Second Best, Third Worst, Fourth in Line

STEVEN FINCH

May 1, 2022

Abstract. We investigate decomposable combinatorial labeled structures more fully, focusing on the exp-log class of type $a = 1$ or $1/2$. For instance, the modal length of the second longest cycle in a random $n$-permutation is $(0.2350...)n$, whereas the modal length of the second smallest component in a random $n$-mapping is $2$ (conjecturally, given $n \geq 434$). As in earlier work, our approach is to establish how well existing theory matches experimental data and to raise open questions.

Given a combinatorial object with $n$ nodes, our interest is in

- the size of its $r$th longest cycle or largest component,
- the size of its $r$th shortest cycle or smallest component

where $r \geq 2$. If the object has no $r$th component, then its $r$th largest/smallest components are defined to have length 0. The case $r = 1$ has attracted widespread attention [1, 2]. Key to our prior study were recursive formulas [3, 4] for $L_{k,n}$ and $S_{k,n}$, the number of $n$-objects whose largest and smallest components, respectively, have exactly $k$ nodes, $1 \leq k \leq n$. Different algorithms shall be used here. As before, an $n$-object is chosen uniformly at random. For simplicity, we discuss here only $n$-permutations and $n$-mappings (from $\{1,2,\ldots,n\}$ to $\{1,2,\ldots,n\}$). Let $c_n$ be the number of $n$-objects that are connected, i.e., who possess exactly one component:

$$c_n = \begin{cases} (n - 1)! & \text{for permutations,} \\ n! \sum_{j=1}^{n} \frac{n^{n-j-1}}{(n-j)!} & \text{for mappings.} \end{cases}$$

The total number of $n$-permutations and $n$-mappings is $n!$ and $n^n$, respectively. For fixed $n$, the sequences $\{L_{k,n} : 1 \leq k \leq n\}$ and $\{S_{k,n} : 1 \leq k \leq n\}$ constitute probability mass functions (upon normalization) for $r = 1$. Until recently, calculating analogous sequences for $r \geq 2$ seemed inaccessibly difficult.

The new algorithms, due to Heinz [5], accept as input the integer $n$ and an ordered $r$-tuple $\ell$ of nonnegative integers, which may include infinity. We write $\ell$ as a list $\{i_1,i_2,\ldots,i_r\}$. Given a positive integer $j$, define $\ell^j$ to be the list obtained by

---

Copyright © 2022 by Steven R. Finch. All rights reserved.
(i) appending $\ell$ with $j$,

(ii) sorting the $(r + 1)$-tuple in ascendent order, and

(iii) removing its first element.

Define $\ell_j$ in the same way as $\ell^j$ except for a revised final step:

(iii') removing its last element.

Note that the lengths of $\ell^j$ and $\ell_j$ are always equal to the length of $\ell$. Let $p[n, \ell]$ and $q[n, \ell]$ denote row polynomials in $x$ and $y$ associated with large and small components. The algorithms are based on recursions

$$p[n, \ell] = \begin{cases} \sum_{j=1}^{n} c_j p[n - j, \ell^j] \binom{n - 1}{j - 1} & \text{if } n > 0, \\ x^i & \text{if } n = 0; \end{cases}$$

$$q[n, \ell] = \begin{cases} \sum_{j=1}^{n} c_j q[n - j, \ell_j] \binom{n - 1}{j - 1} & \text{if } n > 0, \\ y^{i_r} & \text{if } n = 0 \text{ and } i_r < \infty, \\ y^0 & \text{if } n = 0 \text{ and } i_r = \infty. \end{cases}$$

A computer algebra software package (e.g., Mathematica or Maple) makes exact integer calculations for ample $n$ of $p[n, \ell]$ and $q[n, \ell]$ feasible. These are demonstrated for $n = 4$ in the next section, for the sake of concreteness.

Permutations belong to the exp-log class of type $a = 1$, whereas mappings belong to the exp-log class of type $a = 1/2$. Explaining the significance of the parameter $a > 0$ would take us too far afield [6]. Let

$$E(x) = \int_{x}^{\infty} \frac{e^{-t}}{t} dt = -\text{Ei}(-x), \quad x > 0$$

be the exponential integral. Define [7, 8, 9, 10, 11, 12]

$$\ell G_a(r, h) = \frac{\Gamma(a + 1)a^{r-1}}{\Gamma(a + h)(r - 1)!} \int_{0}^{\infty} x^{h-1} E(x)^{r-1} \exp[-a E(x) - x] dx,$$
which are related to the \(h\)th moment of the \(r\)th largest/smallest component size (in this paper, rank \(r = 2, 3\) or \(4\); height \(h = 1\) or \(2\)). While moment formulas are unerring for \(L\), they are not so for \(S\). While \(sG_a\) is flawless for permutations (and for what are called cyclations [I3]), a correction factor \(\sqrt{2}\) is needed for mappings.

For fixed \(n\) and \(r\), the coefficient sequences associated with polynomials

\[
p[n, \{0, 0, \ldots, 0\}], \quad 0 \leq k \leq \lfloor n/r \rfloor;
\]

\[
q[n, \{\infty, \infty, \ldots, \infty\}], \quad 0 \leq k \leq n - r + 1
\]

constitute probability mass functions (upon normalization). These have corresponding means \(L\mu_{n,r}, s\mu_{n,r}\) and variances \(L\sigma_{n,r}^2, s\sigma_{n,r}^2\) given in the tables. We also provide the median \(L\nu_{n,r}\) and mode \(L\vartheta_{n,r};\) evidently \(s\nu_{n,r}\) and \(s\vartheta_{n,r}\) are bounded for permutations as \(n \to \infty\) (the trend of \(s\nu_{n,r}\) is less clear for mappings). In table headings only, the following notation is used:

\[
\bar{L}\mu_{n,r} = \frac{L\mu_{n,r}}{n}, \quad \bar{L}\sigma_{n,r}^2 = \frac{L\sigma_{n,r}^2}{n^2}, \quad \bar{L}\nu_{n,r} = \frac{L\nu_{n,r}}{n}, \quad \bar{L}\vartheta_{n,r} = \frac{L\vartheta_{n,r}}{n}, \quad s\nu_{n,r} = \frac{s\nu_{n,r}}{n},
\]

\[
s\mu_{n,r} = \begin{cases} 
\frac{s\mu_{n,r}}{\ln(n)^r} & \text{if } a = 1, \\
\frac{s\mu_{n,r}}{n^{1/2} \ln(n)^{r-1}} & \text{if } a = 1/2,
\end{cases}
\]

\[
s\sigma_{n,r}^2 = \begin{cases} 
\frac{s\sigma_{n,r}^2}{n \ln(n)^{r-1}} & \text{if } a = 1, \\
\frac{s\sigma_{n,r}^2}{n^{3/2} \ln(n)^{r-1}} & \text{if } a = 1/2.
\end{cases}
\]

When \(r = 1\), the mode \(L\vartheta_{n,1}\) is provably 1/2 in the limit as \(n \to \infty\) for permutations (it is 1 for mappings). This limit is more interesting when \(r = 2\), as will soon be seen.

1. **Calculs à la Heinz**

As promised, we exhibit some hand calculations. It is easy to show directly that \(p[3, \{0, 0\}] = 2 + 4x\) for permutations and \(17 + 10x\) for mappings (see Section 3 of [I]). More generally, \(p[3, \{0, 0\}] = c_3 + c_1(c_1^2 + 3c_2)x\). Let us compute \(p[4, \{0, 0\}]\) using Heinz’s algorithm. From

\[
\begin{align*}
p[2, \{1, 1\}] &= c_1p[1, \{1, 1\}]\binom{4}{1} + c_2p[0, \{1, 2\}]\binom{4}{1} \\
&= c_1^2p[0, \{1, 1\}]\binom{6}{1} + c_2x^1 = (c_1^2 + c_2)x,
\end{align*}
\]
we have
\[ p[0, \{1, 3\}] = x \]

Also, from
\[ p[2, \{0, 2\}] = c_1 p[1, \{1, 2\}] \binom{4}{1} + c_2 p[0, \{2, 2\}] \binom{3}{1} 
= c_1^2 p[0, \{1, 2\}] \binom{3}{2} + c_2 x^2 = c_1^2 x + c_2 x^2, \]
\[ p[1, \{0, 3\}] = c_1 p[0, \{1, 3\}] \binom{4}{0} = c_1 x, \]
\[ p[0, \{0, 4\}] = x^0 = 1 \]
we deduce
\[ p[4, \{0, 0\}] = c_1 p[3, \{0, 1\}] \binom{3}{0} + c_2 p[2, \{0, 2\}] \binom{3}{1} + c_3 p[1, \{0, 3\}] \binom{3}{2} + c_4 p[0, \{0, 4\}] \binom{3}{3} 
= c_1 \left( c_1^3 + 3c_1c_2 + c_3 \right) x + 3c_2 (c_1^2 x + c_2 x^2) + 3c_3 (c_1 x) + c_4 
= c_4 + c_1 \left( c_1^3 + 6c_1c_2 + 4c_3 \right) x + 3c_2 x^2 \]
\[ = \begin{cases} 6 + 15x + 3x^2 & \text{for permutations,} \\ 142 + 87x + 27x^2 & \text{for mappings} \end{cases} \]
completing the argument.

It is likewise easy to show that \( q[3, \{0, 0\}] = 2 + y + 3y^2 \) for permutations and \( 17 + y + 9y^2 \) for mappings. More generally, \( q[3, \{0, 0\}] = c_3 + c_1^2 y + 3c_1c_2 y^2 \). Let us compute \( q[4, \{0, 0\}] \) using Heinz’s algorithm. From
\[ q[2, \{1, 1\}] = c_1 q[1, \{1, 1\}] \binom{4}{1} + c_2 q[0, \{1, 1\}] \binom{4}{1} 
= c_1^2 q[0, \{1, 1\}] \binom{3}{1} + c_2 y^1 = (c_1 + c_2) y, \]
\[ q[1, \{1, 2\}] = c_1 q[0, \{1, 1\}] \binom{4}{0} = c_1 y, \]
\[ q[0, \{1, 3\}] = y^3 \]
we have
\[ q[3, \{1, \infty\}] = c_1 q[2, \{1, 1\}] \binom{4}{0} + c_2 q[1, \{1, 2\}] \binom{4}{1} + c_3 q[0, \{1, 3\}] \binom{4}{2} 
= c_1 \left( c_1^2 + c_2 \right) y + 2c_2 (c_1 y) + c_3 y^3 = (c_1^3 + 3c_1c_2) y + c_3 y^3. \]
Also, from
\[
q[2, \{2, \infty\}] = c_1 q[1, \{1, 2\}]_0^{(1)} + c_2 q[0, \{2, 2\}]_0^{(1)} = c_1^2 q[0, \{1, 1\}]_0^0 + c_2 y^2 = c_1 y + c_2 y^2,
\]
\[
q[1, \{3, \infty\}] = c_1 q[0, \{1, 3\}]_0^0 = c_1 y^3,
\]
\[
q[0, \{4, \infty\}] = y^0 = 1
\]
we deduce
\[
q[4, \{\infty, \infty\}] = \begin{cases}
\begin{align*}
c_1 & \left( c_3 + 3c_1 c_2 \right) y + c_3 y^3 + c_2 \left( c_1^2 y + c_2 y^2 \right) + 3c_3 (c_1 y^3) + c_4 \\
c_1 & \left( c_1^2 + 6c_2 \right) y + 3c_2 y^2 + 4c_3 c_3 y^3
\end{align*}
\end{cases}
\]
\[
= \begin{cases}
\begin{align*}
6 + 7y + 3y^2 + 8y^3 & \quad \text{for permutations,} \\
142 + 19y + 27y^2 + 68y^3 & \quad \text{for mappings}
\end{align*}
\end{cases}
\]
completing the argument.

2. Modes & Medians

The mode of a continuous distribution is the location of its highest peak; the median is its 50th percentile. The length \( \Lambda_r \) of the \( r \)th longest cycle in a random \( n \)-permutation has cumulative probability

\[
\lim_{n \to \infty} \Pr \{ \Lambda_r < x \cdot n \} = \rho_r \left( \frac{1}{x} \right)
\]

where \( \rho_r(x) \) is the \( r \)th order Dickman function [14]:

\[
x \rho'_1(x) + \rho_1(x - 1) = 0 \quad \text{for} \quad x > 1, \quad \rho_1(x) = 1 \quad \text{for} \quad 0 \leq x \leq 1;
\]

\[
x \rho'_r(x) + \rho_r(x - 1) = \rho_{r-1}(x - 1) \quad \text{for} \quad x > 1, \quad \rho_r(x) = 1 \quad \text{for} \quad 0 \leq x \leq 1
\]

and \( r = 2, 3, 4, \ldots \). For notational simplicity, let us write \( \varphi = \rho_1 \) and \( \psi = \rho_2 \). Observe that \( \rho_r \) should not be confused with a different generalization \( \rho_a \) discussed in [11][15].

From

\[
\varphi'(x) = -\frac{\varphi(x - 1)}{x}, \quad x > 1
\]

we have

\[
\varphi' \left( \frac{1}{x} \right) = -\frac{\varphi \left( \frac{1}{x} - 1 \right)}{x}, \quad 0 < x < 1
\]
hence the density \( f(x) \) is

\[
\frac{d}{dx} \varphi \left( \frac{1}{x} \right) = -x \varphi \left( \frac{1}{x} - 1 \right) \left( -\frac{1}{x^2} \right) = \begin{cases} 
\frac{\varphi \left( \frac{1}{x} - 1 \right)}{x} & \text{if } 0 < x \leq 1/2, \\
\frac{1}{x} & \text{if } 1/2 < x < 1.
\end{cases}
\]

Also, from

\[
\varphi''(x) = \frac{\varphi(x - 1)}{x^2} - \frac{\varphi'(x - 1)}{x} = \frac{\varphi(x - 1)}{x^2} + \frac{\varphi(x - 2)}{x(x - 1)}, \quad x > 1
\]

we have

\[
\varphi'' \left( \frac{1}{x} \right) = \frac{\varphi \left( \frac{1}{x} - 1 \right)}{x^2} + \frac{\varphi \left( \frac{1}{x} - 2 \right)}{x(x - 1)}, \quad 0 < x < 1
\]

hence (by the chain rule for second derivatives)

\[
\frac{d^2}{dx^2} \varphi \left( \frac{1}{x} \right) = \varphi' \left( \frac{1}{x} \right) \frac{2}{x^3} + \frac{1}{x^4} \varphi'' \left( \frac{1}{x} \right)
\]

\[
= -2 \varphi \left( \frac{1}{x} - 1 \right) + \frac{\varphi \left( \frac{1}{x} - 1 \right)}{x^2} + \frac{\varphi \left( \frac{1}{x} - 2 \right)}{x^2(1 - x)}
\]

\[
= \begin{cases} 
\frac{1}{x^2(1 - x)} - \frac{\varphi \left( \frac{1}{x} - 1 \right)}{x^2} > \frac{1}{x(1 - x)} > 0 & \text{if } 1/3 < x \leq 1/2, \\
-\frac{1}{x^2} < 0 & \text{if } 1/2 < x \leq 1
\end{cases}
\]

since the first condition implies \( 3 > 1/x \geq 2 \), i.e., \( 1 > 1/x - 2 \geq 0 \) and the second condition implies \( 2 > 1/x \geq 1 \), i.e., \( 1 > 1/x - 1 \geq 0 \). Thus \( f \) is increasing on the left of \( x = 1/2 \) and \( f \) is decreasing on the right, which implies that the median size of \( \Lambda_1 \) is 1/2.

From

\[
\psi'(x) = \frac{\varphi(x - 1) - \psi(x - 1)}{x}, \quad x > 2
\]

we have

\[
\varphi'(x) - \psi'(x) = -\frac{\varphi(x - 1)}{x} - \frac{\varphi(x - 1) - \psi(x - 1)}{x} = -2 \varphi(x - 1) + \psi(x - 1)
\]
(a lemmata needed shortly) and
\[
\psi'(\frac{1}{x}) = \varphi\left(\frac{1}{x} - 1\right) - \psi\left(\frac{1}{x} - 1\right) \quad \frac{1}{x}, \quad 0 < x < 1/2
\]
hence the density \(g(x)\) is
\[
\frac{d}{dx} \psi\left(\frac{1}{x}\right) = x \left( \varphi\left(\frac{1}{x} - 1\right) - \psi\left(\frac{1}{x} - 1\right) \right) \left( -\frac{1}{x^2} \right)
\]
\[
\psi\left(\frac{1}{x} - 1\right) - \varphi\left(\frac{1}{x} - 1\right)
\]
\[
\psi\left(1 - x\right) \quad x.
\]
Also, from
\[
\psi''(x) = -\frac{\varphi(x - 1) - \psi(x - 1)}{x^2} + \frac{\varphi'(x - 1) - \psi'(x - 1)}{x}
\]
\[
= -\varphi(x - 1) + \psi(x - 1) + \frac{-2\varphi(x - 2) + \psi(x - 2)}{x(x - 1)}
\]
(by the lemmata) we have
\[
\psi''\left(\frac{1}{x}\right) = \frac{-\varphi\left(\frac{1}{x} - 1\right) + \psi\left(\frac{1}{x} - 1\right)}{\frac{1}{x^2}} + \frac{-2\varphi\left(\frac{1}{x} - 2\right) + \psi\left(\frac{1}{x} - 2\right)}{\frac{1}{x} \left(\frac{1}{x} - 1\right)}
\]
hence (by the chain rule for second derivatives)
\[
\frac{d^2}{dx^2} \psi\left(\frac{1}{x}\right) = \psi'\left(\frac{1}{x}\right) \frac{2}{x^3} + \frac{1}{x^4} \psi''\left(\frac{1}{x}\right)
\]
\[
\quad = \frac{\varphi\left(\frac{1}{x} - 1\right) - \psi\left(\frac{1}{x} - 1\right)}{\frac{1}{x}} \frac{2}{x^3}
\]
\[
\quad + \frac{1}{x^4} \left[ -\varphi\left(\frac{1}{x} - 1\right) + \psi\left(\frac{1}{x} - 1\right) + \frac{-2\varphi\left(\frac{1}{x} - 2\right) + \psi\left(\frac{1}{x} - 2\right)}{\frac{1}{x} \left(\frac{1}{x} - 1\right)} \right]
\]
\[
\quad = \frac{\varphi\left(\frac{1}{x} - 1\right) - \psi\left(\frac{1}{x} - 1\right)}{x^2} - \frac{2\varphi\left(\frac{1}{x} - 2\right) - \psi\left(\frac{1}{x} - 2\right)}{x^2 (1 - x)}.
\]
There exists a unique $0 < x_0 < 1/2$ for which this expression $g'(x_0)$ vanishes. Plots of $f(x)$ and $g(x)$ appear in [16] and confirm that $x_0$ is the modal size of $\Lambda_2$. Broadhurst [17] obtained an exact equation for $x_0$, involving Dickman dilogarithms and trilogarithms [18], then applied numerics. We have verified his value $x_0$ by purely floating point methods.

There is comparatively little to say about medians $\xi_r$, defined as solutions of [14, 19]

$$\rho_r \left( \frac{1}{x} \right) = \frac{1}{2}$$

except that $\xi_1 = 1/\sqrt{e}$ is well-known and no closed-form representations for $\xi_r$, $r \geq 2$, seem to exist.

3. Knuth & Trabb Pardo

An alternative to Heinz’s algorithm is one proposed by Knuth & Trabb Pardo [14] for a restricted case. Define $u_r(k, n)$ to be the number of $n$-permutations whose $r^{th}$ longest cycle has $\leq k$ nodes [20]. The following recursive formulas apply for $r = 1$:

$$u_1(k, n) = \begin{cases} \sum_{m=0}^{k-1} \frac{(n-1)!}{(n-1-m)!} u_1(k, n-1-m) & \text{if } n \geq 1 \text{ and } k < n, \\ n! & \text{if } n \geq 1 \text{ and } k \geq n, \\ 1 & \text{otherwise} \end{cases}$$

and for $r \geq 2$:

$$u_r(k, n) = \begin{cases} \sum_{m=0}^{k-1} \frac{(n-1)!}{(n-1-m)!} u_r(k, n-1-m) + \\ \sum_{m=k}^{n-1} \frac{(n-1)!}{(n-1-m)!} u_{r-1}(k, n-1-m) & \text{if } n \geq 1 \text{ and } k < \lfloor n/r \rfloor, \\ n! & \text{if } n \geq 1 \text{ and } k \geq \lfloor n/r \rfloor, \\ 1 & \text{otherwise.} \end{cases}$$

Clearly $u_1(0, n) = \delta_{0,n}$ and $u_1(1, n) = 1$, hence

$$u_2(0, 4) = u_1(0, 3) + 3u_1(0, 2) + 6u_1(0, 1) + 6u_1(0, 0) = 6.$$  

Also $u_2(1, 2) = 2$ and $u_2(1, 3) = 6$, hence

$$u_2(1, 4) = u_2(1, 3) + [3u_1(1, 2) + 6u_1(1, 1) + 6u_1(1, 0)] = 6 + 15 = 21.$$
Finally \( u_2(2, 4) = 24 \). The list

\[
\left\{ u_2(k, 4) \right\}_{k=0}^2 = \{6, 21, 24\} = \{6, 6 + 15, 21 + 3\}
\]

conveys the same information as the polynomial \( p[4, \{0, 0\}] \) did in Section 1, although the underlying calculations differed completely.

A proof is as follows \([14]\). We may think of \( u_r(k, n) \) as counting permutations on \( \{1, \ldots, n\} \) that possess fewer than \( r \) cycles of length exceeding \( k \). Call such a permutation \( (r, n) \)-good. Consider now a permutation \( P \) on \( \{0, 1, \ldots, n\} \). The node 0 belongs to some cycle \( C \) within \( P \) of length \( m + 1 \). Let \( P \setminus C \) denote the permutation which remains upon exclusion of \( C \) from \( P \). Suppose \( 0 \leq m \leq k - 1 \); then \( P \) is \((r, n + 1)\)-good iff \( P \setminus C \) is \((r, n - m)\)-good. Suppose \( k \leq m \leq n \); then \( P \) is \((r, n + 1)\)-good iff \( P \setminus C \) is \((r - 1, n - m)\)-good. Thus the formula

\[
u_r(k, n + 1) = \sum_{m=0}^{k-1} \frac{n!}{(n-m)!} u_r(k, n-m) + \sum_{m=k}^{n} \frac{n!}{(n-m)!} u_{r-1}(k, n-m)
\]

is true because \( n!/(n-m)! \) is the number of possible choices for \( C \).

An analog of this recursion for mappings remains open, as far as is known. Finding the number of possible choices for a component \( C \) containing the node 0 is more complicated than for a cycle containing 0. Each component consists of a cycle with trees attached; each tree is rooted at a cyclic point but is otherwise made up of transient points. We must account for the position of 0 (cyclic or transient?) and the overall configuration (inventory of tree types and sizes?). It would be helpful to learn about progress in enumerating such \( C \) or, if this is impractical, some other procedure for moving forward.

4. Une conjecture correspondante

Short cycles have always presented more analytical difficulties than long cycles; this paper offers no exception. Everything in this section is conjectural only. Define \( v_r(k, n) \) to be the number of \( n \)-permutations whose \( r \)th shortest cycle has \( \geq k \) nodes \([20]\). The following recursive formulas would seem to apply for \( r = 1 \):

\[
v_1(k, n) = \begin{cases} 
  n! & \text{if } n \geq 1 \text{ and } k = 0, \\
  \sum_{m=k-1}^{n-1} \frac{(n-1)!}{(n-1-m)!} v_1(k, n-1-m) & \text{if } n \geq 1 \text{ and } 0 < k \leq n, \\
  0 & \text{if } n \geq 1 \text{ and } k > n, \\
  1 & \text{otherwise}
\end{cases}
\]
and for \( r \geq 2 \):

\[
v_r(k, n) = \begin{cases} 
  n! & \text{if } n \geq 0 \text{ and } k = 0, \\
  \Delta_r(k, n) + \\
  \sum_{m=0}^{k-2} \frac{(n-1)!}{(n-1-m)!} v_{r-1}(k, n-1-m) + \\
  \sum_{m=k-1}^{n-1} \frac{(n-1)!}{(n-1-m)!} v_r(k, n-1-m) & \text{if } n \geq 1 \text{ and } 0 < k \leq n - r + 1, \\
  0 & \text{otherwise.}
\end{cases}
\]

The surprising new term \( \Delta_r(k, n) \) has a simple formula for \( r = 2 \):

\[
\Delta_2(k, n) = (n-1)!H_{n-k}, \quad \text{where} \quad \sum_{i=1}^{j} \frac{1}{i} = H_j, \quad \sum_{i=1}^{j} \frac{1}{i^s} = H_{j,s}
\]

and unexpected recursions for \( r = 3 \) and \( r = 4 \):

\[
\Delta_r(k, n) = \begin{cases} 
  \frac{1}{2}(n-1)! \left( H_{n-1}^2 - H_{n-1,2} \right) & \text{if } r = 3 \text{ and } k = 1, \\
  \frac{1}{6}(n-1)! \left( H_{n-1}^3 - 3H_{n-1}H_{n-1,2} + 2H_{n-1,3} \right) & \text{if } r = 4 \text{ and } k = 1, \\
  \Delta_r(k-1, n) - \frac{\Delta_{r-1}(k, n)}{n-k+1} & \text{if } k \geq 2 \text{ and } n \geq k, \\
  0 & \text{otherwise.}
\end{cases}
\]

The values \( \Delta_r(1, n) \) are unsigned Stirling numbers of the first kind, i.e., the number of \( n \)-permutations that have exactly \( r \) cycles. (Why should these appear here?) A plausibility argument supporting \( v_r \) bears resemblance to the proof underlying \( u_r \). We may think of \( v_r(k, n) \) as counting permutations on \( \{1, \ldots, n\} \) that possess fewer than \( r \) cycles of length surpassed by \( k \). Call such a permutation \((r, n)\)-bad. Let \( P \& C \) (of lengths \( n+1 \) & \( m+1 \)) be as before. Suppose \( 0 \leq m \leq k - 2 \); then \( P \) is \((r, n+1)\)-bad iff \( P \setminus C \) is \((r-1, n-m)\)-bad. Suppose \( k - 1 \leq m \leq n \); then \( P \) is \((r, n+1)\)-bad iff \( P \setminus C \) is \((r, n-m)\)-bad. This would suggest

\[
v_r(k, n+1) = \Delta_r + \sum_{m=0}^{k-2} \frac{n!}{(n-m)!} v_{r-1}(k, n-m) + \sum_{m=k-1}^{n} \frac{n!}{(n-m)!} v_r(k, n-m)
\]

is true with \( \Delta_r = 0 \), but experimental data contradict such an assertion.
Let us illustrate via example, in parallel with Section 3. As preliminary steps, 
$v_1(k, 0) = 1$ and $v_1(n + 1, n) = 0$, hence 
$$v_1(2, 3) = 2v_1(2, 1) + 2v_1(2, 0) = 2, \quad v_1(3, 3) = 2v_1(3, 0) = 2.$$ 
Clearly $v_2(0, 4) = 24$. Also $v_2(n, n) = \delta_{0,n}$ and $v_2(n + 1, n) = v_2(n + 2, n) = 0$, hence 
$$v_2(1, 2) = \Delta_2(1, 2) + [v_2(1, 1) + v_2(1, 0)] = 1 + 0 = 1,$$
$$v_2(1, 3) = \Delta_2(1, 3) + [v_2(1, 2) + 2v_2(1, 1) + 2v_2(1, 0)] = 3 + 1 = 4,$$
$$v_2(1, 4) = \Delta_2(1, 4) + [v_2(1, 3) + 3v_2(1, 2) + 6v_2(1, 1) + 6v_2(1, 0)] = 11 + 7 = 18.$$ 
Finally 
$$v_2(2, 4) = \Delta_2(2, 4) + v_1(2, 3) + [3v_2(2, 2) + 6v_2(2, 1) + 6v_2(2, 0)] = 9 + 2 + 0 = 11,$$
$$v_2(3, 4) = \Delta_2(3, 4) + [v_1(3, 3) + 3v_1(3, 2)] + [6v_2(3, 1) + 6v_2(3, 0)] = 6 + 2 + 0 = 8.$$ 
Again, the list 
$$\{v_2(k, 4)\}_{k=1}^3 = \{18, 11, 8\} = \{24 - 6, 18 - 7, 11 - 3 = 8\}$$
conveys the same information as the polynomial $q[4, \{\infty, \infty\}]$ did in Section 1. Without the nonzero contribution of $\Delta_r(k, n)$, our modification of Knuth & Trabb Pardo would yield results incompatible with Heinz.

5. Permutations

Here are numerical results for $r = 2$:

| $n$  | $L\tilde{\mu}_{n,2}$ | $L\tilde{\nu}_{n,2}$ | $L\tilde{\vartheta}_{n,2}$ | $S\tilde{\mu}_{n,2}$ | $S\tilde{\nu}_{n,2}$ |
|------|------------------------|------------------------|------------------------|------------------------|------------------------|
| 1000 | 0.209685               | 0.012567               | 0.2110                 | 0.2350                 | 0.415946               |
| 1500 | 0.209650               | 0.012562               | 0.2113                 | 0.2353                 | 0.408887               |
| 2000 | 0.209633               | 0.012560               | 0.2115                 | 0.2350                 | 0.404309               |
| 2500 | 0.209623               | 0.012559               | 0.2112                 | 0.2352                 | 0.400976               |

Table 5.1: Statistics for Permute, rank two ($a = 1$)

as well as $S\nu_{n,2} = 2$ for $n > 17$ and $S\vartheta_{n,2} = 1$ for $n > 4$. Also 
$$\lim_{n \to \infty} \frac{L\mu_{n,2}}{n} = L\varGamma(2, 1) = 0.20958087428418581398\ldots,$$
$$\lim_{n \to \infty} \frac{L\sigma_{n,2}}{n^2} = L\varGamma(2, 2) - L\varGamma(2, 1)^2 = 0.01255379063590587814\ldots,$$
\[
\lim_{n \to \infty} \frac{L\nu_{n,2}}{n} = \xi_2 = 0.21172114641298273896..., \\
\lim_{n \to \infty} \frac{L\vartheta_{n,2}}{n} = x_0 = 0.23503964593509109370..., \\
\lim_{n \to \infty} \frac{s\mu_{n,2}}{\ln(n)^2} = \frac{e^{-\gamma}}{2} = 0.28072974178344258491..., \\
\lim_{n \to \infty} \frac{s\sigma^2_{n,2}}{n\ln(n)} = sG_P(2,2) = 1.30720779891056809974....
\]

The final $n\ln(n)$ asymptotic is based on [7, 8], not (inaccurate) Theorem 5 in [6].

Here [22] are numerical results for $r = 3$:

| $n$   | $L\tilde{\mu}_{n,3}$ | $L\tilde{\sigma}^2_{n,3}$ | $L\tilde{\nu}_{n,3}$ | $L\tilde{\vartheta}_{n,3}$ | $S\tilde{\mu}_{n,3}$ | $S\tilde{\sigma}^2_{n,3}$ |
|-------|------------------------|-----------------------------|------------------------|---------------------------|----------------------|-----------------------------|
| 1000  | 0.088357               | 0.004499                    | 0.0750                 | 0.0010                    | 0.155997            | 0.450101                    |
| 1500  | 0.088344               | 0.004497                    | 0.0753                 | 0.0007                    | 0.153079            | 0.468681                    |
| 2000  | 0.088337               | 0.004496                    | 0.0755                 | 0.0005                    | 0.151161            | 0.480325                    |
| 2500  | 0.088333               | 0.004496                    | 0.0756                 | 0.0004                    | 0.149752            | 0.488548                    |

Table 5.2: Statistics for Permute, rank three ($a = 1$)

as well as $s\nu_{n,2} = 7$ for $n > 370$ and $s\vartheta_{n,2} = 2$ for $n > 49$. Also

\[
\lim_{n \to \infty} \frac{L\mu_{n,3}}{n} = LG_1(3,1) = 0.0883160988315363101..., \\
\lim_{n \to \infty} \frac{L\sigma^2_{n,3}}{n^2} = LG_1(3,2) - LG_1(3,1)^2 = 0.00449392318179080474..., \\
\lim_{n \to \infty} \frac{L\nu_{n,3}}{n} = \xi_3 = 0.07584372316630152789..., \\
\lim_{n \to \infty} \frac{L\vartheta_{n,3}}{n} = 0, \\
\lim_{n \to \infty} \frac{s\mu_{n,3}}{\ln(n)^3} = \frac{e^{-\gamma}}{6} = 0.09357658059448086163..., \\
\lim_{n \to \infty} \frac{s\sigma^2_{n,3}}{n\ln(n)^2} = sG_P(3,2) = 0.65360389945528404987....
\]

The final $n\ln(n)^2$ asymptotic is based on [7, 8].

Here [23] are numerical results for $r = 4$:
Table 5.3: Statistics for Permute, rank four \((a = 1)\)

| \(n\) | \(\tilde{L}\mu_{n,4}\) | \(\tilde{L}\sigma_{n,4}^2\) | \(\tilde{L}\nu_{n,4}\) | \(\tilde{L}\vartheta_{n,4}\) | \(\tilde{s}\mu_{n,4}\) | \(\tilde{s}\sigma_{n,4}^2\) |
|------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 1000 | 0.040353        | 0.001586        | 0.0260          | 0.0010          | 0.042215        | 0.118491        |
| 1500 | 0.040351        | 0.001585        | 0.0267          | 0.0007          | 0.041482        | 0.126180        |
| 2000 | 0.040349        | 0.001585        | 0.0265          | 0.0005          | 0.040987        | 0.131244        |
| 2500 | 0.040348        | 0.001585        | 0.0268          | 0.0004          | 0.040618        | 0.134938        |

as well as \(s\nu_{n,4} = 19\) for \(n > 1482\) and \(s\vartheta_{n,4} = 3\) for \(n > 666\). Also

\[
\lim_{n \to \infty} \frac{L\mu_{n,4}}{n} = LG_1(4,1) = 0.04034198873687046287..., \\
\lim_{n \to \infty} \frac{L\sigma_{n,4}^2}{n^2} = LG_1(4,2) - LG_1(4,1)^2 = 0.00158383677354017280..., \\
\lim_{n \to \infty} \frac{L\nu_{n,4}}{n} = \xi_4 = 0.02713839684981404992..., \\
\lim_{n \to \infty} \frac{L\vartheta_{n,4}}{n} = 0, \\
\lim_{n \to \infty} \frac{s\mu_{n,4}}{\ln(n)^4} = e^{-\gamma}/24 = 0.02339414514862021540..., \\
\lim_{n \to \infty} \frac{s\sigma_{n,4}^2}{n \ln(n)^3} = sG_P(4,2) = 0.21786796648509468329... .
\]

The final \(n \ln(n)^3\) asymptotic is based on [7, 8].

6. Mappings

Our modified Knuth & Trabb Pardo algorithm is unavailable in this setting, thus we turn to Heinz’s program. A general observation for \(2 \leq r \leq 4\) is \(L\vartheta_{n,r} = 0\) always. Here [24] are numerical results for \(r = 2\):

| \(n\) | \(L\mu_{n,2}\) | \(L\sigma_{n,2}^2\) | \(L\nu_{n,2}\) | \(s\mu_{n,2}\) | \(s\sigma_{n,2}^2\) | \(s\nu_{n,2}\) |
|------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 100  | 0.166817        | 0.019535        | 0.1300          | 0.680589        | 0.279032        | 0.1200          |
| 200  | 0.168100        | 0.019243        | 0.1400          | 0.718071        | 0.323910        | 0.0750          |
| 300  | 0.168642        | 0.019121        | 0.1433          | 0.737331        | 0.350358        | 0.0567          |
| 400  | 0.168959        | 0.019050        | 0.1450          | 0.749928        | 0.368810        | 0.0450          |

Table 6.1: Statistics for Map, rank two \((a = 1/2)\)

as well as \(s\nu_{n,2} = 19\) for \(n > 443\) and \(s\vartheta_{n,2} = 2\) for \(n > 433\). Let us elaborate on the latter statistic (because it seems surprising at first glance: an extended string of
0s abruptly switches to 2s). If \( \pi_r(k, n) \) denotes the probability that the \( r \)th smallest component of a random \( n \)-mapping has exactly \( k \) nodes, then

\[
\left\{ \pi_2(k, 432) \right\}_{k=0}^4 = \{0.0595400, 0.0532617, 0.0594378, 0.0477544, 0.0387585\},
\left\{ \pi_2(k, 433) \right\}_{k=0}^4 = \{0.0594720, 0.0532614, 0.0594373, 0.0477539, 0.0387581\},
\left\{ \pi_2(k, 434) \right\}_{k=0}^4 = \{0.0594044, 0.0532612, 0.0594369, 0.0477535, 0.0387576\},
\left\{ \pi_2(k, 435) \right\}_{k=0}^4 = \{0.0593369, 0.0532609, 0.0594365, 0.0477530, 0.0387571\}.
\]

The maximum probability clearly is at \( k = 0 \) for \( n \leq 433 \) and then shifts to \( k = 2 \) for \( n \geq 434 \). Also

\[
\lim_{n \to \infty} \frac{L \mu_{n,2}}{n} = L G_{1/2}(2, 1) = 0.17090961985966239214..., \\
\lim_{n \to \infty} \frac{L \sigma_{n,2}^2}{n^2} = L G_{1/2}(2, 2) - L G_{1/2}(2, 1)^2 = 0.01862022330678138872..., \\
\lim_{n \to \infty} \frac{L \nu_{n,2}}{n} = 0.148..., \\
\lim_{n \to \infty} \frac{s \mu_{n,2}}{n^{1/2} \ln(n)} = \sqrt{2} S G_{1/2}(2, 1) = 2.06089224152016653900..., \\
\lim_{n \to \infty} \frac{s \sigma_{n,2}^2}{n^{3/2} \ln(n)} = \sqrt{2} S G_{1/2}(2, 2) = 1.40007638550124502818....
\]

No exact equation (akin to one involving \( \rho_r \) in Section 2) is known for the median of \( L \). An \( r \)th order Dickman function \( \rho_{r,1/2} \) of type \( a = 1/2 \) might be needed. What is responsible for mismatches between data and theory for \( S \)? This may be due to uncertainty about how the correction factor \( \sqrt{2} \) should be generalized from \( r = 1 \) to all \( r \geq 1 \). We believe that the sequence \( S \nu_{n,2} \) is bounded; a proof is not known.

Here \[25\] are numerical results for \( r = 3 \):

| \( n \) | \( L \bar{\mu}_{n,3} \) | \( L \bar{\sigma}_{n,3}^2 \) | \( L \bar{\nu}_{n,3} \) | \( s \bar{\mu}_{n,3} \) | \( s \bar{\sigma}_{n,3}^2 \) | \( s \bar{\nu}_{n,3} \) |
|---|---|---|---|---|---|---|
| 100 | 0.044147 | 0.003902 | 0.126620 | 0.052261 | 0.0700 |
| 150 | 0.045094 | 0.003902 | 0.0067 | 0.133605 | 0.055079 | 0.0867 |
| 200 | 0.045642 | 0.003903 | 0.0100 | 0.138200 | 0.057284 | 0.0850 |
| 250 | 0.046008 | 0.003904 | 0.0120 | 0.141572 | 0.059120 | 0.0880 |

Table 6.2: Statistics for Map, rank three \((a = 1/2)\)

as well as \( s \nu_{n,3} = 24 \) for \( n > 275 \) and \( s \vartheta_{n,3} = 0 \) for \( n \leq 278 \) at least. Also

\[
\lim_{n \to \infty} \frac{L \mu_{n,3}}{n} = L G_{1/2}(3, 1) = 0.04889742536845958914..., \\
\lim_{n \to \infty} \frac{L \sigma_{n,3}^2}{n^2} = L G_{1/2}(3, 2) - L G_{1/2}(3, 1)^2 = 0.01862022330678138872..., \\
\lim_{n \to \infty} \frac{L \nu_{n,3}}{n} = 0.148..., \\
\lim_{n \to \infty} \frac{s \mu_{n,3}}{n^{1/2} \ln(n)} = 2 S G_{1/2}(3, 1) = 2.06089224152016653900..., \\
\lim_{n \to \infty} \frac{s \sigma_{n,3}^2}{n^{3/2} \ln(n)} = 2 S G_{1/2}(3, 2) = 1.40007638550124502818....
\]
\[
\lim_{n \to \infty} \frac{L_{G,n,3}^2}{n^2} = L_{G,1/2}(3,2) - L_{G,1/2}(3,1)^2 = 0.00392148747204257695..., \\
\lim_{n \to \infty} \frac{SG_{n,3}^2}{n^1/2 \ln(n)^2} = \sqrt{2} S_{G,1/2}(3,1) = 1.03044612076008326950..., \\
\lim_{n \to \infty} \frac{S_{\sigma_{n,3}^2}}{n^{3/2} \ln(n)^3} = \sqrt{2} S_{G,1/2}(3,2) = 0.70003819275062251409....
\]

The median of \(L\) is unknown and mismatches worsen. It is certainly possible that the sequence \(L_{\nu,n,3}\) might be bounded; the trend of \(S_{\nu,n,3}\) is ambiguous. There are presently insufficient data to render judgement.

Here \([26]\) are numerical results for \(r = 4\):

| \(n\) | \(L_{\mu,n,4}\) | \(L_{\sigma_{n,4}^2}\) | \(S_{\mu,n,4}\) | \(S_{\sigma_{n,4}^2}\) |
|------|----------------|----------------|----------------|----------------|
| 100  | 0.011968       | 0.000710       | 0.015300       | 0.007424       |
| 125  | 0.012324       | 0.000717       | 0.016032       | 0.007682       |
| 150  | 0.012585       | 0.000722       | 0.016606       | 0.007877       |
| 175  | 0.012787       | 0.000726       | 0.017077       | 0.008034       |

Table 6.3: Statistics for Map, rank four \((a = 1/2)\)
as well as \(L_{\nu,n,4} = 0, S_{\nu,n,4} = 0, S_{\theta,n,4} = 0\) for \(n \leq 183\) at least. Also

\[
\lim_{n \to \infty} \frac{L_{\mu,n,4}}{n} = L_{G,1/2}(4,1) = 0.01514572139988693564..., \\
\lim_{n \to \infty} \frac{L_{G,n,4}^2}{n^2} = L_{G,1/2}(4,2) - L_{G,1/2}(4,1)^2 = 0.00077636923173854484..., \\
\lim_{n \to \infty} \frac{S_{\mu,n,4}}{n^{1/2} \ln(n)^2} = \sqrt{2} S_{G,1/2}(4,1) = 0.34348204025336108983..., \\
\lim_{n \to \infty} \frac{S_{\sigma_{n,4}^2}}{n^{3/2} \ln(n)^3} = \sqrt{2} S_{G,1/2}(4,2) = 0.23334606425020750469....
\]

Again, the median of \(L\) is unknown and mismatches worsen. Although both sequences \(L_{\nu,n,4}\) and \(S_{\nu,n,4}\) seem to be bounded (only 0s observed), we sense that they are still in transience and substantially more data will be required to reach steady state.

7. Acknowledgements

I am indebted to Alois Heinz for providing the algorithms underlying \(p[n, \ell]\) and \(q[n, \ell]\), and to David Broadhurst for calculating \(L_{\tilde{\nu},n,2}, L_{\tilde{\nu},n,3}, L_{\tilde{\nu},n,4}, L_{\tilde{\theta},n,2}\) to high precision as \(n \to \infty\) (permutations only). Many thanks are owed to Jean-Francois Alcover for translating Heinz’s concise Maple code to a form I could understand. The volunteers who edit and maintain OEIS, the creators of Mathematica, as well as administrators of the MIT Engaging Cluster, earn my gratitude every day. A sequel to this paper will be released soon \([27]\).
REFERENCES
[1] S. R. Finch, Permute, Graph, Map, Derange. arXiv:2111.05720
[2] S. R. Finch, Rounds, Color, Parity, Squares, arXiv:2111.14487.
[3] S. W. Golomb and P. Gaal, On the number of permutations of n objects with greatest cycle length k, Adv. in Appl. Math. 20 (1998) 98–107; MR1488234.
[4] D. Panario and B. Richmond, Exact largest and smallest size of components, Algorithmica 31 (2001) 413–432; MR1855258.
[5] A. P. Heinz, private communications (2021).
[6] D. Panario and B. Richmond, Smallest components in decomposable structures: exp-log class, Algorithmica 29 (2001) 205–226; MR1887304.
[7] L. A. Shepp and S. P. Lloyd, Ordered cycle lengths in a random permutation, Trans. Amer. Math. Soc. 121 (1966) 340–357; MR0195117.
[8] T. Shi, Cycle lengths of θ-biased random permutations, B.S. thesis, Harvey Mudd College, 2014, http://scholarship.claremont.edu/hmc_theses/65/.
[9] P. Flajolet and A. M. Odlyzko, Random mapping statistics, Advances in Cryptology - EUROCRYPT ’89, ed. J.-J. Quisquater and J. Vandewalle, Lect. Notes in Comp. Sci. 434, Springer-Verlag, 1990, pp. 329–354; MR1083961.
[10] X. Gourdon, Combinatoire, Algorithmique et Géométrie des Polynômes, Ph.D. thesis, École Polytechnique, 1996.
[11] R. Arratia, A. D. Barbour and S. Tavaré, Logarithmic Combinatorial Structures: a Probabilistic Approach, Europ. Math. Society, 2003, pp. 21-24, 52, 87–89, 118; MR2032426.
[12] R. G. Pinsky, A view from the bridge spanning combinatorics and probability, arXiv:2105.13834.
[13] N. Pippenger, Random cyclations, Elec. J. Combin. 20 (2013) R9; arXiv:math/0408031 MR3139394.
[14] D. E. Knuth and L. Trabb Pardo, Analysis of a simple factorization algorithm, Theoret. Comput. Sci. 3 (1976) 321–348; also in Selected Papers on Analysis of Algorithms, CSLI, 2000, pp. 303-339; MR0498355.
[15] M. Omar, D. Panario, B. Richmond and J. Whitely, Asymptotics of largest components in combinatorial structures, *Algorithmica* 46 (2006) 493–503; MR2291966.

[16] S. R. Finch, Extreme prime factors, *Mathematical Constants II*, Cambridge Univ. Press, 2019, pp. 171–172; MR3887550.

[17] D. Broadhurst, Higher-order Dickman functions, private communications (2014).

[18] D. Broadhurst, Dickman polylogarithms and their constants, arXiv:1004.0519.

[19] J.-M. De Koninck, Sur les plus grands facteurs premiers d’un entier, *Monatsh. Math.* 116 (1993) 13–37; MR1239141.

[20] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A330858 and A333726.

[21] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A006231, A155521, A332851, A332906, A349979, and A349980.

[22] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000774, A332852, A332907, A350015, and A350016.

[23] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A332853, A332908, A350273, and A350274.

[24] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A350078 and A350079.

[25] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A350080 and A350081.

[26] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A350275 and A350276.

[27] S. R. Finch, Joint probabilities within random permutations, *forthcoming*.

Steven Finch  
MIT Sloan School of Management  
Cambridge, MA, USA  
steven.finch@harvard.edu