An explicit construction of 4-regular graphs with logarithmic girth

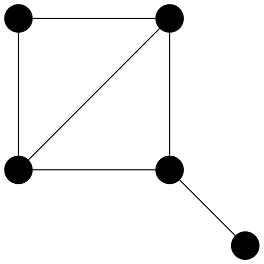
Xavier Povill Claró́s

March 31, 2024

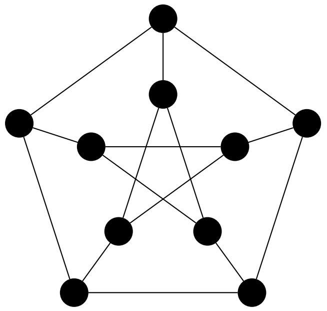
1 Motivation

Given a graph G, we define its *girth*, denoted as $g(G)$, as the minimum length of a cycle in G. Here we assume that cycles are *simple*, that is, they do not repeat edges or vertices. If G does not contain any cycles, we may set $g(G) := +\infty$. [1]

Finding graphs with high girth it's easy (for example, we can take C_n, the cycle graph with n vertices, which has $g(C_n) = n$). However, the question becomes much more interesting once we impose that the resulting graph must be k-regular (i.e. all vertices must have k neighbours). For $k = 2$, one can still use the graph C_n as an example, but as soon as $k \geq 3$, it becomes exceedingly complicated to explicitly construct a graph with high girth.



(a) A graph with girth 3



(b) A 3-regular graph with girth 5

In the previous class we saw that the girth of a k-regular graph, for $k \geq 3$, is at most logarithmic in terms of the number of vertices:

$$g(G) \leq 2 \log_{k-1} \big(|V(G)| \big) + \mathcal{O}(1)$$

where the asymptotic notation $\mathcal{O}(1)$ refers to the limit when $|V(G)| \to \infty$.

It is not known if this bound can be reached, but it has been shown that we can attain a similar bound in which we substitute the 2 in front of the logarithm by a smaller factor. More concretely, we aim to find a family of graphs $\{G_n\}_{n\geq 1}$ which is *growing* (i.e. $|V(G_n)| \to \infty$ as $n \to \infty$), and such that $g(G_n) \geq (1 + o(1))\, C \log_{k-1} \big(|V(G)| \big)$ for a certain constant $C > 0$.

[1] That will not matter though for our set-up, since all k-regular graphs have at least one cycle, provided $k >= 2$.

As early as 1963, Erdös and Sachs proved that there existed one such family for $C = 1$, but their methods were non-constructive, so they could not construct a family of graphs that exhibited this behaviour. Twenty years later, in 1982, Margulis gave the construction for $k = 4$ that we will explain today, which achieves a weaker constant of $C = \dfrac{2}{3\log_3(1+\sqrt{2})} \cong 0.831$ but has an explicit and simple construction. The same idea could also be adapted for other values of k, though with a smaller constant.

In 1983, Biggs and Hoare improved this to $C = 4/3$, but only for the case of cubic graphs (i.e. $k = 3$). This was extended in 1988 to arbitrary k by Lubotzky, Phillips and Sarnak, in their celebrated construction of the so-called *Ramanujan graphs*, which not only have high girth, but also exhibit lots of interesting expanding and spectral properties. This bound remains the best-known to date for general values of k.

2 Description of the construction

In order to properly state the construction given by Margulis, we must first introduce some group theory background:

Definition 2.1. Given a group G, we say that a subset of elements $S \subseteq G$ is *symmetric* if, for any $x \in S$, we have that $x^{-1} \in S$.

Definition 2.2 (Cayley graph)**.** Given a group G (which may be infinite) and a finite and symmetric subset $S \subseteq G$, we define the *Cayley graph* $\mathcal{G}(G, S)$ as the graph that has one vertex for every element of the group (i.e. $V(\mathcal{G}) := G$) and that has an edge between $x, y \in G$ if there exists an $s \in S$ such that $y = xs$.

Remark. The symmetry condition on S is required so that the graph we obtain is undirected. Indeed, according to the definition, we have that $(x, y) \in E(\mathcal{G})$ if there exists an $s \in S$ such that $y = xs$. That implies that $x = ys^{-1}$ so, in order to have edge (y, x) in the graph too, we need $s^{-1} \in S$.

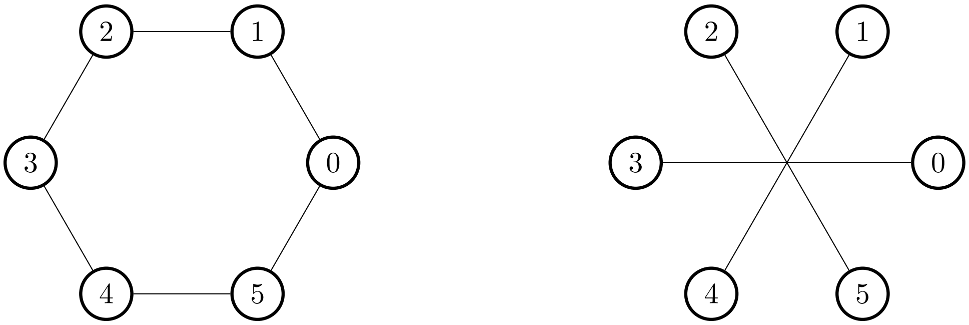
We can see some examples of Cayley graphs in figure 2. Note that taking different sets S can greatly affect the shape of the graph, even if the underlying group is the same.

The construction will be the Cayley graph of a certain group of matrices:

Definition 2.3. The *special linear group* of degree n over a field K, denoted by $\mathrm{SL}_n(K)$, is the group of $n \times n$ matrices with coefficients on K which have determinant 1. The group operation is defined to be matrix multiplication.

We will work with the following two particular cases:

- $\mathrm{SL}_2(\mathbb{F}_q) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \,:\, a, b, c, d \in \mathbb{F}_q, \text{ and } ad - bc = 1 \right\}$
- $\mathrm{SL}_2(\mathbb{Z}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \,:\, a, b, c, d \in \mathbb{Z}, \text{ and } ad - bc = 1 \right\}$



(a) Cayley graph with $G = \mathbb{Z}/6\mathbb{Z}$ and $S = \{1, 5\}$ (b) Cayley graph with $G = \mathbb{Z}/6\mathbb{Z}$ and $S = \{3\}$

Figure 2

Recall that, for q a prime power, $\mathbb{F}_q$ is the unique finite field with q elements.

We are now finally ready to describe the construction of Margulis:

Theorem 2.4. *For an odd prime p, let $G_p := \mathcal{G}(\mathrm{SL}_2(\mathbb{F}_p), S_p)$, where $S_p := \{A, A^{-1}, B, B^{-1}\}$ and*

$$A := \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

Then, $\{G_p\}_p$ is a growing family of 4-regular connected graphs with girth

$$g(G_p) \geq \frac{1}{3\log_3(1+\sqrt{2})} \log_3(|V(G_p)|) + \mathcal{O}(1)$$

We delay the proof of the girth until later, since it requires additional tools, but we already know enough to prove the 4-regularity and the connectedness:

Lemma 2.5. *For an odd prime p, G_p is 4-regular and connected.*

Proof. Let us first prove the graph is connected. We claim that the Cayley graph is connected if, and only if, the set S generates the whole group. For the forward implication, if the graph is connected then there's a path that goes from the identity matrix 1 to an arbitrary matrix $X \in \mathrm{SL}_2(\mathbb{F}_p)$. By the definition of the Cayley graph, moving along an edge is equivalent to multiplying by an element from S_p. Thus, for any $X \in \mathrm{SL}_2(\mathbb{F}_p)$ there is a finite sequence $s_1, \ldots, s_k \in S_p$ such that $1 \cdot s_1 \cdot s_2 \cdots s_k = X$, and hence $X \in \langle S_p \rangle$.

For the backwards implication, if $\langle S_p \rangle = SL_2(\mathbb{F}_p)$, then for any $X, Y \in \mathrm{SL}_2(\mathbb{F}_p)$, $X^{-1}Y \in \langle S_p \rangle$, so there exist $s_1, \ldots, s_k \in S_p$ such that $s_1 \cdot s_2 \cdots s_k = X^{-1}Y$. Then, we can get from vertex X to vertex Y of the graph by the path given by the edges labelled with $s_1, s_2, \ldots, s_k$. That means that there exists a finite path connecting every pair of vertices, so the graph is connected.

Using that, we only need to show that S_p generates the whole group. Note that the powers of A and B take the following form:

$$A^{\pm k} = \begin{pmatrix} 1 & \pm 2k \\ 0 & 1 \end{pmatrix}, \quad B^{\pm k} = \begin{pmatrix} 1 & 0 \\ \pm 2k & 1 \end{pmatrix},$$

The coefficients of the matrices are taken modulo p, which is odd, so we can get any triangular matrix with 1's at the diagonal:

$$\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^b = \left(\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}^{\frac{p+1}{2}}\right)^b = A^{\left(\frac{p+1}{2}\right)b}$$
$$\begin{pmatrix}1 & 0\\ c & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}^c = \left(\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}^{\frac{p+1}{2}}\right)^c = B^{\left(\frac{p+1}{2}\right)c}$$

Then, we construct a general matrix the following way:

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

The equality can be easily verified by multiplying the matrices and then using that $ad - bc = 1$, so $b = \frac{ad-1}{c}$. However, the above construction fails if $c = 0$ (because then it has no inverse in $\mathbb{F}_p$, so we can not divide by it). To circumvent this issue, note that if $c = 0$ then $d \neq 0$ (otherwise the determinant would be 0). Therefore, we can express a matrix with $c = 0$ as a product of matrices with $c \neq 0$, and we apply the previous construction to those:

$$\begin{pmatrix}a & b\\ 0 & d\end{pmatrix} = \begin{pmatrix}a+b & b\\ d & d\end{pmatrix}\begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix}$$

We have shown how to get a general matrix from multiplying elements from S_p, so that completes the proof of the connectedness. It remains to show that the graph is 4-regular, but that is much simpler. Notice that a Cayley graph is $|S|$-regular by definition, since multiplying to the right by two different elements can not give us the same result (we are in a group, so every element has an inverse). The 4 matrices A, A^{-1}, B, B^{-1} are different in any $\mathbb{F}_p$, because if we take two of these matrices and make the difference between each pair of corresponding coefficients, we get at least one of the following differences: $2, -2, 4, -4$, none of which divisible by an odd prime. Hence, any vertex in G_p has $|S_p| = 4$ neighbors. □

It only remains to prove that the G_p have high girth. In order to do so, we will need to introduce some algebraic tools.

3 Free groups and the ping-pong lemma

Definition 3.1. Let G be a group, with generating set S. A *word* on G (with respect to S) is a finite product of elements from S. That is, a word is an expression of the form $w = g_1 g_2 \dots g_t$, where $t \geq 1$ and $g_i \in S$ for all $i \in [t]$.

We say that a word is *reduced* if there are no two consecutive elements which cancel each other (i.e. one of which is the inverse of the other).

We say that a group G is *free with basis* S, if there is no reduced word that is trivial (i.e. equal to 1).

Intuitively, a group is free if there is no way to multiply several elements to get 1 unless they directly cancel by pairs of inverses. For example, $\mathrm{SL}_2(\mathbb{F}_q)$ is not free with respect to the basis $S = \{A, B, A^{-1}, B^{-1}\}$, since we have the trivial reduced word $w = \underbrace{A \cdot A \cdots A}_{p \text{ times}} = 1$.

However, if we do not consider reduction modulo p, the resulting subgroup is indeed free:

Lemma 3.2. *The subgroup $\langle A, A^{-1}, B, B^{-1} \rangle \leq \mathrm{SL}_2(\mathbb{Z})$ is free with respect to the basis $S = \{A, A^{-1}, B, B^{-1}\}$.*

Proof. The proof is based on a method called the *Ping-Pong Lemma*, though we will not state it in full generality because it would require further background in group theory, which is not necessary for this particular application.

The idea is to consider the elements of our group, which are 2×2 matrices, as linear operators on the vector space $\mathbb{R}^2$. Consider the following two subsets of $\mathbb{R}^2$:

$$X := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,:\, |x| > |y| \right\}$$
$$Y := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \,:\, |x| < |y| \right\}$$

Observe that multiplying by a power of A moves the vectors from Y to X, while multiplying by a power of B moves the vectors from X to Y. More formally, we have that for any $k \neq 0$ and for any $v \in Y$, $A^k v \in X$, and analogously, for any $v \in X$, $B^k v \in Y$. We will only prove the first claim, as the second one follows by an analogous argument.

Let $k \neq 0$ and let $v = (x, y)^T \in Y$ (that is, $|y| > |x|$). Then,

$$A^k y = \begin{pmatrix} 1 & 2k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2ky \\ y \end{pmatrix}$$

And, using the triangular inequality,

$$|x + 2ky| \geq |2ky| - |-x| = 2\,|k|\,|y| - |x| > (2\,|k| - 1)\,|y| \geq |y|$$

Hence, $A^k v \in X$.

Why are these claims useful? Suppose we have a reduced word w, which we will suppose for now that starts and ends with a power of A. Then, the word has the form $w = A^{k_1} B^{k_2} \dots A^{k_t}$, where $t \geq 1$ and $k_1, \dots, k_t \neq 0$. If the word is trivial (i.e. $w = 1$), then $A^{k_1} B^{k_2} \dots A^{k_t} v = v$ for any $v \in \mathbb{R}^2$. However, suppose we take $v \in Y$. By the previous claims we have:

$$\begin{aligned} v &\in Y \\ A^{k_t} v &\in X \\ B^{k_{t-1}} A^{k_t} v &\in Y \\ A^{k_{t-2}} B^{k_{t-1}} A^{k_t} v &\in X \\ &\vdots \\ A^{k_1} B^{k_2} \cdots A^{k_{t-2}} B^{k_{t-1}} A^{k_t} v &\in X \end{aligned}$$

Hence, starting from a $v \in Y$ we end up with $A^{k_1}B^{k_2}\dots A^{k_t}v \in X$. If the word was trivial, then we would have both $v \in X$ and $v \in Y$, which is impossible, since X and Y are disjoint, so we reach a contradiction.

This last argument proves that no reduced word starting and ending with a power of A can be trivial. The cases in which the reduced word does not start or end with a power of A can be easily reduced to this case by conjugating by an appropriate power of A. Indeed, take $k \neq 0$ such that $|k| \neq |k_0|$ and $|k| \neq |k_t|$. Then, consider the word $\tilde{w} := A^k w A^{-k}$. It's easy to see that $\tilde{w} = 1 \iff w = 1$, and that (after cancelling out everything that can be cancelled out) $\tilde{w}$ reduces to a reduced word which starts and ends with a power of A. Therefore, by the previous result $\tilde{w} \neq 1$, and hence $w \neq 1$. □

4 Matrix norms

The final ingredient we will need in the proof is some properties of matrix norms.

Definition 4.1. Let $||\cdot||$ be the usual Euclidean norm for vectors of $\mathbb{R}^2$. Then, given a real matrix $T \in \mathcal{M}^{2\times 2}(\mathbb{R})$, we define its *operator norm* as

$$||\,T\,|| := \sup_{v\neq 0} \frac{||\,Tv\,||}{||\,v\,||}$$

Intuitively, the operator norm $||\,T\,||$ measures how much we can lengthen a vector $v \in \mathbb{R}^2$ by applying T, in terms of the original length of v.

Lemma 4.2. *Let $T \in \mathcal{M}^{2\times 2}(\mathbb{R})$. The operator norm satisfies the following properties:*

1. *For all $v \in \mathbb{R}^2$, $||\,Tv\,|| \leq ||\,T\,|| \cdot ||\,v\,||$*
2. *For all other $\tilde{T} \in \mathcal{M}^{2\times 2}(\mathbb{R})$, $||\,T\tilde{T}\,|| \leq ||\,T\,|| \cdot ||\,\tilde{T}\,||$*
3. *$||\,T\,|| = ||\,T^t\,||$*
4. *$||\,TT^t\,|| = ||\,T\,||^2$*
5. *$||\,T\,|| \geq \max\limits_{i,j} |T_{ij}|$*
6. *If T is symmetric, then $||\,T\,|| = \max\limits_{\lambda\in\mathrm{Spec}(T)} |\lambda|$*

Proof.

1. If $v = 0$, then both sides are zero, so the inequality is satisfied. Otherwise,

$$\frac{||\,Tv\,||}{||\,v\,||} \leq \sup_{w\neq 0} \frac{||\,Tw\,||}{||\,w\,||} = ||\,T\,||$$

2. Using the previous one, we have:

$$|| T\tilde{T} || = \sup_{v\neq 0} \frac{|| T\tilde{T}v ||}{|| v ||} \leq \sup_{v\neq 0} \frac{|| T || \cdot || \tilde{T}v ||}{|| v ||} = || T || \cdot || \tilde{T} ||$$

3. Using Cauchy-Schwarz, we have that

$$|| T^t v ||^2 = v^t T T^t v \leq || v || \cdot || T T^t v || \leq || v || \cdot || T || \cdot || T^t v ||$$

Hence, $|| T^t v || \leq || T || \cdot || v ||$. Now, plugging that into the definition of $|| T^t ||$ we obtain

$$|| T^t || = \sup_{v\neq 0} \frac{|| T^t v ||}{|| v ||} \leq || T ||$$

Then, by symmetry, we also have that $|| T || = || (T^t)^t || \leq || T^t ||$, so $|| T || = || T^t ||$.

4. Using 2 and 3, we have that $|| T^t T || \leq || T^t |||| T || = || T ||^2$. For the reverse inequality, we use Cauchy-Schwarz and 1:

$$|| T ||^2 = \sup_{v\neq 0} \frac{|| Tv ||}{|| v ||} = \sup_{v\neq 0} \frac{v^t T^t T v}{|| v ||^2} \leq \sup_{v\neq 0} \frac{|| v || \cdot || T^t T v ||}{|| v ||^2} \leq$$

$$\leq \sup_{v\neq 0} \frac{|| v || \cdot || T^t T || \cdot || v ||}{|| v ||^2} = || T^t T ||$$

5. Let e_i be the i-th vector of the canonical basis. Then, note that $e_i^t T e_j = T_{ij}$. Therefore, for all $i, j \in \{1, 2\}$,

$$|T_{ij}| = \left| e_i^t T e^j \right| \leq || e_i || \cdot || T e_j || = || T e_j || \leq || T || \cdot || e_j || = || T ||$$

6. Let λ be an eigenvalue of T, with eigenvector v. Then,

$$|| T || = \sup_{w\neq 0} \frac{|| Tw ||}{|| w ||} \geq \frac{|| Tv ||}{|| v ||} = \frac{|| \lambda v ||}{|| v ||} = |\lambda|$$

Hence, $|| T || \geq \max_{\lambda \in \mathrm{Spec}(T)} |\lambda|$. To prove the converse, we use that since T is symmetric there exists an orthonormal basis of eigenvectors $\{u_1, u_2\}$. Therefore, for any $v \in \mathbb{R}^2$ there exist $\alpha_1, \alpha_2 \in \mathbb{R}$ such that $\alpha_1 u_1 + \alpha_2 u_2 = v$. Using that basis, we can express the operator norm as:

$$|| T || = \sup_{v\neq 0} \frac{|| Tv ||}{|| v ||} = \sup_{\substack{\alpha_1, \alpha_2 \in \mathbb{R} \\ \alpha_1 u_1 + \alpha_2 u_2 \neq 0}} \frac{|| T(\alpha_1 u_1 + \alpha_2 u_2) ||}{|| \alpha_1 u_1 + \alpha_2 u_2 ||} = \sup_{\substack{\alpha_1, \alpha_2 \in \mathbb{R} \\ \alpha_1 u_1 + \alpha_2 u_2 \neq 0}} \frac{|| \alpha_1 \lambda_1 u_1 + \alpha_2 \lambda_2 u_2 ||}{|| \alpha_1 u_1 + \alpha_2 u_2 ||}$$

Now, note that for any $c_1, c_2 \in \mathbb{R}$ we have that $|| c_1 u_1 + c_2 u_2 || = \sqrt{(c_1 u_1 + c_2 u_2)^t (c_1 u_1 + c_2 u_2)} = \sqrt{c_1^2 + c_2^2}$, using the orthonormality of the basis. Hence,

$$|| T || = \sup_{\substack{\alpha_1, \alpha_2 \in \mathbb{R} \\ \alpha_1 u_1 + \alpha_2 u_2 \neq 0}} \frac{\sqrt{\alpha_1^2 \lambda_1^2 + \alpha_2^2 \lambda_2^2}}{\sqrt{\alpha_1^2 + \alpha_2^2}} \leq \sup_{\substack{\alpha_1, \alpha_2 \in \mathbb{R} \\ \alpha_1 u_1 + \alpha_2 u_2 \neq 0}} \frac{\max\{|\lambda_1|, |\lambda_2|\} \cdot \sqrt{\alpha_1^2 + \alpha_2^2}}{\sqrt{\alpha_1^2 + \alpha_2^2}} =$$

$$= \max\{|\lambda_1|, |\lambda_2|\}$$

□

With the previous properties, it is now easy to compute the norm of A, A^{-1}, B and B^{-1}. First, note that by property 4, $||\,A\,|| = \sqrt{||\,A^tA\,||}$. By a simple computation, we find that A^tA has characteristic polynomial $x^2 - 6x + 1$, which has roots $x = 3 \pm \sqrt{8}$, which will be the eigenvalues of A^tA. Therefore, using property 6, we have that $||\,A\,|| = \sqrt{3+\sqrt{8}} = 1 + \sqrt{2}$. Since $B = A^t$, by property 3 it has the same norm, and similarly one can find that A^{-1} and B^{-1} also have norm $1 + \sqrt{2}$.

5 Finishing up the proof

We want to show that the family of graphs we constructed have no short cycles. Fix an odd prime p, and let $v_1v_2 \dots v_g$ be a simple cycle (i.e. with no repeated vertices) of length g in the graph G_p.

By definition of G_p, two vertices X and Y are adjacent if there exists a matrix $M \in \{A, A^{-1}, B, B^{-1}\}$ such that $Y = XM$. Thus, taking the corresponding matrix from each edge of the cycle, we obtain a sequence of g matrices $M_1, \dots, M_g \in \{A, A^{-1}, B, B^{-1}\}$ such that $X = XM_1M_2 \dots M_g$ for a certain $X \in \mathrm{SL}_2(\mathbb{F}_p)$. Multiplying by X^{-1}, we obtain that

$$M_1M_2 \dots M_g = 1$$

Besides, note that no two adjacent matrices in this product can be the inverse of one another, as that would correspond to going through the same edge twice in a row (which is forbidden in a simple cycle). Hence, the above is a trivial reduced word in $\mathrm{SL}_2(\mathbb{F}_p)$.

Now comes the key idea of the proof. Lift the above product to $\mathrm{SL}_2(\mathbb{Z})$ (i.e. consider the same product of matrices but now in the integers, without taking modulo p). We saw in Lemma 3.2 that $\langle A, A^{-1}, B, B^{-1} \rangle$ is a free subgroup of $\mathrm{SL}_2(\mathbb{Z})$, so no reduced word written with these 4 matrices can be trivial. Hence, $M_1M_2 \dots M_g \neq 1$ in $\mathrm{SL}_2(\mathbb{Z})$.

That means that one of the coefficients of $M_1M_2 \dots M_g$ is equal to 1 (or 0) in $\mathbb{F}_p$ but not in $\mathbb{Z}$. If it was one of the diagonal coefficients, then that would mean that it is of the form $kp \pm 1$, with $k \neq 0$, so its absolute value would be at least $|kp \pm 1| \geq p - 1$. If it was one of the off-diagonal coefficients, then that would mean that it is of the form kp, with $k \neq 0$, so its absolute value would be at least $|kp| \geq p$.

In any of the two cases, we have that the matrix $M := M_1 \dots M_g$ has $||\,M\,|| \geq \max\limits_{i,j} |M_{ij}| \geq p-1$.

On the other hand, we saw that each of the matrices $M_i \in \{A, A^{-1}, B, B^{-1}\}$ have norm $||\,M_i\,|| = 1 + \sqrt{2}$, so

$$||\,M\,|| = ||\,M_1 \dots M_g\,|| \leq ||\,M_1\,|| \cdots ||\,M_g\,|| = (1+\sqrt{2})^g$$

Putting both bounds together, we obtain that

$$p - 1 \leq ||\,M\,|| \leq (1+\sqrt{2})^g$$

Then, taking $\log_3$ at both sides we obtain the desired lower bound on the girth:

$$\frac{1}{\log_3(1+\sqrt{2})} \log_3 (p-1) \leq g$$

To write it in terms of the number of vertices of the graph, we recall from the previous class that $|V(G_p)| = |\mathrm{SL}_2(\mathbb{F}_p)| = (p^2-1)p \leq (p-1)^3/4$. Therefore, $\log_3(|V(G_p)|) = \mathcal{O}(1) + \frac{1}{3}\log_3(p-1)$, so we obtain the following bound:

$$\frac{1}{3\log_3(1+\sqrt{2})} \log_3(|V(G_p)|) + \mathcal{O}(1) \leq g(G_p)$$

Remark. We can slightly alter the previous argument to achieve the better constant $C = \frac{2}{3\log_3(1+\sqrt{2})}$. The idea is that we break the product $M_1 \dots M_g = 1$ into two parts, so that

$$M_1 \dots M_{g/2} = M_{g/2+1}^{-1} \dots M_g^{-1}$$

Here we have supposed by simplicity that g is even, but the same argument would work for g odd. Therefore,

$$M_1 \dots M_{g/2} - M_{g/2+1}^{-1} \dots M_g^{-1} = 0$$

Using the same freeness argument as before, we get that one of the coefficients of the matrix in the left-hand-side must be 0 in $\mathbb{F}_p$ but not in $\mathbb{Z}$. Hence,

$$p \leq ||\, M_1 \dots M_{g/2} - M_{g/2+1}^{-1} \dots M_g^{-1} \,|| \leq ||\, M_1 \dots M_{g/2} \,|| + ||\, M_{g/2+1}^{-1} \dots M_g^{-1} \,||$$

By the pidgeonhole principle, one of the terms on the right must be at least $p/2$. Assume it is the first one (it does not matter for what follows). Then,

$$\frac{p}{2} \leq ||\, M_1 \dots M_{g/2} \,|| \leq (1+\sqrt{2})^{g/2}$$

so, taking logarithms on both sides,

$$\frac{1}{3}\log_3(|V(G_p)|) + \mathcal{O}(1) \leq \frac{g}{2}\log_3(1+\sqrt{2})$$

and rearranging:

$$g \geq \frac{2}{3\log_3(1+\sqrt{2})} \log_3(|V(G_p)|) + \mathcal{O}(1)$$