Singularity Analysis, Hadamard Products, and Tree Recurrences

James Allen Fill[1] Philippe Flajolet[2]
Nevin Kapur[1]

[1] The Johns Hopkins University, Department of Mathematical Sciences, 34th and Charles Streets, Baltimore, MD 21218-2682 (USA)
{jimfill,nevin}@jhu.edu, http://www.mts.jhu.edu/~{fill,kapur}/

[2] ALGORITHMS Project, INRIA, F-78153 Rocquencourt (France)
Philippe.Flajolet@inria.fr, http://algo.inria.fr/flajolet/

June 12, 2003

Abstract

We present a toolbox for extracting asymptotic information on the coefficients of combinatorial generating functions. This toolbox notably includes a treatment of the effect of Hadamard products on singularities in the context of the complex Tauberian technique known as singularity analysis. As a consequence, it becomes possible to unify the analysis of a number of divide-and-conquer algorithms, or equivalently random tree models, including several classical methods for sorting, searching, and dynamically managing equivalence relations.

Keywords: Singularity analysis, Hadamard products, tree recurrences, asymptotic expansions, generating functions, divide-and-conquer, shape functional, generalized polylogarithm, contour integration, moment pumping
2000 Mathematics Subject Classification: 05A16 (Primary); 40E99, 68W40 (Secondary)

This study was motivated by a desire to unify the analysis of a number of algorithms and data structures of computer science. By analysis we mean here (precise) average-case analysis of cost functions as introduced by Knuth and illustrated in the collection [41] as well as in his monumental series, *The Art of Computer Programming* (see especially [39, 40]). In the first part of this paper (Section 1 and 2), we consider a major paradigm of algorithmic design, the "divide-and-conquer" principle, which is closely related to families of random trees and associated "tree recurrences". The basic framework is described in Section 1, while lead examples are introduced in Section 2 below. Our treatment rests on combinatorial generating functions.

The central part of this paper (Sections 3 and 4) is devoted to the process of extracting coefficients, at least asymptotically, from generating functions. Singularities have long been recognized to contain highly useful information in this regard, and we start by recalling in Section 3 the basic principles of the complex Tauberian approach known as "*singularity analysis*". Applications to algorithms and trees require, in particular, techniques for coping with generating functions that may be constructed by a tower of several transformations. Here, we develop the theory of composition of singularities under Hadamard products in Section 4. (The reader only interested in complex-analytic aspects can jump directly to Sections 3 and 4.)

The final part (Sections 5 and 6) returns to the original problem of analysing divide-and-conquer algorithms, taking full advantage of the analytic results of previous sections. Tree recurrences and first moments form the subject of Section 5, where full asymptotic expansions are derived for expectations of costs. Section 6 describes possible extensions of the basic framework to the determination of variances and higher moments as well as to some other random tree models.

1 Introduction

"Divide-and-Conquer" is a major principle of algorithmic design in computer science. An instance (I) of a problem to be solved is first split into smaller subproblems (I', I'') that are solved recursively by the same process; the partial solutions are then woven back to yield a solution to the original problem. The abstract scheme is then of the form:

$$\begin{array}{lll} \text{solve}(I) & := & (I', I'') := \text{split}(I); \\ & & J' := \text{solve}(I');\ J'' := \text{solve}(I''); \\ & & \text{return weave}(J', J''). \end{array} \tag{1}$$

(Problems of size smaller than a certain threshold are treated directly without any recursive call.) Algorithms resorting to the scheme (1) include classical sorting methods (mergesort, quicksort, radix-exchange sort), data structures based on trees (binary search trees, digital trees known as "tries", quadtrees for multidimensional search, union–find trees) as well as various methods used in computational geometry, distributed computation, and communication theory. We refer the reader to classical books on data structures, algorithms, and analysis of algorithms for details, for instance, [10, 31, 35, 40, 47, 48, 57, 58, 60, 62].

In general, a class of probabilistic models $\mathfrak{M}_n$ indexed by the size n of the problem instance is assumed to reflect the nature of data fed to the algorithm. A cost function—typically, the number of certain operations performed by the algorithm—then becomes a random variable X_n whose form is induced by $\mathfrak{M}_n$ and the particular divide-and-conquer algorithm considered. The problem is then to obtain characteristics of X_n, for instance its mean, higher moments, or even distributional information. The asymptotic limit $n \to \infty$ is usually considered, since an important phenomenon of "asymptotic simplification" is to be expected in a large number of situations.

Under natural conditions, a recurrence that closely mimics the recursive structure of (1) relates the random variables X_n:

$$X_n = t_n + X_{K_n} + \widetilde{X}_{n-a-K_n}. \tag{2}$$

The interpretation is as follows: t_n is a quantity[1], called the "toll", that represents the cost incurred by splitting the initial instance and weaving back the final solution; K_n is the (random) size of the first subproblem, in which case, the second subproblem has a size that is the complement of K_n to $n-a$, for some small constant a (usually, $a=0$ or $a=1$), which is specific to the algorithm considered. The random variables of type X and K are assumed to be independent, as are the two X-sequences X and $\widetilde{X}$ on the right in (2), and a subproblem of size k is assumed to satisfy model $\mathfrak{M}_k$—this property is sometimes called "randomness preservation" and is satisfied by many cases of algorithmic interest. A direct asymptotic treatment of the recursive relation (2) binding random variables is sometimes feasible; see the (metric) "contraction method" surveyed by Rösler and Rüschendorf [54] and applied by Neininger [50] to a subset of the problems discussed here.

Turning to average-case analysis, the expected cost $f_n := \mathbb{E}(X_n)$ satisfies a *recurrence* that is directly implied by (2):

$$f_n = t_n + \sum_k p_{n,k}(f_k + f_{n-a-k}). \tag{3}$$

with the *splitting probabilities* $p_{n,k} := \Pr(K_n = k)$ being determined by the model $\mathfrak{M}_n$ used. Trees are naturally associated with recursive procedures, and, accordingly, the recurrence (3) can be viewed as associated with a random tree model of the following form: the root has size a, the left subtree has size k with probability $p_{n,k}$, and the right subtree has the remaining quantity $n-a-k$ as size. Then (3)is interpreted as giving the expectation of a cost function over the tree structure that is induced by the family of tolls, t_n. For this reason, a recurrence having the form (3) is called a *tree recurrence.* Tree recurrences are the main object of study of this paper.

One way to view the tree recurrence (3) is as a linear transformation on sequences

$$(f_n) = \mathcal{K}\left[(t_n)\right], \tag{4}$$

that takes a toll sequence (t_n) and returns the corresponding average-cost sequence (f_n). The functional $\mathcal{K}$ is fully determined by the splitting probabilities $p_{n,k}$. A classical approach to the derivation of explicit forms consists in introducing *generating functions* (GFs). Fix a sequence of *normalization constants* ω_n (that are problem-specific) and define the generating functions

$$f(z) := \sum f_n \omega_n z^n, \qquad t(z) := \sum t_n \omega_n z^n.$$

[1]Some analyses require a randomly varying toll. For mean value analysis, the distinction between deterministic and stochastic tolls is, however, immaterial.

Then, the transformation $\mathcal{K}$ induces another linear transformation $\mathcal{L}$ on GFs:

$$f(z) = \mathcal{L}\left[t(z)\right]. \tag{5}$$

With an adequate choice of the constants ω_n, explicit forms of f_n can often be obtained, provided at least that toll sequences are of a simple enough form.

Our main objective is to develop generating-function methods by which one can quantify the way the *asymptotic form* of (expected) costs relates to properties of the *toll sequence*. It is known that asymptotic properties of number sequences (as the index n tends to infinity) are closely related to the nature of the *singularities* of the corresponding generating functions. This suggests that we examine the way the operator $\mathcal{L}$ operates on scales of singular functions and view it as a "*singularity transformer*". Informally, there is a transformation $\widehat{\mathcal{L}}$, induced by $\mathcal{L}$ and acting on an asymptotic scale of functions singular at some fixed point z_0. Using $\mathrm{Sing}(f(z))$ to denote the expansion of $f(z)$ at the singularity z_0, one has

$$\mathrm{Sing}(f(z)) = \widehat{\mathcal{L}}\left[t(z)\right] \tag{6}$$

Under fairly general conditions, there is a tight coupling between singular expansions of a generating function and the asymptotic form of its coefficients. The outcome of this process, justified by *singularity analysis* [24, 51], is a direct relation written figuratively as

$$\mathrm{Asympt}((f_n)) = \widehat{\mathcal{K}}\left[(t_n)\right], \tag{7}$$

where $\widehat{\mathcal{K}}$ depends on (t_n) via the structure of its generating function $t(z)$.

The path we follow in this paper is the one given by (4)–(7), which is then globally summarized by the following diagram:

$$\begin{array}{ccc} (t_n) \xrightarrow{\mathcal{K}} (f_n) & & (t_n) \xrightarrow{\widehat{\mathcal{K}}} \mathrm{Asympt}((f_n)) \\ \Downarrow & & \Uparrow \\ t(z) \xrightarrow{\mathcal{L}} f(z) & \Longrightarrow & t(z) \xrightarrow{\widehat{\mathcal{L}}} \mathrm{Sing}(f(z)) \end{array} \tag{8}$$

We propose to develop a collection of generic tools that supplement the basic singularity analysis framework of Flajolet and Odlyzko [24]. In particular, we discuss in the next sections the action on singularities of differential and integral operators, as well as of Hadamard products. As a result, the way $\mathcal{L}$ operators associated with many recurrences transform singularities can be analyzed precisely. This in turn yields a fairly general classification of the asymptotic growth phenomena associated to a variety of classical tree recurrences, including the ones of binary search trees, binary trees, and union–find trees, which will serve here as guiding examples.

Most of the existing computer science literature is devoted to the "deterministic" divide-and-conquer recurrences that correspond to a splitting size K_n that is deterministic, depending on n alone—typically, $K_n \equiv \lfloor n/2 \rfloor$. In such a case, the probability distribution $(p_{n,k})_{k=0}^{n}$ is supported at a single point. The main asymptotic order of f_n is then given by what Cormen, Rivest, and Leiserson have termed "master theorems": see [10, 31, 58]. (Usually, the finer characteristics of the asymptotic regime involve fractal fluctuations [22, 58].) What we consider here instead are methods for dealing with "stochastic" divide-and-conquer recurrences, where K_n is a random variable (dependent on n) with support spread over a whole subinterval in $(0, n)$. This stochastic case is discussed by Roura in [55]: Roura's arguments are based on elementary real analysis, so that they are of quite a wide scope, but his estimates are by nature mostly confined to first-order asymptotics. In this article we show that, in the many cases of practical interest where some strong complex-analytic structure is present, full asymptotic expansions can be derived. Our treatment is somewhat parallel in spirit to that of Knuth and Pittel whose inspiring work [42] provided one of the initial motivations[2] for the present study. An additional benefit of the complex-analytic approach is that it often gives access to variances and higher moments, in which case the limit distribution of costs can be identified.

2 Some "special" tree recurrences

In this section, we briefly review some tree recurrences that are of special interest in combinatorial mathematics and analysis of algorithms.

2.1 The binary search tree recurrence

One of the simplest model of random trees is defined as follows: To determine a tree T_n of size $n \geq 1$, take a root and append to it a left subtree of size k and a right subtree of size $n-k$, where k is uniformly distributed over the set of permissible values $\{0, 1, \ldots n-1\}$; a tree of size 0 is the empty tree. In earlier notations, this process corresponds to

$$p_{n,k} \equiv \Pr(K_n = k) := \frac{1}{n}, \qquad \text{for} \quad k = 0, 1, \ldots n-1. \tag{9}$$

As is well known, the model defined by (9) corresponds to random trees defined by either the binary search tree data structure or the quicksort algorithm [40, 47, 48, 58, 62]. The corresponding tree recurrence (3) is then

$$f_n = t_n + \frac{2}{n} \sum_{k=0}^{n-1} f_k, \tag{10}$$

with $f_0 := t_0$.

[2] See also Pittel's interesting recent article [53] which appeared while our own work was still in progress.

Ordinary GFs determined by the choice of coefficients $\omega_n = 1$ for all n are then

$$f(z) := \sum_{n\geq 0} f_n z^n, \qquad t(z) := \sum_{k\geq 0} t_n z^n,$$

and standard rules for the manipulation of GFs translate (10) into a linear integral equation

$$f(z) = t(z) + 2\int_0^z f(w)\,\frac{dw}{1-w}.$$

Differentiation yields the ordinary differential equation

$$f'(z) = t'(z) + \frac{2}{1-z} f(z),$$

which is then solved by the variation-of-constants method:

$$f(z) = \mathcal{L}[t(z)], \qquad \text{where} \quad \mathcal{L}[t(z)] := (1-z)^{-2}\int_0^z (\partial_w t(w))\,(1-w)^2\,dw. \quad (11)$$

In (11) we have assumed without loss of generality the initial conditions $t_0 = f_0 = 0$ (thanks to linearity and the fact that the transform of $t_n \equiv \delta_{n,0}$ is $n+1$). The notation ∂_w borrowed from differential algebra is used to denote derivatives whenever the operator nature of transformations is to be stressed.

It is instructive to follow what Greene and Knuth call the "repertoire" approach [31]. This consists in building a repertoire of the ($\mathcal{K}$ or $\mathcal{L}$) transforms of basic tolls, then trying to determine the effect of a new toll by expressing it in the basis of known tolls. What is convenient here is the class of tolls

$$t_n^\alpha := \binom{n+\alpha}{\alpha} = \frac{(\alpha+1)(\alpha+2)\cdots(\alpha+n)}{n!}, \qquad \text{i.e.,} \quad t^\alpha(z) = (1-z)^{-\alpha-1}.$$

Then, by (11) one finds, for $\alpha \neq 1$,

$$\begin{aligned} f^\alpha(z) &= \frac{\alpha+1}{\alpha-1}\left[(1-z)^{-\alpha-1} - (1-z)^{-2}\right] \\ f_n^\alpha &= \frac{\alpha+1}{\alpha-1}\left[\binom{n+\alpha}{\alpha} - \binom{n+1}{1}\right], \end{aligned}$$

while $\alpha = 1$ leads to

$$f^1(z) = \frac{2}{(1-z)^2}\log\frac{1}{1-z}, \qquad f_n^1 = 2(n+1)\,(H_{n+1} - 1)\,,$$

with $H_n := 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ the nth harmonic number.

Stirling's formula implies asymptotically, for α not a negative integer,

$$\binom{n+\alpha}{\alpha} \sim \frac{n^\alpha}{\Gamma(\alpha+1)},$$

with Γ the Euler gamma function. Then what goes on is summarized by the following table:

$$
\begin{array}{ll}
t_n = \binom{n+\alpha}{\alpha},\ \alpha > 1 & f_n \sim \frac{\alpha+1}{\alpha-1}\frac{n^\alpha}{\Gamma(\alpha+1)} \\
t_n = n+1 & f_n \sim 2n\log n \\
t_n = \binom{n+\alpha}{\alpha},\ 0 < \alpha < 1 & f_n \sim \frac{1+\alpha}{1-\alpha} n.
\end{array}
\qquad (12)
$$

The discontinuity in the asymptotic regime of f at $\alpha = 1$, where a logarithm appears, is noticeable. Also, the tolls in the scale satisfying $t_n \ll n$ are seen to induce costs that all collapse to linear functions.

A full discussion of the binary search tree recurrence necessitates determining the effect of toll functions like $\sqrt{n}$, $\log n$, and $1/n^2$, a task which is not entirely elementary. By the remarks of the introduction, this involves determining the singularities of the corresponding generating functions *and*, in view of (11), making explicit the way singular expansions get composed under differentiation and integration (Section 3). This subject will then be taken up again in Section 5.1; the particular case of the toll $t_n = \log n$ is of special importance and will be treated in detail there.

2.2 The uniform binary tree recurrence

This recurrence is of the form ($n \geq 0$, with the convention $f_0 := t_0$)

$$f_n = t_n + \sum_{k=0}^{n-1} \frac{C_k C_{n-1-k}}{C_n} \left(f_k + f_{n-k}\right), \qquad \text{with} \quad C_n := \frac{1}{n+1}\binom{2n}{n}, \qquad (13)$$

a Catalan number. It corresponds to the uniform model of binary trees, where all the C_n binary trees with n internal nodes are taken with equal likelihood. Indeed, the number of trees of size n satisfies the recurrence

$$C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k} \quad (n \geq 1), \qquad C_0 = 1, \qquad (14)$$

as seen from a root decomposition. The quantity $p_{n,k} = C_k C_{n-1-k}/C_n$ is then the probability that a tree of size n has left and right subtrees of respective sizes k and $n-1-k$.

The GF of Catalan numbers satisfies a relation that is the image of the recurrence (14), namely, $C(z) = 1 + zC(z)^2$, so that

$$C(z) = \frac{1}{2z}\left(1 - \sqrt{1-4z}\right). \qquad (15)$$

In order to solve (13) by generating functions, one should use as normalization constants the quantities $\omega_n = C_n$, and introduce

$$t(z) := \sum_{n\geq 0} t_n C_n z^n, \qquad f(z) := \sum_{n\geq 0} f_n C_n z^n. \qquad (16)$$

Then (13) translates into a linear algebraic equation,

$$f(z) = t(z) + 2zC(z)f(z),$$

from which the form of the $\mathcal{L}$ operator immediately results:

$$f(z) = \mathcal{L}[t(z)], \qquad \text{where} \quad \mathcal{L}[t(z)] = \frac{1}{\sqrt{1-4z}} t(z). \tag{17}$$

This form makes it possible to analyze directly only a restricted collection of tolls, for instance, ones of the form $t_n^r := (n+1)n \cdots (n-r+2)$ (by differentiation), or $t_n^{-r} = 1/((n+2)(n+3)\cdots(n+r+1))$ (by integration). However, tolls of such simple forms as $\sqrt{n}$, H_n, and $\log n$, are left out of the scale of the $t_n^{\pm r}$.

Define the *Hadamard product* of two entire series or two functions analytic at the origin, a and b, as their termwise product,

$$a(z) \odot b(z) = \sum_{n \geq 0} a_n b_n z^n, \qquad \text{if} \quad a(z) = \sum_{n \geq 0} a_n z^n, \quad b(z) = \sum_{n \geq 0} b_n z^n. \tag{18}$$

Then, from (16) and (17), the cost functional is expressed by the modified transformation (of $\mathcal{L}$ type)

$$f(z) = \frac{\tau(z) \odot C(z)}{\sqrt{1-4z}}, \quad \text{where} \quad f(z) = \sum_n f_n C_n z^n, \quad \tau(z) := \sum_n t_n z^n. \tag{19}$$

This now relates the *ordinary* generating function $\tau(z)$ of the tolls and the *normalized* generating function $f(z)$ of the costs (with the $\omega_n = C_n$ normalization) via a Hadamard product.

Determining the way costs get transformed under this model then necessitates a way to combine singular expansions under Hadamard products. This is the central part of our article; see Section 4, where a general theorem is stated. The "critical" value for tolls at which a discontinuity in the induced costs manifests itself is now at $t_n = \sqrt{n}$, and[3]

$t_n = n^\alpha$, $\alpha > 1/2$	$f_n = \Theta(n^{\alpha+\frac{1}{2}})$	(20)
$t_n = n^{1/2}$	$f_n = \Theta(n \log n)$	
$t_n = n^\alpha$, $0 < \alpha < 1/2$	$f_n = \Theta(n)$.	

This phenomenon observed in [28, Prop. 2] [of which (20) above corrects a few misprints] neatly distinguishes the binary Catalan model from the binary search tree model, as seen by comparing (20) to (12). A proof accompanied by complete expansions will be given in the application section: see Section 5.2 below.

[3]The notation $x = \Theta(y)$ expresses the inequalities $c_1 y < x < c_2 y$ for some constants c_1, c_2 satisfying $0 < c_1 < c_2 < +\infty$.

2.3 The union–find tree recurrence

By a result attributed to Cayley, there are $U_n = n^{n-2}$ "free" unrooted trees (i.e., labeled connected acyclic graphs) on n nodes, and, accordingly, $T_n = n^{n-1}$ rooted trees. Consider the model in which initially each unrooted tree of size n is taken with equal likelihood. Choose an edge at random amongst any of the possible $n-1$ edges of the tree, orient it in a random way, then cut it. This separates the tree into an ordered pair of smaller trees that are now rooted. Continue the process with each of the resulting subtrees, discarding the root. Assume that the cost incurred by selecting the edge and splitting the tree is t_n. Then the total cost incurred when starting from a random unrooted tree and recursively splitting it till the completely disconnected graph is obtained satisfies the recurrence ($n \geq 1$)

$$f_n = t_n + \sum_{k=1}^{n-1} p_{n,k}(f_k + f_{n-k}), \quad \text{where} \quad p_{n,k} = \binom{n}{k} \frac{k^{k-1}(n-k)^{n-k-1}}{2(n-1)\, n^{n-2}}. \tag{21}$$

(Proof: There are n^{n-1} rooted trees on n nodes and the binomial coefficient takes care of relabellings.) The recurrence (21) has been studied in great detail by Knuth and Pittel in [42], an article that largely motivated our study. In fact, there are good algorithmic reasons for considering the recurrence (21): if time is reversed, then the recursion describes the evolution of a random graph from totally disconnected to tree-like, when successive edges are added at random. The latter is exactly the probabilistic model involved in the "union–find" (or equivalence-finding) algorithm [10, 57, 62], for which detailed analyses had been provided by Knuth and Schönhage [43] in 1978[4]. (Note that this model is not the same as the simply generated family of Cayley trees.)

Let $T(z), U(z)$ be the exponential generating functions of the sequences (T_n), (U_n), that is,

$$T(z) = \sum_{n=1}^{\infty} n^{n-1} \frac{z^n}{n!}, \qquad U(z) = \sum_{n=1}^{\infty} n^{n-2} \frac{z^n}{n!},$$

It is a well-known fact of combinatorics that $T(z)$ satisfies the functional relation $T(z) = ze^{T(z)}$, and one has $U = T - (T^2/2)$; see [30, 39, 59]. Define now the generating functions

$$t(z) = \sum_{n\geq 1} t_n (n^{n-1} - n^{n-2}) \frac{z^n}{n!}, \qquad f(z) = \sum_{n \geq 1} f_n n^{n-1} \frac{z^n}{n!},$$

where the normalization constants for $f(z)$ are $\omega_n = n^{n-1}/n!$ and, for convenience, a marginally different normalization, $\omega'_n = n^{n-2}(n-1)/n!$, has been introduced in the case of $t(z)$. Then the recurrence (21) has the form of a

[4]Precisely, the model is known as the "random spanning tree model". The derivation of our equation (22) closely mimics Section 11 of [43].

binomial convolution, so that the cost GF $f(z)$ satisfies

$$f(z) - \int_0^z f(w)\,\frac{dw}{w} = t(z) + f(z)T(z).$$

By differentiation, this last relation transforms into a linear differential equation of the first order, itself readily solved by the variation-of-constants method. Assuming (without loss of generality) the initial condition $t_1 = f_1 = 0$, the solution found is

$$f(z) = \frac{T(z)}{1-T(z)} \int_0^z \partial_w t(w)\,\frac{dw}{T(w)}. \tag{22}$$

In terms of the ordinary generating function of costs, namely,

$$\tau(z) := \sum_{n\geq 2} t_n z^n,$$

equation (22) can be rephrased as an integral transform involving a Hadamard product, namely,

$$f(z) = \frac{1}{2}\frac{T(z)}{1-T(z)} \int_0^z \partial_w \left(\tau(w) \odot T(w)^2\right)\,\frac{dw}{T(w)}. \tag{23}$$

The dominant singularity at $z = e^{-1}$ of the Cayley tree function T is well known to be of the square root type. Then the integral transform (23) operates in a way that combines a Hadamard product and ordinary products, as well as integration and differentiation. This subject will be resumed in Section 5.3, after general theorems have been established by which one can cope with such situations. The final conclusions turn out to be qualitatively similar to what was observed for the Catalan model in (20).

3 Singular expansions, differentiation, and integration

Singularities of generating functions encode very precise information regarding the asymptotic behaviour of coefficients. In this section, we first recall in Subsection 3.1 the principles of a process by which this information can be extracted: this is the singularity analysis framework of [24, 51]. We then prove that functions amenable to singularity analysis are closed under integration and differentiation; see Subsection 3.2. These operations have already been seen to intervene in the analysis of some of the major tree recurrences.

3.1 Basics of singularity analysis

Singularity analysis deals with functions that have isolated singularities on the boundary of their disc of convergence and are consequently continuable to wider areas of the complex plane. The case of a unique dominant singularity suffices for the applications treated here. (In addition, the case of finitely many dominant singularities is easily reduced to this situation by using composite contours and cumulating contributions arising from individual singularities.) Given the obvious scaling rule,

$$[z^n]f(z) = \rho^{-n} f(\rho z),$$

one may restrict attention, whenever necessary, to the case where the singularity is at 1. The scaling rule shows that the position of the singularity [at ρ for $f(z)$] introduces an exponential scaling factor (ρ^{-n}) multiplied by the coefficient of a function singular at 1 [the function $f(\rho z)$].

Definition 1. *A function defined by a Taylor series with radius of convergence equal to* 1 *is* Δ-regular *if it can be analytically continued in a domain*

$$\Delta(\phi, \eta) := \{z : |z| < 1 + \eta, |\operatorname{Arg}(z-1)| > \phi\},$$

for some $\eta > 0$ *and* $0 < \phi < \pi/2$. *A function* f *is said to admit a* singular expansion *at* $z = 1$ *if it is* Δ*-regular and*

$$f(z) = \sum_{j=0}^{J} c_j (1-z)^{\alpha_j} + O(|1-z|^A) \tag{24}$$

uniformly in $z \in \Delta(\phi, \eta)$, *for a sequence of complex numbers* $(c_j)_{0 \le j \le J}$ *and an increasing sequence of real numbers* $(\alpha_j)_{0 \le j \le J}$ *satisfying* $\alpha_j < A$. *It is said to satisfy a singular expansion* "with logarithmic terms" *if, similarly,*

$$f(z) = \sum_{j=0}^{J} c_j\,(L(z))\,(1-z)^{\alpha_j} + O(|1-z|^A), \qquad L(z) := \log\frac{1}{1-z}, \tag{25}$$

where each $c_j(\cdot)$ *is a polynomial.*

Note that, by assumption, the $O(\cdot)$ error term in (24) must hold uniformly in $z \in \Delta(\phi, \eta)$. We also allow in the usual way infinite asymptotic expansions representing an infinite collection of mutually compatible expansions of type (24).

For the sake of notational simplicity, we shall mostly limit our statements to the basic case (24) and briefly comment on how they extend to the logarithmic case (25). The basic theorem is the following:

Theorem 2 (Basic singularity analysis [24]). *If* $f(z)$ *admits a singular expansion of the form (24) valid in a* Δ*-domain, then*

$$[z^n]f(z) = \sum_{j=0}^{J} c_j \binom{n-\alpha_j-1}{-\alpha_j-1} + O(n^{-A-1}). \tag{26}$$

(The proof of this and similar results is based on an extensive use of Hankel contours; see the already cited references.) The last expansion can be rephrased as a standard asymptotic expansion since, for $\alpha \notin \{0, 1, 2 \ldots\}$, one has

$$\binom{n-\alpha-1}{-\alpha-1} \sim \frac{n^{-\alpha-1}}{\Gamma(-\alpha)}\left(1 + \frac{\alpha(\alpha+1)}{2n} + \frac{\alpha(\alpha+1)(\alpha+2)(3\alpha+1)}{24n^2} + \cdots\right),$$

while all the terms corresponding to α a nonnegative integer have an asymptotically null contribution. When logarithmic terms are present in the singular expansion, corresponding logarithmic terms arise in the asymptotic expansion of coefficients. The calculations are conveniently carried out by differentiation with respect to the parameter α:

$$[z^n](1-z)^\alpha L(z)^r = (-1)^r \frac{\partial^r}{\partial \alpha^r}[z^n](1-z)^\alpha = (-1)^r \frac{\partial^r}{\partial \alpha^r}\binom{n-\alpha-1}{-\alpha-1},$$

which yields for instance ($\alpha \notin \{0, 1, 2 \ldots\}$):

$$\begin{aligned}
[z^n](1-z)^\alpha L(z) &= -\frac{\partial}{\partial \alpha}\binom{n-\alpha-1}{-\alpha-1} \\
&= \binom{n-\alpha-1}{-\alpha-1}\left(\frac{1}{-\alpha} + \frac{1}{1-\alpha} + \cdots + \frac{1}{n-1-\alpha}\right) \\
&= \frac{n^{-\alpha-1}}{\Gamma(-\alpha)}\left(\log n - \psi(-\alpha) + O\left(\frac{\log n}{n}\right)\right).
\end{aligned}$$

(Here ψ is the logarithmic derivative of Γ.)

The same proof techniques also make it possible to translate error terms involving logarithmic terms; see [24] for details. In particular, the following transfer holds for A and B real numbers:

$$O((1-z)^A L^B(z)) \qquad \rightsquigarrow \qquad O(n^{-A-1} \log^B n). \tag{27}$$

Finally, we shall make use of a result which renders amenable to singularity analysis generating functions whose coefficients involve powers of n and its logarithms.

Definition 3. *The generalized polylogarithm* $\mathrm{Li}_{\alpha,r}$*, where* α *is an arbitrary complex number and* r *a nonnegative integer is defined for* $|z| < 1$ *by*

$$\mathrm{Li}_{\alpha,r}(z) := \sum_{n \geq 1} (\log n)^r \frac{z^n}{n^\alpha},$$

and the notation Li_α *abbreviates* $\mathrm{Li}_{\alpha,0}$.

In particular, one has $\mathrm{Li}_{1,0}(z) = \mathrm{Li}_1(z) = L(z)$, the usual logarithm, cf. (25). The singular expansion of the polylogarithm, taken from [21], involves the Riemann ζ function:

Theorem 4 (Singularities of polylogarithms [21]). *The function* $\mathrm{Li}_{\alpha,r}(z)$ *is* Δ*–continuable and, for* $\alpha \notin \{1, 2, \ldots\}$*, it satisfies the singular expansion*

$$\mathrm{Li}_{\alpha,0}(z) \sim \Gamma(1-\alpha) w^{\alpha-1} + \sum_{j \geq 0} \frac{(-1)^j}{j!}\zeta(\alpha - j) w^j, \quad w = -\log z = \sum_{\ell=1}^{\infty} \frac{(1-z)^\ell}{\ell}. \tag{28}$$

For $r > 0$, the singular expansion of $\mathrm{Li}_{\alpha,r}$ *is obtained by*

$$\mathrm{Li}_{\alpha,r}(z) = (-1)^r \frac{\partial^r}{\partial\alpha^r} \mathrm{Li}_{\alpha,0}(z),$$

and corresponding termwise differentiation of (28) *with respect to* α.

In particular, for $\alpha < 1$, the main asymptotic term of $\mathrm{Li}_{\alpha,r}$ is

$$\Gamma(1-\alpha)(1-z)^{\alpha-1}L^r(z).$$

Similar expansions hold when α is a positive integer; see [21] for details.

Example 5. *Stirling's formulæ.* The factorial function, is attainable via the form

$$\log n! = \log 1 + \log 2 + \cdots + \log n = [z^n]\frac{1}{1-z} \mathrm{Li}_{0,1}(z),$$

to which singularity analysis can be applied now that we have taken ordinary generating functions. Theorem 4 yields the singular expansion

$$\frac{1}{1-z} \mathrm{Li}_{0,1}(z) \sim \frac{L(z)-\gamma}{(1-z)^2} + \frac{1}{2}\frac{-L(z)+\gamma-1+\log 2\pi}{1-z} + \cdots,$$

from which Stirling's formula can be read off, by Theorem 2:

$$\log n! \sim n\log n - n + \frac{1}{2}\log n + \log\sqrt{2\pi} + \cdots.$$

[Stirling's constant $\log\sqrt{2\pi}$ comes out as $-\zeta'(0)$.] Similarly, the "superfactorial function",

$$S(n) := 1^1 \cdot 2^2 \cdots n^n \equiv \frac{(n!)^{n+1}}{1!\,2!\,\cdots\,n!},$$

satisfies

$$\log S(n) = [z^n]\frac{1}{1-z} \mathrm{Li}_{-1,1}(z),$$

which gives rise to a second-order "Stirling's formula",

$$S(n) \sim n^{\frac{1}{2}n^2+\frac{1}{2}n+\frac{1}{12}} e^{-\frac{1}{4}n^2} A,$$

with

$$A := \exp\left(\frac{1}{12} - \zeta'(-1)\right) = \exp\left(-\frac{\zeta'(2)}{2\pi^2} + \frac{\log(2\pi)+\gamma}{12}\right).$$

(This last expansion, originally due to Glaisher, Jeffery, and Kinkelin, goes back to the 1860s and it can be established by Euler–Maclaurin summation; see Finch's book [20] for context and references.) The systematic character of the derivation given here clearly applies to many similar functions. □

Methods of the last example may be used more generally to determine the Euler–Maclaurin constant relative to sums of the form $\sum(\log n)^r/n^s$. The derivation by singularity analysis is quite systematic and several formulæ of Ramanujan can be obtained in this way, for instance,

$$\lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{\log^k n}{n}-\frac{\log^{k+1}N}{k+1}\right)=A_k,\ \text{with}\ A_k:=\frac{(-1)^k}{k+1}\frac{d^{k+1}}{ds^{k+1}}\left((s-1)\zeta(s)\right)_{s=1},$$

involving the Stieltjes constants A_k. See Berndt's account of the problem in [3, p. 164] and references therein.

3.2 Differentiation and integration

In preparation for our later treatment of Hadamard products, we need a theorem that enables us to differentiate local expansions of analytic functions around a singularity. Such a result cannot of course be unconditionally true; see, for example, (30). However, it turns out that functions amenable to singularity analysis satisfy this property. The statement that follows is an adaptation suited to our needs of well-known differentiability properties of complex asymptotic expansions (see especially Theorem I.4.2 of Olver's book [52, p. 9]).

Theorem 6 (Singular differentiation). *If $f(z)$ is Δ-regular and admits a singular expansion near its singularity in the sense of (24), then for each integer $r>0$, $\frac{d^r}{dz^r}f(z)$ is also Δ-regular and admits an expansion obtained through term-by-term differentiation:*

$$\frac{d^r}{dz^r}f(z)=(-1)^r\sum_{j=0}^{J}c_j\frac{\Gamma(\alpha_j+1)}{\Gamma(\alpha_j+1-r)}(1-z)^{\alpha_j-r}+O(|1-z|^{A-r}).$$

Proof. Clearly, all that is required is to establish the effect of differentiation on error terms, which is expressed symbolically as

$$\frac{d}{dz}O(|1-z|^A)=O(|1-z|^{A-1}).$$

By iteration, only the case of a single differentiation ($r=1$) needs to be considered.

Let $g(z)$ be a function that is regular in a domain $\Delta(\phi,\eta)$ where it is assumed to satisfy $g(z)=O(|1-z|^A)$ for $z\in\Delta$. Choose a subdomain $\Delta':=\Delta(\phi',\eta')$, where $\phi<\phi'<\frac{\pi}{2}$ and $0<\eta'<\eta$. By elementary geometry, for any sufficiently small $\kappa>0$, the disc of radius $\kappa|z-1|$ centered at a value $z\in\Delta'$ lies entirely in Δ; see Figure 1. We fix such a small value κ and let $\gamma(z)$ represent the boundary of that disc oriented positively.

The starting point is Cauchy's integral formula

$$g'(z)=\frac{1}{2\pi i}\int_C g(w)\frac{dw}{(w-z)^2}, \tag{29}$$

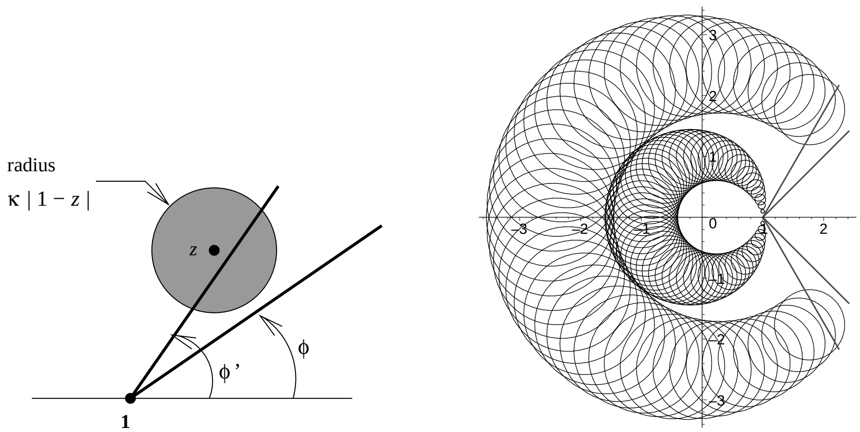

Figure 1: The contour $\gamma(z)$ used in the proof of the differentiation theorem: (left) the basic geometry; (right) two sets of circles $\gamma(z)$.

a direct consequence of the residue theorem. Here C should encircle z while lying inside the domain of regularity of g, and we opt for the choice $C \equiv \gamma(z)$. Then trivial bounds applied to (29) give:

$$\begin{aligned} |g'(z)| &= O\left(\|\gamma(z)\| \cdot |1-z|^A |1-z|^{-2}\right) \\ &= O\left(|1-z|^{A-1}\right). \end{aligned}$$

The estimate involves the length of the contour, $\|\gamma(z)\|$, which is $O(|1-z|)$ by construction, as well as the bound on g itself, which is $O(|1-z|^A)$ since all points of the contour are themselves at a distance exactly of the order of $|1-z|$ from 1. □

For instance, taking

$$g(z) = \cos\log\left(\frac{1}{1-z}\right) \quad \text{and} \quad g'(z) = -\frac{1}{1-z}\sin\log\left(\frac{1}{1-z}\right),$$

we correctly predict that $g(z) = O(1) \Rightarrow g'(z) = O(|1-z|^{-1})$. On the other hand, the apparent paradox given by the pair

$$g(z) = \cos\left(\frac{1}{1-z}\right) \qquad \text{and} \qquad g'(z) = -\frac{1}{(1-z)^2}\sin\left(\frac{1}{1-z}\right), \tag{30}$$

is resolved by observing that in no nondegenerate sector around $z = 1$ do we have $g(z) = O(1)$.

It is also well known that integration of asymptotic expansions is usually easier than differentiation. Here is a statement custom-tailored to our needs.

Theorem 7 (Singular integration). *Let $f(z)$ be Δ-regular and admit a Δ-expansion near its singularity in the sense of* (24). *Then $\int_0^z f(t)\,dt$ is also Δ-regular. Assume that none of the quantities α_j and A equals -1.*

(i) If $A < -1$, then the singular expansion of $\int f$ is

$$\int_0^z f(t)\,dt = -\sum_{j=0}^{J} \frac{c_j}{\alpha_j + 1}(1-z)^{\alpha_j+1} + O\left(|1-z|^{A+1}\right). \tag{31}$$

(ii) If $A > -1$, then the singular expansion of $\int f$ is

$$\int_0^z f(t)\,dt = -\sum_{j=0}^{J} \frac{c_j}{\alpha_j + 1}(1-z)^{\alpha_j+1} + L_0 + O\left(|1-z|^{A+1}\right),$$

where the "integration constant" L_0 has the value

$$L_0 := \sum_{\alpha_j < -1} \frac{c_j}{\alpha_j + 1} + \int_0^1 \left[f(t) - \sum_{\alpha_j < -1} c_j (1-t)^{\alpha_j}\right] dt.$$

Remark. The case where either some α_j or A is -1 is easily treated by the additional rules

$$\int_0^z (1-t)^{-1}\,dt = L(z), \qquad \int_0^z O(|1-t|^{-1})\,dt = O(L(z)).$$

Similar rules consistent with elementary integration are applicable for powers of logarithms: they are derived from the easy identities (for $\alpha \neq -1$)

$$\int_0^z (1-t)^\alpha L^r(t)\,dt = (-1)^r \frac{\partial^r}{\partial \alpha^r} \int_0^z (1-t)^\alpha\,dt = (-1)^{r+1} \frac{\partial^r}{\partial \alpha^r} \frac{(1-z)^{\alpha+1}}{\alpha+1},$$

for r a positive integer. Furthermore, the corresponding O–transfers hold true. (The proofs are simple modifications of the one given below for the basic case.)

Proof. The basic technique consists in integrating, term by term, the singular expansion of f. We let $r(z)$ be the remainder term in the expansion of f, that is,

$$r(z) := f(z) - \sum_{j=0}^{J} c_j (1-z)^{\alpha_j}.$$

By assumption, throughout the Δ-domain one has, for some positive constant K,

$$|r(z)| \le K|1-z|^A.$$

(*i*) *Case* $A < -1$. By straight-line integration between 0 and z, one finds (31), as soon as it has been established that

$$\int_0^z r(t)\,dt = O\left(|1-z|^{A+1}\right).$$

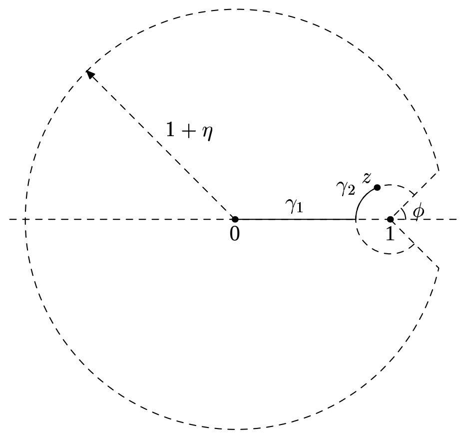

Figure 2: The contour used in the proof of the integration theorem.

By Cauchy's integral formula, we can choose any path of integration that stays within the region of analyticity of r. We choose the contour $\gamma := \gamma_1 \cup \gamma_2$, shown in Figure 2. Then[5]

$$\begin{aligned}\left|\int_\gamma r(t)\,dt\right| &\le \left|\int_{\gamma_1} r(t)\,dt\right| + \left|\int_{\gamma_2} r(t)\,dt\right| \\ &\le K\int_{\gamma_1} |1-t|^A\,|dt| + K\int_{\gamma_2} |1-t|^A|\,|dt| \\ &= O(|1-z|^{A+1}).\end{aligned}$$

Both integrals are $O(|1-z|^{A+1})$: for the integral along γ_1, this results from explicitly carrying out the integration; for the integral along γ_2, this results from the trivial bound $O(\|\gamma_2\|(1-z)^A)$.

(*ii*) *Case* $A > -1$. We let $f_-(z)$ represent the "divergence part" of f that gives rise to nonintegrability:

$$f_-(z) := \sum_{\alpha_j < -1} c_j (1-z)^{\alpha_j}.$$

Then with the decomposition $f = [f - f_-] + f_-$, integrations can be performed separately. First, one finds

$$\int_0^z f_-(t)\,dt = -\sum_{\alpha_j<-1} \frac{c_j}{\alpha_j+1}(1-z)^{\alpha_j+1} + \sum_{\alpha_j<-1} \frac{c_j}{\alpha_j+1}.$$

[5]The symbol $|dt|$ designates the differential line element (often denoted by ds) in the corresponding curvilinear integral.

Next, observe that the asymptotic condition guarantees the existence of $\int_0^1$ applied to $[f - f_-]$, so that

$$\int_0^z [f(t) - f_-(t)]\, dt = \int_0^1 [f(t) - f_-(t)]\, dt + \int_1^z [f(t) - f_-(t)]\, dt.$$

The first of these two integrals is a constant that contributes to L_0. As to the second integral, term-by-term integration yields

$$\int_1^z [f(t) - f_-(t)]\, dt = - \sum_{\alpha_j > -1} \frac{c_j}{\alpha_j + 1}(1-z)^{\alpha_j+1} + \int_1^z r(t)\, dt.$$

The remainder integral is finite, given the growth condition on the remainder term, and, upon carrying out the integration along the rectilinear segment joining 1 to z, trivial bounds show that it is indeed $O(|1-z|^{A+1})$. □

4 Hadamard products and transformation of singularities

In this section we propose to examine the way singular expansions get composed under Hadamard products defined at (18). The Hadamard product is a bilinear form. So if we have a set of functions admitting known singular expansions, we need to establish their composition law, and this will give composition rules for finite terminating expansions (Subsection 4.1). In order to extend this to asymptotic expansions with error terms, we need to establish a theorem providing the shape of

$$O(|1-z|^A) \odot O(|1-z|^B).$$

This is the more demanding part of the analysis, which is the subject of Subsection 4.2. Finally, in Subsection 4.3, we provide a summary statement, Theorem 11, to the effect that the class of functions amenable to singularity analysis is closed under Hadamard products and that the composition of singular expansions is effectively computable.

4.1 Composition of singular elements

The composition rule for polylogarithms is trivial, since

$$\mathrm{Li}_{\alpha,r}(z) \odot \mathrm{Li}_{\beta,s}(z) = \mathrm{Li}_{\alpha+\beta,r+s}(z).$$

However, polylogarithms do not have a simple composition rule with respect to ordinary products. We next turn to the composition rule for the basis formed by functions of the form $(1-z)^a$, where a may be any real number. From the expansion

$$(1-z)^a = 1 + \frac{-a}{1}z + \frac{(-a)(-a+1)}{2!}z^2 + \cdots \qquad (32)$$

around the origin, we get through term-by-term multiplication

$$(1-z)^a \odot (1-z)^b = {}_2F_1[-a,-b;1;z]. \tag{33}$$

Here ${}_2F_1$ represents the classical *hypergeometric function* of Gauss defined by

$${}_2F_1[\alpha,\beta;\gamma;z] = 1 + \frac{\alpha\beta}{\gamma}\frac{z}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}\frac{z^2}{2!} + \cdots . \tag{34}$$

From the transformation theory of hypergeometrics, see e.g. [34, p. 163], we know that, in general, hypergeometric functions can be expanded in the vicinity of $z = 1$ by means of the $z \mapsto 1-z$ transformation. Instantiation of this transformation with $\gamma = 1$ yields

$$\begin{aligned}{}_2F_1[\alpha,\beta;1;z] &= \frac{\Gamma(1-\alpha-\beta)}{\Gamma(1-\alpha)\Gamma(1-\beta)}\,{}_2F_1[\alpha,\beta;\alpha+\beta;1-z]\\ &+ \frac{\Gamma(\alpha+\beta-1)}{\Gamma(\alpha)\Gamma(\beta)}(1-z)^{-\alpha-\beta+1}\,{}_2F_1[1-\alpha,1-\beta;2-\alpha-\beta;1-z]. \end{aligned}\tag{35}$$

In other words, we can state the following proposition:

Proposition 8. *When a, b, and $a+b$ are not integers, the Hadamard product*

$$(1-z)^a \odot (1-z)^b$$

has an infinite Δ-expansion with exponent scale

$$\{0,1,2,\ldots\} \cup \{a+b+1, a+b+2, \ldots\},$$

namely,

$$(1-z)^a \odot (1-z)^b \sim \sum_{k\geq 0} \lambda_k^{(a,b)} \frac{(1-z)^k}{k!} + \sum_{k\geq 0} \mu_k^{(a,b)} \frac{(1-z)^{a+b+1+k}}{k!},$$

where the coefficients λ and μ are given by

$$\begin{aligned}\lambda_k^{(a,b)} &= \frac{\Gamma(1+a+b)}{\Gamma(1+a)\Gamma(1+b)}\frac{(-a)^{\overline{k}}(-b)^{\overline{k}}}{(-a-b)^{\overline{k}}},\\ \mu_k^{(a,b)} &= \frac{\Gamma(-a-b-1)}{\Gamma(-a)\Gamma(-b)}\frac{(1+a)^{\overline{k}}(1+b)^{\overline{k}}}{(2+a+b)^{\overline{k}}}.\end{aligned}$$

Here $x^{\overline{k}}$ is defined when k is a nonnegative integer as $x(x+1)\cdots(x+k-1)$.

Remark. The case where either a or b is an integer poses no difficulty: one has

— $(1-z)^a \odot g(z)$ is a polynomial if $a = m$, where $m \in \mathbb{Z}_{\geq 0}$;

— $(1-z)^a \odot g(z)$ is a derivative if $a = -m$ where $m \in \mathbb{Z}_{>0}$, since

$$(1-z)^{-m} \odot g(z) = \frac{1}{(m-1)!} \partial_z^{m-1} \left(z^{m-1} g(z) \right),$$

and this case is covered by singular differentiation, Theorem 6.

Notice that Proposition 8 remains valid in these two cases with the natural convention that $1/\Gamma(-j) = 0$ when $j \in \mathbb{Z}_{\geq 0}$.

The case where $a+b \in \mathbb{Z}$ needs transformation formulæ that extend (35) and are found explicitly in the books by Abramowitz and Stegun [1, pp. 559–560] and by Whittaker and Watson [63, §14.53].

Remark. The case of expansions with logarithmic terms is covered by "differentiation under the integral sign", as we now explain. Consider for instance the Hadamard product

$$\left[(1-z)^{-\alpha} L(z)\right] \odot (1-z)^{-\beta} = \frac{\partial}{\partial_\alpha} {}_2F_1[\alpha, \beta; 1; z],$$

where we assume for convenience that none of α, β, $\alpha + \beta$ is an integer. For any fixed β and any fixed z, with, say, $z \in (0,1)$, both sides of (35) represent analytic functions of α. Thus, their derivatives with respect to α are identical as functions of α. This induces a transformation formula, originally valid in the stated z-range, which involves modified hypergeometric functions (these have additional ψ-factors in their coefficients) obtained from the fundamental ${}_2F_1$ function by differentiation with respect to some of the parameters. The modified functions then do exist in extended regions of the complex z–plane as shown by taking the classical Barnes representations in terms of contour integrals (see, e.g., [63, §14.5]) and then differentiating under the integral sign. The net effect of this discussion is that the fundamental transformation (35) supports differentiation with respect to α, β and that the formally derived transformations provide analytically valid composition formulæ for Hadamard products

$$[(1-z)^{-\alpha} L^k(z)] \odot [(1-z)^{-\beta} L^\ell(z)] \tag{36}$$

of the base functions.

In practice, for all the cases described above, one may often proceed as follows: *(i)* take advantage of the a priori *existence* of a singular expansion of $f \odot g$, with $f(z) = (1-z)^a$, $g(z) = (1-z)^b$ or some of their derivatives, that is valid for z in a Δ-region (here the slit complex plane); *(ii)* compute an asymptotic expansion of the coefficients of $f \odot g$ by multiplication of the asymptotic expansions of f_n and g_n as obtained via singularity analysis; *(iii)* reconstruct a singular function that matches asymptotically $f_n g_n$ by using singularity analysis in the reverse direction. In Subsection 4.3, this process is formalized by the "Zigzag Algorithm" and illustrated by the return of Pólya's drunkard.

Globally, we are facing a situation where polylogarithms are simple for Hadamard products and relatively complicated for ordinary products, with the dual situation occurring in the case of power functions. Each particular situation is likely to dictate whether calculations are best expressed in a basis of standard singular functions like $\{(1-z)^a L(z)^k\}$ or with polylogarithms, $\{\mathrm{Li}_{\alpha,k}(z)\}$.

4.2 Composition of error terms

We now examine how $O(\cdot)$ terms get composed under Hadamard products. The task is easier when the resulting function gets large at its singularity as shown by Proposition 9. Fortunately, thanks to the results of Section 3 regarding differentiation and integration, all cases can be reduced to this one: see Proposition 10 below.

The starting point is a general integral formula due to Hadamard for $(f \odot g)(z)$, where

$$f(z) = \sum_{n \geq 0} f_n z^n \qquad \text{and} \qquad g(z) = \sum_{n \geq 0} g_n z^n.$$

Assume that f and g are analytic in the unit disc and let z be a complex number satisfying $|z| < 1$. Consider the integral

$$I = \frac{1}{2\pi i} \int_{\gamma_0} f(w) g\left(\frac{z}{w}\right) \frac{dw}{w}, \tag{37}$$

taken (counterclockwise) along a contour γ_0 which is simply a circle of radius ρ centered at the origin such that $|z| < \rho < 1$. In this way, both factors in the integrand are analytic functions of w along the contour. Evaluating the integral (37) by expanding the functions, we find

$$I = \sum_{n \geq 0} f_n g_n z^n.$$

This is the classical formula of Hadamard for Hadamard products,

$$(f \odot g)(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} f(w) g\left(\frac{z}{w}\right) \frac{dw}{w}, \tag{38}$$

valid, by analyticity, for any simple contour $\mathcal{C}$ such that each $w \in \mathcal{C}$ satisfies $|z| < |w| < 1$.

Proposition 9. *Assume that $f(z)$ and $g(z)$ are Δ-regular in $\Delta(\psi_0, \eta)$ and that*

$$f(z) = O(|1-z|^a) \;\text{ and }\; g(z) = O(|1-z|^b), \quad z \in \Delta(\psi_0, \eta),$$

where a and b satisfy $a + b + 1 < 0$. Then the Hadamard product $(f \odot g)(z)$ is regular in a (possibly smaller) Δ-domain, call it Δ', where it admits the expansion

$$(f \odot g)(z) = O(|1-z|^{a+b+1}). \tag{39}$$

Proof. We first observe[6] that $f \odot g$ is continuable to *certain* points z such that $|z| > 1$. (Precisely, as shown below, it admits a continuation in a Δ-domain.)

[6] This part of the argument is an adaptation to our needs of a famous result first due to Hadamard regarding the continuation of Hadamard products; see for instance the description in [4, Vol. II, p. 300] or [14, Sec. 88]. Accordingly, we limit ourselves to a succinct discussion only meant to set the stage for the precise estimates starting at (40).

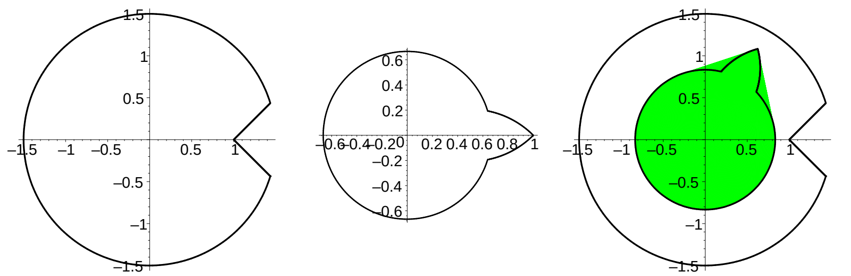

Figure 3: The geometry of Hadamard domains: (left) boundary of a Δ-domain ($1+\eta = 1.5$); (middle) boundary of Δ^{-1}; (right) an allowable domain in $\Delta \cap z\Delta^{-1}$ for application of Hadamard's formula is the unshaded subset of Δ ($|z| = 1.25$).

Indeed, because of the analytic continuation properties of f and g, both $f(w)$ and $g(z/w)$ are analytic functions of w in the domain $\Delta \cap (z\Delta^{-1})$, where Δ^{-1} denotes $\{w^{-1} : w \in \Delta\}$; see Figure 3 for a rendering. In other words, the allowed domain of values of w is Δ stripped of the internal domain $(z\Delta^{-1})^c$, where $(\cdot)^c$ represents complementation. Fix then some z_1 outside the unit disc but within Δ, and choose a simple contour γ_1 inside both Δ and $z_1\Delta^{-1}$. Let $I(z)$ be the integral of (37) and (38) taken along this fixed contour γ_1. (The feasibility of finding a suitable γ_1 is suggested by Figure 3, at least when $|z_1|$ remains close enough to 1 and z_1 is to the left of 1; a particular contour adapted to the case where z_1 is close to 1 and possibly to its right will be constructed explicitly in the proof below.) Now, when z moves radially along the segment $(0, z_1)$, the quantity $I(z)$ defines an analytic function of z that does coincide with $(f \odot g)(z)$ as soon as $|z| \le 1$ [this results from the "standard" formula (38)]. Thus analytic continuation of $f \odot g$, from within the unit disc to some z_1 lying outside of the unit disc is granted. The argument shows at the same time that Hadamard's formula (38) remains a valid representation of $f \odot g$ along such a contour γ_1 or any of its deformations legally granted by analyticity.

We next turn to estimating the growth at its singularity of $h := f \odot g$. It suffices to prove the estimate (39) on h for z belonging to a restricted domain $\Delta' := \Delta(\psi_1, \eta_1)$, where we shall take

$$\eta_1 = c_1\eta, \qquad \left(\frac{\pi}{2} - \psi_1\right) = c_1\left(\frac{\pi}{2} - \psi_0\right), \tag{40}$$

for some small positive constant c_1. Notice also that it suffices to establish the estimate of (39) for

$$|z - 1| < \eta_1 = c_1\eta \tag{41}$$

with $z \in \Delta'$, since h, being analytic in the rest of Δ', is certainly bounded there.

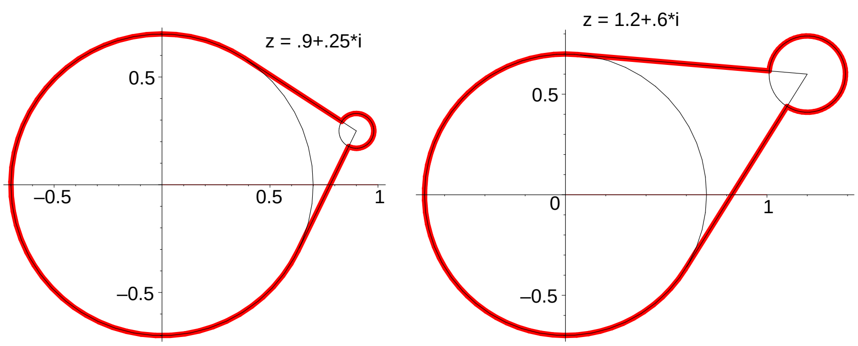

Figure 4: The geometry of the contour γ.

The main geometric objects from which the contour is built are as follows. First consider the circle centered at the origin

$$C_0 := \{w : |w| = R\}, \qquad R := 1 - c_2\eta \tag{42}$$

for some small constant c_2 (independent of z). Set $\delta = |z-1|$, which is the main parameter governing the scaling of the contour γ. We also consider the circle

$$C_z := \{w : |w-z| = c_3\delta\}, \tag{43}$$

for some small positive c_3. Finally, the contour γ includes parts of the two tangents T, T' to the circle C_0 issuing from z; see Figure 4. The contour is then precisely specified as

$$\gamma = \gamma_0 \cup \gamma_T \cup \gamma_z \cup \gamma_{T'},$$

where γ_T is the segment of T formed of points in between C_0 and z that are exterior to C_0 and C_z, and similarly for $\gamma_{T'}$. The component γ_0 is the part of the circle C_0 that lies on the "southwest" of 0 and joins with T, T'; the component γ_z is the part of the circle C_z that lies on the "northeast" of z and joins with T, T'. The constants c_1, c_2, c_3 are to be specified later and they can be taken as small as needed.

The fundamental constraint to be satisfied is that γ should lie entirely within $\Delta \cap (z\Delta^{-1})$ when z stays within Δ': for $w \in \gamma$, this ensures simultaneously $w \in \Delta$ and $z/w \in \Delta$, hence the validity of the Hadamard integral (38). By a priori choosing c_1 (which limits z) and c_3 (which controls the radius of C_z) both small enough, the condition $\gamma \subseteq \Delta$ is granted by elementary geometry. (E.g., the circle C_z will not extend too much to the right of $\Re(w) = 1$ and will therefore be "compatible" with the indentation of Δ at 1.) Next, one should have $\gamma \cap (z\Delta^{-1})^c = \emptyset$. This requires in particular choosing the radius R in (42) larger than $z(1+\eta)^{-1}$, which is at most $(1+c_1\eta)/(1+\eta)$ since z has been restricted to $|z-1| < c_1\eta$ by (41). This geometric condition expressed as

$$\frac{1+c_1\eta}{1+\eta} < 1 - c_2\eta \tag{44}$$

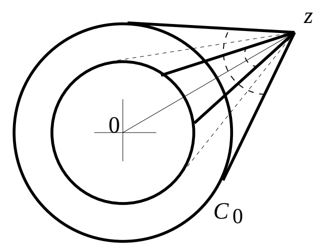

Figure 5: Apex avoidance condition: The angle at z of contour γ is constructed to be larger than the angle at z of the apex of $(z\Delta^{-1})^c$.

is granted as soon as c_1, c_2 are both taken small enough. [E.g., it suffices that both c_1, c_2 be less than $\frac{1}{2}(1+\eta)^{-1}$.] We henceforth assume these smallness conditions on c_1, c_2, c_3 to be satisfied. Finally, the contour should avoid the apex[7] of the domain $(z\Delta^{-1})^c$. Define the "viewing angle" of a point P exterior to a circle C as the angle betwen the two tangents to C issuing from P. For a circle of radius r and a point at distance d from the center, this angle is $2\arcsin(r/d)$. In particular the point z itself views the circle C_0 of radius R under the angle $2\arcsin(R/|z|)$, and this viewing angle is bounded from below by

$$2\arcsin\left(\frac{R}{1+c_1\eta}\right) = 2\arcsin\left(\frac{1-c_2\eta}{1+c_1\eta}\right),$$

since the farthest z can get from the origin is by assumption $1+c_1\eta$. It then suffices to choose c_1, c_2 so that

$$2\arcsin\left(\frac{1-c_2\eta}{1+c_1\eta}\right) > 2\psi_0 \tag{45}$$

(e.g., decide $c_2 = c_1$, then decrease $c_1 = c_2$ until the inequality in (45) is satisfied) in order to ensure that the angle under which z views the circle C_0 exceeds $2\psi_0$. Since C_0 encloses the inner disc of $(z\Delta^{-1})^c$ with which it is concentric, and since the angle at z of the apex of $(z\Delta^{-1})^c$ is $2\psi_0$, there results that the angle at z between γ_T and $\gamma_{T'}$ encompasses the apex of $(z\Delta^{-1})^c$; see Figure 5. In this way, the apex of $(z\Delta^{-1})^c$ is avoided.

Last, for λ any of the four contours of which γ is comprised, let $I(\lambda)$ be the integral of (37) taken along contour λ. The circular arc γ_z has all its points at a distance $c_3\delta$ from z, so that there

$$|1-w| = \Theta(\delta), \quad |z-w| = \Theta(\delta), \quad f(w)g\left(\frac{z}{w}\right) = O\left(\delta^{a+b}\right).$$

Therefore, by trivial bounds,

$$I(\gamma_z) = O\left(\delta^{a+b+1}\right). \tag{46}$$

[7]By the "apex" of $(z\Delta^{-1})^c$, we mean the complement in $(z\Delta^{-1})^c$ of the largest circular disc centered at the origin which is contained in $(z\Delta^{-1})^c$.

On the other hand, along γ_0 the functions $f(w)$ and $g(z/w)$ stay away from their singularities, so that

$$I(\gamma_0) = O(1). \tag{47}$$

There remains only to estimate the contribution along the two connecting segments γ_T and $\gamma_{T'}$. The two situations are similar (upon interchanging the roles of a and b). It is then easily seen that the contribution along the ray stemming from z is bounded from above by a multiple of an integral of the form

$$\int_{c_3\delta}^{+\infty} t^a |t - z_0|^b \, dt \tag{48}$$

where z_0 is a complex number at a distance $\Theta(\delta)$ from the real line. (The quantity t parameterizes the tangent line T or T'.) The last integral is $O(\delta^{a+b+1})$ as results from the change of variables $t = \delta\tau$. Consequently, one finds

$$I(\gamma_T) + I(\gamma_{T'}) = O\left(\delta^{a+b+1}\right). \tag{49}$$

Putting together all the estimates of (46), (47), (49) yields the desired result. □

Remark. The proof technique of Proposition 9 tolerates the presence of logarithmic factors, in which case it suffices to develop the corresponding estimates for the basic integral (48). We find in this way, when $a+b+1<0$, the estimate

$$O\left(|1-z|^a |L^k(z)|\right) \odot O\left(|1-z|^b |L^\ell(z)|\right) = O\left(|1-z|^{a+b+1} |L^{k+\ell}(z)|\right).$$

The contour γ used in the proof is also susceptible to many variations. For instance, one may deform it slightly to include a "hook" near $w = 1$, in which case the modified contour may be used to estimate more finely the singular behaviour of Hadamard products.

We can then extend the asymptotic range covered by Proposition 9 as follows.

Proposition 10. *Assume that $f(z)$ and $g(z)$ are Δ-regular and that for $z \in \Delta$,*

$$f(z) = O(|1-z|^a) \quad \text{and} \quad g(z) = O(|1-z|^b).$$

(i) If $k < a+b+1 < k+1$ for some integer $-1 \le k < \infty$, then for $z \in \Delta'$:

$$(f \odot g)(z) = \sum_{j=0}^{k} \frac{(-1)^j}{j!} (f \odot g)^{(j)}(1)(1-z)^j + O(|1-z|^{a+b+1}).$$

(ii) If $a+b+1$ is a nonnegative integer then for $z \in \Delta'$:

$$(f \odot g)(z) = \sum_{j=0}^{a+b} \frac{(-1)^j}{j!} (f \odot g)^{(j)}(1)(1-z)^j + O(|1-z|^{a+b+1} |L(z)|).$$

Proof. Let $\partial = \partial_z$ denote the operator $\frac{d}{dz}$ and let ϑ denote the Euler operator $z\partial$. Observe that

$$\vartheta(f \odot g) = (\vartheta f) \odot g = f \odot (\vartheta g),$$

which yields

$$\vartheta^{k+1}(f \odot g) = (\vartheta^{k+1} f) \odot g.$$

The differentiation properties of Theorem 6 imply [with $k := a+b+1$ in Case (ii)] that $\vartheta^{k+1} f(z)$ is $O(|1-z|^{a-k-1})$. Thus, Proposition 9 applies, to the effect that

$$\left(\vartheta^{k+1}(f \odot g)\right)(z) = O\left(|1-z|^{a+b-k}\right).$$

On the other hand, the operator ϑ^{-1} is (for h in the image of ϑ)

$$\left(\vartheta^{-1} h\right)(z) := P_0 + \int_0^z h(t)\,\frac{dt}{t},$$

for some integration constant P_0. It is then possible to recover $h = f \odot g$ through successive integrations, by making use of Theorem 7.

Case *(i)*. By definition of k, one has $-1 < a + b - k < 0$. Repeated integrations then show that

$$(f \odot g)(z) = P(z) + O\left(|1-z|^{a+b+1}\right), \tag{50}$$

for some polynomial $P(z)$ of degree k that encapsulates the sequence of integration constants. Equation (50) yields qualitatively the form of the statement. The polynomial $P(z)$ is then automatically determined as the first $(k+1)$ terms of the Taylor expansion of $f \odot g$ at 1, which is precisely what our assertion expresses.

Case *(ii)*. In this case, the first integration step requires integrating a term $O(|1-z|^{-1})$, which leads to the logarithmic form of the statement. (See also the comments following Theorem 7.) □

4.3 Composition rules

At this stage, we can summarize the state of affairs regarding Hadamard products by the following general statement.

Theorem 11 (Hadamard composition of singularities). *Let $f(z)$ and $g(z)$ be two functions that are Δ–regular with expansions of the type* (24)*:*

$$f(z) = \sum_{m=0}^{M} c_m (1-z)^{\alpha_m} + O(|1-z|^A), \qquad g(z) = \sum_{n=0}^{N} d_n (1-z)^{\beta_n} + O(|1-z|^B).$$

Then, the Hadamard product $(f \odot g)(z)$ is also Δ–regular. Its singular expansion is computable by bilinearity, using the composition rules of Proposition 8 and the remarks thereafter, with error terms provided by Propositions 9 and 10:

$$(f \odot g)(z) = \sum_{m,n} c_m d_n [(1-z)^{\alpha_m} \odot (1-z)^{\beta_n}] + P(1-z) + O\left(|1-z|^C\right),$$

where $C := 1 + \min(\alpha_0 + B, A + \beta_0)$ and P is a polynomial of degree less than C.

The polynomial P is accessible via the Taylor expansion of $h - h_{\rm sing}$, where $h_{\rm sing}$ represents the sum of all the elements in the asymptotic expansion of $h := f \odot g$ at $z = 1$ that are singular. This theorem then validates the following algorithm, which is often helpful in computations done by hand when composing functions under Hadamard products.

> **"Zigzag" Algorithm.** [Computes the singular expansion of $f \odot g$ up to $O(|1-z|^C)$.]
>
> **1.** Use singularity analysis to determine separately the asymptotic expansions $\mathrm{Asympt}(f_n)$, $\mathrm{Asympt}(g_n)$ of $f_n = [z^n]f(z)$ and $g_n = [z^n]g(z)$ into descending powers of n.
> **2.** Perform the resulting product and reorganize as $\mathrm{Asympt}(f_n g_n)$.
> **3.** Choose a basis $\mathcal{B}$ of singular functions, for instance, the standard basis $\mathcal{B} = \{(1-z)^\alpha L(z)^k\}$, or the polylogarithm basis, $\mathcal{B} = \{\mathrm{Li}_{\beta,k}(z)\}$. Construct a function $H(z)$ expressed in terms of $\mathcal{B}$ whose singular behaviour is such that the asymptotic form of its coefficients, $\mathrm{Asympt}(H_n)$, is compatible with $\mathrm{Asympt}(f_n g_n)$ up to the needed error terms.
> **4.** Output the singular expansion of $f \odot g$ as the quantity $H(z) + P(z) + O(|1-z|^C)$, where P is a polynomial in $(1-z)$ of degree less than C.

The reason for the addition of a polynomial in Step **4**, is that integral powers of $(1-z)$ do not leave a trace in coefficient asymptotics since their contribution is asymptotically null. (An example of such "hidden" analytic terms already appears in the composition rule for powers given in Proposition 8.) The Zigzag Algorithm is then principally useful for determining the divergent part of expansions. If needed, the coefficients in the polynomial P can be expressed as values of the function $f \odot g$ and its derivatives at 1 once it has been stripped of its nondifferentiable terms. (This is analogous to the situation prevailing in Proposition 10.)

Example 12. *The return of Pólya's drunkard.* In the d–dimensional lattice $\mathbb{Z}^d$ of points with integer coordinates, the drunkard performs a random walk starting from the origin with steps in $\{-1,+1\}^d$, each taken with equal likelihood. The probability that the drunkard is back at the origin after $2n$ steps is

$$q_n^{(d)} = \left(\frac{1}{2^{2n}}\binom{2n}{n}\right)^d, \qquad (51)$$

since the walk is a product d independent 1–dimensional walks. The probability that $2n$ is the epoch of the *first* return to the origin is the quantity $p_n^{(d)}$, which is determined implicitly by

$$\left(1 - \sum_{n=1}^{\infty} p_n^{(d)} z^n\right)^{-1} = \sum_{n=0}^{\infty} q_n^{(d)} z^n, \qquad (52)$$

as results from the convolution equations expressing the decomposition of loops into primitive loops. In terms of the associated ordinary generating functions P and Q, this relation thus reads as $(1-P(z))^{-1} = Q(z)$.

The asymptotic analysis of the q_n's is straightforward; the one of the p_n's is more involved and is of interest in connection with recurrence and transience of the random walk; see, e.g., [15, 44]. The Hadamard closure theorem provides a direct access to this problem. Define

$$\lambda(z) := \sum_{n\geq 0} \frac{1}{2^{2n}}\binom{2n}{n} z^n \equiv \frac{1}{\sqrt{1-z}}.$$

Then, Equations (51) and (52) imply:

$$P(z) = 1 - \frac{1}{\lambda(z)^{\odot d}}, \qquad \text{where} \quad \lambda(z)^{\odot d} := \lambda(z) \odot \cdots \odot \lambda(z) \ (d \text{ times}).$$

The singularities of $P(z)$ are found to be as follows.

$\underline{d=1}$: No Hadamard product is involved and

$$P(z) = 1 - \sqrt{1-z}, \qquad \text{implying} \quad p_n^{(1)} = \frac{1}{n 2^{2n-1}}\binom{2n-2}{n-1} \sim \frac{1}{2\sqrt{\pi n^3}}.$$

(This agrees with the classical combinatorial solution expressed in terms of Catalan numbers.)

$\underline{d=2}$: By the Hadamard closure theorem, the function $Q(z) = \lambda(z) \odot \lambda(z)$ admits a priori a singular expansion at $z = 1$ that is composed solely of elements of the form $(1-z)^\alpha$ possibly multiplied by integral powers of the logarithmic function $L(z)$. From a computational standpoint (cf. the Zigzag Algorithm), it is then best to start from the coefficients themselves,

$$q_n^{(2)} \sim \left(\frac{1}{\sqrt{\pi n}} - \frac{1}{8\sqrt{\pi n^3}} + \cdots\right)^2 \sim \frac{1}{\pi}\left(\frac{1}{n} - \frac{1}{4n^2} + \cdots\right),$$

and reconstruct the only singular expansion that is compatible, namely

$$Q(z) = \frac{1}{\pi} L(z) + K + O((1-z)^{1-\epsilon}),$$

where $\epsilon > 0$ is an arbitrarily small constant and K is fully determined as the limit as $z \to 1$ of $Q(z) - \pi^{-1}L(z)$. Then it can be seen that the function P is Δ–continuable. (Proof: Otherwise, there would be complex poles arising from zeros of the function Q on the unit disc, and this would entail in $p_n^{(2)}$ the presence of terms oscillating around 0, a fact that contradicts the necessary positivity of probabilities.) The singular expansion of $P(z)$ at $z = 1$ results immediately from that of $Q(z)$:

$$P(z) \sim 1 - \frac{\pi}{L(z)} + \frac{\pi^2 K}{L^2(z)} + \cdots .$$

so that, by the extension of Theorem 2 to arbitrary powers of logarithms as given in [24, 51], one has

$$\begin{aligned}
p_n^{(2)} &= \frac{\pi}{n\log^2 n} - 2\pi\frac{\gamma+\pi K}{n\log^3 n} + O\left(\frac{1}{n\log^4 n}\right) \\
K &= 1 + \sum_{n=1}^{\infty}\left(16^{-n}\binom{2n}{n}^2 - \frac{1}{\pi n}\right) \\
&\doteq 0.88254240061060637358582 57.
\end{aligned}$$

(See the study by Louchard *et al.* [46, Sec. 4] for somewhat similar calculations.)

$d = 3$: This case is easy since $Q(z)$ remains finite at its singularity $z = 1$ where it admits an expansion in powers of $(1-z)^{1/2}$, to the effect that

$$q_n^{(3)} \sim \left(\frac{1}{\sqrt{\pi n}} - \frac{1}{8\sqrt{\pi n^3}} + \cdots\right)^3 \sim \frac{1}{\pi^{3/2}}\left(\frac{1}{n^{3/2}} - \frac{3}{8n^{5/2}} + \cdots\right).$$

The function $Q(z)$ is a priori Δ-continuable and its singular expansion can be reconstructed from the form of coefficients:

$$Q(z) \underset{z\to 1}{\sim} Q(1) - \frac{2}{\pi}\sqrt{1-z} + O(|1-z|),$$

leading to

$$P(z) = \left(1 - \frac{1}{Q(1)}\right) - \frac{2}{\pi Q^2(1)}\sqrt{1-z} + O(|1-z|).$$

By singularity analysis, the last expansion gives

$$\begin{aligned}
p_n^{(3)} &= \frac{1}{\pi^{3/2}Q^2(1)}\frac{1}{n^{3/2}} + O\left(\frac{1}{n^2}\right) \\
Q(1) &= \frac{\pi}{\Gamma\left(\frac{3}{4}\right)^4} \doteq 1.3932039296856768591842463.
\end{aligned}$$

A complete asymptotic expansion in powers $n^{-3/2}, n^{-5/2}, \dots$ can be obtained by the same devices. In particular this improves the error term above to $O(n^{-5/2})$. The explicit form of $Q(1)$ results from its expression as the generalized hypergeometric ${}_3F_2[\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; 1, 1; 1]$, which evaluates by Clausen's theorem and Kummer's identity to the square of a complete elliptic integral. (See the papers by Larry Glasser for context, for instance [29]; nowadays, Maple and Mathematica even provide this value automatically).

Higher dimensions are treated similarly, with logarithmic terms surfacing in asymptotic expansions for all even dimensions. □

We observe that, without the developments of the present paper, the precise asymptotic structure of such sequences is not obvious. Methods of the last example may be used to provide a rigorous setting to certain asymptotic enumeration results stated by physicists, where back-and-forth equivalences between singular expansions of functions and asymptotic expansions of coefficients are often used without much justification. See for instance the works of Guttmann and collaborators [5, 32] and Chyzak's numeric-symbolic study [8] relative to special self-avoiding polygons.

5 Applications: first moments

Thanks to the extended singularity analysis toolkit, we are now in a position to analyze the tree recurrences that were introduced in Section 2. For each of the three models, two types of tolls are to be considered:

$$t_n^\alpha := n^\alpha \text{ (with } \alpha > 0), \qquad t_n^{\log} = \log n,$$

and we assume in both cases that $t_0 = 0$. The corresponding ordinary generating functions are the polylogarithm $\mathrm{Li}_\alpha \equiv \mathrm{Li}_{\alpha,0}$ and the specific $\mathrm{Li}_{0,1}$, whose singular expansions have been already recalled as Theorem 4. In each case, a linear transform $\mathcal{L}$ relates the generating function of costs, $f(z)$, to a generating function of tolls, either $t(z)$ (normalized) or $\tau(z)$ ("raw"). Theorems on composition of singularities make it possible to follow step by step the elementary operations of which $\mathcal{L}$ is composed and determine the effect of the $\mathcal{L}$ transform on singularities in a systematic manner. Given that computations are "automatic", we will mostly focus our discussion on main terms and on the global shape of singular expansions, leaving some of the details as exercises to the reader—or better, to a computer algebra engine.

The net outcome in each of the three tree models under consideration is the following: for large tolls, the cost is driven by the toll itself; for small tolls, the cost is of linear growth and, in a sense, "freely" caused by the recursion itself, that is, driven by the cumulation of costs due to small subtrees; in between, there is a threshold value of the toll where a "resonance" takes place between the toll and the recursion, leading to the emergence of a logarithmic factor. Such facts parallel what is familiar in the context of inhomogeneous linear differential equations, where either the free regime or the forced regime dominates, with logarithmic terms being created precisely by resonances.

5.1 The binary search tree recurrence

For the binary search tree model, there is an integral transform $\mathcal{L}$ that relates the ordinary generating function of tolls, $t(z)$, and the ordinary generating function of the induced costs, $f(z)$: it is given by (11) according to which $f(z) = \mathcal{L}[t(z)]$, where (with $t_0 = f_0 = 0$)

$$\mathcal{L}[t(z)] = \frac{1}{(1-z)^2} \int_0^z t'(w)(1-w)^2 \, dw. \tag{53}$$

Consequently, the computation is entirely *mechanical*[8] and it only needs the theorems relating to integration, differentiation, and polylogarithms (Theorems 6 and 7) in conjunction with basic singularity analysis (Theorem 2). Our

[8] In the MAPLE system for symbolic computations about two dozen instructions suffice to implement calculations, once use is made of Bruno Salvy's package equivalent dedicated to the asymptotic analysis of coefficients of generating functions [56]. It suffices to program the polylogarithm expansions (Theorem 4), use the system capabilities for series expansions, differentiation, and integration (Theorems 6 and 7), and conclude by an appeal to Salvy's program that implements the basic transfers of Theorem 2.

derivation below constitutes an alternative to parallel results by Neininger [50], Chern *et al.* [6, 7], and Fill and Kapur [19], who employ elementary but perhaps less transparent methods (typically, the approximation of discrete sums by integrals).

Theorem 13. *Under the binary search tree model, the expected values of the costs induced by tolls of type t_n^α ($\alpha > 0$) and $t_n^{\log}$ admit full asymptotic expansions in descending powers of n and integral powers of $\log n$. The main terms are summarized by the following table:*

Toll (t_n)		Cost (f_n)	
n^α	$(2 < \alpha)$	$\frac{\alpha+1}{\alpha-1}n^\alpha$	$+O(n^{\alpha-1})$
n^2		$3n^2$	$-6n\log n + (10-6\gamma)n + O(\log n)$
n^α	$(1 < \alpha < 2)$	$\frac{\alpha+1}{\alpha-1}n^\alpha$	$+K_\alpha n + O(n^{\alpha-1})$
n		$2n\log n$	$+2(\gamma-1)n + 2\log n + 2\gamma + 1 + O\left(\frac{1}{n}\right)$
n^α	$(0 < \alpha < 1)$	$K_\alpha n$	$+\frac{\alpha+1}{\alpha-1}n^\alpha + K_\alpha + o(1)$
$\log n$		$K_0' n$	$-\log n + (K_0' - 2) - \frac{1}{2n} + \frac{1}{9n^2} + O\left(\frac{1}{n^3}\right)$.

Proof. For the case α a nonnegative integer, the integration can be carried out in finite terms since the generating function of tolls is rational. For instance, the case $\alpha = 1$ corresponds to the well-known analysis of Quicksort and binary search tree algorithms [40, 47, 58, 62].

For t_n^α, it suffices to examine the effect of the $\mathcal{L}$ transform on singular elements of the form $c(1-z)^\beta$; e.g., for the main term corresponding to $t_n = n^\alpha$, we should take $\beta = -\alpha - 1$. The $\mathcal{L}$ transformation reads as a succession of operations, "differentiate, multiply by $(1-z)^2$, integrate, multiply by $(1-z)^{-2}$"—all are covered by our previous theorems. The chain on any particular singular element starts as

$$c(1-z)^\beta \xrightarrow{\partial} c\beta(1-z)^{\beta-1} \xrightarrow{\times(1-z)^2} c\beta(1-z)^{\beta+1}.$$

At this stage, integration intervenes. Assume that $\beta + 1 \neq -1$. (Otherwise, a logarithm appears.) According to Theorem 7, and ignoring integration constants for the moment, integration gives

$$c\beta(1-z)^{\beta+1} \xrightarrow{\int} -c\frac{\beta}{\beta+2}(1-z)^{\beta+2} \xrightarrow{\times(1-z)^{-2}} -c\frac{\beta}{\beta+2}(1-z)^\beta.$$

Then this singular element corresponds to a contribution

$$-c\frac{\beta}{\beta+2}\binom{n-\beta-1}{-\beta-1},$$

which is of order $O(n^{-\beta-1})$. (The treatment of logarithmic terms is entirely similar.)

The derivation above has left aside the determination of the integration constants. These are given by the second case of Theorem 7, which provides in particular access to the constants K_α and K'_0. The constant term in the asymptotic expansion of the integral is of the form

$$\mathbf{K}[t] := \int_0^1 \left[t'(w)(1-w)^2 - \left(t'(w)(1-w)^2\right)_-\right] dw,$$

where f_- represents the sum of the singular terms in f having exponent < -1, as in the proof of Theorem 7. In the singular expansion of $f(z)$, this integration constant gets further multiplied by $(1-z)^{-2}$; the resulting linear term in the asymptotic expansion of f_n is then plainly

$$\mathbf{K}[t] \cdot (n+1).$$

In particular, if the growth of t_n is smaller than n, then, the divergence part is absent and $\mathbf{K}[t]$ reduces to

$$\mathbf{K}[t] = \int_0^1 t'(w)(1-w)^2\, dw = 2\sum_{n=1}^{\infty} \frac{t_n}{(n+1)(n+2)},$$

as follows from expanding the integrand around 0 and integrating the resulting series. This yields the following values for $\alpha < 1$:

$$K_\alpha = 2\sum_{n=1}^{\infty} \frac{n^\alpha}{(n+1)(n+2)}, \qquad K'_0 = 2\sum_{n=1}^{\infty} \frac{\log n}{(n+1)(n+2)}, \tag{54}$$

while for $1 < \alpha < 2$,

$$K_\alpha = 2\sum_{n=1}^{\infty} \frac{n^\alpha - \Gamma(\alpha+1)\binom{n+\alpha}{\alpha}}{(n+1)(n+2)}.$$

The theorem is finally established. □

Remark. The slowly convergent series expressions of K_α, K'_0 can be rephrased as definite integrals, thanks to Mellin transform techniques. The starting point is the easy formal identity,

$$\sum_{n\geq 1} c_n n^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty \left(\sum_{n\geq 1} c_n e^{-nx}\right) x^{s-1}\, dx. \tag{55}$$

The constant K_α with $\alpha < 1$ corresponds to $s = 1-\alpha$ and $c_n = n/[(n+1)(n+2)]$, for which the integrand admits of closed form since

$$\sum_{n=1}^{\infty} \frac{nz^n}{(n+1)(n+2)} = \frac{1}{z^2}\left[(2-z)L(z) - 2z\right].$$

From there, the constant K_0' is attained as $\frac{d}{d\alpha}K_\alpha\big|_{\alpha=0}$. A final change of variables $x = -\log t$ then yields an integral representation for "Fill's first logarithmic constant" (γ is Euler's constant):

$$\begin{aligned} K_0' &= -\gamma - 2\int_0^1 [(t-2)\log(1-t) - 2t]\left(\log\log\frac{1}{t}\right)\frac{dt}{t^3} \qquad (56) \\ &= 1.20356491674961033428628333814873131775552838577096. \end{aligned}$$

The last estimate to 50D improves on the earlier 3D evaluation of Fill [16]. The cost induced by $t^{\log}$ is of particular interest as it is precisely the entropy of the distribution of binary search trees; see the account and first estimates in the book by Cover and Thomas [11, p. 72–74], as well as pointers to self-organizing search in Fill's article [16]. In his doctoral dissertation [38, Section 5.1], Kapur has extended the methods and estimates to m-ary search trees.

5.2 The uniform binary tree recurrence

This section examines the uniform binary tree model that surfaces recurrently in combinatorics. Here, we put on a firm basis a classification of the expected costs corresponding the tolls t_n^α and $t_n^{\log}$ which was outlined (with several typographical errors) in an article by Flajolet and Steyaert [28]. The particular case of the toll $t_n = n$ has, like for binary search trees, a dignified history as it corresponds to path length in Catalan trees and to area under Dyck paths, whose first distributional analyses go back to Louchard and Takács [45, 61].

Our starting point is (19) according to which the generating function of costs $f(z) = \sum C_n f_n z^n$ normalized by the Catalan numbers C_n and the ordinary generating function of costs $\tau(z) = \sum t_n z^n$ are related by $f(z) = \overline{\mathcal{L}}[\tau(z)]$, where

$$\overline{\mathcal{L}}[\tau(z)] = \frac{1}{\sqrt{1-4z}}\left(\tau(z)\odot C(z)\right), \qquad (57)$$

with

$$C(z) = \sum_{n\geq 0} C_n z^n = \frac{1}{2z}\left(1-\sqrt{1-4z}\right), \qquad C_n = \frac{1}{n+1}\binom{2n}{n}.$$

We state:

Theorem 14. *Under the uniform binary tree model, the expected values of the costs induced by tolls of type t_n^α ($\alpha > 0$) and $t_n^{\log}$ admit full asymptotic expansions in descending powers of n and integral powers of $\log n$. The main* *terms are summarized by the following table:*

Toll (t_n)		Cost (f_n)	
n^α	$(\frac{3}{2} < \alpha)$	$\frac{\Gamma(\alpha-\frac{1}{2})}{\Gamma(\alpha)} n^{\alpha+\frac{1}{2}}$	$+O(n^{\alpha-\frac{1}{2}})$
$n^{3/2}$		$\frac{1}{\Gamma(3/2)} n^2$	$+O(n\log n)$
n^α	$(\frac{1}{2} < \alpha < \frac{3}{2})$	$\frac{\Gamma(\alpha-\frac{1}{2})}{\Gamma(\alpha)} n^{\alpha+\frac{1}{2}}$	$+O(n)$
$n^{1/2}$		$\frac{1}{\sqrt{\pi}} n\log n$	$+O(n)$
n^α	$(0 < \alpha < \frac{1}{2})$	$\overline{K}_\alpha n$	$+O(1)$
$\log n$		$\overline{K}'_0 n$	$-2\sqrt{\pi}n^{1/2}+O(1)$.

Proof. For the tolls t_n^α, all that is required is to determine the singular expansion of

$$\tau(z)\odot C(\frac{z}{4}) = \sum_{n=1}^{\infty} \frac{n^\alpha}{(n+1)}\binom{2n}{n}\left(\frac{z}{4}\right)^n.$$

(For convenience, the singularity has been scaled to 1.) We use the Zigzag Algorithm presented in Subsection 4.3. The known asymptotic expansion of the Catalan numbers is

$$4^{-n}C_n \sim \frac{1}{\sqrt{\pi}}n^{-3/2}\left(1-\frac{9}{8n}+\frac{145}{128n^2}-\cdots\right).$$

Multiply this by n^α to get the expansion of $n^\alpha C_n 4^{-n}$. The terms now involve the scale $\{n^{\alpha-\frac{3}{2}}, n^{\alpha-\frac{5}{2}}, \ldots\}$. Assume that α is not a half-integer [i.e., $\alpha \notin (\frac{1}{2}\mathbb{Z})\setminus\mathbb{Z}$]; see below for the contrary case. Then the basis of functions $\mathcal{B} = \{(1-z)^{-\alpha+k+\frac{1}{2}}\}$, where k ranges over the integers, has the property that the coefficients of its generic element are $O(n^{\alpha-k-\frac{3}{2}})$; in particular,

$$[z^n](1-z)^{-\alpha+\frac{1}{2}} \sim \frac{n^{\alpha-\frac{3}{2}}}{\Gamma(\alpha-\frac{1}{2})}\left(1+\frac{(2\alpha-1)(2\alpha-3)}{8n}+\cdots\right).$$

We can thus find a singular function $H(z)$ whose coefficients match asymptotically those of $\tau(z)\odot C(z/4)$, which is of the form

$$H(z) = \frac{\Gamma(\alpha-\frac{1}{2})}{\sqrt{\pi}}(1-z)^{-\alpha+\frac{1}{2}}\left(1+c_1(1-z)+c_2(1-z)^2+\cdots\right),$$

for some effectively computable sequence (c_j). The singular expansion of $\tau(z)\odot C(z/4)$ is then the sum of the expansion of H above and of a power series in $(1-z)$, call it $P(z)$, that can be determined according to the principles of Section 4.

The singular expansion of $f(z/4)$ is that of $H(z)+P(z)$ divided by $\sqrt{1-z}$, so that, by transfer, we get

$$[z^n]f\left(\frac{z}{4}\right) \sim \frac{\Gamma(\alpha-\frac{1}{2})}{\sqrt{\pi}\Gamma(\alpha)}n^{\alpha-1}\left(1+\frac{c_1'}{n}+\frac{c_2'}{n^2}+\cdots\right)+[z^n]\frac{P(z)}{\sqrt{1-z}},$$

for some sequence (c_j'), where the "hidden" analytic part $P(z)$ arises from the "hidden" analytic component in $\tau(z)\odot C(z/4)$. After dividing by $C_n 4^{-n}$, one finds finally:

$$f_n \sim \frac{\Gamma(\alpha-\frac{1}{2})}{\Gamma(\alpha)}n^{\alpha+\frac{1}{2}}\left(1+\frac{c_1''}{n}+\cdots\right)+R_n, \tag{58}$$

where the "hidden" remainder term R_n is of the form

$$R_n \sim d_{-1}n + d_0 + \frac{d_1}{n} + \cdots .$$

This last estimate provides all the entries in the table above, whenever α is not a half-integer, as it suffices to merge the two expansions of (58). In addition, when $0<\alpha<\frac{1}{2}$, the series defining $t(z/4)$ converges at the singularity 1. Thus, the dominant asymptotic term of $f(z/4)$ is $t(1/4)/\sqrt{1-z}$, that is,

$$f\left(\frac{z}{4}\right) \sim \frac{\overline{K}_\alpha}{\sqrt{1-z}}, \qquad \overline{K}_\alpha := \sum_{n=1}^{\infty}\frac{n^\alpha}{n+1}\frac{1}{4^n}\binom{2n}{n}.$$

When α is a half-integer, logarithmic terms appear due to the presence of inverse integral powers of n in the coefficients of $t(z/4)$, but the derivation is otherwise similar. For instance at $\alpha=\frac{1}{2}$, one has

$$4^{-n}\sqrt{n}C_n \sim \frac{1}{\sqrt{\pi}}\frac{1}{n}+O\left(\frac{1}{n^2}\right),$$

which shows that

$$t(z/4) = H(z)+P_0+O((1-z)^{1-\epsilon}), \qquad H(z)=\frac{1}{\sqrt{\pi}}L(z),$$

resulting in the stated estimate.

Finally, when $t_n=\log n$, we have $\tau(z)=\mathrm{Li}_{0,1}(z)=O(|1-z|^{-1-\epsilon})$ for any $\epsilon>0$. Thus, by Proposition 10(i),

$$(\tau\odot C)(z/4)=\overline{K}_0'+O\left(|1-z|^{\frac{1}{2}-\epsilon}\right), \qquad \overline{K}_0' := \sum_{n=1}^{\infty}(\log n)\frac{C_n}{4^n}.$$

Singularity analysis and the estimate for C_n yield $f_n=\overline{K}_0'n+O\left(n^{\frac{1}{2}+\epsilon}\right)$. Carrying higher-order terms, we get the mean of the shape functional,

$$\mu_n = \overline{K}_0'n - 2\sqrt{\pi}n^{1/2}+O(1), \tag{59}$$

which agrees with the estimate in Theorem 3.1 of [16] and improves the remainder estimate. □

The Mellin technique of (55) is once more applicable to the determination of "Fill's second logarithmic constant" $\overline{K}_0'$. It provides the value:

$$\begin{aligned}\overline{K}_0' &:= \sum_{k\geq 1} \frac{\log k}{(k+1)4^k}\binom{2k}{k}. \\ &= -\gamma - \int_0^1 \frac{1}{\sqrt{1-t}(1+\sqrt{1-t})^2}\left(\log\log\frac{1}{t}\right)\,dt, \\ &= 2.0254384677765738877135187391417652470652930617658.\end{aligned}$$

The subject of costs on binary trees is considered in greater depth in [18] by applying the techniques developed in this paper. There, some higher-order estimates, asymptotics for moments of each order, and limiting distributions are derived when the toll sequence is either n^α or $\log n$.

Our methods can also be used to treat more generally the case of all simple families of trees in the sense of Meir and Moon [49], of which Catalan trees are a special case. This generalization is the subject of ongoing work.

5.3 The union–find tree recurrence

In this subsection, we revisit the Knuth–Pittel–Schönhage recurrence corresponding to the destruction of free labelled trees and dually to the management of equivalence relations [42, 43]. The main result of this section is essentially a rephrasing of the main results of Knuth and Pittel in [42], to which we add the possibility of determining complete asymptotic expansions. Like before, the starting point is the integral transform (23) (adjusted for the fact that $t_1^\alpha = f_1^\alpha = 1 \neq 0$), which relates the ordinary generating function of tolls $\tau(z)$ to the normalized generating function of costs $f(z)$ via $f(z) = \mathcal{L}[\tau(z)]$, where

$$\mathcal{L}[\tau(z)] = t_1 z T'(z) + \frac{1}{2}\frac{T(z)}{1-T(z)}\int_0^z \partial_w\left(\tau(w)\odot T^2(w)\right)\frac{dw}{T(w)}. \tag{60}$$

There $T(z)$ is the Cayley tree function whose singular expansion at the (unique) dominant singularity $z = e^{-1}$ is well known: one has the shape

$$T(z) \sim 1 - \sqrt{2}(1-ez)^{1/2} + c_1(1-ez) + \cdots \tag{61}$$

as $z \to e^{-1}$ in any sector of angle $< 2\pi$; see also [42, Eq. (3.16)]. (The paper by Corless *et al.* [9] is a definitive reference regarding the tree function.) As noted earlier, the case of union–find tree recurrences combines all the composition results developed in this paper.

Theorem 15. *Under the union–find tree recurrence model, the expected values of the costs induced by tolls of type t_n^α ($\alpha > 0$) and $t_n^{\log}$ admit full asymptotic expansions in descending powers of n and integral powers of $\log n$. The main* *terms are summarized by the following table:*

Toll (t_n)		Cost (f_n)	
n^α	$(\frac{3}{2} < \alpha)$	$\dfrac{\Gamma(\alpha - \frac{1}{2})}{\sqrt{2}\Gamma(\alpha)} n^{\alpha+\frac{1}{2}}$	$+O(n^{\alpha-\frac{1}{2}})$
$n^{3/2}$		$\dfrac{1}{\sqrt{2}\Gamma(3/2)} n^2$	$+O(n \log n)$
n^α	$(\frac{1}{2} < \alpha < \frac{3}{2})$	$\dfrac{\Gamma(\alpha - \frac{1}{2})}{\sqrt{2}\Gamma(\alpha)} n^{\alpha+\frac{1}{2}}$	$+O(n)$
$n^{1/2}$		$\dfrac{1}{\sqrt{2\pi}} n \log n$	$+O(n)$
n^α	$(0 < \alpha < \frac{1}{2})$	$(1 + \frac{1}{2}\widehat{K}_\alpha)n$	$+O(n^{\alpha+\frac{1}{2}})$
$\log n$		$\frac{1}{2}\widehat{K}'_0 n$	$+O(\sqrt{n})$.

Proof. We shall content ourselves with indicating the way full asymptotic expansions can be determined within the generating function framework. (Detailed computations are left as an exercise for the reader.) In what follows, we set $Z = (1-z)$ and let $\mathcal{A}$ denote an unspecified entire series in powers of Z, not necessarily the same at each occurrence. For instance, one may summarize diversely the expansion (61) of $T(z/e)$ as

$$T(z/e) \sim 1 - \sqrt{2}Z^{1/2} + Z\mathcal{A} + Z^{3/2}A \sim \mathcal{A} + \mathcal{A}Z^{1/2},$$

and so on. We shall also let $\mathcal{N}$ denote generically a series in descending powers of $1/n$.

We consider first the case of the toll t_n^α and assume for simplicity that α is not a half-integer: $\alpha \notin (\frac{1}{2}\mathbb{Z}) \setminus \mathbb{Z}$. The polylogarithm expansions grant us a priori that the generating function $\tau(z)$ lies in the class of functions amenable to singularity analysis, with

$$\tau(z) \sim Z^{-\alpha-1}\mathcal{A} + \mathcal{A}.$$

Therefore, the Hadamard product $(\tau(z) \odot T^2(z/e))$ is also amenable. The coefficients of the latter function are of the form $n^{\alpha-\frac{3}{2}}\mathcal{N}$, as follows from the fact that $[z^n]\tau(z) = n^\alpha$ and $[z^n]T^2(z/e) \sim n^{-3/2}\mathcal{N}$ (by the singular expansion of T^2). Thus, converting back this information to the function, we find

$$\tau(z) \odot T^2(z/e) \sim Z^{-\alpha+\frac{1}{2}}\mathcal{A} + \mathcal{A}, \qquad \partial_z(\tau(z) \odot T^2(z/e)) \sim Z^{-\alpha-\frac{1}{2}}\mathcal{A} + \mathcal{A}.$$

What we have done here is to apply the Zigzag Algorithm of Section 4 and the differentiation theorem. Then multiplication by $1/T(z/e) \sim \mathcal{A} + Z^{1/2}\mathcal{A}$ shows that

$$\frac{1}{T(z/e)}\partial_z[\tau(z) \odot T^2(z/e)] \sim Z^{-\alpha-\frac{1}{2}}\mathcal{A} + Z^{-\alpha}\mathcal{A} + \mathcal{A} + Z^{\frac{1}{2}}\mathcal{A}.$$

Integration of this last expansion corresponds to increasing all exponents by 1. Finally one should multiply by $T(z/e)(1-T(z/e))^{-1}$ which is of type $Z^{-1/2}\mathcal{A}+\mathcal{A}$. This completes our handling of the second term on the right in (60). Also,

$$\frac{z}{e}T'(z/e) \sim Z^{-\frac{1}{2}}\mathcal{A}+\mathcal{A}.$$

The end result is then

$$f(z/e) \sim Z^{-\alpha}\mathcal{A}+Z^{-\alpha+\frac{1}{2}}\mathcal{A}+Z^{-\frac{1}{2}}\mathcal{A}+\mathcal{A}.$$

The dominant term is $Z^{-\alpha}$ when $\alpha > \frac{1}{2}$ whereas it is $Z^{-1/2}$ when $\alpha < \frac{1}{2}$.

At the same time, it is a simple task to trace the coefficients of main terms. For $\alpha > \frac{1}{2}$, the main term of $f(z/e)$ is $Z^{-\alpha}$, and one finds successively

$$\begin{aligned} \tau(z)\odot T^2(z/e) &\sim \sqrt{\frac{2}{\pi}}\Gamma(\alpha-\tfrac{1}{2})(1-z)^{-\alpha+\frac{1}{2}}, \\ f(z/e) &\sim \frac{\Gamma(\alpha-\frac{1}{2})}{2\sqrt{\pi}}(1-z)^{-\alpha}, \end{aligned}$$

where the last equation implies, via singularity analysis, an estimate of expected costs:

$$f_n^{\alpha} \sim \frac{\Gamma(\alpha-\frac{1}{2})}{\sqrt{2}\Gamma(\alpha)}n^{\alpha+\frac{1}{2}}.$$

For $\alpha < \frac{1}{2}$, the main term is $Z^{-1/2}$ and its coefficient is seen to arise from both terms on the right in (60): we have

$$f^{\alpha}(z/e) \sim \frac{1}{\sqrt{2}}\left(1+\frac{\widehat{K}_{\alpha}}{2}\right)(1-z)^{-1/2},$$

where $\widehat{K}_{\alpha}=\widehat{K}[n^{\alpha}]$ and the functional $\widehat{K}$ is

$$\widehat{K}[t] := \int_0^{1/e} \partial_w(\tau(w)\odot T^2(w))\,\frac{dw}{T(w)}. \tag{62}$$

Error terms can be similarly traced: in the case of $f^{\alpha}(z/e)$, it is of type $Z^{-\alpha}$ if $0<\alpha<\frac{1}{2}$, of type $Z^{-1/2}$ if $\frac{1}{2}<\alpha<1$, and so on. The end results are summarized in the statement of the theorem.

For half-integer α, a logarithmic term appears. For instance, in the case $\alpha=\frac{1}{2}$, this fact is associated to the shape of the coefficients

$$[z^n](\tau(z)\odot T^2(z/e)) \sim \frac{-2\sqrt{2}}{\Gamma(-1/2)}\,\frac{1}{n}+\frac{1}{n^2}\mathcal{N},$$

resulting in a singular expansion with a logarithmic term:

$$\tau(z)\odot T^2(z/e) = -\sqrt{\frac{2}{\pi}}\log Z + c + Z\mathcal{A} + (Z\log Z)\mathcal{A},$$

for some c.

For the logarithmic toll [note that now $t_1 = 0$, so that the first term in (60) does not contribute] we have $\tau(z) = \mathrm{Li}_{0,1}(z)$, the integral in (60) is convergent and, in the same way as for the case $\alpha < 1/2$, we get

$$f(z) = \frac{\widehat{K}'_0}{2\sqrt{2}}(1-ez)^{-1/2} + O(|1-ez|^{-\epsilon}),$$

which implies

$$f_n = \frac{1}{2}\widehat{K}'_0 n + O(n^{\frac{1}{2}+\epsilon}),$$

with $\widehat{K}'_0 = \widehat{K}[\log n]$ and $\widehat{K}$ defined at (62). □

It is of interest to compare our approach to that of Knuth and Pittel [42]. These authors use what is fundamentally a "repertoire approach", based on the transforms of two types of tolls, the Dirac tolls δ_{mn} and another family related to "tree polynomials". Their methods do not clearly appear to be extendible to the extraction of sublinear terms in asymptotic expansions. At the same time, their developments require appreciably more involved and perhaps less transparent computations.

6 Perspectives

In this concluding section, we discuss at a fairly informal and abstract level applications of the extended singularity analysis toolkit developed in the present paper in two further directions: the determination of higher-order moments for our basic models, and the treatment of tree recurrences which are more complex than the ones present in our lead examples. (Some of our examples below may accordingly involve nonbinary tree models.)

6.1 Higher moments and limit distributions

Let us return to the general framework of Section 1. There, the random cost X_n is related to costs X_{K_n} and $\widetilde{X}_{n-a-K_n}$ by the fundamental recursion (2). Raising both members of (2) to some integral power s yields

$$X_n^s = X_{K_n}^s + \widetilde{X}_{n-a-K_n}^s + \sum_{\substack{s_1+s_2+s_3=s\\ s_2,s_3\neq s}} \binom{s}{s_1,\, s_2,\, s_3} t_n^{s_1} X_{K_n}^{s_2} \widetilde{X}_{n-a-K_n}^{s_3}, \tag{63}$$

where we have made use of the multinomial expansion and have isolated the two sth powers. Take expectations with respect to the model $\mathfrak{M}_n$ and set $\mu_n^{(s)} := \mathbb{E}(X_n^s)$. The recursion on sth moments becomes, thanks to independence of the X and $\widetilde{X}$ sequences on the right in (63),

$$\mu_n^{(s)} = \sum_k p_{n,k}\left(\mu_k^{(s)} + \mu_{n-a-k}^{(s)}\right) + r_n^{(s)}, \tag{64}$$

where

$$r_n^{(s)} := \sum_{\substack{s_1+s_2+s_3=s\\ s_2,s_3\neq s}} \binom{s}{s_1, s_2, s_3} t_n^{s_1} \sum_k p_{n,k}\mu_k^{(s_2)}\mu_{n-a-k}^{(s_3)}.$$

This calculation shows that the sequence of sth moments for any fixed s satisfies the same type of recurrence as the first moments, save for a more complicated toll $(r_n^{(s)})$ that involves moments of the smaller orders $0, 1, \ldots, s-1$. Define the normalized generating functions

$$\mu^{(s)}(z) := \sum_n \mu_n^{(s)}\omega_n z^n, \qquad r^{(s)}(z) := \sum_n r_n^{(s)}\omega_n z^n,$$

with the normalization sequence $\omega_n \equiv 1$ for binary search trees and $\omega_n = C_n$ for uniform binary trees. Then the relation (64) is solved in terms of generating functions by an $\mathcal{L}$-transform as

$$\mu^{(s)}(z) = \mathcal{L}\left(r^{(s)}\right), \quad r^{(s)}(z) = \sum_{\substack{s_1+s_2+s_3=s\\ s_2,s_3\neq s}} \binom{s}{s_1, s_2, s_3} \tau^{\odot s_1}(z)\odot Q(\mu^{(s_2)}(z), \mu^{(s_3)}(z)), \tag{65}$$

where Q is for the binary search tree model and Catalan model, respectively,

$$Q^{\mathrm{BST}}(a(z), b(z)) = \int_0^z a(t)b(t)\,dt, \qquad Q^{\mathrm{Cat}}(a(z), b(z)) = za(z)b(z). \tag{66}$$

The $\mathcal{L}$ transform is given in (53) and (57) for the respective cases; the case of the union–find tree model [where $\omega_n = n^{n-1}/n!$ is used for $\mu^{(s)}$ and $\omega_n' = n^{n-2}(n-1)/n!$ is used for $r^{(s)}$] is similar but more complicated—see (60) for $\mathcal{L}$, while

$$Q^{UF}(a(z), b(z)) = \frac{1}{2}a(z)b(z).$$

As seen in the previous section, these $\mathcal{L}$ transforms involve only integration, differentiation, and ordinary and Hadamard products—all are operations that preserve the character of being Δ-regular and admitting complete asymptotic expansions at the dominant singularity. We then have a general result:

Theorem 16. *For any of the binary search tree, uniform binary tree, or union–find model, and for any integer $s > 0$, the sth moment of the cost function associated to a toll $t_n^{\log}$ or t_n^{α} admits a complete descending expansion in powers of n (possibly with logarithmic terms).*

Proof. The proof is simply an induction on the order s of the moments. We establish by induction the stronger property that the generating functions $\mu^{(s)}(z)$ are Δ-regular and admit complete asymptotic expansions in powers of $(1-z)$, possibly with logarithmic terms, after rescaling the singularity to be at 1. The property is true for $s = 1$ by results of the previous section. If the property is assumed to be true through order $s-1$, then the tolls $r^{(s)}(z)$ are Δ-regular and admit of complete asymptotic expansions at their singularity: this results from closure theorems of Sections 3 and 4. Next, the $\mathcal{L}$-transform is applied and, again by closure theorems, the property of $r^{(s)}(z)$ is seen to extend to $\mu^{(s)}(z)$. Thus the singular structure of $\mu^{(s)}(z)$ is fully characterized. It then suffices to apply basic singularity analysis in order to recover the existence of full asymptotic expansion of the moments $\mu_n^{(s)} = \frac{1}{\omega_n}[z^n]\mu^{(s)}(z)$. □

The process of extracting moments one after the other has been nicknamed "moment pumping" in the article [25], where it was used to determine the shape of the moments of total displacement in linear hashing tables. It had been employed earlier by Louchard and Takács in order to characterize moments of path length in trees and of area under excursions [45, 61], in a way largely similar to what has been described here in more general terms. In favorable cases, a pattern regarding the asymptotic shape of moments may emerge. In such cases (possibly centering of the random variable is required), the limiting distribution of costs becomes accessible through its moments, thanks to the moment convergence theorem. Instances are found in the already cited papers [25, 45, 61]. Fill's paper [16] provides another example (although it is based on direct recurrence manipulations rather than generating functions) to the effect that the logarithmic toll $t_n^{\log}$ gives rise to asymptotically Gaussian costs under the binary search tree model. Yet other examples, often based on direct recurrence manipulations, are provided by the recent independent studies of Hwang and Neininger [37] and of Fill and Kapur [19]. Clearly, a "metatheorem" similar to Theorem 16 is possible for varieties of increasing trees in the sense of Bergeron-Flajolet-Salvy [2] (generalizing the BST model). For simply generated families of trees in the sense of Meir and Moon [49] (generalizing the Catalan model), asymptotics of moments as well as limiting distributions have been derived by Fill and Kapur [17] as part of a broader project joint with Svante Janson. The union–find tree model can be generalized to other families of trees, and the techniques of the present paper can again be applied; this is the subject of ongoing research by the authors.

6.2 Differential models

Many tree recurrences associated to comparison-based searching and multidimensional retrieval problems generalizing binary search trees, once translated into generating functions, lead to integral equations of the form

$$\mathbf{\Phi}[f](z) = t(z), \tag{67}$$

where $\mathbf{\Phi}$ is a linear integral operator involving coefficients in $\mathbb{C}(z)$, that is, rational function coefficients. Here, as in our lead examples, $f(z)$ is a generating function of expected costs and $t(z)$ is a toll generating function. By successive differentiations, this transforms into a linear differential equation of the form

$$\mathbf{\Delta}[f](z) = \widetilde{t}(z), \tag{68}$$

where $\widetilde{t}(z)$ is a modified toll generating function and is an elementary variant of $t(z)$. We shall let d denote the order of the differential equation (68).

The description above corresponds to the situation already encountered with the binary search tree recurrence, representing the easy case of a differential order equal to 1. Other known cases include the m–ary search tree studied by Mahmoud and Pittel (see the account in [47]) and others [6, 19], quicksort with median-of-sample partitioning and locally balanced trees [40, 58], quadtrees [23, 36] as well as multidimensional search trees also known as k-d-trees [26]. (A valuable survey of a class of problems leading to Euler equations appears in [7].) For instance, in the case of 2-dimensional quadtrees the operator is given in [36] as

$$\mathbf{\Phi}[f](z) = f(z) - 4\int_0^z \left[\int_0^y f(x)\,\frac{dx}{1-x}\right]\,\frac{dy}{y(1-y)},$$

which leads to a second order differential equation,

$$z(1-z)\partial_z^2 f(z) + (1-2z)\partial_z f(z) - \frac{4}{1-z}f(z) = \widetilde{t}(z),$$

where $\widetilde{t}(z) = \partial_z[z(1-z)t'(z)]$.

The variation-of-constants technique applies to equations of order greater than 1 as well as to linear systems. It may then be used to express f as a linear integral transform involving a set $\{h_j\}$ of solutions to the homogeneous equation $\mathbf{\Delta}h = 0$, as we know explain. Indeed, let the linear differential equation (68) be put into the form of a system

$$\partial_z \mathbf{y}(z) = \mathbf{A}\mathbf{y}(z) + \mathbf{b}(z), \tag{69}$$

where $\mathbf{y}$ is the d–dimensional vector $\mathbf{y} = (f, f', \ldots, f^{(d-1)})$, $\mathbf{A} = \mathbf{A}(z)$ is a $d \times d$ matrix of functions [here, by assumption, all in $\mathbb{C}(z)$], and $\mathbf{b} = (\widetilde{t}, \widetilde{t}\,', \ldots, \widetilde{t}^{(d-1)})$; see [33, vol II, §9.3] for the reduction. Recall that a fundamental matrix $\mathbf{W}$ for the system (69) is by definition a nonsingular $d \times d$ matrix whose columns each satisfy the homogenous system $\partial_z \mathbf{y}(z) = \mathbf{A}\mathbf{y}(z)$. Then the general solution to the inhomogeneous system (69) is, by the classical "variation-of-constants" formula,

$$\mathbf{y}(z) = \mathbf{W}(z)\cdot\mathbf{W}^{-1}(z_0)\cdot\mathbf{y}(z_0) + \mathbf{W}(z)\cdot\int_{z_0}^z \mathbf{W}^{-1}(x)\cdot\mathbf{b}(x)\,dx; \tag{70}$$

see once more [33, vol II]. (The initial conditions at some z_0 are assumed to be known.) This provides the solution to (67) as

$$f(z) = \mathcal{L}[\widetilde{t}(z)],$$

with $\mathcal{L}$ a *linear integral transform* that involves polynomially the elements of a fundamental matrix $\mathbf{W}$ as well as the inverse of the Wronskian $\det \mathbf{W}$. (The case of Euler equations is somewhat simpler, as it is fully explicit [7, 19].) For instance, the case of 2-dimensional quadtrees leads to a still explicit form [36], namely, $f(z) = \mathcal{L}[\widetilde{t}(z)]$ where

$$\mathcal{L}[e(z)] = \frac{1+2z}{(1-z)^2}\int_0^z \frac{(1-y)^3}{y(1+2y)^2}\left[\int_0^y \frac{1+2x}{(1-x)^2}e(x)\,dx\right]\,dy.$$

From here onward we suppose for simplicity that $\omega_n \equiv 1$, so that f is an ordinary generating function, though our discussion extends readily to more general normalization constants. Call a system *dominantly regular*[9] if it is singular at 1 (i.e., if the matrix $\mathbf{A}$ has a pole at 1, but at no other point in $|z| \leq 1$ except possibly 0) and if the pole of $\mathbf{A}$ at 1 is simple—the latter case is known as a singularity "of the first kind". All the classical examples listed above and generalizing binary search trees satisfy this condition. We then have:

Theorem 17. *Let a tree recurrence be expressed by a differential system that is dominantly regular. Then the expectations of costs induced by the tolls t_n^α and $t_n^{\log}$ admit complete asymptotic expansions in descending powers of n, possibly with logarithmic terms.*

Proof. First, we observe that for any tree recurrence, the cost induced by an eventually increasing nonnegative toll $t_n \to +\infty$ is at least t_n (by the very nature of the tree recurrence) and at most $O(nt_n)$ (by induction). Thus, for the tolls under consideration, the generating function of costs, $f(z)$, has radius of convergence exactly equal to 1. We also observe that the values of f and its derivatives at some point z_0 such that $|z_0| < 1$ are well-defined. We may adopt for instance $z_0 = \frac{1}{2}$ in the variation-of-constants formula.

By the classical theory of singularities of the first kind, each of the column vectors of matrix $\mathbf{W}$ is analytic for z in a neighborhood of 1 slit along the ray $[1, +\infty)$. There, as z tends to 1, it admits a representation as a finite combination of terms of the form

$$(1-z)^\alpha L(z)^k R(1-z),$$

where α is an algebraic number (a root of the indicial equation), k an integer, and R is analytic at 0. Thus, each element of $\mathbf{W}$ is amenable to singularity analysis as it is Δ-continuable and admits a bona fide expansion near 1.

By formula (70), there remains to discuss the elements of $\mathbf{W}^{-1}$. By the cofactor rule, the elements of $\mathbf{W}^{-1}$ involve polynomially the elements of $\mathbf{W}$ divided by the Wronskian determinant $\det \mathbf{W}(z)$. It is a well-known fact (see §9.3 of [33, vol II]) that the Wronskian is expressible in terms of the system alone and one has

$$[\det \mathbf{W}(z)]^{-1} = [\det \mathbf{W}(z_0)]^{-1} \exp\left(-\int_{z_0}^{z} \operatorname{tr} \mathbf{A}(x)\, dx\right).$$

[Here $\operatorname{tr}(\cdot)$ denotes the matrix trace operator.] By the dominant regularity assumption, the trace is here a rational function with at most a simple pole at 1, so that its integral is either analytic at 1 or logarithmic. In either case, $(\det \mathbf{W}(z))^{-1}$ is of singularity analysis type, and so are consequently all the

[9] The term "dominantly regular" evokes the fact that the condition concerns the *dominant* singularity of the solution function, where the singularity of the system is of the so-called "*regular*" type (first-kind implies regular singularity by a well-known theorem; see [33, vol II, Theorem 9.4d]).

elements of the inverse of the fundamental matrix $\mathbf{W}^{-1}(z)$. By the singular integration and singular differentiation theorems of Section 3, there results that the integral transform (70) preserves for functions the character of being amenable to singularity analysis. Since the toll generating functions are of singularity analysis type, basic singularity analysis is applicable to $f(z)$. The result follows. □

In principle, higher moments will also become accessible to singularity analysis once the nonlinear integral forms Q extending Q^{BST} of (66) have been worked out. We are however not aware of existing research in this direction, despite the fact that the splitting probabilities are known in a number of cases (see, e.g., [23] for quadtrees). There is interest in these questions, as partly heuristic recent work by Majumdar and collaborators (see, e.g., [12] for the type of method employed and succinct developments) indicates the probable existence of phase transitions in the number of internal nodes of d-dimensional quadtrees for large enough d ($d \geq d_c = 9$ is suggested) in a way similar to what is already well established for the size of m-ary search trees [6, 19, 47].

As a final note, we'd like to mention digital trees, which were recognized to be amenable to treatment by *ordinary* (rather than the more customary exponential) generating functions in [27]. Techniques of the present paper would most likely be usable in such a context, in particular as regards tolls of the form n^α and $\log n$. A partial classification of cost functions along these lines has already been given by Derfel and Vogl in [13].

Acknowledgements. The second author is grateful to Don Knuth for a summer invitation to Stanford in 1988, during which his work on the subject was started. This work was supported in part by the Future and Emerging Technologies programme of the EU under contract number IST-1999-14186 (ALCOM–FT Project).

The first and third authors' research is supported by NSF grant DMS–9803780 and DMS–0104167, and by The Johns Hopkins University's Acheson J. Duncan Fund for the Advancement of Research in Statistics.

References

[1] ABRAMOWITZ, M., AND STEGUN, I. A. *Handbook of Mathematical Functions.* Dover, 1973. A reprint of the tenth National Bureau of Standards edition, 1964.

[2] BERGERON, F., FLAJOLET, P., AND SALVY, B. Varieties of increasing trees. In *CAAP'92* (1992), J.-C. Raoult, Ed., vol. 581 of *Lecture Notes in Computer Science*, pp. 24–48. Proceedings of the 17th Colloquium on Trees in Algebra and Programming, Rennes, France, February 1992.

[3] BERNDT, B. C. *Ramanujan's Notebooks, Part I.* Springer Verlag, 1985.

[4] BIEBERBACH, L. *Lehrbuch der Funktionen Theorie.* Teubner, Leipzig, 1931. In two volumes. Reprinted by Johnson, New York, 1968.

[5] BOUSQUET-MÉLOU, M., AND GUTTMANN, A. J. Three-dimensional self-avoiding convex polygons. *Physical Review E. Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics. Third Series 55*, 6, part A (1997), R6323–R6326.

[6] CHERN, H.-H., AND HWANG, H.-K. Phase changes in random m-ary search trees and generalized quicksort. *Random Structures & Algorithms 19*, 3-4 (2001), 316–358. Analysis of algorithms (Krynica Morska, 2000).

[7] CHERN, H.-H., HWANG, H.-K., AND TSAI, T.-H. An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms. *Journal of Algorithms 44*, 1 (2002), 177–225. Analysis of algorithms.

[8] CHYZAK, F. Staircase polygons, a simplified model for self-avoiding walks. *Studies in Automatic Combinatorics I* (1997). 26 pages. Published electronically by Algorithms Project, INRIA.

[9] CORLESS, R. M., GONNET, G. H., HARE, D. E. G., JEFFREY, D. J., AND KNUTH, D. E. On the Lambert W function. *Advances in Computational Mathematics 5* (1996), 329–359.

[10] CORMEN, T. H., LEISERSON, C. E., AND RIVEST, R. L. *Introduction to Algorithms.* MIT Press, New York, 1990.

[11] COVER, T. M., AND THOMAS, J. A. *Elements of information theory.* John Wiley & Sons Inc., New York, 1991. A Wiley-Interscience Publication.

[12] DEAN, D. S., AND MAJUMDAR, S. N. Phase transition in a random fragmentation problem with applications to computer science. *Journal of Physics A: Mathematical and General 35* (2002), L501–L507.

[13] DERFEL, G., AND VOGL, F. Asymptotic analysis of binomial recurrences. In *Topics in functional differential and difference equations (Lisbon, 1999)*, vol. 29 of *Fields Inst. Commun.* Amer. Math. Soc., Providence, RI, 2001, pp. 93–104.

[14] DIENES, P. *The Taylor Series.* Dover, New York, 1958. A reprint of the first Oxford University Press edition, 1931.

[15] DOYLE, P. G., AND SNELL, J. L. *Random walks and electric networks.* Mathematical Association of America, Washington, DC, 1984.

[16] FILL, J. A. On the distribution of binary search trees under the random permutation model. *Random Structures & Algorithms 8*, 1 (1996), 1–25.

[17] FILL, J. A., AND KAPUR, N. An invariance princple for simply generated families of trees. Tech. rep., Department of Mathematical Sciences, The Johns Hopkins University, 2003.

[18] FILL, J. A., AND KAPUR, N. Limiting distributions of addititive functionals on Catalan trees, 2003. Submitted for publication. Available at `http://eprints.mts.jhu.edu/archive/00000033/`.

[19] FILL, J. A., AND KAPUR, N. Transfer theorems and asymptotic distributional results for m-ary search trees, 2003. Submitted for publication. Available at `http://eprints.mts.jhu.edu/archive/00000032/`.

[20] FINCH, S. *Mathematical Constants.* Cambridge University Press, New-York, 2003. In press.

[21] FLAJOLET, P. Singularity analysis and asymptotics of Bernoulli sums. *Theoretical Computer Science 215*, 1-2 (1999), 371–381.

[22] FLAJOLET, P., AND GOLIN, M. Mellin transforms and asymptotics: The mergesort recurrence. *Acta Informatica 31* (1994), 673–696.

[23] FLAJOLET, P., GONNET, G., PUECH, C., AND ROBSON, J. M. Analytic variations on quadtrees. *Algorithmica 10*, 7 (Dec. 1993), 473–500.

[24] FLAJOLET, P., AND ODLYZKO, A. M. Singularity analysis of generating functions. *SIAM Journal on Algebraic and Discrete Methods 3*, 2 (1990), 216–240.

[25] FLAJOLET, P., POBLETE, P., AND VIOLA, A. On the analysis of linear probing hashing. *Algorithmica 22*, 4 (Dec. 1998), 490–515.

[26] FLAJOLET, P., AND PUECH, C. Partial match retrieval of multidimensional data. *Journal of the ACM 33*, 2 (1986), 371–407.

[27] FLAJOLET, P., AND RICHMOND, B. Generalized digital trees and their difference–differential equations. *Random Structures & Algorithms 3*, 3 (1992), 305–320.

[28] FLAJOLET, P., AND STEYAERT, J.-M. A complexity calculus for recursive tree algorithms. *Mathematical Systems Theory 19* (1987), 301–331.

[29] GLASSER, M. L. A Watson sum for a cubic lattice. *Journal of Mathematical Physics 13* (1972), 1145–1146.

[30] GOULDEN, I. P., AND JACKSON, D. M. *Combinatorial Enumeration.* John Wiley, New York, 1983.

[31] GREENE, D. H., AND KNUTH, D. E. *Mathematics for the analysis of algorithms*, second ed. Birkhauser, Boston, 1982.

[32] GUTTMANN, A. J., AND PRELLBERG, T. Staircase polygons, elliptic integrals, heun functions, and lattice green functions. *Physical Review E 47*, 4 (1993), R2233–2236.

[33] HENRICI, P. *Applied and Computational Complex Analysis.* John Wiley, New York, 1977. 3 volumes.

[34] HENRICI, P. *Applied and computational complex analysis. Vol. 2.* Wiley Interscience [John Wiley & Sons], New York, 1977. Special functions—integral transforms—asymptotics—continued fractions.

[35] HOFRI, M. *Analysis of Algorithms: Computational Methods and Mathematical Tools.* Oxford University Press, 1995.

[36] HOSHI, M., AND FLAJOLET, P. Page usage in a quadtree index. *BIT 32* (1992), 384–402.

[37] HWANG, H.-K., AND NEININGER, R. Phase change of limit laws in the quicksort recurrence under varying toll functions. *SIAM Journal on Computing 31*, 6 (2002), 1687–1722 (electronic).

[38] KAPUR, N. *Additive functionals on random search trees.* PhD thesis, Department of Mathematical Sciences, The Johns Hopkins University, April 2003.

[39] KNUTH, D. E. *The Art of Computer Programming*, 3rd ed., vol. 1: Fundamental Algorithms. Addison-Wesley, 1997.

[40] KNUTH, D. E. *The Art of Computer Programming*, 2nd ed., vol. 3: Sorting and Searching. Addison-Wesley, 1998.

[41] KNUTH, D. E. *Selected papers on analysis of algorithms.* CSLI Publications, Stanford, CA, 2000.

[42] KNUTH, D. E., AND PITTEL, B. A recurrence related to trees. *Proceedings of the American Mathematical Society 105*, 2 (Feb. 1989), 335–349.

[43] KNUTH, D. E., AND SCHÖNHAGE, A. The expected linearity of a simple equivalence algorithm. *Theoretical Computer Science 6* (1978), 281–315.

[44] LAWLER, G. F. *Intersections of random walks.* Birkhäuser Boston Inc., Boston, MA, 1991.

[45] LOUCHARD, G. The Brownian excursion: a numerical analysis. *Computers and Mathematics with Applications 10*, 6 (1984), 413–417.

[46] LOUCHARD, G., SCHOTT, R., TOLLEY, M., AND ZIMMERMANN, P. Random walks, heat equations and distributed algorithms. *Journal of Computational and Applied Mathematics 53* (1994), 243–274.

[47] MAHMOUD, H. M. *Evolution of Random Search Trees.* John Wiley, New York, 1992.

[48] MAHMOUD, H. M. *Sorting, A Distribution Theory.* Wiley-Interscience, New York, 2000.

[49] MEIR, A., AND MOON, J. W. On the altitude of nodes in random trees. *Canadian Journal of Mathematics 30* (1978), 997–1015.

[50] NEININGER, R. On binary search tree recursions with monomials as toll functions. *Journal of Computational and Applied Mathematics 142*, 1 (2002), 185–196.

[51] ODLYZKO, A. M. Asymptotic enumeration methods. In *Handbook of Combinatorics*, R. Graham, M. Grötschel, and L. Lovász, Eds., vol. II. Elsevier, Amsterdam, 1995, pp. 1063–1229.

[52] OLVER, F. W. J. *Asymptotics and Special Functions.* Academic Press, 1974.

[53] PITTEL, B. Normal convergence problem? Two moments and a recurrence may be the clues. *The Annals of Applied Probability 9*, 4 (1999), 1260–1302.

[54] RÖSLER, U., AND RÜSCHENDORF, L. The contraction method for recursive algorithms. *Algorithmica 29*, 1/2 (2001), 3–33.

[55] ROURA, S. An improved master theorem for divide-and-conquer recurrences. In *Automata, languages and programming (Bologna, 1997)* (Berlin, 1997), Springer, pp. 449–459.

[56] SALVY, B. *Asymptotique automatique et fonctions génératrices.* Ph. D. thesis, École Polytechnique, 1991.

[57] SEDGEWICK, R. *Algorithms*, second ed. Addison–Wesley, Reading, Mass., 1988.

[58] SEDGEWICK, R., AND FLAJOLET, P. *An Introduction to the Analysis of Algorithms.* Addison-Wesley Publishing Company, 1996.

[59] STANLEY, R. P. *Enumerative Combinatorics*, vol. II. Cambridge University Press, 1998.

[60] SZPANKOWSKI, W. *Average-Case Analysis of Algorithms on Sequences.* John Wiley, New York, 2001.

[61] TAKÁCS, L. On a probability problem connected with railway traffic. *Journal of Applied Mathematics and Stochastic Analysis 4*, 1 (1991), 1–27.

[62] VITTER, J. S., AND FLAJOLET, P. 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.

[63] WHITTAKER, E. T., AND WATSON, G. N. *A Course of Modern Analysis*, fourth ed. Cambridge University Press, 1927. Reprinted 1973.