The Annals of Probability
1992, Vol. 20, No. 3, 1567–1591

THE CYCLE STRUCTURE OF RANDOM PERMUTATIONS[1]

BY RICHARD ARRATIA AND SIMON TAVARÉ

University of Southern California

The total variation distance between the process which counts cycles of size $1, 2, \ldots, b$ of a random permutation of n objects and a process $(Z_1, Z_2, \ldots, Z_b)$ of independent Poisson random variables with $\mathbb{E}Z_i = 1/i$ converges to 0 if and only if $b/n \to 0$. This Poisson approximation can be used to give simple proofs of limit theorems and bounds for a wide variety of functionals of random permutations. These limit theorems include the Erdös–Turán theorem for the asymptotic normality of the log of the order of a random permutation, and the DeLaurentis–Pittel functional central limit theorem for the cycle sizes.

We give a simple explicit upper bound on the total variation distance to show that this distance decays to zero superexponentially fast as a function of $n/b \to \infty$. A similar result holds for derangements and, more generally, for permutations conditioned to have given numbers of cycles of various sizes. Comparison results are included to show that in approximating the cycle structure by an independent Poisson process the main discrepancy arises from independence rather than from Poisson marginals.

1. Introduction. Random permutations play an important role in many areas of probability and statistics. Most people meet them (implicitly at least) in the context of the so-called hat-check problem: n mathematicians drop off their hats at a restaurant before having a meal. After the meal, the hats are returned at random. How many mathematicians get back their own hat? Feller (1968) is the most accessible reference to this. The return of the hats induces a random permutation of $1, \ldots, n$: label the mathematicians $1, \ldots, n$, and assign to j the label of the mathematician whose hat was returned to j. The solution to the simplest hat-check problem is then seen to be the distribution of the number of singleton cycles of this random permutation. Denoting this number by $C_1(n)$, an inclusion–exclusion argument shows that for $k = 0, 1, \ldots, n$,

$$\mathbb{P}(C_1(n) = k) = \frac{1}{k!} \sum_{l=0}^{n-k} \frac{(-1)^l}{l!}. \tag{1}$$

It follows immediately from (1) that

$$\mathbb{P}(C_1(n) = k) \to \frac{e^{-1}}{k!}, \qquad n \to \infty,$$

so that $C_1(n)$ converges in distribution to a random variable Z_1 having the

Received October 1990; revised May 1991.

[1]The authors were supported in part by NSF Grants DMS-88-15106, DMS-88-03284 and DMS-90-05833, and NIH grant GM 41746.

AMS 1980 *subject classifications*. 60C05, 60F17, 60B15, 60G18, 05A05, 05A16.

Key words and phrases. Poisson process, inclusion–exclusion, total variation, exponential generating functions, derangements, conditional limit theorems.

Poisson distribution with mean 1. Using properties of alternating series with decreasing terms, it can be seen that for $k = 0,1,\ldots,n$,

$$\frac{1}{k!}\left(\frac{1}{(n-k+1)!} - \frac{1}{(n-k+2)!}\right) \leq \left|\mathbb{P}(C_1(n) = k) - \mathbb{P}(Z_1 = k)\right| \leq \frac{1}{k!(n-k+1)!},$$

as given in David and Barton [(1962), page 105]. It follows that

$$\frac{2^{n+1}}{(n+1)!}\frac{n}{n+2} \leq \sum_{k=0}^{n}\left|\mathbb{P}(C_1(n) = k) - \mathbb{P}(Z_1 = k)\right| \leq \frac{2^{n+1}-1}{(n+1)!}. \tag{2}$$

Since

$$\mathbb{P}(Z_1 > n) = \frac{e^{-1}}{(n+1)!}\left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots\right) < \frac{1}{(n+1)!},$$

we see from (2) that the total variation distance $d_1(n)$ between the law of $C_1(n)$ and the law of Z_1, defined by

$$d_1(n) \equiv \tfrac{1}{2}\sum_{k=0}^{\infty}\left|\mathbb{P}(C_1(n) = k) - \mathbb{P}(Z_1 = k)\right|,$$

satisfies the inequalities

$$\frac{2^n}{(n+1)!}\frac{n}{n+2} \leq d_1(n) \leq \frac{2^n}{(n+1)!} \tag{3}$$

for $n = 1,2,\ldots$. Therefore the rate of convergence to the Poisson probabilities is superexponential in n as $n \to \infty$.

Let $C_j \equiv C_j(n)$ be the number of cycles of length j in a random permutation of $\{1,2,\ldots,n\}$. Cauchy's formula for the probability law of $(C_1,\ldots,C_n)$ is given by

$$P(C_1 = a_1,\ldots,C_n = a_n) = \prod_{j=1}^{n}\left(\frac{1}{j}\right)^{a_j}\frac{1}{a_j!}, \tag{4}$$

for nonnegative integers $a_1,\ldots,a_n$ satisfying $\sum_{i=1}^{n} i a_i = n$. Asymptotically, the finite-dimensional distributions of $(C_1,\ldots,C_n)$ are those of a Poisson process on $\mathbb{N}$, as the following result of Goncharov (1944) and Kolchin (1971) shows.

THEOREM 1. *For $i = 1,2,\ldots$, let $C_i(n)$ denote the number of cycles of length i in a random n-permutation. The process of cycle counts converges in distribution to a Poisson process on $\mathbb{N}$ with intensity i^{-1}. That is, as $n \to \infty$,*

$$(C_1(n), C_2(n), \ldots) \Rightarrow (Z_1, Z_2, \ldots),$$

where the Z_i, $i = 1, 2, \ldots$, are independent Poisson-distributed random variables with

$$\mathbb{E}(Z_i) = \frac{1}{i}.$$

PROOF. For nonnegative integers $m_1, \ldots, m_k$, it is known that

$$\mathbb{E}\left(\prod_{j=1}^{k} C_j^{[m_j]}\right) = \left(\prod_{j=1}^{k} \left(\frac{1}{j}\right)^{m_j}\right) 1\left\{\sum_{j=1}^{k} jm_j \le n\right\}, \tag{5}$$

where we write $x^{[r]} = x(x-1)\cdots(x-r+1)$. See, for example, Watterson (1974a). Equation (5) may be expressed as

$$\mathbb{E}\left(\prod_{j=1}^{k} C_j^{[m_j]}\right) = \mathbb{E}\left(\prod_{j=1}^{k} Z_j^{[m_j]}\right) 1\left\{\sum_{j=1}^{k} jm_j \le n\right\}. \tag{6}$$

We may now fix k and integers $m_1, \ldots, m_k \ge 0$ and let $n \to \infty$ in (6). The result follows from the method of moments. □

REMARK. It follows from (5) that C_i and C_j are uncorrelated if $i + j \le n$, whereas $\mathbb{E}C_iC_j \equiv 0$ if $i + j > n$.

It is the purpose of this paper to provide explicit bounds on how close the distribution of $(C_1, C_2, \ldots)$ is to that of $(Z_1, Z_2, \ldots)$, the independent Poisson components of Theorem 1. Specifically, for $1 \le b \le n$ we will estimate $d_b \equiv d_b(n)$, the total variation distance between the law of $(C_1, \ldots, C_b)$ and the law of $(Z_1, \ldots, Z_b)$, defined by

$$\begin{aligned} d_b &\equiv \| \mathscr{L}(C_1, \ldots, C_b) - \mathscr{L}(Z_1, \ldots, Z_b) \| \\ &= \sup_{A \subseteq \mathbb{Z}_+^b} \left| \mathbb{P}((C_1, \ldots, C_b) \in A) - \mathbb{P}((Z_1, \ldots, Z_b) \in A) \right|, \end{aligned} \tag{7}$$

where $\mathbb{Z}_+ \equiv \{0, 1, \ldots\}$. We will prove that $d_b(n) \to 0$ if and only if $b = o(n)$ and that if $b/n \to 0$, then $d_b(n) \to 0$ superexponentially fast relative to n/b. The result $d_b(n) \le 2b/n$ was proved by Diaconis and Pitman (1986) and independently by Barbour (1990). We are indebted to Diaconis and Freedman (1980) for showing the value of considering the total variation distance between a growing number of coordinates of a dependent process and the corresponding restriction of its limiting, independent process.

2. Estimating total variation distance. We will need some further notation in this section. Define $\mathbf{a} = (a_1, \ldots, a_b) \in \mathbb{Z}_+^b$, $\mathbf{C}_b = (C_1, \ldots, C_b)$, $\mathbf{Z}_b = (Z_1, \ldots, Z_b)$ and, for integers $0 \le l \le m$,

$$T_{lm} = (l+1)Z_{l+1} + (l+2)Z_{l+2} + \cdots + mZ_m,$$

with $T_{mm} \equiv 0$.

The key observation, due to Watterson (1974a, b), is that the probability law (4) of $(C_1, C_2, \dots)$ is simply related to that of $(Z_1, Z_2, \dots)$:

$$\mathbb{P}(\mathbf{C}_b = \mathbf{a}) = \mathbb{P}(\mathbf{Z}_b = \mathbf{a} \mid T_{0n} = n). \tag{8}$$

Relation (8) may be verified directly from (4). Shepp and Lloyd (1966) used a similar relation: If Z_i are Poisson with parameter $\mathbb{E}Z_i = x^i/i$, for $0 < x < 1$, and $T = T_{0\infty} = \sum_{i=1}^{\infty} iZ_i$, then $(C_1, \dots, C_n) = (Z_1, \dots, Z_n \mid T = n)$ in distribution. See formula (30) for a general discussion of the role of the factor x^i in $\mathbb{E}Z_i$.

Our strategy for bounding the total variation distance d_b is outlined in the next two lemmas. The first task is to find a simple explicit expression for d_b. We will start from the following equivalent definition of d_b:

$$d_b = \tfrac{1}{2} \sum_{\mathbf{a} \in \mathbb{Z}_+^b} \left| \mathbb{P}(\mathbf{C}_b = \mathbf{a}) - \mathbb{P}(\mathbf{Z}_b = \mathbf{a}) \right|. \tag{9}$$

REMARK. The following lemma, although elementary, is significant in that it reduces a total variation distance between processes to a simpler total variation distance between two random variables.

LEMMA 1. *For* $1 \le b \le n$,

$$d_b = \left\| \mathscr{L}(T_{0b}) - \mathscr{L}(T_{0b} \mid T_{0n} = n) \right\|. \tag{10}$$

PROOF. Writing $L\mathbf{a} = \sum_{j=1}^{b} ja_j$, the right-hand side of (9) is

$$\tfrac{1}{2} \sum_{r=0}^{\infty} \sum_{\mathbf{a}:\, L\mathbf{a} = r} \left| \mathbb{P}(\mathbf{C}_b = \mathbf{a}) - \mathbb{P}(\mathbf{Z}_b = \mathbf{a}) \right|,$$

which, with the help of (8), may be written

$$\tfrac{1}{2} \sum_{r=0}^{\infty} \sum_{\mathbf{a}:\, L\mathbf{a} = r} \left| \mathbb{P}(\mathbf{Z}_b = \mathbf{a} \mid T_{0n} = n) - \mathbb{P}(\mathbf{Z}_b = \mathbf{a}) \right|. \tag{11}$$

Using the independence of the Z_i, note that if $\sum_{j=1}^{b} ja_j = r$,

$$\begin{aligned} \mathbb{P}(\mathbf{Z}_b = \mathbf{a} \mid T_{0n} = n) &= \frac{\mathbb{P}(\mathbf{Z}_b = \mathbf{a}, T_{0n} = n)}{\mathbb{P}(T_{0n} = n)} \\ &= \frac{\mathbb{P}(\mathbf{Z}_b = \mathbf{a}, T_{bn} = n - r)}{\mathbb{P}(T_{0n} = n)} \\ &= \frac{\mathbb{P}(\mathbf{Z}_b = \mathbf{a})\mathbb{P}(T_{bn} = n - r)}{\mathbb{P}(T_{0n} = n)}. \end{aligned}$$

Factoring out the common factor $\mathbb{P}(\mathbf{Z}_b = \mathbf{a})$, we see then that (11) may be

written

$$d_b = \frac{1}{2}\sum_{r=0}^{\infty}\sum_{\mathbf{a}:\, L\mathbf{a}=r}\mathbb{P}(\mathbf{Z}_b=\mathbf{a})\left|\frac{\mathbb{P}(T_{bn}=n-r)}{\mathbb{P}(T_{0n}=n)}-1\right|$$

$$= \frac{1}{2}\sum_{r=0}^{\infty}\mathbb{P}(T_{0b}=r)\left|\frac{\mathbb{P}(T_{bn}=n-r)}{\mathbb{P}(T_{0n}=n)}-1\right| \tag{12}$$

$$= \frac{1}{2}\sum_{r=0}^{\infty}\left|\mathbb{P}(T_{0b}=r)-\mathbb{P}(T_{0b}=r\,|\,T_{0n}=n)\right|,$$

establishing the lemma. □

We will apply the following lemma with $R = T_{0b}$, $S = T_{bn}$, $p = \mathbb{P}(T_{0n} = n)$ and $c = \mathbb{P}(T_{bn} = n)$.

LEMMA 2. *Let R and S be independent discrete random variables, with densities g and f, respectively, and suppose n is some constant such that $p \equiv \mathbb{P}(R + S = n) > 0$. Let $c > 0$ be any constant. Then the total variation distance between the law of R, and the law of R conditional on $R + S = n$, satisfies the inequality*

$$\|\mathscr{L}(R) - \mathscr{L}(R\,|\,R+S=n)\| = \frac{1}{2}\sum_r g_r\left|\frac{f_{n-r}}{p}-1\right|$$

$$\le \sum_r g_r\left|\frac{f_{n-r}}{c}-1\right|.$$

PROOF. First we have

$$\left|\frac{p}{c}-1\right| = \left|\sum_r g_r\left(\frac{f_{n-r}}{c}-1\right)\right| \le \sum_r g_r\left|\frac{f_{n-r}}{c}-1\right|.$$

Second, $|g_r - g_r f_{n-r}/p| \le |g_r - g_r f_{n-r}/c| + |g_r f_{n-r}/c - g_r f_{n-r}/p|$. Summing over r, we get

$$\sum_r g_r\left|\frac{f_{n-r}}{p}-1\right| \le \sum_r g_r\left|\frac{f_{n-r}}{c}-1\right| + p\left|\frac{1}{c}-\frac{1}{p}\right| \le 2\sum_r g_r\left|\frac{f_{n-r}}{c}-1\right|.$$

Observe that for $c = p$ the inequality gives away a factor of two. □

Lemmas 3 through 6 are used to show that the density of T_{bn} is nearly constant, corresponding to an upper bound on $|f_{n-r}/c - 1|$. Lemmas 7 and 8 apply large deviation theory to bound the probability that T_{0b} is large, corresponding to an upper bound on g_r. These two upper bounds have matching decay rates, so that in our proof of Theorem 2 the terms with r between 1 and n contribute equally to the final upper bound on total variation

distance. A special argument is required for the case $r = n$; $\mathbb{P}(T_{bn} = 0)/\mathbb{P}(T_{bn} = n) \geq b$, which can be large compared to n/b. This motivates the more detailed large deviation estimate in the second part of Lemma 8.

LEMMA 3. *Fix b and n, and define*

$$f_k = \mathbb{P}(T_{bn} = k).$$

Then f_j, $j \geq 0$, satisfy the recursion

$$kf_k = (k-1)f_{k-1} + f_{k-b-1} - f_{k-n-1}, \tag{13}$$

where f_j is defined to be 0 if $j < 0$.

PROOF. Let $f(z) \equiv \sum_{j=0}^{\infty} f_j z^j$ be the probability generating function of T_{bn}. Then

$$f(z) = \mathbb{E}(z^{T_{bn}}) = \exp\left(\sum_{j=b+1}^{n} \frac{1}{j}(z^j - 1) \right).$$

We will write this in the form $f(z) = \exp(\sum_{j=0}^{\infty} a_j z^j)$, where

$$a_0 = -\sum_{j=b+1}^{n} \frac{1}{j},$$

$$a_j = \begin{cases} 0, & j = 1, 2, \ldots, b, n+1, n+2, \ldots, \\ \dfrac{1}{j}, & j = b+1, \ldots, n. \end{cases}$$

Since $f'(z) = (\sum j a_j z^{j-1}) f(z)$, the coefficients f_j and a_j are related by the recursion [cf. Pourahmadi (1984)]

$$f_{k+1} = \sum_{j=0}^{k} \left(1 - \frac{j}{k+1}\right) a_{k+1-j} f_j, \qquad k = 0, 1, 2, \ldots, \tag{14}$$

where

$$f_0 = \exp(a_0).$$

It follows that $f_1 = f_2 = \cdots = f_b = 0$ and, for all k,

$$f_{k+1} = \frac{1}{k+1} \left(\sum_{j=k+1-n}^{k-b} f_j \right),$$

empty sums being interpreted as zero. Multiplying by $(k+1)$ and differencing in k shows that

$$(k+1) f_{k+1} = k f_k + f_{k-b} - f_{k-n},$$

completing the proof of the lemma. □

REMARK. The special case $b = 0$ of Lemma 3 shows that $\mathbb{P}(T_{0n} = k)$ is constant for $k = 0, 1, \ldots, n$, as noted by Watterson (1974a).

LEMMA 4. *If*

$$\frac{f_i}{f_k} \in [1 - r, 1 + s], \qquad i = k - b, k - b + 1, \ldots, k,$$

then

$$\frac{f_i}{f_k} \in \left[1 - \frac{rb}{k}, 1 + \frac{sb}{k}\right], \qquad i = k, k + 1, \ldots k + b,$$

for $2b < k \leq n - b$.

PROOF. Use the recursion (13) and the hypothesis of the lemma to see that for $j = 0, 1, \ldots, b$,

$$\begin{aligned} f_{k+j} &= \frac{1}{k+j}\left(kf_k + \sum_{l=k-b}^{k-b+j-1} f_l\right) \\ &\geq \frac{1}{k+j}(kf_k + j(1-r)f_k) \\ &= \left(1 - \frac{jr}{j+k}\right) f_k. \end{aligned}$$

Since $0 \leq j \leq b$, it follows that

$$\frac{j}{k+j} \leq \frac{b}{k+b} \leq \frac{b}{k},$$

so that

$$f_{k+j} \geq \left(1 - \frac{rb}{k}\right) f_k, \qquad j = 0, 1, \ldots, b.$$

An analogous argument establishes that for $j = 0, 1, \ldots, b$,

$$\begin{aligned} f_{k+j} &\leq \frac{1}{k+j}(kf_k + j(1+s)f_k) \\ &= \left(1 + \frac{js}{k+j}\right) f_k \\ &\leq \left(1 + \frac{bs}{k}\right) f_k, \end{aligned}$$

completing the proof. □

LEMMA 5. *For $2b < k \le n - b$ define*

$$M(k) = \max_{i,j \in [k-b,k]} \left(\frac{f_i}{f_j} - 1 \right).$$

Then

$$M(k+b) \le M(k) \frac{b}{k}.$$

PROOF. Define $r \ge 0$ and $s \ge 0$ by

$$(1-r) f_k = \min_{i \in [k-b,k]} f_i$$

and

$$(1+s) f_k = \max_{i \in [k-b,k]} f_i.$$

Then $M(k) = (1+s)/(1-r) - 1 = (r+s)/(1-r)$. Now apply Lemma 4 with these values of r and s to get

$$\begin{aligned} M(k+b) &\le \left(1 + \frac{sb}{k}\right) \Big/ \left(1 - \frac{rb}{k}\right) - 1 \\ &= \frac{b(s+r)}{k - rb} \\ &= M(k) \frac{b(1-r)}{k - rb} \\ &\le M(k) \frac{b}{k}. \end{aligned}$$

□

LEMMA 6. *For integers $1 \le b \le n$ and any integer l with $1 < l < n/b$, the density $f_k \equiv \mathbb{P}(T_{bn} = k)$ satisfies*

$$\max_{lb < i,j \le n} \frac{f_i}{f_j} - 1 \le \frac{b}{b+1} \frac{b}{2b+1} \cdots \frac{b}{lb+1} < \frac{1}{l!}.$$

PROOF. Direct calculation for $b + 1 \le k \le 2b + 1$ gives

$$\begin{aligned} f_k &= \mathbb{P}(Z_k = 1, Z_i = 0 \text{ for all } b < i \le n \text{ with } i \ne k) \\ &= k^{-1} \mathbb{P}(Z_i = 0 \text{ for all } b < i \le n) \\ &= f_0 / k. \end{aligned}$$

Thus the function $M(k)$ in Lemma 5 satisfies

$$M(2b+1) = \frac{f_{b+1}}{f_{2b+1}} - 1 = \frac{2b+1}{b+1} - 1 = \frac{b}{b+1}.$$

Applying Lemma 5 with $k = 2b + 1$ yields

$$M(3b + 1) \le M(2b + 1)\frac{b}{2b + 1} = \frac{b}{b + 1}\frac{b}{2b + 1},$$

and repeated applications of Lemma 5 yield

$$\max_{lb<i,\, j\le lb+b+1} \frac{f_i}{f_j} - 1 \equiv M((l + 1)b + 1) \le \frac{b}{b + 1}\frac{b}{2b + 1}\cdots\frac{b}{lb + 1}.$$

Now the recursion $kf_k = (k - 1)f_{k-1} + f_{k-b-1}$ of Lemma 3, valid for $k \le n$, implies that if $f_i \in [c, d]$ for $i = k, k + 1, \ldots, k + b$, then $f_i \in [c, d]$ for $i = k, k + 1, \ldots, n$. Therefore

$$\max_{lb<i,\, j\le n} \frac{f_i}{f_j} \le \max_{lb<i,\, j\le lb+b+1} \frac{f_i}{f_j},$$

which completes the proof of this lemma. □

LEMMA 7. *For* $\beta \ge 0$,

$$\log \mathbb{E}e^{\beta T_{0b}} \le e^{b\beta}. \tag{15}$$

PROOF. Recalling the fact that $T_{0b} = \sum_{j=1}^{b} jZ_j$, where the Z_j are independent Poisson random variables with $\mathbb{E}Z_j = j^{-1}$, we see that

$$\begin{aligned}
\log \mathbb{E}e^{\beta T_{0b}} &= \sum_{j=1}^{b} \frac{1}{j}(e^{\beta j} - 1) \\
&= \sum_{j=1}^{b} \int_0^{\beta} e^{jx}\, dx \\
&\le b\int_0^{\beta} e^{bx}\, dx \\
&= e^{b\beta} - 1,
\end{aligned}$$

from which the result follows. □

LEMMA 8. *For* $x \ge 0$,

$$\mathbb{P}(T_{0b} \ge bx) \le \inf_{\beta \ge 0}(\mathbb{E}e^{\beta T_{0b} - \beta bx}) \le \left(\frac{x}{e}\right)^{-x}, \tag{16}$$

and

$$\mathbb{P}(T_{0b} \ge bx; T_{0n} = n) \le \mathbb{P}(T_{0n} = n) \inf_{\beta \ge 0}(\mathbb{E}e^{\beta T_{0b} - \beta bx}). \tag{17}$$

PROOF. For any $\beta \ge 0$, Markov's inequality shows that

$$\mathbb{P}(T_{0b} \ge bx) = \mathbb{P}(e^{\beta T_{0b}} \ge e^{\beta bx}) \le \mathbb{E}e^{\beta T_{0b}}/e^{\beta bx}.$$

Since β was arbitrary, if follows that

$$\mathbb{P}(T_{0b} \geq bx) \leq \inf_{\beta \geq 0} (\mathbb{E} e^{\beta T_{0b} - \beta bx}).$$

Next, use (15) to establish

$$\begin{aligned}
\log\Big(\inf_{\beta \geq 0} (\mathbb{E} e^{\beta T_{0b} - \beta bx}) \Big) &= \inf_{\beta \geq 0} \log(\mathbb{E} e^{\beta T_{0b} - \beta bx}) \\
&\leq \inf_{\beta \geq 0} (e^{b\beta} - \beta bx) \\
&= \inf_{\tau \geq 0} (e^{\tau} - \tau x) \\
&= -x \log \frac{x}{e},
\end{aligned}$$

the last line following from elementary calculus. This establishes (16).

Under $\mathbb{P}$, the Z_i are independent Poisson random variables with $\mathbb{E}_{\mathbb{P}} Z_i = i^{-1}$. Let $\mathbb{Q}$ be a new measure under which the Z_i are still independent Poisson random variables, but with means given by

$$\mathbb{E}_{\mathbb{Q}} Z_i = \begin{cases} \dfrac{e^{\beta i}}{i}, & 1 \leq i \leq b, \\ \dfrac{1}{i}, & b < i \leq n, \end{cases}$$

for some $\beta > 0$. Then

$$\begin{aligned}
\frac{d\mathbb{Q}}{d\mathbb{P}} &= \prod_{i=1}^{b} e^{\beta i Z_i} \exp\Big(-\frac{1}{i}(e^{\beta i} - 1) \Big) \\
&= e^{\beta T_{0b}} \exp\Big(-\sum_{i=1}^{b} \frac{1}{i}(e^{\beta i} - 1) \Big) \\
&= \frac{1}{\mathbb{E}_{\mathbb{P}} e^{\beta T_{0b}}} e^{\beta T_{0b}}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathbb{P}(T_{0b} \geq bx; T_{0n} = n) &= \int 1(T_{0b} \geq bx; T_{0n} = n)\, d\mathbb{P} \\
&= \int 1(T_{0b} \geq bx; T_{0n} = n) e^{-\beta T_{0b}} \mathbb{E}_{\mathbb{P}} e^{\beta T_{0b}}\, d\mathbb{Q} \\
&\leq \int 1(T_{0b} \geq bx; T_{0n} = n) e^{-\beta bx} \mathbb{E}_{\mathbb{P}} e^{\beta T_{0b}}\, d\mathbb{Q} \\
&= e^{-\beta bx} \mathbb{E}_{\mathbb{P}} e^{\beta T_{0b}} \mathbb{Q}(T_{0b} \geq bx; T_{0n} = n) \\
&\leq e^{-\beta bx} \mathbb{E}_{\mathbb{P}} e^{\beta T_{0b}} \mathbb{Q}(T_{0n} = n).
\end{aligned}$$

Finally, we will estimate $\mathbb{Q}(T_{0n} = n)$. One way to realize $\mathbb{Q}$ is as follows: Let

Z_i^*, $1 \le i \le b$, be independent Poisson random variables with mean $\mathbb{E}Z_i^* = (e^{\beta i} - 1)/i$, and let Z_i, $1 \le i \le n$, be independent Poisson random variables with $\mathbb{E}Z_i = 1/i$, independent of the Z_i^*. Define $W = \sum_{i=1}^{b} iZ_i^*$. Then

$$\begin{aligned}\mathbb{Q}(T_{0n} = n) &= \mathbb{P}(T_{0n} + W = n)\\ &= \sum_{k=0}^{n} \mathbb{P}(T_{0n} = k)\mathbb{P}(W = n - k)\\ &= \mathbb{P}(T_{0n} = n) \sum_{k=0}^{n} \mathbb{P}(W = n - k)\\ &\le \mathbb{P}(T_{0n} = n),\end{aligned}$$

the last equality following from the observation (made in the remark after Lemma 3) that $\mathbb{P}(T_{0n} = k)$ is constant for $k = 0, 1, \ldots, n$. This completes the proof of the lemma. □

THEOREM 2. *As $n \to \infty$, $d_b(n) \to 0$ if and only if $b/n \to 0$. If $n/b \to \infty$, $d_b(n)$ decays superexponentially fast as a function of n/b. In fact, for all $1 \le b < n$,*

$$d_b(n) \le F(n/b),$$

where

$$F(x) \equiv \sqrt{2\pi m}\,\frac{2^{m-1}}{(m-1)!} + \frac{1}{m!} + 3\left(\frac{x}{e}\right)^{-x},$$

with $m \equiv \lfloor x \rfloor$, so that $\log F(x) \sim -x \log x$ as $x \to \infty$.

REMARK. Formula (3) shows that as $n \to \infty$, $F(n)/d_1(n) \sim n^{5/2}\sqrt{\pi/2}$. Thus, relative to the requirement that $F(n/b)$ be an upper bound on $d_b(n)$ for all $1 \le b \le n$, our $F(x)$ is suboptimal by at most a factor of $x^{5/2}\sqrt{\pi/2}$ as $x \to \infty$.

PROOF OF THEOREM 2. First we will show that if $b/n \ge \varepsilon > 0$ for all n, b, then $\liminf_{n\to\infty} d_b(n) > 0$. From the definition of $d_b(n)$ and the condition $\sum_{i=1}^{b} iC_i \le n$, note that

$$\begin{aligned}d_b(n) &\ge \mathbb{P}(T_{0b} > n)\\ &\ge \mathbb{P}\left(\sum_{b/2<i\le b} iZ_i > n\right)\\ &\ge \mathbb{P}\left(\frac{b}{2}\sum_{b/2<i\le b} Z_i > n\right)\\ &\ge \mathbb{P}\left(\sum_{b/2<i\le b} Z_i > \frac{2}{\varepsilon}\right)\\ &\to \mathbb{P}\left(\text{Poisson}(\log 2) > \frac{2}{\varepsilon}\right)\\ &> 0.\end{aligned}$$

If $b/n \geq \varepsilon > 0$ for infinitely many n, we may apply the argument above to an appropriate subsequence to establish the "only if" part of the theorem.

Let $m \equiv \lfloor n/b \rfloor$. Write $\delta_k \equiv |f_k/f_n - 1|$, where f is the density of T_{bn}. Using Lemmas 1 and 2, we have

$$
\begin{aligned}
d_b(n) &= \frac{1}{2}\sum_{k=0}^{\infty} \mathbb{P}(T_{0b} = k)\left|\frac{\mathbb{P}(T_{bn} = n-k)}{\mathbb{P}(T_{0n} = n)} - 1\right| \\
&\leq \sum_{k=0}^{\infty} \mathbb{P}(T_{0b} = k)\left|\frac{\mathbb{P}(T_{bn} = n-k)}{\mathbb{P}(T_{bn} = n)} - 1\right| \\
&= \sum_{k=-\infty}^{n} \mathbb{P}(T_{0b} = n-k)\delta_k \qquad (18) \\
&= \mathbb{P}(T_{0b} > n) + \sum_{k=1}^{n} \mathbb{P}(T_{0b} = n-k)\delta_k + \mathbb{P}(T_{0b} = n)\delta_0.
\end{aligned}
$$

A bound for the first term in (18) is obtained from (16):

$$\mathbb{P}(T_{0b} > n) \leq \left(\frac{n}{be}\right)^{-n/b}.$$

If $n \geq k > lb$, where $l \geq 1$ is an integer, Lemma 6 shows that $\delta_k \leq 1/l!$. For $k = 1, 2, \ldots, b$ we have $f_k = 0$ so that $\delta_k = 1$. Thus, for $l = 0, 1, \ldots, m$, the terms of the sum in (18) with $k = lb+1, lb+2, \ldots, lb+b$ and $k \leq n$ have $\delta_k \leq 1/l!$. Therefore

$$
\begin{aligned}
&\sum_{k=1}^{n} \mathbb{P}(T_{0b} = n-k)\delta_k \\
&\quad= \sum_{l=0}^{m-1}\sum_{k=lb+1}^{(l+1)b} \mathbb{P}(T_{0b} = n-k)\delta_k + \sum_{k=mb+1}^{n} \mathbb{P}(T_{0b} = n-k)\delta_k \\
&\quad\leq \sum_{l=0}^{m-1} \frac{1}{l!}\mathbb{P}(n-(l+1)b \leq T_{0b} < n-lb) + \frac{1}{m!}\sum_{k=mb+1}^{n} \mathbb{P}(T_{0b} = n-k) \\
&\quad\leq \sum_{l=0}^{m-1} \frac{1}{l!}\mathbb{P}(T_{0b} \geq n-(l+1)b) + \frac{1}{m!} \\
&\quad\leq \sum_{l=0}^{m-1} \frac{1}{l!}\mathbb{P}(T_{0b} \geq (m-l-1)b) + \frac{1}{m!} \\
&\quad\leq \sum_{l=0}^{m-1} \frac{1}{l!}\left(\frac{m-l-1}{e}\right)^{-(m-l-1)} + \frac{1}{m!} \\
&\quad\leq \sum_{l=0}^{m-l} \frac{1}{l!}\frac{\sqrt{2\pi(m-1)}}{(m-l-1)!} + \frac{1}{m!} \\
&\quad\leq \sqrt{2\pi m}\sum_{l=0}^{m-1} \frac{1}{l!(m-l-1)!} + \frac{1}{m!} \\
&\quad= \sqrt{2\pi m}\,\frac{2^{m-1}}{(m-1)!} + \frac{1}{m!}.
\end{aligned}
$$

The fourth-from-last inequality follows from (16), and the third-from-last from the fact that $k! < \sqrt{2\pi(k+1)}\,(k/e)^k$.

Finally, we estimate the rightmost term in (18). A direct calculation (compare the proof of Lemma 6) shows that $f_0/f_n \in [b+1, 2b+1]$, so that $b \le \delta_0 \le 2b$. Here, it would be too crude to use the bound $\mathbb{P}(T_{0b} = n) \le \mathbb{P}(T_{0b} \ge n) \le (n/b)^{-n/b}$ since n/b may go to infinity arbitrarily slowly compared to b. Instead, we argue that the $k = 0$ term is

$$\begin{aligned}\mathbb{P}(T_{0b} = n)\left|\frac{\mathbb{P}(T_{bn} = 0)}{\mathbb{P}(T_{bn} = n)} - 1\right| &< \mathbb{P}(T_{0b} = n)\frac{\mathbb{P}(T_{bn} = 0)}{\mathbb{P}(T_{bn} = n)} \\ &= \frac{\mathbb{P}(T_{0b} = n, T_{0n} = n)}{\mathbb{P}(T_{bn} = n)} \\ &\le \left(\frac{n}{be}\right)^{-n/b}\frac{\mathbb{P}(T_{0n} = n)}{\mathbb{P}(T_{bn} = n)},\end{aligned}$$

using (17). For the last factor we have

$$\begin{aligned}\frac{\mathbb{P}(T_{0n} = n)}{\mathbb{P}(T_{bn} = n)} &= \frac{\mathbb{P}(T_{0n} = 0)}{\mathbb{P}(T_{bn} = n)} \\ &\le (2b+1)\frac{\mathbb{P}(T_{0n} = 0)}{\mathbb{P}(T_{bn} = 0)} \qquad (19) \\ &= (2b+1)\mathbb{P}(T_{0b} = 0) \\ &< 2, \quad \text{for any } b \ge 1. \qquad \square\end{aligned}$$

3. Conditioned permutations. The proof of Theorem 2 is robust enough to yield a similar result for the cycle structure of random permutations given some fixed cycle conditions of the form $C_i(n) = c_i$ for $i \in J$. For example, the case of a randomly selected derangement is described by $J = \{1\}$, $c_1 = 0$. Gaussian limits for this situation are described in Flajolet and Soria (1990).

THEOREM 3. *For $1 \le b < n$, $J \subset \{1, \dots, b\}$ and nonnegative integers c_i, $i \in J$, satisfying $\sum_{i \in J} i c_i = s \le n$, consider the set of permutations on n objects having c_i cycles of length i for $i \in J$. If this set is nonempty, the cycle structure of a permutation chosen uniformly from this set satisfies*

$$\begin{aligned}d^* &\equiv \left\| \mathscr{L}((C_1, \dots, C_b) \mid C_i = c_i, \forall\, i \in J) - \mathscr{L}(Z_1^*, \dots, Z_b^*) \right\| \\ &\le F\left(\frac{n-s}{b}\right) + 2be\left(\frac{n-s}{be}\right)^{-(n-s)/b}.\end{aligned}$$

Here F is given in Theorem 2 and the Z_i^* are mutually independent with $Z_i^* \equiv c_i$ if $i \in J$ and $Z_i^* \equiv Z_i$, which is Poisson-distributed with mean $1/i$, if $i \notin J$.

PROOF. Analogous to (8) we have for $\mathbf{a} \in \mathbb{Z}_+^b$,

$$\mathbb{P}(\mathbf{C}_b = \mathbf{a} \mid C_i = c_i, \forall\, i \in J) = \mathbb{P}(\mathbf{Z}_b^* = \mathbf{a} \mid T_{0n}^* = n),$$

where $T_{lm}^* = \sum_{l<i\le m} iZ_i^*$. As in Lemma 1 we have

$$d^* = \frac{1}{2}\sum_{r=0}^{\infty} \mathbb{P}(T_{0b}^* = r)\left|\frac{\mathbb{P}(T_{bn}^* = n-r)}{\mathbb{P}(T_{0n}^* = n)} - 1\right|.$$

Observe that since $T_{bn}^* \equiv T_{bn}$, which is independent of T_{0b}^*, Lemma 2 yields the following analog of (18):

$$d^* \le \sum_{k=-\infty}^{n} \mathbb{P}(T_{0b}^* = n-k)\delta_k, \tag{20}$$

where, as before,

$$\delta_k = \left|\frac{\mathbb{P}(T_{bn} = k)}{\mathbb{P}(T_{bn} = n)} - 1\right|.$$

The $k = 0$ term of (20) requires special handling. For all other terms, the estimate in Theorem 2 was based on (16), namely, $\mathbb{P}(T_{0b} \ge n-k) \le ((n-k)/be)^{-(n-k)/(be)}$. The effect on this upper bound of replacing $\mathbf{Z}$ by $\mathbf{Z}^*$, that is, conditioning on $Z_i = c_i$ for $i \in J$, is to replace n by $n-s$, where $s = \sum_{i\in J} ic_i$. This follows because

$$\begin{aligned}
\mathbb{P}(T_{0b}^* \ge n-k) &= \mathbb{P}\Big(\sum_{i\le b,\, i\notin J} iZ_i \ge n-k-s\Big) \\
&\le \mathbb{P}\Big(\sum_{i\le b} iZ_i \ge n-k-s\Big) \\
&= \mathbb{P}(T_{0b} \ge n-k-s).
\end{aligned}$$

For the $k = 0$ term of (20), the bound used before, based on (17), is destroyed by the extra conditioning. In its place we have

$$\begin{aligned}
\mathbb{P}(T_{0b}^* = n)\delta_0 &< \mathbb{P}(T_{0b}^* \ge n)\frac{\mathbb{P}(T_{bn} = 0)}{\mathbb{P}(T_{bn} = 0)} \\
&\le \mathbb{P}(T_{0b} \ge n-s)\frac{\mathbb{P}(T_{bn} = 0)}{\mathbb{P}(T_{bn} = n)} \\
&= \frac{\mathbb{P}(T_{0b} \ge n-s)}{\mathbb{P}(T_{0b} = 0)}\,\frac{\mathbb{P}(T_{0n} = 0)}{\mathbb{P}(T_{bn} = n)} \\
&< \left(\frac{n-s}{be}\right)^{-(n-s)/b} 2be,
\end{aligned}$$

where the last inequality follows from the fact that $\mathbb{P}(T_{0b} = 0) = \exp(-(1 + \frac{1}{2} + \cdots + 1/b)) > e^{-1}/b$, combined with (19).

This completes the proof. □

4. Independence and Poisson marginals. Theorem 2 says that the process of counts of cycles of sizes 1 through b in a random permutation of n objects may be approximated, in total variation, by its limiting process, which has independent coordinates, if and only if b/n tends to zero. This is essentially a result about the *dependence* within the cycle counting process, rather than about the particular choice of cycle lengths one through b or the Poisson distribution of the limit.

4.1. The next two theorems show that the "if and only if $b/n \to 0$" conclusion of Theorem 2 is preserved even if we consider any b distinct cycle lengths, not just the b smallest, and even if we enlarge the class of comparison processes to include all processes with independent coordinates.

THEOREM 4. *Let $\mathcal{I}_n \subseteq \{1, 2, \ldots, n\}$ be an index set of size $b \equiv b_n = |\mathcal{I}_n|$. The total variation distance between the law of a family $(C_i(n), i \in \mathcal{I}_n)$ of cycle sizes and the corresponding family $(Z_i, i \in \mathcal{I}_n)$, where the Z_i are independent Poisson random variables with $\mathbb{E}Z_i = 1/i$, namely,*

$$d(n) \equiv d(\mathcal{I}_n, n) \equiv \sup_{A \subseteq \mathbb{Z}_+^b} \left| \mathbb{P}((C_i, i \in \mathcal{I}_n) \in A) - \mathbb{P}((Z_i, i \in \mathcal{I}_n) \in A) \right|,$$

tends to zero if and only if $b/n \to 0$. This is a consequence of the following bounds, valid for all $1 \le b \le n$:

$$\mathbb{P}(\text{Poisson}(b/(2n)) > 2n/b) \le d(\mathcal{I}_n, n) \le \inf_{y>1} \left(F(y) + 2y\frac{b}{n} \right),$$

where $F(\cdot)$, given in Theorem 2, has $F(y) \to 0$ as $y \to \infty$.

PROOF. Recall that the total variation distance between two random elements X and Y is the infimum of $\mathbb{P}(X \neq Y)$ over all couplings of X and Y on the same probability space. By compactness, there are maximal couplings, under which the total variation distance is $\mathbb{P}(X \neq Y)$. Now fix $y > 1$, and define $m = \lfloor n/y \rfloor$. Choose a maximal coupling of $(C_1, \ldots, C_m)$ with $(Z_1, \ldots, Z_m)$, and extend this to a coupling of the processes $(C_1, \ldots, C_n)$ and $(Z_1, \ldots, Z_n)$. Under this coupling,

$$\begin{aligned} d(n) &\le \mathbb{P}((C_i, i \in \mathcal{I}_n) \neq (Z_i, i \in \mathcal{I}_n)) \\ &\le \mathbb{P}((C_i : i \in \mathcal{I}_n, i \le m) \neq (Z_i : i \in \mathcal{I}_n, i \le m)) \\ &\quad + \sum_{i \in \mathcal{I}_n : i > m} \mathbb{P}(C_i \neq Z_i). \end{aligned}$$

To bound the second term above, recall that $\mathbb{E}C_i = \mathbb{E}Z_i = 1/i$, so that

$$\begin{aligned}\sum_{i\in\mathscr{I}_n:\, i>m} \mathbb{P}(C_i \neq Z_i) &\leq \sum_{i\in\mathscr{I}_n:\, i>m} \mathbb{P}(|C_i - Z_i| \geq 1) \\ &\leq \sum_{i\in\mathscr{I}_n:\, i>m} \frac{2}{i} \\ &\leq \frac{2y}{n}|\mathscr{I}_n| \\ &= 2y\frac{b}{n}.\end{aligned}$$

To bound the first term, note that by the construction of the coupling, it must be at most d_m. Hence we see from Theorem 2 that

$$\begin{aligned}d(n) &\leq F\left(\frac{n}{\lfloor n/y\rfloor}\right) + 2y\frac{b}{n} \\ &\leq F(y) + 2y\frac{b}{n}.\end{aligned}$$

It follows that

$$d(n) \leq \inf_{y>1}\left(F(y) + 2y\frac{b}{n}\right).$$

Since $F(y) \to 0$ as $y \to \infty$, we see that $d(n) \to 0$ if $b/n \to 0$.

For the converse, note that $|\{j \in \mathscr{I}_n\colon j \geq b/2\}| \geq b/2$. Define the Poisson random variable N by

$$N = \sum_{j\in\mathscr{I}_n:\, j\geq b/2} Z_j,$$

and note that

$$\mathbb{E}N = \sum_{j\in\mathscr{I}_n:\, j\geq b/2} \frac{1}{j} \geq \frac{b}{2}\frac{1}{n}.$$

Since $\{N > 2n/b\} \subseteq \{\sum_{j\in\mathscr{I}_n} jZ_j > n\}$, we see that

$$\begin{aligned}d(n) &\geq \mathbb{P}\left(\sum_{j\in\mathscr{I}_n} jZ_j > n\right) \\ &\geq \mathbb{P}\left(N > \frac{2n}{b}\right) \\ &\geq \mathbb{P}\left(\text{Poisson}\left(\frac{b}{2n}\right) > 2n/b\right).\end{aligned}$$

If, for fixed $\varepsilon > 0$, $|\mathscr{I}_n| \geq \varepsilon n$ for infinitely many n, then this argument may be applied along a suitably chosen subsequence to complete the proof. □

THEOREM 5. *If b/n is bounded away from zero, then no process with independent coordinates can successfully approximate $(C_i(n),\ i \in \mathscr{I}_n)$ in the total variation distance. More precisely, for all $\varepsilon > 0$,*

$$\min\left(\frac{\varepsilon}{4}, \mathbb{P}\left(\text{Poisson}\left(\frac{\varepsilon}{8}\right) > \frac{4}{\varepsilon}\right)\right) \leq \liminf_{n\to\infty} \min_{|\mathscr{I}_n| \geq \varepsilon n, \mathbf{Y}} \left\| \mathscr{L}(C_i, i \in \mathscr{I}_n) - \mathscr{L}(Y_i, i \in \mathscr{I}_n) \right\|,$$

where the minimum is taken over all choices of $\mathscr{I}_n \subseteq \{1, 2 \ldots n\}$ of size at least εn and over all choices of mutually independent random variables Y_i.

PROOF. Assume that $\mathscr{I}_n$ and $\mathbf{Y}$ are given, with $|\mathscr{I}_n|/n \geq \varepsilon > 0$. We may assume that the Y_i are nonnegative integer-valued, since for general $\mathbf{Y}$ the act of replacing with zero all values outside of $\{0, 1, 2, \ldots\}$ preserves the mutual independence of $\mathbf{Y}$ and does not increase the total variation distance to $(C_i(n),\ i \in \mathscr{I}_n)$. Let $A = \{i \in \mathscr{I}_n\colon \mathbb{P}(Y_i > 0) < 1/(2n)\}$.

Suppose for the first case that $|A| \geq \varepsilon n/2$, and let B consist of the $\lceil \varepsilon n/2 \rceil$ smallest elements of A. In a random permutation of $\{1, 2, \ldots, n\}$, the size of the cycle containing the element 1 is uniformly distributed over $\{1, 2, \ldots, n\}$, hence

$$\mathbb{P}\left(\sum_{i\in B} C_i > 0\right) \geq \frac{|B|}{n} \geq \frac{\varepsilon}{2}.$$

On the other hand,

$$\mathbb{P}\left(\sum_{i\in B} Y_i > 0\right) \leq \sum_{i\in B} \mathbb{P}(Y_i > 0) \leq \frac{|B|}{2n}.$$

Thus in this case, the total variation distance is at least $\mathbb{P}(\Sigma_{i\in B} C_i > 0) - \mathbb{P}(\Sigma_{i\in B} Y_i > 0) \geq \varepsilon/4$.

In the second case, suppose that $|A| < \varepsilon n/2$, so that

$$D \equiv \left\{ i \in \mathscr{I}_n\colon i \geq \frac{\varepsilon n}{4}, \mathbb{P}(Y_i \geq 1) \geq \frac{1}{2n} \right\}$$

has size $|D| \geq \varepsilon n/4$. Then we have

$$\begin{aligned}
\mathbb{P}\left(\sum_{i\in D} iY_i > n\right) &\geq \mathbb{P}\left(\frac{\varepsilon n}{4} \sum_{i\in D} Y_i > n\right) \\
&= \mathbb{P}\left(\sum_{i\in D} Y_i > \frac{4}{\varepsilon}\right) \\
&\geq \mathbb{P}\left(\text{Binomial}\left(\frac{1}{2n}, \left\lceil \frac{\varepsilon n}{4} \right\rceil\right) > \frac{4}{\varepsilon}\right) \\
&\to \mathbb{P}\left(\text{Poisson}\left(\frac{\varepsilon}{8}\right) > \frac{4}{\varepsilon}\right) > 0. \qquad \square
\end{aligned}$$

4.2. The following conjecture embodies the idea that independence, rather than Poisson marginals, is the dominant contribution to the distance $d_b(n)$ between the cycle counting process $(C_1, \dots, C_b)$ and its limiting Poisson process $(Z_1, \dots, Z_b)$.

CONJECTURE 1. *As* $n \to \infty$, *for any sequence* $b = b(n)$ *with* $1 \le b(n) \le n$, $d_b(n) \sim \min\{\|\mathscr{L}(C_1(n), \dots, C_b(n)) - \mathscr{L}(Y_1, \dots, Y_b)\|$: Y_i *independent*$\}$.

Note that trivially, the left-hand side is greater than or equal to the right-hand side, since the Poisson process $\mathbf{Z}$ is among the candidates $\mathbf{Y}$. Also, the process $(C_1'(n), \dots, C_b'(n))$, which has independent coordinates and the same marginals as the cycle counting process, is one of the candidates for $(Y_1, \dots, Y_b)$, but in general this choice does not attain the minimum.

4.3. We will now estimate the total variation distance $d_b'(n)$ between $(C_1', \dots, C_b')$ and $(Z_1, \dots, Z_b)$. Recall Goncharov's (1944) result about the marginal distribution of C_j, derivable via inclusion–exclusion:

$$\mathbb{P}(C_j = k) = \frac{j^{-k}}{k!} \sum_{m=0}^{\lfloor n/j \rfloor - k} (-1)^m \frac{j^{-m}}{m!}, \qquad k = 0, 1, \dots, \left\lfloor \frac{n}{j} \right\rfloor. \tag{21}$$

Arguments similar to those used to derive (3) establish the following upper bound on the total variation distance between C_j and Z_j:

$$d(\{j\}, n) \le \left(\frac{1}{j}\right)^{r+1} \frac{2^r}{(r+1)!}, \qquad r = \left\lfloor \frac{n}{j} \right\rfloor. \tag{22}$$

Indeed, for each fixed j, $d(\{j\}, n)$ is asymptotic to the right-hand side of (22).

Since $\|\mathscr{L}(C_j') - \mathscr{L}(Z_j)\| = d(\{j\}, n)$, we see from (22) that for $1 \le b \le n$,

$$\begin{aligned} d_b'(n) &\equiv \|\mathscr{L}(C_1', \dots, C_b') - \mathscr{L}(Z_1, \dots, Z_b)\| \\ &= 1 - \prod_{j=1}^{b} \left(1 - \|\mathscr{L}(C_j') - \mathscr{L}(Z_j)\|\right) \\ &\le \sum_{j=1}^{b} \|\mathscr{L}(C_j') - \mathscr{L}(Z_j)\| \\ &\le \sum_{j=1}^{b} \left(\frac{1}{j}\right)^{r+1} \frac{2^r}{(r+1)!}, \qquad r = \left\lfloor \frac{n}{j} \right\rfloor. \end{aligned}$$

Now define

$$\mathscr{B}_r = \left\{j: \frac{n}{r+1} < j \le \frac{n}{r}\right\} = \left\{j: \left\lfloor \frac{n}{j} \right\rfloor = r\right\}.$$

Then

$$d'_b(n) \le \sum_{r=\lfloor n/b \rfloor}^{n} \left(\sum_{j \in \mathscr{B}_r} \left(\frac{1}{j}\right)^{r+1} \frac{2^r}{(r+1)!} \right).$$

But

$$\begin{aligned}
\sum_{j \in \mathscr{B}_r} \left(\frac{1}{j}\right)^{r+1} \frac{2^r}{(r+1)!} &\le \frac{2^r}{(r+1)!} \left(\frac{r+1}{n}\right)^{r+1} |\mathscr{B}_r| \\
&\le \frac{2^r}{(r+1)!} \left(\frac{r+1}{n}\right)^{r+1} \left(\frac{n}{r(r+1)} + 1\right) \\
&\le \frac{2^r}{\sqrt{2\pi(r+1)}} \left(\frac{e}{n}\right)^{r+1} \left(\frac{n}{r(r+1)} + 1\right) \\
&\le \left(\frac{2e}{n}\right)^r \frac{2e}{\sqrt{2\pi}} \frac{1}{r} \sqrt{\frac{b}{n}},
\end{aligned}$$

the last inequality following from the observation that

$$\frac{1}{n\sqrt{r+1}} \left(\frac{n}{r(r+1)} + 1\right) \le \frac{2}{r} \sqrt{\frac{b}{n}}.$$

It follows that

$$d'_b(n) \le \frac{2e}{\sqrt{2\pi}} \sqrt{\frac{b}{n}} \sum_{r=\lfloor n/b \rfloor}^{n} \frac{1}{r} \left(\frac{2e}{n}\right)^r. \tag{23}$$

Equation (23) may be used to establish the following result, the proof of which is omitted.

PROPOSITION 1. *For $n \ge 6$,*

$$d'_b(n) \le F'_n\left(\frac{n}{b}\right),$$

where

$$F'_n(x) = \frac{2e}{\sqrt{2\pi x}} \frac{1}{m} \left(\frac{2e}{n}\right)^m \left(1 - \frac{2e}{n}\right)^{-1}, \qquad m \equiv \lfloor x \rfloor.$$

REMARK. For $b = n$, the bound in Proposition 1 simplifies to

$$d'_n(n) \le \frac{2\sqrt{2}\, e^2}{\sqrt{\pi}\,(1 - 2e/n)} \frac{1}{n} = O(n^{-1}).$$

The contrasting behavior of $d_n(n)$ is given in Section 5.2.

Observe that as n and $n/b \to \infty$,

$$\frac{F_n'(n/b)}{F(n/b)} \sim \frac{2\sqrt{2}e^2}{\sqrt{\pi}}\left(\frac{n}{b}\right)^{-5/2} b^{-\lfloor n/b \rfloor} \to 0.$$

Comparison with Table 1 in Section 5.3 shows that $d_b(n) \gg F_n'(n/b)$, so that the dominant contribution arises from dependence versus independence, as opposed to matching the marginals.

5. Complementary results.

5.1. The result of this paper is that the cycle counts for sizes up to b are asymptotically independent if and only if $b/n \to 0$. For the process of indicators of cycles of size at most b, which not only counts the cycles but also says which elements form each cycle, the boundary for successful approximation by an independent process is $b = \sqrt{n}$. This is shown using the Chen–Stein method in Arratia, Goldstein and Gordon (1990).

5.2. Lemma 1 and the remark following Lemma 3 show that

$$\begin{aligned} d_n(n) &= 1 - \mathbb{P}(T_{0n} = n) \\ &= 1 - \exp\left(-\sum_{j=1}^{n} \frac{1}{j}\right), \end{aligned} \tag{24}$$

so that as $n \to \infty$,

$$d_n(n) = 1 - \frac{e^{-\gamma}}{n}(1 + o(1)),$$

where γ is Euler's constant.

5.3. Here is the argument to show that if $b/n \to \beta \in [0, 1]$, then $d_b(n) \to H(\beta)$, where $H(0) = 0$, $H(1) = 1$ and H is strictly increasing in $[0, 1]$. It is a trivial consequence of Theorem 2 that if this limit $H(\beta)$ exists, then for $0 < \beta \leq 1$, $H(\beta) \leq F(1/\beta)$. It would be interesting to study the limiting behavior of $H(\beta)$ as $\beta \to 0$.

To evaluate $H(\beta)$ we look at the joint limit distribution of $(1/n)T_{0n}$ and $(1/n)T_{bn}$ with $b = \lfloor \beta n \rfloor$. More generally the process $[(1/n)T_{0,\lfloor \beta n \rfloor};\ 0 \leq \beta < \infty]$ converges in $\mathbb{D}[0, \infty)$ to a process $(X_\beta;\ 0 \leq \beta < \infty)$ which should play a fundamental role in understanding random permutations. There is a simple explicit construction of this limit process, as follows.

Let $\cdots T_{-1} < T_0 < 0 < T_1 < T_2 < \cdots$ be the jump times of a Poisson process on $\mathbb{R}$ with intensity 1, and let $J_i = e^{-T_i}$. Equivalently, let $\ldots, U_{-1}, U_0, U_1, U_2, \ldots \in (0, 1]$ be independent and uniformly distributed, and let $J_1 = U_1$, $J_2 = U_1U_2$, $J_3 = U_1U_2U_3, \ldots, J_0 = 1/U_0$, $J_{-1} = 1/(U_0U_{-1}), \ldots$.

For $\beta \geq 0$ define

$$X_\beta = \sum_i J_i 1(J_i \leq \beta). \tag{25}$$

This gives a process $(X_\beta;\ 0 \leq \beta < \infty)$ having independent increments, and sample paths which are right-continuous step functions, with $X_0 = 0$ and $X_\beta - X_{\beta-} \in \{0, \beta\}$ for all β. Note that $\mathbb{E}X_1 = 1$, $\operatorname{var}(X_1) = \frac{1}{2}$ and $X_\beta =_d \beta X_1$ for all $\beta \geq 0$.

If $b/n \to \beta \in [0,1]$, then $((1/n)T_{0b}, (1/n)T_{bn}) \to_d (X_\beta, X_1 - X_\beta)$. This follows from the observation that

$$\mathbb{E}\exp\left(\frac{u}{n}T_{bn}\right) = \exp\left(\sum_{j=b+1}^{n} \frac{1}{j}\left(\exp\left(\frac{ju}{n}\right) - 1\right)\right)$$

$$\to \exp\left(\int_\beta^1 \frac{1}{x}(\exp(ux) - 1)\,dx\right).$$

A local limit argument and Lemma 1 then show that if $b/n \to \beta \in (0,1]$, then

$$d_b(n) \to H(\beta) \equiv \|\mathscr{L}(X_\beta) - \mathscr{L}(X_\beta | X_1 = 1)\|. \tag{26}$$

Let g_β be the density of X_β, and let h_β be the density of the continuous part of $X_1 - X_\beta$. Note that $\mathbb{P}(X_1 - X_\beta = 0) = \beta$. The limiting total variation distance may also be written in the form

$$H(\beta) = \frac{1}{2}\left(\int_0^\infty g_\beta(x)\left|\frac{h_\beta(1-x)}{g_1(1)} - 1\right|dx + \beta\frac{g_\beta(1)}{g_1(1)}\right).$$

The last term above corresponds to the $r = n$ term of (12).

We have $g_\beta(x) = g_1(x/\beta)/\beta$ and $g_1(x) = e^{-\gamma}$ for $0 \leq x \leq 1$, and

$$g_1'(x) = \frac{-g_1(x-1)}{x}. \tag{27}$$

An explicit expression for the density of X_β may be found in Kolchin [(1986), page 54 ff]. Finally, the density h_β of the continuous part of $X_1 - X_\beta$ satisfies

$$h_\beta(x) = \begin{cases} 0, & x < \beta, \\ \beta/x, & \beta \leq x \leq 2\beta, \end{cases}$$

$$h_\beta'(x) = \frac{h_\beta(x-\beta) - h_\beta(x) - h_\beta(x-1)}{x}, \qquad x > \beta. \tag{28}$$

5.4. It is interesting to compare the exact values of d_b with the upper bound from Theorem 2. The special case $b = 1$ is discussed in the introduction, where upper and lower bounds for $d_1(n)$ are given in equation (3).

In Table 1, we give values of $d_b(n)$ and the upper bound $F(n/b)$ from Theorem 2. Because of numerical inaccuracy in the calculation of $d_b(n)$ using

TABLE 1
Exact and estimated total variation distances[a]

	n/b	8	10	12	14	20	30	40	50
	$F(n/b)$	1.81_{-1}*	1.12_{-2}	4.46_{-4}	1.23_{-5}	4.83_{-11}	8.34_{-22}	4.27_{-34}	1.64_{-47}
b	1	5.86_{-4}	2.20_{-5}	5.75_{-7}	1.11_{-8}	1.88_{-14}	1.23_{-25}	3.13_{-38}	6.99_{-52}
	2	5.14_{-5}	7.35_{-7}	7.09_{-9}	4.82_{-11}	2.90_{-18}	6.20_{-32}	3.93_{-47}	2.07_{-63}
	5	6.85_{-6}	4.26_{-8}	1.72_{-10}	4.81_{-13}	1.85_{-21}	2.56_{-37}	8.55_{-55}	1.81_{-73}
	10	3.01_{-6}	1.35_{-8}	3.84_{-11}	7.48_{-14}	8.99_{-23}	1.51_{-39}	5.29_{-58}	1.06_{-77}
	20	1.94_{-6}	7.24_{-9}	1.70_{-11}	2.71_{-14}	1.72_{-23}	8.93_{-41}	8.91_{-60}	4.76_{-80}

[a]Upper bound $F(n/b)$ for d_b from Theorem 2. Body of table is exact d_b from Lemma 2 and REDUCE code.
*The notation a_b means $a \times 10^b$

conventional programming languages, the values of d_b were computed using very high precision arithmetic in Version 3.3 of the computer algebra package REDUCE [Hearn (1987)]. The algorithm uses recursions analogous to (14) to compute the densities of T_{0b} and T_{bn}; it then calculates d_b using (12). The code is given below:

```
array f(1001),a(1001),ft0b(1001),ftbn(1001);
on numval;
on bigfloat;
procedure fprob(l,m,n);
begin for j:=0:n do a(j):=0;
      for j:=0:n do f(j):=0;
      a(0):=for j:=l+1:m sum -1/j;
      for j:=l+1:m do a(j):=1/j;
      f(0):=e**a(0);
      for k:=0:n-1 do f(k+1):=
       (for v:=1:k+1 sum v*a(v)*f(k+1-v))/(k+1);
end;
procedure dbn(b,n);
begin scalar ft0nn,tv;
      fprob(0,b,n);
      for j:=0:n do ft0b(j):=f(j);
      fprob(b,n,n);
      for j:=0:n do ftbn(j):=f(j);
      ft0nn:=e**(for j:=1:n sum -1/j);
      tv:=for r:=0:n sum ft0b(r)*abs(ftbn(n-r)/
      ft0nn-1);
      tv:=(tv+1-(for j:=0:n sum ft0b(j)))/2;
      return tv
end;
precision 85;
```

The structure of Table 1 suggests the following conjectures:

CONJECTURE 2. *For fixed b, $d_b(n)$ is a decreasing function of n.*

CONJECTURE 3. *For fixed n/b, $d_b(n)$ is a decreasing function of b and n.*

If both these conjectures were proved, it would follow that, for all $n \geq b \geq 1$, $d_b(n) \leq d_b(b\lfloor n/b \rfloor) \leq d_1(\lfloor n/b \rfloor)$, with sharp bounds on d_1 given by (3).

6. Discussion.

6.1. The techniques above are useful for analyzing other combinatorial structures, such as random partitions of a set and random mappings. Suppose the number of "M-structures" on a set of size i is m_i. Consider "M-assemblies" on $\{1, \ldots, n\}$, in which the set $\{1, \ldots, n\}$ is partitioned, and an M-structure is given for each block of the partition [Joyal (1981)]. For example, if the M-structure is a cyclic permutation, then the M-assembly is a permutation, represented by its cycle decomposition. For an M-assembly π, let $C_i(\pi)$ be the number of blocks of size i, so that $C_1 + 2C_2 + \cdots + nC_n \equiv n$. The analog of Cauchy's formula for the number of M-assemblies with a given block structure is

$$\left|\{\pi\colon (C_1(\pi), \ldots, C_n(\pi)) = (a_1, \ldots, a_n)\}\right| = n! \prod_{i=1}^{n} \left(\frac{m_i}{i!}\right)^{a_i} \frac{1}{a_i!} 1\left(\sum_{j=1}^{n} j a_j = n\right),$$

so that if an M-assembly on $\{1, \ldots, n\}$ is chosen uniformly at random and Z_i are independent Poisson random variables with parameters

$$\mathbb{E}Z_i = \frac{m_i}{i!},$$

then

$$\mathbb{P}((C_1, \ldots, C_n) = \mathbf{a}) = \mathbb{P}\left((Z_1, \ldots, Z_n) = \mathbf{a} \middle| \sum_{j=1}^{n} jZ_j = n\right). \tag{29}$$

Furthermore, for any $x > 0$, the conditional probability on the right-hand side is unchanged if the Poisson parameters are changed to

$$\mathbb{E}Z_i = \frac{m_i x^i}{i!}, \tag{30}$$

where x may depend on n but not on i.

Permutations is the case in which the M-structure is a cyclic permutation, with $m_i = (i-1)!$ and $m_i/i! = 1/i$. The collection of all n^n mappings of a set of n elements into itself is the case in which $m_i = (i-1)!\sum_{k=0}^{i-1} i^k/k!$. Using $x = e^{-1}$ in (30) yields $\mathbb{E}Z_i \sim 1/(2i)$ as $i \to \infty$ [see Arratia and Tavaré (1992b)]. The Ewens sampling formula [Ewens (1972)] is a nonuniform measure on permutations that satisfies (29) with $\mathbb{E}Z_i = \theta/i$ for fixed $\theta > 0$ [see Watterson

(1974a, b), and Arratia, Barbour and Tavaré (1992)]. Random partitions of a set is the case $m_i \equiv 1$. This can be compared to a Poisson process by using $x = x(n)$, the solution of $xe^x = n$ in (30) [see Arratia and Tavaré (1992b)].

The analog of Lemma 1 in this general setting is worth recording as the following lemma.

LEMMA 9. *If the distribution of* $(C_1, \ldots, C_n)$ *is equal to that of* $(Z_1, \ldots, Z_n)$ *conditional on* $T_{0n} = n$, *where* $T_{0n} \equiv Z_1 + 2Z_2 + \cdots + nZ_n$, *then, for any set* $B \subset \{1, 2, \ldots, n\}$, *the total variation distance for the processes restricted to* B *is equal to the total variation distance between the random variables* $T_B \equiv \sum_{i \in B} iZ_i$ *and* T_B *conditioned on* $T_{0n} = n$:

$$(31) \quad \| \mathscr{L}(C_i, i \in B) - \mathscr{L}(Z_i, i \in B) \| = \| \mathscr{L}(T_B) - \mathscr{L}(T_B | T_{0n} = n) \|.$$

PROOF. Proceed exactly as in the proof of Lemma 1. Note that the only condition used is the mutual independence of the Z_i, and that Poisson marginals are not needed for the Z_i. □

6.2. The explicit bound in Theorem 2 effectively allows the small cycle sizes to be decoupled into independent Poisson random variables. This decoupling provides elementary proofs of limit theorems and bounds for a variety of functionals of random permutations. The details of these and other applications appear in Arratia and Tavaré (1992a). Among these limit theorems are Goncharov's (1944) result that the number of cycles of a random n-permutation is asymptotically normal, with mean and variance $\log n$, and its functional version [DeLaurentis and Pittel (1985)] that considers the cycles of size at most n^t as a process in t, $0 \le t \le 1$. Analogous results for random mappings appear in Hansen (1989), and for the Ewens sampling formula in Hansen (1990). Another application concerns the celebrated Erdös–Turán theorem, which states that the log of the order of a random n-permutation is asymptotically normal, with mean $(\log n)^2/2$ and variance $(\log n)^3/3$. DeLaurentis and Pittel (1985) give a short proof of the Erdös–Turán theorem using their functional limit theorem.

Acknowledgments. Richard Arratia would like to thank the Courant Institute for its hospitality during the fall, 1988. We thank Gian-Carlo Rota for teaching us about M-assemblies, and we thank an anonymous referee for helpful comments.

REFERENCES

ARRATIA, R., BARBOUR, A. D. and TAVARÉ, S. (1992). Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.* **2**.

ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1990). Poisson approximation and the Chen–Stein method. *Statist. Sci.* **5** 403–434.

ARRATIA, R. and TAVARÉ, S. (1992a). Limit theorems for combinatorial structures via discrete process approximations. *Random Structures and Algorithms*. In press.

ARRATIA, R. and TAVARÉ, S. (1992b). Independent process approximations for random combinatorial structures. *Adv. in Math.* To appear.

BARBOUR, A. D. (1990). Comment on "Poisson Approximation and the Chen-Stein Method" by R. Arratia, L. Goldstein and L. Gordon. *Statist. Sci.* **5** 425–427.

DAVID, F. N. and BARTON, D. E. (1962). *Combinatorial Chance*. Griffin, London.

DELAURENTIS, J. M. and PITTEL, B. (1985). Random permutations and Brownian motion. *Pacific J. Math.* **119** 287–301.

DIACONIS, P. and FREEDMAN, D. (1980). Finite exchangeable sequences. *Ann. Probab.* **8** 745–764.

DIACONIS, P. and PITMAN, J. W. (1986). Unpublished lecture notes. Statistics Department, Univ. California, Berkeley.

ERDÖS, P. and TURÁN, P. (1967). On some problems of statistical group theory. III. *Acta Math. Acad. Sci. Hungar.* **18** 309–320.

EWENS, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoret. Population Biol.* **3** 87–112.

FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications* **1**, 3d Ed. Wiley, New York.

FLAJOLET, P. and SORIA, M. (1990). Gaussian limiting distributions for the number of components in combinatorial structures. *J. Combin. Theory Ser. A* **53** 165–182.

GONCHAROV, V. L. (1944). Some facts from combinatorics. *Izv. Akad. Nauk SSSR, Ser. Mat.* **8** 3–48. [See also: On the field of combinatory analysis. *Transl. Amer. Math. Soc.* **19** (1962) 1–46.]

HANSEN, J. C. (1989). A functional central limit theorem for random mappings. *Ann. Probab.* **17** 317–332. [Correction (1991) **19** 1393–1396.]

HANSEN, J. C. (1990). A functional central limit theorem for the Ewens sampling formula. *J. Appl. Probab.* **27** 28–43.

HEARN, A. C. (1987). REDUCE-3 User's Manual, Version 3.3. Rand Corporation Publication CP78.

JOYAL, A. (1981). Une théorie combinatoire des séries formelles. *Adv. in Math.* **42** 1–82.

KOLCHIN, V. F. (1971). A problem of the allocation of particles in cells and cycles of random permutations. *Theory Probab. Appl.* **16** 74–90.

KOLCHIN, V. F. (1986). *Random Mappings*. Optimization Software, Inc., New York.

POURAHMADI, M. (1984). Taylor expansion of $\exp(\sum_{k=0}^{\infty} a_k z^k)$ and some applications. *Amer. Math. Monthly* **91** 303–307.

SHEPP, L. A. and LLOYD, S. P. (1966). Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.* **121** 340–357.

WATTERSON, G. A. (1974a). Models for the logarithmic species abundance distributions. *Theoret. Population Biol.* **6** 217–250.

WATTERSON, G. A. (1974b). The sampling theory of selectively neutral alleles. *Adv. in Appl. Probab.* **6** 463–488.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTHERN CALIFORNIA
LOS ANGELES, CALIFORNIA 90089-1113