Profiles of permutations

Michael Lugo

Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104
mlugo@math.upenn.edu

Submitted: Apr 2, 2009; Accepted: Jul 22, 2009; Published: Aug 7, 2009
Mathematics Subject Classification: 05A15, 05A16, 60C05

Abstract

This paper develops an analogy between the cycle structure of, on the one hand, random permutations with cycle lengths restricted to lie in an infinite set S with asymptotic density σ and, on the other hand, permutations selected according to the Ewens distribution with parameter σ. In particular we show that the asymptotic expected number of cycles of random permutations of $[n]$ with all cycles even, with all cycles odd, and chosen from the Ewens distribution with parameter 1/2 are all $\frac{1}{2}\log n + O(1)$, and the variance is of the same order. Furthermore, we show that in permutations of $[n]$ chosen from the Ewens distribution with parameter σ, the probability of a random element being in a cycle longer than γn approaches $(1-\gamma)^\sigma$ for large n. The same limit law holds for permutations with cycles carrying multiplicative weights with average σ. We draw parallels between the Ewens distribution and the asymptotic-density case and explain why these parallels should exist using permutations drawn from weighted Boltzmann distributions.

1 Introduction

In this paper we study the cycle structure of random permutations in which the lengths of all cycles are constrained to lie in some infinite set S, and permutations may be made more or less likely to be chosen through multiplicative weights placed on their cycles. Cycle structures viewed in this manner are a special case of certain measures on S_n which are conjugation-invariant and assign a weight to each element of S_n based on its cycle structure.

Definition 1.1. *Let $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots)$ be an infinite sequence of nonnegative real numbers. Then the weight of the permutation $\pi \in S_n$, with respect to $\vec{\sigma}$, is*

$$w_{\vec{\sigma}}(\pi) = \prod_{i=1}^{n} \sigma_i^{c_i(\pi)}$$

where $c_i(\pi)$ is the number of cycles of length i in π.

Informally, each cycle in a permutation receives a weight depending on its length, and the weight of a permutation is the product of the weights of its cycles. The sequence $\vec{\sigma}$ is called a *weighting sequence*.

For each positive integer n, let $(\Omega^{(n)}, \mathcal{F}^{(n)})$ be a probability space defined as follows. Take $\Omega^{(n)} = S_n$, the set of permutations of $[n]$, and let $\mathcal{F}^{(n)}$ be the set of all subsets of S_n. Endow $(\Omega^{(n)}, \mathcal{F}^{(n)})$ with a probability measure $\mathbb{P}^{(n)}_{\vec{\sigma}}$ for each weighting sequence $\vec{\sigma}$ as follows. Let $\mathbb{P}^{(n)}_{\vec{\sigma}}(\pi) = w_{\vec{\sigma}}(\pi)/\sum_{\pi' \in S_n} w_{\vec{\sigma}}(\pi')$; that is, each permutation has probability proportional to its weight. Extend $\mathbb{P}^{(n)}_{\vec{\sigma}}$ to all subsets of S_n by additivity. To streamline the notation, we will sometimes write $\mathbb{P}_{\vec{\sigma}}(\pi)$ for $\mathbb{P}^{(n)}_{\vec{\sigma}}(\pi)$. The sum of the weights of $\vec{\sigma}$-weighted permutations of $[n]$ is

$$\sum_{\pi \in S_n} w_{\vec{\sigma}}(\pi) = n![z^n] \exp\left(\sum_{k \geqslant 1} \sigma_k z^k/k\right)$$

by the exponential formula for labelled combinatorial structures.

We fix some notation. Define the random variable $X^{(n)}_k : \Omega^{(n)} \to \mathbb{Z}^+$ by setting $X^{(n)}_k(\pi)$ equal to the number of k-cycles in the permutation π. Let $X^{(n)}(\pi) = \sum_{k=1}^n X^{(n)}_k(\pi)$ be the total number of cycles. We will often suppress π and (n) in the notation, and we will write (for example) $\mathbb{P}_{\vec{\sigma}}(X_1 = 1)$ as an abbreviation for $\mathbb{P}_{\vec{\sigma}}(\{\pi : X^{(n)}_1(\pi) = 1\})$. Let $Y_k = kX_k$. We define Y_k in order to simplify the statement of some results.

This model incorporates various well-known classes of permutations, including generalized derangements (permutations in which a finite set of cycle lengths is prohibited), and the Ewens sampling formula from population genetics [8], which corresponds to the weighting sequence $(\sigma, \sigma, \sigma, \ldots)$. If $\vec{\sigma}$ is a 0-1 sequence with finitely many 1s, then this model specializes to random permutations of which all cycle lengths lie in a finite set. These have a fascinating structure studied by Benaych-Georges [4] and Timashev [25]; a typical permutation of $[n]$ with cycle lengths in a finite set S has about $\frac{1}{k}n^{k/\max S}$ k-cycles, for each k in S. In particular, most cycles are of length $\max S$, which may be unexpected at first glance. Analytically, this situation is studied via the asymptotics of $[z^n]e^{P(z)}$ where P is a polynomial, as done by Wilf [28]. Yakymiv [29] has studied the case, alluded to by Bender [5], in which $\vec{\sigma}$ is a sequence of 0s and 1s with a fixed density σ of 1s; the behavior of such permutations is in broad outline similar to that of the Ewens sampling formula with parameter σ. An "enriched" version of the model has been studied by Ueltschi and coauthors [15, 26]. In their model, permutations are endowed with a spatial structure. Each element of the ground set of the permutation is a point in the plane, and weights involve distances between points. Their "simple model of random permutations with cycle weight" [26, Sec. 2] is the model used here, where $\sigma_i = e^{-\alpha_i}$.

There are other combinatorially interesting conjugation-invariant measures on S_n, including permutations with all cycle lengths distinct [16], and permutations with kth roots for some fixed k [10, 22]. However the generating functions counting these classes are not exponentials of "nice" functions and thus different techniques are required.

Throughout this paper, we often implicitly assume that permutations under the uniform measure on S_n are the "primitive" structure, and weighted permutations are a perturbation of these. Here we follow Arratia *et al.* in [1, 2], in embracing a similar philosophy and viewing the permutation as the archetype of a class of "logarithmic combinatorial structures", and Flajolet and Soria's definition of functions of logarithmic type [14].

It will be convenient to use bivariate generating functions which count permutations by their size and number of cycles. In general, we take $F(z,u) = \sum_{n,k} f_{n,k} \frac{z^n}{n!} u^k$ to be the bivariate generating function, exponential in z and ordinary in u, of a combinatorial class $\mathcal{F}$, where $f_{n,k}$ is the number of objects in $\mathcal{F}$ of size n and with a certain parameter equal to k. In our case n will be the number of elements of a permutation, and k the total number of cycles or the number of cycles of a specified size. Then $[z^n] \left.\frac{\partial}{\partial u} F(z,u)\right|_{u=1} / [z^n]F(z,1)$ gives the expected value of the parameter k for an object of size n selected uniformly at random. The following lemma will frequently be useful, as it reduces the bivariate analysis to a univariate analysis.

Lemma 1.2. *Let $f(z)$ be the exponential generating function of permutations with weight sequence $\vec{\sigma}$. Then the expected number of k-cycles in a permutation chosen according to the measure $\mathbb{P}_{\vec{\sigma}}^{(n)}$ is*

$$\mathbb{E}_{\vec{\sigma}}^{(n)} X_k = \frac{\sigma_k}{k} \frac{[z^{n-k}]f(z)}{[z^n]f(z)}.$$

Proof. The bivariate generating function counting the cycles of such permutations is

$$\sigma_1 z + \sigma_2 \frac{z^2}{2} + \cdots + \sigma_{k-1} \frac{z^{k-1}}{k-1} + u\sigma_k \frac{z^k}{k} + \sigma_{k+1} \frac{z^{k+1}}{k+1} + \cdots$$

and this can be rewritten as $(u-1)\frac{\sigma_k z^k}{k} + \sum_{j \geqslant 1} \frac{\sigma_j z^j}{j}$. Thus, from the exponential formula, the bivariate generating function counting such permutations is

$$P(z,u) = \exp\left((u-1)\frac{\sigma_k z^k}{k} + \sum_{j\geqslant 1} \frac{\sigma_j z^j}{j} \right).$$

The expected number of cycles in a random permutation is $[z^n]P_u(z,1)/[z^n]P(z,1)$, giving the result. $\square$

The structure of this paper is as follows. In Section 2 we give exact formulas and asymptotic series (Propositions 2.2 and 2.3) for the mean and variance of the number of cycles of permutations chosen from the Ewens distribution. We also consider the average number of k-cycles in such permutations of $[n]$ for fixed k (Propositions 2.4 and 2.5) and for $k = \alpha n$ (Proposition 2.6). An "integrated" version of these results, Theorem 2.7, is one of the main results; this is a limit law for the probability that a random element of a weighted permutation is in a cycle within a certain prescribed range of lengths. In Section 3 we derive similar results for permutations in which all cycle lengths have the same parity. In addition, we determine the mean and variance of the number of cycles of such permutations (Theorem 3.6 treats the odd case, and Theorem 3.8 treats the even case). In Section 4 we explore connections to the generation of random objects by Boltzmann sampling. The main theorem of this section, Theorem 4.3, states that the

Boltzmann-sampled permutations of a certain class of approximate size n, including the Ewens and parity-constrained cases, have their number of cycles distributed with mean and variance approximately a constant multiple of $\log n$.

2 The Ewens sampling formula and Bernoulli decomposition

The *Ewens distribution* [8] on permutations of $[n]$ with parameter σ gives to each permutation π probability proportional to $\sigma^{X(\pi)}$. This corresponds to the weighting sequence $\vec{\sigma} = (\sigma, \sigma, \sigma, \ldots)$; we will write $\mathbb{P}_\sigma^{(n)}, \mathbb{E}_\sigma^{(n)}$ for $\mathbb{P}_{\vec{\sigma}}^{(n)}, \mathbb{E}_{\vec{\sigma}}^{(n)}$, and call a random permutation selected in this manner a *σ-weighted permutation.* In this section we derive formulas for the mean and variance of the number of cycles of permutations chosen from the Ewens distribution. Note that the number of cycles can be decomposed into a sum of independent Bernoulli random variables. Similar decompositions are due to Arratia et al. in [2, Sec. 5.2] for general σ, and Feller [9, (46)] for $\sigma = 1$; the fact that the number of cycles is normally distributed is seen in [14, Example 1]. Thus this section is largely expository; the proofs are provided for the purpose of comparison with other proofs to be given below. The asymptotic series for $\mathbb{E}_\sigma^{(n)}$ and $\mathbb{V}_\sigma^{(n)}$ appear to be new.

Theorem 2.1. *[20, Exercise 3.2.3] The distribution of the random variable X under the measure $\mathbb{P}_\sigma^{(n)}$ is that of the sum $\sum_{k=1}^n Z_k$, where the Z_k are independent random variables and Z_k has the Bernoulli distribution with mean $\sigma/(\sigma + k - 1)$.*

Proof. The generating function of permutations of $[n]$ counted by their number of cycles is $\sum_{k=1}^n S(n,k)u^k = u(u+1)(u+2)\cdots(u+n-1)$, where $S(n,k)$ are the Stirling cycle numbers. Replacing u with σu and normalizing gives the probability generating function for the number of cycles,

$$\sum_{k=1}^{n} S(n,k)\sigma^k u^k = \frac{\sigma u}{\sigma}\frac{\sigma u + 1}{\sigma + 1}\cdots\frac{\sigma u + n - 1}{\sigma + n - 1},$$

and each factor is the probability generating function for a Bernoulli random variable. $\square$

Combinatorially, we can envision this Bernoulli decomposition as follows. We imagine forming a permutation of $[n]$ by placing the elements $1, \ldots, n$ in cycles in turn. When the element k is inserted, with probability $\sigma/(\sigma + k - 1)$ it is placed in a new cycle, and with probability $1/(\sigma + k - 1)$ it is placed after any of $1, 2, \ldots, k-1$ in the cycle containing that element. Then the probability of obtaining any permutation with c cycles is $\sigma^c/(\sigma(\sigma+1)\cdots(\sigma+n-1))$, which is exactly the measure given to this permutation by $\mathbb{P}_\sigma^{(n)}$. This is an instance of the Chinese Restaurant Process [20, Sec. 3.1].

From this decomposition into Bernoulli random variables, we can derive formulas for the mean and variance of the number of cycles under the measure $\mathbb{P}_\sigma^{(n)}$. In particular we note that since X is a sum of Bernoulli random variables with small mean, the variance of

X is very close to its mean. Let ψ denote the digamma function $\psi(z) = \Gamma'(z)/\Gamma(z)$; this has an asymptotic series $\psi(z) = \log z - \frac{1}{2}z^{-1} - \frac{1}{12}z^{-2} + O(z^{-4})$ as $z \to \infty$. Let $H_n = \sum_{k=1}^{n} \frac{1}{k}$ be the nth harmonic number and let $\gamma = 0.57721\ldots$ be the Euler-Mascheroni constant.

Proposition 2.2. *The expected number of cycles of a random σ-weighted permutation of $[n]$ is $\mathbb{E}_\sigma^{(n)} X = \sigma(\psi(n+\sigma) - \psi(\sigma))$; in particular if σ is a positive integer we have*

$$\mathbb{E}_\sigma^{(n)} X = \sigma \log n + (\sigma\gamma - \sigma H_{\sigma-1}) + (\sigma^2 - \sigma/2)n^{-1} + O(n^{-2}). \tag{1}$$

Proof. From Theorem 2.1 we have

$$\mathbb{E}_\sigma^{(n)} X = \sum_{k=1}^{n} \frac{\sigma}{\sigma+k-1} = \sigma \sum_{k=1}^{n} \frac{1}{\sigma+k-1}.$$

Now, $\psi(z+1) - \psi(z) = 1/z$; thus

$$\begin{aligned} \psi(n+\sigma) - \psi(\sigma) &= (\psi(n+\sigma) - \psi(n+\sigma-1)) + \cdots + (\psi(\sigma+1) - \psi(\sigma)) \\ &= \frac{1}{n+\sigma-1} + \frac{1}{n+\sigma-2} + \cdots + \frac{1}{\sigma} \\ &= \sum_{k=1}^{n} \frac{1}{\sigma+k-1}. \end{aligned}$$

This proves that $\mathbb{E}_\sigma^{(n)} X = \sigma(\psi(n+\sigma) - \psi(\sigma))$. The asymptotic series follows from that for $\psi(z)$ where we have used the fact that $\psi(n) = H_{n-1} - \gamma$ when n is a positive integer. $\square$

Proposition 2.3. *The variance of the number of cycles of a random σ-weighted permutation of $[n]$ is*

$$\sigma^2 (\psi'(n+\sigma) - \psi'(\sigma)) + \sigma(\psi(n+\sigma) - \psi(\sigma)); \tag{2}$$

this has an asymptotic series,

$$\mathbb{V}_\sigma^{(n)} X = \sigma \log n + (-\sigma^2\psi'(\sigma) - \sigma\psi(\sigma)) + \frac{4\sigma^2 - 1}{2} n^{-1} + O(n^{-2}) \tag{3}$$

The proof is similar to that of the previous proposition, noting that the variance of a Bernoulli random variable with mean p is $p - p^2$.

From (3) we can also derive for integer σ the explicit formula (not involving ψ)

$$\mathbb{V}_\sigma^{(n)} X = -\sigma^2 \sum_{j=\sigma}^{\sigma+n-1} \frac{1}{j^2} + \sigma (\log n + \gamma - H_{\sigma-1}) + O(1/n)$$

which holds as $n \to \infty$. It suffices to show that

$$\psi'(n+\sigma) - \psi'(\sigma) = -\sum_{j=\sigma}^{\sigma+n-1} \frac{1}{j^2}. \tag{4}$$

To see this, recall the identity $\psi(x+1)-\psi(x)=1/x$; differentiating gives $\psi'(x+1)-\psi'(x)=-1/x^2$. Summation over $x=\sigma,\sigma+1,\ldots,\sigma+n-1$ gives (4).

Finally, we recall a normal distribution result for the total number of cycles [2, (5.22)]. Let $\hat{X}=\frac{X-\sigma\log n}{\sqrt{\sigma\log n}}$ be the standardization of X. Then $\lim_{n\to\infty}\mathbb{P}_\sigma^{(n)}(\hat{X}\leqslant x)=\Phi(x)$, where $\Phi(x)$ is the cumulative distribution function of a standard normal random variable. This follows from Theorem 2.1 and the Lindeberg-Feller central limit theorem.

We have thus far looked at the total number of cycles of σ-weighted permutations. These distributions, suitably scaled, are continuous in the large-n limit. In contrast, looking at each cycle length separately, we approach a discrete distribution. More specifically, the number of k-cycles of σ-weighted permutations of $[n]$, for large n, converges in distribution to $\mathcal{P}(\sigma/k)$, where $\mathcal{P}(\lambda)$ denotes a Poisson random variable with mean λ; here we consider how quickly $\mathbb{E}_\sigma^{(n)}X_k$ approaches σ/k. Recall that X_k is a random variable, with $X_k(\pi)$ the number of k-cycles of a permutation π.

Proposition 2.4. *[3, (37)][27] The average number of k-cycles in a σ-weighted permutation of $[n]$ is*

$$\mathbb{E}_\sigma^{(n)}X_k=\frac{\sigma}{k}\frac{(n)_k}{(n+\sigma-1)_k} \tag{5}$$

where $(n)_k=n(n-1)\ldots(n-k+1)$ is the “falling power”.

We provide a new proof in terms of generating functions.

Proof. The bivariate generating function counting σ-weighted permutations by their size and number of k-cycles is $P(z,u)=(1-z)^{-u\sigma}\exp(\sigma(u-1)z^k/k)$. The mean number of k-cycles is given by

$$\frac{[z^n]\,\partial_u P(z,u)|_{u=1}}{[z^n]P(z,1)}=\frac{[z^n]\frac{\sigma z^k}{k}(1-z)^{-u\sigma}}{[z^n](1-z)^{-\sigma}}=\frac{\sigma}{k}\frac{[z^{n-k}](1-z)^{-\sigma}}{[z^n](1-z)^{-\sigma}}$$

and the binomial formula gives (5). ☐

When σ is an integer, a combinatorial proof can be obtained by considering σ-weighted permutations as permutations where each cycle is colored in one of σ colors.

Proposition 2.5. *There is an asymptotic series for $\mathbb{E}_\sigma^{(n)}X_k$,*

$$\mathbb{E}_\sigma^{(n)}X_k=\frac{\sigma}{k}\left(1-\frac{(\sigma-1)k}{n}+O(n^{-2})\right).$$

Proof. The numerator and denominator of (5) are polynomials in n of degree k; write the two highest-degree terms of each explicitly and divide. ☐

Proposition 2.6. *Fix $0<\alpha\leqslant 1$. The expected number of elements in αn-cycles of a random σ-weighted permutation satisfies, as $n\to\infty$,*

$$\mathbb{E}_\sigma^{(n)}Y_{\alpha n}=\sigma(1-\alpha)^{\sigma-1}+O(n^{-1})$$

(Here we have assumed for simplicity that αn is an integer.)

Proof. Let $\beta = 1-\alpha$. We have from Proposition 2.4 that

$$\mathbb{E}_\sigma^{(n)} Y_{\alpha n} = \sigma \frac{(n)_{\alpha n}}{(n+\sigma-1)_{\alpha n}} = \sigma \frac{n!(\beta n + \sigma - 1)!}{(\beta n)!(n+\sigma-1)!} = \sigma \frac{n!}{(n+\sigma-1)!} \frac{(\beta n+\sigma-1)!}{(\beta n)!}$$

We now note that $(n+r)!/n! = n^r(1+O(n^{-1}))$, for constant r as $n \to \infty$, from Stirling's formula. Applying this twice with $r = \sigma - 1$ gives the result. □

It would be of interest to determine the limiting distribution of the number of cycles with length between γn and δn for constants γ and δ. There can be at most $\lfloor \gamma^{-1} \rfloor$ such cycles, so this random variable is supported on $0, 1, \ldots, \lfloor \gamma^{-1} \rfloor$. Thus to determine the limiting distribution it suffices to determine the 0th through $\lfloor \gamma^{-1} \rfloor$th moments of this random variable. The $\sigma = 1$ case will be treated in [19].

We can essentially integrate the result of Proposition 2.5 to determine the number of elements in cycles with normalized length in a specified interval. However, this can be done in a more general framework. Following [13, Thm. VI.1] and [12, Thm. 1], for constants $R > 1$ and $\phi > 0$ we define a Δ-domain as a set of the form

$$\Delta(\phi, R) = \{z : |z| < R, z \neq 1, |\arg(z-1)| > \phi\}.$$

Theorem 2.7. *Let $\sum_k \sigma_k z^k/k = \sigma \log \frac{1}{1-z} + K + o(1)$ be analytic in its intersection with some Δ-domain, for some constants σ and K. Then the probability that a uniformly chosen random element of a random $\vec{\sigma}$-weighted permutation of $[n]$ lies in a cycle of length between γn and δn approaches $(1-\gamma)^\sigma - (1-\delta)^\sigma$ as $n \to \infty$.*

Note that analyticity in the slit plane suffices; this is the case $\phi = 0$. We begin by stating two lemmas needed in the proof.

Lemma 2.8. *Let $\{\sigma_k\}_{k=1}^\infty$ be a sequence of nonnegative real numbers with mean σ. Fix constants $0 \leqslant \gamma < \delta < 1$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=\lceil \gamma n \rceil}^{\lfloor \delta n \rfloor} \sigma_k \left(1 - \frac{k}{n}\right)^{\sigma-1} = (1-\gamma)^\sigma - (1-\delta)^\sigma.$$

Proof. We rewrite the sum as an integral,

$$\sum_{k=\lceil \gamma n \rceil}^{\lfloor \delta n \rfloor} \sigma_k \left(1 - \frac{k}{n}\right)^{\sigma-1} = \int_{\gamma n}^{\delta n} \left(1 - \frac{k}{n}\right)^{\sigma-1} d\mu(k)$$

where $\mu(x) = \sum_{j=1}^{\lfloor x \rfloor} \sigma_j$. Integrating by parts gives

$$(1-\delta)^{\sigma-1}\mu(\delta n) - (1-\gamma)^{\sigma-1}\mu(\gamma n) - \int_{\gamma n}^{\delta n} \mu(k) d\left(1 - \frac{k}{n}\right)^{\sigma-1}. \tag{6}$$

Differentiation allows us to rewrite the integral in (6) as a Riemann integral,

$$\int_{\gamma n}^{\delta n} \mu(k) d\left(1-\frac{k}{n}\right)^{\sigma-1} = \frac{1-\sigma}{n}\int_{\gamma n}^{\delta n} \mu(k)\left(1-\frac{k}{n}\right)^{\sigma-2} dk. \tag{7}$$

Let $\tau(k)=\mu(k)-\sigma k$. Then the integral on the right-hand side of (7) becomes

$$\frac{1-\sigma}{n}\left(\int_{\gamma n}^{\delta n} \sigma\left(1-\frac{k}{n}\right)^{\sigma-2} dk + \int_{\gamma n}^{\delta n} \tau(k)\left(1-\frac{k}{n}\right)^{\sigma-2}\right) dk. \tag{8}$$

We perform the first integral in (8) and note that $\mu(\delta n)\sim\sigma\cdot\delta n, \mu(\gamma n)\sim\sigma\cdot\gamma n$ in (6). This gives

$$\frac{1}{n}\sum_k\left(1-\frac{k}{n}\right)^{\sigma-1} \sim (1-\gamma)^\sigma-(1-\delta)^\sigma+\frac{1-\sigma}{n^2}\int_{\gamma n}^{\delta n}\tau(k)\left(1-\frac{k}{n}\right)^{\sigma-2} dk. \tag{9}$$

So it suffices to show that the final term in (9) is negligible, i. e.

$$\int_{\gamma n}^{\delta n}\tau(k)\left(1-\frac{k}{n}\right)^{\sigma-2} dk = o(n^2).$$

Since $\{\sigma_k\}_{k=1}^\infty$ has mean σ, we have $\sum_{k=1}^n \sigma_k = \sigma n + o(n)$. Thus $\tau(k)=o(n)$. On $[\gamma n, \delta n]$, $(1-k/n)^{\sigma-2}$ is bounded. So the integrand above is $o(n)$, and the integral is $o(n^2)$ as desired. □

Lemma 2.9. *Say $[z^n]P(z)=Cn^{\sigma-1}(1+o(1))$ uniformly in n, for some positive constants C,σ. Then*

$$\sum_{k=\lceil\gamma n\rceil}^{\lfloor\delta n\rfloor}\sigma_k\frac{[z^{n-k}]P(z)}{[z^n]P(z)}\sim\sum_{k=\lceil\gamma n\rceil}^{\lfloor\delta n\rfloor}\sigma_k\left(1-\frac{k}{n}\right)^{\sigma-1}$$

as $n\to\infty$, for any $0\leqslant\gamma<\delta<1$.

Proof. From the hypothesis that $[z^n]P(z)\sim Cn^{\sigma-1}$, we get

$$\frac{[z^{n-k}]P(z)}{[z^n]P(z)}\sim\frac{C(n-k)^{\sigma-1}}{Cn^{\sigma-1}}=\left(1-\frac{k}{n}\right)^{\sigma-1}$$

uniformly as $n,k\to\infty$ with $0\leqslant k<\delta n$. Therefore

$$\begin{aligned}\sum_{k=\lceil\gamma n\rceil}^{\lfloor\delta n\rfloor}\sigma_k\frac{[z^{n-k}]P(z)}{[z^n]P(z)} &= \sum_{k=\lceil\gamma n\rceil}^{\lfloor\delta n\rfloor}\sigma_k\left(1-\frac{k}{n}\right)^{\sigma-1}(1+o(1))\\ &= \sum_{k=\lceil\gamma n\rceil}^{\lfloor\delta n\rfloor}\sigma_k\left(1-\frac{k}{n}\right)^{\sigma-1}+\sum_{k=\lceil\gamma n\rceil}^{\lfloor\delta n\rfloor}\sigma_k\cdot o(1)\cdot\left(1-\frac{k}{n}\right)^{\sigma-1}.\end{aligned}$$

The first sum in the previous equation is $\Theta(n)$. The second sum has $\Theta(n)$ terms; since $(1-k/n)^{\sigma-1}$ and σ_k can both be bounded above on the interval $[\gamma n, \delta n]$ each term is $o(1)$. Thus the second sum is $o(n)$. So

$$\begin{aligned}
\sum_{k=\lceil \gamma n \rceil}^{\lfloor \delta n \rfloor} \sigma_k \frac{[z^{n-k}]P(z)}{[z^n]P(z)} &= \sum_{k=\lceil \gamma n \rceil}^{\lfloor \delta n \rfloor} \left(\sigma_k \left(1 - \frac{k}{n}\right)^{\sigma-1} \right) + o(n) \\
&= \sum_{k=\lceil \gamma n \rceil}^{\lfloor \delta n \rfloor} \left(\sigma_k \left(1 - \frac{k}{n}\right)^{\sigma-1} \right) (1+o(1))
\end{aligned}$$

as desired. □

Proof of Theorem 2.7. This probability can be written as

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=\lceil \gamma n \rceil}^{\lfloor \delta n \rfloor} \sigma_k \frac{[z^{n-k}]P(z)}{[z^n]P(z)}.$$

Now, recall

$$\sum_k \sigma_k z^k/k = \sigma \log \frac{1}{1-z} + K + o(1)$$

by hypothesis. Thus the generating function $P(z)$ of $\vec{\sigma}$-weighted permutations is

$$P(z) = \exp\left(\sum_k \sigma_k z^k/k\right) = \exp\left(\sigma \log \frac{1}{1-z} + K + o(1)\right) = (1-z)^{-\sigma} e^K (1+o(1)).$$

Applying the Flajolet-Odlyzko transfer theorem [12], $[z^n]P(z) = Cn^{\sigma-1}(1+o(1))$ for some positive real constant C. Thus $P(z)$ satisfies the hypotheses of Lemma 2.9. Applying that lemma, we see that this sum is asymptotic to $n^{-1}\sum_{k=\lceil \gamma n \rceil}^{\lfloor \delta n \rfloor} \sigma_k (1-k/n)^{\sigma-1}$; the desired result then follows from Lemma 2.8. □

The hypotheses, and hence the conclusions, of Theorem 2.7 hold for many weight sequences $\sigma_1, \sigma_2, \ldots$ with $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \sigma_k = \sigma$; that is, for weight sequences averaging σ. In particular, we have the following special case.

Corollary 2.10. *Fix constants $0 \leqslant \gamma \leqslant \delta \leqslant 1$. Let $p_\sigma(n; \gamma, \delta)$ be the probability that the element 1, in a σ-weighted permutation of $[n]$, lies in a cycle of length in the interval $[\gamma n, \delta n]$. Then*

$$\lim_{n\to\infty} p_\sigma(n; \gamma, \delta) = (1-\gamma)^\sigma - (1-\delta)^\sigma.$$

Proof. We have the cycle generating function $\sum_{k=1}^\infty \sigma z^k/k = \sigma \log \frac{1}{1-z}$; apply Theorem 2.7. □

For example, setting $\sigma = 1/2, \gamma = 0.99, \delta = 1$, we see that for large n, 10% of elements of 1/2-weighted permutations are in cycles of length at least $0.99n$. If we define the "co-length" of a cycle of a permutation to be the number of elements *not* in that cycle, a cleaner statement of the theorem becomes possible. The proportion of elements of σ-weighted permutations in cycles of co-length at most ζn is ζ^σ.

It would be desirable to replace the condition in the hypothesis of Theorem 2.7 with the less restrictive

$$\sum_k \frac{\sigma_k z^k}{k} = \sigma \log \frac{1}{1-z} \cdot (1 + o(1));$$

it seems likely that this suffices to prove a limit law but the proof does not easily adapt to that case.

3 Permutations with all cycle length of the same parity

This section is devoted to results on random permutations in which all cycle lengths have the same parity; that is, they are either all even or all odd. We adopt the notation $\mathbb{P}_e^{(n)}$ for the family of measures $\mathbb{P}_{\vec{\sigma}}^{(n)}$ where $\vec{\sigma} = (0, 1, 0, 1, \ldots)$, and similarly $\mathbb{P}_o^{(n)}$ for the family with $\vec{\sigma} = (1, 0, 1, 0, \ldots)$; these are the measures corresponding to permutations with all cycle lengths even and with all cycle lengths odd, respectively.

The results obtained here resemble those for the Ewens sampling formula with parameter 1/2. A heuristic explanation for this phenomenon is as follows. Let us produce a permutation of $[n]$ from the Ewens distribution with parameter 1/2 by first picking a permutation π uniformly at random from S_n, and then flipping a fair coin for each cycle of π. If all the coins come up heads we keep π; otherwise we "throw back" the permutation π and repeat this process until we have a trial in which all coins come up heads. The number and normalized size of cycles of permutations obtained in this manner should be similar to those of permutations with all cycle lengths even, since for large permutations the parity constraint is essentially equivalent to a coin flip.

Proposition 3.1. *The expected number of elements in k-cycles of a permutation of $[n]$ with all cycle lengths even is*

$$\mathbb{E}_e^{(n)} Y_k = \frac{n(n-2)\cdots(n-k+2)}{(n-1)(n-3)\cdots(n-k+1)}$$

if k is even, and 0 if k is odd.

Proof. By Lemma 1.2, we have $\mathbb{E}_e^{(n)} Y_k = [z^{n-k}](1-z^2)^{-1/2}/[z^n](1-z^2)^{-1/2}$; we apply the binomial theorem and simplify. □

For example, when $n = 10$ we have

$$\begin{aligned}\left(\mathbb{E}_e^{(10)} Y_2, \mathbb{E}_e^{(10)} Y_4, \ldots, \mathbb{E}_e^{(10)} Y_{10}\right) &= (10/9, 80/63, 32/21, 128/63, 256/63) \\ &\approx (1.11, 1.27, 1.52, 2.03, 4.06)\end{aligned}$$

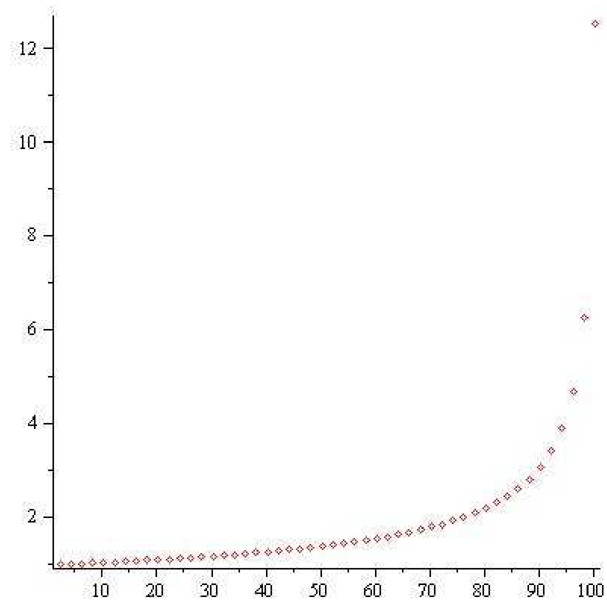

Figure 1: $\mathbb{E}_e^{(100)} Y_k$ for $k = 2, 4, \ldots, 100$.

and we observe that most entries are in the longer cycles. For $n = 100$ this is illustrated in the figure above.

Proposition 3.2. *The expected number of elements in k-cycles of a random permutation of $[n]$ with all cycle lengths odd is*

$$\mathbb{E}_o^{(n)} Y_k = \begin{cases} \frac{n(n-2)\cdots(n-k+1)}{(n-1)(n-3)\cdots(n-k)}, & n \text{ even} \\ \frac{(n-1)(n-3)\cdots(n-k+2)}{(n-2)(n-4)\cdots(n-k+1)}, & n \text{ odd} \end{cases}$$

Proof. The generating function of permutations with all cycle lengths odd, counted by their number of cycles, is $P(z,u) = ((1+z)/(1-z))^{u/2}$. We use Lemma 1.2 to see that the mean number of elements in k-cycles is given by $\mathbb{E}_o^{(n)} Y_k = [z^{n-k}]\sqrt{\frac{1+z}{1-z}}/[z^n]\sqrt{\frac{1+z}{1-z}}$. We recall that $[z^n]\sqrt{\frac{1+z}{1-z}}$ is $\frac{(n-1)!!^2}{n!}$ if n is even and $\frac{n!!(n-2)!!}{n!}$ if n is odd; substituting and simplifying gives the result. $\square$

By similar methods, we can obtain formulas for the exact number of permutations of $[n]$ with all cycle lengths divisible by a, and the exact expected number of k-cycles in such permutations for integers k which are divisible by a. These permutations have exponential generating function $(1-z^a)^{-1/a}$. Permutations with all cycle lengths congruent to k mod a for some nonzero k are more difficult to deal with, as it appears that the generating function cannot be written in an elementary form except when a is even and $k = a/2$. (See [23, Sec. 5.0.3] for the relevant generating functions.)

Proposition 3.3. *(a) The number of elements of k-cycles in a permutation of $[n]$ with all cycle lengths odd, for fixed odd k, satisfies $\mathbb{E}_o^{(n)} Y_k = 1 + \frac{k+1}{2n} + O(n^{-2})$ as n approaches ∞ through even values, and $\mathbb{E}_o^{(n)} Y_k = 1 + \frac{k-1}{2n} + O(n^{-2})$ as n approaches ∞ through odd values.*

(b) The number of elements of k-cycles in a permutation of $[n]$ with all cycle lengths even, for fixed even k, satisfies $\mathbb{E}_e^{(n)} Y_k = 1 + \frac{k}{2n} + O(n^{-2})$ as n approaches infinity through even values.

Proof. To prove (a), from Proposition 3.2 we have previous formulas for $\mathbb{E}_o^{(n)} Y_k$ depending on the parity of n. These are fractions which have numerators and denominators which are polynomials in n; we can write out the two highest-degree terms of each polynomial and simplify. To prove (b) we proceed similarly from Proposition 3.1. □

Note that the expected number of elements in k-cycles of permutations with all cycle lengths even (or odd) approaches 1 as n gets large, if k has the appropriate parity. Assume we are dealing with permutations with all cycle lengths even. Naively, we might add the limits of the expected number of elements in 2-cycles, 4-cycles, ..., n-cycles, and expect to get n. But these are each 1; their sum is $n/2$. Since each element is in a cycle, we must have $\sum_{k=1}^{n} \mathbb{E}_e^{(n)} Y_k = n$. The difficulty is that the convergence of $\mathbb{E}_e^{(n)} Y_k$ as $n \to \infty$ is not uniform over k. Under the correct scaling, then, subtler phenomena can be seen; the "missing" elements end up disproportionately in the longer cycle lengths for permutations with all cycle lengths even. We note that similar phenomena of nonuniform convergence have previously been observed in random mappings, for example in [11].

Proposition 3.4. *Fix $\epsilon \in (0,1)$. The expected number of elements in k-cycles in a random permutation of $[n]$ with all cycle lengths even satisfies uniformly*

$$\mathbb{E}_e^{(n)} Y_k \to \left(1 - \frac{k}{n}\right)^{1/2}$$

as $k, n \to \infty$ with $0 < k/n < 1 - \epsilon$.

Proof. The result of Proposition 3.1 can be rewritten in terms of factorials as

$$\mathbb{E}_e^{(n)} Y_k = 2^k \left(\frac{(n/2)!}{((n-k)/2)!} \right)^2 \frac{(n-k)!}{n!}$$

and by Stirling's approximation and routine simplifications, we have

$$\mathbb{E}_e^{(n)} Y_k = \left(1 - \frac{k}{n}\right)^{-1/2} \frac{1 + \frac{1}{4n} + O(n^{-2})}{1 + \frac{1}{4(n-k)} + O((n-k)^{-2})}. \tag{10}$$

Let $n, k \to \infty$ with $0 < k/n < 1 - \epsilon$. Then we have $1/(4(n-k)) \in [(4n)^{-1}, (4\epsilon n)^{-1})]$, and so $O((n-k)^{-2}) = O(n^{-2})$. Furthermore $1/(4(n-k)) = O(n^{-1})$, with the constant implicit in the O-notation being $(4\epsilon)^{-1}$. Therefore

$$\mathbb{E}_e^{(n)} Y_k = \left(1 - \frac{k}{n}\right)^{-1/2} (1 + O(n^{-1}))$$

uniformly, as $k, n \to \infty$ with $0 < k/n < 1 - \epsilon$. $\square$

Furthermore, we can essentially integrate the result of Proposition 3.4 to determine the cumulative distribution function of the length of the cycle containing a random element of a random permutation with all cycle lengths even (or odd). This is the content of the next theorem.

Theorem 3.5. *Fix constants $0 \leqslant \gamma \leqslant \delta \leqslant 1$. Let $p_e(n;\gamma,\delta)$ be the probability that 1 is contained in a cycle of length between γn and δn of a permutation chosen uniformly at random from all permutations of $[n]$ with all cycle lengths even. Then*

$$\lim_{n\to\infty} p_e(n;\gamma,\delta) = \sqrt{1-\gamma} - \sqrt{1-\delta}.$$

Since the measure $\mathbb{P}_e$ is invariant under conjugation, this is the probability that an element of $[n]$ chosen uniformly at random is in a cycle of length between γn and δn in a random permutation of $[n]$ with all cycle lengths even.

Proof. Note that

$$\sum_{2|k} \frac{z^k}{k} = \frac{1}{2}\log\frac{1+z}{1-z} = \frac{1}{2}\log\frac{1}{1-z} + \log 2 + o(1)$$

and apply Theorem 2.7. $\square$

The same is true for permutations with all cycle lengths odd; like those with all cycle lengths even they fall in the "$\sigma = 1/2$ class".

We now move to consider the mean and variance of the total number of cycles of all lengths.

Theorem 3.6. *The mean number of cycles of a randomly chosen permutation of $[n]$ with all cycle lengths odd is, as $n \to \infty$,*

$$\frac{1}{2}\log n + \frac{\gamma + 3\log 2}{2} \pm \frac{\gamma + \log n}{8n} + O\left(\frac{\log n}{n^2}\right)$$

where we take the $+$ sign if n is odd and the $-$ sign if n is even. The variance of the number of cycles is, as $n \to \infty$,

$$\frac{1}{2}\log n + \frac{\gamma + 3\log 2 - 4\pi^2}{8} + O\left(\frac{\log^2 n}{n}\right).$$

Proof. We have the exponential generating function counting such permutations by size and number of cycles, $\left(\frac{1+z}{1-z}\right)^{u/2}$. We can differentiate to obtain the mean and variance of the number of cycles. These are given by

$$\mu_n := \frac{[z^n]\frac{1}{2}\sqrt{\frac{1+z}{1-z}}\log\frac{1+z}{1-z}}{[z^n]\sqrt{\frac{1+z}{1-z}}}, \sigma_n^2 := \frac{[z^n]\frac{1}{4}\sqrt{\frac{1+z}{1-z}}\log^2\frac{1+z}{1-z}}{[z^n]\sqrt{\frac{1+z}{1-z}}} + \mu_n - \mu_n^2.$$

Let $f_r(z) = \sqrt{\frac{1+z}{1-z}} \log^r \frac{1+z}{1-z}$ for $r = 0, 1, 2$ and let $a_r(n) = [z^n] f_r(z)$ for $r = 0, 1, 2$. Then we have

$$\mu_n = \frac{a_1(n)}{2a_0(n)}, \sigma_n^2 = \frac{a_2(n)}{4a_0(n)} + \mu_n - \mu_n^2 \tag{11}$$

and we need to find asymptotic series for the $a_r(n)$ as $n \to \infty$. We observe that $a_0(n)$ is the number of permutations of $[n]$ with all cycle lengths odd, which is $(n-1)!!^2/n!$ if n is even and $n!!(n-2)!!/n!$ if n is odd; Stirling's formula gives an asymptotic expansion, depending on the parity of n. To find a series for $a_1(n)$ as $n \to \infty$, we expand $f_1(z) = \sqrt{\frac{1+z}{1-z}}$ in a series with terms which are half-integral powers of $1 - z$. From this we derive a series for $\sqrt{\frac{1+z}{1-z}} \log \frac{1+z}{1-z}$ with terms of the form $(1 - z)^{i-1/2} L^j$ where $L = \log 1/(1 - z)$. The function being expanded is analytic in the complex plane slit along the real half-line $\{z \in \mathbb{R} : z \geqslant 1\}$; by [13, Thm. VI.3], an error of $O((1 - z)^{i-1/2} L)$ in the series for $f_1(z)$ leads to an error $O(n^{-i-1/2} \log n)$ in the series for $a_1(n)$. We can thus transfer an asymptotic expansion for $f_1(z)$ near $z = 1$ to give an expansion for $a_1(n)$ as $n \to \infty$, and similarly for $f_2(z)$ and $a_2(n)$. Combining these series as specified by (11) gives the result. □

We observe that this is $(\log 2) + o(1)$ more than the number of cycles of a permutation of $[n]$ with all cycle lengths even.

The following two results give a decomposition of the number of cycles of permutations with all cycle lengths even into a sum of Bernoulli random variables.

Theorem 3.7. *The generating function of permutations of $[2n]$ with all cycle lengths even, counted by their number of cycles, is*

$$p_{2n}(u) = [u(u + 2)(u + 4) \cdots (u + (2n - 2))] \cdot (2n - 1)!! \tag{12}$$

Proof. The bivariate generating function for permutations with all cycle lengths even, counted by their size and number of cycles, is $(1 - z^2)^{-u/2}$. Let $p_k(u)$ be the desired generating function. Then we have

$$(1 - z^2)^{-u/2} = p_0(u) + p_1(u)z + p_2(u)\frac{z^2}{2!} + \cdots$$

and it is clear that p_k is the zero polynomial for odd k. For even k, the binomial theorem gives

$$(1 - z^2)^{-u/2} = 1 + \binom{-u/2}{1}(-z^2) + \binom{-u/2}{2}(-z^2)^2 + \cdots$$

and so we have $p_{2n}(u) = (2n)!\binom{-u/2}{n}$ by comparing coefficients; this can be expanded to give the expression above. □

A combinatorial proof is also possible. Recall that we can write a permutation π of $[n]$ in terms of its *inversion table*, a sequence of integers $a_1, a_2, \ldots, a_n$, with $a_i = |\{j : j < i, \pi(j) > \pi(i)\}|$. The number of zeros in the sequence $(a_1, \ldots, a_n)$ is the number of left-to-right maxima of π. The "fundamental correspondence" between permutations written in cycle notation and in one-line notation takes permutations with k cycles to those with k left-to-right maxima; furthermore, permutations with all cycle lengths even are taken to those with all left-to-right maxima in odd positions, and conversely. Thus it suffices to show that $p_n(u)$ is the generating function of permutations with all left-to-right maxima in odd positions, counted by their number of maxima; this is done by considering the inversion table.

Theorem 3.8. *The number of cycles C_n of a random permutation of $[2n]$ with all cycle lengths even, as $n \to \infty$, is asymptotically normally distributed with*

$$\mathbb{E}(C_n) = \frac{1}{2}\log n + \left(\frac{1}{2}\gamma + \log 2\right) + O(n^{-1}) \tag{13}$$

and

$$\mathbb{V}(C_n) = \frac{1}{2}\log n + \left(\frac{1}{2}\gamma + \log 2 - \frac{\pi^2}{8}\right) + O(n^{-1}) \tag{14}$$

Proof. Let $n = 2m$. From Theorem 3.7, we have $C_m = \sum_{k=1}^m X_{m,k}$ where the $X_{m,k}$ are independent Bernoulli random variables with $\mathbb{P}(X_{m,k} = 1) = 1/(2k-1)$. The formula (13) for the expectation follows from the asymptotic series for the harmonic numbers. The variance is given by

$$\mathbb{V}C_m = \sum_{k=1}^m \left(\frac{1}{2k-1} - \left(\frac{1}{2k-1}\right)^2\right) = \mathbb{E}C_m - \sum_{k=1}^m \frac{1}{(2k-1)^2}.$$

and we need to consider the second sum. We have $\sum_{j=1}^m \frac{1}{j^2} = -\psi'(m+1) + \pi^2/6$, so

$$\mathbb{V}C_m = \mathbb{E}C_m - \left(-\psi'(2n+1) + \frac{1}{4}\psi'(n+1) + \frac{\pi^2}{8}\right).$$

But $\psi'(m) = O(m^{-1})$, so in fact we get $\mathbb{V}C_m = \mathbb{E}C_m - \frac{\pi^2}{8} + O(1/m)$; from this and (13) we get (14). Asymptotic normality follows from the Lindeberg-Feller central limit theorem [7]. $\square$

There is not such a simple decomposition for permutations with all cycle lengths odd. However, it appears that the polynomials counting permutations of $[n]$ with all cycle lengths odd by their number of cycles have only pure imaginary roots. If this is true, then the number of cycles of a random permutation of $[n]$ with all cycle lengths odd can be decomposed into a sum of $\lfloor n/2 \rfloor$ independent $\{0, 2\}$-valued random variables, plus 1 if n is odd. It may be of interest to study the zeros of these polynomials.

4 Boltzmann sampling

We have at this point seen substantial similarities between permutations with all cycle lengths having the same parity and permutations with cycle weights 1/2. This suggests that an average of weights is in some sense a more fundamental parameter than the individual weights. This has been anticipated by the notion of a *function of logarithmic type* [14], which has been used in the study of permutations [17]. Let $\Delta_0(\rho,\eta) = \{z : |z| < \rho+\eta, z \notin [\rho,\rho+\eta]\}$. A function $G(z)$ is called *logarithmic* if it is of the form

$$G(z) = a \log \frac{1}{1-z/\rho} + R(z)$$

for some constant multiplier a and function $R(z)$, where $R(z)$ is analytic in Δ_0 and satisfies $R(z) = K + o(1)$ for some constant K as $z \to \rho$ in Δ_0, and ρ is the unique dominant singularity of G on its circle of convergence. In [14, Prop. 1] structures having components enumerated by a function of logarithmic type $G(z)$ are considered; for such structures of size n, the expected number of cycles is $a \log n + O(1)$, as is the variance. However, the structures considered in this paper have not all had components counted by functions of logarithmic type. For example, the components of permutations with all cycle lengths even are counted by the exponential generating function $\frac{1}{2}\log\frac{1}{1-z^2}$, which has singularities at $z = \pm 1$ and thus does not have a unique dominant singularity.

The following conjecture, in the light of these averaging phenomena, seems natural. It is supported by Theorem 4.3, an analogous result on "Boltzmannized" permutations.

Conjecture 4.1. *Let $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots)$ be a sequence of nonnegative real numbers with mean α, that is, with $\lim_{n\to\infty} \frac{1}{n}\left(\sum_{k=1}^n \sigma_k\right) = \alpha$. Then permutations of $[n]$ selected according to the weights $\vec{\sigma}$ have an asymptotically Gaussian number of cycles as $n \to \infty$, with mean and variance asymptotic to $\alpha \log n$.*

Definition 4.2. *Let $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots)$ be a weighting sequence, and let x be a positive real parameter. Let $|\pi|$ denote the size of the ground set of a permutation π. Then we define the $\vec{\sigma}$-weighted Boltzmann measure with parameter x on permutations, a probability measure on $\bigcup_{k=0}^{\infty} S_k$, by*

$$\mathbb{P}_{\vec{\sigma},x}(\pi) = \frac{w_{\vec{\sigma}}(\pi)\cdot \frac{x^{|\pi|}}{|\pi|!}}{\exp\left(\sum_{k\geqslant 1}\sigma_k x^k/k\right)}$$

(See Definition 1.1 for the weight $w_{\vec{\sigma}}(\pi)$.)

For any choice of $\vec{\sigma}$ and x, $\mathbb{P}_{\vec{\sigma},x}$ is a probability measure. It suffices to show that $\mathbb{P}_{\vec{\sigma},x}$ has total mass 1. But $\sum_{k\geqslant 1}\sigma_k x^k/k$ is the weighted generating function of cycles, and we can apply the exponential formula. Thus $\mathbb{P}_{\vec{\sigma},x}$ is a straightforward weighted generalization of the Boltzmann measure on labelled objects studied in [6, 10]. We also retain the formulas from [6, Thm 2.1] for the expected size and the variance of the size of the objects chosen according to this measure. Let $C(x)$ be the exponential generating function of a labelled combinatorial class, and N the size of a random object chosen from that class according to the Boltzmann measure. Then

$$\mathbb{E}_{\vec{\sigma},x}(N) = \frac{x\frac{d}{dx}C(x)}{C(x)}, \mathbb{E}_{\vec{\sigma},x}(N^2) = \frac{\left(x\frac{d}{dx}\right)^2 C(x)}{C(x)}.$$

We now assemble a sequence of lemmas. These lemmas will be used to prove the following theorem, which is the main result of this section.

Theorem 4.3. *Let $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots)$ be a weighting sequence with mean α. Let $x = x(\mu)$ be chosen so that $\mathbb{E}_{\vec{\sigma},x}(N) = \mu$. Let X be a random variable denoting the number of cycles of a permutation. Then $\mathbb{E}_{\vec{\sigma},x}(X) = \mathbb{V}_{\vec{\sigma},x}(X) \sim \alpha \log \mu$ as $x \to 1^-$ or $\mu \to \infty$.*

The main analytic result needed follows.

Lemma 4.4. *[21, Exercise I.88] Let $b_0, b_1, \ldots$ be positive real numbers, such that $\sum_{n=0}^{\infty} b_n$ is divergent, and $\sum_{k\geqslant 0} b_k t^k$ is convergent for $0 \leqslant t < 1$. Then*

$$\lim_{n\to\infty} \frac{a_0 + a_1 + \cdots + a_n}{b_0 + b_1 + \cdots + b_n} = s \text{ implies } \lim_{t\to 1^-} \frac{\sum_{k\geqslant 0} a_k t^k}{\sum_{k\geqslant 0} b_k t^k} = s.$$

Lemma 4.5. *Let $\sigma_1, \sigma_2, \sigma_3, \ldots$ be a sequence of real numbers such that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sigma_k = \alpha.$$

Then $\sum_{k=1}^{\infty} \sigma_k x^k = \frac{\alpha}{1-x} + o((1-x)^{-1})$.

Proof. Apply Lemma 4.4 with $a_k = \sigma_k, b_k = 1$. □

Lemma 4.6. *Let $\{\sigma_k\}_{k=1}^{\infty}$ be a sequence of nonnegative real numbers, bounded above, such that $\sum_{k=1}^{n} \sigma_k \sim \alpha n$ as $n \to \infty$, for some constant $\alpha > 0$. Then $\sum_{k=1}^{n} \frac{\sigma_k}{k} \sim \alpha \log n$ as $n \to \infty$.*

Proof. We begin by showing that if $\int_1^n f(x)dx \sim n$ as $n \to \infty$ for some function f such that $f(x)/x$ is integrable on $[1, \infty)$, then $\int_1^n \frac{f(x)}{x}\, dx \sim \log n$ as $n \to \infty$. Let $F(x) = \int_1^n f(x)\, dx$. We integrate $\int_1^n \frac{f(x)}{x}\, dx$ by parts, getting

$$\int_1^n \frac{f(x)}{x}\, dx = \frac{F(n)}{n} - \frac{F(1)}{1} + \int_1^n \frac{F(x)}{x^2}\, dx.$$

Clearly $F(1) = 0$, and $F(n) \sim n$ by assumption, so

$$\int_1^n \frac{f(x)}{x}\, dx = 1 + \int_1^n \frac{F(x)}{x^2}\, dx + o(1)$$

Since $F(x) \sim x$ as $x \to \infty$, the integrand satisfies $F(x)/x^2 \sim 1/x$, and so

$$\int_1^n \frac{F(x)}{x^2}\, dx \sim \int_1^n \frac{1}{x}\, dx = \log n,$$

proving the claim.

Now, we need to check that this statement about integrals translates into an analogous one about sums. Let $\{\sigma_k\}_{k=1}^{\infty}$ be as in the hypothesis, and let $f(x) = \sigma_{\lfloor x \rfloor}$. Then we want to show that

$$\int_1^{n+1} \frac{f(x)}{x}\, dx - \sum_{k=1}^{n} \frac{\sigma_k}{k} = o(\log n)$$

as $n \to \infty$. We have

$$\begin{aligned}\int_1^{n+1} \frac{f(x)}{x}\,dx - \sum_{k=1}^n \frac{f(k)}{k} &= \sum_{k=1}^n \left(\int_k^{k+1} \frac{f(x)}{x}\,dx - \frac{f(k)}{k}\right) \\ &= \sum_{k=1}^n f(k)\left(\log\left(1+\frac{1}{k}\right) - \frac{1}{k}\right)\end{aligned}$$

and so, since $|\log(1+1/k) - 1/k| \leqslant 1/2k^2$ for positive integer k,

$$\left|\int_1^{n+1} \frac{f(x)}{x}\,dx - \sum_{k=1}^n \frac{f(k)}{k}\right| \leqslant \left|\sum_{k=1}^n \frac{f(k)}{2k^2}\right|.$$

Since $\{f(k)\}_{k=1}^\infty$ is bounded, the sum on the right-hand side is convergent. We have $\int_1^{n+1} f(x)/x\,dx \sim \log n$, so $\sum_{k=1}^n f(k)/k \sim \log n$ as well. Thus we have proven the lemma for $\alpha = 1$. Multiplying through by α gives the desired result. $\square$

Lemma 4.7. *Let $\sigma_1, \sigma_2, \sigma_3, \ldots$ be a sequence of positive real numbers such that $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n \sigma_k = \alpha$. Then*

$$\sum_{k=1}^\infty \frac{\sigma_k x^k}{k} = \alpha \log \frac{1}{1-x} + o\left(\log\frac{1}{1-x}\right) \tag{15}$$

Proof. Applying Lemma 4.6 to the hypothesis, $\lim_{n\to\infty} \frac{1}{\log n}\sum_{k=1}^n \frac{\sigma_k}{k} = \alpha$. We apply Lemma 4.4 with $a_k = \sigma_k/k, b_k = 1/k$. This gives us

$$\lim_{n\to\infty} \frac{\sum_{k=0}^n \sigma_k/k}{1+H_n} = \lim_{x\to 1^-} \frac{\sum_{k\geqslant 1} \sigma_k x^k/k}{\sum_{k\geqslant 1} x^k/k}.$$

Now, $\sum_{k\geqslant 1} x^k/k = \log(1/(1-x))$, and $1 + H_n \sim \log n$, so we have

$$\lim_{n\to\infty} \frac{1}{\log n}\sum_{k=0}^n \frac{\sigma_k}{k} = \lim_{x\to 1^-} \frac{\sum_{k\geqslant 1} \sigma_k x^k/k}{\log(1/(1-x))}.$$

Thus the right-hand side here has value α, proving (15). $\square$

Proof of Theorem 4.3. Note that for the Boltzmann measure with parameter x and weight sequence $\vec{\sigma}$, we have $C(x) = \exp\left(\sum_{k\geqslant 1} \sigma_k x^k/k\right)$. Thus the distribution of sizes N under this measure has expectation

$$\mathbb{E}_{\vec{\sigma},x}(N) = x\frac{d}{dx}\left(\sum_{k\geqslant 1} \frac{\sigma_k x^k}{k}\right) = \sum_{k\geqslant 1} \sigma_k x^k$$

Furthermore, the Boltzmann distribution $\mathbb{P}_{\vec{\sigma},x}$ can be obtained by taking $\mathcal{P}(\sigma_k x^k/k)$ cycles of length k for each $k \geqslant 1$ and forming uniformly at random a permutation with the resulting cycle type. The mean and variance of the number of cycles chosen from the distribution $\mathbb{P}_{\vec{\sigma},x}$ is thus exactly $\sum_{k\geqslant 1}\sigma_k x^k/k$. Since we have $\mathbb{E}_x(N) \sim \alpha/(1-x)$ as $x \to 1^-$ by Lemma 4.5, we can solve for x to see that $1 - \frac{\alpha}{\mathbb{E}_x(N)} \sim x$ as $x \to 1^-$. Therefore

$$\sum_{k\geqslant 1}\sigma_k x^k/k \sim \alpha \log \frac{1}{1-x} \sim \alpha \log \frac{1}{1-\left(1-\frac{\alpha}{\mathbb{E}_x(N)}\right)} = \alpha \log \frac{\mathbb{E}_x(N)}{\alpha} \sim \alpha \log \mathbb{E}_x(N)$$

which is the desired result. $\square$

It would be desirable to translate Theorem 4.3 into a result about permutations of a fixed size selected uniformly at random; this is one possible way of proving Conjecture 4.1. Note that $\mathbb{P}_{\vec{\sigma},x}$ is a mixture of the various $\mathbb{P}_{\vec{\sigma}}^{(n)}$. It is often possible to prove results about a family of measures P_λ, parametrized by λ, which are mixtures of well-understood measures $\mathbb{P}^{(n)}$, where we draw from $\mathbb{P}^{(n)}$ with probability $e^{-\lambda}\lambda^n/n!$; this goes by the name of analytic de-Poissonization [18], [24, Ch. 10]. Informally, we pick from $\mathbb{P}^{(N)}$ where N is Poisson with parameter λ. In the case described here we can get results on permutations chosen from $\mathbb{P}^{(N)}$ where N is the size of objects from a Boltzmann distribution; thus techniques of "analytic de-Boltzmannization" will be necessary to achieve this goal.

Acknowledgement. I thank Robin Pemantle, my Ph.D. thesis advisor, for very helpful suggestions and valuable support. I thank the anonymous referee for remarks concerning the proof of Theorem 2.7.

References

[1] R. Arratia, A. D. Barbour, and S. Tavaré. Random combinatorial structures and prime factorizations. *Notices Amer. Math. Soc.* 44 (1997), 903-910.

[2] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach.* European Mathematical Society, 2003.

[3] R. Arratia and S. Tavaré. Limit theorems of combinatorial structures via discrete process approximation. *Random Structures Algorithms* 3 (1992), 321-345.

[4] F. Benaych-Georges. Cycles of random permutations with restricted cycle lengths. Preprint, arXiv:0712.1903.

[5] E. A. Bender. Asymptotic methods in enumeration. *SIAM Rev.* 16 (1974), 485-515.

[6] P. Duchon, P. Flajolet, G. Louchard and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combin. Probab. Comput.* 13 (2004), 577-625.

[7] R. Durrett. *Probability: Theory and Examples.* Second edition. Duxbury, 1995.

[8] W. Ewens. The sampling theory of selectively neutral alleles. *Theoret. Population Biol.* 3 (1972), 87-112.

[9] W. Feller. The fundamental limit theorems in probability. *Bull. Amer. Math. Soc.* 51 (1945), 800-832.

[10] P. Flajolet, E. Fusy, X. Gourdon, D. Panario, and N. Pouyanne. A hybrid of Darboux's method and singularity analysis in combinatorial asymptotics. *Electron. J. Combin.*, 13 (2006), R103 (35 pages).

[11] P. Flajolet and A. Odlyzko. Random mapping statistics. *Lecture Notes in Computer Science* 434 (1990), 329-354.

[12] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.* 3 (1990), 216-240.

[13] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009.

[14] P. Flajolet and M. Soria. Gaussian limiting distributions for the number of components in combinatorial structures. *J. Combin. Theory Ser. A* 153 (1990), 165-182.

[15] D. Gandolfo, J. Ruiz, D. Ueltschi, On a model of random cycles. *J. Stat. Phys.* 129 (2007), 663-676.

[16] D. H. Greene and D. E. Knuth. *Mathematics for the Analysis of Algorithms.* Third edition. Birkhauser, 1990.

[17] J. C. Hansen. Order statistics for decomposable combinatorial structures. *Random Structures Algorithms* 5 (1994), 517-533.

[18] P. Jacquet and W. Szpankowski. Analytical depoissonization and its applications. *Theoret. Comput. Sci.* 201 (1998), 1-62.

[19] M. Lugo. The number of cycles of specified normalized length in permutations. In preparation.

[20] J. Pitman. *Combinatorial stochastic processes.* Springer, 2006. *Springer lecture notes in mathematics*, vol. 1875.

[21] G. Pólya and G. Szegő. *Problems and Theorems in Analysis I.* Springer, 1998. Corrected reprint of 1978 edition. Translated from German by Dorothee Aeppli.

[22] N. Pouyanne. On the number of permutations admitting an m-th root. *Electron. J. Combin.* 9(1) (2002), R3 (12 pages).

[23] V. N. Sachkov. *Probabilistic methods in combinatorial analysis.* Cambridge University Press, 1997.

[24] W. Szpanowski. *Average-case analysis of algorithms on sequences.* Wiley, 2001.

[25] A. N. Timashev. Random permutations with cycle lengths in a given finite set. *Discrete Math. Appl.* 18 (2008), 25-39.

[26] D. Ueltschi and V. Betz. Spatial random permutations with small cycle weights. Preprint, arXiv:0812.0569.

[27] G. A. Watterson. Models for the logarithmic species abundance distributions. *Theoret. Population Biol.* 6 (1974) 217-250.

[28] H. Wilf. The asymptotics of $e^{P(z)}$ and the number of elements of each order in S_n, *Bull. Amer. Math. Soc.* 15 (1986), 228-232.

[29] A. L. Yakymiv. On permutations with cycle lengths in a random set. *Discrete Math. Appl.* 10 (2000), 543-551.