ON THE DIAMETER OF CAYLEY GRAPHS OF FINITE GROUPS

YAN SHUO TAN

Abstract. In this paper, we investigate the notion of the diameter of a group with respect to a given set of generators. In particular, we will give bounds for the diameter of the symmetric group S_n with respect to a transposition and a long cycle. We also bound the maximum diameter of S_n with respect to any set of generators. To prepare for the latter task, we will look at the maximum order of a group element in S_n. Finally, we will conclude with a conjecture on the true bound for the diameter of all non-abelian finite simple groups, and discuss some recent progress made on the subject.

Contents

1. Introduction

What is the maximum number of moves required to solve a Rubik's Cube puzzle from any starting position? Since there are $8! \cdot 3^7 \cdot 12!/2 \cdot 2^{11}$ possible permutations, and only 12 options for each move, a simple counting argument tells us that this number, known as *God's Number*, is at least 18. More accurate estimates, however, are much more elusive, and it required 35 CPU years of computer time to prove that the actual value of God's Number is 20. Now, since Rubik's Cube permutations form a group, there is an associated Cayley graph. It turns out that God's Number is the diameter of this Cayley graph, which raises the question: Can we easily estimate the diameters of Cayley graphs of groups that we know more about? In this paper, we shall use elementary combinatorics and group theory to provide some asymptotic bounds.

2. Preliminaries

In this section, we will formalize the notion of group diameter, and review some pertinent concepts from discrete mathematics.

Definition 2.1. Let G be a group and S a set of generators. The *Cayley Graph* $\Gamma(G, S)$ is a directed graph constructed as follows:

- The vertex set $V(\Gamma)$ is G.
- The edge set $E(\Gamma)$ consists of all ordered pairs (g, gs) such that g is in G and s is in S.

We are interested in the undirected version of the Cayley graph, where the edge set consists instead of the unordered pairs $\{g, gs\}$ such that $g \in G$ and $s \in S$. From now on, any reference to Cayley graphs assumes this definition.

Definition 2.2. Let Γ be a (undirected) graph, and v, w two vertices in $V(\Gamma)$. The *distance* between v and w, denoted $d(v, w)$, is the number of edges in a shortest path connecting them.

Definition 2.3. Let Γ be a (undirected) graph. The *diameter* of Γ is the maximum distance between any two vertices in $V(\Gamma)$.

One suspects that the diameter of a Cayley graph relates strongly to the structure of its underlying group, and indeed this is true. We introduce two more definitions to make the connection clear.

Definition 2.4. Let G be a group and S a set of generators. Let g be an element in G. The *length* of g in S is the minimum word length expressing g as a product of elements of $S \cup S^{-1}$. We denote it by $\text{length}(g, S)$.

Definition 2.5. Let G be a group and S a set of generators. The diameter of G with respect to S is defined to be the maximum length of its elements. In other words,

$$\text{diam}(G, S) = \max_{g \in G} \text{length}(g, S).$$

We now see that the group theoretic definition of diameter corresponds nicely to the graph theoretic one.

Proposition 2.6. *Let G be a group, S a set of generators, and $\Gamma(G, S)$ its associated Cayley graph. Then* $\text{diam}(G, S) = \text{diam}\Gamma(G, S)$.

Proof. There is a one-to-one correspondence between words in $S \cup S^{-1}$ and walks in $\Gamma(G, S)$, where the word length is the same as the walk length. In particular, the shortest word corresponds to the shortest path, which implies

$$d(g, h) = \text{length}(g^{-1}h, S).$$

Setting k to be $g^{-1}h$, it follows that

$$\text{diam}\Gamma(G, S) = \max_{g,h \in G} d(g, h) = \max_{k \in G} \text{length}(k, S) = \text{diam}(G, S).$$

□

In view of this proposition, the following is a simple consequence of the graph theoretic triangle inequality:

Corollary 2.7. *Let G be a group, S a set of generators, then for all $g, h \in G$,* $\text{length}(gh, S) \leq \text{length}(g, S) + \text{length}(h, S)$.

Let us now introduce some convenient notation and conventions for dealing with permutation groups.

Notation 2.8. We use $[n] = \{1, 2, \ldots, n\}$ to denote the domain of a permutation group acting on n elements.

Notation 2.9. If π is a permutation, and $x \in [n]$, then $x^\pi = \pi(x)$. In accordance with the exponent notation, we will compose permutations on the right. In other words, given permutations π and σ, $\pi\sigma$ is the permutation that sends x to $(x^\pi)^\sigma = \sigma(\pi(x))$.

Notation 2.10. Suppose we have two permutations π and σ. Then $\pi^\sigma = \sigma^{-1}\pi\sigma$.

Notation 2.11. Let c be a cyclic permutation. If $c(i_1) = i_2$, $c(i_2) = i_3$, $\ldots$ $c(i_r) = i_1$, then we denote c by $(i_1 i_2 \ldots i_r)$.

Within a family of groups, the diameter of a group is often a function of its order (or degree in the case of permutation groups). We are interested in studying the rate of growth of these functions, and we formalize rate of growth comparisons with the following concepts.

Definition 2.12. Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions. We say that

(1) $\sim g$, or f is *asymptotically equal* to g, if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$,
(2) $f = o(g)$, or f is *"little oh"* of g, if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$,
(3) $f = O(g)$, or f is *"big oh"* of g, if $\left|\frac{f(n)}{g(n)}\right|$ is bounded,
(4) $f = \Omega(g)$, or f is *"omega"* of g, if $g = O(f)$, and
(5) $f = \Theta(g)$, or f is *"theta"* of g, if $f = O(g)$ and $f = \Omega(g)$.

The following properties of the asymptotic equality relation follow easily from the definitions. They show that under some conditions, asymptotic equality behaves like regular equality in the sense that it is preserved by arithmetic operations as well as taking logarithms and series limits.

Proposition 2.13. *Let $f, g, h, k : \mathbb{N} \to \mathbb{R}$ be functions such that $f \sim g$ and $h \sim k$. Then $f \cdot h \sim g \cdot k$.*

Proof.

$$\lim_{n\to\infty} \frac{f(n)h(n)}{g(n)k(n)} = \lim_{n\to\infty} \frac{f(n)}{g(n)} \cdot \lim_{n\to\infty} \frac{h(n)}{k(n)} = 1.$$

□

Corollary 2.14. *Given the same f, g, h, k, if $h(n)k(n) \neq 0$ for all sufficiently large n, then $f/g \sim k/h$.*

Proposition 2.15. *Given the same f, g, h, k, if $f(n)h(n) > 0$ for all sufficiently large n, then $f + h \sim g + k$.*

Proof. First, observe that $f(n)h(n) > 0$ implies that all four functions have the same sign eventually, and we can assume without loss of generality that they are all positive. Now suppose $a, b, c, d > 0$ and $a/b < c/d$, then we have

$$\frac{b}{a} \cdot \frac{a+c}{b+d} = \frac{ab+bc}{ab+ad} > \frac{ab+ad}{ab+ad} = 1 = \frac{bc+cd}{bc+cd} > \frac{ad+cd}{bc+cd} = \frac{d}{c} \cdot \frac{a+c}{b+d},$$

which implies that $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$. Hence,

$$\min\left\{\frac{f(n)}{g(n)}, \frac{h(n)}{k(n)}\right\} < \frac{f(n)+h(n)}{g(n)+k(n)} < \max\left\{\frac{f(n)}{g(n)}, \frac{h(n)}{k(n)}\right\},$$

so $\lim_{n\to\infty} \frac{f(n)+h(n)}{g(n)+k(n)} = 1$ by the squeeze theorem. □

Proposition 2.16. *Let $f, g, h : \mathbb{N} \to \mathbb{R}$ be functions such that $f \sim g + h$ and $h = o(f)$ or $h = o(g)$. Then $f \sim g$.*

Proof. If $h(n) = o(f(n))$, then

$$1 = \lim_{n\to\infty} \frac{g(n)+h(n)}{f(n)} = \lim_{n\to\infty} \frac{g(n)}{f(n)} + \lim_{n\to\infty} \frac{h(n)}{f(n)} = \lim_{n\to\infty} \frac{g(n)}{f(n)}.$$

On the other hand, if $h(n) = o(g(n))$,

$$1 = \lim_{n\to\infty} \frac{g(n)+h(n)}{f(n)} = \lim_{n\to\infty} \frac{g(n)}{f(n)} \cdot \lim_{n\to\infty} \left(1 + \frac{h(n)}{g(n)}\right) = \lim_{n\to\infty} \frac{g(n)}{f(n)}.$$

□

Proposition 2.17. *Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions such that $f \sim g$ and $\lim_{n\to\infty} f(n) = \lim_{n\to\infty} g(n) = \infty$. If we define $h, k : \mathbb{N} \to \mathbb{R}$ by $h(n) = \sum_{i=1}^{n} f(i)$ and $k(n) = \sum_{i=1}^{n} g(i)$, then $h \sim k$.*

Proof. Fix $\epsilon > 0$ and pick N such that for all $n \geq N$,

$$\left|\frac{f(n)}{g(n)} - 1\right| = \left|\frac{f(n) - g(n)}{g(n)}\right| < \epsilon. \tag{2.18}$$

Then

$$\left|\frac{\sum_{k=N}^{n} f(k)}{\sum_{k=N}^{n} g(k)} - 1\right| = \left|\frac{\sum_{k=N}^{n} (f(k) - g(k))}{\sum_{k=N}^{n} g(k)}\right| \leq \frac{\sum_{k=N}^{n} |f(k) - g(k)|}{\sum_{k=N}^{n} g(k)}.$$

We want the quantity on the right to be less than ϵ. But that is equivalent to

$$\sum_{k=N}^{n} |f(k) - g(k)| < \sum_{k=N}^{n} \epsilon \cdot g(k),$$

which follows from (2.18). Since we can always remove a finite number of terms from an infinite sum which is tending to infinity, while n was arbitrary, the result follows. □

Proposition 2.19. *Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions such that $f \sim g$ and both f and g are bounded away from 1 (i.e. there exists a constant c such that $f(n) > 1 + c$ or $f(n) < 1 - c$ for large n). Then $\ln f \sim \ln g$.*

Proof. Fix $0 < \epsilon < 1$. Then by hypothesis, for large n, we have $1 - \epsilon < \frac{f(n)}{g(n)} < 1 + \epsilon$. Since log is a monotone function, this implies $\ln(1-\epsilon) < \ln f(n) - \ln g(n) < \ln(1+\epsilon)$, or $\ln(1 - \epsilon) + \ln g(n) < \ln f(n) < \ln(1 + \epsilon) + \ln g(n)$. Dividing by $\ln g(n)$ yields

$$\frac{\ln(1-\epsilon)}{\ln g(n)} + 1 < \frac{\ln f(n)}{\ln g(n)} < \frac{\ln(1+\epsilon)}{\ln g(n)} + 1.$$

We know that $\ln g(n)$ is bounded away from 0. On the other hand, we can choose ϵ arbitrarily small by making n large enough. Hence, the result follows from the squeeze theorem. □

Remark 2.20. Note that $f \sim g$ need not imply that $\ln f \sim \ln g$ if f and g are not bounded away from 1. For instance, let $f(n) = n^{1/n}$ and $g(n) = n^{1/n^2}$.

The following proposition gives a handy characterization of asymptotic equality.

Proposition 2.21. *Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions. Then $f \sim g$ if and only if $f = g \cdot (1 + o(1))$.*

Proof.

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1 \iff \lim_{n\to\infty} \left(\frac{f(n)}{g(n)} - 1\right) = 0 \iff \frac{f(n)}{g(n)} - 1 = o(1).$$

□

Remark 2.22. If $f, g : \mathbb{N} \to \mathbb{R}$ are functions such that $\ln f \sim g$, we say that $f = g^{1+o(1)}$.

Using the asymptotic equality relation, we can give a nice analytic formula for otherwise unwieldy step functions. The canonical example of this is the Prime Number Theorem, which was proved independently by both Hadamard and de la Valle-Poussin in 1896. We shall make extensive use of this result later in the paper.

Theorem 2.23 (The Prime Number Theorem)**.** *Let $\pi(x)$ be the number of primes less than or equal to $x \in \mathbb{N}$. Then $\pi(x) \sim \frac{x}{\ln x}$.*

Corollary 2.24. *Let p_k be the k-th prime. Then $p_k \sim k \ln k$.*

Proof. By the PNT,

$$k \sim \frac{p_k}{\ln p_k} \iff p_k \sim k \ln p_k. \tag{2.25}$$

Taking log of both sides, we get

$$\ln p_k \sim \ln(k \ln p_k) = \ln k + \ln(\ln p_k) \sim \ln k, \tag{2.26}$$

where the last asymptotic equality follows from Proposition 2.16. Finally, we combine (2.25) and (2.26) using Proposition 2.13 to conclude that $p_k \sim k \ln k$. □

3. The diameter of S_n given some generators

To get a taste of some of the techniques used to calculate the diameter of groups, it helps to first look at an example. A natural place to start seems to be permutation groups, whose structure is relatively well-understood. So let us try giving asymptotic bounds for the symmetric group S_n given the set of generators $\{(12), (12\ldots n)\}$.

Given any group G, one method for obtaining an upper bound for its diameter (with respect to a fixed set of generators) is to define an algorithm that takes an arbitrary element $g \in G$ and finds a word for g of length less than cn^2, where c is an absolute constant independent of n. In this particular case, we will decompose a permutation into transpositions and then generate each of these transpositions one by one.

Observe that every permutation has a unique cycle decomposition. If we can prove that permutations of the same cycle decomposition form a conjugacy class, then, in particular, we will know that all transpositions are conjugates of (12), and we can then bound the length of an arbitrary transposition by finding the length of the permutation it has to be conjugated with to obtain (12). Hence, let us begin with the following useful lemma:

Lemma 3.1. *Let $\sigma, \pi \in S_n$. Then $\sigma^{-1}\pi\sigma$ has the same cycle structure as π. That is, if we write $\pi = (i_{1_1} i_{1_2} \dots i_{1_{r(1)}}) \cdots (i_{k_1} i_{k_2} \dots i_{k_{r(k)}})$, then we have*

$$\sigma^{-1}\pi\sigma = (i^\sigma_{1_1} i^\sigma_{1_2} \dots i^\sigma_{1_{r(1)}}) \cdots (i^\sigma_{k_1} i^\sigma_{k_2} \dots i^\sigma_{k_{r(k)}}).$$

Proof. Every permutation is the product of a series of transpositions, and we can write $\sigma = \tau_1\tau_2\cdots\tau_s$. Then $\sigma^{-1}\pi\sigma = \tau_s \dots \tau_2\tau_1\pi\tau_1\tau_2\cdots\tau_s$, and it suffices to prove the lemma for conjugation of π by a transposition τ. Suppose τ swaps a and b. Then we have 3 cases.

Case 1: $a^\pi = b$ and $b^\pi = a$. Clearly $\tau\pi\tau = \pi$, so there is nothing to prove here.

Case 2: $a^\pi = b$ and $b^\pi \neq a$. Since π fixes neither a nor b, we get

$$\begin{array}{ccccccc}
a & \overset{\tau}{\longmapsto} & b & \overset{\pi}{\longmapsto} & b^\pi & \overset{\tau}{\longmapsto} & b^\pi \\
b & \overset{\tau}{\longmapsto} & a & \overset{\pi}{\longmapsto} & b & \overset{\tau}{\longmapsto} & a \\
a^{\pi^{-1}} & \overset{\tau}{\longmapsto} & a^{\pi^{-1}} & \overset{\pi}{\longmapsto} & a & \overset{\tau}{\longmapsto} & b \\
c & \overset{\tau}{\longmapsto} & c & \overset{\pi}{\longmapsto} & c^\pi & \overset{\tau}{\longmapsto} & c^\pi,
\end{array}$$

where the last mapping holds for any $c \in [n]$ other than a, b, or $a^{\pi^{-1}}$.

Case 3: $a^\pi \neq b$ and $b^\pi \neq a$. Possibly, π may fix either a or b, but assume first that it fixes neither. We then have

$$\begin{array}{ccccccc}
a & \overset{\tau}{\longmapsto} & b & \overset{\pi}{\longmapsto} & b^\pi & \overset{\tau}{\longmapsto} & b^\pi \\
b & \overset{\tau}{\longmapsto} & a & \overset{\pi}{\longmapsto} & a^\pi & \overset{\tau}{\longmapsto} & a^\pi \\
a^{\pi^{-1}} & \overset{\tau}{\longmapsto} & a^{\pi^{-1}} & \overset{\pi}{\longmapsto} & a & \overset{\tau}{\longmapsto} & b \\
b^{\pi^{-1}} & \overset{\tau}{\longmapsto} & b^{\pi^{-1}} & \overset{\pi}{\longmapsto} & b & \overset{\tau}{\longmapsto} & a \\
c & \overset{\tau}{\longmapsto} & c & \overset{\pi}{\longmapsto} & c^\pi & \overset{\tau}{\longmapsto} & c^\pi.
\end{array}$$

On the other hand, if π fixes a, then $\tau\pi\tau$ clearly fixes b, while $b^{\pi^{-1}}$ is still mapped to a, and a to b^π by above. If π fixes b, we conclude analogously. □

We can now prove

Proposition 3.2. *Let S be $\{(12), (12\dots n)\}$. Then* $\operatorname{diam}(S_n, S) = O(n^2)$.

Proof. We first give a bound for the number of steps required to generate an arbitrary transposition (ab), where, without loss of generality, we can assume $a < b$. Let $\pi = (12)$, $\sigma = (12\dots n)$, $b = a+k$, and consider $\pi^{(\sigma\pi)^{k-1}\sigma^{a-1}}$, i.e. the conjugate of π by $(\sigma\pi)^{k-1}\sigma^{a-1}$. We have

$$\begin{array}{c}
1 \overset{(\sigma\pi)^{k-1}}{\longmapsto} 1 \overset{\sigma^{a-1}}{\longmapsto} a \\
2 \overset{(\sigma\pi)^{k-1}}{\longmapsto} k+1 \overset{\sigma^{a-1}}{\longmapsto} a+k = b
\end{array}$$

It follows from the lemma that $\pi^{(\sigma\pi)^{k-1}\sigma^{a-1}} = (ab)$.

Now,

$$\begin{aligned}
\operatorname{length}(\pi^{(\sigma\pi)^{k-1}\sigma^{a-1}}, S) &\leq 2 \cdot \operatorname{length}((\sigma\pi)^{k-1}\sigma^{a-1}, S) + \operatorname{length}(\pi, S) \\
&\leq 2 \cdot (2(k-1) + a - 1) + 1 \\
&\leq 2 \cdot (2n + n) + 1 \\
&\leq 5n.
\end{aligned}$$

Let $\rho \in S_n$. It is clear that ρ is the product of at most $n-1$ transpositions. Hence,

$$\operatorname{length}(\rho, S) \leq (n-1) \cdot 5n = O(n^2),$$

and since ρ was arbitrary, we conclude that $\mathrm{diam}(S_n, S) = O(n^2)$. □

In the above proof, we used a well-known set of generators, namely the set of all transpositions, as a stepping stone in our algorithm for generating S_n. We then calculated the diameter of S_n given S by multiplying the maximum length of a transposition and the maximum number of transpositions required to generate any permutation. It will be helpful to extract and formalize this principle for use later.

Definition 3.3. Let S be a set of generators for a group G, and T be a subset of G. The *length* of T with respect to S is the maximum length of the elements in T with respect to S.

Lemma 3.4. *Let S and T be sets of generators for a group G. Then* $\mathrm{diam}(G, S) \leq \mathrm{diam}(G, T) \cdot \mathrm{length}(T, S)$.

One way to obtain a lower bound on the diameter of a group G is to define a potential function $f : G \to \mathbb{Z}_{\geq 0}$ such that $f(e)$ is small and $|f(gs) - f(g)|$ is bounded for all elements $g \in G$ and generators $s \in S$. If we can find an element h such that $f(h)$ is large, we can then use the triangle inequality to give a lower bound on the length of h.

In our case, an example of such a function is $f(\pi) = \sum_{1 \leq i < j \leq n} \epsilon_{\pi,ij}$, where $\epsilon_{\pi,ij} = 1$ if $\pi(i) > \pi(j)$ and 0 otherwise. This function counts the number of "inverted pairs", and it is clear that $f(e) = 0$ while $f(g) = \binom{n}{2}$ for some g. Moreover, $|f(\pi s) - f(\pi)| \leq 2 \cdot (n-2) + 1$, when s is either σ or π giving a lower bound of $\Omega(n)$. This bound, however, is suboptimal, and to obtain one that is tight, we have to resort to a different trick.

We count not the number of "inverted pairs", but the number of "inverted triples". Let us make $\mathbb{Z}/n\mathbb{Z}$ the vertex set of a directed graph $\Gamma(V, E)$ where $E = \{(x, x+1) : x \in \mathbb{Z}/n\mathbb{Z}\}$. Then every ordered triple of distinct numbers $i, j, k \in \mathbb{Z}/n\mathbb{Z}$ has one of two orientations. If the path from i to k passes through j, we call the orientation of (i, j, k) *clockwise*. Conversely, if the path from i to k does not pass through j, we say that (i, j, k) has an *anti-clockwise* orientation.

Lemma 3.5. *Let us define a function $\phi : S_n \to \mathbb{Z}$ by*

$$\phi(\rho) = \sum_{1 \leq i < j < k \leq n} \delta_{\rho,ijk},$$

where $\delta_{\rho,ijk} = 1$ if $(\rho(i), \rho(j), \rho(k))$ has an anti-clockwise orientation and 0 otherwise. Let $\pi = (12), \sigma = (12 \ldots n)$, and γ be defined as follows:

γ	1	2	3	...	n-2	n-1	n
	n	n-1	n-2	...	3	2	1

Then the following are true:

(1) $\phi(e) = 0$,
(2) $\phi(\gamma) = \binom{n}{3}$,
(3) $\forall \rho \in S_n, \quad \phi(\rho\sigma) - \phi(\rho) = \phi(\rho\sigma^{-1}) - \phi(\rho) = 0$, *and*
(4) $\forall \rho \in S_n, \quad |\phi(\rho\pi) - \phi(\rho)| \leq n - 2$.

Proof. Fix $i < j < k$. Then (i, j, k) is clearly clockwise so $\delta_{e,ijk} = 0$ and (1) is obvious. On the other hand, $\gamma(i) > \gamma(j) > \gamma(k)$ so $\delta_{\gamma,ijk} = 1$, and (2) follows from there being $\binom{n}{3}$ triples. Now, observe that $\sigma(i) = i+1$, $\sigma(j) = j+1$, and $\sigma(k) \equiv k+1$ mod n. This means that σ is a graph isomorphism, and $(\sigma\rho(i), \sigma\rho(j), \sigma\rho(k))$ has

the same orientation as $(\rho(i), \rho(j), \rho(k))$. The same argument applies for σ^{-1}, so it also does not change the orientation of any triple. Hence, $\phi(\rho\sigma) = \phi(\rho) = \phi(\rho\sigma^{-1})$, which gives us (3).

It remains to prove (4), but it is clear that the orientation of $(\rho(i), \rho(j), \rho(k))$ is swapped by π if and only if $(\rho(i), \rho(j), \rho(k))$ contains both 1 and 2. There are $n-2$ such triples, so the difference between $\phi(\rho\pi)$ and $\phi(\rho)$ is at most $n-2$. □

The following proposition is now an easy corollary of the previous lemma.

Proposition 3.6. *Let S be as in Proposition 3.2. Then* $\operatorname{diam}(S_n, S) = \Omega(n^2)$.

Proof. Let γ be as defined in Lemma 3.5, and let $\rho_0\rho_1\cdots\rho_k$ be a word for γ, where $\rho_0 = e$, and ρ_i is a generator for $1 < i \leq k$. Then we have

$$\begin{aligned}
\phi(\gamma) &= \sum_{i=1}^{k} \phi(\rho_0\rho_1\cdots\rho_i) - \phi(\rho_0\rho_1\cdots\rho_{i-1}) \\
&\leq \sum_{i=1}^{k} |\phi(\rho_0\rho_1\cdots\rho_i) - \phi(\rho_0\rho_1\cdots\rho_{i-1})| \\
&= \sum_{\rho_i=\pi} |\phi(\rho_0\rho_1\cdots\rho_i) - \phi(\rho_0\rho_1\cdots\rho_{i-1})| \\
&\leq \sum_{\rho_i=\pi} (n-2) \\
&\leq k\cdot(n-2).
\end{aligned}$$

But since $\phi(\gamma) = \binom{n}{3} = \frac{n(n-1)(n-2)}{6}$, this implies that $k \geq \frac{n(n-1)}{6} = \Omega(n^2)$. □

Finally, combining Propositions 3.2 and 3.6 gives us

Theorem 3.7. *Let S be $\{(12), (12\ldots n)\}$. Then* $\operatorname{diam}(S_n, S) = \Theta(n^2)$.

4. Landau's function

In this section, we shall prove the following theorem:

Theorem 4.1 (Landau)**.** *Let $g : \mathbb{N} \to \mathbb{N}$, be defined for every n to be the largest order of an element of the symmetric group S_n. Then*

$$\ln(g(n)) \sim \sqrt{n\ln(n)}$$

or equivalently,

$$g(n) = e^{\sqrt{n\ln n}(1+o(1))}.$$

This function $g(n)$, called *Landau's function*, is named after Edmund Landau, who proved the above theorem in a 1903 paper [1]. Let us first try to develop an intuition for what exactly it measures before embarking on the proof of the theorem.

Now, any permutation π can be written as the product of disjoint cycles $c_1, c_2, \ldots, c_k$ of length $r_1, r_2, \ldots, r_k$. If we count the singleton cycles, we see that the lengths of these cycles sum to n. Also, $\pi^t = 1$ if and only if $c_i^t = 1$ for all i, which is only the case if $r_i \mid t$. Hence,

$$\operatorname{ord}(\pi) = \min\{t : \pi^t = 1\} = \min\{t : \forall i, r_i \mid t\} = \operatorname{lcm}\{r_1, \ldots, r_k\}.$$

In other words, $g(n)$ is the value of the largest least common multiple of any partition of n. Given small values of n, it is easy to compute the corresponding

values of $g(n)$ by looking at all possible partitions of n. The table for $n < 20$ (taken from [2]) is presented below:

n	$g(n)$	cycle lengths
2	2	2
3	3	3
4	4	4
5	6	2,3
6	6	1,2,3 or 6
7	12	3,4
8	15	3,5
9	20	4,5
10	30	2,3,5
11	30	1,2,3,5 or 5,6
12	60	3,4,5
13	60	1,3,4,5
14	84	3,4,7
15	105	3,5,7
16	140	4,5,7
17	210	2,3,5,7
18	210	1,2,3,5,7 or 5,6,7
19	420	3,4,5,7

Looking at all possible partitions, however, becomes impossible as n tends to infinity, while a simple analytic formula for $g(n)$ seems unlikely in light of the table above. Indeed, for each n, it is difficult to even identify a candidate permutation having order $g(n)$. We thus turn our investigation on its head and ask: If we have a permutation of order m, what is the minimum number of elements it must act upon?

Proposition 4.2. *Let $\prod_{i=1}^{k} p_i^{r_i}$ be the prime power factorization of m, and let $s(m) = \sum_{i=1}^{k} p_i^{r_i}$. Then there exists a permutation on n elements having order m if and only if $s(m) \leq n$.*

To prove Proposition 4.2, we will need the following lemma.

Lemma 4.3. *Let $a_1, a_2, \ldots, a_k$ be positive integers and let $m = \operatorname{lcm}\{a_1, \ldots, a_k\}$. Then $s(m) \leq \sum_{i=1}^{k} a_i$.*

Proof. First, we note that we can assume each $a_i > 1$, or else we can remove a_i to get a list of numbers with a smaller sum and the same lowest common multiple. Next, suppose $\gcd(a_j, a_l) = d > 1$. Then $\operatorname{lcm}\{a_1, \ldots, a_j, \frac{a_l}{d}, \ldots, a_k\} = \operatorname{lcm}\{a_1, \ldots, a_j, a_l, \ldots, a_k\}$ but $\sum_{i \neq l} a_i + \frac{a_l}{d} < \sum_{i=1}^{k} a_i$. Hence we can assume that $a_1, a_2, \ldots, a_k$ are relatively prime. Now suppose one of the a_i is not the power of a prime, i.e. there are two distinct primes p and q such that $p|a_i$ and $q|a_i$. Let t be the highest power of p that divides a_i. Then

$$a_i = \frac{a_i}{p^t}(p^t - 1) + \frac{a_i}{p^t} \geq 2(p^t - 1) + \frac{a_i}{p^t} \geq p^t + \frac{a_i}{p^t}$$

but $\operatorname{lcm}\{a_1, \ldots, a_i, \ldots, a_k\} = \operatorname{lcm}\{a_1, \ldots, \frac{a_i}{p^t}, p^t, \ldots, a_k\}$. □

Proof of Proposition 4.2. Suppose $s(m) \leq n$. Then we can pick a permutation having cycles with lengths $p_i^{r_i}$ and this permutation clearly has order m. Conversely, suppose $s(m) > n$ and there exists a permutation $\pi \in S_n$ of order m. Let $a_1, \ldots, a_k$ be the lengths of the cycles of π. Then since $\operatorname{lcm}\{a_1, \ldots, a_k\} = m$, Lemma 4.3 tells us that $\sum_{i=1}^{k} a_i > n$, giving a contradiction. □

Corollary 4.4. *For all $n \in \mathbb{N}$, $g(n) = \max_{s(m) \leq n} m$.*

It is clear from the above discussion that order $g(n)$ is attained by a permutation whose nontrivial cycles have lengths that are the powers of distinct primes. But

even though we have narrowed down candidate permutations of order $g(n)$, it is still not straightforward to establish a clear relationship between n and the partition of n that gives rise to $g(n)$. Hence, we consider instead a permutation of suboptimal order, and show that it is a good approximation.

We now define two new functions $\theta, \nu : \mathbb{N} \to \mathbb{R}$ whose asymptotic values will help us later.

Definition 4.5. Let $x \in \mathbb{N}$. Then

(1) $\theta(x) = \sum_{p \leq x} \ln p = \ln \left(\prod_{p \leq x} p \right)$

(2) $\nu(x) = \sum_{p \leq x} p,$

where both sums are taken over all the primes p less than x.

Remark 4.6. The function $\theta(x)$ is well studied and known as the first Chebyshev function.

Lemma 4.7. *We have $\theta(x) \sim x$.*

Proof. The upper bound is easy. We have

$$\sum_{p \leq x} \ln p \leq \sum_{p \leq x} \ln x = \pi(x) \ln x \sim \frac{x}{\ln x} \cdot \ln x = x.$$

On the other hand, for any $\epsilon > 0$,

$$\sum_{p \leq x} \ln p \geq \sum_{x^{1-\epsilon} < p \leq x} \ln p \geq \sum_{x^{1-\epsilon} < p \leq x} \ln x^{1-\epsilon} = \big(\pi(x) - \pi(x^{1-\epsilon})\big) \ln x^{1-\epsilon}.$$

But

$$\frac{\pi(x^{1-\epsilon})}{\pi(x)} \sim \frac{x^{1-\epsilon}}{(1-\epsilon)\ln x} \cdot \frac{\ln x}{x} = \frac{x^{1-\epsilon}}{1-\epsilon} = o(1)$$

by Corollary 2.14, so by Propositions 2.13 and 2.16,

$$\big(\pi(x) - \pi(x^{1-\epsilon})\big) \ln x^{1-\epsilon} \sim \pi(x) \ln x^{1-\epsilon} \sim \frac{x}{\ln x} \cdot (1-\epsilon) \ln x = x(1-\epsilon).$$

Since ϵ was arbitrary, the result follows. □

Remark 4.8. Lemma 4.7 is equivalent to the Prime Number Theorem.

Lemma 4.9. *We have $\nu(x) \sim \frac{x^2}{2 \ln x}$.*

Proof. Using Corollary 2.23 and Proposition 2.17, we know that

$$\sum_{p \leq x} p = \sum_{k=1}^{\pi(x)} p_k \sim \sum_{k=1}^{\lceil \frac{x}{\ln x} \rceil} p_k \sim \sum_{k=1}^{\lceil \frac{x}{\ln x} \rceil} k \ln k.$$

Given n, define $f : \mathbb{R} \to \mathbb{R}$ by $f(t) = t \ln t$, and let $P_1 = \{1, 2, \ldots, \lceil \frac{x}{\ln x} \rceil\}$. Then $\sum_{k=1}^{\lceil \frac{x}{\ln x} \rceil} k \ln k$ is a Riemann upper sum for f on $[1, \lceil \frac{x}{\ln x} \rceil]$ with partition P_1, and we

have that

$$\begin{aligned}\sum_{k=1}^{\lceil \frac{x}{\ln x} \rceil} k \ln k &\geq \int_1^{\frac{x}{\ln x}} t \ln t \, \mathrm{d}t \\ &= \left[\frac{1}{2}t^2 \ln t - \frac{1}{4}t^2\right]_1^{\frac{x}{\ln x}} \\ &= \frac{1}{2}\frac{x^2}{(\ln x)^2} \ln\left(\frac{x}{\ln x}\right) - \frac{1}{4}\left(\frac{x}{\ln x}\right)^2 + \frac{1}{4} \\ &\sim \frac{x^2}{2 \ln x}.\end{aligned}$$

The last asymptotic equality sign is due to Propositions 2.13 and 2.16, and the fact that $\ln x \sim \ln(\frac{x}{\ln x})$.

If we let $P_2 = \{2, 3, \ldots, \lceil \frac{x}{\ln x} \rceil + 1\}$, then $\sum_{k=1}^{\lceil \frac{x}{\ln x} \rceil} k \ln k$ is a Riemann lower sum for f on $[2, \lceil \frac{x}{\ln x} \rceil + 1]$, which implies that

$$\begin{aligned}\sum_{k=1}^{\lceil \frac{x}{\ln x} \rceil} k \ln k &\leq \int_2^{\frac{x}{\ln x}+2} t \ln t \, \mathrm{d}t \\ &= \left[\frac{1}{2}t^2 \ln t - \frac{1}{4}t^2\right]_2^{\frac{x}{\ln x}+2} \\ &= \frac{1}{2}(\frac{x}{\ln x} + 2)^2 \ln(\frac{x}{\ln x} + 2) - \frac{1}{4}(\frac{x}{\ln x} + 2)^2 - 2\ln 2 + 1 \\ &\sim \frac{x^2}{2 \ln x}.\end{aligned}$$

Once again, we made use of Propositions 2.13 and 2.16. Finally, by the squeeze theorem, it follows that $\sum_{k=1}^{\lceil \frac{x}{\ln x} \rceil} k \ln k \sim \frac{x^2}{2 \ln x}$. □

Using the above two lemmas, we can now prove

Proposition 4.10. *Fix $n \in \mathbb{N}$, and let $k(n) = \max\{k : \sum_{i=1}^{k} p_i \leq n\}$. Let $f(n) = \prod_{i=1}^{k(n)} p_i$. Then $\ln f(n) \sim \sqrt{n \ln n}$.*

Proof. Fix $n \in \mathbb{N}$. Recall that $k(n)$ is defined such that $\sum_{i=1}^{k(n)} p_i \leq n < \sum_{i=1}^{k(n)+1} p_i$, so by Lemma 4.7, we have

$$\ln f(n) = \sum_{i=1}^{k(n)} \ln p_i \sim p_{k(n)}. \tag{4.11}$$

On the other hand, Lemma 4.10 tells us

$$n \sim \sum_{i=1}^{k(n)} p_i \sim \frac{1}{2}\frac{p_{k(n)}^2}{\ln p_{k(n)}}. \tag{4.12}$$

Taking log of both sides, this becomes

$$\ln n \sim 2 \ln p_{k(n)} - \ln(2 \ln p_{k(n)}) \sim 2 \ln p_{k(n)}. \tag{4.13}$$

Combining (4.12) and (4.13),

$$n \sim \frac{p_{k(n)}^2}{\ln n} \iff p_{k(n)} \sim \sqrt{n \ln n}, \tag{4.14}$$

and now (4.11) and (4.14) together give $\ln f(n) \sim \sqrt{n \ln n}$. □

It easy to see that $f(n)$ is the order of a permutation whose cycle lengths are $p_1, \ldots, p_{k(n)}$. From the above discussion, we know that $f(n) \leq g(n)$. In fact, the two functions are asymptotically equal. We will prove this via two lemmas due to Shah and Miller [2].

Lemma 4.15 (Shah)**.** *Given $n \in \mathbb{N},\ n > 1$, let*

- $\mathscr{F}_n = \{p \text{ prime} : p \mid f(n)\} \cup \{p_{k(n)+1}\} = \{p_1, \ldots, p_{k(n)+1}\}$
- $\mathscr{G}_n = \{p \text{ prime} : p \mid g(n)\}$.

Then

$$\sum_{p \in \mathscr{G}_n} \ln p < 2 + \ln f(n) + \ln p_{k(n)+1}.$$

Proof. Recall that if $\prod_{i=1}^{k} p_i^{r_i}$ is the prime power factorization of m, then $s(m) = \sum_{i=1}^{k} p_i^{r_i}$. Proposition 4.2 tells us that $s(g(n)) \leq n$, from which we get

$$\sum_{p \in \mathscr{G}_n} p \leq s(g(n)) \leq n < \sum_{p \in \mathscr{F}_n} p$$

Subtracting $\sum_{p \in \mathscr{G}_n \cap \mathscr{F}_n} p$ from both sides, this becomes

$$\sum_{p \in \mathscr{G}_n \setminus \mathscr{F}_n} p < \sum_{p \in \mathscr{F}_n \setminus \mathscr{G}_n} p. \tag{4.16}$$

Now, $\frac{\ln x}{x}$ is a decreasing function for $x > e$. Since for all $p \in \mathscr{G}_n \setminus \mathscr{F}_n$, $3 \leq p_{k(n)+1} < p$, we have $\ln p < p \cdot \frac{\ln p_{k(n)+1}}{p_{k(n)+1}}$. Analogously, for all $p \in \mathscr{F}_n \setminus \mathscr{G}_n$ except for 2, $3 \leq p \leq p_{k+1}$ implies that $\ln p > p \cdot \frac{\ln p_{k(n)+1}}{p_{k(n)+1}}$. Using (4.16), we get the following string of inequalities:

$$\begin{aligned}
\sum_{p \in \mathscr{G}_n \setminus \mathscr{F}_n} \ln p &< \sum_{p \in \mathscr{G}_n \setminus \mathscr{F}_n} p \cdot \frac{\ln p_{k(n)+1}}{p_{k(n)+1}} \\
&< \sum_{\substack{p \in \mathscr{F}_n \setminus \mathscr{G}_n \\ p \neq 2}} p \cdot \frac{\ln p_{k(n)+1}}{p_{k(n)+1}} + 2 \\
&< \sum_{\substack{p \in \mathscr{F}_n \setminus \mathscr{G}_n \\ p \neq 2}} \ln p + 2.
\end{aligned}$$

Adding back $\sum_{p \in \mathscr{G}_n \cap \mathscr{F}_n} \ln p$ to both sides, we get

$$\sum_{p \in \mathscr{G}_n} \ln p < \sum_{\substack{p \in \mathscr{F}_n \\ p \neq 2}} \ln p + 2 < 2 + \ln f(n) + \ln p_{k(n)+1}.$$

□

Lemma 4.17 (Miller)**.** *Let p^e be a prime power such that $e > 1$. If p^e divides $g(n)$, then $p^e \leq 2p_{k(n)+1}$.*

Proof. Let $q = p_j$ be the smallest prime not dividing $g(n)$. Then $\sum_{i=1}^{j-1} p_i \leq n < \sum_{i=1}^{k(n)+1} p_i$, which implies that $j \leq k(n)+1$. Hence it is sufficient to prove that $p^e \leq 2q$. Now suppose, for contradiction, that $p^e > 2q$, and let $n \in \mathbb{N}$ be such that $q^{n-1} < p < q^n$. Note that $q^n = q^{n-1} \cdot q < pq$. If we let $m = \frac{q^n}{p} \cdot g(n)$, then $m > g(n)$ and

$$s(m) = s(g(n)) + (q^n - p^e + p^{e-1}).$$

We claim that the quantity in parentheses on the right is negative. If $p < q$, then $n = 1$ and

$$q - p^e + p^{e-1} \leq q - \frac{p^e}{2} < q - \frac{2q}{2} = 0.$$

If $p > q$, we have

$$q^n - p^e + p^{e-1} < pq - p(p-1) \leq pq - pq = 0,$$

where the first inequality holds because $e > 1$ and $p \geq 2$. In either case, $s(m) < s(g(n))$ but $m > g(n)$, which contradicts Corollary 4.4. □

Corollary 4.18. *Given the conditions of Lemma 4.17,* $p \leq \sqrt{2p_{k(n)+1}}$.

With these two lemmas, we can now prove

Proposition 4.19. *We have* $\ln f \sim \ln g$.

Proof. Let $\prod_{i=1}^{j} q_i^{e_i}$ be the prime factorization of $g(n)$. Then $\ln g(n) = \sum_{i=1}^{j} e_i \ln q_i$. Without loss of generality, we can let t be such that $e_i > 1$ for $1 \leq i \leq t$ and $e_{t+1} = e_{t+2} = \ldots = e_j = 1$. Then $t \leq q_t \leq \sqrt{2p_{k(n)+1}}$ by Corollary 4.19, and we have

$$\begin{aligned}
\ln g(n) &= \sum_{i=1}^{j} e_i \ln q_i = \sum_{i=1}^{t} e_i \ln q_i + \sum_{i=t+1}^{j} \ln q_i \\
&\leq \sum_{i=1}^{t} e_i \ln q_i + \sum_{p \in \mathscr{G}_n} \ln p \\
&\leq \left(\ln(2p_{k(n)+1}) \cdot \sqrt{2p_{k(n)+1}}\right) + (2 + \ln f(n) + \ln p_{k(n)+1}),
\end{aligned}$$

where we get the last equality by applying the bounds given by Lemma 4.15 and 4.17. We also know trivially that $f(n) \leq g(n)$, so after dividing everything by $\ln f(n)$, we get

$$1 \leq \frac{\ln g(n)}{\ln f(n)} < \frac{\ln(2p_{k(n)+1}) \cdot \sqrt{2p_{k(n)+1}}}{\ln f(n)} + \frac{2}{\ln f(n)} + 1 + \frac{\ln p_{k(n)+1}}{\ln f(n)}.$$

By Lemma 4.7, $\ln f(n) \sim p_{k(n)+1} - 1$, so, we have

$$\lim_{n\to\infty} \frac{\ln(2p_{k(n)+1}) \cdot \sqrt{2p_{k(n)+1}}}{\ln f(n)} = \lim_{n\to\infty} \frac{\ln p_{k(n)+1}}{\ln f(n)} = 0.$$

The result follows. □

When we combine this result with Proposition 4.10, we see that we have completed the proof of Theorem 4.1.

5. The maximum diameter of S_n and A_n

In Section 3, we considered the diameter of S_n with respect to a given set of generators. Since the diameter of a group varies depending on the generators we consider, it will be interesting to know what the diameter is on average and how big the diameter can get in the worst case scenario. We shall investigate the latter question with respect to S_n, reproducing a proof of an upper bound due to Babai and Seress [3].

Definition 5.1. Let G be a group. Define $\text{diam}_{max}G$ to be the maximum of $\text{diam}(G, S)$ taken over all sets of generators S.

Theorem 5.2 (Babai - Seress)**.** *Let G be S_n or A_n. (For the rest of this section, G will denote either S_n or A_n.) Then*

$$\text{diam}_{max}(G) \leq e^{\sqrt{n \ln n}(1+o(1))}.$$

This proof, like that of Proposition 3.2, is algorithmic. We define a procedure for generating any group element from an arbitrary set of generators and show that the length of the word we get has the upper bounds we want. Recall that in Proposition 3.2, we used transpositions as "stepping stones". In this proof, however, we shall let the 3-cycles play that role.

Definition 5.3. The *support* of a permutation π, denoted $supp(\pi)$, is the set of elements displaced by π.

Definition 5.4. The *degree* of a permutation π is the size of its support, i.e. $deg(\pi) = |supp(\pi)|$.

Lemma 5.5. *Let S be the set of all 3-cycles. Then $diam(A_n, S) \leq n$.*

Proof. Fix $\pi \in G$ and consider its cycle decomposition. Let $c = (i_1, \ldots, i_k)$ be a cycle in π. Clearly, $c = (i_k i_{k-1})(i_{k-1} i_{k-2}) \cdots (i_2 i_1)$. Hence, if k is odd, c is the product of an even number of transpositions, and conversely, if k is even, c is the product of an odd number of transpositions.

If c is an odd degree cycle so that $k = 2m + 1$, one can check that we can write $c = \overbrace{(i_{k-2} i_{k-1} i_k)(i_{k-4} i_{k-3} i_{k-2}) \cdots (i_1 i_2 i_3)}^{m \text{ terms}}$, which implies that $\text{length}(c, S) \leq m \leq \deg(c)$.

Now, since π is an even permutation, it contains an even number of even degree cycles. Hence, if $k = 2m$ for some m, we can pair c with another even degree cycle $c' = (j_1 \ldots j_l)$, where $l = 2q$ for some q. We can then write cc' as the product of even degree cycles $(i_1 \ldots i_{k-1} j_1 i_k)$ and $(j_1 i_k j_2 \ldots j_l)$. Since these cycles are of degree $2m + 1$ and $2q + 1$ respectively, $\text{length}(cc', S) \leq m + q \leq \deg(c) + \deg(c')$.

Let $\pi = c_1 c_2 \cdots c_r$ be the cycle decomposition of π. Corollary 2.7 shows that

$$\text{length}(\pi, S) \leq \sum_{i=1}^{r} \text{length}(c_i, S) \leq \sum_{i=1}^{r} \deg(c_i) \leq n$$

□

Notice that if we have a single 3-cycle γ, we can obtain all other 3-cycles by taking conjugates of γ. We can thus bound the diameter of G with respect to the length of γ. Proving this first requires the following definition and lemma.

Definition 5.6. Let G be either A_n or S_n. Suppose $T \subset G$ and $1 \leq k \leq n$. We say that T is a *k-transitive* subset if for every pair of k-tuples $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$, where $x_i \neq x_j$ and $y_i \neq y_j$ for $i \neq j$, there exists a permutation $\pi \in T$ such that $x_i^\pi = y_i$.

Lemma 5.7. *Let G be either A_n or S_n. For any set S of generators and any k such that $1 \leq k \leq n-2$, there exists a k-transitive subset $R_k \subset G$ such that* $\text{length}(R_k, S) \leq n^k$.

Proof. Let Ω_k be the set of all k-tuples of elements in $\{1, \ldots, n\}$, and consider the action of S on this set. We represent this in the form of a directed graph $\Gamma(V, E)$, where the vertex set $V = \Omega_k$ and the edge set E consists of all ordered pairs $((x_1, \ldots, x_k), (y_1, \ldots, y_k))$ such that there exists a permutation $s \in S$ with $(x_1, \ldots, x_k)^s = (y_1, \ldots, y_k)$. Now

$$|V| = n(n-1)\cdots(n-k+1) \leq n^k.$$

Since G acts transitively on Ω_k, and S generates G, Γ is connected, and the distance between any two vertices is at most $|V|$. This distance corresponds to $\text{length}(R_k, S)$. □

Proposition 5.8. *Let G be either A_n or S_n. Suppose S is a set of generators for G, and γ is any 3-cycle in G. Then* $\text{diam}(G, S) \leq 1 + 2n^4 + n \cdot \text{length}(\gamma, S)$.

Proof. We can obtain any 3-cycle from γ by conjugating it with an element from a 3-transitive set, R_3. Recall that $\text{length}(R_3, S) \leq n^3$, and that any even permutation is the product of at most n 3-cycles. Thus, for an arbitrary even permutation π,

$$\text{length}(\pi, S) \leq n \cdot (\text{length}(R_3, S) + \text{length}(\gamma) + \text{length}(R_3, S)) \leq 2n^4 + n \cdot \text{length}(\gamma, S),$$

so the proposition holds for A_n. To extend the result to S_n, we just need to show that the bound also holds for odd permutations. If S generates S_n, it must contain an odd permutation σ. Let $\pi \in S_n$ be an odd permutation, then $\sigma^{-1}\pi$ is an even permutation, so

$$\text{length}(\pi, S) \leq \text{length}(\sigma, S) + \text{length}(\sigma^{-1}\pi, S) \leq 1 + 2n^4 + n \cdot \text{length}(\gamma, S).$$

□

It remains to find out how to reach a 3-cycle. The following proposition gives us a way of achieving this:

Proposition 5.9. *Let σ and π be permutations such that $|\text{supp}(\sigma) \cap \text{supp}(\pi)| = 1$. Then the commutator $[\sigma, \pi] = \sigma^{-1}\pi^{-1}\sigma\pi$ is a 3-cycle.*

Proof. Let $x \in \text{supp}(\sigma) \cap \text{supp}(\pi)$. Then we have

$$\begin{array}{ccccccccc}
x & \overset{\sigma^{-1}}{\longmapsto} & x^{\sigma^{-1}} & \overset{\pi^{-1}}{\longmapsto} & x^{\sigma^{-1}} & \overset{\sigma}{\longmapsto} & x & \overset{\pi}{\longmapsto} & x^\pi \\
x^{\sigma^{-1}} & \overset{\sigma^{-1}}{\longmapsto} & x^{\sigma^{-2}} & \overset{\pi^{-1}}{\longmapsto} & x^{\sigma^{-2}} & \overset{\sigma}{\longmapsto} & x^{\sigma^{-1}} & \overset{\pi}{\longmapsto} & x^{\sigma^{-1}} \\
x^\sigma & \overset{\sigma^{-1}}{\longmapsto} & x & \overset{\pi^{-1}}{\longmapsto} & x^{\pi^{-1}} & \overset{\sigma}{\longmapsto} & x^{\pi^{-1}} & \overset{\pi}{\longmapsto} & x \\
x^{\pi^{-1}} & \overset{\sigma^{-1}}{\longmapsto} & x^{\pi^{-1}} & \overset{\pi^{-1}}{\longmapsto} & x^{\pi^{-2}} & \overset{\sigma}{\longmapsto} & x^{\pi^{-2}} & \overset{\pi}{\longmapsto} & x^{\pi^{-1}} \\
x^\pi & \overset{\sigma^{-1}}{\longmapsto} & x^\pi & \overset{\pi^{-1}}{\longmapsto} & x & \overset{\sigma}{\longmapsto} & x^\sigma & \overset{\pi}{\longmapsto} & x^\sigma.
\end{array}$$

Since $[\sigma, \pi]$ clearly does not affect any other elements, we conclude that $[\sigma, \pi] = (x^\pi x^\sigma x)$ □

It follows immediately from this proposition that if we have a permutation π with degree $t < n$, and a t-transitive subset R_t, then there exists a permutation $\sigma \in R_t$ such that $[\sigma, \pi]$ is a 3-cycle. Furthermore, we have $\text{length}([\sigma, \pi], S) \leq 2 \cdot (\text{length}(\pi, S) + \text{length}(R_t, S))$. However, the bound on $\text{length}(R_t, S)$ given by Lemma 5.7 grows exponentially in t, which is larger than the bound we want (recall that we want $\text{diam}(G, S) = e^{\sqrt{n \ln n}(1+o(1))}$ so we need to reduce the size of t.

We shall achieve this using two tricks. First, we observe that if π contains cycles that are the lengths of distinct primes, we can raise π to a power to kill off most of these cycles and yet not arrive at the identity permutation.

Definition 5.10. Let π be a permutation and $x \in \text{supp}(\pi)$. Then the period of x with respect to π is the smallest positive integer m such that $\pi^m(x) = x$.

Proposition 5.11. *Let π be a permutation of degree n, and suppose $p_1, \ldots, p_k$ are distinct primes such that*

- $p_1, \ldots, p_k \mid \text{ord}(\pi)$*, and*
- $\prod_{i=1}^{k} p_i \geq n^s$ *for some positive integer s.*

Then there exists $m \in \mathbb{N}$ such that $2 \leq deg(\pi^m) < \frac{n}{s}$.

Proof. First, let us write $\text{ord}(\pi) = \prod_{i=1}^{k} p_i^{r_i} \cdot L$, where $\gcd(L, p_i) = 1$. If we define, for each p_i, a number $l_i = \frac{\text{ord}(\pi)}{p_i^{r_i}}$, then l_i is the largest factor of $\text{ord}(\pi)$ coprime with p_i. We claim that one of these l_i gives the value of m we want.

Recall that a permutation can be partitioned into disjoint cycles. If we consider a point $x \in [n]$, then the period of x is clearly the length of the cycle in π that acts upon it. Hence, if we define $A(x) = \{i : p_i \mid \text{period}(x)\}$, then this set contains the indices of all primes that divide the length of that cycle. Observe that for each x,

$$\prod_{i \in A(x)} p_i \leq n. \tag{5.12}$$

If x is fixed by π^{l_i}, then $p_i \nmid \text{period}(x)$, so $i \notin A(x)$. Since $\deg(\pi^{l_i})$ counts the number of elements not fixed by π^{l_i}, its value is just the number of x such that $i \in A(x)$. We call this number N_i. To prove our claim, we need to show that there exists an i such that $N_i \leq \frac{n}{s}$.

Let $R(i, x)$ be the indicator function of the relation "$i \in A(x)$". Then (5.12) implies

$$\sum_{i \in [k]} R(i, x) \ln p_i = \sum_{i \in A(x)} \ln p_i \leq \ln n. \tag{5.13}$$

Also, our hypothesis $\prod_{i=1}^{k} p_i \geq n^s$ implies

$$\sum_{i=1}^{k} \ln p_i \geq s \ln n. \tag{5.14}$$

Observing that $\sum_{x\in[n]} R(i,x) = N_i$, we can compute the weighted average of the N_i.

$$\frac{\sum_{i=1}^{k} N_i \ln p_i}{\sum_{i=1}^{k} \ln p_i} = \frac{\sum_{i=1}^{k}\sum_{x\in[n]} R(i,x)\ln p_i}{\sum_{i=1}^{k} \ln p_i} = \frac{\sum_{x\in[n]}\left(\sum_{i=1}^{k} R(i,x)\ln p_i\right)}{\sum_{i=1}^{k} \ln p_i}.$$

Since the quantity in parentheses is bounded above by $\ln n$ by (5.13) while the denominator is bounded below by $s\ln n$ by (5.14), the last quantity on the right is less than or equal to $\frac{n\ln n}{s\ln n} = \frac{n}{s}$. Therefore, there exists an $N_i \leq \frac{n}{s}$, as we wanted. □

The previous "degree-reducing" trick only works if π contains cycles of the right length. In general, this is not the case. We can, however, modify π at low cost to inject the necessary cycles into it. In the process we increase the degree of π by a factor of 2, but this is more than compensated by the reduction that follows upon application of Proposition 5.11.

Proposition 5.15. *Let G be either A_n or S_n. Let S be a set of generators for G, and suppose $\pi \in G$ has degree $k \geq 2$. Let $d \in \mathbb{N}$ be such that $d \leq \frac{k}{3}$ and $d = d_1 + \ldots + d_r$, where $d_i \in \mathbb{N}$. Then there exists $\lambda \in G$ such that*

(1) $\deg(\lambda) \leq 2k$,
(2) λ includes cycles of length $d_1, \ldots, d_r$,
(3) $\text{length}(\lambda, S) \leq 2\cdot\text{length}(\pi, S) + 2\cdot\text{length}(R_{2d}, S)$.

Proof. Let $B \subset \text{supp}(\pi)$ be such that $|B| = d$ and $B \cap B^\pi = \varnothing$. We can choose such a subset because we can pick every other element in each cycle to be in B. Then select $\tau \in R_{2d}$ such that $\tau|_B$ is a product of cycles of length d_i, but τ fixes each point of B^π. Such a permutation exists because R_{2d} is $2d$-transitive by definition.

Let $\lambda = \pi\tau\pi^{-1}\tau^{-1}$. Now it is clear that the degree of a product of two permutations is less than or equal to the sum of their individual degrees. By Lemma 3.3, $\deg(\tau\pi^{-1}\tau^{-1}) = \deg(\pi)$, and hence, we have that $\deg(\lambda) \leq \deg(\tau\pi^{-1}\tau^{-1}) + \deg(\pi^{-1}) = 2k$. Since $\lambda|_B = \tau^{-1}|_B$, λ also satisfies (2), while (3) is true because of Corollary 2.7. □

Finally, we are ready to prove the main theorem of this section.

Proof of Theorem 5.2. Let p_i denote the i-th prime number, and let

- $\phi(n,s) = \min\left\{k : \prod_{i=1}^{k} p_i > n^s\right\}$,
- $\psi(n,s) = \sum_{i=1}^{\phi(n,s)} p_i$.

If we set $s = \ln n$, then we have $\phi(n,2s) \sim \frac{(\ln n)^2}{\ln\ln n}$, and $\psi(n,2s) \sim \frac{2(\ln n)^4}{\ln\ln n}$ (c.f. [4]). We let $t = 3\cdot\psi(n,2s)$. Now, observe that if we have a permutation $\sigma \in G$ of degree at most t, then we can select $\rho \in R_t$ such that $|\text{supp}(\sigma) \cap \text{supp}(\rho)| = 1$.

Then Proposition 5.9 tells us that $[\sigma, \rho]$ is a 3-cycle, and

$$\begin{aligned} \text{length}([\sigma,\rho], S) &\leq 2 \cdot (\text{length}(\sigma, S) + \text{length}(R_t, S)) \\ &\leq 2 \cdot \text{length}(\sigma, S) \cdot \text{length}(R_t, S) \\ &\leq 2n^t \cdot \text{length}(\sigma, S). \end{aligned}$$

Now, the idea is to reach a permutation of degree at most t by iteratively transforming an initial permutation π using Proposition 5.15 and 5.11. Suppose at round j of this process we have a permutation $\pi_j \in G$ with degree $m > t = 3 \cdot \psi(n, 2s)$. We then apply Proposition 5.15 to transform π_j into π'_j, with $d = \psi(n, 2s)$ and $d_i = p_i$. Observe that $\deg(\pi'_j) = 2m$. Since, π'_j now satisfies the hypotheses for Proposition 5.11, we can transform it once again into π_{j+1} of degree $< \frac{2m}{2s} = \frac{m}{s}$.

Let us compute the cost of each step. Recall that $t = 3d$, so by Proposition 5.15 and Lemma 5.7,

$$\begin{aligned} \text{length}(\pi'_j, S) &\leq 2 \cdot (\text{length}(\pi_j, S) + \text{length}(R_{2d}, S)) \\ &\leq 2 \cdot \text{length}(\pi_j, S) + 2n^{2d} \\ &\leq 2n^t \cdot \text{length}(\pi_j, S). \end{aligned}$$

Also, recall that Landau's function $g(m)$ gives the maximum order of permutations of degree less than or equal to m. This is an upper bound on $\text{length}(\pi_{j+1}, \pi'_j)$. Hence, using Lemma 3.5, we know that

$$\begin{aligned} \text{length}(\pi_{j+1}, S) &\leq \text{length}(\pi_{j+1}, \pi'_j) \cdot \text{length}(\pi'_j, S) \\ &\leq g(m) \cdot 2n^t \cdot \text{length}(\pi_j, S). \end{aligned}$$

Next, we want to find out the number of iterations, l we need to undergo before the algorithm terminates. But since the degree of our permutation is reduced by a factor of $\ln n$ during each round, even in the worst case scenario, this number l need only satisfy

$$\frac{n}{(\ln n)^l} \leq t \sim \frac{6(\ln n)^4}{\ln(\ln n)} \iff (\ln n)^l \gtrsim \frac{n \ln(\ln n)}{6(\ln n)^4}.$$

This means we require

$$\begin{aligned} l &\gtrsim \ln\left(\frac{n\ln(\ln n)}{6(\ln n)^4}\right) \cdot \frac{1}{\ln(\ln n)} \\ &= \ln n + \ln(\ln(\ln n)) - \ln 6 - 4\ln(\ln n) - \ln(\ln n) \\ &\sim \ln n. \end{aligned}$$

Hence, we take $l = \ln n$, and we have

$$\begin{aligned} \text{length}(\pi_l, S) &\leq g(n)g(n/s)\cdots g(n/s^l) \cdot (2n^t)^l \\ &= \left(\prod_{i=0}^{l} e^{\sqrt{n/(\ln n)^i \cdot \ln(n/(\ln n)^i)}(1+o(1))}\right) \cdot (2n^t)^l. \end{aligned}$$

Taking log of both sides, this becomes

$$\begin{aligned}
\ln \text{length}(\pi_l, S) &\lesssim \sum_{i=0}^{\ln n} \sqrt{\frac{n}{(\ln n)^i} \cdot \ln\left(\frac{n}{(\ln n)^i}\right)} + \ln 2 \ln(\ln n) + \frac{2(\ln n)^6}{\ln(\ln n)} \\
&\sim \sum_{i=0}^{\ln n} \sqrt{\frac{n}{(\ln n)^i} \cdot \ln\left(\frac{n}{(\ln n)^i}\right)} \\
&\leq \sqrt{n \ln n} + \sqrt{\frac{n}{\ln n} \ln\left(\frac{n}{\ln n}\right)} + \sqrt{\frac{n}{(\ln n)^2} \ln\left(\frac{n}{(\ln n)^2}\right)} \\
&\quad + \ln n \cdot \sqrt{\frac{n}{(\ln n)^3} \ln\left(\frac{n}{(\ln n)^3}\right)} \\
&\sim \sqrt{n \ln n},
\end{aligned}$$

which implies that $\text{length}(\pi_l, S) = e^{\sqrt{n \ln n}(1+o(1))}$. Obviously, multiplying this value by n or $2n^t$ does not change the asymptotics, so by Propositions 5.9 and 5.8, we have $\text{diam}(G, S) \leq e^{\sqrt{n \ln n}(1+o(1))}$. □

6. Outlook

The diameter of abelian groups can be very large. Suppose G is cyclic of order $2n+1$, and g is a generator of G, then g^n cannot be expressed as a word in $\{g, g^{-1}\}$ of length less than n. On the other hand, non-abelian simple groups are believed to have small diameter. In a later paper [5], Babai and Seress made the following conjecture:

Conjecture 6.1. *If G is a non-abelian finite simple group of order N, then*

$$\text{diam}_{max}(G) < (\ln N)^C$$

for some absolute constant C.

In particular, the true diameter of A_n is believed to be much smaller than what we proved in the previous section. This result, however, has remained elusive, and almost no progress was made in general until 2008 when Helfgott proved the conjecture for $SL_2(\mathbb{Z}/p\mathbb{Z})$ [6] and hence $PSL_2(\mathbb{Z}/p\mathbb{Z})$.

His proof relies on the following key proposition:

Proposition 6.2. *Let p be a prime. Let A be a subset of $SL_2(\mathbb{Z}/p\mathbb{Z})$ not contained in any proper subgroup.*

1. *Assume that $|A| < p^{3-\delta}$ for some fixed $\delta > 0$. Then*
$$|A \cdot A \cdot A| > c|A|^{1+\epsilon}$$
where $c > 0$ and $\epsilon > 0$ depend only on δ.
2. *Assume that $|A| > p^\delta$ for some fixed $\delta > 0$. Then there is an integer $k > 0$, depending only on δ, such that every element of $SL_2(\mathbb{Z}/p\mathbb{Z})$ can be expressed as a product of at most k elements of $A \cup A^{-1}$.*

Let S be a set of generators for $SL_2(\mathbb{Z}/p\mathbb{Z})$. Then $|(S \cdot S) \cup S| \geq |S| + 1$, so if we let $A = \bigcup_{i=1}^{d} S^i$, then $\text{length}(A, S) = d$ and $|A| \geq d$. Now, we set $\delta = 3/2$. We then apply part (1) of the proposition to A iteratively, until we reach a set $A' = A^{3e}$ of size $|A'| > p^{3/2}$ so that we can invoke part (2) of the proposition.

Obviously, this algorithm bounds $\mathrm{diam}_{max}(SL_2(\mathbb{Z}/p\mathbb{Z}))$ by $3edk$, where d and k are absolute constants not depending on p, so we need to obtain a bound for e in terms of p, and an easy computation shows that this is polylogarithmic in p. In other words we have

Theorem 6.3. *Let p be a prime and $G = SL_2(\mathbb{Z}/p\mathbb{Z})$. Then $diam_{max}(G) = O((\ln p)^c)$.*

Helfgott extended this result to $SL_3(\mathbb{Z}/p\mathbb{Z})$ in a more recent paper [7] using similar methods. Meanwhile, Pyber and Szabó managed to generalize Proposition 6.2 to produce the following theorem [8]:

Theorem 6.4. *Let L be a finite simple group of Lie type of rank r and A a generating set of L. Then either $A^3 = L$ or*

$$|A^3| \gg |A|^{1+\epsilon}$$

where ϵ and the implied constant depend only on r.

This proves that the conjecture holds much more generally for all finite simple groups of Lie type of bounded rank.

Acknowledgments. I would like to thank Peter May for organizing the REU program, László Babai for introducing me to discrete mathematics and finite group theory, and my mentors Robin Walters and Aaron Marcus for supervising the writing of my paper.

References

[1] E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, Bd. I, Teubner, Leipzig, 1909.
[2] W. Miller, The maximum order of an element of a finite symmetric group, The American Mathematical Monthly, Vol. 94, No. 6 (Jun. - Jul., 1987), 497-506.
[3] L. Babai and Á. Seress, On the diameter of Cayley graphs of the symmetric group, Journal of Combinatorial Theory, Series A 49 (1988), 175-179.
[4] G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 5th ed., Oxford Univ. Press (Clarendon), London/New York, 1979.
[5] L. Babai and Á. Seress, On the diameter of permutation groups, European Journal of Combinatorics, 13 (1992), 231-243.
[6] H. A. Helfgott, Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$, Annals of Mathematics, 167 (2008), 601-623.
[7] H. A. Helfgott, Growth in $SL_3(\mathbb{Z}/p\mathbb{Z})$, Journal of the European Mathematical Society, Vol. 13, Issue 3 (2011), 761-851.
[8] L. Pyber and E Szabó, Growth in finite simple groups of Lie type, arXiv:1005.1858v2