Expander Graphs and Kazhdan's Property (T)

Giles Gardam

An essay submitted in partial fulfillment of
the requirements for the degree of
B.Sc. (Honours)

Pure Mathematics
University of Sydney

October 2012

Contents

Introduction

Expanders are highly connected sparse graphs of fundamental interest in mathematics and computer science.

While it is relatively simple to prove the existence of expanders using probabilistic arguments in the style of Erdős, their explicit construction is difficult. The first explicit construction was given by Margulis in [Mar75], and employed Kazhdan's Property (T) from the representation theory of semisimple groups. The aim of this essay is to present an essentially complete exposition of the theory leading to the Margulis construction, while concurrently developing the most important and illustrative general theory of both expander graphs and Property (T). We also present an application of expanders to derandomization.

Chapter 1 introduces expander graphs under several common definitions which we show to be equivalent, and gives some first properties of expanders. We also show the existence of expanders, and consider some algebraic obstructions to groups giving expanders via their Cayley graphs.

Chapter 2 then gives an equivalent algebraic definition of expansion, in terms a graph's adjacency matrix, and shows that expanders for this definition are expanders for the combinatorial definitions of the preceding chapter. We consider random walks on an expander graph, and study the relationship between the graph spectrum and the behaviour of such random walks. The algebraic definition allows us to give a straightforward application of expanders to complexity theory, namely the derandomization of algorithms, where one improves the accuracy of a randomized computational procedure with very little need for additional randomness.

Property (T) is then introduced in Chapter 3. We develop some first consequences of the definition, demonstrated by many examples of groups which do not have Property (T) (with the help of amenability). After proving that all compact groups have Property (T) we consider Kazhdan sets, which can be used to give an alternative definition of Property (T) and are essential for the construction of expanders. We then study lattices in groups with Property (T); finite generation of these lattices was the original motivation for the formulation of Property (T). Finally, we give the proofs that certain non-compact groups have Property (T), which will be used in Chapter 4 to give constructions of expanders.

This work is mostly based on the books by Lubotzky [Lub94] and de la Harpe and Valette [dlHV89], with the chapter on random walks and derandomization based on the survey paper [HLW06] of Hoory, Linial and Wigderson. The more recent and much larger work of Bekka, de la Harpe and Valette on Property (T) in English, [BdlHV08], is also referred to occasionally.

We present the material so as to be accessible for students of pure mathematics at honours level. Efforts have been made in the selection of content so as to give a broad yet gentle introduction to the topics of Chapters 1, 2 and 3, while still leading fairly directly to a

proof of the construction of expanders in Chapter 4. The proof that these constructions are expanders is complete except for classical results regarding lattices, and relative property (T) of the pair $(\mathbb{R}^2 \rtimes \mathrm{SL}(2,\mathbb{R}), \mathbb{R}^2)$ for which we only give a proof outline. There are several other parts of this essay where we call upon powerful results which we cannot prove here, but these are not required for the construction of expanders and are used only to develop more general theory.

The books [dlHV89] and especially [Lub94] are written at an advanced level, and leave much unsaid. As well as filling in the details and correcting some erroneous arguments, our treatment includes many examples, remarks and original figures that give context to the non-specialist, motivate results and proofs, and address potential barriers to understanding. We also avoid developing some technical machinery — such as weak containment of representations, induced representations, the Fell topology, and functions of positive type — as although these would be very useful topics for a detailed study of Property (T), that is not our main purpose.

In particular, in [Lub94] the proof of existence of expanders, Theorem 1.45 in the present essay, contains several 'easy to see' claims which are actually erroneous in certain cases and so the statement has been corrected here. The proof that compact groups have Property (T) in Section 3.3 differs from the literature in order to require fewer external results and to illustrate Property (T) more concretely. The theory of random walks, including an application of expanders to derandomization, which forms Chapter 2, leads naturally to the spectral theory of graphs and gives an interesting background to Margulis's construction in Chapter 4. We also present the results of original computations investigating growth in the groups $\mathrm{SL}(2,\mathbb{Z}/p\mathbb{Z})$ in Section 1.8.

Acknowledgements

I would like to express my gratitude to several people. First and foremost I thank my supervisor, Dr Anne Thomas, for her assistance throughout this project. She has been incredibly helpful, patient and encouraging, and this essay has benefited enormously from her meticulous editing. I am indebted to my family and Lisa for their love and support. I would like to acknowledge the support of Donald Jamieson, St Andrew's College, The School of Mathematics and Statistics and The University of Sydney throughout my degree. Many thanks also to Dr Neil Saunders for helpful comments on the manuscript, and to Dr Emma Carberry for coordinating the pure mathematics honours programme.

Chapter 1

Expander Graphs

This chapter introduces some common definitions of expanders and establishes their equivalence, as well as proving that expanders have various properties. We also establish the existence of expander graphs, and demonstrate some algebraic properties that are necessary conditions for groups to have Cayley graphs that are expanders.

We begin with Section 1.1 covering the necessary background knowledge from graph theory. Section 1.2 then gives our main definition of expanders (Definition 1.14), which makes precise the slogan that expanders are highly connected sparse graphs. This is followed by some examples and first results. Section 1.3 examines the relationship between expansion and diameter, concluding with a proof that expanders have logarithmic diameter (Corollary 1.31). We then introduce in Section 1.4 some alternative definitions of expansion and show their equivalence (Corollary 1.37). We develop a model of random regular graphs in Section 1.5 in order to say in Theorem 1.45 that, under this model, almost all regular graphs are expanders. This establishes the existence of expanders. Section 1.6 introduces Cayley graphs as a useful means to construct regular graphs via group theory, with several examples. We prove in Section 1.7 some negative results for expansion of certain families of Cayley graphs in Theorems 1.51 and 1.53. This algebraic result demonstrates the difficulty of constructing expanders. Finally in Section 1.8 we present the results of original computations that explore expansion in the Cayley graphs of $\mathrm{SL}(2, \mathbb{Z}/p\mathbb{Z})$ and in particular their diameters, which are the subject of an open problem due to Lubotzky, Problem 1.58.

1.1. Graph Theory Background

A graph is a simple abstract representation of a collection of objects and their relationships. There are several different flavours of graphs; we begin with the following definition, and introduction variations afterwards as required.

Definition 1.1 (Graph)**.** A *graph* $X = (V, E)$ is a pair of sets V and E such that $E \subseteq V \times V$. The elements of V are called *vertices*, and the elements of E are called *edges*. A graph with the edges being ordered pairs of vertices is *directed*. If we instead consider the edges to be unordered pairs of elements of V, then it is *undirected*. Edges of the form (v, v) are called *loops*. If a graph is undirected and has no loops, we call it *simple*. If there is an edge (v, w), we say that the vertices v and w are *adjacent* and that the edge (v, w) is *incident* to the vertices v and w.

We often sketch graphs like polygons, or perhaps with curved rather than straight edges, but it is important to remember that the geometry of a particular sketch is not important (at least not in general, and not for our purposes here).

Examples 1.2. The following undirected graphs are shown in Figure 1.1.

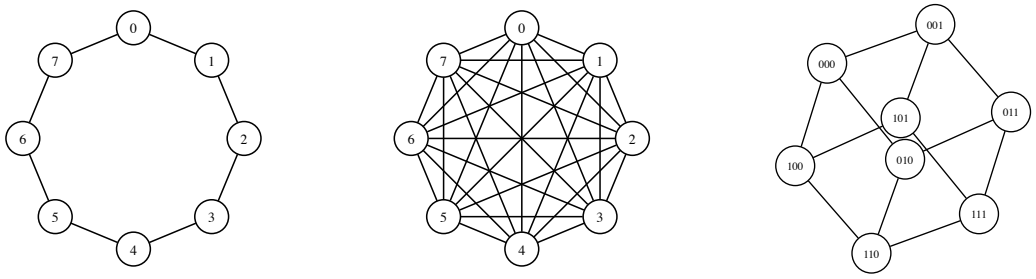

FIGURE 1.1. The graphs C_8, K_8 and Q_3, which all have 8 vertices.

a) The cycle graph on $n \geq 3$ vertices, denoted C_n, is a graph with vertices $V = \mathbb{Z}/n\mathbb{Z} = \{0, 1, 2, \ldots, n-1\}$ and edges joining each i to $i+1$.
b) The complete graph on n vertices, denoted K_n, has an edge joining every pair of distinct vertices. It has $\binom{n}{2} = \frac{n(n-1)}{2}$ edges, which is maximal for an undirected graph.
c) The n^{th} hypercube graph, denoted Q_n, on 2^n vertices, has vertex set $\{0,1\}^n$, and edges joining those n-tuples that differ in exactly one coordinate. It can be identified with the 1-skeleton (that is, the vertices and edges) of the n-dimensional cube, so in particular the graph Q_3 can be identified with the 1-skeleton of the familiar cube (which has 8 vertices and 12 edges in both the geometric and graph-theoretic senses).

It will be necessary sometimes to loosen our definition of a graph to the more general notion of a multigraph, where two vertices can be connected by multiple edges. We will indicate when this generalisation is required, but most of the examples we encounter will be graphs rather than multigraphs. A *multiset* is a generalized set where the elements are allowed to appear more than once (that is, they can have arbitrary positive integer multiplicity).

Definition 1.3 (Multigraph)**.** A *multigraph* $X = (V, E)$ is a pair consisting of a set V and a multiset E such that the elements of E are all members of $V \times V$.

If a graph is simple, it is implicit that it is not a multigraph.

Expanders are "sparse graphs", in a sense that is made precise by the definition of degree. See Remark 1.6 below.

Definition 1.4 (Degree, Regular)**.** The *degree* of a vertex is the number of edges of X incident to it. If all vertices in a graph X have degree k, we say X is a *k-regular* graph. A graph is *regular* if it is k-regular for some non-negative integer k.

Examples 1.5. All cycle graphs are 2-regular, the complete graph K_n is $(n-1)$-regular and the hypercube graph Q_n is n-regular.

Remark 1.6. There is no precise, universally agreed-upon definition of "sparseness" for graphs. It means roughly speaking that the total number of edges is much less than the number of unordered pairs on n vertices, that is $\binom{n}{2}$, which for a simple graph on n vertices is the maximum number of edges possible. To give one possible precise definition, for an

infinite family of graphs (V_i, E_i) we could say that the total number of edges grows sub-quadratically (that is, $|E_i|/|V_i|^2 \to 0$ as $i \to \infty$), but for expanders we will require degree to be constant, so that the total number of edges is $O(|V|)$. (Readers unfamiliar with big-oh notation should refer to Appendix A.)

Remark 1.7. While the theory of infinite graphs is interesting, covering such rich topics as Cayley graphs for infinite groups and automorphism groups of trees, from now on we restrict attention almost exclusively to finite graphs, as these are the graphs for which expansion is defined.

There are many different connectivity notions for graphs, which form an essential part of graph theory. The edges connecting vertices are the fundamental structure that graphs have beyond simply being sets, so in some sense the connectivity properties are the *raison d'être* of graphs.

The most basic connectivity property we could ask of a graph is simply called connectedness.

Definition 1.8. A *walk* between two vertices v and w in a graph is a sequence of vertices

$$v = v_0,\ v_1,\ \ldots,\ v_{l-1},\ v_l = w$$

such that v_i is adjacent to v_{i+1} for $i = 0, 1, \ldots, l-1$. The *length* of such a walk is l. A walk is called a *path* if the vertices are pairwise distinct, that is, $v_i \neq v_j$ whenever $i \neq j$. Two vertices v and w in a graph $X = (V, E)$ are said to be *connected* if there is a path joining them (or, equivalently, if there is a walk joining them). If all pairs of vertices $v, w \in V$ are connected, we say that the graph X is *connected*.

If a graph is not connected, it is often convenient to discuss its connected components. For the definition of connected components, we note that in an undirected graph, the relation of connectedness is an equivalence relation on the set V of vertices: symmetry holds by taking a path of length zero, reflexivity follows from taking a reversed path, and transitivity follows from concatenation of paths.

Definition 1.9. A *connected component* of an undirected graph is an equivalence class of vertices under the equivalence relation of connectedness.

Connected graphs have a natural metric, usually referred to as distance. For unconnected graphs, we can consider two unconnected vertices to be at distance $+\infty$ from each other, but this is not actually a metric because the distances in a metric space must be finite by definition (although with the natural addition in $\mathbb{R}_0^+ \cup \{+\infty\}$ the axioms for a metric still hold).

Definition 1.10. The *distance* between two vertices v and w in a graph, denoted $d(v, w)$, is the length of the shortest path joining them. The distance from v to a subset $A \subseteq V$ of the vertices of the graph is defined to be

$$d(v, A) = \min_{w \in A} d(v, w).$$

Since the distance function on a graph is a "local" property of the graph (pertaining only to the two vertices directly, and to the vertices joining them in a shortest path indirectly), it is often natural to discuss its global maximum.

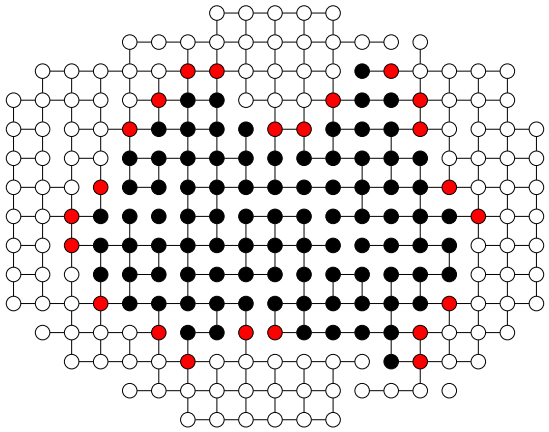

FIGURE 1.2. The boundary of a set of vertices in a graph.

Definition 1.11. The *diameter* of a graph $X = (V, E)$, denoted $\operatorname{diam} X$, is the maximal distance between two vertices, that is,

$$\operatorname{diam} X = \max_{v,w \in V} d(v, w).$$

Good connectivity properties are of course desirable in many networks that can be modelled by graphs, from the very practical examples, such as telecommunications networks, to the theoretical, such as transitions between different inputs to pseudorandom algorithms. If a graph has a small diameter, we might say that it is 'efficient', since any pair of vertices has a short path connecting the two vertices. However, this property alone does not capture enough information about the connectivity of the graph. Most importantly, it is possible for a graph to have a very small diameter but not be what we might call 'robust', in the sense that the removal of a small number of vertices or edges might disconnect the graph (see Figure 1.5 on page 10 for an example). Encapsulating both of these connectivity notions is the boundary of a set of vertices, which will appear in the definition of expander graphs.

Definition 1.12. Let $V = (V, E)$ be a graph and $A \subseteq V$ a subset of its vertices. The *boundary* of A, denoted ∂A, is the set of vertices in the complement of A which are adjacent to at least one vertex in A, that is,

$$\partial A = \{v \in V \mid d(v, A) = 1\}.$$

We will see in Proposition 1.30 exactly how large boundaries imply small diameters.

Figure 1.2 shows a set of vertices A (in black) together with its boundary vertices ∂A (in red). The graph depicted was chosen to have its very particular planar form to illustrate the geometric nature of the boundary. In general, graphs with such a regular geometric structure make poor expanders (as we will explore in Section 1.7).

We need one more graph-theoretic definition.

Definition 1.13. A graph is *bipartite* if its set of vertices V can be partitioned into disjoint sets A and B such that the following holds: all edges are between a vertex in A and a vertex in B. Equivalently, no edge joins a pair of vertices in A or a pair of vertices in B.

1.2. Introduction to Expander Graphs

In this section we give definitions of expander graphs and expander families, illustrated by some examples.

Throughout the rest of this chapter, we will adopt the following convention:

$$X = (V, E) \text{ is a } k\text{-regular graph on } n \text{ vertices}$$

for positive integers n and k.

There are many different definitions of expander graphs, but they are all equivalent up to a change of constant (possibly after some transformation). We will present the most common definitions and prove they are equivalent.

Our first and most important definition of expanders is as follows.

Definition 1.14 (Expander)**.** Let $c > 0$. A finite k-regular graph $X = (V, E)$ on n vertices is called an (n, k, c)-*expander* if for every subset of vertices $A \subseteq V$ with $|A| \leq \frac{n}{2}$,

$$|\partial A| \geq c|A|. \tag{1.15}$$

If n and k are understood, then we will simply call X a c-*expander*.

Remark 1.16. Equation (1.15) means that sets A which are 'not too large' cannot have boundaries that are very small relative to A. A condition like $|A| \leq \frac{n}{2}$ must be imposed, because otherwise as A becomes a large proportion of V, ∂A can only be very small, and consequently c would need to be very small for large graphs (in fact, if we allowed $A = V$ this would force $c = 0$).

Lemma 1.17. Let X be a graph. Then X is a c-expander for some $c > 0$ if and only if X is connected.

Proof. If a graph X is connected, then $|\partial A| > 0$ whenever $0 < |A| < n$, so since there are only finitely many $A \subseteq V$, X will be a c-expander for

$$c(X) := \min_{0<|A|\leq \frac{n}{2}} \frac{|\partial A|}{|A|} > 0. \tag{1.18}$$

If on the other hand X is not connected, letting A be a connected component of minimal size in X we then have $\partial A = \emptyset$ and $|A| \leq \frac{n}{2}$ (since there must be at least 2 connected components). □

Remark 1.19. Although Lemma 1.17 completely classifies which graphs are expanders when taken by themselves, we wish to find *large* sparse graphs with the expansion coefficient c bounded away from zero uniformly. It is applications of expanders that first motivated the following definition. However, we can also motivate it by the observation it allows us to move from the definition of single expander graphs, something which requires choosing an arbitrary c, to families of expanders which can be described more 'qualitatively', that is, without imposing some constant a priori.

Definition 1.20 (Family of expanders)**.** Let k be a positive integer. Let $(X_i) = ((V_i, E_i))$ be a sequence of k-regular graphs such that $|V_i| \to \infty$ as $i \to \infty$. We say that (X_i) is an *expander family* if there is some $c > 0$ such that each graph in this sequence is a c-expander.

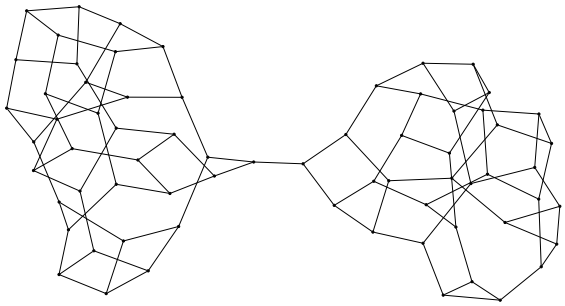

FIGURE 1.3. A graph X with poor expansion.

Figure 1.3 presents a graph X which is a manifestly poor expander. If we take the set A in Definition 1.14 to be the 'cluster' of vertices on the left, then we have $|\partial A| = 1$, but A comprises half of the vertices. Since X is connected, it will be a c-expander for some sufficiently small positive c. However, the best possible c will be very small for a graph of this size, and we can imagine that if we had a sequence of graphs of increasing size with similar structure (only one edge connecting the two 'sides' of the vertex set) then their respective constants c would tend to 0. (Unsurprisingly, the technical term for an edge, like the one joining the two sides in Figure 1.3, whose deletion disconnects the graph, is a *bridge*.)

Example 1.21. The sequence (K_n) of complete graphs is not an expander family because their degree is not constant.

Lemma 1.22. The sequence of cycle graphs (C_n) with $n \geq 3$ is not an expander family.

Proof. Take A in Definition 1.14 to be the vertices on a path of length $\lfloor \frac{n}{2} \rfloor$. Then if C_n is a c-expander, we must have

$$c \leq \frac{|\partial A|}{|A|} = \frac{2}{\lfloor \frac{n}{2} \rfloor} = O(1/n)$$

as $n \to \infty$. □

Corollary 1.23. A family of k-regular expanders must have $k \geq 3$.

Proof. As expanders are connected, they must have degree at least 2, except for the case of graphs on 2 vertices (degree 1 graphs have the form of an even number of vertices connected in disjoint pairs). However, up to isomorphism, the only connected 2-regular graph on $n \geq 3$ vertices is the cycle graph C_n. □

Remark 1.24. As is common practice, we will abuse nomenclature and say that graphs themselves are expanders to mean that they form a family of expanders, and likewise when saying that graphs are not expanders (even though, as noted above, all connected graphs are automatically c-expanders for some sufficiently small c, and could in fact be realised as a member of an expander family by simply adding them to any existing expander family). For example, we say that cycle graphs are not expanders.

Remark 1.25. It turns out that there are in fact expanders of degree 3 (and in certain models of random regular graphs, almost all 3-regular graphs are expanders [HLW06, Theorem 4.16]). As a constructive example, for each prime p we construct a graph whose vertex set

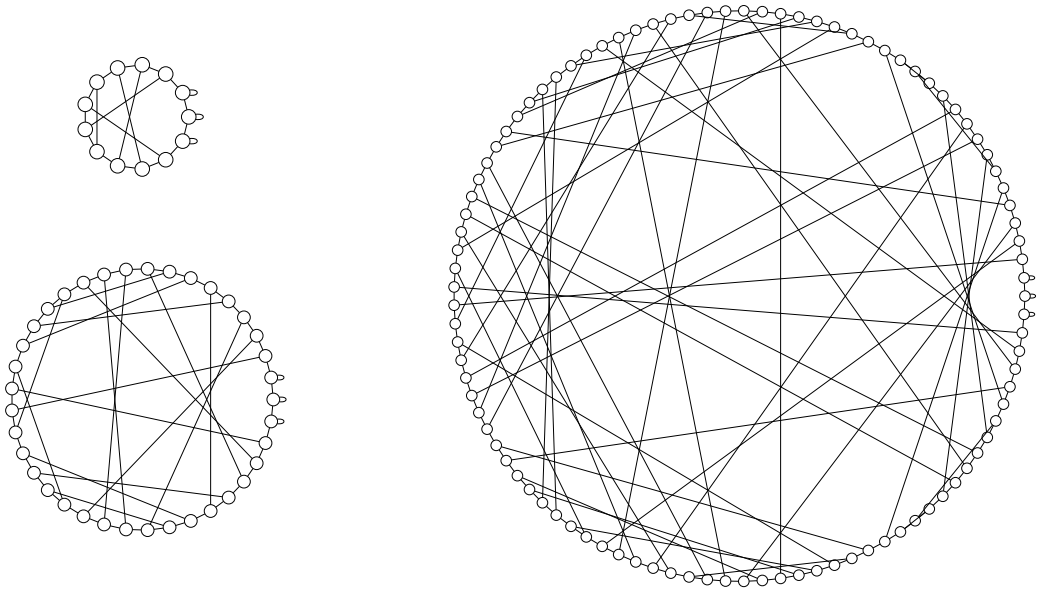

FIGURE 1.4. 3-regular expanders on 13, 37 and 97 vertices.

is $F_p = \mathbb{Z}/p\mathbb{Z}$, with an edge joining each vertex x to $x-1, x+1$ and x^{-1} (taking 0 to be its own inverse). This graph is illustrated in Figure 1.4 for the sequence $p = 13, 37, 97$. These graphs are a family of expanders, indexed by the primes p. However the proof of expansion depends on the Selberg 3/16 Theorem, a deep result from number theory (see [Lub94, p.53] for details). We need to defer to Selberg's theorem because $\mathrm{SL}(2,\mathbb{Z})$ does *not* have Property (T), and so we cannot use Margulis's construction directly (Proposition 4.1, see also Remark 4.7). Note that while this is an explicit construction from a mathematical point of view, it is described in [HLW06, p.453] as being only 'mildly explicit' (in a precise sense defined in that paper) since there is no known efficient deterministic algorithm to generate large primes.

Remark 1.26. An alternative way of defining expander families to Definition 1.20 would be to require that

$$\liminf_{i\to\infty} c(X_i) > 0$$

where c is as defined by Equation (1.18). Naturally, it is possible that a sequence (X_i) of graphs is *not* an expander family for the reason that there is no uniform lower bound on the $c(X_i)$, even though the sequence of graphs (X_i) still has a subsequence that *is* an expander family (consider for example the sequence of constants $0, 1, 0, 1, \ldots$). In this essay, whenever we claim that a sequence of graphs is not an expander family, we will in fact show specifically that $\lim_{i\to\infty} c(X_i) = 0$, so that no subsequence gives an expander family.

Remark 1.27. There is a straightforward upper bound on the expansion constant c in the inequality (1.15). As $\partial A \subseteq V \setminus A$, we have $|\partial A| \leq n - |A|$. Thus for any $A \subseteq V$ of an (n,k,c)-expander, if $|A| = \lfloor n/2 \rfloor$ then we have

$$c \leq \frac{|\partial A|}{|A|} \leq \frac{n - \lfloor n/2 \rfloor}{\lfloor n/2 \rfloor} = \begin{cases} 1 & \text{if } n \text{ is even;} \\ 1 + \frac{2}{n-1} & \text{if } n \text{ is odd.} \end{cases}$$

In a complete graph K_n, the inequalities become equalities, and for any smaller A, that is $|A| < \lfloor n/2 \rfloor$, we have $|\partial A|/|A| > c$. Thus the upper bound on c is tight.

1.3. Diameter in Expanders

We have already mentioned that good expansion guarantees that a graph is both 'robust' and 'efficient', which respectively mean that it is not easily disconnected and that any pair of vertices has a small distance between them. In this section we pursue the second of those notions in detail. We first define balls and spheres in graphs, which will be used to study growth, and show that the boundary of a ball is a sphere. After this setup, we show in Proposition 1.30 and Corollary 1.31 that expanders have logarithmic diameter, and discuss the significance of this result.

Definition 1.28 (Balls and spheres)**.** Let $X = (V, E)$ be a connected, undirected graph with metric d as defined in Definition 1.10. Let $v \in V$ be a vertex of X, and r a non-negative integer. The *ball of radius r centred at v* is defined to be

$$B_r(v) = \{w \in V \mid d(v, w) \leq r\}.$$

The *sphere of radius r centred at v* is defined to be

$$S_r(v) = \{w \in V \mid d(v, w) = r\}.$$

It is immediate from this definition that the balls partition into spheres:

$$B_r(v) = \bigcup_{r'=0}^{r} S_{r'}(v).$$

The following result, however, does require proof (for example, it is not true for arbitrary metric spaces whose distances are integral).

Lemma 1.29. Let v be a vertex of an undirected graph X, and $r \in \mathbb{N}$. Then

$$\partial B_r(v) = S_{r+1}(v).$$

Proof. First let $w \in \partial B_r(v)$. By the definition of boundary, $w \notin B_r(v)$, so $d(v, w) > r$, and there is a $w' \in B_r(v)$ such that $d(w', w) = 1$. Then by the triangle inequality,

$$d(v, w) \leq d(v, w') + d(w', w) \leq r + 1.$$

Thus $d(v, w) = r + 1$, so $w \in S_{r+1}(v)$.

Now let $w \in S_{r+1}(v)$. Then there is a path $w, w', \ldots, v$ in X of length $r + 1$ from w to v, and for this w' we have $d(w, w') = 1$. This path also gives a path of length r from w' to v, so that $w' \in B_r(v)$. As $B_r(v)$ and $S_{r+1}(v)$ are disjoint, $d(w, B_r(v)) > 0$. Thus $d(w, B_r(v)) = 1$, that is, $w \in \partial B_r(v)$. □

Proposition 1.30. Let $X = (V, E)$ be an (n, k, c)-expander. Then

$$\operatorname{diam}(X) \leq \frac{2}{\log(1 + c)} \log n.$$

The proof we now present follows [KS11, p.97].

Proof. Let $v_1, v_2 \in V$ be two arbitrary vertices of X. Because X is an expander it is connected. So any vertex $w \in V$ is joined to v_1 by a path, which must have length at most $n-1$ (by definition, a path consists of distinct vertices). So $d(v_1, w) \leq n-1$. Hence $B_{n-1}(v_1) = V$ and thus

$$\{v_1\} = B_0(v_1) \subseteq B_1(v_1) \subseteq B_2(v_1) \subseteq \cdots \subseteq B_{n-1}(v_1) = V.$$

Since $|V| > \frac{n}{2}$ we can then define r_1 to be the least positive integer r such that

$$|B_r(v_1)| > \frac{n}{2}.$$

Define r_2 similarly, with respect to v_2. For any $r < r_1$ we have $|B_r(v_1)| \leq \frac{n}{2}$, so by Lemma 1.29 and the expansion of X we have

$$\begin{aligned} |B_{r+1}(v_1)| &= |B_r(v_1)| + |S_{r+1}(v_1)| \\ &= |B_r(v_1)| + |\partial B_r(v_1)| \\ &\geq (1+c)|B_r(v_1)|. \end{aligned}$$

As $|B_0(v_1)| = 1$, a trivial induction gives

$$|B_{r_1}(v_1)| \geq (1+c)^{r_1}.$$

Since $|B_{r_1}(v_1)| \leq n$, we now have

$$r_1 \leq \log_{1+c} n = \frac{\log n}{\log(1+c)}$$

and similarly the same upper bound holds for r_2.

Because $|B_{r_1}(v_1)| + |B_{r_2}(v_2)| > n$, these two balls cannot be disjoint, that is, there exists w such that $d(v_1, w) \leq r_1$ and $d(v_2, w) \leq r_2$. Thus

$$d(v_1, v_2) \leq d(v_1, w) + d(w, v_2) \leq \frac{2}{\log(1+c)} \log n.$$

As v_1 and v_2 were arbitrary, we conclude that

$$\operatorname{diam}(X) \leq \frac{2}{\log(1+c)} \log n.$$ □

Corollary 1.31. (Expanders have logarithmic diameter) Let $(X_i) = ((V_i, E_i))$ be a family of expanders. Then

$$\operatorname{diam}(X_i) = O(\log |V_i|).$$

□

Remark 1.32. This corollary is a very simple example of the kind of results that follow from the definition of an expander family. In this way, it validates the definition of expander families. It is only because we have a fixed lower bound on the expansion constants of all the graphs X_i that we can make such a statement: the big-oh notation hides a constant which depends on the constant of expansion c.

Remark 1.33. The diameter of expanders is optimal in the sense that any family of graphs where the vertices have uniformly bounded degree must have at least logarithmic diameter. The proof of this is very similar to the proof of Proposition 1.30 above, noting that if k is an upper bound on degree then $|\partial A| \leq (k+1)|A|$ for all $A \subseteq V$. One might wonder if the

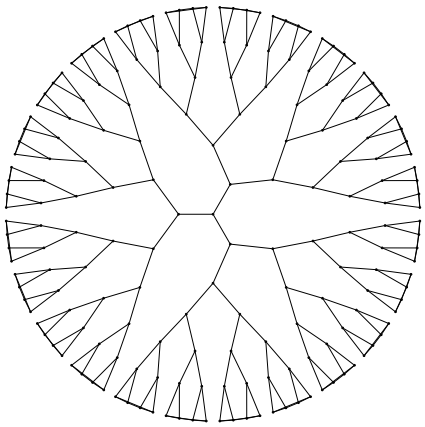

FIGURE 1.5. The ball of radius 6 in the 3-regular tree.

converse to Corollary 1.31 holds, that is, whether a family of graphs of logarithmic diameter is necessarily a family of expanders. The converse however is false, and it is not too difficult to find families of graphs with logarithmic diameter that are not expanders. (Indeed, since the problem of constructing expanders turns out to be so difficult, constructing such counter-examples is essentially the same problem as finding large regular graphs with logarithmic diameter.) A very natural way to construct such examples is to take symmetric trees of constant degree for internal vertices and join the leaves to fill them out to be regular. As the simplest example, we can take balls of radius r inside the 3-regular tree, joining the leaf nodes in local cycles of length 4, as illustrated for $r = 6$ in Figure 1.5. The number of vertices in the balls grows exponentially in r, whereas the diameters of the balls are $2r$, so that they have logarithmic diameter. On the other hand, if we take A in Definition (1.14) as one particular branch at the root, comprising roughly $1/3$ of the vertices but with boundary ∂A containing only the root, we see that these graphs will not form a family of expanders. So in summary, expansion is strictly stronger than logarithmic (that is, optimal) diameter.

1.4. Alternative Definitions of Expansion

In this section, we consider 'edge expansion', the expansion of a graph in terms of the number of edges connecting a subset of its vertices to the rest of the graph, as opposed to Definition 1.14 which we might call 'vertex expansion'. After this, we consider a definition of expansion for bipartite graphs. We will show that these definitions are all equivalent, up to suitable change of constants and transformations.

Definition 1.34 (Edge expansion)**.** Let $X = (V, E)$ be a finite graph. Define the *Cheeger constant* (or *isoperimetric constant*) of X, denoted $h(X)$, by

$$h(X) = \min_{A \sqcup B = V} \frac{|E(A, B)|}{\min(|A|, |B|)}$$

where the (outer) minimum runs over all partitions $V = A \sqcup B$, and $E(A, B)$ is the set of edges between vertices in A and vertices in B.

Remark 1.35. We call $h(X)$ the Cheeger constant in analogue with the Cheeger constant of a Riemannian manifold M, the minimal ratio of the area of a hypersurface that divides M into two disjoint pieces to the smaller of the volumes of those pieces (see [Cha84, p.95] for further details). We will see in the next chapter that the analogy is not superficial, and that bounds relating the Cheeger constant of a manifold to the eigenvalues of its Laplacian have their counterparts for the Cheeger constant of a finite graph and the eigenvalues of its adjacency matrix (or, equivalently, the graph Laplacian).

We could now define an expander family to be a sequence (X_i) of graphs such that the sequence $(h(X_i))$ of Cheeger constants is bounded away from zero. Naturally, we would like to know if this is equivalent to our previous definition of an expander family (Definition 1.20). The following proposition relates the different expansion constants of a single graph, which will give us the desired equivalence of definitions of expander families.

Proposition 1.36 (Bounds between expansion constants)**.** For a k-regular graph X,

$$\frac{h(X)}{k} \leq c(X) \leq h(X).$$

Proof. Suppose $V = A \sqcup B$ is a partition of the vertices. Without loss of generality, we may assume that $|A| \leq |B|$ and thus $|A| \leq \frac{n}{2}$.

Since to each vertex $b \in \partial A \subseteq B$ there corresponds at least one edge $(a, b) \in E(A, B)$, we have $|\partial A| \leq |E(A, B)|$.

As X is k-regular, for each vertex $b \in \partial A$ there are at most k edges incident to b, and in particular there can be at most k edges of the form $(a, b) \in E(A, B)$. Thus $|E(A, B)| \leq k|\partial A|$.

Putting these together, we have

$$\frac{1}{k} \cdot \frac{|E(A,B)|}{\min(|A|,|B|)} \leq \frac{|\partial A|}{|A|} \leq \frac{|E(A,B)|}{\min(|A|,|B|)}.$$

As this holds for any partition with $|A| \leq \frac{n}{2}$, taking minima over all such partitions gives

$$\frac{1}{k} h(X) \leq c(X) \leq h(X). \qquad \square$$

Corollary 1.37 (Equivalence of vertex and edge expansion)**.** Vertex and edge expansion are equivalent, that is, a family of k-regular graphs is a family of vertex expanders (in the sense of Definition 1.14) if and only if it is a family of edge expanders (that is, the Cheeger constants as defined in Definition 1.34 are bounded away from zero). $\square$

Remark 1.38. Note that the equivalence of vertex and edge expansion depends very essentially on the fact that the degree of vertices is bounded (actually a constant k). For example, the trees of diameter 2, which are also known as the *star* graphs, in which one central vertex is joined to all the other vertices, have good edge expansion but poor vertex expansion.

Our third and final combinatorial definition of expanders is for bipartite graphs. While we will only use this definition for the proof of existence below, it is useful in general, and theoretical computer scientists are mostly interested in bipartite expanders. The following definition is taken from [Lub94, p.2].

Definition 1.39 (Bipartite expanders)**.** Let $c > 0$. An (n, k, c)-bipartite expander $X = (V, E)$ is a bipartite, k-regular graph with $V = I \sqcup O$ such that the edges go from I to O, $|I| = |O| = n$ and for any $A \subset I$ with $|A| \leq \frac{n}{2}$ we have

$$|\partial A| \geq (1 + c)|A|.$$

(The two sets I and O are named for *input* and *output*.)

Compared with Definition 1.14, the idea is that since $\partial A \subseteq O$ which is disjoint from I, rather than requiring the boundary ∂A to be not too small relative to A, we instead require it to be *larger* by some fixed proportion.

We turn now to the equivalence of Definitions 1.39 and 1.14. Here we are using the term 'equivalence' rather loosely; there is implicit transformation of graphs involved.

Moving from vertex expansion (Definition 1.14) to bipartite vertex expansion (Definition 1.39) is straightforward. One simply takes the extended bipartite double cover.

Definition 1.40 (Bipartite double cover)**.** Let $X = (V, E)$ be a graph. The *bipartite double cover* of X is a graph whose vertex set is

$$\cup_{i \in \{0,1\}} V = \{v_0 : v \in V\} \cup \{v_1 : v \in V\}$$

and which has two edges, of the form $\{v_0, w_1\}$ and $\{v_1, w_0\}$, corresponding to each edge $\{v, w\} \in E$.

The *extended bipartite double cover* of X is obtained by taking the bipartite double cover, and adding an edge $\{v_0, v_1\}$ for each vertex $v \in V$ (which we may think of joining each vertex to its twin).

Remark 1.41. The bipartite double cover is indeed a 2-sheeted cover of the graph X as a topological space. This is also true for the bipartite double cover of a multigraph X.

Transforming a bipartite expander in the sense of Definition 1.39 into an expander as in Definition 1.14 is more difficult. It requires identifying vertices in I with vertices in O. A *perfect matching* in a bipartite graph is a partitioning of the vertices into pairs of adjacent sets (or alternatively, a subset E' of the edges such that each vertex is incident to precisely one edge $e \in E'$). Hall's 'Marriage Theorem' can be phrased as follows:

Theorem 1.42 (Hall 1935, [Die00, Theorem 2.1.2])**.** Let X be a bipartite graph with edges going between vertex sets I and O, where $|I| = |O|$. There exists a perfect matching in X if and only if $|\partial A| \geq |A|$ for all $A \subseteq I$.

Using this theorem (which we shall not prove),associate a distinct neighbour $w \in O$ to every $v \in I$. We note that applying the theorem does not actually need the expansion property, only the regularity of the graph (and $|I| = |O|$ of course): the $k|A|$ edges leaving A must hit at least $|A|$ vertices in B, as each vertex in B is incident to at most k of these edges. By identifying these pairs of the matching, we get a graph on n vertices.

Remark 1.43. The perfect matching whose existence is guaranteed by Hall's Theorem is not canonical. Moreover, different identifications can result in non-isomorphic quotient graphs, even for small cases such as the cycle graph C_4.

Proposition 1.44. If a graph is a c-expander, then its extended double cover is a c-bipartite expander. If a graph is a c-bipartite expander, then the graph formed by identifying vertices in a perfect matching is a c-expander.

1.5. Existence of Expanders

Pinsker first showed the existence of expander graphs in [Pin73]. In the style of Erdős—who used probabilistic methods to establish the existence of various combinatorial objects — Pinsker considered a model of random regular graphs, and showed that the probability that such a random graph is an expander is non-zero for sufficiently large n. In fact, the probability tends to 1. In this section we follow the example of Pinsker and develop a model of random regular graphs, and use it to prove the existence of expanders.

This line of thinking raises a natural but difficult question: *how should one model a random regular graph?*

Perhaps a nice way to do this would be to consider each isomorphism class of k-regular graphs on n vertices to be equiprobable. However, there is not even a known formula for the number of such isomorphism classes in general! So we model a k-regular bipartite graph on $2n$ vertices as follows. The vertices are labelled $I = \{v_1, \ldots, v_n\}$ and $O = \{w_1, \ldots, w_n\}$. We take k permutations $\pi_1, \ldots, \pi_k \in S_n$, drawn uniformly (each permutation is chosen with probability $\frac{1}{n!}$) and independently. Then for each $1 \leq i \leq n$ and $1 \leq j \leq k$ we create an edge joining v_i and $w_{\pi_j(i)}$. (The random graphs generated are very likely to be multigraphs: even for $n = 2$ we are just asking about the probability that a random permutation is a derangement, which is approximately $1/e$.)

The following theorem is adapted from [Lub94, Proposition 1.2.1] (which in turn follows the presentation in Sarnak's book [Sar90, pp.64-65]). That proposition claims the result to hold true for $k = 5$, however the bound on probability actually diverges to infinity in that case, so the proofs are erroneous in the case $k = 5$. Moreover, it is claimed that a certain function $R(t)$, which we will define below, is decreasing for $1 \leq t < \frac{n}{3}$. This claim is not true, even after ignoring the small variations owing to the parity of t. (In fact, if $k = 6$ then the minimum of $R(t)$ is at $\frac{n}{5}$ approximately.) A proof with all the details would be too long to include here; the reader is referred to [HLW06, pp.478-481]. We sketch the proof from Lubotzky's book.

Theorem 1.45 (Existence of expander families)**.** Let $k \geq 6$ be a positive integer and $c = \frac{1}{2}$. Then the probability that a random k-regular multigraphs on n vertices, drawn from the model described above, is a c-expander tends to 1 as $n \to \infty$. In particular, families of c-expanders exist.

Proof sketch. Consider sets $A \subseteq I$ with $|A| = t \leq \frac{n}{2}$ and $B \subseteq O$ with $|B| = m = \lfloor \frac{3}{2}t \rfloor$. Let $P(t)$ be the probability that for a random k-regular bipartite expander, $\partial A \subseteq B$. We compute that

$$P(t) = \left(\frac{m!(n-t)!}{(m-t)!n!} \right)^k.$$

Let $Q(t)$ be the number of choices of such A and B. Then

$$Q(t) = \binom{n}{t}\binom{n}{m}.$$

Put $R(t) = Q(t)P(t)$. Then a very crude upper bound on the probability that a random k-regular bipartite graph on n vertices is *not* a $c = \frac{3}{2}$ expander is

$$P_n = \sum_{1 \le t \le \frac{n}{2}} R(t).$$

It now remains to show that $P_n \to 0$ as $n \to \infty$.

For small values of t, the probability $P(t)$ is large whereas $Q(t)$ is small. The opposite is true for values more of the order of $\frac{n}{3} \le t \le \frac{n}{2}$. One can verify that $R(t)$ is roughly decreasing for small values, so the maximum occurs at $R(1)$. For the large values, we compare $R(t)$ to $R(\frac{n}{2})$ which we can approximate to within a constant factor by Stirling's formula (Example A.8). Then for all t we have

$$R(t) \le R(1) + R\left(\frac{n}{3}\right) + R\left(\frac{n}{2}\right) = o\left(\frac{1}{n}\right)$$

so that $P_n \to 0$. □

1.6. Cayley Graphs

Cayley graphs give a means to construct graphs from groups. The construction of expanders due to Margulis, which is the main objective of this essay and will be presented in Chapter 4, is a family of Cayley graphs. There are many reasons why it is natural to use Cayley graphs to construct expanders. As well as being regular graphs by definition, they enable us to construct large graphs in an effective and concise manner; it can be much easier to describe the group than the graph. Before describing some previously-encountered graphs as Cayley graphs, we give the definition and some first remarks.

Definition 1.46 (Cayley Graph)**.** The *Cayley graph* of a group G with respect to a generating set $S \subseteq G$ is the directed graph whose vertex set is G and whose edges are given by (g, gs) for each $g \in G$ and $s \in S$. It is denoted by $\mathrm{Cay}(G, S)$.

Remark 1.47. The Cayley graph will be simple (that is, loop-free) if and only if $1 \notin S$.

Since for any distinct $s, s' \in S$ we have $gs \neq gs'$, it follows that the edges of $\mathrm{Cay}(G, S)$ are distinct (so it is not a multigraph).

Because S generates G, every $g \in G$ can be written as a word $s_1 s_2 \cdots s_l$ in the generators S and their inverses, and is reachable from 1_G by the path

$$1_G,\ s_1,\ s_1 s_2,\ s_1 s_2 s_3,\ \ldots,\ s_1 s_2 \cdots s_l = g.$$

Thus $\mathrm{Cay}(G, S)$ is connected.

Finally, we say that $S \subseteq G$ is *symmetric* if

$$S = S^{-1} := \{s^{-1} \mid s \in S\}.$$

If S is symmetric, then we can identify $\mathrm{Cay}(G, S)$ with an undirected graph. This is because to each directed edge (g, gs) there corresponds a reversed edge

$$(gs, (gs)s^{-1}) = (gs, g).$$

(This is still true if $s^{-1} = s$.) This undirected graph is $|S|$-regular. Usually, we will have S a symmetric generating set which does not include the identity, so that the graph $\mathrm{Cay}(G, S)$ is a simple, connected, undirected graph.

Remark 1.48. As expanders are finite graphs, we are only interested in the Cayley graphs of finite groups (note that the number of vertices in $\mathrm{Cay}(G, S)$ is $|G|$). These finite groups might, however, be obtained as subgroups or quotients of infinite groups, so it is certainly not the case that we will consider only finite groups in this essay.

Remark 1.49. It is arguably more common to consider left group actions. The reason we define the Cayley graph by right-multiplication of the vertices $g \in G$ by the generators $s \in S$ is so that the action of G on itself by left-multiplication induces a graph isomorphism (the element $h \in G$ maps any edge (g, gs) to the edge (hg, hgs)).

Examples 1.50. The families of graphs from Examples 1.2, which were illustrated in Figure 1.1 on page 2, can be constructed as Cayley graphs as follows.

a) The Cayley graph for the group $\mathbb{Z}/n\mathbb{Z}$ with respect to the generating set $S = \{1, -1\}$ is the cycle graph C_n.
b) The group $\mathbb{Z}/n\mathbb{Z}$ taken with the generating set $S = \{1, 2, \ldots, n-1\}$ consisting of all non-zero elements has the Cayley graph K_n.
c) The Cayley graph for the direct product of n groups of order 2,

$$G = \bigoplus_{i=1}^{n} \mathbb{Z}/2\mathbb{Z},$$

with the n generators

$$S = \{(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$$

is the n^{th} hypercube graph Q_n.

As previously mentioned, none of these is an expander.

We will now turn to some general negative results for expansion of Cayley graphs.

1.7. Some Negative Results for Cayley Graph Expansion

As expanders resemble random graphs, we might expect that groups which are very uncomplicated would not give expanders. In this section we will prove some results which establish the veracity of this intuition to some extent, first in the case that our particular notion of 'uncomplicated' is 'abelian'.

Theorem 1.51 (Abelian groups do not give expanders)**.** Let (G_i) be a family of abelian groups with respective generating sets S_i of constant cardinality k. Then the graphs $\mathrm{Cay}(G_i, S_i)$ are not a family of expanders.

Proof. We may assume without loss of generality that S_i is symmetric, because we can take $S_i \cup S_i^{-1}$ otherwise (possibly taking some generators more than once so as to get $2k$-regular Cayley graphs). Let $S = \{s_1, \ldots, s_k\}$. Then the balls in the graph can be written

$$B_m(0) = \left\{ a_1 s_1 + \cdots + a_k s_k \;\middle|\; a_i \geq 0, \sum_{i=1}^{k} a_i \leq m \right\}.$$

So the size of $B_m(0)$ is bounded by the number of ways of writing an integer $l \leq m$ as the sum of non-negative integers $a_1, \ldots, a_k$, summed over all possible l. For each fixed l, this is simply making an unordered selection of l objects of k types, with repetition allowed. (The

fact that this selection is unordered is very important, and is where the fact that the group is abelian enters.) It is well-known that this is

$$\binom{l+k-1}{k-1}.$$

We can derive this result by noting that such selections of l objects of k types are in bijection with sequences of $l+k-1$ symbols, $k-1$ of which are $-$'s and which delimit the blocks comprising some of the l $\circ$'s, the blocks encoding how many of that particular item appears in the selection. For example, the selection $(1,2,0,1)$ is represented by the sequence $\circ-\circ\circ--\circ$.

We have

$$\binom{l+k-1}{k-1} \leq \frac{(l+k-1)^{k-1}}{(k-1)!} \leq \frac{2^k}{(k-1)!} l^{k-1}$$

for $l \geq k-1$.

We then have that

$$|B_m(0)| \leq \sum_{l=0}^{m} \binom{l+k-1}{k-1} \leq \sum_{l=1}^{k-1} \binom{l+k-1}{k-1} + \sum_{l=k}^{m} \frac{2^k}{(k-1)!} l^{k-1} \leq \frac{2^k}{(k-1)!} m^k + C$$

for a constant C. Thus $|B_m(0)| = O(m^k)$, where the implied constant depends only on k, and not on the particular abelian group whose Cayley graph we are considering. However, the balls in an expander must grow exponentially (until the comprise at least half the vertices, which will only be an problem for finitely many of the Cayley graphs) as seen in the proof of Proposition 1.30. Thus $\mathrm{Cay}(G_i, S_i)$ is not a family of expanders. □

One can generalize Theorem 1.51 to solvable groups with Theorem 1.53. Solvable groups are a class of groups which can be understood as being 'approximately abelian'. We first recall the definition of solvable groups.

Definition 1.52. A group G is *solvable* if its derived series $G^{(n)}$, defined recursively by $G^{(1)} = G$ and $G^{(n+1)} = [G^{(n)}, G^{(n)}]$, terminates at the trivial group after finitely many steps. That is, there exists a minimal positive integer l called the *derived length* of G such that $G^{(l)} = \{1\}$.

Theorem 1.53. Let (G_i) be a sequence of finite groups with respective generating sets S_i of constant cardinality k. Let l be a positive integer. Suppose that for all n, we have that G_i is solvable with derived length at most l. Then the graphs $\mathrm{Cay}(G_i, S_i)$ are not a family of expanders.

The theorem appears in Krebs and Shaheen [KS11, Theorem 4.47], and is originally due to Lubotzky and Weiss [LW93, Corollary 3.3] (who give it as a corollary to a non-expansion result for quotients of a finitely generated amenable group). We do not prove this theorem as it requires some inheritance results on Cayley graph expansion for subgroups and quotients, which despite not being difficult we do not develop here so as not to stray too far from our path towards the constructions of Chapter 4.

Example 1.54. Finite dihedral groups do not give expanders, as they have derived series of length 2.

Remark 1.55. Theorem 1.53 is essentially the only result known about when it is impossible to choose generating sets to construct a family of expanders as the Cayley graphs of a given family of groups [Kas09]. We will see that in contrast the Margulis construction does work for any fixed set of generators of the group with Property (T) that is used.

The preceding remark on generators raises the question of whether expansion is in fact a group property; Lubotzky and Weiss presented the problem as follows [LW93, Problem 1.1].

Problem 1.56. Let (G_i) be a family of finite groups, with $\langle S_i \rangle = \langle S_i' \rangle = G_i$ and $|S_i|, |S_i'| \leq k$ for all i. Does the fact that $(\mathrm{Cay}(G_i, S_i))$ is an expander family imply the same for $(\mathrm{Cay}(G_i, S_i'))$?

This question was answered in the negative by Alon, Lubotzky and Wigderson in [ALW01]. Their counterexample is beyond the scope of this essay. As noted in [HLW06, p.536], the more recent paper [Kas07] provides a simpler counterexample. In this paper Kassabov answered what had been a big open problem for decades, by demonstrating that for certain generating sets, the alternating groups A_n and symmetric groups Sym_n give expanders. However, one can show that for instance, the generating sets $S_n = \{(1\ 2), (1\ 2\ \ldots\ n)^{\pm 1}\}$ do *not* make $\mathrm{Cay}(\mathrm{Sym}_n, S_n)$ expanders.

Remark 1.57. Expander families are defined by some texts to have uniformly bounded, rather than constant, degree. For instance, the original presentation of Problem 1.56 by Lubotzky and Weiss was for uniformly bounded degree. However, this does not change the problem essentially. In one direction, graphs with constant degree trivially have uniformly bounded degree. In the other direction, if a family of expanders has uniformly bounded degree, we can add edges to obtain constant degree graphs (or, perhaps necessarily, constant degree multigraphs), which only possibly improves their expansion properties.

As a technicality, if the bound on degree is k, we can generally just add an edge between vertices of degree less than k until each graph is k-regular, possibly introducing loops. However, when nk is odd we will have to go to a $(k+1)$-regular graph (the handshaking lemma requires that the sum of degrees is even, being twice the number of edges).

Since we are mostly concerned with Cayley graphs, we will restrict attention to regular graphs.

1.8. Expansion in $\mathrm{SL}(2, \mathbb{Z}/p\mathbb{Z})$

It is a corollary of the Selberg 3/16 Theorem that the Cayley graphs of $\mathrm{SL}(2, \mathbb{Z}/p\mathbb{Z})$ have logarithmic diameter. We present here the results of original computations to determine the diameters of these groups, which reveal a surprisingly well-behaved growth, in light of the very sophisticated result used to get the logarithmic bound. Lubotzky gave the following problem.

Problem 1.58 ([Lub94, Problem 8.1.2])**.** Does there exist a polynomial time algorithm (polynomial in $\log p$) which expresses an element of $\mathrm{SL}(2, \mathbb{Z}/p\mathbb{Z})$ (say $\begin{pmatrix} 1 & \frac{p-1}{2} \\ 0 & 1 \end{pmatrix}$) as a short word (say less than $1000 \log p$) in $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$?

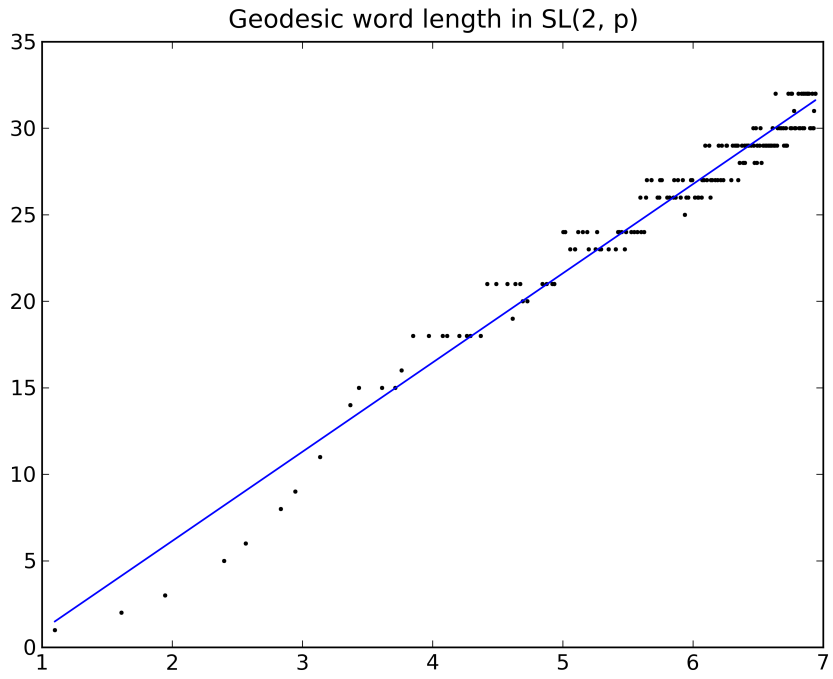

FIGURE 1.6.

Larsen gave a randomized algorithm to construct short word representations, but of length $O(\log p \log\log p)$ rather than $O(\log p)$ [Lar03], which so far as we know is the only work on this problem [Lub12].

The lengths of the shortest word representation of $\begin{pmatrix} 1 & \frac{p-1}{2} \\ 0 & 1 \end{pmatrix}$ in $\{A^{\pm 1}, B^{\pm 1}\}$ of Problem 1.58 are plotted in Figure 1.6, for all primes up to the order of 1000. The results for the diameter of the entire Cayley graphs was similar, albeit with a little more fluctuation.

Chapter 2

Random Walks on Expanders

In this chapter we show that expanders can be characterised by the property that a random walk on their vertices converges to the limit distribution quickly. This allows efficient pseudorandom sampling, which gives a means to reduce the probability of error for a randomized algorithm just as rapidly as repeated random sampling, but with very little additional use of the random resource.

We begin with a discussion of random walks in Section 2.1. This leads naturally into a study of spectral graph theory in Section 2.2. With this background, we can give a motivated definition of spectral expansion, and relate it to the equivalent combinatorial expansion of the previous chapter. We then conclude the chapter with the application of expanders to derandomization in Section 2.3.

Throughout this chapter, all graphs are assumed to be undirected.

2.1. Random Walks and the Graph Spectrum

In this section we introduce random walks on graphs and the adjacency matrix.

A very useful mathematical concept is that of a random walk. On a graph, a random walk moves between adjacent vertices at random.

Definition 2.1 (Random walk, [HLW06, Definition 3.1])**.** A random walk on a finite graph $X = (V, E)$ is a discrete-time stochastic process $(X_0, X_1, \dots)$ taking values in V. The vertex X_0 is sampled from some initial distribution on V, and X_{i+1} is chosen uniformly at random from the neighbours of X_i.

Remark 2.2. Perhaps the first question to ask about a random walk is what its long term behaviour is. In order to study this, we will need to know how the probability distribution of the walk evolves with time.

Let $X = (V, E)$ be a k-regular graph on n vertices. Suppose that x is a *probability distribution vector* that describes a random vertex on the graph at some point in a random walk, that is, $x = (x_1, \dots, x_n)$ where each $x_i \geq 0$ and $\sum_{i=1}^{n} x_i = 1$, and x_i is the probability that vertex v_i is chosen. The probability that walk will be at vertex v_i at the next time step is

$$\frac{1}{k} \sum_{\{v_i, v_j\} \in E} x_j$$

We see that this is a linear operator on $\mathbb{R}^n$. This motivates the following definition.

Definition 2.3 (Adjacency Matrix)**.** The adjacency matrix $A(X)$ of a graph X is defined as follows. Label the vertices $v_1, \ldots, v_n$. Then the (i, j) entry of the matrix is

$$A(X)_{ij} = \begin{cases} 1 & \text{if } v_i \text{ and } v_j \text{ are joined by an edge;} \\ 0 & \text{otherwise.} \end{cases}$$

More generally, for a multigraph X, $A(X)_{ij}$ is the number of edges between v_i and v_j.

Remark 2.4. As we have the running assumption in this chapter that graphs are undirected, the adjacency matrix will be symmetric.

(The interpretation of A as a linear operator still works for non-regular graphs, and for multigraphs, we just need to normalize each column individually.)

For convenience, we will often prefer to discuss the *normalized adjacency matrix* $\hat{A} = \frac{1}{k}A$. Then $\hat{A}_{ij}$ is the probability that a random walk at vertex j will step to the vertex i, so $\hat{A}$ is precisely the linear operator described in Remark 2.2. The probability that the random walk will be at vertex i one time step after the distribution x is

$$(\hat{A}x)_i = \sum_{j=1}^{n} \hat{A}_{ij} x_j.$$

It is not hard to see that if $x = u = \frac{1}{k}(1, 1, \ldots, 1)$, the uniform distribution, then $\hat{A}x = x$, that is, x is an eigenvector of $\hat{A}$ with eigenvalue 1. The uniform distribution is *stationary*. We might imagine that a random walk will always tend towards this stable distribution. To understand whether this is the case, that is, whether the sequence $x, \hat{A}x, \hat{A}^2x, \ldots$ will always tend towards u, we need to study the other eigenvalues of $\hat{A}$.

Throughout the rest of the chapter:

Let $u = (\frac{1}{n}, \ldots, \frac{1}{n})$ denote the uniform distribution probability vector.

2.2. Spectral Expansion

In this section we introduce the graph spectrum and study its most basic properties, and relate the spectral gap to the combinatorial expansion of the graph.

Recall the Real Spectral Theorem [Axl97, Theorem 7.13].

Theorem 2.5. Suppose that V is a real inner-product space and T is a linear operator on V. Then V has an orthonormal basis consisting of eigenvectors of T if and only if T is self-adjoint.

In the language of matrices, this means that since the adjacency matrix $A(X)$ of an undirected graph X on n vertices is symmetric, it will have n real eigenvalues (counting multiplicities) and that the corresponding eigenvectors are orthogonal. We can study a graph via its corresponding eigenvalues, that is, the spectrum of the graph.

Remark 2.6. Spectral graph theory is an area of mathematical research in its own right. Many useful properties of a graph can be inferred from its spectrum, as we shall soon see. Similar matrices have the same spectrum: if v is a λ-eigenvector of A and P is invertible, then $P^{-1}v$ is a λ-eigenvector for $P^{-1}AP$ since

$$(P^{-1}AP)(P^{-1}v) = P^{-1}Av = P^{-1}(\lambda v) = \lambda(P^{-1}v).$$

In particular, this means that the spectrum of a matrix is invariant under conjugation by a permutation matrix. Thus it makes sense to talk about the spectrum of a graph, since it is independent of the particular labelling $v_1, \ldots, v_n$ of the vertices used to write down a particular adjacency matrix (as changing from the matrix corresponding to one particular labelling to another is achieved by conjugation by a permutation matrix).

Lemma 2.7. Let X be a k-regular graph on n vertices, with adjacency matrix A. Let the spectrum of A be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then

a) $\lambda_1 = k$;
b) $\lambda_2 = k$ if and only if X is not connected; and
c) $\lambda_n \geq -k$ with equality if and only if X has a bipartite graph as one of its connected components.

Proof. Let $u = (\frac{1}{n}, \ldots, \frac{1}{n})$. The entry $(Au)_i$ is the sum over all vertices v_j of $\frac{1}{n}$ times the number of edges between v_i and v_j. Since v_i has degree k, we see that $(Au)_i = \frac{k}{n}$, and $Au = ku$.

Since the sum of all entries in A is nk, every eigenvalue of A must have absolute value at most k. If X is not connected, then A will have invariant subspaces corresponding to different connected component, and a suitably chosen (to be orthogonal to u) linear combination of characteristic functions for the spaces will give a second k-eigenvector. Conversely, supposing that we a second k-eigenvector v, as u and v are orthogonal and v is non-zero, v must have both positive and negative entries. But then the vertices corresponding to the maximal (positive) entry in v can only be joined to other such vertices, and this gives a disconnection of X.

Similarly, if X is bipartite with edges between A and B, the regularity of the graph implies $|A| = |B|$ and then the vector $(v_i) = 1_A - 1_B$ will be a $(-k)$-eigenvector. Conversely, with a $(-k)$-eigenvector v orthogonal to u, we can conclude that all vertices corresponding to the maximal entry of v must be joined only to the vertices corresponding to the minimal (negative) entry of v, and vice versa, so the graph has a bipartite connected component. □

Proposition 2.8. Let X be a regular graph. If X is connected and non-bipartite, then any random walk on X tends to the uniform distribution.

Proof. By Lemma 2.7, $u = v_1$ is a eigenvector corresponding to the eigenvalue $\lambda_1 = 1$ of $\hat{A}$, and all other eigenvalues satisfy $|\lambda| < 1$. As the eigenvectors form a basis, any initial probability vector can be written as

$$p = u + a_2 v_2 + \cdots + a_n v_n$$

where the coefficient of u must be 1 since $\langle p, u \rangle = \|u\|^2$. Then

$$\hat{A}^s p = u + \lambda_2^s a_2 v_2 + \cdots + \lambda_n^s a_n v_n \to u$$

as $s \to \infty$. □

The rate of convergence in the above proposition depends on the 'spectral gap' of the adjacency matrix, $k - \max\{|\lambda_2|, |\lambda_n|\}$. It turns out that having a large spectral gap is equivalent to being a combinatorial expander (in the sense of Definition 1.14). With this in mind, we give a definition of spectral expanders.

Definition 2.9 (Spectral expansion)**.** Let X be a k-regular graph on n vertices. Let the spectrum of X be $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then X is a (n, k, α)-*expander* if $|\lambda_2|, |\lambda_n| \leq \alpha\lambda_1$.

Remark 2.10. From Lemma 2.7, we know a regular graph will be an (n, k, α)-expander for some $\alpha < 1$ if and only if the graph is connected and not bipartite. This is exactly like Lemma 1.17 that classifies graphs that are c-expanders for $c > 0$ as connected graphs, and again we note that we require infinite families for which the spectral gap is uniformly bounded away from zero.

Theorem 2.11. Let $X = (V, E)$ be a finite, connected, k-regular graph and let λ be its second eigenvalue. Then

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

Proposition 2.12 (Spectral expansion implies combinatorial expansion)**.** Let X, k and λ be as in Theorem 2.11. Then

$$h(X) \geq \frac{k-\lambda}{2}.$$

Proof. Consider a partition $V = A \sqcup B$ with $|A| \leq |B|$ such that $h(X) = |E(A, B)|/|A|$. The vector (v_i) defined by

$$v_i = \begin{cases} b & \text{if } v_i \in A \\ -a & \text{if } v_i \in B \end{cases}$$

will be orthogonal to u. The rest of the proof is left as a straightforward computational exercise of relating $E(A, B)$ to λ via considering $Av - kv$ (it is similar to a part of the proof of Proposition 4.1). □

Proposition 2.13 (Combinatorial expansion implies spectral expansion)**.** Let X, k and λ be as in Theorem 2.11. Then

$$\sqrt{2k(k-\lambda)} \geq h(X).$$

Proof. See [HLW06, pp.475-477]. □

Remark 2.14. Although *qualitatively* graphs are combinatorial expanders (as in Definition 1.14) if and only if they are spectral expanders (as in Definition 2.9), there is no direct *quantitative* relationship between the particular expansion constants, namely the Cheeger constant and the spectral gap. By this we mean that although there are the bounds of Theorem 2.11, neither is a function of the other. Indeed, there are efficient algorithms to compute the eigenvalues of a matrix, but computing the Cheeger constant is co–**NP**–hard (as first proved by Blum et al. in [BKV81]). Another way in which combinatorial and spectral expansion differ quantitatively is in terms of extremal properties. The best expansion for combinatorial expanders is an open problem [Lub94, Problem 10.1.1], but there is an upper bound on the spectral gap, which is attained by the *Ramanujan graphs* of Lubotzky–Phillips–Sarnak [LPS88] and independently Margulis [Mar88]. The bound is as follows.

Proposition 2.15 (Alon–Boppona, [Lub94, Proposition 4.2.6])**.** Let X_n be a family of k-regular graphs, where k is fixed and n is the number of vertices of the graphs, and $n \to \infty$. Then

$$\limsup_{n\to\infty} \lambda_1(X_n) \leq k - 2\sqrt{k-1}.$$

2.3. Efficient Error Reduction for RP

This section is dedicated to the application of expanders that allows us to derandomize algorithms, that is, to alter a randomized procedure to use the random resource less. This hinges on Theorem 2.16, which gives a quantitative basis to the slogan that 'a random walk on an expander resembles independent sampling'. Note that this is a result in linear algebra, and we will only consider the graph through its adjancency matrix. Accordingly, we use v_i to denote vector components, rather than vertices of a graph.

An important implication is that a random walk on an expander is very unlikely to stay confined in a particular subset of the vertices. In order to quantify this, let (B, s) denote the event that a random walk is confined to B over s time steps, that is, that $V_1, V_2, \ldots, V_t \in B$. A result due to Ajtai–Komlós–Szemerédi and Alon–Deige–Wigderson–Zuckerman is the following. We follow [HLW06, pp.462-463] closely in the following presentation.

Theorem 2.16. Let G be an (n, k, α)-graph and $B \subset V$ with $|B| = \beta n$. Then the probability of the event (B, s) is bounded by

$$\Pr[(B, s)] \leq (\beta + \alpha)^s.$$

Lemma 2.17. Let P denote the orthogonal projection onto B. The probability of the event (B, s) is given by

$$\Pr[(B, s)] = \|(P\hat{A})^s Pu\|_1.$$

Proof. The matrix entry $\hat{A}_{xy}$ is the probability that a random walk at vertex x will step to vertex y, and so the entry $(P\hat{A})_{xy}$ is the probability that a walk at x will step to y *and* that $y \in B$. Thus the probability that a walk of length s starting at x will be confined to to B and end at vertex y is the (x, y) entry of $(P\hat{A})^t$. So finally, summing over all possible terminal vertices y for a uniformly at random x gives that the probability of the event (B, s) is $\|(P\hat{A})^s Pu\|_1$. □

Lemma 2.18. For any vector v,

$$\|P\hat{A}Pv\|_2 \leq (\beta + \alpha)\|v\|_2. \tag{2.19}$$

Proof. We may assume that $Pv = v$, or equivalently that v is supported on B, as otherwise replacing v with Pv will leave the left-hand side unchanged while only possibly decreasing the right-hand side. We can also assume similarly that all the components $v = (v_i)$ satisfy $v_i \geq 0$, because replacing each component v_i with its absolute value $|v_i|$ will leave the right-hand side unchanged while only possible increasing the left-hand side, as each contribution

$$\left(\sum_{j=1}^{n} (P\hat{A}P)_{ij} v_j\right)^2 \leq \left(\sum_{j=1}^{n} (P\hat{A}P)_{ij} |v_j|\right)^2$$

since all matrix entries $(P\hat{A}P)_{ij}$ are non-negative (the same being true of both P and $\hat{A}$). Now since Equation (2.19) holds for $v = 0$ and both sides are linear, we may assume that in fact $\sum_{i=1}^{n} v_i = 1$. We can thus write $v = u + z$ where $\langle u, z \rangle = 0$, that is, $\sum_{i=1}^{n} z_i = 0$ (and $u = (\frac{1}{n}, \ldots, \frac{1}{n})$ as defined above). Since $Pv = v$ and u is a 1-eigenvector for $\hat{A}$ (Lemma 2.7), we have

$$P\hat{A}Pv = P\hat{A}u + P\hat{A}z = Pu + P\hat{A}z.$$

So by the triangle inequality,

$$\|P\hat{A}Pv\|_2 \leq \|Pu\|_2 + \|P\hat{A}z\|_2.$$

We now prove that $\|Pu\|_2 \leq \beta\|v\|_2$ and $\|P\hat{A}z\| \leq \alpha\|v\|_2$, which together imply the claim. The component $(Pu)_i$ is $\frac{1}{n}$ if $i \in B$ and 0 otherwise, and thus

$$\|Pu\|_2^2 = |B| \cdot \left(\frac{1}{n}\right)^2 = \frac{\beta}{n}. \tag{2.20}$$

Since v is supported on B by assumption, $\langle Pu, v\rangle = \sum_{i=1}^{n} \frac{1}{n} v_i = \frac{1}{n}$. Now the Cauchy–Schwartz inequality (Lemma 3.12) gives

$$\frac{1}{n} = \langle Pu, v\rangle \leq \|Pu\|_2 \|v\|_2$$

so that multiplying both sides by $n\|Pu\|_2$ and substituting (2.20) leaves

$$\|Pu\|_2 \leq \beta \|v\|_2 .$$

For the other term, since u and z are orthogonal, z is a linear combination of eigenvectors of $\hat{A}$ with corresponding eigenvalues of absolute value at most α, so $\|\hat{A}z\|_2 \leq \alpha\|z\|$. It is immediate from the definition of P as a projection that $\|P\hat{A}z\|_2 \leq \|\hat{A}z\|_2$. Since $v = u + z$ with u and z orthogonal, $\|z\|_2 \leq \|v\|_2$. Putting all these inequalities together we get $\|P\hat{A}z\|_2 \leq \alpha\|v\|_2$. □

Proof of Theorem 2.16. Since $P^2 = P$, which is both easy to check and true in general of projections, we have by Lemma 2.18 that

$$\|(P\hat{A})^s Pu\|_2 = \|(P\hat{A}P)(P\hat{A})^{s-1}Pu\|_2 \leq (\beta+\alpha)\|(P\hat{A})^{s-1}Pu\|_2.$$

Thus a trivial induction gives

$$\|(P\hat{A})^s Pu\|_2 \leq (\beta+\alpha)^s\|u\|_2.$$

Now

$$\begin{aligned}\|(P\hat{A})^s Pu\|_1 &\leq \sqrt{n}\|(P\hat{A})^s Pu\|_2 \\ &\leq \sqrt{n}(\beta+\alpha)^s\|u\|_2 \\ &= (\beta+\alpha)^s.\end{aligned}$$

□

Thus a random walk on an expander resembles random sampling. This does not mean, however, that one can use a expander to sample a single random element of a set using fewer random bits. Recall from 1.33 that expanders have logarithmic diameter, which is optimal, so it still takes logarithmically many random bits to determine a single random vertex from the whole graph. However, when we repeatedly draw random elements they *appear* to be sampled both independently and uniformly at random, where the illusion is sufficient for many purposes.

Many important algorithms depend essentially on randomness (at least to the extent that there are no known deterministic algorithms that have comparable performance).

Example 2.21. The Miller–Rabin primality test [Rab80] was the first efficient algorithm to test the primality of positive integers. We let $\mathcal{L} = \{2, 3, 5, 7, 11, \dots\}$ denote the *language* of prime numbers. The algorithm tests, for a given positive integer x, whether $x \in \mathcal{L}$.

The algorithm relies on a number-theoretic result that follows from Fermat's little theorem. Let p be a prime, d be the largest odd factor of $p-1$, and let 2^k be the greatest power of 2 dividing $p-1$, so that $p - 1 = 2^k d$. By Fermat's little theorem, $a^{2^k d} \equiv 1 \pmod{p}$. Since $a^2 - 1 = (a-1)(a+1)$, modulo a prime p the only square roots of 1 are ± 1. Thus we compute $a^{2^{k-1}d} \pmod{p}$, which is a square root of 1 and hence either 1 or -1. If it is 1, then we can compute $a^{2^{k-1}d}$, which should again be a square root of 1.

We can continue repeating this until either we have $a^{2^l d} \equiv -1 \pmod{p}$ for some $l \leq k$, or we finish at $a^d \equiv 1 \pmod{p}$. However, this will only necessariy happen if p is a prime. If this does *not* happen for some $a \in \{1, 2, \dots, p-1\}$, we call a a witness to the compositeness of p. The algorithm picks such an a at random, and tests for this property by computing those powers of a (which is computationally efficient).

The effectiveness of the algorithm follows from the theorem of Rabin from his paper where he randomized the ideas of Miller. If p is not a prime, then at least $1/4$ of the possible a's are witnesses to the compositeness of p. Thus, if we can sample such an a at random, the algorithm will correctly identify that p is composite with high probability. Otherwise, it will report that p might be prime.

It was only in 2002 that Agrawal–Kayal–Saxena gave a deterministic polynomial-time algorithm to test primality [AKS04]. It is still much slower in practice than the Miller–Rabin test.

What is not clear from the example is that sampling random elements is non-trivial. To quote [HLW06, p.446]:

> The importance of minimizing the number of random bits may not be evident, but we can assure the reader that it is a basic theoretical problem and, moreover, that getting your hands on good random bits is a nontrivial practical problem.

Definition 2.22. Let the set $\mathcal{L}$ be a language, and suppose that there exists a randomized algorithm A to determine membership of $\mathcal{L}$ with the following properties. To determine whether an input x belongs to $\mathcal{L}$, A samples a random k-bit string r and computes in polynomial time a boolean function $A(x, r)$. Furthermore, if $x \in \mathcal{L}$ then $A(x, r) = 1$ for all $r \in \{0,1\}^k$, and that $A(x, r) = 0$ when $x \notin \mathcal{L}$ for all but $\beta 2^k$ inputs r with $\beta < 1$. Then the language $\mathcal{L}$ is in RP.

Algorithm 2.23. Let A be a randomized algorithm for $\mathcal{L}$ as described in the above definition. To determine membership of $\mathcal{L} \in \mathsf{RP}$, sample uniformly at random $r_0 \in \{0,1\}^k$. Take a $(2^k, d, \alpha)$ expander graph on vertex set $\{0,1\}^k$, and form a random walk of length s. Then return `true` if $A(x, r_i) = 1$ for $i = 0, 1, \dots, s$ and `false` otherwise.

Proposition 2.24. The algorithm fails on $x \notin \mathcal{L}$ with probability at most $(\alpha + \beta)^s$, and always succeeds on $x \in \mathcal{L}$.

Proof. For $x \notin \mathcal{L}$, the algorithm will fail precisely if each random input $r_i \in B$, the set of bad inputs, where $|B| = \beta 2^k$. The result thus follows from Theorem 2.16. It is clear that if $x \in \mathcal{L}$ then the algorithm will always correctly return `true`. □

Remark 2.25. One can get a similar result for algorithms that can err on both sides, that is, algorithms that might incorrectly determine $x \notin \mathcal{L}$ when in fact $x \in \mathcal{L}$. One does this by taking a random walk on the expander, and taking the majority answer of $A(x, r_i)$. Bounding the probability of error is however much more difficult in this case.

Remark 2.26. The algorithm achieves a probability of error that decays exponentially in s using only k random bits to sample r_0 and $s \log d$ random bits to sample $r_1, \ldots, r_s$. That is, the number of random bits needed is $k + O(s)$.

Remark 2.27. In many ways this application of expanders justifies the definition of an expander family. The uniform bound on the spectral gap for a family whose size grows to infinity is necessary to be able to derandomize an algorithm for all possible input sizes. More particularly, the degree of a graph determines how many random bits we need to sample for each time step of the random walk. We need this to be constant in order to achieve an decay of the probability of error that is exponential in the number of additional random bits required.

Remark 2.28. An issue overlooked in [HLW06] is that in order to have an exponentially decaying bound on the probability error, we require $\alpha + \beta < 1$. If one has a very good spectral expander, such as a Ramanujan graph (Remark 2.14), then $\alpha \in (0, 1)$ will not be very close to 1. With expanders of spectral expansion α much closer to 1, we can artificially decrease β by considering all pairs of random inputs $(r, r') \in \{0, 1\}^{2k}$, or even triples etc., which will effectively replace β with β^2 or β^3 respectively. This comes at a cost of requiring more randomness, but still gives exponential decay.

Chapter 3

Kazhdan's Property (T)

Property (T) was introduced by Kazhdan in his seminal and remarkably short paper [Kaz67]. It was used to demonstrate that a large class of lattices in semisimple Lie groups are finitely generated. We present this result in Corollary 3.90.

Property (T) is defined in terms of unitary representations, so we recall the necessary background on Hilbert spaces, group representations and topological groups in Section 3.1. After this we give a definition of Property (T) in Section 3.2 in terms of invariant vectors, along with first results and some easy non-examples that are afforded by the theory of amenability. Section 3.3 is dedicated to an original proof that compact groups have Property (T). Kazhdan sets can be used to phrase one of many alternative definitions of Property (T), and in Section 3.4 we study these sets and their relation to generation in Kazhdan groups. We then turn to the theory of lattices, which historically are closely tied with Property (T), in Section 3.5. Section 3.6 gives examples, with proof, of non-compact Lie groups which have Property (T), which will be used to construct expanders in Chapter 4.

3.1. Unitary Representations of Locally Compact Groups

This section recalls the background needed to define a unitary representation of a locally compact groups: Hilbert spaces, group representations and topological groups. We give examples which we will need later in this chapter.

Our presentation of Hilbert spaces will follow the standard text by Conway [Con90].

Definition 3.1 (Inner product)**.** Let $\mathcal{H}$ be a vector space over $\mathbb{C}$. An *inner product* on $\mathcal{H}$ is a function $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ such that for all $\alpha, \beta \in \mathbb{C}$ and $x, y, z \in \mathcal{H}$, the following are satisfied:

a) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;
b) $\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$;
c) $\langle x, x \rangle \geq 0$ with equality if and only if $x = 0$; and
d) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

A vector space $\mathcal{H}$ over $\mathbb{C}$ together with an inner product on $\mathcal{H}$ will be called an *inner product space*. Usually the particular inner product will be implicit, and we refer to $\mathcal{H}$ by itself as an inner product space.

Example 3.2. The space $\mathbb{C}^n$ has the inner product

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = z_1 \overline{w_1} + \cdots + z_n \overline{w_n}.$$

An inner product endows a vector space with a norm. We present this result without proof (see [Con90, Corollary 1.5]).

Corollary 3.3 (Norm)**.** Let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space $\mathcal{H}$ and define $\|x\| = \sqrt{\langle x, x \rangle}$ for all x in $\mathcal{H}$. Then $\|\cdot\|$ is a norm for $\mathcal{H}$, that is:

a) $\|x + y\| \leq \|x\| + \|y\|$ for all x, y in $\mathcal{H}$;
b) $\|\alpha x\| = |\alpha| \, \|x\|$ for all α in $\mathbb{C}$ and $x \in \mathcal{H}$; and
c) $\|x\| \geq 0$ with equality if and only if $x = 0$.

A normed vector space V has a natural metric structure, given by defining $d(x, y) = \|x - y\|$ for all x, y in V. The metric space structure then gives a topology on V. We are now able to define the class of inner product spaces with which we will work from now on.

Definition 3.4 (Hilbert space)**.** Let $\mathcal{H}$ be an inner product space with norm $\|\cdot\|$ given by its inner product. Then $\mathcal{H}$ is called a *Hilbert space* if it is complete with respect to the norm topology on $\mathcal{H}$.

Remark 3.5. One can also study inner product spaces over $\mathbb{R}$, but for purposes of developing theory around Property (T) will consider only Hilbert spaces over $\mathbb{C}$. From now on, we adopt the following convention:

$$\mathcal{H} \text{ is a Hilbert space over } \mathbb{C}.$$

Example 3.6 (Finite-dimensional Hilbert spaces)**.** The space $\mathbb{C}^n$ with the inner product defined in Example 3.2 is a finite-dimensional Hilbert space. Moreover, any n-dimensional Hilbert space can be identified with $\mathbb{C}^n$ (that is, they are isometrically isomorphic, which is the natural equivalence for Hilbert spaces).

Example 3.7 (L^2 spaces)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space, where the measure μ is defined on the σ-algebra $\mathcal{A}$ of subsets of X. Let

$$\mathcal{L}^2 = \left\{ f : X \to \mathbb{C} \;\middle|\; \int_X |f(x)|^2 d\mu(x) < \infty \right\}$$

be the vector space of square-integrable complex functions on X. If we quotient out by the equivalence relation

$$f \sim g \Leftrightarrow \int_X |f(x) - g(x)|^2 d\mu(x) = 0$$

then we get the Hilbert space $L^2(X)$. The inner product is

$$\langle f, g \rangle = \int_X f(x) \overline{g(x)} d\mu(x).$$

Remark 3.8. When dealing with $L^2(X)$, we will consistently abuse notation and refer to its elements as functions f, rather than equivalence classes of functions $[f]$. However, we will not make any statements about such a function f that are not true when we perturb f on a set of measure 0.

Example 3.9 (Direct sum of Hilbert spaces)**.** For a family $\{\mathcal{H}_i \mid i \in I\}$ of Hilbert spaces, we can define the direct sum

$$\oplus_{i \in I} \mathcal{H}_i$$

to be the complex vector space of $(x_i)_{i \in I}$ such that

$$\sum_{i \in I} \|x_i\|^2 < \infty.$$

When given the inner product

$$\langle (x_i), (y_i) \rangle = \sum_{i \in I} \langle x_i, y_i \rangle$$

this is a Hilbert space.

Many properties follow immediately from the definition of an inner product space. We prove here a few that we will need later.

Lemma 3.10 (Parallelogram rule)**.** For $x, y \in \mathcal{H}$ we have

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 .$$

Proof. From the definition of the norm and linearity of the inner product we have

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle \end{aligned}$$

and

$$\begin{aligned} \|x-y\|^2 &= \langle x-y, x-y \rangle \\ &= \langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 - \langle x, y \rangle - \langle y, x \rangle. \end{aligned}$$

Summing these gives the desired equality. □

Lemma 3.11 (Apollonius)**.** Let $a, b, p \in \mathcal{H}$ and let $m = \frac{1}{2}(a+b)$. Then

$$\|p-m\|^2 + \frac{1}{4}\|a-b\|^2 = \frac{1}{2}\|p-a\|^2 + \frac{1}{2}\|p-b\|^2 .$$

Proof. Let $x = p-a$, $y = p-b$ in the Parallelogram Rule (Lemma 3.10). Then

$$\|2p-a-b\|^2 + \|b-a\|^2 = 2\|p-a\|^2 + 2\|p-b\|^2 .$$

As $\|2p-a-b\|^2 = \|2(p-m)\|^2 = 4\|p-m\|^2$, the result follows. □

We quote without proof a well-known and useful proposition [Con90, Proposition 1.4].

Proposition 3.12 (Cauchy–Schwarz Inequality)**.** Let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space $\mathcal{H}$. Then

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$$

for all x and y in $\mathcal{H}$. Moreover, equality occurs if and only if there are scalars $\alpha, \beta \in \mathbb{C}$, not both zero, such that $\alpha x = \beta y$.

Definition 3.13 (Topological group)**.** A *topological group* G is a group together with a Hausdorff topology on G such that the maps

$$G \times G \to G : (g, h) \mapsto gh \text{ and } G \to G : g \mapsto g^{-1}$$

are continuous.

A *locally compact topological group* is a topological group such that the topology is locally compact, that is, each point has a compact neighbourhood.

Remark 3.14. The continuity conditions in Definition 3.13 can be thought of as the requirement that the algebraic and topological structures on G be compatible. Nonetheless, we note that local compactness is a defined as a purely topological property that has no direct relation to the group structure (although the continuity requirement enforces an indirect relation, so for instance it suffices that *any* element of the group has a compact neighbourhood since G acts on itself transitively by homeomorphisms).

Remark 3.15. Not all authors require a topological group G to be Hausdorff. However, since the axioms imply that translation and inversion are homeomorphisms of G, it is a simple exercise to show that if G satisfies the T_0 separation axiom then it is actually T_2 (Hausdorff): if U is an open neighbourhood of x but not y, then $yU^{-1}x$ is an open neighbourhood of y but not x. (In fact, the topology will be $T_{3\frac{1}{2}}$, that is, completely regular Hausdorff.) Stated differently, if a topological group is *not* Hausdorff, then the elements of G are not topologically distinguishable in general, so it is a coarse topology. So the convenient requirement that the topology be Hausdorff is rather mild.

Examples 3.16. The following are several familiar topological groups. They are all locally compact.

a) Any group is a topological group when endowed with the discrete topology. (Any finite topological group is generally assumed to have the discrete topology.)
b) The additive groups $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are topological groups under the usual topology (given by the Euclidean metric).
c) The general linear group
$$\mathrm{GL}(n,\mathbb{R}) = \{T : \mathbb{R}^n \to \mathbb{R}^n \mid T \text{ is invertible}\}$$
is a topological group. The topology is defined by identifying $\mathrm{GL}(n,\mathbb{R})$ with a subset of $\mathbb{R}^{n^2}$, as each operator T can be identified with an $n \times n$ matrix. (The Euclidean metric topology on $\mathbb{R}^{n^2}$ is then used.) The special linear group $\mathrm{SL}(2,\mathbb{R})$ is a also a topological group, since it is a subgroup of $\mathrm{GL}(2,\mathbb{R})$.

Definition 3.17. A *group representation* of a group G on a vector space V is a group homomorphism
$$\pi : G \to \mathrm{GL}(V)$$
of G into the general linear group on V, that is, the group of invertible linear operators on V. A representation π is *irreducible* if V has no non-trivial subspace W that is invariant under the action of G, that is, no subspace $\{0\} \neq W \subsetneq V$ such that
$$\pi(g)w \in W$$
for all $g \in G$ and $w \in W$.

We will mostly be interested in unitary representations. These are defined on Hilbert spaces and are required to preserve the geometric structure, that is, each $\pi(g)$ must be a surjective isometry. Furthermore, we demand a continuity condition.

Definition 3.18 (Isometry)**.** A linear map $T : \mathcal{H} \to \mathcal{H}'$ between inner product spaces is an *isometry* if it preserves the inner product, that is,
$$\langle Tu, Tv\rangle = \langle u, v\rangle$$
for all $u, v \in \mathcal{H}$.

Definition 3.19 (Unitary group)**.** Let $\mathcal{H}$ be a Hilbert space over $\mathbb{C}$. The *unitary group* on $\mathcal{H}$, denoted $\mathcal{U}(\mathcal{H})$, is the set of all (linear) surjective isometries $T : \mathcal{H} \to \mathcal{H}$.

Lemma 3.20. The unitary group $\mathcal{U}(\mathcal{H})$ under composition of maps is indeed a group.

Proof. It is clear that the identity on $\mathcal{H}$ is a surjective isometry. If $S, T \in \mathcal{U}(\mathcal{H})$, then ST is surjective and

$$\langle STu, STv\rangle = \langle S(Tu), S(Tv)\rangle = \langle Tu, Tv\rangle = \langle u, v\rangle$$

for all $u, v \in \mathcal{H}$, so that $ST \in \mathcal{U}(\mathcal{H})$. Finally, each $T \in \mathcal{U}(\mathcal{H})$ is an isometry, so $\|Tv\| = \|v\|$ for all $v \in \mathcal{H}$ and thus T is injective. So $T : \mathcal{H} \to \mathcal{H}$ is a linear bijection, and has an inverse T^{-1}. Its inverse is an isometry since

$$\langle T^{-1}u, T^{-1}v\rangle = \langle TT^{-1}u, TT^{-1}v\rangle = \langle u, v\rangle$$

for all $u, v \in \mathcal{H}$. □

We are now able to define unitary representations [BdlHV08, Definition A.1.1].

Definition 3.21 (Unitary representation)**.** A *unitary representation* of a topological group G on a Hilbert space $\mathcal{H}$ is a group homomorphism $\pi : G \to \mathcal{U}(\mathcal{H})$ which is *strongly continuous* in the sense that the mapping

$$G \to \mathcal{H} : g \mapsto \pi(g)\xi$$

is continuous for every vector ξ in $\mathcal{H}$.

Remark 3.22. Since any unitary operator is a bounded linear function, the map

$$\mathcal{H} \to \mathcal{H} : \xi \mapsto \pi(g)\xi$$

is continuous for each $g \in G$. We can actually say more: the evaluation map $G \times \mathcal{H} \to \mathcal{H}$ is jointly continuous.

Lemma 3.23. Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be a strongly continuous unitary representation. Then the map

$$G \times \mathcal{H} \to \mathcal{H} : (g, \xi) \mapsto \pi(g)\xi.$$

is continuous.

Proof. Let $g_0 \in G, \xi_0 \in \mathcal{H}$ and $\varepsilon > 0$. Since each $\pi(g) \in \mathcal{U}(\mathcal{H})$ is an isometry, if $\|\xi - \xi_0\| < \varepsilon/2$ then $\|\pi(g)\xi - \pi(g)\xi_0\| < \varepsilon/2$. By the continuity of $g \mapsto \pi(g)\xi_0$, there is an open neighbourhood $U \ni g_0$ such that for all $g \in U$ we have

$$\|\pi(g)\xi_0 - \pi(g_0)\xi_0\| < \varepsilon/2.$$

By the triangle inequality for the norm on $\mathcal{H}$, we then have that if $(g, \xi) \in U \times B(\xi_0, \varepsilon/2)$ then $\|\pi(g)\xi - \pi(g_0)\xi_0\| < \varepsilon$. Thus the map $G \times \mathcal{H} \to \mathcal{H} : (g, \xi) \mapsto \pi(g)\xi$ is continuous. □

> From now on we will assume that all unitary representations are strongly continuous.

Remark 3.24. A way of viewing the strong continuity required in Definition 3.21, which is is perhaps more sophisticated, would be to give the group $\mathcal{U}(\mathcal{H})$ the strong operator topology, and then require that representation $\pi : G \to \mathcal{U}(\mathcal{H})$ be a homomorphism of *topological* groups (that is, a continuous group homomorphism). However, for the sake of simplicity, we will not pursue this perspective further.

It is natural to ask whether a group has any non-trivial unitary representations. In a fashion similar to how Cayley's Theorem shows that every group G embeds in the symmetric group $\mathrm{Sym}(G)$, we can make the group G act on a vector space that is indexed by G. As we need a Hilbert space for a unitary representation, we take $L^2(G)$ to be that vector space.

Definition 3.25 (Left-regular representation)**.** Let G be a locally compact group. The *left-regular representation* of G, denoted λ_G, is defined as follows. For each $g \in G$ we define $\lambda_G(g) : L^2(G) \to L^2(G)$ by

$$(\lambda_G(g)f)(x) = f(g^{-1}x)$$

for all $x \in G$.

Proposition 3.26. Let G be a locally compact group. Then λ_G, the left-regular representation of G, is a unitary representation of G.

Proof. We use the notation $g \cdot f$ for $\lambda_G(g)f$. Let $f \in L^2(G)$ and $g_1, g_2 \in G$. For each $x \in G$ we have

$$\begin{aligned}((g_1g_2) \cdot f)(x) &= f((g_1g_2)^{-1}x) \\ &= f(g_2^{-1}(g_1^{-1}x)) \\ &= (g_2 \cdot f)(g_1^{-1}x) \\ &= (g_1 \cdot (g_2 \cdot f))(x)\end{aligned}$$

so that $(g_1g_2) \cdot f = g_1 \cdot (g_2 \cdot f)$. As the identity of g acts as the identity on $L^2(G)$, λ_G is a group homomorphism.

If $f_1, f_2 \in L^2(G)$, $g \in G$ and $c \in \mathbb{C}$ then

$$\begin{aligned}(g \cdot (f_1 + f_2))(x) &= (f_1 + f_2)(g^{-1}x) \\ &= f_1(g^{-1}x) + f_2(g^{-1}x) \\ &= (g \cdot f_1)(x) + (g \cdot f_2)(x) \\ &= (g \cdot f_1 + g \cdot f_2)(x)\end{aligned}$$

and similarly $g \cdot (cf) = c(g \cdot f)$, so each $\lambda_G(g)$ is linear. Thus λ_G is a group representation, and so we immediately have that each $\lambda_G(g)$ is surjective as it has an inverse $\lambda_G(g^{-1})$.

Now, to see that each λ_G is an isometry, by the left-invariance of the Haar measure we have

$$\begin{aligned}\langle g \cdot f_1, g \cdot f_2 \rangle &= \int_G f_1(g^{-1}x)\overline{f_2(g^{-1}x)}d\mu(x) \\ &= \int_G f_1(g^{-1}gy)\overline{f_2(g^{-1}gy)}d\mu(gy) \\ &= \int_G f_1(y)\overline{f_2(y)}d\mu(y) \\ &= \langle f_1, f_2 \rangle.\end{aligned}$$

Showing the continuity of this representation is a technical result that is left to the reader. It requires the approximation (in the L^2 sense) of function by continuous functions with compact support. □

Definition 3.27 (Right-regular representation)**.** Let G be a locally compact group. The *right-regular* representation of G is defined for each $f \in L^2(G)$ by

$$(\rho_G(g)f)(x) = f(xg)$$

for all $x \in G$.

We omit the proof that ρ_G is a unitary representation, as it is almost identical to the proof for λ_G.

Example 3.28 (Direct sum of representations)**.** If $\pi_i : G \to \mathcal{U}(\mathcal{H}_i)$ are unitary representations, then we can form the *direct sum* of these representations on the direct sum of the Hilbert spaces $\mathcal{H}_i$ by defining

$$\pi(g)((x_i)_{i\in I}) = \bigoplus_{i\in I} \pi_i(g)(x_i).$$

It follows immediately from the definition of the Hilbert space $\oplus_{i\in I}\mathcal{H}_i$ that this is also unitary, and continuity follows similarly.

3.2. Property (T)

We now have the necessary background to define Property (T). The definition is rather involved and technical, so we first give some definitions regarding invariant and almost invariant vectors.

Definition 3.29 (Invariant vectors)**.** Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be a (strongly continuous) unitary representation of a locally compact group G. For a given $\varepsilon > 0$ and $K \subseteq G$ compact, we say that a unit vector $\xi \in \mathcal{H}$ is (ε, K)-*invariant* if

$$\sup\{\|\pi(g)\xi - \xi\| : g \in K\} < \varepsilon. \tag{3.30}$$

We say that π *has almost invariant vectors* if, for all (ε, K), there exists an (ε, K)-invariant unit vector. Finally, we say that π has *non-zero invariant vectors* if there exists $\eta \in \mathcal{H}$ with $\eta \neq 0$ such that $\pi(g)\eta = \eta$ for all $g \in G$.

Remark 3.31. We restrict to unit vectors in the definition of invariant vectors because otherwise by scaling down any given ξ to have sufficiently small norm we could always satisfy Equation (3.30). Alternatively, we could require that

$$\sup\{\|\pi(g)\xi - \xi\| : g \in K\} < \varepsilon \|\xi\|$$

with no restriction on $\xi \in \mathcal{H}$. We could then normalize such a ξ if desired (indeed, the strict inequality means that $\xi \neq 0$).

We illustrate the definition of invariant vectors with the following lemmas which we will need later.

Lemma 3.32. If $\xi \in \mathcal{H}$ is (ε, K)-invariant, then it is $(\varepsilon, K \cup K^{-1})$-invariant.

Proof. For each $g \in K$, since $\pi(g^{-1})$ is an isometry we have

$$\|\pi(g)\xi - \xi\| = \left\|\pi(g^{-1})(\pi(g)\xi - \xi)\right\| = \left\|\pi(g^{-1})\xi - \xi\right\|.$$

□

Lemma 3.33. Suppose that $\xi \in \mathcal{H}$ is (ε, K)-invariant. Let $n \in \mathbb{N}$, and

$$K^n = \{k_1 \cdots k_n \mid k_1, \ldots, k_n \in K\}.$$

Then ξ is $(n\varepsilon, K^n)$-invariant.

Proof. Let

$$\delta = \sup\{\|\pi(g)\xi - \xi\| : g \in K\} < \varepsilon.$$

For $k = k_1 \cdots k_n \in K^n$ we have by the triangle inequality that

$$\begin{aligned}
\|\pi(k)\xi - \xi\| &\leq \|\pi(k_1 \cdots k_n)\xi - \pi(k_1 \cdots k_{n-1})\xi\| + \|\pi(k_1 \cdots k_{n-1})\xi - \xi\| \\
&= \|\pi(k_1 \cdots k_{n-1})(\pi(k_n)\xi - \xi)\| + \|\pi(k_1 \cdots k_{n-1})\xi - \xi\| \\
&= \|\pi(k_n)\xi - \xi\| + \|\pi(k_1 \cdots k_{n-1})\xi - \xi\| \\
&\leq \delta + \|\pi(k_1 \cdots k_{n-1})\xi - \xi\|.
\end{aligned}$$

Proceeding by induction we get $\|\pi(k)\xi - \xi\| \leq n\delta$, so since $k \in K^n$ was arbitrary, the vector ξ is $(n\varepsilon, K^n)$-invariant. □

Definition 3.34 (Property (T))**.** A locally compact group G has *Property (T)*, or is a *Kazhdan group*, if any (strongly continuous) unitary representation of G which has almost invariant vectors has a non-zero invariant vector.

Remark 3.35. The T in Property (T) stands for 'trivial'. One can equip the unitary dual of G, consisting of equivalence classes of irreducible unitary representations of G and denoted by $\hat{G}$, with a natural topology called the Fell topology. Property (T) then amounts to saying that the trivial representation is an isolated point in the Fell topology on $\hat{G}$. However, we will not consider $\hat{G}$ and the Fell topology further, preferring instead the more 'hands on' invariant vectors definition.

Example 3.36. The additive group of integers $\mathbb{Z}$ does not have Property (T). We do not have to look very far to find a unitary representation of $\mathbb{Z}$ which has almost invariant vectors but not invariant vectors: the left-regular representation is one such representation. Let $\varepsilon > 0$ and $K \subseteq \mathbb{Z}$ be compact. Since $\mathbb{Z}$ is a discrete group, K is finite. Thus K is bounded, so there exists $c \in \mathbb{Z}^+$ such that for all $g \in K$ we have $|g| \leq c$. To get an (ε, K)-invariant vector in $L^2(\mathbb{Z})$, we just need to find a function that is sufficiently 'wide'. Let $d > c$ be a positive integer, and define

$$f(x) = \begin{cases} 1/(2d+1)^{1/2} & \text{if } |x| \leq d \\ 0 & \text{otherwise.} \end{cases}$$

where the factor of $(2d+1)^{1/2}$ is chosen to make f a unit vector. Let $g \in \mathbb{Z}$ with $|g| \leq c$, and suppose without loss of generality that $g \geq 0$ (since $\|g \cdot f - f\| = \|f - g^{-1} \cdot f\|$ in general). Then

$$(g \cdot f - f)(x) = \begin{cases} 1/(2d+1)^{1/2} & \text{if } d < x \leq d + g \\ -1/(2d+1)^{1/2} & \text{if } -d \leq x < -d + g \\ 0 & \text{otherwise.} \end{cases}$$

so that

$$\|g \cdot f - f\|^2 = \frac{2g}{2d+1} \leq \frac{2c}{2d+1}.$$

So for sufficiently large d, the vector f will be (ε, K)-invariant.

However, the left-regular representation $\lambda_{\mathbb{Z}}$ does *not* have non-zero invariant vectors. If

$$\|g \cdot f - f\| = 0$$

then

$$\sum_{x \in \mathbb{Z}} |f(x-g) - f(x)|^2 = 0$$

so that $f(x-g) = f(x)$ for all x. If this holds even just for $g = 1$, then f must be constant (when applying a similar argument to a different group such as $\mathbb{R}$, we would need the fact that an invariant vector f has $\|g \cdot f - f\| = 0$ for *all* g). Since $\mathbb{Z}$ does not have finite measure, this implies that the square-summable f must in fact be zero.

With the Heine–Borel Theorem and a little extra care surrounding the Lebesgue measure, we can extend the same argument to see that $\mathbb{R}$ does not have Property (T). One could similarly work with $\mathbb{Z}^d$ or $\mathbb{R}^d$.

Remark 3.37. Property (T) and amenability can be considered to be opposites. We will not develop any of the theory of amenability in this essay; the curious reader is referred to Appendix G of [BdlHV08]. Briefly, a group is amenable if it has a left-invariant mean, that is, a non-negative linear functional Λ of norm 1 defined on the essentially-bounded measurable functions on G such that $\Lambda(g \cdot f) = \Lambda(f)$ for all $g \in G$, $f \in L^\infty(G)$.

We assume Propositions 3.38 and 3.39 below without proof, as they will allow us to develop rapidly some interesting background theory for Property (T). (Note however that these results and the subsequent results that use them will not be required for the constructions of expanders given in Chapter 4.)

Proposition 3.38 ([BdlHV08, Theorem G.2.1])**.** Every abelian group is amenable.

Proposition 3.39 ([Lub94, Corollary 3.1.6])**.** Every amenable locally compact group with Property (T) is compact.

Remark 3.40. As compact groups are amongst the easy first examples for both amenability and Property (T) (see [BdlHV08, Example G.1.5] and Proposition 3.52 below respectively), we might say that any group which is both amenable and Kazhdan is trivially so.

Remark 3.41. The reason that Proposition 3.39 appears in [Lub94] as a Corollary is that one of many equivalent definitions of amenability is that the left-regular representation of that group has almost invariant vectors. If the group also has Property (T), then the left-regular representation has a non-zero invariant vector. Similarly to in Example 3.36, the Haar measure of the whole group must be finite. This implies that the group is compact.

Remark 3.42. Local compactness is required for Proposition 3.39, as for instance the unitary group $\mathcal{U}(\mathcal{H})$ on an infinite-dimensional separable Hilbert space $\mathcal{H}$ is amenable and has Property (T) (as was shown by Bekka in [Bek03]), but $\mathcal{U}(\mathcal{H})$ is *not* compact. Indeed, if a group is not even locally compact then it cannot possibly be compact.

We now use the 'black box' Propositions 3.38 and 3.39 to develop several results, each of which is demonstrated by an example of a group which consequently does not have Property (T).

Corollary 3.43. Let G be an infinite discrete abelian group. Then G does not have Property (T).

Proof. By Proposition 3.38, G is amenable. If G were moreover Kazhdan, then it would be compact by Proposition 3.39. However, a discrete group is compact if and only if it is finite. □

Example 3.44. The free abelian group $\mathbb{Z}^d$ does not have Property (T).

Proposition 3.45. Let G_1 and G_2 be topological groups and $\phi : G_1 \to G_2$ a continuous homomorphism with dense image. If G_1 has Property (T), then so does G_2.

Proof. Let $\pi : G_2 \to \mathcal{U}(\mathcal{H})$ be a unitary representation of G_2 that has almost invariant vectors. We can pull back π to a unitary representation of G_1, letting

$$\psi = \pi \circ \phi : G_1 \to \mathcal{U}(\mathcal{H}).$$

As both π and ϕ are group homomorphisms, so is ψ. For each $\xi \in \mathcal{H}$, the map $g \mapsto \psi(g)\xi$ is the composition of the continuous maps $g_1 \mapsto \phi(g_1)$ and $g_2 \mapsto \pi(g_2)\xi$, hence continuous. Thus ψ is indeed a unitary representation of G_1.

For any $\varepsilon > 0$ and $K \subseteq G_1$ compact, the image $\phi(K) \subseteq G_2$ is compact. As π has almost invariant vectors, there exists a unit vector $\xi \in \mathcal{H}$ such that

$$\sup\{\|\pi(g_2)\xi - \xi\| : g_2 \in \phi(K)\} < \varepsilon$$

and thus

$$\sup\{\|\psi(g_1)\xi - \xi\| : g_1 \in K\} < \varepsilon.$$

Thus ψ has almost invariant vectors, so since G_1 is Kazhdan there is a non-zero $\eta \in \mathcal{H}$ such that $\psi(g_1)\eta = \eta$ for all $g_1 \in G_1$. So $\pi(g_2)\eta = \eta$ for all $g_2 \in \phi(G_1)$. Since $\phi(G_1)$ is dense in G_2 by assumption and $g_2 \mapsto \pi(g_2)\eta$ is continuous, $\pi(g_2)\eta = \eta$ for all $g_2 \in G_2$. Thus π has an almost invariant vector. The unitary representation π was arbitrary, so G_2 has Property (T). □

Corollary 3.46. Let G be a Kazhdan group with closed normal subgroup N. Then G/N has Property (T).

Proof of Corollary 3.46. Since the quotient map is a continuous surjection, this follows immediately from Proposition 3.45. □

Remark 3.47. The requirement that N be closed is to ensure that the quotient space is Hausdorff.

Example 3.48. Any non-abelian free group F_d on $d \geq 2$ generators does not have Property (T). If F_d were a Kazhdan group, then the abelianisation $F_d/[F_d, F_d] \cong \mathbb{Z}^d$ would have Property (T) also. We saw in Example 3.44 that this is not the case.

Corollary 3.49. Let G_1 and G_2 be locally compact topological groups, and suppose that G_1 has Property (T) and G_2 is amenable. If $\phi : G_1 \to G_2$ is a continuous group homomorphism, then $\phi(G_1)$ is relatively compact.

Proof. We have $\phi(G_1) < G_2$, and so the continuity of the group operations gives that $\overline{\phi(G_1)} < G_2$. By Proposition 3.45, since $\phi(G_1)$ is dense in $\overline{\phi(G_1)}$, the latter has Property (T). Now [BdlHV08, Corollary G.3.4] says that closed subgroups of amenable locally compact

groups are amenable. Moreover, since $\overline{\phi(G_1)}$ is a closed subgroup of a locally compact group, it is itself locally compact (this holds more generally for topological spaces [Mun00, Corollary 29.3]). Thus $\overline{\phi(G_1)}$ is an amenable locally compact group with Property (T), hence compact by Proposition 3.39. □

Example 3.50. The general linear group $\mathrm{GL}(n, \mathbb{R})$ does not have property (T). This is because $\det : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}^*$ is a surjective map onto a non-compact abelian group.

Remark 3.51. We noted in Remark 3.42 that we can only conclude that an amenable group with Property (T) is compact if that group is locally compact. This is why we need to take the closure of $\phi(G_1)$ in Corollary 3.49: a subgroup of a locally compact group is not necessarily locally compact, even if it is the continuous image of a locally compact group (similarly, closure is needed for subgroups to inherit amenability). For example, the image of $\phi : \mathbb{Z} \to \mathbb{T} : n \mapsto e^{in}$ is not locally compact. One way to see this is to note that it is the countable union of closed sets but has empty interior, contradicting Theorem 3.70, the Baire Category Theorem.

3.3. Compact Groups Have Property (T)

This section is dedicated to the proof of the following proposition.

Proposition 3.52. All compact groups have Property (T).

The proof we present in this section differs from standard proofs, such as in [dlHV89], [Lub94] and [BdlHV08], in ways and for reasons we now explain. Those three texts have proofs that use an orbit of an almost invariant vector to construct one particular vector, in such a way that since the action preserves the orbit it fixes that vector.

Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be a unitary representation of a compact group G, and let $\varepsilon > 0$. Suppose that $\xi \in \mathcal{H}$ is an (ε, G)-invariant vector. The texts [dlHV89] and [Lub94] both prove Proposition 3.52 with a similar method to the proof that compact groups are amenable: they use the fact that a compact group has finite Haar measure in an essential way (refer to Definition 3.76 of Haar measure). One can use the measure to 'average' the translates $\pi(g)\xi$ over G in order to obtain an invariant vector (by the invariance of the Haar measure), that is, one takes

$$\eta = \int_G \pi(g)\xi dg.$$

We only require a very mild choice of ε to ensure that $\eta \neq 0$. Let $\varepsilon = \sqrt{2}$. Then

$$\sup\{\|\pi(g)\xi - \xi\| : g \in G\} = \sqrt{2 - 2\delta}$$

for some $\delta > 0$. Now each $\langle \pi(g)\xi, \xi \rangle \geq \delta > 0$ so that $\langle \eta, \xi \rangle > 0$, and thus $\eta \neq 0$. (This demonstrates why it is convenient to have strict inequality in Equation (3.30).)

The more recent [BdlHV08] has a rather geometric approach, which considers the closure $\mathcal{C}$ of the convex hull of the orbit $\pi(G)\xi$. One then shows that the vector $\eta_0 \in \mathcal{C}$ of minimal norm is a non-zero invariant vector, since the action of G on the Hilbert space leaves the convex hull invariant and preserves the norm. This requires the non-trivial theorem that the convex hull of a compact set in a Hilbert space is relatively compact.

We instead take an iterative approach with a constructive flavour that better gives a sense for how a compact group acts on a space, and how to find almost invariant vectors

that are close to each other. Another advantage is that we avoid the use of some perhaps unnecessarily powerful tools, namely the Haar measure of a compact group and a theorem about convex hulls in (possibly infinite-dimensional) Hilbert spaces.

The idea of the proof we present is to take an almost invariant vector as the starting point, and iteratively construct a *convergent* sequence of (ε, G)-invariant unit vectors for $\varepsilon \to 0$. This idea is similar to the Bruhat–Tits Fixed-point Theorem [dlHV89, Proposition 3.9], but we consider a different orbit at each point in the sequence. For any unitary representation with almost invariant vectors it is possible to construct a sequence, except that it will not in general be convergent: in an infinite-dimensional Hilbert space the unit ball is not compact, so an arbitrary sequence of unit vectors will not contain a convergent subsequence.

We now prove the following key lemma, which enables us to find for an almost invariant vector v another almost invariant vector $\hat{m}$ so that $\hat{m}$ is a better approximation of being invariant, with $\hat{m}$ close to v.

Lemma 3.53. Let $c = \frac{1}{\sqrt{2}}$ and $r = \sqrt{\frac{6}{7}}$. Let v be an (ε, G)-invariant vector, with $\varepsilon \leq c$. Then there exists a $(r\varepsilon, G)$-invariant unit vector $\hat{m}$ such that $\|v - \hat{m}\| < \varepsilon$.

Proof. Let $\delta = \sup\{\|\pi(g)v - v\| : g \in G\} < \varepsilon$. By the strong continuity of the unitary representation, the continuous map

$$G \to \mathbb{R} : g \mapsto \|\pi(g)v - v\|$$

must attain its maximum on compact G, so let $g_0 \in G$ be such that $v' = \pi(g_0)v$ and $\|v' - v\| = \delta$.

Let $m = (v + v')/2$. Our hope is that m will be more "central" and thus closer to the entire orbit $\pi(G)v$ than v is. The geometry behind this result is illustrated by the *vesica piscis* in Figure 3.1. The two black dots correspond to v and v'. The entire orbit lies within δ of each of v and v', and hence must be contained in the shaded region, all of which lies at a distance significantly smaller than δ to m.

By Lemma 3.11, for each $p \in \pi(G)v$ we have

$$\begin{aligned} \|p - m\|^2 &= \frac{1}{2}\|p - v\|^2 + \frac{1}{2}\|p - v'\|^2 - \frac{1}{4}\|v - v'\|^2 \\ &\leq \frac{1}{2}\delta^2 + \frac{1}{2}\delta^2 - \frac{1}{4}\delta^2 \\ &= \frac{3}{4}\delta^2. \end{aligned} \tag{3.54}$$

Now there are two outstanding issues with m: its orbit is different from the orbit $\pi(G)v$, and it is not a unit vector.

For the first issue, all the points on the orbit $\pi(G)m$ have the form

$$\pi(g)m = \frac{\pi(g)v + \pi(g)v'}{2}$$

and in particular are the midpoints of vectors a, b on the orbit $\pi(G)v$. Now by Lemma 3.11 we see that m will be at least as close to the midpoint of a and b as it is to more distant of a and b:

$$\|m - \pi(g)m\|^2 \leq \frac{1}{2}\|m - a\|^2 + \frac{1}{2}\|m - b\|^2 \leq \frac{3}{4}\delta^2 \tag{3.55}$$

by (3.54).

For the second issue, as we assumed $\varepsilon \leq 1/\sqrt{2}$, and $\delta < \varepsilon$, we can use Lemma 3.11 again, this time with $p = 0$, to get a lower bound on the norm of m:

$$\|m\|^2 = \frac{1}{2} + \frac{1}{2} - \frac{1}{4}\delta^2 > 1 - \frac{1}{8} \tag{3.56}$$

so that

$$\|m\| > \frac{\sqrt{7}}{2\sqrt{2}}. \tag{3.57}$$

Combining (3.55) and (3.57) we have for all $g \in G$ that

$$\|\pi(g)m - m\| / \|m\| \leq \frac{\sqrt{3}}{2} \cdot \frac{2\sqrt{2}}{\sqrt{7}} = \frac{\sqrt{6}}{\sqrt{7}}.$$

As $g \in G$ was arbitrary, after normalising $\hat{m} = m/\|m\|$ we have that $\hat{m}$ is $(r\varepsilon, G)$-invariant.

It is easy to verify that $\|v - \hat{m}\| \leq \delta$. By the triangle inequality, since $\|v - m\| = \frac{\delta}{2}$, it suffices to show that $\|m - \hat{m}\| \leq \frac{\delta}{2}$. From Equation (3.56), we see that

$$\|m - \hat{m}\| = 1 - \sqrt{1 - \left(\frac{\delta}{2}\right)^2}.$$

Now

$$\sqrt{1 - \frac{\delta}{2}} \leq \sqrt{1 + \frac{\delta}{2}}$$

so that

$$1 - \frac{\delta}{2} \leq \sqrt{1 - \frac{\delta}{2}} \cdot \sqrt{1 + \frac{\delta}{2}} = \sqrt{1 - \left(\frac{\delta}{2}\right)^2} = 1 - \|m - \hat{m}\|$$

which completes the proof. □

Remark 3.58. Compactness was not very essential to the above proof. It enters in the following proof mainly in the fact that by definition of almost invariant vectors, we must have an (ε, G)-invariant vector, and then in a continuity argument. Weaking the choice of r in the above lemma, we could have instead taken $v'\pi(G)v$ such that $\|v - v'\|$ is very close to the supremum δ, say, at least $\frac{99}{100}\delta$.

Proof of Proposition 3.52. As π has almost invariant vectors, we can find a $(1/\sqrt{2}, G)$-invariant unit vector v_0. Repeatedly applying Lemma 3.53, we can construct a sequence $v_0, v_1, v_2, \ldots$ of unit vectors, such that each v_n is $(r^n/\sqrt{2}, G)$-invariant, and

$$\|v_{n+1} - v_n\| \leq r^n/\sqrt{2}.$$

For $m \leq n$, the triangle inequality gives

$$\|v_n - v_m\| \leq \sum_{i=m}^{n-1} r^i/\sqrt{2} < \sum_{i=m}^{\infty} r^i/\sqrt{2} = \frac{r^m}{\sqrt{2}(1-r)} \to 0$$

as $m \to \infty$. Hence (v_i) is a Cauchy sequence, which converges to a limit $v \in \mathcal{H}$ by the completeness of the Hilbert space $\mathcal{H}$. Since each v_i is a unit vector, v is a unit vector.

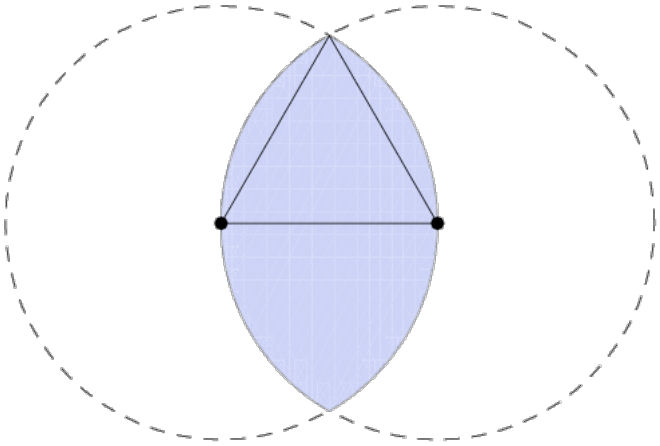

FIGURE 3.1. The geometry of the orbit of an almost invariant vector.

Source: Wolfram MathWorld

To finish the proof, we will use the following technical result.

Lemma 3.59. Let X and Y be topological spaces with X compact. Suppose that $f : X \times Y \to \mathbb{R}$ is continuous. Then

$$g(y) = \sup_{x \in X} f(x, y)$$

defines a continuous function $g : Y \to \mathbb{R}$.

Proof. This is proved in Appendix B. □

Since the evaluation map $G \times \mathcal{H} \to \mathcal{H}$ is continuous (Lemma 3.23), the map

$$G \times \mathcal{H} \to \mathbb{R} : (g, \xi) \mapsto \|\pi(g)\xi - \xi\|$$

is continuous. Thus

$$d : \mathcal{H} \to \mathbb{R} : \xi \mapsto \sup_{g \in G} \|\pi(g)\xi - \xi\|$$

is continuous, by Lemma 3.59.

As d is continuous and $\lim_{n\to\infty} d(v_n) = 0$, we have

$$d(v) = 0$$

so v is a non-zero invariant vector as required. □

Remark 3.60. A useful mental model for the preceding proof is as follows. Take $G = \mathbb{Z}/3\mathbb{Z}$ acting on $\mathbb{R}^3$ by rotations of multiples of $2\pi/3$ around the z-axis. Let $\xi_0 = (\frac{3}{5}, 0, \frac{4}{5})$ be an almost invariant vector. Pick either of the 2 other points on its G-orbit to be ξ_0'. Then the 'projection' ξ_1 of the midpoint $(\xi_0 + \xi_0')/2$ back onto the sphere will be closer to the north pole than ξ_0 is. Continuing this process, we construct a sequence of vectors which converges to the north pole, which is an invariant vector.

Remark 3.61. The fact that a compact group G is necessarily Kazhdan highlights the importance of the topological structures, namely, the topology of the group G and the strong continuity of the representation $\pi : G \to \mathcal{U}(\mathcal{H})$. Any group can be made a compact group, hence a Kazhdan group, by endowing it with the trivial topology (although this does violate our requirement from Definition 3.13 that the topology be Hausdorff). However, one would very rarely have cause to consider a group with such a topology. Rather than going by Proposition 3.52, Property (T) comes immediately as the strong continuity in fact requires that any unitary representation of a group with the trivial topology be trivial.

3.4. Kazhdan Sets and Generation

In this section we introduce Kazhdan sets and Kazhdan pairs. These allow us to give in Corollary 3.68 an equivalent definition of Property (T) that is quantitative in a way that is independent of any particular unitary representation. This will be essential to the Margulis construction in Chapter 4. Property (T) and Kazhdan sets are related to generating sets, as we will see in Proposition 3.67.

Definition 3.62 (Kazhdan Sets, [BdlHV08, Definition 1.1.3])**.** Let G be a topological group. A subset K of G is a *Kazhdan set* if there exists $\varepsilon > 0$ with the following property: every unitary representation $\pi : G \to \mathcal{U}(\mathcal{H})$ which has an (ε, K)-invariant vector also has a non-zero invariant vector. In this case, $\varepsilon > 0$ is called a *Kazhdan constant* for G and K, and (ε, K) is called a *Kazhdan pair* for G.

Remark 3.63. We note a few immediate consequences of the above definition. Suppose that $\varepsilon' < \varepsilon$ and $K' \supset K$. Any (ε', K')-invariant vector is clearly (ε, K)-invariant. Thus if (ε, K) is a Kazhdan pair, then (ε', K') is an Kazhdan pair. Hence a group G will have a Kazhdan pair if and only if (ε, G) is a Kazhdan pair for some $\varepsilon > 0$.

The following proposition is what [BdlHV08] calls 'the first spectacular application of Property (T)'. We adapt their treatment.

Proposition 3.64 ([BdlHV08, Theorem 1.3.1])**.** Let G be locally compact group with Property (T). Then G is compactly generated.

Proof. Let $\mathcal{C}$ be the set of all open and compactly generated subgroups of G. Since G is locally compact, every element $g \in G$ has a compact neighbourhood Q. Then $\langle Q \rangle$ is a compactly generated subgroup of G, which is moreover open since we can write $\langle Q \rangle = \cup_{q \in \langle Q \rangle} q \operatorname{Int}(Q)$ so that it the union of open sets. Thus

$$G = \bigcup_{H \in \mathcal{C}} H.$$

For any $H \in \mathcal{C}$, since H is open the quotient space G/H is discrete. Let $\ell^2(G/H)$ denote the Hilbert space of square-summable functions on G/H (we write ℓ^2 rather than L^2 since the space is discrete). The group G acts on the quotient space G/H, and this gives rise to the quasi-regular representation $\lambda_{G/H}$ of G on $\ell^2(G/H)$, similar to the left-regular representation of Definition 3.25. (This representation however will not be faithful, since for instance all $g \in H$ give the same action on G/H.) Define the Dirac delta function $\delta_H : G/H \to \mathbb{C}$ by

$$\delta_H(gH) = \begin{cases} 1 \text{ if } gH = H \\ 0 \text{ otherwise.} \end{cases}$$

Then δ_H is H-invariant. We let π be the direct sum (Example 3.28) of these representations:

$$\pi = \bigoplus_{i \in I} \lambda_{G/H}.$$

This representation has almost invariant vectors. Let $Q \subseteq G$ be compact. Since the $H \in \mathcal{C}$ cover G, we have

$$Q \subseteq H_1 \cup \cdots \cup H_n$$

for some $H_1, \ldots, H_n \in \mathcal{C}$. Letting K be the subgroup generated by $H_1 \cup \cdots \cup H_n$, and K_i be a compact generating set for each H_i. Then K is generated by the compact set $K_1 \cup \cdots \cup K_n$, so $K \in \mathcal{C}$. We also have $Q \subseteq K$. We can view $\delta_K : G/K \to \mathbb{C}$ as a unit vector in $\oplus_{H \in \mathcal{C}} \ell^2(G/H)$, by considering all other coordinates to be zero. Then δ_K is K-invariant, so for all $x \in Q$ we have

$$\|\pi(x)\delta_K - \delta_K\| = 0.$$

So for an arbitrary compact $Q \subseteq G$, and any $\varepsilon > 0$, we have found an (ε, Q)-invariant vector. As G has Property (T), there must then be some non-zero invariant vector

$$\xi = \bigoplus_{H \in \mathcal{C}} \xi_H \in \bigoplus_{H \in \mathcal{C}} \ell^2(G/H).$$

So now we can let $H \in \mathcal{C}$ be some compactly generated open subset of G such that $\xi_H \neq 0$. As ξ is $\pi(G)$-invariant, we must have that ξ_H is $\lambda_{G/H}(G)$-invariant. Since the action of G on G/H is transitive, ξ must be constant. As it is non-zero, this implies that G/H is finite. Let $S = \{g_1, g_2, \ldots, g_m\}$ be a set of representatives for the cosets of G/H, and H' a compact generating set for H. Then G is compactly generated by $H \cup S$. □

Corollary 3.65. Let G be a discrete group with Property (T). Then G is finitely generated.

Proof. This is immediate, as a subset of a discrete group is compact if and only if it is finite. □

Proposition 3.66. Let G be a topological group. If $\varepsilon_0 > 0$, $K_0 \subseteq G$ is compact and (ε_0, K_0) is a Kazhdan pair for G, then G has Property (T).

Proof. If a unitary representation π of G has almost invariant vectors, setting $\varepsilon = \varepsilon_0$, $K = K_0$ in Definition 3.29, it must have an (ε_0, K_0)-invariant vector. Since (ε_0, K_0) is a Kazhdan pair, it follows that π has a non-zero invariant vector. As π was arbitrary, G has Property (T). □

Proposition 3.66 shows that if a group has a Kazhdan pair, then it has Property (T). The converse holds.

Proposition 3.67 ([dlHV89, Proposition 1.15])**.** Let G be a locally compact group with Property (T), and let K be a compact generating set for G. Then there exists $\varepsilon > 0$ such that (ε, K) is a Kazhdan pair for G.

Corollary 3.68. A locally compact group has Property (T) if and only if it has a compact Kazhdan set.

Proof of Corollary 3.68. This follows immediately from Propositions 3.64, 3.66 and 3.67. □

We now work towards the proof that any compact generating set is a Kazhdan set, Proposition 3.67. An essential part of the proof is the following.

Proposition 3.69. Let G be a locally compact group with a compact subset K and a compact generating set Q. Then there exists an integer N such that every element of K is a word of length at most N in $Q \cup Q^{-1}$.

In order to prove this proposition, we will use a few theorems in topology. We note however that for discrete topological spaces these theorems are trivial, so that this is not a gap in our proof that the constructions in Chapter 4 are expanders (which only requires Proposition 3.67 for discrete groups).

Theorem 3.70 (Baire Category Theorem, [Mun00, Ex. 48.3])**.** Any Hausdorff locally compact space is a Baire space, that is, the countable union of closed sets with empty interior has empty interior.

Theorem 3.71 (Tychonoff, [Mun00, Theorem 37.3])**.** Let X and Y be compact topological spaces. Then $X \times Y$ is compact.

Lemma 3.72. Let Q be a compact subset of a Hausdorff topological space X. Then Q is closed.

Proof. We show that Q^c is open. Let $x \in Q^c$ be arbitrary. Since X is Hausdorff, for each $y \in Q$ we can find open sets $A_y \ni x, B_y \ni y$. Then $\{B_y \mid y \in Q\}$ forms an open cover of Q, and by compactness it has a finite subcover. Denote this subcover by $B_1, \dots, B_k$, and let $A_1, \dots, A_k$ be the corresponding disjoint open sets that contain x. Now the finite intersection $A = \cap_{i=1}^k A_i$ is an open set containing x, and it is disjoint from all the B_i, and hence disjoint from Q. That is, $A \subseteq Q^c$ is an open neighbourhood of x. Since Q was arbitrary, this shows that Q^c is open. □

We now give a proof following [BdlHV08] with all details explained.

Proof of Proposition 3.69. Recall that if H_1 and H_2 are compact subsets of a topological group G then

$$H_1 H_2 = \{h_1 h_2 \in G : h_1 \in H_1, h_2 \in H_2\}$$

is compact, since it is the image of $H_1 \times H_2$ under the continuous group multiplication map.

Let $\widetilde{Q} = Q \cup Q^{-1} \cup \{e\}$. Then $\widetilde{Q}^n = \{q_1 \cdots q_n | q_1, \dots, q_n \in \widetilde{Q}\}$ defines a sequence of compact sets. This sequence is nested, since $e \in \widetilde{Q}$ implies that $\widetilde{Q}^{n-1} \subseteq \widetilde{Q}^n$, and

$$G = \bigcup_{n=1}^{\infty} \widetilde{Q}^n$$

since Q is a generating set for G. Since G is Hausdorff by the definition of a topological group, Lemma 3.72 implies that each $\widetilde{Q}^n$ is closed. By Theorem 3.70, we know that some $\widetilde{Q}^m$ must have non-empty interior, since their countable union G clearly has non-empty interior. Let $\widetilde{Q}^m$ contain a neighbourhood U of some $x \in G$, and let $U' = x^{-1}U$. Then $U' \subseteq \widetilde{Q}^{2m}$.

Now for any $g \in G$, as Q is a generating set we have $g \in \widetilde{Q}^n$ for some n, and thus $gU' \subseteq \widetilde{Q}^{n+2m}$ so that $\widetilde{Q}^{n+2m}$ is a neighbourhood of g. Thus $(\mathrm{int}(\widetilde{Q}^n))_{n \in \mathbb{N}}$ forms an open cover of G, and in particular, an open cover of K. Thus there is some $N \in \mathbb{N}$ such that $K \subseteq \widetilde{Q}^N$, that is, every element of K is the product of at most N elements of $Q \cup Q^{-1}$. □

Remark 3.73. We note that the above proof uses the structure of G as a topological group in an essential way. The space $\mathbb{R}_0^+$ with the usual topology is locally compact and Hausdorff, but not a topological group. The nested sequence $Q_n = \{0\} \cup [\frac{1}{n}, n]$ exhausts the whole space by compact sets, yet no Q_n covers the compact set $[0, 1]$, in contrast to the $\widetilde{Q}^n$ of the above proof.

We now have the all the ingredients for the proof.

Proof or Proposition 3.67. We take the approach of [dlHV89].

We argue by proof by contradiction: suppose that for all $\varepsilon > 0$ there is a unitary representation of G which has an (ε, K)-invariant vector but no non-zero invariant vectors. In particular, let π_n be a sequence of such representations with $(\frac{1}{n^2}, K)$-invariant vectors, denoted ξ_n. By Lemma 3.33, this implies that each ξ_n is $(\frac{1}{n}, K^n)$-invariant. We now take the direct sum of representations

$$\pi = \oplus_{n \geq 1} \pi_n.$$

Then π has almost invariant vectors, since any $\frac{1}{n} < \varepsilon$ and $K^n \supseteq Q$ for sufficiently large n, by Proposition 3.69. So since G has Property (T) there must be some $\eta = (\eta_n)_{n \geq 1}$ that is $\pi(G)$-invariant. Then for some n we have non-zero $\eta_n \in \mathcal{H}_n$, but η_n is in particular $\pi_n(G)$-invariant, contradicting the choice of π_n as a representation with no non-zero invariant vectors. □

Corollary 3.74. Let G be a discrete Kazhdan group with a finite symmetric generating set S. Then S is a Kazhdan set for G. □

3.5. Lattices and Property (T)

Kazhdan introduced Property (T) in order to prove that lattices in certain Lie groups are finitely generated. The main result of this section is that if a group has Property (T) then all its lattices do as well, which then implies that they are finitely generated. The fact that lattices inherit Property (T) is essential to the construction of expanders in Chapter 4 using discrete Kazhdan groups; proofs that those discrete groups have Property (T) that do not embed them as lattices in Lie groups with Property (T) have only recently been given.

Definition 3.75 (Left-invariant measure)**.** A measure μ on a σ-algebra $\mathcal{A}$ of subsets of a group G is *left-invariant* if for all $A \in \mathcal{A}$ and $g \in G$ we have $gA \in \mathcal{A}$ and

$$\mu(gA) = \mu(A).$$

The following theorem is a deep result in analysis. The reader is referred to [BdlHV08, Section A.3] for further details.

Definition 3.76 (Haar measure)**.** A left-invariant Borel regular measure on a locally compact group is called a *Haar measure*.

Theorem 3.77 (Haar, Weil)**.** Every locally compact group admits a Haar measure. Moreover, the Haar measure is unique up to a positive scalar.

Definition 3.78 (Lattice)**.** Let G be a locally compact group with Haar measure μ. A *lattice* in G is a discrete subgroup $\Gamma < G$ with finite covolume, that is, such that the quotient space G/Γ admits a finite volume G-invariant measure. By abuse of notation, we write $\mu(G/\Gamma) < \infty$.

Example 3.79. The discrete subgroup $\mathbb{Z}^n$ is a lattice in the additive group $\mathbb{R}^n$. Since $\mathbb{Z}^n \lhd \mathbb{R}^n$, the quotient space is actually a group, namely $\mathbb{T}^n$, the n-dimensional torus constructed by identifying opposite faces of the n-dimensional unit cube. We can give that fundamental domain $[0,1]^n$ the Lebesgue measure (which is a Haar measure on $\mathbb{R}^n$), and this induces an invariant measure on the quotient space. It can be proved that all lattices in $\mathbb{R}^n$ are isomorphic as groups to $\mathbb{Z}^n$.

Example 3.80. The subgroup $\mathbb{Z}^n \times \{0\}$ is *not* a lattice in $\mathbb{R}^{n+1}$, although it is a discrete subgroup. As it is a normal subgroup, the fact that is does not have finite covolume actually follows from that fact that the quotient $\mathbb{T}^n \times \mathbb{R}$ is a non-compact topological group and thus does not have finite Haar measure.

Example 3.81. For $n \geq 2$, the group $\mathrm{SL}(n,\mathbb{Z})$ is a lattice in $\mathrm{SL}(n,\mathbb{R})$. We will not prove this classical result in detail in this essay; the reader is referred to [BM00, pp.144-146] for a proof. This is the only non-elementary result required for the construction of expanders in Chapter 4 for which we do not give a proof.

The idea of the proof that $\mathrm{SL}(n,\mathbb{Z})$ is a lattice in $\mathrm{SL}(n,\mathbb{R})$ is as follows. It is clear that $\mathrm{SL}(n,\mathbb{Z})$ is a discrete subgroup of $\mathrm{SL}(n,\mathbb{R})$. By the celebrated Iwasawa decomposition, every element g of $\mathrm{SL}(n,\mathbb{R})$ can be expressed uniquely in the form $g = kan$ where $k \in \mathrm{SO}(n)$, a is a diagonal matrix and n is an upper-triangular matrix with 1s on the diagonal. It can be shown that every coset in the quotient space $\mathrm{SL}(n,\mathbb{R})/\mathrm{SL}(n,\mathbb{Z})$ has a representative in a particular *Siegel set* $\mathcal{S}_{t,C} \subseteq \mathrm{SL}(n,\mathbb{R})$, for $t = \frac{2}{\sqrt{3}}$ and $C = \frac{1}{2}$, of elements whose Iwasawa decomposition has a particular form. In the Siegel set $\mathcal{S}_{t,C}$, the diagonal entries of a are positive and the ratio between consecutive entries is bounded by t, and in k the diagonals entries are 1 while the off-diagonal entries have absolute value bounded by C. It can be shown that any Siegel set has finite measure, and thus the quotient space has a left-invariant finite measure.

We recall the definition of the (external) semidirect product.

Definition 3.82. Let N and H be groups with a homomorphism $\phi : H \to \mathrm{Aut}(N)$. The *semidirect product* of N and H with respect to ϕ, denoted by $N \rtimes_\phi H$, is the set $N \times H$ together with the group multiplication

$$(n_1, h_1)(n_2, h_2) = (n_1 \phi(h_1)(n_2), h_1 h_2).$$

Example 3.83. $\mathrm{SL}(2,\mathbb{R})$ acts on $\mathbb{R}^2$ by matrix multiplication, so we can form the semidirect product $\mathbb{R}^2 \rtimes \mathrm{SL}(2,\mathbb{R})$. It is a general result that if $\Lambda < N$, $\Gamma < H$ are lattices and a continuous action of H on N restricts to an action of Γ on Λ, then $\Lambda \rtimes \Gamma$ is a lattice in $N \rtimes H$. So from Examples 3.79 and 3.81, we have that $\mathbb{Z}^2 \rtimes \mathrm{SL}(2,\mathbb{Z})$ is a lattice in $\mathbb{R}^2 \rtimes \mathrm{SL}(2,\mathbb{R})$.

Remark 3.84. It is very important to note that in general G/Γ is only a quotient space and not a quotient group, since Γ is not necessarily a normal subgroup of G. We still have a natural action of G on the quotient space, but there is no sensible well-defined group multiplication on the quotient. While it is the case that $\mathbb{Z}^n \lhd \mathbb{R}^n$ and $\mathbb{R}^n/\mathbb{Z}^n \cong \mathbb{T}^n$, $\mathrm{SL}(n,\mathbb{Z})$ is not normal in $\mathrm{SL}(n,\mathbb{R})$. For example,

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & \frac{1}{4} \\ 0 & 1 \end{pmatrix} \notin \mathrm{SL}(2,\mathbb{Z}).$$

Remark 3.85. Kazhdan introduced Property (T) to show that certain lattices are finitely generated. Since lattices in $\mathbb{R}^n$ are isomorphic to $\mathbb{Z}^n$, it is clear that they are finitely generated. However, for $G = \mathrm{SL}(3,\mathbb{R})$ for instance, it is not obvious that all lattices are finitely generated.

Theorem 3.86. Let Γ be a lattice in a locally compact group G. Suppose that G has Property (T). Then Γ has Property (T) also.

Remark 3.87. The converse of Theorem 3.86 holds also, namely, if a lattice in a group has Property (T) then that group has Property (T). For proofs of both directions, we refer the reader to [BdlHV08, pp.60-62].

Example 3.88. The group $\mathrm{SL}(2,\mathbb{R})$ does not have Property (T). One can use the Ping-Pong Lemma (Klein's criterion) to show that

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$$

generate a subgroup F_2 of $\mathrm{SL}(2,\mathbb{R})$ which is free of rank 2. (One uses the natural action of $\mathrm{SL}(2,\mathbb{R})$ on $\mathbb{R}^2$, with the two subsets of $\mathbb{R}^2$ for the lemma as $\{(x,y) \in \mathbb{R}^2 : |x| > |y|\}$ and $\{(x,y) \in \mathbb{R}^2 : |x| < |y|\}$. See [dlH00, p.26] for further details.) Since F_2 is finite index in $\mathrm{SL}(2,\mathbb{Z})$, it follows from Example 3.81 that F_2 is a lattice in $\mathrm{SL}(2,\mathbb{R})$.

Remark 3.89. The fact that Γ has finite covolume in G was used essentially in the proof of Theorem 3.86. Note that since we can easily embed $\mathrm{SL}(2,\mathbb{R})$ in $\mathrm{SL}(3,\mathbb{R})$, by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for example, the latter also contains a discrete subgroup that is free on 2 generators. However, this subgroup is not a lattice because it does not have finite covolume. In fact, since we will see in Section 3.6 that $\mathrm{SL}(3,\mathbb{R})$ has Property (T), we know by Theorem 3.86 that it is impossible to find a discrete subgroup that is free and has finite covolume.

Another example, which may be familiar from the theory of amenability, is that there is a subgroup that is free of rank 2 in $\mathrm{O}(3)$. However, this subgroup is not a lattice, and is actually not even a discrete subgroup (since $\mathrm{O}(3)$ is compact, its only discrete subgroups are finite).

Corollary 3.90. Any lattice Γ in a Kazhdan group G is finitely generated.

Proof. By definition of a lattice (Definition 3.78), Γ is discrete. Since G has Property (T), so does Γ (Theorem 3.86). Now by Corollary 3.65, Γ is finitely generated. □

Remark 3.91. Topological spaces which are not second-countable are rather uncommon, and certainly the Lie groups that Kazhdan considered are second-countable.

3.6. Some Non-compact Kazhdan Groups

In this section we prove that $\mathrm{SL}(3,\mathbb{R})$ has Property (T). To prove this we use relative property (T), defined below, for the pair $(\mathbb{R}^2 \rtimes \mathrm{SL}(2,\mathbb{R}), \mathbb{R}^2)$; for this result we only sketch the proof. Corollaries to these are that $\mathrm{SL}(2,\mathbb{Z})$ has Property (T) and $(\mathbb{Z}^2 \rtimes \mathrm{SL}(2,\mathbb{Z}), \mathbb{Z}^2)$ has relative property (T), both of which are used in Chapter 4 to construct expanders.

Definition 3.92 (Relative Property (T))**.** Let G be a locally compact group and H a closed subgroup of G. We say that the pair (G, H) has relative property (T) if whenever a (strongly continuous) unitary representation of G has almost invariant vectors, then it has a non-zero vector which is fixed by H.

Remark 3.93. A group G has Property (T) if and only if the pair (G, G) has relative property (T). Thus relative property (T) is a generalisation of Property (T).

Remark 3.94. This property was introduced by Margulis who used it for the construction of expanders and for the resolution of the Ruziewicz problem for $n \geq 3$, but it is implicit in Kazhdan's original paper [Kaz67]. It is not to be confused with another variant on Property (T) that has a relative definition: Property (τ). See Remark 4.7 for a discussion of Property (τ).

Proposition 3.95. Let $H = \mathbb{R}^2 \rtimes \mathrm{SL}(2, \mathbb{R})$, that is, the semidirect product with the standard action of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{R}^2$. Then $(H, \mathbb{R}^2)$ has relative property (T).

Proofs of this proposition are very technical. For instance, in [Lub94], the following are used: Mackey's theorem, induced representations, the fact that nilpotent groups are amenable, and Hulanicki's characterisation of amenability. The reader is encouraged to refer to [BdlHV08, Corollary 1.4.13] for a more recent and accessible proof. It hinges on the following theorem, due to Shalom [Sha99].

Theorem 3.96 ([BdlHV08, Theorem 1.4.5])**.** Let G be a locally compact group and N an abelian closed normal subgroup. Assume that the Dirac measure at the unit character 1_N of N is the unique mean on the Borel subsets of $\hat{N}$ which is invariant under the action of G on $\hat{N}$ dual to the conjugation action. Then the pair (G, N) has relative property (T).

Proof sketch for Proposition 3.95. Let m be an $\mathrm{SL}(2, \mathbb{R})$-invariant mean on $\mathbb{R}^2$, and let

$$\Omega = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \setminus \{(0,0)\} \;\middle|\; |y| \geq |x| \right\}.$$

Let

$$g_n = \begin{pmatrix} 1 & 3n \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}(2, \mathbb{R}).$$

Then the sets $g_n\Omega$ "fan out" and are disjoint. Since $m(g_n\Omega) = m(\Omega)$, and $m(\cup_n g_n\Omega) \leq m(\mathbb{R}^2 \setminus \{(0,0)\})$, if follows that $m(\Omega) = 0$. Then the mean m must the Dirac measure at $(0,0)$. The result then follows by Theorem 3.96. □

We now prove a series of lemmas which allows us to work towards giving a proof that $\mathrm{SL}(3, \mathbb{R})$ has Property (T), assuming $(\mathbb{R}^2 \rtimes \mathrm{SL}(2, \mathbb{R}), \mathbb{R}^2)$ has relative property (T). Some details will only be sketched, owing to space restrictions.

Lemma 3.97 (Mautner)**.** Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be a unitary representation. Let $x \in G$ and suppose that there exists a sequence $(y_i)_i$ in G such that $\lim_{i\to\infty} y_i x y_i^{-1} = e$, the identity of G. If $\xi \in \mathcal{H}$ is fixed by y_i for all i, then ξ is fixed by x.

Proof. As ξ is fixed by each $\pi(y_i)$ we can write

$$\begin{aligned}\|\pi(x)\xi - \xi\| &= \left\|\pi(x)\pi(y_i^{-1})\xi - \pi(y_i^{-1})\xi\right\| \\ &= \left\|\pi(y_i^{-1})\pi(x)\pi(y_i^{-1})\xi - \xi\right\| \\ &= \left\|\pi(y_i^{-1}xy_i^{-1})\xi - \xi\right\|.\end{aligned}$$

Recall that the strong continuity of the unitary representation π gives that the map $G \to \mathcal{H} : g \mapsto \pi(g)\xi$ is continuous. Since $\lim_{i\to\infty} y_i x y_i^{-1} = e$ and $\pi(e)\xi = \xi$, this implies that

$$\lim_{i\to\infty} \left\|\pi(y_i^1 x y_i)\xi - \xi\right\| = 0$$

and thus $\|\pi(x)\xi - \xi\| = 0$, that is, $\pi(x)\xi = \xi$. □

Lemma 3.98. Let $G = \mathrm{SL}(2,\mathbb{R})$. Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be a unitary representation, and let

$$N = \left\{ \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \middle| \, t \in \mathbb{R} \right\}.$$

If $\xi \in \mathcal{H}$ is $\pi(N)$-invariant, then it is $\pi(G)$-invariant.

Proof. We proceed to get invariance under larger and larger classes of elements of $\mathrm{SL}(2,\mathbb{R})$. Let

$$A = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \middle| \, \lambda \in \mathbb{R} \right\}.$$

We will first show that ξ is $\pi(A)$-invariant. Define a function

$$\phi : G \to \mathbb{C} : x \mapsto \langle \pi(x)\xi, \xi \rangle.$$

This function is continuous, since the representation is strongly continuous. Let

$$a = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$$

and define the sequence

$$g_n = \begin{pmatrix} 0 & -n \\ n^{-1} & 0 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R}).$$

Now we have

$$\begin{pmatrix} 1 & n\lambda \\ 0 & 1 \end{pmatrix} g_n \begin{pmatrix} 1 & n\lambda^{-1} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ n^{-1} & \lambda^{-1} \end{pmatrix}.$$

Then since ξ is $\pi(N)$-invariant, we have

$$\phi(g_n) = \phi\left(\begin{pmatrix} \lambda & 0 \\ n^{-1} & \lambda^{-1} \end{pmatrix} \right)$$

Now since ϕ is continuous, taking the limit as $n \to \infty$ gives

$$\lim_{n\to\infty} \phi(g_n) = \phi(a).$$

Thus $\phi(a)$ does not depend on a, so since we can take a to be the identity, it follows that

$$\phi(a) = \phi(e) = \|\xi\|^2.$$

But $\phi(a) = \langle \pi(a)\xi, \xi \rangle$, and so since $\pi(a)$ is an isometry it follows that $\pi(a)\xi = \xi$.

Let

$$N^- = \left\{ \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \middle| \, t \in \mathbb{R} \right\}.$$

We now show that ξ is $\pi(N^-)$ invariant. This follows immediately from Mautner's Lemma 3.97, since

$$\begin{pmatrix} n & 0 \\ 0 & n^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \begin{pmatrix} n^{-1} & 0 \\ 0 & n \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ xn^{-2} & 1 \end{pmatrix}$$

and ξ is $\pi(A)$-invariant.

Let S denote the subgroup of $\mathrm{SL}(2,\mathbb{R})$ generated by N, A and N^-. We now show that $S = \mathrm{SL}(2,\mathbb{R})$. Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$ with $d \neq 0$. We can write

$$\begin{pmatrix} 1 & bd^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d^{-1} & 0 \\ 0 & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ cd^{-1} & 1 \end{pmatrix} = \begin{pmatrix} d^{-1}(1+bc) & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

since $ad - bc = 1$, so $g \in S$. As

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

if $d = 0$ then $bc \neq 0$, so

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c & d \\ -a & -b \end{pmatrix} \in S$$

by the previous conclusion. □

Lemma 3.99. Let $G = \mathrm{SL}(3,\mathbb{R})$. Let $\pi : G \to \mathcal{U}(\mathcal{H})$ be a unitary representation, and let

$$J = \left\{ \begin{pmatrix} 1 & 0 & s \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \middle| \, s,t \in \mathbb{R} \right\}.$$

If $\xi \in \mathcal{H}$ is $\pi(J)$-invariant, then it is $\pi(G)$-invariant.

Proof. Let

$$E_1 = \left\{ \begin{pmatrix} a & 0 & b \\ 0 & 1 & 0 \\ c & 0 & d \end{pmatrix} \in G \middle| \, a,b,c,d \in \mathbb{R} \right\}$$

and

$$E_2 = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix} \in G \middle| \, a,b,c,d \in \mathbb{R} \right\}.$$

Then E_1 and E_2 generate a dense subgroup of G. Also, letting $N_i = E_i \cap J$ for $i = 1,2$, we get subgroups

$$N_1 = \left\{ \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in G \middle| \, t \in \mathbb{R} \right\}$$

and

$$N_2 = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \in G \,\middle|\, t \in \mathbb{R} \right\}$$

which are both isomorphic to the N of Lemma 3.98. Indeed, from Lemma 3.98 we see that since ξ is fixed by J it is fixed by N_i and hence by E_i. Since E_1, E_2 together generate a dense subgroup of G and π is strongly continuous, it follows that ξ is $\pi(G)$-invariant. □

Theorem 3.100. The group $\mathrm{SL}(3,\mathbb{R})$ has Property (T).

Proof. Suppose that $\pi : \mathrm{SL}(3,\mathbb{R}) \to \mathcal{U}(\mathcal{H})$ is a unitary representation which has almost invariant vectors. Let

$$H = \left\{ \begin{pmatrix} a & b & r \\ c & d & s \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{SL}(3,\mathbb{R}) \right\} \cong J \rtimes \mathrm{SL}(2,\mathbb{R}) \cong \mathbb{R}^2 \rtimes \mathrm{SL}(2,\mathbb{R})$$

If we restrict π to H it still will have almost invariant vectors. So by relative property (T) for $(\mathbb{R}^2 \rtimes \mathrm{SL}(2,\mathbb{R}), \mathbb{R}^2)$, Proposition 3.95, the restriction has a non-zero vector ξ invariant under $\mathbb{R}^2 \cong J$. Now by Lemma 3.99, ξ is invariant under G. Thus $\mathrm{SL}(3,\mathbb{R})$ has Property (T). □

Remark 3.101. The above proof actually works for $\mathrm{SL}(3,\mathbb{K})$ where $\mathbb{K}$ is any non-discrete locally compact topological field.

Corollary 3.102. The group $\mathrm{SL}(3,\mathbb{Z})$ has Property (T).

Proof. This follows immediately from Theorems 3.100 and 3.86, in light of the fact that $\mathrm{SL}(n,\mathbb{Z})$ is a lattice in $\mathrm{SL}(n,\mathbb{R})$ (Example 3.81). □

Similarly, one can show the following.

Corollary 3.103. The pair $(\mathbb{Z}^2 \rtimes \mathrm{SL}(2,\mathbb{Z}), \mathbb{Z}^2)$ has relative property (T).

Chapter 4

Constructions of Expanders

In this chapter we present constructions of expanders using Property (T), originally due to Margulis [Mar75]. Construction 4.10 is a simple construction using the fact that $\mathrm{SL}(3,\mathbb{Z})$ has Property (T). Construction 4.12 is the original construction of Margulis, and is slightly more complicated for the fact that it uses relative property (T) of the pair $(\mathbb{Z}^2 \rtimes \mathrm{SL}(2,\mathbb{Z}), \mathbb{Z}^2)$.

4.1. Kazhdan Expanders

Property (T) can be understood as a pushing property: if a unitary representation of a Kazhdan group does not have non-zero invariant vectors then it does not have almost invariant vectors, that is, there is some ε such that all unit vectors are moved by at least ε by some element of a Kazhdan set. It is an essential result that there exists a universal ε that suffices for *any* unitary representation of the group, as studied in Section 3.4.

The construction of expanders using Property (T) boils down to the following result.

Proposition 4.1 ([Lub94, Proposition 3.3.1])**.** Let Γ be a discrete Kazhdan group. Suppose that (N_i) be a family of finite index normal subgroups of Γ such that the indices $|\Gamma/N_i| \to \infty$ as $i \to \infty$. Suppose furthermore that $S \subset \Gamma$ is a finite symmetric (that is, $S^{-1} = S$) generating set for Γ, and let S_i be the image of S in Γ/N_i (under the quotient map). Then the family $\mathrm{Cay}(\Gamma/N_i, S_i)$ is a family of (n_i, k, c)-expanders for some $c > 0$, $k = |S|$ and $n_i = |\Gamma/N_i|$.

Proof. By Corollary 3.74, there exists $\varepsilon > 0$ such that any unitary representation of Γ with an (ε, S)-invariant vector has a non-zero invariant vector. When expressed in the contrapositive, this means that if a unitary representation $\pi : \Gamma \to \mathcal{U}(\mathcal{H})$ does *not* have a non-zero invariant vector, then no unit vector is (ε, S)-invariant, and thus each $\xi \in \mathcal{H}$ satisfies

$$\sup\{\|\pi(s)\xi - \xi\| : s \in S\} \geq \varepsilon \|\xi\| .$$

As S is finite, for all $\xi \in \mathcal{H}$ there exists $s \in S$ such that

$$\|\pi(s)\xi - \xi\| \geq \varepsilon \|\xi\| .$$

Fix a particular N_i and S_i, and let $V_i = \Gamma/N_i$. Consider the representation of Γ on $\mathcal{H} = L^2(V_i)$ defined by

$$(g \cdot f)(x) = f(xg)$$

for all $f \in L^2(V_i)$, $x \in V_i$. (This is the pull back onto Γ of the right-regular representation of V_i.) Since V_i is discrete, if a function $f \in \mathcal{H}$ is invariant, then

$$(g \cdot f)(N_i) = f(N_i g)$$

for all $g \in \Gamma$, where $e \in V_i$ is the identity. The action of Γ on V_i by right-multiplication is transitive, so we can make the argument $N_i g$ any element of V_i. Thus f is constant.

So we consider the subspace

$$\mathcal{H}_0 = \left\{ f : V_i \to \mathbb{C} \;\middle|\; \sum_{x \in V_i} f(x) = 0 \right\}.$$

The only constant function $f \in \mathcal{H}_0$ is zero. Thus the unitary representation $\pi : \Gamma \to \mathcal{U}(\mathcal{H}_0)$ given by the right action of Γ on V_i does *not* have non-zero invariant vectors.

Now let $V_i = A \sqcup B$ where $|A| \leq |B|$, and write $a = |A|$ and $b = |B|$. The 'characteristic function' for A in $\mathcal{H}_0$ is

$$f_A(x) = \begin{cases} b & \text{if } x \in A \\ -a & \text{if } x \in B. \end{cases}$$

By the discussion above, since (ε, S) is a Kazhdan pair for Γ, there is some $s \in S$ such that

$$\|s \cdot f_A - f_A\| \geq \varepsilon \|f_A\|. \tag{4.2}$$

We can easily evaluate both sides of this inequality. Since

$$(s \cdot f_A)(x) = \begin{cases} b & \text{if } xs \in A \\ -a & \text{if } xs \in B. \end{cases}$$

we see that

$$(s \cdot f_A - f_A)(x) = \begin{cases} a+b & \text{if } x \in B \text{ and } xs \in A \\ -a-b & \text{if } x \in A \text{ and } xs \in B \\ 0 & \text{otherwise .} \end{cases}$$

Let $E_s(A, B)$ denote the set of edges between A and B that are due to the generator s. Then $|E_s(A,B)| = |\{x \in B \mid xs \in A\} \cup \{x \in A \mid xs \in B\}|$, or half that in the case that $s^2 = 1$. In either case, we have that

$$|E_s(A,B)| \geq \frac{1}{2} \|s \cdot f_A - f_A\|^2 / (a+b)^2. \tag{4.3}$$

One the other hand,

$$\|f_A\|^2 = |A|b^2 + |B|a^2 = ab(a+b). \tag{4.4}$$

Putting together Equations (4.2), (4.3) and (4.4) we have that

$$|E(A,B)| \geq |E_s(A,B)| \geq \frac{1}{2}\varepsilon^2 ab(a+b)/(a+b)^2 = \frac{\varepsilon^2}{2} ab/(a+b).$$

Since we assumed $|A| \leq |B|$, we have $b/(a+b) \geq \frac{1}{2}$, so this gives

$$\frac{|E(A,B)|}{\min\{|A|,|B|\}} \geq \frac{\varepsilon^2}{4}.$$

As the partition $V = A \sqcup B$ was arbitrary, we can conclude that

$$h(\mathrm{Cay}(V_i, S_i)) \geq \frac{\varepsilon^2}{4}.$$

As ε is independent of N_i, S_i (we chose $\varepsilon > 0$ so that (ε, S) is a Kazhdan pair for $\mathrm{SL}(3,\mathbb{Z})$), it follows that the Cayley graphs form a family of expanders as required. □

Remark 4.5. There are several parallels between the above proof and the proof that spectral expanders are combinatorial expanders, Proposition 2.12. In the same way that u is always a 1-eigenvector for $\hat{A}$, we saw that a constant f is always invariant. Considering the eigenvectors orthogonal to u is essentially the same as considering the invariant subspace $\mathcal{H}_0$.

Remark 4.6. At first it might seem unusual that we achieve the expansion of the above construction by considering only the edges $E_s(A, B)$ owing to one particular generator s for each choice of the partition $V = A \sqcup B$. However, since there are only k generators, some generator will always contribute at least $\frac{1}{k}$ of the edges $E(A, B)$ of a cut, so for a Cayley graph expander one can always just show that $|E_s(A, B)|$ is bounded from below by some constant multiple of $\min\{|A|, |B|\}$.

Remark 4.7. This construction does not use the full power of Property (T), as highlighted by the fact that we are only considering finite-dimensional representations. In fact, all that is required is that any representation of Γ which factors through a finite quotient of Γ and has almost invariant vectors has a non-zero invariant vector. A group Γ has *Property* (τ) *relative to* $\mathcal{L}$, a particular family of finite index normal subgroups, if any unitary representation of Γ with almost invariant vectors that factors through some quotient Γ/N with $N \in \mathcal{L}$ has a non-zero invariant vector. The group has *Property* (τ) if it has Property (τ) with respect to the family of all finite index normal subgroups. The family of Cayley graphs $\mathrm{Cay}(\Gamma/N, S)$ for $N \in \mathcal{L}$ will be a family of expanders if and only if Γ has Property (τ) with respect to $\mathcal{L}$. Property (τ) is an interesting topic of study because although it is weaker than Property (T), consequently more groups have Property (τ). The Selberg 3/16 Theorem implies that $\mathrm{SL}(2, \mathbb{Z})$ has Property (τ) with respect to the family of principle congruence subgroups, that is, the kernels of the natural maps $\mathrm{SL}(2, \mathbb{Z}) \to \mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$. The reader is referred to the book [LZ03] by Lubotzky and Zuk for a detailed treatment of Property (τ).

Remark 4.8. Another comment that we can make on the above proof is that we did not actually use the fact that the N_i are normal in Γ. In order to call the graphs Cayley graphs, this is required. However, we can generally consider 'Schreier graph' which is like a Cayley graph, except that the vertices are cosets of Γ/H.

Lemma 4.9. Define

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix}$$

Then $S = \{A, A^{-1}, B, B^{-1}\}$ generate $\mathrm{SL}(3, \mathbb{Z})$.

Construction 4.10. Let $S = \{A^{\pm 1}, B^{\pm 1}\}$, where A and B are as in Lemma 4.9. We define a family of graphs by taking the Cayley graphs $\mathrm{Cay}(\mathrm{SL}(3, \mathbb{Z}/p\mathbb{Z}), S)$ with those generators.

Proposition 4.11. Construction 4.10 is a family of expanders.

Proof. For any prime p, let $\phi : \mathrm{SL}(3, \mathbb{Z}) \to \mathrm{SL}(3, \mathbb{Z}/p\mathbb{Z})$ be the natural homomorphism defined by mapping each entry a_{ij} of $a \in \mathrm{SL}(3, \mathbb{Z})$ to $a_{ij} \pmod p$. Because the map $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ is a ring homomorphism, and the determinant is defined only in terms of addition and multiplication of matrix entries, the image of any element of $\mathrm{SL}(3, \mathbb{Z})$ will indeed be in $\mathrm{SL}(3, \mathbb{Z}/p\mathbb{Z})$. The group $\mathrm{SL}(3, \mathbb{Z}/p\mathbb{Z})$ is finite, so the kernels of those homomorphisms define

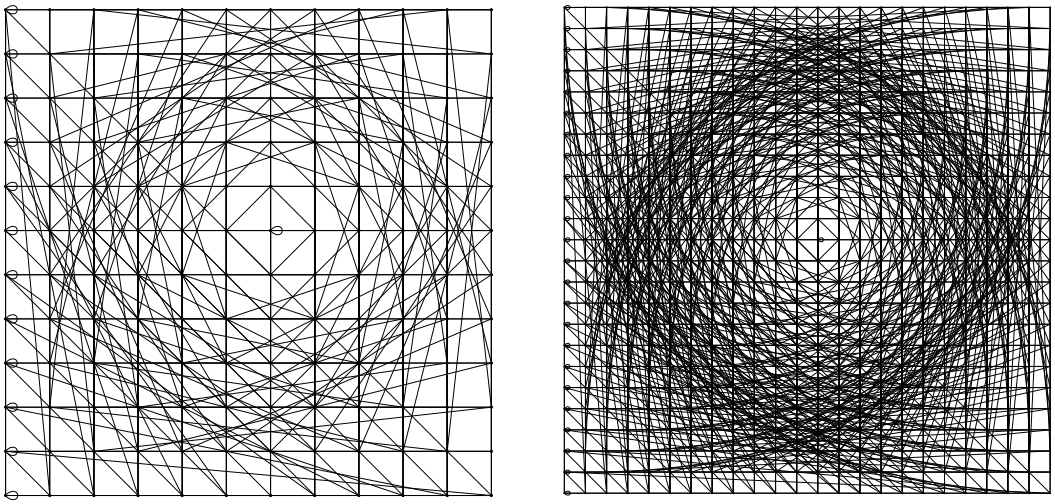

FIGURE 4.1. The Margulis construction for $m = 12$ and $m = 24$.

finite index normal subgroups N_i that can be applied to Proposition 4.1. This completes the proof. □

4.2. Margulis's Construction of Expanders

We also give the original construction of Margulis. It is illustrated for $m = 12$ and $m = 24$ in Figure 4.1.

Construction 4.12 (Margulis, [Mar75])**.** Let m be a positive integer and $V_m = (\mathbb{Z}/m\mathbb{Z})^2$. Define a graph on the set V_m by connecting every $(a, b) \in V_m$ to $\sigma_1(a, b) = (a + 1, b)$, $\sigma_2(a, b) = (a, b + 1)$, $\sigma_3(a, b) = (a, a + b)$, and $\sigma_4(a, b) = (-b, a)$.

Proposition 4.13. The graphs (V_m) in Construction 4.12 are a family of expanders.

Proof. We cannot apply Proposition 4.1 directly because $\Gamma = \mathbb{Z}^2 \rtimes \mathrm{SL}(\mathbb{Z}, 2)$ does not have Property (T), but only relative property (T) with respect to the subgroup $\mathbb{Z}^2$. The group Γ acts naturally on V_m by affine transformations:

$$\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} p & q \\ r & s \end{pmatrix}\right) \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} pa + qb + x \\ ra + sb + y \end{pmatrix}.$$

To see that this is indeed a group action, note that the identity $(0, 1)$ of $\mathbb{Z}^2 \rtimes \mathrm{SL}(\mathbb{Z}, 2)$ does indeed act as the identity, and for arbitrary $(t_1, r_1), (t_2, r_2) \in \mathbb{Z}^2 \rtimes \mathrm{SL}(2, \mathbb{Z})$ and $x \in (\mathbb{Z}/m\mathbb{Z})^2$ we have

$$\begin{aligned}
(t_1, r_1) \cdot (t_2, r_2) \cdot x &= (t_1, r_1) \cdot (r_2 x + t_2) \\
&= r_1(r_2 x + t_2) + t_1 \\
&= r_1 r_2 x + (r_1 t_2 + t_1) \\
&= (r_1 r_2, r_1 t_2 + t_1) \cdot x \\
&= ((r_1, t_1)(r_2, t_1)) \cdot x
\end{aligned}$$

by the definition of the semidirect product (where we took the standard action of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{Z}^2$.

We can express the σ_i as the actions of particular elements of the two groups used to construct the semidirect product $\mathbb{Z}^2 \rtimes \mathrm{SL}(2,\mathbb{Z})$:

$$\sigma_1 = \begin{pmatrix}1\\0\end{pmatrix} \in \mathbb{Z}^2,\ \sigma_2 = \begin{pmatrix}0\\1\end{pmatrix} \in \mathbb{Z}^2,\ \sigma_3 = \begin{pmatrix}1 & 1\\0 & 1\end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}),\ \sigma_3 = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}).$$

It is obvious that σ_1 and σ_2 generate $\mathbb{Z}^2$, and not too hard to see that σ_3 and σ_4 generate $\mathrm{SL}(2,\mathbb{Z})$ (essentially a special case of Lemma 4.9), so that together they generate the entire semidirect product $\mathbb{Z}^2 \rtimes \mathrm{SL}(2,\mathbb{Z})$.

The action of Γ on V_m gives rise to a unitary representation of Γ on $L^2(V_m)$, by

$$(g \cdot f)(x) = f(g \cdot x)$$

for all $x \in V_m$, $f \in L^2(V_m)$. Now, similarly to in the proof of Proposition 4.1, the only $\mathbb{Z}^2$-invariant vectors are the constant functions, since the action of $\mathbb{Z}^2$ on V_m is transitive. So if we quotient out by the constant functions, considering the subspace of functions that sum over V_m to zero, there is some ε such that each unit vector is moved at least ε by one of the four generators. As in Proposition 4.1, by considering characteristic functions, it follows that these graphs form a family of expanders. □

Remark 4.14. An interesting point of comparison between Construction 4.10 and Margulis's construction is that the latter does not require the generation of large primes. Moreover, given any particular vertex in these expanders, there is a polynomial time algorithm (in fact, a linear time algorithm) to compute its neighbours. In the terminology of [HLW06, p.453], it is thus a 'very explicit' construction, as opposed to a 'mildly explicit' construction.

Appendix A

Asymptotic Representations

The following definitions, closely following Knuth [Knu06, pp.107-110], are useful throughout the essay.

Definition A.1 (Big-oh notation)**.** Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions. We say that $f(n)$ is *big-oh of* $g(n)$ *as* n *tends to infinity*, and write $f = O(g)$ or $f(n) = O(g(n))$, if there are positive constants M and n_0 such that

$$|f(n)| \leq M|g(n)|$$

for all $n \geq n_0$.

Example A.2. Since $1^2 + 2^2 + \ldots + n^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$, it follows that:

$$1^2 + 2^2 + \cdots + n^2 = O(n^4); \tag{A.3}$$

$$1^2 + 2^2 + \cdots + n^2 = O(n^3); \text{ and} \tag{A.4}$$

$$1^2 + 2^2 + \cdots + n^2 = \frac{1}{3}n^3 + O(n^2). \tag{A.5}$$

Equation (A.3) is crude but true. Equation (A.4) is a stronger statement, and equation (A.5), which we can read as saying that

$$1^2 + 2^2 + \cdots + n^2 - \frac{1}{3}n^2 = O(n^2),$$

is stronger yet.

Definition A.6 (Asymptotic equivalence)**.** Let $f, g : \mathbb{N} \to \mathbb{R}$ be two functions. We say that f and g are *asymptotically equivalent*, and write $f \sim g$, if

$$\lim_{n\to\infty} f(n)/g(n) = 1.$$

Remark A.7. Asymptotic equivalence in the above Definition A.6 does indeed give an equivalence relation on the functions $\mathbb{N} \to \mathbb{R}$.

Example A.8. Stirling's approximation is that

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

A weaker approximation would be to say that, for example,

$$n! = O((n/e)^{n+1}).$$

Appendix B

A Lemma in Topology

Lemma B.1. Let X and Y be topological spaces with X compact. Suppose that $f : X \times Y \to \mathbb{R}$ is continuous. Then

$$g(y) = \sup_{x \in X} f(x, y)$$

defines a continuous function $g : Y \to \mathbb{R}$.

Proof. First note that for any fixed $y \in Y$, the map $x \mapsto f(x, y)$ is continuous. Thus the compactness of X gives that $\sup_{x \in X} f(x, y) < \infty$, and that moreover the maximum is attained by at least one $x \in X$.

Consider an arbitrary $y_0 \in Y$ and let $a = g(y_0)$. Let $\varepsilon > 0$ be arbitrary. We will show that there exists an open set $V \ni y_0$ in Y such that

$$g(V) \subseteq (a - \varepsilon, a + \varepsilon).$$

Let $x_0 \in X$ be a point such that $f(x_0, y_0) = a$ (such an x_0 must exist by compactness of X as noted above). By the continuity of $y \mapsto f(x_0, y)$, there exists an open set $V^- \ni y_0$ such that $f(x_0, y) \in (a-\varepsilon, a+\varepsilon)$ for all $y \in V^-$. From the definition of g, we have $g(y) \geq f(x_0, y)$ for all $y \in Y$. Thus $g(y) > a - \varepsilon$ for all $y \in V^-$.

The more difficult part of the proof is to find an open V^+ such that $g(y) < a + \varepsilon$ for all $y \in V^+$ (that is, to show that g is upper semi-continuous). This difficulty arises because in a neighbourhood of y_0, $g(y)$ could be determined by x different from x_0. (One such counterexample where X is not compact is $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x, y) \mapsto \sin(xy)$ near $y = 0$, which has a jump discontinuity from 0 to 1.) We need to use the compactness of X to rule out such pathological behaviour.

For each $x \in X$, by the continuity of f we can find open sets $U_x \ni x$ and $V_x \ni y_0$ such that $f(U_x \times V_x) \subseteq (f(x, y_0) - \varepsilon, f(x, y_0) + \varepsilon)$. In particular, $f(u, v) < a + \varepsilon$ for all $u \in U_x, v \in V_x$. Now we define

$$W_x = \{x \in X : f(x, y) < a + \varepsilon \text{ for all } y \in V_x\}.$$

Because U_x is open and $x \in U_x \subseteq W_x$, we have $x \in \operatorname{int}(W_x)$ for each x. Thus $\cup_{x \in X} \operatorname{int}(W_x)$ is an open cover for X. Since X is compact, there is a finite subcover

$$X = \cup_{i=1}^{n} \operatorname{int}(W_{x_i}).$$

Now put $V^+ = \cap_{i=1}^{n} V_{x_i}$ which is the finite intersection of open sets, hence open. (Note that since $y_0 \in V_x$ for all $x \in X$, $y_0 \in V^+$ also.) So for each $x \in X$ there is some x_i for which $x \in W_{x_i}$, so that we have $f(x, y) < a + \varepsilon$ for all $y \in V^+ \subseteq V_{x_i}$. Thus $g(y) < a + \varepsilon$ for all $y \in V^+$ (using compactness of X).

Finally, we can take $V = V^- \cap V^+$ as a neighbourhood of y_0 such that $g(V) \subseteq (a-\varepsilon, a+\varepsilon)$, as required. □

References

[AKS04] M. Agrawal, N. Kayal, and N. Saxena. PRIMES is in P. *Annals of mathematics*, pages 781–793, 2004.

[ALW01] N. Alon, A. Lubotzky, and A. Wigderson. Semi-direct product in groups and zig-zag product in graphs: connections and applications. In *Foundations of Computer Science, 2001. Proceedings. 42nd IEEE Symposium on*, pages 630–637. IEEE, 2001.

[Axl97] S. Axler. *Linear algebra done right*. Springer, 1997.

[BdlHV08] B. Bekka, P. de la Harpe, and A. Valette. *Kazhdan's Property (T)*, volume 11. Cambridge University Press, 2008.

[Bek03] B. Bekka. Kazhdan's Property (T) for the unitary group of a separable Hilbert space. *Geometric and Functional Analysis*, 13:509–520, 2003.

[BKV81] M. Blum, R.M. Karp, O. Vorneberger, C.H. Papadimitriou, and M. Yannakakis. The complexity of testing whether a graph is a superconcentrator. *Inform. Process. Lett.*, 13(4):164–167, 1981.

[BM00] B. Bekka and M. Mayer. *Ergodic theory and topological dynamics of group actions on homogeneous spaces*, volume 269. Cambridge University Press, 2000.

[Cha84] I. Chavel. *Eigenvalues in Riemannian geometry*, volume 115. Academic press, 1984.

[Con90] J.B. Conway. *A course in functional analysis*, volume 96. Springer, 1990.

[Die00] R. Diestel. Graph theory. *Graduate texts in Mathematics*, 173:801–811, 2000.

[dlH00] P. de la Harpe. *Topics in geometric group theory*. University of Chicago Press, 2000.

[dlHV89] P. de la Harpe and A. Valette. *La propriété (T) de Kazhdan pour les groupes localement compacts*, volume 115. Société mathémathique de France, 1989.

[HLW06] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–562, 2006.

[Kas07] M. Kassabov. Symmetric groups and expander graphs. *Inventiones mathematicae*, 170(2):327–354, 2007.

[Kas09] M. Kassabov. Constructions of expanders using group theory. Computer Science / Discrete Math Lecture, Institute for Advanced Study, 2009.

[Kaz67] D.A. Kazhdan. Connection of the dual space of a group with the structure of its close subgroups. *Functional analysis and its applications*, 1(1):63–65, 1967.

[Knu06] D.E. Knuth. *The art of computer programming*. Addison-Wesley, 2006.

[KS11] M. Krebs and A. Shaheen. *Expander Families and Cayley Graphs: A Beginner's Guide*. Oxford University Press, USA, 2011.

[Lar03] M. Larsen. Navigating the Cayley graph of $\mathrm{SL}_2(\mathbb{F}_p)$. *International Mathematics Research Notices*, 2003(27):1465–1471, 2003.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[Lub94] A. Lubotzky. *Discrete Groups, Expanding Graphs and Invariant Measures*. Birkhäuser, 1994.

[Lub12] A. Lubotzky. Personal communication, 2012.

[LW93] A. Lubotzky and B. Weiss. Groups and expanders. *DIMACS series Vol. 10*, 1993.

[LZ03] A. Lubotzky and A. Zuk. On property (τ). *To appear*, 2003.

[Mar75] G. Margulis. Explicit constructions of concentrators. *Probl. of Inform. Transm.*, 1975.

[Mar88] G. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Probl. of Inform. Transm.*, 24(1):51–60, 1988.

[Mun00] J. Munkres. *Topology*. Pearson, 2000.

[Pin73] M. Pinsker. On the complexity of a concentrator. In *7th Annual Teletraffic Conference*, pages 318/1–318/4, Stockholm, 1973.

[Rab80] M.O. Rabin. Probabilistic algorithm for testing primality. *Journal of number theory*, 12(1):128–138, 1980.

[Sar90] P. Sarnak. *Some Applications of Modular Forms*. Cambridge University Press, 1990.

[Sha99] Y. Shalom. Invariant measures for algebraic actions, Zariski dense subgroups and Kazhdan's property (T). *Transactions of the American Mathematical Society*, 351(8):3387–3412, 1999.