DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Algebraic Constructions of Ramanujan Graphs and Applications to Error Correcting Codes

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Polatajko, Daniel Brian. 2019. Algebraic Constructions of Ramanujan Graphs and Applications to Error Correcting Codes. Bachelor's thesis, Harvard College.
Citable link	https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37364612
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Algebraic Constructions of Ramanujan Graphs and Applications to Error Correcting Codes

Daniel Polatajko

Advised by Michael Mitzenmacher and Lauren Williams

March 25, 2019

danielpolatajko@college.harvard.edu

Contents

1 Abstract

This thesis gives an exposition on explicit constructions of Ramanujan graphs, paying close attention to the first such construction of Lubotzky, Phillips and Sarnak in [16], and discussing some of the interesting combinatorial, algebraic, and number theoretic techniques used to construct such graphs. This construction is particularly noteworthy mathematically because it combines far-reaching and seemingly unrelated topics in mathematics. M. Ram Murty describes Ramanujan graphs as "[fusing] diverse branches of pure mathematics, namely, number theory, representation theory, and algebraic geometry" in a survey paper on the topic [20]. The interest in such constructions is, however, not purely mathematical. Ramanujan graphs are a subset of a class of important graphs known as expanders, which colloquially have the property of being highly connected, but using relatively few edges to establish those connections. These have found myriad applications in many fields of mathematics, computer science, and beyond. We will discuss expansion properties in graphs in general, before demonstrating the optimality of one specific expansion property of Ramanujan graphs, and then move on to a discussion of the usefulness of expander graphs in some familiar topics of computer science. Particularly, we will focus on an application of expander graphs in constructing error correcting codes, and derive some results showing that error correcting codes constructed from Ramanujan graphs have desirable asymptotic behaviour.

2 Introduction to Expanders

2.1 A motivating example

Expander graphs are, among many things, a solution to a fundamental problem in the theory of networks. A good way to think of the problem is as follows: consider the construction of a new telephone network. One constraint that we might put on the construction is that we want the network to have a high degree of connectivity. The reason for this is that if a few lines in the network go down, we don't want any two users of the network to become totally disconnected from one another. However, laying lines will be expensive, so we want to achieve this high degree of connectivity using as few edges in the network as possible. Lying at the heart of this optimisation problem are expander graphs. Let us consider a simple example of the above problem, in a world where we need to connect n individuals. We could create a network that is a ring, where each individual is connected to their two immediate neighbours as viewed in the two-dimensional plane, but it is not hard to see that in this network removing any two edges will cause a disconnect between two groups of people, regardless of the size of n, which is not ideal for a communication network. In contrast, we could connect every person to every other person in the network, creating a complete graph. In this case, $n-1$ lines would have to go down in order to remove even a single person from the network, but we would have used $\frac{n}{2}(n+1)$ edges to achieve this, and thus the scalability of the network would be much weaker. Therefore, we are interested in finding graphs in the middle ground, where the number of edges that must be removed to disconnect the network scales with n, but the number of edges in the network scales

only linearly with n, which can be ensured by having graphs where each vertex has constant degree. Such graphs are described as regular or d-regular, and are exactly the types of graphs that this thesis discusses. There are a few mathematical ways to define expanders, as will be displayed later in this thesis, but naively expanders can be described as graphs that have a high degree of connectivity but are relatively sparse. We will also see that expander graphs find applications in other, less intuitive areas of computer science, such as error correcting codes and pseudo-random generators.

2.2 Mathematical definitions

We begin with some mathematical definitions of expanders and introduce the notion of a Ramanujan graph. There are three basic notions of how to measure the expansion properties of a graph, although this thesis will deal largely only with the last of these three presentations. These definitions intuitively measure how well connected a graph is, which does not entirely encapsulate our notion of expanders, but as we move on to construct expanders explicitly we will choose these graphs in such a way that the number of edges scales linearly with the number of vertices, and as such we will only be interested in optimising these connectedness properties, thus we can identify them will the notion of expansion. Note that we are in the context of graph theory where $G = (V, E)$ is an undirected graph with vertex set V and edge set E. If not otherwise stated, $n = |V|$ and $m = |E|$. We also assume the notion of adjacency matrices on the reader.

Definition 2.1 (Vertex Expansion)**.** *The vertex expansion (or vertex isoperimetric number) h_{out} of a graph G on n vertices is given as*

$$h_{out}(G) = \min_{0<|S|\leq \frac{n}{2}} \frac{|\partial_{out}(S)|}{|S|}$$

where $\partial_{out}(S)$ is defined as the set of vertices which have at least one edge adjacent to an element of S.

Next we consider an analogue of vertex expansion:

Definition 2.2 (Edge expansion)**.** *The edge expansion (or Cheeger constant) $c(G)$ of a graph G on n vertices is given as*

$$c(G) = \min_{S\subset V,|S|\leq \frac{|V|}{2}} \frac{|\partial(S)|}{|S|}$$

where $\partial(S)$ is the boundary of S, that is the set of edges (u, v) such that at least one of u, v are in S.

We can consider the Cheeger constant (which is the more commonly used of these two formulations of expansion) on some familiar graphs. Firstly, we consider the ring graphs on n vertices, where each vertex has degree 2, forming an n-cycle. The Cheeger constant can

easily be calculated as the number of edges coming out of any connected set of vertices is 2, so we need simply pick the largest such set to find the desired minimum:

$$c(G_n) = \min_{1 \leq |S| \leq \frac{n}{2}} \frac{|\partial(S)|}{|S|} = \frac{2}{\frac{n}{2}} = \frac{4}{n}$$

It is intuitively clear that rings are not very good expanders, as they are not highly connected at all, and we can see that as $n \to \infty$, $c(G_n) \to 0$. Contrast this with the complete graphs on n vertices which have

$$c(C_n) = \min_{1 \leq |S| \leq \frac{n}{2}} \frac{|\partial(S)|}{|S|}$$

$$= \min_{1 \leq |S| \leq \frac{n}{2}} \frac{|S| \cdot |V \setminus S|}{|S|} = \frac{n}{2}$$

Clearly, the Cheeger constant is much larger than the ring graphs, and in fact scales linearly with n. This, combined with our understanding that complete graphs are very highly connected, allows us to intuit that having a larger Cheeger constant corresponds to a graph being a better expander (providing we are talking about graphs where the number of edges scales linearly with n, as we will be the case for the graphs studied in the remainder of this thesis, but which is not the case for the complete graphs).

We now move on to the definition of the spectral expansion of a graph, which is the definition that will prove useful in our study of Ramanujan graphs. This definition only applies to a specific class of graphs, which we define first.

Definition 2.3 (d-regularity)**.** *A graph G is said to be d-regular if every vertex $v \in G$ has degree d.*

Definition 2.4 (Spectral Gap)**.** *Given G a d-regular graph on n vertices, let A be its $n \times n$ adjacency matrix. Let the eigenvalues of A be denoted by $\{\lambda_1, ..., \lambda_n\}$, where the set is ordered from largest to smallest. Then the spectral gap of G is defined as*

$$s(G) = \lambda_1 - \lambda_2$$

where λ_1 and λ_2 are the largest and second largest eigenvalues respectively.

We will mostly be dealing with d-regular graphs for the remainder of this thesis. It is a well known result that the largest eigenvalue of the adjacency matrix of a graph G is bounded above by the maximum degree of any vertex of G, and it can furthermore be shown in the case of a d-regular graph that the largest eigenvalue is exactly d. So henceforth when considering the spectral gap we will substitute $\lambda_1 = d$, and usually abbreviate λ_2 to simply λ.

The spectral gap gives us another tool for determining how effective an expander a graph is. In particular, the larger the spectral gap, the better expansion the graph will have. In order to see this, we can relate spectral expansion to the more intuitive notions of expansion via Cheeger's inequality.

Proposition 2.1 (Cheeger's inequality [5])**.** *For G a d-regular graph, let $s(G)$ be the spectral gap and $c(G)$ be the edge expansion. Then we have:*

$$2c(G) \geq s(G) \geq \frac{c^2(G)}{2d}$$

Equivalently, we have:

$$\frac{s(G)}{2} \leq c(G) \leq \sqrt{2d \cdot s(G)}$$

In particular, by bounding the spectral gap between two expressions in terms of the edge expansion, we can see that good spectral expanders are good edge expanders and vice versa. In all three definitions of expansion, a graph G is a better expander if its corresponding expansion number is larger.

Having defined the spectral expansion of a graph, we can now introduce the notion of a Ramanujan graph.

Definition 2.5 (Ramanujan Graph)**.** *A d-regular graph G is called a Ramanujan graph if the second largest eigenvalue of the adjacency matrix, λ, satisfies the property*

$$\lambda \leq 2\sqrt{d-1}$$

It is obvious that having a small second eigenvalue implies a large spectral gap, as we have $s(G) = d - \lambda$. We will see in the next section that, due to a theorem of Alon and Boppana, that this bound is asymptotically optimal, in that infinite families of graphs $\{G_n\}$ have the property that

$$\lim_{n\to\infty} \lambda \geq 2\sqrt{d-1}$$

This is the reason that it is so interesting, both mathematically and practically, to construct infinite families of Ramanujan graphs; they are optimal spectral expanders.

3 Preliminary Results

As mentioned in the previous section, it is possible to show that Ramanujan graphs are optimal spectral expanders. In order to see this, we will prove the Alon-Boppana theorem, but first we will need a few short lemmas:

Lemma 3.1. *The following limit holds:*

$$\lim_{n\to\infty} \left(\frac{1}{n}\right)^{\frac{1}{n}} = 1$$

Proof. We can write

$$x = \left(\frac{1}{n}\right)^{\frac{1}{n}}$$

and take logs on each side for

$$\log(x) = \frac{1}{n}\left(\log\left(\frac{1}{n}\right)\right)$$

$$\log(x) = \frac{\log(1) - \log(n)}{n}$$

Then we take limits, and we know that n grows faster than $\log(n)$ hence

$$\lim_{n\to\infty} \frac{\log(1) - \log(n)}{n} = 0$$

which gives that

$$\lim_{n\to\infty} \log(x) = 0$$

thus implying

$$\lim_{n\to\infty} x = 1$$

giving the required limit. □

Lemma 3.2. *The following limit holds:*

$$\lim_{n\to\infty} \binom{2n}{n}^{\frac{1}{2n}} = 2$$

Proof. It is known that the following sum holds:

$$\sum_{i=0}^{2n} \binom{2n}{i} = 2^{2n}$$

We also know that $\binom{2n}{n}$ is the largest binomial coefficient of degree $2n$, thus we can say the following:

$$\frac{2^{2n}}{2n+1} \leq \binom{2n}{n} \leq 2^{2n}$$

from which we derive:

$$\left(\frac{2^{2n}}{2n+1}\right)^{\frac{1}{2n}} \leq \binom{2n}{n}^{\frac{1}{2n}} \leq \left(2^{2n}\right)^{\frac{1}{2n}}$$

$$\Rightarrow 2\left(\frac{1}{2n+1}\right)^{\frac{1}{2n}} \leq \binom{2n}{n}^{\frac{1}{2n}} \leq 2$$

We can now take limits as $n \to \infty$, applying Lemma 3.1 to the left hand side (noting that the +1 in the denominator does not affect the asymptotic behaviour thus allowing us to apply the lemma) for

$$\lim_{n\to\infty} 2\left(\frac{1}{2n+1}\right)^{\frac{1}{2n}} \leq \lim_{n\to\infty} \binom{2n}{n}^{\frac{1}{2n}} \leq \lim_{n\to\infty} 2$$

$$\Rightarrow 2 \leq \lim_{n\to\infty} \binom{2n}{n}^{\frac{1}{2n}} \leq 2$$

and thus

$$\lim_{n\to\infty} \binom{2n}{n}^{\frac{1}{2n}} = 2$$

as required. □

We now move on to a slightly more complicated lemma to do with counting walks in d-regular trees. We remind ourselves at this point that trees are defined as connected, acyclic graphs, and walks are any series of contiguous edges that can be used to traverse between two points (that is, allowing repeated use of single edges).

Lemma 3.3. *Given the infinite d-regular tree T^d, let $\delta'(2l)$ be the number of walks in T^d from a a vertex v to itself of length $2l$ such that the vertex v is encountered only at the beginning and end of the walk. Note that in such a case we are interested only in even length walks, as in a tree it is not possible to have an odd length walk from a vertex to itself. Then the following holds:*

$$\delta'(2l) = \frac{1}{l}\binom{2l-2}{l-1} d(d-1)^{l-1}$$

Proof. We want to consider the possible walks of length $2l$ from a vertex v to itself such that the first instance of encountering v (after the beginning of the walk) is at the end. We can see that in order to do this, we must essentially take l steps away from v and l steps towards it (away from and towards being well defined as T^d is a tree). The first of the steps away can be chosen from any of d edges emanating from v. The subsequent $l-1$ steps away can be chosen from any of $d-1$ edges (as one step at each vertex is goes back towards v). The final condition is that we must not meet the source vertex v until the end, which is to say that the walks must always have more outsteps than insteps until the last step. Thus the problem is equivalent to counting the number of ways of summing together $l-1$ 1's and $l-1$ -1's such that the total after adding each subsequent summand is still non-negative (the first and last step of the walk are not included in this counting method, as the first step is counted by the coefficient of d and the last step is determined).

To do this, we first consider that the number of possible permutations of $l-1$ 1's and $l-1$ -1's is

$$\binom{2(l-1)}{l-1}$$

as once we have chosen the locations of all the 1's on the sequence, without loss of generality, we have determined the locations of all the -1's. Now we consider the sequences $(a_1, ..., a_{2l-2})$ such that condition of having no partial sums less than zero is violated. Clearly, the smallest such partial sum would be odd in length and of the following form

$$(a_1, ..., a_{2k}, a_{2k+1}),\ a_{2k+1} = -1,\ \sum_{i=1}^{2k} a_i = 0$$

Then consider the effect of reversing the sign of these $2k+1$ first elements of such a sequence. This would create a new sequence $(b_1, ..., b_{2l-2})$, and it is not too hard to see that this must contain l 1's and $l-2$ -1's; as the sum up to a_{2k} was 0, the effect on the total numbers of 1's and -1's is only the change of a_{2k+1} from a -1 to 1.

We want to show that this transformation forms a bijection between the sequences with negative partial sums and the sequences of l 1's and $l-2$ -1's. Let us call this transformation f, and suppose $f(x) = f(y)$ for some x, y sequences of the desired form. Then consider the first point k in $f(x)$ such that the partial sum up to $f(x)_k$ is 1. Then we can change $f(x)_k$ to a -1, and also change the parity of all x_i, $i < k$, and observe that this is essentially an inverse to f; it will yield the original x where k is the coordinate of the first partial sum that is negative. But we can apply the same inverse to $f(y)$ to observe that $x = y$, thus f is injective. For surjectivity, we simply notice that as there are l 1's and $l-2$ -1's in any sequence in the range, there must be some k for $k < 2l-2$ such that the partial sum up to k is positive, and thus we can apply the inverse to find an element of the domain with a negative partial sum. Thus f is a bijection, so we simply need to count the number of possible sequences of l 1's and $l-2$ -1's, which we can see is $\binom{2l-2}{l}$. This gives the total number of possible sequences of $l-1$ 1's and -1's such that the partial sums are non-negative as

$$\binom{2l-2}{l-1} - \binom{2l-2}{l} = \frac{(2l-2)!}{(l-1)!(l-1)!} - \frac{(2l-2)!}{l!(l-2)!} = \frac{(2l-2)!}{(l-1)!(l-2)!}\left(\frac{1}{l} - \frac{1}{l-1}\right)$$

$$= \frac{(2l-2)!}{(l-1)!(l-2)!}\frac{1}{l(l-1)} = \frac{1}{l}\binom{2l-2}{l-1}$$

We then remind ourselves that there are d choices for the first move away from v, and $d-1$ choices for every other move away from v, therefore we can calculate the number of walks from v to itself of length $2l$ in T^d as

$$\delta'(2l) = \frac{1}{l}\binom{2l-2}{l-1} d(d-1)^{l-1}$$

□

We now come to the main motivating theorem, which shows that Ramanujan graphs are optimal in terms of their bound on the spectral gap. This theorem is due to Alon and

Boppana (see [1]), but we follow the proof given by Lubotzky, Phillips and Sarnak [16], as it is far more succinct and deals only with the tools and language we have established in this paper. One caveat of using the shorter presentation of [16] is that the proof applies only connected, d-regular graphs, whereas the theorem itself applies to any d-regular graph. Firstly, however, we need a definition.

Definition 3.1 (Universal cover)**.** *Given an infinite family of graphs $\{G_n\}$, the universal cover (or universal covering graph) is the acyclic, connected graph that covers G_n for all $n \in \mathbb{N}$. By covers, we mean that for each G_n, then the intersection of the covering and G_n contains all n vertices and is still connected, thought does not necessarily contain every edge of G_n. For a family of d-regular graphs, the universal cover is the infinite d-regular tree T^d.*

Theorem 3.1 (Alon-Boppana)**.** *Given some infinite family of d-regular graphs $\{G_n\}$, let A_n be the adjacency matrix of G_n and let λ denote the second-largest largest eigenvalue (the largest is d). Note n is the number of vertices of G_n. Then the following holds*

$$\lim_{n\to\infty} \lambda \geq 2\sqrt{d-1}$$

Proof. We know that A_n^l gives a matrix that has in its (i,j)'th position the number of possible paths from i to j of length l. Then if we let $\{\lambda_1, ...\lambda_n\}$ be the eigenvalues of A_n (which we henceforth shorten to A), ordered from smallest to largest, we can use a result of linear algebra relating the trace to the eigenvalues for

$$\sum_{i=1}^{n} \lambda_i^l = \sum_{i=1}^{n} A_{i,i}^l$$

We remember that the universal cover of a family of d-regular graphs is the d-regular tree, T^d. Let $\delta(l)$ be the number of length l walks from vertex v to itself in T^d (we note that this independent of the choice of v, and therefore v does not appear in the notation). Then as T^d is a universal cover of $\{G_n\}$, the following inequality holds for each $1 \leq i \leq n$:

$$A_{i,i}^l \geq \delta(l)$$

We can then use the above inequality to substitute in for the following:

$$\sum_{i=1}^{n} \lambda_i^l \geq n\delta(l)$$

We know that two of the eigenvalues are $\pm d$, so let these be eigenvalues λ_1, λ_n and observe that we can remove them from the inequality as follows:

$$\sum_{i=2}^{n-1} \lambda_i^l \geq (n-2)\delta(l) - 2d^l$$

$$\Rightarrow \lambda^l \geq \delta(l) - \frac{2d^l}{n-2}$$

where the second line comes from the fact that $\lambda \geq \lambda_i$ for $2 \leq i \leq n-1$.

We then let $\delta'(l)$ be the number of walks from some vertex v to itself in T^d such that the first time the walk reaches v (other than the start of the walk) is at length exactly l, as in Lemma 3.3. We note that $\delta(l) \geq \delta'(l)$. As proven in the lemma, we have

$$\delta'(2l) = \frac{1}{l}\binom{2l-2}{l-1}d(d-1)^{l-1}$$

which we substitute in for

$$\lambda^{2l} \geq \frac{1}{l}\binom{2l-2}{l-1}d(d-1)^{l-1} - \frac{2d^{2l}}{n-2}$$

We then note that as we are taking the limit for $n \to \infty$, and that the above inequality should hold for all possible walk lengths in the graphs G_n, so it must hold as we take the limit $l \to \infty$ as well. We also note that we can omit the $\frac{2d^l}{n-2}$ term as this will disappear as $n \to \infty$, and that as $d > d-1$, we have $(d-1)^l < d(d-1)^{l-1}$. So we rewrite and take roots of the order $\frac{1}{2l}$ to derive the expression:

$$\lambda \geq \left(\frac{1}{l}\binom{2l-2}{l-1}\sqrt{(d-1)}^{2l}\right)^{\frac{1}{2l}}$$

$$\Rightarrow \lambda \geq (\frac{1}{l})^{2l} \cdot \binom{2l-2}{l-1}^{\frac{1}{2l}} \cdot \sqrt{d-1}$$

and thus applying Lemma 3.1 to the first expression and Lemma 3.2 to the second, letting $l \to \infty$ with n, we derive

$$\lim_{n\to\infty} \lambda \geq 2\sqrt{d-1}$$

as required. □

4 Explicit Constructions of Ramanujan Graphs

For the purposes of constructing networks with strong expansion properties in real world applications, it is necessary to find explicit formulations of the expander graphs which obtain such properties. Having shown that Ramanujan graphs are optimal spectral expanders, it is therefore interesting to try and find ways to explicitly construct infinite families of Ramanujan graphs that may be translated into practical applications. There have been many such constructions given (see [3] or [19] for example), but we will follow one of the first such constructions given in a 1988 paper of Luboztky, Phillips and Sarnak [16].

4.1 Construction from Cayley Graphs

Definition 4.1 (Cayley graph)**.** *Let G be a group and S be some symmetric set of elements of G. Then the Cayley graph generated by (G, S) is given as the graph where the vertices correspond to elements of G, and given $g, h \in G$, the edge (g, h) exists if and only if there exists some $s \in S$ such that $gs = h$.*

The graphs which we will show are Ramanujan graphs will be Cayley groups based on the following types of groups:

Definition 4.2 (General linear group)**.** *Given some n-dimensional vector space V over a field (or commutative ring) $\mathbb{F}$, the general linear group $GL(V)$, or $GL(n, \mathbb{F})$ (which is the notation we will adopt throughout this thesis) is the group formed by the $n \times n$ invertible matrices with entries in $\mathbb{F}$ with the usual operation of matrix multiplication.*

Definition 4.3 (Projective linear group)**.** *The projective general linear group $PGL(n, \mathbb{F})$ is the group given by the following quotient:*

$$PGL(n, \mathbb{F}) = GL(n, \mathbb{F})/Z(n, \mathbb{F})$$

where $Z(n, \mathbb{F})$ is the group of scalar transformations in the vector space V given by n and $\mathbb{F}$. Intuitively this can be thought of as the group of $n \times n$ invertible matrices with entries in $\mathbb{F}$ with determinant ± 1, when in reality it is the equivalence classes given by associating each such matrix with all of its scalar multiples.

Definition 4.4 (Special linear group)**.** *The special linear group $SL(n, \mathbb{F})$ is the subgroup of the general linear group formed by every element of determinant 1. The projective special linear group $PSL(n, \mathbb{F})$ is given by the following quotient*

$$SL(n, \mathbb{F})/SZ(n, \mathbb{F})$$

where $SZ(n, \mathbb{F})$ is the set of scalar transformations of determinant 1. This can be thought of as the group of $n \times n$ invertible matrices with determinant 1, where M and $-M$ are associated. We note again that PGL and PSL can be defined for a ring R rather than a field $\mathbb{F}$, and we will see that we will choose the ring of integers $\mathbb{Z}$ for some of our treatment of these groups in this thesis.

There are two slightly different cases we must consider for graph construction. We choose two non-equal primes p, q such that $p, q \equiv 1 \bmod 4$. Firstly, if p is a quadratic residue mod q, then we will construct a Cayley graph based on the projective special linear group $PSL(2, \mathbb{Z}/q\mathbb{Z})$, and if p is a quadratic non-residue mod q, then we will instead use the projective general linear group $PGL(2, \mathbb{Z}/q\mathbb{Z})$.
We first consider how to construct the generating set for the Cayley graphs. We will choose a specific generating set of $p+1$ elements that are obtained as the solutions to the following equation:

$$p = a_0^2 + a_1^2 + a_3^2 + a_4^2 \tag{1}$$

such that $a_0 > 0$ and a_0 odd, and a_1, a_2, a_3 are even. Using the following well known theorem of Jacobi (see [9] for a proof), we see that the equation without the constraints has $8(p+1)$ solutions.

Theorem 4.1 (Jacobi's four square theorem)**.** *The number of representations of a natural number n as the sum of four squares is given by*

$$r_4(n) = 8 \sum_{d|n,\ 4\nmid d} d$$

We will use the tools provided by the integer quaternions, defined below, to choose the $p+1$ solutions of Equation 1 which will form the generating set for the Cayley graph.

Definition 4.5 (Quaternions)**.** *The integral quaternions are defined as*

$$\mathbb{H}(\mathbb{Z}) = \{\alpha = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} | a_i \in \mathbb{Z}\}$$

where $\pm 1, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}$ are the quaternion units (the group of units will be denoted $\mathbb{H}^\times$), which satisfy

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$$

We refer henceforth to the space of integral quaternions simply as the quaternions, and shorten the symbol to $\mathbb{H}$, clarifying the generating ring only when it is different from $\mathbb{Z}$. The quaternions also have an associated norm which is of the form

$$|\alpha| = \alpha \cdot \bar{\alpha}$$

where $\bar{\alpha} = a_0 - a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k}$ and is called the conjugate of α

We note now that group of quaternion units acts faithfully on the ring of quaternions, that is, for any $a, b \in \mathbb{H}^\times$ there is at least one $x \in \mathbb{H}$ such that $ax \neq bx$. We return now to Equation 1, and in particular consider the constraints we have placed. We first stated that $a_0 > 0$ and odd. Because $p \equiv 1 \bmod 4$, one and only one of the a_i must be odd. The factor of 8 in Jacobi's sum of four squares theorem is in fact derived from the 8 different choices of quaternion unit, and thus we can think of the $8(p+1)$ solutions as being $p+1$ sets of 8 solutions, one for each the positive and negative of each quaternion unit. We then consider the action of the units on $\alpha = (a_0, a_1, a_2, a_3)$. Then there is only one $e \in \mathbb{H}^\times$ such that $e\alpha \equiv 1 \bmod 2$ and $a_0 > 0$ for each of the $p+1$ solution sets. So we choose precisely the $p+1$ solutions given by this method, and these will be used to form the generating set for the Cayley graph that we will show has the Ramanujan property.

Let α be one of the $p+1$ solutions described above. Choose some $i \in \mathbb{Z}$ such that $i^2 \equiv -1 \bmod q$ Then we form the following element of $PGL(2, \mathbb{Z}/q\mathbb{Z})$ to be part of the generating set:

$$\alpha' = \begin{bmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{bmatrix}$$

This gives $p+1$ elements of $PGL(2, \mathbb{Z}/q\mathbb{Z})$. As mentioned above, there are two different cases we must consider. Firstly, if p is a quadratic residue mod q, then we form the Cayley graph of $PSL(2, \mathbb{Z}/q\mathbb{Z})$ with the generating set described above. In this case, we have $p+1$ regular graph on $n = \frac{q(q^2-1)}{2}$ vertices. If p is a quadratic non-residue mod q, then we instead use $PGL(2, \mathbb{Z}/q\mathbb{Z})$ which instead gives a $p+1$-regular graph on $n = q(q^2-1)$ vertices. We will call such graphs $G^{p,q}$. This completes the construction of the graphs which in the next section we will show to be Ramanujan graphs.

5 $G^{p,q}$ are Ramanujan Graphs

Having now seen the construction of the graphs $G^{p,q}$, it remains to show that these are in fact Ramanujan graphs. Again, this is following the work of [16]. We begin by offering an outline for the proof. We want to show that the Cayley graphs defined above satisfy the Ramanujan property, which means we need to sufficiently bound their eigenvalues. We begin by some spectral analysis of the d-regular tree, which we know to be the universal cover of the family of graphs $G^{p,q}$, allowing us to relate the eigenvalues of the graphs to counts of certain periodic functions on the tree. Then, a study of the quaternions, applied to finding integer solutions to quadratic forms, allows us to create a series of groups, which we can then show are isomorphic to $PGL(2, \mathbb{Z}/q\mathbb{Z})$ or $PSL(2, \mathbb{Z}/q\mathbb{Z})$, the groups used to construct $G^{p,q}$. This allows us to relate the number of solutions of these specific quadratic forms to the study of the periodic functions on the graphs $G^{p,q}$, which we finally use to put the desired bound on the spectral gap of the graphs.

The analysis relies on a study of the numbers $r_Q(n)$, which are defined as follows:

Definition 5.1 (Representations)**.** *Given $n \in \mathbb{N}$ and $Q : \mathbb{Z}^k \to \mathbb{N}$ some function in k variables of the integers, the number of representations on n by Q is denoted $r_Q(n)$ and is the number of solutions to $Q(x) = n$ for $x \in \mathbb{Z}^k$*

We will be concerned with a particular choice of Q, which is the following quadratic form in four variables. Note that the q referred to here is the same one as referred to in $G^{p.q}$:

$$Q(x_1, x_2, x_3, x_4) = x_1^2 + 4q^2x_2^2 + 4q^2x_3^2 + 4q^2x_4^2$$

The work of Igusa in [13] shows the following identity (which was a conjecture of Ramanujan and the namesake for Ramanujan graphs):

$$r_Q(n) = C(p^k) + O(p^{k(\frac{1}{2}+\epsilon)}), \forall\ \epsilon > 0 \tag{2}$$

where O is being used as the usual asymptotic notation. $C(p^k)$ is an expression in terms of the divisors of p^k that will be explicitly derived later in the analysis.

5.1 Periodic functions on the d-regular tree T^d

We now must take a slight detour to discuss some results which take us towards the desired bound on λ. In order to do this, first consider once again the d-regular tree T^d. Let I be a group of isometries of T^d such that the action of I on T^d is free (that is, only the identity sends an element to itself), and we stipulate also that $n = |T^d/I| < \infty$. This leads to the following definition:

Definition 5.2 (Periodic function space on a graph)**.** *The space $L^2(T^d/I)$ is the space of functions on T^d such that for all $f \in L^2(T^d/I)$ we have the following periodicity:*

$$f(ix) = f(x)\ \forall\ x \in T^d,\ \forall\ i \in I$$

We now define the notion of a Laplacian matrix, which is closely related to the notion of an adjacency matrix of a graph:

Definition 5.3 (Graph Laplacian)**.** *The Laplacian of a graph G is given by $D - A$, where A is the adjacency matrix and D is the matrix formed by putting the degree of vertex i at place $D_{i,i}$ and zeroes everywhere else. In the case of d-regular graphs, as we are studying, D is simply dI where I is the identity matrix.*

Elements of $L^2(T^d/I)$ are invariant under the action of the Laplacian of T^d, which we denote Δ, which allows us to construct a orthonormal basis for $L^2(T^d/I)$ in terms of the eigenvalues. We let v_i for $0 \leq i \leq n-1$ be such that the followings two conditions hold:

$$\Delta v_i = \lambda_i, i \geq 1$$

and

$$v_0 = \frac{1}{\sqrt{n}}$$

We now introduce the notion of a point-pair function:

Definition 5.4 (Point-pair function)**.** *The point-pair function $k_l : T^d \times T^d \to \mathbb{C}$ is defined in terms of some $l \in \mathbb{N}$ such that*

$$k_l(x,y) = \begin{cases} 1, & d(x,y) = l \\ 0, & otherwise \end{cases}$$

We then are interested in a function in terms of the point-pair function defined as follows:

$$K_l(x,y) = \sum_{i \in I} k_l(ix, y)$$

This function intuitively gives the number of elements of the orbit Ix for some $x \in T^d$ that are exactly distance l away from some $y \in T^d$.

Remark 5.1. *$K_l(x,y)$ is symmetric in its variables.*

Proof. Suppose for some $i \in I$, $d(ix, y) = l$. Then since i is an isometry, $d(x,y) = d(ix, iy)$. Because of the metric on a tree, we have that

$$d(ix,y) = \begin{cases} d(x,y) + d(ix,x) \\ d(x,y) - d(ix,x) \end{cases}$$

as there is only one path between any two vertices of T^d, therefore i either takes l edges closer to y, or l edges further away. By a similar logic

$$d(x,iy) = \begin{cases} d(x,y) + d(y,iy) \\ d(x,y) - d(y,iy) \end{cases}$$

where the plus or minus is the same parity in both cases. As there is only a single path between any two points in T^d,

$$d(ix, x) = d(y, iy)$$

and therefore

$$d(x, iy) = d(ix, y) = l$$

This in turn implies that

$$K(x, y) = \sum_{i \in I} k(ix, y) = \sum_{i \in I} k(y, ix) = \sum_{i \in I} k(iy, x) = K(y, x)$$

and thus K is symmetric in its variables. $\square$

Remark 5.2. *$K_l(x, y)$ considered as a function of either single variable is an element of $L^2(T^d/I)$.*

Since $K_l(x, y)$ is symmetric in its variables, and in either variable can be considered a function in $L^2(T^d/I)$, we can decompose it into a function of the basis of $L^2(T^d/I)$ given above. This allows us to express the function as

$$K_l(x, y) = \sum_{j=0}^{n-1} h_j(l) v_j(x) v_j(y)$$

where the h_j are some as yet undefined functions that satisfy the equivalence.

We use the expansion given above to create a new function as follows:

$$L_t(x, y) = \sum_{0 \leq r \leq \frac{t}{2}} K_{t-2r}(x, y)$$

which we can think of intuitively as the number of elements of the orbit Ix that are distances $\leq t$ away from y, and which specifically are distances of the form $t-2r$ for some r away from y.

The same authors of the proof we are following have, in another publication [17], given an explicit form of the $h_j(l)$ allowing us to finally write the following:

$$L_t(x, y) = (d-1)^{\frac{t}{2}} \sum_{j=0}^{n-1} \frac{\sin(t+1)\theta_j}{\sin \theta_j} v_j(x) v_j(y)$$

where the θ_j are defined such that

$$\lambda_j = 2\sqrt{d-1} \cos \theta_j$$

Thus we can see that the eigenvalues satisfy $\lambda \leq 2\sqrt{d-1}$ if the θ_j are real for all j, as this would imply $\cos \theta_j \in [-1, 1]$, and the next steps of our analysis will be to frame the above work in the context of the graphs we have constructed in Section 4.

5.2 A return to Quaternions

In order to obtain two suitable groups such that we can apply the above analysis to the graphs constructed in Section 4, we must return to the language of quaternions. We first introduce the notion of words and reduced words.

Definition 5.5 (Group words)**.** *If G is some group, then a word in G is some product for $g_i \in G$ of the form*

$$g_1^{p_1} g_2^{p_2} ...$$

That is, a word is simply a repeated product of elements of G under the group operation. This yields the notion of a reduced word, which is simply a word in G such that all instances of $g \cdot g^{-1}$ for $g \in G$ are removed, that is the expression is fully simplified in terms of the group operation.

Definition 5.6 (Prime quaternions)**.** *A quaternion is called prime (or irreducible) if and only if its norm is prime. Note that this is not the same as prime in the standard sense of the word, as the quaternions do not have unique factorisation like the integers, but such primes are irreducible.*

We now consider the following lemma, which allows us to express all quaternions with norm p^k in terms of the $p + 1$ solutions we chose as the generating set for our Cayley graphs in Section 4.

Lemma 5.1. *Let $S = \{s_1, ..., s_{p+1}\}$ be the $p + 1$ solutions of Equation 1. For every $\alpha \in \mathbb{H}$ such that $|\alpha| = p^k$, we can find a unique representation of the form*

$$\alpha = ep^r R_m(S)$$

where e is a quaternion unit, we have $2r + m = k$ and R_m is a reduced word of length m in the elements of S.

Proof. As p is an odd prime, $|\alpha| = p^k$ must also be odd and therefore we can find a factorisation of α as

$$\alpha = \gamma\beta, |\gamma| = p, |\beta| = p^{k-1}$$

Since S is formed of the elements of the integer quaternions that satisfy $s \in \mathbb{H}, |s| = p$, for all such γ we can find some $s_i \in S$ such that $\gamma = \epsilon s_i$ for $\epsilon \in \mathbb{H}$ a unit. This allows us to write

$$\alpha = \beta\epsilon s_i$$

Now we note that $|\beta\epsilon| = p^{k-1}$ as ϵ is a unit, so we can essentially repeat the process for

$$\alpha = \beta'\epsilon s_i s_j$$

Inductively, this allows us to write

$$\alpha = \epsilon s_i ... s_k$$

We now note that for each $s \in S$, we have also that $\bar{s} \in S$, as clearly if s is a solution to Equation 1, then so too is $\bar{s}$. Therefore, we can reduce the $s_i...s_k$ expression to a reduced word, where we cancel every $s\bar{s}$ pair as $s\bar{s} = |s| = p$, which gives us

$$\alpha = \epsilon p^r R_m(S)$$

where $R_m(S)$ is a reduced word of length m in the elements of S. It is clear that $2r + m = k$ from the process by which we reduced α to the elements of S, and this completes the existence of such a representation.

For the uniqueness argument we first consider how many reduced words can possibly be created of length m. To do this, we note that in a reduced word in S, the first element of the word can be any element of S, and every subsequent element must not be the inverse of the element before it. Thus, because $|S| = p + 1$, the number of possible reduced words of length m in S is $(p+1)p^{m-1}$. Using this, we can calculate all possible representations of

$$\alpha = \epsilon p^r R_m(S)$$

for each possible r, m. We can see that there are

$$8(\sum_{0 \leq r < \frac{k}{2}} (p+1)p^{k-2r-1} + C)$$

possible representations, where $C = 0$ if k is odd and $C = 1$ if k is even. This is because these are all the possible values of m, and the corresponding number of reduced words for each such m (noting that ϵ and p^r are determined by the choice of m and so don't contribute to the calculation of the number of representations). We can expand this sum and notice that:

$$\sum_{0 \leq r < \frac{k}{2}} (p+1)p^{k-2r-1} = 1 + p + ...p^k = \frac{p^{k+1} - 1}{p - 1}$$

which we recognise as the familiar sum of divisors formula for a prime power. Therefore we can write the number of possible representations of $\alpha = \epsilon p^r R_m(S)$ as

$$8(\sum_{d|p^k} d)$$

But then we have seen that this is precisely the number of $\alpha \in \mathbb{H}$ such that $|\alpha| = p^k$, and thus there is a bijection between each such α and its representation as $\epsilon p^r R_m(S)$, and thus the representations are unique. $\square$

An immediate corollary of this, using also our early work which stated that the elements of S are the only solutions of $|\alpha| = p$ where $\alpha \equiv 1 \bmod 2$, is the following:

Corollary 5.1. *Every $\alpha \in \mathbb{H}$ such that $|\alpha| = p^k$ and $\alpha \equiv 1$ mod 2 has a unique representation as*

$$\alpha = \pm p^r R_m(S)$$

We will use the above theory of representations of element of $\mathbb{H}$ to construct groups to which we can apply to our earlier work on periodic functions. We begin by considering the set

$$\Lambda(2) = \{\alpha \in \mathbb{H} | \alpha \equiv 1 \bmod 2, |\alpha| = p^v, v \in \mathbb{Z}\}$$

and applying the following group structure: given two elements $\alpha, \beta \in \Lambda(2)$, we say $\alpha \sim \beta$ if and only if there exists $v_1, v_2 \in \mathbb{Z}$ such that

$$\pm p^{v_1}\alpha = p^{v_2}\beta$$

This is an equivalence relation, and by identifying elements belonging to the same equivalence class we obtain a group structure on $\Lambda(2)$ that corresponds to the usual notion of multiplication in $\mathbb{H}$ such that $[\alpha][\beta] = [\alpha\beta]$. From Corollary 5.1, we can see that the action of $\Lambda(2)$ on S is free. We now consider the following proposition relating free group actions to Cayley graphs.

Proposition 5.1. *Given some group G and a set S, then the Cayley graph X formed by G with respect to S is a tree if and only if G is a free group with generating set S.*

Proof. We first briefly show that the relevant Cayley graph is connected if and only if S generates G. First suppose that the graph X is connected. Then for every pair of vertices $u, v \in X$ corresponding to group elements $u, v \in G$, there is a set of edges connecting u and v. By the definition of a Cayley graph, an edge between two vertices exists if and only if there is $s \in S$ such that $su = v$. Thus there is some set $\{s_1, ..., s_k\}$ such that $s_1...s_k u = v$. In particular, let $u = 1$, and notice that therefore every element $v \in G$ is a product of elements in S, thus S generates G.

For the other direction, S generates G then choose two vertices $u, v \in G$. Then there must exist some $w \in G$ such that $uw = v$, but k can be represented as $s_1...s_k$, thus there are a series of edges in X connecting u to v, and hence X is connected.

We now show that if G is free with respect to S, X is also acyclic as well as being connected, and thus is a tree. First, suppose that there is some cycle in X. Then we have $u, v \in X$ such that there the edge (u, v) exists, but there is also some other non-trivial path from u to v, call it $\{u, w_1, ..., w_k, v\}$. Each of these edges corresponds to the multiplication of a series of elements of S, so we have $s_1...s_i u = v$ and $s_{i+1}u = v$ as well. Furthermore, $s_1 \neq s_{i+1}$, as in such a case the paths would share the first edge and be trivially equivalent, and what is more, $s_j \neq s_{j+1}^{-1}$ as this would imply a repeated edge. But then

$$s_1...s_i u = v, s_{i+1}u = v \Rightarrow s_1...s_i = s_{i+1}$$

but S is a basis of G and thus s_{i+1} is only equal to $s_1...s_i$ if $s_1...s_i$ reduces to s_{i+1} by inverses, which contradicts the assertion that $s_1 \neq s_{i+1}$ and $s_j \neq s_j \neq s_{j+1}^{-1}$, thus there can be no cycle in X. The result can also be shown in the reverse direction, but as it is not needed for our purposes, the proof is omitted. It is very similar to the above proof and can be found in [6], from which the given proof was drawn. □

The proposition implies that the Cayley graph constructed by $\Lambda(2)$ with respect to S is the $p+1$-regular tree. The $p+1$ regular tree is a universal cover of the finite $p+1$ regular graph we are looking for, and so we must find a suitable subgroup by which to quotient $\Lambda(2)$ to make the corresponding Cayley graph the correct finite subgraph. To construct this subgroup, take some $k \in \mathbb{N}$ such that $\text{GCD}(k,p) = 1$ and then consider all the $[\alpha] \in \Lambda(2)$ such that $2k|a_i$ for $\alpha = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$. Such $[\alpha]$ create a subgroup $\Lambda(2k) \subset \Lambda(2)$. The following remark shows that it is in fact a normal subgroup with finite index.

Remark 5.3. $\Lambda(2k) \subset \Lambda(2)$ *is a normal subgroup with finite index.*

Proof. Let $\mathbb{H}(\mathbb{Z}/2k\mathbb{Z})$ be the quaternions with coefficients in $\mathbb{Z}/2k\mathbb{Z}$ rather than $\mathbb{Z}$, and let $\mathbb{H}(\mathbb{Z}/2k\mathbb{Z})^*$ be the invertible elements of $\mathbb{H}(\mathbb{Z}/2k\mathbb{Z})$. Consider now the subgroup $H \subset \mathbb{H}(\mathbb{Z}/2k\mathbb{Z})^*$ defined by

$$H = \{\alpha | a_0 \neq 0\}$$

We can define a homomorphism

$$\varphi : \Lambda(2) \to \mathbb{H}(\mathbb{Z}/2k\mathbb{Z})^*, [\alpha] \mapsto (\alpha \bmod 2k)H$$

We can see that the homomorphism is well-defined simply by the definition, and furthermore we can see that if $2k|a_i, 0 \leq i \leq 3$, then $\alpha \bmod 2k = 0$ and thus we have

$$[\alpha] \in \Lambda(2k) \Rightarrow \varphi([\alpha]) = 0H$$

Furthermore, if $\varphi([\alpha]) = 0$ then we have $\alpha \bmod 2k = 0$, and thus $2k|a_i, 0 \leq i \leq 3$, which is to say that $[\alpha] \in \Lambda(2k)$, that is, $\Lambda(2k)$ is the kernel of φ, and it is well-known that the kernel of a homomorphism is a normal subgroup, hence $\Lambda(2k) \subset \Lambda(2)$ is normal. $\square$

We are particularly interested in the case where $k = q$, and we can form the quotient group $\Lambda(2)/\Lambda(2q)$. We will show that this group is isomorphic to $PGL(2, \mathbb{Z}/q\mathbb{Z})$ when p is a quadratic non-residue modulo q and isomorphic to $PSL(2, \mathbb{Z}/q\mathbb{Z})$ when p is a quadratic residue modulo q, and thus with the generator set S can be used to form the Cayley graphs $G^{p,q}$ which we constructed in Section 4. We will next construct a homomorphism explicitly and exhibit the above isomorphisms, but first we will require some more definitions:

Definition 5.7 (Quadratic form)**.** *A quadratic form is a multivariate polynomial in any number of variables where every term in the polynomial has degree exactly 2. We have already seen some examples such as:*

$$a_0^2 + a_1^2 + a_2^2 + a_3^2$$

If we consider a quadratic form on n variables as an $n \times n$ matrix, then we can define the discriminant of the quadratic form as the determinant of the corresponding matrix.

It will be helpful also to demonstrate a few properties of the spaces $PGL(2, \mathbb{Z}/q\mathbb{Z})$ and $PSL(2, \mathbb{Z}/q\mathbb{Z})$

Remark 5.4. *The space $PSL(2, \mathbb{Z}/q\mathbb{Z})$ is simple group. This is implied by the more general result that for any n and any field F, provided that $|\mathbb{F}| > 3$, $PSL(n, \mathbb{F})$ is simple (and in the cases where $|\mathbb{F}| \leq 3$, the result also holds providing $n > 2$). A proof of this result is given in [11].*

Proposition 5.2. *Any element $A \in PGL(2, \mathbb{Z}/q\mathbb{Z})$ with a square determinant corresponds to an element of $PSL(2, \mathbb{Z}/q\mathbb{Z})$*

Proof. We consider the homomorphism $\det : PGL(2, \mathbb{Z}/q\mathbb{Z}) \to \mathbb{Z}/2\mathbb{Z}$ where det is our usual notion of determinant. We can see that this homomorphism is well-defined as if we have $A, B \in PGL(2, \mathbb{Z}/q\mathbb{Z})$, then

$$\det(AB) = \det(A)\det(B)$$

Now we consider the subgroup of $PGL(2, \mathbb{Z}/q\mathbb{Z})$ given by squaring every element. Then every element A in this subgroup, call it G, satisfies $\det(A) = 1$, therefore corresponds to ane element of $PSL(2, \mathbb{Z}/q\mathbb{Z})$, which implies

$$G \subset PSL(2, \mathbb{Z}/q\mathbb{Z})$$

Furthermore, it is not hard to see that it is a normal subgroup. Calling the subgroup G, take $A \in G$ and $B \in PGL(2, \mathbb{Z}/q\mathbb{Z})$, and consider BAB^{-1}. Then

$$\det(BAB^{-1}) = \det(B)\det(A)\det(B^{-1}) = \det(B)\det(B^{-1})\det(A) = \det(A)$$

therefore, $BAB^{-1} \in G$ also. Therefore G is a normal subgroup of $PGL(2, \mathbb{Z}/q\mathbb{Z})$. But by the above remark that $PSL(2, \mathbb{Z}/q\mathbb{Z})$ is simple, and therefore the fact that $G \subset PSL(2, \mathbb{Z}/q\mathbb{Z})$ and G clearly has more than one element, gives that $G \simeq PSL(2, \mathbb{Z}/q\mathbb{Z})$. □

The above allows us to consider the following proposition.

Proposition 5.3. *Consider the following homomorphism:*

$$\varphi : \Lambda(2) \to PGL(2, \mathbb{Z}/q\mathbb{Z})$$

$$[\alpha] \mapsto \alpha \ mod\ q \mapsto \begin{bmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{bmatrix}$$

where $i^2 \equiv -1\ mod\ q$. The the image of φ is $PGL(2, \mathbb{Z}/q\mathbb{Z})$ when p is a quadratic non-residue modulo q and $PSL(2, \mathbb{Z}/q\mathbb{Z})$ when p is a quadratic residue modulo q.

Proof. It is clear by the definition of the polynomial that the image is a contained in $PGL(2, \mathbb{Z}/q\mathbb{Z})$. Since the index of $PSL(2, \mathbb{Z}/q\mathbb{Z})$ is 2, it therefore suffices to show that

$$PSL(2, \mathbb{Z}/q\mathbb{Z}) \subset \mathrm{Im}(\varphi)$$

as this implies that the image is either $PSL(2, \mathbb{Z}/q\mathbb{Z})$ or $PGL(2, \mathbb{Z}/q\mathbb{Z})$.

We first consider the determinant of one of the elements of $\mathrm{Im}(\varphi)$, call it A. Clearly, as $i^2 \equiv -1 \bmod q$, we have

$$\det(A) = a_0^2 + a_1^2 + a_2^2 + a_3^2$$

We can therefore think of the homomorphism as decomposing into two homomorphisms as follows:

$$\Lambda(2) \to \mathbb{H}(\mathbb{Z}/q\mathbb{Z})^*/Z \to PGL(2, \mathbb{Z}/q\mathbb{Z})$$

where Z is, as above, the subset of $\mathbb{H}(\mathbb{Z}/q\mathbb{Z})$ with $a_0 > 0$. Because of the form of the determinant above, each element in $\mathbb{H}(\mathbb{Z}/q\mathbb{Z})^*$ (that is, the invertible elements) corresponds to an element of $PGL(2, \mathbb{Z}/q\mathbb{Z})$; the latter part of the decomposition is an isomorphism. Therefore we just need to consider the first part of the decomposition.

To show the proposition, we consider some $\beta \in \mathbb{H}(\mathbb{Z}/q\mathbb{Z})$ such that $|\beta| \equiv 1 \bmod q$ and show that there is a corresponding element of $\Lambda(2)$ under φ. If $\beta = b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, then we can define some γ as follows:

$$\gamma = c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}, c_0 \equiv b_0 \bmod q, 2c_i \equiv b_i \bmod q \text{ for } i \geq 1$$

which in turn implies that

$$c_0^2 + 4c_1^2 + 4c_2^2 + 4c_3^2 \equiv 1 \bmod q$$

From the work of [18], we can say a few things about quadratic forms of this type. Take some quadratic form $f(x_0, ..., x_n)$ on $n \geq 4$ variables with integer coefficients and discriminant d. If we take some integer g such that $\mathrm{GCD}(g, 2d) = 1$ and some other integer k (which must be sufficiently large, but for our purposes we can allow it to be arbitrarily large therefore we can ignore this part of the argument) such that $\mathrm{GCD}(g, 2kd) = 1$, then we can say that if

$$GCD(b_0, ..., b_n, g) = 1, f(b_1, ..., b_n) \equiv k \bmod g$$

then there also exists a solution $(a_0, ..., a_n)$ such that

$$(a_0, ..., a_n) \equiv (b_0, ..., b_n) \bmod g, f(a_0, ..., a_n) = k$$

In our case, we have

$$f(x_0, x_1, x_2, x_3) = x_0^2 + 4x_1^2 + 4x_2^2 + 4x_3^2$$

and $m = p^k$, $g = q$, and $(b_0, b_1, b_2, b_3) = (c_0, c_1, c_2, c_3)$. Therefore, if we let k be sufficiently large and require $p^k \equiv 1 \bmod q$, which we can enforce as p is a generator mod q, then we can find some (a_0, a_1, a_2, a_3) such that

$$(a_0, a_1, a_2, a_3) \equiv (c_0, c_1, c_2, c_3) \bmod q, f(a_0, a_1, a_2, a_3) = p^k$$

Therefore, we have $\alpha = a_0 + 2a_1\mathbf{i} + 2a_2\mathbf{j} + 2a_3\mathbf{k}$ with the property that $|\alpha| = p^k$, $\alpha \equiv \beta \bmod q$ and as $a_0 \equiv 1 \bmod 2$, we have $\alpha \equiv 1 \bmod 2$, such that $\alpha \in \Lambda(2)$. Therefore, we have demonstrated that for each $\beta \in \mathbb{H}(\mathbb{Z}/q\mathbb{Z})^*/Z$, we can find a corresponding $\alpha \in \Lambda(2)$ via φ and thus $PSL(2, \mathbb{Z}/q\mathbb{Z}) \subset \mathrm{Im}(\varphi)$ as required. The fact that the image is $PSL(2, \mathbb{Z}/q\mathbb{Z})$ when p is a quadratic residue mod q and $PGL(2, \mathbb{Z}/q\mathbb{Z})$ comes from the fact that these are the only two subgroups of $PGL(2, \mathbb{Z}/q\mathbb{Z})$ containing $PSL(2, \mathbb{Z}/q\mathbb{Z})$ and Proposition 5.2 where we demonstrated that matrices with square determinants are in $PSL(2, \mathbb{Z}/q\mathbb{Z})$. □

With the above proposition, we can apply the first isomorphism theorem, given that $\Lambda(2q)$ is the kernel of φ, to see that

$$\Lambda(2)/\Lambda(2q) \simeq \begin{cases} PGL(2,\mathbb{Z}/q\mathbb{Z}), & p \text{ is a quadratic non-residue mod } q \\ PSL(2,\mathbb{Z}/q\mathbb{Z}), & p \text{ is a quadratic residue mod } q \end{cases}$$

This is crucial as our earlier work on periodic function on the d-regular tree requires two groups to quotient and observe periodic functions upon. What is more, we can see that our construction of φ is such that the elements of the set S of the $p+1$ solutions we chose as generators are mapped to the elements of $PGL(2,\mathbb{Z}/q\mathbb{Z})$ that allow us to construct our Cayley graphs $G^{p,q}$. This allows us to identify the graphs $G^{p,q}$ with the groups $\Lambda(2)/\Lambda(2q)$. We now apply the analysis of periodic functions to $\Lambda(2)/\Lambda(2q)$.

5.3 Tieing it all together

We return to the end of our analysis on periodic functions on the d-regular tree. In that analysis, we used T^d and quotiented by some group I, but we can apply the same analysis where we substitute T^d for $\Lambda(2)$ and I for $\Lambda(2q)$. Bearing in mind that Proposition 5.3 implies that $\Lambda(2)/\Lambda(2q)$ can be identified with $G^{p,q}$, any insight we derive on the spectrum of the Laplacian of $\Lambda(2)/\Lambda(2q)$ is directly applicable to $G^{p,q}$. With that in mind, we remind ourselves of

$$K_l(x,x) = |\{y \in \Lambda(2q) : d(yx,x) = l\}| = |\{y \in \Lambda(2q) : d(x^{-1}yx,e) = l\}|$$

for $e \in \Lambda(2)$ a unit. We remind ourselves of Remark 5.3, which states that $\Lambda(2q) \subset \Lambda(2)$ is a normal subgroup, thus we can say

$$K_l(x,x) = |\{y \in \Lambda(2q) : d(ye,e) = l\}|$$

that is, $K_l(x,x)$ is independent of the choice of $x \in \Lambda(2)$, thus for all $x \in \Lambda(2)$ we have

$$K_l(x,x) = K_l(e,e)$$

which in turn implies

$$L_t(x,x) = L_t(e,e)$$

We want to establish a relation between the spectrum of the Laplacian on the graph and $r_Q(n)$, the number of solutions of a quadratic form Q for some n. We choose the following quadratic form:

$$Q(x_0,x_1,x_2,x_3) = x_0^2 + 4q^2x_1^2 + 4q^2x_2^2 + 4q^2x_3^2$$

Thus we have that $r_Q(p^k)$ is equal to the number of $\alpha \in \mathbb{H}$ such that $|\alpha| = p^k$ and $2q|a_1,a_2,a_3$. We can count these using the theory of reduced words established in Corollary 5.1, which shows that every such α has a representation in the form

$$\alpha = \pm p^r R_m(S)$$

where $2r + m = k$. Moving back to the language of graph theory, we can think of each element of the reduced word intuitively as an edge (as they are the generators of the Cayley graph), and in this sense think of the reduced word as describing a path along the graph. By considering this identification, we see that we can describe $r_Q(p^k)$ in the following way:

$$r_Q(p^k) = 2 \sum_{0 \leq r < \frac{k}{2}} |\{\alpha \in \Lambda(2) : d(\alpha \cdot e, e) = k - 2r\}|$$

But we recognise the right hand side and thus can write the following:

$$r_Q(p^k) = 2L_k(e, e)$$

We now recall the results of Section 5.1, which established a relationship between $L_k(e, e)$ and the spectrum of the Laplacian, thus allowing us to write

$$r_Q(k) = \frac{2p^{\frac{k}{2}}}{n} \sum_{i=0}^{n-1} \frac{\sin(k+1)\theta_i}{\sin \theta_i}$$

thereby establishing the desired connection between the spectrum of the graph $G^{p,q}$ and $r_Q(p^k)$, as the θ_i are defined in terms of the eigenvalyes of the Laplacian on the graph.

Our next step is to consider Equation 2, and use it to prove that the θ_i must be real-valued. The equation was as follows:

$$r_Q(p^k) = C(p^k) + O(p^{k(\frac{1}{2}+\epsilon)})$$

Here $C(p^k)$ is what is known as a singular series, dependent on p, q and k. A digression into a discussion of such series does not shed much light on the argument presented here and so is omitted, but thanks to the work of Hecke [8], it can be shown that such series have the following form:

$$C(p^k) = \sum_{d|p^k} dF(d)$$

where F is a periodic function with period $4q^2$. This leads to the following short lemma:

Lemma 5.2. *If $F : \mathbb{N} \to \mathbb{C}$ is some periodic function and has the property that*

$$\sum_{d|p^k} dF(n) = o(p^k), k \to \infty$$

then

$$\sum_{d|p^k} dF(d) = 0 \ \forall \ k$$

Proof. We let

$$x_k = \sum_{d|p^k} dF(d)$$

and observe that

$$\frac{x_k}{p^k} - \frac{x_{k-1}}{p^{k-1}} \cdot \frac{1}{p} = F(p^k)$$

as the only divisors of p^k are the powers of p. Now as $x_k = o(p^k)$ as $k \to \infty$, p^k is asymptotically at least as large, and so the expression

$$\frac{x_k}{p^k} - \frac{x_{k-1}}{p^{k-1}} \cdot \frac{1}{p} \to 0, k \to \infty$$

but then F is periodic so if it goes to 0, it is subsequently 0 for all k, as required. □

As we know $C(p^k)$ takes this form, we can write

$$\sum_{d|p^k} dF(d) + O(p^{k(\frac{1}{2}+\epsilon)}) = \frac{2p^{\frac{k}{2}}}{n} \sum_{i=0}^{n-1} \frac{\sin(k+1)\theta_i}{\sin\theta_i}$$

We now distinguish between the two possible sets of eigenvalues for the different graphs depending on whether p is a quadratic residue mod q or not. This leads to the following proposition:

Proposition 5.4. *The graph $G^{p,q}$ where p is a quadratic non-residue mod q is bipartite, and when p is a quadratic residue mod q it is not.*

Proof. First we consider the case where p is a quadratic non-residue. In such a case, we observe that edges are only formed between those elements of $A, B \in PGL(2, \mathbb{Z}/q\mathbb{Z})$ such that for some $s \in S$, $sA = B$. Every element of S has determinant p, as we have seen before, and therefore if A has square determinant, and A has an edge to B, then B has a non-square determinant, since p is not a square and thus $\det(A) \cdot p$ is also non-square. Those elements with square determinants are the subgroup $PSL(2, \mathbb{Z}/q\mathbb{Z})$, and those with non-square are the complement, therefore there are no edges within elements of either subgroup, only between the two, and thus this subgroup partitions the group $PGL(2, \mathbb{Z}/q\mathbb{Z})$ into a bipartite graph.

We now show that when p is a quadratic residue, the graph is not bipartite. Suppose that such a partition exists. Then $PSL(2, \mathbb{Z}/q\mathbb{Z})$ is partitioned into two sets S_1, S_2. Whichever of these sets contains the identity is in fact a subgroup of $PSL(2, \mathbb{Z}/q\mathbb{Z})$. But then this is an index 2 subgroup, which must be normal, which contradicts the fact that $PSL(2, \mathbb{Z}/q\mathbb{Z})$ is simple, established in Remark 5.4. □

With this in mind, there are two different cases for the possible eigenvalues of $G^{p,q}$. We first consider the case where p is a quadratic residue mod q. In such a case, only one of the eigenvalues can take the maximal value $p+1$. First, we give a quick analysis of the θ_i that is due to [17]. We can partition the eigenvalues in their form

$$\lambda_i = 2\sqrt{p}\cos\theta_i$$

as

$$x_i = \sqrt{p}e^{i\theta_i}, x_i' = \sqrt{p}e^{-i\theta_i}$$

With this, it isn't hard to see that $x_i + x_i' = \lambda_i$, but we can also consider

$$\frac{x_i^{k+1} - x_i^{k+1'}}{x_i - x_i'} = p^{\frac{k}{2}}\frac{e^{(k+1)i\theta_i} - e^{-(k+1)i\theta_i}}{e^{i\theta_i} - e^{-i\theta_i}} = p^{\frac{k}{2}}\frac{2i(\sin(k+1)\theta_i)}{2i(\sin\theta_i)} = p^{\frac{k}{2}}\frac{\sin(k+1)\theta_i}{\sin\theta_i}$$

From Equation 2, which is as follows

$$C(p^k) + O(p^{k(\frac{1}{2}+\epsilon)}) = \frac{2p^{\frac{k}{2}}}{n}\sum_{i=0}^{n-1}\frac{\sin(k+1)\theta_i}{\sin\theta_i}$$

we can consider the case for $\lambda_0 = p+1$, which corresponds to the first summand thus giving

$$C(p^k) + O(p^{k(\frac{1}{2}+\epsilon)}) = \frac{2}{n}\left(\frac{p^{k+1}-1}{p-1}\right) + o(p^k)$$

since because this eigenvalue is largest, its corresponding summand is also largest and therefore $o(p^k)$ bounds the other summands. But then we remember Lemma 5.2 which tells us that, as $C(p^k)$ satsifies

$$C(p^k) = \sum_{d|p^k} dF(d)$$

for some periodic F, then $C(p^k)$ is either identically 0 or

$$C(p^k) = \frac{2}{n}\left(\frac{p^{k+1}-1}{p-1}\right)$$

We know by the definition of the singular series that $C(p^k)$ is not 0, and thus the relationship above holds, allowing us to write

$$\frac{2p^{\frac{k}{2}}}{n}\sum_{i=1}^{n-1}\frac{\sin(k+1)\theta_i}{\sin\theta_i} = O(p^{k(\frac{1}{2}+\epsilon)})$$

$$\Rightarrow \sum_{i=1}^{n-1}\frac{\sin(k+1)\theta_i}{\sin\theta_i} = O(p^{\epsilon k})$$

As this holds $\forall\ \epsilon > 0$, we need to be able to arbitrarily bound the size of the summands on the left. We can use an expansion for the summands as follows:

$$\frac{\sin(k+1)\theta_i}{\sin\theta_i} = \frac{e^{i(k+1)\theta_i}}{e^{i\theta_i}}\frac{\sin(k+1)\theta_i}{\sin\theta_i} = \frac{1-e^{2i(k+1)\theta_i}}{1-e^{2i\theta_i}}$$

$$= 1 + e^{2i\theta_i} + ... + e^{2i(k+1)\theta_i}$$

It is clear from this expression that if the θ_i are all real this implies each of these summands is bounded in the range $[-1, 1]$, and what is more, complex values for the θ_i would leave the sum unbounded. This in turn implies that the θ_i are all real. But as we saw before, this implies that $|\lambda_i| \leq 2\sqrt{p}$ and as $d = p+1$, $|\lambda_i| \leq 2\sqrt{d-1}$ and thus $G^{p,q}$ is a Ramanujan graph.

Finally, we consider the case where p is a quadratic non-residue mod q. In this case, as we showed, the graph $G^{p,q}$ is bipartite. The eigenvalues of bipartite graphs always appear in pairs (one positive, one negative), therefore both $p+1$ and $-(p+1)$ are eigenvalues. The only difference in the analysis then is that the two summands corresponding to these eigenvalues are together equal to

$$\frac{4}{n}\left(\frac{p^{k+1}-1}{p-1}\right)$$

and thus $C(p^k)$ instead takes this value. But then the same comparison to $O(p^{\epsilon k})$ holds for the other eigenvalues, giving again that the remaining θ_i are real and thus all the other eigenvalues satisfy $\lambda \leq 2\sqrt{d-1}$, thus in this case as well $G^{p,q}$ are Ramanujan graphs. Thus we have demonstrated an explicit construction for an infinite family of Ramanujan graphs. □

6 Applications of Expanders and Error Correcting Codes

At the beginning of this thesis we briefly touched on some applications of expander graphs, particularly referring to their power to construct well connected networks that are economical in terms of the number of edges present in the network. The motivating example in Section 2 gave a strong impetus for our interest in finding explicit constructions of particularly good expanders such as Ramanujan graphs. There are many other useful applications of such graphs to areas of computer science, such as pseudo-random number generators ([14]) and cryptographic hash functions ([4]), and one recent application where Ramanujan graphs in particular can be shown to be useful is the field of error correcting codes.

6.1 Introduction to Error Correcting Codes

Error correcting codes are a method of encoding binary information first demonstrated by Richard Hamming [7], building on earlier theoretical work by Claude Shannon [22], that contain a series of bits that are not related to the information contained in the encoded message, known as a redundant, that provides a mechanism for the code to correct any errors

that may have occurred during transmission.

An easy way to see this is to consider a very simplistic example. Suppose that we are trying to encode some piece of binary information, that is, we can represent the message with either a 0 or a 1. However, we are transmitting our message over a noisy channel so we must use some sort of error correction or risk our message being corrupted along the way. A very simple method we can use is to simply repeat our message n times. That is, we send a message that is composed of a binary string of either n 0's or n 1's. When the the code is received, some bits of the code may have been corrupted by noise on the channel, so the receiver will take whichever symbol appears more frequently in the received string to be the intended message. This idea of a repetition code can be scaled to larger messages as well, but it relies merely on the hope that only a few of the bits in the string will be corrupted. While this is not a very sophisticated error correcting code (nor is it very powerful, or widely used), we will demonstrate in the next few sections that more complex error correcting codes can in fact be very efficient and certain constructions behave in such a way as to put asymptotic bounds on their effectiveness.

6.2 Applications of Error Correcting Codes

Error correcting codes have found many applications in modern society [15]. One good example is in computer memory. Data is stored in computer memory in the form of silicon chips which are susceptible to unavoidable errors due to radiation. Thus it makes sense to employ an error correcting code when storing and retrieving information to combat any changes that may have occurred in the physical state of the chip, and it is in fact a Hamming code that is frequently used to encode and decode data stored in computer memory for exactly this reason. Another application of error correcting codes, albeit to a technology that is becoming more and more outmoded, is data storage in compact discs. For quite similar reasons to computer memory, it makes a lot of sense to use error correcting codes when storing information in this format. To this end, the error correcting codes which are used are known as Reed-Solomon codes, and are particularly interesting in the context of this paper as they are constructed using similar algebraic tools to the codes we will construct from expander graphs in the next section. A full description of the implementation of these codes in CDs can be found in [10]

Another example of error correcting codes impacting modern technologies is the encoding of information being sent to and from spacecraft. When sending data to and from spacecraft, the physical limitation of the speed of light is such that there is a significant overhead in terms of the time it takes to send and receive communications, as the distances over which communication occurs are so large. Therefore, if error were to occur on the communication channel, retransmission would be a costly way to correct such an error, and what is more, errors are more likely to occur the larger the transmission distance. For this reason, error correcting codes are often used to transmit data between spacecraft and Earth, and have

been used for many significant achievements in space exploration such as transmission of photographs of the surface of Mars, which were transmitted with codes known as Reed-Solomon codes. In this regard, error correcting codes may become increasingly important. If we entertain some possible futures that may sound like science fiction, such as human colonisation of exoplanets or discovery of extraterrestrial intelligent life, then long distance communication of this type will become commonplace, and will only be possible thanks to error correcting codes. Thus it will be increasinly important for us to construct efficient and robust error correcting codes, and the constructions from expander graphs given below allow us to achieve this.

6.3 Mathematical definitions

We now introduce the mathematical notation associated with error correcting codes and codes in general, before showing how error correcting codes can be constructed by graphs, and why expanders make particularly good codes. Let us begin with the definition of codes and codewords:

Definition 6.1 (Codes)**.** *A code* $\mathbf{C}$ *is a set of* n*-length binary strings, each of which is known as a codeword. We can usually think of codes as subsets of* $\mathbb{F}_2^n$*.*

When we think of error correcting codes, and codes in general, we need to employ some sort of metric by which we can quantify firstly how much a codeword has changed during transmission, and secondly how well the error-correction has fixed it.

Definition 6.2 (Hamming distance)**.** *The Hamming distance between two codewords* $x, y \in \mathbb{F}_2^n$ *is defined as follows:*

$$d(x, y) = |\{i : x_i \neq y_i\}|$$

Intuitively, this is just the sum of the digits of the codewords that are different, and provides a useful metric for comparing codewords.

With this metric, we can construct the notion of a code distance.

Definition 6.3 (Code distance)**.** *Given a code* $\mathbf{C}$*, the code distance is defined as*

$$\rho(\mathbf{C}) = \min_{x,y \in \mathbf{C}, x \neq y} d(x, y)$$

This also yields the notion of relative distance, which is given by

$$\delta(\mathbf{C}) = \frac{\rho(\mathbf{C})}{n}$$

where n *is the length of codewords in* $\mathbf{C}$*. Intuitively, it is better to have a larger distance, as if codewords are further away from each other within the codes then it is less likely that errors will take codewords to other valid codewords, and it means that errors can be more easily detected and corrected in error correcting codes. For that reason, we will want to place an asymptotic lower bound on distance that is greater than 0.*

Intuitively, it is better to have a larger distance, as if codewords are further away from each other within the codes then it is less likely that errors will take codewords to other valid codewords, and it means that errors can be more easily detected and corrected in error correcting codes. For that reason, we will want to place an asymptotic lower bound on distance that is greater than 0. We are also concerned with how much of our code is taken up by the redundant, leading to the next definition.

Definition 6.4 (Code rate)**.** *There is also the notion of a rate of a code, which is intuitively how much of the code is being used to encode the redundant. Therefore, for a code of with words of length n, that encode k bits of useful information, we can define*

$$r(\mathbf{C}) = \frac{k}{n}$$

With the notions of relative distance and rate we can analyse the asymptotic behaviour of a code in these two metrics to judge whether an error correcting code is successful or not.

Definition 6.5 (Asymptotically good codes)**.** *A family of codes $\mathbf{C_n}$ that are based on the same construction is called asymptotically good if there exist constants $d, r > 0$ such that*

$$\delta(\mathbf{C}) > d, r(\mathbf{C}) > r$$

for all $n \in \mathbb{N}$.

Now that we have introduced the mathematical language of codes, and metrics by which we can judge the quality of a code, we next move to the construction of error correcting codes from expander graphs, and show that better expansion leads to better codes.

6.4 Expander Codes

We now move on to the construction of error correcting codes from expander graphs. The construction requires us to construct a bipartite graph from our base graph, so we now demonstrate a method for such a construction.

Definition 6.6 ((a, b)-regular graphs)**.** *An (a, b)-regular bipartite graph is a bipartite graph G, where every vertex in one side of the partition has degree a and every vertex in the other side of the partition has degree b.*

Definition 6.7 (Edge-vertex incidence graphs)**.** *The edge-vertex incidence graph of a d-regular graph $G - (V, E)$ is the graph G' generated by taking the vertex set as $V' = V \cup E$ and the edge set is given by*

$$E' = \{(e, v) \in E \times V | v \in e\}$$

That is, the graph forms edges between vertices and edges of the original graph G if and only if the vertex v is an endpoint of e in G. This constructs a $(2, d)$-regular bipartite graph with n vertices on one side, which correspond to the original vertices of G and each have degree d, and $\frac{dn}{2}$ vertices on the other side, corresponding to the edges of the original graph G and each with degree 2.

Given this method for constructing unbalanced bipartite graphs, we can now construct error correcting codes based on such graphs. The construction also relies on fixed length linear codes, defined below.

Definition 6.8 (Linear block codes)**.** *A linear block code* $\mathbf{C}$ *is a code with length* n *and rank* k *is a code such that each codeword has length* n*, and the space formed by the codewords under usual vector addition is a rank* k *subspace of the vector space* $\mathbb{F}_2^n$*. When we think of the minimum distance of such a code, we can appeal to the idea of weight. The weight of a codeword is the number of non-zero entries it has. It is not hard to see that the minimum distance of this code is equivalent to the minimum weight non zero codeword, as the codewords form a complete subspace.*

Definition 6.9 (Expander codes)**.** *Let* G *be a* $(2, d)$*-regular bipartite graph with* n *vertices on one side and* $\frac{dn}{2}$ *vertices on the other. We call the* $\frac{dn}{2}$ *vertices the variables, the set of which is denoted* $A = \{v_1, ..., v_{\frac{dn}{2}}\}$*, and the* n *vertices constraints, with set denoted* $B = \{c_1, ..., c_n\}$*. Then define a function* $f(i, j)$ *such that for some* $c_j \in B$*, the set* $\{v_{f(j,1)}, ..., v_{f(j,d)}\}$ *are the* d *neighbours of* c_j *that are in* A*. We then take some fixed length linear block code* $\mathbf{C}'$ *with codeword length* d*. Then the expander code* $\mathbf{C}(G, \mathbf{C}')$ *is the code with words of length* n *where words* $(x_1, ..., x_n)$ *are in the code if and only if* $(x_{f(i,1)}, ..., x_{f(i,d)})$ *is a codeword of* $\mathbf{C}'$ *for all* $1 \leq i \leq n$*.*

The above definition is complicated and can take a couple of reads to fully parse, but essentially the idea is to build arbitrarily long codewords that have fixed relative distance by constraining the possible codewords such that subsets of the letters of the word form codewords in a good code of fixed length, and in this case the graph gives us a mechanism by which to apply these constraints to the letters of the resultant codewords. The particular construction that we use is to take an expander graph G on n vertices (not necessarily bipartite), form the edge-vertex incidence graph from G, call it G', and then let the $\frac{dn}{2}$ vertices on one side be the variables, and let the n vertices on the other side be the constraints. We will see that the distance of a code generated by this method is in fact dependent on how good an expander the graph G is.

Lemma 6.1 (Alon-Chung [2])**.** *Let* $G = (V, E)$ *be a* d*-regular graph with adjacency matrix* A*, on* n *vertices. Let* λ *be the second largest eigenvalue as usual. Take some subset* $S \subset V$*, with the stipulation that* $|S| = \alpha \cdot n$ *for some* $0 < \alpha < 1$*. Let* $e(S)$ *denote the number of edges in* G *that have both endpoints in* S*. Then*

$$e(S) \leq \frac{dn}{2}(\alpha^2 + \frac{\lambda}{d}\alpha(1 - \alpha))$$

Proof. We begin by constructing a vector $v \in \mathbb{R}^n$ defined such that $v_i = \frac{-1}{|S|}$ if $i \in S$, and $v_j = \frac{1}{n-|S|}$ if $j \notin S$. It is clear that

$$\sum_{i=1}^{n} v_i = |S| \times \frac{-1}{|S|} + (n - |S|)\frac{1}{n - |S|} = -1 + 1 = 0$$

We consider then that G is d-regular, hence its largest eigenvalue is d, and it is easy to see that the associated eigenvector is the all ones vector. To see that, let w be the eigenvector and consider that Aw is the vector of all d's as each row in A has exactly d ones and $n-d$ zeros, but this is the same as dw. Therefore, v is orthogonal to w. This implies that

$$|Av \cdot v| \leq \lambda(v \cdot v)$$

We then observe that

$$Av \cdot v = 2 \sum_{(i,j)\in E} v_i v_j = d \sum_{i-1}^{n} (v_i)^2 - \sum_{(i,j)\in E} (v_i - v_j)^2$$

It is not hard to see from the definition of v_i that

$$\sum_{i=1}^{n} v_i^2 = |S| \frac{1}{|S|^2} + (n - |S|) \frac{1}{(n-|S|)^2} = \frac{1}{|S|} + \frac{1}{n-|S|}$$

When we consider

$$\sum_{(i,j)\in E} (v_i - v_j)^2$$

we can easily see that if $i, j \in S$, then $v_i - v_j = 0$ and if $i, j \in V \setminus S$ then $v_i - v_j = 0$ also. So we are only interested in the edges between S and $V \setminus S$, so let e' denote the number of such edges. Then for each such edge, we have

$$(v_i - v_j)^2 = (\frac{-1}{|S|} - \frac{1}{n-|S|})^2 = (\frac{1}{n-|S|} + \frac{1}{|S|})^2$$

thus giving

$$|e'(\frac{1}{|S|} + \frac{1}{n-|S|})^2 - d(\frac{1}{|S|} + \frac{1}{n-|S|})| \leq \lambda(\frac{1}{|S|} + \frac{1}{n-|S|})$$

We can then substitute $|S| = \alpha n$ for

$$|e'(\frac{1}{\alpha n} + \frac{1}{(1-\alpha)n})^2 - d(\frac{1}{\alpha n} + \frac{1}{(1-\alpha)n})| \leq \lambda(\frac{1}{\alpha n} + \frac{1}{(1-\alpha)n})$$

$$|e'(\frac{1}{\alpha(1-\alpha)n})^2 - d(\frac{1}{\alpha(1-\alpha)n})| \leq \lambda(\frac{1}{\alpha(1-\alpha)n})$$

$$|e' - d(\alpha(1-\alpha)n)| \leq \lambda(\alpha(1-\alpha)n)$$

We can then characterise e' in terms of $e(S)$ by using the knowledge that G is a d-regular graph and thus

$$2e(S) + e' = d|S| = d\alpha n \Rightarrow e' = d\alpha n - 2e(S)$$

so we substitute in for

$$|d\alpha n - 2e(S) - d(\alpha(1-\alpha)n)| \leq \lambda(\alpha(1-\alpha)n)$$

thus

$$|e(S) - \frac{1}{2}d\alpha^2 n| \leq \frac{1}{2}\lambda(\alpha(1-\alpha)n)$$

which finally gives

$$e(S) \leq \frac{dn}{2}(\alpha^2 + \frac{\lambda}{d}\alpha(1-\alpha))$$

as required. □

At first glance, it is not clear why this lemma helps us, but we can now move to the following theorem, which establishes a connection between the minimum relative distance of an expander code and the spectrum of the adjacency matrix of the graph upon which the code is based.

Theorem 6.1 (Sipser-Spielman [23])**.** *Let $\mathbf{C}'$ be some linear code with block length d, rate r and minimum relative distance δ_0. Let G be a d-regular graph with adjacency matrix A, which has second largest eigenvalue λ as usual, and let G' be the edge-vertex incidence graph of G. Then then code $\mathbf{C}(G', \mathbf{C}')$ has minimum relative distance greater than*

$$\left(\frac{\delta_0 - \frac{\lambda}{d}}{1 - \frac{\lambda}{d}}\right)^2$$

Proof. Since there are $\frac{dn}{2}$ variables in G' and n constraints, we can think of these in terms of edges and variables in G in order to apply Lemma 6.1, such that if we take some subset S of the variables, then if

$$|S| = \frac{dn}{2}(\alpha^2 + \frac{\lambda}{d}(1-\alpha)\alpha), 0 < \alpha < 1$$

the lemma implies that at least αn of the constraints will be neighbours of elements of S in G', as a subset of vertices with size αn in G will have this many edges in the induced subgraph. As the variables side of G' is 2-regular, each variable has 2 neighbours, and thus we can calculate the average number of variables per constraint as follows:

$$\frac{2(\frac{dn}{2}(\alpha^2 + \frac{\lambda}{d}(1-\alpha)\alpha)}{\alpha n} = d(\alpha + \frac{\lambda}{d}(1-\alpha))$$

Then, by the definition of an expander code, the codewords are only valid if the subwords generated by each constraint are valid codewords in $\mathbf{C}'$, therefore if

$$d(\alpha + \frac{\lambda}{d}(1-\alpha)) < d\delta_0$$

the corresponding word in $\mathbf{C}'$ will be below the minimum distance of $\mathbf{C}'$ and thus be invalid, hence the word of relative weight

$$\alpha^2 + \frac{\lambda}{d}(1-\alpha)\alpha$$

is not a valid codeword in $\mathbf{C}(G', \mathbf{C})$ (the factor of $\frac{dn}{2}$ disappears as we are referring to relative weight). We can solve the inequality above for

$$\alpha < \frac{\delta_0 - \frac{\lambda}{d}}{1 - \frac{\lambda}{d}}$$

Now clearly for $0 < \alpha < 1$, we have

$$\alpha^2 < \alpha^2 + \frac{\lambda}{d}(\alpha - \alpha^2)$$

thus the minimum relative distance is necessarily larger than α^2. We can then observe that this implies that the minimum relative weight of a codeword in $\mathbf{C}(G', \mathbf{C}')$ is greater than

$$\left(\frac{\delta_0 - \frac{\lambda}{d}}{1 - \frac{\lambda}{d}}\right)^2$$

as required. □

This equips us with the tools to show that the expander graphs we constructed earlier can create asymptotically good codes:

Corollary 6.1. *Expander codes based on the $G^{p,q}$ are asymptotically good codes.*

Proof. Suppose we are constructing a code based on $\mathbf{C}'$ such that the obtained code is $\mathbf{C}(G^{p,q'}, \mathbf{C}')$. Then let r be the rate of $\mathbf{C}'$. We can see that the constraints in $G^{p,q'}$ impose $(1-r)d$ restrictions on the variables, as each constraint has degree d, and r of these are bits of information, therefore $(1-r)d$ are the redundant bits that enforce the rate of the obtained code. Using this, we can count the number of linear restrictions that are applied to the variables by the redundancy of $\mathbf{C}'$, which we see is $n(1-r)d$ as there are n constraints. There are $\frac{dn}{2}$ variables, so we see that

$$\frac{dn}{2} - n(1-r)d = dn(r - \frac{1}{2})$$

are the number of bits that correspond to information in $\mathbf{C}(G^{p,q'}, \mathbf{C}')$, and thus the relative rate is

$$\frac{dn(r - \frac{1}{2})}{\frac{dn}{2}} = 2r - 1$$

thus

$$r(\mathbf{C}(G^{p,q'}, \mathbf{C}')) = 2r - 1$$

and therefore is bounded above 0, regardless of how large the graphs $G^{p,q}$ become.

For the minimum relative distance, we observe that by Theorem 6.1, for a code $\mathbf{C}'$ with minimum relative distance δ_0, the minimum relative distance of $\mathbf{C}(G^{p,q'}, \mathbf{C}')$ is

$$\left(\frac{\delta_0 - \frac{\lambda}{d}}{1 - \frac{\lambda}{d}}\right)^2$$

where d is the degree of $G^{p,q}$ and λ is the second largest eigenvalue. We then observe that the construction of $G^{p,q}$ is such that $d = p + 1$ and $\lambda \leq 2\sqrt{p}$, thus we can observe that

$$p \to \infty \Rightarrow \frac{2\sqrt{p}}{p+1} \to 0 \Rightarrow \frac{\lambda}{d} \to 0$$

In fact, we can easily observe the following:

$$2\sqrt{p} = p + 1 \Rightarrow (p-1)^2 = 0$$

such that for all values of $p > 1$, $\frac{2\sqrt{p}}{p+1}$ is decreasing as p increases. As p is a prime congruent to 1 mod 4, the smallest possible value of p is 5, and so if we wanted a formal lower bound for the minimum relative distance of $\mathbf{C}(G^{p,q'}, \mathbf{C}')$, we could compute this value explicitly, but this would involve the introduction of an explicit linear block code $\mathbf{C}$ and we have already demonstrated that $\mathbf{C}(G^{p,q'}, \mathbf{C}')$ have relative rate and distance bounded above 0, regardless of the size of the $G^{p,q}$ used to construct them, and thus it is shown that expander codes based on Ramanujan graphs are asymptotically good codes. □

7 Conclusion

We have demonstrated the explicit construction of an infinite family of Ramanujan graphs, and further shown that Ramanujan graphs can be used to develop asymptotically good error correcting codes. Therefore, we have developed an explicit mechanism by which to develop arbitrarily large asymptotically good error correcting codes, some applications of which we have mentioned. There are other interesting applications of expander graphs to topics of computer science, such as pseudo-random generators, in fact, the interested reader may be tempted to look into results establishing an equivalence between certain pseudo-random generators, expander graphs and error correcting codes.

I would like to thank my advisors Michael Mitzenmacher and Lauren Williams for their assistance in writing this thesis; they have both been indispensible resources to me throughout. I would also like to thank Hector Pasten-Vasquez for first introducing me to Ramanujan graphs and the main paper of Lubotzky, Phillips and Sarnak [16].

References

[1] Alon, Noga. "Eigenvalues and expanders." Combinatorica 6.2 (1986): 83-96.

[2] Alon, Noga, and Fan RK Chung. "Explicit construction of linear sized tolerant networks." Discrete Mathematics 72.1-3 (1988): 15-19.

[3] Ballantine, Cristina, et al. "Explicit construction of Ramanujan bigraphs." Women in Numbers Europe. Springer, Cham (2015): 1-16.

[4] Charles, Denis X., Kristin E. Lauter, and Eyal Z. Goren. "Cryptographic hash functions from expander graphs." Journal of Cryptology 22.1 (2009): 93-113.

[5] Chung, F. R. K. "Laplacians of graphs and Cheeger inequalities". Bolyai Mathematical Society. (1996): 157-172

[6] Gaster, Gabriel. "Acting Freely." (2006).

[7] Hamming, Richard W. "Error detecting and error correcting codes." The Bell System Technical Journal 29.2 (1950): 147-160.

[8] Hecke, Erich. "Analytiscbe arithmetik der positiven quadratic formen". Collected Works (1959): 789-898.

[9] Hirschhorn, Michael. "A Simple Proof of Jacobi's Four-Square Theorem". Proceedings of The American Mathematical Society. (1987): 436-438.

[10] Hoeve, H., J. Timmermans and L. B. Vries. "Error correction and concealment in the Compact Disc system." Philips tech. Rev. 40.6 (1982): 166-172.

[11] Holt, Derek. University of Warwick. http://homepages.warwick.ac.uk/ mareg/lnqsimp.pdf

[12] Hoory, Shlomo, Nathan Linial, and Avi Wigderson. "Expander graphs and their applications." Bulletin of the American Mathematical Society 43.4 (2006): 439-561.

[13] Igusa, Jun-Ichi. Fibre Systems of Jacobian Varieties: (III. Fibre Systems of Elliptic Curves). American Journal of Mathematics, vol. 81, no. 2, (1959): 453476

[14] R. Impagliazzo and D. Zuckerman, "How to recycle random bits," 30th Annual Symposium on Foundations of Computer Science. (1989): 248-253.

[15] Key, J. D. "Some error correcting codes and their applications." Applied Mathematical Modeling: A Multidisciplinary Approach (1999).

[16] Lubotzky, A., Phillips, R. & Sarnak, P. "Ramanujan graphs." Combinatorica (1988) 8: 261.

[17] Lubotzky, A., et al. Hecke Operators and Distributing Points on the Sphere I. Communications on Pure and Applied Mathematics, vol. 39, no. S1, (1986): S149S186.

[18] Mališsev. " On the representation of integers by positive definite forms". Proceedings of the Steklov Institute of Mathematics. 65 (1962)

[19] A. Margulis, G. (1988). "Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators." Problems of Information Transmission (1988): 24.

[20] Murty, M. Ram. "Ramanujan graphs." Journal-Ramanujan Mathematical Society 18.1 (2003): 33-52.

[21] Schellwat, Holger. "Network Construction by Group Representation". University College of Orebro. (1994)

[22] Shannon, Claude Elwood. "A mathematical theory of communication." Bell System Technical Journal 27.3 (1948): 379-423.

[23] Sipser, Michael, and Dan Spielman. "Expander codes". IEEE Transactions on Information Theory 42.6 (1996): 1710-1722