The probability of generating the symmetric group when one of the generators is random

By LÁSZLÓ BABAI (Chicago) and THOMAS P. HAYES (Berkeley)

To the memory of Edit Szabó

Abstract. A classical result of JOHN DIXON (1969) asserts that a pair of random permutations of a set of $n$ elements almost surely generates either the symmetric or the alternating group of degree $n$.

We answer the question, "For what permutation groups $G \leq S_n$ do $G$ and a random permutation $\sigma \in S_n$ almost surely generate the symmetric or the alternating group?" Extending Dixon’s result, we prove that this is the case if and only if $G$ fixes $o(n)$ elements of the permutation domain.

The question arose in connection with the study of the diameter of Cayley graphs of the symmetric group.

Our proof is based on a result by Łuczak and Pyber on the structure of random permutations.

1. Introduction

By a random element of a nonempty finite set $S$ we mean an element chosen uniformly from $S$. A random permutation is a random element of the symmetric group $S_n$. A random pair of permutations is a random...
element of the set \( S_n \times S_n \). Our permutations always act on a domain of size \( n \). We consider the asymptotic behavior of random permutations as \( n \to \infty \).

Let \( \{E_n\} \) be a sequence of events. We say that \( E_n \) holds with high probability if \( \lim_{n \to \infty} P(E_n) = 1 \). Synonymously, we say that \( E_n \) occurs almost surely.

Dixon’s classical result states that with high probability, a random pair of permutations generates either \( A_n \) or \( S_n \) [Di1] (cf. [BW], [Ba]).

We strengthen this result, showing that one random permutation is enough as long as the other generators do not share more than \( o(n) \) fixed points (i.e., the fraction of fixed points in the permutation domain tends to zero). By a fixed point of a permutation group \( G \leq S_n \) we mean an element of the permutation domain fixed by all elements of \( G \).

**Theorem 1.** Let \( G \leq S_n \) be a given permutation group with \( o(n) \) fixed points. Let \( \sigma \in S_n \) be chosen at random. Then with high probability, \( G \) and \( \sigma \) generate either \( A_n \) or \( S_n \).

**Remark 2.** As usual, the precise meaning of such an asymptotic statement involving a \( o(n) \) bound is that for every \( \epsilon > 0 \) there exists \( \delta > 0 \) and a threshold \( n_0 \) such that for every \( n \geq n_0 \), if \( G \leq S_n \) has fewer than \( \delta n \) fixed points then the probability that \( G \) and a random \( \sigma \in S_n \) generate \( A_n \) or \( S_n \) is at least \( 1 - \epsilon \).

Of course if \( G \not\leq A_n \) then the result means that \( G \) and \( \sigma \) almost surely generate \( S_n \); and if \( G \leq A_n \) then with probability approaching \( 1/2 \), the group they generate is \( A_n \), and also with probability approaching \( 1/2 \) they generate \( S_n \).

This question arose in connection with the study of the diameter of Cayley graphs of the symmetric group [BH], [BBS], [BS]. It can also be viewed as a contribution to the “statistical group theory” initiated by Erdős and Turán in 1965 [ET].

We also observe that Theorem 1 is tight in the sense that the \( o(n) \) bound on the number of fixed points is necessary.

**Proposition 3.** If \( G \leq S_n \) has \( f \) fixed points then the probability that the group generated by \( G \) and a random permutation has a fixed point is \( \geq f/2n \).
2. Relation to Dixon’s Theorem

To see that Theorem 1 implies Dixon’s result, we only need to note that with high probability, a random \( \sigma \in S_n \) has \( o(n) \) fixed points. In fact much more is true: the number of fixed points is “almost bounded” in the following sense:

**Observation 4.** If \( \omega_n \to \infty \) arbitrarily slowly, then with high probability, a random \( \sigma \in S_n \) has at most \( \omega_n \) fixed points.

This follows from the fact that the probability that \( \sigma \) has \( \geq k \) fixed points is at most \( 1/k! \). Indeed, let \( d_n \) denote the probability that \( \sigma \in S_n \) is fixed-point-free (\( \sigma \) is a “derangement”). It is well known that \( d_n \to 1/e \).

**Observation 5.** The probability that a random permutation \( \sigma \in S_n \) has exactly \( k \) fixed points is \( d_n - k/k! \sim 1/e^k \). (The asymptotic equality holds uniformly for all \( k \) as long as \( n-k \) goes to infinity.)

In other words, the distribution of the number of fixed points of a random permutation is asymptotically Poisson with expected value 1.

3. The fixed-point-free case

The proof of Theorem 1 will be based on the following powerful result by Łuczak and Pyber.

**Theorem 6 ([LP]).** Let \( \sigma \in S_n \) be a random permutation. Then with high probability, \( \sigma \) does not belong to any transitive subgroup of \( S_n \) other than \( A_n \) or \( S_n \).

So to prove Theorem 1, we only need to show that \( G \) and \( \sigma \) generate a transitive subgroup with high probability. This will be established in Theorem 13 below. First we consider the case when \( G \) has no fixed point (Corollary 9).

We recall some terminology. Let us consider the symmetric group \( \text{Sym}(\Omega) \) acting on the permutation domain \( \Omega \), where \( |\Omega| = n \). Let \( G \leq \text{Sym}(\Omega) \) be a permutation group acting on \( \Omega \). We say that \( x, y \in \Omega \) belong to the same orbit of \( G \) if \( x^\tau = y \) for some \( \tau \in G \). The equivalence classes
of this relation are the *orbits* of \( G \) or \( G \)-orbits; they partition \( \Omega \). If \( A \subseteq \Omega \) is an orbit then \(|A|\) is called the *length* of this orbit. We say that \( G \) is *transitive* if \( \Omega \) is a single orbit (of length \( n \)). An element \( x \in G \) is a *fixed point* of \( G \) if \( \{x\} \) is an orbit (of length 1). We denote the set of fixed points of \( G \) by \( \text{fix}(G) \).

**Lemma 7.** Let \( G \leq S_n \) be a permutation group with \( t \geq 2 \) orbits, each of length \( \geq k \geq 2 \). Let \( \sigma \in S_n \) be chosen at random. Then the probability that \( G \) and \( \sigma \) generate a transitive group is greater than

\[
1 - \frac{t}{\binom{n}{k}} - \delta(n, k, t),
\]

(1)

where

\[
\delta(n, k, t) = \begin{cases} 
0 & \text{if } k > n/4; \\
\left(\frac{t}{n}\right)(1 + O(1/n)) & \text{if } k \leq n/4.
\end{cases}
\]

(2)

Here the constant hidden in the \( O(1/n) \) term is absolute.

**Proof.** Let \(|\Omega| = n\) and \( G \leq \text{Sym}(\Omega) \). Observe that \( k \leq n/2 \) and \( t \leq n/k \).

Let \( q(G) \) denote the probability that \( G \) and \( \sigma \) do not generate a transitive group.

Let \( \Pi = \Pi(G) = (A_1, \ldots, A_t) \) be the partition of \( \Omega \) into \( G \)-orbits. We refer to the \( A_i \) as the *blocks* of the partition \( \Pi \).

Let \( B \subset \Omega \). Let \( p_B \) denote the probability that \( B \) is invariant under \( \sigma \). Clearly, \( p_B = \frac{1}{\binom{|B|}{k}} \). Using the union bound,

\[
q(G) \leq \sum_{r=1}^{t-1} \sum_{B \in \mathcal{I}_r} p_B,
\]

(3)

where \( \mathcal{I}_r \) denotes the set of those unions \( B \) of \( r \) blocks of \( \Pi \) which satisfy \(|B| \leq n/2\). So \(|\mathcal{I}_r| \leq \binom{t}{r} \). Moreover, for \( B \in \mathcal{I}_r \), we have \( rk \leq n/2 \).

Therefore

\[
q(G) \leq \sum_{r=1}^{[n/2k]} \binom{t}{r} \frac{1}{\binom{n}{rk}} \leq \frac{t}{\binom{n}{k}} + \delta(n, k, t).
\]

(4)

The last inequality is vacuously true if \( k > n/6 \); the case \( k \leq n/6 \) is the content of the next proposition. \( \square \)
Proposition 8. Suppose $2 \leq k \leq n/6$ and $tk \leq n$. Then
\[
\sum_{r=3}^{\lfloor n/2k \rfloor} \binom{t}{r} \binom{n}{rk} = O \left( \frac{t^2}{n^{n/2k}} \right).
\]

Proof. Let $a_r = \binom{t}{r}$ and $b_r = \binom{n}{rk}$ and let $S(n, k, t) := \sum_{r=3}^{\lfloor n/2k \rfloor} (b_2 a_r)/(a_2 b_r)$. Our claim is that $nS(n, k, t)$ is bounded (for all $n$, $k$, $t$ satisfying the given constraints).

We observe that
\[
\binom{t}{r}^k \leq \binom{tk}{rk} \leq \binom{n}{rk}.
\]
Further we observe that for $r \geq 64$ and $rk \leq n/2$ we have
\[
\binom{n}{rk} > \binom{n}{2k}^4.
\]
Indeed,
\[
\binom{n}{64k} > \binom{n}{64k}^{64k} > \binom{en}{2k}^{6k} > \binom{n}{2k}^4.
\]
Combining inequalities (6) and (7) we obtain, for $r \geq 64$, that
\[
\frac{b_2 a_r}{b_r} < \frac{1}{b_2} \leq \frac{1}{\binom{t}{r}} < \frac{1}{n^2}.
\]
It follows that
\[
S_1(n, k, t) := \sum_{r=64}^{\lfloor n/2k \rfloor} \frac{b_2 a_r}{a_2 b_r} < \frac{1}{n}.
\]
It remains to bound the sum
\[
S_2(n, k, t) := \sum_{r=3}^{m} \frac{b_2 a_r}{a_2 b_r},
\]
where $m = \min\{63, \lfloor n/2k \rfloor \}$.

Obviously,
\[
S_2(n, k, t) \leq \sum_{r=3}^{m} \frac{b_2 a_m}{b_3} < \frac{n^{64}b_2}{b_3}.
\]
Now
\[
\frac{b_2}{b_3} < \left( \frac{3k}{n-2k} \right)^k.
\] (13)

Since \( k \leq n/6 \), the right hand side is less than \((3/4)^k\); so we obtain the estimate \( S_2(n,k,t) < n^{64}/(3/4)^k \leq 1/n \) if \( k \geq 65 \log n/\log(4/3) \).

Assume now that \( k < 65 \log n/\log(4/3) \). It follows that for large enough \( n \) we have \( 3k/(n-2k) < 1/\sqrt{n} \) and so \( S_2(n,k,t) < n^{64}b_2/b_3 < n^{64}n^{-k/2} \leq 1/n \) assuming \( k \geq 130 \).

Now let us assume \( k \leq 129 \). Then
\[
(b_2a_r)/(a_2b_r) = \Theta(t^{r-2}/n^{k(r-2)}) = O(n^{-(k-1)(r-2)}) = O(1/n),
\] (14)
proving that \( S_2(n,k,t) = O(1/n) \). \( \square \)

**Corollary 9.** Let \( G \leq S_n \) be a permutation group with no fixed points. Let \( \sigma \in S_n \) be chosen at random. Then the probability that \( G \) and \( \sigma \) do not generate a transitive group is less than \( 1/n + O(1/n^2) \).

### 4. Projections

Next we define a projection operator, introduced in [BH], a useful tool for extending results about fixed-point-free groups to the general case. While a direct proof of Theorem 1 would be somewhat shorter, we find that separating the fixed-point-free case and then arriving at the general conclusion via the projection machinery provides greater insight and a general methodology.

We take a subset \( T \) of the permutation domain \( \Omega \) and a permutation \( \sigma \in \text{Sym}(\Omega) \) and assign to it a permutation \( \sigma_T \in \text{Sym}(T) \). Informally, \( \sigma_T \) is obtained by deleting those orbits of \( \sigma \) which lie entirely outside \( T \) and contracting those segments of the remaining orbits which lie outside \( T \).

The formal definition follows.

**Definition 10.** For \( T \subseteq \Omega \), we define the projection \( \text{pr}_T : \text{Sym}(\Omega) \to \text{Sym}(T) \), as follows. Let \( \sigma \in \text{Sym}(\Omega) \). We set \( \sigma_T = \text{pr}_T(\sigma) \) and define \( \sigma_T \).

For \( i \in T \), let \( k \) denote the smallest positive integer such that \( i^{\sigma_k} \in T \). Set \( i^{\sigma_T} = i^{\sigma_k} \).
We now observe two basic facts about projections.

**Observation 11.** Let $T \subseteq \Omega$. The projection map $\text{pr}_T : \text{Sym}(\Omega) \to \text{Sym}(T)$ is uniform, i.e., for all $\tau \in \text{Sym}(T)$, the size of $\text{pr}_T^{-1}(\tau)$ is the same ($|\Omega|!/|T|!$).

**Proof.** Let $\tau \in \text{Sym}(T)$. Let $\lambda : \Omega \setminus T \to \Omega$ be an injection. It is easy to see that there is a unique $\sigma \in \text{Sym}(\Omega)$ such that $\sigma|_{\Omega \setminus T} = \lambda$ and $\sigma_T = \tau$. Indeed, if $i^\sigma = j$ then (a) if $j$ is not in the range of $\lambda$ then let $i^\sigma = i^\tau$; (b) if $j = \ell^\lambda$ for some $\ell \in \Omega \setminus T$ then let $k$ be the largest integer such that $j = m^\lambda$ for some $m \in \Omega \setminus T$ and set $i^\sigma = m$. These are the only possible choices under the given constraints. We conclude that $|\text{pr}_T^{-1}(\tau)|$ is equal to the number of injections $\lambda$ regardless of the choice of $\tau$. □

**Observation 12.** Let $\sigma \in \text{Sym}(\Omega)$ and let $T \subseteq \Omega$. Let $G \leq \text{Sym}(T)$ where $\text{Sym}(T)$ is viewed as a subgroup of $\text{Sym}(\Omega)$. Then the orbits of the subgroup of $\text{Sym}(T)$ generated by $G$ and $\sigma_T$ are precisely the intersection of $T$ with those orbits of the subgroup of $\text{Sym}(\Omega)$ generated by $G$ and $\sigma$ which have non-empty intersection with $T$.

**Proof.** Clear. □

**Theorem 13.** Let $G \leq S_n$ be a given permutation group with $f \leq n/2$ fixed points. Let $\sigma \in S_n$ be chosen at random. Then the probability that $G$ and $\sigma$ do not generate a transitive group is less than $(f+1)(1/n + O(1/n^2))$.

In particular, if $G$ has $o(n)$ fixed points then $G$ and $\sigma$ generate a transitive group with high probability.

**Proof.** Let $A = \text{fix}(G)$; so $|A| = f$. The probability that a subset $B \subseteq A$ is invariant under $\sigma$ is, as before, $p_B = 1/\binom{n}{|B|}$. Let $i(A)$ denote the probability that such an invariant nonempty subset exists. By the union bound,

$$i(A) \leq \sum_{\emptyset \neq B \subseteq A} p_B = \sum_{r=1}^{f} \frac{\binom{f}{r}}{\binom{n}{r}} = \frac{f}{n} + O \left( \left( \frac{f}{n} \right)^2 \right). \quad (15)$$

Let now $H$ denote the group generated by $G$ and $\sigma$ and let $R = \Omega \setminus A$ (the domain where $G$ actually acts). Let $\sigma_R$ be the projection of $\sigma$ to $R$ (see Definition 10). By Observation 12, two elements $x, y \in R$ belong to the same orbit under $H$ if and only if they belong to the same orbit of
the group generated by $G$ and $\sigma_R$. Observing further that $\sigma_R$ is uniformly distributed in $\text{Sym}(R)$ (Observation 11) we conclude, using Corollary 9, that the probability that not all elements of $R$ are in the same orbit under $H$ is $\leq 1/(n - f) + O(1/(n - f)^2) = 1/n + O((f + 1)/n^2)$.

Finally, the probability that $H$ is not transitive is at most the sum of this quantity and $i(A)$, which in turn is $(f + 1)/n + O((f + 1)/n^2)$. □

5. Case: many fixed points

We now prove Proposition 3. Let $A$ be a subset a size $f$ of the permutation domain of size $n$. Let $\sigma$ be a random permutation. Let $p(f)$ denote the probability that $\sigma$ fixes at least one element of $A$.

**Claim 14.**

$$p(f) \geq \frac{f}{2n}. \quad (16)$$

**Proof.** The probability that a given point is fixed by $\sigma$ is $1/n$; the probability that a given pair of points is fixed by $\sigma$ is $1/n(n - 1)$. Hence, by Bonferroni’s Inequalities (truncated Inclusion-Exclusion),

$$p(f) \geq \frac{f}{n} - \frac{f(f - 1)}{n(n - 1)} = \frac{f}{n} \left(1 - \frac{f - 1}{2(n - 1)}\right) \geq \frac{f}{2n}. \quad (17)$$

□

To prove Proposition 3, we apply the Claim to the set of fixed points of $G$. □

6. Open problems

LUCZAK and PYBER [LP] do not provide an explicit bound on the probability that a random permutation belongs to a transitive group other than $S_n$ or $A_n$ (Theorem 6); this probability presumably goes to zero rather slowly. The first problem we propose is to estimate this rate.

The second problem is to find a proof of Theorem 1 which is independent of the Luczak–Pyber Theorem and provides a faster rate of convergence. Specifically, we propose the following
The probability of generating the symmetric group... 279

**Conjecture 15.** There exists $c > 0$ such that for all permutation groups $G \leq S_n$ if $G$ has no fixed point then the probability that $G$ together with a random permutation does not generate $A_n$ or $S_n$ is $O(n^{-c})$.

In this connection we should mention that the probability that a random pair of permutations does not generate $S_n$ or $A_n$ is $1/n + O(1/n^2)$ [Ba]. The full asymptotic expansion of this probability was recently given by Dixon [Di2].

It is a long standing conjecture that all Cayley graphs of $S_n$ and $A_n$ have polynomially bounded diameters ([KMS], [BS]). In [BH], the authors prove that for almost all pairs of permutations $\sigma, \tau \in S_n$, the Cayley graph of the group $G$ generated by $\sigma$ and $\tau$ has polynomially bounded ($O(n^c)$) diameter. (Note that by Dixon’s result, $G$ is almost surely $S_n$ or $A_n$.) It is our hope that Theorem 1 will help extend this result to the case when only $\sigma$ is random; $\tau$ is a given permutation with few fixed points.\footnote{We can almost prove this already. The only case still eluding us is when $\tau$ has few fixed points but has constant order.}

**References**


