3 Pseudorandom Graphs

Chapter Highlights

- Equivalent notions of graph quasirandomness
- Role of graph eigenvalues in pseudorandomness
- Expander mixing lemma
- Eigenvalues of abelian Cayley graphs and the Fourier transform
- Quasirandom groups and representations theory
- Quasirandom Cayley graphs and Grothendieck's inequality
- Alon–Boppana bound on the second eigenvalue of a d-regular graph

In the previous chapter on the graph regularity method, we saw that every graph can be partitioned into a bounded number of vertex parts so that the graph looks "random-like" between most pairs of parts. In this chapter, we dive further into how a graph can be random-like.

Pseudorandomness is a concept prevalent in combinatorics, theoretical computer science, and in many other areas. It specifies how a non-random object can behave like a truly random object.

Example 3.0.1 (Pseudorandom generators)**.** Suppose you want to generate a random number on a computer. In most systems and programming languages, you can do this easily with a single command (e.g., `rand()`). The output is not actually truly random. Instead, the output came from a *pseudorandom generator*, which is some function/algorithm that takes a *seed* as input, and passes it through some sophisticated function, so that there is no practical way to distinguish the output from a truly random object. In other words, the output is not actually truly random, but for all practical purposes the output cannot be distinguished from a truly random output.

Example 3.0.2 (Primes)**.** In number theory, the prime numbers behave like a random sequence in many ways. The celebrated *Riemann hypothesis* and its generalizations give quantitative predictions about how closely the primes behave in a certain specific way like a random sequence. There is also something called *Cramér's random model* for the primes that allows one to make predictions about the asymptotic density of certain patterns in the primes (e.g., how many twin primes up to N are there?). Empirical data support these predictions, and they have been proved in certain cases. Nevertheless, there are still notorious open problems such as the twin prime and Goldbach conjectures. Despite their pseudorandom behavior, the primes are not random!

Example 3.0.3 (Normal numbers)**.** It is very much believed that the digits of π behave in a random-like way, where every digit or block of digits appear with frequency similar to that of a truly random number. Such numbers are called *normal*. It is widely believed that numbers such as $\sqrt{2}$, π, and e are normal, but proofs remain elusive. Again, the digits of π are deterministic, not random, but they are believed to behave pseudorandomly. On the other hand, nearly all real numbers are normal, with the exceptions occupying only a measure zero subset of the reals.

Coming back to graph theory, in an **Erdős–Rényi random graph**, every edge occurs independently with some probability. Now, given some specific graph (perhaps an instance of the random graph, or perhaps generated via some other means), we can ask whether this graph, for the purpose of some intended application, behaves similarly to that of a typical random graph. What are some useful ways to *measure* the pseudorandomness of a graph? This is the main theme that we explore in this chapter.

3.1 Quasirandom Graphs

Here are several natural notions of how a graph (or rather, a sequence of graphs) can look random. The main theorem of this section says that, surprisingly, these notions are all equivalent. This result is due to Chung, Graham, and Wilson (1989), who coined the term **quasirandom graphs**. Similar ideas also appeared in the work of Thomason (1987). These results had an important impact in the field.

Theorem 3.1.1 (Quasirandom graphs)

Let $p \in [0,1]$ be fixed. Let (G_n) be a sequence of graphs with G_n having n vertices and $(p + o(1))\binom{n}{2}$ edges (here $n \to \infty$ along some subsequence of integers, i.e., is allowed to skip integers). Denote G_n by G. The following properties are all equivalent:

DISC (discrepancy) $e(X,Y) = p\,|X|\,|Y| + o(n^2)$ for all $X, Y \subset V(G)$.

DISC' $e(X) = p\binom{|X|}{2} + o(n^2)$ for all $X \subset V(G)$.

COUNT For every graph H, the number of labeled copies of H in G is $(p^{e(H)} + o(1))n^{v(H)}$.

(Here a labeled copy of H is the same as an injective map $V(H) \to V(G)$ that sends every edge of H to an edge of G. The rate that the $o(1)$ goes to zero is allowed to depend on H.)

C_4 (4-cycle) The number of labeled 4-cycles is at most $(p^4 + o(1))n^4$.

CODEG (codegree) Letting $\mathrm{codeg}(u,v)$ denote the number of common neigh-

bors of u and v,

$$\sum_{u,v \in V(G)} \left|\text{codeg}(u,v) - p^2 n\right| = o(n^3).$$

EIG (eigenvalue) If $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of the adjacency matrix of G, then $\lambda_1 = pn + o(n)$ and $\max_{i \neq 1} |\lambda_i| = o(n)$.

Definition 3.1.2 (Quasirandom graphs)
We say a sequence of graphs is **quasirandom** (at edge density p) if it satisfies the above conditions for some constant $p \in [0, 1]$.

Remark 3.1.3 (Single graph vs. a sequence of graphs)**.** Strictly speaking, it does not make sense to say whether a *single* graph is quasirandom, but we will abuse the definition as such when it is clear that the graph we are referring to is part of a sequence.

Remark 3.1.4 ($\mathbf{C_4}$ condition)**.** The $\mathbf{C_4}$ condition is surprising. It says that the 4-cycle density, a single statistic, is equivalent to all the other quasirandomness conditions.

We will soon see below in Proposition 3.1.14 that the $\mathbf{C_4}$ can be replaced by the equivalent condition that the number of labeled 4-cycles is $(p^4 + o(1))n^4$ (rather than at most this quantity).

Remark 3.1.5 (Checking quasirandomness)**.** The discrepancy conditions are hard to verify since they involve checking exponentially many sets. The other conditions can all be checked in time polynomial in the size of the graph. So the equivalence gives us an algorithmically efficient way to certify the discrepancy condition.

Remark 3.1.6 (Quantiative equivalences)**.** Rather than stating these properties for a sequence of graphs using a decaying error term $o(1)$, we can state a quantitative quasirandomness hypothesis for a specific graph using an error tolerance parameter ϵ. For example, we can restate the discrepancy condition as follows.

DISC(ϵ): For all $X, Y \subset V(G)$, $|e(X,Y) - p\,|X|\,|Y|| < \epsilon n^2$.

Similar statements can be made for other quasirandom graph notions. The proof below shows that these notions are equivalence up to a polynomial change in ϵ, i.e., for each pair of properties, **Prop1**(ϵ) implies **Prop2**($C\epsilon^c$) for some constants $C, c > 0$.

Examples of quasirandom graphs

First let us check that random graphs are quasirandom (hence justifying the name).

Recall the following basic tail bound for a sum of independent random variables.

Theorem 3.1.7 (Chernoff bound)
Let X be a sum of m independent Bernoulli random variables (not necessarily identically distributed). Then for every $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}X| \geq t) \leq 2e^{-t^2/(2m)}$$

Proposition 3.1.8
Let $p \in [0, 1]$ and $\epsilon > 0$. With probability at least $1 - 2^{n+1}e^{-\epsilon^2 n^2}$, the Erdős–Rényi random graph $\mathbf{G}(n, p)$ has the property that for every vertex subset X,

$$\left| e(X) - p\binom{|X|}{2} \right| \leq \epsilon n^2$$

Proof. Applying the Chernoff bound to $e(X)$, we see that

$$\mathbb{P}\left(\left| e(X) - p\binom{|X|}{2} \right| > \epsilon n^2 \right) \leq 2\exp\left(\frac{-(\epsilon n^2)^2}{2\binom{|X|}{2}} \right) \leq 2\exp\left(-\epsilon^2 n^2\right).$$

The result then follows by taking a union bound over all 2^n subsets X of the n-vertex graph. □

Applying the Borel–Cantelli lemma with the above bound, we obtain the following consequence.

Corollary 3.1.9 (Random graphs are quasirandom)
Fix $p \in [0, 1]$. With probability 1, a sequence of random graphs $G_n \sim \mathbf{G}(n, p)$ is quasirandom at edge density p.

It would be somewhat disappointing if the only interesting example of quasirandom graph were actual random graphs. Fortunately we have more explicit constructions. In the rest of the chapter, we will see several constructions using Cayley graphs on groups. A notable example, which we will prove in Section 3.3, is that the Paley graph is quasirandom.

Example 3.1.10 (Paley graph). Let $p \equiv 1 \pmod 4$ be a prime. Form a graph with vertex set $\mathbb{F}_p$, with two vertices x, y joined if $x - y$ is a quadratic residue. Then this graph is quasirandom at edge density $1/2$ as $p \to \infty$. (By a standard fact from elementary number theory, since $p \equiv 1 \pmod 4$, -1 is a quadratic residue, and hence $x - y$ is a quadratic residue if and only if $y - x$ is. So the graph is well defined.)

In Section 3.4, we will show that for certain sequence of groups, every sequence of Cayley graphs on them is quasirandom provided that the edge densities converge.

We will call such groups *quasirandom*. We will later prove the following important example.

Example 3.1.11 (PSL$(2, p)$)**.** Let p be a prime. Let $S \subset \mathrm{PSL}(2, p)$ be a subset of non-zero elements with $S = S^{-1}$. Let G be the Cayley graph on $\mathrm{PSL}(2, p)$ with generator S, meaning that the vertices are elements of $\mathrm{PSL}(2, p)$, and two vertices x, y are adjacent if $x^{-1}y \in S$. Then G is quasirandom as $p \to \infty$ as long as $|S|/p^3$ converges.

Finally, here is an explicit construction using finite geometry. We leave it as an exercise to verify its quasirandomness using the conditions given earlier.

Example 3.1.12. Let p be a prime. Let $S \subset \mathbb{F}_p \cup \{\infty\}$. Let G be a graph on vertex set $\mathbb{F}_p^2$ where two points are joined if the slope of the line connecting them lies in S. Then G is quasirandom as $p \to \infty$ as long as $|S|/p$ converges.

Exercise 3.1.13. Prove that the construction in Example 3.1.12 is quasirandom.

Proof of equivalence of graph quasirandomness conditions

We will now start to prove Theorem 3.1.1. Let us begin with a warm-up on how to use apply the Cauchy–Schwarz inequality in graph theory since it will come up several times in the proof (we will revisit this topic in Section 5.2).

The following statement says that the 4-cycle density is always roughly at least as much as random. Later in Chapter 5, we will see Sidorenko's conjecture, which says that all bipartite graphs have this property.

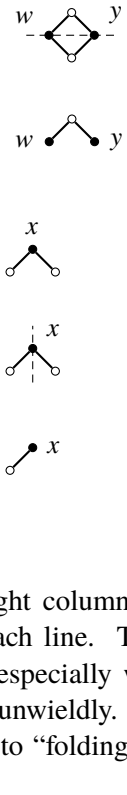
As a consequence, the $\mathbf{C_4}$ condition is equivalent to saying that the number of labeled 4-cycles *is* $(p^4 + o(1))n^4$ (rather than *at most*).

Proposition 3.1.14 (Minimum 4-cycle density)
Every n-vertex graph with at least $pn^2/2$ edges has at least p^4n^4 labeled closed walks of length 4.

Remark 3.1.15. Since all but $O(n^3)$ such closed walks use four distinct vertices, the above statement implies that the number of labeled 4-cycles is at least $(p^4 - o(1))n^4$.

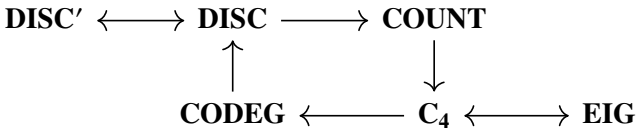
Proof. The number of closed walks of length 4 is

$$
\begin{aligned}
|\{(w,x,y,z) \text{ closed walk}\}| &= \sum_{w,y} |\{x : w \sim x \sim y\}|^2 \\
&\geq \frac{1}{n^2}\left(\sum_{w,y} |\{x : w \sim x \sim y\}|\right)^2 \\
&= \frac{1}{n^2}\left(\sum_{x} |\{(w,y) : w \sim x \sim y\}|\right)^2 \\
&= \frac{1}{n^2}\left(\sum_{x} (\deg x)^2\right)^2 \\
&\geq \frac{1}{n^4}\left(\sum_{x} \deg x\right)^4 \\
&= (2e(G))^4/n^4 \geq p^4 n^4
\end{aligned}
$$



Here both inequality steps are due to Cauchy–Schwarz. On the right column is a pictorial depiction of what is being counted by the inner sum on each line. These diagrams are a useful way to keep track of the graph inequalities, especially when dealing with much larger graphs, where the algebraic expressions get unwieldly. Note that each application of the Cauchy–Schwarz inequality corresponds to "folding" the graph along a line of reflection. □

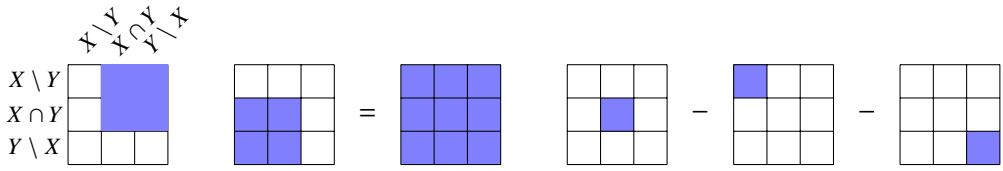
We shall prove the equivalences of Theorem 3.1.1 in the following way:



Proof that **DISC** *implies* **DISC′**. Take $Y = X$ in **DISC**. (Note that $e(X,X) = 2e(X)$ and $\binom{|X|}{2} = |X|^2/2 - O(n)$.) □

Proof that **DISC′** *implies* **DISC**. We have the following "polarization identity", together with a proof by picture (recall $2e(X) = e(X,X)$):

$$e(X,Y) = e(X \cup Y) + e(X \cap Y) - e(X \setminus Y) - e(Y \setminus X).$$



If **DISC′** holds, then the right-hand side above equals to

$$p\binom{|X \cup Y|}{2} + p\binom{|X \cap Y|}{2} + p\binom{|X \setminus Y|}{2} + p\binom{|Y \setminus X|}{2} + o(n^2) = p\,|X|\,|Y| + o(n^2),$$

where the final step applies the polarization identity again, this time on the complete graph. So we have $e(X,Y) = p\,|X|\,|Y| + o(n^2)$ thereby confirming **DISC**. □

Proof (deferred) that **DISC** *implies* **COUNT**. This is essentially a counting lemma. In Section 2.6 we proved a version of the counting lemma but for lower bounds. The same proof can be modified to a two-sided bound. We will see another proof of a counting lemma (Theorem 4.5.1) in the next chapter on graph limits, which gives us a convenient language to set up a more streamlined proof. So we will defer this proof until then. □

Proof that **COUNT** *implies* $\mathbf{C_4}$. $\mathbf{C_4}$ is a special case of **COUNT**. □

Proof that $\mathbf{C_4}$ *implies* **CODEG**. Assuming $\mathbf{C_4}$, we have

$$\sum_{u,v} \mathrm{codeg}(u,v) = \sum_{x \in G} \deg(x)^2 \geq \frac{1}{n}\left(\sum_{x \in G} \deg(x)\right)^2 = \frac{1}{n}\left(pn^2 + o(n^2)\right)^2 = p^2n^3 + o(n^3).$$

We also have (below the $O(n^3)$ error term is due to walks of length 4 that use repeated vertices)

$$\begin{aligned}\sum_{u,v} \mathrm{codeg}(u,v)^2 &= \#\text{ labeled } C_4 + O(n^3)\\ &\leq p^4n^4 + o(n^4).\end{aligned}$$

Thus, by the Cauchy–Schwarz inequality,

$$\begin{aligned}\frac{1}{n^2}\left(\sum_{u,v}\left|\mathrm{codeg}(u,v) - p^2n\right|\right)^2 &\leq \sum_{u,v}\left(\mathrm{codeg}(u,v) - p^2n\right)^2\\ &= \sum_{u,v}\mathrm{codeg}(u,v)^2 - 2p^2n\sum_{u,v}\mathrm{codeg}(u,v) + p^4n^4\\ &\leq p^4n^4 - 2p^2n \cdot p^2n^3 + p^4n^4 + o(n^4)\\ &= o(n^4).\end{aligned}$$

□

Remark 3.1.16. These calculations share the spirit of the *second moment method* in probabilistic combinatorics. The condition $\mathbf{C_4}$ says that the variance of the codegree of two random vertices is small.

Exercise 3.1.17. Show that if we modify the **COEG** condition to

$$\sum_{u,v\in V(G)} \Big(\operatorname{codeg}(u,v) - p^2 n\Big) = o(n^3),$$

then it would not be enough to imply quasirandomness.

Proof that **CODEG** *implies* **DISC**. We first show that the codegree condition implies the concentration of degrees:

$$\begin{aligned}
\frac{1}{n}\Big(\sum_u |\deg u - pn|\Big)^2 &\le \sum_u (\deg u - pn)^2 \\
&= \sum_u (\deg u)^2 - 2pn \sum_u \deg u + p^2 n^3 \\
&= \sum_{x,y} \operatorname{codeg}(x,y) - 4pn\, e(G) + p^2 n^3 \\
&= p^2 n^3 - 2p^2 n^3 + p^2 n^3 + o(n^3) \\
&= o(n^3). \qquad (3.1.1)
\end{aligned}$$

Now we bound the expression in **DISC**. We have

$$\begin{aligned}
\frac{1}{n}\,|e(X,Y) - p\,|X|\,|Y||^2 &= \frac{1}{n}\Big(\sum_{x\in X} (\deg(x,Y) - p\,|Y|)\Big)^2 \\
&\le \sum_{x\in X} (\deg(x,Y) - p\,|Y|)^2 .
\end{aligned}$$

The above Cauchy–Schwarz step turned all the summands nonnegative, which affords us the next step, expanding the domain of summation from X to all of $V = V(G)$. Continuing,

$$\begin{aligned}
&\le \sum_{x\in V} (\deg(x,Y) - p\,|Y|)^2 \\
&= \sum_{x\in V} \deg(x,Y)^2 - 2p\,|Y| \sum_{x\in V} \deg(x,Y) + p^2 n\,|Y|^2 \\
&= \sum_{y,y'\in Y} \operatorname{codeg}(y,y') - 2p\,|Y| \sum_{y\in Y} \deg y + p^2 n\,|Y|^2 \\
&= |Y|^2\, p^2 n - 2p\,|Y|\cdot|Y|\, pn + p^2 n\,|Y|^2 + o(n^3) \qquad \text{[by \textbf{CODEG} and (3.1.1)]} \\
&= o(n^3). \qquad \square
\end{aligned}$$

Finally, let us consider the **graph spectrum**, i.e., the multiset of eigenvalues of the graph adjacency matrix, accounting for eigenvalue multiplicities. Eigenvalues are core to the study of pseudorandomness and they will play a central role in the rest of this chapter.

In this book, when we talk about the **eigenvalues of a graph**, we always mean the eigenvalues of the adjacency matrix of the graph. In other contexts, it may be useful to consider other related matrices, such as the Laplacian matrix, or a normalized adjacency matrix.

We will generally only consider real symmetric matrices, whose eigenvalues are always all real (Hermitian matrices also have this property). Our usual convention is to list all the eigenvalues in order (including multiplicities): $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. We refer to λ_1 as the **top eigenvalue** (or *largest eigenvalue*), and λ_i as the ***i*-th eigenvalue** (or the *i-th largest eigenvalue*). The second eigenvalue plays an important role. We write $\lambda_i(A)$ for the i-th eigenvalue of the matrix A and $\lambda_i(G) = \lambda_i(A_G)$ where A_G is the adjacency matrix of G.

Remark 3.1.18 (Linear algebra review)**.** For every $n \times n$ real symmetric matrix A with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, we can choose an eigenvector $v_i \in \mathbb{R}^n$ for each eigenvalue λ_i (so that $Av_i = \lambda_i v_i$) and such that $\{v_1, \ldots, v_n\}$ is an orthogonal basis of $\mathbb{R}^n$ (this is false for general non-symmetric matrices).

The **Courant–Fischer min-max theorem** is an important characterization of eigenvalues in terms of a variational problem. Here we only state some consequences most useful for us. We have

$$\lambda_1 = \max_{v \in \mathbb{R}^n \setminus \{0\}} \frac{\langle v, Av \rangle}{\langle v, v \rangle}.$$

Once we have fixed a choice of an eigenvector v_1 for the top eigenvalue λ_1, we have

$$\lambda_2 = \max_{\substack{v \perp v_1 \\ v \in \mathbb{R}^n \setminus \{0\}}} \frac{\langle v, Av \rangle}{\langle v, v \rangle}.$$

In particular, if G is a d-regular graph, then the all-1 vector, denoted $\mathbf{1} \in \mathbb{R}^{v(G)}$, is an eigenvector for the top eigenvalue d.

The **Perron–Frobenius theorem** tells us some important information about the top eigenvector and eigenvalue of a nonnegative matrix. For every connected graph G, the top eigenvector is simple (i.e., multiplicity one), so that $\lambda_i < \lambda_1$ for all $i > 1$. We also have $|\lambda_i| \leq \lambda_1$ for all i (one has $\lambda_n = -\lambda_1$ if and only if G is bipartite; see Remark 3.1.22 below). Also, the top eigenvector v_1 (which is unique up to scalar multiplication) has all coordinates positive.

If G has multiple connected components $G_1, \ldots, G_k$, then the eigenvalues of G (with multiplicities) are obtained by taking a multiset union of the eigenvalues of its connected components. An orthogonal system of eigenvectors can also be derived as such, by

extending each eigenvector of G_i to an eigenvector of G via padding the eigenvector by zeros outside the vertices of G_i.

Here is a useful formula:

$$\operatorname{tr} A^k = \lambda_1^k + \cdots + \lambda_n^k.$$

When A is the adjacency matrix of a graph G, $\operatorname{tr} A^k$ counts the number of closed walks of length k. In particular, $\operatorname{tr} A^2 = 2e(G)$.

Proof that **EIG** *implies* $\mathbf{C_4}$. Let A denote the adjacency matrix of G. The number of labeled 4-cycles is within $O(n^3)$ of the number of closed walks of length 4, and the latter equals

$$\operatorname{tr} A^4 = \lambda_1^4 + \cdots + \lambda_n^4 = p^4 n^4 + o(n^4) + \sum_{i=2}^{n} \lambda_i^4.$$

Since $\operatorname{tr} A^2 = 2e(G) \le n^2$, we have

$$\sum_{i=2}^{n} \lambda_i^4 \le \max_{i \ne 1} \lambda_i^2 \cdot \sum_{i=1}^{n} \lambda_i^2. = o(n^2) \cdot \operatorname{tr} A^2 = o(n^4).$$

So $\operatorname{tr} A^4 \le p^4 n^4 + o(n^4)$. □

Remark 3.1.19. A rookie error would be to bound $\sum_{i \ge 2} \lambda_i^4$ by $n \max_{i \ge 2} \lambda_i^4 = o(n^5)$, but this would not be enough. (Where do we save in the above proof?) We will see a similar situation later in Chapter 6 when we discuss the Fourier analytic proof of Roth's theorem.

Lemma 3.1.20 (Top eigenvalue and average degree)
The top eigenvalue of the adjacency matrix of a graph is always at least its average degree.

Proof. Let $\mathbf{1} \in \mathbb{R}^n$ be the all-1 vector. By the Courant–Fischer min-max theorem, the adjacency matrix A of the graph G has top eigenvalue

$$\lambda_1 = \sup_{\substack{x \in \mathbb{R}^n \\ x \ne 0}} \frac{\langle x, Ax \rangle}{\langle x, x \rangle} \ge \frac{\langle \mathbf{1}, A\mathbf{1} \rangle}{\langle \mathbf{1}, \mathbf{1} \rangle} = \frac{2e(G)}{v(G)} = \operatorname{avgdeg}(G).$$ □

Proof that $\mathbf{C_4}$ *implies* **EIG**. Again writing A for the adjacency matrix,

$$\sum_{i=1}^{n} \lambda_i^4 = \operatorname{tr} A^4 = \#\{\text{closed walks of length } 4\} \le p^4 n^4 + o(n^4).$$

On the other hand, by Lemma 3.1.20 above, we have $\lambda_1 \ge pn + o(n)$. So we must have $\lambda_1 = pn + o(n)$ and $\max_{i \ge 2} |\lambda_i| = o(n)$. □

This completes all the implications in the proof of Theorem 3.1.1.

Additional remarks

Remark 3.1.21 (Forcing graphs)**.** The $\mathbf{C_4}$ hypothesis says that having 4-cycle density asymptotically the same as random implies quasirandomness. Which other graphs besides C_4 have this property?

Chung, Graham, and Wilson (1989) called a graph F **forcing** if every graph with edge density $p + o(1)$ and F-density $p^{e(F)} + o(1)$ (i.e., asymptotically the same as random) is automatically quasirandom. Theorem 3.1.1 implies that C_4 is forcing. It remains an open problem to determine which graphs are forcing. The **forcing conjecture** says that F is forcing if and only if G is bipartite and not a tree (Skokan and Thoma 2004; Conlon, Fox, and Sudakov 2010). We will revisit this conjecture in Chapter 5 where we will reformulate it using the language of graphons.

More generally, one says that a family of graphs $\mathcal{F}$ is forcing if having F-density being $p^{e(F)} + o(1)$ for each $F \in \mathcal{F}$ implies quasirandomness. So $\{K_2, C_4\}$ is forcing. It seems to be a difficult problem to classify forcing families.

Even though many other graphs can potentially play the role of the 4-cycle, the 4-cycle nevertheless occupies an important role in the study of quasirandomness. The 4-cycle comes up naturally in the proofs, as we will see below. It also is closely tied to other important pseudorandomness measurements such as the Gowers U^2 uniformity norm in additive combinatorics.

Let us formulate a **bipartite analogue** of Theorem 3.1.1 since we will need it later. It is easy to adapt the above proofs to the bipartite version—we encourage the readers to think about the differences between the two settings.

Remark 3.1.22 (Eigenvalues of bipartite graphs)**.** Given a bipartite graph G with vertex bipartition $V \cup W$, we can write its adjacency matrix as

$$A = \begin{pmatrix} \mathbf{0} & B \\ B^\intercal & \mathbf{0} \end{pmatrix} \tag{3.1.2}$$

where B is an $|V| \times |W|$ matrix with rows indexed by V and columns indexed by W. The eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of A always satisfy

$$\lambda_i = \lambda_{n+1-i} \qquad \text{for every } 1 \leq i \leq n.$$

In other words, the eigenvalues are symmetric around zero. One way to see this is that if $x = (v, w)$ is an eigenvector of A, where $v \in \mathbb{R}^V$ is the restriction of x to the first $|V|$ coordinates, and w is the the restriction of x to the last $|W|$ coordinates, then

$$\begin{pmatrix} \lambda v \\ \lambda w \end{pmatrix} = \lambda x = Ax = \begin{pmatrix} \mathbf{0} & B \\ B^\intercal & \mathbf{0} \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} Bw \\ B^\intercal v \end{pmatrix},$$

so that

$$Bw = \lambda v \qquad \text{and} \qquad B^\intercal v = \lambda w.$$

Then the vector $x' = (v, -w)$ satisfies

$$Ax' = \begin{pmatrix} \mathbf{0} & B \\ B^\intercal & \mathbf{0} \end{pmatrix} \begin{pmatrix} v \\ -w \end{pmatrix} = \begin{pmatrix} -Bw \\ B^\intercal v \end{pmatrix} = \begin{pmatrix} -\lambda v \\ \lambda w \end{pmatrix} = -\lambda x'.$$

So we can pair each eigenvalue of A with its negation.

Exercise 3.1.23. Using the notation from (3.1.2), show that the positive eigenvalues of the adjacency matrix A coincide with the positive singular values of B (the singular values of B are also the positive square roots of the eigenvalues of $B^\intercal B$).

Theorem 3.1.24 (Bipartite quasirandom graphs)

Fix $p \in [0,1]$. Let $(G_n)_{n\geq 1}$ be a sequence of bipartite graphs G_n. Write G_n as G, with vertex bipartition $V \cup W$. Suppose $|V|, |W| \to \infty$ and $|E| = (p + o(1))\,|V|\,|W|$ as $n \to \infty$. The following properties are all equivalent:

DISC $e(X,Y) = p\,|X|\,|Y| + o(n^2)$ for all $X \subset V$ and $Y \subset W$.

COUNT For every bipartite graph H with vertex bipartition (S,T), the number of labeled copies of H in G with S embedded in V and T embedded in W is $(p^{e(H)} + o(1))\,|V|^{|S|}\,|W|^{|T|}$.

C_4 The number of closed walks of length 4 in G starting in V is at most $(p^4 + o(1))\,|V|^2\,|W|^2$.

Left-CODEG $\sum_{x,y\in V} \left|\mathrm{codeg}(x,y) - p^2\,|W|\right| = o(|V|^2\,|W|)$.

Right-CODEG $\sum_{x,y\in W} \left|\mathrm{codeg}(x,y) - p^2\,|V|\right| = o(|V|\,|W|^2)$.

EIG The adjacency matrix of G has top eigenvalue $(p+o(1))\sqrt{|X|\,|Y|}$ and second largest eigenvalue $o(\sqrt{|X|\,|Y|})$.

The bipartite discrepancy condition **DISC** is equivalent to being an $o(1)$-regular pair (Definition 2.1.2, Exercise 2.1.20).

Remark 3.1.25 (Bipartite double cover). Theorem 3.1.24 implies the non-bipartite version Theorem 3.1.1, since every graph G can be transformed into a bipartite graph $G \times K_2$ (a graph tensor power) whose two vertex parts are both copies of $V(G)$. Each edge $u \sim v$ of G lifts to two edges $(u,0) \sim (v,1)$ and $(u,1) \sim (u,0)$ in $G \times K_2$. An example is shown below.

G $\qquad$ $G \times K_2$

It is not hard to check G satisfies each property in Theorem 3.1.1 if and only if $G \times K_2$ satisfies the corresponding bipartite property in Theorem 3.1.24 (exercise).

Like earlier, random bipartite graphs are bipartite quasirandom. The proof (omitted) is essentially the same as Proposition 3.1.8 and Corollary 3.1.9.

Proposition 3.1.26 (Random bipartite graphs are typically quasirandom)
Fix $p \in [0,1]$. With probability 1, a sequence of bipartite random graphs $G_n \sim \mathbf{G}(n,n,p)$ (obtained by keeping every edge of $K_{n,n}$ with probability p independently) is quasirandom in the sense of Theorem 3.1.24.

Remark 3.1.27 (Sparse graphs)**.** We stated quasirandom properties so far only for graphs of constant order density (i.e., p is a constant). Let us think about what happens if we allow $p = p_n$ to depend on n and decaying to zero as $n \to \infty$. Such graphs are sometimes called *sparse* (although some other authors reserve the word "sparse" for bounded degree graphs). Theorems 3.1.1 and 3.1.24 as stated do hold for a constant $p = 0$, but the results are not as informative as we would like. For example, the error tolerance on the **DISC** is $o(n^2)$, which does not tell us much since the graph already has much fewer edges due to its sparseness anyway.

To remedy the situation, the natural thing to do is to adjust the error tolerance relative to the edge density $p = p_n \to 0$. Here are some representative examples (all of these properties should also depend on p):

SparseDISC $|e(X,Y) - p|X||Y|| = o(pn^2)$ for all $X, Y \subset V(G)$.
SparseCOUNT$_H$ The number of labeled copies of H is $(1 + o(1))p^{e(H)}n^{v(H)}$.
SparseC$_4$ The number of labeled 4-cycles is at most $(1 + o(1))p^4n^4$.
SparseEIG $\lambda_1 = (1 + o(1))pn$ and $\max_{i \neq 1} |\lambda_i| = o(pn)$.

Warning: these sparse pseudorandomness conditions are *not* all equivalent to each other. Some of the implications still hold (the reader is encouraged to think about which ones). However, some crucial implications such as the counting lemma fail quite miserably. For example:

SparseDISC does not imply **SparseCOUNT**.

Indeed, suppose $p = n^{-c}$ for some constant $1/2 < c < 1$. In a typical random graph $\mathbf{G}(n,p)$, the number of triangles is close to $\binom{n}{3}p^3$, while the number of edges is close to $\binom{n}{2}p$. We have $p^3n^3 = o(pn^2)$ as long as $p = o(n^{-1/2})$, so there are significantly fewer triangles than there are edges. Now remove an edge from every triangle in this random graph. We will have removed $o(pn^2)$ edges, a negligible fraction of the $(p + o(1))\binom{n}{2}$ edges, and this edge removal should not significantly affect **SparseDISC**. However, we have changed the triangle count significantly as a result.

Fortunately, this is not the end of the story. With additional hypotheses on the sparse graph, we can sometimes salvage a counting lemma. *Sparse counting lemmas* play an

important role in the proof of the Green–Tao theorem on arithmetic progressions in the primes, as we will explain in Chapter 9.

The next three exercises ask you to prove additional equivalent quasirandomness properties. It is easy to verify that the quasirandom graphs indeed satisfy each of the properties below.

Exercise 3.1.28* (Quasirandomness through fixed sized subsets)**.** Fix $p \in [0,1]$. Let (G_n) be a sequence of graphs with $v(G_n) = n$ (here $n \to \infty$ along a subsequence of integers).

1. Fix a single $\alpha \in (0,1)$. Suppose
$$e(S) = \frac{p\alpha^2 n^2}{2} + o(n^2) \qquad \text{for all } S \subset V(G) \text{ with } |S| = \lfloor \alpha n \rfloor .$$
Prove that G is quasirandom.
2. Fix a single $\alpha \in (0, 1/2)$. Suppose
$$e(S, V(G) \setminus S) = p\alpha(1-\alpha)n^2 + o(n^2) \qquad \text{for all } S \subset V(G) \text{ with } |S| = \lfloor \alpha n \rfloor .$$
Prove that G is quasirandom. Furthermore, show that the conclusion is false for $\alpha = 1/2$.

Exercise 3.1.29 (Quasirandomness and regularity partitions)**.** Fix $p \in [0,1]$. Let (G_n) be a sequence of graphs with $v(G_n) \to \infty$. Suppose that for every $\epsilon > 0$, there exists $M = M(\epsilon)$ so that each G_n has an ϵ-regular partition where all but ϵ-fraction of vertex pairs lie between pairs of parts with edge density $p + o(1)$ (as $n \to \infty$). Prove that G_n is quasirandom.

Exercise 3.1.30* (Triangle counts on induced subgraphs)**.** Fix $p \in (0,1]$. Let (G_n) be a sequence of graphs with $v(G_n) = n$. Let $G = G_n$. Suppose that for every $S \subset V(G)$, the number of triangles in the induced subgraph $G[S]$ is $p^3\binom{|S|}{3} + o(n^3)$. Prove that G is quasirandom.

Exercise 3.1.31* (Perfect matchings)**.** Prove that there are constant $\beta, \epsilon > 0$ such that for every positive even integer n and real $p \geq n^{-\beta}$, if G is an n-vertex graph where every vertex has degree $(1 \pm \epsilon)pn$ (meaning within ϵpn of pn) and every pair of vertices has codegree $(1 \pm \epsilon)p^2 n$, then G has a perfect matching.

3.2 Expander Mixing Lemma

We dive further into the relationship between graph eigenvalues and its pseudorandomness properties. We focus on d-regular graphs since they occur often in practice

(e.g., from Cayley graphs), and they are also cleaner to work with. Unlike the previous section, the results here are effective for any value of d (not just when d is on the same order as n).

As we saw earlier, the magnitudes of eigenvalues are related to the pseudorandomness of a graph. In a d-regular graph, the top eigenvalue is always exactly d. The following condition says that all other eigenvalues are bounded by λ in absolute value.

Definition 3.2.1 ((n, d, λ)-graph)
An **(n, d, λ)-graph** is an n-vertex, d-regular graph whose adjacency matrix eigenvalues $d = \lambda_1 \geq \cdots \geq \lambda_n$ satisfy

$$\max_{i \neq 1} |\lambda_i| \leq \lambda.$$

Remark 3.2.2 (Notation). Rather than saying, e.g., "an $(n, 7, 6)$-graph," we prefer to say "an (n, d, λ)-graph with $d = 7$ and $\lambda = 6$" for clarity as the name "(n, d, λ)" is quite standard and recognizable.

Remark 3.2.3 (Linear algebra review). The **operator norm** of a matrix $A \in \mathbb{R}^{m \times n}$ is defined by

$$\|A\| = \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|Ax|}{|x|} = \sup_{\substack{x \in \mathbb{R}^n \setminus \{0\} \\ y \in \mathbb{R}^m \setminus \{0\}}} \frac{\langle y, Ax \rangle}{|x|\,|y|}.$$

Here $|x| = \sqrt{\langle x, x \rangle}$ denotes the length of vector x. The operator norm of A is the maximum ratio that A can amplify the length of a vector by. If A is a real symmetric matrix, then

$$\|A\| = \max_i |\lambda_i(A)|.$$

For general matrices, the operator norm of A equals the largest singular value of A.

Here is the main result of this section.

Theorem 3.2.4 (Expander mixing lemma)
If G is an (n, d, λ)-graph, then

$$\left| e(X, Y) - \frac{d}{n} |X|\,|Y| \right| \leq \lambda \sqrt{|X|\,|Y|} \qquad \text{for all } X, Y \subset V(G).$$

On the left-hand side, $(d/n)\,|X|\,|Y|$ is the number of edges that one should expect between X and Y purely based on the edge density d/n of the graph and the sizes of X and Y. Note that unlike the discrepancy condition (**DISC**) from quasirandom graphs (Theorem 3.1.1), the error bound on the right-side hand depends on the sizes of X and Y. We can apply the expander mixing lemma to small subsets X and Y and still obtain useful estimates on $e(X, Y)$, unlike the dense quasirandom graph conditions.

Proof. Let J be the $n \times n$ all-1 matrix. Since the all-1 vector $\mathbf{1} \in \mathbb{R}^n$ is an eigenvector of A_G with eigenvalue d, we see that $\mathbf{1}$ is an eigenvector of $A_G - \frac{d}{n}J$ with eigenvalue 0. Any other eigenvector v of A_G, with $v \perp \mathbf{1}$, satisfies $Jv = 0$, and thus v is also an eigenvector of $A_G - \frac{d}{n}J$ with the same eigenvalue as in A_G. Therefore, the eigenvalues of $A_G - \frac{d}{n}J$ are obtained by taking the eigenvalues of A_G then replacing one top eigenvalue d by zero. All the other eigenvalues of $A_G - \frac{d}{n}J$ are therefore at most λ in absolute value, so $\left\|A_G - \frac{d}{n}J\right\| \le \lambda$. Therefore,

$$\begin{aligned}\left|e(X,Y) - \frac{d}{n}|X|\,|Y|\right| &= \left|\left\langle \mathbf{1}_X, \left(A_G - \frac{d}{n}J\right)\mathbf{1}_Y\right\rangle\right| \\ &\le \left\|A_G - \frac{d}{n}J\right\| |\mathbf{1}_X|\,|\mathbf{1}_Y| \\ &\le \lambda\sqrt{|X|\,|Y|}.\end{aligned}$$

□

Exercise 3.2.5. Prove the following strengthening the expander mixing lemma.

Theorem 3.2.6 (Expander mixing lemma – slightly strengthened)
If G is an (n, d, λ)-graph, then

$$\left|e(X,Y) - \frac{d}{n}|X|\,|Y|\right| \le \frac{\lambda}{n}\sqrt{|X|\,(n-|X|)\,|Y|\,(n-|Y|)} \qquad \text{for all } X, Y \subset V(G).$$

We also have a bipartite analogue (the nomenclature used here is less standard). Recall from Remark 3.1.22 that the eigenvalues of a bipartite graph are symmetric around zero.

Definition 3.2.7 (Bipartite-(n, d, λ)-graph)
An **bipartite-(n, d, λ)-graph** is a d-regular bipartite graph with n vertices in each part, such that its second largest eigenvalue is at most λ.

Exercise 3.2.8. Show that G is an (n, d, λ)-graph if and only if $G \times K_2$ is a bipartite-(n, d, λ)-graph.

Theorem 3.2.9 (Bipartite expander mixing lemma)
Let G be a bipartite-(n, d, λ)-graph with vertex bipartition $V \cup W$. Then

$$\left|e(X,Y) - \frac{d}{n}|X|\,|Y|\right| \le \lambda\sqrt{|X|\,|Y|} \qquad \text{for all } X \subset V \text{ and } Y \subset W.$$

Exercise 3.2.10. Prove Theorem 3.2.9.

Remark 3.2.11. The following partial converse to the expander mixing lemma was shown by Bilu and Linial (2006). The extra log factor turns out to be necessary.

Theorem 3.2.12 (Converse to expander mixing lemma)
There exists an absolute constant C such that if G is a d-regular graph, and β satisfies

$$\left| e(X,Y) - \frac{d}{n}|X|\,|Y| \right| \le \beta\sqrt{|X|\,|Y|} \qquad \text{for all } X, Y \subset V(G),$$

then G is an (n, d, λ)-graph with $\lambda \le C\beta\log(2d/\beta)$.

Cheeger's inequality: edge expansion vs. spectral gap

The **spectral gap** is defined to be the difference between the two most significant eigenvalues, i.e., $\lambda_1 - \lambda_2$ for the adjacency matrix of a graph. This quantity turns out to be closely related to expansion in graphs. We define the **edge-expansion ratio** of a graph $G = (V, E)$ to be the quantity

$$h(G) := \min_{\substack{S \subset V \\ 0 < |S| \le |V|/2}} \frac{e_G(S, V \setminus S)}{|S|}.$$

In other words, a graph with edge-expansion ratio at least h has the property that for every nonempty subset of vertices S with $|S| \le |V|/2$, there are at least $h\,|S|$ edges leaving S.

Cheeger's inequality, stated below, tells us that among d-regular graphs for a fixed d, having spectral gap bounded away from zero is equivalent to having edge-expansion ratio bounded away from zero. Cheeger (1970) originally developed this inequality for Riemannian manifolds. The graph theoretic analogue was proved by Dodziuk (1984), and independently by Alon and Milman (1985) and Alon (1986).

Theorem 3.2.13 (Cheeger's inequality)
Let G be an n-vertex d-regular graph with adjacency matrix spectral gap $\kappa = d - \lambda_2$. Then its edge-expansion ratio $h = h(G)$ satisfies

$$\kappa/2 \le h \le \sqrt{2d\kappa}.$$

The two bounds of Cheeger's inequality are tight up to constant factors. For the lower bound, taking G to be the skeleton of the d-dimensional cube with vertex set $\{0,1\}^d$ gives $h = 1$ (achieved by the $d-1$ dimensional subcube) and $\kappa = 2$. For the upper bound, taking G to be an n-cycle gives $h = 2/(n/2) = \Theta(1/n)$ while $d = 2$ and $\kappa = 2 - 2\cos(2\pi/n)) = \Theta(1/n^2)$.

We call a family of d-regular graphs **expanders** if there is some constant $\kappa_0 > 0$ so that each graph in the family has spectral gap $\geq \kappa_0$; by Cheeger's inequality, this is equivalent to the existence of some $h_0 > 0$ so that each graph in the family has edge expansion ratio $\geq h_0$. Expander graphs are important objects in mathematics and computer science. For example, expander graphs have rapid mixing properties, which are useful for designing efficient Monte Carlo algorithms for sampling and estimation.

The following direction of Cheeger's inequality is easier to prove. It is similar to the expander mixing lemma.

Exercise 3.2.14 (Spectral gap implies expansion)**.** Prove the $\kappa/2 \leq h$ part of Cheeger's inequality.

The other direction, $h \leq \sqrt{2d\kappa}$, is more difficult and interesting. The proof is outlined in the following exercise.

Exercise 3.2.15 (Expansion implies spectral gap)**.** Let $G = (V, E)$ be a connected d-regular graph with spectral gap κ. Let $x = (x_v)_{v \in V} \in \mathbb{R}^V$ be an eigenvector associated to the second largest eigenvalue $\lambda_2 = d - \kappa$ of the adjacency matrix of G. Assume that $x_v > 0$ on at most half of the vertex set (or else we replace x by $-x$). Let $y = (y_v)_{v \in V} \in \mathbb{R}^V$ be obtained from x by replacing all its negative coordinates by zero.

(a) Prove that

$$d - \frac{\langle y, Ay \rangle}{\langle y, y \rangle} \leq \kappa.$$

Hint: Recall that $\lambda_2 x_v = \sum_{u \sim v} x_u$.

(b) Let

$$\Theta = \sum_{uv \in E} \left| y_u^2 - y_v^2 \right|.$$

Prove that

$$\Theta^2 \leq 2d(d \langle y, y \rangle - \langle y, Ay \rangle) \langle y, y \rangle .$$

Hint: $y_u^2 - y_v^2 = (y_u - y_v)(y_u + y_v)$. Apply Cauchy–Schwarz.

(c) Relabel the vertex set V by $[n]$ so that $y_1 \geq y_2 \cdots \geq y_t > 0 = y_{t+1} = \cdots = y_n$. Prove

$$\Theta = \sum_{k=1}^{t} (y_k^2 - y_{k+1}^2)\, e([k], [n] \setminus [k]).$$

(d) Prove that for some $1 \leq k \leq t$,

$$\frac{e([k], [n] \setminus [k])}{k} \leq \frac{\Theta}{\langle y, y \rangle}.$$

(e) Prove the $h \leq \sqrt{2d\kappa}$ claim of Cheeger's inequality.

Exercises

Exercise 3.2.16 (Independence numbers)**.** Prove that every independent set in a (n, d, λ)-graph has size at most $n\lambda/(d + \lambda)$.

Exercise 3.2.17 (Diameter)**.** Prove that the diameter of an (n, d, λ)–graph is at most $\lceil \log n / \log(d/\lambda) \rceil$. (The *diameter* of a graph is the maximum distance between a pair of vertices.)

Exercise 3.2.18 (Counting cliques)**.** For each part below, prove that for every $\epsilon > 0$, there exists $\delta > 0$ such that the conclusion holds for every (n, d, λ)-graph G with $d = pn$.

(a) If $\lambda \le \delta p^2 n$, then the number of triangles of G is within a $1 \pm \epsilon$ factor of $p^3 \binom{n}{3}$.

(b*) If $\lambda \le \delta p^3 n$, then the number of K_4's in G is within a $1 \pm \epsilon$ factor of $p^6 \binom{n}{4}$.

3.3 Abelian Cayley Graphs and Eigenvalues

Many important constructions of pseudorandom graphs come from groups.

Definition 3.3.1 (Cayley graph)

Let Γ be a finite group, and let $S \subset \Gamma$ be a subset with $S = S^{-1}$ (i.e., $s^{-1} \in S$ for all $s \in S$) and not containing the identity element. We write $\mathrm{Cay}(\Gamma, S)$ to denote the **Cayley graph** on Γ generated by S, which has elements of Γ as vertices, and

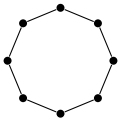
$$g \sim gs \quad \text{for all } g \in \Gamma \text{ and } s \in S.$$

as edges.

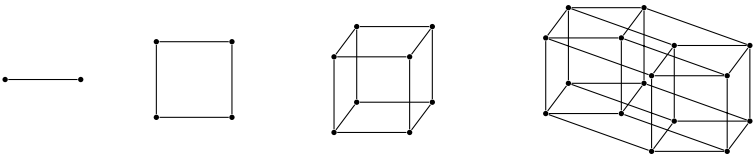
In this section, we only consider abelian groups, specifically $\mathbb{Z}/p\mathbb{Z}$ for concreteness (though everything here generalizes easily to all finite abelian groups). For abelian groups, we write the group operation additively, i.e., $g + s$. So edges join elements whose difference lies in S.

Remark 3.3.2. In later sections when we consider a non-abelian group Γ, one needs to make a choice whether to define edges by left- or right-multiplication (i.e., gs or sg; we chose gs here). It does not matter which choice one makes (as long as one is consistent) since the resulting Cayley graphs are isomorphic (why?). However, some careful bookkeeping is sometimes required to make sure that later computations are consistent with the initial choice.

Example 3.3.3. $\mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, \{-1, 1\})$ is a cycle of length n. The graph for $n = 8$ is shown below.



Example 3.3.4. $\mathrm{Cay}(\mathbb{F}_2^n, \{e_1, \ldots, e_n\})$ is the skeleton of an n-dimensional cube. Here e_i is the i-th standard basis vector. The graphs for $n = 1, 2, 3, 4$ are illustrated below..

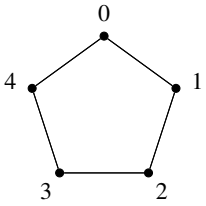


Here is an explicitly constructed family of quasirandom graphs with edge density $1/2 + o(1)$.

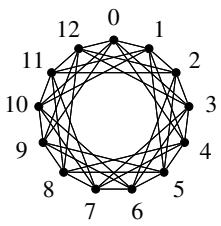
Definition 3.3.5 (Paley graph)

Let $p \equiv 1 \pmod 4$ be a prime. The **Paley graph** of order p is $\mathrm{Cay}(\mathbb{Z}/p\mathbb{Z}, S)$, where S is the set of non-zero quadratic residues in $\mathbb{Z}/p\mathbb{Z}$ (here $\mathbb{Z}/p\mathbb{Z}$ is viewed as an additive group).

Example 3.3.6. The Paley graphs for $p = 5$ and $p = 13$ are shown below.



$\mathrm{Cay}(\mathbb{Z}/5\mathbb{Z}, \{\pm 1\})$



$\mathrm{Cay}(\mathbb{Z}/13\mathbb{Z}, \{\pm 1\})$

Remark 3.3.7 (Quadratic residues)**.** Here we recall some facts from elementary number theory. For every odd prime p, the set $S = \left\{a^2 : a \in \mathbb{F}_p^\times\right\}$ of quadratic residues is a multiplicative subgroup of $\mathbb{F}_p^\times$ with index two. In particular, $|S| = (p-1)/2$. We have $-1 \in S$ if and only if $p \equiv 1 \pmod 4$ (which is required to define a Cayley graph, as the generating set needs to be a symmetric set, i.e., $S = -S$).

We will show that Paley graphs are quasirandom by verifying the **EIG** condition, i.e., all eigenvalues, except the top one, are small. Here is a general formula for computing the eigenvalues of any Cayley graph on $\mathbb{Z}/p\mathbb{Z}$.

Theorem 3.3.8 (Eigenvalues of abelian Cayley graphs on $\mathbb{Z}/n\mathbb{Z}$)
Let n be a positive integer. Let $S \subset \mathbb{Z}/n\mathbb{Z}$ with $0 \neq S$ and $S = -S$. Let

$$\omega = \exp(2\pi i/n).$$

Then we have an orthonormal basis $v_0, \ldots, v_{n-1} \in \mathbb{C}^n$ of eigenvectors of $\mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, S)$ where

$$v_j \in \mathbb{C}^n \text{ has } x\text{-coordinate } \omega^{jx}/\sqrt{n}, \text{ for each } x \in \mathbb{Z}/n\mathbb{Z}.$$

The eigenvalue associated to the eigenvector v_j equals to

$$\lambda_j = \sum_{s \in S} \omega^{js}.$$

In particular, $\lambda_0 = |S|$ and v_0 has all coordinates $1/\sqrt{n}$.

Remark 3.3.9 (Eigenvalues and the Fourier transform)**.** The coordinates of the eigenvectors are shown below.

	$\mathbb{Z}/n\mathbb{Z}$				
	0	1	2	$\cdots$	$n-1$
$\sqrt{n}\, v_0$	1	1	1	$\cdots$	1
$\sqrt{n}\, v_1$	1	ω	ω^2	$\cdots$	ω^{n-1}
$\sqrt{n}\, v_2$	1	ω^2	ω^4	$\cdots$	$\omega^{2(n-1)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\sqrt{n}\, v_{n-1}$	1	ω^{n-1}	$\omega^{2(n-1)}$	$\cdots$	$\omega^{(n-1)^2}$

Viewed as a matrix, this is sometimes known as the **discrete Fourier transform matrix**. We will study the Fourier transform in Chapter 6. These two topics are closely tied. The eigenvalues of an abelian Cayley graph $\mathrm{Cay}(\Gamma, S)$ are precisely the Fourier transform in Γ of the generating set S, up to normalizing factors:

$$\text{eigenvalues of } \mathrm{Cay}(\Gamma, S) \quad \longleftrightarrow \quad \text{Fourier transform } \widehat{1_S} \text{ in } \Gamma.$$

We will say more about this in Remark 3.3.11 below.

Proof. Let A be the adjacency matrix of $\mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, S)$. First we check that each v_j is an eigenvector of A with eigenvalue λ_j. The coordinate of $\sqrt{n}Av_j$ at $x \in \mathbb{Z}/n\mathbb{Z}$ equals to

$$\sum_{s \in S} \omega^{j(x+s)} = \left(\sum_{s \in S} \omega^{js}\right)\omega^{jx} = \lambda_j \omega^{jx}.$$

So $Av_j = \lambda_j v_j$.

Next we check that $\{v_0, \dots, v_{n-1}\}$ is an orthonormal basis. We have the inner product

$$\begin{aligned}\langle v_j, v_k \rangle &= \frac{1}{n}\left(1 \cdot 1 + \overline{\omega^j}\omega^k + \overline{\omega^{2j}}\omega^{2k} + \cdots + \overline{\omega^{(n-1)j}}\omega^{(n-1)k}\right) \\ &= \frac{1}{n}\left(1 + \omega^{k-j} + \omega^{2(k-j)} + \cdots + \omega^{(n-1)(k-j)}\right) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}.\end{aligned}$$

For the $i \neq j$ case, we use that for any m-th root of unity $\zeta \neq 1$, $\sum_{j=0}^{m-1} \zeta^j = 0$. So $\{v_0, \dots, v_{n-1}\}$ is an orthonormal basis. □

Remark 3.3.10 (Real vs complex eigenbases)**.** The adjacency matrix of a graph is a real symmetric matrix, so all its eigenvalues are real, and it always has a real orthogonal eigenbasis. The eigenbasis given in Theorem 3.3.8 is complex, but it can always be made real. Looking at the formulas in Theorem 3.3.8, we have $\lambda_j = \lambda_{n-j}$, and v_j is the complex conjugate of v_{n-j}. So we can form a real orthogonal eigenbasis by replacing, for each $j \notin \{0, n/2\}$, the pair (v_j, v_{n-j}) by $((v_j + v_{n-j})/\sqrt{2}, i(v_j - v_{n-j})/\sqrt{2})$. Equivalently, we can separate the real and imaginary parts of each v_j, which are both eigenvectors with eigenvalue λ_j. All the real eigenvalues and eigenvectors can be expressed in terms of sines and cosines.

Remark 3.3.11 (Every abelian Cayley graph has an eigenbasis independent of the generators)**.** The above theorem and its proof generalizes to all finite abelian groups, not just $\mathbb{Z}/n\mathbb{Z}$. For every finite abelian group Γ, we have a set $\widehat{\Gamma}$ of *characters*, i.e., homomorphisms $\chi : \Gamma \to \mathbb{C}^\times$. Then $\widehat{\Gamma}$ turns out to be a group isomorphic to Γ (one can check this by first writing Γ as a direct product of cyclic groups). For each $\chi \in \widehat{\Gamma}$, define the vector $v_\chi \in \mathbb{C}^\Gamma$ by setting the coordinate at $g \in \Gamma$ to be $\chi(g)/\sqrt{|\Gamma|}$. Then $\{v_\chi : \chi \in \widehat{\Gamma}\}$ is an orthonormal basis for the adjacency matrix of every Cayley graph on Γ. The eigenvalue corresponding to v_χ is $\lambda_\chi(S) = \sum_{s \in S} \chi(s)$. Up to normalization, $\lambda_\chi(S)$ is the Fourier transform of the indicator function of S on the abelian group Γ (Theorem 3.3.8 is a special case of this construction). In particular, this eigenbasis $\{v_\chi : \chi \in \widehat{\Gamma}\}$ depends only on the finite abelian group and not on the generating set S. In other words, we have a *simultaneous diagonalization* for all adjacency matrices of Cayley graphs on a fixed finite abelian group.

If Γ is a non-abelian group, then there does not exist a simultaneous eigenbasis for all Cayley graphs on Γ. There is a corresponding theory of non-abelian Fourier analysis, which uses group representation theory. We will discuss more about non-abelian Cayley graphs in Section 3.4.

Now we apply the above formula to compute eigenvalues of Paley graphs. In particular, the following tells us that Paley graphs satisfy the quasirandomness condition **EIG** from Theorem 3.1.1.

Theorem 3.3.12 (Eigenvalues of Paley graphs)
Let $p \equiv 1 \pmod 4$ be a prime. The adjacency of matrix of the Paley graph of order p has top eigenvalue $(p-1)/2$, and all other eigenvalues are either $(\sqrt{p}-1)/2$ or $(-\sqrt{p}-1)/2$.

Proof. Applying Theorem 3.3.8, we see that the eigenvalues are given by, for $j = 0, 1, \ldots, p-1$,

$$\lambda_j = \sum_{s \in S} \omega^{js} = \frac{1}{2}\left(-1 + \sum_{x \in \mathbb{F}_p} \omega^{jx^2}\right),$$

since each quadratic residue s appears as x^2 for exactly two non-zero x. Clearly $\lambda_0 = (p-1)/2$. For $j \neq 0$, the next result shows that the inner sum on the right-hand side is $\pm\sqrt{p}$ (note that the above sum is real when $p \equiv 1 \pmod 4$ since $S = S^{-1}$ and so the sum equals to its own complex conjugate; alternatively, the sum must be real since all eigenvalues of a symmetric matrix are real). □

Remark 3.3.13. Since the trace of the adjacency matrix is zero, and equals the sum of eigenvalues, we see that the non-top eigenvalues are equally split between $(\sqrt{p}-1)/2$ and $(-\sqrt{p}-1)/2$.

Theorem 3.3.14 (Gauss sum)
Let p be an odd prime, $\omega = \exp(2\pi i/p)$, and $j \in \mathbb{F}_p \setminus \{0\}$. Then

$$\left|\sum_{x \in \mathbb{F}_p} \omega^{jx^2}\right| = \sqrt{p}.$$

Proof. We have

$$\left|\sum_{x \in \mathbb{F}_p} \omega^{jx^2}\right|^2 = \sum_{x,y \in \mathbb{Z}/p\mathbb{Z}} \omega^{j((x+y)^2 - x^2)} = \sum_{x,y \in \mathbb{Z}/p\mathbb{Z}} \omega^{j(2xy+y^2)}.$$

For each fixed y, we have

$$\sum_{x \in \mathbb{Z}/p\mathbb{Z}} \omega^{j(2xy+y^2)} = \begin{cases} p & \text{if } y = 0, \\ 0 & \text{if } y \neq 0. \end{cases}$$

Summing over y yields the claim. □

Remark 3.3.15 (Sign of the Gauss sum)**.** The determination of this sign is a more difficult problem. Gauss conjectured the sign in 1801 and it took him four years to prove it. When j is a nonzero quadratic residue mod p, the inner sum above turns

out to equal $\sqrt{p}$ if $p \equiv 1 \pmod 4$ and $i\sqrt{p}$ if $p \equiv 3 \pmod 4$. When j is a quadratic non-residue, it is $-\sqrt{p}$ and $-i\sqrt{p}$ in the two cases respectively. For a proof, see, e.g., Ireland and Rosen (1990, Section 6.4).

Exercise 3.3.16. Let p be an odd prime and $A, B \subset \mathbb{Z}/p\mathbb{Z}$. Show that

$$\left|\sum_{a\in A}\sum_{b\in B}\left(\frac{a+b}{p}\right)\right| \leq \sqrt{p\,|A|\,|B|}$$

where (a/p) is the Legendre symbol defined by

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } a \equiv 0 \pmod p \\ 1 & \text{if } a \text{ is a nonzero quadratic residue mod } p \\ -1 & \text{if } a \text{ is a quadratic nonresidue mod } p \end{cases}$$

Exercise 3.3.17. Prove that in a Paley graph of order p, every clique has size at most $\sqrt{p}$.

Exercise 3.3.18 (No spectral gap if too few generators)**.** Prove that for every $\epsilon > 0$ there is some $c > 0$ such that for every $S \subset \mathbb{Z}/n\mathbb{Z}$ with $0 \notin S = -S$ and $|S| \leq c \log n$, the second largest eigenvalue of the adjacency matrix of $\mathrm{Cay}(\mathbb{Z}/n\mathbb{Z}, S)$ is at least $(1-\epsilon)\,|S|$.

Exercise 3.3.19*. Let p be a prime and let S be a multiplicative subgroup of $\mathbb{F}_p^\times$. Suppose $-1 \in S$. Prove that all eigenvalues of the adjacency matrix of $\mathrm{Cay}(\mathbb{Z}/p\mathbb{Z}, S)$, other than the top one, are at most $\sqrt{p}$ in absolute value.

3.4 Quasirandom Groups

In the previous section, we saw that certain Cayley graphs on cyclic groups are quasirandom. Note that not all Cayley graphs on cyclic groups are quasirandom, e.g., the Cayley graphs obtained by $\Gamma = \mathbb{Z}/n\mathbb{Z}$ and $S = \{x : |x| \leq n/4\} \subset \mathbb{Z}/n\mathbb{Z}$ are not quasirandom.

In this section, we will see that for certain families of non-abelian groups, *every* Cayley graph on the group is quasirandom, regardless of the Cayley graph generators. Gowers (2008) called such groups **quasirandom groups**, and showed that they are precisely groups with no small non-trivial representations. He came up with this notion while solving the following problem about product-free sets in groups.

Question 3.4.1 (Product-free subset of groups)
Given a group of order n, what is the size of its largest product-free subset? Is it always $\geq cn$ for some constant $c > 0$?

Remark 3.4.2 (Representations of finite groups)**.** We need some basic concepts from group representation theory in this section—mostly just some definitions. Feel free to skip this remark if you have already seen group representations before.

Given a finite group Γ, it is often useful to study its actions as linear transformations on some vector space. For example, if Γ is a cyclic or dihedral group, it is natural to think of elements of Γ as rotations and reflection of a plane, which are linear transformation on $\mathbb{R}^2$. The theory turns out to be much nicer over $\mathbb{C}$ than $\mathbb{R}$ since $\mathbb{C}$ is algebraically closed. We are interested in ways that Γ can be represented as a group of linear transformations acting on some $\mathbb{C}^d$.

A **representation** of a finite group Γ is a group homomorphism $\rho\colon \Gamma \to \mathrm{GL}(V)$, where V is a complex vector space (everything will take place over $\mathbb{C}$) and $\mathrm{GL}(V)$ is the group of invertible linear transformations of V. We sometimes omit ρ from the notation and just say that V is a representation of Γ, and also that Γ **acts** on V (via ρ). For each $g \in \Gamma$ and $v \in V$, we write $gv = \rho(g)v$ for the image of the g-action on v. We write $\dim \rho = \dim V$ for the **dimension** of the representation.

The fact that $\rho\colon \Gamma \to \mathrm{GL}(V)$ is a group homomorphism means that the action of Γ on V is compatible with group operations in Γ in the following sense: if $g, h \in \Gamma$, then the expression ghx does not depend on whether we first apply h to x and then g to hx, or if we first multiply g and h in Γ and then apply their product gh to x.

For example, suppose Γ is a subgroup of permutations of $[n]$, with each element $g \in \Gamma$ viewed as a permutation $g\colon [n] \to [n]$. We can define a representation of Γ on $\mathbb{C}^n$ by letting Γ permute the coordinates: for any $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$, set $gx = (x_{g(1)}, \ldots, x_{g(n)})$. As an element of $\mathrm{GL}(n, \mathbb{C})$, $\rho(g)$ is the $n \times n$ permutation matrix of the permutation g, and $gx = \rho(g)x$ for each $x \in \mathbb{C}^n$.

We say that the representation V of Γ is **trivial** if $gv = v$ for all $g \in \Gamma$ and $v \in V$, and **non-trivial** otherwise.

We say that a subspace W of V is **Γ-invariant** if $gw \in W$ for all $w \in W$. In other words, the image of W under Γ is contained in W (and actually must equal W due to the invertibility of group elements). Then W is a representation of Γ, and we call it a **subrepresentation** of V.

For an introduction to group representation theory, see any standard textbook, e.g., *Linear Representations of Finite Groups* by Serre (1977) is a classic. Also, the lectures notes titled *Representation Theory of Finite Groups, and Applications* by Wigderson (2012) is a friendly introduction with applications to combinatorics and theoretical computer science.

Recall from Definition 3.2.1 that an **(n, d, λ)-graph** is an n-vertex d-regular graph all of whose eigenvalues, except the top one, are at most λ in absolute value.

The main theorem of this section, below, says that a group with no small non-trivial representations always produces quasirandom Cayley graphs (Gowers 2008).

Theorem 3.4.3 (Cayley graphs on quasirandom groups)
Let Γ be a group of order n with no non-trivial representations of dimension less than K. Then every d-regular Cayley graph on Γ is an (n, d, λ)-graph for some $\lambda < \sqrt{dn/K}$.

Remark 3.4.4 (Abelian groups and one-dimensional representations). If Γ is abelian, then it has many one-dimensional non-trivial representations, namely coming its multiplicative characters. For example, of $\Gamma = \mathbb{Z}/n\mathbb{Z}$, then the map $\rho\colon \Gamma \to \mathbb{C}^\times$ sending $g \in \mathbb{Z}/n\mathbb{Z}$ to ω^g, where ω is some non-trivial root of unity, is a non-trivial one-dimensional representation. In fact, one can vary ω over all roots of unity to obtain all non-isomorphic one-dimensional representations of Γ.

So the hypothesis of having no low dimensional non-trivial representations can be viewed as a statement that the group is *highly non-abelian* in some sense.

A representation is **irreducible** if it contains no subrepresentations other than itself and the zero-dimensional subrepresentation. Irreducible representations are the basic building blocks of group representations, and so understanding all irreducible representations of a group is a fundamental objective. Among finite groups, a group is abelian if and only if all its irreducible representations are one-dimensional.

More generally we will prove the result for vertex-transitive groups, of which Cayley graphs is a special case.

Definition 3.4.5 (Vertex-transitive graphs)
Let G be a graph. An **automorphism** of G is a permutation of $V(G)$ that induces an isomorphism of G to itself (i.e., sending edges to edges). Let Γ be a group of automorphisms of G (not necessarily the whole automorphism group). We say that Γ **acts vertex-transitively on** G if for every pair $v, w \in V(G)$ there is some $g \in \Gamma$ such that $gv = w$. We that G is a **vertex-transitive graph** if the automorphism group of G acts vertex-transitively on G.

In particular, every group Γ acts vertex-transitively on its Cayley graph $\mathrm{Cay}(\Gamma, S)$ by left-multiplication: the action of $g \in \Gamma$ sends each vertex $x \in \Gamma$ to $gx \in \Gamma$, which sends each edge (x, xs) to (gx, gxs), for all $x \in \Gamma$ and $s \in S$.

Theorem 3.4.6 (Vertex-transitive graphs and quasirandom groups)
Let Γ be a finite group with no non-trivial representations of dimension less than K. Then every n-vertex d-regular graph that admits a vertex-transitive Γ action is an (n, d, λ)-graph with $\lambda < \sqrt{dn/K}$.

Note that $\sqrt{dn/K} \le n/\sqrt{K}$, so that a sequence of such Cayley graphs is quasirandom (Definition 3.1.2) as long as $K \to \infty$ as $n \to \infty$.

Proof. Let A denote the adjacency matrix of the graph, whose vertices are indexed by $\{1, \ldots, n\}$. Each $g \in \Gamma$ gives a permutation $(g(1), \ldots, g(n))$ of the vertex set, which induces a representation of Γ on $\mathbb{C}^n$ given by permuting coordinates, sending $v = (v_1, \ldots, v_n) \in \mathbb{C}^n$ to $gv = (v_{g(1)}, \ldots, v_{g(n)})$.

We know that the all-1 vector $\mathbf{1}$ is an eigenvector of A with eigenvalue d. Let $v \in \mathbb{R}^n$ be an eigenvector of A with eigenvalue μ such that $v \perp \mathbf{1}$. Since each $g \in \Gamma$ induces a graph automorphism, $Av = \mu v$ implies $A(gv) = \mu gv$ (check this claim! Basically it is because g relabels vertices in an isomorphically indistinguishable way).

Since $\Gamma v = \{gv : g \in \Gamma\}$ is Γ-invariant, its $\mathbb{C}$-span W is a Γ-invariant subspace (i.e., $gW \subset W$ for all $g \in \Gamma$), and hence a subrepresentation of Γ. Since v is not a constant vector, the Γ-action on v is non-trivial. So W is a non-trivial representation of Γ. Hence $\dim W \geq K$ by hypothesis. Every nonzero vector in W is an eigenvector of A with eigenvalue μ. It follows that μ appears as an eigenvalue of A with multiplicity at least K. Recall that we also have an eigenvalue d from the eigenvector $\mathbf{1}$. Thus

$$d^2 + K\mu^2 \leq \sum_{j=1}^{n} \lambda_j(A)^2 = \operatorname{tr} A^2 = nd.$$

Therefore

$$|\mu| \leq \sqrt{\frac{d(n-d)}{K}} < \sqrt{\frac{dn}{K}}.$$

□

The above proof can be modified to prove a bipartite version, which will be useful for certain applications.

Given a finite group Γ and a subset $S \subset \Gamma$ (not necessarily symmetric), we define the **bipartite Cayley graph** $\mathrm{BiCay}(\Gamma, S)$ as the bipartite graph with vertex set Γ on both parts, with an edge joining g on the left with gs on the right for every $g \in \Gamma$ and $s \in S$.

Theorem 3.4.7 (Bipartite Cayley graphs on quasirandom groups)
Let Γ be a group of order n with no non-trivial representations of dimension less than K. Let $S \subset \Gamma$ with $|S| = d$. Then the bipartite Cayley graph $\mathrm{BiCay}(\Gamma, S)$ is a bipartite-(n, d, λ)-graph for some $\lambda < \sqrt{nd/K}$.

In other words, the second largest eigenvalue of the adjacency matrix of this bipartite Cayley graph is less than $\sqrt{nd/K}$.

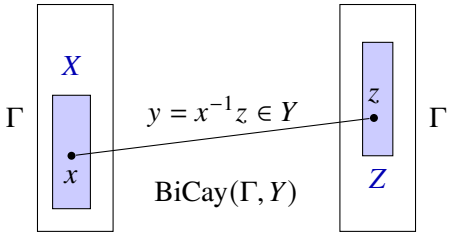
Exercise 3.4.8. Prove Theorem 3.4.7.

As an application of the expander mixing lemma, we show that in a quasirandom group, the number of solutions to $xy = z$ with x, y, z lying in three given sets $X, Y, Z \subset \Gamma$ is close to what one should predict from density alone. Note that the right-hand side expression below is relatively small if K^2 is large compared to $|X|\,|Y|\,|Z|\,/|\Gamma|^3$ (e.g., if X, Y, Z each occupy at least a constant proportion of the group, and K tends to infinity).

Theorem 3.4.9 (Mixing in quasirandom groups)
Let Γ be a finite group with no non-trivial representations of dimension less than K. Let $X, Y, Z \subset \Gamma$. Then

$$\left| |\{(x,y,z) \in X \times Y \times Z : xy = z\}| - \frac{|X|\,|Y|\,|Z|}{|\Gamma|} \right| < \sqrt{\frac{|X|\,|Y|\,|Z|\,|\Gamma|}{K}}.$$

Proof. Every solution to $xy = z$, with $(x, y, z) \in X \times Y \times Z$ corresponds to an edge (x, z) in $\mathrm{BiCay}(\Gamma, Y)$ between vertex subset X on the left and vertex subset Z on the right.



By Theorem 3.4.7, $\mathrm{BiCay}(\Gamma, Y)$ is a bipartite-(n, d, λ)-graph with $n = |\Gamma|$, $d = |Y|$, and some $\lambda < \sqrt{|\Gamma|\,|Y|\,/|K|}$. The above inequality then follows from applying the bipartite expander mixing lemma, Theorem 3.2.9, to $\mathrm{BiCay}(\Gamma, Y)$. □

Corollary 3.4.10 (Product-free sets)
Let Γ be a finite group with no non-trivial representations of dimension less than K. Let $X, Y, Z \subset \Gamma$. If there is no solution to $xy = z$ with $(x, y, z) \in X \times Y \times Z$, then

$$|X|\,|Y|\,|Z| < \frac{|\Gamma|^3}{K}.$$

In particular, every product-free $X \subset \Gamma$ (*product-free* meaning that there is no solution to $xy = z$ with $x, y, z \in X$) has size less than $|\Gamma|\,/K^{1/3}$.

Proof. If there is no solution to $xy = z$, then the left-hand side of the inequality in Theorem 3.4.9 is $|X|\,|Y|\,|Z|\,/|\Gamma|$. Rearranging gives the result. □

The above result already shows that all product-free subsets of a quasirandom group must be small. This sharply contrasts the abelian setting. For example, in $\mathbb{Z}/n\mathbb{Z}$ (written additively), there is a sum-free subset of size around $n/3$ consisting of all group elements strictly between $n/3$ and $2n/3$.

Exercise 3.4.11 (Growth and expansion in quasirandom groups)**.** Let Γ be a finite group with no non-trivial representations of dimension less than K. Let $X, Y, Z \subset \Gamma$. Suppose $|X||Y||Z| \geq |\Gamma|^3/K$. Then $XYZ = \Gamma$ (i.e., every element of Γ can be expressed as xyz for some $(x, y, z) \in X \times Y \times Z$).

Examples of quasirandom groups

Example 3.4.12 (Quasirandom groups)**.** Here are some examples of groups with no small non-trivial representations.

(a) A classic result of Frobenius from around 1900 shows that every non-trivial representation of $\mathrm{PSL}(2, p)$ has dimension at least $(p-1)/2$ for all prime p. A short proof is included below. Jordan (1907) and Schur (1907) computed the character tables for $\mathrm{PSL}(2, q)$ for all prime power q. In particular, we know that every non-trivial representation of $\mathrm{PSL}(2, q)$ has dimension $\geq (q-1)/2$ for all prime power q.

(b) The alternating group A_m for $m \geq 2$ has order $m!/2$, and its smallest non-trivial representation has dimension $m-1 = \Theta(\log n/\log\log n)$. The representations of symmetric and alternating groups have a nice combinatorial description using Young diagrams. See, e.g., Sagan (2001) or Fulton and Harris (1991) for expository accounts of this theory.

(c) Gowers (2008, Theorem 4.7) gives an elementary proof that in every non-cyclic simple group of order n, the smallest non-trivial representation has dimension at least $\sqrt{\log n}/2$.

Recall that the special linear group $\mathrm{SL}(2, p)$ is the group of 2×2 matrices (under multiplication) with determinant 1:

$$\mathrm{SL}(2, p) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{F}_p, ad - bc = 1 \right\}.$$

The projective special linear group $\mathrm{PSL}(2, p)$ is a quotient of $\mathrm{SL}(2, p)$ by all scalars, i.e.,

$$\mathrm{PSL}(2, p) = \mathrm{SL}(2, p)/\{\pm I\}.$$

The following result is due to Frobenius

Theorem 3.4.13 ($\mathrm{PSL}(2, p)$ is quasirandom)
Let p be a prime. Then all non-trivial representations of $\mathrm{SL}(2, p)$ and $\mathrm{PSL}(2, p)$ have dimension at least $(p-1)/2$.

Proof. The claim is trivial for $p = 2$, so we can assume that p is odd. It suffices to prove the claim for $\mathrm{SL}(2, p)$. Indeed, any non-trivial representation of $\mathrm{PSL}(2, p)$

can be made into a representation of $\mathrm{SL}(2,p)$ by first passing through the quotient $\mathrm{SL}(2,p) \to \mathrm{SL}(2,p)/\{\pm I\} = \mathrm{PSL}(2,p)$.

Now suppose ρ is a non-trivial representation of $\mathrm{SL}(2,p)$. The group $\mathrm{SL}(2,p)$ is generated by the elements (Exercise: check!)

$$g = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad h = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

These two elements are conjugate in $\mathrm{SL}(2,p)$ via $z = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ as $gz = zh$. If $\rho(g) = I$, then $\rho(h) = I$ by conjugation, and ρ would be trivial since g and h generate the group. So, $\rho(g) \neq I$. Since $g^p = I$, we have $\rho(g)^p = I$. So $\rho(g)$ is diagonalizable (here we use that a polynomial is diagonalizable if an only if its minimal polynomial has distinct roots, and that the minimal polynomial of $\rho(g)$ divides $X^p - 1$). Since $\rho(g) \neq I$, $\rho(g)$ has an eigenvalue $\lambda \neq 1$. Since $\rho(g)^p = I$, λ is a primitive p-th root of unity.

For every $a \in \mathbb{F}_p^\times$, g is conjugate to

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a^{-1} & 0 \\ 0 & a \end{pmatrix} = \begin{pmatrix} 1 & a^2 \\ 0 & 1 \end{pmatrix} = g^{a^2}.$$

Thus $\rho(g)$ is conjugate to $\rho(g)^{a^2}$. Hence these two matrices have same set of eigenvalues. So λ^{a^2} is an eigenvalue of $\rho(g)$ for every $a \in \mathbb{F}_p^\times$, and by ranging over all $a \in \mathbb{F}_p^\times$, this gives $(p-1)/2$ distinct eigenvalues of $\rho(g)$ (recall that λ is a primitive p-th root of unity). It follows that $\dim \rho \geq (p-1)/2$. □

Applying Corollary 3.4.10 with Theorem 3.4.13 yields the following corollary (Gowers 2008). Note that the order of $\mathrm{PSL}(2,p)$ is $(p^3 - p)/2$.

Corollary 3.4.14 (Product-free subset of $\mathrm{PSL}(2,p)$)
The largest product-free subset of $\mathrm{PSL}(2,p)$ has size $O(p^{3-1/3})$.

In particular, there exist infinitely many groups of order n whose largest product-free subset has size $O(n^{8/9})$.

Before Gowers' work, it was not known whether every order n group has a product-free subset of size $\geq cn$ for some absolute constant $c > 0$ (this was Question 3.4.1, asked by Babai and Sós). Gowers' result shows that the answer is no.

In the other direction, Kedlaya (1997; 1998) showed that every finite group of order n has a product-free subset of size $\gtrsim n^{11/14}$. In fact, he showed that if the group has a proper subgroup H of index m, then there is a product-free subset that is a union of $\gtrsim m^{1/2}$ cosets of H.

Equivalence of quasirandomness conditions

We saw that having no small non-trivial representations is a useful property of groups. Gowers further showed that this group representation theoretic property is equivalent to several other characterizations of the group.

Theorem 3.4.15 (Quasirandom groups)
Let Γ_n be a sequence of finite groups of increasing order. The following are equivalent:

REP The dimension of the smallest non-trivial representation of Γ_n tends to infinity.

GRAPH Every sequence of bipartite Cayley graphs on Γ_n, as $n \to \infty$, is quasirandom in the sense of Theorem 3.1.24.

PRODFREE The largest product-free subset of Γ_n has size $o(|\Gamma_n|)$.
($X \subset \Gamma_n$ is *product-free* if there is no solution to $xy = z$ with $x, y, z \in X$)

QUOTIENT For every proper normal subgroup H of Γ_n, the quotient Γ_n/H is nonabelian and has order tending to infinity as $n \to \infty$.

Let us comment on the various implications.

By Theorem 3.4.7, **REP** implies **GRAPH**. For the converse, we need to construct a non-quasirandom Cayley graph on each group with a non-trivial representation of bounded dimension. One can first construct a weighted analogue of a bipartite Cayley graph with large eigenvalues by appealing to formulas from non-abelian Fourier transform (see Remark 3.4.17 below). And then one can sample a genuine bipartite Cayley graph from the weighted version.

By Corollary 3.4.10, **REP** implies **PRODFREE**. The converse is proved in Gowers (2008) using elementary methods. It was later proved with better polynomial quantitative dependence in Nikolov and Pyber (2011), who proved the following result.

Theorem 3.4.16 (PRODFREE implies REP)
Let Γ be a group with a non-trivial representation of dimension K. Then Γ has a product-free subset of size at least $c\,|\Gamma|\,/K$, where $c > 0$ is some absolute constant.

To see that **REP** implies **QUOTIENT**, note that any non-trivial representation of Γ/H is automatically a representation of Γ after passing through the quotient. Furthermore, every non-trivial abelian group has a non-trivial 1-dimensional representation, and every group of order $m > 1$ has a non-trivial representation of dimension $< \sqrt{m}$. For the proof of the converse, see Gowers (2008, Theorem 4.8). (This implication has an exponential dependence of parameters.)

Remark 3.4.17 (Non-abelian Fourier analysis)**.** (This is an advanced remark and can be skipped over.) Section 3.3 discussed the Fourier transform on finite abelian groups. The

topic of this section can be alternatively viewed through the lenses of the non-abelian Fourier transform. We refer to Wigderson (2012) for a tutorial on the non-abelian Fourier transform from a combinatorial perspective.

Let us give here the recipe for computing the eigenvalues and an orthonormal basis of eigenvectors of $\mathrm{Cay}(\Gamma, S)$.

For each irreducible representation ρ of Γ (always working over $\mathbb{C}$), let

$$M_\rho := \sum_{s \in S} \rho(s),$$

viewed as a $\dim \rho \times \dim \rho$ matrix over $\mathbb{C}$. Then M_ρ has $\dim \rho$ eigenvalues $\lambda_{\rho,1}, \ldots, \lambda_{\rho,\dim \rho}$.

Here is how to list all the eigenvalues of the adjacency matrix of $\mathrm{Cay}(\Gamma, S)$: repeating each $\lambda_{\rho,i}$ with multiplicity $\dim \rho$, ranging over all irreducible representations ρ and all $1 \le i \le \dim \rho$.

To emphasize, the eigenvalues always come in bundles with multiplicities determined by the the dimensions of the irreducible representations of Γ (although it is possible for there to be additional coalescence of eigenvalues).

One can additionally recover a system of eigenvectors of $\mathrm{Cay}(\Gamma, S)$. For each eigenvector v with eigenvalue λ of M_ρ, and every $w \in \mathbb{C}^{\dim \rho}$, set $x^{\rho,v,w} \in \mathbb{C}^\Gamma$ with coordinates

$$x_g^{\rho,v,w} = \langle \rho(g)v, w \rangle$$

for all $g \in \Gamma$. Then x is an eigenvector of $\mathrm{Cay}(\Gamma, S)$ with eigenvalue λ. Now let ρ range over all irreducible representations of Γ, and let v range over an orthonormal basis of eigenvectors of M_ρ (let λ be the corresponding eigenvalue), and let w range over an orthonormal basis of eigenvectors of $\mathbb{C}^{\dim \rho}$, then $x^{\rho,v,w}$ ranges over an orthogonal system of eigenvectors of $\mathrm{Cay}(\Gamma, S)$. The eigenvalue associated to $x^{\rho,v,w}$ is λ.

A basic theorem in representation theory tells us that the regular representation decomposes into a direct sum of $\dim \rho$ copies of ρ ranging over every irreducible representation ρ of Γ. This decomposition then corresponds to a block diagonalization (simultaneously for all S) of the adjacency matrix of $\mathrm{Cay}(\Gamma, S)$ into blocks M_ρ, repeated $\dim \rho$ times, for each ρ. The above statement comes from interpreting this block diagonalization.

The matrix M_ρ, appropriately normalized, is the **non-abelian Fourier transform** of the indicator vector of S at ρ. Many basic and important formulas for Fourier analysis over abelian groups, e.g, inversion and Parseval (which we will see in Chapter 6) have nonabelian analogs.

3.5 Quasirandom Cayley Graphs and Grothendieck's Inequality

Let us examine the following two sparse quasirandom graph conditions (c.f. Remark 3.1.27).

Definition 3.5.1 (Sparse quasirandom graphs)
Let G be an n-vertex d-regular graph. We say that G satisfies property
SparseDISC(ϵ) if $\left|e(X,Y) - \frac{d}{n}|X||Y|\right| \le \epsilon dn$ for all $X, Y \subset V(G)$;
SparseEIG(ϵ) if G is an (n,d,λ)-graph for some $\lambda \le \epsilon d$.

In Section 3.1, we saw that when d grows linearly in n, then these two conditions are equivalent up to a polynomial change in the constant ϵ. As discussed in Remark 3.1.27, many quasirandomness equivalences break down for sparse graphs, meaning $d = o(n)$ here. Some still holds, for example:

Proposition 3.5.2 (SparseEIG implies SPARSEDISC)
Among regular graphs,

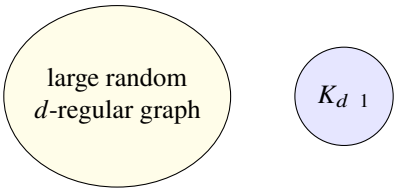
$$\textbf{SparseEIG}(\epsilon) \quad \text{implies} \quad \textbf{SparseDISC}(\epsilon).$$

Proof. In an (n,d,λ) graph with $\lambda \le \epsilon d$, by the expander mixing lemma (Theorem 3.2.4), for every vertex subsets X and Y,

$$\left|e(X,Y) - \frac{d}{n}|X||Y|\right| \le \lambda\sqrt{|X||Y|} \le \epsilon d\sqrt{|X||Y|} \le \epsilon dn.$$

So the graph satisfies **SparseDISC**(ϵ). □

The converse fails badly. Consider the disjoint union of a large random d-regular graph and a K_{d+1} (here $d = o(n)$).



This graph satisfies **SparseDISC**$(o(1))$ since it is satisfied by the large component, and the small component K_{d+1} contributes negligibly to discrepancy due to its size. On the other hand, each connected component contributes a eigenvalue of d (by taking the all-1 vector supported on each component), and so **SparseEIG**(ϵ) fails for any $\epsilon < 1$.

The main result of this section is that despite the above example, if we restrict ourselves to Cayley graphs (abelian or non-abelian), **SparseDISC**(ϵ) and **SparseEIG**(ϵ) are always equivalent up to a linear change in ϵ. This result is due to Conlon and Zhao (2017).

Theorem 3.5.3 (SparseDISC implies SparseEIG for Cayley graphs)
Among Cayley graphs,

SparseDISC(ϵ) implies **SparseEIG**(8ϵ).

As in Section 3.4, we prove the above result more generally for vertex-transitive graphs (see Definition 3.4.5).

Theorem 3.5.4 (SparseDISC implies SparseEIG for vertex-transitive graphs)
Among vertex-transitive graphs,

SparseDISC(ϵ) implies **SparseEIG**(8ϵ).

Grothendieck's inequality

The proof of the above theorem leads us to the following important inequality from functional analysis due to Grothendieck (1953).

Given a matrix $A = (a_{i,j}) \in \mathbb{R}^{m\times n}$, we can consider its $\ell^\infty \to \ell^1$ norm

$$\sup_{\|y\|_\infty \le 1} \|Ay\|_{\ell^1},$$

which can also be written as (exercise: check! Also see Lemma 4.5.3 for a related fact about the cut norm of graphons)

$$\sup_{\substack{x\in\{-1,1\}^m \\ y\in\{-1,1\}^n}} \langle x, Ay\rangle = \sup_{\substack{x_1,\cdots,x_m\in\{-1,1\} \\ y_1,\ldots,y_n\in\{-1,1\}}} \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i y_j. \tag{3.5.1}$$

This quantity is closely related to discrepancy.

One can consider a **semidefinite relaxation** of the above quantity:

$$\sup_{\substack{\|x_1\|,\ldots,\|x_m\|\le 1 \\ \|y_1\|,\ldots,\|y_n\|\le 1}} \sum_{i=1}^m \sum_{j=1}^n a_{i,j} \langle x_i, y_i\rangle, \tag{3.5.2}$$

where the surpremum is taken over vectors $x_1, \ldots, x_m, y_1, \ldots, y_n$ in the unit ball of some real Hilbert space, whose norm is denoted by $\| \|$. Without loss of generality, we can take assume that these vectors lie in $\mathbb{R}^{m+n}$ with the usual Euclidean norm (here $m + n$

dimensions are enough since $x_1, \ldots, x_m, y_1, \ldots, y_n$ span a real subspace of dimension at most $m + n$).

We always have

$$(3.5.1) \leq (3.5.2)$$

by restricting the vectors in (3.5.2) to $\mathbb{R}$. The latter expression (3.5.2) is called a semidefinite relaxation since it can be also written as the supremum of $\sum_{i,j} a_{i,j} M_{i,j}$ over all positive semidefinite matrices M. So (3.5.2) can be efficiently computed using **semidefinite programming**, whereas no efficient algorithm is believed to exist for computing (3.5.1) (Alon and Naor 2006).

Grothendieck's inequality says that this semidefinite relaxation never loses lose more than a constant factor.

Theorem 3.5.5 (Grothendieck's inequality)

There exists a constant $K > 0$ ($K = 1.8$ works) such that for all matrices $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$,

$$\sup_{\|x_i\|, \|y_j\| \leq 1} \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} \langle x_i, y_j \rangle \leq K \sup_{x_i, y_j \in \{\pm 1\}} \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j} x_i y_j,$$

where the left-hand side supremum is taken over vectors $x_1, \ldots, x_n, y_1, \ldots, y_m$ in the unit ball of some real Hilbert space.

Remark 3.5.6. The optimal constant K is known as the **real Grothendieck's constant**. Its exact value is unknown. It is known to lie within $[1.676, 1.783]$. There is also a complex version of Grothendieck's inequality, where the left-hand side uses a complex Hilbert space (and place an absolute value around the final sum). The corresponding **complex Grothendieck's constant** is known to lie within $[1.338, 1.405]$.

We will not prove Grothendieck's inequality here. See Alon and Naor (2006) for three proofs of the inequality, along with algorithmic discussions.

Proof that SparseDISC implies SparseEIG for vertex-transitive graphs

Proof of Theorem 3.5.4. Let G be an n-vertex d-regular graph with a vertex-transitive group Γ of automorphisms. Suppose G satisfies **SparseDISC**(ϵ). Let A be the adjacency matrix of G. Write

$$B = A - \frac{d}{n} J$$

where J is the $n \times n$ all-1 matrix. To show that G is an (n, d, λ)-graph with $\lambda \leq \epsilon d$, it suffices to show that B has operator norm $\|B\| \leq \epsilon d$ (here we are using that G is

d-regular, so the all-1 eigenvector of A with eigenvalue d becomes an eigenvector of B with eigenvalue zero 0).

For any $X, Y \subset V(G)$, the corresponding indicator vectors $x = \mathbf{1}_X \in \mathbb{R}^n$ and $y = \mathbf{1}_Y \in \mathbb{R}^n$ satisfy, by **SparseDISC**(ϵ),

$$|\langle x, By\rangle| = \left| e(X,Y) - \frac{d}{n}|X|\,|Y| \right| \le \epsilon dn.$$

Then, for any $x, y \in \{-1,1\}^n$, we can write $x = x^+ - x^-$ and $y = y^+ - y^-$ with $x^+, x^-, y^+, y^- \in \{0,1\}^n$. Since,

$$\langle x, By\rangle = \langle x^+, By^+\rangle - \langle x^+, By^-\rangle - \langle x^-, By^+\rangle + \langle x^-, By^-\rangle,$$

and each term on the right-hand side is at most ϵdn in absolute value, we have

$$|\langle x, By\rangle| \le 4\epsilon dn \quad \text{for all } x, y \in \{-1,1\}^n. \tag{3.5.3}$$

For any graph automorphism $g \in \Gamma$ and any $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $i \in [n]$, write

$$x^j = \left(\sqrt{\frac{n}{|\Gamma|}} x_{g(j)} : g \in \Gamma \right) \in \mathbb{R}^\Gamma.$$

For every unit vector $x \in \mathbb{R}^n$, the vector $x^j \in \mathbb{R}^\Gamma$ is a unit vector since $x_1^2 + \cdots + x_n^2 = 1$ and the map $g \mapsto g(j)$ is $n/|\Gamma|$-to-1 for each j. Similarly define y^j for any $y \in \mathbb{R}^n$ and $j \in [n]$. Furthermore, $B_{i,j} = B_{g(i),g(j)}$ for any $g \in \Gamma$ and $j \in [n]$ due to g being a graph automorphism.

To prove the operator norm bound $\|B\| \le 8\epsilon d$, it suffices to show that $\langle x, By\rangle \le 8\epsilon d$ for every pair of unit vectors $x, y \in \mathbb{R}^n$. We have

$$\begin{aligned}\langle x, By\rangle = \sum_{i,j=1}^n B_{i,j} x_i y_j &= \frac{1}{|\Gamma|} \sum_{g\in\Gamma} \sum_{i,j=1}^n B_{g(i),g(j)} x_{g(i)} y_{g(j)} \\ &= \frac{1}{|\Gamma|} \sum_{g\in\Gamma} \sum_{i,j=1}^n B_{i,j} x_{g(i)} y_{g(j)} = \frac{1}{n} \sum_{i,j=1}^n B_{i,j} \langle x^i, y^j\rangle \le 8\epsilon d.\end{aligned}$$

The final step follows from Grothendieck's inequality (applied with $K \le 2$) along with (3.5.3). This completes the proof of **SparseEIG**(8ϵ). □

3.6 Second Eigenvalue: Alon–Boppana Bound

The expander mixing lemma tells us that in an (n, d, λ)-graph, a smaller value of λ guarantees stronger pseudorandomness properties. In this chapter, we explore the following natural extremal question.

Question 3.6.1 (Minimum second eigenvalue)
Fix a positive integer d. What is the smallest possible λ (as a function of d alone) such that there exist infinitely many $(n, d, \lambda + o(1))$-graphs, where the $o(1)$ is some quantity that goes to zero as $n \to \infty$?

The following result gives a lower bound on λ (Alon 1986). As we will see later, it turns out to be tight.

Theorem 3.6.2 (Alon–Boppana second eigenvalue bound)
Fix a positive integer d. Let G be an n-vertex d-regular graph. If $\lambda_1 \geq \cdots \geq \lambda_n$ are the eigenvalues of its adjacency matrix, then

$$\lambda_2 \geq 2\sqrt{d-1} - o(1),$$

where $o(1) \to 0$ as $n \to \infty$.

In particular, the Alon–Boppana bound implies that $\max\{|\lambda_2|, |\lambda_n|\} \geq 2\sqrt{d-1} - o(1)$, which can be restated as below.

Corollary 3.6.3 (Alon–Boppana second eigenvalue bound)
For every fixed d and $\lambda < 2\sqrt{d-1}$, there are only finitely many (n, d, λ)-graphs.

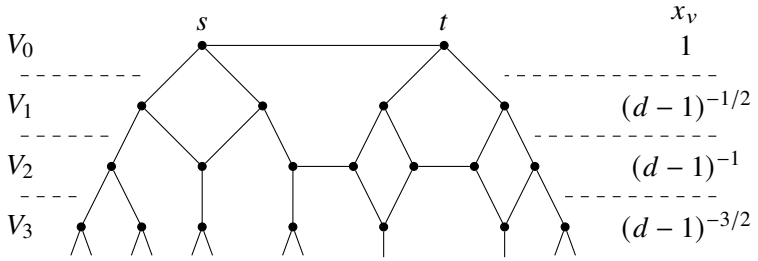
We will see two different proofs. The first proof (Nilli 1991) constructs an eigenvector explicitly. The second proof (only for Corollary 3.6.3) uses the trace method to bound moments of the eigenvalues via counting closed walks.

Lemma 3.6.4 (Test vector)
Let $G = (V, E)$ be a d-regular graph. Let A be the adjacency matrix of G. Let r be a positive integer. Let st be an edge of G. For each $i \geq 0$, let V_i denote the set of all vertices at distance exactly i from $\{s, t\}$ (so that in particular $V_0 = \{s, t\}$). Let $x = (x_v)_{v \in V} \in \mathbb{R}^V$ be a vector with coordinates

$$x_v = \begin{cases} (d-1)^{-i/2} & \text{if } v \in V_i \text{ and } i \leq r, \\ 0 & \text{otherwise, i.e., } \operatorname{dist}(v, \{s, t\}) > r. \end{cases}$$

Then

$$\frac{\langle x, Ax \rangle}{\langle x, x \rangle} \geq 2\sqrt{d-1}\left(1 - \frac{1}{r+1}\right)$$



Proof. Let $L = dI - A$ (this is called the **Laplacian matrix** of G). The claim can be rephrased as an upper bound on $\langle x, Lx\rangle / \langle x, x\rangle$. Here is an important and convenient formula (it can be easily proved by expanding):

$$\langle x, Lx\rangle = \sum_{uv \in E} (x_u - x_v)^2.$$

Since x_v is constant for all v in the same V_i, we only need to consider edges spanning consecutive V_i's. Using the formula for x, we obtain

$$\langle x, Lx\rangle = \sum_{i=0}^{r-1} e(V_i, V_{i+1}) \left(\frac{1}{(d-1)^{i/2}} - \frac{1}{(d-1)^{(i+1)/2}} \right)^2 + \frac{e(V_r, V_{r+1})}{(d-1)^r}$$

For each $i \geq 0$, each vertex in V_i has at most $d-1$ neighbors in V_{i+1}, so $e(V_i, V_{i+1}) \leq (d-1)\,|V_i|$. Thus continuing from above,

$$\begin{aligned}
&\leq \sum_{i=0}^{r-1} |V_i|\,(d-1) \left(\frac{1}{(d-1)^{i/2}} - \frac{1}{(d-1)^{(i+1)/2}} \right)^2 + \frac{|V_r|\,(d-1)}{(d-1)^r} \\
&= \left(\sqrt{d-1} - 1\right)^2 \sum_{i=0}^{r-1} \frac{|V_i|}{(d-1)^i} + \frac{|V_r|\,(d-1)}{(d-1)^r} \\
&= \left(d - 2\sqrt{d-1}\right) \sum_{i=0}^{r} \frac{|V_i|}{(d-1)^i} + \left(2\sqrt{d-1} - 1\right) \frac{|V_r|}{(d-1)^r}.
\end{aligned}$$

We have $|V_{i+1}| \leq (d-1)\,|V_i|$ for every $i \geq 0$, so that $|V_r|\,(d-1)^{-r} \leq |V_i|\,(d-1)^{-i}$ for each $i \leq r$. So continuing,

$$\begin{aligned}
&\leq \left(d - 2\sqrt{d-1} + \frac{2\sqrt{d-1} - 1}{r+1} \right) \sum_{i=0}^{r} \frac{|V_i|}{(d-1)^i} \\
&= \left(d - 2\sqrt{d-1} + \frac{2\sqrt{d-1} - 1}{r+1} \right) \langle x, x\rangle,
\end{aligned}$$

It follows that

$$\frac{\langle x, Ax\rangle}{\langle x, x\rangle} = d - \frac{\langle x, Lx\rangle}{\langle x, x\rangle} \geq \left(2\sqrt{d-1} - \frac{2\sqrt{d-1}-1}{r+1}\right)$$
$$\geq \left(1 - \frac{1}{r+1}\right) 2\sqrt{d-1}. \qquad \square$$

Proof of the Alon–Boppana bound (Theorem 3.6.2). Let $V = V(G)$. Let $\mathbf{1}$ be the all-1's vector, which is an eigenvector with eigenvalue d. To prove the theorem, it suffices to exhibit a nonzero vector $z \perp \mathbf{1}$ such that

$$\frac{\langle z, Az\rangle}{\langle z, z\rangle} \geq 2\sqrt{d-1} - o(1).$$

Let r be an arbitrary positive integer. When n is sufficiently large, there exist two edges st and $s't'$ in the graph with distance at least $2r+2$ apart (indeed, since the number of vertices within distance k of an edge is $\leq 2(1 + (d-1) + (d-1)^2 + \cdots + (d-1)^k)$). Let $x \in \mathbb{R}^V$ be the vector constructed as in Lemma 3.6.4 for st, and let $y \in \mathbb{R}^V$ be the corresponding vector constructed for $s't'$. Recall that x is supported on vertices within distance r from st, and likewise with y and $s't'$. Since st and $s't'$ are at distance at least $2r+2$ apart, the support of x is at distance at least 2 from the support of y. Thus

$$\langle x, y\rangle = 0 \quad \text{and} \quad \langle x, Ay\rangle = 0.$$

Choose a constant $c \in \mathbb{R}$ such that $z = x - cy$ has sum of its entries equal to zero (this is possible since $\langle y, \mathbf{1}\rangle > 0$). Then

$$\langle z, z\rangle = \langle x, x\rangle + c^2 \langle y, y\rangle$$

and so by Lemma 3.6.4

$$\begin{aligned}\langle z, Az\rangle &= \langle x, Ax\rangle + c^2 \langle y, Ay\rangle \\ &\geq \left(1 - \frac{1}{r+1}\right) 2\sqrt{d-1} \left(\langle x, x\rangle + c^2 \langle y, y\rangle\right) \\ &= \left(1 - \frac{1}{r+1}\right) 2\sqrt{d-1} \langle z, z\rangle .\end{aligned}$$

Taking $r \to \infty$ as $n \to \infty$ gives the theorem. $\square$

Remark 3.6.5. The above proof cleverly considers distance from an *edge* rather than from a single vertex. This is important for a rather subtle reason. Why does the proof fail if we had instead considered distance from a vertex?

Now let us give another proof—actually we will only prove the slightly weaker statement of Corollary 3.6.3, which is equivalent to

$$\max\{|\lambda_2|, |\lambda_n|\} \geq 2\sqrt{d-1} - o(1). \tag{3.6.1}$$

As a warmup, let us first prove (3.6.1) with $\sqrt{d} - o(1)$ on the right-hand side. We have

$$dn = 2e(G) = \operatorname{tr} A^2 = \sum_{i=1}^{n} \lambda_i^2 \leq d^2 + (n-1)\max\{|\lambda_2|, |\lambda_n|\}^2.$$

So

$$\max\{|\lambda_2|, |\lambda_n|\} \geq \sqrt{\frac{d(n-d)}{n-1}} = \sqrt{d} - o(1)$$

as $n \to \infty$ for fixed d.

To prove (3.6.1), we consider higher moments $\operatorname{tr} A^k$. This is a useful technique, sometimes called the **trace method** or the **moment method**.

Alternative proof of (3.6.1). The quantity

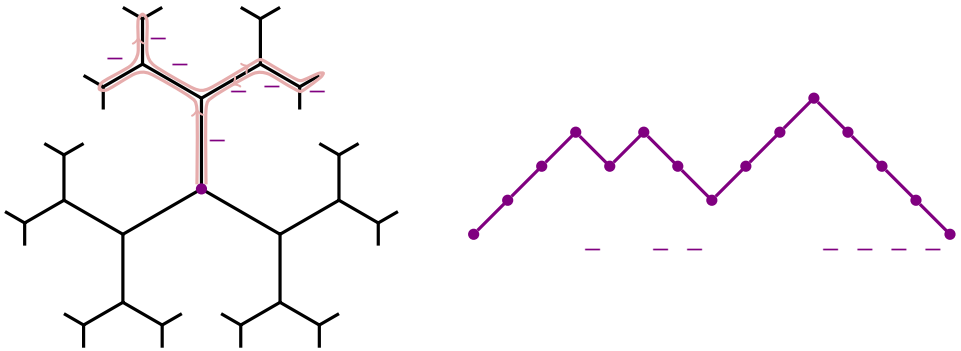
$$\operatorname{tr} A^{2k} = \sum_{i=1}^{n} \lambda_i^{2k}$$

counts the number of closed walks of length $2k$ on G. Let $\mathbb{T}_d$ denote the infinite d-regular tree. Observe that

$$\begin{aligned} &\#\text{ closed length-}2k\text{ walks in } G \text{ starting from a fixed vertex} \\ &\quad \geq \#\text{ closed length-}2k\text{ walks in } \mathbb{T}_d \text{ starting from a fixed vertex.} \end{aligned}$$

Indeed, at each vertex, for both G and $\mathbb{T}_d$, we can label its d incident edges arbitrarily from 1 to d (the labels assigned from the two endpoints of the same edge do not have to match). Then every closed length-$2k$ walk in $\mathbb{T}_d$ corresponds to a distinct closed length-$2k$ walk in G by tracing the same outgoing edges at each step (why?). Note that not all closed walks in G arise this way (e.g., walks that go around cycles in G).

The number of closed walks of length $2k$ on an infinite d-regular graph starting at a fixed root is at least $(d-1)^k C_k$, where $C_k = \frac{1}{k+1}\binom{2k}{k}$ is the k-th Catalan number. To see this, note that each step in the walk is either "away from the root" or "towards the root." We record a sequence by denoting steps of the former type by + and of the latter type by −.



Then the number of valid sequences permuting k +'s and k −'s is exactly counted by the Catalan number C_k, as the only constraint is that there can never be more −'s than +'s up to any point in the sequence. Finally, there are at least $d-1$ choices for where to step in the walk at any + (there are d choices at the root), and exactly one choice for each −.

Thus, the number of closed walks of length $2k$ in G is at least

$$\operatorname{tr} A^{2k} \geq n(d-1)^k C_k \geq \frac{n}{k+1}\binom{2k}{k}(d-1)^k.$$

On the other hand, we have

$$\operatorname{tr} A^{2k} = \sum_{i=1}^{n} \lambda_i^{2k} \leq d^{2k} + (n-1)\max\left\{|\lambda_2|, |\lambda_n|\right\}^{2k}.$$

Thus,

$$\max\left\{|\lambda_2|, |\lambda_n|\right\}^{2k} \geq \frac{1}{k+1}\binom{2k}{k}(d-1)^k - \frac{d^{2k}}{n-1}.$$

The term $\frac{1}{k+1}\binom{2k}{k}$ is $(2-o(1))^{2k}$ as $k \to \infty$. Letting $k \to \infty$ slowly (e.g., $k = o(\log n)$) as $n \to \infty$ gives us $\max\left\{|\lambda_2|, |\lambda_n|\right\} \geq 2\sqrt{d-1} - o(1)$. □

Remark 3.6.6. The infinite d-regular graph $\mathbb{T}_d$ is the universal cover of all d-regular graphs (this fact is used in the first step of the argument). The spectral radius of $\mathbb{T}_d$ is $2\sqrt{d-1}$, which is the fundamental reason why this number arises in the Alon–Boppana bound.

Graphs with $\lambda_2 \approx 2\sqrt{d-1}$

Let us return to Question 3.6.1: what is the smallest possible λ_2 for n-vertex d-regular graphs, with d fixed and n large? Is the Alon–Boppana bound tight? (The answer is yes.)

Alon's second eigenvalue conjecture says that random d-regular graphs match the Alon–Boppana bound. This was proved by Friedman (2008). We will not present the proof, as it is quite a difficult result.

Theorem 3.6.7 (Friedman's second eigenvalue theorem)
Fix positive integer d and $\lambda > 2\sqrt{d-1}$. With probability $1 - o(1)$ as $n \to \infty$ (with n even if d is odd), a uniformly chosen random n-vertex d-regular graph is an (n, d, λ)-graph.

In other words, the above theorem says that random d-random graphs on n vertices satisfy, with probability $1 - o(1)$ (for fixed $d \geq 3$ and $n \to \infty$),

$$\max\{|\lambda_2|, |\lambda_n|\} \leq 2\sqrt{d-1} + o(1).$$

Can we get $\leq 2\sqrt{d-1}$ exactly without an error term? This leads us to one of the biggest open problems of the field.

Definition 3.6.8 (Ramanujan graph)
A **Ramanujan graph** is an (n, d, λ)-graph with $\lambda = 2\sqrt{d-1}$. In other words, it is a d-regular graph whose adjacency matrix has all eigenvalues, except the top one, at most $2\sqrt{d-1}$ in absolute value.

A major open problem is to show the existence of infinite families of d-regular Ramanujan graphs.

Conjecture 3.6.9 (Existence of Ramanujan graphs)
For every positive integer $d \geq 3$, there exist infinitely many d-regular Ramanujan graphs.

While it is not too hard to construct small Ramanujan graphs, e.g., K_{d+1} has eigenvalues $\lambda_1 = d$ and $\lambda_2 = \cdots = \lambda_n = -1$, it is a difficult problem to construct infinitely many d-regular Ramanujan graphs for each d.

The term *Ramanujan graphs* was coined by Lubotzky, Phillips, and Sarnak (1988), who constructed infinite families of d-regular Ramanujan graphs when $d - 1$ is an odd prime. The same result was independently proved by Margulis (1988). The proof of the eigenvalue bounds uses deep results from number theory, namely solutions to the Ramanujan conjecture (hence the name). These constructions were later extended by Morgenstern (1994) whenever $d - 1$ is a prime power. The current state of Conjecture 3.6.9 is given below, and it remains open for all other d, with the smallest open case being $d = 7$.

Theorem 3.6.10 (Existence of Ramanujan graphs)
If $d-1$ is a prime power, then there exist infinitely many d-regular Ramanujan graphs.

All known results are based on explicit constructions using Cayley graphs on $\mathrm{PSL}(2,q)$ or related groups. We refer the reader to the book Davidoff, Sarnak, and Valette (2003) for a gentle exposition of the construction.

Theorem 3.6.7 says that random d-regular graphs are "nearly-Ramanujan." Empirical evidence suggests that for each fixed d, a uniform random n-vertex d-regular graph is Ramanujan with probability bounded away from 0 and 1, for large n.

Conjecture 3.6.11 (A random d-regular graph is likely Ramanujan)
For every $d \geq 3$, there is some $c_d > 0$ so that for all sufficiently large n (with n even if d is odd), a a uniformly chosen random n-vertex d-regular graph is Ramanujan with probability at least c_d.

If this were true, it would prove Conjecture 3.6.9 on the existence of Ramanujan graphs. However, no rigorous results are known in this vein.

One can formulate a bipartite analog.

Definition 3.6.12 (Bipartite Ramanujan graph)
A **bipartite Ramanujan graph** is some bipartite-(n,d,λ)-graph with $\lambda = 2\sqrt{d-1}$.

Given a Ramanujan graph G, we can turn it into a bipartite Ramanujan graph $G \times K_2$. So the existence of bipartite Ramanujan graphs is weaker than of Ramanujan graphs. Nevertheless, for a long time, it was not known how to construct infinite families of bipartite Ramanujan graphs other than using Ramanujan graphs. A breakthrough by Marcus, Spielman, and Srivastava (2015) completely settled the bipartite version of the problem. Unlike earlier construction of Ramanujan graphs, their proof is existential (i.e., non-constructive) and introduces an important technique of *interlacing families of polynomials*.

Theorem 3.6.13 (Bipartite Ramanujan graphs of every degree)
For every $d \geq 3$, there exist infinitely many d-regular bipartite Ramanujan graphs.

Exercise 3.6.14 (Alon–Boppana bound with multiplicity)**.** Prove that for every positive integer d and real $\epsilon > 0$, there is some constant $c > 0$ so that every n-vertex d-regular graph has at least cn eigenvalues greater than $2\sqrt{d-1} - \epsilon$.

Exercise 3.6.15* (Net removal decreases top eigenvalue)**.** Show that for every d and r, there is some $\epsilon > 0$ such that if G is a d-regular graph, and $S \subset V(G)$ is such that every vertex of G is within distance r of S, then the top eigenvalue of the adjacency matrix of $G - S$ (i.e., remove S and its incident edges from G) is at most $d - \epsilon$.

Chapter Summary

- We are interested in quantifying how a given graph can be similar to a random graph.
- **Chung–Graham–Wilson quasirandom graphs theorem** says that several notions are equivalent, notably:
  - **DISC**: edge discrepancy (similar to the ϵ-regular pair condition from the previous chapter),
  - **C_4**: 4-cycle count being close to random, and
  - **EIG**: all eigenvalues (except the largest) small.

  These equivalences only apply to graphs at constant order edge density. Some of the implications break down for sparser graphs.
- An (n, d, λ)**-graph** is an n-vertex d-regular graph all of whose adjacency matrix eigenvalues are $\leq \lambda$ in absolute value except the top one (which must be d). The second eigenvalue plays an important role in pseudorandomness.
- **Expander mixing lemma.** An (n, d, λ)-graph satisfies

$$\left| e(X, Y) - \frac{d}{n} |X| |Y| \right| \leq \lambda \sqrt{|X| |Y|} \qquad \text{for all } X, Y \subset V(G).$$

- The eigenvalues of an abelian Cayley graphs $\mathrm{Cay}(\Gamma, S)$ can be computed via a Fourier transform of 1_S. For example, using a Gauss sum, one can deduce that the Paley graph (generated by quadratic residues) is quasirandom.
- A non-abelian group with no small non-trivial representations is call a **quasirandom group**.
  - Every Cayley graph on a quasirandom group is a quasirandom graph.
  - There are no large **product-free sets** in a quasirandom group.
  - Example of quasirandom group: $\mathrm{PSL}(2, p)$, which has order $(p^3 - p)/2$, and all non-trivial representations have dimension $\geq (p - 1)/2$.
- Among vertex-transitive graphs (which includes all Cayley graphs), the sparse analogs of the discrepancy property (**SparseDISC**) and small second eigenvalue property (**SparseEIG**) are equivalent up to a linear change of the error tolerance parameter. This equivalence is false for general graphs.
  - Proof applies **Grothendieck's inequality**, which says that that semidefinite relaxation of the $\ell^\infty \to \ell^1$ norm (equivalent to the cut norm) gives a constant factor approximation.
- **Alon–Boppana second eigenvalue bound.** Every d-regular graph has second largest eigenvalue $\geq 2\sqrt{d-1} - o(1)$ for the adjacency matrix, with d fixed and the number of vertices $\to \infty$.
  - Two spectral proof methods: (1) constructing a test vector and (2) trace/moment method

- The constant $2\sqrt{d-1}$ is best possible, as a random d-regular graph is typically an (n, d, λ)-graph with $\lambda = 2\sqrt{d-1} + o(1)$ (Friedman's theorem).
- A **Ramanujan graph** is an (n, d, λ)-graph with $\lambda = 2\sqrt{d-1}$. It is conjectured that for every $d \geq 3$, there exist infinitely many d-regular Ramanujan graphs (this is known to hold when $d - 1$ is a prime power). A bipartite version of this conjecture is true.

Further Reading

The survey *Pseudo-random Graphs* by Krivelevich and Sudakov (2006) discusses many combinatorial aspects of this topic.

Expander graphs are a large and intensely studied topic, partly due to many important applications in computer science. Here are two important surveys articles:

- *Expander Graphs and Their Applications* by Hoory, Linial, and Wigderson (2006);
- *Expander Graphs in Pure and Applied Mathematics* by Lubotzky (2012).

For spectral graph theory, see the book *Spectral Graph Theory* by Chung (1997), or the book draft *Spectral and Algebraic Graph Theory* by Spielman.

The book *Elementary Number Theory, Group Theory and Ramanujan Graphs* by Davidoff, Sarnak, and Valette (2003) gives a gentle introduction to the construction of Ramanujan graphs.

The breakthrough by Marcus, Spielman, and Srivastava (2015) constructing bipartite Ramanujan graphs via interlacing polynomials is an instant classic.