X is bipartite, then the eigenvalue $\mu_{n-1} = -k$ also occurs (and only in this case). These eigenvalues k and $-k$, if the second occurs, are called the *trivial* eigenvalues of X. The difference $k - \mu_1 = \mu_0 - \mu_1$ is the *spectral gap* of X.

Exercises on Section 1.1

1. For the complete graph K_n and the cycle C_n, write down the adjacency matrix and compute the spectrum of the graph (with multiplicities). When are these graphs bipartite?
2. Let D_n be the following graph on $2n$ vertices: $V = \mathbb{Z}/n\mathbb{Z} \times \{0, 1\}$; $E = \{\{(i, j), (i + 1, j)\} : i \in \mathbb{Z}/n\mathbb{Z}, j \in \{0, 1\}\} \cup \{\{(i, 0), (i, 1)\} : i \in \mathbb{Z}/n\mathbb{Z}\}$. Make a drawing and repeat exercise 1 for D_n.

3. Show that a graph is bipartite if and only if it has no circuit with odd length.
4. Let X be a finite, k-regular graph. Complete the proof of Proposition 1.1.2 by showing that the multiplicity of the eigenvalue k is equal to the number of connected components of X (Hint: look at the space of locally constant functions on X.)
5. Let X be a finite, simple graph without loop. Assume that, for some $r \geq 2$, it is possible to find a set of r vertices all having the same neighbors. Show that 0 is an eigenvalue of A, with multiplicity at least $r - 1$.
6. Let X be a finite, simple graph without loop, on n vertices, with eigenvalues $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_{n-1}$. Show that $\sum_{i=0}^{n-1} \mu_i = 0$, that $\sum_{i=0}^{n-1} \mu_i^2$ is twice the number of edges in X, and that $\sum_{i=0}^{n-1} \mu_i^3$ is six times the number of triangles in X.
7. Let $X = (V, E)$ be a graph, not necessarily finite. We say that X has *bounded degree* if there exists $N \in \mathbb{N}$, such that, for every $x \in V$, one has $\sum_{y \in V} A_{xy} \leq N$. Show that in this case, for any $f \in \ell^2(V)$, one has
$$\|Af\|_2 = \left(\sum_{x \in V} |(Af)(x)|^2\right)^{1/2} \leq N \cdot \|f\|_2 = N \cdot \left(\sum_{x \in V} |f(x)|^2\right)^{1/2};$$
that is, A is a bounded linear operator on the Hilbert space $\ell^2(V)$ (Hint: use the Cauchy–Schwarz inequality.)

1.2. Inequalities on the Spectral Gap

Let $X = (V, E)$ be a graph. For $F \subseteq V$, we define the *boundary* ∂F of F to be the set of edges with one extremity in F and the other in $V - F$. In other words, ∂F is the set of edges connecting F to $V - F$. Note that $\partial F = \partial(V - F)$.

1.2.1. Definition. The *isoperimetric constant*, or *expanding constant* of the graph X, is
$$h(X) = \inf\left\{\frac{|\partial F|}{\min\{|F|, |V - F|\}} : F \subseteq V, 0 < |F| < +\infty\right\}.$$
Note that, if X is finite on n vertices, this can be rephrased as $h(X) = \min\left\{\frac{|\partial F|}{|F|} : F \subseteq V, 0 < |F| \leq \frac{n}{2}\right\}$.

1.2.2. Definition. Let $(X_m)_{m\geq 1}$ be a family of finite, connected, k-regular graphs with $|V_m| \to +\infty$ as $m \to +\infty$. We say that $(X_m)_{m\geq 1}$ is a *family of expanders* if there exists $\varepsilon > 0$, such that $h(X_m) \geq \varepsilon$ for every $m \geq 1$.

1.2.3. Theorem. Let $X = (V, E)$ be a finite, connected, k-regular graph without loops. Let μ_1 be the first nontrivial eigenvalue of X (as in section 1.1). Then

$$\frac{k - \mu_1}{2} \leq h(X) \leq \sqrt{2k(k - \mu_1)}.$$

Proof. (a) We begin with the first inequality. We endow the set E of edges with an arbitrarily chosen orientation, allowing one to associate, to any edge $e \in E$, its origin e^- and its extremity e^+. This allows us to define the *simplicial coboundary operator* $d : \ell^2(V) \to \ell^2(E)$, where, for $f \in \ell^2(V)$ and $e \in E$,

$$df(e) = f(e^+) - f(e^-).$$

Endow $\ell^2(V)$ with the hermitian scalar product

$$\langle f \mid g \rangle = \sum_{x\in V} \overline{f(x)}\, g(x)$$

and $\ell^2(E)$ with the analogous one. So we may define the adjoint (or conjugate-transpose) operator $d^* : \ell^2(E) \to \ell^2(V)$, characterized by $\langle df \mid g \rangle = \langle f \mid d^*g \rangle$ for every $f \in \ell^2(V)$, $g \in \ell^2(E)$. Define a function $\delta : V \times E \to \{-1, 0, 1\}$ by

$$\delta(x, e) = \begin{cases} 1 & \text{if } x = e^+ \\ -1 & \text{if } x = e^- \\ 0 & \text{otherwise.} \end{cases}$$

Then one checks easily that, for $e \in E$ and $f \in \ell^2(V)$,

$$df(e) = \sum_{x\in V} \delta(x, e)\, f(x);$$

while, for $v \in V$ and $g \in \ell^2(E)$,

$$d^*g(x) = \sum_{e\in E} \delta(x, e)\, g(e).$$

We then define the *combinatorial Laplace operator* $\Delta = d^*d : \ell^2(V) \to \ell^2(V)$. It is easy to check that

$$\Delta = k \cdot \mathrm{Id} - A;$$

in particular, Δ does not depend on the choice of the orientation. For an orthonormal basis of eigenfunctions of A, the operator Δ takes the form

$$\Delta = \begin{pmatrix} 0 & & & \\ & k-\mu_1 & & \bigcirc \\ & & \ddots & \\ \bigcirc & & & k-\mu_{n-1} \end{pmatrix},$$

the eigenvalue 0 corresponding to the constant functions on V. Therefore, if f is a function on V with $\sum_{x\in V} f(x) = 0$ (i.e., f is orthogonal to the constant functions in $\ell^2(V)$), we have

$$\|df\|_2^2 = \langle df \mid df\rangle = \langle \Delta f \mid f\rangle \geq (k-\mu_1)\,\|f\|_2^2.$$

We apply this to a carefully chosen function f. Fix a subset F of V and set

$$f(x) = \begin{cases} |V-F| & \text{if } x \in F \\ -|F| & \text{if } x \in V-F. \end{cases}$$

Then $\sum_{x\in V} f(x) = 0$ and $\|f\|_2^2 = |F|\,|V-F|^2 + |V-F|\,|F|^2 = |F|\,|V-F|\,|V|$. Moreover,

$$df(e) = \begin{cases} 0 \text{ if } e \text{ connects two vertices either in } F \text{ or in } V-F; \\ \pm|V| \text{ if } e \text{ connects a vertex in } F \text{ with a vertex in } V-F. \end{cases}$$

Hence, $\|df\|_2^2 = |V|^2\,|\partial F|$. So the previous inequality gives

$$|V|^2\,|\partial F| \geq (k-\mu_1)\,|F|\,|V-F|\,|V|.$$

Hence,

$$\frac{|\partial F|}{|F|} \geq (k-\mu_1)\,\frac{|V-F|}{|V|}.$$

If we assume $|F| \leq \frac{|V|}{2}$, we get $\frac{|\partial F|}{|F|} \geq \frac{k-\mu_1}{2}$; hence, by definition, $h(X) \geq \frac{k-\mu_1}{2}$.

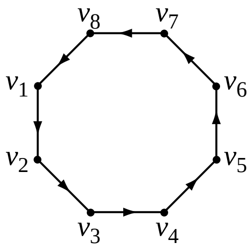
(b) We now turn to the second inequality, which is more involved. Fix a nonnegative function f on V, and set

$$B_f = \sum_{e\in E} |f(e^+)^2 - f(e^-)^2|.$$

Denote by $\beta_r > \beta_{r-1} > \cdots > \beta_1 > \beta_0$ the values of f, and set

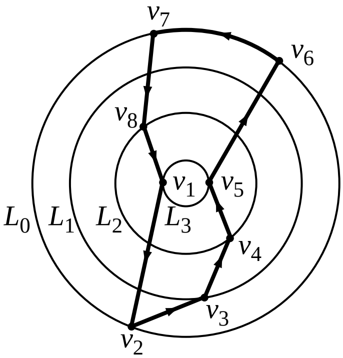
$$L_i = \{x \in V : f(x) \geq \beta_i\} \qquad (i = 0, 1, \ldots, r).$$

Note that $L_0 = V$. (Hence, $\partial L_0 = \emptyset$.) To have a better intuition of what is happening, consider the following example on C_8, the cycle graph with eight vertices.



with $f(v_1) = f(v_5) = 4$, $f(v_2) = f(v_6) = f(v_7) = 1$, $f(v_3) = 2$, $f(v_4) = f(v_8) = 3$, so that $\beta_3 = 4 > \beta_2 = 3 > \beta_1 = 2 > \beta_0 = 1$. Then

$$\begin{aligned}
L_0 &= \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8\};\\
L_1 &= \{v_1, v_3, v_4, v_5, v_8\};\\
L_2 &= \{v_1, v_4, v_5, v_8\};\\
L_3 &= \{v_1, v_5\};\\
\partial L_0 &= \emptyset;\\
\partial L_1 &= \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_5, v_6\}, \{v_7, v_8\}\};\ |\partial L_1| = 4;\\
\partial L_2 &= \{\{v_1, v_2\}, \{v_3, v_4\}, \{v_5, v_6\}, \{v_7, v_8\}\};\ |\partial L_2| = 4;\\
\partial L_3 &= \{\{v_1, v_2\}, \{v_4, v_5\}, \{v_5, v_6\}, \{v_8, v_1\}\};\ |\partial L_3| = 4.
\end{aligned}$$



Geometrically, one can envision the graph broken into level curves as follows: L_0 consists of all vertices on or inside the outer-level curve corresponding to $\beta_0 = 1$; L_1 consists of all vertices on or inside the level curve corresponding to $\beta_1 = 2$; and so forth. Then any ∂L_i consists of those edges that reach "downward" from inside L_i to a vertex with a lower value. From the diagram we see clearly that, for example, $\partial L_2 = \{\{v_1, v_2\}, \{v_3, v_4\}, \{v_5, v_6\}, \{v_7, v_8\}\}$.

Coming back to the general case, we now prove the following result about the number B_f.

First Step. $B_f = \sum_{i=1}^{r} |\partial L_i| (\beta_i^2 - \beta_{i-1}^2)$.

To see this, we denote by E_f the set of edges $e \in E$, such that $f(e^+) \neq f(e^-)$. Clearly $B_f = \sum_{e \in E_f} |f(e^+)^2 - f(e^-)^2|$. Now, an edge $e \in E_f$ connects some vertex x with $f(x) = \beta_{i(e)}$ to some vertex y with $f(y) = \beta_{j(e)}$. We index these two index values so that $i(e) > j(e)$. Therefore,

$$\begin{aligned} B_f &= \sum_{e \in E_f} (\beta_{i(e)}^2 - \beta_{j(e)}^2) \\ &= \sum_{e \in E_f} (\beta_{i(e)}^2 - \beta_{i(e)-1}^2 + \beta_{i(e)-1}^2 - \cdots - \beta_{j(e)+1}^2 + \beta_{j(e)+1}^2 - \beta_{j(e)}^2) \\ &= \sum_{e \in E_f} \sum_{\ell = j(e)+1}^{i(e)} (\beta_\ell^2 - \beta_{\ell-1}^2). \end{aligned}$$

Referring to the diagram of level curves, we see that as a given edge e connects a vertex x, with $f(x) = \beta_{i(e)}$, to a vertex y with $f(y) = \beta_{j(e)}$, it crosses every level curve β_ℓ between those two. In the expression for B_f, this corresponds to expanding the term $\beta_{i(e)}^2 - \beta_{j(e)}^2$ by inserting the zero difference $-\beta_\ell^2 + \beta_\ell^2$ for each level curve β_ℓ crossed by the edge e. This means that, in the previous summation for B_f, the term $\beta_\ell^2 - \beta_{\ell-1}^2$ appears for every edge e connecting some vertex x with $f(x) = \beta_i$ and $i \geq \ell$ to some vertex y with $f(y) = \beta_j$ and $j < \ell$. In other words, it appears for every edge $e \in \partial L_\ell$, which establishes the first step.

Second Step. $B_f \leq \sqrt{2k}\, \|df\|_2 \, \|f\|_2$.

Indeed,

$$\begin{aligned} B_f &= \sum_{e \in E} |f(e^+) + f(e^-)| \cdot |f(e^+) - f(e^-)| \\ &\leq \left[\sum_{e \in E} (f(e^+) + f(e^-))^2 \right]^{1/2} \left[\sum_{e \in E} (f(e^+) - f(e^-))^2 \right]^{1/2} \\ &\leq \sqrt{2} \left[\sum_{e \in E} (f(e^+)^2 + f(e^-)^2) \right]^{1/2} \|df\|_2 \\ &= \sqrt{2k} \left[\sum_{x \in V} f(x)^2 \right]^{1/2} \|df\|_2 = \sqrt{2k}\, \|f\|_2 \, \|df\|_2 \end{aligned}$$

by the Cauchy–Schwarz inequality and the elementary fact that $(a+b)^2 \le 2(a^2+b^2)$.

Third Step. Recall that the *support* of f is $\operatorname{supp} f = \{x \in V : f(x) \ne 0\}$. Assume that $|\operatorname{supp} f| \le \frac{|V|}{2}$. Then, $B_f \ge h(X)\,\|f\|_2^2$.

To see this, notice that $\beta_0 = 0$ and that $|L_i| \le \frac{|V|}{2}$ for $i = 1, \ldots, r$, so that $|\partial L_i| \ge h(X)\,|L_i|$ by definition of $h(X)$. So it follows from the first step that

$$\begin{aligned}
B_f &\ge h(X) \sum_{i=1}^{r} |L_i|\,(\beta_i^2 - \beta_{i-1}^2) \\
&= h(X)\left[|L_r|\,\beta_r^2 + (|L_{r-1}| - |L_r|)\,\beta_{r-1}^2 + \cdots + (|L_1| - |L_2|)\,\beta_1^2\right] \\
&= h(X)\left[|L_r|\,\beta_r^2 + \sum_{i=1}^{r-1} |L_i - L_{i+1}|\,\beta_i^2\right];
\end{aligned}$$

however, since $L_i - L_{i+1}$ is exactly the level set where f takes the value β_i, the term in brackets is exactly $\|f\|_2^2$.

Coda. We now apply this to a carefully chosen function f. Let g be a real-valued eigenfunction for Δ, associated with the eigenvalue $k - \mu_1$. Set $V^+ = \{x \in V : g(x) > 0\}$ and $f = \max\{g, 0\}$. By replacing g by $-g$ if necessary, we may assume $|V^+| \le \frac{|V|}{2}$. (Note that $V^+ \ne \emptyset$ because $\sum_{x\in V} g(x) = 0$ and $g \ne 0$.) For $x \in V^+$, we have (since $g \le 0$ on $V - V^+$)

$$\begin{aligned}
(\Delta f)(x) &= kf(x) - \sum_{y\in V} A_{xy}\, f(y) = kg(x) - \sum_{y\in V^+} A_{xy}\, g(y) \\
&\le kg(x) - \sum_{y\in V} A_{xy}\, g(y) = (\Delta g)(x) = (k - \mu_1)\, g(x).
\end{aligned}$$

Using this pointwise estimate, we get

$$\begin{aligned}
\|df\|_2^2 = \langle \Delta f \mid f\rangle &= \sum_{x\in V^+} (\Delta f)(x)\, g(x) \le (k-\mu_1) \sum_{x\in V^+} g(x)^2 \\
&\le (k-\mu_1)\,\|f\|_2^2.
\end{aligned}$$

Combining the second and third steps, we get

$$h(X)\,\|f\|_2^2 \le B_f \le \sqrt{2k}\,\|df\|_2\,\|f\|_2 \le \sqrt{2k\,(k-\mu_1)}\,\|f\|_2^2,$$

and the result follows by cancelling out. $\square$

From Definition 1.2.2 and Theorem 1.2.3, we immediately deduce the following:

1.2.4. Corollary. Let $(X_m)_{m\geq 1}$ be a family of finite, connected, k-regular graphs without loops, such that $|V_m| \to +\infty$ as $m \to +\infty$. The family $(X_m)_{m\geq 1}$ is a family of expanders if and only if there exists $\varepsilon > 0$, such that $k - \mu_1(X_m) \geq \varepsilon$ for every $m \geq 1$.

This is the spectral characterization of families of expanders: a family of k-regular graphs is a family of expanders if and only if the spectral gap is bounded away from zero. Moreover, it follows from Theorem 1.2.3 that, the bigger the spectral gap, the better "the quality" of the expander.

Exercises on Section 1.2

1. How was the assumption "X has no loop" used in the proof of Theorem 1.2.3?
2. Let X be a finite graph without loop. Choose an orientation on the edges; let d, d^* and $\Delta = d^*d$ be the operators defined in this section. Check that, for $f \in \ell^2(V)$, $x \in V$,

$$\Delta f(x) = \deg(x)\, f(x) - (Af)(x),$$

   where $\deg(x)$ is the *degree* of x, i.e., the number of neighboring vertices of x.
3. Using the example given for a function f on the cycle graph C_8, verify that B_f satisfies the first two steps in the proof of the second inequality of Theorem 1.2.3.
4. Show that the multiplicity of the eigenvalue $\mu_0 = K$ is the number of connected components of X.

1.3. Asymptotic Behavior of Eigenvalues in Families of Expanders

We have seen in Corollary 1.2.4 that the quality of a family of expanders can be measured by a lower bound on the spectral gap. However, it turns out that, asymptotically, the spectral gap cannot be too large. All the graphs in this section are supposed to be without loops.

1.3.1. Theorem. Let $(X_m)_{m\geq 1}$ be a family of connected, k-regular, finite graphs, with $|V_m| \to +\infty$ as $m \to +\infty$. Then,

$$\liminf_{m\to+\infty} \mu_1(X_m) \geq 2\sqrt{k-1}.$$

A stronger result will actually be proved in section 1.4. There is an asymptotic threshold, analogous to Theorem 1.3.1, concerning the bottom of the spectrum. Before stating it, we need an important definition.

1.3.2. Definition. The *girth* of a connected graph X, denoted by $g(X)$, is the length of the shortest circuit in X. We will say that $g(X) = +\infty$ if X has no circuit, that is, if X is a tree.

For a finite, connected, k-regular graph, let $\mu(X)$ be the smallest nontrivial eigenvalue of X.

1.3.3. Theorem. Let $(X_m)_{m\geq 1}$ be a family of connected, k-regular, finite graphs, with $g(X_m) \to +\infty$ as $m \to +\infty$. Then

$$\limsup_{m\to+\infty} \mu(X_m) \leq -2\sqrt{k-1}.$$

Theorems 1.3.1 and 1.3.3 single out an extremal condition on finite k-regular graphs, leading to the main definition.

1.3.4. Definition. A finite, connected, k-regular graph X is a *Ramanujan graph* if, for every nontrivial eigenvalue μ of X, one has $|\mu| \leq 2\sqrt{k-1}$.

Assume that $(X_m)_{m\geq 1}$ is a family of k-regular Ramanujan graphs without loop, such that $|V_m| \to +\infty$ as $m \to +\infty$. Then the X_m's achieve the biggest possible spectral gap, providing a family of expanders which is optimal from the spectral point of view.

All known constructions of infinite families of Ramanujan graphs involve deep results from number theory and/or algebraic geometry. As explained in the Overview, our purpose in this book is to give, for every odd prime p, a construction of a family of $(p+1)$-regular Ramanujan graphs. The original proof that these graphs satisfy the relevant spectral estimates, due to Lubotzky-Phillips, and Sarnak [42], appealed to Ramanujan's conjecture on coefficients of modular forms with weight 2: this explains the chosen terminology. Note that Ramanujan's conjecture was established by Eichler [23].

Exercises on Section 1.3

1. A tree is a connected graph without loops. Show that a k-regular tree T_k must be infinite and that it exists and is unique up to graph isomorphism.

2. Let X be a finite k-regular graph. Fix a vertex x_0 and, for $r < \frac{g(X)}{2}$, consider the ball centered at x_0 and of radius r in X. Show that it is isometric to any ball with the same radius in the k-regular tree T_k. Compute the cardinality of such a ball.
3. Deduce that, if $(X_m)_{m\geq 1}$ is a family of connected k-regular graphs, such that $|V_m| \to +\infty$ as $m \to +\infty$, then

$$g(X_m) \leq (2 + o(1)) \log_{k-1} |V_m|,$$

where $o(1)$ is a quantity tending to 0 as $m \to +\infty$.
4. Show that, if $k \geq 5$, one has actually, in exercise 3,

$$g(X_m) \leq 2 + 2 \log_{k-1} |V_m|.$$

1.4. Proof of the Asymptotic Behavior

In this section we prove a stronger result than that stated in Theorem 1.3.1.

The source of the inequality in Theorem 1.3.1 is the fact that the number of paths of length m from a vertex v to v, in a k-regular graph, is at least the number of such paths from v to v in a k-regular tree. To refine this observation, we count paths without backtracking, and to do this we introduce certain polynomials in the adjacency operator.

Let $X = (V, E)$ be a k-regular, simple graph, with $|V|$ possibly infinite. Recall that we defined a path in X in the Overview. We refine that definition now. A path of length r *without backtracking* in X is a sequence

$$\underline{e} = (x_0, x_1, \ldots, x_r)$$

of vertices in V such that x_i is adjacent to x_{i+1} $(i = 0, \ldots, r-1)$ and $x_{i+1} \neq x_{i-1}$ $(i = 1, \ldots, r-1)$. The origin of $\underline{e}$ is x_0, the extremity of $\underline{e}$ is x_r. We define, for $r \in \mathbb{N}$, matrices A_r indexed by $V \times V$, which generalize the adjacency matrix and which are polynomials in A:

$$(A_r)_{xy} = \begin{array}{l}\text{number of paths of length } r, \text{ without backtracking,}\\ \text{with origin } x \text{ and extremity } y.\end{array}$$

Note that $A_0 = \mathrm{Id}$ and that $A_1 = A$, the adjacency matrix. The relationship between A_r and A is the following:

1.4.1. Lemma.

(a) $A_1^2 = A_2 + k \cdot \mathrm{Id}$.

(b) For $r \geq 2$, $A_1 A_r = A_r A_1 = A_{r+1} + (k-1) A_{r-1}$.

Proof.

(a) For $x, y \in V$, the entry $(A_1^2)_{xy}$ is the number of all paths of length 2 between x and y. If $x \neq y$, such paths cannot have backtracking; hence, $(A_1^2)_{xy} = (A_2)_{xy}$. If $x = y$, we count the number of paths of length 2 from x to x, and, since X is simple, $(A_1^2)_{xx} = k$.

(b) Let us prove that $A_r A_1 = A_{r+1} + (k-1)\,A_{r-1}$ for $r \geq 2$. For $x, y \in V$, the entry $(A_r A_1)_{xy}$ is the number of paths $(x_0 = x, x_1, \ldots, x_r, x_{r+1} = y)$ of length $r+1$ between x and y, without backtracking except possibly on the last step (i.e., $(x_0, x_1, \ldots, x_r)$ has no backtracking). We partition the set of such paths into two classes according to the value of x_{r-1}:

- if $x_{r-1} \neq y$, then the path $(x_0, \ldots, x_{r+1})$ has no backtracking, and there are $(A_{r+1})_{xy}$ such paths;
- if $x_{r-1} = y$, then there is backtracking at the last step, and there are $(k-1)(A_{r-1})_{xy}$ such paths.

We leave the proof of $A_1 A_r = A_{r+1} + (k-1)\,A_{r-1}$ as an exercise. $\square$

From Lemma 1.4.1, we can compute the *generating function* of the A_r's, that is, the formal power series with coefficients A_r. It turns out to have a particularly nice expression; namely, we have the following:

1.4.2. Lemma.

$$\sum_{r=0}^{\infty} A_r\, t^r = \frac{1-t^2}{1 - At + (k-1)\,t^2}.$$

(This must be understood as follows: in the ring $\operatorname{End}\ell^2(V)[[t]]$ of formal power series over $\operatorname{End}\ell^2(V)$, we have

$$\left(\sum_{r=0}^{\infty} A_r\, t^r\right)(\mathrm{Id} - At + (k-1)\,t^2\,\mathrm{Id}) = (1-t^2)\,\mathrm{Id}.)$$

Proof. This is an easy check using Lemma 1.4.1. $\square$

In order to eliminate the numerator $1 - t^2$ in the right-hand side of 1.4.2, we introduce polynomials T_m in A given by

$$T_m = \sum_{0 \leq r \leq \frac{m}{2}} A_{m-2r} \qquad (m \in \mathbb{N}).$$

The generating function of the T_m's is readily computed.

1.4.3. Lemma.

$$\sum_{m=0}^{\infty} T_m\, t^m = \frac{1}{1 - At + (k-1)\, t^2}.$$

Proof.

$$\begin{aligned}\sum_{m=0}^{\infty} T_m\, t^m &= \sum_{m=0}^{\infty} \sum_{0 \le r \le \frac{m}{2}} A_{m-2r}\, t^m = \sum_{r=0}^{\infty} \sum_{m \ge 2r} A_{m-2r}\, t^m \\ &= \sum_{r=0}^{\infty} t^{2r} \sum_{m \ge 2r} A_{m-2r}\, t^{m-2r} = \left(\sum_{r=0}^{\infty} t^{2r}\right)\left(\sum_{\ell=0}^{\infty} A_\ell\, t^\ell\right) \\ &= \frac{1}{1-t^2} \cdot \frac{1-t^2}{1 - At + (k-1)\, t^2} = \frac{1}{1 - At + (k-1)\, t^2}\end{aligned}$$

by Lemma 1.4.2. □

1.4.4. Definition. The *Chebyshev polynomials of the second kind* are defined by expressing $\frac{\sin(m+1)\theta}{\sin\theta}$ as a polynomial of degree m in $\cos\theta$:

$$U_m(\cos\theta) = \frac{\sin(m+1)\,\theta}{\sin\theta} \qquad (m \in \mathbb{N}).$$

For example, $U_0(x) = 1$, $U_1(x) = 2x$, $U_2(x) = 4x^2 - 1, \ldots$. Using trigonometric identities, we see that these polynomials satisfy the following recurrence relation:

$$U_{m+1}(x) = 2x\, U_m(x) - U_{m-1}(x).$$

As in Lemma 1.4.2, from this recurrence relation, we compute the generating function of the U_m's; namely,

$$\sum_{m=0}^{\infty} U_m(x)\, t^m = \frac{1}{1 - 2xt + t^2}.$$

Performing a simple change of variables, we then compute the generating function of the related family of polynomials $(k-1)^{\frac{m}{2}}\, U_m\left(\frac{x}{2\sqrt{k-1}}\right)$:

$$\sum_{m=0}^{\infty} (k-1)^{\frac{m}{2}}\, U_m\left(\frac{x}{2\sqrt{k-1}}\right) t^m = \frac{1}{1 - xt + (k-1)\, t^2}.$$

In comparison to Lemma 1.4.3, we immediately get the following expression for the operators T_m as polynomials of degree m in the adjacency matrix.

1.4.5. Proposition. For $m \in \mathbb{N}$: $T_m = (k-1)^{\frac{m}{2}} U_m\left(\frac{A}{2\sqrt{k-1}}\right)$. $\square$

Assume that $X = (V, E)$ is a finite, k-regular graph on n vertices, with spectrum

$$\mu_0 = k \geq \mu_1 \geq \cdots \geq \mu_{n-1}.$$

In Proposition 1.4.5, we are going to estimate the trace of T_m in two different ways. This will lead to the trace formula for X.

First, working from a basis of eigenfunctions of A, we have, from Proposition 1.4.5,

$$\operatorname{Tr} T_m = (k-1)^{\frac{m}{2}} \sum_{j=0}^{n-1} U_m\left(\frac{\mu_j}{2\sqrt{k-1}}\right).$$

On the other hand, by definition of T_m,

$$\operatorname{Tr} T_m = \sum_{0 \leq r \leq \frac{m}{2}} \operatorname{Tr} A_{m-2r} = \sum_{x \in V} \sum_{0 \leq r \leq \frac{m}{2}} (A_{m-2r})_{xx}.$$

For $x \in V$, denote by $f_{\ell,x}$ the number of paths of length ℓ in X, without backtracking, with origin and extremity x; in other words, $f_{\ell,x} = (A_\ell)_{xx}$. Then we get the trace formula:

1.4.6. Theorem.

$$\sum_{x \in V} \sum_{0 \leq r \leq \frac{m}{2}} f_{m-2r,x} = (k-1)^{\frac{m}{2}} \sum_{j=0}^{n-1} U_m\left(\frac{\mu_j}{2\sqrt{k-1}}\right),$$

for every $m \in \mathbb{N}$.

We say that X is *vertex-transitive* if the group $\operatorname{Aut} X$ of automorphisms of X acts transitively on the vertex-set V. Specifically, this means that for every pair of vertices x and y, there exists $\alpha \in \operatorname{Aut} X$, such that $\alpha(x) = y$. Under this assumption, the number $f_{\ell,x}$ does not depend on the vertex x, and we denote it simply by f_ℓ.

1.4.7. Corollary. Let X be a vertex-transitive, finite, k-regular graph on n vertices. Then, for every $m \in \mathbb{N}$,

$$n \cdot \sum_{0 \leq r \leq \frac{m}{2}} f_{m-2r} = (k-1)^{\frac{m}{2}} \sum_{j=0}^{n-1} U_m\left(\frac{\mu_j}{2\sqrt{k-1}}\right). \quad \square$$

The value of the trace formula 1.4.6 is the following: only looking at the right-hand side (called the spectral side) $(k-1)^{\frac{m}{2}} \sum_{j=0}^{n-1} U_m\left(\frac{\mu_j}{2\sqrt{k-1}}\right)$, it is not obvious that it defines a nonnegative integer. As we shall now explain, the mere positivity of the spectral side has nontrivial consequences. We first need a somewhat technical result about the Chebyshev polynomials.

1.4.8. Proposition. Let $L \geq 2$ and $\varepsilon > 0$ be real numbers. There exists a constant $C = C(\varepsilon, L) > 0$ with the following property: for any probability measure ν on $[-L, L]$, such that $\int_{-L}^{L} U_m\left(\frac{x}{2}\right) d\nu(x) \geq 0$ for every $m \in \mathbb{N}$, we have

$$\nu[2-\varepsilon, L] \geq C.$$

(Thus, ν gives a measure at least C to the interval $[2-\varepsilon, L]$.)

Proof. It is convenient to introduce the polynomials $X_m(x) = U_m\left(\frac{x}{2}\right)$; they satisfy $X_m(2\cos\theta) = \frac{\sin(m+1)\theta}{\sin\theta}$ and the recursion formula $X_{m+1}(x) = x\,X_m(x) - X_{m-1}(x)$. It is clear from the first relation that the roots of X_m are $2\cos\frac{\ell\pi}{m+1}$ ($\ell = 1, \ldots, m$). In particular the largest root of X_m is $\alpha_m = 2\cos\frac{\pi}{m+1}$. The proof is then in several steps.

First Step. For $k \leq \ell : X_k\, X_\ell = \sum_{m=0}^{k} X_{k+\ell-2m}$.

We prove this by induction over k. Since $X_0(x) = 1$ and $X_1(x) = x$, the formula is obvious for $k = 0, 1$. (For $k = 1$, this is nothing but the recursion formula.) Then, for $k \geq 2$, we have, by induction hypothesis,

$$\begin{aligned}
X_k\, X_\ell &= (x\, X_{k-1} - X_{k-2})\, X_\ell \\
&= x\,(X_{k+\ell-1} + X_{k+\ell-3} + \cdots + X_{\ell-k+3} + X_{\ell-k+1}) \\
&\quad - (X_{k+\ell-2} + X_{k+\ell-4} + \cdots + X_{\ell-k+4} + X_{\ell-k+2}) \\
&= (X_{k+\ell} + X_{k+\ell-2}) + (X_{k+\ell-2} + X_{k+\ell-4}) \\
&\quad + \cdots + (X_{\ell-k+4} + X_{\ell-k+2}) + (X_{\ell-k+2} + X_{\ell-k}) \\
&\quad - (X_{k+\ell-2} + X_{k+\ell-4} + \cdots + X_{\ell-k+4} + X_{\ell-k+2}) \\
&= X_{k+\ell} + X_{k+\ell-2} + \cdots + X_{\ell-k+2} + X_{\ell-k}.
\end{aligned}$$

Second Step.

$$\frac{X_m(x)}{x - \alpha_m} = \sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m) \cdot X_i(x).$$

Indeed,

$$\begin{aligned}
&(x - \alpha_m)\left(\sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m)\, X_i(x)\right) \\
&= X_{m-1}(\alpha_m)\, X_1(x) + \sum_{i=1}^{m-1} X_{m-1-i}(\alpha_m)(X_{i+1}(x) + X_{i-1}(x)) \\
&\quad - \sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m)\, \alpha_m\, X_i(x) \\
&= (X_{m-2}(\alpha_m) - X_{m-1}(\alpha_m)\, \alpha_m)\, X_0(x) \\
&\quad + \sum_{i=1}^{m-2} (X_{m-i}(\alpha_m) + X_{m-i-2}(\alpha_m) - \alpha_m\, X_{m-1-i}(\alpha_m))\, X_i(x) \\
&\quad + (X_1(\alpha_m) - \alpha_m\, X_0(\alpha_m))\, X_{m-1}(x) + X_0(\alpha_m)\, X_m(x).
\end{aligned}$$

Now $X_0(\alpha_m) = 1$ and $X_1(\alpha_m) - \alpha_m\, X_0(\alpha_m) = 0$; in the summation $\sum_{i=1}^{m-2}$ all the coefficients are 0, by the recursion formula. Finally, $X_{m-2}(\alpha_m) - X_{m-1}(\alpha_m)\, \alpha_m = -X_m(\alpha_m) = 0$, by definition of α_m.

Third Step. Set $Y_m(x) = \frac{X_m(x)^2}{x - \alpha_m}$; then $Y_m = \sum_{i=0}^{2m-1} y_i\, X_i$, with $y_i \geq 0$.

Indeed, by the second step we have $Y_m = \sum_{i=0}^{m-1} X_{m-1-i}(\alpha_m)\, X_i\, X_m$. Now observe that the sequence $\alpha_m = 2 \cos \frac{\pi}{m+1}$ increases to 2. So for $j < m$: $X_j(\alpha_m) > 0$ (since $\alpha_m > \alpha_j$ and α_j is the largest root of X_j). This means that all coefficients are positive in the previous formula for Y_m. By the first step, each $X_i\, X_m$ is a linear combination, with nonnegative coefficients, of $X_0, X_1, \ldots, X_{2m-1}$, so the result follows.

Fourth Step. Fix $\varepsilon > 0$, $L \geq 2$. For every probability measure ν on $[-L, L]$ such that $\int_{-L}^{L} X_m(x)\, d\nu(x) \geq 0$ for every $m \in \mathbb{N}$, we have $\nu\, [2 - \varepsilon, L] > 0$.

Indeed, assume by contradiction that $\nu\, [2 - \varepsilon, L] = 0$; i.e. the support of ν is contained in $[-L, 2 - \varepsilon]$. Take m large enough to have $\alpha_m > 2 - \varepsilon$. Since $Y_m(x) \leq 0$ for $x \leq \alpha_m$, we then have $\int_{-L}^{L} Y_m(x)\, d\nu(x) \leq 0$. On the other hand, by the third step and the assumption on ν, we clearly have $\int_{-L}^{L} Y_m(x)\, d\nu(x) \geq 0$. So $\int_{-L}^{L} Y_m(x)\, d\nu(x) = 0$, which implies that ν is supported in the finite set F_m of zeroes of Y_m; as before we have $F_m = \{2\cos\frac{\ell\pi}{m+1} : 1 \leq \ell \leq m\}$. But this holds for every m large enough. And clearly, since $m+1$ and $m+2$ are relatively prime, we have $F_m \cap F_{m+1} = \emptyset$, so that supp ν is empty. But this is absurd.

Coda. Fix $\varepsilon > 0, L \geq 2$. Let f be the continuous function on $[-L, L]$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x \leq 2-\varepsilon \\ 1 & \text{if } x \geq 2-\frac{\varepsilon}{2} \\ \frac{2}{\varepsilon}(x-2+\varepsilon) & \text{if } 2-\varepsilon \leq x \leq 2-\frac{\varepsilon}{2}. \end{cases}$$

On $\left[2-\varepsilon, 2-\frac{\varepsilon}{2}\right]$, the function f linearly interpolates between 0 and 1. For every probability measure ν on $[-L, L]$, we then have

$$\nu\,[2-\varepsilon, L] \geq \int_{-L}^{L} f(x)\, d\nu(x) \geq \nu\left[2-\frac{\varepsilon}{2}, L\right].$$

Let $\wp$ be the set of probability measures ν on $[-L, L]$, such that $\int_{-L}^{L} X_m(x)\, d\nu(x) \geq 0$ for every $m \geq 1$. For $\nu \in \wp$, we have by the fourth step $\int_{-L}^{L} f(x)\, d\nu(x) > 0$. But $\wp$ is compact in the weak topology and, since f is continuous, the map

$$\wp \to \mathbb{R}^+ : \nu \mapsto \int_{-L}^{L} f(x)\, d\nu(x)$$

is weakly continuous. By compactness there exists $C(\varepsilon, L) > 0$, such that $\int_{-L}^{L} f(x)\, d\nu(x) \geq C(\varepsilon, L)$ for every $\nu \in \wp$. *A fortiori* $\nu\,[2-\varepsilon, L] \geq C(\varepsilon, L)$, and the proof is complete. (Note that, in the final step, the need for introducing the function f comes from the fact that the map $\wp \to \mathbb{R}^+ : \nu \mapsto \nu\,[2-\varepsilon, L]$ is, *a priori*, not weakly continuous; however, it is bounded below by a continuous function, to which the compactness argument applies.) $\square$

Coming back to the spectra of finite connected, k-regular graphs, we now reach the promised improvement of Theorem 1.3.1: it shows not only that the first nontrivial eigenvalue becomes asymptotically larger than $2\sqrt{k-1}$, but also that a positive proportion of eigenvalues lies in any interval $\left[(2-\varepsilon)\sqrt{k-1}, k\right]$.

1.4.9. Theorem. For every $\varepsilon > 0$, there exists a constant $C = C(\varepsilon, k) > 0$, such that, for every connected, finite, k-regular graph X on n vertices, the number of eigenvalues of X in the interval $\left[(2-\varepsilon)\sqrt{k-1}, k\right]$ is at least $C \cdot n$.

Proof. Take $L = \frac{k}{\sqrt{k-1}} \geq 2$ and $\nu = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\frac{\mu_j}{\sqrt{k-1}}}$ (where δ_a is the Dirac measure at $a \in [-L, L]$, that is, the probability measure on $[-L, L]$ such that $\int_{-L}^{L} f(x)\, d\,\delta_a(x) = f(a)$, for every continuous function f on $[-L, L]$). Then ν is a probability measure on $[-L, L]$, and $\int_{-L}^{L} U_m\left(\frac{x}{2}\right)\, d\nu(x) = \frac{1}{n} \sum_{j=0}^{n-1} U_m\left(\frac{\mu_j}{2\sqrt{k-1}}\right)$ is nonnegative, by the trace formula 1.4.6. So the assumptions of Proposition 1.4.8 are satisfied, and therefore there exists $C = C(\varepsilon, k) > 0$ such that $\nu\,[2-\varepsilon, L] \geq C$. But

$$
\begin{aligned}
\nu\,[2-\varepsilon, L] &= \frac{1}{n} \times (\text{number of } j\text{'s with } 2-\varepsilon \leq \frac{\mu_j}{\sqrt{k-1}} \leq L) \\
&= \frac{1}{n} \times (\text{number of eigenvalues of } X \text{ in } [(2-\varepsilon)\sqrt{k-1},\ k]).
\end{aligned}
$$

□

Continuing this analysis we prove the following:

1.4.10. Theorem. Let $(X_m)_{m\geq 1}$ be a sequence of connected, k-regular, finite graphs for which $g(X_m) \to \infty$ as $m \to \infty$. If $\nu_m = \nu(X_m)$ is the measure on $\left[-\frac{k}{\sqrt{k-1}}, \frac{k}{\sqrt{k-1}}\right]$ defined by

$$
\nu_m = \frac{1}{|X_m|} \sum_{j=0}^{|X_m|-1} \frac{\delta_{\mu_j}(X_m)}{\sqrt{k-1}},
$$

then, for every continuous function f on $\left[-\frac{k}{\sqrt{k-1}}, \frac{k}{\sqrt{k-1}}\right]$,

$$
\lim_{m\to\infty} \int_{\frac{-k}{\sqrt{k-1}}}^{\frac{k}{\sqrt{k-1}}} f(x)\, d\nu_m(x) = \int_{-2}^{2} f(x) \sqrt{4-x^2}\, \frac{dx}{2\pi}.
$$

In other words, the sequence of measures $(\nu_m)_{m\geq 1}$ on $\left[-\frac{k}{\sqrt{k-1}}, \frac{k}{\sqrt{k-1}}\right]$ weakly converges to the measure ν supported on $[-2, 2]$, given by $d\nu(x) = \frac{\sqrt{4-x^2}}{2\pi}\, dx$.

Proof. Set $L = \frac{k}{\sqrt{k-1}}$. Recall that $f_{\ell,x}$ denotes the number of paths of length ℓ, without backtracking, from x to x in X_m. We have that for $n \geq 1$, fixed and m large enough (precisely $g(X_m) > n$):

$$
f_{n-2r,x} = 0
$$

for any $x \in X_m$ and $0 \leq r \leq \frac{n}{2}$. Hence, for m large enough the left-hand side of the equation in Theorem 1.4.6 is zero. Thus, so is the right-hand side, and