# 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 $|\mathcal{IS}_n|/|\mathrm{End}(\mathcal{IS}_n)|$ is asymptotically 0, solving a problem of Schein and Teclezghi. 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 $N_n = \{1, \dots, n\}$. We refer the reader to [GM1, GM2, Li] for the details and standard notation. For $\alpha \in \mathcal{IS}_n$ we denote by $\mathrm{dom}(\alpha)$ the *domain* of $\alpha$, by $\mathrm{im}(\alpha)$ the *range* of $\alpha$, by $\mathrm{rank}(\alpha) = |\,\mathrm{dom}(\alpha)| = |\,\mathrm{im}(\alpha)|$ the *rank* of $\alpha$, and by $\mathrm{def}(\alpha) = n - \mathrm{rank}(\alpha)$ the *defect* of $\alpha$. For $k = 0, 1, \dots, n$ let $R_{n,k}$ denote the cardinality of the set $\{\alpha \in \mathcal{IS}_n \,:\, \mathrm{rank}(\alpha) = k\}$. We immediately have

$$R_{n,k} = \binom{n}{k}^2 \cdot k!, \qquad |\mathcal{IS}_n| = \sum_{i=0}^{k} R_{n,k} = \sum_{i=0}^{k} \binom{n}{k}^2 \cdot k!.$$

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

$$\alpha = \left( \begin{array}{cccc} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{array} \right),$$

where $\mathrm{dom}(\alpha) = \{i_1, \dots, i_k\}$ and $\mathrm{im}(\alpha) = \{j_1, \dots, j_k\}$. 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

$$\alpha = \left( \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 7 & 9 \\ 7 & 4 & 5 & 1 & 10 & 2 & 6 \end{array} \right) \in \mathcal{IS}_{10}$$

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

$$\begin{array}{ccc} 1 & \rightarrow & 7 \\ \uparrow & & \downarrow \\ 4 & \leftarrow & 2 \end{array} \qquad 3 \rightarrow 5 \rightarrow 10 \qquad 9 \rightarrow 6 \qquad 8,$$

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 $\mathrm{dom}(0) = \varnothing$. 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 $|\mathrm{End}(\mathcal{IS}_n)|$. These formulae happen to be enough to show that $|\mathcal{IS}_n|/|\mathrm{End}(\mathcal{IS}_n)| \to 0$, $n \to \infty$, 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.**
$$|\mathcal{IS}_n| \sim \frac{1}{2\sqrt{\pi e}} n^{-1/4} e^{2\sqrt{n}} n! \sim \frac{1}{\sqrt{2e}} \cdot e^{2\sqrt{n}-n} n^{n+1/4}.$$

*Proof.* For $R_{n,k} = \binom{n}{k}^2 \cdot k! = \frac{n!^2}{(n-k)!^2 k!}$ we have the ratio $\frac{R_{n,k+1}}{R_{n,k}} = \frac{(n-k)^2}{k+1}$. 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

$$k = n - x\sqrt{n}, \qquad 0 \leq x \leq \sqrt{n}. \tag{1}$$

Using the Stirling formula we have

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

Using the arguments above we have $\frac{R_{n,k+1}}{R_{n,k}} < \frac{\left(\frac{1}{2}\sqrt{n}\right)^2}{n-\frac{1}{2}\sqrt{n}} < \frac{1}{2}$ 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,

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

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

$$\sum_{k\leq n-3\sqrt{n}} R_{n,k} = O\left(4^{-\sqrt{n}} R_{n,k_2}\right),$$

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

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

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

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:

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

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

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

$$\frac{1}{n!} R_{n,n-x\sqrt{n}} = e^{2\sqrt{n}f(x)+g(x)} \cdot \frac{1+O(n^{-1/2})}{2\pi\sqrt{n}}. \tag{4}$$

Now we have:

$$\frac{1}{n!}\sum_{k=0}^{n} R_{n,k} = \int_0^{n+1} \frac{1}{n!} R_{n,n-\lfloor t\rfloor} dt \sim \int_{\sqrt{n}/4}^{3\sqrt{n}} \frac{1}{n!} R_{n,n-\lfloor t\rfloor} dt = \left[t = \sqrt{n}y\right] =$$
$$= \sqrt{n}\int_{1/4}^{3} \frac{1}{n!} R_{n,n-\lfloor y\sqrt{n}\rfloor} dy = \left[\tilde{y} = \frac{\lfloor y\sqrt{n}\rfloor}{\sqrt{n}}\right] = \sqrt{n}\int_{1/4}^{3} \frac{1+O(n^{-1/2})}{2\pi\sqrt{n}} e^{2\sqrt{n}f(\tilde{y})+g(\tilde{y})} dy \sim$$
$$\sim \frac{e^{2\sqrt{n}}}{2\pi}\int_{1/4}^{3} e^{2\sqrt{n}(f(\tilde{y})-1)+g(\tilde{y})} dy.$$

Write

$$\int_{1/4}^{3} e^{2\sqrt{n}(f(\tilde{y})-1)+g(\tilde{y})} dy = \int_{I_1} e^{2\sqrt{n}(f(\tilde{y})-1)+g(\tilde{y})} dy + \int_{I_2} e^{2\sqrt{n}(f(\tilde{y})-1)+g(\tilde{y})} dy,$$

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

Since $|\tilde{y} - y| < n^{-1/2}$, for $1/4 \le y \le 3$ we have

$$2\sqrt{n}(f(\tilde{y})-1) + g(\tilde{y}) \le -2\sqrt{n}\frac{(\tilde{y}-1)^2}{6} + O(1) = -\frac{\sqrt{n}}{3}(y-1)^2 + O(1).$$

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

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

$$2\sqrt{n}(f(\tilde{y})-1) = 2\sqrt{n}\left(-\frac{1}{2}(\tilde{y}-1)^2 + O(n^{-3/5})\right) = -\sqrt{n}(\tilde{y}-1)^2 + O(n^{-1/10}) =$$
$$= -\sqrt{n}(y-1)^2 + O(n^{1/2-1/5}|\tilde{y}-y| + n^{-1/10}) = -\sqrt{n}(y-1)^2 + o(1),$$

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

Now we calculate again:

$$\frac{1}{n!}\sum_{k=0}^{n} R_{n,k} \sim \frac{e^{2\sqrt{n}}}{2\pi}(X_1+X_2) \sim \frac{e^{2\sqrt{n}}}{2\pi}X_2 \sim \frac{e^{2\sqrt{n}}}{2\pi}\int_{1-n^{-1/5}}^{1+n^{-1/5}} e^{-\sqrt{n}(y-1)^2-1/2} dy \sim$$
$$\sim \frac{e^{2\sqrt{n}-1/2}}{2\pi}\int_{-\infty}^{+\infty} e^{-\sqrt{n}(y-1)^2} dy = \frac{e^{2\sqrt{n}-1/2}}{2\pi}\sqrt{\frac{\pi}{\sqrt{n}}} = \frac{1}{2}\pi^{-1/2}e^{-1/2}n^{-1/4}e^{2\sqrt{n}}.$$

Finally, using the Stirling formula again, we obtain

$$|\mathcal{IS}_n| = \sum_{k=0}^{n} R_{n,k} \sim \frac{1}{2\sqrt{\pi e}} n^{-1/4} e^{2\sqrt{n}} n! \sim \frac{1}{\sqrt{2e}} n^{n+1/4} e^{2\sqrt{n}-n},$$

completing the proof. □

## 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,*
$$\frac{|T_n|}{|\mathcal{IS}_n|} \to 0, \quad n \to \infty.$$

*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
$$\frac{1}{n!}L_n \sim \int_{\sqrt{n}/4}^{3\sqrt{n}} \lfloor t \rfloor \frac{1}{n!}R_{n,n-\lfloor t \rfloor}\,dt.$$
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
$$\frac{1}{n!}L_n \sim \sqrt{n}\int_{\sqrt{n}/4}^{3\sqrt{n}} \frac{1}{n!}R_{n,n-\lfloor t \rfloor}\,dt \sim \sqrt{n}\frac{1}{n!}|\mathcal{IS}_n|.$$
This implies that $L_n \sim \sqrt{n}|\mathcal{IS}_n|$ and completes the proof. □

## 4 An asymptotic for $|\operatorname{End}(\mathcal{IS}_n)|$

In [ST] it is shown that for $n > 6$ the cardinality of the semigroup $\operatorname{End}(\mathcal{IS}_n)$ of all endomorphisms of the semigroup $\mathcal{IS}_n$ equals
$$|\operatorname{End}(\mathcal{IS}_n)| = 3^n + 3\cdot n! + n!\sum_{m=0}^{n}\sum_{k=1}^{\lfloor m/2 \rfloor}\frac{2^{m-3k}}{(n-m)!\cdot(m-2k)!\cdot k!}.$$
On [ST, Page 303] the following problem is formulated:

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

In this section we answer both parts of this problem.

**Theorem 2.** $|\operatorname{End}(\mathcal{IS}_n)| \sim 3n!$.

*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 $X_n/n! \to 0$, $n \to \infty$. 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 $Y = \cup Y_i$, $|Y_i| = 2$, $Y_i \cap Y_j = \varnothing$ 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

$$X_n \leq \sum_{m=0}^{n} \binom{n}{m} \cdot 2^m \cdot 2^m \cdot m!! \leq \left( \sum_{m=0}^{n} \binom{n}{m} \cdot 4^m \right) (2\lceil n/2 \rceil)!! = 5^n 2^{\lceil n/2 \rceil} \lceil n/2 \rceil!.$$

To complete the proof it is enough to show that $5^n 2^{\lceil n/2 \rceil} \lceil n/2 \rceil!/n! \to 0$, $n \to \infty$. Using the Stirling formula we have

$$5^n 2^{\lceil n/2 \rceil} \lceil n/2 \rceil! \leq 5^n 2^{(n+1)/2} \lceil n/2 \rceil! \sim \frac{1}{\sqrt{\pi}} e^{n \ln 5\sqrt{2} - \frac{1}{2} \ln \lceil n/2 \rceil + \lceil n/2 \rceil \ln \lceil n/2 \rceil - \lceil n/2 \rceil},$$

and thus

$$\frac{5^n 2^{\lceil n/2 \rceil} \lceil n/2 \rceil!}{n!} \sim \frac{1}{\sqrt{2}} e^{n \ln 5\sqrt{2} - \frac{1}{2} \ln \lceil n/2 \rceil + \lceil n/2 \rceil \ln \lceil n/2 \rceil - \lceil n/2 \rceil - \frac{1}{2} \ln n - n \ln n + n}.$$

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.**

$$\frac{|\operatorname{End}(\mathcal{IS}_n)|}{|\mathcal{IS}_n|} \to 0, \quad n \to \infty.$$

*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.**

$$X_n \sim \frac{1}{\sqrt{2}} \cdot e^{\frac{1}{2} n \ln n - \frac{1}{2} n + 3\sqrt{n} - \frac{9}{4}}.$$

*Proof.* We can write

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

and we denote $a_k = \frac{2^{-3k}}{(m-2k)!\cdot k!}$, $b_m = \sum_{k=1}^{\lfloor m/2 \rfloor} a_k$, $c_m = \frac{2^m}{(n-m)!} b_k$. 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 $\varepsilon \le x \le C$. 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 $O\left(e^{-c\sqrt{m}} a_{m/2-\sqrt{m}}\right)$, that is relatively very small and can be neglected. Assume now that $1/2 \le x \le 2$. 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:

$$\begin{aligned}
\ln a_k = -3k \ln 2 - \ln(2x\sqrt{m})! - \ln k! = -\frac{3 \ln 2}{2} m + 3 \ln 2 \sqrt{m} x - 2x\sqrt{m} \ln 2 - \\
- 2x\sqrt{m} \ln x - x\sqrt{m} \ln m + 2x\sqrt{m} - \ln(2\pi) - \frac{1}{2} \ln(2x\sqrt{m}) - k \ln k + k - \frac{1}{2} \ln k + o(1) = \\
= -\frac{3 \ln 2}{2} m + (\ln 2 + 2)x\sqrt{m} - 2x\sqrt{m} \ln x - x\sqrt{m} \ln m - \ln(2\pi) - \frac{1}{2} \ln(2x\sqrt{m}) - \\
- k \ln \frac{m}{2} + \frac{2kx}{\sqrt{m}} + \frac{mx^2}{m} + \frac{m}{2} - x\sqrt{m} - \frac{1}{2} \ln \frac{m}{2} + o(1) = \\
= -m \ln 2 + 2x\sqrt{m} - 2x\sqrt{m} \ln x - \frac{1}{2} \ln(4\pi^2 x m^{3/2}) - \frac{1}{2} m \ln m - x^2 + \frac{m}{2} + o(1).
\end{aligned}$$

Further, assuming $x = 1 + m^{-1/4} y$ yields $x \ln x - x = -1 + \frac{1}{2}(x-1)^2 + O\left((x-1)^3\right) = -1 + \frac{y^2}{2\sqrt{m}} + O\left(\frac{y^3}{m^{3/4}}\right)$ and thus

$$\ln a_k = -\frac{1}{2} m \ln m + m\left(\frac{1}{2} - \ln 2\right) - 1 + O(ym^{-1/4}) + 2\sqrt{m} - y^2 + O(y^3 m^{-1/4}) - \frac{3}{4} \ln m - \ln(2\pi).$$

Therefore $k = \frac{m}{2} - \sqrt{m} - m^{1/4} y$ yields

$$a_k = \frac{1}{2\pi} e^{\left(\frac{1}{2} - \ln 2\right)m - \frac{1}{2} m \ln m - \frac{3}{4} \ln m + 2\sqrt{m} - 1} e^{-y^2} \left(1 + O\left(\frac{y + y^3}{m^{1/4}}\right)\right).$$

We can assume that, say, $y = O(m^{1/12})$ and ignore larger $y$. In this way we obtain

$$b_m = \sum_{k=1}^{\lfloor m/2 \rfloor} a_k = \frac{1}{2\pi} e^{\left(\frac{1}{2} - \ln 2\right)m - \frac{1}{2}m\ln m - \frac{3}{4}\ln m + 2\sqrt{m} - 1} m^{1/4} \int_{-\infty}^{\infty} e^{-y^2} dy \,(1 + o(1)) \sim$$

$$\sim \frac{1}{2\sqrt{\pi}} e^{\left(\frac{1}{2} - \ln 2\right)m - \frac{1}{2}m\ln m - \frac{1}{2}\ln m + 2\sqrt{m} - 1}.$$

The latter implies

$$\ln b_m = \left(\frac{1}{2} - \ln 2\right) m - \frac{1}{2} m \ln m - \frac{1}{2}\ln m + 2\sqrt{m} - 1 - \ln(2\sqrt{\pi}) + o(1) \tag{5}$$

and also

$$\ln c_m = \ln b_m + m \ln 2 - \ln((n-m)!). \tag{6}$$

Further, for $m \to \infty$ we compute $\ln \frac{b_{m+1}}{b_m} = \frac{1}{2} - \ln 2 - \frac{1}{2}\ln m - \frac{1}{2} + o(1) = -\ln 2 - \frac{1}{2}\ln m + o(1)$ 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 $m \le n/2$ 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 $M = \lceil n - 2\sqrt{n} \rceil$. 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 $1/2 \le x \le 3$. 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 $1/2 \le x \le 3$ we have, using the Stirling formula, that

$$\ln m = \ln n + \ln\left(1 - \frac{x}{\sqrt{n}}\right) = \ln n - \frac{x}{\sqrt{n}} - \frac{x^2}{2n} + O(n^{-3/2}),$$

$$m \ln m = n \ln n - x\sqrt{n}\ln n - x\sqrt{n} + x^2 - \frac{x^2}{2} + O(n^{-1/2}),$$

$$\ln(n-m)! = \ln(x\sqrt{n})! = x\sqrt{n}\ln x + \frac{1}{2}x\sqrt{n}\ln n - x\sqrt{n} + \frac{1}{2}\ln x + \frac{1}{4}\ln n + \ln\sqrt{2\pi} + o(1),$$

$$\sqrt{m} = \sqrt{n}(1 - x/\sqrt{n})^{1/2} = \sqrt{n} - x/2 + o(1).$$

Hence, using (5) and (6), we obtain

$$\ln c_m = \frac{1}{2}n - \frac{1}{2}x\sqrt{n} - \frac{1}{2}n\ln n + \frac{1}{2}x\sqrt{n}\ln n + \frac{1}{2}x\sqrt{n} - \frac{x^2}{4} - \frac{1}{2}\ln n + 2\sqrt{n} - x - 1 - \ln(2\sqrt{\pi}) -$$
$$- x\sqrt{n}\ln x - \frac{1}{2}x\sqrt{n}\ln n + x\sqrt{n} - \frac{1}{2}\ln x - \frac{1}{4}\ln n - \ln\sqrt{2\pi} + o(1) =$$
$$= \frac{1}{2}n - \frac{1}{2}n\ln n - \frac{3}{4}\ln n + 2\sqrt{n} - 1 - \ln(2^{3/2}\pi) + \sqrt{n}(x - x\ln x) - \frac{x^2}{4} - x - \frac{1}{2}\ln x + o(1).$$

Setting $x = 1 + yn^{-1/4}$ yields

$$\ln c_m = \frac{1}{2}n - \frac{1}{2}n\ln n - \frac{3}{4}\ln n + 3\sqrt{n} - 1 - \ln(2^{3/2}\pi) - \frac{5}{4} - \frac{y^2}{2} + O\left(\frac{y^3}{n^{1/4}}\right) + o(1)$$

and thus

$$\sum_{m=0}^{n} c_m = \exp\left(\frac{1}{2}n - \frac{1}{2}n\ln n - \frac{3}{4}\ln n + 3\sqrt{n} - \frac{9}{4} - \ln(2^{3/2}\pi)\right) n^{1/4} \times$$
$$\times \int_{-\infty}^{+\infty} e^{-y^2/2} dy\,(1 + o(1)) = \frac{1}{2\sqrt{\pi}} e^{\frac{1}{2}n - \frac{1}{2}n\ln n - \frac{1}{2}\ln n + 3\sqrt{n} - \frac{9}{4} + o(1)}.$$

This implies that

$$X_n \sim \frac{n!}{2\sqrt{\pi n}} e^{\frac{1}{2}n - \frac{1}{2}n\ln n + 3\sqrt{n} - \frac{9}{4}} \sim \frac{1}{\sqrt{2}} e^{\frac{1}{2}n\ln n - \frac{1}{2}n + 3\sqrt{n} - \frac{9}{4}},$$

and completes the proof. □

# 5 Some distributions

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

**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 $P(D_n = k) = \frac{R_{n,n-k}}{|\mathcal{IS}_n|}$ by definition and $\frac{R_{n,n-k}}{|\mathcal{IS}_n|} \sim \frac{1}{\sqrt{\pi} n^{1/4}} e^{-z^2}$, follows from (3) and (4). □

**Proposition 2.**
$$P(X_n = k) \sim \frac{1}{\sqrt{n}} e^{-k/\sqrt{n}} \textit{ if } k = o(n^{3/4}),$$
*in particular,*
$$\frac{X_n}{\sqrt{n}} \xrightarrow{d} \exp(1).$$

*Proof.* We have
$$P(X_n = k) = \binom{n}{k} \cdot k! \cdot \frac{|T_{n-k}|}{|\mathcal{IS}_n|} = \frac{|T_{n-k}|/(n-k)!}{|\mathcal{IS}_n|/n!}.$$
Hence, if $k = o(n)$ we have, using Section 2 and Section 3, that
$$P(X_n = k) \sim \frac{(n-k)^{-3/4} e^{2\sqrt{n-k}}}{n^{-1/4} e^{2\sqrt{n}}} \sim \frac{1}{\sqrt{n}} e^{2(\sqrt{n-k}-\sqrt{n})} =$$
$$= \frac{1}{\sqrt{n}} e^{2\sqrt{n}\left((1-k/n)^{1/2}-1\right)} = \frac{1}{\sqrt{n}} e^{-2\sqrt{n}\cdot\frac{k}{2n}+O(k^2/n^{3/2})}$$
and the statement follows. □

**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 $\ln\frac{X_n}{\sqrt{n}} \xrightarrow{d} \ln\exp(1)$ 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. □

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.**
$$d_{TV}\left(C_n, \mathrm{Po}\left(\frac{1}{2}\ln n\right)\right) \to 0, \qquad n \to \infty.$$

*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

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

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

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

Consequently,

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

Since also $d_{TV} \le 1$, we obtain $d_{TV}(C_n, \mathrm{Po}(\ln\sqrt{n})) \le E(f(X_n) \wedge 1)$. 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. $\square$

**Corollary 4.**

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

*Proof.* Follows from Propositions 1 and 3. $\square$

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

$$\mathrm{O}(\sigma) = |\{\sigma^l : l \in \{0, 1, 2, \dots\}\}|, \qquad \mathrm{IO}(\sigma) = |\{\sigma^l : l \in \mathbb{Z}\}|.$$

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

**Proposition 5.**

$$\frac{\ln O_n - \frac{1}{8}\ln^2 n}{\sqrt{\frac{1}{24}\ln^3 n}} \xrightarrow{d} N(0,1), \qquad \frac{\ln I_n - \frac{1}{8}\ln^2 n}{\sqrt{\frac{1}{24}\ln^3 n}} \xrightarrow{d} N(0,1).$$

*Proof.* For $\sigma \in \mathcal{IS}_n$ denote $X(\sigma) = \{i \in \{1, \dots, n\} : \sigma^l(i) = i \text{ for some } l > 0\}$. Then $|X(\sigma)|$ is the stable rank of $\sigma$. Moreover, any $\sigma \in \mathcal{IS}_n$ can be written as a product $\sigma = \sigma_1 \cdot \sigma_2$, where $\mathrm{dom}(\sigma_1) = \{1, \dots, n\}$ and $\sigma_1(i) = \sigma(i)$, $i \in X(\sigma)$, $\sigma_1(i) = i$, $i \notin X(\sigma)$; $\mathrm{dom}(\sigma_2) = \mathrm{dom}(\sigma)$ and $\sigma_2(i) = i$, $i \in X(\sigma)$, $\sigma_2(i) = \sigma(i)$, $i \in \mathrm{dom}(\sigma) \setminus X(\sigma)$. 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

$$\mathrm{O}(\sigma_1) \le \mathrm{O}(\sigma) \le \mathrm{O}(\sigma_1) + n - |X(\sigma)|, \quad \mathrm{O}(\sigma_1) \le \mathrm{IO}(\sigma) \le \mathrm{O}(\sigma_1) + 2(n - |X(\sigma)|). \tag{7}$$

For a random element $\sigma \in \mathcal{IS}_n$, let $\mathrm{O}'_n(\sigma) = \mathrm{O}(\sigma_1)$. Given $X_n = X(\sigma) = k$, this has the same distribution as the order $\tilde{\mathrm{O}}_k$ of a random permutation of length $k$. In [ET] it was shown that, as $k \to \infty$,

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

Hence, as $n \to \infty$,

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

and it follows as in the proof of Proposition 3 that

$$\frac{\ln \mathrm{O}'_n - \frac{1}{8}\ln^2 n}{\sqrt{\frac{1}{24}\ln^3 n}} \xrightarrow{d} N(0,1). \tag{8}$$

In particular, for almost all $\sigma \in \mathcal{IS}_n$ we have that $\mathrm{O}(\sigma_1) \approx n^{(\ln n)/8}$. Since the difference between the left and the right sides of the inequalities in (7) is less than $2n$, in particular is $o(n^{(\ln n)/9})$, 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 $A(z) = \sum_{m \geq 1} z^m = z/(1-z)$ 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 $L_n$:

**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 $E_{L_n}(z) = \frac{z}{(1-z)^2}e^{z/(1-z)}$.*

*Proof.* A fixed chain of length $m$ is contained in exactly $|T_{n-m}|$ elements of $T_n$, and in exactly $|\mathcal{IS}_{n-m}|$ elements of $\mathcal{IS}_n$. This implies that $E_{L^{(n)}}(z) = \frac{z}{1-z}E_{T_n}(z) = \frac{z}{1-z}e^{z/(1-z)}$ and $E_{L_n}(z) = \frac{z}{1-z}E_{\mathcal{IS}_n}(z) = \frac{z}{(1-z)^2}e^{z/(1-z)}$. □

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

**Corollary 5.** *([GM2, Theorem 7(2)]) $|\mathcal{IS}_n| = |T_n| + L^{(n)}$.*

*Proof.* Follows from $E_{\mathcal{IS}_n}(z) = E_{T_n}(z) + E_{L^{(n)}}(z)$. □

**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)]) $|\mathcal{IS}_n| = \frac{1}{n+1}L^{(n+1)}$.*

*Proof.* The statement is equivalent to $zE_{\mathcal{IS}_n}(z) = E_{L^{(n)}}(z)$, 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 $\mathcal{IS}_n$ satisfies $P_n = L^{(n)}$.*

*Proof.* For $x \in \{1, \dots, n\}$, the point $x$ is fixed in exactly $|\mathcal{IS}_{n-1}|$ elements of $\mathcal{IS}_n$, which implies that $E_{P_n}(z) = zE_{\mathcal{IS}_n}(z)$. Further $zE_{\mathcal{IS}_n}(z) = E_{L^{(n)}}(z)$ 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


[BH] A.D.Barbour, P.Hall, On the rate of Poisson convergence. Math. Proc. Cambridge Philos. Soc. 95 (1984), no. 3, 473–480.

[BHJ] A.Barbour, L.Holst, S.Janson, Poisson approximation. Oxford Studies in Probability, 2. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1992. x+277 pp.

[ET] P.Erdös, P.Turán, On some problems of a statistical group theory. I. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 4 1965 175–186 (1965).

[GK] O.G.Ganyushkin, T.V.Kormysheva, Chain decomposition of partial permutations and classes of conjugated elements of the semigroup $\mathcal{IS}_n$, Proc. Kyiv University, Ser. Math.-Mech. 27 (1992), 10-18.

[GM1] O.Ganyushkin, V.Mazorchuk, The full finite Inverse symmetric semigroup $\mathcal{IS}_n$, Preprint 2001:37, Chalmers University of Technology and Göteborg University, Göteborg, 2001.

[GM2] O.Ganyushkin, V.Mazorchuk, Combinatorics and distributions of partial injections, Preprint 2003:14, Uppsala University, Sweden, 2003.

[Go] V.L.Goncharov, Some facts from combinatorics, Izvestia Akad. Nauk. SSSR, Ser. Mat. 8: 3-48 (1944).

[Li] S.Lipscomb, Symmetric inverse semigroups. Mathematical Surveys and Monographs, 46. American Mathematical Society, Providence, RI, 1996.

[ST] B.M.Schein, B.Teclezghi, Endomorphisms of finite symmetric inverse semigroups. J. Algebra 198 (1997), no. 1, 300–310.

[Wi] H.S.Wilf, Generatingfunctionology. Second edition. Academic Press, Inc., Boston, MA, 1994. x+228 pp.



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”