ON INVARIABLE GENERATION OF ALTERNATING GROUPS BY ELEMENTS OF PRIME AND PRIME POWER ORDER

ROBERT M. GURALNICK, JOHN SHARESHIAN, AND RUSS WOODROOFE

Abstract. We verify that every alternating group of degree at most one quadrillion is invariably generated by an element of prime order together with an element of prime power order.

1. Introduction

We say that a finite group G is *invariably generated* by the elements $h, g \in G$ if for every $x, y \in G$ it holds that $\langle h^x, g^y \rangle = G$. This paper is motivated by the following question, asked by the first author in work with Dolfi, Herzog, and Praeger [9].

Question 1.1. *Which finite simple groups are invariably generated by two elements of prime order? Of prime-power order?*

There are applications of invariable generation to computational Galois theory [7, 10], where it can be used to verify that the Galois group of a given polynomial is the full symmetric group. Invariable generation also yields useful fixed-point free actions on certain simplicial complexes that may be derived from a group [33].

Problems related to Question 1.1 have been studied elsewhere. King [22] shows that every finite simple group is generated (not necessarily invariably) by two elements of prime order. Every finite simple group is invariably generated by two elements (of unspecified order) [14, 21]. Invariable generation by a small number of random elements [7, 10, 27, 26] or by a sequence of random elements [21, 24] has received a fair bit of recent attention. Other related work on invariable generation includes [5, 6, 12].

It seems reasonable to believe, and was conjectured in [9], that all but finitely many simple groups of Lie type are invariably generated by two elements of prime order. For alternating groups, this is not the case. The second two authors gave infinitely many counterexamples in [34], by showing that A_n is not invariably generated by elements of prime order whenever $n > 4$ is a power of 2. On the other hand, another result of [34] shows that the set of integers

2000 *Mathematics Subject Classification.* 20D06, 20B30, 11Y99.

Work of the first author was partially supported by NSF Grant DMS-1901595 and Simons Foundation Fellowship 609771. Work of the second author was supported in part by NSF Grant DMS 1518389. Work of the third author is supported in part by the Slovenian Research Agency research program P1-0285 and research projects J1-9108, N1-0160, J1-2451, J3-3003.

Github: `RussWoodroofe/invgen` Dataset doi:10.5281/zenodo.5914767.

n where A_n is so generated has asymptotic density of 1 under assumption of the Riemann Hypothesis (and very close to 1 with no assumption).

In the current paper, we attack the problem of invariable generation of alternating groups by prime-power elements computationally, improving significantly on the prior computational results in [34]. Our main result is as follows.

Theorem 1.2. *For all n between 5 and 10^{15}, the alternating group A_n is invariably generated by an element of prime power order p^a together with an element of prime order r.*

Moreover, the prime r can be chosen for every $n > 24$ in this range so that $r \geq 2\sqrt{n}$, while p^a can be chosen to be one of the three largest prime powers dividing n for every n in this range except for 199,445,521,968 *and* 6,421,990,708,848.

The computation, as we will describe in greater detail below, uses an extension of the segmented sieve of Eratosthenes. A key observation is to notice that while we are sieving, we can simultaneously identify the integers that fail to be power-smooth, while also identifying the largest prime factor of each power-smooth integer.

After initially prototyping code in GAP [11], we implemented the most speed-critical part in C, using the primesieve library [37]. We ran this code using 15 concurrent processes for about 225 hours (or about 3062 process-hours) on a 16 core 3.3 GHz Intel Xeon E5 server, resulting in a list of numbers failing Lemma 3.3 below. We then checked this list of numbers in GAP on a 2.4 GHz Intel Core i5 MacBook Pro from 2019, which took about another 14 hours (single-threaded), and that completed the verification of the theorem.

The condition of invariable generation by an element of prime-power order and an element of prime order may be relaxed in several natural ways, as is studied in detail in [34]; see also [2]. Is every alternating group invariably generated by two elements of prime-power order? Or by any Sylow subgroups taken at two suitably-selected primes? All of these problems appear to remain open (beyond 10^{15}).

The question of generation by any Sylow subgroups taken at two primes is shown in [34, Theorem 1.3] to be equivalent to an elementary question on divisibility of binomial coefficients by the same primes. Working across this equivalence, the following is a straightforward corollary of Theorem 1.2.

Corollary 1.3. *For all n between 5 and 10^{15}, there are primes p and r so that every nontrivial binomial coefficient $\binom{n}{k}$ is divisible by at least one of p, r.*

The results of this paper improve on the computational results in [34] in two main ways: First, we examine a more restrictive generation condition. Second, we make better use of known algorithms and approaches from the computational number theory literature.

We expect to comprehensively address Question 1.1 for sporadic simple groups and for simple groups of Lie type in a forthcoming paper.

This paper is organized as follows. In Section 2, we give the necessary background from number theory and from the theory of alternating and symmetric groups. In Section 3, we give known and new lemmas used in the computation. In Section 4, we present the details

of our computation. In Section 5, we discuss relationships with the number theory literature and possible directions for future research.

Acknowledgements

We thank Steven Xiao at Washington University in St. Louis for giving us access to the computing hardware for our main computation, and Jernej Vičič at the University of Primorska for providing computer hardware used in earlier versions of the computation. Chris Jefferson and Marcus Pfeiffer answered some questions about GAP memory management. Kaisa Matomäki explained several aspects of her work and that of Joni Teräväinen (see Section 5), and of the number theory literature more broadly.

2. Background and notation

Our computation will combine number theory and group theory. A positive integer is *B-smooth* if it has no prime factor greater than B, and *B-power-smooth* if it has no prime-power factor greater than B.

In order for a set of elements S to generate a group G, it is necessary and sufficient that no maximal subgroup of G contains S. The maximal subgroups of A_n are well-known, and come in three flavors.

- The *intransitive maximal subgroups*, which fix a nonempty proper subset of $[n]$ under the natural action. Each such subgroup is isomorphic to a subgroup of index 2 in $S_k \times S_{n-k}$ for some $1 \leq k \leq n-1$, and has index $\binom{n}{k}$ in A_n.
- The *imprimitive maximal subgroups*, which are transitive but fix a partition of $[n]$ under the natural action. Each such subgroup is isomorphic to a subgroup of index 2 in $S_d \wr S_{n/d}$ for some proper divisor d of n, and has index $\binom{n}{d,d,\ldots,d}/(n/d)!$ in A_n.
- The *primitive maximal subgroups*, which satisfy neither of the above. This is the most difficult class to handle, but may be approached with the Classification of Finite Simple Groups and other results.

We discuss the primitive subgroups first. A classic theorem of Jordan is already a useful tool.

Theorem 2.1 (Jordan [20])**.** *No proper primitive subgroup of A_n contains a p-cycle for any prime p with $p \leq n-3$.*

One approach to proving Theorem 2.1 is to note that a primitive subgroup of S_n containing an $n-k$ cycle is $(k+1)$-transitive (see e.g. [8, Exercise 7.4.11]), a strong restriction. Combining this idea with the Classification yields substantial improvements to Theorem 2.1. We will make use of the following such improvement.

Theorem 2.2. *If a proper primitive subgroup H of A_n contains an $(n-2)$-cycle, then $n-1$ is a prime power.*

Proof. As we have observed, the action of H is 3-transitive. The 3-transitive subgroups of A_n are known from the Classification, as is laid out accessibly in [1, table following Theorem 5.3]. In particular, if $n > 24$ and $H \neq A_n$, then H has a minimal normal subgroup that is isomorphic to $PSL(2, q)$, and $n = q + 1$. □

Remark 2.3. A further generalization of Theorem 2.1 and Theorem 2.2 may be found in [19], where the Classification is used to completely characterize the primitive subgroups of A_n that contain a cycle of any length.

Another line of generalization of Theorem 2.1 is to replace the p-cycle with a product of a small number of p-cycles. Praeger did significant work along these lines in [28, 29, 30], shortly before the completion of the Classification Theorem. Later, Liebeck and Saxl in [23] used the Classification to describe completely (for large enough r) the groups containing a product of fewer than r cycles of length r. Although a full statement of their results takes several pages, we use the following consequence.

Theorem 2.4 (Liebeck and Saxl [23, Tables 1 and 2])**.** *Let r be a prime with $\sqrt{n} < r$, and let G be a proper primitive subgroup of A_n. If $y \in G$ is the product of $\lfloor \frac{n}{r} \rfloor$ disjoint r-cycles with $k = n - \lfloor \frac{n}{r} \rfloor \cdot r$ fixed points, then one of the following holds:*

(1) $n = \binom{c}{2}$, where $2 + r \leq c \leq \frac{3}{2}r - \frac{1}{2}$, and $k = \frac{1}{2}(r^2 + r - 2rc - c + c^2)$.
(2) $n = \frac{q^d - 1}{q - 1}$, where r divides $q^i - 1$, and $k = q^{d-i-1} + \cdots + 1 = \frac{q^{d-i}-1}{q-1}$.
(3) $k \leq 2$.
(4) $n = 24$.
(5) n is a prime power.

Remark 2.5. In the case of Theorem 2.4 where $n = \binom{c}{2}$, then

$$\binom{c}{2} = n \geq \frac{1}{2}(r + 2)(r + 1),$$

so $2n \geq (r + 1)^2$. In particular, $\sqrt{2n} - 1 \geq r$.

After handling the primitive subgroups with Theorem 2.4, the intransitive and imprimitive subgroups are dealt with using number theory (after some additional work). We will frequently use the following fact, which may be found for example in [32, after Theorem 7.27].

Lemma 2.6. *If P is a Sylow p-subgroup of A_n, then the orbits of P in the natural action of A_n on $[n]$ agree with the base p representation of n.*

More precisely, if $n = \alpha_0 p^0 + \alpha_1 p^1 + \cdots$, then P has α_0 fixed points, α_1 orbits of order p, α_2 orbits of order p^2, and so forth.

3. Lemmas

In this section, we give several lemmas that will be useful in the computation. First, it is easy to handle powers of primes.

Lemma 3.1. [34, Proposition 1.4B and proof] *If n is a power of the prime p, then A_n is invariably generated by an r-cycle for any prime r with $n/2 < r < n-2$, together with an element of p-power order.*

We now fix the following notation.

Notation 3.2. Throughout the following, n will be a large integer that is not a prime power, p will be a prime and p^a will be a prime-power divisor of n, and r will be another prime that is smaller than n (but which is "fairly large").

The following is very similar to results of [34, Lemma 1.8], and is proved in exactly the same manner.

Lemma 3.3. *If r and p^a are such that $r < n-2 < n < r+p^a$, then A_n is invariably generated by an r-cycle together with any fixed-point-free permutation x having no cycle of length less than p^a.*

Similarly if $r = n-2$ and $n-1$ is not a power of 2.

Proof. An r-cycle y avoids all primitive proper subgroups by Theorems 2.1 and 2.2. Moreover, the orders of the maximal imprimitive subgroups are not divisible by r. This leaves the intransitive maximal subgroups. Since $r + p^a > n$, the support of every cycle of x intersects the support of the unique nontrivial cycle of y, hence the subgroup generated by x and y is transitive. □

We say that $g \in A_n$ or S_n is a *base-p element* if the cycle structure of g agrees with the base p representation of n. That is, if n has base p representation $\alpha_0 p^0 + \alpha_1 p^1 + \cdots$, then a base-$p$ element of A_n has α_0 fixed points, α_1 cycles of length p, α_2 cycles of length p^2, and so forth. Thus, by Lemma 2.6, the orbits of a base-p element coincide with those of a Sylow p-subgroup. It is clear that A_n has a base-p element for every $p \neq 2$, and that A_n has a base-2 element if and only if the digit sum $\alpha_1 + \alpha_2 + \cdots$ is even.

We can take the element x in Lemma 3.3 to be a base-p element when one exists. More broadly, it is not difficult to describe the intransitive and imprimitive maximal subgroups containing a base-p element.

Lemma 3.4. *A base-p element of A_n fixes some set of size i in the natural action if and only if $p \nmid \binom{n}{i}$.*

Proof. The orbit system in the action on $[n]$ of a Sylow subgroup and of a base-p element coincide. The result now follows from the Orbit-Stabilizer theorem. □

There is also a simple condition that suffices to show that a base-p element does not fix a block system, as follows.

Lemma 3.5. *Suppose that the base-p element x of A_n preserves a partition π having e blocks of size d. Then no carry occurs when multiplying d by e in base p. Equivalently, if $d = \sum_{i\geq 0} \delta_i p^i$ and $e = \sum_{i\geq 0} \epsilon_i p^i$ are the base-p representations for d, e, then for each k it holds that $\sum_{0\leq i\leq k} \delta_i \epsilon_{k-i} < p$.*

Proof. The equivalence of the two conclusions is straightforward from definitions.

Let H be the stabilizer in S_n of π. By basic Sylow theory, the subgroup $\langle x \rangle$ is contained in a Sylow p-subgroup P of H, which is contained in a Sylow p-subgroup of S_n. It follows from Lemma 2.6 that the orbits of the three subgroups agree.

An order argument shows that $P \cong P_d \wr P_e$, where P_d and P_e are respectively Sylow p-subgroup of S_d and S_e. (Here, S_d acts on a block of π, while S_e acts on the set of blocks.) It now follows directly (see for example [32, Theorem 7.25]) that for each $k \geq 0$, the action of P on $[n]$ has $\alpha_k = \sum_{0 \leq i \leq k} \delta_i \epsilon_{k-i}$ orbits of size p^k. By definition of the base-p representation of n, we must have $\alpha_k < p$.

This completes the proof for odd p. For even p, we complete the proof by intersecting with A_n. □

Lemma 3.5 is less useful for large primes. For these we need another condition.

Lemma 3.6. *Suppose that $r > \sqrt{n}$, and let the base-r representation of n be $\beta_0 + \beta_1 r$, where $\beta_0 > 0$. If a base-r element x of A_n preserves a partition π having e blocks of size d, then either d or e divides* $\gcd(\beta_0, \beta_1)$.

Proof. Let d and e respectively have base-r representation $\delta_0 + \delta_1 r$ and $\epsilon_0 + \epsilon_1 r$. By a similar argument as in the proof of Lemma 3.5, we have $\beta_0 = \delta_0 \epsilon_0$, $\beta_1 = \delta_0 \epsilon_1 + \delta_1 \epsilon_0$, and at least one of δ_1, ϵ_1 must be 0.

Now if $\delta_1 = 0$, then $d = \delta_0$ divides $\beta_0 = \delta_0 \epsilon_0$ and $\beta_1 = \delta_0 \epsilon_1$; while if $\epsilon_1 = 0$, then $e = \epsilon_0$ similarly divides both β_0 and β_1. □

4. Strategy and details of computation

There are two requirements for verifying Theorem 1.2. We must be able to show that a given A_n is generated by an r-element and a p-power element, and we must compute very quickly.

4.1. **Strategy for checking generation.** The second two authors showed in [34, Section 5] that (a stronger statement than) the condition of Lemma 3.3 holds with high asymptotic density. Indeed, running our code on various ranges suggests that about $O(\sqrt{m})$ numbers in the range from 1 to m fail this condition.

Thus, the strategy for computation will be to find for each n the largest prime-power factor p^a of n and the largest prime r smaller than (or possibly equal to) $n-2$. If $p^a + r > n$, then a fixed-point free element of order p^a and an r-cycle invariably generate. This verifies invariable generation for most values of n.

For the remaining values of n, we seek numbers of the form cr strictly between $n - p^a$ and $n-2$, where r is a prime and c is smaller than $\sqrt{n/2}$. The product g of c cycles of length r is frequently helpful for invariable generation. By Theorem 2.4 and following remark, this g is not in any proper primitive subgroup of A_n, except possibly if $n = \frac{q^d-1}{q-1}$ and $n - c \cdot r = \frac{q^i-1}{q-1}$ for some $i \geq 2$. In particular, in this case $n-1$ and $n - c \cdot r - 1$ have a common prime power divisor, which is easy to detect computationally.

Having eliminated the possibility of a proper primitive subgroup containing g, we must also avoid intransitive and imprimitive subgroups. We let p_1 and p_2 be the primes associated with the largest two odd prime-power divisors of n, and combine g with a base-p_1 or base-p_2 element. (We avoid using 2 here for convenience, so that we can avoid checking whether A_n has a base-2 element.) We consider two primes, as there are situations where a given prime may fail to yield transitivity with one or all possible choices of r.

Example 4.1. Consider $n = 31416 = 2^3 \cdot 3 \cdot 7 \cdot 11 \cdot 17$. There is no prime $r < n$ such that $r + 17 > n$, so we must look at r such that $c \cdot r$ is close to $n - 3$. The prime $r = 7853$ is a useful such choice. However, both the product of four r-cycles and an appropriate base-17 element respect a partition of n into $2 \cdot 7854 + 2 \cdot 7854$. Here, the base 17 representation of n is $12 \cdot 17 + 6 \cdot 17^2 + 6 \cdot 17^3$, and $2 \cdot 7854 = 2 \cdot 7853 + 2 \cdot 1 = 6 \cdot 17 + 3 \cdot 17^2 + 3 \cdot 17^3$. The base 11 representation of n is the less symmetric $n = 7 \cdot 11 + 6 \cdot 11^2 + 1 \cdot 11^3 + 2 \cdot 11^4$, and indeed a base-11 element and a product of 7853-cycles generate A_{31416}.

We check the condition of Lemma 3.5 on each d dividing $\gcd(c, n - c \cdot r)$ (using Lemma 3.6), then check Lemma 3.4 on every i of the form $a \cdot r + b$. Out to 10^{15}, there is always such an r satisfying the desired conditions together with some prime divisor of n.

4.2. **Implementation.** We perform the computation in two phases.

Phase 1: Because most values of n satisfy the condition from Lemma 3.3, it is critical to check this condition quickly. Indeed, almost all of the time spent in computing is spent in verifying this condition.

We must find all or most of the primes out to the upper end of the range we are checking, and also find a large prime-power factor of each integer in the range. It is well-known that the Sieve of Eratosthenes is a fast algorithm for generating primes, but is space-hungry, taking $O(m)$ memory. More space-efficient is the Segmented Sieve of Eratosthenes. The Segmented Sieve uses the fact that each non-prime number n is divisible by some number smaller than $\sqrt{n}$. Thus, to compute primes out to m, we can find all primes smaller than $\sqrt{m}$, then repeatedly check numbers in segments of size $\sqrt{m}$ for divisibility by some small prime. This allows us to list primes out to m using $O(\sqrt{m})$ memory.

We modify the Segmented Sieve of Eratosthenes algorithm slightly to iterate over prime powers smaller than $\sqrt{m}$, rather than just primes. Then, instead of simply marking each n as not prime, we record the largest-seen prime-power divisor. Thus, at the end of sieving, we have found the largest prime-power divisor of each n that is smaller than $\sqrt{m}$. By performing a bit more bookkeeping, we are able to identify the numbers that are not B-power-smooth, for some B larger than the largest expected prime gap. We record B in place of the largest prime-power divisor of the non-B-power-smooth numbers. See [4, Chapter 3.2] for similar variations on the Segmented Sieve of Eratosthenes.

For further practical efficiency on real computer hardware, we use the primesieve library [37]. This library uses segments of length less than $\sqrt{m}$, which has advantages for cache efficiency, but which requires keeping different lists of primes. The library handles the

adjustments for smaller segments transparently, and has other optimizations. We sieve separately for power-smooth numbers and large prime-power divisors.

The upshot is that for each segment, we can quickly calculate the primes in the segment, and the largest prime-power divisors for power-smooth numbers in the segment. Now for each number n in each segment, we add the prime preceding $n-2$ with the largest prime-power divisor (or a large enough placeholder value if n is not power-smooth). If this is smaller than n, we see if $n-2$ is prime and $n-1$ is not a power of 2. If this fails, we record the number (together with its largest prime-power divisor) for Phase 2.

Remark 4.2. It is worth remarking that, for B large enough, we do not expect to ever see a number that passes to Phase 2 which is not B-power-smooth. That is, we expect to see primes on any interval of sufficient length before m. Our code uses $B = 5(\log_2 m)^2$, and gives a warning if it ever does pass a non-B-power-smooth integer on to Phase 2.

The output from Phase 1 in our computation to 10^{15} is publicly available as a dataset on the Zenodo repository [16], and comprises about 26.5 million numbers. The first segment (comprising the 0.8 million fairly power-smooth numbers out to 1 trillion) is also posted as an arXiv ancillary file. Code for both phases of the computation is available on GitHub [15].

Phase 2: At the conclusion of Phase 1, we have a relatively small number of fairly power-smooth numbers, where the specific smoothness threshold varies according to the distribution of primes. Running our code on various ranges suggests that we can expect around $O(\sqrt{m})$ such power-smooth numbers between 25 and m. In Phase 2, we use one of the order r elements suggested by Theorem 2.4 to show invariable generation for these leftover numbers.

Thus, for each n left over from Phase 1, we have already stored the largest prime-power factor. As these values of n are power-smooth, it is efficient to use trial division to find two more large prime-power factors. We keep the largest two odd prime-power divisors.

We then look for integers of the form cr that are close to n, where r is prime and c is small. Let p^a be the largest prime-power factor of n. As there is somewhat less to go wrong with transitivity for smaller values of c, we start by looking at c from 2 to $p^a/2$, and checking for a prime r so that $n - p^a < cr < n - 2$. If smaller values of c do not yield the desired, then we factor each number between $n - p^a$ and $n - 2$, and take r to be the prime factor (if any) that is larger than $2\sqrt{n}$.

We check for each candidate cr that n and $n - cr$ do not have the form $\frac{q^d-1}{q-1}$ and $\frac{q^i-1}{q-1}$ for some common q (avoiding the primitive groups of Theorem 2.4 not already eliminated by requiring $k > 2$). This yields an element of order r which avoids all the primitive subgroups of A_n, and which does not obviously fail to invariably generate A_n with a prime-power element.

We now check primitivity for the subgroup generated by an r-element and an element of order one of the two largest prime power divisors of n. We apply Lemmas 3.4, 3.5, and 3.6. Specifically, we check for each d dividing $\gcd(c, n - c \cdot r)$ that multiplication of d and n/d yields a carry in the associated prime base. Finally, we check for a system of intransitivity over each partition of n having a part $a \cdot r + b$, where $a \leq c$ and $b \leq n - cr$.

Although the checks that we need to do for these numbers are considerably more expensive than those of Phase 1, we only need to perform them on a small collection of numbers.

Remark 4.3. We prototyped the code for Phase 1 in GAP for ease of coding and experimentation. We then ported the code from this phase to C, speeding the computation by a factor of around 20. The code from Phase 2 is not a speed bottleneck, and the GAP library is convenient to use for some of the checks here, so we have kept it in GAP.

Remark 4.4. The average value of c over large ranges has been between 2 and 3 in experiments. Occasional numbers from Phase 2 require a c that is larger, however. See also Section 4.5.

4.3. **Complexity.** Phase 1 is at its core a Segmented Sieve of Eratosthenes, requiring $O(m \log\log m)$ operations and $O(\sqrt{m})$ space. Our use of prime powers instead of primes makes no difference in order of complexity, as prime powers have nearly the same density as primes.

We also store the power-smooth numbers failing Phase 1: these also appear experimentally to take just over $O(\sqrt{m})$ space. Indeed, if we assume Cramer's conjecture [3] (as certainly holds in the ranges in which we are likely to compute), then prime gaps are of at most $O(\log^2 x)$ size. An estimate of Rankin (in [31], see also [13]) gives the number of $\log^2 x$-smooth numbers in $[1, x]$ to be $x^{1/2+O(1/\log\log x)}$. Assuming that the number of $\log^2 x$-smooth numbers on an interval on a length of $\log^2 x$ is typically well-behaved, a estimate for the number of failures is $O(\int_1^m \frac{\log^2 x}{x} \cdot \sqrt{x}\, dx) = O(\sqrt{m} \cdot \log^2 m)$. Moderate optimizations in the use of space are likely possible, but as time is a much bigger bottleneck, we have not pursued this.

Phase 2 does not require any significant additional memory. It also does not appear to take much additional time, although a careful time analysis is more elusive. Let ℓ be the number of integers failing Phase 1. Then for each, we must in the worst case find and check $O(\sqrt{\ell})$ primes r. The checks for a PSL action and for primitivity are not expensive. The check for transitivity on an integer n of the form $c \cdot r + k$ requires examining each number of the form $a \cdot r + b$, where $0 \le a < c/2$ and $0 \le b \le k$.

Although making the analysis careful is difficult here, we give a heuristic argument for the time. By the same argument as in [34, Section 5], we expect there to be a prime in the short interval $[(n - p^a)/2, (n - 3)/2]$ with high asymptotic density. We must also satisfy other transitivity and primitivity tests, but these are also passed with high asymptotic density. Thus, we may assume that $c = 2$, so long as occasional numbers requiring a higher associated c do not cost greatly more time. Assuming Cramer's conjecture, we also have $k < \log^2 n$. Making broad assumptions as above, we have $\ell = O(\sqrt{m} \cdot \log^2 m)$. Assuming that we may find primes quickly, the time is easily subsumed within the $O(m \log\log m)$ time for Phase 1.

This heuristic bears up in practice, where Phase 2 runs quickly. When both phases were implemented in GAP, Phase 2 takes only as much time as that for a few segments in Phase 1.

p	base-p representation (place order is from least to most significant)		
83	n	$=$ $0, 30, 72, 44, 52, 50$	$= 30 \cdot p + 72 \cdot p^2 + 44 \cdot p^3 + 52 \cdot p^4 + 50 \cdot p^5$
	$2rp$	$0, 15, 49, 19, 34, 23$	
	$n - 2rp$	$0, 15, 23, 25, 18, 27$	
53	n	$0, 43, 23, 38, 48, 52, 8$	
	$4rp$	$0, 37, 11, 34, 33, 16, 5$	
	$n - 4rp$	$0, 6, 12, 4, 15, 36, 3$	
47	n	$0, 32, 38, 30, 29, 23, 18$	
	$3rp$	$0, 23, 13, 2, 26, 12, 7$	
	$n - 3rp$	$0, 9, 25, 28, 3, 11, 11$	
29	n	$0, 12, 13, 2, 22, 8, 16, 11$	
		transitive	
11	n	$0, 10, 6, 5, 8, 8, 7, 4, 6, 7, 7$	
	$6rp$	$0, 2, 2, 2, 4, 5, 6, 0, 6, 4, 1$	
	$n - 6rp$	$0, 8, 4, 3, 4, 3, 1, 4, 0, 3, 6$	
7	n	$0, 6, 6, 1, 5, 6, 5, 0, 3, 0, 6, 2, 0, 2$	
		transitive	
3	n	$0, 0, 0, 2, 2, 1, 1, 2, 0, 1, 2, 2, 0, 0, 2, 0, 1, 2, 1, 0, 0, 1, 0, 2$	
		transitive	

TABLE 1. Prime divisors p of $n = 199{,}445{,}521{,}968$, with a system of intransitivity for a base-p element and a base-r element, if any. Here $r = 555{,}558{,}557$.

4.4. **Small integers.** A few small cases included in Theorem 1.2 must be handled separately. Theorem 2.4 has an exception at $n = 24$, but this meets the condition of Phase 1 (with, say, $r = 19$ and $p^a = 2^3$). On the other hand, the numbers $n = 6$ and $n = 12$ cannot be handled by the checks in Phase 1. It is easy to verify using GAP or facts about primitive subgroups that A_6 is generated by a 5-cycle and a base-2 element, while A_{12} is generated by an 11-cycle and a base-3 element.

4.5. **The integer** 199,445,521,968**.** Our computational strategy works without trouble for all integers up to 10^{15} with five exceptions, as follow:

$n_1 = 199{,}445{,}521{,}968$, $n_2 = 5{,}760{,}706{,}652{,}536$, $n_3 = 6{,}421{,}990{,}708{,}848$,
$n_4 = 22{,}062{,}987{,}063{,}208$, $n_5 = 138{,}057{,}417{,}511{,}650$.

For these five integers, we modify the Phase 2 strategy slightly to also examine smaller odd prime-power divisors. All but n_1 and n_3 are generated by an element of large prime r order

(as in Phase 2), together with the base-p element for p the third largest (odd) prime-power divisor. For n_1 and n_3, we must use the fourth largest prime-power divisor for p.

To give some insight, we examine carefully the situation with $n = n_1 = 199{,}445{,}521{,}968 = 2^4 \cdot 3^3 \cdot 7 \cdot 11 \cdot 29 \cdot 47 \cdot 53 \cdot 83$. This value of n has several candidate large primes r that divide a slightly smaller integer. We focus on $n - 5 = 359 \cdot 555{,}558{,}557$, so $r = 555{,}558{,}557$. The corresponding r-element invariably generates A_n together with a base-p element for the odd primes $p = 3, 7$, or 29, but it fails transitivity for the primes $p = 11, 47, 53$, and 83. In Table 1, we show the base-p representation for n for each odd prime divisor n, together with a partition of n giving a system of intransitivity (if any).

Our main computation verified that the primes $p = 83$ and 53 fail to give invariable generation with any candidate value of r. We have verified by similar additional computation that the prime $p = 47$ fails transitivity for with any candidate value of r. However, the primes $p = 3$ and $p = 29$ each yield invariable generation of A_n at several values of $r \geq \sqrt{2n}$ (including $r = 555{,}558{,}557$, as we have already stated).

It is additionally worth noticing that $r = 728{,}209 > \sqrt{2n}$, a divisor of $n - 3$, fails to yield transitivity with any prime divisor of n.

5. Discussion

In light of our computational result, it is reasonable to ask:

Question 5.1. *For every $n \geq 5$, is the alternating group A_n invariably generated by an element whose order is a prime power divisor p^a of n, together with an element of prime order $r > \sqrt{n}$?*

We believe that counterexamples, if any, to the condition of Question 5.1 must be vanishingly rare. Theorem 1.2 says that any counterexample must satisfy $n > 10^{15}$.

It would already be somewhat interesting to give a proof that does not rely on the Riemann Hypothesis that the set of counterexamples to Question 5.1 has asymptotic density 0. (The second two authors gave a conditional proof of such a result with the stronger restriction that $a = 1$ in [34, Section 5].)

Note 5.2. After this paper was arXived, Teräväinen answered the question implicit in the previous paragraph in a strong form in [36], where he gave an explicit upper bound on the density of the set of n for which A_n is not generated by two elements of prime order.

5.1. **Number theoretic ingredients.** The strategy discussed in Section 4.1 suggests that an important ingredient for answering Question 5.1 will be the existence of integers with large prime factors on short intervals. This problem has been studied a certain amount in the number theory literature. Most of the papers on this problem focus on intervals of length close to that of $[x - \sqrt{x}, x]$. For sufficiently large x, a result of Harman [17, Theorem 6.1] shows we may find a number with a prime factor greater than $x^{37/50}$, while one of Jia and Liu [18] yields a number with a prime factor greater than $x^{25/26-\epsilon}$ on a slightly longer interval.

However, intervals of length $\sqrt{x}$ are already longer than we generally require. The set of integers that are $\log^{1+\epsilon} x$-smooth already have asymptotic density 0, as shown by Rankin in [31]. Meanwhile, a prime factor of order $x^{1/2+\epsilon}$ will suffice for use in Theorem 2.4, and $x^{25/26}$ is perhaps overkill.

Recent results of Teräväinen [35] (improving on earlier work by several authors) show that almost all intervals of the form $[x, x + \log^{3.51} x]$ contain a number that is the product of exactly two primes. Indeed, the more technical result [35, Theorem 5] shows that one of these primes must be at least $x/\log^3 x$. Slightly smaller primes (though still larger than $\sqrt{2x}$) on somewhat shorter intervals may be produced from Theorem 4 in the same paper. More recent related work of Matomäki [25] shows that almost all intervals of length $\log^{1+\epsilon} x$ have either a prime or a product of two primes.

5.2. **A heuristic for transitivity.** Since integers n that are $\log^{1+\epsilon} n$-smooth have asymptotic density 0, we can with asymptotic density 1 find r that is at least about $\sqrt{n}$ with $\lfloor \frac{n-3}{r} \rfloor \cdot r + p^a > n$, where p^a is the largest prime-power divisor of n. Indeed, the $k = 2$ and $k = 1$ cases that we avoid discussing in Theorem 2.4 are handled in [23], and can occur only with asymptotic density 0, so may be ignored for the purpose of asymptotic density arguments.

We assume we can find such a prime r. Frequently, we will have $\lfloor \frac{n-3}{r} \rfloor = 1$, and then the desired invariable generation holds by Lemma 3.3.

Otherwise, the main thing that can go wrong is for n to admit an integer partition into two parts $n_1 + n_2$ so that $n_1 = a \cdot r + b$ for $b < p^a$, and so that the base-p representation of n_1 has every digit smaller than the corresponding digit of the base-p representation of n. As seen in Example 4.1 and in Section 4.5, this indeed can happen for certain p. While we do not have a sketch that we know how to make precise, we know that the condition of Lemma 3.3 will fail only for highly power-smooth n. In this situation, the integer n will be divisible by several prime-powers of about the same size. Heuristically, changing the prime slightly will tend to alter the digits in the base-p expression greatly.

The example of Section 4.5 shows the extent to which this heuristic may fail.

References

[1] Peter J. Cameron, *Finite permutation groups and finite simple groups*, Bull. London Math. Soc. **13** (1981), no. 1, 1–22.

[2] Sílvia Casacuberta, *On the divisibility of binomial coefficients*, Ars Math. Contemp. **19** (2020), no. 2, 297–309.

[3] Harald Cramér, *On the order of magnitude of the difference between consecutive prime numbers*, Acta. Arith. **2** (1936), no. 1, 23–46.

[4] Richard Crandall and Carl Pomerance, *Prime numbers: a computational perspective*, second ed., Springer, New York, 2005.

[5] Eloisa Detomi and Andrea Lucchini, *Invariable generation of permutation groups*, Arch. Math. (Basel) **104** (2015), no. 4, 301–309.

[6] ———, *Invariable generation with elements of coprime prime-power orders*, J. Algebra **423** (2015), 683–701.

[7] John D. Dixon, *Random sets which invariably generate the symmetric group*, Discrete Math. **105** (1992), no. 1-3, 25–39.
[8] John D. Dixon and Brian Mortimer, *Permutation groups*, Graduate Texts in Mathematics, vol. 163, Springer-Verlag, New York, 1996.
[9] Silvio Dolfi, Robert M. Guralnick, Marcel Herzog, and Cheryl E. Praeger, *A new solvability criterion for finite groups*, J. Lond. Math. Soc. (2) **85** (2012), no. 2, 269–281, arXiv:1105.0475.
[10] Sean Eberhard, Kevin Ford, and Ben Green, *Invariable generation of the symmetric group*, Duke Math. J. **166** (2017), no. 8, 1573–1590.
[11] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.11.0*, 2020, (`http://www.gap-system.org/`).
[12] Daniele Garzoni and Andrea Lucchini, *Minimal invariable generating sets*, J. Pure Appl. Algebra **224** (2020), no. 1, 218–238.
[13] Andrew Granville, *Smooth numbers: computational number theory and beyond*, Algorithmic number theory: lattices, number fields, curves and cryptography, Math. Sci. Res. Inst. Publ., vol. 44, Cambridge Univ. Press, Cambridge, 2008, pp. 267–323.
[14] Robert Guralnick and Gunter Malle, *Simple groups admit Beauville structures*, J. Lond. Math. Soc. (2) **85** (2012), no. 3, 694–721.
[15] Robert M. Guralnick, John Shareshian, and Russ Woodroofe, *Invgen: Invariable generation of alternating groups*, GitHub RussWoodroofe/invgen, January 2022, GAP and C source code.
[16] ———, *Smooth numbers in large prime gaps*, Zenodo, February 2022, doi:10.5281/zenodo.5914767, dataset.
[17] Glyn Harman, *Prime-detecting sieves*, London Mathematical Society Monographs Series, vol. 33, Princeton University Press, Princeton, NJ, 2007.
[18] Chaohua Jia and Ming-Chit Liu, *On the largest prime factor of integers*, Acta Arith. **95** (2000), no. 1, 17–48.
[19] Gareth A. Jones, *Primitive permutation groups containing a cycle*, Bull. Aust. Math. Soc. **89** (2014), no. 1, 159–165.
[20] Camille Jordan, *Sur la limite du degré des groupes primitifs qui contiennent une substitution donnée*, J. Reine Angew. Math. **79** (1875), 248–258.
[21] W. M. Kantor, A. Lubotzky, and A. Shalev, *Invariable generation and the Chebotarev invariant of a finite group*, J. Algebra **348** (2011), 302–314.
[22] Carlisle S. H. King, *Generation of finite simple groups by an involution and an element of prime order*, J. Algebra **478** (2017), 153–173, arXiv:1603.04717.
[23] Martin W. Liebeck and Jan Saxl, *Primitive permutation groups containing an element of large prime order*, J. London Math. Soc. (2) **31** (1985), no. 2, 237–249.
[24] Andrea Lucchini, *The Chebotarev invariant of a finite group: a conjecture of Kowalski and Zywina*, Proc. Amer. Math. Soc. **146** (2018), no. 11, 4549–4562.
[25] Kaisa Matomäki, *Almost primes in almost all very short intervals*, J. Lond. Math. Soc. **106** (2022), no. 2, 1061–1097, arXiv:2012.11565.
[26] Eilidh McKemmie, *Invariable generation of finite classical groups*, J. Algebra **585** (2021), 592–615.
[27] Robin Pemantle, Yuval Peres, and Igor Rivin, *Four random permutations conjugated by an adversary generate $\mathcal{S}_n$ with high probability*, Random Structures Algorithms **49** (2016), no. 3, 409–428.
[28] Cheryl E. Praeger, *Primitive permutation groups containing an element of order p of small degree, p a prime*, J. Algebra **34** (1975), 540–546.
[29] ———, *Primitive permutation groups containing a p-element of small degree, p a prime. II*, J. Algebra **42** (1976), no. 1, 278–293.
[30] ———, *On elements of prime order in primitive permutation groups*, J. Algebra **60** (1979), no. 1, 126–157.

[31] Robert A. Rankin, *The difference between consecutive prime numbers*, J. London Math. Soc. **13** (1938), no. 4, 242–247.
[32] Joseph J. Rotman, *An introduction to the theory of groups*, Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, 1995.
[33] John Shareshian and Russ Woodroofe, *Order complexes of coset posets of finite groups are not contractible*, Adv. Math. **291** (2016), 758–773, arXiv:1406.6067.
[34] ———, *Divisibility of binomial coefficients and generation of alternating groups*, Pacific J. Math. **292** (2018), no. 1, 223–238, arXiv:1505.05143.
[35] Joni Teräväinen, *Almost primes in almost all short intervals*, Math. Proc. Cambridge Philos. Soc. **161** (2016), no. 2, 247–281.
[36] ———, *Almost all alternating groups are invariably generated by two elements of prime order*, preprint, 2022, arXiv:2203.05427.
[37] Kim Walisch, *Primesieve library*, (`https://github.com/kimwalisch/primesieve/ `), 2010–2022.

Department of Mathematics, University of Southern California, 3620 S. Vermont Ave Los Angeles, CA 90089-2532 USA
Email address: `guralnic@usc.edu`

Department of Mathematics, Washington University in St. Louis, St. Louis, MO, 63130
Email address: `jshareshian@wustl.edu`

Univerza na Primorskem, Glagoljaška 8, 6000 Koper, Slovenia
Email address: `russ.woodroofe@famnit.upr.si`
URL: `https://osebje.famnit.upr.si/~russ.woodroofe/`