Systems of Functional Equations

Michael Drmota
Department of Discrete Mathematics, Technical University of Vienna, Wiedner Hauptstrasse 8-10, A-1040, Vienna, Austria

Received 7 February 1996; revised 27 August 1996

ABSTRACT: The aim of this paper is to discuss the asymptotic properties of the coefficients of generating functions which satisfy a system of functional equations. It turns out that under certain general conditions these coefficients are related to the distribution of a multivariate random variable that is asymptotically normal. As an application it turns out that the distribution of the terminal symbols in context-free languages is typically asymptotically normal. © 1997 John Wiley & Sons, Inc. *Random Struct. Alg.*, **10**, 103–124 (1997)

Key Words: generating function; functional equation; asymptotic expansion; singularity analysis

1. INTRODUCTION

Let $\mathscr{Y}$ be a set of combinatorial objects, i.e., every element $o \in \mathscr{Y}$ has a size $|o|$ such that the numbers $y_n = |\{o \in \mathscr{Y}: |o| = n\}$ are finite for every nonnegative integer n. Especially, if $\mathscr{Y}$ has a recursive description, then the generating function $y(x) = \sum_{o \in \mathscr{Y}} x^{|o|} = \sum_{n \geq 0} y_n x^n$ satisfies a functional equation quite frequently.

In order to give a first example and to motivate the topic of this paper, let us consider the system of planted plane trees. They can be recursively characterized in the following way: A planted plane tree contains a root which is followed by a finite number $k \geq 0$ of planted plane trees. Hence the corresponding generating function $y(x)$ satisfies the functional equation

$$y(x) = x + xy(x) + xy(x)^2 + \cdots = \frac{x}{1 - y(x)},$$

Contract grant sponsor: Austrian Science Foundation; contract grant number: P10187-PHY

which gives

$$y(x) = \frac{1-\sqrt{1-4x}}{2}$$

and

$$y_n = [x^n]y(x) = \frac{1}{n}\binom{2n-2}{n-1},$$

which are the Catalan numbers. (As usual, $[x^n]y(x)$ denotes the nth coefficient of the power series $y(x)$.)

This easy problem gets a little bit more involved if one is not only interested in the numbers y_n of planted plane trees of size n but in the numbers y_{nk} of planted plane trees of size n with exactly k leaves. Similarly as above the corresponding generating function $y(x,z) = \sum_{n,k} y_{nk} x^n z^k$ satisfies the functional equation

$$y(x,z) = xz + xy(x,z) + xy(x,z)^2 + \cdots = xz + \frac{xy(x,z)}{1-y(x,z)}.$$

Thus

$$y(x,z) = \frac{\sqrt{xz - x + 1 - \sqrt{(xz-x+1)^2 - 4xz}}}{2}$$

is explicit, but this representation is not useful to obtain proper representations for y_{nk}. However, we can use Lagrange's inversion formula. From

$$z = y\cdot\left(\frac{x}{1-\frac{x}{1-y}}\right)^{-1},$$

we obtain

$$[z^k]y(x,z) = \frac{1}{k}[v^{k-1}]x^k\left(1-\frac{x}{1-v}\right)^{-k}$$

and consequently

$$\begin{aligned} y_{nk} &= [x^n z^k]y(x,z) \\ &= \frac{1}{k}[v^{k-1}x^n]x^k\left(1-\frac{x}{1-v}\right)^{-k} \\ &= \frac{1}{k}[v^{k-1}]\binom{n-1}{n-k}\frac{1}{(1-v)^{n-k}} \\ &= \frac{1}{k}\binom{n-1}{n-k}\binom{n-2}{k-1} = \frac{1}{n-1}\binom{n-1}{n-k}\binom{n-1}{k}. \end{aligned}$$

Hence, by Stirling's approximation formula

$$y_{nk} \sim \frac{\pi}{n^2} 2^{2n-1} e^{-4(k-n/2)^2/n}, \tag{1.1}$$

i.e., these numbers admit a Gaussian limiting distribution.

It is also possible to obtain a functional equation for the generating function $y(x,z)$ for the numbers y_{nk} of planted plane trees of size n and k nodes of outdegree d, where $d \geq 0$ is a fixed nonnegative integer:

$$y(x,z) = \frac{x}{1 - y(x,z)} - xy(x,z)^d + xzy(x,z)^d.$$

However, there is generally neither an explicit representation for $y(x,z)$ nor a method (similar to the above Lagrange inversion) to obtain proper exact formula for y_{nk}. Nevertheless, there is a general theorem [4] saying that if $y = y(x,z)$ satisfies a functional equation of the form $y = F(x,y,z)$ (with certain regularity conditions), then $y(x,z)$ can be represented in the form

$$y(x,z) = g(x,z) - h(x,z)\sqrt{1 - x/f(z)} \tag{1.2}$$

with proper analytic functions $g(x,z)$, $h(x,z)$, and $f(z)$, which leads to asymptotic expansions for the coefficients y_{nk} of $y(x,z)$, which is of type (1.1), and to asymptotic normality.

The aim of this paper is to discuss generating functions that satisfy a system of functional equations and to obtain (multivariate) asymptotic expansions for the coefficients. Systems of functional equations naturally appear in more involved tree enumeration problems, e.g., every word in an unambiguous context-free language can be uniquely represented by its derivation tree. The number of functional equations for the generating functions of the corresponding counting problem (i.e., counting the number of words of length n with given distribution of terminal symbols) equals the number of nonterminal symbols (see Section 3.2).

As already mentioned, if the generating function $y(x,z)$ satisfies a single functional equation (with some additional assumptions), then the problem is already solved in [4]. We will present a more systematic approach than in [4]. The first step is to use the implicit function theorem and the Weierstrass preparation theorem to provide a representation of the form (1.2). The next step is to apply a transfer lemma by Flajolet and Odlyzko [7] to obtain an asymptotic expansion for $y_n(z) = [x^n]y(x,z)$. Finally the coefficient y_{nk} is represented by Cauchy's formula, and a standard saddle point method is applied to get an asymptotic expansion for y_{nk} which is always of a form similar to (1.1) and is related to a Gaussian limiting distribution. Therefore, the main problem is to reduce a system of functional equations to a single one. Of course, this is always possible by using an elimination procedure. However, there are cases (see [6]) where the final limiting distribution is not normal. Nevertheless, we will provide a sufficient condition such that a proper elimination process terminates at a single functional equation for which the above concept can be applied.

This paper is organized in the following way. In Section 2 we state and discuss our main theorem. In Section 3 several examples and applications (such as context-free languages and tree enumeration problems) are treated. Section 4 is devoted to a systematic treatment of a single functional equation. Finally, Section 5 describes the reduction of a system of functional equations to a single functional equation.

2. RESULTS

Let $\mathbf{F}(x,\mathbf{y},\mathbf{z}) = (F_1(x,\mathbf{y},\mathbf{z}),\ldots,F_N(x,\mathbf{y},\mathbf{z}))'$ a vector[1] of functions $F_j(x,\mathbf{y},\mathbf{z})$, $1 \le j \le N$, with complex variables x, $\mathbf{z} = (z_1,\ldots,z_k)'$, $\mathbf{y} = = (y_1,\ldots,y_N)'$, which are analytic around 0 and satisfy $F_j(0,\mathbf{0},\mathbf{0}) = 0$, $1 \le j \le N$. We will be interested in the analytic solution $\mathbf{y} = \mathbf{y}(x,\mathbf{z}) = (y_1(x,\mathbf{z}),\ldots,y_N(x,\mathbf{z}))'$ of the functional equation

$$\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z}) \tag{2.1}$$

with $\mathbf{y}(0,\mathbf{0}) = \mathbf{0}$, i.e, the (unknown) functions $y_j = y_j(x,\mathbf{z})$, $1 \le j \le N$ satisfy the system of functional equations

$$\begin{aligned} y_1 &= F_1(x, y_1, y_2, \ldots, y_N, \mathbf{z}),\\ y_2 &= F_2(x, y_1, y_2, \ldots, y_N, \mathbf{z}),\\ &\vdots\\ y_N &= F_N(x, y_1, y_2, \ldots, y_N, \mathbf{z}). \end{aligned}$$

Before stating Theorem 1 we have to introduce some notations. The dependency graph $G_{\mathbf{F}} = (V, E)$ of a system of functional equations $\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z})$ consists of vertices $V = \{y_1, y_2, \ldots, y_N\}$ and a directed edge (y_i, y_j) is contained in E if and only if $F_i(x,\mathbf{y},\mathbf{z})$ really depends on y_j.

In order to calculate specific parameters (a vector μ and a matrix σ^2) it is also necessary to solve the system of functional equations

$$\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z}), \tag{2.2}$$

$$0 = \det\big(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x,\mathbf{y},\mathbf{z})\big), \tag{2.3}$$

where $\mathbf{z} = (z_1,\ldots,z_k)'$ is the vector of variables and $x = x(\mathbf{z})$ and $\mathbf{y} = \mathbf{y}(\mathbf{z})$ are the unknown functions; $\mathbf{I}$ denotes the identity matrix. The parameters of interest are

$$\mu = -\frac{x_{\mathbf{z}}(\mathbf{1})}{x(\mathbf{1})} \quad \text{and} \quad \sigma^2 = -\frac{x_{\mathbf{zz}}(\mathbf{1})}{x(\mathbf{1})} + \mu'\mu + \operatorname{diag}(\mu), \tag{2.4}$$

where

$$x_{\mathbf{z}} = \left(\frac{\partial x}{\partial z_1}, \ldots, \frac{\partial x}{\partial z_k}\right)$$

[1]We always identify a k-dimensional vector $\mathbf{a} = (a_1,\ldots,a_k)'$ with a $(k \times 1)$-matrix, i.e., a column. Furthermore, if A is a matrix, then A' denotes the transposed matrix.

denotes the vector of first partial derivatives

$$x_{\mathbf{zz}} = \left(\frac{\partial^2 x}{\partial z_i \, \partial z_j} \right)_{1 \le i, j \le k}.$$

the matrix of second partial derivatives of the solution $x = x(\mathbf{z})$.

In order to simplify the formulation of the following theorem we will make the following restriction on $\mathbf{F}$. We will call the system of functional equations $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z})$ of simple type if there exist $(k+1)$-dimensional cones $C_j \subseteq \mathbf{R}^{k+1}$, $1 \le j \le N$, (centered at $\mathbf{0}$) such that, for every j, all Taylor coefficients $y_{j,n\mathbf{m}}$ (where n and $\mathbf{m} = (m_1, \dots, m_k)'$ are nonnegative integers resp. integral lattice points) of the solutions

$$y_j(x, \mathbf{z}) = \sum_{n, \mathbf{m}} y_{j, n\mathbf{m}} x^n \mathbf{z}^{\mathbf{m}} \qquad (1 \le j \le N)$$

of (2.1) with $(n, \mathbf{m}) \in C_j$ are nonzero provided that $n, m_1, \dots, m_k$ are sufficiently large. (As usual $\mathbf{z}^{\mathbf{m}}$ denotes $z_1^{m_1} \cdots z_k^{m_k}$.)

In applications we are interested in the coefficients $y_{j,n\mathbf{m}}$ have a combinatorial interpretation. Therefore, it is more or less a combinatorial question to decide whether $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z})$ is of simple type or not. We also want to note that there are special cases (see [12]) where it is quite easy to find sufficient conditions for a single functional equation to be of simple type.

Our main result is stated in the following theorem.

Theorem 1. *Let* $\mathbf{F}(x, \mathbf{y}, \mathbf{z}) = (F_1(x, \mathbf{y}, \mathbf{z}), \dots, F_N(x, \mathbf{y}, \mathbf{z}))'$ *be analytic functions around* $x = 0$, $\mathbf{z} = (z_1, \dots, z_k)' = \mathbf{0}$, $\mathbf{y} = (y_1, \dots, y_N)' = \mathbf{0}$ *such that all Taylor coefficients are nonnegative, that* $\mathbf{F}(0, \mathbf{y}, \mathbf{z}) \equiv \mathbf{0}$, *that* $\mathbf{F}(x, \mathbf{0}, \mathbf{z}) \not\equiv \mathbf{0}$, *and that there exists* j *with* $\mathbf{F}_{y_j y_j}(x, \mathbf{y}, \mathbf{z}) \not\equiv \mathbf{0}$. *Furthermore, assume that the system* $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z})$ *is of simple type and that the region of convergence of* $\mathbf{F}$ *is large enough such that there exists a nonnegative solution* $x = x_0$, $\mathbf{y} = \mathbf{y}_0$ *of the system of equations*

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{1}),$$
$$0 = \det\bigl(\mathbf{I} - \mathbf{F}_{\mathbf{y}}(x, \mathbf{y}, \mathbf{1})\bigr),$$

inside it.

If the dependency graph $G_{\mathbf{F}} = (V, E)$ *of the system*

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z}) \tag{2.5}$$

in the unknown functions

$$\mathbf{y} = \mathbf{y}(x, \mathbf{z}) = \bigl(y_1(x, \mathbf{z}), \dots, y_N(x, \mathbf{z})\bigr)'$$

is strongly connected and if the matrix σ^2 [*defined in* (2.4)] *is regular then the Taylor coefficients of*

$$y_j(x, \mathbf{z}) = \sum_{n, \mathbf{m}} y_{j, n\mathbf{m}} x^n \mathbf{z}^{\mathbf{m}} \qquad (1 \le j \le N)$$

are asymptotically given by

$$y_{j,n\mathbf{m}} = \frac{a_j x_0^{-n}}{\sqrt{2^{k+2}\pi^{k+1}n^{k+3}}}\left(\exp\left(-\frac{1}{2n}(\mathbf{m}-\mu n)'(\sigma^2)^{-1}(\mathbf{m}-\mu n)\right)+\mathscr{O}(n^{-1/2})\right) \tag{2.6}$$

uniformly for all $n, \mathbf{m}$ *for which* $y_{j,n\mathbf{m}} \neq 0$; *the numbers* a_j, $1 \le j \le N$, *are given by* $a_j = |b_j|$, $1 \le j \le N$, *where* $\mathbf{b} = (b_1, \dots, b_N)'$ *is a solution of*

$$\begin{aligned} \left(\mathbf{I} - \mathbf{F_y}(x_0, \mathbf{y}_0, \mathbf{1})\right)\mathbf{b} &= \mathbf{0}, \\ \mathbf{b}'\mathbf{F_{yy}}(x_0, \mathbf{y}_0, \mathbf{1})\mathbf{b} &= -2\mathbf{F}_x(x_0, \mathbf{y}_0, \mathbf{1}). \end{aligned} \tag{2.7}$$

Remark 1. Theorem 1 can also be interpreted in the following sense. Let $\mathbf{X}_n = (X_{1n}, \dots, X_{kn})'$ be a sequence of discrete random vectors with distribution

$$\mathbf{Pr}\{\mathbf{X}_n = \mathbf{m}\} = \frac{y_{1,n\mathbf{m}}}{y_{1,n}},$$

where

$$y_{1,n} = \sum_{\mathbf{m}} y_{1,n\mathbf{m}}$$

is the coefficient of $y_1(x, \mathbf{1})$. Then the asymptotic expansion (2.6) is exactly a local limit theorem for $\mathbf{X}_n$, i.e., $\mathbf{X}_n$ is asymptotically normal with expected value and covariance matrix

$$\mathbf{EX}_n = \mu n + \mathscr{O}(1) \qquad \text{and} \qquad \mathbf{VX}_n = \sigma^2 n + \mathscr{O}(1).$$

(Compare with [2,4].)

Remark 2. The assumptions on $\mathbf{F}$ of Theorem 1 are quite natural. For example, if $\mathbf{F}(x, \mathbf{0}, \mathbf{z}) \equiv \mathbf{0}$ then the only analytic solution of (2.1) is $\mathbf{y}(x, \mathbf{z}) \equiv \mathbf{0}$. Furthermore, if $\mathbf{F}_{y_j y_j} \equiv 0$ for all j, then (2.1) is a linear system in $\mathbf{y}$. Hence by Cramer's rule $y_j(x, \mathbf{z})$ is a quotient of analytic functions in x, $\mathbf{z}$, and our methods will not apply. However, if $f(\mathbf{z}_0)$ denotes the smallest positive zero of the denominator of $y_j(x, \mathbf{z}_0)$ (which is assumed to be simple), then $y_j(x, \mathbf{z}_0)$ has a polar singularity which varies in $\mathbf{z}_0$. At this point we can apply results of [1,2] and obtain essentially the same asymptotic expansions for $y_{j,n\mathbf{m}}$ expansions for $y_{j,n\mathbf{m}}$ as in Theorem 1.

Remark 3. The assumption that the dependency graph $G_{\mathbf{F}} = (V, E)$ of $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z})$ is strongly connected means that there is no subsystem of equations that can be solved independently from the others. Of course, this is a restriction. However, if this assumption is not satisfied, several cases may appear; in fact Theorem 1 need not remain true. In [6] the simplest case of two functional equations is discussed in detail. The corresponding limiting distribution is either discrete, or χ^2, or normal.

Remark 4. If the matrix σ^2 is singular, then (2.6) is not applicable. However, a singular matrix σ^2 indicates that there are (quite) strict linear correlations between those n and $m_1,\dots,m_k$ for which $y_{j,n\mathbf{m}} \neq 0$. For example, in Section 3.2 we will count specific words of length n with m_1 letters a, m_2 letters b, and m_3 letters c. Hence we always have $n = m_1 + m_2 + m_3$, and we can easily check that the corresponding matrix σ^2 is singular. But, if we set $z_3 \equiv 1$, which means that we only count words of length n with m_1 letters a and m_2 letters b then we can apply Theorem 1. A similar procedure can be used in general. If σ^2 is a nonzero matrix, then there always is a maximal symmetric submatrix $\bar{\sigma}^2$ which is regular. Suppose that $\bar{\sigma}^2$ corresponds to columns resp. lines $l \in K \subseteq \{1,\dots,k\}$, then we only have to set $z_{l'} \equiv 1$ for $l' \notin K$, and we can apply Theorem 1 for this reduced problem. Sometimes it is now possible to recover the full information about the initial distribution (as in the above-mentioned example discussed in Section 3.2). In any case (even in the case of singular σ^2) the (full) vector of mean value $\mu = -x_{\mathbf{z}}(\mathbf{1})/x(\mathbf{1})$ has a natural interpretation. The expected values of $\mathbf{X}_n$ (introduced in Remark 1) are always given by $\mathbf{E}\mathbf{X}_n = \mu n + \mathcal{O}(1)$ and can be easily calculated.

Remark 5. The asymptotic expansion (2.6) is only relevant for those $\mathbf{m}$ which are close to μn. However, it is possible to provide a proper asymptotic expansion for all sufficiently large values of $\mathbf{m}$ if the region of convergence of $\mathbf{F}$ is large enough. If $c_1,\dots,c_k > 0$ are positive numbers and if we apply Theorem 1 to the functions $\tilde{y}_j(x, z_1,\dots,z_k) = y_j(y, c_1 z_1,\dots,c_k z_k)$, $1 \le j \le N$, which satisfy the system of equations

$$\tilde{y}_j(x,\mathbf{z}) = F\left(x, \tilde{y}_1(x,\mathbf{z}),\dots,\tilde{y}_N(x,\mathbf{z}), c_1^{-1} z_1,\dots,c_k^{-1} z_k\right) \qquad (1 \le j \le N),$$

then we obtain an asymptotic expansion for $y_{j,n\mathbf{m}} c_1^{m_1} \cdots c_k^{m_k}$ for those $\mathbf{m}$ close to

$$\tilde{\mu} = \mu(c_1,\dots,c_k) = -\left(c_j x_{z_j}(c_1,\dots,z_k)/x(c_1,\dots,c_j)\right)_{1 \le j \le N}.$$

It turns out (see [1,2,4]) that all resulting asymptotic expansions for $y_{j,n\mathbf{m}}$ are uniform for $c_1,\dots,c_k > 0$ which are contained in a compact set (provided all calculations needed can be worked out inside the region of convergence of $\mathbf{F}$). Observe that those positive real numbers resp. vectors $n', \mathbf{m}'$ which satisfy $\mathbf{m}' = \mu(c_1,\dots,c_k) n'$, in which $c_1,\dots,c_k$ vary in some range describe a cone. This justifies the restriction on systems of functional equations $\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z})$ to be of simple type.

3. APPLICATIONS

3.1. Tree Enumeration Problems

As already indicated in the Introduction tree enumeration problems are frequently related to functional equations of the form $\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z})$. Let us start with a series of examples concerning independent subsets of trees.

We consider simply generated famines $\mathcal{T}$ of rooted trees ([10]). This means that there exists a power series $\varphi(t) = 1 + c_1 t + c_2 t^2 + \cdots$ with $c_j \ge 0$ such that the generating function $y(x) = \Sigma_{n \geq 0} y_n x^n$ of the numbers

$$y_n = \sum_{T \in \mathcal{T}, |T| = n} \omega(T),$$

where

$$\omega(T) = \prod_{j \geq 0} c_j^{D_j(T)}$$

is the "weight" of a rooted tree T and $D_j(T)$ denotes the number of nodes of T with out-degree j, satisfies the functional equation

$$y(x) = x\varphi(y(x)).$$

For example, if $\varphi(t) = 1/(1-t)$, then every rooted tree T has weight $\omega(T) = 1$, and we are just considering planted plane trees as in the Introduction, or if $\varphi(t) = 1 + t^2$, then every binary tree T has weight $\omega(T) = 1$ and any other tree has weight $\omega(T) = 0$. If fact, y_n may be interpreted as the (weighted) number of rooted trees of a special type.

A subset I (of the vertex set) of a graph is called independent if no two elements of I are adjacent. Let $I_{1,m}(T)$ denote the number of independent subsets of a rooted tree T of size m such that the root is contained in the independent subset. If we set

$$y_{1,nm} = \sum_{T \in \mathcal{T}, |T| = n} \omega(T) I_{1,m}(T),$$

$$y_{2,nm} = \sum_{T \in \mathcal{T}, |T| = n} \omega(T) I_{2,M}(T),$$

then $y_{nm} = y_{1,nm} + y_{1,nm}$ is the (weighted) number of independent subsets of size m in trees of size n. In [9] it is shown that the generating functions $y_1 = y_1(x,z) = \Sigma_{n,m} y_{1,nm} x^n z^m$, $y_2 = y_2(x,z) = \Sigma_{n,m} y_{2,nm} x^n z^m$ satisfy the system of functional equations

$$y_1 = xz\varphi(x\varphi(y_1 + y_2)),$$

$$y_2 = x\varphi(y_1 + y_2).$$

The situation is quite similar if one is interested in the number of maximal independent subsets. A subset I (of the vertex set) of a graph is called maximal independent if I is independent and if every vertex which is not contained in I is adjacent to I. As above, we can introduce generating functions $y_1 = y_1(x,z)$, $y_2 = y_2(x,z)$, which count the number of maximal independent subsets in simply generated families of trees. Here we have (see [11])

$$y_1 = xz\varphi(x\varphi(y_1 + y_2)),$$

$$y_2 = x\varphi(y_1 + y_2) - x\varphi(y_2).$$

In [3] several notions weaker than the notion of maximal independence are considered. For example, a subset I (of the vertex set) of a rooted tree T is called

2-independent, if I is independent and every vertex v which is not contained in I has distance ≤ 2 to I_v, where I_v denotes the restriction of I to the path connecting the root and v and to the subtree of T rooted at v. Here the corresponding functions $y_1 = y_1(x,z)$ and $y_2 = y_2(x,z)$ satisfy

$$\begin{aligned} y_1 &= xz\varphi(x\varphi(y_1 + x\varphi(y_1 + y_2))), \\ y_2 &= x\varphi(y_1 + y_2) - x\varphi(y_2 - x\varphi(y_1 + y_2) + x\varphi(y_2)). \end{aligned}$$

In any of these cases we can apply Theorem 1. (Note that these systems of equations are of simple type if and only if $\gcd\{l > 0\colon c_l \neq 0\} = 1$.) We also want to mention that the first two systems of equations can be reduced to one equation if one considers $y = y_1 + y_2$. We obtain

$$y = xz\varphi(x\varphi(y)) + x\varphi(y),$$

resp.

$$y = xz\varphi(x\varphi(y)) + x\varphi(y) - x\varphi(y - xz\varphi(x\varphi(y))).$$

However, the third system of equations cannot be reduced to a single equation in such a simple way.

Next we want to describe a very general situation, where systems of functional equations appear. We consider rooted trees with N different types of nodes and a node of type j is followed by m_i nodes of type i, $1 \leq i \leq N$, with “weight” $c_{j,m_1,\ldots,m_N} \geq 0$. Here the generating functions

$$y_j(x, z_1, \ldots, z_N) = \sum_{m_1,\ldots,m_N} y_{j,m_1,\ldots,m_n} x^{m_1 + \cdots + m_N} z_1^{m_1} \cdots z_N^{m_N}$$

of the “weighted” numbers $y_{j,n,m_1,\ldots,m_N}$ of those trees of size n with m_i nodes of type i, $1 \leq i \leq N$, and with root type j satisfy the functional equations

$$y_j(x, \mathbf{z}) = xz_j \sum_{m_1,\ldots,m_N} c_{j,m_1,\ldots,m_N} y_1(x,\mathbf{z})^{m_1} \cdots y_N(x,\mathbf{z})^{m_N} \qquad (1 \leq j \leq N). \quad (3.1)$$

This kind of relations has already been established in [6]. However, the main scope of [6] was to determine asymptotic expansions for the number of trees in the case of two types of nodes and where the dependency graph is not strongly connected.

Note that we cannot apply Theorem 1 directly to the system (3.1) since all numbers $y_{j,nm_1\ldots,m_N}$ that are nonzero satisfy $m_1 + \cdots + m_N = n$. (Compare with Remark 4 and with the last paragraph of Section 3.2.) However, if we set $z_l \equiv 1$ for specific l, if the above system is of simple type, and if the dependency graph is strongly connected, then we can apply Theorem 1.

3.2. Context-Free Languages

We also want to mention another application of Theorem 1 to recursively defined objects, namely, to context-free languages. Our aim is to determine the distribution of the terminal symbols.

For this purpose let us consider the following easy example of a context-free grammar

$$G = (N, \Sigma, P, S)$$

with nonterminal symbols $N = \{S, T\}$, terminal symbols $\Sigma = \{a, b, c\}$, rules $P = \{S \to aSbS,\ S \to bT,\ T \to bS,\ T \to cT,\ T \to a\}$, and start symbol S. The corresponding context-free language $L(G)$ consists of all words which can be generated from S by using the rules P, e.g., the following derivation leads to a member of $L(G)$:

$$\begin{aligned} S &\to aSbS \\ &\to abTbaSbS \\ &\to abcTbaaSbSbbT \\ &\to abcabaabTbbTbba \\ &\to abcabaababbcTbba \\ &\to abcabaababbcabba. \end{aligned}$$

Obviously, the above grammar is unambiguous; any word of $l(G)$ has a unique derivation tree. Hence, if

$$s = s(u_1, u_2, u_3) = \sum_{l_1, l_2, l_3} s_{l_1 l_2 l_3} u_1^{l_1} u_2^{l_2} u_3^{l_3}$$

denotes the generating function of the numbers $s_{l_1 l_2 l_3}$ of those words in $L(G)$ with l_1 terminal symbols a, l_2 terminal symbols b, and l_3 terminal symbols c and

$$t = t(u_1, u_2, u_3) = \sum_{l_1, l_2, l_3} t_{l_1 l_2 l_3} u_1^{l_1} u_2^{l_2} u_3^{l_3}$$

denotes the corresponding generating function of the context-free language $L(G')$ of the grammar $G' = (N, \Sigma, P, T)$ (i.e., the start symbol S of G is replaced by T), then $s = s(u_1, u_2, u_3)$ and $t = t(u_1, u_2, u_3)$ satisfy the relations

$$\begin{aligned} s &= u_1 u_2 s^2 + u_2 t, \\ t &= u_2 s + u_3 t + u_1. \end{aligned}$$

We will now apply Theorem 1 to the above context-free language. Set $y_1 = y_1(x, z_1, z_2) = s(xz_1, xz_2, x)$ and $y_2 = y_2(x, z_1, z_2) = t(xz_1, xz_2, x)$. Then y_1, y_2 satisfy the system of equations

$$\begin{aligned} y_1 &= x^2 z_1 z_2 y_1^2 + x z_2 y_2, \\ y_2 &= x z_2 y_1 + x y_2 + x z_1 \end{aligned}$$

and the coefficient $y_{1, n m_1 m_2}$ of $y_1(x, z_1, z_2)$ is exactly the number of words of length n with m_1 terminal symbols a and m_2 terminal symbols b. Furthermore, all assumptions of Theorem 1 are satisfied. Especially we obtain $x_0 = 0.4658229\ldots$ and $\mu = (0.22723\ldots, 0.53813\ldots)$. This means that an average word consists of 22.7% terminal symbols a and of 53.8% terminal symbols b.

Note that it is not useful to consider the functions $y_1 = y_1(x, z_1, z_2, z_3) = s(xz_1, xz_2, xz_3)$ and $y_2 = y_2(x, z_1, z_2, z_3) = t(xz_1, xz_2, xz_3)$. Here the coefficient

$y_{1,nm_1m_2m_3}$ of $y_1(x, z_1, z_2, z_3)$ is the number of words of length n with m_1 terminal symbols a, m_2 terminal symbols b, and m_3 terminal symbols c. But we always have $m_1 + m_2 + m_3 = n$, which means that m_1, m_2, m_3 cannot vary independently if n is fixed. In this case we cannot apply Theorem 1. The corresponding matrix σ^2 is singular (compare with Remark 4 and with the last paragraph of Section 3.1.).

4. THE IMPLICIT FUNCTION THEOREM REVISITED

The most important tool in the proof of Theorem 1 is the following proposition which describes the structure of the dominating singularity of the solution $y = y(x, \mathbf{z})$ of a functional equation $y = F(x, y, \mathbf{z})$.

Proposition 1. *Let us suppose that $F(x, y, \mathbf{z})$ is an analytic function in x, $\mathbf{z} = (z_1, \ldots, z_k)'$, and y such that $F(0, \mathbf{y}, \mathbf{z}) \equiv 0$, that $F(x, \mathbf{0}, \mathbf{z}) \not\equiv 0$, and that all Taylor coefficients of F around 0 are real and nonnegative. Then the unique solution $y = y(x, \mathbf{z})$ of the functional equation*

$$y = F(x, y, \mathbf{z}) \tag{4.1}$$

with $y(0, \mathbf{z}) = 0$ is analytic around 0 and has nonnegative Taylor coefficients around 0.

Furthermore, if we assume that the region of convergence of $F(x, y, \mathbf{z})$ is large enough such that there exist nonnegative solutions $x = x_0$ and $y = y_0$ of the system of equations

$$y = F(x, y, \mathbf{1}),$$
$$1 = F_y(x, y, \mathbf{1}),$$

with $F_x(x_0, y_0, \mathbf{1}) \neq 0$ and $F_{yy}(x_0, y_0, \mathbf{1}) \neq 0$, then there exist functions $f(\mathbf{z})$, $g(x, \mathbf{z})$, $h(x, \mathbf{z})$ which are analytic around $x = x_0$, $\mathbf{z} = \mathbf{1}$ such that $y(x, \mathbf{z})$ is analytic for $|x| < x_0$ and $|z_j| \le 1$, $1 \le j \le k$, and has a representation of the form

$$y(x, \mathbf{z}) = g(x, \mathbf{z}) - h(x, \mathbf{z})\sqrt{1 - \frac{x}{f(\mathbf{z})}} \tag{4.2}$$

locally around $x = x_0$, $\mathbf{z} = \mathbf{1}$. We have $f(\mathbf{1}) = x_0$, $g(f(\mathbf{z}), \mathbf{z}) = y(f(\mathbf{z}), \mathbf{z})$, and

$$h(f(\mathbf{z}), \mathbf{z}) = \sqrt{\frac{2f(\mathbf{z})F_x(f(\mathbf{z}), g(f(\mathbf{z}), \mathbf{z}), \mathbf{z})}{F_{yy}(f(\mathbf{z}), g(f(\mathbf{z}), \mathbf{z}), \mathbf{z})}}\,.$$

If $\mathbf{z}$ is real and close to $\mathbf{1}$, then $f(\mathbf{z})$ is the radius of convergence of the power series $x \mapsto y(x, \mathbf{z})$. Moreover, if $\arg(x - f(\mathbf{z})) \neq 0$, then (4.2) provides a local analytic continuation of $y(x, \mathbf{z})$.

Remark 6. Note that the assumptions $F_x(x_0, y_0, \mathbf{1}) \neq 0$ and $F_{yy}(x_0, y_0, \mathbf{1}) \neq 0$ are really necessary to obtain a representation of the form (4.2). If $F_x(x_0, y_0, \mathbf{1}) = 0$, then $F(x, y, \mathbf{z})$ (and $y(x, \mathbf{z})$) would not depend on x. Furthermore, if $F_{yy}(x_0, y_0, \mathbf{1})$

$= 0$, then F is of the form

$$F(x,y,\mathbf{z}) = yF_1(x,\mathbf{z}) + F_2(x,\mathbf{z}); \tag{4.3}$$

then

$$y(x,\mathbf{z}) = \frac{F_2(x,\mathbf{z})}{1 - F_1(x,\mathbf{z})}, \tag{4.4}$$

which is explicit and surely not of the form (4.2). However, representation of the form (4.4) (where $F_1(x,\mathbf{z}) \not\equiv 0$) usually leads to the same asymptotic expansions for the coefficients of $y(x,\mathbf{z})$ as is the case covered by Proposition 1; compare with Remark 2.

Proof. First, we show that there exists a unique (analytic) solution $y = y(x,\mathbf{z})$ of $y = F(x,y,\mathbf{z})$ with $y(0,\mathbf{z}) = 0$. Since $F(0,y,\mathbf{z}) = 0$, it follows that the functional mapping

$$y(x,\mathbf{z}) \mapsto F(x, y(x,\mathbf{z}), \mathbf{z})$$

is a contraction for small x. Thus the iteratively defined functions $y_0(x,\mathbf{z}) \equiv 0$ and

$$y_{n+1}(x,\mathbf{z}) = F(x, y_n(x,\mathbf{z}), \mathbf{z}) \qquad (n \geq 0)$$

converge uniformly to a limit function $y(x,\mathbf{z})$, which is the unique solution of (4.1). By definition it is clear that $y_n(x,\mathbf{z})$ is an analytic function around 0 and has real and nonnegative Taylor coefficients. Consequently, the uniform limit $y(x,\mathbf{z})$ is analytic, too, with nonnegative Taylor coefficients.

It is also possible to use the implicit function theorem. Since

$$F_y(0,0,\mathbf{z}) = 0 \neq 1,$$

there exists a solution $y = y(x,\mathbf{z})$ of (4.1) that is analytic around 0.

However, it is very useful to know that all Taylor coefficients of $y(x,\mathbf{z})$ are nonnegative. Namely, it follows that if $y(x,\mathbf{z})$ is regular at $x_0, \mathbf{z}_0$ which are real and positive, then $y(x,\mathbf{z})$ is regular for all $x, \mathbf{z}$ with $|x| \leq x_0$ and $|\mathbf{z}| \leq \mathbf{z}_0$. Therefore, we will now suppose that x_0 and $\mathbf{z}_0$ are real and positive.

For a moment let $\mathbf{z}_0 \leq \mathbf{1}$ be fixed. Let $f_0(\mathbf{z}_0)$ denote the radius of convergence of $y(x,\mathbf{z}_0)$. It is well known that $x_0 = f_0(\mathbf{z}_0)$ is a singular value of $y(x,\mathbf{z}_0)$. The mapping

$$x \mapsto F_y(x, y(x,\mathbf{z}_0), \mathbf{z}_0)$$

is strictly increasing for real and nonnegative x as long as $y(x,\mathbf{z}_0)$ is regular. Note that $F_y(0, y(0,\mathbf{z}_0), \mathbf{z}) = 0$. As long as $F_y(x, y(x,\mathbf{z}_0), \mathbf{z}_0) < 1$, it follows from the implicit function theorem that $y(x,\mathbf{z}_0)$ is regular even in a neighborhood of x. Hence there exists a finite limit point x_0 such that $\lim_{x \to x_0^-} y(x,\mathbf{z}_0) = y_0$ is finite and satisfies $F_y(x_0, y_0, \mathbf{1}) = 1$. If $y(x,\mathbf{z}_0)$ were regular at $x = x_0$, then

$$y_x(x_0,\mathbf{z}_0) = F_x(x_0, y(x,\mathbf{z}_0), \mathbf{z}_0) + F_y(x_0, y(x_0,\mathbf{z}_0), \mathbf{z}_0) y_x(x_0,\mathbf{z}_0)$$

would imply $F_x(x_0, y(x_0, \mathbf{z}_0), \mathbf{z}_0) = 0$, which is surely not true. Thus $y(x, \mathbf{z}_0)$ is singular at $x = x_0$ [i.e., $x_0 = f_0(\mathbf{z}_0)$ is the radius of convergence] and $y(x_0, \mathbf{z}_0)$ is finite.

Now, let us consider the equation $y - F(x, y, \mathbf{z}) = 0$ around $x = x_0$, $y = y_0$, $\mathbf{z} = \mathbf{1}$. We have $1 - F_y(x_0, y_0, \mathbf{1}) = 0$ but $-F_{yy}(x_0, y_0, \mathbf{1}) \neq 0$. Hence by the Weierstrass preparation theorem (see [8]) there exist functions $H(x, y, \mathbf{z})$, $p(x, \mathbf{z})$, $q(x, \mathbf{z})$, which are analytic around $x = x_0$, $y = y_0$, $\mathbf{z} = \mathbf{1}$, and satisfy $H(x_0, y_0, \mathbf{1}) \neq 1$, $p(x_0, \mathbf{1}) = q(x_0, \mathbf{1}) = 0$, and

$$y - F(x, y, \mathbf{z}) = H(x, y, \mathbf{z})\big((y - y_0)^2 + p(x, \mathbf{z})(y - y_0) + q(x, \mathbf{z})\big)$$

locally around $x = x_0$, $y = y_0$, $\mathbf{z} = \mathbf{1}$. Since $F_x(x_0, y_0, \mathbf{1}) \neq 0$ we also have $q_x(x_0, \mathbf{1}) \neq 0$. This means that any analytic function $y = y(x, \mathbf{z})$ which satisfies $y = F(x, y, \mathbf{z})$ in a subset of a neighbourhood of $x = x_0$, $\mathbf{z} = \mathbf{1}$ with $x_0, \mathbf{1}$ on its boundary is given by

$$y(x, \mathbf{z}) = y_0 - \frac{p(x, \mathbf{z})}{2} \pm \sqrt{\frac{p(x, \mathbf{z})^2}{4} - q(x, \mathbf{z})}\,.$$

Since $p(x_0, \mathbf{1}) = 0$ and $q_x(x_0, \mathbf{1}) \neq 0$, we have

$$\frac{\partial}{\partial x}\left(\frac{p(x, \mathbf{1})^2}{4} - q(x, \mathbf{1})\right)_{x = x_0} \neq 0,$$

too. Again by the Weierstrass preparation theorem there exist functions $K(x, \mathbf{z})$ and $r(\mathbf{z})$ which are analytic around $x = x_0$, $\mathbf{z} = \mathbf{1}$ such that $K(x_0, \mathbf{1}) \neq 0$, $r(\mathbf{1}) = 0$, and

$$\frac{p(x, \mathbf{z})^2}{4} - q(x, \mathbf{z}) = K(x, y)((x - x_0) + r(\mathbf{z}))$$

locally around $x = x_0$, $\mathbf{z} = \mathbf{1}$. This finally leads to a local representation of $y = y(x, \mathbf{z})$ of the kind

$$y(x, \mathbf{z}) = g(x, \mathbf{z}) - h(x, \mathbf{z})\sqrt{1 - \frac{x}{f(\mathbf{z})}}\,, \tag{4.5}$$

in which $g(x, \mathbf{z})$, $h(x, \mathbf{z})$, and $f(\mathbf{z})$ are analytic around $x = x_0$, $\mathbf{z} = \mathbf{1}$ and satisfy $g(x_0, \mathbf{1}) = y_0$, $h(x_0, \mathbf{1}) < 0$, and $f(\mathbf{1}) = x_0$.

Our starting point was to show that for positive real $\mathbf{z}_0$ the radius of convergence $x_0 = f_0(\mathbf{z}_0)$ of $y(x, \mathbf{z}_0)$ can be extracted from the system of equations $y = F(x, y, \mathbf{z}_0)$, $1 = F_y(x, y, \mathbf{z}_0)$. Thus $f_0(\mathbf{z}_0) = f(\mathbf{z}_0)$ for positive real $\mathbf{z}_0$. Similarly we can show that $x = f_0(\mathbf{z})$ is a singular point of $y(x, \mathbf{z})$ if $\mathbf{z}$ is close to the reals, where $x = f_0(\mathbf{z})$ (and $y = g_0(\mathbf{z})$) are the solutions of the system of equations

$$\begin{aligned} y &= F(x, y, \mathbf{z}_0), \\ 1 &= F_y(x, y, \mathbf{z}_0) \end{aligned} \tag{4.6}$$

Thus $f_0(\mathbf{z}) = f(\mathbf{z})$ even in this case. Therefore, we can calculate the derivatives of $f(\mathbf{z})$ by implicit differentiation.

Furthermore, it is clear that $g(f(\mathbf{z}), \mathbf{z}) = g_0(f(\mathbf{z})) = y(f(\mathbf{z}), \mathbf{z}))$, which gives

$$g(x,\mathbf{z}) = g_0(f(\mathbf{z})) + g_{0,x}(f(\mathbf{z}),\mathbf{z})(x - f(\mathbf{z})) + O\big((x - f(\mathbf{z})^2\big).$$

In order to calculate $h(f(\mathbf{z}), \mathbf{z})$, we use Taylor's theorem

$$\begin{aligned} 0 &= F(x,\mathbf{z}, y(x,\mathbf{z})) \\ &= F_x(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))(x - f(\mathbf{y})) \\ &\quad + \tfrac{1}{2}F_{yy}(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))(y(x,\mathbf{z}) - g_0(\mathbf{z}))^2 + \cdots \qquad (4.7) \\ &= F_x(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))(x - f(\mathbf{y})) \\ &\quad + \tfrac{1}{2}F_{yy}(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z})) \times h(f(\mathbf{z}),\mathbf{z})^2(1 - x/f(\mathbf{z})) + O\big(|x - f(\mathbf{z})|^{3/2}\big) \end{aligned}$$

and by comparing the coefficients of $(x - f(\mathbf{z}))$ we immediately obtain

$$h(f(\mathbf{z}),\mathbf{z}) = \sqrt{\frac{2f(\mathbf{z})F_x(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))}{F_{yy}(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))}}. \qquad \blacksquare$$

Corollary. *Suppose that $F(x, y, \mathbf{z})$ satisfies the same assumptions as in Proposition* 1. *Then*

$$y(x,\mathbf{z}) = g(f(\mathbf{z}),\mathbf{z}) - \sqrt{\frac{2f(\mathbf{z})F_x(f(\mathbf{z}), g(f(\mathbf{z}),\mathbf{z}),\mathbf{z})}{F_{yy}(f(\mathbf{z}), g(f(\mathbf{z}),\mathbf{z}),\mathbf{z})}}\sqrt{1 - \frac{x}{f(\mathbf{z})}}$$
$$+ g_x(f(\mathbf{z}),\mathbf{z})(x - f(\mathbf{z})) + O\big(|1 - x/f(\mathbf{z})|^{3/2}\big). \qquad (4.8)$$

Proof. Since $h(x, \mathbf{z}) = h(f(\mathbf{z}), \mathbf{z}) + O(|1 - x/f(\mathbf{z})|)$, (4.8) immediately follows from Proposition 1. $\blacksquare$

Thus we have succeeded in determining the exact behavior of $y(x, \mathbf{z})$ (considered as a function in x) near its real singularity $x = f(\mathbf{z})$ at its radius of convergence if $\mathbf{z}$ is positive, real, and close to $\mathbf{1}$. Next we will show that if we assume that $y = F(x, y, \mathbf{z})$ is of simple type, then there are no other singularities on the circle $|x| = f(\mathbf{z})$.

Lemma 1. *Suppose that $F(x, y, \mathbf{z})$ satisfies the same assumptions as in Proposition* 1 *and is of simple type. Then for $\mathbf{z}$ sufficiently close to $\mathbf{1}$ the radius of convergence of the power series $x \mapsto y(x, \mathbf{z})$ is $|f(\mathbf{z})|$ and there are no other singularities on the circle of convergence $|x| = |f(\mathbf{z})|$ than $x = f(\mathbf{z})$. Furthermore, there exists $\varepsilon > 0$ such that $y(x, \mathbf{z})$ can be analytically continued to the region $|x| \le |f(\mathbf{z})| + \varepsilon$,* $\arg(x - f(\mathbf{z})) \neq 0$.

Proof. Suppose that $\mathbf{z}$ is real and positive and let $y(x, \mathbf{z})$ be represented as a power series

$$y(x,\mathbf{z}) = \sum_n y_n(\mathbf{z})x^n.$$

Then by assumption $y_n(\mathbf{z}) > 0$ for $n \ge n_0$. Hence $|y(x,\mathbf{z})| < y(|x|,\mathbf{z})$ if $x \ne |x|$ and consequently

$$\left|F_y(x, y(x,\mathbf{z}),\mathbf{z})\right| < F_y(f(\mathbf{z}), y(f(\mathbf{z}),\mathbf{z}),\mathbf{z}) = 1$$

for $|x| = f(\mathbf{z})$, $x \ne f(\mathbf{z})$. (Compare with [12].) Thus, by the implicit function theorem there are no other singularities on the circle $|x| = f(\mathbf{z})$.

Note that the local expansion (4.5) is also valid for nonreal $\mathbf{z}$. It is now an easy exercise to show that $|f(\mathbf{z})|$ is the radius of convergence $y(x,\mathbf{z})$ (considered as a function in x) if $\mathbf{z}$ is sufficiently close to the reals. Obviously, $x = f(\mathbf{z})$ is a singular point of $y(x,\mathbf{z})$. Furthermore, by continuity we also obtain $|F_y(x,\mathbf{z}, y(x,\mathbf{z}))| < 1$ for $|x| \le |f(\mathbf{z})|$ and $|x - f(\mathbf{z})| \ge \delta$, where $\delta > 0$ is sufficiently small. Thus $x = f(\mathbf{z})$ is the only singularity of $y(x,\mathbf{z})$ on the circle $|x| = |f(\mathbf{z})|$ and it is regular for $|x| < |f(\mathbf{z})|$. Furthermore, the implicit function theorem implies that $y(x,\mathbf{z})$ has an analytic continuation at $x = f(\mathbf{z})$. Since the range $|x| = |f(\mathbf{z})|$, $|x - f(\mathbf{z})| \ge \delta$ is compact there surely exists $\varepsilon > 0$ such that $y(x,\mathbf{z})$ is analytic for $|x| \le |f(\mathbf{z})| + \varepsilon$, $\arg(x - f(\mathbf{z})) \ne 0$. ■

Now we are in a position to apply a transfer lemma of Flajolet and Odlyzko [7].

Lemma 2. *Let*

$$F(x) = \sum_{n \ge 0} a_n x^n$$

be analytic in a region

$$\Delta = \{x : |x| < x_0 + \eta, \ |\arg(x - x_0)| > \vartheta\},$$

in which x_0 and η are positive real numbers and $0 < \vartheta < \pi/2$. Furthermore, suppose that there exists a real number $\alpha \notin \{0, -1, -2, \dots\}$ such that

$$F(x) = \mathscr{O}\left((1 - x/x_0)^{-\alpha}\right) \qquad (x \in \Delta).$$

Then

$$a_n = \mathscr{O}\left(z_0^{-n} n^{\alpha - 1}\right).$$

Corollary. *Suppose that $F(x,\mathbf{y},\mathbf{z})$ satisfies the same assumptions as in Proposition* 1 *and is of simple type. Set $y_n(\mathbf{z}) = [x^n]y(x,\mathbf{z})$, i.e., $y(x,\mathbf{z}) = \Sigma_{n \ge 0} y_n(\mathbf{z}) x^n$. Then*

$$y_n(\mathbf{z}) = \sqrt{\frac{2f(\mathbf{z})F_x(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))}{F_{yy}(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))}} \binom{\frac{1}{2}}{n} (-f(\mathbf{z}))^{-n} + O\left(f(\mathbf{z})^{-n} n^{-5/2}\right)$$

$$= \sqrt{\frac{f(\mathbf{z})F_x(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))}{2\pi F_{yy}(f(\mathbf{z}),\mathbf{z}, g_0(\mathbf{z}))}} f(\mathbf{z})^{-n} n^{-3/2} + O\left(f(\mathbf{z})^{-n} n^{-5/2}\right) \qquad (4.9)$$

uniformly for $\mathbf{z}$ *which are sufficiently close to* $\mathbf{1}$.

What remains is to apply a multivariate saddle point method (compare with [2] and [5]) to extract the coefficient $y_{n\mathbf{m}}$ of

$$\mathbf{y}_n(\mathbf{z}) = \sum_{\mathbf{m}} y_{n\mathbf{m}} \mathbf{z}^{\mathbf{m}}.$$

Proposition 2. *Suppose that $F(x, \mathbf{y}, \mathbf{z})$ satisfies the same assumptions as in Proposition 1 and is of simple type. If the vector μ and the matrix σ^2 are given by*

$$\mu = \left(- \frac{\partial \log f(e^{u_1}, \dots, e^{u_k})}{\partial u_j} \bigg|_{\mathbf{u}=\mathbf{0}} \right)_{1 \le j \le k}$$

and by

$$\sigma^2 = \left(- \frac{\partial^2 \log f(e^{u_1}, \dots, e^{u_k})}{\partial u_i \, \partial u_j} \bigg|_{\mathbf{u}=\mathbf{0}} \right)_{1 \le i, j \le k}$$

and if σ^2 is regular, then

$$y_{n\mathbf{m}} = \sqrt{\frac{x_0 F_x(x_0, y_0, \mathbf{1})}{2\pi F_{yy}(x_0, y_0, \mathbf{1})}} \frac{x_0^{-n}}{\sqrt{2^k \pi^k n^{k+3}}}$$
$$\times \left(\exp\left(- \frac{1}{2n} (\mathbf{m} - \mu n)' (\sigma^2)^{-1} (\mathbf{m} - \mu n) \right) + O(n^{-1/2}) \right). \quad (4.10)$$

Remark 7. Note that the preceding representation of μ and σ^2 corresponds with the definition given in (2.4). Since

$$\frac{\partial \log f(e^{u_1}, \dots, e^{u_k})}{\partial u_j} = \frac{f_{z_j}(e^{u_1}, \dots, e^{u_k}) e^{u_j}}{f(e^{u_1}, \dots, e^{u_k})}$$

and $f(\mathbf{z}) = x(\mathbf{z})$ we obtain

$$\mu = - \frac{x_{\mathbf{z}}(\mathbf{1})}{x(\mathbf{1})}.$$

Similarly σ^2 can be calculated.

Sketch of the Proof. The basic idea is to extract the coefficient $y_{n\mathbf{m}}$ of $y_n(\mathbf{z}) = \sum_{n \ge 0} y_{n\mathbf{m}} \mathbf{z}^{\mathbf{m}}$ by Cauchy's formula

$$y_{n\mathbf{m}} = \frac{1}{(2\pi)^k} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} y_n(e^{it_1}, \dots, e^{it_k})^n e^{-i(m_1 t_1 + \cdots + m_k t_k)} \, dt_1 \cdots dt_k.$$

It turns out that the major part of this integral comes from that part of integration, where t_l, $1 \le l \le k$, are very close to 0. Since

$$f(e^{it_1}, \dots, e^{it_k})^{-1} \sim f(\mathbf{1})^{-1} e^{i\mu' \mathbf{t} - \frac{1}{2} \mathbf{t}' \sigma^2 \mathbf{t}},$$

in which $\mathbf{t} = (t_1, \ldots, t_k)$, we can approximate the integrand by

$$y_n(e^{it_1}, \ldots, e^{it_k})^n e^{-i(m_1 t_1 + \cdots + m_k t_k)}$$
$$\sim n^{-3/2} \sqrt{\frac{f(\mathbf{1}) F_x(f(\mathbf{1}), g_0(\mathbf{1}), \mathbf{1})}{2\pi F_{yy}(f(\mathbf{1}), g_0(\mathbf{1}), \mathbf{1})}}\, f(\mathbf{1})^{-n} e^{i(\mu n - \mathbf{m})'\mathbf{t} - \frac{n}{2}\mathbf{t}'\sigma^2 \mathbf{t}}.$$

Hence, (4.10) follows almost directly. A detailed proof of a related problem can be found in [5]. ■

Remark 8. It is also very easy to derive asymptotic normality of $\mathbf{X}_n$ (from Remark 1) in terms of a weak limit theorem directly from (4.9) without using (4.10) (see [1, 2]).

5. REDUCTION TO A SINGLE EQUATION

In this final section we will show that the assumptions of Theorem 1 assure that we can apply the concept presented in Section 4 so that the asymptotic expansion (2.6) follows immediately from Proposition 2 [see also Remark 9 and Lemma 3 in order to evaluate the parameters needed in (2.6)].

Proposition 3. *Let* $\mathbf{y} = \mathbf{y}(x, \mathbf{z}) = (y_1(x, \mathbf{z}), \ldots, y_N(x, \mathbf{z}))'$ *be the solution of the system of functional equations* $\mathbf{y} = F(x, \mathbf{y}, \mathbf{z})$ *satisfying the same assumptions as in Theorem* 1. *If the dependency graph* $G_{\mathbf{F}}$ *of* $\mathbf{y} = F(x, \mathbf{y}, \mathbf{z})$ *is strongly connected, then there exists a function* $f(\mathbf{z})$ *and functions* $g_j(x, \mathbf{z}), h_j(x, \mathbf{z})$ $(1 \le j \le N)$ *which are analytic around* $x = x_0$, $\mathbf{z} = 1$ *such that*

$$y_j(x, \mathbf{z}) = g_j(x, \mathbf{z}) - h_j(x, \mathbf{z}) \sqrt{1 - \frac{x}{f(\mathbf{z})}}$$

locally around $x = x_0$, $\mathbf{z} = \mathbf{1}$ *with* $\arg(x - f(\mathbf{z})) \ne 0$.

Proof. For the sake of shortness and transparency we just discuss the case of three functional equations. The general case of N functional equations can be treated along the same lines.

We consider a system of functional equations for the (unknown) functions $y_j = y_j(x, \mathbf{z})$, $1 \le j \le 3$:

$$\begin{aligned} y_1 &= F(x, y_1, y_2, y_3, \mathbf{z}), \\ y_2 &= G(x, y_1, y_2, y_3, \mathbf{z}), \\ y_3 &= H(x, y_1, y_2, y_3, \mathbf{z}), \end{aligned} \tag{5.1}$$

where $\mathbf{z} = (z_1, \ldots, z_k)'$ is a k-dimensional complex vector,

$$F(0, y_1, y_2, y_3, \mathbf{z}) = G(0, y_1, y_2, y_3, \mathbf{z}) = H(0, y_1, y_2, y_3, \mathbf{z}) = 0,$$

and that F, G, H are analytic functions in $x, \mathbf{z}, y_1, y_2, Y_3$ are real and nonnegative and the regions of convergence are large enough in the sense that the following calculations can be worked out inside them. Furthermore, we may assume that

$$F_{y_1^2} \not\equiv 0, \qquad G_{y_2^2} \not\equiv 0, \qquad H_{y_3^3} \not\equiv 0.$$

For example, if $F_{y_1^2} \equiv 0$ then F has the form

$$F(x, y_1, y_2, y_1, \mathbf{z}) = y_1 f(x, y_2, y_3, \mathbf{z}) + g(x, y_2, y_3, \mathbf{z})$$

and y_1 can explicitly represented by

$$y_1 = \frac{g(x, y_2, y_3, \mathbf{z})}{1 - f(x, y_2, y_3, \mathbf{z})}.$$

Inserting this into the second and third equation of (5.1) yields a reduction of the number of equations.

Finally the condition that the dependency graph is connected may be translated into

$$F_{y_2} \not\equiv 0, \qquad G_{y_3} \not\equiv 0, \qquad H_{y_1} \not\equiv 0, \tag{5.2}$$

into

$$F_{y_2} \not\equiv 0, \qquad G_{y_1} \not\equiv 0, \qquad G_{y_3} \not\equiv 0, \qquad H_{y_3} \not\equiv 0, \tag{5.3}$$

or into conditions symmetric to (5.2) or (5.3).

The basic idea to extract the solutions y_1, y_2, y_3 of (5.1) is quite simple. With help of the implicit function theorem we can use the third equation of (5.1) to obtain $y_3 = y_3(x, y_1, y_2, \mathbf{z})$, where y_1 and y_2 are considered as additional variables. Next we insert this solution into the second equation of (5.1) and extract $y_2 = y_2(x, y_1, \mathbf{z})$, where y_1 is considered as an additional variable. Finally we insert $y_2(x, y_1, \mathbf{z})$ and $y_2(x, y_1, y_2(x, y_1, \mathbf{z}), \mathbf{z})$ into the first equation of (5.1) and obtain just one equation for one unknown function $y_1 = y_1(x, \mathbf{z})$. At this point we will be able to describe the singularity structure of $y_1(x, \mathbf{z})$ and hence that of $y_2(x, \mathbf{z}) = y_2(x, y_1(x, \mathbf{z}), \mathbf{z})$ and that of $y_3(x, \mathbf{z}) = y_3(x, y_1(x, \mathbf{z}), y_2(x, \mathbf{z}), \mathbf{z})$.

From $y_3 = H(x, y_1, y_2, y_3, \mathbf{z})$ we immediately obtain that its solution $y_3 = y_3(x, y_1, y_2, \mathbf{z})$ [with $y_3(0,0,0,0) = 0$] is analytic around 0 and all Taylor coefficients are real and nonnegative (see section 4). Furthermore, it can be analytically continued as long as

$$\frac{\partial}{\partial y_3}(y_3 - H(x, y_1, y_2, y_3, \mathbf{z})) = 1 - H_{y_3} \neq 0.$$

In this region we obtain by implicit differentiation

$$\frac{\partial y_3}{\partial y_1} = \frac{H_{y_1}}{1 - H_{y_3}} \qquad \text{and} \qquad \frac{\partial y_3}{\partial y_2} = \frac{H_{y_1}}{1 - H_{y_3}}.$$

Note that the Taylor coefficients of these functions are real and nonnegative, too.

Next consider the equation $y_2 = G(x, y_1, y_2, y_3(x, y_1, y_2, \mathbf{z}), \mathbf{z})$. Similarly its solution $y_2 = y_2(x, y_1, \mathbf{z})$ (with $y_2(0,0,0) = 0$) is analytic around 0 and its Taylor coefficients are real and non-negative. It can be analytically continued as long as

$$\frac{\partial}{\partial y_2}\left(y_2 - G(x, y_1, y_2, y_3(x, y_1, y_2, \mathbf{z}), \mathbf{z})\right)$$
$$= 1 - G_{y_2} - G_{y_3}\frac{\partial y_3}{\partial y_2} = 1 - G_{y_2} - G_{y_3}\frac{H_{y_2}}{1 - H_{y_3}} \neq 0.$$

Again by implicit differentiation

$$\frac{\partial y_2}{\partial y_1} = \frac{G_{y_1} + G_{y_3}\frac{\partial y_3}{\partial y_1}}{1 - G_{y_2} - G_{y_3}\frac{\partial y_3}{\partial y_2}} = \frac{G_{y_1} + G_{y_3}\frac{H_{y_1}}{1 - H_{y_3}}}{1 - G_{y_2} - G_{y_3}\frac{H_{y_2}}{1 - H_{y_3}}}.$$

As mentioned above, the next step is to consider the equation

$$y_1 = F\left(x, y_1, y_2(x, y_1, \mathbf{z}), y_3(x, y_1, y_2(x, y_1, \mathbf{z}), \mathbf{z}), \mathbf{z}\right). \tag{5.4}$$

Its solution $y_1 = y_1(x, \mathbf{z})$ is exactly the unknown function we are looking for. [Obviously, if we insert this function $y_1(x, \mathbf{z})$ into $y_2(x, y_1, \mathbf{z})$, we obtain the unknown function $y_2(x, \mathbf{z}) = y_2(x, y_1(x, \mathbf{z}), \mathbf{z})$. Similarly $y_3(x, \mathbf{z}) = y_3(x, y_1(x, \mathbf{z}), y_2(x, \mathbf{z}), \mathbf{z})$.]

By the implicit function theorem the solution $y_1(x, \mathbf{z})$ is analytic as long as

$$J(x, \mathbf{z}) = F_{y_1} + F_{y_2}\frac{\partial y_2}{\partial y_1} + F_{y_3}\left(\frac{\partial y_3}{\partial y_2} + \frac{\partial y_3}{\partial y_2}\frac{\partial y_2}{\partial y_1}\right)$$
$$= F_{y_1} + F_{y_2}\frac{G_{y_1} + G_{y_3}\frac{H_{y_1}}{1 - H_{y_3}}}{1 - G_{y_2} - G_{y_3}\frac{H_{y_2}}{1 - H_{y_3}}}$$
$$+ F_{y_3}\frac{H_{y_1}}{1 - H_{y_3}} + F_{y_3}\frac{H_{y_2}}{1 - H_{y_3}}\frac{G_{y_1} + G_{y_3}\frac{H_{y_1}}{1 - H_{y_3}}}{1 - G_{y_2} - G_{y_3}\frac{H_{y_2}}{1 - H_{y_3}}}$$
$$\neq 1.$$

Clearly we also have to check that $H_{y_3} \neq 1$ and that $G_{y_2} + G_{y_3}H_{y_2}/(1 - H_{y_3}) \neq 1$. However, this will follow almost automatically as we will see in a moment.

For a moment, let $\mathbf{z} = (z_1, \ldots, z_k) = (1, \ldots, 1)$ be fixed and suppose that x is real and nonnegative. Since $J(0, \mathbf{1}) = 0$ and the mapping $x \mapsto J(x, \mathbf{1})$ is strictly increasing there surely exists a minimal solution $x = x_0$ of the equation $J(x, \mathbf{1}) = 1$. (Note that we have now used the assumption that the regions of convergence of F, G, and H

are large enough.) Furthermore, it follows that

$$H_{y_3} = H_{y_3}(x_0, y_1(x_0, \mathbf{1}), y_2(x_0, \mathbf{1}), y_3(x_0, \mathbf{1}), \mathbf{1}) < 1 \tag{5.5}$$

and that

$$G_{y_2} + G_{y_3} \frac{H_{y_2}}{1 - H_{y_3}} < 1. \tag{5.6}$$

This can be checked in the following way. It is clear that (5.5) can only fail if a term of the form

$$\frac{1}{1 - H_{y_3}}$$

does not appear in $J(x, \mathbf{1})$ resp. (5.6) can only fail if a term of the form

$$\frac{1}{1 - G_{y_2} - G_{y_3} \frac{H_{y_2}}{1 - H_{y_3}}}$$

does not appear in $J(x, \mathbf{1})$. However, it is easy to check that this cannot occur if the corresponding dependency graph is strongly connected.

Hence we are in a situation as described in Section 4. There exist analytic functions $g_1(x, \mathbf{z})$, $h_1(x, \mathbf{z})$, $f(\mathbf{z})$ [around $x = x_0$, $\mathbf{z} = 1$ satisfying $f(\mathbf{1}) = x_0$] such that $y_1(x, \mathbf{z})$ admits the representation

$$y_1(x, \mathbf{z}) = g_1(x, \mathbf{z}) - h_1(x, \mathbf{z}) \sqrt{1 - \frac{x}{f(\mathbf{z})}} . \tag{5.7}$$

Moreover, as mentioned above, (5.7) yields similar representations for $y_2(x, \mathbf{z})$ and $y_3(x, \mathbf{z})$. From (5.6) it follows that $y_2(x, y_1, \mathbf{z})$ is analytic around $x = x_0$, $y_1 = y_1(x_0, \mathbf{1}) = g_1(x_0, \mathbf{1})$, $\mathbf{z} = \mathbf{1}$. Hence by Taylor's theorem

$$\begin{aligned} y_2(x, \mathbf{z}) &= y_2(x, y_1(x, \mathbf{z}), \mathbf{z}) \\ &= y_2\left(x, g_1(x, \mathbf{z}) - h_1(x, \mathbf{z}) \sqrt{1 - \frac{x}{f(\mathbf{z})}}, \mathbf{z}\right) \\ &= g_2(x, \mathbf{z}) - h_2(x, \mathbf{z}) \sqrt{1 - \frac{x}{f(\mathbf{z})}}, \end{aligned}$$

in which $g_2(x, \mathbf{z})$ and $h_2(x, \mathbf{z})$ are analytic functions (around $x = x_0$, $\mathbf{z} = \mathbf{1}$). A similar representation holds for $y_3(x, \mathbf{z})$:

$$y_3(x, \mathbf{z}) = g_3(x, \mathbf{z}) - h_3(x, \mathbf{z}) \sqrt{1 - \frac{x}{f(\mathbf{z})}} . \qquad \blacksquare$$

Remark 9. In the preceding proof the function $f(\mathbf{z})$ has been extracted by considering the functional equation (5.4). We also want to mention that (if the dependency graph is strongly connected) if is also possible to obtain $f(\mathbf{z})$ by using the following system of equations:

$$y_1 = F(x, y_1, y_2, y_3, \mathbf{z}),$$
$$y_2 = G(x, y_1, y_2, y_3, \mathbf{z}),$$
$$y_3 = H(x, y_1, y_2, y_3, \mathbf{z}),$$
$$0 = \begin{vmatrix} F_{y_1} - 1 & F_{y_2} & F_{y_3} \\ G_{y_1} & G_{y_2} - 1 & G_{y_3} \\ H_{y_1} & H_{y_2} & H_{y_3} - 1 \end{vmatrix},$$

where $\mathbf{z}$ is considered as the only variable and $y_1 = y_1(\mathbf{z})$, $y_2 = y_2(\mathbf{z})$, $y_3 = y_3(\mathbf{z})$, and $x = x(\mathbf{z}) = f(\mathbf{z})$ are the unknown functions. This is also generally true. The system (2.2), (2.3), i.e., $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z})$, $0 = \det(\mathbf{I} - \mathbf{F_y}(x, \mathbf{y}, \mathbf{z}))$, which was used to define μ and σ^2 in Section 2 can always be used to extract $x = x(\mathbf{z}) = f(\mathbf{z})$.

Set $D(x, \mathbf{y}, \mathbf{z}) = \det(\mathbf{I} - \mathbf{F_y}(x, \mathbf{y}, \mathbf{z}))$. Then by implicit differentiation

$$\mathbf{y_z} = \mathbf{F}_x x_{\mathbf{z}} + \mathbf{F_z} + \mathbf{F_y} \mathbf{y}_z, \tag{5.8}$$

$$0 = D_x x_{\mathbf{z}} + D_{\mathbf{z}} + D_{\mathbf{y}} \mathbf{y_z}. \tag{5.9}$$

If the dependency graph is strongly connected it turns out that the matrix

$$\begin{pmatrix} \mathbf{I} - \mathbf{F_y} & -\mathbf{F}_x \\ D_{\mathbf{y}} & D_x \end{pmatrix}$$

is regular. Hence the above system (5.8), (5.9) can be used calculate $x_{\mathbf{z}}$ and $\mathbf{y_z}$. Similarly we obtain $x_{\mathbf{zz}}$. Note that if we are only interested in μ (which means that we only have to calculate $x_{\mathbf{z}}$), then there is a very quick way of calculation. Since $D(x, \mathbf{y}, \mathbf{z}) = \det(\mathbf{I} - \mathbf{F_y}) = 0$, there always exists a vector $\mathbf{a} = \mathbf{a}(\mathbf{z})$ such that $\mathbf{a}'(\mathbf{I} - \mathbf{F_y}) = \mathbf{0}$. (In fact, if the dependency graph is strongly connected then $\mathbf{I} - \mathbf{F_y}$ has rank $N - 1$, i.e., $\mathbf{a}$ is unique up to a nonzero factor.) Hence we obtain

$$x_{\mathbf{z}} = -\frac{\mathbf{a}'\mathbf{F_z}}{\mathbf{a}'\mathbf{F}_x}$$

from (5.8).

Finally, we can now prove that (2.7) has a solution $\mathbf{b}$.

Lemma 3. *By assuming the assumptions of Theorem* 1 *the system* (2.7) *has exactly two solutions* $\mathbf{b} = \pm(a_1, \dots, a_N)'$, *where* $a_j > 0$, $1 \le j \le N$, *and* $y_j(x, \mathbf{1})$ *are locally represented by*

$$y_j(x, \mathbf{1}) = c_j - a_j\sqrt{1 - x/x_0} \qquad (1 \le j \le N) \tag{5.10}$$

for certain positive real numbers c_j, $1 \le j \le N$.

Proof. By applying Taylor's theorem to $F_j(x, \mathbf{y}, \mathbf{1})$ and inserting (5.10) [as in (4.7)] we see that $\mathbf{a} = (a_1, \dots, a_N)'$ and $-\mathbf{a}$ satisfy (2.7). Moreover, since $\mathbf{I} - \mathbf{F}_y(x(\mathbf{z}), \mathbf{y}(\mathbf{z}), \mathbf{z})$ has rank $N - 1$ (by the connectedness of $G_{\mathbf{F}}$) these are the only solutions. ■

ACKNOWLEDGMENTS

The author is indebted to Philippe Flajolet for valuable discussions on the topic of this paper and to an anonymous referee for pointing out several improvements of the presentation.

REFERENCES

[1] E. A. Bender, Central and local limit theorems applied to asymptotic enumeration, *J. Combinat. Theory. Ser. A*, **15**, 91–111 (1973).

[2] E. A. Bender and L. B. Richmond, Central and local limit theorems applied to asymptotic enumeration. II: Multivariate generating function, *J. Combinat. Theory. Ser. B*, **34**, 255–265 (1983).

[3] M. Drmota, On generalized Fibonacci numbers of trees, in *Applications of Fibonacci Numbers*, Kluwer, Dordrecht, 1990, Vol. 3, pp. 63–76.

[4] M. Drmota, Asymptotic distributions and a multivariate Darboux method in enumeration problems, *J. Combinat. Theor. Ser. A*, **76**, 169–184 (1994).

[5] M. Drmota, A bivariate asymptotic expansion of coefficients of powers of generating functions, *Eur. J. Combinat.*, **14**, 139–152 (1994).

[6] M. Drmota and V. Vatutin, Limiting distributions in branching processes with two types of particles, in *Classical and Modern Branching Processes*, *Proc. IMA 1994*, Springer-Verlag, New York, to appear.

[7] Ph. Flajolet and A. Odlyzko, Singularity analysis of generating functions, *SIAM J. Discrete Math.*, **3**, 216–240 (1990).

[8] L. Kaup and B. Kaup, *Holomorphic Functions of Several Variables*, Studies in Mathematic, de Gruyter, Berlin, 1983.

[9] P. Kirschenhofer, H. Prodinger, and R. F. Tichy, Fibonacci numbers of graphs III: Planted plane trees, in *Fibonacci Numbers and Their Applications*, Reidel, Dordrecht/Boston, 1986, pp. 105–120.

[10] A. Meir and J. W. Moon, On the altitude of nodes in random trees, *Can. J. Math.* **30**, 997–1015 (1978).

[11] A. Meir and J. W. Moon, On maximal independent sets of nodes in trees, *J. Graph. Theor.* **12**, 265–283 (1988).

[12] A. Meir and J. W. Moon, On an asymptotic method in enumeration, *J. Combinat. Theor. Ser. A*, **51**, 77–89 (1989).