GRAPHS WITH LARGE GIRTH AND LARGE CHROMATIC NUMBER

CHEUK TO TSUI

Abstract. This paper investigates graphs that have large girth and large chromatic number. We first give a construction of a family of graphs that do not contain cycles of length 3, but have unbounded chromatic number. Next, we prove the existence of graphs that have large girth and large chromatic number. With existence proven, we give an explicit construction of such graphs.

Contents

1. Introduction

Finding a lower bound for the chromatic number of a given graph is, in general, difficult to do. There are few techniques that show that a certain number of colors are *not* enough to color a graph. Because of this, it is interesting to consider how we may construct graphs with large chromatic number. One way to do so is to include many cycles; since every even cycle requires at least 2 colors and every odd cycle requires at least 3, it is easy to construct graphs with large chromatic number as long as cycles are permitted (take the complete graphs, for instance). But the matter becomes more complicated when we do not permit cycles below a given length, and want to construct graphs with large girth and large chromatic number. In the remainder of Section 1, we define the terminology that allows us to begin addressing this problem. In section 2, we address the simplified problem of showing that there are triangle-free graphs with arbitrarily large chromatic number. In section 3, we show more generally that there exist graphs that have girth $\geq l$ and chromatic number $\geq k$ for any k and l. In section 4, we explicitly construct graphs with large girth and chromatic number, show the basic method of finding a lower bound for their girth, and find a lower bound for their chromatic number.

Definition 1.1. A *graph* $G = (V, E)$ consists of a vertex set V and an edge set E, where elements of the edge set E are unordered pairs $\{v_i, v_j\}$, with $v_i \neq v_j$ and $v_i, v_j \in V$.

Definition 1.2. Two vertices v_i and v_j are *adjacent* if there is an edge between them, that is, $\{v_i, v_j\} \in E$. We can also write $v_i \sim v_j$.

Definition 1.3. A *cycle* is a sequence of vertices that begins and ends at the same vertex, such that successive vertices are adjacent and no vertex except the first repeats. The number of distinct vertices in the sequence is called the *length* of the cycle. A cycle of length k is denoted C_k.

Definition 1.4. The *girth* of a graph is the length of the shortest cycle contained in it. If a graph contains no cycles, its girth is defined to be ∞.

Definition 1.5. A graph G is *triangle-free* if it does not contain a cycle of length 3.

Definition 1.6. A set of vertices S is *independent* if no two vertices in S are adjacent.

Definition 1.7. The *independence number* $\alpha(G)$ of a graph G is the maximum size of an independent vertex set.

Definition 1.8. A (proper) *k-coloring* of a graph is a function $f : V \to \{1, 2, ..., k\}$ such that if v_i and v_j are adjacent, then $f(v_i) \neq f(v_j)$.

Definition 1.9. The *chromatic number* $\chi(G)$ of a graph G is the minimum value of k such that a k-coloring exists.

2. Triangle-Free Graphs with Large Chromatic Number

Before addressing the general question of constructing graphs with large girth and chromatic number, we first consider the simpler problem of constructing triangle-free graphs with large chromatic number.

Theorem 2.1. *For any positive integer k, there exists a triangle-free graph with chromatic number k.*

Proof. For $k = 1$, a single point will work. For $k = 2$, two adjacent vertices will work, and for $k = 3$, C_5 will work. For chromatic numbers larger than 3, we give a construction that, when applied to a triangle-free graph, produces a new triangle-free graph that increases the chromatic number by 1.

Construction 2.2 (Mycielski. See [1])**.** Given a graph G, label its vertices $v_i, i \in \{1, 2, ..., n\}$. Create n new vertices u_i and an additional vertex w, and draw an edge between each u_i and w. Then, if $\{v_i, v_j\}$ is an edge, draw the two edges $\{u_i, v_j\}$ and $\{v_i, u_j\}$. Call this new graph G'.

First, we show that if G is triangle-free, then G' is triangle-free. Suppose there exists a C_3 in G'. Vertex w cannot be a vertex of a C_3, because it is adjacent only to all the u_i, which are not pairwise adjacent. There are also no C_3's containing two of the u_i, since no two u_i are adjacent. Therefore, the C_3 must contain two adjacent v_i and one of the u_i. Let the vertices be v_j, v_k and u_l. They must be pairwise adjacent, meaning that $\{v_j, u_l\}$ and $\{v_k, u_l\}$ are edges of G'. By the construction, this would be the case only if $\{v_j, v_l\}$ and $\{v_k, v_l\}$ were edges of G. But this would mean that v_j, v_k, v_l, v_j would be a C_3 in G, which contradicts the fact that G is triangle-free. Thus there exists no C_3 in G'.

Now, we show that $\chi(G') = \chi(G) + 1$. Let $\chi(G) = k$. This means that the v_i must be colored by at least k colors. We claim that in any coloring of G', the u_i must also be colored by at least k colors. Suppose that the u_i can be colored by at most $k-1$ colors. so at least one of the v_i are colored by color k, while none of the u_i are. Observe that for all i, v_i and u_i are adjacent to exactly the same v_j's in G. Take the v_i that are colored by color k, and change them to the color of the corresponding u_i. This new coloring of the v_i would still be valid, because each v_i has the same neighbours as the corresponding u_i, meaning that none of the v_j's adjacent to each v_i have the same color as the corresponding u_i. But this new coloring (of the v_i) uses at most $k-1$ colors, contradicting the fact that G must be colored by at least k colors. Thus, the u_i must be colored by at least k colors as well. Since w is adjacent to all the u_i, it must be colored by a $(k+1)$st color. Therefore, $\chi(G') \geq \chi(G) + 1$. Furthermore, G' can be colored by $\chi(G) + 1$ colors: First color the v_i with k colors, then color each u_i the same color as the corresponding v_i. Then color w with a $(k+1)$st color. So $\chi(G') = \chi(G) + 1$.

Therefore, to obtain a triangle-free graph with chromatic number $k > 3$, we iteratively apply the construction to C_5 $k-3$ times.

□

3. Existence of Graphs with Large Girth and Large Chromatic Number

Theorem 2.1 shows that the chromatic number of a graph can be made arbitrarily large, even if it does not contain cycles of length 3. Now we consider the more general question of whether it is possible to avoid cycles below a given length, yet still have an arbitrarily large chromatic number. As preparation, we present one probabilistic result and define some more notation.

Lemma 3.1 (Markov's Inequality). *If X is a non-negative discrete random variable and $a > 0$, then*

$$P[X \geq a] \leq \frac{\boldsymbol{E}[X]}{a}.$$

Proof. We have

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{x \geq 0} xP(x) \\
&= \sum_{x \geq a} xP(x) + \sum_{x < a} xP(x) \\
&\geq \sum_{x \geq a} xP(x) \\
&\geq a \sum_{x \geq a} P(x) \\
&= aP[X \geq a].
\end{aligned}$$

□

Definition 3.2 ([2, p. 38])**.** We write $G \sim G(n,p)$ if G is a random graph on n vertices chosen by picking each pair of vertices as an edge randomly and independently with probability p.

Definition 3.3 (little-oh notation)**.** We write $a_n = o(b_n)$ if

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

Remark 3.4. $a_n = o(1)$ means that $\lim_{n\to\infty} a_n = 0$.

Theorem 3.5 (Erdös 1959. See [2, pp. 38-39])**.** *For all k, l there exists a graph G with girth(G) $> l$ and $\chi(G) > k$.*

Proof. Let $\theta < 1/l$, and let $G \sim G(n,p)$ with $p = n^{\theta-1}$. Let X be the number of cycles of length at most l in G, and let X_i be the number of cycles of size exactly i. By linearity of expectation, we have

$$E[X] = \sum_{i=3}^{l} E[X_i].$$

To determine E[X], we count the possible number of cycles of length i. There are $\binom{n}{i}$ ways to choose the vertices of the cycle, and $\frac{(i-1)!}{2}$ ways to order them (we can start with any vertex). We must divide by 2, because the reverse of any ordering gives the same cycle. So the total number of possible cycles is

$$\binom{n}{i}\frac{(i-1)!}{2} = \frac{(n)_i}{2i},$$

where $(n)_i = (n)(n-1)...(n-i+1)$ (falling factorial).

The probability of each of these cycles occuring is p^i, so by linearity of expectation, $E[X_i] = \frac{(n)_i}{2i}p^i$. This, along with the bound $(n)_i \leq n^i$, gives us

$$E[X] = \sum_{i=3}^{l} \frac{(n)_i}{2i}p^i \leq \sum_{i=3}^{l} \frac{n^i}{2i} n^{i(\theta-1)} = \sum_{i=3}^{l} \frac{n^{\theta i}}{2i} \leq n^{\theta l} \sum_{i=3}^{l} \frac{1}{2i}.$$

Now, since $\theta l < 1$ by assumption, $(n^{\theta l}\sum_{i=3}^{l}\frac{1}{2i})/n = n^{\theta l-1}\sum_{i=3}^{l}\frac{1}{2i}$ goes to 0 as n goes to ∞. So $E[X] = o(n)$. Using this fact and Markov's Inequality, we get

$$P[X \geq n/2] \leq E[X]/(n/2) = o(1).$$

We now use the independence number $\alpha(G)$ to estimate the chromatic number $\chi(G)$. Let us compute $P[\alpha(G) \geq x]$. Any set of x vertices are independent if all $\binom{x}{2}$ edges between them are not drawn, which occurs with probability $(1-p)^{\binom{x}{2}}$. Thus, the probability $P[\alpha(G) \geq x]$ that at least one set of x vertices is independent is at most $\binom{n}{x}(1-p)^{\binom{x}{2}}$ by the union bound. By the bounds $\binom{n}{x} \leq n^x$ and $(1-p)^x \leq e^{-px}$, we obtain

$$P[\alpha(G) \geq x] \leq \binom{n}{x}(1-p)^{\binom{x}{2}} \leq [ne^{-p(x-1)/2}]^x.$$

We want to set x such that $P[\alpha(G) \geq x]$ goes to 0 as n gets large. To do this, we can require that $ne^{-p(x-1)/2} < \frac{1}{\sqrt{n}}$. This is equivalent to

$$n^{\frac{3}{2}} < e^{p(x-1)/2}$$
$$\ln n < \frac{p(x-1)}{3}$$
$$\frac{3}{p}\ln n < x-1$$
$$1+\frac{3}{p}\ln n < x.$$

As n gets large, the constant 1 becomes negligible, so we can simply set $x = \lceil \frac{3}{p}\ln n \rceil$ to ensure $P[\alpha(G) \geq x] = o(1)$. We can verify this: We have, using $\lceil \frac{3}{p}\ln n \rceil < \frac{3}{p}\ln n + 1$,

$$\begin{aligned} P[\alpha(G) \geq x] \leq [ne^{-p(x-1)/2}]^{x} &< [ne^{-p(\frac{3}{p}\ln n)/2}]^{\frac{3}{p}\ln n+1} \\ &= [n^{-\frac{1}{2}}]^{\frac{3}{n^{\theta-1}}\ln n+1} \\ &= n^{-\frac{3}{2}n^{1-\theta}\ln n+\frac{1}{2}} \\ &= o(1). \end{aligned}$$

Now let n be so large that $P[X \geq n/2] < 0.5$ and $P[\alpha(G) \geq x] < 0.5$, which is possible since both are $o(1)$. Then there exists a graph G with less than $n/2$ cycles of length at most l and with $\alpha(G) < 3n^{1-\theta}\ln n$. Removing a vertex from each cycle of length at most l leaves a graph G^* with at least $n/2$ vertices and girth greater than l. We also have $\alpha(G^*) \leq \alpha(G)$, since every independent set of G^* is also one of G. This allows us to give a lower bound for the chromatic number, using the following lemma.

Lemma 3.6. *For any graph G, $\chi(G)\alpha(G) \geq n$.*

Proof. Let $\chi(G) = k$ and $\alpha(G) = a$. Since all the vertices colored by the same color form an independent set, there can be at most a vertices of each color. G can be k-colored, and k colors can color at most ka vertices, so $ka \geq n$. □

Denote the number of vertices of G^* by $|G^*|$. We have, by the lemma,

$$\chi(G^*) \geq \frac{|G^*|}{\alpha(G^*)} \geq \frac{n/2}{3n^{1-\theta}\ln(n)} = \frac{n^{\theta}}{6\ln n}.$$

Since $\frac{n^{\theta}}{6\ln n}$ is unbounded, we can make $\chi(G^*)$ larger than k by choosing a large enough n. Thus G^* has the desired properties.

□

4. Construction of Graphs with Large Girth and Large Chromatic Number

The probabilistic argument given in the proof of Theorem 3.5 proves the existence of graphs with large girth and large chromatic number. The next question to ask is how such graphs can be explicitly constructed. We address this question in this section by constructing graphs called Ramanujan graphs. We then give the basic idea behind obtaining a lower bound for the girth of these graphs, and obtain a lower bound for the chromatic number. First, we state some definitions from group theory.

Definition 4.1. A *group* $(G, \cdot)$ consists of a set of elements G and an operation $\cdot$ with the following properties:

Closure: For all a, b in G, ab is in G.
Associativity: For all a, b, c in G, $a(bc) = (ab)c$.
Identity: There exists an element 1_G in G such that for a in G, $1_G a = a 1_G = a$.
Inverse: For all a in G, there exists an element a^{-1} in G such that

$$a \cdot a^{-1} = a^{-1} \cdot a = 1_G$$

Definition 4.2. A subset H of G is a *subgroup* of G, denoted $H \leq G$, if H forms a group under the same operation.

Definition 4.3. A subgroup N of G is a *normal subgroup*, denoted $N \lhd G$, if for all a in G,

$$a^{-1}Na = N.$$

Definition 4.4. The *general linear group of degree n* $GL(n, \mathbb{F})$ is the group of $n \times n$ nonsingular matrices over a field $\mathbb{F}$, with the matrix multiplication operation.

Definition 4.5. The *special linear group of degree n* $SL(n, \mathbb{F})$ is the group of $n \times n$ matrices with determinant 1 over a field $\mathbb{F}$, with the matrix multiplication operation.

Definition 4.6. Given a subgroup $H \leq G$, and $x \in G$, define
$xH = \{xh : h \in H\}$ and $Hx = \{hx : h \in H\}$. A subset of G that is in the form xH for some $x \in G$ is a *left coset of H*. A subset of G in the form Hx for some $x \in G$ is a *right coset of H*.

Definition 4.7. Given a group G and $N \lhd G$, we denote by G/N the set of all (right) cosets of N. This set is called the *quotient set*. Define an operation $*$ by

$$Na * Nb = Nab.$$

Then $(G/N, *)$ is called the *quotient group of N in G*.

Example 4.8 (Modulo d residue classes)**.** Examples of quotient groups are the modulo d residue classes, denoted $\mathbb{Z}/d\mathbb{Z}$, where $d\mathbb{Z} = \{dk : k \in \mathbb{Z}\}$ is a subgroup of $\mathbb{Z}$.

Definition 4.9. The *center* of a group G, denoted $Z(G)$, is defined as $Z(G) = \{a \in G : \forall b \in G, ab = ba\}$.

The following result will be used in the construction.

Proposition 4.10 (Center of the general linear group)**.** *The center of* $GL(n, \mathbb{F})$ *is the set of scalar matrices except for the zero matrix, i.e.* $\{\lambda I : \lambda \in \mathbb{F} \backslash \{0\}\}$.

The next result is necessary for the following two definitions.

Proposition 4.11. *The center of a group is a normal subgroup.*

Definition 4.12. The *projective general linear group of degree n* is defined as $PGL(n, \mathbb{F}) = GL(n, \mathbb{F})/Z(GL(n, \mathbb{F}))$.

Definition 4.13. The *projective special linear group of degree n* is defined as $PSL(n, \mathbb{F}) = SL(n, \mathbb{F})/Z(SL(n, \mathbb{F}))$.

The following definition provides a way to generate a graph from a group and a subset.

Definition 4.14. The *Cayley graph* of a group G with respect to a subset S, denoted $\Gamma(G, S)$, is defined as follows:
$V(\Gamma) = G$
For all $g \in G$, $g \sim gh$ if and only if $h \in S$.

We state the following theorem without proof.

Theorem 4.15 (Jacobi's Four Squares Theorem. See [4, p. 262])**.** *The number of ways to express a positive integer n as a sum of 4 squares is*

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

Let $X = X_{n,k}$ be a k-regular graph on n vertices. Let A_X be its adjacency matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. We know that $|\lambda_i| \leq k$ for $1 \leq i \leq n$. Let $\lambda(X)$ denote the absolute value of the largest eigenvalue (in absolute value) of A_X that is not equal to $\pm$ k.

Definition 4.16. $X_{n,k}$ is called a *Ramanujan graph* if

$$\lambda(X) \leq 2\sqrt{k-1}.$$

We now construct Ramanujan graphs that will be shown to have large girth and chromatic number.

Construction 4.17 ([4])**.** Choose two prime numbers p, q that are congruent to 1 mod 4, and choose an integer i such that $i^2 \equiv -1 \pmod{q}$. By Jacobi's Four Squares Theorem, the number of solutions $\alpha = (a_0, a_1, a_2, a_3)$ to the equation

$${a_0}^2 + {a_1}^2 + {a_2}^2 + {a_3}^2 = p$$

is $8(p+1)$. It can be proven that $p+1$ of the solutions satisfy the conditions that a_0 is positive and odd, and a_1, a_2, a_3 are even. For each of these solutions α, associate the matrix

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

We can then form the Cayley graph of $PGL(2, \mathbb{Z}/q\mathbb{Z})$ with respect to the set of these $p + 1$ matrices. The next two lemmas give the number of vertices that this Cayley graph has.

Lemma 4.18. $|GL(2, \mathbb{Z}/q\mathbb{Z})| = (q^2 - q)(q^2 - 1)$

Proof. The size of $GL(2, \mathbb{Z}/q\mathbb{Z})$ is the number of 2 by 2 nonsingular matrices with each entry an integer between 0 to $q - 1$ inclusive. There are $q^2 - 1$ ways to choose the first column, since a column with all 0s cannot occur in a nonsingular matrix. The second column can be anything except for the multiples of the first column, of which there are q, so the second column can be chosen in $q^2 - q$ ways. Thus, there are a total of $(q^2 - q)(q^2 - 1)$ matrices in $GL(2, \mathbb{Z}/q\mathbb{Z})$. □

Lemma 4.19. $|PGL(2, \mathbb{Z}/q\mathbb{Z})| = q(q^2 - 1)$

Proof. The elements of $PGL(2,\mathbb{Z}/q\mathbb{Z})$ are the cosets of $Z(G)$ in $GL(2,\mathbb{Z}/q\mathbb{Z})$, which are the subsets in the form $\{\lambda A : \lambda \in \{1,2,...,q-1\}\}$, where $A \in GL(2,\mathbb{Z}/q\mathbb{Z})$. Thus, each coset contains $q-1$ elements of $GL(2,\mathbb{Z}/q\mathbb{Z})$, and there are

$$|GL(2,\mathbb{Z}/q\mathbb{Z})|/(q-1) = q(q^2-1)$$

cosets in $PGL(2,\mathbb{Z}/q\mathbb{Z})$ by the previous lemma. □

It can be similarly shown that $|SL(2,\mathbb{Z}/q\mathbb{Z})| = q(q^2-1)$ and $|PSL(2,\mathbb{Z}/q\mathbb{Z})| = q(q^2-1)/2$.

From the two above lemmas, we can see that the Cayley graph has $n = q(q^2-1)$ vertices. It can also be shown that the graph is $(p+1)$-regular. If p is not a perfect square modulo q, then this graph can be shown to be connected. However, consider the case when p is a (non-zero) perfect square modulo q. The determinants of the generators $\tilde{\alpha}$ are

$$(a_0+ia_1)(a_0-ia_1)-(-a_2+ia_3)(a_2+ia_3) = {a_0}^2+{a_1}^2+{a_2}^2+{a_3}^2 = p,$$

which is a perfect square by assumption. This means that each $\tilde{\alpha}$ is a scalar multiple of a matrix in $PSL(2,\mathbb{Z}/q\mathbb{Z})$, i.e. a matrix with determinant 1. So all the generators lie in the subgroup $PSL(2,\mathbb{Z}/q\mathbb{Z})$. But then the component of the Cayley graph of $PGL(2,\mathbb{Z}/q\mathbb{Z})$ with respect to the set of generators that contains the identity matrix will only include vertices corresponding to matrices that are in $PSL(2,\mathbb{Z}/q\mathbb{Z})$. Thus, if p is a perfect square mod q, the Cayley graph is not connected.

Because of this, if p is not a perfect square mod q, we define $X^{p,q}$ to be the Cayley graph of $PGL(2,\mathbb{Z}/q\mathbb{Z})$ with respect to the $p+1$ generators. If p is a (non-zero) perfect square mod q, we define $X^{p,q}$ to be the Cayley graph of $PSL(2,\mathbb{Z}/q\mathbb{Z})$ with respect to the generators, which can be shown to be connected, and in addition, non-bipartite. It can be shown (see [4]) that the $X^{p,q}$ are Ramanujan graphs (i.e. they satisfy Definition 4.16).

The next step is to find a lower bound on the girth of the Ramanujan graphs $X^{p,q}$. The result is given below.

Theorem 4.20 ([4])**.** *If p is not a perfect square mod q, then $girth(X^{p,q}) \geq 4log_pq - log_p4$. If p is a non-zero perfect square mod q, then $girth(X^{p,q}) \geq 2log_pq$.*

Since log_pq is can be made arbitrarily large, the girth of $X^{p,q}$ can be made as large as desired.

The proof of Theorem 4.20 is omitted, but the basic idea behind obtaining a lower bound is as follows: A cycle is formed by two different walks from one vertex to another. We can represent a walk as the multiplication of a sequence of matrices. Using the matrix norm, we can find how long two walks must at least be for them to start and end at the same vertex (thus forming a cycle). This will then give us a bound on the girth. We illustrate this idea by estimating the girth of a simpler Cayley graph X_p that is defined below.

4.1. **Lower Bound on Girth of X_p.** [5]

If H is a subset of a group G, define a *word W over H* to be a finite sequence $f_1, f_2, ..., f_n$ such that for all $1 \leq i \leq n$, either f_i or ${f_i}^{-1}$ is in H. Define a word W to be *reduced* if $f_{i+1} \neq {f_i}^{-1}$ for all $1 \leq i \leq n-1$.

Definition 4.21. A map $\varphi : G \to L$ for groups G and L, is called a *homomorphism* if for all $x, y \in G$, we have

$$\varphi(xy) = \varphi(x)\varphi(y).$$

Consider the matrices $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$. It is known that there is no nontrivial multiplicative relation between A and B in $SL(2,\mathbb{Z})$, which means that any two reduced words over the set $\{A, B\}$ will define different matrices in $SL(2,\mathbb{Z})$, where the matrix defined by a word is just $f_1 f_2 \cdots f_n$.

Given a prime p, denote the group $SL(2,\mathbb{Z}/p\mathbb{Z})$ by G_p. Define the homomorphism φ_p from $SL(2,\mathbb{Z})$ to G_p which maps each matrix $X \in SL(2,\mathbb{Z})$ to the matrix $\varphi_p(X)$, obtained by taking each entry of X mod p. Then, define $A_p = \varphi_p(A)$, $B_p = \varphi_p(B)$, ${A_p}^{-1} = \varphi_p(A^{-1})$, and ${B_p}^{-1} = \varphi_p(B^{-1})$. Note that this notation is justified, since we have $\varphi_p(A)\varphi_p(A^{-1}) = \varphi_p(AA^{-1}) = I$, and so $\varphi_p(A)^{-1} = \varphi_p(A^{-1})$. Now define the set $\mathscr{U}_p = \{A_p, B_p, {A_p}^{-1}, {B_p}^{-1}\}$ and denote the Cayley graph $\Gamma(G_p, \mathscr{U}_p)$ by X_p.

The goal is to find a lower bound for $girth(X_p)$. To do so, we will find a lower bound for $d(X_p)$, which we define as the smallest integer such that there exist two walks in X_p of lengths $\leq d(X_p)$ starting at $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ that end at the same vertex. In this context, a walk means a sequence of adjacent vertices $x_0, x_1, ..., x_k$ such that $x_{i-1} \neq x_{i+1}$. We have the following relationship between $girth(X_p)$ and $d(X_p)$:

Lemma 4.22. $girth(X_p) \geq 2d(X_p) - 1$.

Proof. Let $D(X_p)$ be the largest integer such that any two walks in X_p of lengths $\leq D(X_p)$ starting at I end at different vertices. By the definitions of $d(X_p)$ and $D(X_p)$, $d(X_p) = D(X_p) + 1$. Let g_I be the length of the shortest cycle in X_p that contains I. We show that either $g_I = 2D(X_p) + 1$ or $g_I = 2D(X_p) + 2$.

Suppose that $g_I \leq 2D(X_p)$. Then there exists a cycle $x_1 = I, x_2, ..., x_k, x_1$, where $k \leq 2D(X_p)$. But the walks $x_1, x_2, ..., x_{\lceil k/2 \rceil}$ and $x_1, x_k, x_{k-1}, ..., x_{\lceil k/2 \rceil}$ both have length $\leq D(X_p)$, start at I and have a common endpoint, which contradicts the definition of $D(X_p)$.

Now suppose that $g_I \geq 2D(X_p) + 3$. Then any two walks of length $\leq D(X_p) + 1$ beginning at I must end at different vertices, otherwise they will form a cycle of length at most $2D(X_p) + 2$. But $D(X_p)$ is the largest integer with this property, so this is a contradiction.

Therefore, $g_I = 2D(X_p) + 1$ or $g_I = 2D(X_p) + 2$. Equivalently, $g_I = 2d(X_p) - 1$ or $g_I = 2d(X_p)$. Now, X_p is vertex-transitive, so this result applies to all vertices of X_p, that is, the length of the shortest cycle that contains any vertex is either $2d(X_p) - 1$ or $2d(X_p)$. So $girth(X_p) \geq 2d(X_p) - 1$ as desired. □

Suppose that we have two walks $P = (p_0, p_1, ..., p_r)$ and $S = (s_0, s_1, ..., s_t)$ that both start at $I = p_0 = s_0$ and end at the same vertex $p_r = s_t$. Since $X_p = \Gamma(G_p, \mathscr{U}_p)$, we have $p_i = p_{i-1}v_i$ and $s_j = s_{j-1}w_j$, $1 \leq i \leq r$, $1 \leq j \leq t$, where $v_i, w_j \in \mathscr{U}_p$. From this, we obtain $p_i = v_1 \cdot ... \cdot v_i$ and $s_j = w_1 \cdot ... \cdot w_j$. We have $p_r = s_t$, so

$$v_1 \cdot ... \cdot v_r = w_1 \cdot ... \cdot w_t. \tag{4.23}$$

Now define the word $\tilde{V} = (\tilde{v_1}, ..., \tilde{v_r})$ by

$$\tilde{v_i} = \begin{cases} A & \text{if } v_i = A_p \\ B & \text{if } v_i = B_p \\ A^{-1} & \text{if } v_i = {A_p}^{-1} \\ B^{-1} & \text{if } v_i = {B_p}^{-1} \end{cases}$$

and $\tilde{W} = (\tilde{w_1}, ..., \tilde{w_r})$ similarly. Since P and S are walks, we have $p_{i+1} \neq p_{i-1}$, but $p_{i+1} = p_{i-1}v_i v_{i+1}$, implying that $v_{i+1} \neq {v_i}^{-1}$. So $V = (v_1, ..., v_r)$, and hence $\tilde{V}$, is a reduced word. The same is true for $W = (w_1, ..., w_t)$ and hence $\tilde{W}$. But P and S are different walks and both begin at I, so V and W are different, and so $\tilde{V}$ and $\tilde{W}$ are different. Since $\tilde{V}$ and $\tilde{W}$ are different reduced words over $\{A, B\}$, we have

$$\tilde{v_1} \cdot ... \cdot \tilde{v_r} \neq \tilde{w_1} \cdot ... \cdot \tilde{w_t} \tag{4.24}$$

because there is no nontrivial multiplicative relation between A and B. But by (4.23), we have

$$\varphi_p(\tilde{v}_1 \cdot ... \cdot \tilde{v}_r) = \varphi_p(\tilde{w}_1 \cdot ... \cdot \tilde{w}_t).$$

This means that the entries of the matrix $\tilde{v}_1 \cdot ... \cdot \tilde{v}_r$ are equal mod p to the entries of the matrix $\tilde{w}_1 \cdot ... \cdot \tilde{w}_t$. It follows that all entries of $\tilde{v}_1 \cdot ... \cdot \tilde{v}_r - \tilde{w}_1 \cdot ... \cdot \tilde{w}_t$ are divisible by p. In addition, it follows from (4.24) that at least one of the entries must be nonzero. Next, we have a result regarding the norm of a matrix, where the norm of a square matrix M is defined by

$$\|M\| = \sup_{x \neq 0} \frac{\|Mx\|}{\|x\|},$$

and the norm of the column vector $x = (x_1, x_2, ..., x_n)^T$ is $\sqrt{{x_1}^2 + {x_2}^2 + ... + {x_n}^2}$.

Lemma 4.25. *If $M = (m_{ij})$ is a $n \times n$ square matrix and $m_{ij} = k$ for some $k \in \mathbb{R}$, then $\|M\| \geq |k|$.*

Proof. It suffices to find a unit vector x in $\mathbb{R}^n$ with $\|Mx\| \geq |k|$. Let

$$x = (0, 0, \cdots, 1, 0, \cdots, 0)^T,$$

where the 1 is in the jth column. Then Mx is just the jth column of M, which has norm

$$\|Mx\| = \sqrt{{m_{1j}}^2 + ... + {m_{ij}}^2 + ... + {m_{nj}}^2} \geq \sqrt{{m_{ij}}^2} = |m_{ij}| = |k|.$$

□

It follows from the lemma that

$$\|\tilde{v}_1 \cdot ... \cdot \tilde{v}_r - \tilde{w}_1 \cdot ... \cdot \tilde{w}_t\| \geq p.$$

Using the fact that $\|M\| + \|L\| \geq \|M + L\|$ for all $n \times n$ matrices M and L, we get

$$\|\tilde{v}_1 \cdot ... \cdot \tilde{v}_r\| + \|\tilde{w}_1 \cdot ... \cdot \tilde{w}_t\| \geq \|\tilde{v}_1 \cdot ... \cdot \tilde{v}_r - \tilde{w}_1 \cdot ... \cdot \tilde{w}_t\| \geq p.$$

This implies

$$\max\{\|\tilde{v}_1 \cdot ... \cdot \tilde{v}_r\|, \|\tilde{w}_1 \cdot ... \cdot \tilde{w}_t\|\} \geq p/2 \tag{4.26}$$

The norms of the matrices A, B, A^{-1} and B^{-1} are $\alpha = 1 + \sqrt{2}$ (calculation omitted). The matrix norm is submultiplicative, meaning $\|M\|\|L\| \geq \|ML\|$ for all $n \times n$ matrices M and L. Thus by (4.26), we have $\alpha^{\max\{r,t\}} \geq p/2$, or

$$\max\{r, t\} \geq \log_\alpha(p/2). \tag{4.27}$$

What (4.27) shows is that given two walks that start at I and have a common endpoint, one of them must have length at least $\log_\alpha(p/2)$. Recall that $d(X_p)$ is defined as the smallest integer such that there exist two walks with length $\leq d(X_p)$ that start at I and have a common endpoint. Thus we must have

$$d(X_p) \geq \log_\alpha(p/2).$$

By Lemma 4.22, we have

$$girth(X_p) \geq 2\log_\alpha(p/2) - 1.$$

Thus we have obtained a lower bound for the girth of X_p. This is the basic idea behind finding a lower bound for the girth of the constructed Ramanujan graphs $X^{p,q}$, which is given in Theorem 4.20.

4.2. **Lower Bound on Chromatic Number of $X^{p,q}$.** We now find a lower bound on the chromatic number of $X^{p,q}$ by using a relationship between the chromatic number of a graph and the eigenvalues of its adjacency matrix.

Theorem 4.28 ([6]). *Let $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$ be the eigenvalues of a graph G, and let $k = \chi(G)$. Then*

$$\lambda_1 + ... + \lambda_{k-1} \leq -\lambda_n.$$

Proof. Since $k = \chi(G)$, we can color G with colors $\{1, 2, ..., k\}$. Let m_i be the number of points with color i. By labelling the m_1 vertices with color 1 $1, 2, ..., m_1$, the m_2 vertices with color 2 $m_1 + 1, m_1 + 2, ..., m_1 + m_2$ and so on, the adjacency matrix A of G can be written in the form

$$\begin{bmatrix} 0 & A_{12} & \cdots & A_{1k} \\ A_{21} & 0 & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & 0 \end{bmatrix},$$

where A_{ij} has dimension $m_i \times m_j$ and $A_{ji} = {A_{ij}}^T$.

Let $\mathbf{v}$ be an eigenvector corresponding to λ_n. We can break $\mathbf{v}$ into pieces $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$ in $\mathbb{R}^{m_1}$, $\mathbb{R}^{m_2}$, …, $\mathbb{R}^{m_k}$ respectively. Let us define

$$\mathbf{w}_i = \left.\begin{bmatrix} |\mathbf{v}_i| \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right\} m_i \text{ entries}, \quad \mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_k \end{bmatrix}.$$

Let B_i be any orthogonal matrix such that

$$B_i\mathbf{w}_i = \mathbf{v}_i$$

for $i = 1, ..., k$. Note that such an orthogonal matrix always exists, since we can construct one by setting the first column of the matrix to be $\mathbf{v}_i/\|\mathbf{v}_i\|$, extending $\mathbf{v}_i/\|\mathbf{v}_i\|$ to an orthonormal basis of $\mathbb{R}^{m_i}$ using the Gram-Schmidt process, and setting the other columns to be the other vectors in the orthonormal basis. Now let

$$B = \begin{bmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_k \end{bmatrix}.$$

Then

$$B\mathbf{w} = \begin{bmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_k \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_k \end{bmatrix} = \begin{bmatrix} B_1\mathbf{w}_1 \\ B_2\mathbf{w}_2 \\ \vdots \\ B_k\mathbf{w}_k \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_k \end{bmatrix} = \mathbf{v},$$

and

$$B^{-1}AB\mathbf{w} = B^{-1}A\mathbf{v} = \lambda_n B^{-1}\mathbf{v} = \lambda_n \mathbf{w},$$

which means that $\mathbf{w}$ is an eigenvector of $B^{-1}AB$. $B^{-1}AB$ is in the form

$$\begin{bmatrix} 0 & {B_1}^{-1}A_{12}B_2 & \cdots & {B_1}^{-1}A_{1k}B_k \\ {B_2}^{-1}A_{21}B_1 & 0 & \cdots & {B_2}^{-1}A_{2k}B_k \\ \vdots & \vdots & \ddots & \vdots \\ {B_k}^{-1}A_{k1}B_1 & {B_k}^{-1}A_{k2}B_2 & \cdots & 0 \end{bmatrix}.$$

Since the B_i are orthogonal, we have ${B_i}^{-1}A_{ij}{B_j}^T = {B_j}^T A_{ji} {B_i}^{-1}T = {B_j}^{-1}A_{ji}B_i$, so $B^{-1}AB$ is symmetric.

Form a $k \times k$ symmetric submatrix, call it D, by choosing the entry in the upper left corner of each of the k^2 submatrices ${B_i}^{-1}A_{ij}B_j$ ($A_{ii} = \mathbf{0}$). Consider the vector

$$\mathbf{u} = \begin{bmatrix} |\mathbf{v}_1| \\ |\mathbf{v}_2| \\ \vdots \\ |\mathbf{v}_k| \end{bmatrix}.$$

We claim that $\mathbf{u}$ is an eigenvector of D. We can obtain D from $B^{-1}AB$ by deleting the rows and columns that do not contain the entries in the upper left corner of each submatrix ${B_i}^{-1}A_{ij}B_j$. Let $\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_k$ be the columns of $B^{-1}AB$ that contain the "upper left corner entries", and let $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_k$ be the rows. The non-zero entries $|\mathbf{v}_i|$ of $\mathbf{w}$ are in the rows $\mathbf{r}_i$. Thus, the vector $B^{-1}AB\mathbf{w} = \lambda_n \mathbf{w}$ is just

$$(\star) \qquad |\mathbf{v}_1|\,\mathbf{c}_1 + |\mathbf{v}_2|\,\mathbf{c}_2 + ... + |\mathbf{v}_k|\,\mathbf{c}_k.$$

Let $\mathbf{c}'_1, \mathbf{c}'_2, ..., \mathbf{c}'_k$ be the columns $\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_k$ with all the rows not $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_k$ deleted. We write $D = \begin{bmatrix} \mathbf{c}'_1 & \mathbf{c}'_2 & \cdots & \mathbf{c}'_k \end{bmatrix}$, and $D\mathbf{u}$ is equal to

$$|\mathbf{v}_1|\,\mathbf{c}'_1 + |\mathbf{v}_2|\,\mathbf{c}'_2 + ... + |\mathbf{v}_k|\,\mathbf{c}'_k.$$

This expression is equal to the vector obtained by taking the rows $\mathbf{r}_i$ of $(\star)$, which is

$$\begin{bmatrix} \lambda_n|\mathbf{v}_1| \\ \lambda_n|\mathbf{v}_2| \\ \vdots \\ \lambda_n|\mathbf{v}_k| \end{bmatrix} = \lambda_n \mathbf{u}.$$

So $D\mathbf{u} = \lambda_n \mathbf{u}$, and λ_n is an eigenvalue of D.

Let $\mu_1 \leq \mu_2 \leq ... \leq \mu_k$ be the eigenvalues of D. D has all 0's in its main diagonal, so we have $\text{tr}(D) = 0$, which implies $\mu_1 + \mu_2 + ... + \mu_k = 0$, using the fact that the trace of a matrix is equal to the sum of its eigenvalues. This is equivalent to $\mu_1 + ... + \mu_{k-1} = -\mu_k$. Now λ_n is an eigenvalue of D, so $\lambda_n \leq \mu_k$. By the Interlacing Theorem, we have $\lambda_i \leq \mu_i$ for $i = 1, ..., k-1$. Therefore,

$$\lambda_1 + ... + \lambda_{k-1} \leq \mu_1 + ... + \mu_{k-1} = -\mu_k \leq -\lambda_n.$$

□

Corollary 4.29 (Hoffman. See [4, p. 276])**.** *If the eigenvalues of the adjacency matrix of a graph G are $\lambda_1 \leq \ldots \leq \lambda_n$, then*

$$\chi(G) \geq 1 - \frac{\lambda_n}{\lambda_1}.$$

Proof. We have

$$(\chi(G) - 1)\lambda_1 \leq \lambda_1 + ... + \lambda_{\chi(G)-1} \leq -\lambda_n$$

by Theorem 4.28. Since λ_n is greater than or equal to the average degree of G, which is positive (except when G is an empty graph), and $\lambda_1 + \ldots + \lambda_n = 0$ by the trace, we must have $\lambda_1 < 0$. Dividing by λ_1 gives the result. □

We now use Corollary 4.29 to obtain a lower bound on the chromatic number of constructed Ramanujan graphs. Let $X_{n,k}$ be a non-bipartite Ramanujan graph with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$ (note the change in order from the statement of the corollary). Since $X_{n,k}$ is regular, we know that $\lambda_1 = k$. Since $\lambda_n = -k$ if and only if $X_{n,k}$ is bipartite, we have $\lambda_n \neq -k$. By Definition 4.16, $|\lambda_n| \leq 2\sqrt{k-1}$, so $-\lambda_n \leq 2\sqrt{k-1}$ and $\frac{1}{-\lambda_n} \geq \frac{1}{2\sqrt{k-1}}$. By Corollary 4.29, we have

$$\chi(X_{n,k}) \geq 1 - \frac{\lambda_1}{\lambda_n} \geq \frac{k}{2\sqrt{k-1}}. \tag{4.30}$$

If we let p be a perfect square mod q, then $X^{p,q}$ is Ramanujan and non-bipartite, and is $(p+1)$-regular on $q(q^2-1)/2$ vertices. So $X^{p,q}$ can be written, in the notation of Definition 4.16, as $X_{q(q^2-1)/2,p+1}$. Therefore, by (4.30), its chromatic number can be made arbitrarily large. Along with Theorem 4.20, this shows that the Ramanujan graphs constructed in Construction 4.17 have large girth and large chromatic number.

Acknowledgments. I thank Prof. Babai for suggesting the topic for this paper, providing and discussing the material which the paper is based on, and giving help whenever needed. I also thank Yi Guo for her suggestions on the writing and formatting of the paper and help on the mathematical content. Finally, I thank Prof. May for organizing the REU and providing helpful comments for revision.

References

[1] Gabriel Lugo. Mycielski's Theorem. http://www.math.uri.edu/ eaton/mycielski.pdf
[2] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. John Wiley & Sons, Inc. 2000.
[3] Jiří Matoušek and Jan Vondrák. *The Probabilistic Method Lecture Notes*. Department of Applied Mathematics, Charles University. 2008.
[4] A. Lubotzky, R. Phillips and P. Sarnak, Ramanujan Graphs, *Combinatorica* **8**, 1988, 261-277.
[5] G.A. Margulis. Graphs without short cycles, *Combinatorica* **2**, 1982, 71-78.
[6] László Lovász. *Combinatorial Problems and Exercises*. Akadémiai Kiadó. 1979.