On the diameter of the symmetric group: polynomial bounds

László Babai*, Robert Beals†, Ákos Seress‡

Abstract

We address the long-standing conjecture that all permutations have polynomially bounded word length in terms of any set of generators of the symmetric group. The best available bound on the maximum required word length is exponential in $\sqrt{n \log n}$. Polynomial bounds on the word length have previously been established for very special classes of generating sets only.

In this paper we give a polynomial bound on the word length under the sole condition that one of the generators fix at least 67% of the domain. Words of the length claimed can be found in Las Vegas polynomial time.

The proof involves a Markov chain mixing estimate which permits us, apparently for the first time, to break the "element order bottleneck."

As a corollary, we obtain the following average-case result: for a $1 - \delta$ fraction of the pairs of generators for the symmetric group, the word length is polynomially bounded. It is known that for almost all pairs of generators, the word length is less than $\exp(\sqrt{n \ln n}(1 + o(1)))$.

1 Introduction

Let G be a finite group and S a set of generators of G. We consider the undirected Cayley graph $\Gamma(G, S)$ which has G as its vertex set; the pairs $\{\{g, gs\} : g \in G, s \in S\}$ are the edges. Let $\mathrm{diam}(G, S)$ denote the diameter of $\Gamma(G, S)$.

The diameter of Cayley graphs has been studied in a number of contexts, including interconnection networks, expanders, puzzles such as Rubik's cube and Rubik's rings, card shuffling and rapid mixing, random generation of group elements, combinatorial group theory. Even and Goldreich [EG] proved that finding the diameter of a Cayley graph of a permutation group is NP-hard even for the basic case when the group is an elementary abelian 2-group (every element has order 2). Jerrum [Je] proved that for directed Cayley graphs of permutation groups, to find the directed distance between two permutations is PSPACE-hard. No approximation algorithm is known for distance in and the diameter of Cayley graphs of permutation groups. Strikingly, the question of the diameter of the Rubik's cube Cayley graph appears to be wide open (cf. [Ko]). We refer to [BHKLS] for more information about the history of the diameter problem and related results and to the survey [He] for applications of Cayley graphs to interconnection networks.

For $G = S_n$ and $G = A_n$ we conjecture that $\mathrm{diam}(G, S)$ is polynomially bounded in n for all sets S of generators:

Conjecture 1.1. **[BS1]** *There exists a constant C such that for all n and all sets S of generators of $G = S_n$ or A_n, $\mathrm{diam}(G, S) \leq n^C$.*

This conjecture would imply a quasipolynomial upper bound on the diameters of all Cayley graphs of all *transitive* permutation groups [BS2].

The Conjecture has been verified for very special classes of generating sets. Driscoll and Furst [DF] proved an $O(n^2)$ bound for the case when all generators are cycles of bounded lengths and McKen-

*Department of Computer Science, University of Chicago, 1100 East 58th Street, Chicago, IL 60637, and Mathematical Institute of the Hungarian Academy of Science. Email: laci@cs.uchicago.edu. This research was partially supported by the NSF.

†IDA Center for Communications Research - Princeton, 805 Bunn Drive, Princeton, NJ 08540. Email: beals@idaccr.org.

‡Department of Mathematics, Ohio State University, 231 W. 18th Avenue, Columbus, OH 43210. Email: akos@math.ohio-state.edu. Partially supported by the NSF.

zie [McK] gave a polynomial bound for the case when the generators have bounded degree. (The **degree** of a permutation is the number of elements actually moved.) We note that McKenzie's exponent is $\Theta(k)$ where k is the bound on the degree.

We extend these results to a large class of generating sets in the hope that this may lead to settling the Conjecture.

THEOREM 1.1. *Let ϵ be a positive constant. Let S be a set of generators for $G = S_n$ or $G = A_n$. Assume that S contains an element of degree $\leq n/(3+\epsilon)$. Then $\mathrm{diam}(G, S) = O(n^C)$ where C is an absolute constant (does not depend on ϵ). Moreover, between any pair of elements of G, a path of length $O(n^C)$ can be found in Las Vegas polynomial time.*

We did not attempt to optimize the exponent C; without effort, our proof yields a diameter upper bound of $n^7(\log n)^{O(1)}$. The bound reduces to $n^6(\log n)^{O(1)}$ if the number of generators is bounded.

The best upper bound known to hold for the diameter of *all* Cayley graphs of A_n and S_n is $\exp(\sqrt{n \ln n}(1 + o(1)))$. Except for the function implicit in the $o(1)$ term, this bound is the same as the largest order of elements of S_n [La]; indeed one of the operations used in the proof of the diameter bound is raising an element to an arbitrary power, which makes the order of elements a bottleneck.

Theorem 1.1 seems to be the first result to overcome the "element order bottleneck." To achieve this, we have to avoid taking powers of permutations as a degree-reduction tool. Instead we increase the degree-reducing power of the commutator operation by applying it to carefully chosen conjugates τ and $\pi^{-1}\tau\pi$; the conjugating permutation π is obtained as the result of a polynomial-length random walk.

Our worst-case result brings us close to proving a polynomial upper bound for the "typical case." First we need to recall Dixon's celebrated result: almost all pairs of permutations in S_n generate either S_n or A_n [Dix] (cf. [Bo, Ba2]).

For random pairs of generators, the best diameter bound to-date states that the diameter is almost always at most $n^{\ln n(1/2+o(1))}$ [BH]. This expression also describes the order of almost all permutations [ET], so this result again suffers from the "element order bottleneck." In 99% (but not in almost all) cases we are now able to overcome this obstacle.

COROLLARY 1.1. *Fix $k \geq 2$ and $\delta > 0$. Let G be either S_n or A_n. Let S be a random subset of size k in G and let G_1 be the group generated by S. Then, with probability $> 1-\delta$ we have $\mathrm{diam}(G_1, S) \leq n^C$ for a constant $C = C(\delta)$.*

2 Proof of the worst-case bound: a Markov chain argument

We shall use the following fact. We explain the concepts involved after the statement.

FACT 2.1. *Let X be a connected undirected graph on n vertices, with maximum degree Δ and diameter d. Then the strong mixing time for a lazy random walk on X (starting from any vertex) is $O(d\Delta n \log n)$.*

A **lazy random walk** on the graph means a stochastic process where a particle moves from vertex to vertex; at each step, with probability 1/2 the particle does not move; the remaining probability 1/2 is split evenly between the neighbors. On a connected graph, lazy random walks are *ergodic* and therefore the distribution of the particle approaches the stationary distribution.

In the stationary distribution, each vertex is visited with a frequency proportional to its degree. So for regular graphs (as will always be the case in our applications), the stationary distribution is uniform.

The **strong mixing time** of a random walk means the number of steps needed for the random walk to reach a distribution close to the stationary distribution in the following sense: each state x is reached with probability $p(x)\mathrm{e}^{\pm\epsilon}$ where $p(x)$ is the stationary probability. Here $0 < \epsilon \leq 1/2$; the influence of the specific value of ϵ on the strong mixing time estimates is a factor of $\log(1/\epsilon)$ which we can ignore as long as ϵ is bounded away from zero.

The Fact remains true if we allow loops and multiple edges. The Fact follows from a result of Landau and Odlyzko [LO] stating that the eigenvalue gap of the transition matrix is at least $(d(\Delta+1)n)^{-1}$. The **eigenvalue gap** refers to the quantity $1-\max|\lambda|$ where the maximum is taken over all eigenvalues $\lambda \neq 1$ of the transition matrix. Note that 1 is a simple eigenvalue and the lazyness of the walk guarantees that $|\lambda| < 1$ for all other eigenvalues.

By standard arguments, the Landau-Odlyzko eigenvalue gap implies an $O(d\Delta n \log n)$ strong mixing time.

The proof of Theorem 1.1 is based on Lemma 2.2 below. We first state Lemma 2.1 in order to illustrate the technique on a simple instance.

LEMMA 2.1. *Let $G = \langle S \rangle$ be a transitive permutation group of degree n. Let $\delta > 0$ be fixed. Let $A \subset [n]$, $|A| = k$. Then there exists a word π of length $O(n^2|S| \log n)$ over $S \cup S^{-1}$ such that*

$$|A \cap A^\pi| \leq (k^2/n)(1+\delta).$$

Note that if π is chosen at random uniformly from G then the expected size of $A \cap A^\pi$ is exactly k^2/n (cf. [BE]). The content of the Lemma is that we can produce an intersection nearly this small by a permutation π which is a short word in the generators. We shall use a short random walk to find π.

First we note that without loss of generality, here and later we may assume $|S| < 2n$ by throwing away redundant generators. (It is shown in [Ba1] that the longest subgroup chain in S_n has length $< 2n$).

Proof. Consider the connected undirected graph Γ (possibly with loops and multiple edges) consisting of the vertex set $[n]$ and the edges $\{i, i^\sigma\}$, where $i \in [n]$ and $\sigma \in S \cup S^{-1}$. Note that this graph is regular of degree $2|S|$. This graph is also connected because G is transitive.

Let s denote the strong mixing time of the lazy random walk on Γ. So the lazy random walk on Γ, starting from any vertex, is ϵ-nearly uniformly distributed over $[n]$ after s steps, i. e., each vertex has probability $(1/n)\mathrm{e}^{\pm\epsilon}$ chance to be visited at time s. Let us choose $\epsilon := \ln(1+\delta)$, so $\epsilon \approx \delta$. We have $s = O(n^2|S| \log n)$ by the Fact.

Let π be a lazy random word of length s over $S \cup S^{-1}$. Then for any $x \in [n]$,

$$\mathrm{Prob}(x^\pi \in A) \leq (k/n)\mathrm{e}^\epsilon = (k/n)(1+\delta).$$

Adding these probabilities over all $x \in A$, we find that

$$E(|A \cap A^\pi|) \leq (k^2/n)(1+\delta).$$

Choose π with at most average value of $|A \cap A^\pi|$. ∎

A permutation group G is *t-transitive* if for any pair of ordered t-tuples of distinct elements of the permutation domain, $(x_1, \ldots, x_t)$ and $(u_1, \ldots, u_t)$, there exists $\sigma \in G$ such that for all i, $x_i^\sigma = u_i$.

LEMMA 2.2. *Let $G = \langle S \rangle$ be a $(t+1)$-transitive permutation group of degree n. Fix $\delta > 0$. Let $(x_1, \ldots, x_t)$ and $(u_1, \ldots, u_t)$ be t-tuples of distinct elements of $[n]$. Let $A \subset [n]$, $|A| = k$. Then there exists a permutation $\pi \in G$ expressible as a word of length $O(tn^{2t+2}|S| \log n)$ over $S \cup S^{-1}$ such that*

(a) $\quad |A \cap A^\pi| \leq t' + (k^2/n)(1+\delta)$;

(b) $\quad x_i^\pi = u_i \quad (i = 1, ..., t)$,

where t' is the number of pairs (x_i, u_i) which both belong to A.

Note that in our application below, $t = 2$.

Proof. Let Γ be the following graph: the vertices of Γ are the ordered $(t+1)$-tuples of distinct elements of $[n]$; an edge corresponds to a transition between vertices under some member of $S \cup S^{-1}$. This is a connected undirected graph.

Let s be the mixing time of the lazy random walk on Γ, starting from any vertex with parameter $\epsilon = \ln(1+\delta)/2$, so $\epsilon \approx \delta/2$. By the Fact, $s = O(tn^{2t+2}|S| \log n)$ since Γ is a regular graph of degree $2|S|$ with $n(n-1)\cdots(n-t) < n^{t+1}$ vertices).

Let π be a lazy random word of length s over $S \cup S^{-1}$. Let B denote the event that π satisfies condition (b). Then

$$\mathrm{Prob}(B) = (1/n(n-1)\cdots(n-t+1))\mathrm{e}^{\pm\epsilon},$$

and for any $x \notin \{x_1, \ldots, x_t\}$ and $u \notin \{u_1, \ldots, u_t\}$,

$$\mathrm{Prob}(B \text{ and } x^\pi = u) = (1/n(n-1)\cdots(n-t))\mathrm{e}^{\pm\epsilon}.$$

Therefore for the conditional probability we have

$$\mathrm{Prob}(x^\pi = u | B) = (1/(n-t))\mathrm{e}^{\pm 2\epsilon}.$$

Hence, as in the proof of Lemma 2.1, we obtain the following bound on the conditional expectation of $|A \cap A^\pi|$:

$$E(|A \cap A^\pi||B) \le t' + ((k-t')^2/n)\mathrm{e}^{2\epsilon}$$
$$= t' + ((k-t')^2/n)(1+\delta).$$

Therefore we may choose π to satisfy B as well as the inequality

$$|A \cap A^\pi| \le t' + ((k-t')^2/n)(1+\delta).$$

■

For a permutation σ we use $\mathrm{supp}(\sigma)$ to denote the **support** of σ, i. e., the set of elements moved by σ. Note that $|\mathrm{supp}(\sigma)|$ is the degree of σ.

PROPOSITION 2.1. *Let σ, τ be permutations and $A = \mathrm{supp}(\tau)$. Let $x \in A$ and $y = x^\tau$. If $x^{\sigma^{-1}} \in A$ and $y^{\sigma^{-1}} \notin A$ then τ and τ^σ do not commute.*

Proof. Let $\phi = \tau^\sigma$. Since $y^{\sigma^{-1}} \notin A$, we have $y^{\sigma^{-1}\tau} = y^{\sigma^{-1}}$, i.e., $y^\phi = y$. In other words, $x^\tau = x^{\tau\phi}$. Assume now that τ and ϕ commute. We infer that $x^\tau = x^{\phi\tau}$ and therefore $x = x^\phi = x^{\sigma^{-1}\tau\sigma}$. Hence $x^{\sigma^{-1}} = x^{\sigma^{-1}\tau}$, i.e., $x^{\sigma^{-1}} \notin A$, contrary assumption. ■

The *proof* of Theorem 1.1 now follows. We use Lemma 2.2 with $A = \mathrm{supp}(\tau)$, $t = 2$, $x_1, u_1 \in A$, $u_2 := u_1^\tau \in A$, and $x_2 \notin A$. Then, by Proposition 2.1, condition (b) implies that τ and τ^π do not commute. Note that we have $t' = 1$.

To prove Theorem 1.1, we use Lemma 2.2 repeatedly, starting from some τ of degree $< n/(3+\epsilon)$. In each round we replace τ by the commutator $[\tau, \tau^\pi]$ where π satisfies conditions (a) and (b) with $A = \mathrm{supp}(\tau)$; we choose x_i and u_i as in the preceding paragraph. Condition (b) guarantees that the new τ is not the identity. Condition (a) implies that the degree of τ decreases rapidly: if the degree of the old τ is k and the degree of the new τ is ℓ then $\ell/n < 3/n + 3(k/n)^2(1+\delta)$ for an arbitrarily small fixed δ. Since at the start, $k/n < 1/(3+\epsilon)$ for a fixed $\epsilon > 0$, in $O(\log\log n)$ rounds, we shall reach $k = 3$.

Each round increases the word length by quadrupling it and adding a fixed polynomially bounded amount ($O(n^6|S|\log n)$ from Lemma 2.2). This keeps the total word length bounded by $n^6|S|(\log n)^{O(1)}$, completing the proof of the diameter bound.

The justification of the algorithmic claim in Theorem 1.1 follows the lines of the proof of the diameter bound. The first step is to discard redundant generators; this can be done in polynomial time using Sims's algorithm ([Si, FHL, Kn], cf. Chaper 4 in [Se]). Let ξ denote the intersection size produced by our random walk method. For any $\delta > 0$, the probability that $\xi > (1+\delta)E(\xi)$ is less than $1/(1+\delta)$ by Markov's inequality. Let us repeat the random walk until $\xi \le (1+\delta)E(\xi)$ holds; it follows from the preceding observation that we shall succeed in an expected $O(1/\delta)$ trials. So the degree reduction will proceed at the same rate as in the proof of existence above. More precisely, we need to consider the conditional expectation of ξ conditioned on $x_i^\pi = u_i$ $(i = 1, 2)$; the probability that this happens is $O(n^{-2})$, so the cost of taking this condition into account is an additional factor of $O(n^2)$ in the running time. ■

3 Average case

LEMMA 3.1. *For every $\epsilon, \delta > 0$ there exists $C = C(\epsilon, \delta)$ such that all but an ϵ fraction of the $n!$ permutations of $[n]$ have a set of at most C cycles whose combined length is greater than $n(1-\delta)$.*

Proof. Follow the argument of the proof of Lemma 3.1 in [BH]; replace the reference to the Central Limit Theorem by Chebyshev's Inequality. ■

Let now σ and τ be two random permutations and let G_1 be the group generated by σ and τ. Setting $\delta = 0.33$, we find that by the Lemma, with probability $> 1-\epsilon$, the permutation σ has cycles of lengths $n_1, \ldots, n_k$ where $k \le C$ for some constant $C = C(\epsilon)$ such that $\sum_i n_i \ge 0.67n$. Let $N = \prod n_i < n^C$ and let $\sigma_1 = \sigma^N$. Now σ_1 has degree $\le 0.33n$, and almost surely $\sigma_1 \ne 1$ because the order of σ is almost surely $n^{\ln n(1/2+o(1))}$ [ET]. So with probability $\ge 1 - \epsilon - o(1)$, the Cayley graph $\Gamma(G_1, \{\sigma, \sigma_1, \tau\})$ has polynomialy bounded diameter. But obviously $\mathrm{diam}\,\Gamma(G_1, \{\sigma, \tau\}) \le N\,\mathrm{diam}\,\Gamma(G_1, \{\sigma, \sigma_1, \tau\})$, and the right-hand side is polynomially bounded. This completes the proof of Corollary 1.1. ■

4 Possible extension to almost all pairs of generators

For a permutation σ, let ℓ_2 denote the total length of its first two cycles. For a random permutation, the expected value of ℓ_2 is $3(n+1)/4$. Let $k_n \to \infty$ arbitrarily slowly. Let us pick a set S of k_n generators at random and let $\overline{\ell_2}$ denote the average of $\ell_2(\sigma)$ over $\sigma \in S$. Then the standard deviation of $\overline{\ell_2}$ is $o(n)$ and therefore, by Chebyshev's inequality, almost surely $\overline{\ell_2} \geq 0.7n$. Therefore, by the argument of the preceding section, $\operatorname{diam}\Gamma(S_n, S)$ is almost surely polynomially bounded.

This argument assumes no more than pairwise independence of the members of S. So if out of a random pair of permutations we could construct k_n pairwise independent permutations as short words, we would have proven that for almost all pairs of generators, the diameter is polynomially bounded.

Pairwise independence is surely too much to expect (in fact one can prove that it cannot be attained by more than three words in the generators), but it seems plausible that if $\sigma, \tau \in S_n$ are independent random permutations then the permutations $\sigma, \sigma\tau, \sigma\tau^2, \ldots, \sigma\tau^{k_n}$ are pairwise "nearly independent." What we would need is a concept of near-independence of pairs of permutations appropriate for this plan to go through.

References

[Ba1] L. Babai: On the length of subgroup chains in the symmetric group. *Communications in Algebra* **14** (1986), 1729–1736.

[Ba2] L. Babai: The probability of generating the symmetric group. *J. Comb. Theory – A* **52** (1989), 148–153.

[BE] L. Babai, P. Erdős: Representation of group elements as short products, in: "Theory and Practice of Combinatorics" (A. Rosa, G. Sabidussi, J. Turgeon, eds.), *Ann. Discr. Math.* **12** (1982), 21-26.

[BH] L. Babai, G. L. Hetyei: On the diameter of random Cayley graphs of the symmetric group. *Combinatorics, Probability, and Computing* **1** (1992), 201-208.

[BHKLS] L. Babai, G. Hetyei, W.M. Kantor, A. Lubotsky, Á. Seress: On the diameter of finite groups. *31st IEEE FOCS,* St. Louis MO, 1990, pp. 857–865.

[BS1] L. Babai, Á. Seress: On the diameter of Cayley graphs of the symmetric group. *J. Combinatorial Theory-A* **49** (1988), 175–179.

[BS2] L. Babai, Á. Seress: On the diameter of permutation groups. *Europ. J. Comb.* **13** (1992), 231–243.

[Bo] J. D. Bovey: The probability that some power of a permutation has small degree. *Bull. London Math. Soc.* **12** (1980), 47–51.

[Dix] J. D. Dixon: The probability of generating the symmetric group. *Math. Z.* **110** (1969), 199–205.

[DF] J. R. Driscoll, M. L. Furst: Computing short generator sequences. *Info. and Comput.* **72** (1987), 117–132.

[ET] P. Erdős, P. Turán: On some problems of a statistical group theory I., *Z. Wahrscheinlichkeitstheorie verw. Geb.* **4** (1965), 175–186.

[EG] S. Even, O. Goldreich: *The minimum length generator sequence is NP-hard. J. Algorithms* **2** (1981), 311–313.

[FHL] M. L. Furst, J. Hopcropft, E. M. Luks: Polynomial time algorithms for permutation groups. In: *Proc. 21st IEEE FOCS,* IEEE Computer Society, 1980, pp. 36–41.

[Je] M. R. Jerrum: The complexity of finding minimum length generator sequences. *Theoretical Computer Science* **36** (1985), 265–289.

[He] M-C. Heydemann: Cayley graphs and interconnection networks. In: *Graph Symmetry (Montreal, 1996),* NATO Adv. Sci. Inst. Ser. C, Math. Phys. Sci. 497, Kluwer Acad. Publ., Dordrecht 1997, pp. 167–224.

[Kn] D. E. Knuth: Notes on efficient representation of perm groups. *Combinatorica* **11** (1991), 57–68.

[Ko] R. E. Korf: Finding optimal solutions to Rubik's Cube using pattern databases. In: *Proc. 14th Nat. Conf. on Artificial Intelligence (AAAI-97)*, Amer. Assoc. for Artificial Intelligence, 1997, pp. 700-705.

[La] E. Landau: *Handbuch der Lehre von der Verteilung der Primzahlen.* Bd. I, Teubner, Leipzig, 1909.

[LO] H. J. Landau, A. M. Odlyzko: Bounds for eigenvalues of certain stochastic matrices. *Linear Algebra and its Appl.* **38** (1981), 5–15.

[McK] P. McKenzie: Permutations of bounded degree generate groups of polynomial diameter. *Info. Proc. Lett.* **19** (1984), 253–254.

[Se] Á. Seress: *Permutation Group Algorithms.* Cambridge Univ. Press, 2003.

[Si] C. C. Sims: Computation with permutation groups. In: *Proc. Second Symposium on Symbolic and Algebraic Manipualtion*, ACM Press 1971, pp. 23–28.