Some remarks on the combinatorics of $\mathcal{IS}_n$

Svante Janson and Volodymyr Mazorchuk

Abstract

We describe the asymptotic behavior of the cardinalities of the finite symmetric inverse semigroup $\mathcal{IS}_n$ and its endomorphism semigroup. This is applied to show that $|\mathcal{IS}_n|/|\text{End}(\mathcal{IS}_n)|$ is asymptotically 0, solving a problem of Schein and Telegzhi. We also apply our results to compute the distributions of elements from $\mathcal{IS}_n$ with respect to certain combinatorial properties, and to compute the generating functions for $|\mathcal{IS}_n|$ and for the number of nilpotent elements in $\mathcal{IS}_n$.

1 Introduction

For $n \in \mathbb{N}$ let $\mathcal{IS}_n$ denote the symmetric inverse semigroup of all partial injections on $N_n = \{1, \ldots, n\}$. We refer the reader to [GM1, GM2, Li] for the details and standard notation. For $\alpha \in \mathcal{IS}_n$ we denote by $\text{dom}(\alpha)$ the domain of $\alpha$, by $\text{im}(\alpha)$ the range of $\alpha$, by $\text{rank}(\alpha) = |\text{dom}(\alpha)| = |\text{im}(\alpha)|$ the rank of $\alpha$, and by $\text{def}(\alpha) = n - \text{rank}(\alpha)$ the defect of $\alpha$. For $k = 0, 1, \ldots, n$ let $R_{n,k}$ denote the cardinality of the set $\{\alpha \in \mathcal{IS}_n : \text{rank}(\alpha) = k\}$. We immediately have

$$R_{n,k} = \binom{n}{k} \cdot k!, \quad |\mathcal{IS}_n| = \sum_{i=0}^{k} R_{n,k} = \sum_{i=0}^{k} \binom{n}{k} \cdot k!.$$

For elements from $\mathcal{IS}_n$ one can use their regular tableaux presentation

$$\alpha = \begin{pmatrix} i_1 & i_2 & \ldots & i_k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix},$$

where $\text{dom}(\alpha) = \{i_1, \ldots, i_k\}$ and $\text{im}(\alpha) = \{j_1, \ldots, j_k\}$. However, sometimes it is more convenient to use the so-called chain (or chart) decomposition of $\alpha$, which is analogous to the cyclic decomposition for usual permutations. We refer the reader to [Li] for rigorous definitions, however, this decomposition is very easy to explain on the following example. The element

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 7 & 9 \\ 7 & 4 & 5 & 1 & 10 & 2 & 6 \end{pmatrix} \in \mathcal{IS}_{10}$$

1
has the following graph of the action on \( \{1, 2, \ldots, 10\} \):

\[
\begin{array}{c c c c c c}
1 & \rightarrow & 7 \\
\uparrow & \downarrow & 3 & \rightarrow & 5 & \rightarrow & 10 & 9 & \rightarrow & 6 & 8,
\end{array}
\]

and hence it is convenient to write it as \( \alpha = (1, 7, 2, 4)[3, 5, 10][9, 6][8] \). We call \((1, 7, 2, 4)\) a cycle and \([3, 5, 10]\) (as well as \([9, 6]\) and \([8]\)) a chain of the element \( \alpha \).

We denote by \( L_n \) the total number of chains in the chain decompositions of all elements in \( |\mathcal{IS}_n| \). Each element of rank \( k \) has defect \( n-k \) and thus contains \( n-k \) chains implying \( L_n = \sum_{k=0}^{n} (n-k) R_{n,k} \). The semigroup \( \mathcal{IS}_n \) contains the zero element \( 0 \), uniquely characterized by the property \( \text{dom}(0) = \emptyset \). We denote by \( T_n \) the set of all nilpotent elements in \( \mathcal{IS}_n \), that is the set of all \( \alpha \in \mathcal{IS}_n \) satisfying \( \alpha^n = 0 \). We also denote by \( L^{(n)} \) the total number of chains in the chain decompositions of all elements in \( T_n \).

In [GM2] various combinatorial relations between \( |\mathcal{IS}_n| \), \( |T_n| \), \( L_n \) and \( L^{(n)} \) were obtained in a purely combinatorial way. The paper [GM2] contains also various estimates of distributions of elements from \( \mathcal{IS}_n \) with respect to certain algebraic properties. These distributions are obtained using several technical lemmas. The most of the technical difficulties in [GM2] arise from the fact that the authors did not have any reasonable asymptotic formula for \( |\mathcal{IS}_n| \) available. The aim of the present paper is to fill this gap. In Section 2 we derive an asymptotic formula for \( |\mathcal{IS}_n| \). In Section 4 we even show that analogous methods can be applied to derive an asymptotic formula for \( |\text{End}(\mathcal{IS}_n)| \). These formulae happen to be enough to show that \( |\mathcal{IS}_n|/|\text{End}(\mathcal{IS}_n)| \to 0 \) as \( n \to \infty \), which solves a problem from [ST]. Our results can be used to recover (in hopefully an easier way) several asymptotic statements from [GM2]. This is done in Section 3. Our results can be also used to obtain several new statements about the distributions of elements of \( \mathcal{IS}_n \) with respect to such combinatorial properties as the defect, the stable rank, the order etc. This is done in Section 5. Finally, in Section 6 we compute exponential generating functions for \( |\mathcal{IS}_n| \), \( |T_n| \), \( L_n \) and \( L^{(n)} \) and use them to recover various combinatorial results from [GM2].

## 2 An asymptotic for \( |\mathcal{IS}_n| \)

This section is devoted to the proof of the following

**Theorem 1.**

\[ |\mathcal{IS}_n| \sim \frac{1}{2\sqrt{\pi}e} n^{-1/4} e^{2\sqrt{n}n!} \sim \frac{1}{\sqrt{2e}} n^{2\sqrt{n}-n} n^{n+1/4}. \]

**Proof.** For \( R_{n,k} = \binom{n}{k}^2 \cdot k! \) we have the ratio \( \frac{R_{n,k+1}}{R_{n,k}} = \frac{n-k}{k+1} \). Moreover, for large \( n \) we obtain that \( \frac{R_{n,k+1}}{R_{n,k}} \approx 1 \) when \( k \approx n - \sqrt{n} \), hence \( \max_k R_{n,k} \) is achieved for such a \( k \). Note that \( \frac{R_{n,k+1}}{R_{n,k}} \) is decreasing with respect to \( k \). Write

\[ k = n - x\sqrt{n}, \quad 0 \leq x \leq \sqrt{n}. \]
Using the Stirling formula we have

\[
\ln \left( \frac{R_{n,k}}{n!} \right) = n \ln n - n + \frac{1}{2} \ln(2\pi n) - k \ln k + k - \frac{1}{2} \ln(2\pi k) - 2(n - k) \ln(n - k) + 2(n - k) - \ln(2\pi(n - k)) + O \left( \frac{1}{n} + \frac{1}{k} + \frac{1}{n - k} \right). \tag{2}
\]

Using the arguments above we have \( \frac{R_{n,k+1}}{R_{n,k}} < \frac{\left( \frac{1}{3} \sqrt{n} \right)^2}{n} < \frac{1}{2} \) for \( k > n - \frac{1}{2}\sqrt{n} \) and large \( n \). Thus, for \( k \geq k_1 = \lceil n - \frac{1}{2}\sqrt{n} \rceil \) we have \( R_{n,k} \leq 2^{-(k-k_1)} R_{n,k_1} \). In particular,

\[
\sum_{k \geq n - \frac{1}{2}\sqrt{n}} R_{n,k} \leq 2^{2-\frac{1}{2}\sqrt{n}} R_{n,k_1} = O \left( 2^{-\sqrt{n}/4} R_{n,k_1} \right).
\]

Similarly, for \( k \leq n - 2\sqrt{n} \) we have \( \frac{R_{n,k}}{R_{n,k+1}} < \frac{n}{(2\sqrt{n})} = \frac{1}{4} \), and

\[
\sum_{k \leq n - 5\sqrt{n}} R_{n,k} = O \left( 4^{-\sqrt{n}} R_{n,k_2} \right),
\]

where \( k_2 = \lceil n - 2\sqrt{n} \rceil \).

Hence, to estimate \( |\mathcal{I}_n| = \sum_{k=0}^{n} R_{n,k} \) we can ignore \( k \geq n - \frac{1}{2}\sqrt{n} \) and \( k \leq n - 3\sqrt{n} \). We may thus assume that \( \frac{1}{4} \leq x \leq 3 \). For such \( x \) we have:

\[
\ln \left( \frac{R_{n,k}}{n!} \right) = n \ln n - n - (n-x\sqrt{n}) \ln(n-x\sqrt{n}) + n-x\sqrt{n} - \frac{1}{2} \ln \frac{n-x\sqrt{n}}{n} - 2x\sqrt{n} \ln x - 2x\sqrt{n} \ln(\sqrt{n}) + 2x\sqrt{n} - \ln(2\pi x\sqrt{n}) + O(n^{-1/2}) =
\]

\[
= - (n-x\sqrt{n}) \ln \left( 1 - \frac{x}{\sqrt{n}} \right) - 2\sqrt{n} \ln x + x\sqrt{n} - \ln(2\pi x\sqrt{n}) + O(n^{-1/2}) =
\]

\[
x \sqrt{n} - x^2 + \frac{x^2}{2} + x\sqrt{n} - 2\sqrt{n} \ln x - \ln(2\pi x\sqrt{n}) + O(n^{-1/2}) =
\]

\[
2\sqrt{n}(x-x \ln x) - \frac{x^2}{2} - \ln x - \ln(2\pi \sqrt{n}) + O(n^{-1/2}),
\]

where all \( O \) are uniform in \( x \) and \( n \).

Denote \( f(x) = x - x \ln x \) and we have \( f'(x) = -\ln x \), \( f''(x) = -\frac{1}{x} \). Thus \( f(x) \) is concave on \( [0, +\infty) \) with a maximum at \( x_0 = 1 \). As \( f(x_0) = 1 \), we have the following Taylor expansion:

\[
f(x) = 1 - \frac{1}{2} (x-x_0)^2 + O(||x-x_0||^3), \quad 0 \leq x < \infty. \tag{3}
\]

For \( \frac{1}{4} \leq x \leq 3 \) we have \( f''(x) < -\frac{1}{3} \) and thus \( f(x) \leq 1 - \frac{1}{6}(x-x_0)^2 \).

3
Further, let \( g(x) = -\frac{x^2}{2} - \ln x \). Then for all \( \frac{1}{4} \leq x \leq 3 \) such that \( x\sqrt{n} \in \mathbb{Z} \) we have

\[
\frac{1}{n!} R_{n,n-x\sqrt{n}} = e^{2\sqrt{n} f(x) + g(x)} \cdot \frac{1 + O(n^{-1/2})}{2\pi \sqrt{n}}.
\]

(4)

Now we have:

\[
\frac{1}{n!} \sum_{k=0}^{n} R_{n,k} \sim \frac{1}{n!} \int_0^{n+1} \frac{1}{n!} R_{n,n-\lfloor t \rfloor} dt \sim \int_{\sqrt{n}/4}^{3\sqrt{n}} \frac{1}{n!} R_{n,n-\lfloor t \rfloor} dt = \left[ t = \sqrt{n} \hat{y} \right] =
\]

\[
= \sqrt{n} \int_{1/4}^{3} \frac{1}{n!} R_{n,n-\lfloor \sqrt{n} \hat{y} \rfloor} d\hat{y} = \sqrt{n} \int_{1/4}^{3} \frac{1 + O(n^{-1/2})}{2\pi \sqrt{n}} e^{2\sqrt{n} f(\hat{y}) + g(\hat{y})} d\hat{y} \sim \frac{e^{2\sqrt{n}}}{2\pi} \int_{1/4}^{3} e^{2\sqrt{n} f(\hat{y}) - 1 + g(\hat{y})} d\hat{y}.
\]

Write

\[
\int_{1/4}^{3} e^{2\sqrt{n} f(\hat{y}) - 1 + g(\hat{y})} d\hat{y} = \int_{I_1} e^{2\sqrt{n} f(\hat{y}) - 1 + g(\hat{y})} d\hat{y} + \int_{I_2} e^{2\sqrt{n} f(\hat{y}) - 1 + g(\hat{y})} d\hat{y},
\]

where \( I_1 = \{ y \in [1/4, 3] : |y - 1| \geq n^{-1/5} \} \) and \( I_2 = \{ y \in [1/4, 3] : |y - 1| \leq n^{-1/5} \} \) and denote these integrals by \( X_1 \) and \( X_2 \) respectively.

Since \( |\hat{y} \ - \ y| < n^{-1/2} \), for \( 1/4 \leq y \leq 3 \) we have

\[
2\sqrt{n} f(\hat{y}) - 1 + g(\hat{y}) \leq -2\sqrt{n} (\hat{y} - 1)^2 + O(1) = -\frac{\sqrt{n}}{3} (y - 1)^2 + O(1).
\]

Hence \( X_1 = O(e^{-n^{-1/10}/3}) \).

From (3) we also have, uniformly for \( y \in I_2 \), that

\[
2\sqrt{n} f(\hat{y}) - 1 = 2\sqrt{n} \left( -\frac{1}{2} (\hat{y} - 1)^2 + O(n^{-3/5}) \right) = -\sqrt{n} (\hat{y} - 1)^2 + O(n^{-1/10}) = -\sqrt{n} (y - 1)^2 + O(n^{-1/10}) = -\frac{\sqrt{n}}{3} (y - 1)^2 + O(1),
\]

and, similarly, \( g(\hat{y}) = g(1) + O(n^{-1/5}) = -1/2 + o(1) \).

Now we calculate again:

\[
\frac{1}{n!} \sum_{k=0}^{n} R_{n,k} \sim \frac{e^{2\sqrt{n}}}{2\pi} \left( X_1 + X_2 \right) \sim \frac{e^{2\sqrt{n}}}{2\pi} X_2 \sim \frac{e^{2\sqrt{n}}}{2\pi} \int_{1-n^{-1/5}}^{1+n^{-1/5}} e^{-\sqrt{n}(y-1)^2 - 1/2} dy \sim \frac{e^{2\sqrt{n}-1/2}}{2\pi} \int_{-\infty}^{+\infty} e^{-\sqrt{n}(y-1)^2} dy = \frac{e^{2\sqrt{n}-1/2}}{2\pi} \sqrt{\frac{\pi}{\sqrt{n}}} = \frac{1}{2\pi} \frac{1}{\sqrt{2}\sqrt{n}} e^{-1/2} e^{-1/2} n^{-1/5} e^{2\sqrt{n}}.
\]

Finally, using the Stirling formula again, we obtain

\[
|TS_n| = \sum_{k=0}^{n} R_{n,k} \sim \frac{1}{2\sqrt{\pi} e} n^{-1/4} e^{2\sqrt{n} n!} \sim \frac{1}{\sqrt{2} e} n^{n+1/4} e^{2\sqrt{n} n!},
\]

completing the proof. \( \square \)
3 Some applications of Theorem 1

An immediate corollary of Theorem 1 is the following statement, proved in [GM2, Theorem 8]:

Corollary 1.

\[
\frac{|\mathcal{IS}_{n+1}|}{|\mathcal{IS}_n|} \sim n, \quad n \to \infty.
\]

Another corollary is the following reinforcement of [GM2, Theorem 9]:

Corollary 2. \(|T_n| \sim \frac{1}{\sqrt{n}}|\mathcal{IS}_n|\), in particular,

\[
\frac{|T_n|}{|\mathcal{IS}_n|} \to 0, \quad n \to \infty.
\]

\textit{Proof.} From [GM2, Theorem 6] we know that \(|T_n| = \frac{1}{n}L_n\) (see a different proof in Section 6). By the definition, \(L_n = \sum_{k=0}^{n}(n-k)R_{n,k}\). An argument, analogous to that of Theorem 1, yields

\[
\frac{1}{n!}L_n \sim \int_{\sqrt{n}/4}^{3\sqrt{n}/4} |t| \frac{1}{n!}R_{n,n-|t|}dt.
\]

The same estimates as in Theorem 1 show that most of the integral comes from \(y = t/\sqrt{n} = 1 + O(n^{-1/5})\). Hence

\[
\frac{1}{n!}L_n \sim \sqrt{n} \int_{\sqrt{n}/4}^{3\sqrt{n}/4} \frac{1}{n!}R_{n,n-|t|}dt \sim \sqrt{n} \frac{1}{n!} |\mathcal{IS}_n|.
\]

This implies that \(L_n \sim \sqrt{n} |\mathcal{IS}_n|\) and completes the proof. \(\square\)

4 An asymptotic for \(|\text{End}(\mathcal{IS}_n)|\)

In [ST] it is shown that for \(n > 6\) the cardinality of the semigroup \(\text{End}(\mathcal{IS}_n)\) of all endomorphisms of the semigroup \(\mathcal{IS}_n\) equals

\[
|\text{End}(\mathcal{IS}_n)| = 3^n + 3 \cdot n! + n! \sum_{m=0}^{n} \sum_{k=0}^{[m/2]} \frac{2^{m-3k}}{(n-m)! \cdot (m-2k)! \cdot k!}.
\]

On [ST, Page 303] the following problem is formulated:

\textit{Find an asymptotic estimate for }\(|\text{End}(\mathcal{IS}_n)|\ \text{when } n \to \infty.\ \text{Is } |\text{End}(\mathcal{IS}_n)|/|\mathcal{IS}_n|\ \text{approaching 0?}\

In this section we answer both parts of this problem.

\textbf{Theorem 2.} \(|\text{End}(\mathcal{IS}_n)| \sim 3n!|\mathcal{IS}_n|\).
Proof. Set

\[ X_n = n! \sum_{m=0}^{n} \sum_{k=1}^{\lfloor m/2 \rfloor} \frac{2^{m-3k}}{(n-m)! \cdot (m-2k)! \cdot k!} \]

It would be enough to show that \( X_n/n! \to 0 \) as \( n \to \infty \). To do this we remark that \( X_n \)

equals the number of ways to perform the following procedure:

(i) choose \( X \subset N_n \);

(ii) choose \( Y \subset X \) such that \( |Y| = 2k > 0 \);

(iii) decompose \( Y = \cup Y_i, |Y_i| = 2, Y_i \cap Y_j = \emptyset \) for \( i \neq j \), the order of \( Y_i \) is not important;

(iv) Choose \( Z \subset X \setminus Y \).

Now let \( |X| = m, 0 \leq m \leq n \), and note that (i) can be done in \( \binom{n}{m} \) different ways, each of

(ii) and (iv) can be done in at most \( 2^m \) different ways, and, finally, (iii) can be done in at most \( m!! \) different ways. Hence

\[ X_n \leq \sum_{m=0}^{n} \binom{n}{m} \cdot 2^m \cdot m!! \leq \left( \sum_{m=0}^{n} \binom{n}{m} \cdot 4^m \right) (2\lfloor n/2 \rfloor)!! = 5^n 2^{\lfloor n/2 \rfloor} \lfloor n/2 \rfloor !. \]

To complete the proof it is enough to show that \( 5^n 2^{\lfloor n/2 \rfloor} \lfloor n/2 \rfloor !/n! \to 0 \) as \( n \to \infty \). Using the Stirling formula we have

\[ 5^n 2^{\lfloor n/2 \rfloor} \lfloor n/2 \rfloor ! \leq 5^n 2^{(n+1)/2} \lfloor n/2 \rfloor ! \approx \frac{1}{\sqrt{n}} e^{n \ln 5 \sqrt{\pi} - \frac{1}{2} \ln n + n \ln n - n - [n/2] + [n/2] \ln n - [n/2] - [n/2] - \frac{1}{2} \ln n + O(n)} , \]

and thus

\[ \frac{5^n 2^{\lfloor n/2 \rfloor} \lfloor n/2 \rfloor !}{n!} \approx \frac{1}{\sqrt{2}} e^{n \ln 5 \sqrt{\pi} - \frac{1}{2} \ln n + n \ln n - n - [n/2] + [n/2] \ln n - [n/2] - [n/2] - \frac{1}{2} \ln n + O(n)} . \]

Since the exponent is \( -\frac{1}{2} n \ln n + O(n) \), we obtain that the expression approaches 0 for large \( n \). This completes the proof.

Corollary 3.

\[ \frac{|\text{End}(\mathcal{S}_n)|}{|\mathcal{S}_n|} \to 0, \quad n \to \infty. \]

Proof. Follows immediately from the formulae of Theorem 1 and Theorem 2.

Using the methods, analogous to those of Theorem 1, one can even estimate the asymptotic for the “problematic” term \( X_n \) above.

Theorem 3.

\[ X_n \sim \frac{1}{\sqrt{2}} \cdot e^{\frac{1}{2} n \ln n - \frac{1}{2} n + 3 \sqrt{n} \ln n - \frac{9}{8}}. \]
Proof. We can write
\[
n! \sum_{m=0}^{n} \sum_{k=1}^{[m/2]} \frac{2^{m-3k}}{(n-m)! \cdot (m-2k)! \cdot k!} = n! \sum_{m=0}^{n} \frac{2^{m}}{(n-m)!} \sum_{k=1}^{[m/2]} \frac{2^{-3k}}{(m-2k)! \cdot k!},
\]
and we denote \(a_k = \frac{2^{-3k}}{(m-2k)! \cdot k!}\), \(b_m = \sum_{k=1}^{[m/2]} a_k\), \(c_m = \frac{2^{m}}{(n-m)!} b_k\). Remark that \(\frac{a_{k+1}}{a_k} = \frac{2^{-3}}{k+1}(m-2k)(m-2k-1)\) decreases on \([0, m/2]\) and \(a_k\) has on \([0, m/2]\) a unique maximum at \(\approx \frac{m}{2} - \sqrt{m}\). Let \(k = \frac{m}{2} - x\sqrt{m}\), that is \(m - 2k = 2x\sqrt{m}\), where \(\varepsilon \leq x \leq C\). Then we have
\[\frac{a_{k+1}}{a_k} = \frac{1}{8} \cdot \frac{1}{m/2} \cdot 4x^2 m (1 + o(1)) = x^2 + o(1)\]
This implies that \(\sum_{x<1/2} a_k\) and \(\sum_{x>2} a_k\) belong to \(O(e^{-c\sqrt{m} a_{m/2-\sqrt{m}}})\), that is relatively very small and can be neglected. Assume now that \(1/2 \leq x \leq 2\). Taking into account that
\[\ln k = \ln \frac{m}{2} + \ln \left(1 - \frac{2x}{\sqrt{m}}\right) = \ln \frac{m}{2} - \frac{2x}{\sqrt{m}} - \frac{2x^2}{m} + O(m^{-3/2})\]
and using the Stirling formula we obtain the following:
\[
\ln a_k = -3k \ln 2 - \ln(2x\sqrt{m})! - \ln k! = -\frac{3}{2} \ln 2m + 3 \ln 2\sqrt{m}x - 2x\sqrt{m} \ln 2 - \frac{2x\sqrt{m} \ln x - x^2 \ln m + 2x\sqrt{m} - \ln(2\pi) - \frac{1}{2} \ln(2x\sqrt{m}) - k \ln k + k - \frac{1}{2} \ln k + o(1)}{m/2} = \ln \frac{m}{2} + \frac{2kx}{\sqrt{m}} + \frac{mx^2}{m} - \frac{m}{2} - x\sqrt{m} - \frac{1}{2} \ln \frac{m}{2} + o(1) = \ln 2 + 2x\sqrt{m} - 2x\sqrt{m} \ln x - \frac{1}{2} \ln(4\pi^2 x m^{3/2}) - \frac{1}{2} m \ln m - x^2 + \frac{m}{2} + o(1).
\]
Further, assuming \(x = 1 + m^{-1/4} y\) yields \(x \ln x = x = 1 + \frac{x}{2} (1 - x) + O ((x - 1)^3) = -1 + \frac{y^2}{2\sqrt{m}} + O \left(\frac{y^2}{m^{1/2}}\right)\) and thus
\[
\ln a_k = -\frac{1}{2} m \ln m + m (\frac{1}{2} - \ln 2) - 1 + O (ym^{-1/4}) + 2\sqrt{m} - y^2 + O (y^3 m^{-1/4}) - \frac{3}{4} \ln m \ln (2\pi) \leq 0.
\]
Therefore \(k = \frac{m}{2} - \sqrt{m} = m^{1/4} y\) yields
\[
a_k = \frac{1}{2\pi} e^{(\frac{1}{4} \ln 2) m - \frac{1}{2} \ln m - \frac{1}{4} \ln m + 2\sqrt{m} - e^{-y^2} \left(1 + O \left(\frac{y + y^3}{m^{1/4}}\right)\right)}.
\]
We can assume that, say, $y = O(m^{1/12})$ and ignore larger $y$. In this way we obtain
\[
b_m = \sum_{k=1}^{|m/2|} a_k = \frac{1}{2\pi} e^{\left(\frac{1}{4}\ln 2\right) m - \frac{1}{2} \ln m} - \frac{\sqrt{m}}{2\pi} \ln m + \frac{2\sqrt{m} - 1}{2m^{1/4}} \int_{-\infty}^{\infty} e^{-y^2} dy \left(1 + o(1)\right) \sim \frac{1}{2\sqrt{\pi}} e^{\left(\frac{1}{4}\ln 2\right) m - \frac{1}{2} \ln m} - \frac{1}{2\pi} \ln m + \frac{2\sqrt{m} - 1}{2m^{1/4}}.
\]

The latter implies
\[
\ln b_m = \left(\frac{1}{2} - \ln 2\right) m - \frac{1}{2} \ln m - \frac{1}{2} \ln m + 2\sqrt{m} - 1 - \ln(2\sqrt{\pi}) + o(1) \quad (5)
\]
and also
\[
\ln c_m = \ln b_m + m \ln 2 - \ln((n - m)!).
\]

Further, for $m \to \infty$ we compute $\ln \frac{c_{m+1}}{c_m} = \frac{1}{2} - \ln 2 - \frac{1}{2} \ln m - \frac{1}{2} + o(1) = -\ln 2 - \frac{1}{2} \ln m + o(1)$ and also $\ln \frac{c_{m+1}}{c_m} = -\frac{1}{2} \ln m + \ln(n - m) + o(1)$. This gives us that for large $n$ the value of $c_m$ is largest when $\frac{1}{2} \ln m \approx \ln(n - m)$ that is $m \approx n - \sqrt{n}$. In particular, it follows easily that $m \leq n/2$ can be ignored and thus we obtain that $o(1), m \to \infty$, is small even if $n \to \infty$.

Let us now show that even $m < n - 3\sqrt{n}$ can be ignored. If $m < n - 2\sqrt{n}$ then we have $-\frac{1}{2} \ln m + \ln(n - m) > -\frac{1}{2} \ln n + \ln(2\sqrt{n}) = \ln 2$ and thus for large $n$ we derive $\ln \frac{c_{m+1}}{c_m} > 1/2$ and hence $\frac{c_{m+1}}{c_m} > e^{1/2}$. Set $M = [n - 2\sqrt{n}]$. Then $\frac{c_m}{c_M} < e^{-|M-m|/2}$ and thus
\[
\sum_{m < n - 3\sqrt{n}} c_m < e^{-\sqrt{n}/2} \frac{1}{1 - e^{-1/2}} c_M.
\]

The latter implies that all terms with $m < n - 3\sqrt{n}$ can be ignored. Similarly, all terms with $m > n - \sqrt{n}/2$ can be ignored.

Thus we can assume $m = n - x\sqrt{n}$, where $1/2 \leq x \leq 3$. Under such assumption we have $\ln \frac{c_{m+1}}{c_m} = -\frac{1}{2} \ln n + \ln(x\sqrt{n}) + o(1) = \ln x + o(1)$.

For $1/2 \leq x \leq 3$ we have, using the Stirling formula, that
\[
\ln m = \ln n + \ln \left(1 - \frac{x}{\sqrt{n}}\right) = \ln n - \frac{x}{\sqrt{n}} - \frac{x^2}{2n} + O(n^{-3/2}),
\]
\[
m \ln m = n \ln n - x\sqrt{n}\ln n - x\sqrt{n} + x^2 - \frac{x^2}{2} + O(n^{-1/2}),
\]
\[
\ln(\sqrt{n}) = \ln(x\sqrt{n}) = x\sqrt{n}\ln n + \frac{1}{2} x\sqrt{n}\ln n - x\sqrt{n} + \frac{1}{2} \ln x + \frac{1}{4} \ln n + \ln \sqrt{2\pi} + o(1),
\]
\[
\sqrt{m} = \sqrt{n} \left(1 - x/\sqrt{n}\right)^{1/2} = \sqrt{n} - x/2 + o(1).
\]
Hence, using (5) and (6), we obtain
\[
\ln c_n = \frac{1}{2} n - \frac{1}{2} x \sqrt{n} - \frac{1}{2} n \ln n + \frac{1}{2} x \sqrt{n} \ln n + \frac{1}{2} x \sqrt{n} - \frac{x^2}{4} - \frac{1}{2} \ln n + 2 \sqrt{n} - x - 1 - \ln(2 \sqrt{\pi}) - x \sqrt{n} \ln x - \frac{1}{2} x \sqrt{n} \ln n + x \sqrt{n} - \frac{1}{2} \ln n - \frac{1}{4} \ln n - \ln \sqrt{2 \pi} + o(1) = \\
= \frac{1}{2} n - \frac{1}{2} n \ln n - \frac{3}{4} \ln n + 2 \sqrt{n} - 1 - \ln(2^{3/2} \pi) + \sqrt{n} - x \ln x - \frac{x^2}{4} - x - \frac{1}{2} \ln x + o(1).
\]

Setting \( x = 1 + y n^{-1/4} \) yields
\[
\ln c_n = \frac{1}{2} n - \frac{1}{2} n \ln n - \frac{3}{4} \ln n + 3 \sqrt{n} - 1 - \ln(2^{3/2} \pi) - \frac{5}{4} - \frac{y^2}{2} + O \left( \frac{y^3}{n^{1/4}} \right) + o(1)
\]
and thus
\[
\sum_{m=0}^{n} c_m = \exp \left( \frac{1}{2} n - \frac{1}{2} n \ln n - \frac{3}{4} \ln n + 3 \sqrt{n} - \frac{9}{4} - \ln(2^{3/2} \pi) \right) n^{1/4} \times \\
\times \int_{-\infty}^{+\infty} e^{-y^2/2} dy (1 + o(1)) = \frac{1}{2 \sqrt{\pi}} e^{\frac{1}{2} n \ln n - \frac{1}{2} \ln n + 3 \sqrt{n} - \frac{9}{4} + o(1)}.
\]
This implies that
\[
X_n \sim \frac{n^{1/2}}{2 \sqrt{\pi \pi}} e^{-\frac{1}{2} n \ln n - \frac{1}{2} \ln n + 3 \sqrt{n} - \frac{9}{4}} \sim \frac{1}{\sqrt{2}} e^{\frac{1}{2} n \ln n - \frac{1}{2} n + 3 \sqrt{n} - \frac{9}{4}},
\]
and completes the proof. \( \square \)

5 Some distributions

Denote by \( D_n \) the defect of a random element of \( I \mathcal{S}_n \), by \( X_n \) the stable rank of a random element of \( I \mathcal{S}_n \), by \( C_n \) the number of cycles of a random element of \( I \mathcal{S}_n \), and by \( K_n = C_n + D_n \) the total number of components (i.e. cycles and chains) of a random element of \( I \mathcal{S}_n \).

**Proposition 1.** If \( n \to \infty \) and \( \frac{k - \sqrt{n}}{n^{1/4}} \to z \) with \( -\infty < z < \infty \), then
\[
P(D_n = k) \sim \frac{1}{\sqrt{\pi n^{1/4}}} e^{-z^2}.
\]
In particular,
\[
\frac{D_n - \sqrt{n}}{n^{1/4}} \xrightarrow{d} N(0, 1/2).
\]

**Proof.** We have \( P(D_n = k) = \frac{R_{n,n-k}}{|I \mathcal{S}_n|} \) by definition and \( \frac{R_{n,n-k}}{|I \mathcal{S}_n|} \sim \frac{1}{\sqrt{\pi n^{1/4}}} e^{-z^2} \), follows from (3) and (4). \( \square \)
Proposition 2. 

\[ P(X_n = k) \sim \frac{1}{\sqrt{n}} e^{-k/\sqrt{n}} \text{ if } k = o(n^{3/4}), \]

in particular,

\[ \frac{X_n}{\sqrt{n}} \xrightarrow{d} \text{exp}(1). \]

Proof. We have

\[ P(X_n = k) = \binom{n}{k} \cdot k! \cdot \frac{|T_{n-k}|}{|\mathcal{S}_n|} = \frac{|T_{n-k}|/(n-k)!}{|\mathcal{S}_n|/n!}. \]

Hence, if \( k = o(n) \) we have, using Section 2 and Section 3, that

\[ P(X_n = k) \sim \frac{(n-k)^{-3/4} e^{2\sqrt{n-k}}}{n^{1/4} e^{2\sqrt{n}}} \sim \frac{1}{\sqrt{n}} e^{2\sqrt{n-k} - \sqrt{n}} = \frac{1}{\sqrt{n}} e^{2\sqrt{n(1-k/n)^{1/2}} - 1} = \frac{1}{\sqrt{n}} e^{-2\sqrt{n} \frac{n}{n^{3/4}} + O(k^2/n^{3/2})} \]

and the statement follows. \( \square \)

Proposition 3. 

\[ \frac{C_n - \frac{1}{2} \ln n}{\sqrt{\frac{1}{2} \ln n}} \xrightarrow{d} N(0,1). \]

Proof. Given \( X_n \), the number of cycles for the permutational part of size \( X_n \) is approximately \( \ln X_n \). More precisely, by [Go], we have

\[ \frac{C_n - \ln X_n}{\sqrt{\ln X_n}} \xrightarrow{d} N(0,1). \]

Further, we have \( \ln X_n = \frac{1}{2} \ln n + \ln \frac{X_n}{\sqrt{n}} \) and \( \ln \frac{X_n}{\sqrt{n}} \xrightarrow{d} \ln \text{exp}(1) \) by Proposition 2. Hence, in

\[ \frac{C_n - \frac{1}{2} \ln n}{\sqrt{\frac{1}{2} \ln n}} = \frac{\sqrt{\ln X_n}}{\sqrt{\frac{1}{2} \ln n}} \cdot \frac{C_n - \ln X_n}{\sqrt{\ln X_n}} + \frac{\ln X_n - \frac{1}{2} \ln n}{\sqrt{\frac{1}{2} \ln n}} \]

we have \( \frac{\sqrt{\ln X_n}}{\sqrt{\frac{1}{2} \ln n}} \xrightarrow{p} 1 \) and \( \frac{\ln X_n - \frac{1}{2} \ln n}{\sqrt{\frac{1}{2} \ln n}} \xrightarrow{p} 0 \), completing the proof. \( \square \)

More precisely, we can show that \( C_n \) is almost Poisson distributed. Let \( d_{TV} \) denote the total variation distance between two distributions, see e.g. [BHJ].

Proposition 4. 

\[ d_{TV} \left( C_n, \text{Po} \left( \frac{1}{2} \ln n \right) \right) \to 0, \quad n \to \infty. \]
Proof. Let $h_k = \sum_{i=1}^{k} 1/i = \ln k + O(1)$. Given $X_n = k$, the number of cycles is distributed as the number of cycles in a random permutation of length $k$. Using [BH], we obtain

$$d_{TV} \left( \mathcal{L}(C_n|X_n = k), \text{Po}(h_k) \right) \leq \frac{c}{h_k} \leq \frac{c}{\ln k}$$

for some constant $c \leq \pi^2/6$. Further, by [BHJ, Remark 1.1.4], we have

$$d_{TV} \left( \text{Po}(h_k), \text{Po}(\ln \sqrt{n}) \right) \leq \frac{|h_k - \ln \sqrt{n}|}{\sqrt{\ln \sqrt{n}}} \leq \frac{|\ln k - \ln \sqrt{n}| + 1}{\sqrt{\ln \sqrt{n}}}.$$

Consequently,

$$d_{TV} \left( \mathcal{L}(C_n|X_n = k), \text{Po}(\ln \sqrt{n}) \right) \leq f(k) := \frac{\pi^2}{6 \ln k} + \frac{|\ln k - \ln \sqrt{n}| + 1}{\sqrt{\ln \sqrt{n}}}.$$

Since also $d_{TV} \leq 1$, we obtain $d_{TV}(C_n, \text{Po}(\ln \sqrt{n})) \leq E(f(X_n) \wedge 1)$. From the proof of Proposition 3 it follows that $f(X_n) \xrightarrow{p} 0$ and thus $E(f(X_n) \wedge 1) \xrightarrow{p} 0$, completing the proof.

**Corollary 4.**

$$\frac{K_n - \sqrt{n}}{n^{1/4}} \xrightarrow{d} N(0, 1/2).$$

**Proof.** Follows from Propositions 1 and 3.

Recall that for $\sigma \in \mathcal{S}_n$ the order $O(\sigma)$ of $\sigma$ is defined as the cardinality of the monoid, generated by $\sigma$, and the inverse order $\text{IO}(\sigma)$ of $\sigma$ is defined as the cardinality of the inverse monoid, generated by $\sigma$, that is

$$O(\sigma) = |\{\sigma^l : l \in \{0, 1, 2, \ldots\}\}|, \quad \text{IO}(\sigma) = |\{\sigma^l : l \in \mathbb{Z}\}|.$$

Let $O_n$ and $I_n$ denote the order and the inverse order of a random element of $\mathcal{S}_n$ respectively.

**Proposition 5.**

$$\frac{\ln O_n - \frac{1}{8} \ln^2 n}{\sqrt{\frac{1}{2\pi} \ln^3 n}} \xrightarrow{d} N(0, 1), \quad \frac{\ln I_n - \frac{1}{8} \ln^2 n}{\sqrt{\frac{1}{2\pi} \ln^3 n}} \xrightarrow{d} N(0, 1).$$

**Proof.** For $\sigma \in \mathcal{S}_n$ denote $X(\sigma) = \{i \in \{1, \ldots, n\} : \sigma^l(i) = i \text{ for some } l > 0\}$. Then $|X(\sigma)|$ is the stable rank of $\sigma$. Moreover, any $\sigma \in \mathcal{S}_n$ can be written as a product $\sigma = \sigma_1 \cdot \sigma_2$, where $\text{dom}(\sigma_1) = \{1, \ldots, n\}$ and $\sigma_1(i) = \sigma(i), i \in X(\sigma), \sigma_1(i) = i, i \notin X(\sigma)$; $\text{dom}(\sigma_2) = \text{dom}(\sigma)$ and $\sigma_2(i) = i, i \in X(\sigma), \sigma_2(i) = \sigma(i), i \in \text{dom}(\sigma) \setminus X(\sigma)$. It follows immediately from the definition that $\sigma_1 \cdot \sigma_2 = \sigma_2 \cdot \sigma_1$. It is further easy to see (see e.g. [GK]) that

$$O(\sigma_1) \leq O(\sigma) \leq O(\sigma_1) + n - |X(\sigma)|, \quad O(\sigma_1) \leq \text{IO}(\sigma) \leq O(\sigma_1) + 2(n - |X(\sigma)|). \quad (7)$$
For a random element $\sigma \in \mathcal{IS}_n$, let $O'_n(\sigma) = O(\sigma)$. Given $X_n = X(\sigma) = k$, this has the same distribution as the order $O_k$ of a random permutation of length $k$. In [ET] it was shown that, as $k \to \infty$,

$$\frac{\ln \tilde{O}_k - \frac{1}{3} \ln^2 k}{\sqrt{\frac{1}{3} \ln^3 k}} \to N(0, 1).$$

Hence, as $n \to \infty$,

$$\frac{\ln O'_n - \frac{1}{3} \ln^2 X_n}{\sqrt{\frac{1}{3} \ln^3 X_n}} \to N(0, 1),$$

and it follows as in the proof of Proposition 3 that

$$\frac{\ln O'_n - \frac{1}{3} \ln^2 n}{\sqrt{\frac{1}{3} \ln^3 n}} \to N(0, 1). \quad (8)$$

In particular, for almost all $\sigma \in \mathcal{IS}_n$ we have that $O(\sigma) \approx n^{(\ln n)/8}$. Since the difference between the left and the right sides of the inequalities in (7) is less than $2n$, in particular is $o(n^{(\ln n)/8})$, we obtain that, asymptotically, the left and the right sides of inequalities in (7) are the same. Now the necessary statement follows from (8).

\[\square\]

6 Some generating functions

Consider some “objects” consisting of “components”, whose order in the objects is not important. Assume that there are $a_m$ possible components containing exactly $m$ elements. Let $f_n$ denote the total number of objects, which consist of exactly $n$ elements. The following well-known statement can be easily derived for example from [Wi, Theorem 3.4.1]

**Proposition 6.** The exponential generating function for $\{f_n, n \geq 0\}$ is $F(z) = e^{A(z)}$, where $A(z) = \sum_{m=1}^{\infty} \frac{a_m}{m!} z^m$.

Proposition 6 now can be used to compute the exponential generating functions for $|T_n|$, $|\mathcal{IS}_n|$.

**Theorem 4.** 1. The exponential generating function for $a_n = |T_n|$ is $E_{T_n}(z) = e^{z/(1-z)}$.

2. The exponential generating function for $b_n = |\mathcal{IS}_n|$ is $E_{\mathcal{IS}_n}(z) = \frac{1}{1-z} e^{z/(1-z)}$.

**Proof.** For $T_n$ we have that components are chains and $a_m = m!$. Hence $A(z) = \sum_{m \geq 1} \frac{m!}{m!} z^m = z/(1-z)$ and we get $F(z) = e^{z/(1-z)}$.

For $\mathcal{IS}_n$ we have two types of components: cycles and chains, and thus $a_m = m! + (m-1)!$. This gives $A(z) = \frac{1}{1-z} - \ln(1-z)$ and therefore $F(z) = \frac{1}{1-z} e^{z/(1-z)}$. \[\square\]

Analogous arguments can be used to compute the exponential generating function for $L^{(n)}$ and $L_n$: 

12
Theorem 5.  1. The exponential generating function for the sequence $c_n = |L^n|$ is $E_{L^n}(z) = \frac{z}{1-z}e^{z/(1-z)}$.

2. The exponential generating function for $d_n = |L_n|$ is $E_{L_n}(z) = \frac{z}{1-z}e^{z/(1-z)}$.

Proof. A fixed chain of length $m$ is contained in exactly $|T_{n-m}|$ elements of $T_n$, and in exactly $|IS_{n-m}|$ elements of $IS_n$. This implies that $E_{L^n}(z) = \frac{z}{1-z}E_{T_n}(z) = \frac{z}{1-z}e^{z/(1-z)}$ and $E_{L_n}(z) = \frac{z}{1-z}E_{IS_n}(z) = \frac{z}{1-z}e^{z/(1-z)}$. □

Theorem 4 and the first part of Theorem 5 can now be used to derive the following corollaries:

Corollary 5. ([GM2, Theorem 7(2)]) $|IS_n| = |T_n| + L^n$.

Proof. Follows from $E_{IS_n}(z) = E_{T_n}(z) + E_{L^n}(z)$. □

Corollary 6. ([GM2, Theorem 6(1)]) $|T_n| = \frac{1}{n}L_n$.

Proof. For the sequence $n|T_n|$ we have

$$E_{n|T_n|}(z) = zE'_{|T_n|}(z) = \frac{z}{(1-z)^2}e^{z/(1-z)} = E_{L^n}(z)$$

and the statement follows. □

Corollary 7. ([GM2, Theorem 6(2)]) $|IS_n| = \frac{1}{n+1}L^{n+1}$.

Proof. The statement is equivalent to $zE_{IS_n}(z) = E_{L^n}(z)$, which is straightforward. □

We also obtain one relation, which seems to be missing in [GM2].

Corollary 8. The total number $P_n$ of fixed points for all elements from $IS_n$ satisfies $P_n = L^n$.

Proof. For $x \in \{1, \ldots, n\}$, the point $x$ is fixed in exactly $|IS_{n-1}|$ elements of $IS_n$, which implies that $E_{P_n}(z) = zE_{IS_n}(z)$. Further $zE_{IS_n}(z) = E_{L_n}(z)$ by Corollary 7 and the statement follows. □

Acknowledgments

For the second author the research was partially supported by The Swedish Research Council and The Swedish Institute.
References


S.J.: Department of Mathematics, Uppsala University, Box 480, SE 751 06, Uppsala, SWEDEN, e-mail: svante@math.uu.se, web: “http://www.math.uu.se/~svante”

V.M.: Department of Mathematics, Uppsala University, Box 480, SE 751 06, Uppsala, SWEDEN, e-mail: mazor@math.uu.se, web: “http://www.math.uu.se/~mazor”

14