EXPANDER GRAPH THEORY

ABSTRACT

This research paper explores expander graph theory.The first section of the paper focuses on fundamental graph theory and includes basic definitions and proofs that relate to expander graphs and Cayley graphs. The second section of the paper focuses on Cayley graphs and delves into some basic group theory.

The main section contains information about expander graph theory. This section first looks at what the properties and uses of expander graphs are, then moves onto the edge expansion ratio and examples of good and bad expanders. We then further explore the explicit construction of expanders and the applications of spectral graph theory in approximating the edge expansion ratio.

The final section contains an expansion of a Cayley graph family. This expander family is a family created for this paper with inspiration from Wood's paper (referenced in bibliography).

INTRODUCTION

The research of expander graphs is one of the most exciting and fascinating parts of graph theory. The study of expander graphs has only developed in the past 5 decades. This emergence has been driven from many fronts, including its uses in algorithms, coding and cryptography. All of these fields have benefited tremendously from expanders, with the most obvious example coming from the proof of the PCP Theorem in the 1990s. Furthermore, at this moment in time we are only scratching the surface of the potential of expander graphs. Expander graphs could be the tool that revolutionizes computer science, and cryptography with its applications in error-correcting code. In the future, expander graphs could construct error-correcting code that perfectly protects information against noise, allowing for advances in discrete mathematics and computer science that at this moment seem unimaginable. This paper will attempt to unveil the complexities of expander graphs and provide a foundation for future research on the subject.

1. FUNDAMENTAL GRAPH THEORY

To introduce the concept of expander graph theory, we need more knowledge on the field of graph theory. This section will explain essential concepts in graph theory that must be understood in order to lay a foundation for introducing expander graphs.

A *graph* in its most abstract form is a set of objects and a set of connections between groups or pairs of these objects (Maricq, 2014).

Definition 1.1: A *graph* is an ordered pair **G** = **(V, E)** where **V** is a set of *vertices* and **E** is a set of subsets **e** ⊂ **V** of *cardinality* 1 or 2 (Maricq, 2014).

Cardinality is the number of elements in a set. The *order* of a graph **G** is the *cardinality* of its vertex set, and the *size* of a graph is the *cardinality* of its edge set.

A *directed graph* or *digraph*, **G = (V, E),** is a graph in which the elements of **E** or the edges have a direction. This means that they go from one vertex to another in one way. An *undirected graph*, **G = (V, E)**, is a graph in which the elements of **E** or the edges have no direction. This means that they don't go from one vertex to another. A graph is *simple* if it is undirected and there are only one or fewer edges between two vertices. Further, there should be no edges that loop and connect any vertex to itself. On the other hand, a *multigraph*, **G = (V, E)**, allows for multiple edges between vertex pairs and allows for a looping edge that connects a vertex to itself.

Another type of graph is a *bipartite graph.* A *bipartite graph* is a graph in which the elements of the vertex set can be partitioned into two separate sets in a way such that every edge in the graph connects a vertex in one set to the other. The two sets are the *partite sets* (Harris, 2008).

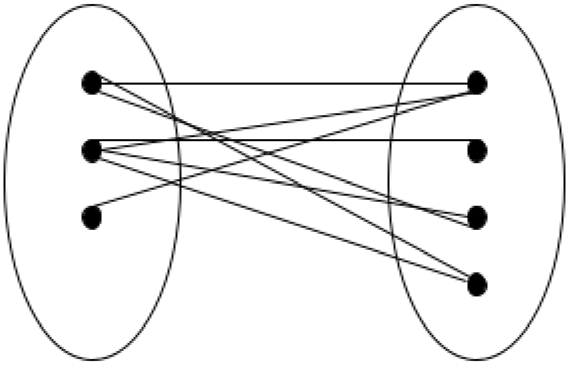

Figure 1.1: An example of a bipartite graph.

Another concept in graph theory is *regularity*.

Definition 1.2: A graph **G = (V, E)** is considered ***k**-regular* if all the vertices in the graph are contained in exactly **k** edges, not including loops. For a directed graph to be **k**-regular, it must also have an equal number of edges leaving and approaching each vertex.

The degree of a vertex is the number of edges in which the vertex is contained in. For example, the degree of $\mathbf{v_n}$ would be the number of edges in **E** attached to $\mathbf{v_n}$ when $\mathbf{v_n} \in \mathbf{V}$. For a **k**-regular graph, the degree of each vertex is **k** and so it can be said that the graph itself has degree **k**.

If a **k**-regular graph is simple and every vertex in the graph is connected to every other vertex, or **n-1** vertices, the graph is called a *complete graph* on **n** vertices. This graph is denoted by $\mathbf{K_n}$.

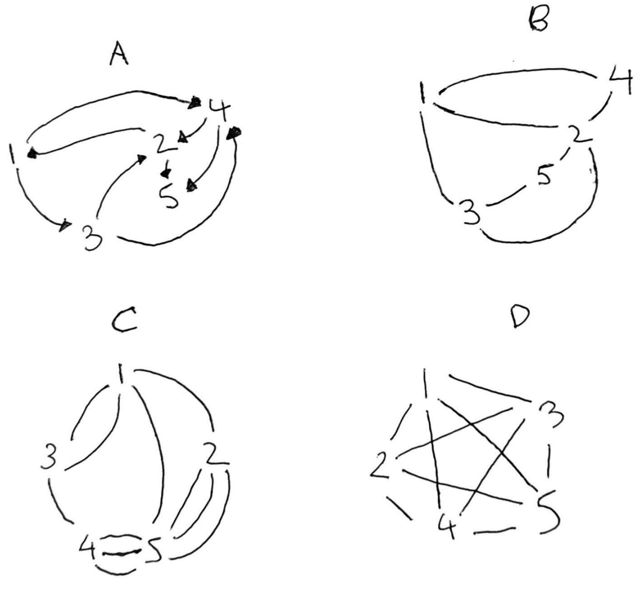

Figure 1.2: A is a directed graph, B is a simple graph, C is a multigraph, D is a ***k****-regular graph (****k****=4).*

A *connected graph* is a different class of graph. Every pair of vertices in an *undirected connected graph* is linked together by one or more edges. *Vertex connectivity* and *edge connectivity* are two characteristics of a connected graph. A connected graph's vertex connectivity is the smallest set of vertices that must remain for **G** to remain connected. **K(G)** can be used to denote vertex connectivity. The number of edges required to make a connected graph G disconnected, or the edge connectivity, is denoted by **λ(G)** (Buckley, 2013).

In graph theory, the concept of a *walk* on graph refers to an alternating sequence of vertices and edges. A vertex is where a walk always finishes and starts. In addition, every edge in a walk must directly connect the vertex before and after the edge in the sequence. If the starting and finishing vertices of a walk are the same vertex and the sequence loops, the walk is said to be *closed*. If the finishing and starting vertices of a walk are different, it is said to be *open* (Maricq, 2014).

In a connected graph **G**, the *distance* between vertex **u** and vertex **v** is defined as the *length* (number of edges) of the shortest **u-v** path in **G**. This distance is represented by **dG(u,v)**. In a disconnected graph, the distance between any two vertices that are in different components is infinite (Harris, 2008).

The maximum distance between a vertex **v** to all other vertices is called the *eccentricity* of the vertex **v**. It is denoted **ecc(v)**. The *radius* of **G**, denoted **rad(G)**, is the minimum eccentricity of any graph vertex in a graph. Similarly, the *diameter* of **G**, denoted **diam(G)**, is the value of the greatest eccentricity of any graph vertex in a graph. The *center* of the graph **G** is the set of vertices, **v**, such that **ecc(v) = rad(G)**. Another similar property of the graph is the graph's *girth*. The *girth* of a graph **G(V, E)** is the length of the shortest cycle in the graph. The *eccentricity* of the vertex **v** is the greatest distance between that vertex and every other vertex. Its symbol is **ecc(v)**. The minimum eccentricity of any vertex in a graph is known as the *radius* of **G**, indicated by the symbol **rad(G)**. A graph's *diameter*, given by the symbol **diam(G)**, is the value of any vertex's maximum eccentricity. The vertex or set of vertices, **v**, such that **ecc(v) = rad(G)**, is the centre of the graph **G**. The length of the shortest cycle in a graph **G(V, E)** is referred to as the graph's *girth*.(Harris, 2008). A *cycle* in a graph is a non-repeating walk in which only first and last vertices are the same.

Proposition 1.1: For all connected graphs **G = (V, E)**, **rad(G) ≤ diam(G) ≤ 2 rad(G)**.

Proof 1.1: By their definitions, we know that **rad(G) ≤ diam(G)**. So the part that we need to prove is only the second inequality. To prove this, let **u** and **v** be vertices in the vertex set such that the distance between the two vertices is equal to the diameter. Additionally, let **c** be a vertex in the center of the graph. Then:

$$\mathbf{diam(G) = d(u,v) \leq d(u,c) + d(c,v) \leq 2\ ecc(c) = 2\ rad(G).}$$

Proposition 1.2: A graph **G** cannot be connected if **G** is a graph of order **n** and size strictly less than **n − 1**.

Proof 1.2: Using induction we can prove that the minimum number of edges needed to have a connected graph is **n − 1**.

Base Case: $\mathbf{K_1}$, has 0 edges and is connected.

Suppose **G** of order **n** requires minimum **n − 1** edges to be connected. Consider such a connected graph with **n - 1** edges, and consider **G + v**, **G** with an additional vertex. If we add no edges, then the new graph is disconnected, so we must add one edge. Since **G** is connected by the inductive hypothesis, adding an edge between any **u** ∈ **G** and **u** will make **G + v** connected, since for the path **u** to **x,** where **x** ∈ **G**, the path **+v** is the path **v** to **x**.

Proposition 1.3: A graph with a closed odd walk must always also have an odd cycle.

Proof 1.3: This can also be proved by induction. This can be proven by induction on the length of the odd closed walk.

Base Case: The length 3 odd closed walk is just a length 3 cycle.

Suppose odd closed walks with lengths up to **2n - 1** contain odd cycles.

Let $\mathbf{W = v_1 , v_2 , \ldots, v_{2n+1} = v_1}$ be a length **2n + 1** closed walk. If no vertices in the walk repeat, then we are done, the odd walk is an odd cycle. Otherwise, let **L** be the smallest number not **1** such that $\mathbf{v_1}$ repeats, and let $\mathbf{v_1 = v_k}$, where $\mathbf{l < k}$. Then we have two closed walks in **W**, $\mathbf{v_1, \ldots, v_L = v_k , \ldots, v_{2n+1} = v_1}$ and $\mathbf{v_L , \ldots, v_k = v_L}$**.** The lengths of these two walks must add up to **2n + 1**, thus one of them must be an odd length closed walk, which by the inductive hypothesis must contain an odd length cycle.

In this paper, we are looking at expander graphs and expander graph theory. To simplify some of the concepts and relations in expander graph theory we will use and work with **k**-regular undirected multigraphs for most of the paper. Expander graphs are normally looked at as a family of graphs rather than individual graphs, and so we will be working with families of graphs. Simply put, a *family* of graphs is a set of graphs which share one or more characteristics (Maricq, 2014).

In expander graph theory and graph theory, the terms *dense* and *sparse* can be used to describe graphs. In general terms, a graph with many edges is *dense*, while a graph with fewer edges is *sparse*. A formal definition, which further applies this idea of density to families of graphs, is as follows:

Definition 1.3: Consider a sequence of graphs $\{G_i\}_{i \in N} = \{(V_i, E_i)\}_{i \in N}$ where the number of vertices in **G_i** increases as **i** increases. Then:

- If $|E_i| = \Theta(|V_i|^2)$ for all **i**, then $\{G_i\}_{i \in N}$ is a family of dense graphs.
- If $|E_i| = \Theta(|V_i|)$ for all **i**, then $\{G_i\}_{i \in N}$ is a family of sparse graphs.

Here **Θ** represents the asymptotic behavior of $|E_i|$ (Maricq, 2014)**.**

This will mean that the function will not deviate as **i** tends to infinity, and the trends will stay constant. This definition can be interpreted to mean that for a graph to be classified as *sparse* in expander graph theory, the number of vertices in the graph should be at least in the same order as the number of edges in the graph. And for a graph to be classified as *dense* in expander graph theory, the number of vertices in the graph squared should be in the same order as the number of edges in the graph (Preiss, 1999).

2. CAYLEY GRAPHS

In the second section of this paper we will look into *Cayley graphs*. A *Cayley graph* is a graph that contains the structure of a *group*. A *group* is a set of objects with a rule of combination. Cayley Graphs are very useful when looking into expander graph theory as the expansions in expander graphs are easier to define with a base knowledge of what Cayley Graphs are and how they function.

Definition 2.1: A Cayley graph is a graph **C(G,S)**, where **G** is a group and **S** is the generating set of the group.

To define what the *generating set* of a *group* is, let **G** be a group and **S** ⊆ **G** (**S** is a *subset* of **G)**. A set **S** is a *subset* of group **G** if all the objects in set **S** are also in the group **G**. The smallest *subgroup* of **G** that also contains all the objects in **S** is **(S)**. A *subgroup* of a group is a subset of the group that also forms a group with the same rule of combination. It's like a group within a group.

Group **G**'s subgroups are all included in set **H**. The intersection of all subgroups that contain **S** in **H** of **G** is hence the same way **(S)**. Additionally, all of the products of the integer powers of **S**'s constituent elements make up the subgroup **(S)**. These are also known as the *words* in **S**. When **(S) = G, S** is referred to as a *generating set* of **G**. Thus, we state that **S** generates **G**. (Sherman-Bennet, 2016). In other words, if every component of **G** can be represented as a word in **S**, then **S** generates **G**. *Generators* are the components that make up a generating set.

Case 2.1: The group of *units* **U(Z_9)** is the group of all integers relatively prime to 9 under multiplication **mod 9 (U9 = {1, 2, 4, 5, 7, 8})**. All arithmetic here is done modulo **9**.

Here, **7** is not a generator of **U(Z_9)**, since
{7^i mod 9 | i ∈ N }= { 7, 4, 1},
while **2** is, since
{2^i mod 9 | i ∈ N }= {1, 2, 4, 5, 7, 8}.

Case 2.2: For **n > 2** the symmetric group of degree **n** is not *cyclic*. This means that unlike Case 2.1, the group cannot be generated by a single element. For groups that are not cyclic we have to find different ways to generate the group. For this group, the group is generated by the two permutations **(1 2)** and **(1 2 3 ... n)** (Wikipedia, 2022).

A symmetric group on a set is a group containing all one-to-one and onto functions on the set, from the set to itself, with the rule of combination being the composition of the functions.

For example, using *left-to-right composition* when we take **n = 3** for the symmetric group of degree **n**:

e = (1 2)(1 2)
(1 2) = (1 2)
(1 3) = (1 2)(1 2 3)
(2 3) = (1 2 3)(1 2)
(1 2 3) = (1 2 3)
(1 3 2) = (1 2)(1 2 3)(1 2).

Another way to look at this is to consider how the Cayley graph **C(G,S)** is built for a group **G** and a subset **S** of that group. If and only if there is a **s** ∈ **S** such that **g_2 = g_1** ∘ **s**, then **V** includes a vertex **v(g)** associated with each element **g** ∈ **G**, and **E** contains the directed edge (**v(g_1) , v(g_2)**) (Petit, 2009) (Maricq, 2014).

For example, this would be the *Cayley Graph* **C(G,S)** where **G** is the *additive group modulo **8*** and **S = {1,2}**. *Additive group modulo 8* is the group where you take all integers under **8** and add each integer to all the integers under **8**. All the answers must also be under **8** as the group is *modular* so will *wrap around*. This is shown by the table:

+	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7	0
2	2	3	4	5	6	7	0	1
3	3	4	5	6	7	0	1	2

4	4	5	6	7	0	1	2	3
5	5	6	7	0	1	2	3	4
6	6	7	0	1	2	3	4	5
7	7	0	1	2	3	4	5	6

The integers **1** and **2** are the *generating set* of the *Cayley Graph* as both are the only integers that can be used as *words* in **S** to generate the *group*.

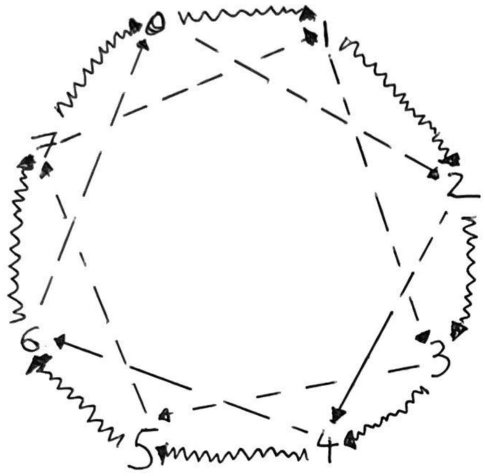

Figure 2.1: An example of a Cayley Graph C(G,S) where G is the additive group modulo 8 and S = {1,2}.

Certain aspects of the generating set **S** of **G** determine certain properties that **C(G,S)** will hold:

- The graph is **k-regular** when **|S| = k** when only considering out-degree or in-degree in a directed graph.
- If the generating set of the Cayley graph is symmetric, the graph is undirected. A set is symmetric if $\mathbf{s} \in \mathbf{S} \Longleftrightarrow \mathbf{s}^{-1} \in \mathbf{S}$ (only applicable for graphs which are not cyclic).

The elements in the generating set, **S**, are referred to as the Cayley graph's *graph generators*. A *vertex-transitive* graph is a graph which has the property of there being some automorphism such that, given any two vertices the first vertex will map to the next vertex. Cayley graphs are *vertex transitive* graphs as for any $\mathbf{g_1}$, $\mathbf{g_2} \in \mathbf{G}$ the mapping $\mathbf{v_x} \rightarrow \mathbf{v_{g(2)g(1)^{-1}x}}$ is a graph automorphism that sends $\mathbf{g_1}$ to $\mathbf{g_2}$ (Maricq, 2014).

Case 2.3: In the Cayley Graph C(G,S) where G is the additive group modulo 8 and S = {1,2}, when each vertex is mapped to the vertex **+1**, the edge and vertex connectivity are preserved and the new graph is an automorphism of order **1** and is **{(),(0,..,7)}**. This means that this Cayley Graph is vertex transitive.

An *automorphism* of a graph is a type of mapping of a graph in which the edge and vertex connectivity remains constant and the graph is mapped onto itself. An automorphism could be the mapping operation or the graph created by the mapping operation.

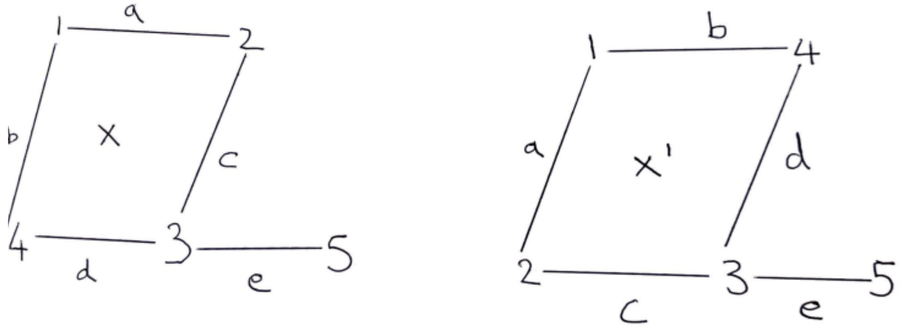

*Figure 2.2: For example, the group **Aut(X)** is of order **2** and is {(),(2,4)}.This automorphism has been applied on Graph **X**, and Graph **X'** is an automorphism of **X**.*

Theorem 2.1: Cayley's theorem is that, there is a group of permutations for every group to be *isomorphic* to.

A group *isomorphism* is a function that sets up a one-to-one correspondence between the elements of two separate groups. The function does this in a way that follows a series of group operations. Two groups are called *isomorphic* when there exists an isomorphism between them.

Theorem 2.2: Cayley's better theorem is a slightly different theorem that states every finitely produced group can be *faithfully* represented as a *connected, directed, locally finite* graph's *symmetry* group. A *faithful* representation of a group is a linear representation (in a vector space) in which the group's components, such as its vertices and edges, are shown and represented by unique linear mappings. Here, locally finite means that for any vertex **e**, there are finitely many edges $\mathbf{v_1}, \ldots, \mathbf{v_n}$ with $\mathbf{v_k} \in$ Ends **(e)**.

We will now prove Cayley's better theorem using Nicol's proof:

Proof 2.1: Let G be a finitely generated group with generating set $\mathbf{S = \{s_1, \ldots, s_n\}}$. We can prove this theorem by constructing a graph, **C(G,S)**. The vertices of **C(G,S)** will be the elements of **G**. For each $\mathbf{g \in G, s \in S}$, make an edge labeled **s** from the vertex labeled g to the vertex labeled $\mathbf{g_s}$. Since **G** is finitely generated, this graph is locally finite. Since **S** generates **G**, this graph is connected. **G** is directed. Let **G** act on the graph by left multiplication. In every group **G**, *left multiplication* is an action of **G** on **G**: $\mathbf{g \cdot x = gx}$ for all **g, x** in **G**. This action is free and transitive. That is, for any $\mathbf{g \in G}$, **g** will send the vertex labeled **h** to the vertex labeled $\mathbf{g_h}$ (Nicol, 2008).

3. EXPANDER GRAPHS

Expander graphs are a type of graph that hold the seemingly contradictory property of being not only *well-connected* but also being *sparse*, like a random graph. Unlike random graphs, they are also *explicitly constructible*. Almost every graph is an expander, but what differentiates good expanders from other graphs is their *edge expansion ratio*. This will be discussed later in the section.

Since the study of expander graphs started in the 1970s, they have become an integral part of both discrete mathematics and theoretical computer science. This is because expander graphs have a very wide range of uses in both subjects, due to the properties outlined above. For example:

- Expander graphs can reduce the need for randomness. Good expanders can be used to improve the desired probability to make a probabilistic algorithm work. They do this by reducing the amount of random bits needed for the algorithm to work (Nielson, 2005).

- Expander graphs are also able to find helpful error-correcting codes. This is because good expanders can be used to find error-correcting codes that can protect information against noise. Surprisingly, the error-correcting codes that expanders construct are efficiently encodable and decodable. The code constructed by expanders even have a non-zero rate of transmission. This is surprising for information theorists because finding codes with these properties used to be nigh impossible. In fact, finding error-correcting code with those properties was one of the main goals of coding theory for years after Shannon's work on coding and information theory back in the 1940s (Nielson, 2005).

- An additional fresh proof for the PCP Theorem was provided using expander graphs. The PCP Theorem is among the most significant in computer science. According to this theorem, a randomised polynomial-time proof verifier exists for any language **L** in **NP**. This verifier only has to examine a certain number of bits in a claimed proof that t $\mathbf{x} \in \mathbf{L}$ **or** $\mathbf{x} \notin \mathbf{L}$ in order to decide (with a high chance of success) whether the proof is accurate or not. Although this theorem was initially "given a result" in the 1990s, expanders recently provided a fresh proof for it. (Nielson, 2005).

There are lots of definitions for expander graphs. One of these definitions is:
Definition 3.1: An expander graph is a graph in which every *subset* **S** of vertices is connected to many vertices in the complementary set $\bar{\mathbf{S}}$ of vertices (Tao, 2011).
So for example, a graph with no edges is not considered an expander graph, as if you split the graph into a set of vertices, the vertices in this set will not be connected to any vertices in the complementary set of vertices. This will be constant no matter the set of vertices you take, therefore this graph cannot be considered an expander.

Useful properties of expander graphs not outlined above include having a *high chromatic number* and a low *diameter*. As discussed in our Pioneer course, the chromatic number of a graph is the minimum number of colors needed to color the vertices in such a way that no two adjacent vertices have the same color.

However, that was just one definition of what an expander graph is. In fact, expander graphs have been previously defined in a diverse range of other ways. Something to keep in mind is that every time we consider expander graphs, we only consider undirected graphs. Furthermore, we generally define the graphs in the normal sense of **G = (V, E)** or sometimes in the Cayley graph sense of **C(G, S)** (here **S** does not mean the same thing as the **S** previously used in definition 3.1).

Another way of thinking about the properties of an expander graph is, consider a graph **G = (V, E)**. Now take the magnitude of set of vertices to go to infinity, or **|V|** → ∞, and the graph is also always **k**-regular (as we covered in the first section). This will mean that as **|V|** increases and goes to infinity and **k** stays constant, the graph becomes *sparse*. The property of being *highly connected* can be considered in two ways. One way is in terms of edge expansion, and the other is in terms of vertex expansion (explored later in the section). Informally, expander graphs are considered to be highly connected when all the subsets, **S** ⊆ **V : 0 < |S| ≤ |V|/2**, have many edges (edge expansions) or vertices (vertex expansion) on its boundary (O’Donnell, 2013).

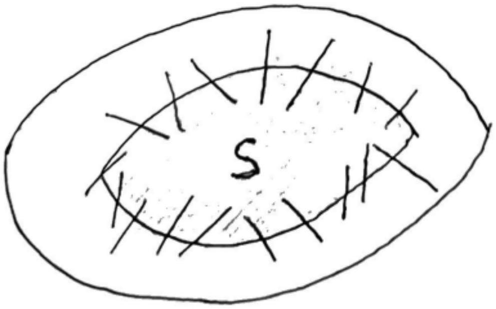

Figure 3.1: An example of a way of picturing expansion. The lines coming from ***S*** *represent edges from vertices in* ***S*** *to those in* $\overline{\textbf{S}}$*(complement). The two ways of picturing this are, count the number of edges leaving* ***S*** *or count the number of neighbors that* ***S*** *has.*

Informally, *vertex expansion* considers the number of vertices in the *complement* of **S** that can be reached with one step from a vertex in **S**. *Edge expansion* on the other hand considers the number of edges going from **S** to the *complement* of **S**.

When researching expander graphs, one of the most important properties to consider is the *expansion parameters* of an expander graph. In this paper we will only look at the *edge expansion ratio* in the *expansion parameters*, but there is also an *expansion parameter* for vertex expansion. Whenever we talk about *expansion parameters* in this paper we are talking about the *expansion parameter* for edge expansion. This is as the *edge expansion ratio* is more useful for when we apply expander graphs in cryptography.

The *edge expansion ratio* of a graph **G = (V, E)** on **n** vertices is given by:

$$h(G) = \min_{S \subset V : 1 \leq |S| \leq \frac{n}{2}} \frac{|\partial S|}{|S|}.$$

To understand this formula, take a graph **G = (V, E)**, that has **n** vertices. First, consider a subset **S** of the vertices **V** in **G**, and its **S¯** (complement of **S**). New notation in this formula is **∂S.** This denotes the *edge boundary* of **S**. The edge boundary of a subset is the set of edges **(v, w)** ∈ **E** such that **v** ∈ **S** and **w** ∉ **S**. Informally, this means the set of single edges which connect vertices inside the subset to vertices outside the subset (Maricq, 2014). For the *edge expansion ratio* we have to consider the subset which fulfills the conditions defined in the formula and gives the minimum. The *expansion parameter* for **G** is denoted by **h(G).** The *expansion parameter*, or **h(G)**, is also sometimes referred to as the *Cheeger constant* of the graph **G** (Nielson, 2005).

A helpful feature of the *edge expansion ratio* of an expander is that it also gives a sense of the connectivity of the expander. A high value of the *edge expansion ratio*, or **h(G)**, means that the minimum value for the ratio of the size of the edge boundary of the subset and the size of the subset is high. We can define the subsets as having a size of less than or equal to half the number of vertices, or **|S| ≤ n/2**. This inequality tells us that the subsets of the vertices that comprise less than half of the total number of vertices will be well-connected to larger subsets of vertices, therefore implying that the expander can be considered *well-connected*.
To further explore the *edge expansion ratio* of a graph, take this example.

Case 3.1: Consider a complete graph on **n** vertices, denoted **K_n**. The graph **K_n** is an example of a graph which is *well-connected* but not *sparse*. Every subset of the vertices in the graph is connected to every vertex in the complement. The size of the edge boundary is therefore the size of the subset times the complement of the subset, **|∂S| = |S| * |S¯|**. The complement of this graph can be considered **(n − |S|)**, therefore **|∂S| = |S|(n − |S|)** (Maricq 2014). This leads to the *edge expansion ratio* being:

$$h\left(K_n\right) = \min_{S \subset V : 1 \leq |S| \leq \frac{n}{2}} \left(n - |S|\right) = \left[\frac{n}{2}\right].$$

The graph **K_n** is a deviant. The calculation of the *edge expansion ratio* for almost all graphs is very complex and not nearly as simple as it was for **K_n**. This is due to the difficulties in finding which subset gives the minimum value. Calculating the minimum values of **|S|** and **|∂S|** is extremely difficult and also very time consuming.

Consider a graph **G = (V, E)** on **n** vertices. The reason that finding the minimum value of **|∂S|/|S|** for that graph could be hard is that the number of subsets that fit the criteria is very large. The number of subsets, denoted by **N** is given by:

$$N = \sum_{k=1}^{\frac{n}{2}} \binom{n}{k}.$$

The number of subsets denoted by **N** is calculated by considering the following. The binomial coefficients follow the relation of $\binom{n}{k} = \binom{n}{n-k}$, and the number is equal to the complete number of subsets **S** of the vertices of **G** of size between **1** and **n/2**. Therefore we can say that **N** can be approximated to be half of the number of all the subsets of **V**. The number of all the subsets of set of size **n** is $\mathbf{2^n}$. This means that $\mathbf{N \approx 2^{n-1}}$. Therefore, as the number of vertices, **n**, increases, the number of subsets we have to consider in the calculation to find **h(G)** increases exponentially. This means that if **n** is very large, as it normally is in complex graphs studied in expander graph theory, the time taken in computing **h(G)** also increases exponentially (Maricq 2014). This makes computing **h(G)** unreasonably difficult. To combat this, mathematicians have found ways of approximating **h(G)** that are comparatively easier. One of these ways involves the *spectrum* of a graph and *spectral graph theory*.

Spectral graph theory is, in simple terms, the study of using matrices and graphs in conjunction. In spectral graph theory, a graph can be expressed by its *adjacency matrix*.

Also, as stated previously in the first section, in this part of the paper we are only looking at and working with **k**-regular, undirected graphs. The main reason for doing this in the paper is that doing this simplifies relating the expansion parameter **h(G)** to the eigenvalues of the adjacency matrix of **G**. But first, we will define what an adjacency matrix is.

Definition 3.2: An *adjacency matrix* is a square matrix used to represent a finite graph. The rows and columns of the adjacency matrix are labeled by the vertices of **V**. For vertices **v** and **w** the entry $\mathbf{A(G)_{vw}}$ is defined to be **1** if **(v, w)** is an edge, and **0** if it is not an edge (Nielson, 2005). To further define what an adjacency matrix is in a slightly different way, take a graph **G = (V, E)**. Let this graph have **n** vertices. The adjacency matrix $\mathbf{A_G}$ can be defined as a **n*n** matrix, and let the elements of the matrix be determined by the following equation:

$$\mathbf{a(i, j) = |\{(v_i, v_j) \in E\}|}.$$

The graph **G** is undirected. This means that the edges connecting vertices are unordered pairs of vertices, meaning that **a(i, j) = a(j, i)** for all $\mathbf{1 \leq i \leq n}$ and for all $\mathbf{1 \leq j \leq n}$. Therefore, $\mathbf{A_G}$ is a *symmetric matrix*. A symmetric matrix will always have only real *eigenvalues*. Thus, the *eigenvalues* of $\mathbf{A_G}$ are all real. This means that all the values, $\mathbf{\lambda_m(G)}$, that satisfy the relation $\mathbf{A_G v_m = \lambda_m(G) v_m}$ are *real*. The notation $\mathbf{v_m}$ denotes the *eigenvector* of the adjacency matrix that corresponds to the *eigenvalue* $\mathbf{\lambda_m}$. In the case the graph is symmetric, the eigenvector of m will not be equal to 0. The eigenvectors will also form an *orthonormal* basis for $\mathbf{R}^V$ (**R** is the set of all real numbers) .

The *spectrum* of a graph is the set of eigenvalues of the adjacency matrix of the graph, $\mathbf{A_G}$. The eigenvalues of a **k**-regular graph hold interesting properties. To explore these properties, we have to consider a **k**-regular undirected multigraph.

Case 3.2: Take a graph **G = (V, E)** and let $\mathbf{A_G}$ be its corresponding adjacency matrix. Let $\mathbf{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n}$ be the real eigenvalues of the graph's adjacency matrix. Then:

- The largest eigenvalue will be equal to **k** and the last eigenvalue is more than or equal to **−k**.
- If the second largest eigenvalue is equal to **k**, then the graph is disconnected.
- If the smallest eigenvalue is equal to **-k** then at least one of the connected components of the graph is bipartite (Maricq, 2014).

Let us return to the *edge expansion ratio* of an expander. The **h(G)** of an expander is closely related to the second largest eigenvalue of the adjacency matrix, $\boldsymbol{\lambda_2}$ of $\mathbf{A_G}$. To explore this relation, we will first define what a *spectral gap* is. The *spectral gap* of a graph is the difference between the largest eigenvalue and the second largest eigenvalue, $\boldsymbol{\lambda_1 - \lambda_2}$, and this *spectral gap* is denoted Δ**(G)**. We can use this to create a relationship between **h(G)** and $\boldsymbol{\lambda_2}$ (Nielson, 2005).

The edge expansion ratio **h(G)** for a **k**-regular graph is related to the spectral gap, Δ**(G)**, by:

$$\frac{\Delta(G)}{2} \leq h(G) \leq \sqrt{2k\Delta(G)}\,.$$

The inequality is also referred to as *Cheeger's Inequality*.

The largest eigenvalue, $\boldsymbol{\lambda_1}$, is equal to **k** for a **k**-regular graph. Therefore, *Cheeger's inequality* can also be defined as:

$$\frac{k-\lambda_2}{2} \leq h(G) \leq \sqrt{2k\left(k-\lambda_2\right)}\,.$$

The above helps us achieve estimates for the edge expansion ratio for a **k**-regular graph far easier than the original equation to calculate the value (Maricq, 2014). This allows us to easily differentiate in the quality of the different families of graphs as expanders and allows for easier further research into expander graphs.

Going back to exploring expander graphs, when we explore expander graph theory, we do not usually take singular graphs. This is the same for when we usually explore the properties of expanders. Instead of looking at singular graphs, we look at families of graphs. Therefore, we defined *sparsity* and *density* in terms of graph families. When looking into families of expander graphs, we will follow two definitions, one that considers vertex expansion and one that considers edge expansion. Both definitions are not concrete.

Definition 3.3: Given a fixed positive integer **k**, a family of expander graphs is a family of **k**-regular graphs in which, if $\mathbf{G_n}$ is a member with the vertex set **V = [n]**, the $\mathbf{G_n}$ has good edge expansion for all $\mathbf{S \subseteq V : 0 < |S| \leq n/2}$ **:**

$$\mathbf{Pr}_{u\text{~}v}\,[\mathbf{u} \in \mathbf{S}, \mathbf{v} \,/\!\in \mathbf{S}] \geq \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon > 1}$. It is possible to alter the constraint of **S** to be no more than half of the vertices in favour of another fixed percentage (O'Donnell, 2013).

Definition 3.4: Given a fixed positive integer **k**, a family of expander graphs is a family of **k**-regular graphs in which, if $\mathbf{G_n}$ is a member with the vertex set **V = [n]**, the $\mathbf{G_n}$ has good vertex expansion for all $\mathbf{S \subseteq V : 0 < |S| \leq n/2}$ **:**

$$\mathbf{Pr}_{u\text{~}v}\,[\mathbf{u} \in \mathbf{S}, \mathbf{v} \,/\!\in \mathbf{S}] \geq \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon > 1}$. It is possible to alter the constraint of **S** to be no more than half of the vertices in favour of another fixed percentage (O'Donnell, 2013).

Expander graph families of **k**-regular graphs also follow the rule that a sequence of **k**-regular graphs $\mathbf{\{G_i\}_{i \in N}}$ of size increasing with **i** is a *family of expander graphs* if and only if there exists $\boldsymbol{\alpha > 0}$ such that $\mathbf{h(G_i) \geq \alpha}$ for all **i** (Hoory, 2006). Otherwise, the family cannot be considered an expander graph family.

When considering ***k**-regular expander graphs*, we are usually considering families of these graphs that satisfy these properties. A reason for this is that families of expanders allow us to construct arbitrarily large graphs which have the properties of being *sparse* and *well-connected*. Another class of graph that has these properties are random graphs. However, random graphs are not useful as they are hard to explicitly construct in groups without decreasing randomness. Therefore, the third major property of an expander graph is that they are able to be constructed explicitly. Families of expander graphs improve on this property.

Before we look at an example of a family of expander graphs, we will explore an example of a construction of a family of graphs in which the graphs in the family are not good expanders.

Case 3.3: To construct the family of graphs:

Take four vertices that link all but the diagonals of a square's four corners. Then, carry out the same procedure with a smaller, 90°-rotated square inside the first one. By joining the two nearest vertices at each corner, join the larger and smaller squares together. Inside the first square, create a smaller square. Repeat this procedure over and over.

$\mathbf{F_N = \{ 4^8, 6^{4N}\}}$ is the order of the degrees. This applies when **N** is a natural number, and the number of vertices with each degree is indicated by the superscripts. Therefore, to continue linking square graphs inside square graphs from$\mathbf{F_0 \rightarrow F_1 \rightarrow F_2 \rightarrow \ldots \rightarrow F_N}$, we will start with 2 squares for $\mathbf{F_0}$ in this family.

An example of a construction is $\mathbf{F_2}$:

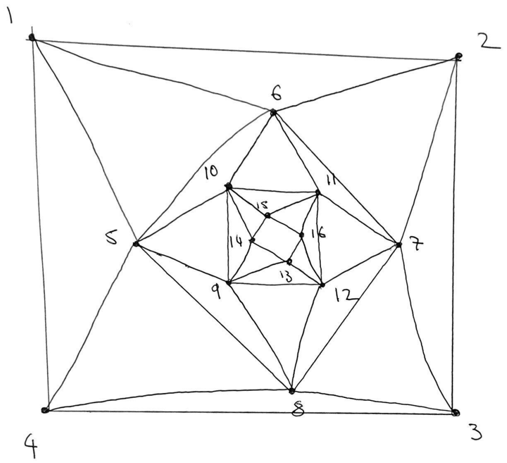

Figure 3.2: Graph $\mathbf{F_2}$

However, this family of graphs is not a good expander family.

Proof 3.1: If the vertices of **F_N** are numbered **1, 2, 3, 4, …, 4(n+2)** in order of construction, the way they are in the diagram, then for any **k**, the boundary of the vertex set **Sk={1, 2, …,4 k}** consists of **8** edges joining the vertices **{4k−3, 4k−2, 4k−1, 4k}** to the vertices **{4k+1, 4k+2, 4k+3, 4k+4}**. Take **k** to be about **n/2** for maximum effect; by any measure, the boundary of **S** is constant while **$|S| \approx 2n$**.

So **F_N** has vertex expansion and edge expansion both on the order of **$O(n^{-1})$**, which means that the quality of this family of graphs as expanders is low.

We will now give an example of a family of expanders.

Case 3.3: A family of expander's s the Lubotzky-Phillips-Sarnak expander. The LPS expander family can be depicted as such:

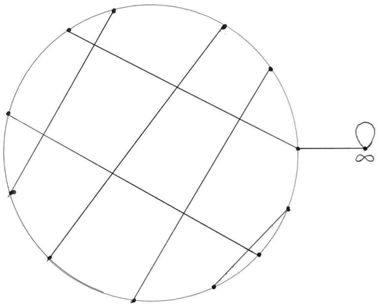

Figure 3.3: LPS expander

The construction we will be describing for this graph is far simpler than the original construction of the expander family, as the mathematical level and time required for the original construction is beyond the scope of this paper. This construction is inspired by the simplified construction of this expander family by Ryan O'Donnell in his Autumn 2005 lecture on 'The PCP Theorem and Hardness of Approximation'.

First consider a graph **G = (V,E)**. Let **$V = Z_p \cup$ {∞}**, and **p** is prime. We extend the multiplicative inverse by defining **0** to **1** as the special point, and∞ **+ x = ∞** for all **$x \in V$**. Consider **V** as a **p**-element field defined by addition and multiplication modulo **p** (which we examined in the Cayley graph section). Connect any vertex **x** to **x + 1**, **x − 1** and **x^{-1}**. With an absolute constant **λ_0**., this results in a **3**-regular graph with the second-largest eigenvalue, **$\lambda_2 \leq \lambda_0 < 3$**. Figure 3.3 depicts the graph's structure. The graph is a cycle with a matching between the edges. Without significantly affecting the expansion, the extra point with a self-loop that is introduced by the point ∞ can be eliminated together with the zero and one connected to **p − 1**. We can delete them and obtain a simple graph that is a cycle on **p − 3** nodes because **1** and **p − 1** also have self loops for their inverses (instead of matching edges, though this isn't depicted in the diagram). We can use this construction for different prime numbers and the different graphs produced are the members of the family of expander graphs (O'Donnell, 2005).

Even though expander graphs having explicit construction is important, it is more important to make sure that the expander graph is a good expander. To do this we must be able to find the *edge expansion ratio*. The ratio helps in determining how good the expander is. This is because the *edge expansion ratio* also helps us interpret the uses of the expander graph in performing operations and also in determining the usefulness of the graph in random walks which are useful in cryptography and creating cryptographic hash functions. Hence we focused earlier on spectral graph theory and easier ways of calculating the *edge expansion ratio*.

4. EXPANSION IN CAYLEY GRAPHS

In the final section of this paper, we will investigate how expander families can be formed for groups. We will explore this concept by looking into the expansion of Cayley graphs to construct a family of expanders. The construction of the Cayley expander family that we will use in this section, however, is not actually a prototypical expander family. It has not been proven to be an expander family. Even though it has the properties of being well-connected and explicitly constructable, the problem with the expansion in this Cayley graph is that the behavior of the graph as **n** increases does not show that there are no large variations, and it is beyond the scope of this paper to prove that as **n** increases the graph will continue to be considered *sparse*.

To construct the expander family:

Consider the cyclic group $\mathbf{Z_n}$ and let $\mathbf{S = \{1\}}$. (Here we are considering $\mathbf{Z_n}$ to be the set of integers $\mathbf{0 \leq i < n}$ under modular addition). Then if $\mathbf{u} \in \mathbf{Z_n}$, its neighborhood in $\mathbf{G(Z_n, S)}$ must be $\mathbf{\{u - 1, u + 1}\}$. Hence $\mathbf{G(Z_n, S)}$ is isomorphic to the cycle $\mathbf{C_n}$.

This construction creates a family of Cayley graphs, and changing the value of **n** creates new graphs which are a part of the family.

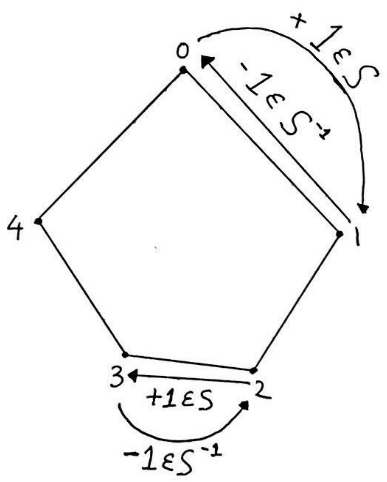

Figure 4.1: Cayley expander Graph with n = 5

This is a case of using this construction with a **n** value of 5. We can use other values of **n** to create other members of the family; all in all allowing us to create a family of expander graphs using Cayley generation of graphs. This particular Cayley expander family was inspired by a Cayley graph constructed in the paper 'Cayley Graph Theory' written by Kenan Wood.

The spectral gap in a Cayley graph also has some special properties. One of these special properties is the relationship the spectral gap has with the diameter. Take a Cayley graph **C(G, S)**. Let the group of this graph be a finite group, and let the generating set, or **S**, be a finite and symmetric set. Then the spectral gap of the graph will be greater than or equal to the inverse of double the diameter. This relationship helps improve estimates of the edge expansion ratios of Cayley graphs.

CONCLUSION

In this paper, we have effectively introduced the concept of expander graphs, discussed the properties that make an expander graph a good expander, given examples of previous expander families in published work and have given an example of a Cayley expander family. We were unable to prove the Cayley expander family was a good expander as the value of **n** increases and were unable to calculate the edge expansions ratio as an algebraic value as **n** increases, as this was beyond the scope of this paper.

We also explored spectral graph theory and how the eigenvalues of a graph can be used to determine the edge expansions ratio of an expander. However, we were unable to explore examples of the uses of spectral graph theory in families of graphs. We further explored the properties of an expander and how these properties are special, however we were unable to fully explore the applications of these properties in real life.

Expander graphs can be applied in cryptography using spectral graph theory and Cayley graph notation. The properties of the group, subgroup and adjacency matrix all become more useful and have wider real world applications when looking into expander hashes. In the future, I will investigate the applications of expander graphs in random walks and expander hashes. Furthermore, I will conduct cryptanalysis of expander and Cayley hashes.

Bibliography:

Preiss, Bruno (1999), *Data structures and algorithms with object oriented design patterns in c*++, Wiley

Buckley, Fred, Lewinter, Marty (2013) *Introductory Graph Theory with Applications*, Waveland Press

Wood, Kenan (2022), *Cayley Graph Theory*, Davidson College

Harris, John M. Hirst, Jeffry L. Mossinghoff, Michael J. (2008), *Combinatorics and Graph Theory, Second Edition*, Springer

Nielson, Michael A. (2005), *Introduction to expander graphs*, The University of Queensland Brisbane https://michaelnielsen.org/blog/archive/notes/expander_graphs.pdf

Maricq, Aleksander (2014), *Applications of Expander Graphs in Cryptography* https://www.whitman.edu/Documents/Academics/Mathematics/2014/maricqaj.pdf

O’Dennell, Ryan (2005), *CSE 533: The PCP Theorem and Hardness of Approximation,* Lecture 1, Washington

Tao, Terrence (2011), *Basic theory of expander graphs*, 254B, Notes 1
https://terrytao.wordpress.com/2011/12/02/245b-notes-1-basic-theory-of-expander-graphs/

Weisstein, Eric W (2022), *Graph Automorphism*, From MathWorld- A Wolfram Web Resource
https://mathworld.wolfram.com/GraphAutomorphism.html

Petit, Christophe (2009), *On graph-based cryptographic hash functions*, University College London Thesis Library

Hoory, Shlomo, Linial, Nathan, Wigdersib, Avi (2006), *Expander Graphs And Their Applications*, Bulletin Of The American Mathematical Society

Sherman-Bennett, Melissa U. (2016), *On Groups and Their Graphs*, Berkeley Thesis Library

Brandstädt, Andreas (2012), *On the complexity of some packing and covering problems in certain graph classes*, University of Rostock

O'Dennell, Ryan (2013), *Expanders*, Lecture 12, Carnegie Mellon

Nicol, Andrew (2008), *What Is... A Cayley Graph?*, Ohio State University Math Department

Kamble, Avinash J. Rithe, Shital, Pratham, Harshada (2022), *A Review on Graphs arising from Finite Groups*, International Journal of Mathematics And its Applications

Lavror, Misha (2019), *Explicit construction*, Math Stack Exchange
https://math.stackexchange.com/questions/3133874/explicit-construction-and-proving-or-disproving-expander-graph-for-this-family

Wikipedia (2022), *Generating set of a group*, Wikimedia Foundation
https://en.wikipedia.org/wiki/Generating_set_of_a_group

Wang, Chih-Hung (2011), *Finite Fields*, Introduction to Number Theory, Information Security and Management