Discrete Mathematics 66 (1987) 59–77
North-Holland

SOME METHODS FOR COMPUTING COMPONENT DISTRIBUTION PROBABILITIES IN RELATIONAL STRUCTURES

Kevin J. COMPTON*

Dept. of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 49019, U.S.A.

Received 13 December 1985
Revised 9 October 1986

Introduction

Component distribution problems arise frequently in computer science and combinatorics. In analyzing algorithms or enumerating combinatorial classes one often needs to compute the probability that a random structure will have certain properties having to do with size and number of components. Many results of this sort in the literature rely on estimating growth of coefficients in generating series. These include Darboux's method (see Bender [2]; examples include Hanlon [7], and Palmer and Schwenk [14] who use a special case of the method that they credit to Pòlya), the saddle point method in combination with Laplace's method for estimating integrals (as in Hayman [9]; see Bender [2] for applications), and the Lagrange Inversion Theorem, usually applied to classes of trees (see [6, 8, 13]).

We present simpler, more easily applicable methods for determining the probability that a random relational structure will have a given number of components, and for determining the expected number of components. A *relational structure* is a set together with some relations (such as edge relations, order relations, etc.) on the set. By *random* we mean that structures of the same size (i.e., cardinality of underlying set) are chosen from some set $\mathcal{S}$ of relational structures. We are concerned with the asymptotic probabilities of properties as the size of structures becomes unbounded. We consider both *labeled* structures (each structure on some fixed underlying set having equal weight) and *unlabeled* structures (each isomorphism type having equal weight).

We require that $\mathcal{S}$ be closed under disjoint unions and components. This means that structures in $\mathcal{S}$ may be uniquely decomposed into disjoint unions of connected structures in $\mathcal{S}$. In [4] and [5] we investigated probabilities of general properties for such classes. We obtain stronger results here in the restricted domain of component distribution properties.

* Research supported by NSF Grant DCR-84 18524.

Let a_n be the number of labeled structures of size n in $\mathcal{S}$ and c_n the number of connected labeled structures of size n in $\mathcal{S}$. Then the *exponential generating series* $a(x) = \sum_{n\geq 0} (a_n/n!)x^n$ and $c(x) = \sum_{n\geq 0} (c_n/n!)x^n$ are formally related by

$$a(x) = e^{c(x)} \tag{1}$$

(by convention, $a_0 = 1$, $c_0 = 1$). If b_n is the number of unlabeled structures of size n in $\mathcal{S}$, d_n the number of connected unlabeled structures of size n in $\mathcal{S}$, then the *ordinary generating series* $b(x) = \sum_{n\geq 0} b_n x^n$ and $d(x) = \sum_{n\geq 0} d_n x^n$ are formally related by

$$b(x) = \prod_{i\geq 1} (1 - x^i)^{-d_i} = \exp \sum_{i\geq 1} d(x^i)/i \tag{2}$$

(by convention $b_0 = 1$, $d_0 = 0$).

We will not prove these and other familiar combinatorial facts (see Bender and Goldman [3], Goulden and Jackson [6], Harary and Palmer [8], or Compton [4] for terminology and proofs). Our main results proceed from (1) and (2) and may be viewed as findings about series satisfying these relationships.

We briefly summarize our main results.

Theorem 4 gives expressions for the probability that a random labeled structure has precisely m components, and for the expected number of components, in cases where these quantities exist and some higher derivative of $a(x)$ diverges at its radius of convergence.

Theorem 5 is the analogous theorem for unlabeled structures; there some higher order derivative of $b(x)$ is required to diverge at its radius of convergence.

Theorem 7 (for the labeled case) and 8 (for the unlabeled case) give sufficient conditions for the probability of having precisely m component and for the expected number of components to exist when the appropriate generating series ($a(x)$ or $b(x)$) *diverges* at its radius of convergence.

Theorems 10 and 11 give sufficient conditions (again in the labeled and unlabeled cases) for these quantities to exist when the appropriate generating series *converges* at its (positive) radius of convergence.

Other investigators have considered related problems.

Wright, in [19] and [20], takes (1) and (2) as his starting point to find conditions under which the probability of connectivity is 1. One of his results is that for the probability of connectivity to be 1 in labeled structures, $a(x)$ must have radius of convergence 0, and in unlabeled structures, $b(x)$ must have radius of convergence 0. In [18] he extends these results to obtain asymptotic expansions. See also Bender [1] for sufficient conditions that the probability of connectivity be 1. Since our results concern only series with positive radii of convergence, they are, in some sense, complementary.

Some classes to which our results apply are different species of labeled forests, investigated by Moon [13]; unlabeled forests, investigated by Palmer and Schwenk [14]; and unlabeled unit interval graphs, investigated by Hanlon [7].

Both [7] and [14] rely on Darboux's method (see Bender [2]) or some variant to obtain asymptotic expressions for coefficients of generating series. Our methods are more general since they apply in every case where Darboux's method applies and many others.

Katz [10] and Kruskal [11] consider component distribution problems for unary functions, and Shepp and Lloyd [17] for permutations. Our methods give only probabilities, not asymptotic expressions when probabilities are 0, so are not general enough to yield their results.

Möhring [12] has studied similar issues concerning the probability of indecomposability in relations. His techniques, however, do not seem to pertain here.

In a subsequent paper we will present some applications of the ideas here to analysis of algorithms.

1. Preliminaries

We shall use the falling factorial notation $(n)_i = n(n-1)\cdots(n-i+1)$. Thus, $\binom{n}{i} = (n)_i/i!$ and if $\alpha(x) = \sum_{n\geq 0} \alpha_n x^n$, then the formal rth derivative is $\alpha^{(r)}(x) = \sum_{n\geq 0} (n+r)_r \alpha_{n+r} x^n$ (sometimes also written $D^r[\alpha(x)]$).

Let $\mathcal{S}$ be a class of relational structures closed under disjoint unions and components. (We assume throughout that all classes of structures considered are closed under isomorphism.) Suppose $X \subseteq \mathcal{S}$. Let c_n be the number labeled structures of size n (i.e., with domain some fixed set of size n) in X. Let d_n be the number of unlabeled structures of size n (i.e., the number of isomorphism types with domain of cardinality n) in X. The *exponential generating series* for X is $\sum_{n\geq 0} (c_n/n!)x^n$; the *ordinary generating series* for X is $\sum_{n\geq 0} d_n x^n$. Now we specify that $\mathcal{S}$ itself has exponential generating series $\sum_{n\geq 0} (a_n/n!)x^n$ and ordinary generating series $\sum_{n\geq 0} b_n x^n$. Define $\mu_n(X) = c_n/a_n$, $\nu_n(X) = d_n/b_n$, and $\mu(X) = \lim_{n\to\infty} \mu_n(X)$, $\nu(X) = \lim_{n\to\infty} \nu_n(X)$ whenever these limits exist.

Suppose that $\boldsymbol{X}$ is a function from $\mathcal{S}$ to $\mathbb{N} = \{0, 1, \ldots\}$. (We assume always that functions are constant on isomorphism classes.) Define $\{\boldsymbol{X} = m\}$, $m \in \mathbb{N}$, to be the class of structures mapped to m by $\boldsymbol{X}$ and $\{\boldsymbol{X} \leq m\}$ to be the class of structures mapped to an integer not greater than m. When we apply μ_n, ν_n, μ, or ν to $\{\boldsymbol{X} = m\}$ or $\{\boldsymbol{X} \leq m\}$ we delete the braces: e.g. $\mu_n(\boldsymbol{X} = m)$.

We define the *expected value* of $\boldsymbol{X}$ with respect to μ_n to be

$$\mathrm{E}(\boldsymbol{X}, \mu_n) = \sum_{m\geq 0} m\mu_n(\boldsymbol{X} = m),$$

and with respect to ν_n to be

$$\mathrm{E}(\boldsymbol{X}, \nu_n) = \sum_{m\geq 0} m\nu_n(\boldsymbol{X} = m).$$

Also, put $\mathrm{E}(\boldsymbol{X}, \mu) = \lim_{n\to\infty} \mathrm{E}(\boldsymbol{X}, \mu_n)$ and $\mathrm{E}(\boldsymbol{X}, \nu) = \lim_{n\to\infty} \mathrm{E}(\boldsymbol{X}, \nu_n)$ whenever these limits exist.

Henceforth, let $\boldsymbol{C}: \mathcal{S} \to \mathbb{N}$ be the map that takes a structure to the number of its components. It is well known (see the sources cited in the introduction) that if $\{\boldsymbol{C} = 1\}$ has exponential generating series $c(x) = \sum_{n\geq 0} (c_n/n!)x^n$, then $\{\boldsymbol{C} = m\}$ has exponential generating series $c(x)^m/m!$. Summing over all m gives (1). (By convention $\{\boldsymbol{C} = 0\}$ has exponential generating series 1.) Suppose $e_n = a_n\mathrm{E}(\boldsymbol{C}, \mu)$. Then

$$\sum_{n\geq 0} (e_n/n!)x^n = \sum_{m\geq 0} mc(x)^m/m! = c(x)a(x). \tag{3}$$

To do the same sort of thing in the unlabeled case we require the Pòlya cycle indicator function for S_m, the symmetric group on m elements. $Z(S_m; x_1, x_2, \ldots, x_m)$ is a function of m variables; it satisfies the following formal relationship (see Harary and Palmer [8].)

$$\exp \sum_{i\geq 1} x_i y^i/i = \sum_{m\geq 0} Z(S_m; x_1, x_2, \ldots, x_m)y^m \tag{4}$$

(by convention $Z(S_0) = 1$). The reader unfamiliar with the Pòlya cycle indicator function may take this as a definition.

If $\{\boldsymbol{C} = 1\}$ has ordinary generating series $d(x) = \sum_{n\geq 0} d_n x^n$, then $\{\boldsymbol{C} = m\}$ has ordinary generating series

$$Z_m(x) = Z(S_m; d(x), d(x^2), \ldots, d(x^m)).$$

Substituting $d(x^i)$ for x_i in (4) we have

$$\exp \sum_{i\geq 1} d(x^i)y^i/i = \sum_{m\geq 0} Z_m(x)y^m. \tag{5}$$

If we set $y = 1$ we have (2). If we differentiate (5) with respect to y and set coefficients of like powers of y equal we have

$$mZ_m(x) = \sum_{1\leq i\leq m} d(x^i)Z_{m-i}(x). \tag{6}$$

If we differentiate (5) with respect to x and set coefficients of like powers of y equal we have

$$Z'_m(x) = \sum_{1\leq i\leq m} x^{i-1}d'(x^i)Z_{m-i}(x). \tag{7}$$

Let $f_n = b_n\mathrm{E}(\boldsymbol{C}, \nu)$. Then

$$\sum_{n\geq 0} f_n x^n = \sum_{m\geq 0} mZ_m(x) = \Big(\sum_{i\geq 1} d(x^i)\Big)b(x), \tag{8}$$

by (2) and (6).

As we noted in the introduction, computations of $\mu(C=m)$, $\nu(C=m)$, $\mathrm{E}(C, \mu)$, and $\mathrm{E}(C, \nu)$ have relied on asymptotic methods such as Darboux's method (see Bender [2]) and Hayman's generalization of Stirling's formula (see Hayman [9]). Our simpler methods are partial converses to the following simple extension of Abel's theorem on the value of a power series at its radius of convergence.

Proposition 1. *Suppose* $\alpha(x)=\sum_{n\geqslant 0}\alpha_n x^n$, $\beta(x)=\sum_{n\geqslant 0}\beta_n x^n$, $\gamma(x)=\sum_{n\geqslant 0}\gamma_n x^n$, $\gamma(x)=\alpha(x)\beta(x)$. *If*

(i) $\alpha(x)$ *has radius of convergence* R, $0<R\leqslant\infty$,
(ii) $\lim_{x\to R}\alpha(x)=\infty$, *and*
(iii) $\lim_{n\to\infty}\gamma_n/\alpha_n=L$,

then $\lim_{x\to R}\beta(x)=L$.

The proof is easy and will be omitted. Proposition 1 is the logical sum of Problems I.85 and I.94 in Pòlya and Szegö [15]. They attribute Problem I.85 to Cesàro.

We now state the partial converses to Proposition 1. They will be our main asymptotic techniques.

Proposition 2. *Suppose* $\alpha(x)=\sum_{n\geqslant 0}\alpha_n x^n$, $\alpha_n\geqslant 0$, $\beta(x)=\sum_{n\geqslant 0}\beta_n x^n$, $\beta_n\geqslant 0$, $\gamma(x)=\sum_{n\geqslant 0}\gamma_n x^n=\alpha(x)\beta(x)$. *If*

(i) $\lim_{n\to\infty}\alpha_{n-1}/\alpha_n=R$, $0<R<\infty$, *and*
(ii) $\beta(x)$ *has radius of convergence greater than* R,

then $\lim_{x\to\infty}\gamma_n/\alpha_n=\beta(R)$.

Proposition 3. *Suppose* $\alpha(x)=\sum_{n\geqslant 0}\alpha_n x^n$, $\alpha_n\geqslant 0$, $\beta(x)=\sum_{n\geqslant 0}\beta_n x^n$, $\gamma(x)=\sum_{n\geqslant 0}\gamma_n x^n=\alpha(x)\beta(x)$. *If*

(i) $\lim_{n\to\infty}\alpha_{n-1}/\alpha_n=R$, $0<R<\infty$,
(ii) $\beta(x)$ *is absolutely convergent at* $x=R$, *and for some real* $K>0$ *we have for all large* n *and* $k\leqslant n$ *that* $\alpha_k R^k\leqslant K\alpha_n R^n$,

then $\lim_{n\to\infty}\gamma_n/\alpha_n=\beta(R)$.

Proposition 2 is Theorem 2 of Bender [2]. He remarks that his proof is "standard and uninteresting". The proof of Proposition 3 is nearly identical and the same remark applies. Proposition 2 appears as Problem I.178 in Pòlya and Szegö [15]. They attribute it to Schur.

2. Main Results

Theorem 4, the first main result, tells us, under certain hypotheses, what $\mu(C=m)$ and $\mathrm{E}(C=\mu)$ must be, should they exist. Note that by the theorem, whenever the hypotheses hold and $\mu(\boldsymbol{C}, m)$ is nonzero for each $m \geq 1$, then

$$\mathrm{E}(\boldsymbol{C}, \mu) = \sum_{m \geq 0} m\mu(\boldsymbol{C} = m).$$

Thus, it is reasonable to refer to $\mathrm{E}(\boldsymbol{C}, \mu)$ as the *expected value* of $\boldsymbol{C}$ with respect to μ. If $\mu(\boldsymbol{C} = m) = 0$ for all $m \geq 1$, we can still make $\mathrm{E}(\boldsymbol{C}, \mu)$ an expected value by taking $\mu(\boldsymbol{C} = \infty) = 1$. Thus

$$\mathrm{E}(\boldsymbol{C}, \mu) = \sum_{0 \leq m \leq \infty} m\mu(\boldsymbol{C} = m).$$

In the remainder of the paper an expression of the form $\gamma(R)/\alpha(R)$ will be interpreted as $\lim_{x \to R} \gamma(x)/\alpha(x)$ when $\alpha(x)$ diverges at R.

Theorem 4. *Suppose that $\mathcal{S}$ has exponential generating series $a(x) = \exp c(x)$, where*

(i) *$a(x)$ has radius of convergence R, $0 < R \leq \infty$, and*
(ii) *for some nonnegative integer r $\lim_{x \to R} a^{(r)}(x) = \infty$.*

For every $m \geq 1$, if $\mu(\boldsymbol{C} = m)$ exists, then $\mu(\boldsymbol{C} = m) = c(R)^{m-1}/((m-1)!\, a(R))$. If $\mathrm{E}(\boldsymbol{C}, \mu)$ exists, then $\mathrm{E}(\boldsymbol{C}, \mu) = 1 + c(R)$.

Proof. Let r be the least integer satisfying hypothesis (ii). By Proposition 1, if $\mu(\boldsymbol{C} = m)$ exists, then

$$\mu(\boldsymbol{C} = m) = \lim_{x \to R} \frac{D^r[c(x)^m/m!]}{a^{(r)}(x)}.$$

If $r = 0$ the value of this expression is 0, which is the desired value.

If $r > 0$, then

$$\begin{aligned} D^r[c(x)^m/m!] &= D^{r-1}[c'(x)c(x)^{m-1}/(m-1)!] \\ &= \sum_{1 \leq i \leq r} \binom{r-1}{i-1} c^{(i)}(x) D^{r-i}[c(x)^{m-1}/(m-1)!] \end{aligned} \tag{9}$$

and

$$a^{(r)}(x) = D^{r-1}[c'(x)a(x)] = \sum_{1 \leq i \leq r} \binom{r-1}{i-1} c^{(i)}(x) a^{(r-i)}(x) \tag{10}$$

In both these sums the only term that diverges as x approaches R occurs when $i = r$. Thus, eliminating all but the dominant terms we have

$$\mu(\boldsymbol{C} = m) = \lim_{x \to R} \frac{c^{(r)}(x)c(x)^{m-1}/(m-1)!}{c^{(r)}(x)a(x)} = \frac{c(R)^{m-1}}{(m-1)!\, a(R)}.$$

Now if $\mathrm{E}(\boldsymbol{C}, \mu)$ exists, then by (3) and Proposition 1

$$\mathrm{E}(\boldsymbol{C}, \mu) = \lim_{x \to R} \frac{D^r[c(x)a(x)]}{a^{(r)}(x)}$$

If $r=0$ this is ∞, the desired value. If $r>0$, then

$$D^r[c(x)a(x)] = D^{r-1}[c'(x)(1+c(x))a(x)]$$

$$= \sum_{1\leq i\leq r} \binom{r-1}{i-1} c^{(i)}(x) D^{r-i}[(1+c(x))a(x)]. \tag{11}$$

The dominant term occurs when $i=r$ so

$$\mathrm{E}(\boldsymbol{C}, \mu) = \lim_{x\to R} \frac{c^{(r)}(x)(1+c(x))a(x)}{c^{(r)}(x)a(x)} = 1 + c(R). \quad \square$$

The next theorem is the unlabeled analogue of Theorem 4. Once again note that the conclusion shows it is reasonable to refer to $\mathrm{E}(\boldsymbol{C}, \nu)$ as the *expected value* of $\boldsymbol{C}$ with respect to ν.

Theorem 5. *Suppose that $\mathcal{S}$ has ordinary generating series $b(x) = \exp \sum_{i\geq 1} d(x^i)/i$, where*

(i) *$b(x)$ has radius of convergence R, $0<R\leq 1$, and*

(ii) *for some nonnegative integer r $\lim_{x\to R} b^{(r)}(x) = \infty$.*

For every $m\geq 1$, if $\nu(\boldsymbol{C}=m)$ exists, then $\nu(\boldsymbol{C}=m) = Z_{m-1}(R)/d(R)$. If $\mathrm{E}(\boldsymbol{C}, \nu)$ exists, then $\mathrm{E}(\boldsymbol{C}, \nu) = 1 + \sum_{i\geq 1} d(R^i)$. (Recall that $Z_m(x)$ is given by (5).)

Proof. Let r be the least integer satisfying hypothesis (ii). By Proposition 1 and the remarks in Section 1, if $\nu(\boldsymbol{C}=m)$ exists, then

$$\nu(\boldsymbol{C}=m) = \lim_{x\to R} \frac{Z_m^{(r)}(x)}{b^{(r)}(x)}.$$

If $r=0$ the value of this expression is 0, the desired value.

If $r>0$, then $0<R<1$ and, from (7),

$$Z_m^{(r)}(x) = D^{r-1}\Big[\sum_{1\leq i\leq m} x^{i-1} d'(x^i) Z_{m-i}(x)\Big]$$

$$= \sum_{1\leq j\leq r} \binom{r-1}{j-1} d^{(j)}(x) Z_{m-1}^{(r-j)}(x) + D^{r-1}\Big[\sum_{2\leq i\leq m} x^{i-1} d'(x^i) Z_{m-i}(x)\Big]. \tag{12}$$

Now when $i\geq 2$, $x^{i-1}d'(x^i)$ has radius of convergence $R^{1/i} > R$. Thus if we apply Leibniz' rule to the subexpression $D^{r-1}[\ldots]$ of (12), all the resulting terms are bounded as x approaches R; examining the sum that forms the initial part of (12) we see that the only term that diverges as x approaches R occurs when $j=r$. Also, from (2),

$$b^{(r)}(x) = D^{r-1}\Big[\Big(\sum_{i\geq 1} x^{i-1} d'(x^i)\Big) b(x)\Big]$$

$$= \sum_{1\leq j\leq r} \binom{r-1}{j-1} d^{(j)}(x) b^{(r-j)}(x) + D^{r-1}\Big[\Big(\sum_{i\geq 2} x^{i-1} d'(x^i)\Big) b(x)\Big]. \tag{13}$$

It is easy to show that $\sum_{i\geq 2} x^{i-1}d(x^i)$ is analytic for $|x| < R^{\frac{1}{2}}$, so the same argument as above shows that the only part of (13) that diverges as x approaches r is the $j = r$ term in the first sum. Hence, by Proposition 1,

$$\mu m(C = m) = \lim_{x\to R} \frac{d^{(r)}(x) Z_{m-1}(x)}{d^{(r)}(x) b(x)} = \frac{Z_{m-1}(R)}{b(R)}.$$

If $\mathrm{E}(C, v)$ exists, then by (8) and Proposition 1

$$\mathrm{E}(C, v) = \lim_{x\to R} \frac{D^r[(\sum_{i\geq 1} d(x^i)) b(x)]}{b^{(r)}(x)}.$$

If $r = 0$ this is ∞, the desired value. If $r > 0$,

$$\begin{aligned} D^r\Big[\Big(\sum_{i\geq 1} d(x^i)\Big) b(x)\Big] &= D^{r-1}\Big[\Big(\sum_{i\geq 1} x^{i-1} d'(x^i)\Big(i + \sum_{n\geq 1} d(x^n)\Big)\Big) b(x)\Big] \\ &= \sum_{1\leq j\leq r} \binom{r-1}{j-1} d^{(j)}(x) D^{r-j}\Big[\Big(1 + \sum_{n\geq 1} d(x^n)\Big) b(x)\Big] \\ &\quad + D^{r-1}\Big[\Big(\sum_{i\geq 2} i x^{i-1} d'(x^i) + \sum_{i\geq 2} x^{i-1} d'(x^i) \sum_{n\geq 1} d(x^n)\Big) b(x)\Big]. \end{aligned} \tag{14}$$

Notice that $\sum_{i\geq 2} i x^i d'(x^i)$ and $\sum_{i\geq 2} x^{i-1} d'(x^i)$ are analytic for $|x| < R^{\frac{1}{2}}$ and $\sum_{i\geq 1} d(x^i) = d(x) + \sum_{i\geq 2} d(x^i)$. Therefore, if we apply Leibniz' rule again to the second part of (14) we see that all the terms converge as x approaches R. In the first sum the only term that diverges as x approaches R occurs when $j = r$. Consequently,

$$\mathrm{E}(C, v) = \lim_{x\to R} \frac{d^{(r)}(x)(1 + \sum_{i\geq 1} d(x^i)) b(x)}{d^{(r)}(x) b(x)} = 1 + \sum_{i\geq 1} d(R^i). \quad \square$$

Theorems 4 and 5 tell us what the probability of having m components and the expected number of components must be when the hypotheses are met *and* these values are known to exist. The remainder of the paper is devoted to establishing sufficient conditions for the existence of these values.

First we examine conditions in the case where the appropriate generating series diverges at its radius of convergence. The following lemma is required.

Lemma 6. *Let* $\alpha(x) = \sum_{n\geq 0} \alpha_n x^n$, $\beta(x) = \sum_{n\geq 0} \beta_n x^n$, α_n, $\beta_n \geq 0$, *and* $\gamma(x) = \sum_{n\leq 0} \gamma_n x^n = \alpha(x)\beta(x)$. *If*

(i) $\lim_{n\to\infty} \alpha_{n-1}/\alpha_n = R$, $0 < R \leq \infty$,

(ii) $\lim_{x\to R} \alpha(x) = \infty$, *and*

(iii) $\lim_{n\to\infty} \beta_{n-1}/\beta_n = S \geq R$,

then $\lim_{n\to\infty} \gamma_{n-1}/\gamma_n = R$.

Proof. Since $\gamma_n = \sum_{0\leqslant i\leqslant n} \beta_i \alpha_{n-i}$,

$$R\gamma_n - \gamma_{n-1} = \sum_{0\leqslant i\leqslant n-1} \beta_i(R\alpha_{n-i} - \alpha_{n-i-1}) + R\beta_n\alpha_0, \tag{15}$$

Now if we can show that the right side is $o(\gamma_n)$, then $R\gamma_n - \gamma_{n-1} = o(\gamma_n)$ and the conclusion follows. By (i), $R\alpha_n - \alpha_{n-1} = o(\alpha_{n-1})$ so for $\varepsilon > 0$ there is an N such that $|R\alpha_n - \alpha_{n-1}| \leqslant \varepsilon\alpha_{n-1}$ when $n \geqslant N$. Thus, (15) is bounded in absolute value by

$$\varepsilon \sum_{0\leqslant i\leqslant n-N} \beta_i\alpha_{n-i-1} + \sum_{1\leqslant i\leqslant N} \beta_{n-i} |R\alpha_i - \alpha_{i-1}| + R\beta_n\alpha_0.$$

The first term of this expression is not greater than $\varepsilon\gamma_{n-1}$.

The remainder of the proof divides into two cases. First, suppose $S = \infty$. By (i) there is an $M \geqslant N$ such that $\alpha_M > 0$. Let

$$K = \frac{\sum_{1\leqslant i<N} |R\alpha_i - \alpha_{i-1}| + R\alpha_0}{\varepsilon\alpha_M}.$$

By (iii) for large enough n,

$$\frac{\beta_{n-M-1}}{\beta_{n-M}} \geqslant \max(1, K).$$

Thus, for large n,

$$\begin{aligned}\sum_{1\leqslant i<N} \beta_{n-i} |R\alpha_i - \alpha_{i-1}| + R\beta_n\alpha_0 &\leqslant \beta_{n-N}\Big(\sum_{1\leqslant i<N} |R\alpha_i - \alpha_{i-1}| + R\alpha_0\Big)\\ &\leqslant \beta_{n-M}(K\varepsilon\alpha_M) \leqslant \varepsilon\beta_{n-M-1}\alpha_M\\ &\leqslant \varepsilon \sum_{0\leqslant i\leqslant n-1} \alpha_i\beta_{n-i-1} \leqslant \varepsilon\gamma_{n-1}.\end{aligned}$$

Therefore, (15) is, in absolute value, not greater than $2\varepsilon\gamma_{n-1}$. By making ε small, we see that $R\gamma_n - \gamma_{n-1} = o(\gamma_{n-1})$. Now suppose $S < \infty$. Since $\lim_{n\to\infty} \beta_{n-i}/\beta_n = S^i$,

$$\lim_{n\to\infty} \Big(\sum_{1\leqslant i<N} \beta_{n-i} |R\alpha_i - \alpha_{i-1}| + R\beta_n\alpha_0\Big)\Big/\beta_n < \infty,$$

$$\lim_{n\to\infty} \gamma_{n-1}/\beta_n = \lim_{n\to\infty} \sum_{0\leqslant i<n-1} \alpha_i\Big(\frac{\beta_{n-i-1}}{\beta_n}\Big) = \infty.$$

The latter follows from (ii). We see then that for large n,

$$\sum_{1\leqslant i<N} \beta_{n-i} |R\alpha_i - \alpha_{i-1}| + R\beta_n\alpha_0 \leqslant \varepsilon\gamma_{n-1}$$

as in the first case. Again, $R\gamma_n = \gamma_{n-1} = o(\gamma_{n-1})$. □

Remark. Lemma 6 holds if the hypotheses are modified so that $\gamma(x) = \alpha(x)\beta(x^i)$, $i \geqslant 1$. The proof requires only minor modifications.

The next theorem combines two results: one says that when the exponential generating series $a(x)=\exp c(x)$ for $\mathcal{S}$ diverges at its radius of convergence then $\mu(C=m)$ and $\mathrm{E}(C,\mu)$ exist if the ratio test applies to $a(x)$; the other makes the same assertion for $c(x)$. Often one of these series will be easier to test than the other. The existence of $\mu(C=m)$ when the ratio test applies to $a(x)$ follows from Theorem 6.1 of Compton [4], but we give a proof that does not require the machinery developed there.

Notice that any series $a(x)$ that is admissible in the sense of Hayman [9] satisfies hypotheses (i) and (ii) of Theorem 7. Asymptotic expressions for the coefficients of such series are obtained by saddle point integration and Laplace's method. Theorem 7 eliminates the need for these steps when computing component distributions.

Theorem 7. *Suppose that $\mathcal{S}$ has exponential generating series $a(x)=\sum_{n\geq 0}(a_n/n!)x^n=\exp c(x)$, $c(x)=\sum_{n\geq 0}(c_n/n!)x^n$. If*

(i) *either $\lim_{n\to\infty} na_{n-1}/a_n=R$ or $\lim_{n\to\infty} nc_{n-1}/c_n=R$ for some R, $0<R\leq\infty$, and*

(ii) *$\lim_{x\to R} c(x)=\infty$ (or equivalently, $\lim_{x\to R} a(x)=\infty$),*

then for all $m\geq 1$, $\mu(C=m)=0$, and $\mathrm{E}(C,\mu)=\infty$.

Proof. Suppose $\lim_{n\to\infty} na_{n-1}/a_n=R$ and $R<\infty$. Define C_r to be the function mapping a structure to the number of its components of size not greater than r. Thus, $\mu_n(C=m)\leq\mu_n(C\leq m)\leq\mu_n(C_r\leq m)$. The exponential generating series for the class of connected structures in $\mathcal{S}$ of cardinality not greater than r is

$$c_{\leq r}(x)=\sum_{1\leq n\leq r}\frac{c_n}{n!}x^n.$$

The exponential generating series for $\{C_r\leq m\}$ is

$$\Big(\sum_{i\leq m} c_{\leq r}(x)^i/i!\Big)\exp(c(x)-c_{\leq r}(x)).$$

Dividing this by $a(x)=\exp c(x)$ gives

$$\Big(\sum_{i\leq m} c_{\leq r}(x)^i/i!\Big)/\exp c_{\leq r}(x),$$

which has radius of convergence ∞. By Proposition 2

$$\mu(C_r\leq m)=\frac{\sum_{i\leq m} c_{\leq r}(R)^i/i!}{\exp c_{\leq r}(R)}.$$

Now by letting r approach ∞ we see that $0\leq\mu(C=m)\leq\mu(C_r\leq m)=0$.

Suppose that $\lim_{n\to\infty} na_{n-1}/a_n=R$ and $R=\infty$. The exponential generating series for $\{C=m\}$ is $c(x)^m/m!$. Let $c_{m,n}$ be the coefficient of x^n in this series; i.e.,

$$c(x)^m/m! = \sum_{n\geq 0} \frac{c_{m,n}}{n!} x^n.$$

Since $c(x)^{m+1}/(m+1)! = (c(x)/(m+1))(c(x)^m/m!)$, we have

$$c_{m+1,n} = \sum_{0\leq i\leq n} \frac{\binom{n}{i} c_i c_{m,n-i}}{m+1}. \tag{16}$$

Pick r such that $c_r > 0$. Then

$$a_n \geq c_{m+1,n} \geq \frac{\binom{n}{r} c_r c_{m,n-r}}{m+1},$$

from which it follows that

$$\frac{((n)_r a_{n-r}/a_n)^{-1}(m+1)r!}{c_r} \geq \frac{c_{m,n-r}}{a_{n-r}} = \mu_{n-r}(C=m).$$

Taking the limit as n approaches ∞ we have $\mu(C=m)=0$.

Suppose now that $\lim_{n\to\infty} nc_{n-1}/c_n = R$. From (16)

$$\frac{c_{m+1,n}}{c_{m,n}} = \sum_{0\leq i\leq n} \frac{(c_i/i!)((n)_i c_{m,n-i}/c_{m,n})}{m+1}.$$

Now by Lemma 6 it follows easily that the ratio test applies to $c(x)^m/m!$; i.e., that $\lim_{n\to\infty} nc_{m,n-1}/c_{m,n} = R$. Hence, $\lim_{n\to\infty} (n)_i c_{m,n-i}/c_{m,n} = R^i$. Because $\sum_{i\geq 0} (c_i/i!)R^i$ diverges, it follows from the equation displayed above that $\lim_{n\to\infty} c_{m+1,n}/c_{m,n} = \infty$. But $\mu_n(C=m) = c_{m,n}/a_n \leq (c_{m+1,n}/c_{m,n})^{-1}$ so $\mu(C=m)=0$.

If $\mu(C=m)=0$ for all $m\geq 1$, it is clear that $\mathrm{E}(C,\mu)=\infty$. □

The next theorem is the analogue for the unlabeled case.

Theorem 8. *Suppose that $\mathcal{S}$ has ordinary generating series $b(x) = \sum_{n\geq 0} b_n x^n = \exp \sum_{i\geq 1} d(x^i)/i$, $d(x) = \sum_{n\geq 0} d_n x^n$. If*

(i) *either $\lim_{n\to\infty} b_{n-1}/b_n = R$ or $\lim_{n\to\infty} d_{n-1}/d_n = R$ for some R, $0<R\leq 1$, and*

(ii) *$\lim_{x\to R} b(x) = \infty$ (or equivalently, either $\lim_{x\to R} d(x) = \infty$ or $R=1$),*

then for all $m\geq 1$, $\nu(C=m)=0$, and $\mathrm{E}(C,\nu)=\infty$.

Proof. The proof is nearly identical to that of Theorem 7. When $\lim_{n\to\infty} b_{n-1}/b_n = R$ and $R<\infty$ we show that $\{C_r \leq m\}$ has ordinary generating series

$$\Big(\sum_{i\leq m} Z(S_i; d_{\leq r}(x), \ldots, d_{\leq r}(x^i))\Big)\exp \sum_{i\leq 1} (d(x^i) - d_{\leq r}(x^i))/i,$$

where $d_{\leqslant r}(x)=\sum_{0\leqslant i\leqslant r} d_i x^i$. This case proceeds as before. In the cases where $\lim_{n\to\infty} b_{n-1}/b_n = R$ and $R<\infty$ and $\lim_{n\to\infty} d_{n-1}/d_n = R$, we use the identity

$$Z_{m+1}(x)=\sum_{1\leqslant i\leqslant m+1}\frac{d(x^i)Z_{m-i+1}(x)}{m+1},$$

derived from (6), rather than the identity used to derive (16). The arguments go through as before if we observe that the coefficients of $Z_{m+1}(x)$ majorize those of $d(x)Z_m(x)$ and that, by the remark following Lemma 6, the ratio test applies to $Z_{m+1}(x)$ if it applies to $d(x)$. □

Examples. Let $\mathcal{S}$ be the class of functional diagraphs (see Goulden and Jackson [6]). The number a_n of labeled structures of size n is n^n. It is easy to show by the ratio test that $a(x)=\sum_{n\geqslant 0}(n^n/n!)x^n$ (here $0^0=1$) has radius of convergence $R=1/\mathrm{e}$. Also, $\lim_{n\to R} a(x)=\infty$. Therefore, by Theorem 7, $\mu(\boldsymbol{C}=m)=0$ for each $m\geqslant 1$, and $\mathrm{E}(\boldsymbol{C},\mu)=\infty$. This example points out the deficiency of Theorem 7.

We would like a result that gives asymptotic estimates for $\mu(\boldsymbol{C}=m)$ and $\mathrm{E}(\boldsymbol{C},\mu)$. In particular, for functional diagraphs Katz [10] shows that $\mu_n(\boldsymbol{C}=1)\sim(\frac{1}{2}\pi)^{\frac{1}{2}}n^{-\frac{1}{2}}$ and Kruskal [11] shows that $\mathrm{E}(\boldsymbol{C},\mu_n)\sim\frac{1}{2}\log n$.

Let $\mathcal{S}$ be the class of acyclic diagraphs in which each vertex has outdegree and indegree at most one. There is precisely one unlabeled connected structure of each finite size, so the class has ordinary generating series

$$b(x)=\exp\Big(\sum_{i\geqslant 1} d(x^i)/i\Big),$$

where $d(x)=\sum_{n\geqslant 1}x^n=x/(1-x)$. Clearly the ratio test applies to $d(x)$ to show that its radius of convergence is 1. Also, $\lim_{x\to 1} d(x)=\infty$. Therefore, by Theorem 5, $\nu(\boldsymbol{C}=m)=0$ for $m\geqslant 1$ and $\mathrm{E}(\boldsymbol{C},\nu)=\infty$. To find a strengthening of Theorem 5 that produces asymptotic estimates for $\mu(\boldsymbol{C}=1)$ may be a formidable task because it would give, in this example, the Hardy–Ramanujan estimate for the partition function.

The final results of this section concern classes with generating series that converge at the radius of convergence. They use the following technical lemma.

Lemma 9. *Let* $\alpha(x)=\sum_{n\geqslant 0}\alpha_n x^n$, $\beta(x)=\sum_{n\geqslant 0}\beta_n x^n$, $\delta(x)=\sum_{n\geqslant 0}\delta_n x^n$, α_n, β_n, $\delta_n\geqslant 0$, *and* $N>0$. *Suppose*

(i) $\lim_{n\to\infty}\delta_{n-1}/\delta_n=R$, $0<R\leqslant\infty$,

(ii) $\lim_{x\to R}\delta(x)<\infty$,

(iii) *there are constants* K_α, K_β *such that for* $n\geqslant N$, $\alpha_n\leqslant K_\alpha\delta_n$, $\beta_n\leqslant K_\beta\delta_n$, *and*

(iv) *there is a nonnegative integer* $r<\frac{1}{4}N$ *and real* $K>0$ *such that if*

$$\delta^{(r)}(x)=\sum_{n\geqslant 0}\delta_{r,n}x^n,$$

then, for $n\geqslant N$ *and all* $k\leqslant n$, $\delta_{r,k}R^k\leqslant K\delta_{r,n}R^n$.

If $s, t > 0$, $s + t \leqslant r$, and $\gamma(x) = \sum_{n\geqslant 0} \gamma_r x^n = \alpha^{(s)}(x)\beta^{(t)}(x)$, then $\gamma_n = o(\delta_{n,r})$.

If $\gamma(x) = \sum_{n\geqslant 0} \gamma_n x^n = \alpha(x)\beta(x)$, $\lim_{n\to\infty} \alpha_n/\delta_n = L_\alpha$, and $\lim_{n\to\infty} \beta_n/\delta_n = L_\beta$, then $\gamma_n \leqslant K(K_\beta\alpha(R) + K_\alpha\beta(R))3^r\delta_n$ for $n \geqslant N$, and $\lim_{n\to\infty} \gamma_n/\delta_n = L_\beta\alpha(R) + L_\alpha\beta(R)$.

Proof. Suppose $\gamma(x) = \alpha^{(s)}(x)\beta^{(t)}(x)$. Then

$$\gamma_n = \sum_{0\leqslant j\leqslant n} (s+j)_s\alpha_{s+j}(n+t-j)_t\beta_{n+t-j}.$$

Divide both sides by $\delta_{r,n}$ and break the summation into parts $1 \leqslant j < \frac{1}{2}n$, $\frac{1}{2}n \leqslant j \leqslant n$. After a few manipulations we have

$$\begin{aligned}\frac{\gamma_n}{\delta_{r,n}} &= \sum_{0\leqslant j<\frac{1}{2}n} \alpha_{s+j}\frac{(s+j)_s}{(n-j)_{r-t}}\frac{\beta_{n+t-j}}{\delta_{n+t-j}}\frac{\delta_{r,n+t-r-j}}{\delta_{r,n}} \\ &\quad + \sum_{0\leqslant j\leqslant\frac{1}{2}n} \beta_{t+j}\frac{(t+j)_t}{(n-j)_{r-s}}\frac{\alpha_{n+s-j}}{\delta_{n+s-j}}\frac{\delta_{r,n+s-r-j}}{\delta_{r,n}}. \end{aligned} \tag{17}$$

Now for $n \geqslant N > 4s$, $j \leqslant \frac{1}{2}n$,

$$\frac{(s+j)_s}{(n-j)_{r-t}} \leqslant \frac{s+j}{n-j} \leqslant \cdots \leqslant \frac{j+1}{n-j-s+1} \leqslant 2^s.$$

Similarly, when $n \geqslant N > 4t$, $(t+j)_t/(n-j)_{r-s} \leqslant 2^t$. Also, $\beta_{n+t-j}/\delta_{n+t-j} < K_\beta$, $\alpha_{n+s-j}/\delta_{n+s-j} < K_\alpha$, $\delta_{r,n+t-r-j}/\delta_{r,n} < KR^{r-t+j}$, $\delta_{r,n+s-r-j}/\delta_{r,n} < KR^{r-s+j}$ for large n. Thus, if we extend the ranges of summation in (17) to $0 \leqslant j < \infty$ by adding zero terms, the sums will be majorized by

$$\sum_{j\geqslant 0} 2^s K_\beta K\alpha_{s+j}R^{r-t+j}, \qquad \sum_{j\geqslant 0} 2^t K_\alpha K\beta_{t+j}R^{r-s+j},$$

respectively. But $\alpha(R)$ and $\beta(R)$ converge so these sums converge. We may apply the Dominated Convergence Theorem (see [16]) to eq. (17) to show that $\lim_{n\to\infty} \gamma_n/\delta_{n,r} = 0$ since for j fixed, $(s+j)_s/(n-j)_{r-t}$ and $(t+j)_t/(n-j)_{r-s}$ approach 0 as n diverges to ∞.

Suppose now that $\gamma(x) = \alpha(x)\beta(x)$. Then

$$\gamma_n = \sum_{0\leqslant j\leqslant n} \alpha_j\beta_{n-j}.$$

Divide both sides by δ_n and again break the sum into two parts. After a few manipulations we have

$$\frac{\gamma_n}{\delta_n} = \sum_{0\leqslant j<\frac{1}{2}n} \alpha_j\frac{(n)_r}{(n-j)_r}\frac{\beta_{n-j}}{\delta_{n-j}}\frac{\delta_{r,n-r-j}}{\delta_{r,n-r}} + \sum_{0\leqslant j\leqslant\frac{1}{2}n} \beta_j\frac{(n)_r}{(n-j)_r}\frac{\alpha_{n-j}}{\delta_{n-j}}\frac{\delta_{r,n-r-j}}{\delta_{r,n-r}}.$$

Now for $n \geqslant N > 4r$ and $j \leqslant \frac{1}{2}n$, $(n)_r/(n-j)_r < 3^r$. Also, $\beta_{n-j}/\delta_{n-j} < K_\beta$, $\alpha_{n-j}/\delta_{n-j} < K_\alpha$, and $\delta_{r,n-r-j}/\delta_{r,n-r} < KR_j$ for $n \geqslant N$. Again we extend the range of summation to $0 \leqslant j < \infty$; observe that the sums are majorized by

$$\sum_{j\geqslant 0} 3^r K K_\beta \alpha_j R^j = 3^r K K_\beta \alpha(R), \qquad \sum_{j\geqslant 0} 3^r K K_\alpha \beta_j R^j = 3^r K K_\alpha \beta(R),$$

respectively. This gives the bound

$$\frac{\gamma_n}{\delta_n} < 3^r K(K_\beta \alpha(R) + K_\alpha \beta(R)),$$

when $n \geqslant N$, which was to be proved. Applying the Dominated Convergence Theorem as before shows that

$$\lim_{n\to\infty} \frac{\gamma_n}{\delta_n} = L_\beta \alpha(R) + L_\alpha \beta(R). \quad \square$$

Theorems 10 and 11, which follow, play the same role for classes with generating series convergent at the radius of convergence R as Theorems 7 and 8 did for classes with generating series divergent at R. The hypotheses for Theorems 10 and 11 are somewhat stronger, but the theorems should be considered more powerful results because they give asymptotic estimates.

Notice that any series $a(x)$ that has a finite number of singularities, all algebraic, on its circle of convergence, and has the dominant singularity at the radius of convergence, satisfies conditions (i) and (ii) of Theorem 10. Asymptotic expressions for coefficients of such series are obtained by Darboux's method (see Bender [2]). Theorem 10 eliminates the need for this step when computing component distributions.

Theorem 10. *Suppose that $\mathcal{S}$ has exponential generating series $a(x) = \sum_{n\geqslant 0} (a_n/n!)x^n = \exp c(x)$, $c(x) = \sum_{n\geqslant 0} (c_n/n!)x^n$. If*

(i) $\lim_{n\to\infty} na_{n-1}/a_n = R$, $0 < R \leqslant \infty$,

(ii) *for some nonnegative integer r and real $K > 0$ we have for large enough n and $r \leqslant k \leqslant n$ that $(a_k/(k-r)!)R^k \leqslant K(a_n/(n-r)!)R^n$,*

then $\mu(C = m) = c(R)^{m-1}/((m-1)!\,a(R))$ and $\mathrm{E}(C, \mu) = 1 + c(R)$. If all coefficients a_j are replaced by c_j in hypotheses (i) *and* (ii), *the conclusion still holds.*

Proof. If $\lim_{x\to R} a(x) = \infty$, then the conclusion follows by Theorem 7. Thus, assume $\lim_{x\to R} a(x) < \infty$.

As in formulas (9) and (10)

$$D^r[c(x)^m/m!] = \sum_{1\leqslant i\leqslant r} \binom{r-1}{i-1} c^{(i)}(x) D^{r-i}[c(x)^{m-1}/(m-1)!],$$

$$a^{(r)}(x) = \sum_{1\leqslant i\leqslant r} \binom{r-1}{i-1} c^{(i)}(x) a^{(r-i)}(x).$$

Now in Lemma 9 put $\alpha(x)=c(x)$, $\beta(x)=c(x)^{m-1}/(m-1)!$, $\delta(x)=a(x)$, $s=i$, $t=r-i$, where $i<r$. Then the coefficients of $c^{(i)}(x)D^{r-i}[c(x)^{m-1}/(m-1)!]$ become vanishingly small with respect to the corresponding coefficients of $a^{(r)}(x)$. Similarly, putting $\alpha(x)=c(x)$, $\beta(x)=\delta(x)=a(x)$, $s=i$, $t=r-i$, $i<r$, we have that the coefficients of $c^{(i)}(x)a^{(r-i)}(x)$ become vanishingly small with respect to the coefficients of $a^{(r)}(x)$. Therefore, in the sums displayed above only the $i=r$ terms need be considered when we seek to show that the ratio of coefficients of $D^r[c(x)^m/m!]$ to corresponding coefficients of $a^{(r)}(x)$ approaches a nonzero limit. Now

$$\frac{c^{(r)}(x)c(x)^{m-1}/(m-1)!}{c^{(r)}(x)a(x)}=\frac{c(x)^{m-1}}{(m-1)!}e^{-c(x)}$$

The series for $c(x)^{m-1}\exp(-c(x))/(m-1)!$ is majorized in absolute value by $c(x)^{m-1}\exp c(x)/(m-1)!$ which is absolutely convergent at $x=R$. Hence, Proposition 3 shows that

$$\mu(\boldsymbol{C}=m)=\frac{c(R)^{m-1}}{(m-1)!\,a(R)}.$$

To show that $\mathrm{E}(\boldsymbol{C},\mu)$ exists and find its value, put $\alpha(x)=c(x)$, $\beta(x)=a(x)$, $\delta(x)=a(x)$ so that $\gamma(x)=c(x)a(x)$, $K_\alpha=K_\beta=1$, $L_\alpha=1/a(R)$, $L_\beta=1$. Thus, from (3)

$$\mathrm{E}(\boldsymbol{C},\mu)=\lim_{n\to\infty}\gamma_n/\delta_n=1+c(R).$$

Suppose that hypotheses (i) and (ii) are true of coefficients c_j rather than a_j. Setting $c(x)^m/m!=\sum_{n\geq 0}(c_{m,n}/n!)x^n$ and N large enough to guarantee that $N\geq 5r$ and

$$\frac{c_m}{(m-r)!}R^m\leq K\left(\frac{c_n}{(n-r)!}\right)R^n,$$

when $n\geq N$, we have by induction, using Lemma 9, that

$$\frac{c_{m,n}}{c_n}\leq\frac{(K3^rc(R))^{m-1}}{(m-1)!}$$

when $n\geq N$, and

$$\lim_{n\to\infty}c_{m,n}/c_n=\frac{c(R)^{m-1}}{(m-1)!}.$$

But $a_n/c_n=\sum_{m\geq 1}c_{m,n}/c_n$. This sum is majorized by the convergent sum $\sum_{m\geq 1}(K3^rc(R))^{m-1}/(m-1)!=\exp(K3^rc(R))$ when $n\geq N$. The Dominated Convergence Theorem is called to service once more to establish that $\lim_{n\to\infty}a_n/c_n=\exp c(R)=a(R)$. Now from this it follows easily that the coefficients a_j satisfy hypothesis (i) and (ii) so the conclusion follows from the earlier argument. □

Theorem 11. *Suppose that $\mathscr{S}$ has ordinary generating series $b(x)=\sum_{n\geqslant 0} b_n x^n = \exp\sum_{i\geqslant 1} d(x^i)/i$, $d(x)=\sum_{n\leqslant 0} d_n x^n$. If*
(i) $\lim_{n\to\infty} b_{n-1}/b_n = R$, $0<R\leqslant\infty$, *and*
(ii) *for some nonnegative integer r and real $K>0$ we have for large enough n and $r\leqslant m\leqslant n$ that $(m)_r b_m R^m \leqslant K(n)_r b_n R^n$,*

then $\nu(C=m)=Z_{m-1}(R)/b(R)$, for all $m\geqslant 1$, and $\mathrm{E}(C,\nu)=1+\sum_{i\geqslant 1} d(R^i)$. If all coefficients b_j are replaced by d_j in hypotheses (i) *and* (ii), *the conclusion still holds. ($Z_m(x)$ is given by eq.* (5).)

Proof. If $\lim_{x\to R} b(x)=\infty$ (which is always the case when $R=1$), then the result follows from Theorem 8. Thus, assume $\lim_{x\to R} b(x)<\infty$ and $R<1$.

The proof parallels that of Theorem 10. Rather than formulas (9) and (10) it relies on formulas (11) and (12). The presence of sums like $\sum_{i\leqslant 2} x^{i-1} d'(x^i)$ in these formulas poses a difficulty not encountered in the proof of Theorem 10. However, these sums are analytic for $|x|<R^{\frac{1}{2}}$ so they and all their derivatives have coefficients vanishingly small with respect to the coefficients of like powers of x in $b(x)$, hence may be ignored (to see this, note that by (ii) $b_n\geqslant K'n^{-r}R^{-n}$ for some $K'>0$ and large n while coefficients of a sum with radius of convergence greater than R must be $o((R+\varepsilon)^{-n})$ for some $\varepsilon>0$. The details in this first part of the proof are now easily checked.

Suppose that hypotheses (i) and (ii) are true of coefficients d_j rather than b_j. We know

$$b(x)=\exp d(x)\exp\sum_{i\geqslant 2} d(x^i)/i.$$

Let $\sum_{n\geqslant 0} e_n x^n=\exp d(x)$, $\sum_{n\geqslant 0} f_n x^n=\exp\sum_{i\geqslant 2} d(x^i)/i$. Then, as in the proof of Theorem 10, $e_n\leqslant K_e d_n$ for large n, where $K_e=\exp(K3^r d(R))$, and $\lim_{n\to\infty} e_n/d_n=\exp d(R)$. Also, since $\exp\sum_{i\geqslant 2} d(x^i)/i$ has radius of convergence $R^{\frac{1}{2}}>R$, there is a constant K_f such that $f_n\leqslant K_f d_n$, and $\lim_{n\to\infty} f_n/d_n=0$. Thus, by Lemma 9.

$$\lim_{n\to\infty} b_n/d_n=\exp d(R)\exp\sum_{n\geqslant 2} d(R^i)/i=b(R).$$

Now the coefficients b_j can be seen to satisfy hypotheses (i) and (ii), and the conclusion follows. □

Examples. Let $\mathscr{S}$ be the class of rooted forests, $a(x)$ the exponential generating series for $\mathscr{S}$, $a(x)=\exp c(x)$, where $c(x)$ is the exponential generating series for rooted trees. One of the oldest results in enumerative combinatorics is that $c(x)=xa(x)=x\exp c(x)$. From this one can show that $c_n=n^{n-1}$ so that $\lim_{n\to\infty} nc_{n-1}/c_n=1/\mathrm{e}=R$ and

$$\frac{c_k}{(k-2)!}R^k \leq A\left(\frac{c_n}{(n-2)!}\right)R^n$$

for some A, all large n and $k \leq n$, by Stirling's formula. Also, $c(x)$ is seen to be the functional inverse of xe^{-x} so $c(R) = 1$. By Theorem 10

$$\mu(C = m) = 1/(e(m-1)!), \qquad E(C, \mu) = 2.$$

Let $\mathscr{S}$ be the class of unrooted forests, $\bar{a}(x) = \sum_{n\geq 0} (\bar{a}_n/n!)x^n$ the exponential generating series for $\mathscr{S}$, $\bar{a}(x) = \exp \bar{c}(x)$, where $\bar{c}(x) = \sum_{n\geq 0} (\bar{c}_n/n!)x^n$ is the exponential generating series for unrooted trees. Clearly $c_n = n\bar{c}_n$ (c_n is as in the previous paragraph) so $\bar{c}'(x) = c(x)/x$ ($c(x)$ is taken from the previous example). Differentiate both sides of the equation $c(x) = x \exp c(x)$ and substitute $c(x)/x$ for $\exp c(x)$. We obtain $c(x)/x = (1 - c(x))c'(x)$. Hence

$$\bar{c}(x) = \int_0^x (1 - c(t))c'(t)\, dt = c(x) - \tfrac{1}{2}c(x)^2.$$

From these facts it follows that $\lim_{n\to\infty} n\bar{c}_{n-1}/\bar{c}_n = 1/e = R$, that

$$\left(\frac{\bar{c}_k}{(k-1)!}\right)R^k \leq A\left(\frac{c_n}{(n-1)!}\right)R^n$$

for some A, all large n and $k \leq n$, and that $c(R) = \frac{1}{2}$. By Theorem 10 $\mu(C = m) = 1/(2^{m-1}(m-1)!\, e^{\frac{1}{2}})$, $E(C, \mu) = \frac{3}{2}$. The latter was shown in Moon [13].

Let $\mathscr{S}$ be the class of unit interval graphs, $b(x)$ the ordinary generating series for $\mathscr{S}$, $b(x) = \exp \sum_{i\geq 1} d(x^i)/i$. Hanlon [7] shows that

$$d(x) = \frac{2x+1}{4(1-4x^2)^{\frac{1}{2}}} - \tfrac{1}{4}(1-4x)^{\frac{1}{2}}.$$

It follows that $\lim_{n\to\infty} d_{n-1}/d_n = \frac{1}{4}$ and that

$$(k)_2 d_k \leq A(n)_2 d_n$$

for some A, all large n, and $k \leq n$. By Theorem 11, $\nu(C = m) = Z_{m-1}(\frac{1}{4})/b(\frac{1}{4})$. (Hanlon shows this when $m = 1$) and $E(C, \nu) = 1 + \sum_{i\geq 1} d(1/4^i)$.

Palmer and Schwenk [14] derive the expressions for $\nu(C = m)$ and $E(C, \nu)$ in the cases of unrooted forests, rooted forests, and planted forests. Their methods are general enough to apply in all cases where $b(x)$ has a finite number of singularities, all algebraic, on its circle of convergence, with its dominant singularity at the radius of convergence (i.e., in those cases where Darboux's method works). In fact, the same generality can be obtained in the labeled case. However, when Darboux's method does not apply, another approach—viz., Theorem 10 and 11—must be used.

Let $\mathscr{S}$ consist of forests of rooted trees with the restriction that the root may not occur at the centroid of a tree with $2^k + 1$ vertices when $k \geq 1$. (A centroid is a vertex for which the sum of distances to leaves is minimal; in trees with an odd number of vertices, there is a unique centroid.) The fraction of rooted trees with $n = 2^k + 1$ vertices having their root at the centroid is $1/n$. Thus

$$c(x) = \sum_{n \geq 0} (n^{n-1}/n!)x^n - \sum_{n=2^k+1} (n^{n-2}/n!)x^n.$$

Now it is still no problem to estimate the growth of coefficients c_n—in fact, $c_n \sim n^{n-1}$—but Darboux's method no longer applies to $a(x) = \exp(c(x))$ because, by the Hadamard Gap Theorem (see [16]), $\sum_{n=2^k+1} (n^{n-2}/n!)x^n$ has the circle $|z| = 1/e$ as a natural boundary and $\sum_{n \geq 0} (n^{n-1}/n!)x^n$ has only one singularity on $|z| = 1/e$ so no cancellation occurs. However, Theorem 10 does apply and from it $\mu(\boldsymbol{C} = m)$ and $\mathrm{E}(\boldsymbol{C}, \mu)$ may be computed.

3. Conclusion

Proposition 2 is often used in combinatorial enumeration. Proposition 3 and Lemma 9, which are not much more difficult than Proposition 2, are not often used, even though they will work in many cases where more sophisticated techniques fail. They will frequently shorten and simplify arguments when generating series converge at their radius of convergence (this pertains not just to component distribution problems).

We close with some related problems.

(1) As noted, Wright [19] shows that if $\mu(\boldsymbol{C} = 1) = 1$, then $a(x)$ has radius of convergence 0. Is it true that when $\mu(\boldsymbol{C} = 1)$ exists and $a(x)$ has radius of convergence 0 that $\mu(\boldsymbol{C} = 1) = 1$? Is the analogous theorem true in the unlabeled case?

(2) Can Theorems 7 and 8 be strengthened to give asymptotic expressions for $\mu(\boldsymbol{C} = m)$, etc., in terms of $a(x)$ and $b(x)$?

(3) If $\mu(\boldsymbol{C} = 1)$ exists, then does $\mu(\boldsymbol{C} = m)$ exist for $m > 1$? Does $\mathrm{E}(\boldsymbol{C}, \mu)$? What happens in the unlabeled case?

References

[1] E.A. Bender, An asymptotic expansion for coefficients of some formal power series, J. London Math. Soc. (2) 9 (1974) 451–458.

[2] E.A. Bender, Asymptotic methods in enumeration, SIAM Rev. 16 (1974) 485–515.

[3] E.A. Bender and J.R. Goldman, Enumerative uses of generating functions, Indiana Univ. Math. J. 20 (1971) 753–765.

[4] K.J. Compton, A logical approach to asymptotic combinatorics I: first order properties, Adv. in Math., to appear.

[5] K.J. Compton, A logical approach to asymptotic combinatorics II: monadic second order properties, Adv. in Math., to appear.

[6] I.P. Goulden and D.M. Jackson, Combinatorial Enumeration (Wiley, New York, 1983).

[7] P. Hanlon, Counting interval graphs, Trans. Amer. Math. Soc. 272 (1982) 383–426.

[8] F. Harary and E. Palmer, Graphical Enumeration (Academic Press, New York, 1973).
[9] W.K. Hayman, A generalization of Stirling's formula, J. Reine Angew. Math. 196 (1956) 67–95.
[10] L. Katz, The problem of indecomposability of random mapping functions, Ann. Math. Statistics 26 (1955) 512–517.
[11] M.D. Kruskal, The expected number of components under a random mapping function, Amer. Math. Monthly 61 (1954) 392–397.
[12] R.H. Möhring, On the distribution of locally undecomposable relations and independence systems, Methods Oper. Res. 42 (1981) 33–48.
[13] J.W. Moon, Counting labelled trees, Canad. Math. Congress Monograph 1 (1970).
[14] E.M. Palmer and A.J. Schwenk, On the number of trees in a random forest, J. Combin. Theory Ser. B 27 (1979) 109–121.
[15] G. Pòlya and Szegö, Problems and Theorems in Analysis I (Springer, Berlin, 1970).
[16] W. Rudin, Real and Complex Analysis (McGraw-Hill, New York, 1966).
[17] L.A. Shepp and S.P. Lloyd, Ordered cycle lengths in random permutations, Trans. Amer. Math. Soc. 121 (1966) 340–357.
[18] E.M. Wright, Asymptotic relations between enumerative functions in graph theory, Proc. London Math. Soc. (3) 20 (1970) 558–572.
[19] E.M. Wright, A relationship between two sequences, Proc. London Math. Soc. (3) 17 (1967) 296–304.
[20] E.M. Wright, A relationship between two sequences III, J. London Math. Soc. 43 (1968) 720–724.