The Annals of Probability
1983, Vol. 11, No. 2, 428–441

ON DISTRIBUTIONS RELATED TO TRANSITIVE CLOSURES OF RANDOM FINITE MAPPINGS

BY BORIS PITTEL[1]

The Ohio State University

For f, a random single-valued mapping of an n-element set X into itself, let f^{-1} be the inverse mapping, and f^* be such that $f^*(x) = \{f(x)\} \cup f^{-1}(x)$, $x \in X$. For a given subset $A \subset X$, introduce three random variables $\xi(A) = |\hat{f}(A)|$, $\eta(A) = |\hat{f}^{-1}(A)|$, and $\zeta(A) = |\hat{f}^*(A)|$, where $\hat{f}, \hat{f}^{-1}, \hat{f}^*$ stand for transitive closures of f, f^{-1}, f^*. The distributions of $\xi(A)$ and $\zeta(A)$ are obtained. ($\eta(A)$ was earlier studied by J. D. Burtin.) For large n, the asymptotic behavior of those distributions is studied under various assumptions concerning $m = |A|$. For instance, it is shown that $\xi(A)$ is asymptotically normal with mean $(2mn)^{1/2}$ and variance $n/2$, and $(n - \zeta(A))(n/m)^{-1}$ is asymptotically $\mathscr{U}^2/2$ ($\mathscr{U}$ being the standard normal variable), provided $m \to \infty$, $m = o(n)$. The results are interpreted in terms of epidemic processes on random graphs introduced by I. Gertsbakh.

1. Introduction. Let $X = \{x_1, \cdots, x_n\}$ be a finite set, and let F be the set of all single-valued mappings f or X into itself; clearly $|F| = n^n$. Introducing the uniform distribution on F we obtain the random mapping f which assumes all its values with the same probability n^{-n}. Thus all conceivable characteristics of f become random variables. There are fairly many results [1–17] concerning various characteristics of a random digraph G_f representing f. (It is the graph having X as a set of its vertices such that, for each x_i, $x_j \in X$, an arc goes from x_i to x_j iff $x_j = f(x_i)$.)

Gertsbakh [3] suggested using G_f for models of epidemic processes: namely, X is considered as a population; a contagious disease is initially confined to a subset (subpopulation) $A \subset X$, and it is then spread to other vertices (individuals) along arcs of G_f, the latter being interpreted as the description of acquaintance-type relations in the population. There were described three versions regarding whether the infection is spread only forward, i.e. according to the orientation of arcs, or only backward, or in both directions. For the former two, Gertsbakh proved that, asymptotically almost certainly, no significant fraction of the population will be infected if $|A| = o(\sqrt{n})$, $n \to \infty$. For the latter, an essential portion of X will be infected, provided $|A| \to \infty$ however slowly. Clearly, these are asymptotic properties of the distributions of the random variables $\xi(A)$, $\eta(A)$, and $\zeta(A)$ described as follows. Let f^{-1} be the inverse mapping, and let f^* be such that

$$f^*(x) = \{f(x)\} \cup f^{-1}(x), \quad x \in X.$$

Denote $\hat{f}, \hat{f}^{-1}$, and $\hat{f}^*$ the transitive closures of f, f^{-1}, and f^*, so that, for instance,

$$\hat{f}(x) = \{x\} \cup \{f(x)\} \cup \{f(f(x))\} \cup \cdots, \quad x \in X.$$

Then

$$\xi(A) = |\hat{f}(A)|, \quad \eta(A) = |\hat{f}^{-1}(A)|, \quad \zeta(A) = |\hat{f}^*(A)|.$$

It is interesting to note that, for $|A| = 1$, the distributions of $\xi(A)$ and $\eta(A)$, and their asymptotics, have long been known (Harris [4], Rubin and Sitgreaves [14]). For $|A| = m$

Received December 1981; revised July 1982.

[1] Research supported by the NSF under grant MCS 8002966.

AMS 1980 *subject classification.* Primary 60C05; secondary 60F05, 60J80, 05C30.

Key words and phrases. Random mappings and graphs, enumeration, generating functions, distributions, limit laws, epidemic processes.

> 1, Burtin [1] obtained a surprisingly simple formula for the distribution of $\eta(A)$. The derivation is based on his nontrivial result concerning a more general random mapping. (Another characteristic of this mapping, the total number of components, was studied by Ross [13].) It made it possible to carry out a detailed asymptotical analysis of the distribution of $\eta(A)$ under diverse assumptions regarding m, n. For instance, $m = [\sqrt{n}]$ was proven to be the threshold function for this *inverse* epidemic process: namely, $\eta(A)/n$ tends to 0 or 1 in probability regarding whether $m = o(\sqrt{n})$ or $\sqrt{n} = o(m)$. (J. D. Burtin obtained these results in 1977, several months before his untimely death.)

We shall derive exact formulas for two other distributions, namely those of $\xi(A)$ and $\zeta(A)$. The derivation involves solving two enumeration problems: let $L(\mu, \nu)$, respectively $M(\mu, \nu)$, count the total number of mappings f of a $\mu + \nu$-element set $\mathscr{X}$ into itself such that

$$\xi(\mathscr{A}) = \mu + \nu, \quad \text{respectively} \quad \zeta(\mathscr{A}) = \mu + \nu,$$

for a given μ-element set $\mathscr{A}$ of $\mathscr{X}$. Then (1) and (2) hold whenever $|A| = m$.

$$P(\xi(A) = m + s) = \binom{n-m}{s} L(m, s) n^{-m-s}, \tag{1}$$

$$P(\zeta(A) = m + s) = \binom{n-m}{s} M(m, s)(n - m - s)^{n-m-s} n^{-n}. \tag{2}$$

It turns out (Lemmas 1, 2) that

$$L(\mu, \nu) = (\mu + \nu) \sum_{j=0}^{\nu} (-1)^{\nu-j} \binom{\nu}{j} (\mu + j)^{\mu+\nu-1}, \tag{3}$$

$$M(\mu, \nu) = (\mu + \nu)^{\mu+\nu} - \sum_{j=1}^{\nu} \binom{\nu}{j} j^{j-1} (\mu + \nu - j)^{\mu+\nu-j}. \tag{4}$$

(A similar number for $\eta(A)$ can be shown to be the number of forests on $\mu + \nu$ vertices, which consists of μ disjoint trees such that some μ fixed vertices belong to distinct trees, times $(\nu + \mu)^{\mu}$. By Cayley's formula for forests, this number is therefore $\mu(\nu + \mu)^{\nu+\mu-1}$. This relation can be used for an alternative derivation of Burtin's formula for the distribution of $\eta(A)$.)

Using (1–4), we shall prove the following two theorems about asymptotical properties of $\xi(A)$ and $\zeta(A)$ for large n.

THEOREM 1. (a) *Let m $(= |A|)$ be fixed; then*

$$P(\xi(A) n^{-1/2} \le x) \to (2^{m-1}(m-1)!)^{-1} \int_0^x y^{2m-1} \exp(-y^2/2)\, dy, \quad x > 0. \tag{5}$$

(b) *Let $m \to \infty$ and $n - m \to \infty$; then*

$$P((\xi(A) - a)/\sqrt{b} \le x) \to (2\pi)^{-1/2} \int_{-\infty}^x \exp(-y^2/2)\, dy, \quad -\infty < x < \infty, \tag{6}$$

where a is the root of the equation

$$\exp(a/n) = (n - m)(n - a)^{-1}, \tag{7}$$

and

$$b = \frac{(n - a)[a^2(n - m) + a\,mn - mn^2]}{(n - m)a^2}. \tag{8}$$

In particular,

$$a \sim \sqrt{2mn}, \quad b \sim n/2 \quad \textit{if} \quad m = o(n),$$

$$a - m \sim (n - m)(1 - e^{-1}), \quad b \sim (n - m)(1 - e^{-1})e^{-1} \quad \textit{if} \quad n - m = o(n).$$

If $m \sim xn$, $x \in (0, 1)$, then $a \sim \alpha n$, $b \sim \beta n$ where α, β are determined from

$$e^{\alpha} = (1 - x)(1 - \alpha)^{-1}, \quad \beta = \frac{(1 - \alpha)[\alpha^2(1 - x) + \alpha x - x]}{(1 - x)\alpha^2}. \tag{9}$$

(c) *Finally, let $n - m$ be fixed. Then $\xi(A) - m$ is asymptotically binomial with parameter $1 - e^{-1}$.*

Note. Loosely speaking, this Theorem indicates that in the direct epidemic process the number of all eventually infected individuals is of order $(mn)^{1/2}$. It also shows that the distribution of $\xi(A) - m$, i.e. of the number of *newly* infected points, is concentrated asymptotically at values of order $o(n)$, unless m/n is bounded away both from 0 and 1. In the latter case, the distribution of $\xi(A) - m$ is concentrated around $a - m$, a being determined by (7). It is easy to prove the following: let z_0 be the root of an equation $ze^z = 1$; then $z_0 \approx 0.57$ and

$$\max_m(a - m) = (a - m)|_{m=m_0} \sim (z_0 + z_0^{-1} - 2) \cdot n \approx 0.33\, n, \quad m_0 \sim (2 - z_0^{-1})n \approx 0.24n.$$

THEOREM 2. (a) *Let m be fixed; then*

$$P(\zeta(A)n^{-1} \le x) \to c_m \int_0^x (1 - y)^{-1/2} \cdot y^{m-1}\, dy, \quad x \in [0, 1], \tag{10}$$

$$c_m = 2^{-1}(2m - 1)!!/(2(m - 1))!!, \tag{11}$$

so that $\zeta(A)n^{-1}$ is asymptotically beta-distributed with parameters m and ½.
(b) *Let $m \to \infty$, $m = o(n)$; then*

$$P((n - \zeta(A))(n/m)^{-1} \le x) \to \int_0^x (\pi y)^{-1/2} e^{-y}\, dy, \quad x > 0, \tag{12}$$

so, as a consequence,

$$p \lim n^{-1}\zeta(A) = 1.$$

(c) *Let $mn^{-1} \to \alpha \in (0, 1)$; then*

$$P(n - \zeta(A) = k) \to (1 - \rho)k^k\gamma^k/k!, \quad k = 0, 1, 2, \cdots, \tag{13}$$

where $\gamma = (1 - \alpha)e^{-1}$ and ρ is the root of the equation

$$\rho = (1 - \alpha)\exp(\rho - 1). \tag{14}$$

Note. Thus, for the two-sided epidemic process, an essential fraction of the elements are eventually infected even if only one element is initially infected. (The expected value of this fraction approaches to ⅔ as $n \to \infty$.) Furthermore, with high probability, nearly all the elements will be infected provided m grows with n however slowly. Quite crudely, the number of elements the spreading infection will not reach is of the order n/m.

2. Evaluation of $L(\nu, \mu)$, $M(\nu, \mu)$. Let f be a mapping of a finite set into itself. It is well known, and can be easily proven, that each component of the digraph G_f has just one cycle. (The vertices constituting cycles are called cyclic.) The arcs which do not belong to cycles form trees each having exactly one cyclic point, and the arcs of each tree are oriented toward its cyclic point. It is natural to consider a cyclic point as the root of the tree it belongs to. A mapping is called connected if its digraph has a single component.

Denote $t(n)$ the total number of rooted trees on n labelled vertices, $C(n)$ the number of connected mappings and $F(n)(= n^n)$ the number of *all* mappings of $\{1, \cdots, n\}$ into itself. (By Cayley's formula, $t(n) = n^{n-1}$.) The exponential generating functions

$$t(x) = \sum_{n=1}^{\infty} t(n)x^n/n!, \quad F(x) = \sum_{n=1}^{\infty} F(n)x^n/n!, \quad C(x) = \sum_{n=1}^{\infty} C(n)x^n/n!$$

are known to satisfy

(15) $$t(x) = x\exp(t(x)),$$ (Pólya [8])

(16) $$F(x) = t(x)(1 - t(x))^{-1},$$ (Riordan [11])

(17) $$1 + F(x) = \exp(C(x)).$$

Let $\mathscr{X} = (x_1, \cdots, x_{\mu+\nu})$, $\mathscr{A} = (x_1, \cdots, x_\mu)$. Denote (see Introduction)

$$L(\mu, \nu) = |\{f : \mathscr{X} \to \mathscr{X} \mid \xi(\mathscr{A}) = \mu + \nu\}|,$$

$$M(\mu, \nu) = |\{f: \mathscr{X} \to \mathscr{X} \mid \zeta(\mathscr{A}) = \mu + \nu\}|.$$

LEMMA 1.

(18) $$M(\mu, \nu) = (\mu + \nu)^{\mu+\nu} - \sum_{j=1}^{\nu} \binom{\nu}{j} j^{j-1}(\mu + \nu - j)^{\mu+\nu-j}, \quad \mu + \nu \geq 1.$$

PROOF. Notice first that $\zeta(\mathscr{A}) = \mu + \nu$ iff each component of G_f has at least one element of $\mathscr{A}$. Hence

(19) $$M(\mu, \nu) = \sum_{k\geq 1} (k!)^{-1} \sum_{\mu_1+\cdots+\mu_k=\mu;\nu_1+\cdots+\nu_k=\nu;\mu_s\geq 1,\nu_s\geq 0} \binom{\mu}{\mu_1, \cdots, \mu_k} \cdot \binom{\nu}{\nu_1, \cdots, \nu_k} \prod_{s=1}^{k} C(\mu_s + \nu_s).$$

So, introducing the exponential generating function of $M(\mu, \nu)$,

$$\begin{aligned}\sum_{\mu+\nu\geq 1} M(\mu, \nu)x^\mu y^\nu/\mu!\nu! &= \sum_{k\geq 1} (k!)^{-1} \sum_{\mu_s\geq 1;\nu_s\geq 0} \prod_{s=1}^{k} C(\mu_s + \nu_s)x^{\mu_s}y^{\nu_s}/\mu_s!\nu_s! \\ &= \sum_{k\geq 1} (k!)^{-1} \left[\sum_{\mu\geq 1;\nu\geq 0} C(\mu + \nu)x^\mu y^\nu/\mu!\nu!\right]^k \\ &= \exp\left[\sum_{\mu\geq 1;\nu\geq 0} C(\mu + \nu)x^\mu y^\nu/\mu!\nu!\right] - 1.\end{aligned}$$

Here

$$\begin{aligned}\sum_{\mu\geq 1;\nu\geq 0} C(\mu + \nu)x^\mu y^\nu/\mu!\nu! &= \sum_{n=1}^{\infty} C(n)/n!\left[\sum_{\mu+\nu=n;\mu\geq 1} \binom{n}{\mu} x^\mu y^\nu\right] \\ &= \sum_{n=1}^{\infty} C(n)/n![(x + y)^n - y^n] = C(x + y) - C(y).\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{\mu+\nu\geq 1} M(\mu, \nu)x^\mu y^\nu/\mu!\nu! &= \exp(C(x + y))/\exp(C(y)) - 1 \\ &\qquad \text{(by (17))} \\ &= (1 + F(x + y))/(1 + F(y)) - 1 \\ &\qquad \text{(by (16))} \\ &= (1 + F(x + y))(1 - t(y)) - 1.\end{aligned}$$

Hence

$$\begin{aligned}M(\mu, \nu) &= \mu!\nu! \operatorname{coeff}_{x^\mu y^\nu}[(1 + F(x + y)) \cdot (1 - t(y)) - 1] \\ &= \mu!\nu!\left[\operatorname{coeff}_{x^\mu y^\nu} \sum_{j=0}^{\infty} j^j(x + y)^j/j! - \operatorname{coeff}_{x^\mu y^\nu} \sum_{j_1\geq 1;j_2\geq 1} j_1^{j_1-1} \cdot j_2^{j_2} \cdot y^{j_1}(x + y)^{j_2}/j_1!j_2!\right] \\ &= \mu!\nu!\left[(\mu + \nu)^{\mu+\nu}\binom{\mu + \nu}{\mu}\Big/(\mu + \nu)! - \sum_{j_1+j_2=\mu+\nu;j_1\geq 1,j_2\geq\mu} j_1^{j_1-1} \cdot j_2^{j_2}\binom{j_2}{\mu}\Big/j_1!j_2!\right] \\ &= (\mu + \nu)^{\mu+\nu} - \sum_{j=1}^{\nu} \binom{\nu}{j} j^{j-1}(\mu + \nu - j)^{\mu+\nu-j}.\end{aligned}$$

Note. For a partial check, let $\mu = 0$ and $\nu \geq 1$ in (18). Then clearly $M(0, \nu) = 0$ and we get

$$\nu^\nu = \sum_{j=1}^{\nu} \binom{\nu}{j} j^{j-1}(\nu - j)^{\nu-j},$$

which is one of Abel's identities, [12].

LEMMA 2.

$$L(\mu, \nu) = (\mu + \nu) \sum_{j=0}^{\nu} (-1)^{\nu-j}\binom{\nu}{j}(\mu + j)^{\mu+\nu-1}. \tag{20}$$

PROOF. (1) Let $t(\mu, \nu)$ be the number of all rooted trees on $\mu + \nu$ vertices $x_1, \cdots, x_{\mu+\nu}$ such that all leaves belong to $\mathscr{A} = (x_1, \cdots, x_\mu)$. This would imply $t(0, \nu) = 0$ for all $\nu \geq 1$; however for notational convenience in the sums below we define $t(0, 1) = 1$. Introduce the exponential generating function $t(x, y) = \sum_{\mu+\nu\geq 1} t(\mu, \nu)x^\mu y^\nu/\mu!\nu!$ of these trees. We want to show that

$$t(x, y) = (x + y)\exp(t(x, y) - y). \tag{21}$$

To begin, notice that

$$t(\mu, 0) = t(\mu), \quad \mu \geq 1, \quad \text{and} \quad t(0, 1) = 1, \quad t(0, \nu) = 0 \quad \text{for} \quad \nu \geq 2. \tag{22}$$

Let $\mu \geq 1$, $\nu \geq 1$. A feasible tree has a certain root, some $k \geq 1$ adjacent vertices and the correspondent k feasible subtrees rooted at them, so that the sth subtree has some μ_s (resp. ν_s) elements of $\mathscr{A}$ (resp. $\mathscr{A}^c$), $1 \leq s \leq k$, subject to restriction

$$\mu_s + \nu_s \geq 1, \quad (\mu_s, \nu_s) \neq (0, 1). \tag{23}$$

Therefore

$$t(\mu\ \nu) = \sum_{k=1}^{\infty} (k!)^{-1}\Bigg[\mu \cdot \sum_{\mu_1+\cdots+\mu_k=\mu-1;\nu_1+\cdots+\nu_k=\nu;(23)} \binom{\mu - 1}{\mu_1, \cdots, \mu_k}\binom{\nu}{\nu_1, \cdots, \nu_k} \prod_{s=1}^{k} t(\mu_s, \nu_s)$$
$$+ \nu \cdot \sum_{\mu_1+\cdots+\mu_k=\mu;\nu_1+\cdots+\nu_k=\nu-1;(23)} \binom{\mu}{\mu_1,\cdots, \mu_k}\binom{\nu - 1}{\nu_1, \cdots, \nu_k} \prod_{s=1}^{k} t(\mu_s, \nu_s)\Bigg].$$

By (22),

$$\begin{aligned} t(x, y) &= t(x) + y + \sum_{\mu\geq 1;\nu\geq 1} t(\mu, \nu)x^\mu y^\nu/\mu!\nu! \\ &= t(x) + y + \sum_{k=1}^{\infty} (k!)^{-1}[x \cdot \sum_{\mu_1+\cdots+\mu_k\geq 0;\nu_1+\cdots+\nu_k\geq 1;(23)} \prod_{s=1}^{k} t(\mu_s, \nu_s)x^{\mu_s}y^{\nu_s}/\mu_s!\nu_s! \\ &\quad + y \cdot \sum_{\mu_1+\cdots+\mu_k\geq 1;\nu_1+\cdots+\nu_k\geq 0;(23)} \prod_{s=1}^{k} t(\mu_s, \nu_s)x^{\mu_s}y^{\nu_s}/\mu_s!\nu_s!]. \end{aligned} \tag{24}$$

Denoting two sums inside the square brackets $\sum_1$ and $\sum_2$, we have

$$\begin{aligned} \textstyle\sum_1 &= \sum_{\mu_s+\nu_s\geq 1;(\mu_s,\nu_s)\neq(0,1)} \prod_{s=1}^{k} t(\mu_s, \nu_s)x^{\mu_s}y^{\nu_s}/\mu_s!\nu_s! - \sum_{\mu_s\geq 1} \prod_{s=1}^{k} t(\mu_s, 0)x^{\mu_s}/\mu_s! \\ &= [\textstyle\sum_{\mu+\nu\geq 1;(\mu,\nu)\neq(0,1)} t(\mu, \nu)x^\mu y^\nu/\mu!\nu!]^k - [\sum_{\mu\geq 1} t(\mu)x^\mu/\mu!]^k \\ &= (t(x, y) - y)^k - t^k(x), \end{aligned} \tag{25}$$

and, similarly,

$$\textstyle\sum_2 = (t(x, y) - y)^k - (\sum_{\nu\geq 2} t(0, \nu)y^\nu/\nu!)^k = t^k(x, y). \tag{26}$$

By (24–26),

$$\begin{aligned} t(x, y) &= t(x) + y + (x + y) \cdot \textstyle\sum_{k=1}^{\infty} (t(x, y) - y)^k/k! - x\sum_{k=1}^{\infty} t^k(x)/k! \\ &= (x + y)\exp(t(x, y) - y) + t(x) - x\exp(t(x)) \\ &= (x + y)\exp(t(x, y) - y), \qquad \text{(by (15))} \end{aligned}$$

(21) is proven.

(2) Let $C(\mu, \nu)$ be the number of all *connected* mappings f of $\mathscr{X} = (x_1, \cdots, x_{\mu+\nu})$ into itself such that $\xi(\mathscr{A}) = \mu + \nu$, $\mathscr{A} = (x_1, \cdots, x_\mu)$. The graph G_f of any such mapping f consists of an *oriented* cycle and a forest of disjoint trees rooted at, and oriented toward, cyclic points. Obviously, the leaves of each tree must belong to $\mathscr{A}$. Let $\mu \geq 1$, $\nu \geq 1$; then

$$C(\mu, \nu) = \sum_{k\geq 1} k^{-1} \cdot \sum_{\mu_1+\cdots+\mu_k=\mu;\nu_1+\cdots+\nu_k=\nu;\mu_s+\nu_s\geq 1} \binom{\mu}{\mu_1, \cdots, \mu_k} \cdot \binom{\nu}{\nu_1, \cdots, \nu_k} \prod_{s=1}^{k} t(\mu_s, \nu_s), \tag{27}$$

$t(\cdot, \cdot)$ is defined in the preceding section. (Really, G_f has some $k \geq 1$ cyclic points, so that elements of $\mathscr{A}$ and $\mathscr{A}^c$ are distributed among k disjoint trees whose roots form an oriented cycle. Besides, the leaves of each tree must belong to $\mathscr{A}$.) Also

$$C(\mu, 0) = C(\mu), \mu \geq 1, \quad \text{and} \quad C(0, \nu) = 0, \nu \geq 1. \tag{28}$$

In view of (27), (28),

$$\begin{aligned} C(x, y) &= \textstyle\sum_{\mu+\nu\geq 1} C(\mu, \nu)x^\mu y^\nu/\mu!\, \nu! = C(x) + \sum_{\mu\geq 1;\nu\geq 1} C(\mu, \nu)x^\mu y^\nu/\mu!\, \nu! \\ &= C(x) + \textstyle\sum_{k\geq 1} k^{-1} \cdot \sum_{\mu_1+\cdots+\mu_k\geq 1;\nu_1+\cdots+\nu_k\geq 1;\mu_s+\nu_s\geq 1} \prod_{s=1}^{k} t(\mu_s, \nu_s)x^{\mu_s}y^{\nu_s}/\mu_s!\, \nu_s!. \end{aligned}$$

Now,

$$\begin{aligned} \textstyle\sum_{\mu_1+\cdots+\mu_k\geq 1;\nu_1+\cdots+\nu_k\geq 1;\mu_s+\nu_s\geq 1} &= \textstyle\sum_{\mu_s+\nu_s\geq 1;1\leq s\leq k} - \sum_{\mu_s\geq 1,\nu_s=0;1\leq s\leq k} - \sum_{\mu_s=0,\nu_s\geq 1;1\leq s\leq k} \\ &= t^k(x, y) - t^k(x) - y^k. \end{aligned}$$

Hence

$$\begin{aligned} C(x, y) &= C(x) + \textstyle\sum_{k\geq 1} k^{-1} \cdot [t^k(x, y) - t^k(x) - y^k] \\ &= \ln[(1 - y)(1 - t(x, y))^{-1}] + C(x) - \ln[(1 - t(x))^{-1}], \end{aligned}$$

and, by (16), (17),

$$C(x, y) = \ln[(1 - y)(1 - t(x, y))^{-1}]. \tag{29}$$

(3) Introduce $L(x, y) = \sum_{\mu+\nu\geq 1} L(\mu, \nu)x^\mu y^\nu/\mu!\, \nu!$, the generating function of all (not necessarily connected) mappings such that $\xi(A) = \mu + \nu$. We have then

$$1 + L(x, y) = \exp(C(x, y)), \tag{30}$$

a relation which is of the same kind as (17) and can be proven in essentially the same way as (17), [11].

(4) By (29), (30),

$$L(x, y) = \exp(C(x, y)) - 1 = (1 - y)(1 - t(x, y))^{-1} - 1.$$

Besides, according to (21),

$$t(x, y) = z \exp(t(x, y)), \qquad z = (x + y)e^{-y},$$

and, by (15), (16),

$$(1 - t(x, y))^{-1} = 1 + t(x, y)(1 - t(x, y))^{-1} = 1 + F(z).$$

So

$$L(x, y) = (1 - y)(1 + F(z)) - 1 = (1 - y) \textstyle\sum_{j\geq 0} j^j (x + y)^j e^{-jy}/j! - 1,$$

and

$$L(\mu, \nu) = \mu!\ \nu!\ [\text{coeff}_{x^\mu y^\nu} \textstyle\sum_{j\geq 0} j^j(x+y)^j e^{-jy}/j! - \text{coeff}_{x^\mu y^{\nu-1}} \sum_{j\geq 0} j^j(x+y)^j e^{-jy}/j!],$$

which after some elementary transformations yields

$$L(\mu, \nu) = (\mu + \nu) \textstyle\sum_{j=0}^{\nu} (-1)^{\nu-j}\binom{\nu}{j}(\mu + j)^{\mu+\nu-1}.$$

Note. For a partial check of the last relation, choose $\mu = 1$. It is easy to see that $L(1, \nu) = (\nu + 1)!$. On the other hand,

$$(\nu + 1) \textstyle\sum_{j=0}^{\nu} (-1)^{\nu-j}\binom{\nu}{j}(1 + j)^{\nu} = (\nu + 1) \sum_{k=0}^{\nu} \binom{\nu}{k}\left[\sum_{j=0}^{\nu} (-1)^{\nu-j}\binom{\nu}{j} j^k\right]$$

$$= (\nu + 1) \textstyle\sum_{k=0}^{\nu} \binom{\nu}{k} A(k, \nu),$$

where $A(k, \nu)$ is the number of distributions of k balls among ν cells leaving none of them empty, [12]. So

$$(\nu + 1) \textstyle\sum_{k=0}^{\nu} \binom{\nu}{k} A(k, \nu) = (\nu + 1)\binom{\nu}{\nu} A(\nu, \nu) = (\nu + 1) \cdot \nu! = (\nu + 1)!,$$

too.

3. Distributions of $\xi(A)$, $\zeta(A)$ and their asymptotics. Recall that

$$P(\xi(A) = m + s) = \binom{n-m}{s} L(m, s) n^{-(m+s)}, \quad 0 \leq s \leq n - m, \tag{31}$$

$$P(\zeta(A) = m + s) = \binom{n-m}{s} M(m, s)(n - m - s)^{n-m-s}/n^n, \quad 0 \leq s \leq n - m, \tag{32}$$

$|A| = m$.

PROOF OF THEOREM 1. By Lemma 2, we obtain

$$L(m, s) = (m + s) \textstyle\sum_{j=0}^{s} (-1)^{s-j}\binom{s}{j}(m + j)^{m+s-1}$$

$$= (m + s)(d^{m+s-1}/dx^{m+s-1})\left[\textstyle\sum_{j=0}^{s}(-1)^{s-j}\binom{s}{j} e^{(m+j)x}\right]_{|x=0}$$

$$= (m + s)(d^{m+s-1}/dx^{m+s-1})[e^{mx}(e^x - 1)^s]_{|x=0}.$$

Then, by Cauchy's formula,

$$L(m, s) = (2\pi i)^{-1}(m + s)! \int_C e^{mz}(e^z - 1)^s z^{-(m+s)}\, dz, \tag{33}$$

where C stands for a counterclockwise oriented contour in the complex plane surrounding the origin.

(a) Let m be fixed and $s \to \infty$. In (33) choose C the circle of the radius s^{-1}. Set $z = s^{-1}e^{i\varphi}$, $\varphi \in [-\pi, \pi)$; then

$$e^{mz} = 1 + o(1),$$

$$(e^z - 1)^s = z^s(1 + z/2 + O(z^2))^s = s^{-s}\exp(is\varphi + e^{i\varphi}/2 + o(1)),$$

uniformly over $\varphi \in [-\pi, \pi)$. Hence

$$L(m, s) = (m + s)!(2\pi i)^{-1} \int_{-\pi}^{\pi} s^{-s}\exp(is\varphi + e^{i\varphi}/2)s^{m+s}e^{-i(m+s)\varphi}is^{-1} \cdot e^{i\varphi}(1 + o(1))\, d\varphi$$

$$= (m + s)!s^{m-1}(2\pi)^{-1} \int_{-\pi}^{\pi} \exp(e^{i\varphi}/2)e^{-i(m-1)\varphi} \cdot (1 + o(1))\, d\varphi$$

$$\sim (m + s)!s^{m-1}(2\pi)^{-1} \int_{-\pi}^{\pi} \exp(e^{i\varphi}/2)e^{-i(m-1)\varphi}\, d\varphi$$

$$\cdot (\exp(e^{i\varphi}/2) = \textstyle\sum_{k=0}^{\infty} e^{ik\varphi}/2^k k!)$$

$$= (m + s)!s^{m-1}/2^{m-1}(m - 1)!,$$

or

(34) $$L(m, s) \sim (m + s)!s^{m-1}/2^{m-1}(m - 1)!, \quad s \to \infty.$$

Let $n \to \infty$, $s \to \infty$ so that for some $c > 0$

(35) $$x = (m + s)n^{-1/2} \le c.$$

Then, by Stirling's formula and (31), (34), uniformly over x satisfying (35), there is

$$P(\xi(A)n^{-1/2} = x) \sim (2^{m-1}(m - 1)!s)^{-1} \cdot \exp[m \ln s + (n - m)\ln(n - m)$$

$$- (n - m - s)\ln(n - m - s) - s \ln s + (s + m)\ln(s + m) - (s + m) - (m + s)\ln n]$$

(expanding logarithms)

$$\sim (2^{m-1}(m - 1)!)^{-1}x^{2m-1}\exp(-x^2/2) \cdot n^{-1/2}.$$

Therefore, for a fixed $x > 0$,

$$P(\xi(A)n^{-1/2} \le x) \to (2^{m-1}(m - 1)!)^{-1} \cdot \int_0^x y^{2m-1}\exp(-y^2/2)\, dy.$$

(b) Let now $m \to \infty$ and $n - m \to \infty$. In view of (33) and Stirling's formula, (31) becomes: for $n - u \to \infty$, $u - m \to \infty$,

(36) $$P(\xi(A) = u) = \left(\frac{(n - m)u}{(n - m)(u - m)}\right)^{1/2} \cdot (2\pi i)^{-1} \int_C \exp(h(z, u))\, dz,$$

(37) $$\begin{aligned} h(z, u) = {} & (n - m)\ln(n - m) - (n - u)\ln(n - u) \\ & - (u - m)\ln(u - m) + u \ln u - u \\ & - u \ln n + mz + (u - m)\ln(e^z - 1) - u \ln z. \end{aligned}$$

To estimate the integral in (36), we shall apply the saddle-point method. To this end, let us choose C the circle of the radius $r = r(u)$ which is the root of the equation

(38) $$\partial h/\partial z = m + (u - m)e^z(e^z - 1)^{-1} - uz^{-1} = 0.$$

As for u, it will suffice to consider the case when it is close to a, the root of the equation

(39) $$\partial h/\partial u\,|_{z=r(u)} = \ln[(n - u)u(e^r - 1)/(u - m)nr] = 0.$$

Combination of (38), (39) shows after some manipulations that

(40) $$\exp(a/n) = (n - m)/(n - a),$$

(41) $$\rho = r(a) = a/n,$$

and, somewhat unexpectedly, that

$$h(\rho, a) = 0. \tag{42}$$

It can be proven that

$$a \begin{cases} = (2mn)^{1/2}(1 + o(1)), & \text{if } m = o(n), \\ \in [c_1 n, c_2 n], \quad 0 < c_1 \le c_2 < 1 & \text{if none of } m, n - m \text{ is } o(n), \\ = m + (1 - e^{-1})(n - m)(1 + o(1)), & \text{if } n - m = o(n). \end{cases} \tag{43}$$

Let u in (36) be such that

$$|u - a| \le c(n - m)^{1/2}, \tag{44}$$

$c > 0$ and fixed. It follows then from (38), (41), and (43) that

$$r = \begin{cases} \rho(1 + O(m^{-1/2})), & \text{if } m = o(n), \\ \rho(1 + O(n^{-1/2})), & \text{if neither } m \text{ nor } n - m \text{ is } o(n), \\ \rho(1 + O((n - m)^{1/2} \cdot n^{-1})), & \text{if } n - m = o(n). \end{cases} \tag{45}$$

Setting $z = re^{i\varphi}$ in (36), we get

$$\begin{gathered} \int_C \exp(h(z, u))\, dz = ir \exp(h(r, u)) \int_{-\pi}^{\pi} \exp(H(\varphi))\, d\varphi, \\ H(\varphi) = h(re^{i\varphi}, u) - h(r, u) - i\varphi. \end{gathered} \tag{46}$$

Here, since

$$|e^z - 1| \le (e^{|z|} - 1)\exp[(Rl z - |z|)/2],$$

we have (see also (37), (41), (43))

$$\begin{aligned} Rl H(\varphi) &\le mr(\cos\varphi - 1) + (u - m)r(\cos\varphi - 1)/2 \\ &\le - c_1 ur\varphi^2, \qquad c_1 > 0. \end{aligned} \tag{47}$$

Expanding $H(\varphi)$ in powers of φ shows that

$$\begin{aligned} H(\varphi) &= - i\varphi + i\varphi r[m + (u - m)e^r(e^r - 1)^{-1} - ur^{-1}] - (\varphi^2/2)\alpha(\tilde{z}, u) \\ &= - i\varphi - (\varphi^2/2)\alpha(\tilde{z}, u), \end{aligned} \qquad \text{(by (38))} \tag{48}$$

where

$$\alpha(z, u) = z[u + (u - m)(e^z - 1 - ze^z)(e^z - 1)^{-2}],\ \tilde{z} = \tilde{r}e^{i\tilde{\varphi}},\ |\tilde{\varphi}| \in [0, |\varphi|]. \tag{49}$$

Since

$$-0.5 \le (e^z - 1 - ze^z)(e^z - 1)^{-2} \le 0, \quad z \ge 0,$$

$$0.5\, ru \le \alpha(r, u) \le ru; \tag{50}$$

so, in particular, (see (41), (43), (45))

$$\lim_{n\to\infty} \alpha(r, u) = \infty. \tag{51}$$

Using (47 − 51), we can prove the existence of $\varepsilon = \varepsilon(n)$ such that $\varepsilon \to 0$, $\varepsilon\, (\alpha(r, u))^{1/2} \to \infty$, for which

$$\begin{aligned} \int_{-\pi}^{\pi} \exp(H(\varphi))\, d\varphi &\sim \int_{-\varepsilon}^{\varepsilon} \exp(H(\varphi))\, d\varphi \sim (\alpha(r, u))^{-1/2} \int_{-\varepsilon(\alpha(r,u))^{1/2}}^{\varepsilon(\alpha(r,u))^{1/2}} \exp(-y^2/2)\, dy \\ &\sim (2\pi/\alpha(r, u))^{1/2}. \end{aligned} \tag{52}$$

Hence (see (36), (46))

$$P(\xi(A) = u) \sim (2\pi b(u))^{-1/2}\exp(h(u)),\ h(u) = h(r(u), u), \tag{53}$$

(54) $$b(u) = \frac{(n-u)(u-m)\alpha(r,u)}{(n-m)ur^2}.$$

To simplify (53), observe first that, according to (40 − 45) and (49),

(55) $$b(u) \sim b(a) = \frac{(n-a)[a^2(n-m) + amn - mn^2]}{(n-m)a^2} = O(n-m).$$

Second, expanding $h(u)$ in powers of $u - a$ and invoking (38), (39) and (42), we get

(56) $$h(u) = (u-a)^2 h''(\tilde{u})/2, \quad \tilde{u} = \theta a + (1-\theta)u, \quad \theta \in [0, 1].$$

To evaluate h'', notice that (see (38), (39))

$$\begin{aligned} h'(u) &= \partial h(r,u)/\partial u + (\partial h(r,u)/\partial r)\cdot r'(u) \\ &= \partial h(r,u)/\partial u = \ln[(n-u)u(e^r - 1)/(u-m)nr] \\ &= \ln[(n-u)/n(1-r)], \\ r' &= -\frac{\partial^2 h}{\partial r \partial u} \Big/ \frac{\partial^2 h}{\partial u^2} = m\{u(e^r-1)[(u-m)e^r(e^r-1)^{-2} - ur^{-2}]\}^{-1}, \end{aligned}$$

so $h''(u) = -(n-u)^{-1} + r'(1-r)$.
Two last relations, together with (40), (41), imply that

$$-h''(a) = a^2(n-m)\{(n-a)[a^2(n-m) + amn - mn^2]\}^{-1},$$

or, (see (55)), that

$$-h''(a) = b^{-1}(a). \quad (!)$$

With the aid of (41) and (43–45), it can be shown then that (56) is replaced by

$$h(u) = -b^{-1}(a)\cdot(u-a)^2/2 + o(1), \quad |u-a| \le c(n-m)^{1/2}.$$

Putting it together with (53), (55) we conclude that

$$P(\xi(A) = u) \sim (2\pi b)^{-1/2}\exp(-b^{-1}(u-a)^2/2), \quad b = b(a),$$

uniformly over these u's. As a direct consequence of this *local* limit property and the estimate $b = O(n-m)$, it follows then: for fixed $x_1 < x_2$,

$$P(x_1 \le (\xi(A) - a)/b^{1/2} \le x_2) \to (2\pi)^{-1/2}\int_{x_1}^{x_2} \exp(-y^2/2)\,dy.$$

(c) Let $n - m = k$ be fixed. By (20), (31): for $0 \le s \le k$,

$$\begin{aligned} P(\xi(A) = m+s) &= \binom{k}{s} n^{-(n-k)-s}(n-k+s)\textstyle\sum_{j=0}^{s} (-1)^{s-j}\binom{s}{j}(n-k+j)^{n-k+s-1} \\ &= \binom{k}{s}\textstyle\sum_{j=0}^{s} (-1)^{s-j}\binom{s}{j}(1-(k-j)n^{-1})^n + o(1), \end{aligned}$$

so

$$\begin{aligned} \lim_{n\to\infty} P(\xi(A) = m+s) &= \binom{k}{s}\textstyle\sum_{j=0}^{s} (-1)^{s-j}\binom{s}{j}e^{-(k-j)} \\ &= \binom{k}{s}(1-e^{-1})^s(e^{-1})^{k-s}. \end{aligned}$$

Proof of Theorem 2. Let m be fixed. By (19),

(57) $$\begin{aligned} M(m,s) &= \textstyle\sum_{k=1}^{m} M_k(m,s), \\ M_k(m,s) &= m!s!(k!)^{-1}\cdot \textstyle\sum_{m_1+\cdots+m_k=m;\, s_1+\cdots+s_k=s;\, m_t\ge 1,\, s_t\ge 0} \prod_{t=1}^{k} C(m_t+s_t)/m_t!\,s_t!. \end{aligned}$$

To estimate $M_k(m, s)$ for $s \to \infty$, introduce $\tilde{M}_k(m, s)$ the sum similar to $M_k(m, s)$ except that $s_t \geq [s^{1/2}]$, $t = 1, \cdots, k$; obviously $M_k \geq \tilde{M}_k$. Since

$$C(\nu) = (\pi/2)^{1/2}\nu^{\nu-1/2}(1 + o(1)),\ \nu \to \infty,$$

(Katz [5], and Rényi[10]), and Stirling's formula, we have: for $s_t \geq [s^{1/2}]$,

$$\begin{aligned} C(m_t + s_t)/m_t!s_t! &\gtrsim \frac{(\pi/2)^{1/2}(m_t + s_t)^{m_t+s_t-1/2}}{m_t!(2\pi s_t)^{1/2}(s_t/e)^{s_t}} \\ &= (1/(2m_t!))[s_t(m_t + s_t)]^{-1/2}(m_t + s_t)^{m_t}(1 + m_t/s_t)^{s_t}\cdot e^{s_t} \\ &\gtrsim (1/(2m_t!))\, s_t^{m_t-1} \exp(m_t + s_t). \end{aligned}$$

(Note that $A \gtrsim B$ means that $\lim A/B \geq 1$.) Hence

$$\begin{aligned} \tilde{M}_k(m, s) \gtrsim m!s! \exp(m + s)(2^k k!)^{-1} &\textstyle\sum_{m_1+\cdots+m_k=m;m_j\geq1} \prod_{j=1}^k (m_j!)^{-1} \\ &\cdot\textstyle\sum_{s_1+\cdots+s_k=s;s_t\geq[s^{1/2}]} \prod_{t=1}^k s_t^{m_t-1}. \end{aligned} \tag{58}$$

Denote the innermost sum $\sum_k(\vec{m}, s)$. A standard argument shows that

$$\begin{aligned} \textstyle\sum_k(\vec{m}, s) &\gtrsim s^{m-1} \int \cdots \int_{x_1+\cdots+x_{k-1}\leq1;x_t\geq0} \left(\prod_{t=1}^{k-1} x^{m_t-1}\right)\left(1 - \sum_{t=1}^{k-1} x_t\right)^{m_k-1} dx_1 \cdots dx_{k-1} \\ &= s^{m-1}[(m-1)!]^{-1}\textstyle\prod_{t=1}^k (m_t - 1)!. \end{aligned}$$

Therefore, (see (58)),

$$\begin{aligned} \tilde{M}_k(m, s) &\gtrsim ms! \exp(m + s)s^{m-1}f(m, k), \\ f(m, k) &= (2^k k!)\cdot \textstyle\sum_{m_1+\cdots+m_k=m;m_j\geq1} \prod_{j=1}^k m_j^{-1}, \end{aligned} \tag{59}$$

and, consequently,

$$\begin{aligned} \tilde{M}(m, s) &= \textstyle\sum_{k=1}^m \tilde{M}_k(m, s) \gtrsim ms! \exp(m + s)s^{m-1}f(m), \\ f(m) &= \textstyle\sum_{k=1}^m f(m, k). \end{aligned} \tag{60}$$

To evaluate $f(m)$, introduce $F(z) = \sum_{m=1}^\infty f(m)z^m$. By (59), (60),

$$\begin{aligned} F(z) &= \textstyle\sum_{m=1}^\infty z^m\left(\sum_{k=1}^m (2^k k!)^{-1}\cdot\sum_{m_1+\cdots+m_k=m;m_j\geq1} \prod_{j=1}^m m_j^{-1}\right) \\ &= \textstyle\sum_{k=1}^\infty (2^k k!)^{-1}\cdot\sum_{m_1\geq1,\ldots,m_k\geq1} \prod_{j=1}^k z^{m_j}/m_j \\ &= \textstyle\sum_{k=1}^\infty (2^k k!)^{-1}[\ln(1 - z)^{-1}]^k \\ &= \exp[(1/2)\ln(1 - z)^{-1}] - 1 = (1 - z)^{-1/2} - 1. \end{aligned}$$

Therefore

$$f(m) = \operatorname{coeff}_{z^m}[(1 - z)^{-1/2} - 1] = (-1)^m\binom{-1/2}{m} = (2m - 1)!!/2^m m!. \tag{61}$$

Combination of (57), (60), and (61) yields an estimate

$$\begin{aligned} M(m, s) &\gtrsim s! \exp(m + s)s^{m-1}c_m, \\ c_m &= (1/2)[(2m - 1)!!]/[2(m - 1)!!]. \end{aligned} \tag{62}$$

Now, consider $P(\zeta(A) = \nu)$ for $\nu \to \infty$, $n - \nu \to \infty$. By (32), (62),

$$P(\zeta(A) = \nu) \gtrsim c_m s! \exp(m + s)s^{m-1}\binom{n - m}{s} (n - m - s)^{n-m-s}/n^n, \qquad s = \nu - m.$$

Here

$$s!\binom{n-m}{s} = (n-m)!/(n-\nu)! \gtrsim (1-\nu/n)^{-1/2}$$

$$\cdot \exp[(n-m)\ln(n-m) - (n-\nu)\ln(n-\nu) + m - \nu]$$

$$\sim (1-\nu/n)^{-1/2}\exp[(n-m)\ln n - (n-\nu)\ln(n-\nu) - \nu],$$

and, after cancellations,

$$\begin{aligned} P(\zeta(A)=\nu) &\gtrsim c_m(1-\nu/n)^{-1/2}s^{m-1}\exp(-m\ln n) \\ &\gtrsim (n^{-1})c_m(1-\nu/n)^{-1/2}(\nu/n)^{m-1} \\ &= g(x_\nu)\Delta x_\nu, \qquad x_\nu = \nu/n, \quad \Delta x_\nu = n^{-1}, \end{aligned} \tag{63}$$

where

$$g(x) = c_m(1-x)^{-1/2}x^{m-1}, \quad x \in [0, 1).$$

The last estimate is uniform over ν such that

$$\nu \leq (1-\varepsilon)n, \quad \varepsilon \in (0, 1) \text{ and fixed.}$$

Observe that $g(x)$ is the density of the beta-distribution G_0 with parameters m and ½. Let G be the limit of a subsequence of the distributions of $\zeta(A)/n$. By (63),

$$G(x_2) - G(x_1) \geq G_0(x_2) - G_0(x_1),$$

whenever $x_1, x_2 \in (0, 1)$ are continuity points of G. This inequality implies that $G(x) \equiv G_0(x)$, $x \in [0, 1]$. Hence the distribution of $\zeta(A)/n$ converges to G_0 as $n \to \infty$ *in any way*.

To study the case $m \to \infty$, we shall need the following.

LEMMA 3. *For* $\mu \geq 1$,

$$M(\mu, \nu) \geq (\mu+\nu)^{\mu+\nu}[1 - t(u)], \tag{64}$$

$$t(x) = \textstyle\sum_{j=1}^{\infty} t(j)x^j/j!, \quad u = \nu(\mu+\nu-1)^{\mu+\nu-1}/(\mu+\nu)^{\mu+\nu} < e^{-1}, \tag{65}$$

where $t(x)$ *is the exponential generating function of rooted trees*, $t(j) = j^{j-1}$.

PROOF OF LEMMA 3. By Lemma 1,

$$M(\mu, \nu) = (\mu+\nu)^{\mu+\nu} - \textstyle\sum_{j=1}^{\nu} (j^{j-1}/j!)(\nu)_j(\mu+\nu-j)^{\mu+\nu-j}. \tag{66}$$

First, let us prove that

$$g(j) = (\nu)_j(\mu+\nu-j)^{\mu+\nu-j} \leq (\mu+\nu)^{\mu+\nu}\cdot u^j, \quad 0 \leq j \leq \nu. \tag{67}$$

It is trivially true for $j = 0$, whence it suffices to show that $u^{-j}g(j)$ decreases as j increases. We have

$$u^{-(j+1)}g(j+1)/u^{-j}g(j) = u^{-1}\Psi(\nu-j),$$

$$\Psi(z) = z(\mu+z-1)^{\mu+z-1}/(\mu+z)^{\mu+z}.$$

By taking the logarithmic derivative of Ψ it can be shown that $\Psi(z)$ is an increasing function, provided $\mu \geq 1$. Therefore, by the definition of $g(j)$ and u,

$$u^{-(j+1)}g(j+1)/u^{-j}g(j) \leq u^{-1}g(1)/g(0) = 1.$$

Second, it is well known that the series for $t(x)$ converges provided $|x| < e^{-1}$. Observe that, by monotonicity of Ψ,

$$u = \Psi(\nu) < \lim_{z\to\infty} \Psi(z) = e^{-1}.$$

Hence, by (65–67),

$$M(\mu, \nu) \geq (\mu + \nu)^{\mu+\nu}[1 - \textstyle\sum_{j=1}^{\nu} (j^{j-1}/j!)u^j]$$
$$> (\mu + \nu)^{\mu+\nu}[1 - \textstyle\sum_{j=1}^{\infty} (j^{j-1}/j!)u^j]$$
$$= (\mu + \nu)^{\mu+\nu}[1 - t(u)].$$

(b) Suppose $m \to \infty$, $m = o(n)$. Let $k \to \infty$, $k \leq c(n/m)$. By (32) and Lemma 3,

$$(68)\qquad \begin{aligned} P(n - \zeta(A) = k) &= \binom{n-m}{k} M(m, n-k-m)k^k/n^n \\ &> [1 - t(u)]\binom{n-m}{k}(n-k)^{n-k}k^k/n^n, \end{aligned}$$

where

$$(69)\qquad u = (n-k-m)(n-k-1)^{n-k-1}/(n-k)^{n-k}.$$

Exponentiating the right side of (69), and expanding the resulting logarithms, shows that

$$(70)\qquad u = e^{-1}[1 - (m/n)(1 + O(m^{-1}))].$$

Also, it easily follows from (15) that, for $x \to e^{-1}$ from the left,

$$1 - t(x) \sim (2(1 - ex))^{1/2},$$

whence, (see (68)),

$$(71)\qquad \begin{aligned} P(n - \zeta(A) = k) &\gtrsim (2m/n)^{1/2}\binom{n-m}{k}(n-k)^{n-k}k^k/n^n \\ &\qquad\qquad \text{(by Stirling's formula)} \\ &\gtrsim (m/\pi kn)^{1/2}\exp[W(m) - W(0)], \\ W(x) &= (n-x)\ln(n-x) - (n-x-k)\ln(n-x-k). \end{aligned}$$

But

$$(72)\qquad \begin{aligned} W(m) - W(0) &= W'(0)m + W''(\theta)m^2/2 \\ &\qquad\qquad (\theta \in [0, m]) \\ &= m\ln(1 - k/n) + O(m^2k/n^2) \\ &= -m(k/n) + O(mk^2/n^2 + m^2k/n^2) \\ &\qquad\qquad (k \leq c(n/m)) \\ &= -m(k/n) + O(m^{-1} + mn^{-1}) \\ &= -m(k/n) + o(1). \end{aligned}$$

In view of (71), (72),

$$P(n - \zeta(A) = k) \gtrsim q(y_k)\Delta y_k, \quad y_k = k(n/m)^{-1}, \quad \Delta y_k = (n/m)^{-1},$$
$$q(y) = (\pi y)^{-1/2}e^{-y}.$$

Let us notice that $q(y)$ is the density of the distribution of $\mathcal{U}^2/2$ where $\mathcal{U}$ is the standard normal variable. Therefore, as in the item (a) of the current proof, we conclude that

$$P((n - \zeta(A))(n/m)^{-1} \leq x) \to \int_0^x \varphi(y)\,dy, \quad x > 0.$$

(c) Finally, suppose $m \sim \alpha n$, $\alpha \in (0, 1)$. Let k be fixed. According to (68), (69),

$$P(n - \zeta(A) = k) \gtrsim (1 - t(u))(k^k/k!)(1 - k/n)^{n-k}$$
$$\cdot\exp[(n-m)\ln(n-m) - (n-m-k)\ln(n-m-k) - k\ln n - k],$$

$$u \sim (1 - m/n)(1 - 1/(n-k))^{n-k} \to (1-\alpha)e^{-1}.$$

Furthermore,

$$\begin{aligned}&(n-m)\ln(n-m) - (n-m-k)\ln(n-m-k) - k\ln n - k\\ &= (n-m)\ln(n-m) - (n-m-k)[\ln(n-m) - k/(n-m)] - k\ln n - k + o(1)\\ &= k\ln(1-m/n) + o(1) = k\ln(1-\alpha) + o(1),\end{aligned}$$

whence

$$P(n - \zeta(A) = k) \gtrsim (1-\rho)k^k\gamma^k/k!, \quad k \geq 0,$$

$$\gamma = (1-\alpha)e^{-1}, \quad \rho = t(\gamma).$$

But $\{(1-\rho)k^k\gamma^k/k!\}_{k=0}^{\infty}$ is a probability distribution. Really, by (16),

$$\textstyle\sum_{k=0}^{\infty} (1-\rho)k^k\gamma^k/k! = (1-t(\gamma))(1+F(\gamma)) = (1-t(\gamma))(1-t(\gamma))^{-1} = 1.$$

Therefore, like two times before, we infer that

$$\lim_{n\to\infty} P(n - \zeta(A) = k) = (1-\rho)k^k\gamma^k/k!, \quad k \geq 0.$$

Acknowledgement. I thank my friend Ilya Gertsbakh for convincingly showing to me how interesting the random mappings are. Our discussions on the subject of this paper were of great help to me. I am also grateful to the editor and a referee for many suggestions concerning the exposition.

REFERENCES

[1] Burtin, Y. D. (1980). On a simple formula for random mappings and its applications. *J. Appl. Probability* **17** 403–414.

[2] Frank, H. and Frisch, I. (1971). *Communication, Transmission and Transportation Networks.* Addison-Wesley, Reading, Mass.

[3] Gertsbakh, I. B. (1977). Epidemic processes on a random graph: some preliminary results. *J. Appl. Probability* **14** 427–438.

[4] Harris, B. (1960). Probability distributions related to random mappings. *Ann. Math. Statist.* **31** 1045–1062.

[5] Katz, L. (1955). Probability of indecomposability of a random mapping function. *Ann. Math. Statist.* **26** 512–517.

[6] Kolchin, V. F. (1976). A problem of the allocation of particles in cells and random mappings. *Theory Probability Appl.* **21** 48–63.

[7] Kruskal, J. B. (1954). The expected number of components under random mapping functions. *Amer. Math. Monthly* **61** 392–397.

[8] Pólya, G. (1937). Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und Chemische Verbindungen. *Acta Mathematica* **68** 145–254.

[9] Proskurin, G. V. (1973). On the distribution of the number of vertices in strata of a random mapping. *Theory Probability Appl.* **18** 803–808.

[10] Rényi, A. (1959). On connected graphs, I. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **4** 73–85.

[11] Riordan, J. (1962). Enumeration of linear graphs for mappings of finite sets. *Ann. Math. Statist.* **33** 178–185.

[12] Riordan, J. (1968). *Combinatorial Identities.* Wiley, New York.

[13] Ross, S. M. (1981). A random graph. *J. Appl. Probability* **18** 309–315.

[14] Rubin, H., and Sitgreaves, R. (1954). Probability distributions related to random transformations of a finite set. Technical report 19 A, Applied Mathematics and Statistics Laboratory, Stanford University.

[15] Sachkov, V. N. (1973). Random mappings with bounded height. *Theory Probability Appl.* **1** 120–130.

[16] Stepanov, V. E. (1969). Limit distributions of certain characteristics of random mappings. *Theory Probability Appl.* **14** 612–626.

[17] Stepanov, V. E. (1971). Random mappings with a single attracting centre. *Theory Probability Appl.* **16** 155–161.

Department of Mathematics
The Ohio State University
Columbus, Ohio 43210