Representing Random Permutations as the Product of Two Involutions

Charles Burnette
Department of Mathematics
Drexel University
Philadelphia, PA 19104-2875
cdb72@drexel.edu

Eric Schmutz
Department of Mathematics
Drexel University
Philadelphia, PA 19104-2875
Eric.Jonathan.Schmutz@drexel.edu

August 22, 2015

Abstract

An involution is a permutation that is its own inverse. Given a permutation $\sigma$ of $[n]$, let $N_n(\sigma)$ denote the number of ways to write $\sigma$ as a product of two involutions of $[n]$. If we endow the symmetric groups $S_n$ with uniform probability measures, then the random variables $N_n$ are asymptotically lognormal.

The proof is based upon the observation that, for most permutations $\sigma$, $N_n(\sigma)$ can be well approximated by $B_n(\sigma)$, the product of the cycle lengths of $\sigma$. Asymptotic lognormality of $N_n$ can therefore be deduced from Erdős and Turán’s theorem that $B_n$ is itself asymptotically lognormal.
1 Introduction

An involution is a permutation that is its own inverse, i.e. a permutation whose cycle lengths are all less than or equal to two. If $\sigma$ is a permutation of $[n]$, let $N_n(\sigma)$ be the number of ordered pairs of involutions $\tau_1, \tau_2$ of $[n]$ such that $\sigma = \tau_2 \circ \tau_1$. The goal of this paper is to determine the asymptotic distribution of the random variable $N_n$ for uniform random permutations $\sigma$.

Let $T_n$ be the set of all involutions of $[n]$. The cardinalities $|T_n|, n = 1, 2, 3, \ldots$ have been extensively investigated and form OEIS Sequence A000085 [25]. See also Amdeberhan and Moll [1] for more recent work. Of particular importance for this paper is an asymptotic formula that was derived by Chowla, Herstein, and Moore [8]:

$$|T_n| \sim \frac{1}{\sqrt{2}} \left( \frac{n}{e} \right)^{n/2} e^{n^{1/2} - 1/4}.$$ (1.1)

Related approximations appear in Moser and Wyman [19], [20].

Vivaldi and Roberts [22] studied the random permutations that are obtained by multiplying random involutions with various restrictions on their fixed points. However the product of two uniformly random involutions is not a uniformly random permutation. For example the identity permutation is generated with probability $1/|T_n|$, which is much larger than $1/n!$. Thus $N_n$ is clearly not constant.

Let $I_{\tau_2, \tau_1}(\sigma) = 1$ if $\tau_2 \circ \tau_1 = \sigma$ (and $I_{\tau_2, \tau_1}(\sigma) = 0$ otherwise), so that

$$N_n = \sum_{\tau_1, \tau_2} I_{\tau_2, \tau_1}.$$ (1.2)

Using this representation and Stirling's formula, it is straightforward to estimate the average number of factorizations [17]:

$$\mathbb{E}_n(N_n) = \frac{1}{n!} \sum_{\tau_1, \tau_2} \sum_{\sigma} I_{\tau_2, \tau_1}(\sigma) = \frac{|T_n|^2}{n!} \sim \frac{e^{2\sqrt{n}}}{\sqrt{8\pi n}}.$$ (1.3)

Our results show that the average in (1.3) is misleadingly large; if $n$ is large, then for most permutations $\sigma \in S_n$, one has

$$e^{(\frac{1}{2} - \epsilon) \log^2 n} < N_n(\sigma) < e^{(\frac{1}{2} + \epsilon) \log^2 n}.$$  

Another consequence of the sum of indicators representation (1.2) is that $\max_\sigma N_n(\sigma) = |T_n|$. The unique permutation that attains the maximum is the identity permutation that fixes all $n$ points. At the other extreme, for $n \geq 2$, $\min_\sigma N_n(\sigma) = n - 1$. The minimum is attained only by the $\frac{n!}{n-1}$ permutations that have a cycle of length $n - 1$. These two extremal results are stated on page 161 of Lugo’s thesis [17] and are also proved later in [7]. Lugo also conjectured, but did not prove, that $N_n$ is asymptotically lognormal.
There is an extensive literature on formulas for the number of ways to write a permutation as the product of two or more permutations with various restrictions on the conjugacy classes of the factors of the product. Without trying to review that literature, we refer readers to [13], [14] as possible starting points. For asymptotic problems, even an explicit formula can be quite useless if it is too complicated. However, as the authors in [13] and [14] point out, formulas with non-negative terms tend to be more tractable. In this paper, we make use of one such formula:

\[ N_n(\sigma) = \prod_{k=1}^{n} \sum_{j=0}^{\lfloor c_k/2 \rfloor} \frac{k^{c_k - j} c_k!}{2^j j!(c_k - 2j)!}, \quad (1.4) \]

where \( c_k = c_k(\sigma) \) denotes the number of cycles of length \( k \) that \( \sigma \) has. As far as we know, the first complete proofs of (1.4) are in Petersen and Tenner [21] and Lugo [17].

We use the formula (1.4) to prove that, for most permutations \( \sigma \), \( N_n(\sigma) \) can be well approximated by \( B_n(\sigma) = \prod_k k^{c_k} \), the product of the cycle lengths of \( \sigma \). The random variable \( B_n \) has been studied by many authors, beginning with the work of Erdős and Turán [10], [11]. Asymptotic lognormality of \( N_n \) will be deduced from the known fact that \( B_n \) is asymptotically lognormal.

2 Factorizations

This section is more or less expository: we discuss the known factorization (1.4). For each integer \( x \), let \( \overline{x} = x - n \lfloor \frac{x}{n} \rfloor \) denote the integer remainder when \( x \) is divided by \( n \). (The positive integer \( n \) will be clear from context.) Yang, Ellis, Mamakani, and Ruskey [28] proved the following lemma.

**Lemma 2.1.** There are exactly \( n \) ways to factor the \( n \)-cycle \( \sigma = (0,1,\ldots,n-1) \) as the product of two involutions of \( \{0,1,2,\ldots,n-1\} \). The \( n \) factorizations are \( \sigma = I_k \circ I_{k-1} \), \( 1 \leq k \leq n \), where \( I_k(x) = k - x \) is the integer remainder when \( k - x \) is divided by \( n \).

Our notational preference for modular arithmetic is influenced by page 158 of [12], where the setting is different but the factorization is similar. In [28], the proof of lemma 2.1 is quite short, elementary, and easy to read. As we show in proposition 2.4 below, the proof of lemma 2.1 can be adapted to the product of two \( m \) cycles, and therefore can be used as the basis for an alternative proof of (1.4). Corresponding lemmas appear in [17] and [21], but the derivations there are based on a graph theoretical insight and appear to be different from the proof that is presented here.

For any permutation \( \sigma \), we can apply lemma 2.1 separately to each of the cycles of \( \sigma \). Therefore a consequence of lemma 2.1 is that the product of the cycle lengths is a lower bound:

\[ N_n(\sigma) \geq B_n(\sigma). \quad (2.1) \]
This inequality is not sharp because, in the factorization $\sigma = \tau_2 \circ \tau_1$, there is no requirement that the cycles of $\sigma$ are invariant under the involutions $\tau_1$ and $\tau_2$. For example, we can write $\sigma = (1,2,3)(4,5,6)$ as $\tau_2 \circ \tau_1$, where $\tau_2 = (1,4)(2,6)(3,5)$ and $\tau_1 = (1,6)(2,5)(3,4)$. Both involutions "exchange" the elements of $\{1,2,3\}$ with those of $\{4,5,6\}$. The next lemma asserts that there are no other possibilities.

**Lemma 2.2.** Suppose $\mathcal{O}$ is the set of points on a cycle of $\sigma$, and that $\sigma = \tau_2 \circ \tau_1$ is a factorization of $\sigma$ into two involutions. Then $\tau_1(\mathcal{O}) = \tau_2(\mathcal{O})$, and $\tau_1(\mathcal{O})$ is the set of points on a cycle of $\sigma$ of length $|\mathcal{O}|$.

**Proof.** Because each $\tau_i$ is a bijection, it is clear that $|\tau_1(\mathcal{O})| = |\tau_2(\mathcal{O})| = |\mathcal{O}|$.

Suppose $y_1, y_2$ are points in $\tau_1(\mathcal{O})$. We need to verify that $y_1$ and $y_2$ are on the same cycle of $\sigma$. Let $x_1, x_2$ be the their preimages on $\mathcal{O} : \tau_1(n) = y_i$, $i = 1, 2$. Because $x_1$ and $x_2$ are on the same cycle $\mathcal{O}$, we have $x_2 = \sigma^\ell(x_1)$ for some $\ell$. But then $y_2 = \tau_1(\sigma^\ell(x_1)) = \tau_1 \circ (\tau_2 \circ \tau_1)^\ell(x_1) = (\tau_1 \circ \tau_2)^\ell \circ \tau_1(x_1) = -\ell(y_1)$. Thus $y_1$ and $y_2$ are on the same cycle, and $\tau_1(\mathcal{O})$ is a single cycle of length $|\mathcal{O}|$.

Finally, note that $\tau_2 = \sigma \circ \tau_1$. If $x \in \mathcal{O}$, then the set of points on the cycle of $\sigma$ that contains $\tau_2(x)$ is $\{v : v = \sigma^t \circ \tau_2(x)$ for some $t \in \mathbb{Z}\}$ = $\{v : v = \sigma^{t+1} \circ \tau_1(x)$ for some $t \in \mathbb{Z}\}$, and the latter set is the set of points on the cycle of $\sigma$ that contains $\tau_1(x)$. This proves that $\tau_1(\mathcal{O}) = \tau_2(\mathcal{O})$; the two involutions both map $\mathcal{O}$ to the same cycle. □

**Definition 2.3.** Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two distinct sets of points on cycles of $\sigma$. Two involutions $\tau_1, \tau_2$ exchange $\mathcal{O}_1$ and $\mathcal{O}_2$ provided that $\sigma = \tau_2 \circ \tau_1$ and $\tau_1(\mathcal{O}_1) = \tau_2(\mathcal{O}_1) = \mathcal{O}_2$.

**Lemma 2.4.** If $\sigma = (0,1,2,\ldots,n-1)(n,n+1,n+2,\ldots,2n-1)$, then there are precisely $n$ ways to write $\sigma$ as a product of two involutions of $\{0,1,\ldots,2n-1\}$ that exchange the two cycles of $\sigma$.

**Example:** If $n = 5$, then one of the five factorizations is $(0,1,2,3,4)(5,6,7,8,9) = J_3 \circ J_2$, where $J_3 = (0,8)(1,7)(2,6)(3,5)(4,9)$ and $J_2 = (0,7)(1,6)(2,5)(3,9)(4,8)$.

**Proof.** Let $X = \{0,1,\ldots,2n-1\}$. For integral $k$, define $J_k$ to be the involution whose $n$ transpositions are $(x, n + \overline{k} - x)$, $x = 0,1,2,\ldots,n-1$. Note that $J_k(x) = J_{k+\overline{n}}(x)$, so we are free to calculate the index $k$ modulo $n$. Also note that if $y = n + \overline{k} - x$, then $J_k(y) = x$. Hence it is straightforward to verify that, for any integer $k$, $\sigma = J_k \circ J_{k-1}$. Since there are $n$ choices for $\overline{k}$, this proves that there at least $n$ of the factorizations.

Now suppose $\sigma = S \circ T$ for some involutions $S$ and $T$ on $X$, and suppose $S$ and $T$ exchange the two cycles of $\sigma$. Because $S$ exchanges the cycles of $\sigma$, there must be some $k$ for which $S(0) = n + \overline{k}$. To prove the lemma, it suffices to prove that $S = J_k$ and $T = J_{k-1}$.

We use induction to show that, for $0 \leq i < n$, $S(i) = n + \overline{k} - i$ and $T(i) = n + \overline{k-1} - i$.

For the base case $i = 0$, we already have $S(0) = n + \overline{k}$. Note that $T(n + \overline{k-1}) = S^2 \circ T(n + \overline{k-1}) = S \circ \sigma(n + \overline{k-1}) = S(n + \overline{k}) = 0$. Therefore $T(0) = n + \overline{k-1}$. This completes the base case $i = 0$. 

4
Now let $0 < i < n-1$, and assume the inductive hypothesis. Since $i = \sigma(i-1) = ST(i-1)$, we have

$$S(i) = S^2 \circ T(i-1) = T(i-1) = n + \overline{k-1} - (i-1) = n + \overline{k-i}.$$  

Similarly

$$T(n + \overline{k-i-1}) = S^2 \circ T(n + \overline{k-1} - 1) = S \circ \sigma(n + \overline{k-i-1}) = S(n + \overline{k-i}) = i.$$  

Therefore

$$T(i) = n + \overline{k-1} - i.$$  

For non-negative integers $m$ and $k$ define

$$V_m(k) = \sum_{j=0}^{[m/2]} \frac{k^{-j} m!}{2^j j! (m - 2j)!} \frac{H_{m}(i \sqrt{k})}{(i \sqrt{k})^m}, \quad (2.2)$$

where $H_{m}$ is the “probabilists’ Hermite polynomial” $H_{m}(x) = m! \sum_{r=0}^{[m/2]} \frac{(-1)^r}{r! (m - 2r)!} \frac{x^{m-2r}}{2^r}$. We thank Victor Moll for pointing out this connection with the Hermite polynomials. A less general version appears as equation 2 of Moser and Wyman [19].

**Theorem 2.5.** (Lugo, Petersen, Tenner) If $c_k(\sigma)$ denotes the number of $k$-cycles that $\sigma \in S_n$ has, then

$$N_n(\sigma) = B_n(\sigma) \prod_{k=1}^{n} V_{c_k}(k)$$

**Proof.** By lemma 2.2, any involution factorization of $\sigma$ exchanges some number of pairs of cycles of the same size, and leaves the rest fixed. For each $j \leq \lfloor c_k/2 \rfloor$, there are precisely $\frac{k^j}{2^j j!(c_k - 2j)!}$ ways to match $j$ pairs of $k$-cycles for swapping, leaving the remaining $c_k - 2j$ $k$-cycles to be fixed. Once the $j$ pairs have been specified, lemmas 2.1 and 2.4 show that there are $k^j \cdot k^{c_k - 2j}$ ways to factor the $k$-cycles. Hence, the total number of factorizations of $\sigma$ is

$$\prod_{k=1}^{n} \sum_{j=0}^{\lfloor c_k/2 \rfloor} \frac{k^{j} \cdot c_k}{2^j j!(c_k - 2j)!} = \prod_{k=1}^{n} k^{c_k} V_{c_k}(k).$$  

\[\square\]

### 3 Approximation by $B_n$

Let $T_n(\sigma)$ be the order of $\sigma$ as an element of the symmetric group, i.e. the least common multiple of the cycle lengths. The asymptotic distribution of $T_n$ was deduced from that of $B_n$. (See equation 14.4 of [10], section 7 of [6], and lemma 2 of [4].) A similar strategy is used in this paper. The goal of this section is to prove that $B_n$ can serve as proxy for $N_n$.

The following deterministic lemma supplies a sufficient condition on $\sigma$ that, when satisfied, imposes a bound on the error of the approximation.
**Lemma 3.1.** Suppose $\xi \geq 1$ and that, for every integer $k > \xi$, we have $c_k(\sigma) \leq 1$. Also assume that, for every positive integer $k$, $c_k(\sigma) \leq \xi$. Then there is a constant $c > 0$, not dependent on $\sigma$ nor $\xi$, such that $B_n(\sigma) \leq N_n(\sigma) \leq B_n(\sigma) \cdot (c\xi)^\xi$.

**Proof.** We already have the lower bound (See equation 2.1). Observe that $V_0(k) = 1$ and $V_1(k) = 1$ for all $k \in [n]$. For $2 \leq m < \xi$ and $1 \leq k \leq \xi$, a very crude bound for $V_m(k)$ suffices. For example, by Stirling’s formula we see that for $2 \leq m < \xi$,

$$V_m(k) \leq \frac{m!}{\left\lfloor \frac{m}{2} \right\rfloor \sum_{j=0}^{\left\lceil \frac{\xi}{2} \right\rceil} (-1)^j \frac{1}{j!} k^j} \leq \frac{m!}{e^k} < cm^m,$$

where $c$ is a positive constant independent of $k$ and $m$. By assumption $c_k(\sigma) \leq \xi$ for all $k \leq \xi$. Therefore

$$N_n(\sigma) \leq B_n(\sigma) \cdot \left( \prod_{1 \leq k \leq \xi} V_{c_k(\sigma)}(k) \right) \leq B_n(\sigma) \cdot (c\xi)^\xi.$$

\[\square\]

Clearly $B_n(\sigma)$ is not always a good approximation for $N_n(\sigma)$. For example, if $\sigma$ is the identity permutation with $n$ cycles of length one, then $\log B_n(\sigma) = 0$ and $\log N_n(\sigma) \sim \frac{n}{2} \log n$. There is a tradeoff when applying lemma 3.1. The parameter $\xi = \xi(n)$ must be sufficently large so that most permutations satisfy the hypotheses. However the larger $\xi$ is, the cruder the bound. The next two lemmas make this precise.

**Lemma 3.2.** If $\xi = \xi(n) \to \infty$ as $n \to \infty$, and if $\mathbb{P}_n$ is the uniform probability measure on $S_n$, then $\mathbb{P}_n(c_k \geq 2$ for some $k \geq \xi) = O\left(\frac{1}{\xi}\right)$.

**Proof.** For any choice of $\xi$, Boole’s inequality implies that

$$\mathbb{P}_n(c_k \geq 2$ for some $k \geq \xi) \leq \sum_{k \geq \xi} \mathbb{P}_n(c_k \geq 2) = \sum_{k=\left\lceil \frac{\xi}{2} \right\rceil}^{\left\lceil \frac{\xi}{2} \right\rceil} [1 - \mathbb{P}_n(c_k = 0) - \mathbb{P}_n(c_k = 1)].$$

(3.1)

It is well known that the probabilities $\mathbb{P}_n(c_k = j)$ can be calculated using the Principle of Inclusion Exclusion, and that the alternating inequalites yield upper and lower bounds. (See also chapter 5 of Sachkov [24] for the “generatingfunctionological” approach). Thus

$$\mathbb{P}_n(c_k = 0) = \sum_{j=0}^{\left\lceil \frac{\xi}{2} \right\rceil} (-1)^j \frac{1}{j!k^j} \geq 1 - \frac{1}{k}$$

(3.2)

and
\[ P_n(c_k = 1) = \frac{1}{k} \sum_{j=0}^{\left\lfloor \frac{n}{k} - 1 \right\rfloor} (-1)^j \frac{1}{j! k^j} \geq \frac{1}{k} \left( 1 - \frac{1}{k} \right). \tag{3.3} \]

Putting (3.2) and (3.3) into (3.1), we get

\[ P_n(c_k \geq 2 \text{ for some } k \geq \xi) \leq \sum_{k=\lfloor \xi \rfloor}^{\left\lfloor \frac{n}{2} \right\rfloor} \frac{1}{k^2} = O\left( \frac{1}{\xi} \right). \]

\[ \square \]

The second hypothesis is even more likely to hold.

**Lemma 3.3.** If \( \xi = \xi(n) \to \infty \), then \( P_n(c_k \geq \xi \text{ for some } k \leq \xi) = O((\xi \xi + \xi / n)). \)

**Proof.** Let \( Z_k, \xi \leq k \leq n \) be a sequence of independent Poisson\((1/k)\) random variables. By theorem 4 of [5], \( P_n(c_k \leq \xi \text{ for all } k \leq \xi) = Pr(Z_k \leq \xi \text{ for all } k \leq \xi) + O(\xi / n) \). Standard estimates using Markov’s inequality and moment generating functions shows that this probability is small:

\[ \Pr(Z_k \geq \xi) = \Pr(e^{Z_k} \geq e^{\xi}) \leq \frac{\text{E}(e^{Z_k})}{e^{\xi}} = \frac{e^{\frac{1}{k}(e-1)}}{e^{\xi}} < \frac{8}{e^\xi}. \]

Therefore

\[ \Pr(Z_k \leq \xi \text{ for all } k \leq \xi) \geq \left( 1 - \frac{8}{e^\xi} \right)^{\xi} = 1 - O\left( \frac{\xi}{e^\xi} \right). \]

\[ \square \]

## 4 The Asymptotic Lognormality of \( N \)

It is well known that \( B_n \) is asymptotically lognormal.

**Lemma 4.1.** (Erdős and Turán) For any real number \( x \),

\[ \lim_{n \to \infty} P_n(\log B_n(\sigma) \leq \mu_n + x\sigma_n) = \Phi(x) \]

where \( \mu_n = \frac{1}{2} \log^2 n, \sigma_n^2 = \frac{1}{3} \log^3 n, \) and \( \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \)

**Remark 4.2.** The first proof lemma 4.1 is in the work of Erdős and Turán [11]. Alternative proofs, as well as stronger and more general results have been proved using quite varied techniques. See, for example, [2], [3], [4], [9], [18].
Theorem 4.3. $\mathbb{P}_n(\log N_n(\sigma) \leq \mu_n + x\sigma_n) = \Phi(x) + o(1)$.

Proof. Because $N_n(\sigma) \geq B_n(\sigma)$ for all $\sigma \in S_n$, one direction is an immediate consequence of lemma 4.1.

$$\mathbb{P}_n(\log N_n \leq \mu_n + x\sigma_n) \leq \mathbb{P}_n(\log B_n \leq \mu_n + x\sigma_n) = \Phi(x) + o(1). \quad (4.1)$$

For the other direction, we use the continuity of $\Phi$ and the bound $N_n(\sigma) \leq (c\xi)^\xi B_n(\sigma)$ from Lemma 3.1, which, due to lemma 3.2 and lemma 3.3, holds with probability $1 - O(\frac{1}{\xi} + \frac{\xi}{n} + \frac{\xi}{e\xi})$.

In more detail, let $\epsilon > 0$ be a fixed but arbitrarily small positive number. We can choose $\delta > 0$ so that $|\Phi(x) - \Phi(a)| < \epsilon$ whenever $|x - a| < \delta$. If we choose $\xi = \sqrt{\log n}$, then we have $\log((c\xi)^\xi) = o(\sigma_n)$. Therefore we can choose $N_\epsilon$ so that, for all $n \geq N_\epsilon$, $\log((c\xi)^\xi) < \frac{\delta \sigma_n}{2}$.

But then

$$\mathbb{P}_n(\log N(\sigma) \leq \mu_n + x\sigma_n) \geq \mathbb{P}_n(\log B(\sigma) + \log((c\xi)^\xi) \leq \mu_n + x\sigma_n) \quad (4.2)$$

$$\geq \mathbb{P}_n\left(\log B(\sigma) + \frac{\delta \sigma_n}{2} \leq \mu_n + x\sigma_n\right) \quad (4.3)$$

$$= \mathbb{P}_n\left(\log B(\sigma) \leq \mu_n + \left(x - \frac{\delta}{2}\right)\sigma_n\right) \quad (4.4)$$

$$= \Phi(x - \frac{\delta}{2}) + o(1) > \Phi(x) - \epsilon + o(1) \quad (4.5)$$

Yet $\epsilon > 0$ was arbitrary, and so $\mathbb{P}_n(\log N(\sigma) \leq \mu_n + x\sigma_n) \geq \Phi(x) + o(1)$.

References


