ENUMERATIONS OF TREES AND FORESTS RELATED TO BRANCHING PROCESSES AND RANDOM WALKS

Jim Pitman

Technical Report No. 482

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

May 29, 1997

Abstract

In a Galton-Watson branching process with offspring distribution $(p_0, p_1, \ldots)$ started with k individuals, the distribution of the total progeny is identical to the distribution of the first passage time to $-k$ for a random walk started at 0 which takes steps of size j with probability p_{j+1} for $j \geq -1$. The formula for this distribution is a probabilistic expression of the Lagrange inversion formula for the coefficients in the power series expansion of $f(z)^k$ in terms of those of $g(z)$ for $f(z)$ defined implicitly by $f(z) = zg(f(z))$. The Lagrange inversion formula is the analytic counterpart of various enumerations of trees and forests which generalize Cayley's formula kn^{n-k-1} for the number of rooted forests labeled by a set of size n whose set of roots is a particular subset of size k. These known results are derived by elementary combinatorial methods without appeal to the Lagrange

formula, which is then obtained as a byproduct. This approach unifies and extends a number of known identities involving the distributions of various kinds of random trees and random forests.

1 Introduction

Let $(Z_0, Z_1, \ldots)$ be a *Galton-Watson branching process with offspring distribution* $(p_i, i = 0, 1, \ldots)$, where $p_i \geq 0, \sum_i p_i = 1$. Interpret p_i as the probability that an individual has i children, and Z_g as the number of individuals in the gth generation of a population starting from some initial number Z_0. For each $g \geq 0$, it is assumed that given the evolution of the population up to the gth generation, the Z_g individuals in the gth generation have independent random numbers of children distributed according to (p_i). These children are the Z_{g+1} members of the $(g+1)$th generation. See [30, 33, 5] for background. Let

$$S_n := X_1 + \cdots + X_n \tag{1}$$

be the sum of n independent random variables X_j with common distribution (p_i). From the description of $(Z_0, Z_1, \ldots)$, this sequence is a Markov chain with state space $\{0, 1, 2, \ldots\}$ and time-homogeneous transition probabilities

$$P(Z_{g+1} = m \mid Z_g = n) = P(S_n = m) \tag{2}$$

where for each n the sequence $(P(S_n = m), m = 0, 1, \ldots)$ is the n-fold convolution of the sequence $(p_i, i = 0, 1, \ldots)$ with itself. Two elementary expressions for $P(S_n = m)$ in terms of (p_i) are recalled in formulae (7) and (8) below. Let

$$\#\mathcal{F}_k := \sum_{g=0}^{\infty} Z_g \ \in \{1, 2, \ldots, \infty\} \tag{3}$$

represent the *total progeny* in the branching process given $Z_0 = k$ individuals to start with. Here $\#\mathcal{F}_k$ stands for the number of individuals in the random family $\mathcal{F}_k$ generated by the branching process. In Section 4, $\mathcal{F}_k$ will be defined as a *random family forest*, that is a collection of k *random family trees*, one for each of the k initial individuals. The following theorem presents a remarkable formula for the distribution of $\#\mathcal{F}_k$ restricted to positive integers. This formula was discovered by Otter [43] for $k = 1$ and extended to all $k \geq 1$ by Dwass [17]:

Theorem 1 (Otter-Dwass formula) *For all $n, k = 1, 2, \ldots$*

$$P(\#\mathcal{F}_k = n) = \frac{k}{n} P(S_n = n - k). \tag{4}$$

There is another setting where the same array of probabilities arises:

Theorem 2 (Kemperman's formula [34],[35, (7.15)]) *Let T_{-k} be the least $n \geq 1$ such that $S_n - n = -k$, with the convention that $T_{-k} = \infty$ is there is no such n. Then for all $1 \leq k \leq n$*

$$P(T_{-k} = n) = \frac{k}{n} P(S_n = n - k). \tag{5}$$

The consequence of (4) and (5), that $\#\mathcal{F}_k$ has the same distribution as T_{-k} for each $k = 1, 2, \ldots$, can be understood in terms of Kendall's [36] interpretation of the branching process (Z_n) and the random walk $(S_n - n)$ in terms of a queueing process, which is recalled in Section 5.

The first approach to these formulae was by the method of probability generating functions. Let

$$g(z) := \sum_{i=0}^{\infty} p_i z^i \tag{6}$$

be the probability generating function derived from the offspring distribution (p_i) Two elementary expressions for $P(S_n = m)$ are

$$P(S_n = m) \;=\; \text{coefficient of } \; z^m \text{ in } g(z)^n \tag{7}$$

$$= \sum_{\substack{\Sigma n_i = n \\ \Sigma i n_i = m}} \binom{n}{n_0, \cdots, n_n} \prod_{i \geq 0} p_i^{n_i} \tag{8}$$

where the sum ranges over the finite set of all sequences of non-negative integers $(n_i, i \geq 0)$ with $\Sigma n_i = n$ and $\Sigma i n_i = m$. Here

$$\binom{n}{n_0, \ldots, n_n} := \frac{n!}{n_0! \cdots n_n!} \tag{9}$$

is the multinomial coefficient which is the number of sequences $(i_j, 1 \leq j \leq n)$ of *type* $(n_i, i \geq 0)$, that is sequences (i_j) such that $\#\{j : i_j = i\} = n_i$ for all $i \geq 0$ and hence $\sum_j i_j = \sum_i i n_i$. Let

$$h_k(z) := \sum_n P(\#\mathcal{F}_k = n) z^n$$

be the generating function of the distribution of $\#\mathcal{F}_k$ restricted to finite values. Since the branching process started with k individuals can be constructed from k independent copies of the branching process started with one individual, the random variable $\#\mathcal{F}_k$ is distributed like the sum of k independent copies of $\#\mathcal{F}_1$. Similarly, $\#\mathcal{F}_1$ given $Z_1 = k$ has the same distribution as $1 + \#\mathcal{F}_k$. It follows that

$$h_k(z) = h(z)^k \text{ where } h(z) = zg(h(z)). \tag{10}$$

In particular, it is well known [30] that the *extinction probability* $h_k(1) = P(\#\mathcal{F}_k < \infty)$ equals q^k where q is the least non-negative root of $q = g(q)$, and that $q = 1$ or $q < 1$ according as $\mu \leq 1$ or $\mu > 1$, where $\mu := \sum_i ip_i$ is the mean of the offspring distribution of X_j, and it is assumed that $p_1 < 1$.

This method of determining the distribution of $\#\mathcal{F}_k$, which dates back to a 1944 report of Hawkins and Ulam [31], was used also by Otter [43] and Good [24]. If $\hat{h}_k(z)$ is defined in the setting of Kemperman's formula by

$$\hat{h}_k(z) := \sum_n P(T_{-k} = n) z^n$$

then the same relation (10) can be derived for $\hat{h}_k(z)$ instead of $h_k(z)$. According to a classical result of Lagrange the equation for $h(z)$ in (10) has a unique analytic solution in a neighbourhood of 0. It follows that $\hat{h}_k(z) = h_k(z) = h(z)^k$ for this unique $h(z)$. The Otter-Dwass and Kemperman formulae can now be read from the power series expansion of $h(z)^k$ provided by

Theorem 3 (Lagrange inversion formula [11]) *Let $g(z)$ be analytic in a neigbourhood of 0 with $g(0) \neq 0$. Then the equation $h(z) = zg(h(z))$ has a unique analytic solution in a neighbourhood of 0 such that*

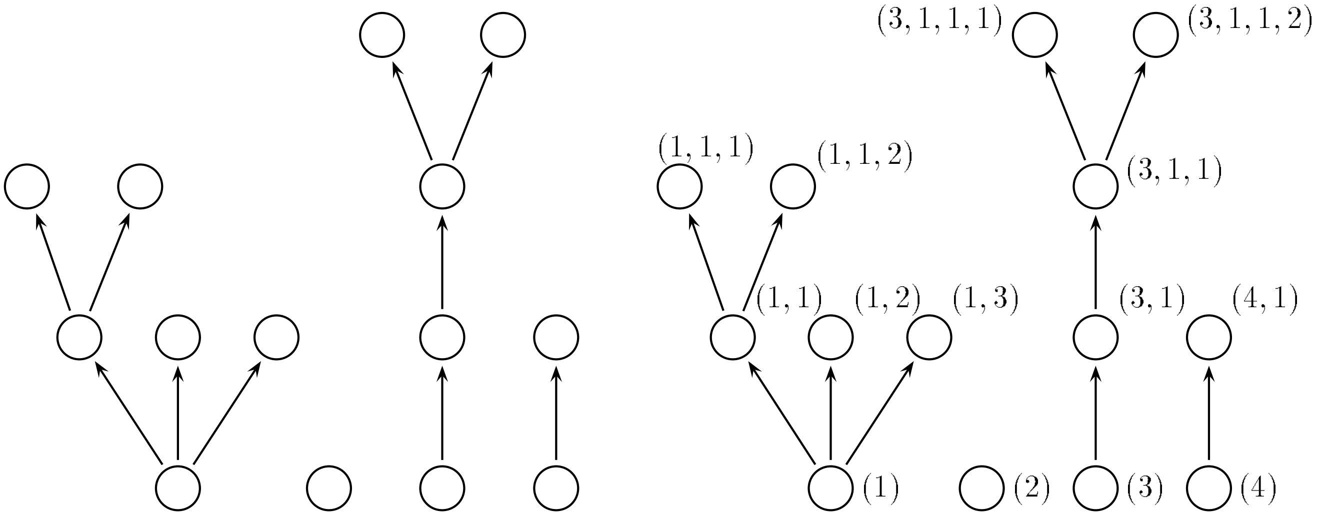
$$\text{coefficient of } z^n \text{ in } h(z)^k = \frac{k}{n}\left(\text{coefficient of } z^{n-k} \text{ in } g(z)^n\right). \tag{11}$$

While stated here in an analytic form, it is well known [59, §5.4] that the Lagrange inversion formula can be formulated as an identity of formal power series. From this perspective, it is clear that both sides of (11) are polynomials in the first $n+1$ coefficients $p_0, \ldots, p_n$ say of $g(z) := \sum_n p_n z^n$. As observed by Wendel [67], to prove (11) it is enough to establish this identity of polynomials in $p_0, \ldots, p_n$ for $p_i \geq 0$ and $\sum_{i=1}^n p_i \leq 1$, and that is precisely the content of each of the preceding probabilistic theorems. Thus the Otter-Dwass formula and Kemperman's formula are probabilistic expressions of the Lagrange inversion formula.

It is known to combinatorialists [53, 40, 10, 59] that the Lagrange inversion formula is the analytic counterpart to various enumerations of trees and forests which trace back to Cayley [8], and that these enumerations can also be interpreted, by a suitable bijection between forests and lattice paths, as enumerations of lattice paths. This paper offers an elementary approach to this circle of ideas by development of the combinatorial and probabilistic results without appeal to the Lagrange inversion formula, which is then a byproduct as just indicated.

The term *forest* will be used here for a *finite rooted forest*, that is a directed graph with a finite number of vertices, each of whose connected components is a tree with edges directed away from its root vertex. A forest with vertex set V is said to be *labeled by* V. For vertices v and w of a forest $\mathbf{f}$ write $v \xrightarrow{\mathbf{f}} w$ to show that (v, w) is a directed edge of $\mathbf{f}$. The *number of children* or *out-degree* v in the forest $\mathbf{f}$ is $C(v, \mathbf{f}) := \#\{w : v \xrightarrow{\mathbf{f}} w\}$. In a *plane forest* $\mathbf{f}$ with k component trees, the set of roots of the tree components is ordered, as is the set of children of v for each vertex v of $\mathbf{f}$. Regard a plane forest with k root vertices as a collection of family trees, one for each of k initial individuals, with each vertex in the forest corresponding to an individual, and with the order of the roots and the orders of children corresponding to the order of birth of individuals. A plane forest is often depicted without labels as on the left side of Figure 1, and called an *unlabeled plane forest.* However, there is a natural way to identify each vertex of a plane forest by a finite sequence of non-negative integers which indicates the location of the vertex in the forest. So, following the convention of [30] for labeling family trees, the set of vertices of a plane forest will be identified as a subset of the set of all finite sequences of integers, as illustrated on the right side of Figure 1.

Figure 1



An individual in the gth generation of a family forest (= vertex at height g in the plane forest) is identified by a sequence of $g+1$ integers, for instance $(3,7,4)$ to indicate a second generation individual who is the 4th child of the 7th child of the 3rd root individual.

The *type* of a forest $\mathbf{f}$ is the sequence of non-negative integers (n_i), where n_i is the number of vertices of $\mathbf{f}$ with i children. Let $1 \le k \le n$ and let (n_i) be a sequence of non-negative integers with

$$\sum_i n_i = n \text{ and } \sum_i i n_i = n-k. \tag{12}$$

A forest of type (n_i) has n vertices and $n-k$ non-root vertices, hence k root vertices and k tree components.

Theorem 4 (Enumeration of plane forests by type) [18, 19, 43, 54] *For $1 \le k \le n$ and (n_i) subject to (12) the number $N^{plane}(n_0, n_1, \dots)$ of plane forests of type (n_i) with k tree components and n vertices is*

$$N^{plane}(n_0, n_1, \dots) = \frac{k}{n}\binom{n}{n_0, \dots, n_n}. \tag{13}$$

This enumeration is due to Erdélyi and Etherington [18, 19] and Otter [43] for $k = 1$. An equivalent enumeration in terms of words instead of trees appears in Raney [53, Thm. 2.2] for $k \geq 1$. Stanley [59, Thm. 3.10 of Ch. 5] gives two proofs of (13) based on a bijection between forests and lattice paths, and an enumeration of lattice paths by consideration of cyclic shifts as in Section 5 of this paper. The corresponding result for labeled forests is:

Theorem 5 (Enumeration of labeled forests by type) [59, Cor. 3.5] *Let $[n] := \{1, \ldots, n\}$. For $1 \leq k \leq n$ and (n_i) subject to (12), the number $N^{[n]}(n_0, n_1, \ldots)$ of forests labeled by $[n]$ of type (n_i) with k tree components and n vertices is*

$$N^{[n]}(n_0, n_1, \ldots) = \frac{k}{n}\binom{n}{k}\frac{(n-k)!}{\prod_{i\geq 0}(i!)^{n_i}}\binom{n}{n_0, \ldots, n_n}. \tag{14}$$

The enumeration of plane forests by type is simpler than its companion for labeled forests. But the result for labeled forests has the following simpler equivalent:

Theorem 6 (Enumeration of labeled forests by out-degree sequence) [51],[59, Thm. 3.4] *For all sequences of non-negative integers $(c_1, \ldots, c_n)$ with $\sum_i c_i = n-k$ the number $N(c_1, \ldots, c_n)$ of forests $\mathbf{f}$ with vertex set $[n]$ in which vertex i has c_i children for each $i \in [n]$ (and hence $\mathbf{f}$ has k tree components) is*

$$N(c_1, \ldots, c_n) = \frac{k}{n}\binom{n}{k}\binom{n-k}{c_1, \ldots, c_n}. \tag{15}$$

In view of the multinomial theorem, the enumeration (15) amounts to the following identity of polynomials in n commuting variables $x_i, 1 \leq i \leq n$:

$$\sum_{\mathbf{f}\in\mathbf{F}_{k,n}}\prod_{i=1}^{n} x_i^{C(i,\mathbf{f})} = \frac{k}{n}\binom{n}{k}(x_1 + \cdots + x_n)^{n-k} \tag{16}$$

where the sum is over the set $\mathbf{F}_{k,n}$ of all forests with k tree components labeled by $[n]$, and $C(i, \mathbf{f})$ is the number of children of i in the forest $\mathbf{f}$. Take the x_i to be identically 1 in (16) to recover the well known enumeration

$$\#\mathbf{F}_{k,n} = k\binom{n}{k}n^{n-k-1} \tag{17}$$

which is equivalent to Cayley's [8] formula

$$\#\{\text{forests with root set } [k] \text{ and vertex set } [n]\ \} = kn^{n-k-1}. \qquad (18)$$

In particular, for $k = 1$ the number of rooted trees labeled by $[n]$ is n^{n-1}. Equivalently, the number of unrooted trees labeled by $[n]$ is n^{n-2}. For various other approaches to these formulae of Cayley, see [42, 51, 52, 57, 59, 63]. The multinomial expansion over rooted trees obtained from the case $k = 1$ of (16) is a variant of the multinomial expansion over unrooted trees indicated by Cayley [8] for small n and formulated and proved for all n by Rényi [55]. But throughout this paper, all trees and forests are assumed to be rooted.

The rest of this paper is organized as follows. Section 2 offers a simple proof of Theorem 6. The equivalence of Theorems 4, 5 and 6 is argued in Section 3. The equivalence of Theorem 4 and the Otter-Dwass formula is explained in Section 4. (See also Kolchin [Chapter 2.1, Lemma 3][39] for another proof of the Otter-Dwass formula, by showing that both sides satisfy a recursion which has a unique solution.) Section 5 reviews the connection between the Otter-Dwass formula for branching processes and Kemperman's formula for random walks via the interpretation of both processes in terms of queues pointed out by Kendall [36]. As shown by Takács [60, 61], Kemperman's formula is both generalized and simplified by consideration of random walks of a fixed length n whose distribution of steps is invariant under cyclic shifts. The ubiquitous factor of k/n in each of the six theorems above finds its most intuitive explanation in this context: the k/n in Kemperman's formula represents the conditional probability of the event $(T_{-k} = n)$ given $(S_n - n = -k) \cap A$ for any event A determined by the first n increments of the walk $(S_n - n)$ such that A is invariant under cyclic shifts of these increments. Section 6 reviews the well known representation of the uniform distribution on plane forests with k trees and n vertices as the distribution of a Galton-Watson forest of k trees with geometric offspring distribution conditioned on a total progeny of n. Section 7 presents a similar result which explains a number of known identities relating the uniform distribution on the set of forests of k trees labeled by $[n]$ to the distribution of the random forest generated a Galton-Watson branching process with a Poisson offspring distribution.

2 Enumeration of labeled forests by out-degree sequence.

Proof of Theorem 6. For a forest $\mathbf{f}$ with vertex set $[n]$ and $i \in [n]$ let $J_i(\mathbf{f}) := \{j \in [n] : i \stackrel{\mathbf{f}}{\to} j\}$ be the set of children of i in $\mathbf{f}$. So $\mathbf{f}$ is determined by the sequence of disjoint sets $J_1(\mathbf{f}), \dots, J_n(\mathbf{f})$, and vice-versa. Given a sequence of disjoint subsets $J_1, \dots, J_n$ of $[n]$, for each $m \in [n]$ let $\mathbf{f}_m$ be the relation on $[m] \cup (\cup_{i=1}^m J_i)$ defined by

$$i \stackrel{\mathbf{f}_m}{\to} j \text{ iff } i \in [m] \text{ and } j \in J_i. \tag{19}$$

There exists a forest $\mathbf{f}$ labeled by $[n]$ such that $J_i(\mathbf{f}) = J_i$ for all $i \in [n]$ if and only if the J_i are such that $\mathbf{f}_m$ defined by (19) is a forest with vertex set $[m] \cup (\cup_{i=1}^m J_i)$ for every $m \in [n]$; then $\mathbf{f} = \mathbf{f}_n$ and $\mathbf{f}_m$ is the restriction of $\mathbf{f}$ to $[m] \cup (\cup_{i=1}^m J_i)$ for each $m \in [n]$. It follows that for each sequence of non-negative integers (c_i) with $\sum_i c_i = n - k$, the number

$$N(c_1, \dots, c_n) := \#\{\mathbf{f} \in \mathbf{F}_{k,n} : \#J_i(\mathbf{f}) = c_i \text{ for all } i \in [n]\} \tag{20}$$

is the number of ways to choose a sequence of subsets $(J_1, \dots, J_n)$ of $[n]$ such that $\#J_m = c_m$ and the relation $\mathbf{f}_m$ defined by (19) is a forest with vertex set $[m] \cup (\cup_{i=1}^m J_i)$ for every $m \in [n]$. Clearly, J_1 can be any of the $\binom{n-1}{c_1}$ subsets of $[n] - \{1\}$ of size c_1. For $m \in [n-1]$ make the inductive hypothesis that sets $J_1, \dots, J_m$ of sizes $c_1, \dots, c_m$ have been chosen so that $\mathbf{f}_m$ is a forest. Which choices of J_{m+1} of size c_{m+1} make $\mathbf{f}_{m+1}$ a forest? There are two cases to consider. Either
(i) $m+1 \notin \cup_{i=1}^m J_i$: then J_{m+1} can be any subset of $[n] - (\cup_{i=1}^m J_i) - \{m+1\}$; or
(ii) $m+1 \in \cup_{i=1}^m J_i$: then J_{m+1} can be any subset of $[n] - (\cup_{i=1}^m J_i) - \{r_m\}$ where $r_m \notin \cup_{i=1}^m J_i$ is the root of the tree component of $\mathbf{f}_m$ which contains $m+1$. Either way, and regardless of what sets $J_1, \dots, J_m$ of sizes $c_1, \dots, c_m$ were previously chosen to make $\mathbf{f}_m$ a forest, the number of possible choices of J_{m+1} which make $\mathbf{f}_{m+1}$ a forest is

$$\binom{n - \sum_{i=1}^m c_i - 1}{c_{m+1}}.$$

Consequently, by induction

$$N(c_1, \dots, c_n) = \binom{n-1}{c_1}\binom{n-c_1-1}{c_2}\cdots\binom{n-\sum_{i=1}^{n-1} c_i - 1}{c_n}$$

and this expression simplifies easily to yield (15). □

3 Enumeration of forests by type

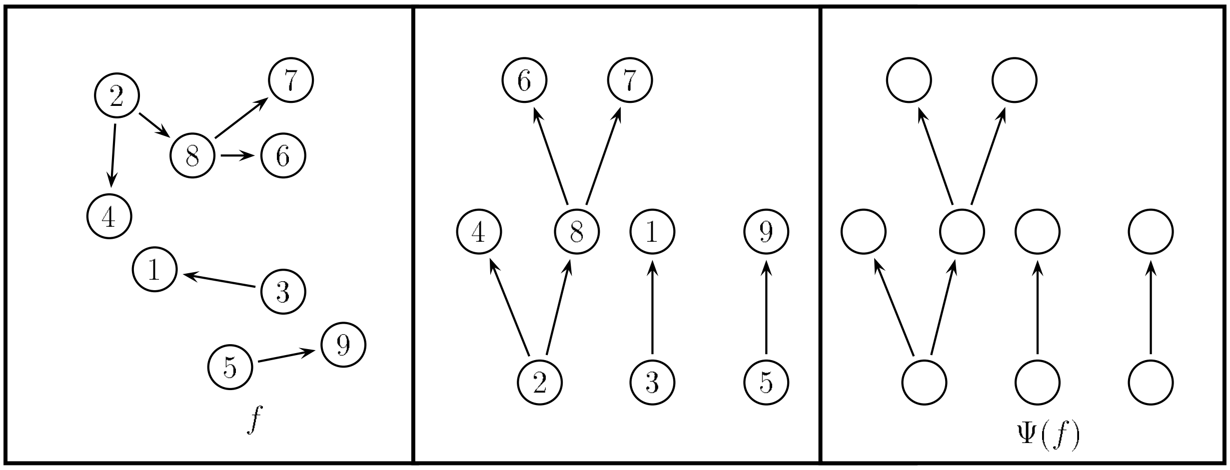
Proof of Theorem 5. Let (n_i) subject to (12) be a possible type sequence for a forest of k trees with n vertices. According Theorem 6, for each particular sequence (c_j) of type (n_i) the number of forests in which j has c_j children for all $j \in [n]$ is

$$\frac{k}{n}\binom{n}{k}\frac{(n-k)!}{c_1!\dots c_n!} = \frac{k}{n}\binom{n}{k}\frac{(n-k)!}{\prod_{i\geq 0}(i!)^{n_i}}. \tag{21}$$

But the number of sequences of non-negative integers (c_j) such that (c_j) has type (n_i) is just the multinomial coefficient (9). The number of forests labeled by $[n]$ of type (n_i) is the product of the number in (21) and this multinomial coefficient. Thus Theorem 6 implies Theorem 5, and vice versa. □

Proof of Theorem 4. The enumeration (13) for plane forests corresponds to the enumeration (14) for labeled forests via the following map Ψ, which transforms a forest $\mathbf{f}$ labeled by $[n]$ into a plane forest $\Psi(\mathbf{f})$ of the same type. To define $\Psi(\mathbf{f})$, first place the components of $\mathbf{f}$ in order of their root labels, then recursively for each $g \geq 0$ place the children of each vertex of $\mathbf{f}$ at height g in order of their labels. Finally, delabel to obtain the plane forest $\Psi(\mathbf{f})$. See Figure 2.

Figure 2



For a fixed plane forest $\mathbf{f}^\circ$ with n vertices and k root vertices, and type (n_i), let $\{v_1, \ldots, v_n\}$ be a listing of the vertices in some arbitrary order. Let B_0 be the set of roots of $\mathbf{f}^\circ$, and B_i the set of children of v_i for $i = 1, \ldots, n$. The sets $B_i, 0 \leq i \leq n$, some of which must be empty, are disjoint with union $V(\mathbf{f}^\circ)$, the set of vertices of $\mathbf{f}^\circ$. As discussed earlier, $V(\mathbf{f}^\circ)$ is regarded as a subset of the set of finite sequences of positive integers. Each non-empty B_i has a linear ordering, and each $\mathbf{f} \in \Psi^{-1}(\mathbf{f}^\circ)$ corresponds to a unique bijection from $V(\mathbf{f}^\circ)$ to $[n]$ which is increasing on each non-empty B_i. It follows that $\#\Psi^{-1}(\mathbf{f}^\circ)$, the number of forests $\mathbf{f}$ labeled by $[n]$ such that $\Psi(\mathbf{f}) = \mathbf{f}^\circ$ is the multinomial coefficient

$$\#\Psi^{-1}(\mathbf{f}^\circ) = \binom{n!}{k, \#B_1, \ldots, \#B_n} = \frac{n!}{k! \prod_{v \in V(\mathbf{f}^\circ)} C(v, \mathbf{f}^\circ)!} = \frac{n!}{k! \prod_{i \geq 0} (i!)^{n_i}} \qquad (22)$$

where $C(v, \mathbf{f}^\circ)$ is the number of children of v in the forest $\mathbf{f}^\circ$, and (n_i) is the type of $\mathbf{f}^\circ$ and of every $\mathbf{f} \in \Psi^{-1}(\mathbf{f}^\circ)$. It follows that the number $N^{plane}(n_0, n_1, \ldots)$ of plane forests of type (n_i) and the corresponding number $N^{[n]}(n_0, n_1, \ldots)$ of forests of type (n_i) labeled by $[n]$ are related by

$$\frac{N^{[n]}(n_0, n_1, \ldots)}{N^{plane}(n_0, n_1, \ldots)} = \frac{n!}{k! \prod_{i \geq 0} (i!)^{n_i}}. \qquad (23)$$

Hence the equivalence of Theorems 4 and 5. $\square$

4 The Otter-Dwass formula

Proof of the equivalence of Theorems 1 and 4. Following Otter [43] and subsequent authors [30, 37, 13], regard a Galton-Watson process started with $Z_0 = k$ individuals as generating a collection of k family trees, which combine to form a *random family forest* $\mathcal{F}_k$. On the event $(\#\mathcal{F}_k < \infty)$ the random family forest $\mathcal{F}_k$ can be defined in an elementary way as a random element with values in the countable set $\mathbf{F}$ of all plane forests. The *distribution of* $\mathcal{F}_k$ is then the sub-probability distribution on $\mathbf{F}$ defined by the formula [43]

$$P(\mathcal{F}_k = \mathbf{f}) = \prod_{v \in V(\mathbf{f})} p_{C(v,\mathbf{f})} = \prod_{i \geq 0} p_i^{n_i(\mathbf{f})} \qquad \forall k \geq 1,\ \mathbf{f} \in \mathbf{F}_k^{plane} \tag{24}$$

where $V(\mathbf{f})$ is the set of vertices of $\mathbf{f}$, the number of vertices of $\mathbf{f}$ with i children is denoted $n_i(\mathbf{f})$, and $\mathbf{F}_k^{plane}$ is the set of plane forests with k root vertices. This distribution on $\mathbf{F}_k^{plane}$ has total mass $P(\#\mathcal{F}_k < \infty) \leq 1$. For each $k = 1, 2, \ldots$ the probability of the event $(\#\mathcal{F}_k = n)$ is obtained by summing the expression (24) over all plane forests $\mathbf{f}$ of k trees with a total of n vertices. The terms in this sum can be classified by the type (n_i) of the forest $\mathbf{f}$. Since the number of vertices in a forest of type (n_i) is $\sum_i n_i$ and the number of root vertices is $n - \sum_i n_i$, the result is

$$P(\#\mathcal{F}_k = \mathbf{n}) = \sum_{\substack{\Sigma n_i = n \\ \Sigma i n_i = n-k}} N^{plane}(n_0, n_1, \ldots) \prod_{i \geq 0} p_i^{n_i} \tag{25}$$

where the sum is over all possible types of a sequence with length n and sum $n - k$. Now fix n and k and regard the probability displayed in (25), and the probability $P(S_n = m)$ displayed in (8) for $m = n - k$, as functions of the sequence $(p_0, \ldots, p_n)$. Since each probability is a polynomial in $(p_0, \ldots, p_n)$, the Otter-Dwass formula (4) is an identity of polynomials whose coefficient identity is the enumeration (13) of plane forests by type. □

Lagrangian distributions. Recall that $h(z)$ determined by the offspring generating function $g(z)$ via (10) or (11) is the probability generating function of the total progeny of a Galton-Watson branching process started with one individual. For Z_0 with an arbitrary distribution with probability generating function $f(z) := \sum_n P(Z_0 = n) z^n$ it is evident by conditioning on Z_0 that the unconditional distribution of the total progeny is that determined by the generating function $f(h(z))$. This distribution is known as the *Lagrangian distribution* derived from the distribution of Z_0 and the offspring distribution. See [16, 17, 44, 9, 58] regarding Lagrangian distributions and their applications, and [25, 26, 27, 23] for the multivariate extension of Lagrange's expansion and its relation to multi-type branching processes.

5 Random walks, queues, and branching processes.

The following interpretation of the random walk $(S_n - n)$ and the branching process (Z_n) in terms of queuing theory is due to Kendall [36]. See also [25, 60, 62, 64, 22]. Suppose customers arrive and wait for service in a queue with a single server. The time the server is working consists of alternating busy and idle periods, each busy period consisting of one or more service periods, one per customer. Let X_j denote the number of customers arriving during the jth service period, so S_n represents the number of customers arriving by the end of the nth service period. Suppose at time zero there are $k \geq 1$ customers already present in the queue. For $0 \leq n \leq T_{-k}$ the number of customers in the queue just after the end of the end of the nth service period is $k + S_n - n$. So T_{-k} represents the number of customers served during the first busy period, and $P(T_{-k} = n)$ is the probability that the server has to deal with n customers before taking a break, given k customers in the queue to start with. The queuing process defines a random family $\mathcal{F}_k$, starting with k individuals, such that $\#\mathcal{F}_k = T_{-k}$. Each individual j in $\mathcal{F}_k$ represents a different customer, with j' the child of j if j' arrives during the service period of j. Assuming that the numbers X_j are independent with common distribution (p_i), the random family derived from the queue defines a Galton-Watson branching process (Z_n) with offspring distribution (p_i). This argument explains why the total progeny of the branching process (Z_n) started with k individuals has the same distribution as the first passage time T_{-k} of the random walk $(S_n - n)$. The argument can be made more precise by setting up an appropriate bijection between the following two sets: the set of walk paths starting at $(0, 0)$ and first reaching $-k$ at time n by a sequence of integer increments with no increment less than -1, and the set of plane forests of k trees and n vertices. See [59, 22] for details and further developments.

In the setting of Theorem 2, by definition

$$(T_{-k} = n) := (\forall_{m=1}^{n-1} S_m - m \neq -k, S_n - n = -k) \subseteq (S_n - n = -k) \quad (26)$$

where $S_n - n = Y_1 + \cdots + Y_n$ for $Y_j = X_j - 1$, derived from independent X_j with common distribution (p_i). So Kemperman's formula (5) can be restated as follows:

$$P(T_{-k} = n \,|\, S_n - n = -k) = \frac{k}{n}. \quad (27)$$

That is to say, given that the random walk $(S_m - m)$ is at $-k$ at time $m = n$, the chance that the walk first reached $-k$ at time n is k/n. The work of Takács [60, 61] shows that this form of Kemperman's result can be generalized as follows. See [60, Thm 1 of §4 and Thm 5 of §28] for two other essentially equivalent formulations related to the classical ballot theorem. The basic idea of considering cyclic shifts traces back to Dvoretsky and Motzkin [14]. See also [21, 20, 15, 12] for closely related results and further references.

For a sequence of integers $\mathbf{y} := (y_1, \ldots, y_n)$ and $i \in [n]$ let $\mathbf{y}^{(i)}$ denote the ith *cyclic shift* of $\mathbf{y}$, that is the sequence whose jth term is y_{i+j} with addition modulo n. Call a set of sequences A *cyclically invariant* iff $\mathbf{y} \in A$ implies $\mathbf{y}^{(1)} \in A$, in which case $\mathbf{y}^{(i)} \in A$ for all $i \in [n]$. Call a sequence of random variables $\mathbf{Y}_n := (Y_1, \ldots, Y_n)$ *cyclically exchangeable* if $(Y_2, \ldots, Y_n, Y_1)$ has the same distribution as $(Y_1, \ldots, Y_n)$. For integer valued Y_i this is equivalent to

$$P(\mathbf{Y}_n = \mathbf{y}) = P(\mathbf{Y}_n = \mathbf{y}^{(i)}) \quad (28)$$

for each sequence of integers $\mathbf{y}$ and all $i \in [n]$.

Theorem 7 (Takács [60, 61]) *Suppose that* $\mathbf{Y}_n := (Y_1, \ldots, Y_n)$ *is a cyclically exchangeable sequence of random variables, let* $S_m^- = Y_1 + \cdots + Y_m$*, and let* $(T_{-k} = n)$ *be the event that the walk* (S_m^-) *first reaches* $-k$ *at time* n*. Let* $\mathbb{N}_- := \{-1, 0, 1, 2, \ldots\}$*. Then for every cyclically invariant subset* A *of* $\mathbb{N}_-^n$*,*

$$P(T_{-k} = n \,|\, S_n^- = -k, \mathbf{Y}_n \in A) = \frac{k}{n}. \quad (29)$$

Proof. Every cyclically invariant A decomposes as the union of some collection of cyclic orbits, where for a sequence $\mathbf{y}$ the *cyclic orbit of* $\mathbf{y}$ is the set $A_{\mathbf{y}} := \{\mathbf{z} : \mathbf{z} = \mathbf{y}^{(i)} \text{ for some } i \in [n]\}$. So it suffices to prove (29) in the case $A = A_{\mathbf{y}}$ for an arbitrary $\mathbf{y} \in \mathbb{N}_-^n$. This special case is a consequence of following elementary lemma. □

For a sequence $\mathbf{y}$, let $t_m = y_1 + \cdots + y_m$ and call the sequence of partial sums $(t_1, \ldots, t_n)$ the *walk with steps* $\mathbf{y}$. Say the walk *first reaches b at time n* if $t_i \neq b$ for $i < n$ and $t_n = b$.

Lemma 8 (Wendel [67, §3]) *Let $\mathbf{y} \in \mathbb{N}_-^n$ be such that $y_1 + \cdots + y_n = -k$ for some $1 \leq k \leq n$, and let $\mathbf{y}^{(i)}$ be the ith cyclic shift of $\mathbf{y}$. Then there are exactly k distinct $i \in [n]$ such that the walk with steps $\mathbf{y}^{(i)}$ first reaches $-k$ at time n.*

For different formulations of the lemma, which show how it generalizes the classical ballot theorem, see [60, Thm. 3 and Thm. 4 of §2]. While Theorem 7 was stated in a probabilistic way, the lemma reveals its combinatorial essence. Thus Theorem 7 can be formulated in purely combinatorial terms as follows:

Corollary 9 *Let A be a cyclically invariant subset of $\mathbb{N}_-^n$ such that $y_1 + \cdots + y_n = -k$ for every $\mathbf{y} \in A$. Let A^* be the set of all $\mathbf{y} \in A$ such that the walk with steps $\mathbf{y}$ first reaches $-k$ at time n. Then the fraction of elements of A that are elements of A^* equals k/n.*

To illustrate the corollary, let $I := \{a, a+1, \ldots, b\}$ be a finite set of consecutive integers with $a \leq -1$. Given a sequence $(n_i, i \in I)$ of non-negative integers with $\sum_i n_i = n$ and $\sum_i i n_i = -k$, consider the set A of all sequences $\mathbf{y} \in I^n$ of type (n_i), so

$$\#A = \binom{n}{n_a, \ldots, n_b}. \tag{30}$$

For $a \leq -2$ there is no simple formula for $\#A^*$, the number of sequences $\mathbf{y}$ of type (n_i) such that the walk with steps $\mathbf{y}$ first reaches $-k$ at time n. But in the special case $a = -1$ Corollary 9 implies

$$\#A^* = \frac{k}{n}\binom{n}{n_{-1}, \ldots, n_b}. \tag{31}$$

Implicit in the previous description of the branching process derived from a queuing process is a bijection between the set of walks with step sequence in A^*, as enumerated by (31), and the set of plane forests of k trees in which n_{i+1} vertices have i children, as enumerated by Theorem 4.

6 Uniform random plane forests

For $1 \leq k \leq n$ let $\mathbf{F}_{k,n}^{plane}$ denote the set of all plane forests of k trees with a total of n vertices, and let $\mathcal{F}_{k,n}^{plane}$ be a uniformly distributed random element of $\mathbf{F}_{k,n}^{plane}$. For $k = 1, 2, \ldots$ and $0 < p < 1$ let $\mathcal{G}_{k,p}$ be a Galton-Watson forest of k trees with the geometric(p) offspring distribution $p_i := p(1-p)^i$. The general product formula (24) gives

$$P(\mathcal{G}_{k,p} = \mathbf{f}) = \prod_{v \in V(\mathbf{f})} p(1-p)^{C(v,\mathbf{f})} = p^n (1-p)^{n-k} \quad \forall\, \mathbf{f} \in \mathbf{F}_{k,n}^{plane}. \tag{32}$$

Since this probability is the same for all $\mathbf{f} \in \mathbf{F}_{k,n}^{plane}$, the distribution of $\mathcal{G}_{k,p}$ given $(\#\mathcal{G}_{k,p} = n)$ is uniform on $\mathbf{F}_{k,n}^{plane}$, as observed in [29, 48]. Symbolically

$$\mathcal{F}_{k,n}^{plane} \stackrel{d}{=} (\mathcal{G}_{k,p} \,|\, \#\mathcal{G}_{k,p} = n) \tag{33}$$

where $\stackrel{d}{=}$ denotes equality of distributions. The Otter-Dwass formula implies

$$P(\#\mathcal{G}_{k,p} = n) = \frac{k}{n} P(S_{n,p} = n-k) = \frac{k}{n} \binom{2n-k-1}{n-k} p^n (1-p)^{n-k} \tag{34}$$

where $S_{n,p}$ is the sum of n independent geometric random variables, and the second equality in (34) is read from the negative binomial formula [21, VI.8] for the distribution of $S_{n,p}$. Compare (32) and (34) to obtain

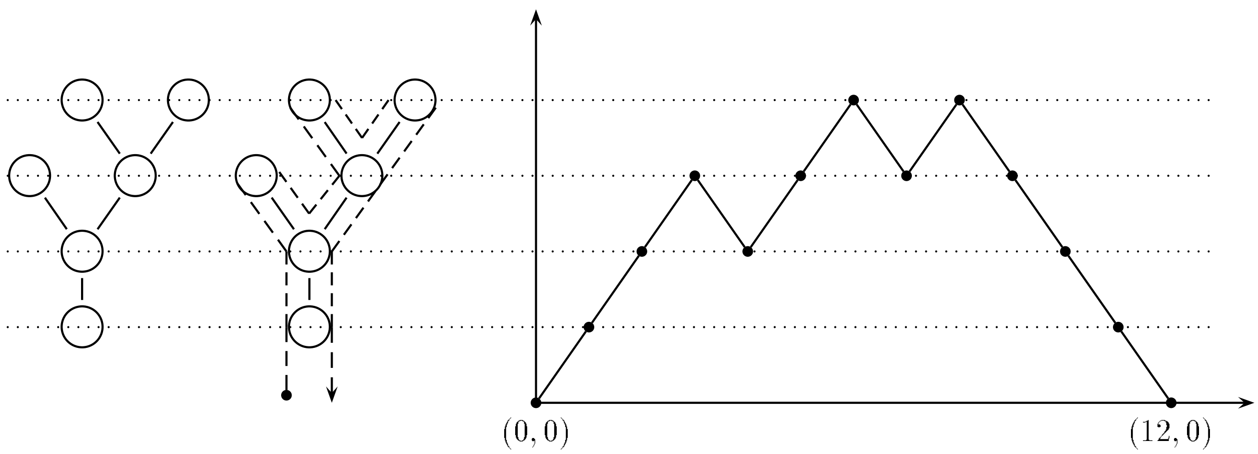
$$\#\mathbf{F}_{k,n}^{plane} = \frac{k}{n} \binom{2n-k-1}{n-k}. \tag{35}$$

This argument simplifies and corrects a similar derivation of Pavlov [48]. In particular the number of plane trees with n vertices is

$$\#\mathbf{F}_{1,n}^{plane} = \frac{1}{n} \binom{2n-2}{n-1}$$

which is the $(n-1)$th Catalan number [7, 32]. This is also the number of *lattice excursions of length* $2n$, that is sequences $(s_j, 0 \leq j \leq 2n)$ where $s_0 = s_{2n} = 0$, and $s_j > 0$ and $s_{j+1} - s_j \in \{-1, +1\}$ for all $0 \leq j \leq 2n-1$. As observed by Harris [29], there is a natural bijection between $\mathbf{F}_{1,n}^{plane}$ and the set of lattice excursions of length $2n$. See Figure 3.

Figure 3



This bijection extends to a bijection between $\mathbf{F}_{k,n}^{plane}$ and the set of non-negative lattice walk paths from $(0,0)$ to $(0,2n)$ with increments of ± 1 and exactly k returns to 0. By another bijection, the number of plane forests with n vertices equals the number of plane trees with $n+1$ vertices:

$$\# \bigcup_{k=1}^{n} \mathbf{F}_{k,n}^{plane} = \# \mathbf{F}_{1,n+1}^{plane}.$$

In view of (35) this yields the identity

$$\sum_{k=1}^{n} \frac{k}{n} \binom{2n-k-1}{n-k} = \frac{1}{n+1} \binom{2n}{n}. \tag{36}$$

The sum is the nth Catalan number, which is the the number of non-negative lattice walk paths from $(0,0)$ to $(0,2n)$ with increments of ± 1. The kth term of the sum is the number of such paths with k returns to zero.

Similarly, the number of ways to pick a plane forest with n vertices and assign each tree component a sign ± 1 equals the number of lattice paths from $(0,0)$ to $(0,2n)$ with increments of ± 1, that is

$$\sum_{k=1}^{n} \frac{k}{n} \binom{2n-k-1}{n-k} 2^k = \binom{2n}{n}. \tag{37}$$

In agreement with the result of Feller [21, Thm. 2 of §III.7], the kth term in the sum (37) divided by 2^{2n} is the probability that a simple symmetric random walk, started at 0 and moving with increments of ± 1, returns to 0 for the kth time after $2n$ steps.

The bijection between trees and lattice excursions implies that the large n asymptotic distribution of many functionals of a random tree $\mathcal{T}_n$ with uniform distribution on $\mathbf{F}_{1,n}^{plane}$ can be read from the asymptotic distribution of a functional of a uniformly distributed random lattice excursion of length $2n$, which is typically the distribution of a corresponding functional of a Brownian excursion [2, 13]. As shown by Aldous [2, 3], the same holds for any random plane tree $\mathcal{T}_n$ with $\mathcal{T}_n$ distributed like $\mathcal{T}$ given $\#\mathcal{T} = n$ where $\mathcal{T}$ is a Galton-Watson tree whose offspring distribution has mean 1 and finite variance. In particular, due to the result of the next section, this conclusion applies to $\mathcal{T}_n$ derived from a random tree with uniform distribution on the set $\mathbf{T}_n$ of all n^{n-1} rooted trees labeled by $[n]$. The general class of distributions for a planar tree of size n obtained by conditioning a Galton-Watson tree to be of size n is the class of distributions of "simply generated trees" studied by Meir and Moon [41]. See [2, 3, 1, 13, 28] for further developments.

7 The plane forest derived from a uniform labeled forest

Recall from around (22) the map $\Psi : \mathbf{F}_{k,n} \to \mathbf{F}_{k,n}^{plane}$, where $\mathbf{F}_{k,n}$ is the set of forests with k tree components labeled by $[n]$, and $\mathbf{F}_{k,n}^{plane}$ is the set of plane forests with k tree components and n vertices. For a lighter notation, write $\mathbf{f}^\circ$ instead of $\Psi(\mathbf{f})$, and call $\mathbf{f}^\circ$ the *plane forest derived from* $\mathbf{f}$. So $\mathbf{f}^\circ$ is just $\mathbf{f}$ regarded as a plane forest by giving the set of roots of $\mathbf{f}$ and the sets of children of various vertices of $\mathbf{f}$ the order these sets acquire from the usual ordering of $[n]$. The following theorem strengthens connections discovered Kolchin [38] and Pavlov[46, 47] between the uniform distribution on $\mathbf{F}_{k,n}$ and the distribution of a Galton-Watson forest with the Poisson(μ) offspring distribution $p_i := e^{-\mu}\mu^i/i!$. Kolchin and Pavlov [38, 39, 46, 45, 47, 50, 48, 49] exploited these connections to derive the asymptotic distributions of functionals of a uniform random forest of k trees labeled by $[n]$, such as the numbers of trees of various sizes and the maximum tree size, as $n \to \infty$

for various ranges of k. The case $k = 1$ of the theorem is implicit in the discussion of Aldous [2, 3].

Theorem 10 *For $\mu \in (0, \infty)$ let $\mathcal{P}_{k,\mu}$ be a Galton-Watson forest with the Poisson(μ) offspring distribution, for $1 \le k \le n$ let $\mathcal{F}_{k,n}$ have uniform distribution on the set of forests of k trees labeled by $[n]$, and let $\mathcal{F}^{\circ}_{k,n}$ be $\mathcal{F}_{k,n}$ regarded as a plane forest. Then $\mathcal{F}^{\circ}_{k,n}$ has the same distribution as $\mathcal{P}_{k,\mu}$ given $(\#\mathcal{P}_{k,\mu} = n)$:*

$$\mathcal{F}^{\circ}_{k,n} \stackrel{d}{=} (\mathcal{P}_{k,\mu} \,|\, \#\mathcal{P}_{k,\mu} = n). \tag{38}$$

Proof. To be more explicit, there is the following formula. For all plane forests $\mathbf{f}$ of k trees with n vertices

$$P(\mathcal{F}^{\circ}_{k,n} = \mathbf{f}) = P(\mathcal{P}_{k,\mu} = \mathbf{f} \,|\, \#\mathcal{P}_{k,\mu} = n) = \frac{n(n-k)!}{kn^{n-k}} \prod_{v \in V(\mathbf{f})} \frac{1}{C(v, \mathbf{f})!}. \tag{39}$$

The first probability in (39) is the number displayed in (22) divided by $\#\mathbf{F}_{k,n}$ in (17), which reduces to the last expression in (39) by cancellation. The second probability reduces similarly, by application of (24) and the consequence of the Otter-Dwass formula (4) that the total progeny in a Poisson-Galton-Watson family forest of k trees has the distribution

$$P(\#\mathcal{P}_{k,\mu} = n) = \frac{k}{n} \frac{(\mu n)^{n-k}}{(n-k)!} e^{-\mu n} \qquad (n = k, k+1, \ldots) \tag{40}$$

known as the *Borel-Tanner distribution* [6, 43, 65, 66]. □

Call a function Φ of forests $\mathbf{f}$ an *invariant* if $\Phi(\mathbf{f}) = \Phi(\mathbf{f}')$ whenever $\mathbf{f}'$ is a relabeling of $\mathbf{f}$, meaning $v \xrightarrow{\mathbf{f}'} w$ iff $\ell(v) \xrightarrow{\mathbf{f}} \ell(w)$ for some bijection ℓ from the vertices of $\mathbf{f}$ to the vertices of $\mathbf{f}'$. For example, the number $Z_h\mathbf{f}$ of vertices of $\mathbf{f}$ at height h is an invariant. So is the matrix $M(\mathbf{f}) := (M_{h,c}(\mathbf{f}), h \ge 0, c \ge 0)$ where $M_{h,c}(\mathbf{f})$ is the number of individuals in generation h of $\mathbf{f}$ that have c children. Since the plane forest $\mathcal{F}^{\circ}_{k,n}$ is by definition a relabeling of the uniform labeled forest $\mathcal{F}_{k,n}$, the identity (38) implies $\mathcal{F}_{k,n}$

$$\Phi(\mathcal{F}_{k,n}) \stackrel{d}{=} (\Phi(\mathcal{P}_{k,\mu}) \,|\, \#\mathcal{P}_{k,\mu} = n) \qquad \forall \text{ invariant } \Phi. \tag{41}$$

For $\Phi = M$ this result is due to Kolchin [38] and Pavlov [47]. The identity (41) is expressed more intuitively by the following construction, suggested by Aldous [2] for $k = 1$. Fix $\mu > 0$ and generate a Poisson-Galton-Watson family forest $\mathcal{P}_{k,\mu}$ starting from k root individuals. Given that $\mathcal{P}_{k,\mu}$ has vertex set V with $\#V = n$, let $\mathcal{P}^*_{k,\mu} \in \mathbf{F}_{k,n}$ be $\mathcal{P}_{k,\mu}$ relabeled by a uniform random permutation $\sigma : V \to [n]$. Then

$$\mathcal{F}_{k,n} \stackrel{d}{=} (\mathcal{P}^*_{k,\mu} \,|\, \#\mathcal{P}_{k,\mu} = n). \tag{42}$$

That is, given that $\mathcal{P}_{k,\mu}$ has n vertices, a random relabeling of $\mathcal{P}_{k,\mu}$ has uniform distribution over the set of all forests of k trees labeled by $[n]$. For some recent applications of this relation between uniform random trees and Poisson-Galton-Watson trees see [51, 56, 4].

Acknowledgments

Thanks to Vlada Limic for preparation of the graphics, to Richard Stanley for providing a draft of [59], to Anthony Pakes for help with the literature, and to Steven Evans for comments on a preliminary draft.

References

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

[2] D.J. Aldous. The continuum random tree I. *Ann. Probab.*, 19:1–28, 1991.

[3] D.J. Aldous. The continuum random tree II: an overview. In M.T. Barlow and N.H. Bingham, editors, *Stochastic Analysis*, pages 23–70. Cambridge University Press, 1991.

[4] D.J. Aldous and J. Pitman. Tree-valued Markov chains derived from Galton-Watson processes. Technical Report 481, Dept. Statistics, U.C. Berkeley, 1997.

[5] K.B. Athreya and P. Ney. *Branching Processes*. Springer, 1972.

[6] E. Borel. Sur l'emploi du théorème de Bernoulli pour faciliter le calcul d'un infinité de coefficients. Application au probleme de l'attente á un guichet. *C.R. Acad. Sci. Paris*, 214:452–456, 1942.

[7] W.G. Brown. Historical note on a recurrent combinatorial problem. *Amer. Math. Monthly*, 72:973–977, 1965.

[8] 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).

[9] P.C. Consul. *Generalized Poisson Distributions.* Dekker, 1989.

[10] R. Cori. Words and Trees. In M. Lothaire, editor, *Combinatorics on Words*, volume 17 of *Encyclopedia of Mathematics and its Applications*, pages 215–229. Addison-Wesley, Reading, Mass., 1983.

[11] L. de Lagrange. Nouvelle méthode pour résoudre des 'equations littérales par le moyen des séries. *Mém. Acad. Roy. Sci. Belles-Lettres de Berlin*, 24, 1770.

[12] N. Dershowitz and S. Zaks. The cycle lemma and its applications. *Europ. J. Combinatorics*, 11:35–40, 1990.

[13] R. Durrett, H. Kesten, and E. Waymire. On weighted heights of random trees. *J. Theoret. Probab.*, 4:223–237, 1991.

[14] A. Dvoretsky and Th. Motzkin. A problem of arrangements. *Duke Math. J.*, 14:305–313, 1947.

[15] M. Dwass. A fluctuation theorem for cyclic random variables. *Ann. Math. Stat.*, 33:1450–1453, 1962.

[16] M. Dwass. A theorem about infinitely divisible distributions. *Z. Wahrsch. Verw. Gebiete*, 9:287–289, 1967.

[17] M. Dwass. The total progeny in a branching process. *J. Appl. Probab.*, 6:682–686, 1969.

[18] A. Erdélyi and I.M.H. Etherington. Some problems of non-associative combinations (2). *Edinburgh Math. Notes*, 32:7–12, 1940.

[19] I.M.H. Etherington. Some problems of non-associative combinations (1). *Edinburgh Math. Notes*, 32:1–6, 1940.

[20] W. Feller. *An Introduction to Probability Theory and its Applications, Vol 2*. Wiley, 1966.

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

[22] J-F. Le Gall and Y. Le Jan. Branching processes in Lévy processes: the exploration process. To appear in *Ann. Probab*, 1997.

[23] I. Gessel. A combinatorial proof of the multivariable Lagrange inversion formula. *J. Combinatorial Theory A*, 45:178–195, 1987.

[24] I.J. Good. The number of individuals in a cascade process. *Proc. Camb. Phil. Soc.*, 45:360–363, 1949.

[25] I.J. Good. Generalizations in several variables of Lagrange's expansion, with applications to stochastic processes. *Proc. Camb. Phil. Soc.*, 56:366–380, 1963.

[26] I.J. Good. The generalization of Lagrange's expansion and the enumeration of trees. *Proc. Camb. Phil. Soc.*, 61:499–517, 1965.

[27] I.J. Good. The Lagrange distributions and branching processes. *SIAM Journal on Applied Mathematics*, 28:270–275, 1975.

[28] W. Gutjahr. Expectation transfer between branching processes and random trees. *Random Structures and Algorithms*, 4:447–467, 1993.

[29] T. E. Harris. First passage and recurrence distributions. *Trans. Amer. Math. Soc.*, 73:471–486, 1952.

[30] T.E. Harris. *The Theory of Branching Processes*. Springer-Verlag, New York, 1963.

[31] D. Hawkins and S.M. Ulam. Theory of Multiplicative Processes, 1. Technical Report LA-171, Los Alamos Scientific Laboratory, 1944. (reprinted in *Analogies between analogies: the mathematical reports of S.M. Ulam and his collaborators*, A.R. Bednarek and F. Ulam editors, University of California Press, Berkeley(1990)).

[32] P. Hilton and J. Pedersen. Catalan numbers, their generalization, and their uses. *Math. Intelligencer*, 13:64–75, 1991.

[33] P. Jagers. *Branching Processes with Biological Applications*. Wiley, 1975.

[34] J.H.B. Kemperman. The general one-dimensional random walk with absorbing barriers. Thesis, Excelsior, The Hague, 1950.

[35] J.H.B. Kemperman. *The Passage Problem for a Stationary Markov Chain*. University of Chicago Press, 1961.

[36] D.G. Kendall. Some problems in the theory of queues. *J.R.S.S. B*, 13:151–185, 1951.

[37] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22:425–487, 1987.

[38] V.F. Kolchin. Branching processes, random trees, and a generalized scheme of arrangements of particles. *Mathematical Notes of the Acad. Sci. USSR*, 21:386–394, 1977.

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

[40] G. Labelle. Une novelle démonstration combinatoire des formules d'inversion de Lagrange. *Adv. in Math.*, 42:217–247, 1981.

[41] A. Meir and J.W. Moon. On the altitude of nodes in random trees. *Canad. J. Math.*, 30:997–1015, 1978.

[42] J.W. Moon. Various proofs of Cayley's formula for counting trees. In F. Harary, editor, *A Seminar on Graph Theory*, pages 70–78. Holt, Rineharte and Winston, New York, 1967.

[43] R. Otter. The multiplicative process. *Ann. Math. Statist.*, 20:206–224, 1949.

[44] A. G. Pakes and T. P. Speed. Lagrange distributions and their limit theorems. *SIAM Journal on Applied Mathematics*, 32:745–754, 1977.

[45] Yu. L. Pavlov. Limit theorems for the number of trees of a given size in a random forest. *Math. USSR Subornik*, 32:335–345, 1977.

[46] Yu. L. Pavlov. The asymptotic distribution of maximum tree size in a random forest. *Theory of Probability and its Applications*, 22:509–520, 1977.

[47] Yu. L. Pavlov. Limit distributions of the height of a random forest. *Theory of Probability and its Applications*, 28:471 – 480, 1983.

[48] Yu. L. Pavlov. Limit distributions of the height of a random forest of plane rooted trees. *Discrete Math. Appl.*, 4:73–88, 1994.

[49] Yu. V. Pavlov. Limit distribution of the number of trees of a given size in a random forest. *Discrete Math. Appl.*, 6:117–133, 1996.

[50] Yu.V. Pavlov. A case of the limit distribution of the maximal volume on a tree in a random forest. *Mathematical Notes of the Acad. Sci. USSR*, 25:387–392, 1979.

[51] J. Pitman. Coalescent random forests. Technical Report 457, Dept. Statistics, U.C. Berkeley, 1996.

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

[53] G.N Raney. Functional composition patterns and power series reversion. *TAMS*, 94:441–451, 1960.

[54] G.N. Raney. A formal solution of $\sum_{i=1}^{\infty} a_i e^{B_i X} = x$. *Canad. J. Math.*, 16:755–762, 1964.

[55] 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.

[56] R.K. Sheth and J. Pitman. Coagulation and branching process models of gravitational clustering. To appear in *Mon. Not. R. Astron. Soc.*, 1997.

[57] P.W. Shor. A new proof of Cayley's formula for counting labelled trees. *J. Combinatorial Theory A.*, 71:154–158, 1995.

[58] M. Sibuya, N. Miyawaki, and U. Sumita. Aspects of Lagrangian probability distributions. *Studies in Applied Probability. Essays in Honour of Lajos Takács (J. Appl. Probab.)*, 31A:185–197, 1994.

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

[60] L. Takács. A generalization of the ballot problem and its application to the theory of queues. *J. Amer. Stat. Assoc.*, 57:154–158, 1962.

[61] L. Takács. *Combinatorial Methods in the Theory of Stochastic Processes.* Robert E. Kreiger Publ. Co., Huntington, New York, 1977.

[62] L. Takács. Ballots, queues and random graphs. *J. Appl. Probab.*, 26:103–112, 1989.

[63] L. Takács. Counting forests. *Discrete Mathematics*, 84:323–326, 1990.

[64] L. Takács. Limit distributions for queues and random rooted trees. *J. Applied Mathematics and Stochastic Analysis*, 6:189–216, 1993.

[65] J.C. Tanner. A problem of interference between two queues. *Biometrika*, 40:58–69, 1953.

[66] J.C. Tanner. A derivation of the Borel distribution. *Biometrika*, 1961:222–224, 1961.

[67] J.G. Wendel. Left continuous random walk and the Lagrange expansion. *Amer. Math. Monthly*, 82:494–498, 1975.