Abel-Cayley-Hurwitz multinomial expansions associated with random mappings, forests, and subsets*

Jim Pitman

Technical Report No. 498

Department of Statistics
University of California
367 Evans Hall # 3860
Berkeley, CA 94720-3860

June 16, 1998

Abstract

Extensions of binomial and multinomial formulae due to Abel, Cayley and Hurwitz are related to the probability distributions of various random subsets, trees, forests, and mappings. For instance, an extension of Hurwitz's binomial formula is associated with the probability distribution of the random set of vertices of a fringe subtree in a random forest whose distribution is defined by terms of a multinomial expansion over rooted labeled forests which generalizes Cayley's expansion over unrooted labeled trees.

Contents

*Research supported in part by N.S.F. Grant DMS97-03961

1 Introduction

This paper offers some developments and interpretations of the binomial and multinomial expansions due to Abel [1], Cayley [14] and Hurwitz [27], by consideration of variously constructed random subsets, trees, forests, and mappings. Study of the various probability distributions involved is also motivated by applications treated elsewhere [13, 22, 42, 44, 43, 45]. A central result of this paper, proved in Section 5, is the following generalization of Cayley's multinomial expansion over trees. See Section 3 for background.

Theorem 1 *For a non-empty subset R of a finite set S let $\mathbf{F}(S,R)$ be the set of all forests of rooted trees labeled by S, with edges directed away from the roots, whose set of root vertices is R. Then there is the following identity of polynomials in variables $x_s, s \in S$:*

$$\sum_{\mathbf{f}\in\mathbf{F}(S,R)} \prod_{s\in S} x_s^{C_s \mathbf{f}} = \left(\sum_{r\in R} x_r\right)\left(\sum_{s\in S} x_s\right)^{|S|-|R|-1} \qquad (1)$$

where $C_s\mathbf{f}$ *is the number of children (out-degree) of* s *in the forest* $\mathbf{f}$*, and* $|A|$ *is the number of elements of a set* A*.*

Take $x_s \equiv 1$ in (1) to recover Cayley's [14] well known formula $|\mathbf{F}(S,R)| = |R|\,|S|^{|S|-|R|-1}$. For $1 \leq k \leq |S| \Leftrightarrow 1$ Let $\mathbf{F}_k(S)$ be the set of all rooted forests of k trees labeled by S. Summing (1) over all subsets R of S of size k yields the cruder identity [42, 45, 54]

$$\sum_{\mathbf{f}\in\mathbf{F}_k(S)} \prod_{s\in S} x_s^{C_s\mathbf{f}} = \binom{|S| \Leftrightarrow 1}{k \Leftrightarrow 1} \left(\sum_{s\in S} x_s \right)^{|S|-k}. \qquad (2)$$

For variables $z_s, s \in S$ and a subset A of S, let $z_A := \sum_{s\in A} z_s$. Let $[n] := \{1, \ldots, n\}$. Hurwitz [27] studied sums of the form

$$H_n^{\gamma,\delta} := H_n^{\gamma,\delta}(x, y; z_s, s \in [n]) := \sum_{A\subseteq[n]} (x + z_A)^{|A|+\gamma}(y + z_{\bar{A}})^{|\bar{A}|+\delta} \qquad (3)$$

for integers γ and δ, where the sum is over all 2^n subsets A of $[n]$, and $\bar{A} := [n] \Leftrightarrow A$. Hurwitz used recurrences to obtain the identities

$$xH_n^{-1,0} = yH_n^{0,-1} = (x + y + z_{[n]})^n, \qquad (4)$$

$$xyH_n^{-1,-1} = (x + y)(x + y + z_{[n]})^{n-1} \qquad (5)$$

which follows easily from (4), and

$$H_n^{0,0} = \sum_{A\subseteq[n]} |A|!(\textstyle\prod_{s\in A} z_s)(x + y + z_{[n]})^{|\bar{A}|}. \qquad (6)$$

As noted by Hurwitz, for $z_s \equiv 1$ these formulae yield evaluations of corresponding *Abel sums* [1]

$$A_n^{\gamma,\delta}(x,y) := \sum_{a=0}^{n} \binom{n}{a} (x + a)^{a+\gamma}(y + n \Leftrightarrow a)^{n-a+\delta}. \qquad (7)$$

For various combinatorial interpretations of these identities and related formulae see [30, 34, 25, 11, 50, 52, 56, 57]. Section 2 presents probabilistic interpretations of the Hurwitz identities (4)-(5)-(6). These interpretations lead to a number of new identities involving other homogeneous polynomials in $2 + n$ commuting variables $x, y, z_1, \ldots, z_n$ defined by sums of products indexed by subsets of $[n]$. Following is a selection of several such identities, with references to their explanations in following sections. These explanations involve the probability distribution of a suitable random subset of $[n]$ derived from a

random forest labeled by a superset of $[n]$ with one or two extra elements. The number of trees in the forest is $k = m+1$ in (8), (9) and (10), and $k = 1$ in (11) and (12). Effectively, these identities are deduced by repeated applications of Theorem 1.
(Theorem 36) For $0 \leq m \leq n$:

$$\sum_{A \subseteq [n]} \binom{|\bar{A}|}{m} x(x+z_A)^{|A|-1}(y+z_{\bar{A}})^{|\bar{A}|-m} = \binom{n}{m}(x+y+z_{[n]})^{n-m}, \tag{8}$$

which reduces to (4) for $m = 0$.
(Corollary 27) For $1 \leq m \leq n$:

$$\sum_{\substack{A \subseteq [n] \\ |A| \leq n-m}} \binom{|\bar{A}| \Leftrightarrow 1}{m \Leftrightarrow 1}(x+z_A)^{|A|}(z_{\bar{A}})^{|\bar{A}|-m} = \binom{n}{m}(x+z_{[n]})^{n-m}. \tag{9}$$

(Corollary 6) For $0 \leq m \leq n$:

$$\sum_{A \subseteq [n]} \binom{|\bar{A}|}{m}(x+z_A)^{|A|}(y+z_{\bar{A}})^{|\bar{A}|-m} = \sum_{A \subseteq [n]} \binom{|\bar{A}|}{m} |A|! \left(\textstyle\prod_{s \in A} z_s\right)(x+y+z_{[n]})^{|\bar{A}|-m} \tag{10}$$

which reduces to (6) for $m = 0$.
(Theorem 33)

$$\sum_{A \subseteq [n]} |A|! \left(\textstyle\prod_{s \in A} z_s\right)(x+z_A)(x+z_{[n]})^{|\bar{A}|-1} = (x+z_{[n]})^n. \tag{11}$$

(Theorem 23)

$$\sum_{A \subseteq [n]} z_A \, x^{|A|-1} z_{[n]}^{n-|A|-1} = (x+z_{[n]})^{n-1}. \tag{12}$$

For $z_s \equiv 1$ these Hurwitz type identities reduce to corresponding identities for Abel sums. For instance, the Abel type identity derived from (9) is

$$\sum_{a=0}^{n-m} \binom{n}{a}\binom{n \Leftrightarrow a \Leftrightarrow 1}{m \Leftrightarrow 1}(x+a)^a(n \Leftrightarrow a)^{n-a-m} = \binom{n}{m}(x+n)^{n-m} \tag{13}$$

for $1 \leq m \leq n$. The Abel type identity derived from (11) is the case $b = 0$ of the telescoping sum

$$\sum_{a=b}^{n}(n)_a(x+a)(x+n)^{n-a-1} = (n)_b(x+n)^{n-b} \quad (0 \leq b \leq n) \tag{14}$$

where $(n)_a := \prod_{i=1}^{a}(n \Leftrightarrow i+1)$, while that derived from (12) reduces easily to the elementary binomial formula.

2 Probabilistic Interpretations

This section presents probabilistic interpretations of the basic expansions, along with a number of results which are proved in later sections. First, a brief review of probabilistic terms as used in this paper. A *probability distribution* on a finite set S is a non-negative real-valued function $p := (p_s, s \in S)$ with $\sum_{s \in S} p_s = 1$. The definition of p is extended to subsets A of S by $p(A) := p_A := \sum_{i \in A} p_s$. Throughout the paper, P denotes a probability distribution on a suitable finite set Ω. A function $X : \Omega \to S$ is called a *random element of* S. The *distribution of* X, denoted $\mathrm{dist}(X)$, is the probability distribution p on S defined by

$$p_s := P(X = s) := P(\{\omega \in \Omega : X(\omega) = s\}) \qquad (s \in S).$$

If elements of S are for instance subsets of another set, or trees, or mappings, a random element X of S may called a *random set*, a *random tree*, or a *random mapping*, as the case may be, whether or not the distribution of X is *uniform*, meaning $P(X = s) = 1/|S|$ for all $s \in S$. Subsets of Ω are called *events*. For an event $B \subseteq \Omega$ with $P(B) > 0$ and a random element X of S, the *conditional distribution of* X *given* B, denoted $\mathrm{dist}(X \,|\, B)$, is the probability distribution p on S defined by

$$p_s := P(X = s \,|\, B) := P(\{\omega \in B : X(\omega) = s\})/P(B) \qquad (s \in S).$$

For further background, and definitions of other probabilistic terms such as independence and expectation, see [23].

Definition 2 Let p be a probability distribution on the interval of integers $[0, n+1] := \{0, 1, \ldots, n, n+1\}$. Say that a random subset V of $[n]$ has the *Hurwitz distribution of index* (γ, δ) *with parameters* $p_0, p_1, \ldots, p_{n+1}$, denoted $H_n^{\gamma,\delta}(p)$, if $P(V = A)$ is proportional to the Ath term of the Hurwitz sum $H_n^{\gamma,\delta}(p_0, p_{n+1}; p_s, s \in [n])$ defined by (3) as A ranges over $2^{[n]}$. Call the distribution of $|V|$ on $[0, n]$ induced by such a random subset V of $[n]$ a *Hurwitz-binomial distribution*, or $H_n^{\gamma,\delta}(p)$*-binomial distribution* to indicate the parameters.

According to (4), a random set V has $H_n^{-1,0}(p)$ distribution if

$$P(V = A) = p_0(p_0 + p_A)^{|A|-1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|} \qquad (A \subseteq [n]). \tag{15}$$

Similarly from (5), V has the $H_n^{-1,-1}(p)$ distribution if

$$P(V = A) = \frac{p_0 p_{n+1}}{(p_0 + p_{n+1})}(p_0 + p_A)^{|A|-1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|-1} \qquad (A \subseteq [n]). \tag{16}$$

These and other similar formulae should be interpreted by continuity in degenerate cases such the case $p_0 = 0$ in (15). So (15) for $p_0 = 0$ means $P(V = \emptyset) = 1$ and $P(V = A) = 0$ for $A \neq \emptyset$. Take $z_s = p_s$ in (4) to see that the existence for each probability distribution p on $[0, n+1]$ of a random subset V of $[n]$ with distribution given by (15) is equivalent to Hurwitz's evaluation (4) of $H^{-1,0}$. A similar remark applies to (16) and Hurwitz's evaluation (5) of $H_n^{-1,-1}$. In the *Abel case*

$$p_1 = x/\Sigma; \; p_{n+1} = y/\Sigma; \; p_i = 1/\Sigma \text{ for } i \in [n] \tag{17}$$

where $\Sigma := x + y + n$ for arbitrary $x, y \geq 0$, the $H_n^{\gamma,\delta}(p)$-binomial distribution on $[0, n]$ is obtained by normalization of the terms of the coresponding Abel sum $A^{\gamma,\delta}(x, y)$ defined by (7). Call this special case of the Hurwitz-binomial distribution an *Abel-binomial distribution* or $A^{\gamma,\delta}(x, y)$-binomial distribution to indicate the parameters. The Abel-binomial distributions $A^{-1,-1}(x, y)$ and $A^{-1,0}(x, y)$ are known in the statistical literature as *quasi-binomial distributions* [20, 19, 18, 15].

The following theorem, proved in Section 4, presents three different constructions of a random set V with the $H_n^{-1,0}(p)$ distribution. The first construction is a probabilistic translation of an identity of enumerator polynomials used by Françon[25] to derive Hurwitz's evaluation of $H^{-1,0}$, while the second can be read from results of Jaworski [31]. Corollaries 14 and 29 give similar constructions of V with the $H_n^{-1,-1}(p)$ distribution. See also Berg and Mutafchiev [8] for a closely related appearance of Abel-binomial distributions in connection with random mappings.

For a mapping M from S to S and $v \in S$ define the set of *predecessors of v induced by M* by

$$\text{pred}(v, M) := \{s \in S : M_s^i = v \text{ for some } i \geq 1\} \tag{18}$$

where $s \mapsto M_s^i$ is the ith iterate of M.

Theorem 3 *If M is a random mapping defined by independent random variables $M_s, s \in S$ with common distribution p on $S := [0, n+1]$, then each of the following random subsets of $[n]$ has the Hurwitz distribution $H^{-1,0}(p)$ on $2^{[n]}$:*
(i) (Françon [25]) *assuming $p_0 p_{n+1} > 0$, the random set* $\text{pred}(0, M)$ *conditionally given that both 0 and $n+1$ are fixed points of M;*
(ii) (Jaworski [31]) *assuming $p_{n+1} = 0$, the random set $[n] \cap \text{pred}(0, M)$;*
(iii) *assuming $p_{n+1} > 0$, the random set* $\text{pred}(0, M)$ *conditionally given that $n+1$ is the unique cyclic point of M.*

Let $\mathcal{D}(M)$ be the usual functional digraph associated with M, with vertex set S and a directed edge (s, M_s) for each $s \in S$. See [39, 24, 25, 41, 35, 2] for background. The set of all cyclic points of M is

$$\text{cyclic}(M) := \{s \in S : s \in \text{pred}(s, M)\}.$$

Call (s, M_s) a *cyclic edge* of $\mathcal{D}(M)$ if $s \in \text{cyclic}(M)$. Let $\mathcal{F}(M)$ be the digraph with vertex set S derived from $\mathcal{D}(M)$ as follows: first delete all cyclic edges, then reverse the direction of all remaining edges. So $\mathcal{F}(M)$ is a forest of rooted trees labeled by S, with edges directed away from cyclic(M), the set of all root vertices of $(\mathcal{F}(M))$. Call $\mathcal{F}(M)$ the *forest derived from M*. Each connected component C of $\mathcal{D}(M)$ contains a unique cycle C_0, and C decomposes further into a collection of tree components of $\mathcal{F}(M)$ whose set of roots is C_0. In case (i) of the previous theorem, the conditioning forces $\text{pred}(0, M) \cup \{0\}$ to be a connected component of $\mathcal{D}(M)$ which is a single tree component of$\mathcal{F}(M)$ rooted at 0, while $n + 1$ is forced to be the unique cyclic point of another component of $\mathcal{D}(M)$. In case (ii) the set $\text{pred}(0, M) \cup \{0\}$ may be either the union of a tree component of $\mathcal{F}(M)$ and a cycle of arbitrary size, or just a subtree of a tree component, according to whether or not $0 \in \text{cyclic}(M)$. In case (iii) the conditioning forces $\mathcal{F}(M)$ to be a tree rooted at $n + 1$, and $\text{pred}(0, M) \cup \{0\}$ is a fringe subtree of this tree. While the three results are obtained by a similar arguments, it does not seem easy to deduce any one from another. Neither does there appear to be any natural generalization from which all three results could be deduced.

As observed in [42], formula (2) amounts to the fact that for each probability distribution p on S, the formula

$$P(\mathcal{F}_k = \mathbf{f}) = \binom{|S| \Leftrightarrow 1}{k \Leftrightarrow 1}^{-1} \prod_{s \in S} p_s^{C_s \mathbf{f}} \qquad (\mathbf{f} \in \mathbf{F}_k(S)) \tag{19}$$

defines the probability distribution of a random rooted forest $\mathcal{F}_k$ of k trees labeled by S.

Definition 4 For a probability distribution p on S, call a random forest $\mathcal{F}_k$ of k rooted trees labeled by S a *p-forest*, and for $k = 1$ a *p-tree*, if the distribution of $\mathcal{F}_k$ is given by (19).

If p is uniform on S, the distribution of a p-forest of k trees is uniform on $\mathbf{F}_k(S)$. Several natural constructions of a p-tree for general p are reviewed in Section 3. As shown in [42], a p-forest of k trees is obtained by deleting $k \Leftrightarrow 1$ edges picked uniformly at random from the $|S| \Leftrightarrow 1$ edges of a p-tree. The following proposition records a characterization of the distribution of a p-forest which follows immediately from the definition.

Proposition 5 *A random forest $\mathcal{F}_k$ of k rooted trees labeled by S is a p-forest if and only if both*
(i) *the distribution of the out-degree count vector $C\mathcal{F}_k := (C_s\mathcal{F}_k, s \in S)$ is multinomial with parameters $n \Leftrightarrow k$ and $p = (p_s, s \in S)$, and*
(ii) *for each vector of counts $c \in \{0, 1, 2, \ldots\}^S$ with $\sum_s c_s = n \Leftrightarrow k$, the conditional distribution of $\mathcal{F}_k$ given $C\mathcal{F}_k = c$ is uniform over the set $\mathbf{F}_k(S; c)$ of all forests with the given out-degree counts c, as enumerated by*

$$|\mathbf{F}_k(S; c)| = \frac{(n \Leftrightarrow 1)!}{(k \Leftrightarrow 1)! \prod_{s \in S} c_s!}. \tag{20}$$

For any random forest $\mathcal{F}_k$ of k rooted trees labeled by S, the vector of out-degree counts $(C_s\mathcal{F}_k, s \in S)$ is subject to the constraint $\sum_s C_s\mathcal{F}_k = n \Leftrightarrow k$. Therefore, the expectation of $C_s\mathcal{F}_k$ equals $(n \Leftrightarrow k)p_s$ for some probability distribution p on S. According to the previous proposition, for any given p this is achieved by a p-forest. Section 5 presents a number of enumerations of rooted random forests which arise naturally from the study of p-forests. See also [44].

For distinct vertices u and v of a forest $\mathbf{f}$, a *directed path from u to v in $\mathbf{f}$* is a sequence of edges of $\mathbf{f}$ of the form $(u, s_1), (s_1, s_2), \ldots (s_{m-1}, v)$ for some $m \geq 1$. Write

$$u \overset{\mathbf{f}}{\rightsquigarrow} v \text{ if there is a directed path from } u \text{ to } v \text{ in } \mathbf{f} \text{ and } u \overset{\mathbf{f}}{\not\rightsquigarrow} v \text{ otherwise.} \tag{21}$$

Let $\mathrm{root}(\mathbf{t}_v)$ be the root of the unique tree component $\mathbf{t}_v$ of $\mathbf{f}$ that contains v. By the convention that edges of $\mathbf{f}$ are directed away from the roots, $u \overset{\mathbf{f}}{\rightsquigarrow} v$ if and only if u lies on the unique path from $\mathrm{root}(\mathbf{t}_v)$ to v along edges of $\mathbf{t}_v$. Write $u \overset{\mathbf{f}}{\sim} v$ if there is a path from u to v in the undirected graph obtained by ignoring edge directions of $\mathbf{f}$, that is if $\mathbf{t}_u = \mathbf{t}_v$. Section 6 treats the problem of finding expressions for the *percolation probability* $P(s \overset{\mathcal{F}_k}{\sim} v)$ and the *oriented percolation probability* $P(s \overset{\mathcal{F}_k}{\rightsquigarrow} v)$ for two vertices s and v of a p-forest $\mathcal{F}_k$. See [12, 46] for closely related studies of such percolation probabilities for the digraph of a random mapping, and [26] for a study of such problems for other models of random forests, and applications to reliability of networks. By a relabeling of vertices, the problem of finding $P(s \overset{\mathcal{F}_k}{\sim} v)$ for two arbitrary vertices s and v of a p-forest $\mathcal{F}_k$ is reduced to the case when $S = [0, n+1]$, $s = 0$ and $v = n+1$, as supposed in the following straightforward consequence of Theorem 1:

Corollary 6 *Let $\mathcal{F}_k$ be a p-forest of k trees labeled by $[0, n+1]$. Then*

$$P(0 \overset{\mathcal{F}_k}{\rightsquigarrow} n+1) = \sum_{A \subseteq [n]} \frac{(|\bar{A}|)_{k-1}}{(n+1)_{k-1}} p_0 (p_0 + p_A)^{|A|} (p_{n+1} + p_{\bar{A}})^{|\bar{A}| - (k-1)} \tag{22}$$

where the A*th term equals* $P(0 \overset{\mathcal{F}}{\rightsquigarrow} n+1, \mathcal{V}_k = A)$ *for* $\mathcal{V}_k$ *the random set of all* $v \in [n]$ *such that there exists a directed path from* 0 *to* v *in* $\mathcal{F}_k$ *that does not pass via* $n+1$*. Also*

$$P(0 \overset{\mathcal{F}_k}{\rightsquigarrow} n+1) = \sum_{A \subseteq [n]} \frac{(|\bar{A}|)_{k-1}}{(n+1)_{k-1}} |A|!\, p_0 \textstyle\prod_{s \in A} p_s \tag{23}$$

where the A*th term equals* $P(0 \overset{\mathcal{F}}{\rightsquigarrow} n+1, \mathcal{L}_k = A)$ *for* $\mathcal{L}_k$ *the random set of all* $v \in [n]$ *such that* v *lies on the path which joins* 0 *to the root of its tree component in* $\mathcal{F}_k$*.*

For $k = 1$, Corollary 6 yields Hurwitz's expression (6) for $H^{0,0}$, along with the following probabilistic interpretation: for $\mathcal{T}_n$ a p-tree labeled by $[0, n+1]$

$$P(0 \overset{\mathcal{T}_n}{\rightsquigarrow} n+1) = p_0 H^{0,0}(p_0, p_{n+1}; p_j, j \in [n]). \tag{24}$$

In the Abel case (17) this specializes to give the following probabilistic interpretation of the Abel sum $A_n^{0,0}(x,y)$ as in (7), with an asymptotic expression obtained by a straightfoward integral approximation using the local normal approximation to the binomial distribution [23]:

Corollary 7 *For* $\mathcal{T}_n$ *a* p*-tree with* p *the distribution on* $[0, n+1]$ *defined by* (17)*, the probability that there is a directed path from* 0 *to* $n+1$ *in* $\mathcal{T}_n$ *is*

$$P(0 \overset{\mathcal{T}_n}{\rightsquigarrow} n+1) = \frac{x A_n^{0,0}(x,y)}{(x+y+n)^{n+1}} \sim \sqrt{\frac{\pi}{2}} \frac{x}{\sqrt{n}} \text{ as } n \to \infty. \tag{25}$$

Consequently, for each pair of distinct vertices r *and* s *in a set* S *of* $n+2$ *elements, the number of rooted trees* $\mathbf{t}$ *labeled by* S *such that* r *lies on the path in* $\mathbf{t}$ *from* $\mathrm{root}(\mathbf{t})$ *to* s *is the Abel sum* $A_n^{0,0}(1,1)$*, which is asymptotically equivalent to* $\sqrt{\pi/2}e^2 n^{n+1/2}$ *as* $n \to \infty$*.*

It does not appear that there is any simpler expression for the oriented percolation probability $P(0 \overset{\mathcal{F}_k}{\rightsquigarrow} n+1)$ than the Hurwitz sums provided provided by Theorem 6. Similar Hurwitz sums are obtained in Section 6 for the percolation probability $P(0 \overset{\mathcal{F}_k}{\sim} n+1)$. There are however some closely related probabilities where some remarkable simplifications occur, as indicated in the next theorem, which is proved in Section 5.1. As a preliminary to the theorem, there is the following easy consequence of Theorem 1:

Corollary 8 *Let* $\mathcal{R}_k$ *be the random set of* k *root vertices of a* p*-forest* $\mathcal{F}_k$ *of* k *trees labeled by* S *with* $|S| = n$*. Then*

(i) *for each subset R of S with $|R| = k$*

$$P(\mathcal{R}_k = R) = \binom{n \Leftrightarrow 1}{k \Leftrightarrow 1}^{-1} p_R; \tag{26}$$

(ii) *for each $r \in S$*

$$P(r \in \mathcal{R}_k) = \frac{k \Leftrightarrow 1 + (n \Leftrightarrow k)p_r}{n \Leftrightarrow 1}. \tag{27}$$

Theorem 9 *For $\mathcal{F}_k$ a p-forest of k trees with roots $\mathcal{R}_k$ as in the previous corollary,*
(i) *for all distinct $r, s \in S$ with $P(r \in \mathcal{R}_k) > 0$*

$$P(r \overset{\mathcal{F}_k}{\rightsquigarrow} s \,|\, r \in \mathcal{R}_k) = \frac{(n \Leftrightarrow k)p_r}{k \Leftrightarrow 1 + (n \Leftrightarrow k)p_r}; \tag{28}$$

(ii) *for all distinct $r, s \in S$*

$$P(r \overset{\mathcal{F}_k}{\not\rightsquigarrow} s \text{ and } r \in \mathcal{R}_k) = \frac{k \Leftrightarrow 1}{n \Leftrightarrow 1}. \tag{29}$$

The fact that the probability in (29) does not depend on p is quite surprising. In the terminology of statistical theory [36], for each choice of r and s the indicator of the event in (29) is an *ancillary statistic*. Here k is regarded as fixed and known, and the family of distributions of p-forests on $\mathbf{F}_k(S)$ is regarded as a statistical family parameterized by the underlying probability distribution p on S. Proposition 5 implies that the vector of counts $C\mathcal{F}_k$ is what is known [36] as a *complete sufficient statistic* in the statistical problem of estimating p given an observation of $\mathcal{F}_k$. According to Basu's theorem [36, Thm. 1.5.5] if T is a complete sufficient statistic for a statistical problem $(P_\theta, \theta \in \Theta)$, then every ancillary statistic A is independent of T under P_θ for all $\theta \in \Theta$. Thus formula (29) has the following consequence:

Corollary 10 *Let $1 \leq k \leq n \Leftrightarrow 1$. For S with $|S| = n$, for each choice of non-negative integers $c_v, v \in S$ with $\sum_{v \in S} c_v = n \Leftrightarrow k$, and each choice of two vertices r and s of S, among all forests $\mathbf{f}$ of k trees labeled by S such that v has c_v children in $\mathbf{f}$ for every $v \in S$, the fraction of $\mathbf{f}$ such that both $r \in \mathrm{roots}(\mathbf{f})$ and $r \overset{\mathbf{f}}{\not\rightsquigarrow} s$ equals $(k \Leftrightarrow 1)/(n \Leftrightarrow 1)$.*

This enumeration was actually discovered by the above line of reasoning, which makes an unusual application of ideas of mathematical statistics to enumerative combinatorics. But such a simple result invites a direct combinatorial proof, which is provided is at the end of Section 5.2.

3 Cayley's multinomial expansion

Let $\mathbf{U}(S)$ be the set of unrooted trees labeled by the finite set S. For $i \in S$, $\mathbf{u} \in \mathbf{U}(S)$ let $D_i\mathbf{u}$ be the *degree* of i in $\mathbf{u}$, that is $D_i(\mathbf{u}) := |\{j : i \overset{\mathbf{u}}{\longleftrightarrow} j\}|$ where $\overset{\mathbf{u}}{\longleftrightarrow}$ is the undirected edge relation of $\mathbf{t}$. According to *Cayley's multinomial expansion over unrooted trees* [14, 48]

$$\sum_{\mathbf{u}\in\mathbf{U}(S)} \prod_{s\in S} x_s^{D_s\mathbf{u}-1} = \left(\sum_{s\in S} x_i\right)^{|S|-2}. \tag{30}$$

For $x_i \equiv 1$ this reduces to Cayley's formula $|\mathbf{U}(S)| = |S|^{|S|-2}$. Let $\mathbf{T}(S)$ be the set of all rooted trees labeled by S. Let edges of $\mathbf{t} \in \mathbf{T}(S)$ be directed away from the root of $\mathbf{t}$, denoted $\mathrm{root}(\mathbf{t})$. For $s \in S$, $\mathbf{t} \in \mathbf{T}(S)$ let $C_s\mathbf{t}$ be the *number of children* or *out-degree* of s in $\mathbf{t}$, that is $C_s(\mathbf{t}) := |\{v : s \overset{\mathbf{t}}{\to} v\}|$ where $\overset{\mathbf{t}}{\to}$ is the directed edge relation of $\mathbf{t}$. For $r \in S$ let $\mathbf{T}(S,r)$ be the set of all trees $\mathbf{t} \in \mathbf{T}(S)$ with $\mathrm{root}(\mathbf{t}) = r$. Fix $r \in S$. Multiply both sides of (30) by x_r and use the obvious bijection between $\mathbf{U}(S)$ and $\mathbf{T}(S,r)$ to see that (30) can be rewritten

$$\sum_{\mathbf{t}\in\mathbf{T}(S,r)} \prod_{s\in S} x_s^{C_s\mathbf{t}} = x_r \left(\sum_{s\in S} x_s\right)^{|S|-2} \qquad (r \in S). \tag{31}$$

This is the special case $|R| = 1$ of (1). Sum (31) over all $r \in S$ to obtain the following variant of Cayley's expansion, which is the case $k = 1$ of (2):

$$\sum_{\mathbf{t}\in\mathbf{T}(S)} \prod_{s\in S} x_s^{C_s\mathbf{t}} = \left(\sum_{s\in S} x_s\right)^{|S|-1}. \tag{32}$$

Take $x_s \equiv 1$ to see $|\mathbf{T}(S)| = |S|^{|S|-1}$. Now let p be a probability distribution on S. According to Definition 4, a rooted random tree $\mathcal{T}$ labeled by S is a *p-tree* if

$$P(\mathcal{T} = \mathbf{t}) = \prod_{s\in S} p_s^{C_s\mathbf{t}} \qquad (\mathbf{t} \in \mathbf{T}(S)). \tag{33}$$

Call an unrooted random tree $\mathcal{U}$ labeled by S an *unrooted p-tree* if

$$P(\mathcal{U} = \mathbf{u}) = \prod_{s\in S} p_s^{D_s\mathbf{u}-1} \qquad (\mathbf{u} \in \mathbf{U}(S)). \tag{34}$$

The following lemma summarizes the previous discussion of (30),(31) and (32) in probabilistic terms.

Lemma 11 *Let $(r, \mathbf{u})$ denote the rooted tree obtained by assigning root $r \in S$ to an unrooted tree $\mathbf{u}$ labeled by S. A random tree $\mathcal{T}$ is a p-tree if and only if $\mathcal{T} = (R, \mathcal{U})$ where $\mathcal{U}$ is an unrooted p-tree, the root R of $\mathcal{T}$ has distribution p, and R and $\mathcal{U}$ are independent.*

In the rest of this paper, all trees and forests are assumed to be rooted unless otherwise specified. The next lemma reviews some constructions of a p-tree. Note that once formula (33) has been established for a generic p by any of these constructions, it follows that the sum of the right side of (33) over all $\mathbf{t} \in \mathbf{T}(S)$ equals 1. The various forms (32), (30) and (31) of Cayley's multinomial expansion then follow easily, in that order. Constructions (i) and (ii) yield (33) quite easily by the results cited. Constructions (iii) and (iv) yield (33) up to a constant of proportionality, which must equal 1 by comparison with any of the other constructions. A nicer proof for Construction (iv) is indicated in the next section.

Lemma 12 *Let $X_0, X_1, X_2, \ldots$ be a sequence of independent random variables with common distribution p on S with $|S| = n$.*
(i) *Let $T : S^{n-1} \to \mathbf{T}(S)$ be the bijection defined by the Prüfer code* [47, 17] *such that $T(s_1, \ldots, s_{n-1}) = \mathbf{t}$ with $C_s\mathbf{t}$ equal to the number of j such that $s_j = s$, for every $s \in S$. Then $\mathcal{T} := T(X_1, \ldots, X_{n-1})$ is a p-tree.*
(ii) [42] *Define a coalescing sequence of forests $\mathcal{F}(0), \mathcal{F}(2), \ldots, \mathcal{F}(n-1)$ as follows, by adding edges one by one in such a way that $\mathcal{F}(j)$ has j edges (and hence $n-j$ tree components) for each $1 \leq j \leq n-1$. Let $\mathcal{F}(0)$ be the trivial forest labeled by S with no edges. Given that $\mathcal{F}(0), \ldots, \mathcal{F}(j-1)$ have been defined for some $1 \leq j \leq n-1$, define $\mathcal{F}(j)$ by adding the edge (X_j, Y_j) to $\mathcal{F}(j-1)$, where given X_j and the (X_i, Y_i) for $1 \leq i < j$, the random variable Y_j has uniform distribution on the set of $n-j$ roots of tree components of $\mathcal{F}(j-1)$ other than the component containing X_j. Then $\mathcal{F}(j)$ is a p-forest of $n-j$ trees for every $0 \leq j \leq n-1$. In particular, $\mathcal{F}(n-1)$ is a p-tree.*
(iii) [10, Theorem 1],[37, §6.1] *Assuming $p_s > 0$ for every $s \in S$, let*

$$\mathcal{T} := \{(X_{j-1}, X_j) : j \geq 1, X_j \notin \{X_0, \ldots, X_{j-1}\}\}.$$

Then $\mathcal{T}$ is a p-tree.
(iv) *Let $\mathcal{F}(M)$ be the forest derived from a random mapping M from S to S with independent images M_s distributed according to p. Then for each $r \in S$, $\mathcal{F}(M)$ conditioned to be a single tree rooted at r has the same distribution as a p-tree $\mathcal{T}$ given* $\mathrm{root}(\mathcal{T}) = r$. *Hence, the undirected digraph derived from $\mathcal{F}(M)$, given that $\mathcal{F}(M)$ is a tree, is an unrooted p-tree.*

For a tree $\mathbf{t}$ labeled by S let $V_s(\mathbf{t}) := \{v \in S : s \overset{\mathbf{t}}{\leadsto} v\}$. So $V_s(\mathbf{t})$ is the set of non-root vertices of the *fringe subtree of* $\mathbf{t}$ *rooted at* i, that is the tree $\mathbf{t}_s$ labeled by $\{s\} \cup V_s(\mathbf{t})$ whose edge relation is the restriction to $\{s\} \cup V_s(\mathbf{t})$ of the edge relation of $\mathbf{t}$. See [3] for background and further references to fringe subtrees. If $\mathbf{t}$ is a tree component of the forest $\mathcal{F}(M)$ derived from a mapping M, so $\mathbf{t}$ is rooted at some vertex $r \in \text{cyclic}(M)$, then for each non-root vertex s of $\mathbf{t}$ the set $\text{pred}(s, M)$ of predecessors of s induced by M is identical to $V_s(\mathbf{t})$. In view of Lemma 12(iv), case (iii) of Theorem 3 can be reformulated as follows in terms of trees instead of mappings:

Theorem 13 *Let p be a probability distribution on the set $S := [0, n+1]$ for $n \geq 1$, with $p_{n+1} > 0$. Let $\mathcal{T}$ be a p-tree with root R, and let $V_0(\mathcal{T}) := \{v \in S : 0 \overset{\mathcal{T}}{\leadsto} v\}$ be the set of non-root vertices of the fringe subtree of $\mathcal{T}$ rooted at 0. Given the event $(R = n+1)$ the random set $V_0(\mathcal{T})$ has the Hurwitz distribution $H_n^{-1,0}(p)$ on $2^{[n]}$. That is, for all $A \subseteq [n]$*

$$P(R = n+1, V_0(\mathcal{T}) = A) = p_0(p_0 + p_A)^{|A|-1} p_{n+1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|} \tag{35}$$

where $\bar{A} := [n] \Leftrightarrow A$, and the sum of these probabilities over all subsets A of $[n]$ is

$$P(R = n+1) = p_{n+1}.$$

Proof. Fix an arbitrary subset A of $[n]$. The probability $P(R = n+1, V_0(\mathcal{T}) = A))$ is the sum of $P(\mathcal{T} = \mathbf{t})$ over $\mathbf{t} \in \mathbf{T}^* := \{\mathbf{t} \in \mathbf{T}(S, n+1) : V_0(\mathbf{t}) = A\}$. For $\mathbf{t} \in \mathbf{T}^*$ let $\mathbf{v}$ be the restriction of $\mathbf{t}$ to $A_0 := A \cup \{0\}$ and let $\mathbf{w}$ be the restriction of $\mathbf{t}$ to $A_0^c := S \Leftrightarrow A_0$. Regard $\mathbf{t}$, $\mathbf{v}$ and $\mathbf{w}$ as subsets of S^2. Then $\mathbf{t} = \mathbf{v} \cup \mathbf{w} \cup \{(s, 0)\}$ for some $s \in A_0^c$. Thus there is a bijection between $\mathbf{T}^*$ and $\mathbf{T}(A_0, 0) \times \mathbf{T}(A_0^c, n+1) \times A_0^c$. For $\mathbf{t} \in \mathbf{T}^*$ the probability $P(\mathcal{T} = \mathbf{t})$ can be written in terms of the corresponding $(\mathbf{v}, \mathbf{w}, s)$ as

$$P(\mathcal{T} = \mathbf{t}) := \prod_{i=0}^{n+1} p_i^{C_i \mathbf{t}} = \left(\prod_{i \in A_0} p_i^{C_i \mathbf{v}}\right) \left(\prod_{\ell \in A_0^c} p_\ell^{C_\ell \mathbf{w}}\right) p_s$$

So $P(R = n+1, V_0 = A)$ is the sum of this product over all

$$(\mathbf{v}, \mathbf{w}, s) \in \mathbf{T}(A_0, 0) \times \mathbf{T}(A_0^c, n+1) \times A_0^c.$$

This sum of products factors into a product of three sums, the first two of which can be evaluated using (31), and the third of which is the sum of p_s over A_0^c, that is $p_{n+1} + p_{\bar{A}}$. This yields (35). The evaluation of the sum over all subsets follows from the result of Lemma 11 that R has distribution p. $\square$

The above theorem yields Hurwitz's evaluation (4) of $H_n^{-1,0}$. A corresponding intepretation and evaluation $H_n^{-1,-1}$ is given by the following corollary:

Corollary 14 *For $\mathcal{T}$ a p-tree labeled by $[0, n+1]$ with root R let V_0^* be the random set of all $j \in [n]$ such that there is a directed path from 0 to j in $\mathcal{T}$ that does not pass via $n+1$. Given the event $(R \in \{0, n+1\})$ the random set V_0^* has the Hurwitz distribution $H_n^{-1,-1}(p)$ on $2^{[n]}$. That is, for all $A \subseteq [n]$*

$$P(R \in \{0, n+1\}, V_0^* = A) = p_0(p_0 + p_A)^{|A|-1} p_{n+1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|-1} \tag{36}$$

where $\bar{A} := [n] \Leftrightarrow A$, and the sum of these probabilities over all $A \subseteq [n]$ is

$$P(R \in \{0, n+1\}) = p_0 + p_{n+1}.$$

Proof. The event $(R = n+1, V_0^* = A)$ is identical to the event $(R = n+1, V_0 = A)$, whose probability is given by (35). If $R = 0$ then $V_0^* = V_0(\mathcal{T}) \Leftrightarrow V_{n+1}(\mathcal{T}) \Leftrightarrow \{n+1\}$. It follows that $P(R = 0, V_0^* = A)$ can be evaluated from formula (35) by switching p_0 and p_{n+1} and switching A and $\bar{A}$. That is

$$P(R = 0, V_0^* = A) = p_0(p_0 + p_A)^{|A|} p_{n+1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|-1}. \tag{37}$$

Add (35) and (37) and eliminate the factor of $(p_0 + p_A + p_{\bar{A}} + p_{n+1}) = 1$ to obtain (36). □

4 Random Mappings

Let S and T be two finite sets. Let $M := (M_t, t \in T)$ be a collection of S-valued random variables defined on Ω. Then $M : \Omega \to S^T$, so M may be regarded as random element of S^T. To emphasise that viewpoint, call M a *random mapping* from T to S.

Definition 15 Call M a *p-mapping* from T to S if the M_t are independent random elements of S with common distribution p. That is

$$P(M = (s_t)) = \prod_{t \in T} p_{s_t} \qquad \text{for every } (s_t) \in S^T \ .$$

For each subset B of S^T the formula

$$\Sigma_B(x_s, s \in S) := \sum_{(s_t) \in B} \prod_{t \in T} x_{s_t} \tag{38}$$

defines a polynomial Σ_B in commuting variables $x_s, s \in S$, known as the *enumerator polynomial* of B [25] [17, p. 71]. If M is a p-mapping from T to S then

$$P(M \in B) = \Sigma_B(p_s, s \in S) \qquad (B \subseteq S^T). \tag{39}$$

The addition and multiplication rules for enumerator polynomials [17, p. 72] then match corresponding rules of probability. Any evaluation of the probability of an event defined by p-mapping from T to S as a function of $p = (p_s, s \in S)$ can be interpreted as an evaluation of an enumerator polynomials, and vice versa. The probabilistic expression often appears simpler than the combinatorial one, because replacing the variables x_s by p_s subject to $\sum_s p_s = 1$ usually eliminates some factors of $x_S := \sum_s x_s$. Compare (40) and (41) below for a typical example. If an identity of enumerator polynomials in variables p_s subject to $\sum_s p_s = 1$ is obtained by a probabilistic argument, the factors of x_S can always be recovered at the end by substituting $p_s = x_s/x_S$ in the probabilistic identity and then multiplying both sides by $x_S^{|T|}$. Repeated application of this method allows the various identities (8)-(12) to be deduced from their probabilistic expressions.

4.1 Mappings from S to S

For a mapping $m \in S^S$, let cyclic(m) be the set of cyclic points of m, and let $\mathcal{F}(m)$ be the forest derived from m, as defined after Theorem 3. For $r \in S$ let T_r be the set of mappings m from S to S such that r is the unique cyclic point of m, or, equivalently, $\mathcal{F}(m) \in \mathbf{T}(S, r)$, the set of trees labeled by S with root r. It is elementary and well known [49],[24, (6.7)] that the restriction to T_r of the map $m \to \mathcal{F}(m)$ is a bijection from T_r to $\mathbf{T}(S, r)$. Cayley's multinomial expansion (31) amounts via this bijection to the following formula for the enumerator polynomial of the subset T_r of S^S, obtained in a different way by Françon[25, Prop. 3.1] from the Foata-Fuchs coding of mappings [24]:

$$\Sigma_{T_r}(x_s, s \in S) = x_r^2 \left(\sum_{s \in S} x_s \right)^{|S|-2}. \tag{40}$$

Now let M be a p-mapping from S to S, and apply (39) to see that (40) amounts to

$$P(\mathcal{F}(M) \in \mathbf{T}(S, r)) = p_r^2 \qquad (r \in S). \tag{41}$$

found various equivalents and extensions of (41) by probabilistic arguments. See also [4]. Formula (41) is closely related to results of Burtin [12], Ross [51] and Jaworski [31]. To make this connection, let $\hat{\mathcal{D}}_r$ be the digraph with vertex set S and $\{(M_s, s), s \in S \Leftrightarrow \{r\}\}$

for its set of directed edges. So $\hat{\mathcal{D}}_r$ is derived from the functional digraph $\mathcal{D}(M)$ by first deleting the edge leading out of $r \in S$, then reversing the directions of the remaining $|S| - 1$ edges. Observe that

$$(\mathcal{F}(M) \in \mathbf{T}(S,r)) = (M_r = r) \cap (\hat{\mathcal{D}}_r \in \mathbf{T}(S,r)).$$

and that if $\mathcal{F}(M) \in \mathbf{T}(S,r)$ then $\hat{\mathcal{D}}_r = \mathcal{F}(M)$. Since M_r is independent of $\hat{\mathcal{D}}_r$, and $P(M_r = r) = p_r$, formula (41) is equivalent to

$$P(\hat{\mathcal{D}}_r \in \mathbf{T}(S,r)) = p_r. \tag{42}$$

By definition, $\hat{\mathcal{D}}_r$ has $|S|$ vertices and $|S| - 1$ edges, so

$$(\hat{\mathcal{D}}_r \in \mathbf{T}(S,r)) \Leftrightarrow (\hat{\mathcal{D}}_r \text{ is connected})$$

and (42) therefore amounts to the result of [12, 51] that

$$P(\hat{\mathcal{D}}_r \text{ is connected }) = p_r. \tag{43}$$

The above argument can of course be reversed to deduce the form (31) of Cayley's multinomial expansion from (43).

Proof of Lemma 12 (iv). Suppose that $\mathcal{T}$ is a p-tree. It follows immediately from the definition of $\hat{\mathcal{D}}_r$ that

$$P(\hat{\mathcal{D}}_r = \mathbf{t}) = P(\mathcal{T} = \mathbf{t}) \text{ for each } \mathbf{t} \in \mathbf{T}(S,r). \tag{44}$$

It follows that

$$\operatorname{dist}(\hat{\mathcal{D}}_r \,|\, \hat{\mathcal{D}}_r \in \mathbf{T}(S,r)) = \operatorname{dist}(\mathcal{T} \,|\, \operatorname{root}(\mathcal{T}) = r) \tag{45}$$

and hence that

$$\operatorname{dist}(\mathcal{F}(M) \,|\, \mathcal{F}(M) \in \mathbf{T}(S,r)) = \operatorname{dist}(\mathcal{T} \,|\, \operatorname{root}(\mathcal{T}) = r) \tag{46}$$

as claimed. □

The next lemma was suggested by arguments of Ross [51] and Jaworski [31].

Lemma 16 *Let M_A be a p-mapping from A to S for some non-empty subset A of S. Let $\mathcal{D}(M_A)$ be the associated digraph with vertex set S and edge set $\{(a, M_s), a \in A\}$, and denote the range of M_A by $M_A(A) := \{M_s, s \in A\}$. Then for each $R \subseteq S - A$*

$$P[\mathcal{D}(M_A) \text{ contains no cycles and } M_A(A) \subseteq A \cup R] = p_R(p_R + p_A)^{|A|-1}. \tag{47}$$

Proof. The formula will be established in three steps.
Step I. Suppose that $|A| = |S| \Leftrightarrow 1$ and $R = S \Leftrightarrow A$. Then (47) reduces to formula (42).
Step II. Suppose that $R = S \Leftrightarrow A$. Then the right side of (47) reduces to p_R. If R is empty the result is trivially true, with both sides of (47) equal to 0. So assume R is not empty and let r be some arbitrary element of R. Define $\tilde{M}_A : A \to A \cup \{r\}$ by $\tilde{M}_A(a) = M_A(a)$ if $M_A(a) \in A$ and $\tilde{M}_A(a) = r$ if $M_A(a) \in R$. Then $\tilde{M}_A$ is a $\tilde{p}$-mapping from A to $A \cup \{r\}$ for $\tilde{p}_a = p_a$ for $a \in A$ and $\tilde{p}_r = p_R$. Since $\mathcal{D}(M_A)$ contains no cycles iff $\mathcal{D}(\tilde{M}_A)$ contains no cycles, the conclusion follows from the result of Step I applied to $\tilde{M}_A$.
Step III. General A and R. Let $F_{A,R}$ denote the event that $\mathcal{D}(M_A)$ contains no cycles and $M_A(A) \subseteq A \cup R$). Then

$$P(F_{A,R}) = P(M_A(A) \subseteq A \cup R)\, P(F_{A,R} \,|\, M_A(A) \subseteq A \cup R)$$

$$= (p_R + p_A)^{|A|} \, \frac{p_R}{p_R + p_A}$$

where $P(F_{A,R} \,|\, M_A(A) \subseteq A \cup R)$ is evaluated by the result of Step II applied to M_A given $M_A(A) \subseteq A \cup R$, using the fact that M_A given $M_A(A) \subseteq A \cup R$ is a p'-mapping from A to $A \cup R$ where p' is p conditioned on $A \cup R$. $\Box$

Proof of Theorem 3. Case (i) can be read from the proof of Proposition 3.7 of [25] by the probabilistic interpretation (39) of enumerator polynomials. Case (ii) is implied by the proof of Theorem 3 of Jaworski [31]. In view of (46), case (iii) can be read from Theorem 13, and vice versa. Cases (i) and (iii) of the Theorem can also be obtained probabilistic arguments similar to Jaworski's proof of (ii). Following are details of this approach in case (iii).

Fix $A \subseteq [n]$. Consider the conditional probability of the event $(\mathrm{pred}(0, M) = A)$ given that $\mathrm{cyclic}(M) = \{n+1\}$, where M is a p-mapping from S to S for $S := [0, n+1]$. Note that $\bar{A}$ is the complement of A relative to $[n]$ not S, so $\bar{A} \subseteq [n]$. From (41), the conditioning event has probability p_{n+1}^2, so the problem is to find the probability of the event $(\mathrm{pred}_0(M) = A, \mathrm{cyclic}(M) = \{n+1\})$. Let M_B be M with its domain restricted to $B \subseteq S$. The event $(\mathrm{pred}_0(M) = A, \mathrm{cyclic}(M) = \{n+1\})$ is the intersection the following four events

(I) the event that M_A has range $A \cup \{0\}$ and $\mathcal{D}(M_A)$ has no cycles, which by Lemma 16 has probability $p_0(p_0 + p_A)^{|A|-1}$;

(II) the event $(M_0 \in \bar{A} \cup \{n+1\})$, which has probability $p_{n+1} + p_{\bar{A}}$;

(III) the event that $M_{\bar{A}}$ has range $\bar{A} \cup \{n+1\}$ and $\mathcal{D}(M_{\bar{A}})$ has no cycles, which by Lemma 16 has probability $p_{n+1}(p_{n+1}+p_{\bar{A}})^{|\bar{A}|-1}$;

(IV) the event $(M_{n+1} = n+1)$, which has probability p_{n+1}.

Since these four events are determined by the restrictions of M to four disjoint subsets of $[0, n+1]$, they are independent. The conditional probability in question is therefore the product of these four probabilities, divided by the probability p_{n+1}^2 of the conditioning event. □

4.2 The random set of cyclic points

The following generalization of (41) gives the distribution of the random set of cycles of a p-mapping:

Proposition 17 *Let M be a p-mapping from S to S. Then the distribution of* cyclic(M) *on 2^S is determined by the formula*

$$P(\text{cyclic}(M) = R) = |R|!(\textstyle\prod_{r\in R} p_r)p_R \qquad (R \subseteq S). \tag{48}$$

Proof. With notation of Lemma 16, the event (cyclic$(M) = R$) is the intersection of the events $(M_R(R) = R)$ and ($\mathcal{D}(M_{R^c})$ contains no cycles). The probability of the first event is easily seen to be $|R|!\prod_{r\in R} p_r$, while the probability of the second event is p_R by Lemma 16. Since the two events are independent, the formula (48) follows. □

The special case $x = 0$ of (11) follows from the fact that formula (48) sums to 1 over all subsets R of S for all probability distributions p on S. Sum (48) over all R with $|R| = k$ to see that for $1 \le k \le |S|$ the probability that a p-mapping M from S to S has exactly k cycles is

$$P(|\text{cyclic}(M)| = k) = k! \sum_{|R|=k} \Pi(R)p_R \text{ where } \Pi(R) := \textstyle\prod_{r\in R} p_r \tag{49}$$

and the sum is over all subsets R of S of size k. Jaworski [31, Theorem 2] found the following alternative expression for the same probability:

$$P(|\text{cyclic}(M)| = k) = k! \sum_{|R|=k} \Pi(R) \Leftrightarrow (k+1)! \sum_{|R|=k+1} \Pi(R) \tag{50}$$

which can evidently be recast as

$$P(|\text{cyclic}(M)| \ge k) = k! \sum_{|R|=k} \Pi(R). \tag{51}$$

As a check, the implied equality between the right sides of (49) and (50) is easily verified directly.

See [51, 31] for further results about p-mappings. See [35, 2, 41] for results and further references to the literature for uniform p. The case when all of the p_s but one are equal has also been studied in detail [55, 40, 8]. See also [12, 21, 28, 29, 6, 7, 9, 33, 5] regarding various other models for random mappings.

5 Random Forests

The proof of Theorem 1 is based on the following lemma:

Lemma 18 *For M a p-mapping from S to S the distribution of the associated random forest $\mathcal{F}(M)$ is given by the formula*

$$P(\mathcal{F}(M) = \mathbf{f}) = |R(\mathbf{f})|! \left(\prod_{r \in R(\mathbf{f})} p_r \right) \left(\prod_{s \in S} p_s^{C_s \mathbf{f}} \right) \tag{52}$$

where $R(\mathbf{f})$ is the set of roots of $\mathbf{f}$, and $\mathbf{f}$ ranges over the set of all $(|S|+1)^{|S|-1}$ rooted forests labeled by S.

Proof. For each given forest $\mathbf{f}$ the first factor is the number of permutations of $R(\mathbf{f})$, the second is the probability that the restriction of M to $R(\mathbf{f})$ equals any particular one of these permutations, and the third is the probability that the restriction of M to $S \Leftrightarrow R(\mathbf{f})$ is as dictated by $\mathbf{f}$. □

Proof of Theorem 1. Compare (52) and (48) to obtain (1) for $x_s := p_s$ with $p_s \geq 0$ and $\sum_s p_s = 1$. The usual substitution $p_s = x_s/x_S$ yields (1) for $x_s \geq 0$ with $x_S > 0$, hence the polynomial identity. □

5.1 Distribution of the roots of a p-forest

Recall from Proposition 5 that the vector of out-degree counts $C\mathcal{F}_k$ of a p-forest $\mathcal{F}_k$ has a multinomial distribution with parameters $n \Leftrightarrow k$ and $(p_s, s \in S)$. This observation, combined with the following corollary of Theorem 1, determines the joint distribution of the random vector $C\mathcal{F}_k$ and the random set roots$(\mathcal{F}_k)$, whose marginal distribution was described by formula (26).

Corollary 19 *Let $\mathcal{R}_k$ be the random set of roots of $\mathcal{F}_k$, a p-forest of k trees labeled by S, where $1 \leq k < n = |S|$. Then for each possible vector of counts $c := (c_s, s \in S)$ with $\sum_s c_s = n \Leftrightarrow k$, the conditional distribution of $\mathcal{R}_k$ given $C\mathcal{F}_k = c$ is given by*

$$P(\mathcal{R}_k = R \,|\, C\mathcal{F}_k = c) = \binom{n \Leftrightarrow 1}{k \Leftrightarrow 1}^{-1} \frac{c_R}{(n \Leftrightarrow k)} \qquad (R \subseteq S \text{ with } |R| = k). \tag{53}$$

Proof. Let $\mathbf{F}(S, R; c)$ be the set of all $\mathbf{f} \in \mathbf{F}(S, R)$ such that $C\mathbf{f} = c$. Then from (1)

$$|\mathbf{F}(S, R; c)| = \frac{c_R\,(n \Leftrightarrow k \Leftrightarrow 1)!}{\prod_{s \in S} c_s!} \tag{54}$$

and (53) follows easily from (54) and (20) by canceling factorials. □

Let $C_B\mathcal{F}_k := \sum_{s \in B} C_s\mathcal{F}_k$, the number of vertices of $\mathcal{F}_k$ that are children of some vertex in B. As an immediate consequence of Proposition 5 there is the following analog for a p-forest $\mathcal{F}_k$ of the result of Clarke [16] regarding the binomial distribution of vertex degrees in a uniform unrooted random tree:

$$\mathrm{dist}(C_B\mathcal{F}_k) = \mathrm{binomial}(n \Leftrightarrow k, p_B). \tag{55}$$

As a check, formula (26) for $P(\mathcal{R}_k = R)$ is recovered from (53) as the expectation of the conditional probability, because the binomial$(n \Leftrightarrow k, p_R)$ distribution of $C_R\mathcal{F}_k$ has mean $(n \Leftrightarrow k)p_R$. The following lemma, which is easily checked directly, is a consequence of formula (26):

Lemma 20 *For $1 \leq k < n := |S|$ let S_k be the set of all subsets B of S with $|B| = k$. For each non-negative function $w = (w_s, s \in S)$ with $w_S > 0$, the formula*

$$P(\mathcal{R}_k = R) = \binom{n \Leftrightarrow 1}{k \Leftrightarrow 1}^{-1} \frac{w_R}{w_S} \qquad (R \in S_k) \tag{56}$$

defines the probability distribution of a random element $\mathcal{R}_k$ of S_k.

Call this distribution of $\mathcal{R}_k$ *the distribution on S_k induced by w*. According (26), for a p-forest $\mathcal{F}_k$ the unconditional distribution of $\mathcal{R}_k := \mathrm{roots}(\mathcal{F}_k)$ is the distribution on S_k induced by p. According to (53), the conditional distribution of $\mathcal{R}_k$ given $C\mathcal{F}_k = c$ is then the distribution on S_k induced by c. The probabilities of events determined by

$\mathcal{R}_k$ with the distribution (56) induced by w, can be obtained by summation of (56) over appropriate $R \in S_k$, then applied to $\mathcal{R}_k := \mathrm{roots}(\mathcal{F}_k)$, either unconditionally with $w = p$, or conditionally given $C\mathcal{F} = c$ with $w = c$. To illustrate, for each fixed subset A of S with $|A| = a$, it is easily shown that for $\mathcal{R}_k$ with the distribution on S_k induced by w

$$P(\mathcal{R}_k \supseteq A) = \binom{n \Leftrightarrow 1}{k \Leftrightarrow 1}^{-1} \frac{1}{w_S} \left[\binom{n \Leftrightarrow a}{k \Leftrightarrow a} w_A + \binom{n \Leftrightarrow a \Leftrightarrow 1}{k \Leftrightarrow a \Leftrightarrow 1} (w_S \Leftrightarrow w_A) \right]. \tag{57}$$

Take $A = \{r\}$ and simplify to obtain (27). As a check on (27), take $p_s \equiv 1/n$. Then $\mathcal{F}_k$ has uniform distribution on $\mathbf{F}_k(S)$, and $\mathrm{roots}(\mathcal{F}_k)$ has uniform distribution on S_k. Obviously then $P(r \in \mathrm{roots}(\mathcal{F}_k)) = k/n$, in agreement with (27) for $p_r = 1/n$. The formula (27) implies also the less obvious result that $P(r \in \mathrm{roots}(\mathcal{F}_k)) = k/n$ if $p_r = 1/n$, no matter what the p_s for $s \neq r$. Similarly, (57) yields

$$P(r \in \mathrm{roots}(\mathcal{F}_k) \,|\, C\mathcal{F}_k = c) = \frac{k \Leftrightarrow 1 + c_r}{n \Leftrightarrow 1}. \tag{58}$$

That is, by application of Proposition 5(ii):

Corollary 21 *Among all forests* $\mathbf{f}$ *of* k *trees labeled by* $[n]$ *with a given sequence of out-degrees* $(c_s, s \in [n])$, *the fraction such that* r *is the root of some tree component of* $\mathbf{f}$ *equals* $(k \Leftrightarrow 1 + c_r)/(n \Leftrightarrow 1)$.

As a check, take expectations in (58) and use the fact that c_r is the given value of the binomial$(n \Leftrightarrow k, p_r)$ variable C_r with expectation $(n \Leftrightarrow k)p_r$ to see that the unconditional probability of $(r \in \mathrm{roots}(\mathcal{F}_k))$ is given by (58) with c_r replaced by $(n \Leftrightarrow k)p_r$, as in (27).

5.2 Conditioning on the set of roots.

The proof of Theorem 1 by comparison of (48) and (52) has the following immediate corollary:

Corollary 22 *Let* $\mathcal{F}(M)$ *be the random forest with* $\mathrm{roots}(\mathcal{F}(M)) := \mathrm{cyclic}(M)$ *derived from a p-mapping* M *from* S *to* S. *For each* $1 \leq k \leq |S|$ *let* $\mathcal{F}_k$ *be a p-forest of* k *trees labeled by* S. *Then for each subset* R *of* S *with* $p_R > 0$ *and* $|R| = k$

$$\mathrm{dist}(\mathcal{F}(M) \,|\, \mathrm{cyclic}(M) = R) = \mathrm{dist}(\mathcal{F}_k \,|\, \mathrm{roots}(\mathcal{F}_k) = R). \tag{59}$$

If $\mathcal{F}_R$ *denotes a random forest with the common distribution displayed in* (59) *, then for each forest* $\mathbf{f}$ *labeled by* S *with* $\mathrm{roots}(\mathbf{f}) = R$

$$P(\mathcal{F}_R = \mathbf{f}) = p_R^{-1} \prod_{s \in S} p_s^{C_s \mathbf{f}}. \tag{60}$$

Formula (60) is a probabilistic expression of the polynomial identity (1). It is not claimed, nor is it true for general p, that the distribution of $\mathcal{F}(M)$ given that M has k cyclic points is the same as the distribution of $\mathcal{F}_k$. By inspection of formulae (26) and (48), this is true for p uniform on S, but false otherwise.

The following theorem lays out some facts regarding a random forest $\mathcal{F}_R$ with distribution (60), call it a *p-forest with roots R*. According to the above Corollary, these facts apply both to a p-forest $\mathcal{F}_k$ given roots$(\mathcal{F}_k) = R$, and to $\mathcal{F}(M)$ derived from a p-mapping M given that cyclic$(M) = R$. For a forest $\mathbf{f}$ labeled by S with roots$(\mathbf{f}) = R$, and $v \in S \Leftrightarrow R$ let $M_v(\mathbf{f}) \in S$ be the *mother of v in* $\mathbf{f}$, that is the unique $s \in S$ such that $s \xrightarrow{\mathbf{f}} v$. For $A \subseteq S$ the *restriction of* $\mathbf{f}$ *to* A is the forest $\mathbf{f}^A$ labeled by A whose set of edges is the intersection with $A \times A$ of the set of edges of $\mathbf{f}$. Write $p(\cdot \mid A)$ for the probability distribution p conditioned on A.

Theorem 23 *Let $\mathcal{F}_R$ be a p-forest labeled by S with roots $R \subset S$. Let $\mathcal{H}_1$ be the random set of all children of the root vertices in $\mathcal{F}_R$. Then*
(i) the distribution of $|\mathcal{H}_1| \Leftrightarrow 1$ is binomial$(|S| \Leftrightarrow |R| \Leftrightarrow 1, p_R)$;
(ii) given $|\mathcal{H}_1| = m$ the restriction of $\mathcal{F}_R^{S-R}$ of $\mathcal{F}_R$ to $S \Leftrightarrow R$, whose set of roots is $\mathcal{H}_1$, is a $p(\cdot \mid S \Leftrightarrow R)$-forest of m trees labeled by $S \Leftrightarrow R$;
(iii) the distribution of $\mathcal{H}_1$ is given by the formula

$$P(\mathcal{H}_1 = B) = p_B \, p_{S-R}^{|S-R-B|-1} p_R^{|B|-1} \qquad (B \subseteq S \Leftrightarrow R) \tag{61}$$

(iv) for each non-empty $B \subseteq S \Leftrightarrow R$, conditionally given $\mathcal{H}_1 = B$ the restricted forest $\mathcal{F}_R^{S-R}$ is a $p(\cdot \mid S \Leftrightarrow R)$-forest labeled by $S \Leftrightarrow R$ with roots B, and this restricted forest is independent of the random variables $M_b(\mathcal{F}_R), b \in B$, which are conditionally independent with common distribution $p(\cdot \mid R)$.

Proof. These claims follow easily from formula (60). As a check, the formula in (iii) can be read from Lemma 16, Proposition 17, and the representation of $\mathcal{F}_R$ as $\mathcal{F}(M)$ given cyclic$(M) = R$ for a p-mapping M. Thus $P(\mathcal{H}_1 = H) = P(A)/P(\text{cyclic}(M) = R)$ where A is the event that the restriction of M to $S \Leftrightarrow R \Leftrightarrow H$ has no cycles and that M maps H_1 to R and R to R. $\square$

Corollary 24 *For $\mathcal{F}_R$ a p-forest labeled by S with roots $R \subset S$,*

$$P(r \overset{\mathcal{F}_R}{\rightsquigarrow} s) = p_r / p_R \qquad (r \in R, s \in S \Leftrightarrow R) \tag{62}$$

and for all such r and s the event $(r \overset{\mathcal{F}_R}{\rightsquigarrow} s)$ is independent of the restriction of $\mathcal{F}_R$ to $S \Leftrightarrow R$.

Proof. As in Theorem 23, let $\mathcal{H}_1$ be the random set of children of R in $\mathcal{F}_R$. Given $\mathcal{H}_1 = B$ say let $X \in B$ be the root of the subtree containing s in the restriction of $\mathcal{F}_R$ to $S \Leftrightarrow R$. There is a path from r to s in $\mathcal{F}_R$ if and only if $M_X = r$ where $M_X \in R$ is the mother of X in $\mathcal{F}_R$. But according to part (iv) of Theorem 23, given the restricted forest $\mathcal{F}_R^{S-R}$, which together with s determines X, the random variables M_b for $b \in B$ are independent with common distribution $p(\cdot \mid R)$. Therefore, the conditional distribution of M_X given $\mathcal{F}_R^{S-R}$ is $p(\cdot \mid R)$, as claimed. $\square$

Proof of Theorem 9
(i) This evaluation of $P(r \overset{\mathcal{F}_k}{\rightsquigarrow} s \mid r \in \mathcal{R}_k)$ is obtained by conditioning on $\mathcal{R}_k = R$ and then summing over the $\binom{n-2}{k-1}$ possible choices of R with $r \in R$ and $s \notin R$. By application of Corollary 8 and part (i), the terms of the sum are all equal, hence

$$P(r \overset{\mathcal{F}_k}{\rightsquigarrow} s \mid r \in \mathcal{R}_k) = \binom{n \Leftrightarrow 2}{k \Leftrightarrow 1} \frac{p_r}{p_R} \binom{n \Leftrightarrow 1}{k \Leftrightarrow 1}^{-1} p_R \left(\frac{k \Leftrightarrow 1 + (n \Leftrightarrow k) p_r}{n \Leftrightarrow 1} \right)^{-1}$$

which reduces to (28).
(ii) This follows from (27) and (28) by elementary rules of probability. $\square$

Direct Proof of Corollary 10. Without loss of generality, take $S = [n], r = 1, s = 2$. A forest $\mathbf{f}$ with a given out-degree sequence $(c_1, \ldots, c_n)$ corresponds to a unique sequence of choices of the sets $J_i := \{j : i \xrightarrow{\mathbf{f}} j\}$ of sizes c_i subject to the constraint that $\mathbf{f}$ is a forest. As argued in [45], the set J_1 can be any subset of $[n] \Leftrightarrow \{1\}$ of size c_1. Given J_1, the set J_2 can be any subset of size c_2 of a set of permissible elements of size $n \Leftrightarrow 1 \Leftrightarrow c_1$ that is determined by J_1, and so on. So the number of such forests is

$$\binom{n \Leftrightarrow 1}{c_1} \binom{n \Leftrightarrow 1 \Leftrightarrow c_1}{c_2} \cdots \binom{n \Leftrightarrow 1 \Leftrightarrow \sum_{i=1}^{n-1} c_i}{c_n} = \frac{(n \Leftrightarrow 1)!}{(k \Leftrightarrow 1)! \prod_{i=1}^n c_i!}. \tag{63}$$

This is the identity of coefficients of $\prod_{i=1}^n x_i^{c_i}$ in (2) for $S = [n]$. Consider now the number of these forests $\mathbf{f}$ subject to the additional constraint that $1 \in \mathrm{roots}(\mathbf{f})$ and $1 \overset{\mathbf{f}}{\not\rightsquigarrow} 2$. In selecting the sequence of sets $J_i := \{j : i \xrightarrow{\mathbf{f}} j\}$ at each stage i the additional constraint reduces by exactly 1 the number of vertices from which it is permissible to choose J_i. So the number of forests $\mathbf{f}$ of k trees labeled by $[n]$ with the given out-degree sequence

$(c_1, \ldots, c_n)$ and such that $1 \in \operatorname{roots}(\mathbf{f})$ and $1 \overset{\mathbf{f}}{\not\to} 2$ equals

$$\binom{n \Leftrightarrow 2}{c_1}\binom{n \Leftrightarrow 2 \Leftrightarrow c_1}{c_2} \cdots \binom{n \Leftrightarrow 2 \Leftrightarrow \sum_{i=1}^{n-1} c_i}{c_n} = \frac{(n \Leftrightarrow 2)!}{(k \Leftrightarrow 2)! \prod_{i=1}^{n} c_i!}$$

The ratio of these two numbers yields the fraction $(k \Leftrightarrow 1)/(n \Leftrightarrow 1)$. $\square$

5.3 Level sets

For a random forest $\mathbf{f}$ labeled by S let $\mathcal{L}_h(\mathbf{f})$ denote the random subset of S defined by the random set of all vertices of $\mathbf{f}$ at height h from the root. So $\mathcal{L}_0(\mathbf{f}) = \operatorname{roots}(\mathbf{f})$, and for each $h \geq 1$ the set $\mathcal{L}_h(\mathbf{f})$ is the set of all children of vertices in $\mathcal{L}_{h-1}(\mathbf{f})$. Repeated application of Theorem 23 gives a simple formula for the joint distribution of $(\mathcal{L}_i(\mathcal{F}_R), 1 \leq i \leq h)$ for any fixed h. In particular:

Corollary 25 *Let $\mathcal{F}_R$ be a p-forest labeled by $R \cup [n]$ with root set R, where R is a finite set disjoint from $[n]$ and $p_R > 0$. Then for each sequence of m non-empty subsets $(B_h, 1 \leq h \leq m)$ whose union is $[n]$,*

$$P(\mathcal{L}_h(\mathcal{F}_R) = B_h \text{ for all } 1 \leq h \leq m) = p_R^{|B_1|-1} \prod_{h=2}^{m} p_{B_{h-1}}^{|B_h|}. \tag{64}$$

As usual, there is a corresponding identity of polynomials, in this case the following variant of (12):

$$\sum_{m=1}^{n} \sum_{(B_1,\ldots,B_m)} x^{|B_1|} \prod_{h=2}^{m} z_{B_{h-1}}^{|B_h|} = x(x + z_{[n]})^{n-1} \tag{65}$$

where the inner sum is over all ordered partitions $(B_1, \ldots, B_m)$ of $[n]$.

5.4 Fringe trees

Consider now the distribution of the random set of vertices of the fringe tree of $\mathcal{F}_k$ with root s for $\mathcal{F}_k$ a p-forest labeled by S and $s \in S$. After relabeling S by $[0, n] := \{0\} \cup [n]$ there is no loss of generality in supposing that $S = [0, n]$ and that $s = 0$.

Theorem 26 *Let $V_0(\mathcal{F}_k) \subseteq [n]$ be the set of non-root vertices of the fringe subtree of $\mathcal{F}_k$ rooted at 0, for $\mathcal{F}_k$ a p-forest of k trees labeled by $[0, n]$. Then for $A \subseteq [n]$*

$$P(V_0(\mathcal{F}_k) = A) = \binom{|\bar{A}|}{k \Leftrightarrow 1}\binom{n}{k \Leftrightarrow 1}^{-1} p_0 (p_0 + p_A)^{|A|-1} p_{\bar{A}}^{|\bar{A}|-(k-1)} \tag{66}$$

where $\bar{A} := [n] \Leftrightarrow A$.

Proof. Let $c_{n,k} := \binom{n}{k-1}^{-1}$. From (19), the probability $P(V_0(\mathcal{F}_k) = A)$ is the sum of $P(\mathcal{F}_k = \mathbf{f}) := c_{n,k} \prod_{i=0}^{n} p_i^{C_i \mathbf{f}}$ over all forests $\mathbf{f}$ of k trees labeled by $[0,n]$ such that $V_0(\mathbf{f}) = A$. Regarding each $\mathbf{f}$ as a subset of $[0,n]^2$, there are two kinds of $\mathbf{f}$ to consider.
(i) $(0 \notin \mathrm{roots}(\mathbf{f}), V_0(\mathbf{f}) = A)$. Then $\mathbf{f} = \mathbf{v} \cup \mathbf{g} \cup \{(j,0)\}$ for some tree $\mathbf{v}$ labeled by A with with root 0, some forest $\mathbf{g}$ of k trees labeled by $\bar{A}$ and some edge $(j,0)$ with $j \in \bar{A}$, and

$$P(\mathcal{F}_k = \mathbf{f}) = c_{n,k} \left(\prod_{i \in A_0} p_i^{C_i \mathbf{v}} \right) \left(\prod_{\ell \in \bar{A}} p_\ell^{C_\ell \mathbf{f}} \right) p_j.$$

By application of (1) and (2), the sum of $P(\mathcal{F}_k = \mathbf{f})$ over these these $\mathbf{f}$ is

$$P(0 \notin \mathrm{roots}(\mathcal{F}_k), V_0(\mathcal{F}_k) = A) = c_{n,k}\, p_0 (p_0 + p_A)^{|A|-1} \binom{|\bar{A}| \Leftrightarrow 1}{k \Leftrightarrow 1} p_{\bar{A}}^{|\bar{A}|-k}\, p_{\bar{A}} \tag{67}$$

(ii) $(0 \in \mathrm{roots}(\mathbf{f}), V_0(\mathbf{f}) = A)$. Then $\mathbf{f} = \mathbf{v} \cup \mathbf{g}$ for some tree $\mathbf{v} \in \mathbf{T}(A,0)$, some forest $\mathbf{g}$ of $k \Leftrightarrow 1$ trees labeled by $\bar{A}$. The probability of each such $\mathbf{f}$ is

$$P(\mathcal{F}_k = \mathbf{f}) = c_{n,k} \left(\prod_{i \in A_0} p_i^{C_i \mathbf{v}} \right) \left(\prod_{\ell \in \bar{A}} p_\ell^{C_\ell \mathbf{f}} \right)$$

Summing over these $\mathbf{f}$ gives similarly

$$P(0 \in \mathrm{roots}(\mathcal{F}_k), V_0(\mathcal{F}_k) = A) = c_{n,k}\, p_0 (p_0 + p_A)^{|A|-1} \binom{|\bar{A}| \Leftrightarrow 1}{k \Leftrightarrow 2} p_{\bar{A}}^{|\bar{A}|-(k-1)} \tag{68}$$

Since $\binom{b-1}{k-1} + \binom{b-1}{k-2} = \binom{b}{k-1}$, addition of (67) and (68) gives (66). $\square$

As a consequence of the above calculation there is the following curious formula: for $V_0(\mathcal{F}_k) \subseteq [n]$ the set of non-root vertices of the fringe subtree rooted at 0 of p-forest of k trees labeled by $[0,n]$,

$$P(0 \in \mathrm{roots}(\mathcal{F}_k) \,|\, V_0(\mathcal{F}_k) = A) = \frac{k \Leftrightarrow 1}{n \Leftrightarrow |A|} \qquad (A \subseteq [n]) \tag{69}$$

where $0/0 := 1$.

5.5 Distribution of tree components

Formulae for the distributions of variously defined tree components of a p-forest follow easily from formulae (1), (2) and (19). In particular, there is the following corollary of Theorem 1.

Corollary 27 *Let p be a probability distribution on $[0,n]$, and let $2 \leq k \leq n$. For $\mathcal{F}_k$ with the distribution induced by p on forests of k trees labeled by $[0,n]$, let $W_0(\mathcal{F}_k) \subseteq [n]$ be the random set of all vertices other than 0 in the tree component of $\mathcal{F}_k$ containing 0. Then the distribution of the random subset $W_0(\mathcal{F}_k)$ of $[n]$ is given by the formula*

$$P(W_0(\mathcal{F}_k) = A) = \binom{n}{k \Leftrightarrow 1}^{-1} \binom{|\bar{A}| \Leftrightarrow 1}{k \Leftrightarrow 2} (p_0 + p_A)^{|A|} p_{\bar{A}}^{|\bar{A}|-(k-1)} \qquad (A \subseteq [n], |A| \leq n \Leftrightarrow k+1) \tag{70}$$

where $\bar{A} := [n] \Leftrightarrow A$.

In Theorem 1, take $S = R \cup [n]$, decompose the sum on the left side of (1) according to the partition of S into tree components, factorize over blocks of the partition, and apply Cayley's expansion (31) over trees within each block, to see that (1) implies:

Corollary 28 (Hurwitz's Multinomial Theorem [27]). *Let R be some finite set disjoint from $[n]$. Then there is the following identity of polynomials in $|R| + n$ commuting variables $x_s, s \in R \cup [n]$:*

$$\sum_{(B_r)} \prod_{r \in R} x_r (x_r + x_{B_r})^{|B_r|-1} = x_R (x_R + x_{[n]})^{n-1} \tag{71}$$

where the sum is over all $|R|^n$ choices of disjoint, possibly empty sets $B_r, r \in R$ with $\cup_{r \in R} B_r = [n]$.

Note that (5) is the particular case $k = 2$ of (71). See also [25, 52] for related combinatorial interpretations of this case, and see [53] for some extensions of (71). Part (i) of the following Corollary spells out the probabilistic interpretation of Hurwitz's multinomial theorem in terms of p-forests. Part (ii) corresponds to another identity of Hurwitz which can be read similarly from Theorem 1:

Corollary 29 *Let $\mathcal{F}_R$ be a p-forest labeled by S with roots R, where $S := R \cup [n]$ and R is disjoint from $[n]$. For $r \in R$ let $V_r(\mathcal{F}_R)$ be the random subset of $[n]$ defined by the non-root vertices of the tree component of $\mathcal{F}_R$ containing r. Then*

(i) *for each of $|R|^n$ possible choices of disjoint subsets $(B_r, r \in R)$ whose union is $[n]$*

$$P(V_r(\mathcal{F}_R) = B_r \text{ for all } r \in R) = p_R^{-1} \prod_{r \in R} p_r (p_r + p_{B_r})^{|B_r|-1} \tag{72}$$

(ii) *for each subset B of R the random set $V_B(\mathcal{F}_R) := \cup_{r \in B} V_r(\mathcal{F}_R)$ has the Hurwitz distribution $H_n^{-1,-1}(p^B)$ on $2^{[n]}$, where $p_0^B = p_B, p_{n+1}^B = p_{R-B}$, and $p_s^B = p_s$ for $s \in [n]$.*

For $V_B(\mathcal{F}_R)$ defined as in the previous Corollary, there is the remarkably simple formula

$$E(|V_B(\mathcal{F}_R)|) = np_B/p_R \tag{73}$$

because $V_B(\mathcal{F}_R)$ is the sum of the indicator variables $1(r \stackrel{\mathcal{F}_R}{\rightsquigarrow} s)$ over all $r \in B$ and $s \in [n]$, so formula (62) can be applied to compute:

$$E(|V_B(\mathcal{F}_R)|) = \sum_{r \in B} \sum_{s=1}^{n} P(r \stackrel{\mathcal{F}_R}{\rightsquigarrow} s) = \sum_{r \in B} np_r/p_R = np_B/p_R. \tag{74}$$

On the other hand, Corollary 29(ii) shows that (73) amounts to:

Proposition 30 *Suppose that a random subset V of $[n]$ has the Hurwitz $H_n^{-1,-1}(p)$ distribution. Then the mean of the $H_n^{-1,-1}(p)$-binomial distribution of $|V|$ is*

$$E\left(|V|\right) = n\left(\frac{p_0}{p_0 + p_{n+1}}\right). \tag{75}$$

This formula can also be confirmed as follows. Differentiate Hurwitz's formula (4) with respect to x to obtain

$$\sum_{A \subseteq [n]} y\,|A|\,(x + z_A)^{|A|-1}(y + z_{\bar{A}})^{|\bar{A}|-1} = n(x + y + z_{[n]})^{n-1}. \tag{76}$$

From the definition (16) of the $H_n^{-1,-1}(p)$ distribution,

$$E(|V|) = \sum_{A \subseteq [n]} \frac{|A|\,p_0\,p_{n+1}}{(p_0 + p_{n+1})}(p_0 + p_A)^{|A|-1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|-1}$$

and (75) follows by application of (76) with $x = p_0, y = p_{n+1}$ and $z_s = p_s$ for $s \in [n]$. $\square$

Formula (75) is a generalization of the known result [15] that the Abel $A_n^{-1,-1}(x,y)$-binomial distribution has mean $nx/(x+y)$. The proof of (75) just indicated via (74) provides a probabilistic explanation for this otherwise mysterious exception to the general rule that moments of Abel-binomial distributions are not simple functions of the parameters. See for instance [15] where a complicated expression is obtained for the second factorial moment of the $A_n^{-1,-1}(x, y)$-binomial distribution. In view of this difficulty in the Abel case, it does not seem possible to simplify the Hurwitz sums for higher moments of the $H_n^{-1,-1}(p)$-binomial distribution. For the $H_n^{0,-1}(p)$-binomial distribution, the Hurwitz sum for the mean does not simplify even in the Abel case.

5.6 A Hurwitz multinomial distribution

Riordan [50] considers multinomial forms of Abel's binomial theorem. See Berg and Mutafchiev [8] for the appearance of an Abel-trinomial distribution in the context of random mappings. Corollaries 28 and 29 show that the Hurwitz-multinomial distribution introduced in the following definition is a natural generalization of the the usual multinomial distribution.

Definition 31 For a probability distribution p on $[n] \cup R$ with $p_R > 0$, where R is a finite set disjoint from $[n]$, say that a random vector of non-negative integers $\mathbf{N}_R := (N_r, r \in R)$ has the *Hurwitz* $H_n^{-1,1}(p)$*-multinomial distribution* if for all vectors of non-negative integers $\mathbf{n}_R := (n_r, r \in R)$ with $\sum_r n_r = n$

$$P(\mathbf{N}_R = \mathbf{n}_R) = p_R^{-1} \sum_{(B_r)} \prod_{r \in R} p_r (p_r + p_{B_r})^{n_r - 1} \tag{77}$$

where the sum is over all $n!/(\prod_r n_r!)$ possible choices of disjoint subsets B_r of $[n]$ whose union is $[n]$ with $|B_r| = n_r, r \in R$.

The fact that (77) defines a probability distribution over vectors of non-negative integers $\mathbf{n}_R := (n_r, r \in R)$ with $\sum_r n_r = n$ amounts to Hurwitz's multinomial formula (71). According to Corollary 29, a random vector $\mathbf{N}_R$ with this distribution is obtained by defining N_r to be the size of the tree rooted at r in a p-forest $\mathcal{F}_R$ labeled by $[n] \cup R$ with roots R. The usual multinomial distribution with parameters n and $(p_r, r \in R)$ is recovered by taking $p_s = 0$ for all $s \in [n]$. In the corresponding forest $\mathcal{F}_R$, each vertex $s \in [n]$ is a leaf attached to a root $M_s \in R$ where the M_s are independent with common distribution p.

According to Theorem 23, in the more general model when p can assign positive probability to $[n]$, the restriction $\mathcal{F}_R^{[n]}$ of $\mathcal{F}_R$ to $[n]$ clusters the elements of $[n]$ into a random number K of subtrees such that $K \Leftrightarrow 1$ has binomial$(n \Leftrightarrow 1, p_R)$ distribution.

Given K the forest $\mathcal{F}_R^{[n]}$ formed by these subtrees is a $p(\cdot \mid [n])$-forest of K trees labeled by $[n]$, and each of these subtrees is attached to a root picked independently from R according to $p(\cdot \mid R)$. The size N_r of the tree rooted at r is then the sum of the sizes of those subtrees of $\mathcal{F}_R^{[n]}$ that happen to have r chosen as their root. From this construction of a random vector $\mathbf{N}_R$ with the $H_n^{-1,1}(p)$-multinomial distribution it follows without calculation that this family of multivariate distributions shares with the usual family of multinomial distributions the following basic rule for merging of categories:

Theorem 32 *Suppose that a random vector $\mathbf{N}_R$ has the $H_n^{-1,1}(p)$-multinomial distribution for some probability distribution p on $R \cup [n]$. Let Ψ be a map from R to Q, and let $\mathbf{N}_Q$ be the image of $\mathbf{N}_R$ after merging categories according to Ψ, that is*

$$\mathbf{N}_Q := \left(\sum_{r \in R} N_r 1(\Psi(r) = q), q \in Q \right).$$

Then $\mathbf{N}_Q$ has the $H_n^{-1,1}(p')$-multinomial distribution, where p' is the probability distribution on $Q \cup [n]$ defined by $p'_s = p_s$ if $s \in [n]$ and

$$p'_q = \sum_{r \in R} p_r 1(\Psi(r) = q), q \in Q.$$

5.7 The range of paths in a random tree

Let $\mathcal{U}$ be an unrooted p-tree labeled by S. For each pair of vertices $s, v \in S$ there is a unique path from s to v in $\mathcal{U}$. Let $\mathcal{R}_{s,v}$ denote the *range* of this path, meaning the random set all vertices along the path except s and v. By appropriate relabeling of vertices, to describe the distribution of the random subset $\mathcal{R}_{s,v}$ of S it is enough to consider the case $S := [0, n+1], s = 0, v = n+1$, as in the following theorem.

Theorem 33 *Let $\mathcal{U}$ be an unrooted p-tree labeled by $S := [0, n+1]$, and let $\mathcal{R}$ be the random subset of $[n]$ defined by the set of vertices of $\mathcal{U}$ on the unique path from 0 to $n+1$ in $\mathcal{U}$. Then the distribution of $\mathcal{R}$ on $2^{[n]}$ is given by the formula*

$$P(\mathcal{R} = A) = |A|! \left(\prod_{r \in A} p_r \right) (p_0 + p_{n+1} + p_A) \quad (A \subseteq [n]) \tag{78}$$

Proof. As indicated by Meir and Moon [38] and Joyal [32], given $\mathcal{R} = A$ the unrooted tree $\mathcal{U}$ induces a rooted forest $\mathcal{F}$ labeled by S with $\text{roots}(\mathcal{F}) = \{0\} \cup \{n+1\} \cup A$. It follows from (34) for that for every $A \subseteq [n]$ and every forest $\mathbf{f}$ with roots$(\mathcal{F}) = \{0\} \cup \{n+1\} \cup A$

$$P(\mathcal{R} = A, \mathcal{F} = \mathbf{f}) = |A|! \left(\prod_{r \in A} p_r \right) \prod_{s \in S} p_s^{C_s \mathbf{f}} \tag{79}$$

Sum this formula over all $\mathbf{f} \in \mathbf{F}(S, \{0\} \cup \{n+1\} \cup A)$ and apply (1) to obtain (78) $\square$

The fact that formula (78) defines a probability distribution on $2^{[n]}$ for each probability distribution p on $[0, n+1]$ yields the identity of polynomials (11). Note that there is only one extra variable x in this identity rather than two variables x and y, because of the way formula (78) depends on p_0 and p_{n+1} only through $p_0 + p_{n+1}$.

Consider now the special case $p_0 = p_{n+1} = 0$. Compare (78) and (48) to see that in this case $\mathcal{R}$ has the same distribution as the random set cyclic(M) where M is a p-mapping from $[n]$ to $[n]$, and p is regarded as a distribution on $[n]$ rather than $[0, n+1]$. Let $\mathcal{F}^\circ$ denote the restriction to $[n]$ of the rooted forest $\mathcal{F}$ derived from $\mathcal{U}$ as in Theorem 33. The assumption that $p_0 = p_{n+1} = 0$ implies that with probability 1 both 0 and $n+1$ are *leaves* of $\mathcal{U}$, that is vertices of degree 1. Therefore, the forest $\mathcal{F}^\circ$ labeled by $[n]$ has the same set of roots $\mathcal{R}$ as $\mathcal{F}$. Compare (79) and (52) to see that $\mathcal{F}^\circ$ has the same distribution as $\mathcal{F}(M)$, the forest of tree components generated by the digraph of M.

To explain this coincidence, let $\mathcal{U}_{[n]}$ denote the restriction of $\mathcal{U}$ to $[n]$, which is an unrooted tree labeled by $[n]$ because 0 and $n+1$ are leaves of $\mathcal{U}$. Let $\{0, R_1\}$ be the edge of $\mathcal{U}$ connecting 0 to $R_1 \in [n]$ and define $\{n+1, R_2\}$ similarly. It is easy to see that R_1 and R_2 are independent random elements of $[n]$ with distribution p, independent also of $\mathcal{U}_{[n]}$ which has the distribution on $\mathbf{U}([n])$ induced by p. Clearly, $\mathcal{R}$ is the range of the path from R_1 to R_2 in $\mathcal{U}_{[n]}$, where $\{R_1\} \cup \{R_2\}$ is regarded as part of the path. The coincidence is explained by Joyal's [32] bijection $m \leftrightarrow (\mathbf{u}, r_1, r_2)$ between $[n]^{[n]}$ and $\mathbf{U}([n]) \times [n] \times [n]$, where for a mapping $m : [n] \to [n]$ the corresponding unrooted tree $\mathbf{u}$ and pair of points $(r_1, r_2) \in [n] \times [n]$ are defined as follows. For m with and associated forest of k trees $\mathcal{F}(m)$ rooted at k cyclic points of m let $s_1, \ldots s_k$ be these cyclic points listed in increasing order; for $1 \leq i \leq k$ let $r_1 = m(s_1), r_2 = m(s_k)$, and let $\mathbf{u}$ be the unrooted tree whose edges are the edges of $\mathcal{F}(m)$ (with directions ignored) together with $\{\{m(s_i), m(s_{i+1})\}, 1 \leq i < k\}$. Then it is easily checked that M is a p-mapping form $[n]$ to $[n]$ if and only if the corresponding triple $(\mathcal{U}_{[n]}, R_1, R_2)$ is such that $\mathcal{U}_{[n]}$ is an unrooted p-tree labeled by $[n]$, R_1 and R_2 have distribution p, and these three random elements are independent.

The following corollary spells out one implication of the above argument in terms of a random rooted tree:

Corollary 34 *Let S be a finite set and let M be a random mapping from S to S with the product distribution induced by p on S^S. Let $\mathcal{T}$ be a rooted random tree with the distribution on induced by p on $\mathbf{T}(S)$, let $R := \mathrm{root}(\mathcal{T})$, let V be independent of $\mathcal{T}$ with distribution p on S, and let H be the height of V in $\mathcal{T}$, that is the number of vertices on the path from R to V in $\mathcal{T}$, not counting either R or V. Then $H+1$ has the same distribution as the number of cyclic points of M, as described by formulae* (49) -(51) .

5.8 Spanning subtrees

The path joining two vertices in an unrooted tree $\mathbf{u}$ labeled by S is the subtree of $\mathbf{u}$ spanning a two point subset of S. The proof of Theorem 33 extends easily to yield the following generalization of that result:

Theorem 35 *Let $\mathcal{U}$ be an unrooted p-tree labeled by S, let F be a subset of S of size two or more, and let $\mathcal{U}_F$ denote the subtree of $\mathcal{U}$ spanning F. Then for every unrooted tree $\mathbf{u}$ labeled by a finite subset $V(\mathbf{u})$ of S, such that the set of vertices of $\mathbf{u}$ of degree one is contained in F,*

$$P(\mathcal{U}_F = \mathbf{u}) = \left(\prod_{v \in V(\mathbf{u})} p_v^{D_v \mathbf{u} - 1} \right) p_{V(\mathbf{u})} \tag{80}$$

where $D_v \mathbf{u}$ is the degree of vertex v in the tree $\mathbf{u}$.

6 Percolation probabilities

Consider for two vertices $s, v \in S$ the probability that of the event $(s \overset{\mathcal{F}_k}{\sim} v)$ that s and v lie in the same tree component of a random forest $\mathcal{F}_k$ with the distribution on $\mathbf{F}_k(S)$ induced by p. By a suitable relabeling, it suffices to find a formula for $P(0 \overset{\mathcal{F}_k}{\sim} n+1)$ in the case $S := [0, n+1]$ and $2 \le k \le n+1$. Recall that $V_s(\mathbf{f})$ is the set of vertices of the tree component of $\mathbf{f}$ containing s. A now familiar argument yields

$$P(0 \overset{\mathcal{F}_k}{\sim} n+1) = \sum_{\substack{A \subset [n] \\ |A| \le n-k+1}} \binom{n+1}{k \Leftrightarrow 1}^{-1} (p_0 + p_{n+1} + p_A)^{|A|+1} \binom{|\bar{A}| \Leftrightarrow 1}{k \Leftrightarrow 2} p_{\bar{A}}^{|\bar{A}|-k+1} \tag{81}$$

where Ath term is $P(V_0(\mathcal{F}_k) = V_{n+1}(\mathcal{F}_k) = \{0\} \cup \{n+1\} \cup A)$. Similarly

$$P(0 \overset{\mathcal{F}_k}{\not\sim} n+1) = \sum_{\substack{A \subset [n] \\ |A| \le n-k+2}} \binom{n+1}{k \Leftrightarrow 1}^{-1} (p_0 + p_A)^{|A|} \binom{|\bar{A}|}{k \Leftrightarrow 2} (p_{n+1} + p_{\bar{A}})^{|\bar{A}|-k+2} \tag{82}$$

where the Ath term is $P(V_0(\mathcal{F}_k) = \{0\} \cup A)$. Another expression for the same probability is obtained by switching p_0 and p_{n+1}, since the Ath term is then $P(V_{n+1}(\mathcal{F}_k) = \{n + 1\} \cup A)$. The consequent equality of polynomials in $p_s, s \in S$ is a non-trivial identity, even in the Abel case (17). So is the equality between either of these expressions and 1 minus the right hand expression in (81), where 1 should be replaced by $(\sum_{i=0}^{n+1} p_s)^{n-k+2}$ to obtain the general polynomial identity. In a similar vein, there is the following theorem. In comparing formulae (29) and (85), note that $|S|$ here is $n+2$ instead of n.

Theorem 36 *For each probability distribution p on $[0, n+1]$ and each $1 \leq k \leq n+1$ the formula*

$$Q_{p,k}(A) := \binom{|\bar{A}|}{k \Leftrightarrow 1} \binom{n}{k \Leftrightarrow 1}^{-1} p_0(p_0 + p_A)^{|A|-1}(p_{n+1} + p_{\bar{A}})^{|\bar{A}|-(k-1)} \qquad (A \subseteq [n]) \tag{83}$$

defines a probability distribution on the set $2^{[n]}$ of all subsets of $[n]$; that is

$$\sum_{A \subseteq [n]} Q_{p,k}(A) = 1 \tag{84}$$

for all such p and k. Let $V^*(\mathcal{F}_k)$ be the random set of all $s \in [n]$ such that there is a directed path from 0 to s in $\mathcal{F}_k$ which does not pass via $n+1$, for $\mathcal{F}_k$ a p-forest of k trees labeled by $[0, n+1]$. Then

(i) *the distribution of $V^*(\mathcal{F}_k)$ given $[0 \overset{\mathcal{F}_k}{\rightsquigarrow} n+1$ or $0 \notin \mathrm{roots}(\mathcal{F}_k)]$ is $Q_{p,k}(\cdot)$;*

(ii) *the distribution of $V^*(\mathcal{F}_k)$ given $[0 \overset{\mathcal{F}_k}{\not\rightsquigarrow} n+1$ and $0 \in \mathrm{roots}(\mathcal{F}_k)]$ is $Q_{p,k-1}(\cdot)$;*

(iii) *the unconditional distribution of $V^*(\mathcal{F}_k)$ on $2^{[n]}$ is the mixture of these conditional distributions weighted by the probabilities of the conditioning events, which depend only on n and k and not on p:*

$$P(V^*(\mathcal{F}_k) = A) = \frac{n + 2 \Leftrightarrow k}{n+1} Q_{p,k}(A) + \frac{k \Leftrightarrow 1}{n+1} Q_{p,k-1}(A) \qquad (A \subseteq [n]) \tag{85}$$

Proof. The fact $Q_{p,k}$ is a probability distribution on $2^{[n]}$ is a byproduct of the assertions (i)-(iii), which are verified by application of the basic formulae (1) and (2). As a check, the identity of polynomials (8) corresponding to (84) can be verified as follows. Starting from the case $m = 0$ of (8) due to Hurwitz, replace y by $y + \theta$, expand both sides in powers of θ by the elementary binomial formula, and equate coefficients of θ^m. $\square$

Acknowledgment

Thanks to Richard Stanley for some helpful suggestions.

References

[1] N.H. Abel. Beweis eines Ausdrucks von welchem die Bonomial-Formel einer einzelner Fall ist. *Crelle's J. Reine Angew. Math.*, 1:159–160, 1826.

[2] D. Aldous and J. Pitman. Brownian bridge asymptotics for random mappings. *Random Structures and Algorithms*, 5:487–512, 1994.

[3] D.J. Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1:228–266, 1991.

[4] S. Anoulova, J. Bennies, J. Lenhard, D. Metzler, Y. Sung, and A. Weber. Six ways of looking at Burtin's lemma. Preprint, 1998.

[5] J. Arney and E.A. Bender. Random mappings with constraints on coalescence and number of origins. *Pacific J. Math.*, 103:269–294, 1982.

[6] S. Berg and J. Jaworski. Modified binomial and Poisson distributions with applications in random mapping theory. *Journal of Statistical Planning and Inference*, 18:313 – 322, 1988.

[7] S. Berg and J. Jaworski. Probability distributions related to the local structure of a random mapping. In A. Frieze and T. Luczak, editors, *Random Graphs*, volume 2, pages 1–21. Wiley, 1992.

[8] S. Berg and L. Mutafchiev. Random mappings with an attracting center: Lagrangian distributions and a regression function. *J. Appl. Probab.*, 27:622 – 636, 1990.

[9] S. Berg and K. Nowicki. Statistical inference for a class of modified power series distributions with applications to random mapping theory. *Journal of Statistical Planning and Inference*, 28:247 – 261, 1991.

[10] A. Broder. Generating random spanning trees. In *Proc. 30'th IEEE Symp. Found. Comp. Sci.*, pages 442–447, 1989.

[11] A.Z. Broder. A general expression for Abelian idenities. In L.J. Cummings, editor, *Combinatorics on words*, pages 229–245. Academic Press, New York, 1983.

[12] Y. D. Burtin. On a simple formula for random mappings and its applications. *J. Appl. Probab.*, 17:403 – 414, 1980.

[13] M. Camarri and J. Pitman. Asymptotic distributions in the generalized birthday problem. In preparation, 1997.

[14] A. Cayley. A theorem on trees. *Quarterly Journal of Pure and Applied Mathematics*, 23:376–378, 1889. (Also in *The Collected Mathematical Papers of Arthur Cayley. Vol XIII*, 26-28, Cambridge University Press, 1897).

[15] Ch. A. Charalambides. Abel series distributions with applications to fluctuations of sample functions of stochastic processes. *Communications in Statistics, Part A–Theory and Methods*, 19:317–335, 1990.

[16] L.E. Clarke. On Cayley's formula for counting trees. *J. London Math. Soc.*, 33:471–474, 1958.

[17] L. Comtet. *Advanced Combinatorics.* D. Reidel Pub. Co., Boston, 1974. (translated from French).

[18] P. C. Consul. On some properties and applications of quasi-binomial distribution. *Communications in Statistics, Part A–Theory and Methods*, 19:607, 1990.

[19] P. C. Consul and S. P. Mittal. A new urn model with predetermined strategy. *Biometrical Journal. Journal of Mathematical Methods in Biosciences.*, 17:67 – 76, 1975.

[20] P.C. Consul. A simple urn model dependent upon a predetermined strategy. *Sankhya Ser. B*, 36:391–399, 1974.

[21] C. C. Y. Dorea. Connectivity of random graphs. *J. Appl. Probab.*, 19:880 – 884, 1982.

[22] S.N. Evans and J. Pitman. Construction of Markovian coalescents. Technical Report 465, Dept. Statistics, U.C. Berkeley, 1996. Revised May 1997. To appear in *Ann. Inst. Henri Poincaré.*

[23] W. Feller. *An Introduction to Probability Theory and its Applications,* Vol 1,3rd ed. Wiley, New York, 1968.

[24] D. Foata and A. Fuchs. Réarrangements de fonctions et dénombrement. *J. Comb. Theory*, 8:361–375, 1970.

[25] J. Françon. Preuves combinatoires des identités d'Abel. *Discrete Mathematics*, 8:331–343, 1974.

[26] W. Gutjahr. Connection reliabilities in stochastic acyclic networks. *Random Structures and Algorithms*, 5:57–72, 1994.

[27] A. Hurwitz. Über Abel's Verallgemeinerung der binomischen Formel. *Acta Math.*, 26:199–203, 1902.

[28] J. Jaworski. A random bipartite mapping. *Ann. Discrete Math.*, 28:137–158, 1985.

[29] J. Jaworski. Random mappings with independent choices of images. In J. Jaworski M. Karoński and A. Ruciński, editors, *Random graphs '87*, pages 89–101, Chichester, 1990. Wiley.

[30] J.L.W. Jensen. Sur une identité d'abel et sur autres formules analogues. *Acta Math.*, 26:307–318, 1902.

[31] J.Jaworski. On a random mapping (T, P_j). *J. Appl. Probab.*, 21:186 – 191, 1984.

[32] A. Joyal. Une théorie combinatoire des séries formelles. *Adv. in Math.*, 42:1–82, 1981.

[33] I.B. Kalugin. A class of random mappings. *Proc. Steklov Inst. Math.*, 177:79–110, 1986.

[34] D.E. Knuth. Discussion on Mr. Riordan's paper. In R.C. Bose and T.A. Dowling, editors, *Combinatorial Mathematics and its Applications*, pages 71–91. Univ. of North Carolina Press, Chapel Hill, 1969.

[35] V.F. Kolchin. *Random Mappings.* Optimization Software, New York, 1986. (Translation of Russian original).

[36] E. L. Lehmann. *Theory of Point Estimation.* Wiley, New York, 1983.

[37] R. Lyons and Y. Peres. Probability on trees and networks. Book in preparation, available at http://www.ma.huji.ac.il/ lyons/prbtree.html, 1996.

[38] A. Meir and J.W. Moon. The distance between points in random trees. *J. Comb. Theory*, 8:99–103, 1970.

[39] R. Mullin and G.-C. Rota. On the foundation of combinatorial theory III: Theory of binomial enumeration. In B. Harris, editor, *Graph Theory and its Applications*, pages 167–213. Academic Press, New York, 1970.

[40] L.R. Mutafchiev. On random mappings with a single attracting centre. *J. Appl. Probab.*, 24:258 – 264, 1987.

[41] L. Mutafciev. Probability distributions and asymptotics for some characteristics of random mappings. In W. Grossman et al., editor, *Proc. 4th Pannonian Symp. on Math. Statist.*, pages 227–238, 1983.

[42] J. Pitman. Coalescent random forests. Technical Report 457, Dept. Statistics, U.C. Berkeley, 1996. To appear in *J. Comb. Theory A*. Available via http://www.stat.berkeley.edu/users/pitman.

[43] J. Pitman. The asymptotic behavior of the Hurwitz binomial distribution. Technical Report 500, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

[44] J. Pitman. The multinomial distribution on rooted labeled forests. Technical Report 499, Dept. Statistics, U.C. Berkeley, 1997. Available via http://www.stat.berkeley.edu/users/pitman.

[45] J. Pitman. Enumerations of trees and forests related to branching processes and random walks. In D. Aldous and J. Propp, editors, *Microsurveys in Discrete Probability*, number 41 in DIMACS Ser. Discrete Math. Theoret. Comp. Sci, pages 163–180, Providence RI, 1998. Amer. Math. Soc.

[46] B. Pittel. On distributions related to transitive closures of random finite mappings. *The Annals of Probability*, 11:428 – 441, 1983.

[47] H. Prüfer. Neuer Beweis eines Satzes über Permutationen. *Archiv für Mathematik und Physik*, 27:142–144, 1918.

[48] A. Rényi. On the enumeration of trees. In R. Guy, H. Hanani, N. Sauer, and J. Schonheim, editors, *Combinatorial Structures and their Applications*, pages 355–360. Gordon and Breach, New York, 1970.

[49] J. Riordan. Enumeration of linear graphs for mappings of finite sets. *Ann. Math. Stat.*, 33:178–185, 1962.

[50] J. Riordan. *Combinatorial Identities*. Wiley, New York, 1968.

[51] S. M. Ross. A random graph. *J. Appl. Probab.*, 18:309–315, 1981.

[52] L.W. Shapiro. Voting blocks, reluctant functions, and a formula of Hurwitz. *Discrete Mathematics*, 87:319–322, 1991.

[53] A.J. Stam. Two identities in the theory of polynomials of binomial type. *J. Math. Anal.*, 122:439–443, 1987.

[54] R. Stanley. Enumerative combinatorics, vol. 2. Book in preparation, to be published by Cambridge University Press, 1996.

[55] V.E. Stepanov. Random mappings with a single attracting center. *Theory Probab. Appl.*, 16:155–162, 1971.

[56] V. Strehl. Identities of the Rothe-Abel-Schläfli-Hurwitz-type. *Discrete Math.*, 99:321–340, 1992.

[57] J. Zheng. Multinomial convolution polynomials. *Discrete Math.*, 160:219–228, 1996.