RANDOM PERMUTATIONS WITH CYCLE WEIGHTS

VOLKER BETZ, DANIEL UELTSCHI, YVAN VELENIK

Abstract. We study the distribution of cycle lengths in models of nonuniform random permutations with cycle weights. We identify several regimes. Depending on the weights, the length of typical cycles grows like the total number n of elements, or a fraction of n, or a logarithmic power of n.

Keywords: Random permutations, cycle weights, cycle lengths, Ewens distribution.
2000 Math. Subj. Class.: 60K35.

1. Introduction

We study the cycle distributions in models of weighted random permutations. The probability of a permutation π of n elements is defined by

$$P(\pi) = \frac{1}{h_n n!} \prod_{j \geqslant 1} \theta_j^{r_j(\pi)}, \tag{1.1}$$

where $(\theta_1, \theta_2, \dots) \equiv \boldsymbol{\theta}$ are real nonnegative numbers, $r_j(\pi)$ denotes the number of j-cycles in π (we always have $\sum_j j r_j(\pi) = n$), and h_n is the normalization. We are mainly interested in the distribution of cycle lengths in the limit $n \to \infty$, and in how these lengths depend on the set of parameters $\boldsymbol{\theta}$.

This model was introduced in [4] but variants of it have been studied previously. The case of constant $\theta_j \equiv \theta$ is known as the Ewens distribution. It appears in the study of population dynamics in mathematical biology [5]; detailed results about the number of cycles were obtained by Hansen [9] and by Feng and Hoppe [6]. The distribution of cycle lengths was considered by Lugo [10]. Another variant of this model involves parameters $\theta_j \in \{0, 1\}$, with finitely many 1's [11, 2], or with parity dependence [10]. Notice that the probability P is really a probability on sequences $\boldsymbol{r} = (r_1, r_2, \dots)$ that satisfy $\sum_j j r_j = n$. It is well-known that $\boldsymbol{r}$ are the "occupation numbers" of a partition $\boldsymbol{\lambda}$ of n. That is, if $\boldsymbol{\lambda}$ denotes the partition $\lambda_1 \geqslant \lambda_2 \geqslant \dots$ with $\sum_i \lambda_i = n$, then r_j is the number of λ_i that satisfy $\lambda_i = j$. Thus we are really dealing with random partitions, which is also the original context of the Ewens distribution. The number of permutations that are compatible with occupation numbers $\boldsymbol{r}$ is equal to

$$\frac{n!}{\prod_{j \geqslant 1} j^{r_j} r_j!}.$$

It follows that the marginal of (1.1) on partitions is given by

$$P(\boldsymbol{\lambda}) = \frac{1}{h_n} \prod_{j \geqslant 1} \frac{1}{r_j!} \big(\tfrac{1}{j}\theta_j\big)^{r_j}. \tag{1.2}$$

The formulæ look simpler and more elegant for permutations rather than partitions; that is why we consider the former.

The most studied distribution for random partitions is the Plancherel measure, where the probability of $\boldsymbol{\lambda}$ is proportional to $\frac{1}{n!}(\dim \boldsymbol{\lambda})^2$; the "dimension" $\dim \boldsymbol{\lambda}$ of a partition is defined as the number of Young tableaux in Young diagrams and it does not seem to have an easy expression in terms of $\boldsymbol{r}$. We do not know of any direct relation between weighted random permutations and the Plancherel measure.

Weighted random permutations also appear in the study of large systems of quantum bosonic particles [3], where the parameters $\boldsymbol{\theta}$ depend on such quantities as the temperature, the density, and the particle interactions. The θ_j's are thus forced upon us and they do not necessarily take a nice form. This motivates the present setting, where we only fix the asymptotic behavior of θ_j as $j \to \infty$.

The relevant random variables in our analysis are the lengths $\ell_i = \ell_i(\pi)$ of the cycle containing the index $i = 1, \dots, n$. These random variables are always identically distributed, and obviously not independent. Another relevant random variable is the number of indices belonging to cycles of length between a and b, $N_{a,b}(\pi) = \#\{i = 1, \dots, n : a \leqslant \ell_i(\pi) \leqslant b\}$. Thanks to the relation

$$\frac{1}{n} E(N_{a,b}) = P(\ell_1 \in [a,b]), \tag{1.3}$$

the properties of the distribution of ℓ_1 that we derive below can be translated into properties of the expectation of $N_{a,b}$.

From a statistical mechanics point of view, it is natural to introduce the sequence $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots)$ of parameters such that $\mathrm{e}^{-\alpha_j} = \theta_j$. The model has an important symmetry, which is also a source of confusion. Namely, the probability of the permutation π is left invariant under the transformation

$$\alpha_j \mapsto \alpha_j + cj, \qquad h_n \mapsto \mathrm{e}^{-cn}\, h_n, \tag{1.4}$$

for any constant $c \in \mathbb{R}$. In particular, the case $\alpha_j = cj$ is identical to $\alpha_j \equiv 0$, the case of uniform random permutations.

The general results which we prove in this article rely on various technical assumptions. To keep this introduction simple, we only describe the results in the particular but interesting case $\alpha_j = j^\gamma$.

- The case $\gamma < 0$ is a special case of the model studied in [4], which is close to the uniform distribution.
- In the case $\gamma = 0$, i.e. when $\theta_j \to \theta$ (the Ewens case, asymptotically), we find that $P(\ell_1 > sn) \to (1-s)^\theta$. Thus, almost all indices belong to cycles whose length is a fraction of n. Precise statements and proofs can be found in Section 2.
- The case $0 < \gamma < 1$ is surprising. At first glance we might expect smaller cycles than in the uniform case $\alpha_j \equiv 0$. However, we find that almost all indices belong to a single giant cycle! The symmetry (1.4) is playing tricks upon us indeed. We also prove two additional results: (i) there is a uniformly positive probability that *all* indices belong to a single cycle of lenth n; (ii) in some cases, finite cycles exist with uniformly positive probability. This is explained in details in Section 3.
- The case $\gamma = 1$ corresponds to uniform permutations because of the symmetry (1.4).
- When $\gamma > 1$ the cycles become shorter, and ℓ_1 behaves asymptotically as $(\frac{1}{1-\gamma} \log n)^{1/\gamma}$; see Section 4.

Weighted random permutations clearly show a rich behavior and only a little part has been uncovered so far. The case of negative parameters, $\alpha_j \asymp -j^\gamma$ remains to be explored, and the future will hopefully bring more results regarding concentration properties.

In the case of uniform permutations, it is known that the random variables r_k converge to independent Poisson random variables with parameter $1/k$ in the limit $n \to \infty$ [8, 1]. An open problem is to understand how this generalizes to weighted random permutations.

2. Asymptotic Ewens distribution

In the case of the uniform distribution, it is an easy exercise to show that $P(\ell_1 = a) = 1/n$ for any $a = 1, \dots, n$. It follows that $P(\ell_1 > sn) \to 1 - s$ for any $0 \leqslant s \leqslant 1$. This result was extended to the case of small weights in [4]. We consider here parameters that are close to Ewens weights. A result similar to (a) below has been recently derived by Lugo [10].

Theorem 2.1. *Let $\theta \in \mathbb{R}_+$. We suppose that $\sum_{j=1}^\infty \frac{1}{j}|\theta_j - \theta| < \infty$ if $\theta \geqslant 1$, or that $\sum_{j=1}^\infty |\theta_j - \theta| < \infty$ if $\theta < 1$.*

(a) *The distribution of ℓ_1 satisfies, for $0 \leqslant s \leqslant 1$,*

$$\lim_{n\to\infty} P(\ell_1 > sn) = (1-s)^\theta. \tag{2.1}$$

(b) *The joint distribution of ℓ_1 and ℓ_2 satisfies, for $0 \leqslant s,t \leqslant 1$,*

$$\lim_{n\to\infty} P(\ell_1 > sn, \ell_2 > tn) = \frac{\theta}{1+\theta}(1-s-t)_+^{\theta+1} + \frac{1+\theta(s\vee t)}{1+\theta}(1-s\vee t)^\theta, \tag{2.2}$$

where f_+ denotes the positive part of a function f.

Let us recall a few properties that are satisfied by the normalization factors h_n; we omit the proofs since they are essentially elementary, but more details can be found in [4]. First,

$$P(\ell_1 \in [a,b]) = \frac{1}{n}\sum_{j=a}^{b} \theta_j \frac{h_{n-j}}{h_n}. \tag{2.3}$$

Choosing $[a,b] = [1,n]$, we get

$$h_n = \frac{1}{n}\sum_{j=1}^{n} \theta_j h_{n-j}, \qquad h_0 = 1. \tag{2.4}$$

Next, let $G_h(s) = \sum_{n\geqslant 0} h_n s^n$ be the generating function of the sequence (h_n). Then $G_h(s) = \exp \sum_{j\geqslant 1} \frac{1}{j}\theta_j s^j$. The first step in the proof of Theorem 2.1 is to control the normalization h_n. Here, $(\theta)_n = \theta(\theta+1)\dots(\theta+n-1)$ denotes the ascending factorial.

Proposition 2.2. *Under the assumptions of Theorem 2.1, we have*

$$h_n = C(\boldsymbol{\theta})\frac{(\theta)_n}{n!}(1+o(1)) \qquad \text{with} \qquad C(\boldsymbol{\theta}) = \exp\sum_{j\geqslant 1}\tfrac{1}{j}(\theta_j - \theta).$$

Proof. We have

$$G_h(s) = \exp\Big\{\theta\sum_j \tfrac{1}{j}s^j + \sum_j \tfrac{1}{j}(\theta_j-\theta)s^j\Big\} = (1-s)^{-\theta}\,\mathrm{e}^{u(s)}, \tag{2.5}$$

with

$$u(s) = \sum_{j\geqslant 1}\tfrac{1}{j}(\theta_j-\theta)s^j. \tag{2.6}$$

Notice that $u(1) = \lim_{s\nearrow 1} u(s)$ exists. Let c_j be the Taylor coefficients of $\mathrm{e}^{u(s)}$, i.e. $\mathrm{e}^{u(s)} = \sum c_j s^j$. Then, by Leibniz' rule,

$$h_n = \frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}s^n}G_h(s)\Big|_{z=0} = \frac{(\theta)_n}{n!}\sum_{k\geqslant 0} d_{n,k}c_k, \tag{2.7}$$

with

$$d_{n,k} = \begin{cases} \frac{n(n-1)\dots(n-k+1)}{(\theta+n-1)\dots(\theta+n-k)} & \text{if } k \leqslant n,\\ 0 & \text{otherwise.}\end{cases} \tag{2.8}$$

It is not hard to check that

$$d_{n,k} \leqslant \begin{cases} 1 & \text{if } \theta \geqslant 1,\\ \theta^{-1}k+1 & \text{if } \theta > 0.\end{cases} \tag{2.9}$$

Let $U(s) = \sum \frac{1}{j}|\theta_j - \theta| s^j$ and C_j the Taylor coefficients of $\mathrm{e}^{U(s)}$. It is clear that $|c_j| \leqslant C_j$ for all j. When $\theta \geqslant 1$, the first bound of (2.9) and the dominated convergence theorem imply

$$\lim_{n\to\infty}\sum_{k\geqslant 0} d_{n,k}c_k = \sum_{k\geqslant 0} c_k = \mathrm{e}^{u(1)} = C(\boldsymbol{\theta}). \tag{2.10}$$

When $\theta < 1$, the second bound of (2.9) gives $d_{n,k}|c_k| \leqslant (\theta^{-1}k+1)C_k$. The sequence (kC_k) is absolutely convergent:

$$\sum kC_k = \frac{\mathrm{d}}{\mathrm{d}s}\,\mathrm{e}^{U(s)}\Big|_{s=1} = \mathrm{e}^{U(1)}\,U'(1) = \mathrm{e}^{\sum \frac{1}{j}|\theta_j-\theta|}\sum|\theta_j-\theta| < \infty. \tag{2.11}$$

We again obtain (2.10) by the dominated convergence theorem. □

Proof of Theorem 2.1. We show that, for any $0 < s < t < 1$, we have

$$\lim_{n\to\infty} P(\ell_1 \in [sn, tn]) = (1-s)^\theta - (1-t)^\theta. \tag{2.12}$$

Using Proposition 2.2, we have

$$P(\ell_1 \in [sn,tn]) = \frac{1}{n}\sum_{j=sn}^{tn}\theta_j\frac{h_{n-j}}{h_n} = \frac{\theta}{n}\sum_{j=sn}^{tn}\frac{(\theta)_{n-j}}{(n-j)!}\frac{n!}{(\theta)_n}(1+o(1)). \tag{2.13}$$

Here and throughout this article, when a and b are not integer we use the convention

$$\sum_{j=a}^{b} f(j) = \sum_{j\in[a,b]\cap\mathbb{N}} f(j) = \sum_{j=\lceil a\rceil}^{\lfloor b\rfloor} f(j). \tag{2.14}$$

We now use the identity

$$(\theta)_n = \frac{\Gamma(n+\theta)}{\Gamma(\theta)} \tag{2.15}$$

and the asymptotic

$$\frac{\Gamma(n+\theta)}{n!} = n^{\theta-1}(1+o(1)). \tag{2.16}$$

We get

$$P(\ell_1 \in [sn,tn]) = \frac{\theta}{n}\sum_{j=sn}^{tn}(1-\tfrac{j}{n})^{\theta-1}(1+o(1)). \tag{2.17}$$

As $n\to\infty$, the right side converges to the Riemann integral $\theta\int_s^t(1-\xi)^{\theta-1}\mathrm{d}\xi$, and we obtain the first claim of Theorem 2.1. Let us now turn to the second claim. Let $1 \leqslant a \leqslant b \leqslant n$ and $1 \leqslant c \leqslant d \leqslant n$. Splitting the last term according to whether 1 and 2 belong to the same cycle or to different cycles, we get

$$P\big(\ell_1\in[a,b],\ell_2\in[c,d]\big) = \frac{1}{n(n-1)}\sum_{\substack{j\in[a,b]\\ k\in[c,d]\\ j+k\,\leqslant\, n}}\theta_j\theta_k\frac{h_{n-j-k}}{h_n} + \frac{1}{n(n-1)}\sum_{j\in[a,b]\cap[c,d]}(j-1)\theta_j\frac{h_{n-j}}{h_n}. \tag{2.18}$$

Let $\epsilon > 0$ and set $a = sn$, $c = tn$ and $b = d = n$. We assume, without loss of generality, that $1 \geqslant s \geqslant t \geqslant 0$. Using the above expression, Proposition 2.2 and Eqs (2.15)–(2.16), we deduce that, for n large,

$$\begin{aligned} P\big(\ell_1 \geqslant sn, \ell_2 \geqslant tn\big) = \frac{\theta^2}{n^2}&\sum_{\substack{j\,\geqslant\, sn,k\,\geqslant\, tn\\ j+k\,\leqslant\,(1-\epsilon)n}}\Big(1-\frac{j+k}{n}\Big)^{\theta-1}(1+o_\epsilon(1)) \\ &+\frac{\theta}{n^2}\sum_{sn\,\leqslant\, j\,\leqslant\,(1-\epsilon)n}(j-1)\Big(1-\frac{j}{n}\Big)^{\theta-1}(1+o_\epsilon(1)) + O(\epsilon). \end{aligned} \tag{2.19}$$

Taking first the limit $n\to\infty$ and then the limit $\epsilon\to 0$, the right side of the latter expression is seen to converge to

$$1_{\{s+t\,\leqslant\,1\}}\theta^2\int_{s+t}^1(\xi-s-t)(1-\xi)^{\theta-1}\mathrm{d}\xi + \theta\int_s^1\xi(1-\xi)^{\theta-1}\mathrm{d}\xi, \tag{2.20}$$

and the second claim of Theorem 2.1 follows. □

3. Slowly diverging parameters

This section is devoted to parameters α_j that grow slowly to $+\infty$. The typical case is $\alpha_j = j^\gamma$ with $0 < \gamma < 1$, but our conditions allow more general sequences. As mentioned in the introduction, the system displays a surprising behavior: almost all indices belong to a single giant cycle.

Theorem 3.1. *We assume that* $\frac{\theta_{n-j}\theta_j}{\theta_n} \leqslant c_j$ *for all* n *and for* $j = 1, \dots, \frac{n}{2}$*, with constants* c_j *that satisfy* $\sum_{j \geqslant 1} \frac{c_j}{j} < \infty$*. Then*

$$\lim_{m\to\infty} \lim_{n\to\infty} P(\ell_1 > n - m) = 1.$$

It may be worth recalling that n always denote the number of elements in this article, and that P depends on n. The proof of this theorem can be found later in this section. In the case $\alpha_j = j^\gamma$, we have

$$\frac{\theta_{n-j}\theta_j}{\theta_n} = \mathrm{e}^{-n^\gamma[(1-\frac{j}{n})^\gamma + (\frac{j}{n})^\gamma - 1]} \approx \begin{cases} \mathrm{e}^{-j^\gamma} & \text{if } j \ll n, \\ \mathrm{e}^{-cn^\gamma} & \text{if } j = sn, \end{cases} \tag{3.1}$$

where the constant in the last equation is $c = (1-s)^\gamma + s^\gamma - 1$. It is positive for $0 < \gamma < 1$, and the condition of the theorem is fulfilled.

Let us understand why parameters $\alpha_j = j^\gamma$ favor longer and longer cycles. The heuristics is actually provided by statistical mechanics. Namely, we can write the probability $P(\pi)$ as a Gibbs distribution $\frac{1}{Z}\,\mathrm{e}^{-H(\pi)}$ with "Hamiltonian" $H(\pi) = \sum_{i=1}^n \frac{\alpha_{\ell_i(\pi)}}{\ell_i(\pi)}$. Thus, to each index i that belongs to a cycle of length j is associated an "energy" $\frac{\alpha_j}{j} = j^{\gamma-1}$. Indices in longer cycles have lower energy, so they are favored. This discussion also provides an illustration for the symmetry (1.4); it amounts to shifting the Hamiltonian by a constant and this does not affect the Gibbs distribution.

A natural question in view of Theorem 3.1 is whether finite cycles occur at all, or whether there is exactly one cycle of length n. The next theorem provides a partial answer to this question by giving a sufficient condition for the occurrence of cycles of length 1.

Theorem 3.2. *Suppose the assumptions of Theorem 3.1 hold.*

(a) *There exists* $c > 0$ *such that* $P(\ell_1 = n) = P(r_n = 1) > c$.
(b) *Suppose in addition that* $\theta_1\theta_{n-1}/\theta_n$ *is uniformly bounded away from 0, and consider the weights* $\lambda\theta_j$ *where* λ *is a parameter. There exist* $\lambda_0 = \lambda_0(\boldsymbol{\theta})$ *and* $c > 0$ *such that*

$$P(r_1 \geqslant 1) > c$$

for all n *and for all* $0 < \lambda < \lambda_0$.

This theorem is proved at the end of the section. We first obtain estimates for h_n.

Proposition 3.3. *Under the assumptions of Theorem 3.1 there exists a constant* B *such that, for all* n,

$$1 \leqslant \frac{nh_n}{\theta_n} \leqslant B.$$

The constant B *depends on* $\{c_j\}$*, but it does not explicitly depend on* $\boldsymbol{\theta}$.

Proof. Let $a_n = \frac{nh_n}{\theta_n}$. The relation (2.4) can be written as

$$a_n = 1 + \sum_{j=1}^{n-1} \frac{1}{j} \frac{\theta_{n-j}\theta_j}{\theta_n} a_j. \tag{3.2}$$

The lower bound of the lemma is trivial, contrary to the upper bound. Let us rewrite the relation above as

$$a_n = \begin{cases} 1 + \sum_{j=1}^{\frac{n-1}{2}} \frac{\theta_{n-j}\theta_j}{\theta_n} \Big(\frac{a_j}{j} + \frac{a_{n-j}}{n-j} \Big) & \text{if } n \text{ is odd}, \\ 1 + \sum_{j=1}^{\frac{n}{2}-1} \frac{\theta_{n-j}\theta_j}{\theta_n} \Big(\frac{a_j}{j} + \frac{a_{n-j}}{n-j} \Big) + \frac{2\theta_{n/2}^2}{n\theta_n} a_{n/2} & \text{if } n \text{ is even}. \end{cases} \tag{3.3}$$

We define the sequence (b_n) by the recursion equation

$$b_n = 1 + \sum_{j=1}^{n/2} c_j \Bigl(\frac{b_j}{j} + \frac{b_{n-j}}{n-j} \Bigr). \tag{3.4}$$

It is clear that $a_n \leqslant b_n$ for all n. Next, let m be a number such that

$$\frac{2}{n} \sum_{j=1}^{n/2} c_j + \sum_{j>m/2} \frac{c_j}{j} \leqslant \frac{1}{2} \tag{3.5}$$

for all $n \geqslant m$. Such an m exists because (c_j/j) is summable. We set

$$B = 2 \max_{1 \leqslant j \leqslant m} b_j. \tag{3.6}$$

Notice that B depends on the c_j's but not on the θ_j's. Finally, we introduce another sequence (b'_n) defined by

$$b'_n = \begin{cases} b_n & \text{if } n \leqslant m, \\ 1 + \sum_{j=1}^{n/2} c_j \Bigl(\frac{b'_j}{j} + \frac{2B}{n} \Bigr) & \text{if } n > m. \end{cases} \tag{3.7}$$

It is clear that $b'_n \leqslant \frac{1}{2} B$ for $n \leqslant m$; we now show by induction that $b'_n \leqslant B$ for all n. We have

$$\begin{aligned} b'_n - b'_m &= \frac{2B}{n} \sum_{j=1}^{n/2} c_j - \sum_{j=1}^{m/2} c_j \frac{b'_{m-j}}{m-j} + \sum_{j=\frac{m}{2}+1}^{n/2} c_j \frac{b'_j}{j} \\ &\leqslant \Bigl(\frac{2}{n} \sum_{j=1}^{n/2} c_j + \sum_{j>m/2} \frac{c_j}{j} \Bigr) B. \end{aligned} \tag{3.8}$$

This is less than $\frac{1}{2} B$ by the definition (3.5) of m. Since $b'_m \leqslant \frac{1}{2} B$, we find that $b'_n \leqslant B$ for all n. The final step is to see that $b_n \leqslant b'_n$. This is clear when $n \leqslant m$, and we get it by induction when $n > m$:

$$b_{n+1} = 1 + \sum_{j=1}^{n/2} c_j \Bigl(\frac{b_j}{j} + \frac{b_{n-j+1}}{n-j+1} \Bigr) \leqslant 1 + \sum_{j=1}^{n/2} c_j \Bigl(\frac{b'_j}{j} + \frac{2B}{n+1} \Bigr) = b'_{n+1}. \tag{3.9}$$

We have shown that $a_n \leqslant b_n \leqslant b'_n \leqslant B$ for all n. □

Proof of Theorem 3.1. Using Proposition 3.3, we get

$$P(\ell_1 \leqslant n-m) = \frac{1}{n} \sum_{j=m}^{n-1} \theta_{n-j} \frac{h_j}{h_n} \leqslant B \sum_{j=m}^{n-1} \frac{1}{j} \frac{\theta_{n-j}\theta_j}{\theta_n} \leqslant B \sum_{j=m}^{n/2} \frac{c_j}{j} + B \sum_{j=n/2}^{n-1} \frac{c_{n-j}}{j}. \tag{3.10}$$

The last term goes to zero as $n \to \infty$. The first term goes to zero as $n \to \infty$ and $m \to \infty$. □

Proof of Theorem 3.2. For the first claim, we note that

$$P(\ell_1 = n) = \frac{\theta_n}{n h_n}. \tag{3.11}$$

This is larger than $1/B$ by Proposition 3.3.

We prove now the second claim. Let A_i be the event where i belongs to a cycle of length 1, i.e. $A_i = \{\pi \in \mathcal{S}_n : \pi(i) = i\}$. Then $\cup_{i=1}^n A_i$ is the event where $\ell_i = 1$ for at least an index i, and we have the lower bound

$$\begin{aligned}P(r_1 \geqslant 1) = P(\cup_{i=1}^n A_i) &\geqslant \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) \\ &= nP(A_1) - \tfrac{1}{2}n(n-1)P(A_1 \cap A_2) \\ &= \frac{\lambda\theta_1}{h_n}\big[h_{n-1} - \tfrac{1}{2}\lambda\theta_1 h_{n-2}\big].\end{aligned} \tag{3.12}$$

Using Proposition 3.3, we get

$$P(r_1 \geqslant 1) \geqslant \frac{\lambda\theta_1\theta_{n-1}}{\theta_n}\Big[\frac{n}{B(n-1)} - \frac{n\lambda\theta_1\theta_{n-2}}{2(n-2)\theta_{n-1}}\Big]. \tag{3.13}$$

Notice that $\theta_1\theta_{n-2}/\theta_{n-1} \leqslant c_1$. The lower bound is strictly positive when $\lambda < \frac{2(n-2)}{Bc_1(n-1)}$. □

4. Quickly diverging parameters

Here we treat parameters $\theta_j = \mathrm{e}^{-\alpha_j}$ with α_j diverging quickly, or equivalently θ_j decaying quickly. More precisely, we shall make the following two assumptions: for some $M > 0$, all $k \geqslant 1$, and two coprime numbers $j_1, j_2 \geqslant 4$,

$$\begin{aligned}&0 \leqslant \theta_k \leqslant \frac{\mathrm{e}^{Mk}}{k!}, \\ &\theta_{j_1} > 0,\, \theta_{j_2} > 0.\end{aligned} \tag{4.1}$$

It is necessary to suppose some kind of aperiodicity condition on the set of indices corresponding to nonvanishing coefficients θ_j. This prevents us from prescribing e.g. permutations with only even lengths of cycles. In this case, we have $h_n = 0$ for all odd n, as can be easily seen from the recursion (2.4); Proposition 4.5 below would fail.

Our assumptions allow to get the asymptotics of h_n using the saddle point method. We write down the steps explicitly in order to keep the article self-contained. A slightly shorter path would be to prove that our assumptions imply that e^f, with $f(z) = \sum_{j=0}^\infty \theta_j z^j$, is "Hayman admissible", and to use standard results [7]. That Hayman admissibility holds is implicit in our proof.

We describe general results in Subsection 4.1, relegating proofs to Subsection 4.2. The general results turn out to be somewhat abstract, so we use them to study the particularly interesting class $\alpha_j = j^\gamma$, $\gamma > 1$, in Subsection 4.3.

4.1. **Main properties.** We now describe three general theorems about cycle lengths. They all assume the conditions (4.1), although we do not recall it explicitly.

The first statement concerns the absence of macroscopic cycles.

Theorem 4.1. *For arbitrarily small $\delta > 0$ and arbitrarily large $k > 0$, there exists $n_0 = n_0(\delta, k)$ such that*

$$P\big(\max_{1 \leqslant i \leqslant n} \ell_i \geqslant \delta n\big) \leqslant n^{-k},$$

for all $n \geqslant n_0$.

More precise information about typical cycle lengths can be extracted from the following result. Let r_n be defined by the equation

$$\sum_{j \geqslant 1} \theta_j r_n^j = n. \tag{4.2}$$

Theorem 4.2. *Let $a(n), b(n)$ be such that*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{a(n)} \theta_j r_n^{j+1/2} = 0, \qquad \lim_{n\to\infty} \frac{1}{n}\sum_{j=b(n)}^{n} \theta_j r_n^{j+1/2} = 0.$$

Then

$$\lim_{n\to\infty} P(\ell_1 \in [a(n), b(n)]) = 1.$$

When the information about the coefficients θ_j is sufficiently detailed, some control on r_n is possible and Theorem 4.2 can be used to obtain sharp results. This is exemplified in Subsection 4.3 for the special case $\alpha_j = \alpha(j) = j^\gamma$ with $\gamma > 1$. In such cases, the sum $\sum_{j=1}^\infty \theta_j r_n^{j+1/2}$ (whose value is $r_n^{1/2} n$) is dominated by the terms corresponding to indices j close to the solution $j_{\max}$ of the equation $\alpha'(j) = \log r_n$.

Finally, it is also possible to extract from Theorem 4.2 a general result proving absence of small cycles.

Theorem 4.3.

$$\lim_{n\to\infty} P\Big(\ell_1 \leqslant \frac{\log n}{\log r_n} - \tfrac{3}{4}\Big) = 0.$$

We shall see below that the proof of Theorem 4.3 is straightforward; nonetheless, the result is quite strong. In the case where only finitely many θ_j are nonzero, we find $r_n \sim n^{1/j_0}$, where j_0 is the last index with nonzero θ_j. Thus $\log n/\log r_n \approx j_0$, and we obtain that the probability that $\ell_1 \leqslant j_0 - 1$ is zero. It follows that almost all cycles have length j_0, a fact already observed in [11, 2]. On the other hand, if infinitely many θ_j are nonzero, it is easy to see that $\log n/\log r_n$ diverges. Thus the probability of ℓ_1 being finite goes to zero. To summarize, the only way to force a positive density of indices to lie in finite cycles is to forbid infinite cycles altogether, in which case typical cycles have the maximal length that is allowed.

4.2. **Proofs of the main properties.** We now prove Theorems 4.1 – 4.3. We use the following elementary result, which is a consequence of the first assumption in (4.1).

Lemma 4.4. *Let $f(x) = \sum_{k=0}^\infty c_k x^k$ with Taylor coefficients that satisfy $0 \leqslant c_k \leqslant \mathrm{e}^{Mk} k^{-k}$ for some $M > 0$ and all $k \geqslant 1$. Then for all $\delta > 0$ and all $x \geqslant 0$, we have*

$$f'(x) \leqslant (1+\delta)\,\mathrm{e}^M f(x) + \mathrm{e}^M/\delta.$$

Proof. Let $k_0 = k_0(x) = \lfloor (1+\delta)\,\mathrm{e}^M x \rfloor$. We decompose

$$f'(x) = \sum_{k=1}^\infty c_k k x^{k-1} = \sum_{k=1}^{k_0} c_k k x^{k-1} + R(x).$$

By our assumptions,

$$\begin{aligned} R(x) = \sum_{k=k_0+1}^\infty c_k k x^{k-1} &\leqslant \mathrm{e}^M \sum_{k=k_0+1}^\infty \Big(\frac{x\,\mathrm{e}^M}{k}\Big)^{k-1} \\ &\leqslant \mathrm{e}^M \sum_{k=k_0+1}^\infty \Big(\frac{1}{1+\delta}\Big)^k \leqslant \mathrm{e}^M/\delta. \end{aligned}$$

On the other hand, for the terms up to k_0 we have $k \leqslant k_0 \leqslant (1+\delta) x\,\mathrm{e}^M$, and thus

$$\sum_{k=1}^{k_0} c_k k x^{k-1} \leqslant (1+\delta)\,\mathrm{e}^M \sum_{k=0}^{k_0} c_k x^k \leqslant (1+\delta)\,\mathrm{e}^M f(x).$$

This completes the proof. □

Let us define the functions

$$I_\beta(z) = \sum_{j=1}^\infty j^\beta \theta_j z^j$$

for $\beta \in \mathbb{R}$. $\phi(z) := I_{-1}(z)$ plays a special role, since the generating function of (h_n) is given by $G_h(z) = \exp(\phi(z))$. All I_β are analytic by the first assumption in (4.1), monotone increasing and positive on $\{z > 0\}$ together with all their derivatives, and $I_{\beta+1}(z) = zI'_\beta(z)$. Lemma 4.2 implies that for each $\beta > 0$, there exists C such that for all $z \geqslant 0$, we have

$$I'_\beta(z) \leqslant CI_\beta(z). \tag{4.3}$$

Recall that $r_n = I_0^{-1}(n)$, where I_0^{-1} denote the inverse function.

Proposition 4.5. *We have*

$$h_n = \frac{r_n^{-n}}{\sqrt{2\pi I_1(r_n)}} \, \mathrm{e}^{\phi(r_n)} \, (1 + o(1)).$$

Proof. The condition (4.1) on Taylor coefficients implies that $I_0(z) < \tilde{D}\exp(Cz)$. Then

$$r_n \geqslant c \log n \tag{4.4}$$

for some $c > 0$. On the other hand, r_n diverges more slowly than $n^{1/4}$ since $I_0(x)$ diverges faster than x^4 by (4.1).

For the saddle point method, we use Cauchy's formula and we obtain

$$\begin{aligned} h_n &= \frac{1}{2\pi r^n} \int_{-\pi}^{\pi} \mathrm{e}^{\phi(r\,\mathrm{e}^{\mathrm{i}\gamma}) - n\mathrm{i}\gamma} \,\mathrm{d}\gamma \\ &= \frac{\mathrm{e}^{\phi(r)}}{2\pi r^n} \Big[\int_{-\gamma_0}^{\gamma_0} \mathrm{e}^{\phi(r\,\mathrm{e}^{\mathrm{i}\gamma}) - \phi(r) - n\mathrm{i}\gamma} \,\mathrm{d}\gamma + 2 \int_{\gamma_0}^{\pi} \mathrm{e}^{\phi(r\,\mathrm{e}^{\mathrm{i}\gamma}) - \phi(r) - n\mathrm{i}\gamma} \,\mathrm{d}\gamma \Big], \end{aligned} \tag{4.5}$$

for any $r > 0$ and any $0 < \gamma_0 < \pi$. We choose $r = r_n$ as given by Eq. (4.2), and $\gamma_0 = \gamma_0(n) = r_n^{-(1+\delta)}$, for some $0 < \delta < 1/2$. The leading order of the first term above can be found by expanding $\phi(z) - n \log z$ around $\gamma = 0$. We have

$$\phi(r_n \,\mathrm{e}^{\mathrm{i}\gamma}) - \phi(r_n) - n\mathrm{i}\gamma = \sum_{j \geqslant 1} \frac{\theta_j}{j} r_n^j \big(\mathrm{e}^{\mathrm{i}j\gamma} - 1 - \mathrm{i}j\gamma \big). \tag{4.6}$$

Expanding $\mathrm{e}^{\mathrm{i}j\gamma} - 1 - \mathrm{i}j\gamma = -\frac{1}{2} j^2\gamma^2 + R(j\gamma)$ with $|R(j\gamma)| \leqslant \frac{1}{3!}(j\gamma)^3$, we get

$$\phi(r_n \,\mathrm{e}^{\mathrm{i}\gamma}) - \phi(r_n) - n\mathrm{i}\gamma = -\tfrac{1}{2}\gamma^2 \sum_{j \geqslant 1} j\theta_j r_n^j + A(\gamma) = -\tfrac{1}{2}\gamma^2 I_1(r_n) + A(\gamma), \tag{4.7}$$

with

$$|A(\gamma)| \leqslant \frac{\gamma_0^3}{3!} \sum_{j \geqslant 1} j^2 \theta_j r_n^j = \frac{\gamma^2}{r_n^{1+\delta} 3!} I_2(r_n) \tag{4.8}$$

for all $\gamma \leqslant \gamma_0$. Now by (4.3), we have $I_2(r_n) \leqslant Cr_n I_1(r_n)$. Thus, as $n \to \infty$, the term $A(\gamma)$ is negligible compared to $\gamma^2 I_1(r_n)$ in the first integral, which is therefore given by

$$\int_{-\gamma_0}^{\gamma_0} \mathrm{e}^{-\frac{1}{2}\gamma^2 I_1(r_n)(1+o(1))} \,\mathrm{d}\gamma = \frac{1}{\sqrt{I_1(r_n)}} \int_{-\gamma_0\sqrt{I_1(r_n)}}^{\gamma_0\sqrt{I_1(r_n)}} \mathrm{e}^{-\frac{1}{2}\xi^2(1+o(1))} \,\mathrm{d}\xi = \sqrt{\frac{2\pi}{I_1(r_n)}} (1 + o(1)). \tag{4.9}$$

The last equality is justified by the fact that $\gamma_0(n) I_1(r_n) \geqslant r_n^{-1-\delta} I_0(r_n) \geqslant r_n^{-2} n$, which diverges as $n \to \infty$.

We now turn to the second term in (4.5). We want to show that it is negligible, and we estimate it by replacing the integral by π times the maximum of the integrand. In view of (4.9) it is enough to show that

$$\lim_{n\to\infty} \tfrac{1}{2} \log I_1(r_n) - \mathrm{Re}(\phi(r_n) - \phi(r_n \,\mathrm{e}^{\mathrm{i}\gamma})) = -\infty \tag{4.10}$$

for all $\gamma \in [\gamma_0, \pi]$. For the first term, we have $\log I_1(r_n) \leqslant \log(Cr_n I_0(r_n)) \leqslant \tilde{C} \log n$. For the second term, we have

$$\begin{aligned} \mathrm{Re}(\phi(r_n) - \phi(r_n \,\mathrm{e}^{\mathrm{i}\gamma})) &= \sum_{j \geqslant 1} \frac{1}{j} \theta_j r_n^j (1 - \cos(\gamma j)) \\ &\geqslant \frac{\theta_{j_1}}{j_1} r_n^{j_1} (1 - \cos(\gamma j_1)) + \frac{\theta_{j_2}}{j_2} r_n^{j_2} (1 - \cos(\gamma j_2)), \end{aligned} \tag{4.11}$$

where j_1 and j_2 are picked according to (4.1). The right side is zero at $\gamma = 0$, and it is strictly positive when $\gamma \in (0, \pi]$ (j_1 and j_2 are coprime); so its minimum is taken at γ_0 when n is sufficiently large (recall that $\gamma_0 \to 0$ when $n \to \infty$). Expanding the cosine, we get

$$\mathrm{Re}(\phi(r_n) - \phi(r_n \, \mathrm{e}^{\mathrm{i}\gamma})) \geqslant c' r_n^4 \gamma_0^2 = c' r_n^{2-2\delta} \geqslant cc' (\log n)^{2-2\delta}. \tag{4.12}$$

This dominates the first term of (4.10) since $\delta < 1/2$, and this completes the proof. □

Proof of Theorem 4.1. Clearly,

$$P(\max_i \ell_i > \delta n) \leqslant n P(\ell_1 > \delta n). \tag{4.13}$$

We have $I_1(r_n) \leqslant C^2 r_n^2 \phi(r_n)$ by (4.3), and thus Proposition 4.5 gives $h_n \geqslant C' r_n^{-n-1}$ for n large enough. Since all the h_{n-j}'s are clearly bounded by some $D > 0$, we have by (2.3)

$$n P(\ell_1 > \delta n) \leqslant D r_n^{n+1} \sum_{j=\delta n}^{n} \Big(\frac{\mathrm{e}^M}{j} \Big)^j \leqslant D r_n^{n+1} n \Big(\frac{\mathrm{e}^M}{\delta n} \Big)^{\delta n} \leqslant D n \Big(\frac{\mathrm{e}^M \, r_n^{2/\delta}}{\delta n} \Big)^{\delta n}. \tag{4.14}$$

The statement is trivial (and seen directly from (2.3)) if only finitely many θ_j are nonzero; thus we may assume there are infinitely many nonzero θ_j. Then $I_0(z)$ grows faster at infinity than any power of z, and r_n diverges more slowly than any power of n. The last bracket is less than 1 for n large enough, so that the right side vanishes in the limit $n \to \infty$. □

In order to make more precise statements about the length of typical cycles, we need a better control over the terms appearing in (2.3). By the previous result it suffices to consider the case where j is not too close to n.

Proposition 4.6. *For each $\delta > 0$ there exists C_δ such that, for all $n \in \mathbb{N}$ and all $j < (1-\delta)n$, we have*

$$\frac{h_{n-j}}{h_n} \leqslant C_\delta r_n^{j+1/2}.$$

Proof. By Proposition 4.5 we have

$$\begin{aligned} \frac{h_{n-j}}{h_n} &\approx r_n^j \left(\tfrac{r_n}{r_{n-j}} \right)^{n-j} \left(\tfrac{I_1(r_n)}{I_1(r_{n-j})} \right)^{1/2} \mathrm{e}^{\phi(r_{n-j}) - \phi(r_n)} \\ &= r_n^j \exp\Big(- \big(\phi(r_n) - \phi(r_{n-j}) - (n-j)(\ln(r_n) - \ln(r_{n-j})) \big) \Big) \left(\tfrac{I_1(r_n)}{I_1(r_{n-j})} \right)^{1/2} = \\ &= r_n^j \exp\Big(- \big(\phi(r_n) - \phi(r_{n-j}) - \phi'(r_{n-j}) r_{n-j} \ln(\tfrac{r_n}{r_{n-j}}) \big) \Big) \left(\tfrac{I_1(r_n)}{I_1(r_{n-j})} \right)^{1/2}, \end{aligned} \tag{4.15}$$

when both n and $n-j$ are large. Put $r_{n-j} = x$ and $r_n = x + u$. Since $n \mapsto r_n$ is increasing, we have $u > 0$. The exponent above then has the form

$$\phi(x+u) - \phi(x) - x\phi'(x) \ln(\tfrac{x+u}{x}) = (\phi(x+u) - \phi(x) - \phi'(x) u) + \phi'(x)(u - x \ln(\tfrac{x+u}{x})). \tag{4.16}$$

The first bracket in the right side is greater than $\frac12 u^2 \phi''(x)$, since all derivatives of ϕ are positive on $\mathbb{R}^+$. The second bracket is always positive. Thus for all $n \in \mathbb{N}$ and all $j \leqslant (1-\delta)n$, there exists $C'_\delta > 0$ such that

$$\frac{h_{n-j}}{h_n} \leqslant C'_\delta r_n^j \, \mathrm{e}^{-\frac12 (r_n - r_{n-j})^2 \phi''(r_{n-j})} \left(\tfrac{I_1(r_n)}{I_1(r_{n-j})} \right)^{1/2} \tag{4.17}$$

By (4.3), $I_1(x) = x I_0'(x) \leqslant C x I_0(x)$. We also have $I_0(x) \leqslant I_1(x)$. Since $I_0(r_n) = n$, we get

$$\frac{I_1(r_n)}{I_1(r_{n-j})} \leqslant C r_n \frac{n}{n-j} \leqslant \frac{C}{\delta} r_n. \tag{4.18}$$

This proves the claim. □

Proof of Theorem 4.2. The claims follows immediately from (2.3) and Proposition 4.6. □

Proof of Theorem 4.3. Let $m = \log n / \log r_n - \frac{3}{4}$. We use Eq. (2.3), bounding θ_j by a constant and using Proposition 4.6 for the ratio of normalization factors. Since r_n diverges, we have

$$P(\ell_1 \leqslant m) \leqslant \tfrac{C}{n} \sum_{j=1}^{m} r_n^{j+1/2} = \tfrac{C}{n} r_n^{3/2} \frac{r_n^m - 1}{r_n - 1} \leqslant \tfrac{C'}{n} r_n^{m+1/2}, \tag{4.19}$$

if n is large enough. The right side is equal to $C' r_n^{-1/4}$ and it vanishes in the limit $n \to \infty$. □

4.3. **An explicit example.** In this subsection, we treat explicitly the case $\alpha_j = \alpha(j) = j^\gamma$ with $\gamma > 1$, as an example of application of the previous general results. We first observe that the assumptions (4.1) are trivially satisfied, so that the general results in this section apply.

The main result of this subsection is that typical cycles are of size $(\frac{1}{\gamma-1} \log n)^{1/\gamma}$, to leading order.

Theorem 4.7. *Let $\alpha_j = j^\gamma$, with $\gamma > 1$. Then*

$$\frac{\ell_1}{(\frac{1}{\gamma-1} \log_n)^{1/\gamma}} \to 1 \qquad \textit{in probability.} \tag{4.20}$$

Let us define

$$\Delta(j) = \alpha(j) - \alpha(j_{\max}) - (j - j_{\max}) \log r_n. \tag{4.21}$$

The proof of Theorem 4.7 follows from two simple technical estimates.

Lemma 4.8. *Let $j_{\max} \in \mathbb{R}$ be such that $\alpha'(j_{\max}) = \log r_n$.*

(a) *Assume that $\gamma \geqslant 2$. Then, for all $j \geqslant 1$, there exists $c = c(\gamma) > 0$ such that*

$$\Delta(j) \geqslant c\alpha''(j_{\max})(j - j_{\max})^2. \tag{4.22}$$

(When $j \geqslant j_{\max}$, one can choose $c = \frac{1}{2}$.)

(b) *Assume that $\gamma \in (1, 2)$. Then, for all $1 \leqslant j \leqslant 2j_{\max}$, there exists $c = c(\gamma) > 0$ such that*

$$\Delta(j) \geqslant c\alpha''(j_{\max})(j - j_{\max})^2. \tag{4.23}$$

(When $j \leqslant j_{\max}$, one can choose $c = \frac{1}{2}$.) Moreover, for all $j > 2j_{\max}$, there exists $c = c(\gamma) > 0$ such that

$$\Delta(j) \geqslant cj^\gamma. \tag{4.24}$$

Proof. We start with the case $\gamma \geqslant 2$. First of all, since $j_{\max} = (\alpha')^{-1}(\log r_n)$, we have, for any $j > j_{\max}$,

$$\begin{aligned} \Delta(j) &= \alpha(j) - \alpha(j_{\max}) - (j - j_{\max}) \log r_n = \alpha(j) - \alpha(j_{\max}) - (j - j_{\max})\alpha'(j_{\max}) \\ &= \int_{j_{\max}}^{j} \mathrm{d}s \int_{j_{\max}}^{s} \alpha''(t)\mathrm{d}t \geqslant \tfrac{1}{2}\alpha''(j_{\max})(j - j_{\max})^2, \end{aligned} \tag{4.25}$$

since α'' is an increasing function. Similarly, we have, for any $\frac{1}{2} j_{\max} \leqslant j < j_{\max}$,

$$\Delta(j) = \int_{j}^{j_{\max}} \mathrm{d}s \int_{s}^{j_{\max}} \alpha''(t)\mathrm{d}t \geqslant \tfrac{1}{2}\alpha''(\tfrac{1}{2} j_{\max})(j - j_{\max})^2 = 2^{1-\gamma}\alpha''(j_{\max})(j - j_{\max})^2. \tag{4.26}$$

Finally, for $0 \leqslant j < \frac{1}{2} j_{\max}$, we use

$$\begin{aligned} \Delta(j) = \int_{j}^{j_{\max}} \mathrm{d}s \int_{s}^{j_{\max}} \alpha''(t)\mathrm{d}t &\geqslant \int_{j_{\max}/2}^{j_{\max}} \mathrm{d}s \int_{s}^{j_{\max}} \alpha''(t)\mathrm{d}t \\ &\geqslant \tfrac{1}{2}\alpha''(\tfrac{1}{2} j_{\max}) \tfrac{1}{4} j_{\max}^2 \geqslant 2^{-\gamma-1}\alpha''(j_{\max})(j - j_{\max})^2. \end{aligned} \tag{4.27}$$

Let us now turn to the case $\gamma \in (1, 2)$. The proof is completely similar. When $j \leqslant j_{\max}$, we use (observe that α'' is a decreasing function now)

$$\Delta(j) = \int_{j}^{j_{\max}} \mathrm{d}s \int_{s}^{j_{\max}} \alpha''(t)\mathrm{d}t \geqslant \tfrac{1}{2}\alpha''(j_{\max})(j - j_{\max})^2. \tag{4.28}$$

When $j_{\max} < j \leqslant 2j_{\max}$, we use

$$\Delta(j) = \int_{j_{\max}}^{j} \mathrm{d}s \int_{j_{\max}}^{s} \alpha''(t)\mathrm{d}t \geqslant \tfrac12 \alpha''(2j_{\max})(j-j_{\max})^2 = 2^{\gamma-3}\alpha''(j_{\max})(j-j_{\max})^2. \tag{4.29}$$

Finally, when $j > 2j_{\max}$, we have

$$\Delta(j) = \int_{j_{\max}}^{j} \mathrm{d}s \int_{j_{\max}}^{s} \alpha''(t)\mathrm{d}t \geqslant \tfrac12 \alpha''(j)(j-j_{\max})^2 \geqslant \tfrac18 \alpha''(j) j^2 = \tfrac18 \gamma(\gamma-1) j^\gamma. \tag{4.30}$$

□

Corollary 4.9. *For any $\gamma > 1$, we have, as $n \to \infty$,*

$$j_{\max} = (\frac{1}{\gamma-1}\log n)^{1/\gamma}\,(1+o(1)), \tag{4.31}$$

$$\log r_n = \alpha'(j_{\max}) = \gamma(\frac{1}{\gamma-1}\log n)^{(\gamma-1)/\gamma}\,(1+o(1)), \tag{4.32}$$

$$\mathrm{e}^{-\alpha(j_{\max})}\, r_n^{j_{\max}} = n^{1+o(1)}. \tag{4.33}$$

Proof. We start with the case $\gamma \geqslant 2$. Using the previous lemma, it immediately follows that

$$I_0(r_n) = \sum_{j\geqslant 1} e^{-\alpha(j)} r_n^j \leqslant e^{-\alpha(j_{\max})} r_n^{j_{\max}} \sum_{j\geqslant 1} e^{-c\alpha''(j_{\max})(j-j_{\max})^2} \leqslant C_1\, e^{-\alpha(j_{\max})} r_n^{j_{\max}} \tag{4.34}$$

Since, for $j < j_{\max}$, $\Delta(j) \leqslant \frac12 \alpha''(j_{\max})(j-j_{\max})^2$, we also have

$$I_0(r_n) \geqslant e^{-\alpha(\lfloor j_{\max}\rfloor)} r_n^{\lfloor j_{\max}\rfloor} \geqslant e^{-\frac12 \alpha''(j_{\max})}\, e^{-\alpha(j_{\max})} r_n^{j_{\max}}. \tag{4.35}$$

Using the relation $I_0(r_n) = n$, (4.34) and (4.35) immediately imply the claimed asymptotics.

Let us now turn to the case $\gamma \in (1,2)$. The lemma implies that

$$I_0(r_n) = e^{-\alpha(j_{\max})} r_n^{j_{\max}} \sum_{j\geqslant 1} e^{-\Delta(j)} \leqslant C_2\, e^{-\alpha(j_{\max})} r_n^{j_{\max}} \bigl\{\alpha''(j_{\max})^{-1/2} + \sum_{j>2j_{\max}} e^{-cj^\gamma}\bigr\}. \tag{4.36}$$

Since $j_{\max} \nearrow \infty$ as $n \to \infty$, we see that $\sum_{j>2j_{\max}} e^{-cj^\gamma} \ll \alpha''(j_{\max})^{-1/2}$ and thus that, for large n,

$$I_0(r_n) \leqslant C_3\, \alpha''(j_{\max})^{-1/2}\, e^{-\alpha(j_{\max})} r_n^{j_{\max}}. \tag{4.37}$$

As above, we also have

$$I_0(r_n) \geqslant e^{-\alpha(\lceil j_{\max}\rceil)} r_n^{\lceil j_{\max}\rceil} \geqslant e^{-\frac12 \alpha''(j_{\max})}\, e^{-\alpha(j_{\max})} r_n^{j_{\max}} \geqslant C_4\, e^{-\alpha(j_{\max})} r_n^{j_{\max}}. \tag{4.38}$$

The claimed asymptotics follow as before. □

Proof of Theorem 4.7. Let $\epsilon > 0$. It is sufficient to check that Theorem 4.2 applies with $a(n) = (1-\epsilon)j_{\max}$ and $b(n) = (1+\epsilon)j_{\max}$. It follows from Lemma 4.8 and Corollary 4.9 that

$$\frac1n \sum_{j=b(n)}^{\infty} e^{-\alpha(j)} r_n^{j+1/2} \leqslant n^{o(1)} \sum_{j=b(n)}^{\infty} e^{-c\alpha''(j_{\max})(j-j_{\max})^2}, \tag{4.39}$$

which goes to 0 as $n \to \infty$, since

$$e^{-c\alpha''(j_{\max})(b(n)-j_{\max})^2} = n^{-c\epsilon^2\gamma(1+o(1))}. \tag{4.40}$$

Similarly,

$$\frac1n \sum_{j=1}^{a(n)} e^{-\alpha(j)} r_n^{j+1/2} \leqslant n^{o(1)} \sum_{j=1}^{a(n)} e^{-c\alpha''(j_{\max})(j-j_{\max})^2} \leqslant n^{o(1)} e^{-c\alpha''(j_{\max}) j_{\max}^2 \epsilon^2}, \tag{4.41}$$

which again goes to 0 as $n \to \infty$. □

Acknowledgments: D.U. is grateful to Nick Ercolani and several members of the University of Arizona for many discussions about the Plancherel measure. D.U. also acknowledges the hospitality of the University of Geneva, ETH Zürich, the Center of Theoretical Studies of Prague, and the University of Arizona, where parts of this project were carried forward. V.B. is supported by the EPSRC fellowship EP/D07181X/1. D.U. is supported in part by the grant DMS-0601075 of the US National Science Foundation. Y.V. is supported in part by the Swiss NSF grant #200020-121675.

References

[1] R. Arratia, S. Tavaré, *The cycle structure of random permutations* Ann. Probab. 20, 1567–1591 (1992)
[2] F. Benaych-Georges, *Cycles of random permutations with restricted cycle lengths*, arxiv:0712.1903
[3] V. Betz, D. Ueltschi, *Spatial random permutations and infinite cycles*, Commun. Math. Phys. 285, 469-501 (2009)
[4] V. Betz, D. Ueltschi, *Spatial permutations with cycle weights*, preprint, to appear in Prob. Theory Rel. Fields (2009); arXiv:0812.0569
[5] W. J. Ewens, *The sampling theory of selectively neutral alleles*, Theoret. Populations Biol. 3, 87–112 (1972)
[6] S. Feng, F. M. Hoppe, *Large deviation principles for some random combinatorial structures in population genetics and Brownian motion*, Ann. Appl. Probab. 8, 975–994 (1998)
[7] P. Flajolet, R. Sedgewick, Analytic combinatorics, Cambridge University Press (2009)
[8] V.L. Goncharov, *Some facts from combinatorics*, Izv. Akad. Nauk SSSR, Ser. Mat. 8, 3–48 (1944)
[9] J. C. Hansen, *A functional central limit theorem for the Ewens sampling formula*, J. Appl. Prob. 27, 28–43 (1990)
[10] M. Lugo, *Profiles of permutations* Electr. J. Comb. 16, R99 (2009)
[11] A. N. Timashev, *Random permutations with cycle lengths in a given finite set*, Discrete Math. Appl. 18, 25–39 (2008)

Volker Betz and Daniel Ueltschi
Department of Mathematics
University of Warwick
Coventry, CV4 7AL, United Kingdom
http://www.maths.warwick.ac.uk/∼betz/
http://www.ueltschi.org
E-mail address: v.m.betz@warwick.ac.uk, daniel@ueltschi.org

Yvan Velenik
Section de Mathématiques
Université de Genève
1211 Genève 4, Switzerland
http://www.unige.ch/math/folks/velenik
E-mail address: Yvan.Velenik@unige.ch