

Limiting processes with dependent increments for measures on symmetric group of permutations

Gutti Jogesh Babu, Eugenijus Manstavičius
and
Vytas Zacharovas

Abstract.

A family of measures on the set of permutations of the first n integers, known as *Ewens sampling formula*, arises in population genetics. In a series of papers, the first two authors have developed necessary and sufficient conditions for the weak convergence of a partial sum process based on these measures to a process with independent increments. Under very general conditions, it has been shown that a partial sum process converges weakly in a function space if and only if a related process defined through sums of independent random variables converges. In this paper, a functional limit theory is developed where the limiting processes need not be processes with independent increments. Thus, under Ewens sampling formula, the limiting process of the partial sums of dependent variables differs from that of the associated process defined through the partial sums of independent random variables.

§1. Introduction

In a series of papers [3]-[7] and [16], Babu and Manstavičius have developed functional limit theorems for partial sum processes defined on

Received January 12, 2006.

Revised June 26, 2006.

2000 *Mathematics Subject Classification*. Primary 60C05; Secondary 60F17, 11K65.

Key words and phrases. Cycle, Ewens sampling formula, functional limit theorem, random partitions, tightness, slowly varying function, weak convergence.

Research of Babu was supported in part by NSF grants DMS-0101360 and AST-0434234. Research of Manstavičius was supported in part by National Research Council's Twinning Fellowship. Research of Zacharovas was supported in part by a Lithuanian Science and Studies Foundation stipend.

random permutations. They have obtained necessary and sufficient conditions for the weak convergence of partial sums of dependent variables with respect to the Ewens sampling formula (1.1) to a process with independent increments. In contrast to the previous results, in this paper we examine the cases where the limit processes have dependent increments.

The family of probability measures on the symmetric group \mathbf{S}_n of permutations on $\{1, \dots, n\}$, considered in these papers, is closely related to the Ewens sampling formula (see [12]) in population genetics. The measures are given by

$$(1.1) \quad \nu_{n,\theta}(\bar{k}) := \frac{n!}{\theta_{(n)}} \prod_{j=1}^n \left(\frac{\theta}{j}\right)^{k_j} \frac{1}{k_j!}$$

for the partition

$$(1.2) \quad n = 1k_1 + \dots + nk_n, \quad n \in \mathbf{N},$$

and 0 otherwise, where $\bar{k} := (k_1, \dots, k_n) \in \mathbf{Z}^{+n}$, $\theta > 0$, and $\theta_{(n)} = \theta(\theta+1)\dots(\theta+n-1)$. The quantity $\nu_{n,\theta}(\bar{k})$ can also be viewed as the probability measure of the class of conjugate elements $\sigma \in \mathbf{S}_n$, all having $k_j(\sigma) = k_j$ cycles of length j , $1 \leq j \leq n$. The probability measure $\nu_{n,\theta}$ is induced by the measure $\nu'_{n,\theta}$ on \mathbf{S}_n , that assigns a mass proportional to $\theta^{w(\sigma)}$ for $\sigma \in \mathbf{S}_n$, where $w(\sigma) = k_1(\sigma) + \dots + k_n(\sigma)$ denotes the total number of cycles of σ . This can be seen from

$$\nu'_{\theta}(\sigma) = \theta^{w(\sigma)} \left(\sum_{\tau \in \mathbf{S}_n} \theta^{w(\tau)} \right)^{-1} = \frac{\theta^{w(\sigma)}}{\theta_{(n)}}.$$

Thus, we use this probability measure on \mathbf{S}_n and leave the same notation $\nu_{n,\theta}$ for it. The case $\theta = 1$ corresponds to the uniform probability on \mathbf{S}_n (Haar measure).

It is well known that the asymptotic distribution, as $n \rightarrow \infty$, of $k_j(\sigma)$ for a fixed $j \geq 1$ under $\nu_{n,\theta}$ is Poisson with parameter θ/j . Relation (1.2) makes $k_j(\sigma)$, $1 \leq j \leq n$ a dependent sequence. Nevertheless, the asymptotic value distribution problems of the sums

$$(1.3) \quad h(\sigma) = \sum_{j=1}^n h_j(k_j(\sigma)),$$

called additive functions, with respect to $\nu_{n,\theta}$ have been studied extensively, where $h_j(k)$ is a real double sequence, $k \geq 0$, $j \geq 1$ such that $h_j(0) = 0$ for each j . The first result in the case $\theta = 1$ for the process

defined via the number-of-cycles function $w(\sigma) = k_1(\sigma) + \dots + k_n(\sigma)$ was obtained by DeLaurentis and Pittel [10]. The case of general θ for the function $w(\sigma)$ was examined in [11] and [13]. A short proof of this result is given in Section 2.C of the paper by Arratia, Barbour and Tavaré [1]. Their recent book [2] contain rates of convergence. Convergence of more general partial sum processes to the Brownian motion was investigated by Babu and Manstavičius in [3]. It was shown that an analog of the Lindeberg condition is necessary and sufficient for the weak convergence of the processes. However, by constructing an example it was demonstrated that the Lindeberg condition is not necessary for the one dimensional central limit theorem. These results were extended in [6] to a class of infinitely divisible limit processes that include stable processes. The results were further extended in [7, 16] for arbitrary stochastically continuous limit processes with independent increments.

While discussing functional limit theory, the authors of [2] realize that this is complicated. To stress this point, note the comments following Theorem 8.33 on page 221 of [2]: “Even when (8.95) holds ... the limit theory is complicated. ... there is no universal approximation valid for a wide class of U_j sequences, as was the case with slow growth and Gaussian approximation. For example, take the case in which $EU_j^2 \sim cj^\alpha$ for some $\alpha > 0$ ” In this paper, we consider such complicated process when U_j are constants, instead of assuming them as random variables independent of $k_j(\sigma)$. It is shown that the limiting process need not be a process with independent increments.

We start with some notation and preliminary results in the next section. The main result and some illustrative examples are presented in section §3. The auxiliary lemmas needed in the proof of the main theorem are given in §4. The proof of the main theorem is given in §5.

§2. Preliminaries and notations

Let, as described above, $h_j(k)$ be a real double sequence, $k \geq 0$, $j \geq 1$ such that $h_j(0) = 0$ for each j . Set for brevity $a(j) = h_j(1)$, and $u^* = (1 \wedge |u|) \operatorname{sgn} u$, where $a \wedge b := \min\{a, b\}$. Throughout this paper the limits are taken as $n \rightarrow \infty$ and we assume that the normalizing sequences $\beta(n) > 0$ satisfy $\beta(n) \rightarrow \infty$. The weak convergence of processes (or their probability distributions) and convergence of bounded nondecreasing sequences of functions to a limit function at its continuity points is denoted by \Rightarrow . Together with the symbol $O(\cdot)$ we will use \ll

assuming the same meaning. Define

$$B(y, n) = \sum_{j \leq y} \left(\frac{a(j)}{\beta(n)} \right)^* \frac{1}{j}, \quad A(y, n) = \theta \sum_{j \leq y} \left(\frac{a(j)}{\beta(n)} \right)^* \frac{1}{j}$$

and

$$y(t) := y_n(t) = \max\{l \leq n : B(l, n) \leq tB(n, n)\}, \quad t \in [0, 1].$$

We consider the weak convergence of the process

$$(2.1) \quad \tilde{H}_n := \tilde{H}_n(\sigma, t) = \frac{1}{\beta(n)} \sum_{j \leq y(t)} h_j(k_j(\sigma)) - A(y(t), n), \quad t \in [0, 1]$$

under the measure $\nu_{n, \theta}$, in the space $\mathbf{D}[0, 1]$ endowed with the Skorohod topology (see [9]). As it has been shown in Lemma 5 of [7], if $\beta(n) \rightarrow \infty$, the process \tilde{H}_n and

$$(2.2) \quad H_n := H_n(\sigma, t) = \sum_{j \leq y(t)} a_n(j) k_j(\sigma) - A(y(t), n),$$

where $a_n(j) = a(j)/\beta(n)$, can converge weakly to a limit only simultaneously and the limits coincide. Therefore in what follows we examine only the process H_n .

The corresponding process X_n with independent summands is defined by

$$(2.3) \quad X_n := X_n(t) = \sum_{j \leq y(t)} a_n(j) \zeta_j - A(y(t), n), \quad t \in [0, 1],$$

where ζ_j are independent Poisson random variables with $\mathbf{E}(\zeta_j) = \theta/j$. The characteristic function φ_n^0 of $(X_n(t_1), \dots, X_n(t_s))$, for $0 \leq t_1 < \dots < t_s \leq 1$, is given by

$$\begin{aligned} \varphi_n^0(\lambda_1, \dots, \lambda_s) &= \mathbf{E} \left(\exp \left\{ i \sum_{r=1}^s \lambda_r X_n(t_r) \right\} \right) \\ &= \exp \left\{ \sum_{1 \leq j \leq n} \frac{\theta}{j} \left(e^{i a_n(j) \lambda_j^0} - 1 - i \lambda_j^0 (a_n(j))^* \right) \right\}, \end{aligned}$$

where $\lambda_j^0 = \sum_{r=1}^s \lambda_r \mathbf{1}\{j \leq y(t_r)\}$.

In general, the limiting behavior of the dependent process H_n is different from the corresponding X_n . However, if the normalizing sequence $\{\beta(n)\}$ is slowly varying, then it is possible to compare the limiting properties of H_n and X_n . This is presented below.

Theorem A (Babu and Manstavičius [7]). *In order that $H_n \Rightarrow X$, where X is a process with independent increments such that the distribution of $X(1)$ is non-degenerate, it is necessary and sufficient that the following two conditions are satisfied:*

- (I) *the sequence $\beta(n)$ is slowly varying in the sense of Karamata;*
- (II) *the sequence of functions*

$$\Psi_n(x) := \sum_{\substack{j \leq n \\ a_n(j) \leq x}} \frac{1}{j} a_n(j)^{*2}$$

converges weakly to some non-decreasing function Ψ defined on \mathbf{R} satisfying $0 = \Psi(-\infty) < \Psi(\infty) < \infty$, so that $\Psi_n(\pm\infty) \rightarrow \Psi(\pm\infty)$.

As discussed in [7] Condition II implies $X_n \Rightarrow X$, and under Condition I, $H_n \Rightarrow X$ if and only if $X_n \Rightarrow X$. We also observed in that paper that if

$$(2.4) \quad \sum_{j \leq n} \left(\frac{a(j)}{n^\varepsilon} \right)^{*2} \frac{1}{j} = o(1)$$

for each positive $\varepsilon > 0$, then the limit process of H_n , if it exists has independent increments. Thus, in order to model a limiting process with dependent increments we must avoid the relation (2.4). That can be achieved by taking larger $a(j)$ and normalizing sequences $\beta(n)$.

Stimulated by the investigations in probabilistic number theory [18], we have conjectured in [7] that the process H_n defined via the additive function in the counter examples constructed in [3] or via the functions with $a(j) = j^\rho$ with $\rho > 0$ (see [14]) might converge to processes with dependent increments. We now settle this affirmatively.

It should be stressed that the probabilistic approach, based on the approximation of dependent random variables by independent ones in the total variation distance, used in the earlier papers is not applicable now. Therefore we return to the analytic methods proposed in [14] in the case $\theta = 1$ and applied in [15] to estimate convergence rates. This approach works for $\theta > 1/2$ as well. Nevertheless, on this path, dealing with arbitrary $\theta > 0$ as in [17], we faced serious obstacles. However, in this paper we adopt a simpler version of this analytic approach.

§3. Results and examples

For the main theorem, we assume that

$$(3.1) \quad F_n(x) := \frac{1}{n} \#\{1 \leq j \leq n : a(j) < x\beta(n)\} \text{ and } F_n \Rightarrow F,$$

where $F(x)$ is a non-degenerate probability distribution function and

$$(3.2) \quad \beta(n)n^{-\rho} \text{ is slowly varying in the sense of Karamata,}$$

where $\rho > 0$ is a constant. Note that, under conditions of Theorem A, (3.1) holds with the degenerate probability distribution F with jump of size 1 at $x = 0$.

Throughout this paper, we use the notation \hat{P} to denote the characteristic function of a probability distribution function P . Thus $\hat{F}(\eta) = \int_{\mathbf{R}} e^{i\eta x} dF$.

For the time index function $y(t)$, it would be more natural to take the asymptotic solution to $B(y(t), n) \sim tB(n, n)$, nevertheless we change such time scale by a simple one to one map of $[0, 1]$ on to itself, and simply set

$$(3.3) \quad y(t) = tn, \quad 0 \leq t \leq 1.$$

We now state the main result.

Theorem 3.1. *Let H_n be the process defined in (2.2), where $\beta(n)$ and $y(t)$ are given by (3.2) and (3.3). If (3.1) and Condition II are satisfied, then the process H_n converges weakly in $\mathbf{D}[0, 1]$.*

A crucial part of the proof of Theorem 3.1 involves functions s_n and μ_n defined by

$$(3.4) \quad s_n(u, \eta; z) = \sum_{1 \leq j \leq un} \frac{1}{j} (e^{ia_n(j)\eta} - 1)(e^{-zj/n} - 1),$$

$$(3.5) \quad \mu_n(u, \eta) = \sum_{1 \leq j \leq un} \frac{1}{j} (e^{ia_n(j)\eta} - 1 - i\eta a_n(j)^*),$$

for $\Re z \geq 0$, $\eta \in \mathbf{R}$ and $0 \leq u \leq 1$, and the empty sums are taken to be equal to zero. We shall show in Lemma 4.4 that for each $T > 0$, $s_n(u, \eta; z) \rightarrow S(u, \eta; z)$ uniformly in $0 \leq u \leq 1$ and $0 \leq |\eta|$, $\Re z \leq T$, where

$$(3.6) \quad S(u, \eta; z) = uV(u)(\hat{F}(\eta u^\rho) - 1) - \int_0^u v(\hat{F}(\eta v^\rho) - 1)V'(v)dv,$$

and V' denotes the derivative of $V(x) = (e^{-zx} - 1)/x$ for $x > 0$. Note that

$$S(u, 0; z) = S(0, \eta; z) = 0,$$

Under Condition II, we shall establish in Lemma 4.5, that $\mu_n(u, \eta) \rightarrow \mu(u, \eta)$ for each $0 \leq u \leq 1$ and $\eta \in \mathbf{R}$, where

$$(3.7) \quad \mu(u, \eta) = \int_{\mathbf{R}} (e^{i\eta xu^\rho} - 1 - i\eta(xu^\rho)^*) x^{*-2} d\Psi(x).$$

If (3.3) holds, then the characteristic function φ_n^0 of the process X_n in (2.3) is given by

$$\varphi_n^0(\lambda_1, \dots, \lambda_s) = \exp \left\{ \theta \sum_{r=1}^s (\mu_n(t_r, \xi_r) - \mu_n(t_{r-1}, \xi_r)) \right\},$$

where $t_0 = 0$ and $\xi_r = \lambda_r + \dots + \lambda_s$ for $1 \leq r \leq s$. Thus φ_n^0 is the main multiplicative factor of φ_n in (5.1).

The next Corollary is useful in illustrating examples.

Corollary 1. *Let $\{d(j)\}$ be a sequence of real numbers and $\beta(n) = n^\rho$ for some constant $\rho > 0$. Suppose G_n given by*

$$G_n(x) = \frac{1}{n} \#\{1 \leq j \leq n : d(j) < x\}$$

converges weakly to some probability distribution function G . If Condition II holds with $a_n(j) = d(j)(j/n)^\rho$, $j \geq 1$, then H_n converges weakly in $\mathbf{D}[0, 1]$.

Moreover, in this case,

$$(3.8) \quad S(u, \eta; z) = \int_0^u (\hat{G}(\eta v^\rho) - 1)V(v) dv.$$

Proof. We shall show that $F_n \Rightarrow F$, where

$$(3.9) \quad F_n(x) = \frac{1}{n} \#\{j \leq n : d(j)(j/n)^\rho \leq x\} \text{ and } F(x) = \int_0^1 G(xu^{-\rho}) du.$$

The result then follows by Theorem 3.1.

Towards this goal, observe that for all $0 < u < 1$ and at all the continuity points x of G , the joint distribution function

$$(3.10) \quad \begin{aligned} R_n(x, u) &= \frac{1}{n} \#\{j \leq n : d(j) \leq x, (j/n) \leq u\} \\ &= \frac{1}{n} \#\{j \leq un : d(j) \leq x\} \rightarrow uG(x). \end{aligned}$$

Thus $R_n \Rightarrow UG$, the product of uniform measure and G . Hence by Fubini's Theorem [8, Theorem 18.3], it follows that $F_n \Rightarrow F$ with F defined in (3.9). The relation between the characteristic functions of F and G ,

$$(3.11) \quad \hat{F}(\eta) = \int_{\mathbf{R}} e^{i\eta x} dF(x) = \int_0^1 \hat{G}(\eta v^\rho) dv,$$

yields

$$(3.12) \quad u(\hat{F}(\eta u^\rho) - 1) = \int_0^1 u(\hat{G}(\eta u^\rho y^\rho) - 1) dy = \int_0^u (\hat{G}(\eta y^\rho) - 1) dy.$$

Substituting (3.12) in (3.6) and using integration by parts, we obtain

$$\begin{aligned} S(u, \eta; z) &= V(u) \int_0^u (\hat{G}(\eta v^\rho) - 1) dv - \int_0^u V'(v) \int_0^v (\hat{G}(\eta y^\rho) - 1) dy dv \\ &= \int_0^u V(v) (\hat{G}(\eta v^\rho) - 1) dv. \end{aligned}$$

Q.E.D.

Remark 1. Condition II in Corollary 1 holds if for some $\varepsilon > 0$,

$$(3.13) \quad D(x) = \frac{1}{x} \sum_{j \leq x} |d(j)|^\varepsilon \ll 1.$$

To prove this, without loss of generality we assume that (3.13) holds with $0 < \varepsilon < 1$. Let $\rho > 0$, $b > 1$ and note that

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n \left(\frac{n}{j} (d(j)^2 (j/n)^{2\rho})^* \right)^b &\leq \frac{1}{n} \sum_{j=1}^n \left(\frac{n}{j} (d(j)^2 (j/n)^{2\rho})^{*(\varepsilon/2b)} \right)^b \\ &\leq n^{b-1-\varepsilon\rho} \sum_{j=1}^n |d(j)|^\varepsilon j^{\varepsilon\rho-b}. \end{aligned}$$

If $1 < b < 1 + \varepsilon\rho$, then summation by parts yields,

$$(3.14) \quad \frac{1}{n} \sum_{j=1}^n \left(\frac{n}{j} (d(j)(j/n)^\rho)^{*2} \right)^b \ll D(n) + n^{b-1-\varepsilon\rho} \int_1^n D(u) u^{\varepsilon\rho-b} du \ll 1.$$

The inequality (3.14) assures uniform integrability of $(yu^\rho)^{*2} u^{-1}$ with respect to R_n . Hence, in view of (3.10) and (3.14), we have by [8,

Theorem 25.12] that

$$\begin{aligned} \Psi_n(x) &= \int_{(-\infty, x] \times (0, 1]} (yu^\rho)^{*2} u^{-1} dR_n(y, u) \\ &\rightarrow \int_{-\infty}^x \left(\int_0^1 (yu^\rho)^{*2} u^{-1} du \right) dG(y) \\ &= \frac{1}{\rho} \int_{-\infty}^x \left(\int_0^{|y|} \min(v, 1/v) dv \right) dG(y) \\ &= \frac{1}{2\rho} \int_{-\infty}^x \left(y^{*2} + 2 \log(\max(1, |y|)) \right) dG(y) =: \Psi(x), \end{aligned}$$

at all continuity points x of Ψ . The same arguments also establish $\Psi_n(\pm\infty) \rightarrow \Psi(\pm\infty)$.

We shall now provide examples to illustrate the results.

Example 1. Let $a(j) = j^\rho$, $\beta(n) = n^\rho$ for some constant $\rho > 0$. In this case $F(x) = x^{1/\rho}$ for $0 < x \leq 1$. By (3.11),

$$\hat{F}(\eta) = \int_0^1 e^{i\eta x} \left(\frac{1}{\rho} \right) x^{(1/\rho)-1} dx = \int_0^1 e^{i\eta y^\rho} dy.$$

In this case

$$S(u, \eta; z) = \int_0^u \frac{1}{v} \left(e^{i\eta v^\rho} - 1 \right) \left(e^{-zv} - 1 \right) dv.$$

This can also be treated using the Corollary 1 with $d(j) \equiv 1$. So G has degenerate distribution at 1. In this case, Condition II holds with $\Psi(x) = (1/2\rho) x^{*2}$ for $x > 0$, and $\Psi(x) = 0$ for $x \leq 0$. We also have

$$\mu(u, \eta) = \frac{1}{\rho} \int_0^1 \left(e^{i\eta x u^\rho} - 1 - i\eta x u^\rho \right) x^{-1} dx.$$

Remark 2. The limiting process of H_n in Theorem 3.1 need not be a process with independent increments. If $\rho = \theta = 1$ in Example 1, then we show below that the limiting process X of H_n is not a process with independent increments.

We first establish that $X(t)$ has non-degenerate limiting distribution, if $t > 1/2$. We use properties of the characteristic functions, in particular, the relation between the derivatives of the characteristic function and the moments. Fix $t > 1/2$ and let ϕ denote the characteristic

function of $X(1) - X(t)$. Since $A(y(t), n) = [tn]/n \rightarrow t$, where $[x]$ denotes the largest integer not exceeding x , it follows by Lemma 4.6 that

$$\phi(s) = e^{-is(1-t)} \left(1 + \int_t^1 \frac{1}{u} (e^{isu} - 1) du \right).$$

If the distribution corresponding to ϕ is degenerate at c , then ϕ has all the derivatives and $\phi'(0) = ic$, $\phi''(0) = -c^2$ (see [8, eq. (26.10)]). But

$$\phi'(s) = -i(1-t)\phi(s) + ie^{-is(1-t)} \left(\int_t^1 e^{isu} du \right).$$

So $\phi'(0) = 0$. Here we used Theorem 16.8 of [8], in taking the derivatives under the integral sign. Now the second derivative of ϕ is given by

$$\begin{aligned} \phi''(s) &= -i(1-t)\phi'(s) \\ &+ (1-t)e^{-is(1-t)} \left(\int_t^1 e^{isu} du \right) - e^{-is(1-t)} \left(\int_t^1 ue^{isu} du \right). \end{aligned}$$

So $\phi''(0) = (1-t)^2 - \frac{1}{2}(1-t^2) \neq 0$ as $t > 1/2$. Thus the distribution of $X(1) - X(t)$ is non-degenerate. As $H_n(\sigma, 1) = 0$ for all $\sigma \in \mathbf{S}_n$, we have $X(1) \equiv 0$. Hence the distribution of $X(t)$ is non-degenerate.

Again since $X(1) \equiv 0$, it follows that $X(t)$ and $X(1) - X(t)$ are independent if and only if $X(t)$ is independent of itself. So

$$0 = P(X(t) < x, X(t) > x) = P(X(t) < x)P(X(t) > x)$$

for all x . This is impossible as the distribution of $X(t)$ is shown to be non-degenerate. Consequently, the limiting process of H_n has dependent increments.

We apply Corollary 1 to the next example.

Example 2. Let $0 < \alpha < 2$. Let G denote the distribution function of the stable law with characteristic function ϕ_α given by, $\phi_\alpha(s) = e^{-|s|^\alpha}$. Define $\beta(n) = n^{1/\alpha}$ and

$$d(j) = \begin{cases} G^{-1}(\{j\sqrt{2}\}) & \text{if } |G^{-1}(\{j\sqrt{2}\})| \leq j^{1/\alpha} \\ 0 & \text{otherwise,} \end{cases}$$

where $\{x\}$ denotes the fractional part of x . Clearly G_n given by

$$G_n(x) = \frac{1}{n} \#\{1 \leq j \leq n : d(j) < x\}$$

converges weakly to G (see [4]). So the Corollary 1 is applicable.

It is shown in [4, Equation (12)] that (3.13) holds with $\varepsilon = \alpha/2$. So Condition II holds by Remark 1 with

$$\Psi(x) = \alpha \int_{-\infty}^x \left(\frac{1}{2} y^{*2} + \log(\max(1, |y|)) \right) dG(y).$$

In this case the function S is given by

$$S(u, \eta; z) = \int_0^u \frac{1}{v} \left(\phi_\alpha(\eta v^{1/\alpha}) - 1 \right) (e^{-zv} - 1) dv.$$

By using the symmetry $G(x) = 1 - G(-x)$ and arguments similar to the ones given in Remark 1, we can establish that $A(tn, n) \rightarrow 0$ as $n \rightarrow \infty$. For $\theta = 1$ and $\frac{1}{2} < t \leq 1$, it can now be deduced from (5.1), Lemmas 4.4 and 4.5 that

$$\mathbf{E}e^{i\eta X(t)} = e^{-|\eta|^\alpha} \left(1 + \int_0^{|\eta|^\alpha} \frac{1}{x} (e^x - e^{tx}) dx \right),$$

and from Lemmas 4.5 and 4.6 that

$$\mathbf{E}e^{i\eta(X(1)-X(t))} = 1 + \int_t^1 \frac{1}{u} (e^{-u|\eta|^\alpha} - 1) du.$$

§4. Auxiliary results

We now present preparatory results needed for the proof of Theorem 3.1. First, we consider the mean values of multiplicative functions $g: \mathbf{S}_n \rightarrow \mathbf{C}$ defined via

$$g(\sigma) = \prod_{j=1}^n f(j)^{k_j(\sigma)}, \quad 0^0 := 1,$$

where $|f(j)| \leq 1$, $j \geq 1$ are complex numbers, depending, maybe, on n or other parameters. Its mean value with respect to the measure $\nu_{n,\theta}$ equals

$$\begin{aligned} \left(\sum_{\sigma \in \mathbf{S}_n} \theta^{w(\sigma)} \right)^{-1} \sum_{\sigma \in \mathbf{S}_n} \theta^{w(\sigma)} g(\sigma) &= \frac{1}{\theta_{(n)}} \sum_{\sigma \in \mathbf{S}_n} \theta^{w(\sigma)} g(\sigma) \\ &= \frac{n!}{\theta_{(n)}} \sum_{\substack{k_1, \dots, k_n \geq 0 \\ 1k_1 + \dots + nk_n = n}} \prod_{j=1}^n \left(\frac{1}{j} \theta f(j) \right)^{k_j} \frac{1}{k_j!} \\ (4.1) \qquad \qquad \qquad &=: \frac{n!}{\theta_{(n)}} M_n. \end{aligned}$$

Set

$$L(z) = \sum_{j \leq n} \frac{1}{j} (f(j) - 1) z^j.$$

The asymptotic behavior of M_n is examined in the next Proposition, which is interest on its own. It generalizes the Main Lemma in [14].

Proposition. *Let $2 \leq K \leq \sqrt{n}$ be arbitrary. If*

$$(4.2) \quad -\Re L(1) \leq L < \infty,$$

then there exists $c = c(\theta) > 0$ such that

$$\frac{n!}{\theta_{(n)}} M_n = \frac{\Gamma(\theta)}{2\pi i} \int_{1-Ki}^{1+Ki} \frac{e^z}{z^\theta} \exp \left\{ \theta L \left(e^{-z/n} \right) \right\} dz + O(K^{-c}).$$

The constant in $O(\cdot)$ depends on θ and L only.

The proof of the proposition will be given in Appendix.

Let for brevity, $\mathcal{L}(I)$ be the linear space of real functions ℓ on $I \subset \mathbf{R}$ with $\sup_{t \in I} |\ell(t)| < \infty$.

Lemma 4.1. *Let $h(\sigma, t)$, $t \in I \subset \mathbf{R}$, be a set of real valued additive functions defined by (1.3) via $h_j(k, \cdot) \in \mathcal{L}(I)$, where $k \geq 0$, $h_j(0, t) = 0$, for $j \leq n$, $t \in I$, and $\Xi_n(t) = h_1(\xi_1, t) + \cdots + h_n(\xi_n, t)$. Then*

$$\nu_{n,\theta} \left(\sup_{t \in I} |h(\sigma, t) - a(t)| \geq u \right) \leq C(\theta) \left(P^{\theta \wedge 1} \left(\sup_{t \in I} |\Xi_n(t) - a(t)| \geq u/3 \right) + n^{-\theta} \right).$$

Here $u > 0$ and $a \in \mathcal{L}(I)$ are arbitrary and $C(\theta)$ is a positive constant depending only on θ .

A proof of Lemma 4.1 is given in [3]. Before stating the next lemma, define for $1 \leq l \neq m \leq n$ and $l + m \leq n$,

$$\nu_{n,\theta}^{l,m} := \nu_{n,\theta} (k_l(\sigma) \geq 1, k_m(\sigma) \geq 1) = \frac{n!}{\theta_{(n)}} \sum_{\substack{1k_1 + \cdots + nk_n = n \\ k_l, k_m \geq 1}} \prod_{j=1}^n \left(\frac{\theta}{j} \right)^{k_j} \frac{1}{k_j!}.$$

Lemma 4.2. *Suppose $1 \leq l \neq m \leq n$. If $l + m = n$, then*

$$\nu_{n,\theta}^{l,m} \ll \frac{1}{lm} n^{1-\theta}.$$

If $l + m \leq n - 1$, then

$$\nu_{n,\theta}^{l,m} \ll \frac{1}{lm} \left(1 - \frac{l+m}{n} \right)^{\theta-1}.$$

Thus for $1 \leq l \neq m \leq n$ and $l + m \leq n$,

$$(4.3) \quad lm \nu_{n,\theta}^{l,m} \ll \begin{cases} 1 & \text{if } \theta \geq 1 \\ n^{1-\theta} & \text{if } 0 < \theta < 1. \end{cases}$$

Proof. We use the formula for $\nu_{n,\theta}^{l,m}$ and observe that $1k_1 + \dots + nk_n = n - l - m$ leads to $k_j = 0$ for all $n - l - m + 1 \leq j \leq n$. If $l + m \leq n - 1$, then

$$\begin{aligned} \nu_{n,\theta}^{l,m} &= \frac{\theta^2 n!}{lm \theta_{(n)}} \sum_{1k_1 + \dots + nk_n = n-l-m} \prod_{j=1}^n \binom{\theta}{j}^{k_j} \frac{1}{k_j! (k_l + 1)(k_m + 1)} \\ &\leq \frac{\theta^2 n!}{lm \theta_{(n)}} \sum_{1k_1 + \dots + nk_n = n-l-m} \prod_{j=1}^{n-l-m} \binom{\theta}{j}^{k_j} \frac{1}{k_j!} \\ &= \frac{\theta^2 n!}{lm \theta_{(n)}} \frac{\theta_{(n-l-m)}}{(n-l-m)!} \ll \frac{1}{lm} \left(1 - \frac{l+m}{n}\right)^{\theta-1}, \end{aligned}$$

by the well known inequalities $n^{1-\theta} \ll n!/\theta_{(n)} \ll n^{1-\theta}$ valid for $n \geq 1$.

If $l + m = n$, then $k_l = k_m = 1$ and all other $k_j = 0$. So in this case

$$\nu_{n,\theta}^{l,m} = \frac{\theta^2 n!}{lm \theta_{(n)}} \ll \frac{1}{lm} n^{1-\theta}.$$

This completes the proof.

Q.E.D.

The next result is used in proving the tightness part of Theorem 3.1.

Lemma 4.3. *Let $0 < \delta < 0.02$ and $\mathcal{C} = \{(l, m) : \frac{1}{2}\sqrt{\delta}n \leq l < m \leq \min(n, l + 2\delta n)\}$. Then*

$$\nu_0 = \sum_{(l,m) \in \mathcal{C}} \nu_{n,\theta}^{l,m} \ll \sqrt{\delta} + n^{-\theta}.$$

Proof. Let for $r = 1, 2, 3$,

$$\nu_r = \sum_{(l,m) \in \mathcal{C}_r} \nu_{n,\theta}^{l,m},$$

where $\mathcal{C}_1 = \{(l, m) \in \mathcal{C} : l \leq n/3\}$, $\mathcal{C}_2 = \{(l, m) \in \mathcal{C} : l + m = n, \text{ and } l > n/3\}$, and $\mathcal{C}_3 = \{(l, m) \in \mathcal{C} : l + m \leq n - 1, \text{ and } l > n/3\}$.

If $\theta \geq 1$, then by Lemma 4.2, we have

$$(4.4) \quad \begin{aligned} \nu_0 &\ll \sum_{(l,m) \in \mathcal{C}} \frac{1}{lm} \leq \sum_{(1/2)\sqrt{\delta n} \leq l \leq n} \frac{1}{l} \log(1 + 2\delta n/l) \\ &\ll \delta n \sum_{\sqrt{\delta n} \leq l \leq n} l^{-2} \ll \sqrt{\delta}. \end{aligned}$$

Now suppose $0 < \theta < 1$. If $l \leq n/3$, $m \leq \frac{n}{3} + 2\delta n \leq \frac{n}{2}$, then $l + m \leq \frac{5}{6}n$ and hence by (4.3) $\nu_{n,\theta}^{l,m} \ll 1/(lm)$. Thus as in (4.4) we have $\nu_1 \ll \sqrt{\delta}$. Note that

$$\begin{aligned} \nu_2 &\ll n^{1-\theta} \sum_{(1/3)n < l < n-l \leq n} \frac{1}{l(n-l)} \ll n^{1-\theta} \sum_{(1/3)n < l} l^{-2} \ll n^{-\theta}, \\ \nu_3 &\ll \sum_{\substack{(1/3)n < l, l+m \leq n-1 \\ l < m \leq l+2\delta n}} \frac{1}{lm} \left(1 - \frac{l+m}{n}\right)^{\theta-1} \ll \frac{\delta}{n} \sum_{1 \leq j \leq (n/3)} \left(\frac{j}{n}\right)^{\theta-1} \ll \delta. \end{aligned}$$

The result now follows as $\nu_0 = \nu_1 + \nu_2 + \nu_3$.

Q.E.D.

Lemma 4.4. *Suppose $\beta(n)n^{-\rho}$ is slowly varying function for some constant $\rho > 0$ and (3.1) holds. Then for each $T > 0$, $s_n(u, \eta; z) \rightarrow S(u, \eta; z)$ uniformly in $0 \leq u \leq 1$, $|\eta| \leq T$, $\Re z \geq 0$ and $|z| \leq T$, where s_n and S are defined in (3.4) and (3.6).*

Proof. Let $K_n(t) = 0$ for $0 \leq t < 1$,

$$K_n(t) = \frac{1}{n} \sum_{1 \leq j \leq t} (e^{i\eta a_n(j)} - 1) \quad \text{for } t \geq 1.$$

For any fixed $\delta \in (0, 0.1)$, the properties of regularly varying functions give us

$$\beta(nv)/\beta(n) = v^\rho + o(1)$$

uniformly in $\delta \leq v \leq 1$. As $\hat{F}_n(\eta) \rightarrow \hat{F}(\eta)$ uniformly in $|\eta| \leq T$ (see [8, Exercise 26.15 (b)]),

$$\frac{1}{vn} \sum_{j \leq vn} e^{i\eta a_n(j)} = \hat{F}(\eta v^\rho) + o(1)$$

uniformly in $|\eta| \leq T$, and hence

$$(4.5) \quad K_n(vn) = v(\hat{F}(\eta v^\rho) - 1) + o(1),$$

uniformly in $|\eta| \leq T$ and $\delta \leq v \leq 1$. Thus for any $0 < \varepsilon < 1$ and for all $|\eta| \leq T$, there exists an $n_\varepsilon \geq 1$ such that

$$\sup_{(\varepsilon/8) < v \leq 1} |K_n(vn) - v(\hat{F}(\eta v^\rho) - 1)| < \frac{\varepsilon}{2},$$

for all $n \geq n_\varepsilon$. Since $|K_n(vn)| \leq v$ and $|\hat{F}(\eta v^\rho)| \leq 1$, for all $n \geq 1$, it follows that

$$|K_n(vn) - v(\hat{F}(\eta v^\rho) - 1)| < \frac{\varepsilon}{2} + \frac{3\varepsilon}{8} < \varepsilon,$$

for all $n \geq n_\varepsilon$. Hence (4.5) holds uniformly in $0 \leq v \leq 1$ and $|\eta| \leq T$.

Recall $V(x) = (e^{-zx} - 1)/x$ for $x > 0$ and note that the derivative of V is given by

$$V'(x) = -(e^{-zx} - 1 + zxe^{-zx})x^{-2}.$$

We now apply summation by parts to obtain, for $0 \leq u \leq 1$, that

$$s_n(u, \eta; z) = K_n(un)V(u) - \int_0^u K_n(vn)V'(v) dv.$$

Since $|e^{-z} - 1 + z| \leq \frac{1}{2}|z|^2$, $|e^{-z} - 1| \leq |z|$ for $\Re z \geq 0$, and hence $|V'(x)| \leq 2|z|^2$ for $x > 0$ and $\Re z \geq 0$, the lemma now follows as $K_n(vn) \rightarrow v(\hat{F}(\eta v^\rho) - 1)$ uniformly in $0 \leq v \leq 1$ and $|\eta| \leq T$, Q.E.D.

Lemma 4.5. *Suppose Condition II, (3.1) and (3.2) hold for some constant $\rho > 0$. Then $\mu_n(u, \eta) \rightarrow \mu(u, \eta)$ for each $0 < u \leq 1$ and $\eta \in \mathbf{R}$, where μ_n and μ are defined in (3.5) and (3.7).*

Proof. First note that

$$\mu_n(u, \eta) = \int_{\mathbf{R}} \gamma(xc_n)x^{*-2} d\Psi_{nu}(x) \quad \text{and} \quad \mu(u, \eta) = \int_{\mathbf{R}} \gamma(xu^\rho)x^{*-2} d\Psi(x),$$

where

$$\gamma(x) = e^{i\eta x} - 1 - i\eta x^*.$$

By (3.2) and (3.3), $c_n := \beta(nu)/\beta(n) \rightarrow u^\rho \in (0, 1]$, and hence $c_n < 2$ for all large n . So

$$\sup_x |\gamma(xc_n)|x^{*-2} \ll \sup_x (xc_n)^{*2}x^{*-2} \ll 1.$$

Since $|\gamma(xc_n) - \gamma(xu^\rho)|x^{*-2} \rightarrow 0$ uniformly on compact sets, the result now follows by Condition II. Q.E.D.

Lemma 4.6. *Let $g : \mathbf{S}_n \rightarrow \mathbf{C}$ be a multiplicative function defined via $f(j)$ such that $|f(j)| \leq 1$ and $f(j) = 1$ for all but $j \in J \subset (n/2, n]$. Then*

$$\frac{n!}{\theta_{(n)}} M_n = 1 + \theta \sum_{j \in J} \frac{1}{j} (f(j) - 1) \frac{n!}{\theta_{(n)}} \frac{\theta_{(n-j)}}{(n-j)!}.$$

A proof of the lemma is given in [3].

§5. Proof of Theorem 1

To establish the weak convergence of $H_n(\sigma, \cdot)$ to some limit process X in $\mathbf{D}[0, 1]$, it is sufficient to establish (see [9, Theorem 13.3]) convergence of finite dimensional distributions of H_n to those of X , and the tightness of the sequence of measures $\{\nu_{n, \theta} \cdot H_n^{-1}\}$. The first task will be achieved using analytical methods, and the second one will be established using the relevant criteria [9, §13], which requires a careful analysis of the influence of large cycles on $H_n(\sigma, t)$.

5.1. Convergence of finite dimensional distributions

For arbitrary $s \geq 1$ and $0 \leq t_1 < \dots < t_s \leq 1$, to establish

$$(H_n(\sigma, t_1), \dots, H_n(\sigma, t_s)) \xrightarrow{\nu_{n, \theta}} (X(t_1), \dots, X(t_s)),$$

we shall prove that for any $T > 0$, the characteristic functions

$$\varphi_n(\lambda_1, \dots, \lambda_s) := \frac{1}{\theta_{(n)}} \sum_{\sigma \in \mathbf{S}_n} \theta^{w(\sigma)} \exp \left\{ i \sum_{r=1}^s \lambda_r H_n(\sigma, t_r) \right\}$$

converge uniformly in $|\lambda_1| \leq T, \dots, |\lambda_s| \leq T$. We apply the Proposition with

$$f(j) = \exp \left\{ i a_n(j) \sum_{r=1}^s \mathbf{1}\{j \leq t_r n\} \lambda_r \right\}.$$

Since

$$\sum_{j \leq n} \frac{1}{j} (1 - \Re f(j)) \ll \sum_{j \leq n} \sum_{r \leq s} \frac{1}{j} (\lambda_r^2 + 1) a_n(j)^{s-2} \ll \Psi_n(+\infty) < \infty,$$

(4.2) is satisfied and the Proposition yields,

$$\begin{aligned} \varphi_n(\lambda_1, \dots, \lambda_s) &= \exp \left\{ -i \sum_{r=1}^s \lambda_r A(t_r n, n) \right\} \frac{\Gamma(\theta)}{2\pi i} \\ &\times \int_{1-Ki}^{1+Ki} \frac{e^z}{z^\theta} \exp \left\{ \theta \sum_{j \leq n} \frac{1}{j} e^{-zj/n} \left(\exp \left\{ i a_n(j) \sum_{r=1}^s \mathbf{1}\{j \leq t_r n\} \lambda_r \right\} - 1 \right) \right\} dz \\ &+ O(K^{-c}). \end{aligned}$$

Set $t_0 = 0$ and $\xi_r = \lambda_r + \dots + \lambda_s$ for $1 \leq r \leq s$, then this formula can be rewritten as

$$\begin{aligned} (5.1) \quad \varphi_n(\lambda_1, \dots, \lambda_s) &= \exp \left\{ \theta \sum_{r=1}^s (\mu_n(t_r, \xi_r) - \mu_n(t_{r-1}, \xi_r)) \right\} \\ &\times \frac{\Gamma(\theta)}{2\pi i} \int_{1-Ki}^{1+Ki} \frac{e^z}{z^\theta} \exp \left\{ \theta \sum_{r=1}^s (s_n(t_r, \xi_r; z) - s_n(t_{r-1}, \xi_r; z)) \right\} dz \\ &+ O(K^{-c}), \end{aligned}$$

where s_n and μ_n are defined in (3.4) and (3.5).

The proof of weak convergence of finite dimensional distributions now follows from Lemmas 4.4 and 4.5.

5.2. Tightness

For $x \in \mathbf{D}[0, 1]$ and $0 < \delta < 1$, set

$$\begin{aligned} \Delta_\delta^1(x) &= \sup \{ |x(1) - x(t)| : 1 - \delta \leq t \leq 1 \}, \\ \Delta_\delta^2(x) &= \sup \{ |x(t) - x(0)| : 0 \leq t \leq \delta \}, \quad \text{and} \\ \Delta_\delta^3(x) &= \sup_{\delta \leq t \leq 1-\delta} \{ |x(t) - x(u)| \wedge |x(v) - x(t)| : t - \delta \leq u \leq t \leq v \leq t + \delta \}. \end{aligned}$$

To establish tightness, it is enough to show for arbitrary $\varepsilon > 0$ and $j = 1, 2, 3$, that

$$(5.2) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \nu_{n, \theta}(\Delta_\delta^j(H_n(\sigma, \cdot)) \geq \varepsilon) = 0.$$

Fix $\varepsilon > 0$. Since for sufficiently small $\delta > 0$,

$$\sum_{(1-\delta)n < j \leq n} \frac{1}{j} \theta |a_n(j)|^* \ll \delta \leq \varepsilon/2,$$

we have by Lemma 4.1,

$$\begin{aligned} \nu_{n,\theta}(\Delta_\delta^1(H_n(\sigma, \cdot)) \geq \varepsilon) &\leq \nu_{n,\theta} \left(\sum_{(1-\delta)n < j \leq n} \frac{|h_j(k_j(\sigma))|}{\beta(n)} \geq \varepsilon/2 \right) \\ &\ll P^{\theta \wedge 1} \left(\sum_{(1-\delta)n < j \leq n} \left(\frac{|h_j(\xi_j)|}{\beta(n)} \right)^* \geq \varepsilon/6 \right) + o(1) \\ &\ll \left(\sum_{(1-\delta)n < j \leq n} \frac{1}{j} |a_n(j)|^* \right)^{\theta \wedge 1} + o(1) \ll \delta^{\theta \wedge 1} + o(1). \end{aligned}$$

This proves (5.2) for $j = 1$. Another application of Lemma 4.1 gives,

$$(5.3) \quad \begin{aligned} \nu_{n,\theta}(\Delta_\delta^2(H_n(\sigma, \cdot)) \geq \varepsilon) \\ \ll P^{\theta \wedge 1} \left(\sup_{0 < t \leq \delta} \left| \sum_{j \leq tn} a_n(j) \xi_j - A(tn, n) \right| \geq \varepsilon \right) + o(1). \end{aligned}$$

Since by [8, Theorem 22.4],

$$P \left(\sup_{0 < t \leq \delta} \left| \sum_{j \leq tn} a_n(j)^* \xi_j - A(tn, n) \right| \geq \varepsilon \right) \ll \varepsilon^{-2} \tilde{\Psi}_n(\infty; \delta),$$

we have as $n \rightarrow \infty$,

$$(5.4) \quad \begin{aligned} P \left(\sup_{0 < t \leq \delta} \left| \sum_{j \leq tn} a_n(j) \xi_j - A(tn, n) \right| \geq \varepsilon \right) \\ \leq P \left(\sup_{0 < t \leq \delta} \left| \sum_{j \leq tn} a_n^*(j) \xi_j - A(tn, n) \right| \geq \varepsilon \right) \\ + P(a_n(j) \xi_j \neq a_n^*(j) \xi_j \text{ for some } j \leq \delta n) \\ \ll \varepsilon^{-2} \tilde{\Psi}_n(\infty; \delta) + \sum_{\substack{j \leq \delta n \\ |a_n(j)| \geq 1}} P(\xi_j \geq 1) \\ \ll \tilde{\Psi}_n(\infty; \delta) = o(1) + \tilde{\Psi}(\infty; \delta). \end{aligned}$$

Note that the last term

$$\tilde{\Psi}(\infty; \delta) = \Psi(-\delta^{-\rho}) + \Psi(\infty) - \Psi(\delta^{-\rho}) + \int_{-\delta^{-\rho}}^{\delta^{-\rho}} \delta^{2\rho} \max\{1, u^2\} d\Psi(u) \rightarrow 0$$

as $\delta \rightarrow 0$. Hence by (5.3) and (5.4), the limit (5.2) holds for $j = 2$.

To establish (5.2) for $j = 3$, note that for $0 < \delta < 1/12$,

$$\begin{aligned} \Delta_\delta^3(x) &\leq \Delta_\delta^4(x) \\ &\quad + \sup_{\delta \leq t \leq \sqrt{\delta}} \{|x(t) - x(u)| \wedge |x(v) - x(t)| : t - \delta \leq u \leq t \leq v \leq t + \delta\} \\ &\leq \Delta_\delta^4(x) + 2 \sup\{|x(t)| : 0 \leq t \leq 2\sqrt{\delta}\}, \end{aligned}$$

the last term above is just $\Delta_{2\sqrt{\delta}}^2(x)$ provided $x(0) = 0$, where

$$\Delta_\delta^4(x) = \sup_{\sqrt{\delta} \leq t \leq 1 - \delta} \{|x(t) - x(u)| \wedge |x(v) - x(t)| : t - \delta \leq u \leq t \leq v \leq t + \delta\}.$$

For $\sqrt{\delta} \leq t \leq 1 - \delta$ and $t - \delta \leq u \leq t \leq v \leq t + \delta$, we clearly have

$$\begin{aligned} \theta \sum_{t'n \leq j \leq tn} \frac{1}{j} |a_n(j)|^* + \theta \sum_{tn \leq j \leq t'n} \frac{1}{j} |a_n(j)|^* \\ \ll -\log\left(1 - \frac{\delta}{t}\right) + \log\left(1 + \frac{\delta}{t}\right) \\ \ll \frac{\delta}{t} \leq C\sqrt{\delta}, \end{aligned}$$

for some $C > 0$, and hence

$$\Delta_\delta^4(H_n(\sigma, \cdot)) \leq V_{n,t,\delta}(\sigma) + C\sqrt{\delta},$$

where

$$V_{n,t,\delta}(\sigma) = \sum_{(t-\delta)n \leq j \leq tn} \frac{|h_j(k_j(\sigma))|}{\beta(n)} \wedge \sum_{tn \leq j \leq (t+\delta)n} \frac{|h_j(k_j(\sigma))|}{\beta(n)}.$$

Let $\delta > 0$ be sufficiently small so that $\varepsilon_1 := \varepsilon - C\sqrt{\delta} > 0$. Hence it is enough to show that

$$\Delta_{\delta,n}(\sigma) := \nu_{n,\theta}(\sup\{V_{n,t,\delta}(\sigma) : \sqrt{\delta} \leq t \leq 1 - \delta\} \geq \varepsilon_1) \rightarrow 0,$$

as $n \rightarrow \infty$ and $\delta \rightarrow 0$. If $V_{n,t,\delta}(\sigma) > 0$ for some $t \in [\sqrt{\delta}, 1 - \delta]$, then there exist $l \in [(t - \delta)n, tn]$ and $m \in (tn, (t + \delta)n]$ such that $k_l(\sigma) \geq 1$ and $k_m(\sigma) \geq 1$. Since this $t \leq \delta + l/n$, we further have the bounds $(\sqrt{\delta} - \delta)n \leq l \leq n$ and $l < m \leq l + 2\delta n$. Since $2\delta \leq \sqrt{\delta}$ for small $0 < \delta < 1/4$, we have by Lemma 4.3,

$$\Delta_{\delta,n}(\sigma) \leq \sum \mathbf{1}\{(\sqrt{\delta} - \delta)n \leq l \leq n, l < m \leq l + 2\delta n\} \nu_{n,\theta}^{l,m} \ll \sqrt{\delta} + n^{-\theta}.$$

This completes the proof of Theorem 3.1.

§6. Appendix

In this section we prove the Proposition stated in Section 4. As in Section 4, let $|f(j)| \leq 1$, $j \geq 1$ be complex numbers, depending, maybe, on n or other parameters, $f(j) = 1$ for $j > n$. We first note that the n -th Taylor coefficient of the series

$$M(z) := 1 + \sum_{m=1}^{\infty} M_m z^m = \exp \left\{ \theta \sum_{j=1}^{\infty} \frac{1}{j} f(j) z^j \right\}, \quad |z| < 1,$$

is given by (4.1), and that

$$(6.1) \quad |M_n| \leq \frac{\theta^{(n)}}{n!} = \frac{n^{\theta-1}}{\Gamma(\theta)} (1 + O(n^{-1})).$$

We divide the proof into several lemmas.

Lemma 6.1. *Let $r = e^{-1/n}$, $2 \leq K \leq n$, $0 \leq j \leq n$, and $\theta \neq 1$. Then*

$$\int_{K/n < |\tau| \leq \pi} M(re^{i\tau}) e^{-ij\tau} d\tau \ll ((j+1)^{\theta-1} + n^{\theta-1}) \log K.$$

Proof. Integrating the power series, by (6.1), we obtain

$$\begin{aligned} \int_{K/n < |\tau| \leq \pi} M(re^{i\tau}) e^{-ij\tau} d\tau &= \sum_{m=0}^{\infty} M_m r^m \int_{K/n < |\tau| \leq \pi} e^{i(m-j)\tau} d\tau \\ &= M_j r^j (2\pi - 2K/n) - 2 \sum_{m \geq 0, m \neq j} M_m r^m \frac{\sin \frac{K}{n}(m-j)}{m-j} \\ &\ll (j+1)^{\theta-1} + \sum_{m \geq 0, m \neq j} (m+1)^{\theta-1} r^m \left| \frac{\sin \frac{K}{n}(m-j)}{m-j} \right| \\ &\ll (j+1)^{\theta-1} + (j+1)^{\theta-1} \sum_{1 \leq |m-j| \leq j/2} \left| \frac{\sin \frac{K}{n}(m-j)}{m-j} \right| \\ &+ \sum_{|m-j| > j/2} \frac{(m+1)^{\theta-1}}{|m-j|} r^m. \end{aligned} \tag{6.2}$$

The last sum

$$\begin{aligned}
& \sum_{|m-j|>j/2} \frac{(m+1)^{\theta-1}}{|m-j|} e^{-m/n} \\
& \ll \frac{1}{j+1} \sum_{0 \leq m < j/2} (m+1)^{\theta-1} + \sum_{m > 3j/2} (m+1)^{\theta-2} e^{-m/n} \\
& \ll (j+1)^{\theta-1} + \sum_{3j/2 < m \leq n} (m+1)^{\theta-2} + n^{\theta-2} \sum_{m > n} \left(\frac{m}{n}\right)^{\theta-2} e^{-m/n} \\
& \ll (j+1)^{\theta-1} + n^{\theta-1}.
\end{aligned}$$

The sum in the second term on the last inequality of (6.2) can be estimated as follows:

$$\begin{aligned}
& \sum_{\substack{1 \leq |m-j| \leq j/2 \\ m \neq j}} \left| \frac{\sin \frac{K}{n}(m-j)}{m-j} \right| \leq \sum_{1 \leq |m-j| \leq n/K} \frac{K}{n} + \sum_{n/K \leq |m-j| \leq j/2} \frac{1}{|m-j|} \\
& \ll 1 + \log \left(2 + \frac{j}{2} \frac{K}{n} \right) \ll \log K.
\end{aligned}$$

This completes the proof of Lemma 6.1.

Q.E.D.

Recall that

$$L(z) = \sum_{1 \leq j \leq n} \frac{1}{j} (f(j) - 1) z^j.$$

Instead of (4.2) we will use the bound

$$E(u) := \exp \left\{ 2 \sum_{\substack{j \leq n \\ |f(j)-1| > u}} \frac{1}{j} |f(j) - 1| \right\} \leq \exp \{-4u^{-1} \Re L(1)\}$$

where $u > 0$.

Lemma 6.2 (Lemma 4 of [15]). *Let $r = e^{-1/n}$, $z = re^{i\tau}$, and $|\tau| \leq \pi$. Then, for arbitrary $u > 0$,*

$$\exp\{|L(z) - L(1)|\} = \exp \left\{ \left| \sum_{j \leq n} \frac{1}{j} (f(j) - 1) (z^j - 1) \right| \right\} \ll_u E(u) \left| \frac{1-z}{1-r} \right|^{4u/\pi}$$

where the constant in \ll depends only on u .

Set

$$l(u) := E^\theta(u) \exp\{\theta \Re L(1)\}$$

and observe that Lemma 6.2 yields the estimate

$$(6.3) \quad \exp\{\theta L(z)\} \ll_u l(u) n^{4u\theta/\pi} |1-z|^{4u\theta/\pi}$$

for arbitrary $u > 0$, $z = re^{i\tau}$, and $|\tau| \leq \pi$.

Lemma 6.3. *Let $\theta \neq 1$ and $0 < u < \pi/4$ be fixed. For arbitrary $2 \leq K \leq n$ and $\varepsilon \in [2/n, 1/2]$, we have*

$$\begin{aligned} \frac{n!}{\theta(n)} M_n &= \frac{\Gamma(\theta)}{2\pi i} \int_{1-Ki}^{1+Ki} \frac{e^z}{z^\theta} \exp\{\theta L(e^{-z/n})\} dz \\ &+ O\left(l(u) K^{\theta(4u/\pi-1)} (K^2 n^{-1} + \varepsilon^{-1/2})\right) + O((\varepsilon^\theta + \varepsilon) \log K) \\ &+ O(l(u) n^{-1} K \log K) + O(n^{-1}). \end{aligned}$$

The constant in $O(\cdot)$ depends on θ and u only.

Proof. Set $D_0 = \{z = re^{i\tau} : |\tau| \leq K/n\}$ and $D = \{z = re^{i\tau} : K/n < |\tau| \leq \pi\}$. We use Cauchy's formula

$$M_n = \frac{1}{2\pi i} \left(\int_{D_0} + \int_D \right) \frac{M(z)}{z^{n+1}} dz =: J_0 + J_1.$$

Since by (6.3)

$$(6.4) \quad \begin{aligned} \max_{z \in D} |M(z)| &= \max_{z \in D} |\exp\{\theta L(z)\} (1-z)^{-\theta}| \\ &\ll l(u) n^{4\theta u/\pi} \max_{z \in D} |1-z|^{\theta(4u/\pi-1)} \\ &\ll n^\theta l(u) K^{\theta(4u/\pi-1)}, \end{aligned}$$

integrating by parts, we obtain

$$\begin{aligned} J_1 &= \frac{\theta}{2\pi i n} \int_D \frac{\exp\{\theta L(z)\} L'(z)}{z^n (1-z)^\theta} dz + \frac{\theta}{2\pi i n} \int_D \frac{\exp\{\theta L(z)\}}{z^n (1-z)^{\theta+1}} dz \\ &+ O\left(n^{\theta-1} l(u) K^{\theta(4u/\pi-1)}\right) \\ &=: J_{11} + J_{12} + O\left(n^{\theta-1} l(u) K^{\theta(4u/\pi-1)}\right). \end{aligned}$$

From (6.3) we have

$$\begin{aligned}
 (6.5) \quad J_{12} &\ll l(u)n^{4u\theta/\pi-1} \int_D |1-z|^{-\theta-1+4u\theta/\pi} |dz| \\
 &\ll l(u)n^{4u\theta/\pi-1} \int_{K/n \leq |\tau| \leq \pi} \tau^{-\theta-1+4u\theta/\pi} d\tau \\
 &\ll l(u)n^{\theta-1} K^{\theta(4u/\pi-1)}.
 \end{aligned}$$

For J_{11} , we use Lemma 6.1 and obtain

$$\begin{aligned}
 (6.6) \quad J_{11} &= \frac{\theta}{2\pi i n} \int_D \frac{M(z)}{z^n} L'(z) dz \\
 &= \frac{\theta}{2\pi n} \sum_{j \leq n} (f(j) - 1) r^{j-n} \int_{K/n \leq |\tau| \leq \pi} M(re^{i\tau}) e^{i\tau(j-n)} d\tau \\
 &\ll \frac{1}{n} \left(\sum_{0 \leq j \leq T} + \sum_{T < j \leq n} \right) \left| \int_{K/n \leq |\tau| \leq \pi} M(re^{i\tau}) e^{-ij\tau} d\tau \right| \\
 &=: \frac{1}{n} (I_1 + I_2),
 \end{aligned}$$

where $T = \lceil \varepsilon n \rceil$ and $\varepsilon \in [2/n, 1/2]$ is a parameter. By Lemma 6.1,

$$(6.7) \quad I_1 \ll (T^\theta + Tn^{\theta-1}) \log K.$$

To estimate I_2 , we again use integration by parts and (6.4). Further applying Cauchy's inequality, we obtain

$$\begin{aligned}
 I_2 &\ll \sum_{T < j \leq n} \frac{1}{j} \left| \int_{K/n \leq |\tau| \leq \pi} M'(re^{i\tau}) e^{-i(j-1)\tau} d\tau \right| + \sum_{T < j \leq n} \frac{1}{j} \max_{z \in D} |M(z)| \\
 &\ll T^{-1/2} \left(\sum_{1 \leq j \leq n} \left| \int_{K/n \leq |\tau| \leq \pi} M'(re^{i\tau}) e^{-ij\tau} d\tau \right|^2 \right)^{1/2} \\
 &\quad + l(u)n^\theta K^{\theta(4u/\pi-1)} \log(\varepsilon^{-1}).
 \end{aligned}$$

The integrals under the last sum are just the Fourier coefficients of an appropriate function therefore, via Parseval's identity, we further have

$$(6.8) \quad I_2 \ll T^{-1/2} \left(\int_{K/n \leq |\tau| \leq \pi} |M'(re^{i\tau})|^2 d\tau \right)^{1/2} + l(u)n^\theta K^{\theta(4u/\pi-1)} \log(\varepsilon^{-1}).$$

Estimating the integral in (6.8), we exploit (6.3). The integral can be bounded by

$$(6.9) \quad \int_{K/n \leq |\tau| \leq \pi} \frac{|\exp\{2\theta L(re^{i\tau})\}|}{|1 - re^{i\tau}|^{2\theta+2}} d\tau + \int_{K/n \leq |\tau| \leq \pi} \frac{|\exp\{2\theta L(re^{i\tau})\}| |L'(re^{i\tau})|^2}{|1 - re^{i\tau}|^{2\theta}} d\tau$$

$$\ll l^2(u)n^{8u\theta/\pi} \int_{K/n \leq |\tau| \leq \pi} \frac{d\tau}{|1 - re^{i\tau}|^{2\theta+2-8u\theta/\pi}}$$

$$+ l^2(u)n^{8u\theta/\pi} \int_{K/n \leq |\tau| \leq \pi} \frac{|L'(re^{i\tau})|^2}{|1 - re^{i\tau}|^{2\theta-8u\theta/\pi}} d\tau$$

$$\ll l^2(u)n^{1+2\theta} K^{2\theta(4u/\pi-1)-1}$$

$$+ l^2(u)n^{2\theta} K^{2\theta(4u/\pi-1)} \int_{|\tau| \leq \pi} |L'(re^{i\tau})|^2 d\tau$$

$$\ll l^2(u)n^{2\theta+1} K^{2\theta(4u/\pi-1)}.$$

In the last step, we again used Parseval's identity. Now (6.8) and (6.9) yield

$$(6.10) \quad I_2 \ll l(u)K^{\theta(4u/\pi-1)}n^\theta \varepsilon^{-1/2}.$$

Inserting (6.7) and (6.10) into (6.6), we obtain

$$J_{11} \ll n^{\theta-1} (\varepsilon^\theta + \varepsilon) \log K + n^{\theta-1} l(u) K^{\theta(4u/\pi-1)} \varepsilon^{-1/2}.$$

This and the estimate of J_{12} in (6.5) show that

$$J_1 \ll n^{\theta-1} (\varepsilon^\theta + \varepsilon) \log K + n^{\theta-1} l(u) K^{\theta(4u/\pi-1)} \varepsilon^{-1/2}.$$

So we have arrived at

$$(6.11) \quad M_n = J_0 + O(n^{\theta-1} (\varepsilon^\theta + \varepsilon) \log K + n^{\theta-1} l(u) K^{\theta(4u/\pi-1)} \varepsilon^{-1/2}).$$

Substituting $z = e^{-w/n}$ and applying (6.3), we have

$$\begin{aligned}
J_0 &= \frac{1}{2\pi i n} \int_{1-iK}^{1+iK} \frac{\exp\{w + \theta L(e^{-w/n})\}}{(1 - e^{-w/n})^\theta} dw \\
&= \frac{n^{\theta-1}}{2\pi i} \int_{1-Ki}^{1+Ki} \exp\{w + \theta L(e^{-w/n})\} w^{-\theta} \left(1 + O\left(\frac{w}{n}\right)\right) dw \\
&= \frac{n^{\theta-1}}{2\pi i} \int_{1-Ki}^{1+Ki} \exp\{w + \theta L(e^{-w/n})\} w^{-\theta} dw \\
&\quad + O\left(l(u)n^{\theta-2}K \int_{1-Ki}^{1+Ki} |w|^{\theta(4u/\pi-1)} |dw|\right) \\
&= \frac{n^{\theta-1}}{2\pi i} \int_{1-Ki}^{1+Ki} \exp\{w + \theta L(e^{-w/n})\} w^{-\theta} dw \\
&\quad + O\left(l(u)n^{\theta-2}K^{2+\theta(4u/\pi-1)}\right) + O\left(l(u)n^{\theta-2}K \log K\right).
\end{aligned}$$

This inequality, (6.1) and (6.11) complete the proof of Lemma 6.3.
Q.E.D.

In the case $\theta = 1$, one can repeat these calculations with some minor changes or recall the following result.

Lemma 6.4 (The Main Lemma of [14]). *Let $\theta = 1$ and condition (4.2) be satisfied. Then, for arbitrary fixed $1 < K < n$ and $0 < \delta < 1/2$, we have*

(6.12)

$$M_n = \frac{1}{2\pi i} \int_{1-Ki}^{1+Ki} \frac{e^z}{z} \exp\left\{L(e^{-z/n})\right\} \left(1 + O\left(\frac{K}{n}\right)\right) dz + O(K^{-1/2+\delta}),$$

where the constants implied in the symbols $O(\cdot)$ depend on L and δ only.

Proof of the Proposition. Under condition (4.2), we have $l(u) \leq e^{4\theta L/u} < \infty$. Thus, if $\theta \neq 1$ and $K \leq \sqrt{n}$, it suffices to apply Lemma 6.3 with $u = \pi/8$ and $\varepsilon = K^{-(\theta \wedge 1)/2}$.

If $\theta = 1$, we use (6.3) to estimate the remainder

$$\int_{1-Ki}^{1+Ki} |\exp\{L(e^{-z/n})\}| O(K/n) \frac{|dz|}{|z|} \\ \ll Kn^{4u/\pi-1} \int_{1-Ki}^{1+Ki} |1 - e^{-z/n}|^{4u/\pi} \frac{|dz|}{|z|} \ll K^{1+4u/\pi} n^{-1}.$$

This, for $u = \pi/8$ and $K \leq \sqrt{n}$, together with (6.12) yield the Proposition.

References

- [1] R. Arratia, A. D. Barbour and S. Tavaré, Poisson process approximations for the Ewens sampling formula, *Ann. Appl. Probab.*, **2** (1992), 519–535.
- [2] R. Arratia, A. D. Barbour and S. Tavaré, *Logarithmic Combinatorial Structures: a Probabilistic Approach*, EMS Monogr. Math., EMS, Zürich, 2003.
- [3] G. J. Babu and E. Manstavičius, Brownian motion and random permutations, *Sankhyā Ser. A*, **61** (1999), 312–327.
- [4] G. J. Babu and E. Manstavičius, Random permutations and the Ewens sampling formula in genetics, In: *Probab. Theory and Math. Stat.*, (eds. B. Grigelionis et al.), VSP/TEV, Vilnius/Utrecht, 1999, pp. 33–42.
- [5] G. J. Babu and E. Manstavičius, Limit theorems for random permutations, In: *Paul Erdős and His Mathematics*, János Bolyai Mathematical Society, 1999, pp. 19–22.
- [6] G. J. Babu and E. Manstavičius, Infinitely divisible limit processes for the Ewens sampling formula, *Lithuanian Math. J.*, **42** (2002), 232–242.
- [7] G. J. Babu and E. Manstavičius, Limit Processes with independent increments for the Ewens sampling formula, *Ann. Inst. Stat. Math.*, **54** (2002), 607–620.
- [8] P. Billingsley, *Probability and Measure*, Third Edition, John Wiley & Sons, New York, 1995.
- [9] P. Billingsley, *Convergence of Probability Measures*, Second Edition, John Wiley & Sons, New York, 1999.
- [10] J. M. DeLaurentis and B. G. Pittel, Random permutations and the Brownian motion, *Pacific J. Math.*, **119** (1985), 287–301.
- [11] P. Donnelly, T. G. Kurtz and S. Tavaré, On the functional central limit theorem for the Ewens Sampling Formula, *Ann. Appl. Probab.*, **1** (1991), 539–545.
- [12] W. J. Ewens, The sampling theory of selectively neutral alleles, *Theoret. Population Biology*, **3** (1972), 87–112.

- [13] J. C. Hansen, A functional central limit theorem for the Ewens Sampling Formula, *J. Appl. Probab.*, **27** (1990), 28–43.
- [14] E. Manstavičius, Additive and multiplicative functions on random permutations, *Liet. Mat. Rink.*, **36** (1996), 400–408.
- [15] E. Manstavičius, The Berry–Esseen bound in the theory of random permutations, *Ramanujan J.*, **2** (1998), 185–199.
- [16] E. Manstavičius, Functional limit theorem for sequences of mappings on the symmetric group, In: *Anal. Probab. Methods Number Theory*, (eds. A. Dubickas et al.), TEV, Vilnius, 2002, 175–187.
- [17] E. Manstavičius, Mappings on decomposable combinatorial structures: analytic approach, *Combin. Probab. Comput.*, **11** (2002), 61–78.
- [18] N. M. Timofeev and H. H. Usmanov, On some arithmetic models of random processes, *Dokl. Acad. Sci. Tadzh. SSR*, **24** (1981), 284–287 (Russian).

Gutti Jogesh Babu
Department of Statistics,
319 Thomas Building,
The Pennsylvania State University,
University Park, PA 16802, USA

Eugenijus Manstavičius
Department of Mathematics and Informatics,
Vilnius University,
Naugarduko str. 24,
03225 Vilnius,
Lithuania

Vytas Zacharovas
Department of Mathematics and Informatics,
Vilnius University,
Naugarduko str. 24,
03225 Vilnius,
Lithuania