Institut für Geometrie
Technische Universität Wien
Michael Drmota and Bernhard Gittenberger

Asymptotics in Combinatorics and Applications

Proposal for a Research Area

1 Introduction

1.1 General Remarks

The basic goal of **combinatorics** is more or less **enumeration**. It provides certain tools to make an enumeration problem feasible. Of course, it is not always possible to get explicit expressions for the numbers of interest. Moreover, even if there is an explicit expression, e.g.

$$a_{n,k} = \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} \binom{n+l-k-1}{n-k} \binom{2n+l-2}{k-1},$$

it may be useless for any application. Therefore, one is also interested in the asymptotic behaviour of the numbers of interest.

Mostly, the driving forces in combinatorics are specific applications, e.g. the analysis of certain algorithms – say Quicksort – is more or less a combinatorial problem. Especially in applications the order of magnitude of certain numbers is more important than an explicit form.

Therefore, our main goal is to provide general tools for obtaining **asymptotic expansions** for the solutions of specific combinatorial enumeration problems which are (mostly) motivated by certain applications, e.g. the analysis of algorithms.

This proposal is organized in the following way. In section 1.2 we give an illustrative example of a combinatorial problem on permutations which leads us to our main tool, to **generating functions** and asymptotic methods which use analytic properties of generating functions. In section 1.3 we provide a systematic method how combinatorial enumeration problems can be translated to generating functions. Finally, in section 1.4 we describe two classical analytic methods, the **saddle point method** and **singularity analysis**, which can be used to get asymptotic expansions of coefficients of generating functions from their analytic properties.

In section 2 we present six topics which have been discussed in the literature and where the applicants have worked on, too. Based on this description we formulate certain aims we want to treat within the next years and describe the kind of methods we want to use (section 3). Of course, we have several international contacts and cooperations which are listed in section 4.

It should be further noted that there was a **FWF-project**

Multivariate asymptotische Entwicklungen
P 10187-MAT, Jan. 1995 – July 1997, Oct. 1997 – April 1999

guided by M. Drmota, where similar topics have been discussed. Within this project the following papers [26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 51, 52, 53, 55, 56, 74] have been written. ([37] should become a short book and is still in progress.)

We will frequently refer to these papers in the sequel and we also enclose the final report (**Endbericht**) of this project.

This project was very successful. We treated all proposed problems and extended them considerably, e.g. we additionally treated problems concerning **digital expansions** ([29, 32, 33]). The number theoretic point of view of digital expansions will be mainly discussed in **another FWF-project**

Statistical Properties of Number Systems

which will be a part of the research area **"Numbertheoretical Algorithms"** which has been granted in October 1999. Only one (more combinatorially) aspect concerning digital expansions will be discussed in this proposal, too.

The topics presented in this proposal evolved as continuations of parts of the FWF-project *Multivariate asymptotische Entwicklungen*, P 10187-MAT, at one side and of some new ones at the other side. Please note that there is **no overlap** with the FWF-project *Statistical Properties of Number Systems* (with only one exception concerning a combinatorial question on digitial expansion, see section 3.6.)

1.2 A First Example: Permutations

In this section we want to present an easy example on permutations showing the mighty of generating functions and the usefulness of analytic methods which provide asymptotic expansions of the numbers of interest. We also provide some applications of these asymptotic expansions.

Let $\mathcal{S}_n$ denote the set of permutations of $\{1,2,\ldots,n\}$, i.e., the set of bijective mappings on $\{1,2,\ldots,n\}$. It is well known that $p_n = |\mathcal{S}_n| = n!$. We now recall the fact that every permutation $\pi \in \mathcal{S}_n$ has a unique cyclic representation, i.e., the digraph G_π with vertices $1,2,\ldots,n$ and edges $(i,\pi(i))$, $1 \le i \le n$, consists of cycles. Let $\mathcal{S}_{n,k}$ denote the set of permutations $\pi \in \mathcal{S}_n$ which consist of exactly k cycles, $1 \le k \le n$. The numbers $p_{n,k} = |\mathcal{S}_{n,k}|$ are given by $p_{n,k} = s_{n,k}(-1)^{n+k}$, where $s_{n,k}$ are the Stirling numbers of the first kind, defined by

$$u(u-1)(u-2)\cdots(u-n+1) = \sum_{k=0}^{n} s_{n,k}u^k. \tag{1}$$

They have an explicit representation of the form

$$s_{n,k} = \sum_{0\le j\le h\le n-k} (-1)^{j+h}\binom{h}{j}\binom{n-1+h}{n-k+h}\binom{2n-k}{n-k-h}\frac{(h-j)^{n-k+h}}{h!}.$$

So, $p_{n,k}$ is given by

$$p_{n,k} = \sum_{0\le j\le h\le n-k} (-1)^{j+h+n+k}\binom{h}{j}\binom{n-1+h}{n-k+h}\binom{2n-k}{n-k-h}\frac{(h-j)^{n-k+h}}{h!}.$$

This representation is of course useless if one wants to know the order of magnitude of $p_{n,k}$.

From (1) we get

$$u(u+1)(u+2)\cdots(u+n-1) = \sum_{k=0}^{n} p_{n,k}u^k$$

and consequently

$$p(x,u) = \sum_{n,k\ge 0} p_{n,k}\frac{x^n}{n!}u^k = \frac{1}{(1-x)^u} = \exp\left(u\log\frac{1}{1-x}\right). \tag{2}$$

This kind of function is a generating function for the numbers $p_{n,k}$. (More precisely, it is bivariate generating function which is exponential in n and ordinary in k, see below.) In the next section we point out that this representation can be directly obtained by the fact that every permutation has a unique cyclic representation combined with some universal tools which translate combinatorial constructions to generating functions.

The next step is extract the *coefficient* of u^k

$$[u^k]p(x,u) = \sum_{n\geq 0} p_{n,k}\frac{x^n}{n!} = \frac{1}{k!}\left(\log\frac{1}{1-x}\right)^k,$$

i.e., we are led to get information about the coefficients of powers of a fixed function. Here we can use a theorem of Drmota and Soria [40].

Theorem 1 *Assume that $y(x)$ is analytic in a region $|x| < x_0+\varepsilon$, $x \notin [x_0, x_0+\varepsilon)$ ($\varepsilon > 0$) such that*

$$y(x) = -a\log\left(1-\frac{x}{x_0}\right) + C + \mathcal{O}(|\log^{-2}(1-x/x_0)|)$$

for some real constants $a>0$, C. Then we have

$$[x^n]y(x)^k = \frac{a^k\log^k n}{nx_0^n\Gamma(\frac{k}{\log n})}\exp\left(\frac{C}{a}\frac{k}{\log n}\right)\left(1+\mathcal{O}\left(k\left(\frac{\log\log n}{\log n}\right)^2\right)\right) \tag{3}$$

uniformly for $k = \mathcal{O}((\frac{\log n}{\log\log n})^2)$.

Hence, we obtain

$$p_{n,k} \sim \frac{(\log n)^k}{k!\,n\Gamma(\frac{k}{\log n})} \tag{4}$$

uniformly or $k = \mathcal{O}((\frac{\log n}{\log\log n})^2)$. (Of course, Stirling numbers are very well discussed in the literature, so Moser and Wyman already obtained asymptotic expansions for Stirling numbers, see [80, 81]. The reason why we chose this introductory example is that it perfectly fits to a general concept of combinatorial constructions and generating functions which will described in the next section 1.3.)

Equation (4) is also useful to see which k is *typical* for permutations $\pi \in \mathcal{S}_n$. $p_{n,k}$ is maximal if $k \sim \log n$ and we have a local expansion of the form

$$p_{n,k} \sim \frac{n!}{\sqrt{2\pi\log n}}e^{-\frac{(k-\log n)^2}{2\log n}}$$

if $k - \log n = O(\log n)$. This asymptotic estimate suggests that there is a central limit law behind which is in fact true. By a theorem of Flajolet and Soria [45] it just follows from a representation of the kind (2) that

$$\sum_{k\leq \log n + x\sqrt{\log n}} \frac{p_{n,k}}{n!}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\,dt + o(1).$$

We now turn to a completely different topic, namely to binary search trees. (They will be treated in more detail in sections 2.5 and **??**.) There, for every permutation $\pi \in \mathcal{S}_n$ a binary tree with n internal nodes is constructed. If we assume that every permutation is equally likely then any parameter of these binary trees can be seen as a random variable. For example, let $U_k(n)$ denote the number of external nodes at level k (in a random binary search tree of size n). In what follows we derive a relation for the expected value $\mathbf{E}U_k(n)$. For this purpose we introduce the generating functions

$$y_k(x,u) = \sum_{n\geq 0} \mathbf{E}u^{U_k(n)} x^n.$$

By the recursive structure of binary search trees we observe the following recurrence for $y_k(x,u)$. We have $y_0(x,u) = u + \frac{x}{1-x}$ and

$$\frac{\partial y_{k+1}(x,u)}{\partial x} = y_k(x,u)^2$$

with the initial condition $y_k(0,u) = 1$ for $(k > 0)$.

Obviously the derivative

$$G_k(x) = \left.\frac{\partial y_k(x,u)}{\partial u}\right|_{u=1} = \sum_{n\geq 0} \mathbf{E}U_k(n)\, x^n,$$

is the generating function for $\mathbf{E}U_k(n)$.

By using the above recurrence we directly get

$$G'_{k+1}(x) = \frac{2}{1-x}\, G_k(x)$$

(with $G_0(x) \equiv 1$ and $G_k(0) = 0$ for $k > 0$). The solutions of this recurrence is given by:

$$G_k(x) = \frac{2^k}{k!}\left(\log\frac{1}{1-x}\right)^k$$

Thus we get

$$\mathbf{E}U_k(n) = 2^k p_{n,k}$$

and of course the asymptotic expansion

$$\mathbf{E}U_k(n) \sim \frac{2^k(\log n)^k}{k!\, n\, \Gamma\left(\frac{k}{\log n}\right)}. \tag{5}$$

So we get the *average profile* of binary search trees. Especially, if $k - 2\log n = O(\sqrt{\log n})$ then

$$\mathbf{E}U_k(n) \sim \frac{n}{\sqrt{4\pi\log n}} e^{-\frac{(k-2\log n)^2}{4\log n}}.$$

1.3 Generating Functions and Combinatorial Constructions

We now want to indicate that there are lots of combinatorial problems (including the above example of section 1.2) where we naturally have generating functions.

We shortly introduce the concept of combinatorial structures described in [95]. A combinatorial structure $\mathcal{C}$ is a set of objects $o \in \mathcal{C}$ with size $|o|$ such that the set $\mathcal{C}_n = \{o \in \mathcal{C} \,|\, |o| = n\}$ of objects of size n is finite. For any combinatorial structure we can adjoin an ordinary generating function

$$c(x) = \sum_{o\in\mathcal{C}} x^{|o|} = \sum_{n\geq 0} c_n x^n \tag{6}$$

and an exponential generating function

$$\hat{c}(x) = \sum_{o\in\mathcal{C}} \frac{x^{|o|}}{(|o|)!} = \sum_{n\geq 0} c_n \frac{x^n}{n!}, \tag{7}$$

where $c_n = |\mathcal{C}_n|$ denotes the number of objects o of size $|o| = n$. Usually the ordinary generating function is used for unlabelled structures and the exponential generating function for labelled structures.

The essential point is that there are combinatorial constructions (such as disjoint union, product, sequence of, set of, multiset of, substitution,...) which correspond to simple transformations for the generating functions. Table 1 provides a list of such admissible constructions.

For the sake of shortness we just describe the product in more detail. For unlabelled structures $\mathcal{A}, \mathcal{B}$ we use the set theoretic product $\mathcal{S} = \mathcal{A} \times \mathcal{B}$ and define the size of a pair $(a, b) \in \mathcal{S}$ by $|(a, b)| := |a| + |b|$. Hence

$$\begin{aligned} s(x) &= \sum_{(a,b)\in\mathcal{A}\times\mathcal{B}} x^{|a|+|b|} \\ &= \sum_{a\in A} x^{|a|} \cdot \sum_{b\in A} x^{|b|} \\ &= a(x)b(x) \end{aligned}$$

or equivalently

$$s_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

For labelled structures the situation is a little bit different. If an object $a \in \mathcal{A}$ has size $|a| = k$ and an object $b \in \mathcal{B}$ has size $|b| = m$ then a is labelled with the numbers $1, 2, \ldots, k$ and b with the numbers $1, 2, \ldots, m$. If we form the product $\mathcal{S} = \mathcal{A} * \mathcal{B}$ then we have to consider pairs $(a, b) \in \mathcal{A} \times \mathcal{B}$ but labelled with numbers $1, 2, \ldots, k + m$. The natural way to do this is consider any choice of k labels $1 \le l_1 < l_1 < \cdots < l_k \le k + m$ and label a with $l_1, l_2, \ldots, l_k$ instead of $1, 2, \ldots, k$ and do the same with b and the remaining labels. Hence, we get (in a natural way) $\binom{k+m}{k}$ different labelled pairs (a, b) in $\mathcal{S}$. For the counting problem this means

$$s_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}$$

or equivalently for the exponential generating functions

$$\hat{s}(x) = \hat{a}(x)\hat{b}(x).$$

(For more details see [95].)

EXAMPLE 1 : Let $\mathcal{P}$ denote the set of planted plane rooted trees. In order to count the number of trees of sixe n we will use the ordinary generating function $p(x)$. Planted plane trees can be recursively described as a node followed by a sequence of planted plane trees, or in terms of combinatorial constructions

$$\mathcal{P} = \circ * \texttt{sequence}(\mathcal{P}) = \circ + \circ * \mathcal{P} + \circ * \mathcal{P}^2 + \cdots \tag{8}$$

where $\circ$ stands for the combinatorial structure of a node, with generating function $\circ(x) = x$. This translates to $p(x) = x/(1 - p(x))$ which gives $p(x) = (1 - \sqrt{1 - 4x})/2$. Now it is an easy exercise by using Lagrange's inversion formula or the binomial series expansion to obtain that the number p_n of planted plane trees is given by the Catalan number

$$p_n = [x^n]p(x) = \frac{1}{n}\binom{2n-2}{n-1}. \tag{9}$$

It is easy to extend this expample to count the numbers $p_{n,k}$ of trees with n nodes and k leaves. We only have to mark leaves in this combinatorial construction (20) with an additional variable $[u]$:

$$\mathcal{P} = [u] \circ + \circ * \mathcal{P} + \circ * \mathcal{P}^2 + \cdots$$

	unlabelled case (ogf)	labelled case (egf)
$\mathcal{S} = \mathcal{A} + \mathcal{B}$	$s(x) = a(x) + b(x)$	$\hat{s}(x) = \hat{a}(x) + \hat{b}(x)$
$\mathcal{S} = \mathcal{A} * \mathcal{B}$	$s(x) = a(x) * b(x)$	$\hat{s}(x) = \hat{a}(x) * \hat{b}(x)$
$\mathcal{S} = \texttt{sequence}(\mathcal{A})$	$s(x) = \frac{1}{1-a(x)}$	$\hat{s}(x) = \frac{1}{1-\hat{a}(x)}$
$\mathcal{S} = \texttt{cycle}(\mathcal{A})$	$s(x) = \sum\limits_{k\geq 1} \frac{\varphi(k)}{k} \log \frac{1}{1-a(x^k)}$	$\hat{s}(x) = \log \frac{1}{1-\hat{a}(x)}$
$\mathcal{S} = \texttt{set}(\mathcal{A})$	$s(x) = \exp\left(\sum\limits_{k\geq 1} \frac{(-1)^{k+1}}{k} a(x^k)\right)$	$\hat{s}(x) = \exp(\hat{a}(x))$

Table 1: Admissible constructions

and we obtain for the bivariate genarting function $p(x,u) = \sum\limits_{n,k} p_{n,k} x^n u^k$

$$p(x,u) = ux + x\frac{p(x,u)}{1-p(x,u)}.$$

Here it is again possible to obtain an explicit representation for $p_{n,k}$:

$$p_{n,k} = \frac{1}{n-1}\binom{n-1}{n-k}\binom{n-1}{k}.$$

Example 2 : Let $\mathcal{S}$ denote the set of all permutations (or equivalently the set of all (labelled) digraphs such that every node has indegree and outdegree equal to 1). Hence, in this language

$$\mathcal{S} = \texttt{set}(\texttt{cycle}(\circ))$$

which translates to

$$\hat{s}(x) = \sum_{n\geq 0} s_n \frac{x^n}{n!} = \exp(\log\frac{1}{1-x}) = \frac{1}{1-x}.$$

As above, we can formally mark a parameter in the constructions by a symbol like $[u]$. And this marking directly leads to the bivariate generating function for the number of objects according to their size and the value of the parameter of interest.

For example, if we are interested in the number of cycles in permutations we have to mark the cycles in the combinatorial construction and obtain

$$\mathcal{S} = \texttt{set}([u]\texttt{cycle}(\circ)).$$

Formally this leads to

$$\hat{s}(x,u) = \exp\left(u \log\frac{1}{1-x}\right)$$

which is exactly the generating function $\hat{s}(x,u) = p(x,u)$ from above.

1.4 Asymptotic Methods

As indicated in the previous section there are lots of combinatorial problems for which it is possible to get implicit or explicit representations for generating functions.

We now present two analytic tools for extracting the asymptotic behaviour of coefficents of analytic functions: the saddle point method and singularity analysis.

In what follows we will always assume that the coefficients are non-negative and that we are dealing with convergent power series. Hence Cauchy's formula

$$y_n = [x^n]y(x) = \frac{1}{2\pi i} \int\limits_{|x|=r} \frac{y(x)}{x^{n+1}}\, dx \tag{10}$$

applies for $0 < r < R$, where $R > 0$ denotes the radius of convergence of $y(x)$. Since we have the a-priori information $y_n \ge 0$ we get $\max\limits_{|x|=r} y(x) = y(r)$ and so we can estimate $y_n \le y(r)r^{-n}$. The minimum of $y(r)r^{-n}$ is obtained for a saddle point $\rho_n \in (0, R)$ of $y(x)x^{-n}$. Thus we have $y_n \le \frac{y(\rho_n)}{\rho_n^n}$, where $\rho_n \in (0, R)$ is defined by

$$\mu(\rho_n) = \left[\frac{\partial}{\partial v} \log y(\rho_n e^v)\right]_{v=0} = \frac{\rho_n y'(\rho_n)}{y(\rho_n)} = n. \tag{11}$$

There is a large scale of functions where $y(x)x^{-n}$ is concentrated at the saddle point ρ_n on the circle contour $|x| = \rho_n$ such that a usual saddle point approximation in the sense of Laplace applies:

$$y_n \sim \frac{y(\rho_n)\rho_n^{-n}}{\sqrt{2\pi\sigma^2(\rho_n)}}, \qquad (n \to \infty) \tag{12}$$

where $\sigma^2(\rho)$ is defined by

$$\sigma^2(\rho) = \left[\frac{\partial^2}{\partial v^2} \log y(\rho e^v)\right]_{v=0} = \frac{\rho^2 y''(\rho)}{y(\rho)} - \left(\frac{\rho y'(\rho)}{y(\rho)}\right)^2 + \frac{\rho y'(\rho)}{y(\rho)}. \tag{13}$$

Such functions are called **admissible** in the sense of Hayman [64]. For example, $y(x) = e^x$ or $y(x) = e^{e^x}$ is admissible, and there is a list of construction principles, how to build up admissible functions ([64]). (We will provide a more detailed description of admissible functions in sections 2.3 and 3.3.) For example, if we apply (12) for $y(x) = e^x$ we get Stirling's approximation formula

$$n! \sim n^n e^{-n}\sqrt{2\pi n}.$$

There is also a very important case, where the saddle point method always applies, namely for powers of functions $y(x)^k$, where n and k are of the same order of magnitude (compare with Good [58], Greene and Knuth [62], and Drmota [25])

Theorem 2 *Let $g_n, y_n \ge 0$ be two sequences of non-negative real numbers such that*

$$d = gcd\{n \,|\, y_n > 0\} = 1$$

and that the generating functions

$$y(x) = \sum_{n\ge0} y_n x^n \quad \textit{and} \quad g(x) = \sum_{n\ge0} g_n x^n$$

have positive radius of convergence R resp. R'. Let $0 < a < b$ such that $a > \lim\limits_{r\to0} \frac{ry'(r)}{y(r)}$ and $b < \min\left(\lim\limits_{r\to R} \frac{ry'(r)}{y(r)}, R'\right)$. Then the coefficient

$$y_{nk} = [x^n]g(x)y(x)^k$$

has the following asymptotic behaviour

$$y_{nk} = \frac{g(\rho)y(\rho)^k}{\rho^n\sqrt{2\pi\sigma^2(\rho)k}}\left(1+\mathcal{O}\left(\frac{1}{\sqrt{k}}\right)\right) \quad (k\to\infty) \tag{14}$$

uniformly for $\frac{n}{k}\in[a,b]$*, where* $\rho=\rho_{\frac{n}{k}}$ *is (the saddle point) defined by*

$$\frac{n}{k} = \frac{\rho y'(\rho)}{y(\rho)} \tag{15}$$

and $\sigma^2(\rho)$ *is given by (13).*

There is an another important class of functions which can be treated by analytic methods but contains no admissible functions. Let $y(x)=(1-x)^{-\alpha}$ $(\alpha>0)$ and let us try to apply (12). The saddle point evaluates to $\rho_n=\left(1+\frac{\alpha}{n}\right)^{-1}\sim 1-\frac{\alpha}{n}$. Hence we obtain

$$\frac{y(\rho_n)\rho_n^{-n}}{\sqrt{2\pi\sigma^2(\rho_n)}} \sim \frac{n^{\alpha-1}}{\sqrt{2\pi}\alpha^{\alpha-\frac12}e^{-\alpha}}, \tag{16}$$

whereas the real coefficient satisfies

$$y_n = \binom{\alpha+n-1}{n} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}. \tag{17}$$

Of course, (16) and (17) are different and hence $y(x)=(1-x)^{-\alpha}$ is not admissible. Nevertheless (16) is a proper formula to get a first insight into the behaviour of y_n. (By the way the constant $\sqrt{2\pi}\alpha^{\alpha-\frac12}e^{-\alpha}$ is the first term in Stirling's approximation formula for $\Gamma(\alpha)$.)

However there exist methods to find the right asymptotic behaviour of y_n for functions like $y(x)=(1-x)^{-\alpha}$. Flajolet and Odlyzko [44] developed a very powerful method, called Singularity Analysis. A typical theorem reads like the following one. The essential tool for the proof is the use of a part of an Hankel contour around the singularity.

Theorem 3 *Let* $y(x)$ *be analytic in a region* $|x|<R+\varepsilon$, $|\arg(x-R)|>\phi$ $(\varepsilon>0$, $0\le\phi<\frac{\pi}{2})$ *such that*

$$y(x) \sim \frac{1}{(1-x/R)^\alpha}\left(\log\frac{1}{1-x/R}\right)^\beta\left(\log\left(\frac{R}{x}\log\frac{1}{1-x/R}\right)\right)^\gamma \quad (x\to R)$$

for some real $\alpha\notin\{0,-1,-2,\ldots\}$ *and real* β,γ*. Then*

$$y_n = [x^n]y(x) \sim R^{-n}\frac{n^{\alpha-1}}{\Gamma(\alpha)}(\log n)^\beta(\log\log n)^\gamma \qquad (n\to\infty).$$

The same implication holds if asymptotic equivalence ($\sim$) is replaced by big-$\mathcal{O}$ or by small-o.

We just indicate a simple application of this theorem. If we know that a function $y(x)$ has a local expansion of the kind

$$y(x) = c_0 - c_1\sqrt{1-x} + c_2(1-x) + \mathcal{O}\left((1-x)^{3/2}\right)$$

in a region of the form $|x|<1+\varepsilon$, $|\arg(x-1)|>\phi$, then we immediately get (by using trivial asymptotics for the exact coefficient of $\sqrt{1-x}$ and Theorem 3 for the error term)

$$y_n = [x^n]y(x) = \frac{c_1}{2\sqrt{\pi}n^{3/2}} + \mathcal{O}(n^{-5/2}).$$

2 Status of Research

2.1 Galton-Watson Trees

In section 1.3 we considered the example of planted plane rooted trees $\mathcal{P}$ given by

$$\mathcal{P} = \circ * \texttt{sequence}(\mathcal{P}),$$

We now brievely describe a more general concept, namely simply generated families of trees resp. Galton-Watson trees ([78]).

The starting point is a (formal) power series

$$\varphi(t) = \varphi_0 + \varphi_1 t + \varphi_2 t^2 + \cdots \tag{18}$$

of non-negative numbers $\varphi_0 > 0$, $\varphi_1 \geq 0$, $\varphi_2 \geq 0, \ldots$, which are used to define a weight $\omega(T)$ for any planted plane (rooted) tree T by

$$\omega(T) = \prod_{i \geq 0} \varphi_i^{D_i(T)}, \tag{19}$$

where $D_i(T)$ denotes the number of nodes in T with out-degree i ($D_0(T)$ is the number of leaves, for example). Let $\mathcal{F}$ denote the set of planted plane trees with weight $\omega(T) > 0$ and $\mathcal{F}_n$ be the set of trees $T \in \mathcal{F}$ with $|T| = n$ (i.e. T consists of n nodes). Then the generating function

$$y(x) = \sum_{n \geq 0} y_n x^n \tag{20}$$

of the numbers

$$y_n = \sum_{|T|=n} \omega(T) = \sum_{T \in \mathcal{F}_n} \omega(T) \tag{21}$$

satisfies the (formal) functional equation

$$y = x\varphi(y). \tag{22}$$

Hence $\mathcal{F}$, interpreted as a combinatorial structure can be defined recursively by

$$\mathcal{F} = \circ \times \varphi(\mathcal{F}), \tag{23}$$

where $\circ$ denotes a node and $\varphi(\mathcal{F})$ a substituted structure (compare with Vitter and Flajolet [95]).

For example, let $\varphi_i = 1$ for all $i \geq 0$, then we get $\mathcal{P}$ from above. For $\varphi(t) = 1 + t^2$ we get binary trees, for $\varphi(t) = 1 + t + t^2$ Motzkin trees and so on. Gererally if $\varphi_i \in \{0, 1\}$ we are always counting special kinds of rooted trees admitting a recursive description. It is easy to extend this concept to count special parameters in these families of trees, e.g. the number of leaves, compare with section 1.3.

Alternatively if $\varphi(1) = \sum_{i \geq 0} \varphi_i = 1$ then simply generated families of trees can be considered as realizations of a (Galton-Watson) branching process where $\omega(T)$ denotes the probabilty that T occurs (see Athreya and Ney [4]). Instead of counting a parameter we directly get the probabilistic distribution of a parameter. However, both concepts are equivalent when we are only interested in discrete parameters in $\mathcal{F}_n$ (e.g. the number of leaves) resp. in the conditional distribution of the parameter of interest in realizations of size n, see Kennedy [66].

It should be further noted that the "offspring variance"

$$\sigma^2 = \sigma^2(\tau) = \sum_{i \geq 0} \frac{(i-1)^2 \varphi_i \tau^i}{\varphi(\tau)} = \frac{\tau^2 \varphi''(\tau)}{\varphi(\tau)}, \tag{24}$$

where $\tau > 0$ is the solution of $\tau\varphi'(\tau) = \varphi(\tau)$, plays an important role since it appears very often as a normalizing constant in the analysis of the asymptotic behaviour as we will see in a moment.

Simply generated families of trees (or Galton-Watson trees) have been discussed in the literature very intensively. The basis question is to describe the **asymptotic shape of trees** of size n as $n \to \infty$.

Firstly, we consider the *average depth* of a node in such trees. For this purpose we introduce the *traversing process* of a planted plane rooted tree. This means that we traverse the tree T with n nodes starting from the root and then traversing all subbranches from left to right. So we pass every edge twice and after $2n-2$ steps we are done. Let $X_n(k)$ the depth of the node which is reached after k steps of the traversing procedure. Then Aldous [1] proved the following theorem.

Theorem 4 *Suppose that the offspring variance $\sigma^2 = \sigma^2(\tau)$ exists and is finite. Then the normalized traversing process converges weakly to Brownian excursion $W^+(t)$ of duration 1:*

$$\left(\frac{\sigma}{2\sqrt{n}}X_n([2nt]), 0 \le t \le 1\right) \xrightarrow{w} (W^+(t), 0 \le t \le 1).$$

Aldous' method relies on properties of the *continuum random tree* and is purely probabilistic.

A related process was studied by Gutjahr and Pflug [63] for binary trees ($\varphi(t) = 1 + t^2$). They considered the *contour process*, i.e. the height $\hat{X}_n(k)$ of the k-th leaf (if we use the natural order from left to right.) Their method is based on explicit combinatrial sums. This approach could be generalized by Gittenberger [54] to arbitrary simply generated families of trees (Galton-Watson trees) by using generating functions and analytic methods.

Theorem 5 *Suppose that τ (the soloution of $\tau\varphi'(\tau) = \varphi(\tau)$) is smaller than the radius of convergence of $\varphi(t)$. Then the normalized contour process converges weakly to Brownian excursion $W^+(t)$ of duration 1:*

$$\left(\frac{\sigma}{2\sqrt{n}}X_n\left(\left[\frac{\varphi_0}{\varphi(\tau)}nt\right]\right), 0 \le t \le 1\right) \xrightarrow{w} (W^+(t), 0 \le t \le 1).$$

It is also indicated in [54] that Aldous's result can be reproved by these analytic methods.

The proof relies in Prohorov's theorem which says that weak convergence of stochastic processes on $C[0,1]$ is equivalent to finite dimensional (weak) convergence and tightness. The advantage of this analytic approach (via combinatorial constructions and generating functions) is that one gets more precise results, e.g. one gets error terms in the finite dimensional cases.

We now turn to another paramter of interest describing the shape of simply generated families of trees. For $k \ge 0$ let $L_n(k)$ denote the number of nodes of depth k in a tree of size n. The *process* $(L_n(k), k \ge 0)$ is also called *profile* of these trees. Drmota and Gittenberger [35] proved the following property.

Theorem 6 *Suppose that τ (the soloution of $\tau\varphi'(\tau) = \varphi(\tau)$) is smaller than the radius of convergence of $\varphi(t)$. Then the normalized profile process converges weakly to (total) local time of Brownian excursion $l(t)$:*

$$\left(\frac{2}{\sigma\sqrt{n}}L_n\left(\frac{2}{\sigma t}\sqrt{n}\right), t \ge 0\right) \xrightarrow{w} (l\,(t)\,, t \ge 0).$$

Recall that the (total) local time of Brownian excursion $l(t)$ can be defined by

$$l(t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^1 I_{[t,t+\varepsilon]}(W^+(s))\, ds,$$

where $W^+(s)$ denotes Brownian excursion of duration 1.

This property was stated as a conjecture by Aldous [2, Conjecture 4]. Another conjecture of Aldous [2, Conjecture 7] was proved by Gittenberger [52].

Theorem 7 *Suppose that τ (the soloution of $\tau\varphi'(\tau) = \varphi(\tau)$) is smaller than the radius of convergence of $\varphi(t)$ and suppose that $a(n)$ is a sequence with $a(n) \to \infty$ and $a(n) = o(\sqrt{n})$ as $n \to \infty$ Then the profile process normalized by $a(n)$ converges weakly to a two the occupation densitiy $q_t = \frac{dQ_t}{dt}$ for 2-sided Bessel process $BES(3)$:*

$$\left(\frac{2}{\sigma\, a(n)} L_n \left(\frac{2}{\sigma t} a(n) \right), t \geq 0 \right) \xrightarrow{w} (q_t, t \geq 0).$$

It should be noted that now there are at least two further proofs of Theorem 6, one by Kersting [68] and one by Pitman [85].

The combinatorial-analytic approach developed in [35] can also be applied to related problems, e.g. in [27] Drmota just considered nodes of a given degree, and in [34] Drmota and Gittenberger derived a functional limit theorem for the *strata of random mappings*.

As already mentioned, one advantage of the combinatorial-analytic approach is that one gets more precise information, e.g. one gets (more or less) explicit density representations or representations for characteristic functions. As a side product Drmota and Gittenberger [35] obtained the following property for the local time of Brownian excursion $l(t)$.

Theorem 8 *Let $\bar\phi_{\kappa_1\ldots\kappa_p}(t_1,\ldots,t_p)$ denote the characteristic function of the joint distribution of $(l(\kappa_1),\ldots,l(\kappa_p))$. Then we have*

$$\bar\phi_{\kappa_1\ldots\kappa_p}(t_1,\ldots,t_p) = 1 + \frac{\sqrt{2}}{i\sqrt{\pi}} \int_\gamma f_{\kappa_1,\ldots,\kappa_p,2}(x,t_1,\ldots,t_p) e^{-x}\, dx,$$

where $f_{\kappa_1,\ldots,\kappa_p,2}(x,t_1,\ldots,t_p)$ is given by

$$\begin{aligned} &f_{\kappa_1,\ldots,\kappa_p,\sigma}(x,t_1,\ldots,t_p) = \\ &\qquad \Psi_{\kappa_1\sigma}(x, it_1 + \Psi_{\kappa_2-\kappa_1,\sigma}(\ldots \Psi_{\kappa_{p-1}-\kappa_{p-2},\sigma}(x, it_{p-1} + \Psi_{\kappa_p-\kappa_{p-1},\sigma}(x,it_p))\cdots) \end{aligned}$$

with

$$\Psi_{\kappa\sigma}(x,t) = \frac{t\sqrt{-x}e^{-\kappa\sigma\sqrt{-x/2}}}{\sqrt{-x}e^{\kappa\sigma\sqrt{-x/2}} - t\frac{\sigma}{\sqrt{2}} \sinh\left(\kappa\sigma\sqrt{-\frac{x}{2}}\right)} \tag{25}$$

and γ is the Hankel-like contour.

Similar representation for the densities have been obtained by Gittenberger and Louchard [55]. (In [56] Gittenberger and Louchard also derived representations for the local time density of the reflected Brownian bridge.)

It should be mentioned that only the one dimensional density of $(l(t); t \geq 0)$ had been well studied. There are several representations available in the literature, one by Kennedy [66] and Kolchin [72, Theorem 2.5.6] by the use of the theory of branching processes, one by Takács [90] by a generating function approach (with Hermite polynomials), one by Knight [71], one by Getoor and Sharpe [50] and Louchard [75] and one by Cohen and Hooghiemstra [23]. (Cohen and Hooghiemstra also derived the Laplace transform of the two dimensional densities of local and occupation time.)

2.2 Systems of Functional Equations

In section 1.3 and section 2.1 we discussed the problem of counting simply generated trees, and it turned out that the generating function $y(x)$ of this numbers satisfies a functional equation of the form

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

which determines $y(x)$ uniquely.

Next we want to describe a more general situation, where systems of functional equations appear and the unknown functions have more that one (complex) variable. 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, $1 \le i \le N$, with "weight" $c_{j,m_1,\ldots,m_N} \ge 0$. 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, $1 \le i \le N$, and with root type j satisfy the functional equations

$$y_j(x, \mathbf{z}) = xz_j \sum_{m_1,\ldots,m_N} c_{j,m_1,\ldots,m_N} y_1(x, \mathbf{z})^{m_1} \cdots y_N(x, \mathbf{z})^{m_N} \qquad (1 \le j \le N). \quad (26)$$

The analytic behaviour of the solution $y = y(x, \mathbf{z})$ of one functional equation of the form $y = F(x, y, \mathbf{z})$ (where $\mathbf{z} = (z_1, z_2, \ldots, z_k)$) is very well understood, see e.g. Drmota [24, 26]. The proof of the following theorem essentially relies on the Weierstrass preparation theorem.

Theorem 9 *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 $F(0, \mathbf{z}, y) \equiv 0$, that $F(x, \mathbf{z}, \mathbf{0}) \not\equiv 0$, and that all Taylor coefficients of F around 0 are real and non-negative. Then the unique solution $y = y(x, \mathbf{z})$ of the functional equation*

$$y = F(x, y, \mathbf{z}) \qquad (27)$$

with $y(0, \mathbf{z}) = 0$ is analytic around 0 and has non-negative 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 non-negative solutions $x = x_0$ and $y = y_0$ of the system of equations

$$\begin{aligned} y &= F(x, y, \mathbf{1}), \\ 1 &= F_y(x, y, \mathbf{1}). \end{aligned}$$

with $F_x(x_0, y_0, \mathbf{1}) \ne 0$ and $F_{yy}(x_0, y_0, \mathbf{1}) \ne 0$ 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

$$y(x, \mathbf{z}) = g(x, \mathbf{z}) - h(x, \mathbf{z}) \sqrt{1 - \frac{x}{f(\mathbf{z})}} \qquad (28)$$

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 $\arg(x - f(\mathbf{z})) \ne 0$ then (28) provides a local analytic continuation of $y(x, \mathbf{z})$.

If one is interested in the asmyptotic behaviour of the coefficients of $y(x, \mathbf{z})$ one can use (28) and singularity analysis (compare with Theorem 3 of section 1.4) to obtain

$$[x^n]y(x,\mathbf{z}) \sim \frac{h(x_0,\mathbf{z})}{2\sqrt{\pi}} f(\mathbf{z})^{-n}.$$

So we are precisely in the situation where we can apply Theorem 2 and obtain a corresponding asymptotic expansion for $[x^n\mathbf{z}^\mathbf{m}$ of the form (33).

It is interesting that the approach for one functional equation also applies for a system of functional equations if the *dependency graph* is strongly connected. The dependency graph $G_\mathbf{F} = (V, E)$ of a system of functional equations $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z})$ has vertex set $V = \{y_1, y_2, \dots, y_N\}$ and an 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.

Before we can state the result of Drmota [26] we have to introduce some further notation. In order to calculate specific parameters (a vector μ and a matrix σ^2) it is necessary to solve the system of functional equations

$$\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z}), \tag{29}$$

$$0 = \det(\mathbf{I} - \mathbf{F_y}(x,\mathbf{y},\mathbf{z})), \tag{30}$$

in which $\mathbf{z} = (z_1, \dots, z_k)'$ 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

$$\mu = -\frac{x_\mathbf{z}(\mathbf{1})}{x(\mathbf{1})} \qquad \text{and} \qquad \sigma^2 = -\frac{x_{\mathbf{zz}}(\mathbf{1})}{x(\mathbf{1})} + \mu'\mu + \mathrm{diag}(\mu), \tag{31}$$

where

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

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 non-negative 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 $\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z})$ with $(n, \mathbf{m}) \in C_j$ are non-zero 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 [79]) where it is quite easy to find sufficient conditions for a single functional equation to be of simple type.

Theorem 10 *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 non-negative, that* $\mathbf{F}(0, \mathbf{z}, \mathbf{y}) \equiv \mathbf{0}$, *that* $\mathbf{F}(x, \mathbf{z}, \mathbf{0}) \not\equiv \mathbf{0}$, *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 non-negative solution* $x = x_0, \mathbf{y} = \mathbf{y}_0$ *of the system of equations*

$$\begin{aligned} \mathbf{y} &= \mathbf{F}(x, \mathbf{y}, \mathbf{1}), \\ 0 &= \det(\mathbf{I} - \mathbf{F_y}(x, \mathbf{y}, \mathbf{1})), \end{aligned}$$

inside it.

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

$$\mathbf{y} = \mathbf{F}(x, \mathbf{y}, \mathbf{z}) \tag{32}$$

in the unknown functions

$$\mathbf{y} = \mathbf{y}(x, \mathbf{z}) = (y_1(x, \mathbf{z}), \ldots, y_N(x, \mathbf{z}))'$$

is strongly connected and if the matrix σ^2 *(defined in (31)) is regular then* $y_j(x, \mathbf{z})$, $1 \le j \le N$, *has a singular representation of the form (28) and consequently 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

$$y_{j,n\mathbf{m}} = \frac{a_j x_0^{-n}}{\sqrt{2^{k+2} \pi^{k+1} n^{k+3}}} \left(\exp\left(-\frac{1}{2n} (\mathbf{m} - \mu\, n)' (\sigma^2)^{-1} (\mathbf{m} - \mu\, n) \right) + \mathcal{O}\left(n^{-1/2} \right) \right) \tag{33}$$

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

$$\begin{aligned} (\mathbf{I} - \mathbf{F_y}(x_0, \mathbf{y}_0, \mathbf{1}))\mathbf{b} &= \mathbf{0}, \\ \mathbf{b}' \mathbf{F_{yy}}(x_0, \mathbf{y}_0, \mathbf{1})\mathbf{b} &= -2\mathbf{F}_x(x_0, \mathbf{y}_0, \mathbf{1}). \end{aligned} \tag{34}$$

It should be noted that Lalley [73] and Woods [96] proved similar theorems for solutions $\mathbf{y}(x)$ in one variable x.

The case of a general system (where the dependency graph is not strongly connected) seems to be much more involved (and will be part of this project.) However, there are partial results, e.g. in [41] asymptotic expansions for the number of trees in the case of two types of nodes were established, where the dependency graph is not strongly connected.

2.3 Admissible Functions

In section 1.4 we already mentioned *admissible function* in the sense of Hayman. The exact definition is the following one, see [64].

Definition 1 *A function*

$$y(x) = \sum_{n \ge 0} y_n x^n \tag{35}$$

is called admissible in the sense of Hayman (H-admissible) if it satisfies the following conditions:

1. $y(x)$ *is analytic in* $|x| < R$ *where* $0 < R \le \infty$ *and for some* $R_0 < R$ *we have*

$$y(r) > 0, \quad R_0 < r < R.$$

2. *There exists a function $\delta(z) : (R_0, R) \to (0, \pi)$ such that for $R_0 < r < R$ we have*

$$y\left(re^{i\theta}\right) \sim y(r) \exp\left(i\theta a(r) - \frac{\theta^2}{2} b(r)\right), \quad \textit{as } r \to R,$$

uniformly for $|\theta| \le \delta(r)$, where

$$a(r) = r\frac{y'(r)}{y(r)}$$

and

$$b(r) = ra'(r) = r\frac{y'(r)}{y(r)} + r^2\frac{y''(r)}{y(r)} - r^2\left(\frac{y'(r)}{y(r)}\right)^2.$$

3. *For $R_0 < r < R$ we have*

$$y\left(re^{i\theta}\right) = o\left(\frac{y(r)}{\sqrt{b(r)}}\right), \quad \textit{as } r \to R,$$

uniformly for $\delta(r) \le |\theta| \le \pi$.

4. *$b(r) \to \infty$ as $r \to R$.*

For H-admissible functions Hayman [64] proved the following basic result:

Theorem 11 *Let $y(x)$ be a function defined in (35) which is H-admissible. Then as $r \to R$ we have*

$$y_n = \frac{y(r)}{r^n\sqrt{2\pi b(r)}}\left(\exp\left(-\frac{(a(r)-n)^2}{2b(r)}\right) + o(1)\right), \quad \textit{as } n \to \infty,$$

uniformly in n.

Furthermore, H-admissible functions have the following *nice* closure properties which are closely related to proper combinatorial constructions (and their generating functions), see section 1.3.

Theorem 12

1. *Suppose that $p(x)$ is a polynomial with real coefficients and that all but finitely many coefficients in the power series expansion of $e^{p(x)}$ are positive, then $e^{p(x)}$ is H-admissible in the whole complex plane.*

2. *If $y(x)$ is H-admissible, then $e^{y(x)}$ is H-admissible, too.*

3. *If $y_1(x), y_2(x)$ are H-admissible, then so is $y_1(x)y_2(x)$.*

4. *If $y(x)$ is H-admissible in $|x| < R$ and $p(x)$ is a polynomial with real coefficients and $p(R) > 0$ if $R < \infty$ and positive leading coefficient if $R = \infty$, then $y(x)p(x)$ is H-admissible in $|x| < R$.*

5. *Let $y(x)$ be H-admissible in $|x| < R$ and $f(x)$ an analytic function in this region. Assume that $f(x)$ is real if x is real and that there exists a $\delta > 0$ such that*

$$\max_{|x|=r} |f(x)| = \mathcal{O}\left(y(r)^{1-\delta}\right), \quad \textit{as } r \to R.$$

Then $y(x) + f(x)$ is H-admissible in $|x| < R$.

6. *If $y(x)$ is H-admissible in $|x| < R$ and $p(x)$ is a polynomial with real coefficients, then $y(x) + p(x)$ is H-admissible in $|x| < R$. If $p(x)$ has a positive leading coefficient, then $p(y(x))$ is also H-admissible.*

Drmota and Soria [40] showed that H-admissible functions (combined with so-called alg-log functions) constitute a proper set of functions which naturally appear as generating functions in combinatorial constructions and that there are very general theorems on the kind of limiting distributions of specific parameters in combinatorial constructions.

It is of course a natural problem how this concept generalizes to the multivariate case. There are in fact, several possibilities, e.g. Bender and Richmond [9] introduced the following concept.

Let $\mathbf{x}$ be d-dimensional and let $\mathbf{r}e^{i\boldsymbol{\theta}}$ be the vector whose kth component is $r_k e^{i\theta_k}$. Suppose $f(\mathbf{x})$ has a power series expansion $\sum a_{\mathbf{n}}\mathbf{x}^{\mathbf{n}}$ where $\mathbf{x}^{\mathbf{n}}$ is the product of $x_k^{n_k}$. The lattice $\Lambda_f \subseteq \mathbf{Z}^d$ is the $\mathbf{Z}$-module spanned by the differences of those $\mathbf{n}$ for which $a_{\mathbf{n}} \neq 0$. We *assume* that Λ_f is d-dimensional. Let $d(\Lambda_f)$ be the absolute value of the determinant of a basis of Λ_f. In other words, $d(\Lambda_f)$ is the reciprocal of the density of Λ_f in $\mathbf{Z}^d$.

We say that $f(x) = o_{u(x)}(g(x))$ for x in some set $\mathcal{S}$ if there is a function $\lambda(t) \to 0$ as $t \to \infty$ such that $|f(x)/g(x)| \le \lambda(|u(x)|)$ for all $x \in \mathcal{S}$. The extension to equations involving little-oh expressions is done in the usual manner.

Definition 2 *Let f be a d-variable function that is analytic at the origin and has a fundamental region $\Phi(f)$. When Λ_f is d-dimensional, we say that $f(\mathbf{x})$ is admissible in $\mathcal{R} \subseteq \mathbf{R}_+^d$ with angles Θ if there are a function Θ from $\mathcal{R}$ to open subsets of $\Phi(f)$ containing $\mathbf{0}$ and functions*

$$\mathbf{a} : \mathbf{C}^d \to \mathbf{C}^d \quad \text{and} \quad B : \mathbf{C}^d \to \mathbf{C}^{d\times d}$$

1. *$f(\mathbf{x})$ is analytic whenever $\mathbf{r} \in \mathcal{R}$ and $|x_i| \le r_i$ for all i;*
2. *$B(\mathbf{r})$ is positive definite for $\mathbf{r} \in \mathcal{R}$;*
3. *the diameter of $\Theta(\mathbf{r})$ is $o_u(1)$, where $u = \det B(\mathbf{r})$;*
4. *for $\mathbf{r} \in \mathcal{R}$, $u = \det B(\mathbf{r})$, and $\boldsymbol{\theta} \in \Theta(\mathbf{r})$, we have*
$$f(\mathbf{r}e^{i\boldsymbol{\theta}}) = f(\mathbf{r})\big(1 + o_u(1)\big)\exp\left(i\mathbf{a}(\mathbf{r})'\boldsymbol{\theta} - \boldsymbol{\theta}' B(\mathbf{r})\boldsymbol{\theta}/2\right);$$
5. *For $\mathbf{r} \in \mathcal{R}$, $u = \det B(\mathbf{r})$, and $\boldsymbol{\theta}$ in the complement of $\Theta(\mathbf{r})$ relative to $\Phi(f)$, we have*
$$f(\mathbf{r}e^{i\boldsymbol{\theta}}) = o_u\big(f(\mathbf{r})\big) \,/\, \sqrt{\det B(\mathbf{r})}. \tag{36}$$

Of course, the definition implies the following asymptotic expansion for the coefficients (by an application of the saddle point method).

Theorem 13 *Suppose $f(\mathbf{x})$ is admissible in $\mathcal{R}$. Let $\mathbf{k}$ be any vector such that $[\mathbf{x}^{\mathbf{k}}]f(\mathbf{x}) \neq 0$, let $u = \det B(\mathbf{r})$. Then*

$$[\mathbf{x}^{\mathbf{n}}]\, f(\mathbf{x}) = \frac{d(\Lambda_f) f(\mathbf{r})\mathbf{r}^{-\mathbf{n}}}{(2\pi)^{d/2}\sqrt{\det B(\mathbf{r})}} \left(\exp\left(-\frac{1}{2}(\mathbf{a}(\mathbf{r}) - \mathbf{n})^{\mathrm{t}} B(\mathbf{r})^{-1}(\mathbf{a}(\mathbf{r}) - \mathbf{n})\right) + o_u(1)\right)$$

for $\mathbf{r} \in \mathcal{R}$ and $\mathbf{n} - \mathbf{k} \in \Lambda_f$.

The advantage of this notion is that it is widely applicable – there is an impressing list of examples in [9]. However, one might lose some of the closure properties of the univariate case. In section 3.3 we will try to develope a variation of this definition which also covers many examples and satisfies that $e^{y(\mathbf{x})}$ is admissible if $y(\mathbf{x})$ is admissible.

2.4 Urn Models

Consider a sequence of m urns into which we throw (usually undistinguishable) balls according to some rules. The balls are thrown one at a time and independently.

Assign to each urn U a valuation $Y(U)$ that is a real valued random variable and is additive, i.e. when we allocate balls in batches the final value of $Y(U)$ is the sum of the values for each batch considered separately. Furthermore, let $\mathcal{E}$ be a subset of the set of possible values of $Y(U)$. The object of interest it random variable

$$X = \sum_{i=1}^{m} 1_{Y(U_i)\in\mathcal{E}},$$

i.e., the number of urns U such that $Y(U) \in \mathcal{E}$. More precisely, let $X_m(n)$, $n \geq 0$, be the value of X at the time when exactly n balls have been thrown into the set of m urns.

If we are interested in the number of urns having a specified number of balls, then $Y(U)$ is the number of balls in the urn U, the set $\mathcal{E}$ is the set of the required numbers for a single urn, and is a subset of the natural integers. For example, empty urns (which have been studied by Kolchin [72]) correspond to $\mathcal{E} = \{0\}$, urns with exactly r balls to $\mathcal{E} = \{r\}$, and urns with at most r balls to $\mathcal{E} = \{0, 1, \dots, r\}$.

The reason why these kind of urn models are of interest is that they provide a dynamic model for the size of relational databases, see [46, 48].

We now present a generating function approach to this problem. Let $g(z)$ be the generating function enumerating the allocation of balls into a single urn, and $f(z)$ the function enumerating those allocations such that $Y \in \mathcal{E}$. Furthermore, let us assume that $g(z)$ and $f(z)$ are entire functions. We mark the balls by z and the urns satisfying $Y(U) \in \mathcal{E}$ by x. Then the generating function describing the allocations of balls in the m urns is

$$\Phi_{\mathcal{E},1}(x, z) = (g(z) + (x-1)f(z))^m.$$

The idea is to extract the coefficients $[z^n]\Phi_{\mathcal{E},1}(1, z)$, $[z^n x^k]\Phi_{\mathcal{E},1}(x, z)$ to obtain the probabilities

$$\mathbf{P}\{X_m(n) = k\} = \frac{[z^n x^k]\Phi_{\mathcal{E},1}(x, z)}{[z^n]\Phi_{\mathcal{E},1}(1, z)}.$$

Similarly it is possible to construct the bi-dimensional generating function

$$\Phi_{\mathcal{E},2}(x_1, x_2, z_1, z_2) = (\phi_{\mathcal{E},2}(x_1, x_2, z_1, z_2))^m$$

and get

$$\mathbf{P}\{X_m(n_1) = k_1, X_m(n_1+n_2) = k_2\} = \frac{[x_1^{k_1} x_2^{k_2} z_1^{n_1} z_2^{n_2}]\Phi_{\mathcal{E},2}(x_1, x_2, z_1, z_2)}{[z_1^{n_1} z_2^{n_2}]\Phi_{\mathcal{E},2}(1, 1, z_1, z_2)}.$$

In general, we obtain a recurrent relation for the generating functions $\phi_{\mathcal{E},d}(x_1, \dots, x_d; z_1, \dots, z_d)$ of the form

$$\begin{aligned}\phi_{\mathcal{E},d}(x_1, \dots, x_d; z_1, \dots, z_d) &= K_{d+1}(z_1, \dots, z_d) \\ &+ \sum_{l=1}^{d} x_l K_l(z_1, \dots, z_l) \sum_{r\in\mathcal{E}} \phi_{\mathcal{E}-r,d-l}(x_{l+1}, \dots, x_d; z_{l+1}, \dots, z_d). \qquad (37)\end{aligned}$$

The nature of $\Phi_{\mathcal{E},d}(x_1, \dots, x_d; z_1, \dots, z_d)$ allows a straightforward application of the results of Bender and Richmond [8], which gives directly the convergence towards a Gaussian distribution, e.g. the asymptotic mean of $X_m(n)$ is given by

$$\mathbf{E}X_m(n) = m\,\frac{[z^n]f(z)g(z)^{m-1}}{[z^n]g(z)^m} \sim m\frac{f(\rho)}{g(\rho)},$$

where ρ is the solution of the saddle point equation $zg'(z)/g(z) = n/m$. The normalizing factor is

$$[z^n]g(z)^m = \frac{g(\rho)^m}{\rho^n\sqrt{2\pi\sigma^2 m}}(1+o(1)),$$

with $\sigma^2 = \rho^2 g''(\rho)/g(\rho) - (\rho g'(\rho)/g(\rho))^2 + \rho g'(\rho)/g(\rho)$. (In the frequent case where $g(z) = e^z$ we get $\rho = n/m$ and $s^2 = \rho$).

Drmota, Gardy, and Gittenberger [38] proved a general version of Kolchin's orginial result [72] on emply urns.

Theorem 14 *Let $X_m(\lfloor mt \rfloor)$, $t \ge 0$, be the process associated to allocating balls into urns of a general urn model such that the generating functions describing the allocation process have the shape*

$$\Phi_{\mathcal{E},d}(x_1,\ldots,x_d,z_1,\ldots,z_d) = \phi_{\mathcal{E},d}(x_1,\ldots,x_d,z_1,\ldots,z_d)^m$$

where $\phi_{\mathcal{E},d}$ satisfies a recurrence relation of the form (37) with entire functions K_i. Then the following weak limit theorem holds:

$$Y_m(t) := \frac{X_m(\lfloor mt\rfloor) - \mathbf{E}X_m(\lfloor mt\rfloor)}{\sqrt{m}} \xrightarrow{w} G(t)$$

where $G(t)$ is a centered Gaussian process with continuous sample paths. The covariance function $B_{s,t}$, $s,t \ge 0$, is given by

$$B_{s,t} = B_{t,s} = \left.\frac{\partial^2(\log \lambda_{s,t}(e^{u_1},e^{u_2}))}{\partial u_1 \partial u_2}\right|_{u_1=0,u_2=0} \tag{38}$$

where $s < t$ and

$$\lambda_{s,t}(x_1,x_2) = \frac{\phi_{\mathcal{E},2}(x_1,x_2,\rho_1,\rho_2)}{\rho_1^s \rho_2^{t-s}}$$

where $\rho_1 = \rho_1(x_1,x_2,s,t)$ and $\rho_2 = \rho_2(x_1,x_2,s,t)$ are the saddle points which are defined by the equations in z_1 and z_2

$$z_1\, \partial\phi_{\mathcal{E},2}/\partial z_1 = s\phi_{\mathcal{E},2}, \tag{39}$$
$$z_2\, \partial\phi_{\mathcal{E},2}/\partial z_2 = (t-s)\phi_{\mathcal{E},2}; \tag{40}$$

and

$$B_{s,s} = \left.\frac{\partial^2(\log \lambda_s(e^u))}{\partial^2 u}\right|_{u=0} \tag{41}$$

with

$$\lambda_s(x) = \frac{\phi_{\mathcal{E},1}(x,\rho)}{\rho^s}$$

where $\rho = \rho(x,s)$ is the saddle point defined by the equation in z

$$z\, \partial\phi_{\mathcal{E},1}/\partial z = s\phi_{\mathcal{E},1}. \tag{42}$$

This result means that if you throw balls subsequently into m urns then X follows more or less the expected value. The deviation from the mean is governed by a Gaussian process $G(t)$ which is explicit in terms of the generating functions. (For expample, we can decide whether $G(t)$ is Markov process. $G(t)$ is a Markov process if and only if the covariance can be factorized in the form $B_{s,t} = b_1(s)b_2(t)$, see [82]).

We also want to mention that this result is a considerable extension of [46, 47]. There are of course other aspects considered in the literature, see e.g. [11, 48].

2.5 Binary Search Trees

A binary search tree is a binary tree, in which each (internal) node is associated to a key, where the keys are drawn form some totally ordered set, say $\{1, 2, \ldots, n\}$. The first key is associated to the root. Now, the next key is put to the left child of the root if it is smaller than the key of the root and it is put to the right child of the root if it is larger than the key of the root. In this way we proceed further by inserting key by key. So starting from a permutation of $\{1, 2, \ldots, n\}$ we get a binary tree with n (internal) nodes such that the keys of the left subtree of any given node x are smaller than the key of x and the keys of the right subtee are larger than the key of x.

Binary search trees are widely used to store (totally orderer) data, and many parameters have been discussed in the literature. (The monograph of Mahmoud [76] gives a very good overview of the state of the art.) Usually it is assumed that every permutation of $\{1, 2, \ldots, n\}$ is equally likely and any parameter of binary search trees may be considered as a random variable. (An alternative way of looking at is a Markov chain of trees which describes the evolution of a binary search tree.)

We want to concentrate on three parameters, on the height H_n, the saturation level H_n', and on the profile $U_k(n)$, $V_k(n)$, $k \geq 0$. Here $U_k(n)$ denotes the number of external nodes at level (depth) k and $V_k(n)$ the number of internal nodes at level k.

For example, the probability distribution of the height H_n can be described by

$$\mathbf{P}[H_n \leq h] = \frac{a_{n,h}}{n!},$$

where $a_{n,h}$ denotes the number of permutations $\sigma \in \mathcal{S}_n$ of n elements such that the corresponding binary search tree has height $\leq h$.

In 1986 Devroye [19] proved that the expected value $\mathbf{E}H_n$ satisfies the asymptotic relation $\mathbf{E}H_n \sim c \log n$ (as $n \to \infty$), where $c = 4.31107\ldots > 2$ is the solution of the equation $\left(\frac{2e}{c}\right)^c = e$. (Earlier Pittel [85] had shown that $H_n/\log n \to \gamma$ almost surely as $n \to \infty$, where $\gamma \leq c$, compare also with Robson [87]. Later Devroye [20] proved quantified versions, i.e. $H_n - c\log n = \mathcal{O}(\sqrt{\log n}\,\log\log n)$ in probabilty.) Based on numerical data Robson conjectured that the variance $\mathbf{V}H_n$ is bounded. In fact, he could prove ([88]) that there is an infinite subsequence for which

$$\mathbf{E}|H_n - \mathbf{E}H_n| = \mathcal{O}(1).$$

and that his conjecture is equivalent to the assertion that the expected value of number of noded at level $h = H_n$ is bounded ([89]). The best bounds (before 1999) were given by two completely different methods by Devroye and Reed [22] and by Drmota [28]. They (both) proved

$$\mathbf{E}H_n = c\log n + \mathcal{O}(\log\log n) \tag{43}$$

and

$$\mathbf{V}H_n = \mathbf{E}(H_n - \mathbf{E}H_n)^2 = \mathcal{O}((\log\log n)^2).$$

Devroye and Reed used a more probabilistic method. They *approxmiated* binary search trees by a certain branching process. Drmota's approach is purely combinatorially. He uses generating functions (which will be presented below) and analytic methods.

Eventually, Reed [86] settled Robson's conjecture

$$\mathbf{V}H_n = \mathcal{O}(1).$$

He used an approach related to that of [22], where he could also show that

$$\mathbf{E}H_n = c\log n - \frac{3c}{2(c-1)}\log\log n + \mathcal{O}(1). \tag{44}$$

Few months later Drmota [31] gave a second proof of Robson's conjecture, again purely combinatorially. His result reads as follows.

Theorem 15 *Let $y_h(x)$ $(h \geq 0)$ be polynomials recursively defined by $y_0(x) \equiv 1$ and by*

$$y_{h+1}(x) := 1 + \int_0^x y_h(t)^2\, dt \qquad (h \geq 0)$$

and let e_h $(h \geq 0)$ be defined by

$$e_h := c \cdot \log y_h(1),$$

where $c = 4.31107\ldots > 2$ is the solution of the equation $\left(\frac{2e}{c}\right)^c = e$. Then the expected value of the height H_n of binary search trees of size n is given by

$$\mathbf{E}H_n = \max\{h : e_h \leq c \log n\} + \mathcal{O}(1) \qquad (n \to \infty) \tag{45}$$

and there exists a constant $C > 0$ such that for all integers $k > 0$ and $n \geq 0$

$$\mathbf{E}|H_n - \mathbf{E}H_n|^k \leq C\,(k+1)!\left(1 - \frac{2}{c}\right)^{-k}. \tag{46}$$

Note that the polynomials $y_h(x)$ are exactly the exponential generating functions of the numbers $a_{n,h}$:

$$y_h(x) = \sum_{n \geq 0} \frac{a_{n,h}}{n!}\, x^n.$$

Please note further that Drmota's approach does not provide an explicit asymptotic expansion for $\mathbf{E}H_n$. It is implicit and depends on the (unknown) behaviour of $y_h(1)$.

The fact that $\mathbf{E}H_n = \mathcal{O}(\log n)$ indicates that binary search trees are very well balanced. This observation is reflected by the fact that the saturation level H_n', i.e. the level of the first external node, is relatively large. Biggins [10] proved that

$$\frac{H_n'}{c' \log n} \to 1 \quad a.s.,$$

where $c' = 0.373\ldots < 2$ is the (other) solution of the equation $\left(\frac{2e}{c'}\right)^{c'} = e$.

This result indicates some *duality* between the height H_n and the saturation level H_n' of binary search trees. This seems to be true also from a combinatorial point of view.

Let $b_{n,h}$ denotes the number of permutations $\sigma \in S_n$ of n elements such that the corresponding binary search tree has saturation level $> h$, i.e. $\mathbf{P}[H_n' > n] = b_{n,h}/n!$ Then the exponential generating functions

$$y_h(x) = \sum_{n \geq 0} \frac{b_{n,h}}{n!} x^n = \sum_{n \geq 0} \mathbf{P}[H_n' > n] x^n \tag{47}$$

satisfy the recurrent relation

$$y_{h+1}(x) = \int_0^x y_h(t)^2\, dt.$$

with $y_0(x) = x/(1-x)$. These functions are no longer polynomials. Even, the *converse* functions

$$z_h(x) = \frac{1}{1-x} - z_h(x) = \sum_{n \geq 0} \mathbf{P}[H_n' \leq n] x^n$$

are no polynomials, either. With help of these generating functions Drmota [31] could prove a corresponding theorem for H_n'.

Theorem 16 *The expected value* $\mathbf{E}H'_n$ *of the saturation level of binary search trees is given by*

$$\mathbf{E}H_n = \max\{h : z_h(1-n^{-1}) \le n\} + \mathcal{O}(1) \qquad (n \to \infty) \tag{48}$$

and all centralized moments of H'_n *are bounded:*

$$\mathbf{E}|H_n - \mathbf{E}H_n|^k = \mathcal{O}(1) \qquad (n \to \infty). \tag{49}$$

Finally, let us consider the profile of binary search trees, i.e., $U_k(n)$, denote the number of external nodes of T_n at level k, and $V_k(n)$ the number of internal nodes at level k. For the sake of shortness just formulate the results in terms of $U_k(n)$. The same results hold for $V_k(n)$, too. The following theorem has been proved by Chauvin, Drmota, and Jabbour-Hattab [15].

Theorem 17 *We have a.s.*

$$\frac{U_k(n)}{n/\sqrt{4\pi\log n}} = e^{\frac{(k-2\log n)^2}{4\log n}} + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right)$$

as $n \to \infty$, *where the error term* $\mathcal{O}(1/\sqrt{\log n})$ *is uniform for all* $k \ge 0$.

From this theorem we directly obtain a result for the width.

Corollary 1 *Let* $\overline{U}(n) = \max_{k\ge 0} U_k(n)$. *Then we have*

$$\frac{\overline{U}(n)}{n/\sqrt{4\pi\log n}} = 1 + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right) \quad a.s.$$

as $n \to \infty$.

Interestingly, the proof of this theorem relies on a saddle point method combined with properties of a martingale of analytic functions.

2.6 Sum-of-Digits Function

Let $G = (G_j)_{j\ge 0}$ be a strictly increasing sequence of integers with $G_0 = 1$. Then every non-negative integer n has a (unique) proper G-ary digital expansion

$$n = \sum_{j\ge 0} \epsilon_j(n) G_j$$

with integer digits $\epsilon_j(n) \ge 0$ provided that

$$\sum_{j<k} \epsilon_j(n) G_j < G_k$$

for all $k \ge 0$.

If $G_j = q^j$ (where q is a fixed integer > 1) we obtain the usual q-ary digital expansion.

Generalizing q-ary expansions we also consider the following situation:
There exist non-negative integers $a_1 \ge a_2 \ge \cdots \ge a_r > 0$ such that

$$G_j = \sum_{i=1}^{j} a_i G_{j-i} + 1$$

for $j < r$ and

$$G_j = \sum_{i=1}^{r} a_i G_{j-i}$$

for $j \ge r$.

It is well known that the characterictic polynomial

$$P(u) = u^r - \sum_{i=1}^{r} a_i u^{r-i}$$

is irreducible and that there is a unique root α of maximal modulus which is real and positive and all other roots have modulus < 1, i.e. α is a Pisot number (compare with [60]). Obviously we have

$$G_j \sim C\alpha^j \quad (j \to \infty)$$

for some constant $C > 0$. Furthermore, a sequence of non-negative integers ϵ_j, $j \ge 0$, constitutes the digital expansions of

$$n = \sum_{j\ge 0} \epsilon_j G_j,$$

i.e. $\epsilon_j(n) = \epsilon_j$, $j \ge 0$, if and only if $\epsilon_j = 0$ for all but finitely many j and if

$$(\epsilon_k, \epsilon_{k-1}, \ldots, \epsilon_{k-r+1}) < (a_1, a_2, \ldots, a_r) \qquad \text{for all } k \ge 0,$$

where "$<$" denotes the lexicographic order. ($\epsilon_j = 0$ for $j < 0$.)

let $F : \{0, 1, \ldots, a_1 + 1\}^{L+1} \to \mathbf{R}$ be any given function (for some $L \ge 0$) with $F(0, 0, \ldots, 0) = 0$. Furthermore, set

$$s_F(n) = \sum_{j\ge 0} F\left(\epsilon_{j+L}(n), \epsilon_{j+L-1}(n), \ldots, \epsilon_j(n)\right).$$

This means that we consider a weighted sum over all subsequent digital patterns of length $L + 1$ of the digital expansion of n. For example, for $L = 0$ and $F(\epsilon) = \epsilon$ we just obtain the sum-of-digits function, or if $L = 1$ and $F(\epsilon, \eta) = 1 - \delta_{\epsilon,\eta}$ ($\delta_{x,y}$ denoting the Kronecker delta) then $s_F(n)$ is just counting the number of times that a digit is different from the preceding one etc.

Drmota [29] proved the following property on the distribution of $s_F(n)$. (The analytic methods which rely on generating functions and saddle point approximations are adapted from [**?**] and [39].)

Theorem 18 *Suppose that $G = (G_j)_{j\ge 0}$ satisfies a linear recurrence of the above type and let a function F be as above. Then there exist constants μ, $\sigma \ge 0$ such that*

$$E_N := \frac{1}{N} \sum_{n<N} s_G(n) = \mu \log N + \mathcal{O}(1) \tag{50}$$

and

$$\frac{1}{N} \sum_{n<N} (s_G(n) - E_N)^2 = \sigma^2 \log N + \mathcal{O}(1) \tag{51}$$

as $N \to \infty$. Furthermore, if $\sigma \ne 0$ then for every $\varepsilon > 0$

$$\frac{1}{N} \left| \{ n < N : s_F(n) < \mu \log N + x \cdot \sigma \cdot (\log N)^{\frac{1}{2}} \right|$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt + \mathcal{O}((\log N)^{-1/2+\varepsilon})$$

uniformly for all real x as $N \to \infty$.

If F attains only integer values, it is also possible to derive an asymptotic expansion for the numbers $|\{n < N : s_F(n) = k\}|$, see [29].

Previously, asymptotic and exact formulas for E_N for the sum-of-digits function $s(n)$ are due to Bush [13], Bellman and Shapiro [7], Delange [18], and Trollope [94] for q-ary digital expansions and due to Pethö and Tichy [83] and Grabner and Tichy [60, 61] for G-ary digital expansions with respect to linear recurrences. The asymptotic expansion of E_N has been discussed for several other special cases of $s_F(n)$, too, see [14], [93], [5]. especially the $\mathcal{O}(1)$-term has been analyzed in detail.

Corresponding formulas for higher moments and for the variance can be found in Coquet [17], Kirschenhofer [70], Kennedy and Cooper [67], Grabner, Kirschenhofer, Prodinger, and Tichy [59], and in Dumont and Thomas [42].

The asymptotic distribution of the sum-of-digits function and related problems are discussed in Schmidt [92], Schmid [91], Bassily and Katai [6], and in Dumont and Thomas [43].

The paper of Bassily and Katai [6] is really of significant interest. There it is shown that for any integer polynomial $P(x)$ and any $q > 1$ there exist constants μ, σ such that

$$\begin{aligned}\frac{1}{N}&\left|\{n < N : s_F(P(n)) < \mu \log N + x \cdot \sigma \cdot (\log N)^{\frac{1}{2}}\right| \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt + o(1),\end{aligned}$$

and a corresponding result holds if we just consider primes $p < N$.

Quite recently, Gittenberger and Thuswaldner [57] could prove a similar property for digital expansions of Gaussian integers $\mathbf{Z}[i]$.

A completely different problem has been discussed by Kim [69], namely the distribution of the sum-of-digits function modulo m (with respect to different bases $q_1, \ldots, q_k > 1$.)

Theorem 19 *Let $q_1, \ldots, q_k > 1$ be pairwisely relative prime integers and $m_1, \ldots, m_k$ positive integers such that*

$$\gcd(q_j - 1, m_j) = 1 \qquad (1 \le j \le k).$$

Then we have for all integers $c_1, \ldots, c_k$

$$|\{n < N : s_{q_j}(n) \equiv c_j \bmod m_j \ (1 \le j \le k)\}| = \frac{N}{m_1 m_2 \cdots m_k} + \mathcal{O}(N^{1-\delta})$$

with

$$\delta = \frac{1}{120 k^2 q^2 m^2},$$

where $q = \max\{q_1, \ldots, q_k\}$ and $m = \max\{m_1, \ldots, m_k\}$ and the $\mathcal{O}$-constant depends only on k and q.

Corresponding one-dimensional results are due to Gelfond [49] in the q-ary case and due to Hoit [65] in more general number systems.

3 Aims and Methodology

3.1 Galton-Watson Trees

The analysis of simply generated families of trees (resp. Galton-Watson trees) was a major part of the The FWF-project *Multivariate asymptotische Entwicklungen.* Most results have been obtained within this topic ([27, 30, 35, 51, 52, 53, 55]).

We could solve the major problems. So for this proposal we concentrate just on few problems concerning simply generated families of trees and try to include more problems from other fields.

The proof of Theorem 6 requires finite dimensional (weak) convergence theorems of the form that for any set of non-negative real numbers $\kappa_1 < \kappa_2 < \cdots < \kappa_p$ the random vector

$$\frac{1}{\sqrt{n}}(L_n(\kappa_1\sqrt{n}), L_n(\kappa_2\sqrt{n}), \ldots, L_n(\kappa_p\sqrt{n}))$$

converges weakly to the corresponding d-dimensional distribution of the local time $l(t)$.

We will shortly describe the combinatorial background of the method used in [35]. Let $y(x)$ be the solution of $y = x\varphi(y)$ and for every $k \geq 0$

$$y_k(x,u) = \sum_{n,m} \left(\sum_{|T|=n, L_n(k)=m} \omega(T) \right) x^n u^m.$$

Then it follows by induction that

$$\begin{aligned} y_0(x,u) &= uy(x) \\ y_{i+1}(x,u) &= x\varphi(y_i(x,u)), \quad i \geq 0. \end{aligned} \tag{52}$$

The characteristic function of $\frac{1}{\sqrt{n}}L_n(k)$ is

$$\phi_{kn}(t) = \frac{1}{y_n}[x^n]y_k\left(x, e^{it/\sqrt{n}}\right).$$

In a similar way we get the characteristic function for the higher dimensional distributions: The characteristic function of $\left(\frac{1}{\sqrt{n}}L_n(k_1), \ldots, \frac{1}{\sqrt{n}}L_n(k_p)\right)$ is given by

$$\phi_{k_1\cdots k_p n}(t_1, \ldots, t_p) = \frac{1}{y_n}[x^n]y_{k_1}\left(x, e^{it_1/\sqrt{n}} y_{k_2-k_1}\left(x, \ldots y_{k_p-k_{p-1}}\left(x, e^{it_p/\sqrt{n}}\right)\ldots\right)\right).$$

In order to extract the desired coefficient asymptotically we will use Cauchy's integral formula with a suitably chosen integration contour and approximate the integrand there. Therefore we need a detailed knowledge of the behaviour of the recursion (52):

Lemma 1 *Let x_0 be the point on the circle of convergence of $a(x)$ which lies on the positive real axis. Set $x = x_0\left(1 + \frac{w}{n}\right)$ and $\alpha = z\varphi'(a(z))$. Furthermore assume that $|u-1| = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$ and $\frac{w}{n} \to 0$ in such a way that $|\arg(-w)| < \pi$ and*

$$\left|1 - \sqrt{\frac{-w}{n}}\right| \leq 1 + \frac{C}{\sqrt{n}}$$

are satisfied. Then $y_k(x,u)$ admits the local representation

$$y_k(x,u) = y(x) + \frac{y(x)(u-1)\alpha^k}{\frac{\sqrt{-w/n} + \sigma(\tau - uy(x))/\tau\sqrt{2}}{2\sqrt{-w/n}} + \frac{\sqrt{-w/n} - \sigma(\tau - uy(x))/\tau\sqrt{2}}{2\sqrt{-w/n}}\alpha^k + \mathcal{O}\left(\sqrt{\frac{|x|}{n}}\right)},$$

uniformly for $k = \mathcal{O}(\sqrt{n})$, where $\alpha = x\varphi'(y(x))$

With help of this asymptotic representation it is not too difficult to prove that the characteristic function of $\frac{1}{\sqrt{n}}L_n(\kappa\sqrt{n})$ converges to the corresponding one-dimensional characteristic function of the local time. (You have to set $u = e^{it/\sqrt{n}}$ and a proper contour around the singularity of $y(x)$ in order to extract the coefficient $[x^n]y_k(x,u)$ asymptotically via Cauchy's formula.) So weak convergence can be established. However, we were not able to estabish a local version of this weak convergence result (with the only exception of planted plane trees, where the recurrence for $y_k(x,u)$ can be explicitly solved). Therefore the following problem remains to be solved.

Problem 1 *For a simply generated family of trees derive bivariate asymptotic expansions for the coefficients*

$$[x^n u^k]y_k(x,u) = \sum_{|T|=n, L_n(k)=m} \omega(T),$$

where $k = \mathcal{O}(\sqrt{n})$.

Suggested Method. The technical problem is to extend Lemma 1 to u which are not so close to 1 so that one can also use Cauchy's formula to extract the coefficient of u^k. (For planted plane trees ($\varphi(t) = 1/(1-t)$) there is an explicit expression for $y_k(x,u)$, for which Cauchy's formula applies.) □

There is an interesting corollary of Theorem 6, namely that the (normalized) maximum

$$\frac{1}{\sqrt{n}} \max_{k\ge 0} L_n(k)$$

has a limiting distribution which is (despite of a constant) the same as $\sup_{t\ge 0} l(t)$. So it is natural to pose the following question.

Problem 2 *Prove that*

$$\mathbf{E}\left(\max_{k\ge 0} L_n(k)\right) \sim c\sqrt{n}$$

for some constant $c > 0$.

For rooted labeled trees ($\varphi(t) = e^t$) this problem has been solved by Chassaing and Marckert [16].

Suggested Method. The advantage of analytic methods is that one usually gets the order of magnitude for error terms. So one gets

$$\mathbf{E}\left(\max_{1\le j\le p} L_n(\kappa_j\sqrt{n})\right) \sim c(\kappa_1,\ldots,\kappa_p)\cdot\sqrt{n}$$

(again with an error term). So it seems that the solution is not out of reach, maybe by using also ideas of [16] and local versions of the weak limit theorem. □

Last year Aldous [3] proved a *joint version* of Theorems 4 and 6:

$$\left(\frac{\sigma}{2\sqrt{n}}X_n([2ns]), \frac{2}{\sigma\sqrt{n}}L_n\left(\frac{2}{\sigma t}\sqrt{n}\right); 0\le s\le 1, t\ge 0\right) \xrightarrow{w} (W^+(s), l\,(t)\,; 0\le s\le 1, t\ge 0).$$

It seems that this result can also be obtained by combinatorial and analytic methods. Especially it should be possible to solve the following problem.

Problem 3 *Derive integral representations for finite dimensional characteristic functions and densities of the stochastic field* $(W^+(s), l\,(t)\,; 0\le s\le 1, t\ge 0)$.

Suggested Method. Combine the analytic proofs of Theorems 4 and 6. These (previously given) proofs rely on explicit representations of characteristic functions resp. densities. □

3.2 Systems of Functional Equations

These kinds of questions should be worked out in a chapter of a small book [37] which is already in progress.

In section 2.2 we described that the analytic behaviour of the solution $\mathbf{y} = \mathbf{y}(x,\mathbf{z})$ of the system of functional equations $\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z})$ is well understood if the dependency graph is strongly connected, there are only squareroot singularities. We also indicated that this easy description does not remain true (in general) if the dependency graph is not strongly connected, see [41].

The difference between a system with a strongly connected dependency graph and a system without a strongly connected dependency graph is that in the latter case there are subsystems which can be solved independently of the other parts. Analytically this has the effect that singularities (coming from these subsystems) may *interfere* when they *come together* in another subsystem. For example, the following seems to be true.

Problem 4 *Suppose that $\mathbf{y} = \mathbf{F}(x,\mathbf{y})$ is a system of functional equations, where the solution $\mathbf{y} = \mathbf{y}(x)$ just depend on one variable x. If the coefficients of $\mathbf{F}(x,\mathbf{y})$ (considered as a power series around $(0,\mathbf{0})$) are non-negative and if no subsystem of $\mathbf{y} = \mathbf{F}(x,\mathbf{y})$ is a linear system then every component $y_j(x)$, $1 \le j \le d$, of the solution $\mathbf{y} = \mathbf{y}(x)$ has a Pussieu expansion around its radius of convergence $x_0 = r_j$ with exponent 2^{-m} for some $m \ge 1$.*

Suggested Method. It follows from Theorem 9 that the solution of a single equation $y = F(x,y)$ (if F is not linear) has a square root singularity at the radius of convergence. So, Theorem 10 can be interpreted in a way that a system of equations with a strongly connected dependency graph behaves *like a single equation*, i.e., all components have the same kind of squareroot singularity. Now, if we start by solving subsystems which can be solved independently from the others and for which the corresponding (smaller) dependency graphs are strictly connected then we have (possibly) different squareroot singularities. Then we insert these solutions into the other equations and so we can solve other subsystems. So it might occur that the *old* squareroot singularies and the *new* one *interfere* which give a fourth root singularity. So we can proceed by induction. □

The sitation is surely much more involved if we consider general systems of the form $\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z})$. Theorem 10 says that in the case of a strongly connected dependency graph there always is a *Gaussian law* behind. However, if the dependency graph is not strongly connected it seems to be a really hard problem to give a complete description of all possible cases which can occur.

Problem 5 *Give a description of all possible limiting distributions of the coefficients of generating functions satisfying a system of the form $\mathbf{y} = \mathbf{F}(x,\mathbf{y},\mathbf{z})$.*

As already mentioned, one chapter of [37] will be devoted to systems of functional equations. Of course we will also include certain (extended) examples. In [36] Drmota and Gittenberger proved that in random trees the number of nodes with a given degree have a Gaussian limiting distribution. A node x of degree d (together with ist d neighbours) constitutes a subtree which has the form of a star. So we can also say that the number of occurances of a star with d edges has a Gaussian limiting distribution, or the number of occurancen of this specific pattern has a Gaussian limiting distribution. Therefore it is natural to state the following problem.

Problem 6 *Let T_0 be a fixed tree (pattern). Then the numbers of occurances of T_0 as a subtree (pattern) in random trees have a Gaussian limiting distribution.*

Suggested Method. First of all we have to deal with rooted trees (compare with [36]) and subdivide the set of rooted trees into sets where the root is a specific node of T_0. Of course, the structure of T_0 gives rise to a system of recursive combinatorial constructions. So a translation to generating functions yields a system of functional equations which can be solved by the methods developed before. Of course, the final step from rooted trees to unrooted trees efforts a construction similarly to Otter, see [36]. $\square$

3.3 Multivariate Admissible Functions

Multivariate admissible functions should become a chapter of [37]. We want to give an outline which kind of problems we want to cope with. In section 2.3 we already introduced Bender's and Richmond's concept of multivariate admissible functions. Although their concept is a very powerful tool, it seems that it not sufficient for *all* applications, especially if one needs closure properties (other than the product). Especially we would like to cover the property that $e^{y(\mathbf{x})}$ is admissible if $y(\mathbf{x})$ is admissible. Before doing this we start with an *easy* problem.

Admissible functions $y(x)$ have the nice property that their coefficients $[x^n]y(x)$ can be asymptotically evaluated by a saddle point method. The same is true for a fixed power $y(x)^k$ (since products of admissible functions are admissible, too.) Furthermore, if k is of the same order as n we know by Theorem 2 that $[x^n]y(x)^k$ can be asymptotically determined by (formally) the same procecure as for $[x^n]y(x)$. However, this need not be true, if k is, say, of order $\sqrt{n}$. So we can state our first problem concerning admissible functions.

Problem 7 *Find a proper definition for power-admissible function $y(x)$ which have the property that for all $k = \mathcal{O}(n)$ we uniformly have*

$$[x^n]y(x)^k \sim \frac{y(r)}{r^n\sqrt{2\pi b(r)}},$$

where r is given by $a(r) = n/k$. Prove closure properties of power-admissible function. Is it true that the composition $y(x) = y_1(y_2(x))$ of two entire power-admissible functions is power-admissible, too?

Suggested Method. The problem is not to define power-admissible functions in some way. The real problem is to find a proper definition such that one can prove closure properties and maybe decide the last question. This problem is surely not out of reach. $\square$

We now try to define another notion of multivariate admissibility which we call H-admissibility. (For $d = 1$ this definition reduces to the usual H-admissibility.)

Definition 3 *We will call a function*

$$y(\mathbf{x}) = \sum_{n_1,\ldots,n_d\geq 0} a_{n_1\cdots n_d} x_1^{n_1}\cdots x_d^{n_d} \tag{53}$$

H-admissible if $y(\mathbf{x})$ is analytic in $\{(x_1,\ldots,x_d) : |x_1| < R_1,\ldots,|x_d| < R_d\}$ and has the following properties: ($R := (R_1, R_2,\ldots,R_d)$)

1. *There exists a function $\delta : \mathbf{R}^d \to [-\pi,\pi]$ such that*

$$y\left(\mathbf{r}e^{i\boldsymbol{\theta}}\right) \sim y(\mathbf{r})\exp\left(i\boldsymbol{\theta}\mathbf{a}(\mathbf{r})^t - \frac{\boldsymbol{\theta}B(\mathbf{r})\boldsymbol{\theta}^t}{2}\right), \quad \textit{as } \mathbf{r}\to R, \tag{54}$$

uniformly for $\|\boldsymbol{\theta}\|_{\max} \leq \delta(\mathbf{r})$*, where* $\mathbf{a}(\mathbf{x}) = (a_j(\mathbf{x}))_{j=1,\dots,d}$ *with (subscripts* x_j*, etc. denote the partial derivatives w.r.t.* x_j*, etc.)*

$$a_j(\mathbf{x}) = \frac{x_j y_{x_j}(\mathbf{x})}{y(\mathbf{x})},$$

and $B(\mathbf{x}) = (B_{jk}(\mathbf{x}))_{j,k=1,\dots,d}$ *with*

$$B_{jk}(\mathbf{x}) = \frac{x_j x_k y_{x_j x_k}(\mathbf{x}) + \delta_{jk} x_j y_{x_j}(\mathbf{x})}{y(\mathbf{x})} - \frac{x_j x_k y_{x_j}(\mathbf{x}) y_{x_k}(\mathbf{x})}{y(\mathbf{x})^2},$$

where δ_{jk} *denotes Kronecker's* δ *defined by*

$$\delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}$$

2. *We have*

$$y\left(\mathbf{r}e^{i\boldsymbol{\theta}}\right) = o\left(\frac{y(\mathbf{r})}{\sqrt{\det B(\mathbf{r})}}\right), \quad \text{as } \mathbf{r} \to R, \tag{55}$$

uniformly for $\|\boldsymbol{\theta}\|_{\max} \leq \delta(\mathbf{r})$.

3. $B(\mathbf{r})$ *is positive definite and satisfies*

$$\det B(\mathbf{r}) \to \infty, \quad \text{as } \mathbf{r} \to R,$$

and for all $j = 1, \dots, d$ *we have*

$$B_{jj}(\mathbf{r}) \to \infty, \quad \text{as } r_j \to R_j,$$

uniformly in r_k, $k \neq j$.

4. $y(\mathbf{x})$ *is univariate admissible (which we will call 1-admissible in the sequel) in every variable, i.e. for all* $j = 1, \dots, d$ *the relations*

   (a)

$$y(r_1, \dots, r_{j-1}, r_j e^{i\theta_j}, r_{j+1}, \dots, r_d) \sim y(\mathbf{r}) \exp\left(i\theta_j a_j(\mathbf{r}) - \frac{\theta_j^2 B_{jj}(\mathbf{r})}{2}\right),$$

   as $r_j \to R_j$*, uniformly for* $|\theta_j| \leq \delta(\mathbf{r})$ *and in* r_k, $k \neq j$.

   (b)

$$y(r_1, \dots, r_{j-1}, r_j e^{i\theta_j}, r_{j+1}, \dots, r_d) = o\left(\frac{y(\mathbf{r})}{\sqrt{B_{jj}(\mathbf{r})}}\right), \quad \text{as } r_j \to R_j,$$

   uniformly for $|\theta_j| \geq \delta(\mathbf{r})$ *and in* r_k, $k \neq j$.

   are valid.

Of course, this definition implies the following theorem.

Theorem 20 *Let* $y(\mathbf{x})$ *be a function defined in (53) which is H-admissible. Then as* $\mathbf{r} \to R$ *we have*

$$a_{\mathbf{n}} = \frac{y(\mathbf{r})}{\mathbf{r}^{\mathbf{n}}(2\pi)^{d/2}\sqrt{\det B(\mathbf{r})}}\left(\exp\left(-\frac{1}{2}(\mathbf{a}(\mathbf{r}) - \mathbf{n})^t B(\mathbf{r})^{-1}(\mathbf{a}(\mathbf{r}) - \mathbf{n})\right) + o(1)\right),$$

as $\|\mathbf{n}\| \to \infty$*, uniformly in* $\mathbf{n}$.

Problem 8 *Is is true that $e^{y(\mathbf{x})}$ is H-adissible if $y(\mathbf{x})$ is H-admissible? Is it true that $e^{P(\mathbf{x})}$ is H-admissible if (almost) all coefficients of $e^{P(\mathbf{x})}$ are positive? Is it possible to strengthen the definition of H-admissibility to (multivariate) power-admissibility such that $e^{y(\mathbf{x})}$ is power-adissible if $y(\mathbf{x})$ is power-admissible?*

Suggested Method. Hayman [64] used a series of partly tricky lemmas to prove the closure properties stated in Theorem 12. We have tried to formulate the above definition of multivariate H-admissibility in a way that these lemmas should generalize. So the first question is more or less settled. So there is realistic hope that the other two questions can be treated, too. □

3.4 General Urn Models

We now assume a more general situation. We also have m urn but $d \geq 1$ different kinds of balls we can throw. This means, that there are now d generating functions $g_j(z) = \sum_k a_{j,k} z^j$, $1 \leq j \leq d$, which govern the allocation of these kinds of balls. Now, the parameter we want to count (in each urn) depends on the numbers k_j, $1 \leq j \leq d$, of balls of type j, which have been thrown into this urn and will denoted by $f(k_1, k_2, \ldots, k_d)$. Hence, $\Phi_1(x, z_1, z_2, \ldots, z_d)$ is given by

$$\Phi_1(x, z_1, z_2, \ldots, z_d) = \phi_1(x, z_1, z_2, \ldots, z_d)^m$$

with

$$\phi_1(x, z_1, z_2, \ldots, z_d) = \sum_{k_1, \ldots, k_d} a_{1,k_1} \cdots a_{d,k_d} x^{f(k_1, k_2, \ldots, k_d)} z_1^{k_1} \cdots z_d^{k_d}.$$

In a similar fashion we can also define ϕ_2, ϕ_3, etc.

The difference between this model and the previous one is that we now have different kinds of balls, which means that we have to describe the total number X as a d-dimensional process (or a stochastic field) $X_m(n_1, n_2, \ldots, n_d)$.

Problem 9 *Show (under suitable analyticiy conditions on ϕ_r, $r \geq 1$) that there exists a Gaussian field $G(t_1, \ldots, t_d)$ such that the normalized field $X_m(n_1, n_2, \ldots, n_d)$ converges weakly to $G(t_1, \ldots, t_d)$.*

Suggested Method. It seems that the analytic method presented in [38] can be generalized. Since every function $\Phi_r = (\phi_r)^m$ is a m-th power it is clear that we always have finite dimensional Gaussian distributions. The real problems are to show that the Gaussian field which is suggested by the finite dimensional limiting distributions has a continuous realization (i.e. the covariance function is continuous) and to prove that the normalized field $X_m(n_1, n_2, \ldots, n_d)$ is tight. It looks as if everything can be worked out as in [38] but it is surely much more involved. □

3.5 Binary and m-Ary Search Trees

We now briefely describe the combinatorial concept of [28, 31]. As already indicated, the exponential generating functions $y_h(x)$ of the number $a_{n,h}$ are recursively given by

$$\begin{aligned} y_0(x) &= 1, \\ y'_{h+1}(x) &= y_h(x)^2, \ y_{h+1}(0) = 1. \end{aligned}$$

(Please note that this definition is equivalent to the preceding one.) The idea is to imitate this recurrence.

The key observation is that the following retarded differential equation

$$\Phi'(u) = -\frac{1}{\alpha^2}\Phi(u/\alpha)^2 \tag{56}$$

with the initial condition $\Phi(0) = 1$, where α is an abbreviation for

$$\alpha = e^{1/c} = 1.26\ldots$$

has a unique entire solution with the following properties (see [28]):

1. $\Phi(u)$ is strictly decreasing for (real) $u \geq 0$.
2. $u\Phi(u)$ is strictly increasing for (real) $u \geq 0$.
3. We have
$$1 - u\Phi(u) \sim C_1 \frac{\log u}{u^{c-1}} \tag{57}$$
as $u \to \infty$ for some real constant $C_1 > 0$.

If we define the set of auxiliary functions $\tilde{y}_h(x)$ $(h \geq 0)$ defined by

$$\tilde{y}_h(x) := \alpha^h \Phi(\alpha^h(1-x)), \tag{58}$$

where $\alpha = e^{1/c} = 1.26\ldots$ is as above and $\Phi(u)$ is the analytic solution of (56) with initial condition $\Phi(0) = 1$. Then these functions simulate the recurrence of $y_h(x)$:

1. $\tilde{y}_h(0) < 1$.
2. $1 - \tilde{y}_h(0) \sim \frac{C_1}{c} h \left(\frac{2}{c}\right)^h \quad (h \to \infty)$.
3. $\tilde{y}_h(1) = \alpha^h$.
4. $\tilde{y}_{h+r}(x) \geq \tilde{y}_h(x)$ for all $x \geq 0$ and $r \geq 0$.
5. $\tilde{y}'_{h+1}(x) = \tilde{y}_h(x)^2$.

The only (but important) difference between $y_h(x)$ and $\tilde{y}_h(x)$ is that they do not have the same initial condition. We have $\tilde{y}_h(0) = \alpha^h\Phi(\alpha^h) < 1$. This means that $\tilde{y}_h(x) < y_h(x)$ for $0 \leq x \leq 1$ (since we have $\tilde{y}_0(x) < y_0(x)$ just for $x < 1$).

In Theorem 15 the quantity $e_h = c\log y_h(1)$ was defined. This quantity was chosen in a way that $y_h(1) = \tilde{y}_{e_h}(1)$ and so one might conjecture that $y_h(x)$ is approximated by $\tilde{y}_{e_h}(x)$. In fact it is shown, that $y_h(x) \leq \tilde{y}_{e_h}(x)$ for $x \geq 1$ and $y_h(x) \geq \tilde{y}_{e_h}(x)$ for $x \leq 1$ (see [31]) which is sufficient to prove Theorem 15, i.e. we get proper estimates for the coefficients $a_{n,h}/n!$. However, we can conjecture even a little bit more.

Problem 10 *Is it true that for n, h with $c\log n - e_h = \mathcal{O}(1)$ we uniformly have*

$$[x^n]y_h(x) = [x^n]\tilde{y}_{e_h}(x) + o(1)\ ?$$

Suggested Method. It seems that $y_h(1) = \tilde{y}_{e_h}(1)$ *ensures* that $y_h(x)$ is close to $\tilde{y}_{e_h}(x)$ for x close to 1. So it should be possible to compare their coefficients via Cauchy's formula:

$$[x^n]y_h(x) = \frac{1}{2\pi i}\int_{|x|=1} y_h(x)x^{-n-1}\,dx \sim \frac{1}{2\pi i}\int_{|x|=1} \tilde{y}_{e_h}(x)x^{-n-1}\,dx = [x^n]\tilde{y}_{e_h}(x).\square$$

The next problem concerns the coefficents of $\tilde{y}_h(x)$. It turns out that they could be asymptotically determined if $\Phi(u)$ is a Laplace transform.

Problem 11 *Is $\Phi(u)$, the unique solution of (56) with $\Phi(0) = 1$ the Laplace transform of some function $\Psi(y)$, $y \geq 0$, which is decreasing and satisfies $\Psi(0) = 1$?*

Suggested Method. One can try to solve the integral equation

$$\int_0^y \Psi(z)\Psi(y-z)\,dy = y\Psi(y/\alpha) \tag{59}$$

or use the inverse Laplace transform and some regularity properties of $\Phi(u)$ to ensure that $\Psi(y)$ exists. □

It is the clear that if $\Psi(y)$ exists we get

$$[x^n]\tilde{y}_h(x) = \frac{1}{n!}\int_0^\infty y^n e^{-y}\Psi(y\alpha^{-h})\,dy \approx \Psi(n\alpha^{-h}).$$

if h is close to $c\log n$. So, if both problems can be solved one would obtain asymptotic expressiong for $a_{n,h}/n! = \mathbf{P}[H_n \leq h]$ of the form

$$\mathbf{P}[H_n \leq h] \sim \Psi(n/y_h(1))$$

for $c\log n - e_h = \mathcal{O}(1)$. This means, we would get a very precise estimate for the distribution.

We now turn to m-ary search trees. They are direct generalization of binary search trees, for a description see e.g. [77]. Devroye [21] has proved the following asymptotic proptery concering the height H_n of m-ary search trees. (For simplicity we assume that we fix some $m \geq 2$ and will not indicate it as a subscript or superscript.) Let $c = c_m > 1/(h_m - 1)$ ($h_m = \sum_{i=1}^m 1/i$) be the minimal solution of the equations

$$\beta + c\log(m!) = c\sum_{i=1}^{m-1}\log(\beta+i),$$

where β is linked with c by

$$\frac{1}{c} = \sum_{i=1}^{m-1}\frac{1}{\beta+i}.$$

Then we have

$$\frac{H_n}{c_m\log n} \to 1 \quad a.s.$$

We now indicate how the present approach for binary search trees might be applied for m-ary search trees in order to prove that the variance (and all centralized moments) of H_n are bounded.

Problem 12 *Let $m \geq 3$ be given. Is there an entire solution $\Phi(u)$ with $\Phi(0) = 1$ of the differential equation*

$$\Phi^{(m-1)}(u) = (-1)^m\frac{(m-1)!}{\alpha^m}\Phi\left(\frac{u}{\alpha}\right)^m, \tag{60}$$

where $\alpha = e^{1/c_m}$ such that for $0 \leq j \leq m-1$

$$0 < (-1)^j\Phi^{(j)}(u) < \frac{(j-1)!}{u^j}$$

for (real) $u > 0$?

Suggested Method. One can try to immitate the method of [28], where the case $m = 2$ is solved. The only (but essential) problem is that is seems that for $m > 2$ there is only one choice for $\Phi'(0), \Phi''(0), \ldots, \Phi^{(m-2)}(0)$, for which the solution of (60) has the desired property. Another way would be a direct solution of an integral equation similarly to (59). □

The point is that if this problem can be solved then it follows (almost automatically) that the variance of the height H_n of m-ary search trees is bounded, see [31].

We now turn to another problem, to the contour of m-ary search trees.

Problem 13 *Is there a analogue to Theorem 17 and its Corollary for m-ary search trees?*

Suggested Method. The difference between binary search trees and m-ary search trees is that for binary search trees it is quite easy to obtain an explicit expression for $\mathbf{E}U_k(n)$, see section 1.2. For m-ary search trees this seems to be much more difficult and most probably we have to use more asymptotic methods. However, this problem seems not to be out of reach. □

In section 1.2 it is indicated how we can get explicit (and asymptotic) expressions for $\mathbf{E}U_k(n)$. It is due to H.K. Hwang (personal communication) that it is also possible to get *explicit* representations for $\mathbf{E}U_k(n)(U_k(n)-1)$:

$$\mathbf{E}U_k(n)(U_k(n)-1) = \frac{2^k}{(2\pi i)^2} \oint \oint \frac{w^{-k} \binom{n+\frac{1}{y}-2}{n}}{\Gamma(\frac{1}{y}-1)(1-(1+w)y)\sqrt{1-8wy^2}}\, dy\, dw.$$

Similar integral representations can also be obtained for $\mathbf{E}U_k(n)U_m(n)$.

Problem 14 *Determine the exact asmptotic behaviour of* $\mathbf{E}U_k(n)U_k(n)$ *and* $\mathbf{E}U_k(n)U_m(n)$ *for all values k, m (as $n \to \infty$).*

Suggested Method. Saddle point methods. □

3.6 Digital Expansions

In section 2.6 we collected some facts on the distribution of *digital sums.* Firstly we presented a very general theorem (Theorem 18) on the distribution of $s_F(n)$ for $n \in \{0, 1, 2, \ldots, N-1\}$, where very general number systems are considered. We also mentioned Bassily and Katai's ([6]) result on the distribution of $s(P(n))$, where $P(x)$ in an integral polynomial and $s(n)$ the q-ary sum-of-digits function. Therefore it is natural to consider the following problem.

Problem 15 *Suppose that $G = (G_j)_{j\ge 0}$ satisfies a linear recurrence of the type presented in section 2.6 and let a function F be as above, too. Is it true that for every integral polynomial $P(x)$ the limiting distribution of the numbers $s_F(P(n))$, $0 \le n < N$, is asymptotically normal?*

Suggested Method. Bassily and Katai's method uses estimates for exponential sums in order to show that the limiting distribution of $s(P(n))$ and $s(n)$, $0 \le n < N$, are the same. Estimates for specific exponential sums are quite well understood for number systems, for which the sequence G_j satisfies a linear recurrence, see [60]. So we will try to modify the methods of [60] so that they can be *inserted* into Bassily and Katai's method. □

The most striking result which has been stated in section 2.6 is Kim's result (Theorem 19). It may be interpreted in the way that the sum-of-digits function with respect to different bases $q_1, q_2, \ldots, q_k$ (which are pairwisely relatively prime) behaves asymptotically independent. So one is led to the following problem.

Problem 16 *Suppose that* $q_1, q_2, \ldots, q_k > 1$ *are pairwisely relatively prime integers. Is it true that the corresponding sum-of-digits functions* $s_{q_1}(n), s_{q_2}(n), \ldots, s_{q_k}(n)$ *admit a multivariate Gaussian limit law with no correlation?*

Suggested Method. This problem is surely very difficult. The technical problem is that it is not sufficient to estimate exponential sums but to get the asymptotic leading term. If we interprete $(s_{q_1}(n), s_{q_2}(n), \ldots, s_{q_k}(n))$ as a random vector then the characteristic function of this random vector is exactly the exponential sum

$$\frac{1}{N}\sum_{n=0}^{N} e^{it_1 s_{q_1}(n)+it_2 s_{q_2}(n)+\cdots+it_k s_{q_k}(n)}.$$

Kim's method provides upper bounds for these exponential sums which have the *expected* order of magnitude. However, if we want to apply Levy's theorem in order to prove asymptotic normality we have to prove convergence of characteristic funcions. So we need an additional idea, however, it seems that Kim's method (or a proper variation of it) is a major step to the solution of Problem 16. □

Remark. Problem 16 is the only overlap between the already granted FWF-project *Statistical Problems of Number Systems.*

If one compares Kim's and Hoit's result one may also ask whether Kim's result is also true for more general number systems.

Problem 17 *Suppose that* $G = (G_j)_{j\ge 0}$ *and* $H = (H_j)_{j\ge 0}$ *satisfy linear recurrencex of the type presented in section 2.6 such that the dominanting roots* α *and* β *of the characteristic polynomials are independent, i.e.,* $\log\alpha/\log\beta \notin \mathbf{Q}$. *Let* $s_G(n)$, $s_H(n)$ *denote the corresponding sum-of-digits functions. Is it true that for (almost) all* m_1, m_2 *and all* c_1, c_2

$$|\{n < N : s_G(n) \equiv c_1 \bmod m_1, s_H(n) \equiv c_2 \bmod m_2\}| = \frac{N}{m_1 m_2} + \mathcal{O}(N^{1-\delta})$$

for some $\delta > 0$*?*

Suggested Method. Kim's method relies on estimates for exponential sums and on a very tricky application of Weyl's inequality (for exponential sums) which allows to *separate* the contributions of the different q_i. So we have two problems. The first one is to imitate the *separation procedure* with help of Weyl's inequality and the second one is to consider proper exponential sums concering these general number systems. However, we already mentioned that in Problem 15 that specific exponential sums have been already treated. So it seems that this problem is feasible, too. □

4 International Cooperations

4.1 INRIA

The applicants have frequent contacts to the algorithms group of INRIA leaded by Philippe Flajolet. Special cooperations are planed within the topics *Systems of Functional Equations* and *Binary and* m*-Ary Search Trees.* Furthermore, every year a workshop *Average Case Analysis of Algorithms* is organized by Philippe Flajolet, Helmut Prodinger, Rainer Kemp et al. Drmota has attended all of these workshops and Gittenberger more than half of them. These workshops had turned out to be a steady source of inspriation for many topics which has been treated during the last years and also for this proposal.

4.2 University of Versailles

There is a very close connection to the University of Versailles, e.g. Gittenberger was a visiting professor there in May 1999 and Drmota in September 1999. We want to work together with Daniele Gardy on *General Urn Models* and with Brigitte Chauvin and Jean Jabbour-Hattab on *Binary Search Trees.* There already exist joint papers [15, 38].

4.3 University of Brussels

Both applicants have joint papers with Guy Louchard from Brussels [12, 55, 56]. It is planned to cooperate with Louchard within the topic *Galton-Watson trees* especially for density representations for the joint distrtibution of $(W^+(s), l(t))$.

4.4 Academy of Sciences Taipei

Hsien-Kuei Hwang from the Academy of Sciences Taipei is also a frequent participant of the workshops on the *Average Case Analysis of Algorithms.* He has visited Graz and Vienna in July 1999 and we already started a cooperation within the topic *Binary Search Trees*, see section 3.5.

References

[1] D. J. Aldous, *The continuum random tree I*, Ann. Prob. 19 (1991), 1–28.

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

[3] D. J. Aldous, *Brownian excursion conditioned on ist local time*, Electr. Comm. Probab. **3** (1998), 79–90.

[4] K. B. Athreya and P. E. Ney, *Branching Processes*, Springer, Berlin-Heidelberg-New York, 1972.

[5] G. Barat, R. F. Tichy, and R. Tijdeman, *Digital blocks in linear numeration systems*, Proceedings of the Number Theory Conference, Zakopane, 1997, to appear.

[6] N. L. Bassily and I. Kátai, *Distribution of the values of q-additive functions on polynomial sequences*, Acta Math. Hung. **68** (1995), 353–361.

[7] R. Bellman and H. N. Shapiro, *On a problem in additive number theory*, Ann. Math. **49** (1948), 333–340.

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

[9] E. A. Bender and L. B. Richmond, *Admissible functions and asymptotics for labelled structures by number of components.* Electron. J. Comb. **3** (1996), Research paper R34, 23 p.; J. Comb. **3** (1996), 479–501.

[10] J. D. Biggins, *How fast does a general branching random walk spread?*, in: IMA Vol. Math. Appl. **84** (1997), 19–39.

[11] S. Boucheron and D. Gardy, *An urn model from learning theory*, Random Struct. Alg. **10** (1997), 43–67.

[12] T. Bruss, M. Drmota and G. Louchard, *The Complete Solution of the Competitive Rank Selection Problem,* Algorithmica **22** (1998), 413–447.

[13] L. E. Bush, *An asymptotic formula for the average sum of the digits of integers*, Am. Math. Monthly **47** (1940), 154–156.

[14] D. Cateland, *Suites digitales et suites k-régulières*, Thèse de doctorat, Université de Bordeaux 1 (1992).

[15] B. Chauvin, M. Drmota, and J. Jabbour-Hattab, *The profile of binary search trees*, manuscript.

[16] P. Chassaing and J.-F. Marckert, *Parking functions, empirical processes, and the width of rooted labeled trees*, manuscript.

[17] J. Coquet, *Power sums of digital sums*, J. Number Th. **22** (1986), 161–176.

[18] H. Delange *Sur la fonction sommatoire de la fonction "Somme de Chiffres"*, L'Enseignement math. **21** (1975), 31–77.

[19] L. Devroye, *A note on the height of binary search trees*, J. Assoc. Comput. Mach. **33** (1986), 489–498.

[20] L. Devroye, *Branching processes in the analysis of the height of trees*, Acta Inform. **24** (1987), 277–298.

[21] L. Devroye, *On the height of random m-ary search trees*, Random Struc. Algorithms **1** (1990), 191–203.

[22] L. Devroye and B. Reed, *On the variance of the height of random binary search trees*, SIAM J. Comput. **24** (1995), 1157–1162.

[23] J. W. Cohen and G. Hooghiemstra, *Brownian excursion, the $M/M/1$ queue and their occupation times*, Mathematics of Operations Research 6, 4 (1981), 608–629.

[24] M. Drmota, *Asymptotic distributions and a multivariate Darboux method in enumeration problems,* J. Comb. Theory, Ser. A, **67,** 2 (1994), 169–184.

[25] M. Drmota, *A bivariate asymptotic expansion of coefficients of powers of generating functions,* Europ. J. Combinatorics **15** (1994), pp. 139–152.

[26] M. Drmota, *Systems of functional equations* Random Struct. Algorithms, **10** (1997), 103–124.

[27] M. Drmota, *On nodes of given degree in random trees*, in: Probabilistic Methods in Discrete Mathematics (V.F. Kolchin et al. eds.), VSP, Utrecht, 1997, 31–44.

[28] M. Drmota *An analytic approach to the height of binary search trees,* Algorithmica, to appear.

[29] M. Drmota *The distribution of patterns in digital expansions,* Proceedings, Number Theory Conference, Graz, to appear.

[30] M. Drmota *Analysis of generating functions and probabilities on trees,* Prob. Theory and Math. Stat. (B. Grigelionis et al., eds.), to appear

[31] M. Drmota *An analytic approach to the height of binary search trees. II,* submitted.

[32] M. Drmota and J. Gajdosik, *The parity of the sum-of-digits-function of generalized Zeckendorff representations*, Fibonacci Quarterly **36** (1998), 3–19.

[33] M. Drmota and J. Gajdosik, *The distribution of the sum-of-digits-function*, J. Theor. Nombres Bordx. **10** (1998), 17–32.

[34] M. Drmota and B. Gittenberger, *Strata of random mappings – a combinatorial approach*, Stoch. Proc. Appl. **82** (1999), 157–171.

[35] M. Drmota and B. Gittenberger, *On the profile of random trees*, Random Struct. Algorithms **10** (1997), 421–451.

[36] M. Drmota and B. Gittenberger, *The distribution of nodes of given degree in random trees*, J. Graph Theory **31** (1999), 227–253.

[37] M. Drmota and B. Gittenberger, *Asymptotic Expansions for Coefficients of Multivariate Generating Functions*, in preparation.

[38] M. Drmota, D. Gardy, and B. Gittenberger, *A unified presentation of some urn models*, Algorithmica, to appear.

[39] M. Drmota and M. Skałba, *The parity of the Zeckendorf sum-of-digits-function*, manuscripta math., to appear.

[40] M. Drmota and M. Soria, Marking in combinatorial constructions: generating functions and limiting distributions, *Theoretical Comp. Science 144* (1995), 67–99.

[41] M. Drmota and V. Vatutin, *Limiting distributions in branching processes with two types of particles*, in: Classical and Modern Branching Processes (K.B. Athreya and P. Jagers eds), Springer, New York, 1997, 89–110.

[42] J. M. Dumont and A. Thomas, *Digital sum moments and substitutions*, Acta Arith. **64** (1993), 205–225.

[43] J. M. Dumont and A. Thomas, *Gaussian asymptotic properties of the sum-of-digits functions*, J. Number Th. **62** (1997), 19–38.

[44] P. Flajolet and A. M. Odlyzko, *Singularity analysis of generating functions*, SIAM J. Discrete Math. 3, 2 (1990), 216–240.

[45] P. Flajolet and M. Soria, *Gaussian limiting distributions for the number of components in combinatorial structures*, J. Combinatorial Th. (Ser. A) **53** (1990), 165–182.

[46] D. Gardy, *Normal limiting distributions for projection and semijoin sizes*, SIAM J. Discrete Math. **5** (1992), 219–248.

[47] D. Gardy, *Join sizes, urn models and normal limiting distributions*, Theoretical Computer Science **131** (1994), 375–414.

[48] D. Gardy and G. Louchard, *Dynamic analysis of some relational data bases parameters*, Theoretical Computer Science, **144** (1995), 125-159.

[49] A. O. Gelfond, *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith. **13** (1968), 259–265.

[50] R. K. Getoor and M. J. Sharpe, *Excursions of Brownian motion and Bessel processes*, Z. Wahrsch. verw. Gebiete 47 (1979), 83-106.

[51] B. Gittenberger, *On the contour of random trees*, SIAM J. Discrete Math., im Erscheinen.

[52] B. Gittenberger, *Convergence of branching processes to the local time of a Bessel process*, Random Struct. Algorithms **13** (1998), 423–438.

[53] B. Gittenberger, *The number of descendents in simply generated random trees*, Combinatorics, Probability and Computing, to appear.

[54] B. Gittenberger, *On the contour of random trees*, Siam J. Discrete Math, to apprear.

[55] B. Gittenberger and G. Louchard, *The Brownian excursion multidimensional local time density*, J. Appl. Probab., im Erscheinen.

[56] B. Gittenberger and G. Louchard, *On the local time density of the reflected Brownian bridge*, J. Appl. Math. Stoch. Analysis., to appear.

[57] B. Gittenberger and J. Thuswaldner, *Asymptotic normality of b-additive functions on polynomial sequences in the Gaussian number field*, J. Number Th., to appear.

[58] I. J. Good, *Saddle-Point Methods for the Multinomial Distribution,* Ann. Math. Statistics **28** (1957), pp. 861–881.

[59] P. J. Grabner, P. Kirschenhofer, H. Prodinger, and R. F. Tichy, *On the moments of the sum-of-digits function*, in: Applications of Fibonacci Numbers **5** (1993), 263–271

[60] P. J. Grabner and R. F. Tichy, *Contributions to digit expansions with respect to linear recurrences*, J. Number Th. **36** (1990), 160–169.

[61] P. Grabner and R. F. Tichy, *α-Expansions, linear recurrences, and the sum-of-digits function*, manuscripta math. **70** (1991), 311–324.

[62] D. H. Greene and D. E. Knuth, *Mathematics for the Analysis of Algorithms* (2^{nd} edition), Birkhäuser, Boston-Basel-Stuttgart, 1982.

[63] W. Gutjahr and G. Ch. Pflug, The asymptotic contour process of a binary tree is a Brownian excursion, *Stochastic Processes and their Applications 41* (1992), 69–89.

[64] W. K. Hayman, A generalization of Stirling's formula, Journal f. d. reine u. angewandte Math. 196 (1956), 67–95.

[65] A. Hoit, *The distribution of generalized sum-of-digits functions in residue classes*, J. Number Th., to appear.

[66] D. P. Kennedy, *The Galton-Watson process condidioned on the total progeny,* J. Appl. Prob. **12** (1975), 800–806.

[67] R. E. Kennedy and C. N. Cooper, *An extension of a theorem by Cheo and Yien concerning digital sums*, Fibonacci Q. **29** (1991), 145–149.

[68] G. Kersting, *On the height profile of a conditioned Galton-Watson tree*, manuscript.

[69] D.-H. Kim, *On the joint distribution of q-additive functions in residue classes* , J. Number Theory **74** (1999), 307 – 336.

[70] P. Kirschenhofer, *On the variance of the sum of digits function*, Lecture Notes Math. **1452** (1990), 112–116.

[71] F. B. Knight, *On the excursion process of Brownian motion*, Trans. AMS 258 (1980), 77–86.

[72] V. F. Kolchin, *Random Mappings,* Optimization Software, New York, 1986.

[73] S. P. Lalley, *Finite range random walk on free groups and homogeneous trees*, Ann. Probab. **21** (1993), 2087–2130.

[74] H. Leinfellner *New results on rarefied sums of the Thue-Morse sequence,* Grazer Math. Berichte, to appear.

[75] G. Louchard, *Kac's formula, Levy's local time and Brownian excursion,* J. Appl. Prob. 21 (1984), 479-499.

[76] H. M. Mahmoud, *Evolution of Random Search Trees*, John Wiley & Sons, New York, 1992.

[77] H. M. Mahmoud, *On the average internal path length of m-ary search trees*, Acta Inf. **23** (1986), 111–117.

[78] A. Meir and J. W. Moon, *On the Altitude of Nodes in Random Trees*, Canadian Journal of Mathematics 30 (1978), 997–1015.

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

[80] L. Moser and M. Wyman, *Asymptotic development of the Stirling numbers of the first kind* J. London Math. Soc. **33** (1958), 133–146.

[81] L. Moser and M. Wyman, *Stirling numbers of the second kind,* Duke Math. J. **25** (1958), 29–43.

[82] J. Neveu, *Processus aléatoires gaussiens,* Presses de l'Université de Montréal, 1968.

[83] A. Pethö and R. F. Tichy, *On digit expansions with respect to linear recurrences*, J. Number Th. **33** (1989), 243–256.

[84] J. Pitman, *The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree of forest*, Ann. Prob., to appear.

[85] B. Pittel, *On growing random binary trees*, J. Math. Anal. Appl. **103** (1984), 461–480.

[86] B. Reed, *How tall is a tree*, preprint.

[87] J. M. Robson, *The height of binary search trees*, Austral. Comput. J. **11** (1979), 151–153.

[88] J. M. Robson, *On the concentration of the height of binary search trees.* ICALP 97 Proceedings, LNCS **1256** (1997), 441–448.

[89] J. M. Robson, *Constant bounds on the moments of the height of binary search trees*, Theoret. Comput. Sci, to appear.

[90] L. Takács, *Conditional limit theorems for branching processes*, J. Appl. Math. Stoch. Analysis 4, 4 (1991), 263–292.

[91] J. Schmid, *The joint distribution of the binary digits of integer multiples*, Acta Arith. **43** (1984), 391–415.

[92] W. M. Schmidt, *The joint distribution the digits of certain integer s-tuples*, Studies in pure mathematics, Mem. of P. Turan (1983), 605–622.

[93] G. Tenenbaum, *Sur la non-dérivaibité de fonctions périodiques associées à certaines formules sommatoires*, In: The mathematics of Paul Erdős, I (R. L. Graham and J. Nesestril eds.), Springer, Berlin, 1997, 117–128.

[94] H. Trollope, *An explicit expression for binary digital sums*, Meth. Mag. **41** (1968), 21–25.

[95] J. S. Vitter and P. Flajolet, *Average-Case Analysis of Algorithms and Data Structures*, in Handbook of Theoretical Computer Science, J. van Leeuwen, Ed., vol. A: Algorithms and Complexity. North Holland, 1990, ch. 9, pp. 431–524.

[96] A. R. Woods, *Coloring rules for finite trees, and probabilities of monadic second order sentences.* Random Struct. Algorithms **10** (1997), 453–485.