Eugenijus Manstavičius
Discrete limit laws for additive functions on the symmetric group


Persistent URL: [http://dml.cz/dmlcz/137471](http://dml.cz/dmlcz/137471)

**Terms of use:**

© University of Ostrava, 2005

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these Terms of use.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* [http://project.dml.cz](http://project.dml.cz)
Discrete limit laws for additive functions on the symmetric group

Eugenijus Manstavičius

Abstract. Inspired by probabilistic number theory, we establish necessary and sufficient conditions under which the numbers of cycles with lengths in arbitrary sets possess an asymptotic limit law. The approach can be extended to deal with the counts of components with the size constraints for other random combinatorial structures.

1. Introduction and Results

Many combinatorial and algorithmic problems involve permutations with the cycle length constraints. Assume that these constraints change when the order of the symmetric group increases. Then the following natural question arises: what is the asymptotic behavior of the number of such permutations? If a permutation is taken at random, this can be reformulated as the problem on the asymptotic value distribution of sequences of linear statistics defined in terms of the cycle structure vector. So far (see, for instance, [1] or [6]), numerous results were published on the number of all or different length cycles, short cycles, long cycles, cycles in a fortiori given interval et ceter. We now discuss a general case and establish conditions when arbitrary sequences of the cycle count functions possesses an asymptotic distribution. Until now, the results on the necessity of the conditions are pretty rare, nevertheless some progress has been achieved in the functional limit theorems (see [3] and [10]).

Let $S_n$ be the symmetric group and $\sigma \in S_n$ be a permutation having $k_j(\sigma) \geq 0$ cycles of length $j$, $1 \leq j \leq n$. The structure vector is defined as $k(\sigma) := (k_1(\sigma), \ldots, k_n(\sigma))$. It satisfies the relation

$$1k_1(\sigma) + \cdots + nk_n(\sigma) = n. \quad (1)$$

Received: July 26, 2005.
1991 Mathematics Subject Classification: 60C05, 05D40.
Key words and phrases: random permutation, cycle structure, Poisson distribution, factorial moment.
Set $\nu_n(\ldots) = (n!)^{-1}\#\{\sigma \in S_n: \ldots\}$ for the probability measure on $S_n$ and $E_n f(\sigma)$ for the mean-value of a function $f(\sigma)$ on $S_n$ with respect to this frequency. If $\xi_j$, $j \geq 1$ are independent Poisson random variables (r.vs) given on some probability space $\{\Omega, F, P\}$ with the parameter $E\xi_j = 1/j$, then (see [1], formula (1.15))

$$
\nu_n(\tilde{k}(\sigma) = \tilde{k}) = \prod_{j \leq n} \frac{1}{j^{k_j} k_j!} = P\left((\xi_1, \ldots, \xi_n) = \tilde{k} \mid \sum_{j \leq n} j \xi_j = n\right),
$$

where $\tilde{k} \in \mathbb{Z}_r^n$, and $(k_1(\sigma), \ldots k_n(\sigma), 0, \ldots)^{\nu_n}(\xi_1, \ldots, \xi_n, \xi_{n+1}, \ldots)$ in the sense of convergence of the finite dimensional distributions. Here and in what follows we assume that $n \to \infty$. Despite to that, dealing with the asymptotic value distribution of the linear combinations

$$
h_n(\sigma) := a_{n1} k_1(\sigma) + \cdots + a_{nn} k_n(\sigma), \quad a_{nj} \in \mathbb{R}
$$

we face a lot of obstacles. The main reason is dependence of the summands arising from relation (1). The first case of (3) with $a_{nj} = 1/\log n$ describing the normalized number of cycles of $\sigma$ was examined in the paper [5]. Using an analytic method, the author of the present remark established the central limit theorem under the Lindeberg type condition and some more general limit theorems (see [8]). Section 8.5 of the recent book [1] exposes a general probabilistic approach similar to that cultivated in probabilistic number theory (see [7]). Comparing the results obtained for functions $h_n(\sigma)$ on permutations with that achieved for the additive number theoretical functions (see [4]) we see that, on this path, probabilistic number theory is far ahead from combinatorics.

For permutations, the general problem can be formulated as follows:

Under what conditions the frequencies $V_n(x; h_n, \alpha) := \nu_n(h_n(\sigma) - \alpha(n) < x)$ with some $\alpha(n) \in \mathbb{R}$ weakly converge to a limit distribution law?

Only in the case of degenerated limit law we have (see [12]) the final answer.

To formulate this result, we set $x^* = \min\{1, |x|\} \text{ sign } x$.

**Theorem 1.1** ([12]). Let $h_n(\sigma)$ be defined in (3). The frequencies $V_n(x; h_n, \alpha)$ weakly converge to the degenerated at the point $x = 0$ distribution function if and only if

$$
U_n(h, \lambda) := \sum_{j \leq n} \frac{(a_{nj} - \lambda j)^2}{j} = o(1)
$$

and

$$
\alpha(n) = n\lambda + \sum_{j \leq n} \frac{(a_{nj} - \lambda j)^*}{j} + o(1)
$$

for some sequence $\lambda = \lambda_n \in \mathbb{R}$.

In general, the formulated problem, especially its necessity part, seems to be difficult. Proving sufficiency, one can first apply the total variation estimate

$$
\frac{1}{2} \sum_{k_1, \ldots, k_r \geq 0} |\nu_n(k_1(\sigma) = k_1, \ldots, k_r(\sigma) = k_r) - P(\xi_1 = k_1, \ldots, \xi_r = k_r)| = o(1)
$$
established in [1] for \( r = o(n) \). In the next step, to show that \( a_{n, [n]} k_{[n]}(\sigma) + \cdots + a_{nn} k_n(\sigma) \) tends to zero in probability \( \nu_n \) for each \( 0 < \varepsilon < 1 \), one can use author’s inequality (see [9])

\[
\nu_n (|h_n(\sigma) - x| \geq y) \leq 32e^{2} P (|a_{n1} \xi_1 + \cdots + a_{nn} \xi_n - x| \geq y/3),
\]

valid for arbitrary \( a_{nj}, x \in \mathbb{R} \) and \( y \geq 0 \). This has proved to be useful (see [2] or [3]). The analytic approach exposed in paper [8] goes even further, nevertheless it requires some regularity of \( a_{nj} \) for large \( j \).

The present remark is based upon a few ideas originated in probabilistic number theory, especially upon those proposed by J. Šiaulys in [13]. For brevity, we now examine the case when \( a_{nj} \in \{0, 1\} \). Then \( h_n(\sigma) \) just counts the cycles in sequences of subsets \( J_n: = \{j \leq n: a_{nj} = 1\} \). The first attempt [12] to investigate the case when the Poisson limit law \( \Pi_\mu(x) \) with parameter \( \mu > 0 \) appears for \( V_n(x) := V_n(x; h_n, 0) \) gave an unexpected phenomenon. In what follows we add the star over the sums to replace the condition \( a_{nj} = 1 \).

**Theorem 1.2** ([12]). Let \( h_n(\sigma) \) be defined in (3) with \( a_{nj} \in \{0, 1\} \). The frequencies \( V_n(x) \) weakly converge to \( \Pi_\mu(x) \) if and only if

\[
\sum_{j \leq n}^{*} \frac{1}{j} = \mu + o(1) \quad (8)
\]

and

\[
\sum_{\varepsilon_n < j \leq n}^{*} \frac{1}{j} = o(1) \quad (7)
\]

for each fixed \( 0 < \varepsilon < 1 \).

The necessary condition (7) implies that counting the cycles with lengths in \( [\varepsilon_n, n] \) we can not obtain the Poisson law. By virtue of \( a_{nj} \in \{0, 1\} \), the sum \( a_{n1} \xi_1 + \cdots + a_{nn} \xi_n \) is also the Poisson random variable. It converges in distribution to \( \Pi_\mu(x) \) if and only if condition (6) holds. So, condition (7) is the price we are paying for dependence of the random variables \( k_j(\sigma), j \leq n \). On the other hand (see [10]), the Poisson law can be the limit for a sequence \( h_n(\sigma) \) defined via unbounded sequence \( a_{nj} \) if \( j \) runs only the integers of the interval \( (n/2, n] \). Such possibility is contained in Theorem 1.4 below.

If \( a_{nj} \in \{0, 1\} \), the linear statistics \( h_n(\sigma) \) possesses rather simple expressions for the factorial moments \( E_n h_n(\sigma)(x) \), where \( x(s) := x(x-1) \cdots (x-s+1) \). The next result is based on their analysis. Denote

\[
\gamma_{ns} = \sum_{j_1 \leq n}^{*} \frac{1}{j_1} \sum_{j_2 \leq n}^{*} \frac{1}{j_2} \cdots \sum_{j_s \leq n}^{*} \frac{1}{j_s} \left\{ j_1 + j_2 + \cdots + j_s \leq n \right\}.
\]

**Theorem 1.3.** Let \( h_n(\sigma) \) be defined in (3) with \( a_{nj} \in \{0, 1\} \). The frequencies \( V_n(x) \) weakly converge to a limit law if and only if there exist finite limits

\[
\lim_{n \to \infty} \gamma_{ns} = \gamma_s \quad (8)
\]
for each \( s \in \mathbb{N} \). Moreover, if (8) is satisfied, the characteristic function of the limit distribution is
\[
1 + \sum_{s=1}^{\infty} \frac{\gamma_s}{s!} (e^{it} - 1)^s, \quad t \in \mathbb{R}.
\]

The next problem would be to describe the class of the limit laws appearing in Theorem 1.3. We just observe that it contains the degenerated at zero and arbitrary Poisson law. Moreover, the factorial moments of each limit law satisfy the inequalities
\[
\gamma_s \leq \gamma_r \gamma_{s-r}
\]
for each \( 1 \leq r \leq s - 1 \) and \( s \geq 2 \). Recall that the Poisson law can be reckoned by the relation \( \gamma_s = \gamma_1^m \) for each \( s \geq 1 \). What are the conditions for the convergence of \( V_n(x) \) to a law \( V(x) \) which factorial moment \( \gamma_s \) satisfies the last formula only for some \( s = m \geq 2 \)?

Corollary 1. The weak convergence of \( V_n(x) \) to a law \( V(x) \) with the property \( \gamma_m = \gamma_1^m \) for some \( m \geq 2 \) occurs if and only if condition (6) of Theorem 1.2 with \( \gamma_1 = \mu \) and
\[
\sum_{n/m < j \leq n} \frac{1}{j} = o(1)
\]
are satisfied and there exist finite limits
\[
\lim_{n \to \infty} \sum_{j_1 \leq n/m} \frac{1}{j_1} \sum_{j_2 \leq n/m} \frac{1}{j_2} \cdots \sum_{j_s \leq n/m} \frac{1}{j_s} 1\{j_1 + j_2 + \cdots + j_s \leq n\} = \gamma_s
\]
for each \( s = m + 1, m + 2, \ldots \).

The corollary generalizes Theorem 1.2. Condition (10) shows that, in the case of limit law satisfying \( \gamma_m = \gamma_1^m \) for some \( m \geq 2 \), the main role is played only by the cycles with lengths not exceeding \( n/m \). Going along this path, we can characterize the case of limit distribution with a finite support.

Corollary 2. Assume that \( V(x) \) is a distribution function of a r.v. taking values \( 0, 1, \ldots, m - 1 \). Then \( V_n(x) \) weakly converges to \( V(x) \) if and only if
\[
\sum_{j \leq n/m} \frac{1}{j} = o(1)
\]
and the finite limits
\[
\lim_{n \to \infty} \sum_{n/m < j_1 \leq n} \frac{1}{j_1} \sum_{n/m < j_2 \leq n} \frac{1}{j_2} \cdots \sum_{n/m < j_s \leq n} \frac{1}{j_s} 1\{j_1 + j_2 + \cdots + j_s \leq n\} = \gamma_s
\]
exist for each fixed \( s \in \mathbb{N} \).

Now, the short cycles are negligible by virtue of condition (11). Finally, to give some impression what happens in the case of \( h_n(\sigma) \) defined in (3) with unbounded \( a_{nj} \in \mathbb{Z} \), without proof we include the next result.
Set
\[
\begin{align*}
a_{nj}^{(m)} &= \begin{cases} 
a_{nj} & \text{if } 0 \leq a_{nj} \leq m, \\
m & \text{if } a_{nj} > m, \\
0 & \text{if } a_{nj} < 0
\end{cases}
\end{align*}
\]
and \(h_n^m(\sigma) := a_{n1}^{(m)}k_1(\sigma) + \cdots + a_{nj}^{(m)}k_n(\sigma)\).

Theorem 1.4. The distribution functions \(V_n(x)\) weakly converge to the Poisson limit law \(\Pi_\mu(x)\) if and only if
\[
\sum_{j \leq n \atop a_{nj} \leq -1} \frac{1}{j} = o(1), \quad \lim_{m \to \infty} \limsup_{n \to \infty} \sum_{j \leq n \atop a_{nj} \geq m} \frac{1}{j} = 0,
\]
and
\[
\lim_{m \to \infty} \limsup_{n \to \infty} E_n h_n^m(\sigma)(s) = \lim_{m \to \infty} \liminf_{n \to \infty} E_n h_n^m(\sigma)(s) = \mu^s
\]
for each fixed \(s \in \mathbb{N}\).

Now the expressions of the factorial moments in (13) become fairly complicated, nevertheless their analysis is not hopeless.

2. Auxiliary Lemmas

Lemma 1 ([1]). For arbitrary natural numbers \(j_1 < \cdots < j_m\) and \(l_1, \ldots, l_m\), we have
\[
E_n(k_{j_1}(\sigma)_{l_1} \cdots k_{j_m}(\sigma)_{l_m}) = \mathbf{1}\{j_1l_1 + \cdots + j_ml_m \leq n\} E(\xi_{j_1(l_1)} \cdots \xi_{j_m(l_m)}).
\]

For brevity, we use \(\ll\) as an analog of \(O(\cdot)\).

Lemma 2. Let \(h(\sigma) := h_n(\sigma)\) be as in (3) with \(a_j := a_{nj} \in \{0, 1\}\), \(\gamma_{nl}\) be defined before Theorem 1.3, and \(s \in \mathbb{N}\). Then
\[
E_n h(\sigma)(s) = \gamma_{ns} \leq \gamma_{n1}.
\]

Proof. Let \(z \in \mathbb{C}, |z| \leq 2\), and \(\ell_n(\bar{k}) := 1k_1 + \cdots + mk_m\), where \(\bar{k} = (k_1, \ldots, k_m) \in \mathbb{Z}^+^m\) and \(0 \leq m \leq n\). Grouping the permutations with the same cycle structure and using (3) together with (2), we have
\[
E_n z^h(\sigma) = \frac{1}{n!} \sum_{\sigma \in S_n} z^h(\sigma) = \sum_{\ell_n(\bar{k}) = n} \prod_{j=1}^n z^{a_j} \nu_n(\bar{k}(\sigma) = \bar{k})
= \sum_{\ell_n(\bar{k}) = n} \prod_{j=1}^n \left( \frac{z^{a_j}}{j} \right)^{k_j} \frac{1}{k_j!},
\]
where the summation is extended over \(\bar{k} \in \mathbb{Z}^+^n\) with the property \(\ell_n(\bar{k}) = n\). Set
\[
g_m(z) = \sum_{\ell_m(\bar{k}) = m} \prod_{j=1}^m \left( \frac{z^{a_j}}{j} \right)^{k_j} \frac{1}{k_j!}.
\]
where \( g_0(z) \equiv 1 \) and \( 0 \leq m \leq n \). If \( w \in \mathbb{C} \), \(|w| \leq 1/2\) is another variable, then
\[
\exp \left\{ \sum_{j=1}^{\infty} \frac{z^{a_j}}{j} w^j \right\} = \prod_{j=1}^{\infty} \sum_{k=0}^{\infty} \left( \frac{z^{a_j}}{j} \right)^k \frac{1}{k!} w^j = \sum_{m=0}^{\infty} g_m(z)w^m.
\]
These series as well as the series involving the derivative with respect to \( z \) are uniformly convergent if \(|z| \leq 2\) and \(|w| \leq 1/2\) therefore
\[
\sum_{j=1}^{\infty} \frac{a_j}{j} z^{a_j-1} w^j \sum_{m=0}^{\infty} g_m(z)w^m = \sum_{n=0}^{\infty} g'_n(z)w^n.
\]
Hence
\[
g'_n(z) = \sum_{j=1}^{n} \frac{a_j}{j} z^{a_j-1} g_{n-j}(z) = \sum_{j=1}^{n} \frac{1}{j} g_{n-j}(z)
\]
and
\[
g^{(s)}_n(z) = \sum_{j=1}^{n} \frac{1}{j} g^{(s-1)}_{n-j}(z) \tag{15}
\]
for each \( s \geq 1 \). Using this equality and the agreement \( \gamma_{m0} := 1 \) for each \( m \geq 0 \), by induction we derive the following formula
\[
g^{(s)}_m(1) = \gamma_{ms}, \tag{16}
\]
where \( 0 \leq m \leq n \) and \( s \geq 0 \). For the induction parameter, we take \( r = m + s \geq 0 \). The well known Cauchy’s identity gives the value \( g_m(1) = 1 \) for each \( m \geq 0 \). Hence \( g^{(0)}_m(1) = g_m(1) = \gamma_{m0} \), confirming the first step of induction.

If \( g_{n-j}^{(s-1)}(1) = \gamma_{n-j,s-1} \), where \( 1 \leq j \leq n \) and \( s \geq 1 \), is already established, then by (15)
\[
g^{(s)}_n(1) = \sum_{j=1}^{n} \frac{1}{j} g_{n-j,s-1} = \sum_{j=1}^{n} \frac{1}{j} \sum_{j_1 \leq j} \cdots \sum_{j_{s-1} \leq j} \frac{1}{j_{s-1}} 1\{j_1 + \cdots + j_{s-1} \leq n \}.
\]
This is the equality in (16) for \( n = m \). By the definitions \( g^{(s)}_n(1) = E_n h(\sigma)^{(s)} \), consequently, the equality in (14) is proved. Observe that \( \gamma_{ns} \leq \gamma_{nr} \gamma_{n,s-r} \) for each \( 1 \leq r \leq s-1 \) and \( s \geq 2 \). Hence we obtain the inequality in (14). The lemma is proved.

The main analytic ingredient is an estimate for the concentration function. For the function \( h(\sigma) \) defined via \( a_j \), \( 1 \leq j \leq n \), we denote
\[
Q_n(u) = \sup_{x \in \mathbb{R}} \nu_n(|h(\sigma) - x| < u), \quad u \geq 0,
\]
and \( a_j(\lambda) = a_j - j\lambda \). Set
\[
D_n(u; \lambda) = \sum_{j=1}^{n} \min \left\{ u^2, a_j(\lambda)^2 \right\}, \quad D_n(u) = \min_{\lambda \in \mathbb{R}} D_n(u; \lambda).
\]
Lemma 3 ([11]). We have
\[ Q_n(u) \ll u D_n(u)^{-1/2} \] (17)
with an absolute constant in \( \ll \).

3. Proofs

Proof of Theorem 1.3. Sufficiency. Condition (8) of Theorem 1.3 implies
\[ \gamma_{n1} = \sum_{j \leq n} \frac{1}{j} \to \gamma_1 < \infty. \]
Hence by (14) the factorial moments \( \gamma_{ns} \leq C_1^s \) for some constant \( C_1 > \gamma_1 \) if \( n \) is sufficiently large. Further we use (8) and the expansion
\[ E_n e^{ith_n(\sigma)} = 1 + \sum_{i=1}^{L} \frac{\gamma_{ns}}{s!} (e^{it} - 1)^s + O \left( \frac{\gamma_{nL+1}}{(L+1)!} |e^{it} - 1|^{L+1} \right) \]
\[ = 1 + \sum_{i=1}^{L} \frac{\gamma_i}{i!} (e^{it} - 1)^i + O \left( \frac{(2C_1)^L}{(L+1)!} \right) + o_L(1), \]
where either of the estimates is uniform in \( t \in \mathbb{R} \) and the second one depends on \( L \geq 1 \). Taking now \( n \to \infty \) and later \( L \to \infty \), we complete the proof of convergence of \( V_n(x) \) and obtain the formula of the characteristic function of the limit law.

Necessity. Let \( V_n(x) \) weakly converge to a limit distribution \( V(x) = P(\xi < x) \), where \( \xi \) is a random variable taking values in the set \( \mathbb{Z}^+ \). Hence for the concentration function, we obtain
\[ Q_n(1) \gg \max_{m \in \mathbb{Z}^+} P(\xi = m) \geq c > 0, \]
where the constant \( c \) depends at most on \( \xi \). In what follows we disregard such dependence. Thus, by (17), we have \( D_n(1; \lambda) \ll 1 \) with some \( \lambda = \lambda_n \in \mathbb{R} \). By virtue of \( a_{nj} \in \{0, 1\} \), this leads to
\[ 1 \gg \sum_{j \leq n} \frac{(|\lambda|j - 1)^2}{j} \gg \sum_{2/|\lambda| \leq j \leq n} \frac{1}{j} \gg \log(n|\lambda|) - 1 \]
if \( \lambda \neq 0 \). Hence \( |\lambda| \leq C/n \) with some \( C > 0 \). Using this and \( (x + y)^2 \leq 2(x^2 + y^2) \), we obtain
\[ \gamma_{n1} = D_n(1; 0) \ll D_n(1; \lambda) + \sum_{j \leq n} \frac{(\lambda j)^2}{j} \ll 1 + \sum_{j \leq n} \frac{(Cj/n)^2}{j} \]
\[ \ll 1 + \frac{C_2^2}{n^2} \sum_{j \leq n/C} j + \sum_{n/C < j \leq n} \frac{1}{j} \leq C_2 < \infty. \]

As in the previous part, using (14), we see that \( \sup_n E_n h_n(\sigma)(s) \leq C_2^s \) for each \( s \geq 1 \). Consequently, from the weak convergence of \( V_n(x) \) we obtain convergence of the factorial moments. Applying (14), we complete the proof of (8). Theorem 1.3 is proved.
Proof of Corollary 1. Under its conditions by virtue of (9) we also have $\gamma_s = \gamma_s^1 = \mu_s$ for each $s \leq m$. Thus, they imply (8), hence the sufficiency part follows from Theorem 1.3.

If the limit distribution $V(x)$ exists, then condition (8) is satisfied. Observe that

$$\gamma_{n1} - \gamma_{nm} = \left( \sum_{j \leq n} \frac{1}{j} \right)^m - \sum_{j_1 \leq n} \frac{1}{j_1} \cdots \sum_{j_m \leq n} \frac{1}{j_m} \cdot 1\{j_1 + \cdots + j_m \leq n\} \geq \left( \sum_{n/m < j \leq n} \frac{1}{j} \right)^m.$$  

Now the property of the limit law implies (10). Combining it with (8), we complete the proof of the corollary.

Proof of Corollary 2. If the limit law $V(x)$ exists and is concentrated in the set $\{0, 1, \ldots, m - 1\}$, its factorial moment $\gamma_m = 0$. Consequently, from (8) we obtain

$$o(1) = \gamma_{nm} \geq \left( \sum_{j \leq n/m} \frac{1}{j} \right)^m.$$  

This is condition (11). Further using it, we reduce (8) to the form (12).

The sufficiency is trivial. The corollary is proved.

Acknowledgement. The author cordially thanks an anonymous referee whose comments have helped to improve the exposition of the paper.

References


*Author(s) Address(es):*

DEPARTMENT OF MATHEMATICS AND INFORMATICS, VILNIUS UNIVERSITY, NAUGARDOKO STR. 24, LT-03225 VILNIUS, LITHUANIA

E-mail address: eugenijus.manstavicius@maf.vu.lt