An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms

Hua-Huai Chern[1]
Department of Mathematics and Computer Science Education
Taipei Municipal Teachers College
Taipei 100
Taiwan

Hsien-Kuei Hwang[2]
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

Tsung-Hsi Tsai
Institute of Statistical Science
Academia Sinica
Taipei 115
Taiwan

March 18, 2002

Abstract

Cauchy-Euler differential equations surfaced naturally in a number of sorting and searching problems, notably in quicksort and binary search trees and their variations. Asymptotics of coefficients of functions satisfying such equations has been studied for several special cases in the literature. We study in this paper the most general framework for Cauchy-Euler equations and propose an asymptotic theory that covers almost all applications where Cauchy-Euler equations appear. Our approach is very general and requires almost no background on differential equations. Indeed the whole theory can be stated in terms of recurrences instead of functions. Old and new applications of the theory are given. New phase changes of limit laws of new variations of quicksort are systematically derived. We apply our theory to about a dozen of diverse examples in quicksort, binary search trees, urn models, increasing trees, etc.

Key words. Quicksort, binary search trees, Cauchy-Euler differential equations, asymptotic transfers, method of moments, phase changes, convergence in distributions.

1 Introduction

Cauchy-Euler differential equations. Differential equations of the type

$$(1-z)^r f^{(r)}(z) = c_{r-1}(1-z)^{r-1} f^{(r-1)}(z) + \cdots + c_0 f(z) + \phi(z), \qquad (1)$$

with given initial conditions will be referred to, here and throughout this paper, the *Cauchy-Euler differential equations* (CE equations, for short); see Nagle et al. [88]. Here r is a positive integer, f and ϕ are analytic functions or formal power series, and the c_j's are complex numbers. Such equations

---

[1] Supported by National Science Council of ROC under the Grant `NSC-90-2115-M-133-001`.

[2] The work of this author was partially supported by National Science Council of ROC under the Grants `NSC-90-2118-M-001-034` and `NSC-90-2118-M-001-028`.

are also referred to in the literature as *Euler equations*, *Euler-Cauchy equations*, or *equidimensional equations*; see Bender and Orszag [10], Ince [65]. Equations of this type often appear in analysis of algorithms, notably in analysis of quicksort and search trees. We propose in this paper an asymptotic theory for the coefficients of z^n in the Taylor expansion of f when the c_j's are given and asymptotics of the coefficients of ϕ is known. Our framework covers all previous cases and our method of proof is very general and applies to almost all possible forms of interests of (1). Important applications of our theory to distributional properties of cost measures of a variety of quicksort are also given.

Quicksort. Sir Charles Antony Richard Hoare (knighted in 2000) invented quicksort in the early sixties (see Hoare [56, 57]), which has since become one of the most widely used and studied sorting algorithms. Quicksort was ranked among the top ten algorithms in the 20th century with the greatest influence on the development and practice of science and engineering (see [66]). The *simplicity* and *efficiency* of quicksort has tremendous impact on both the theory and the practice of sorting and related algorithms. On the practical side, a huge number of variants of quicksort have been devised and applied in diverse disciplines, the usefulness of the prototypical design paradigm increasing significantly the popularity of quicksort. On the theoretical side, the probabilistic analysis of quicksort and its variations introduced many intriguing phenomena and challenging problems, the close relationship between quicksort and binary search trees adding further dimensions to the 40-year-old sorting algorithm. More results, connections and references will be given later.

Interestingly, the original paper by Hoare [57] on quicksort remains the most cited paper after 40 years of its invention among papers dealing with quicksort, according to the statistics collected from ISI's (Institute of Scientific Information) SCI and NEC's ResearchIndex databases.

For variants of quicksort (for linear arrays and for sequential machines) before 1975, see Sedgewick [101, 102], and [2, 5, 12, 23, 34, 37, 58, 86, 87, 96, 99, 111, 113, 115, 116, 117] after 1975.

Description of quicksort. Given a sequence of distinct keys (or elements, numbers, etc.), the quicksort proper works as follows. The sorting is complete if there is only one key to be sorted; otherwise, choose a pivot, say x, from the input and then partition the input into two subsets containing keys with values smaller and larger, respectively, than x. Then sort the two subfiles recursively by the same procedure.

Although such a quicksort has poor worst-case performance (quadratic in input size), it is very efficient for most practical purposes, and theoretical analysis under the usual uniform permutation model confirms such a phenomenon (see [72, 102]). Major practical variants of quicksort in the literature concentrate on avoiding the potential worst-case without slowing significantly down the highly efficient inner loops. One of the most frequently used ideas to achieve this is to take the median of a random sample of odd size (even size samples having no gain, as demonstrated by Sedgewick [102] and Hennequin [54]) and the most successful variant is the median-of-three version due partly to simplicity of implementation and file size of most applications.

Another interesting variant (inter alia) was introduced by Bentley and McIlroy [11]: instead of choosing the median of three elements as the pivot, they use the median of three elements each of which is the median of three elements and call this approximate median-of-nine the ninther, after J. W. Tukey [112]. Note that for a random permutation of nine elements, the ninther equals the median with probability 4/7 and is the fourth or sixth smallest with equal probability 3/14; it requires 32/3 comparisons on average to find the ninther, which is smaller than the theoretically best median-finding algorithm for nine elements; see Knuth [72, §5.3.5]. In general, the idea of choosing an easily-computable approximate median is extremely useful and has been fruitfully applied in diverse disciplines ranging from computational statistics to databases, from fuzzy systems to geography (see for example [9, 25, 33, 60, 69, 82, 98, 118]). We will consider several new ways of extending the ninther quicksort and study the phase changes of the limiting distributions of a linear cost measure.

A class of quicksort. If the pivot of quicksort is selected by taking first a random sample of r elements, and choosing the $(j+1)$-st order statistics in this sample with probability p_j, then the moments of the cost measures of this quicksort satisfy the recurrence (assuming, *here and throughout this paper*, the usual uniform permutation model for the input)

$$a_n = \sum_{0\le k<n} \pi_{n,k}(a_k + a_{n-1-k}) + b_n, \tag{2}$$

with suitable initial conditions, where (see [7, 54, 83, 102])

$$\pi_{n,k} = \sum_{0\le j<r} p_j \frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}}, \qquad \sum_{0\le j<r} p_j = 1.$$

Note that the ratio $\binom{k}{j}\binom{n-1-k}{r-1-j}/\binom{n}{r}$ is the probability that the $(j+1)$-st order statistics of a random permutation of r elements is the $(k+1)$-st order statistics of a random permutation of n elements. Also $\sum_{0\le k<n}\pi_{n,k} = 1$. An interpretation of this quicksort in terms of search trees is given in Section 3; another urn model interpretation is given at the end of Section 4.

A quicksort interpretation to CE equations. To see how CE equations are generated from real problems, observe that we can first rewrite (1) in the form

$$f^{(r)}(z) = \sum_{0\le j<r} c_j (1-z)^{-(r-j)} f^{(j)}(z) + g^{(r)}(z),$$

where $g^{(r)}(z) := (1-z)^{-r}\phi(z)$. Then the coefficients $a_n := [z^n]f(z)$ and $b_n := [z^n]g(z)$ satisfy, where $[z^n]f(z)$ denotes the coefficients of z^n in the Taylor expansion of f,

$$a_n = \sum_{0\le k<n} \pi'_{n,k} a_k + b_n, \tag{3}$$

where

$$\pi'_{n,k} := \sum_{0\le j<r} c'_j \frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}}, \qquad c'_j := c_j \frac{j!}{r!}.$$

This is of the same form as (2) and thus gives an "algorithmic interpretation" of CE equations in the case when $c_j = r!(p_j + p_{r-1-j})/j!$.

Order vs coefficients. While the CE equations are usually treated as easy cases in most textbooks on differential equations, the asymptotics of the Taylor coefficients (of functions satisfying CE equations) is more subtle as far as more precise expansions are concerned. This problem also arises for general differential equations. The main problem is that to determine the order of the solution of a differential equation, an O-estimate for the function near the singularity suffices, while to compute the asymptotics of the Taylor coefficients, all constants involved have to be determined in more explicit forms, together with more uniform estimate of the function. Typical examples of such a situation (where explicit determinations of the dominant constants are difficult) include partial-match queries in k-d trees [45], partial-match queries in quadtrees [41], and the number of consecutive records [21].

Methodology. Restricting to CE equations, the *method of linear operator* (see [19, 20, 50, 54, 65, 102]) is especially useful for providing effective forms of the coefficients in the asymptotic expansion of the function (near its dominant singularity); it is also computationally simpler than other approaches such as the method of undetermined coefficients and the method of variation of parameters. We extend and summarize this approach and provide a general framework (cf. [20]). While our approach relies heavily on and exploits fully the decomposability of the CE equations, it is completely elementary (in the sense that it does not use complex analysis) and no analytic properties of the equations are used.

The approach is, although similar, different from that used in our previous paper [20]. It is still elementary but less computational. Another new feature of our approach is that we can easily include the case of zeros of higher orders. All previous approaches for this class of problems rely on the simplicity of the zeros, which is in many cases not obvious to prove; see Sections 3 and 4 for concrete examples. We show that the multiplicity of the dominant zero changes only mildly the form of the asymptotic approximations.

Also the differential equations in this paper are purely "formal" in nature. Convergence of series or unicity of solutions are immaterial in our analysis since our approach does not rely on any analytic properties, but instead manipulates the differential equations in terms of the underlying Taylor coefficients. This means that we can even have series of the form $\sum_n n!z^n$ (without any radius of convergence) as the inhomogeneous part. While we might develop our arguments by completely avoiding differential equations and by using only sequences and recurrences, the resulting proof would be messier, less elegant, and the interplay between analytic approach and algebraic approach less transparent; see Section 2.3 for more detailed discussions.

Binary search trees (BSTs). Quicksort algorithms are divide-and-conquer in nature and have a natural correspondence to binary search trees. Given a sequence of distinct elements, the binary search tree associated with this sequence is constructed by placing the first element of the sequence in the root; the remaining elements are compared to the root element; those smaller elements are directed to the left branch and are constructed recursively as a BST; the larger elements are similarly constructed as the right subtree. See Stephensen [108] for an alternative construction.

Such a construction makes it possible to interpret some cost measures of quicksort in terms of parameters in BSTs and vice versa. For example, the number of comparisons used by quicksort is essentially the path length of BST, and the number of partitioning stages in quicksort corresponds to the number of non-leaf nodes in BSTs. On the other hand, the m-ary search tree, one natural extension of BST to multiway search trees, can be interpreted as quicksort using multiple pivots.

Such a connection is remarkable since it links an algorithm to a data structure, an array to a tree, on the one hand, and a sorting problem to a searching problem, on the other hand. It is this multifaceted feature that makes quicksort and BST so popular, fundamental and prototypical.

While quicksort is closely related to BSTs, they are not identical in all aspects. For example, it does not seem obvious how to construct practical versions of quicksort that correspond to, say hight-balanced (or AVL) BSTs. Also there are many multidimensional extensions of BSTs like k-d trees and quadtrees, but "multidimensional quicksort" is not obvious since there is no total ordering for multivariate data.

Binary increasing trees. An alternative way of constructing a tree structure from a sequence of distinct elements is as follows. Put the smallest element in the root; this splits the sequence into three parts: left part, the smallest element itself, and the right part; construct the left and right parts recursively as (binary) increasing trees. By construction, all paths from a node to the root form a decreasing sequence. Such trees are sometimes called tournament trees, nestling trees, or heap-ordered trees; see Bergeron et al. [13], Burge [15], Françon [46], Kundu [73], Smythe and Mahmoud [106], Stanley [107, pp. 23–25], Vuillemin [114] for further information. These is indeed a bijection between increasing trees and BSTs. Variants of these trees introduce, as those of BSTs, naturally CE equations for recursively defined parameters.

Method of moments. The first use of the method of moments can be traced back to Chebyshev's work on the moment problem and the classical central limit theorems (see [1, 17]). The fundamental theorem for this method is the *moment convergence theorem of Frechet and Shohat* (see [47, 76]): *If* $E(X_n^m) \to \mu_m < \infty$ *for all* $m \geq 0$*, and the sequence* $\{\mu_m\}$ *determines uniquely a distribution, then* X_n *converges in distribution to some random variable* X*, and* $E(X^m) = \mu_m$*.* Such an approach, although primitive in some sense, has been successfully applied to many problems (see, for example, [14, 27, 32, 67]). Applications of this method to analysis of algorithms include: height and path length in binary trees (see [42], [110]), log-product of subtree sizes in BSTs (see [38]), linear probing hashing (see [44]), tries (see [100]), and, more recently, m-ary search trees (see [20, 61]), quicksort under varying "toll functions" (see [62]), quickselect (see [63]), broadcast communication model (see [18]), and bucket digital search trees (see [59]).

The usefulness of the method of moments is further discussed in Section 3, where we propose a simple class of quicksort-type recurrences, with an emphasis on phase changes of the limit laws.

Asymptotic transfers. Most cost measures of quicksort or characteristic parameters of BSTs are recursive in nature. This means that all moments (centered or not) satisfy essentially the same type of recurrence with different "inhomogeneous part" (or "toll function"). Thus all asymptotic estimates needed for the moments can be encapsulated into the so-called "asymptotic transfers" in which we deduce asymptotics of the solution to the recurrence in question from that of the "toll functions".

For example, in the case (3), we assume that the asymptotics of b_n is known to some extent and then derive asymptotics of a_n. The extent of the asymptotics to which b_n should be known depends on applications. Also the tools (analytic or elementary) needed for deriving the asymptotic transfers vary from one problem to another. However, the rough idea is that if the "toll function" is large and dominating, then the analysis is in general easier; the analysis is more delicate if the "toll function" is small since we have to find means of summing all contributions from each individual part.

Phase changes. Random structures often exhibit phase changes under varying parameters and have received much attention in recent computer science literature, notably for NP-complete or NP-hard problems; see for example Dubois et al. [35]. The phenomenon is similar to the usual change of phase of matters (like water).

For quicksort (at large), we have demonstrated two types of phase changes of the limit laws for some cost measures:

1. from normal to non-existence (see [20]);
2. from normal to non-normal (see [62]).

We will refer to the first type of phase change as type I phase change and that of the second the type II phase change. The type I phase change occurs when we increase the sample size or increase the number of pivots used; The type II phase change occurs when we increase the "toll functions" (the cost used in partitioning the input) from small ($O(n^{1/2})$) to large ($\gg n^{1/2}$). Note that statistical physicists also distinguish between two types of phase transitions: type I (discrete) and type II (continuous). We use "phase change" instead of "phase transition" since the varying parameters in our problems are all integers.

In theory, we can produce both types of phase changes by considering the class of (binary) quicksort described above by varying the underlying distribution $\{p_j\}_j$. We will consider random variables of the form

$$X_n \overset{d}{=} X_{I_n} + X^*_{n-1-I_n} + 1 \qquad (n \geq r),$$

with $X_n = 0$ for $0 \leq n < r$, where $P(I_n = k) = \pi_{n,k}$. Here the X_n^*'s and the X_n's are independent and have the same distribution. This class of random variables corresponds to the number of partitioning stages used by the quicksort described above.

The important index is, as before, the second largest zero(s) (in real parts) of the associated indicial equation:

$$\Lambda_r(\vartheta) := \vartheta \cdots (\vartheta + r - 1) - r! \sum_{0 \leq j < r} (p_j + p_{r-1-j}) \frac{\vartheta \cdots (\vartheta + j - 1)}{j!} = 0.$$

We will consider a simple cost measure, the number of partitioning stages used by quicksort, in Section 3. Our results roughly say that

1. if the second largest zero(s) are complex (denoting their real part by α):
   (a) if $\alpha \leq 3/2$, then the limiting law is normal;
   (b) if $3/2 < \alpha < 2$, then the limiting law does not exist;
2. if the second largest zero(s), denoted by α, is real:
   (a) if $\alpha \leq 3/2$, then the limiting law is normal;
   (b) if $3/2 < \alpha < 2$, then the limiting law exists and is not normal.

New variants of quicksort evincing the type I phase changes will be given. For example, if we use the remedian of 3^d elements, defined recursively as the median of three remedians of 3^{d-1} elements, then the limiting distribution of the number of partitioning stages is normal for $d \leq 5$, and does not exist for $d \geq 6$ (note that the sample size is $3^6 = 729$). If $p_j = \binom{r-1}{j}/2^{r-1}$, then the same phase change of the limit laws occurs at $r = 236$. Similarly, if we choose the pivot as the median of $(2s+1)$ medians each of which is the median of $(2t+1)$ elements, then a complete description of the pairs leading to asymptotic normality is given in Table 1.

t	0	1	2	3	4	5	6	7	8,9	10,11	12,...,16	17,..., 25	26,...,58
$0 \leq s \leq ?$	58	25	16	11	9	7	6	5	4	3	2	1	0

Table 1: *The upper bound of s for each fixed t (in quicksort using (s,t)-median) for which the asymptotic normality subsists. The table is symmetric in s and t. Also the degenerate case $(s,t) = (0,0)$ has to be discarded.*

Thus the determination of the limit laws is more or less reduced to the calculation of the real part of the second largest zero(s). While we can easily produce lots of examples with type I phase change, it is not obvious how to generate concrete probability distributions with type II phase change.

Organization of the paper. We start in the next section with the asymptotics of CE equations and extend our previous framework on asymptotic transfers in [20] to cover general CE equations of the form (1). Refinements of some asymptotic approximations under stronger assumptions of the "toll function" (corresponding to non-homogeneous part) are also discussed. We then consider in

Section 3 a class of quicksort and study possible phase changes of limit laws associated with such sorting schemes when we vary the underlying probability distribution of choosing the order statistics. Our proof relies on the method of moments and the asymptotic transfers derived in Section 2. Applications are then discussed in Section 4 and they deal mostly with variants of quicksort and BSTs.

2 Asymptotics of CE equations

In this section, we consider the general, inhomogeneous, high-order, Cauchy-Euler, linear differential equations of the form

$$(1-z)^r f^{(r)}(z) = \sum_{0\le j<r} c_j (1-z)^j f^{(j)}(z) + (1-z)^r g^{(r)}(z), \tag{4}$$

with the initial conditions $f^{(j)}(0) = g^{(j)}(0)$, $0 \le j < r$, which, in terms of $a_n := [z^n]f(z)$ and $b_n := [z^n]g(z)$, can also be written as $a_j = b_j$ for $0 \le j < r$.

Under the sole assumption that the dominant zero of the indicial equation associated with (4) is simple with real part positive, we can automatically infer asymptotics of a_n from that of b_n; precise expressions are given for major constants involved. The assumption is more a technical convenience than a restriction, and our approach can be easily modified in case when the assumption fails; this is discussed in Section 2.4.

2.1 Solutions of general CE equations

Linear operator. The starting point for solving (4) is to introduce the linear differential operator $\vartheta := (1-z)\frac{\mathrm{d}}{\mathrm{d}z}$ and $(1-z)^j \frac{\mathrm{d}^j}{\mathrm{d}z^j} = \vartheta^{\overline{j}} := \vartheta(\vartheta+1)\cdots(\vartheta+j-1)$ for each $j \in \mathbb{N}$. Then we can rewrite (4) as

$$\Lambda_r(\vartheta) f(z) := \phi(z), \quad \text{where} \quad \Lambda_r(\vartheta) := \vartheta^{\overline{r}} - \sum_{0\le j<r} c_j \vartheta^{\overline{j}}. \tag{5}$$

From an analytic point of view, analytic properties of $\phi(z)$ in the inhomogeneous part are indispensable in justifying the unicity of solution, and in solving the equation (either exactly or asymptotically). Our approach proceeds along a very different line; it views the equation as some formal "coefficients transform" that links the Taylor coefficients of $\phi(z)$ to those of $f(z)$; see (3). Thus all our manipulations of the equation are taken at the level of coefficients; see below for more details.

Indicial equation. The indicial equation associated with (5) is $\Lambda_r(\vartheta) = 0$. The zeros of polynomial $\Lambda_r(\vartheta)$, when viewing ϑ as a complex variable, determine the solutions of the homogeneous equation $\Lambda_r(\vartheta)\mathbf{y} = 0$. Let the zeros of $\Lambda_r(\vartheta)$ be arranged in descending order of their real parts as

$$\Re(\lambda_1) \ge \Re(\lambda_2) \ge \Re(\lambda_3) \ge \cdots \ge \Re(\lambda_r).$$

Also let $\mathscr{Z} := \{(\rho_j, m_j)\}_{1\le j\le\kappa}$ denote the set of pairs of distinct zeros and their multiplicities; more precisely,

$$\mathscr{Z} := \left\{(\rho, m) : \Lambda_r^{(j)}(\rho) = 0,\ 0 \le j < m,\ \text{but } \Lambda_r^{(m)}(\rho) \ne 0 \text{ for } \rho \in \mathbb{C},\ m \in \mathbb{N}\right\}.$$

General solutions to homogeneous equations. When $(\rho, m) \in \mathscr{Z}$, the function

$$y(z) = \sum_{0 \le j < m} C_j (1-z)^{-\rho} \log^j (1-z),$$

is a solution to the equation $\Lambda_r(\vartheta)\mathbf{y} = 0$, where the C_j's are constants. Thus the general solution of $\Lambda_r(\vartheta)\mathbf{y} = 0$ is of the form

$$\mathbf{y}(z) = \sum_{(\rho,m)\in\mathscr{Z}} \sum_{0 \le j < m} C_j(\rho)(1-z)^{-\rho} \log^j (1-z),$$

where the $C_j(\rho)$'s are constants.

The fundamental equation. Recall the following simple lemma.

Lemma 1 ([20, Lemma 2]). *The function*

$$Y(z) = y_0(1-z)^{-\rho} + (1-z)^{-\rho} \int_0^z (1-x)^{\rho-1}\phi(x)\,\mathrm{d}x$$

is the exact solution to the first-order differential equation

$$\begin{cases} (\vartheta - \rho)Y(z) = \phi(z), \\ Y(0) = y_0. \end{cases}$$

Analytic solution vs formal solution. Lemma 1 implicitly assumes that the integral is convergent. What happens if the integral diverges? This happens when $\phi(z)$ has singularities on the path of integration or even converges only at $z = 0$. The lemma is to be interpreted in a formal power series sense as follows.

Lemma 2. *Let $\phi(z) = \sum_{n\ge0} \phi_n z^n$ and $Y(z) = \sum_{n\ge0} y_n z^n$ be two formal power series. Then the solution to the recurrence*

$$y_n = \frac{\rho}{n} \sum_{0\le j<n} y_j + \frac{1}{n} \sum_{0\le j<n} \phi_j \qquad (n \ge 1), \tag{6}$$

with y_0 given, satisfies

$$y_n = y_0 \binom{n+\rho-1}{n} + \sum_{0\le k<n} \frac{\phi_k}{k+1} \prod_{k+2\le j\le n} \left(1 + \frac{\rho-1}{j}\right) \qquad (n \ge 0). \tag{7}$$

Proof. By (6),

$$ny_n = \rho \sum_{0\le j<n} y_j + \sum_{0\le j<n} \phi_j \qquad (n \ge 1).$$

Simplifying the difference $ny_n - (n-1)y_{n-1}$ leads to

$$y_n = \frac{n+\rho-1}{n} y_{n-1} + \frac{\phi_{n-1}}{n}.$$

Iterating the recurrence yields (7). ▮

Here and throughout this paper, all differential equations are to be interpreted in this way, although in many cases the differential equations are also analytically meaningful. This is the first step towards our general "asymptotic transfers" without any analytic conditions on the associated generating functions.

Formal solutions. For convenience, define, for $|z| < 1$,

$$I_\rho[\phi](z) = (1-z)^{-\rho} \int_0^z (1-x)^{\rho-1}\phi(x)\,\mathrm{d}x \qquad (\rho \in \mathbb{C}).$$

Then for $m \geq 1$

$$\begin{aligned} I_\rho^{[m]}[\phi](z) &:= I_\rho\left[I_\rho^{[m-1]}[\phi]\right](z) = \Bigg(\underbrace{I_\rho \circ I_\rho \cdots \circ I_\rho}_{m}\Bigg)[\phi](z) \\ &= (1-z)^{-\rho}\int_0^z (1-z_1)^{-1}\int_0^{z_1}\cdots(1-z_{m-1})^{-1}\int_0^{z_{m-1}}(1-z_m)^{\rho-1}\phi(z_m)\,\mathrm{d}z_m\cdots\mathrm{d}z_1. \end{aligned}$$

By applying successively the lemma, we obtain the following extension.

Corollary 1. *The function*

$$Y(z) = (1-z)^{-\rho}\sum_{0\leq j<m}(\vartheta-\rho)^j(Y)(0)\frac{1}{j!}\log^j\frac{1}{1-z} + I_\rho^{[m]}[\phi](z)$$

solves the linear differential equation

$$(\vartheta-\rho)^m Y(z) = \phi(z) \qquad (m \in \mathbb{N}). \tag{8}$$

The expression $I_\rho^{[m]}[\phi](z)$ can be simplified in certain special cases. Here we list some cases that are frequently encountered.

Corollary 2. *Assume $\eta, \rho \in \mathbb{C}$ and $\eta \neq \rho$. Then*

$$I_\rho\left[\frac{(1-z)^{-\eta}}{k!}\log^k\frac{1}{1-z}\right](z) = (1-z)^{-\eta}\sum_{0\leq j\leq k}\frac{(\eta-\rho)^{-k+1+j}}{j!}\log^j\frac{1}{1-z} - \frac{(1-z)^{-\rho}}{(\eta-\rho)^{k+1}}, \tag{9}$$

for $k \in \mathbb{N}$.

Proof. Apply iteratively the relation

$$\begin{aligned} I_\rho\left[(1-z)^{-\eta}\log^k\frac{1}{1-z}\right](z) &= \frac{1}{\eta-\rho}\frac{(1-z)^{-\eta}}{k!}\log^k\frac{1}{1-z} \\ &\quad - \frac{1}{\eta-\rho}I_\rho\left[(1-z)^{-\eta}\log^{k-1}\frac{1}{1-z}\right](z). \quad \blacksquare \end{aligned}$$

Corollary 3. *For $m \in \mathbb{N}$,*

$$I_\rho^{[m]}\left[(1-z)^{-\rho}\right](z) = \frac{1}{m!}(1-z)^{-\rho}\log^m\frac{1}{1-z}, \tag{10}$$

and

$$I_\rho^{[m]}\left[(1-z)^{-1}\right](z) = \frac{(1-z)^{-1}}{(1-\rho)^m} - (1-z)^{-\rho}\sum_{0\leq j<m}\frac{1}{j!}(1-\rho)^{-m+j}\log^j\frac{1}{1-z}. \tag{11}$$

Proof. The first identity (10) is obvious; the second follows from the relation

$$I_\rho^{[m]}\left[(1-z)^{-1}\right] = \frac{1}{(1-\rho)}\left(I_\rho^{[m-1]}\left[(1-z)^{-1}\right] - I_\rho^{[m-1]}[1]\right). \quad \blacksquare$$

Note that $I_\rho^{[m]}[c\phi] = cI_\rho^{[m]}[\phi]$ for any $c \in \mathbb{C}$.

General solutions to inhomogeneous equations. By successively applying (8) to (5) and by incorporating the general solution $\mathbf{y}(z)$ of $\Lambda_r(\vartheta)\mathbf{y} = 0$, we obtain the general solution to (5) as follows.

$$f(z) = \sum_{1\le k\le \kappa} \sum_{0\le j<m_\kappa} C_j(\rho_k)(1-z)^{-\rho_k}\log^j(1-z) + I_{\rho_\kappa}^{[m_\kappa]}\left[I_{\rho_{\kappa-1}}^{[m_{\kappa-1}]}\left[\cdots I_{\rho_1}^{[m_1]}[\phi]\cdots\right]\right](z), \tag{12}$$

where $C_j(\rho_k) \in \mathbb{C}$ are constants.

Main problem: determination of the constants. In theory, the constants $C_j(\rho)$ appearing in (12) can be determined by the initial conditions of the equation $f^{(j)}(0) = j!b_j, 0 \le j < r$. The basic problem is to determine these constants in effective and manageable forms so that they are as transparent as possible. But a straightforward determination of these constants always involves complicated computations, and thus we need a better approach.

2.2 Asymptotic transfers

We propose in this section a framework of asymptotic transfers. The main problem is, as described above, the determination of the constant in the dominant term in (12). We show that a suitable application of the method of linear operator leads to an effective form for the constant, as we did in [20]. Such an approach is very general, yet requiring almost no background on differential equations or matrix computations. It is thus elementary in nature, although we manipulate the calculations via generating functions.

We study recurrences of the form

$$a_n = \sum_{0\le k<n} a_k \sum_{0\le j<r} c_j \frac{j!\binom{k}{j}\binom{n-1-k}{r-1-j}}{r!\binom{n}{r}} + b_n, \tag{13}$$

for $n \ge r$, with $a_n := b_n$ for $0 \le n < r$. Here the c_j's are complex numbers such that $|c_1|+\cdots+|c_{r-1}| \ne 0$. Define

$$\Lambda_r(\vartheta) := \vartheta^{\overline{r}} - \sum_{0\le j<r} c_j \vartheta^{\overline{j}},$$

and arrange the zeros of $\Lambda_r(\vartheta)$ in decreasing order of their real parts:

$$\Re(\lambda_1) \ge \Re(\lambda_2) \ge \cdots \ge \Re(\lambda_r).$$

Theorem 1. *Assume that the zero λ_1 of the indicial polynomial $\Lambda_r(\vartheta)$ is simple and $\Re(\lambda_1) > \Re(\lambda_j)$ for all $2 \le j \le r$. Let $\nu = \Re(\lambda_1) > 0$.*

(i) If

$$b_n = o\left(n^{\nu-1}\right), \quad \text{and} \quad K_0 := \sum_{k\ge 0} \frac{k!\, b_k}{\lambda_1^{\overline{k+1}}} \text{ converges,} \tag{14}$$

then $a_n \sim K_1 n^{\lambda_1-1}$, where

$$K_1 := \frac{1}{\Lambda_r'(\lambda_1)\Gamma(\lambda_1)}\left(\lambda_1^{\overline{r}}K_0 - \sum_{0\le k<r-1}\frac{k!b_k}{\lambda_1^{\overline{k+1}}}\sum_{k<j<r} c_j\lambda_1^{\overline{j}}\right); \tag{15}$$

if $\sum_{0\le j<r} c_j\Gamma(\lambda_1+j)/\Gamma(\lambda_1+r) = 1$ then the conditions (14) are also necessary.

(*ii*) *If* $b_n = K_2\binom{n+\lambda_1-1}{n} + w_n$*, where* w_n *satisfies (14), then*

$$a_n \sim \frac{\lambda_1^{\overline{r}}}{\Lambda_r'(\lambda_1)\Gamma(\lambda_1)}\, n^{\lambda_1-1}\left(K_2 \sum_{0\le k<n} \frac{1}{\lambda_1+k} + K_3\right), \tag{16}$$

where

$$K_3 \;=\; \frac{1}{\lambda_1^{\overline{r}}} \sum_{0\le k<r} \frac{k!b_k}{\lambda_1^{\overline{k+1}}} \sum_{0\le j\le k} c_j \lambda_1^{\overline{j}} - K_2 \sum_{2\le j\le r} \frac{1}{\lambda_1-\lambda_j} + \sum_{k\ge r} \frac{k!\; w_k}{\lambda_1^{\overline{k+1}}}.$$

(*iii*) *If* $b_n \sim K_4 n^{\upsilon} \log^{\beta} n$*, where* $\Re(\upsilon) > \nu - 1$ *and* $\beta \in \mathbb{R}$*, then*

$$a_n \sim \frac{K_4\Gamma(\upsilon+r+1)}{\Lambda_r(\upsilon+1)\Gamma(\upsilon+1)}\, n^{\upsilon}\log^{\beta} n. \tag{17}$$

This theorem is a generalization of Proposition 7 of [20]. Note that K_0 converges iff $\int_0^1 (1-x)^{\lambda_1-1} g(x)\,\mathrm{d}x$ converges, where $g(z) := \sum_{n\ge 0} b_n z^n$, and that $a_n = o(n^{\nu-1})$ when $K_1 = 0$.

2.3 Proof of Theorem 1

Lemmas. Recall that $I_\rho[\phi](z) := (1-z)^{-\rho}\int_0^z (1-x)^{\rho-1}\phi(x)\,\mathrm{d}x$.

For convenience, let $\mathbf{Q}_\tau := \{Q(z) : [z^n]Q(z) = o(n^\tau)\}$, where $\tau \in \mathbb{R}$.

Lemma 3. *If* $Q(z) \in \mathbf{Q}_\tau$*, then* $I_\rho[Q](z) \in \mathbf{Q}_\tau$ *for* $\tau > \Re(\rho) - 1$*.*

Proof. Recall from (7) that

$$[z^n]I_\rho[Q](z) = \sum_{0\le k<n} \frac{[z^k]Q(z)}{k+1} \prod_{k+2\le j\le n}\left(1 + \frac{\rho-1}{j}\right),$$

since $y_0 = 0$ there. Now

$$\begin{aligned}\prod_{k+2\le j\le n}\left(1+\frac{\rho-1}{j}\right) &= \exp\left((\rho-1)\sum_{k+2\le j\le n}\frac{1}{j}\right)\prod_{k+2\le j\le n}\left(1+\frac{\rho-1}{j}\right)e^{-(\rho-1)/j} \\ &= O\left(n^{\Re(\rho)-1}(k+1)^{-\Re(\rho)+1}\right),\end{aligned}$$

uniformly for $0 \le k < n$. By applying this estimate to (7),

$$\begin{aligned}[z^n]I_\rho[Q](z) &= O\left(n^{\Re(\rho)-1}\sum_{1\le k\le n}\frac{|[z^k]Q(z)|}{k+1}\,k^{-\Re(\rho)+1}\right) \\ &= o\left(n^{\Re(\rho)-1}\sum_{1\le k\le n} k^{\tau-\Re(\rho)}\right) \\ &= o(n^\tau).\end{aligned}$$

Thus $I_\rho[Q](z) \in \mathbf{Q}_\tau$. ∎

Lemma 4. *If* $[z^n]Q(z) \sim cn^{\upsilon} \log^{\beta} n$*, where* $c \in \mathbb{C}$ *and* $\Re(\upsilon) > \Re(\rho) - 1$*, then*

$$[z^n]I_\rho[Q](z) \sim \frac{c}{\upsilon + 1 - \rho}\, n^{\upsilon} \log^{\beta} n.$$

If $[z^n]Q(z) \sim cn^{\rho-1} \log^{\beta} n$*, where* $\beta > -1$*, then*

$$[z^n]I_\rho[Q](z) \sim \frac{c}{\beta + 1}\, n^{\rho-1} \log^{\beta+1} n. \tag{18}$$

Proof. By (7) and Stirling's formula for the Gamma function,

$$\begin{aligned} [z^n]I_\rho[Q](z) &= \frac{\Gamma(n+\rho)}{n!} \sum_{0 \le k < n} \frac{k!}{\Gamma(k+\rho+1)}\,[z^k]Q(z) \\ &\sim cn^{\rho-1} \sum_{1 \le k \le n} k^{-\rho+\upsilon} \log^{\beta} k \\ &\sim \frac{c}{\upsilon + 1 - \rho}\, n^{\upsilon} \log^{\beta} n. \end{aligned}$$

The proof implicitly assumes that ρ is not a negative integer. If $\rho = -k$, where $k \in \mathbb{N}$, then the same estimate holds by simply dropping the first k terms in (7).

The proof of (18) is similar. ▮

Proof of Theorem 1. Case (i)**.**

Our starting point is the differential equation (see (4), (5))

$$\Lambda_r(\vartheta) f(z) = (1-z)^r g^{(r)}(z).$$

The proof is easier if one assumes that g admits analytic continuation in some extended region outside the unit disk since the cancellation caused by the factor $(1-z)^r$ can be easily incorporated. The situation becomes more intricate once this condition is dropped; see the discussions at the end of this section.

Contribution from the dominant zero λ_1**.** Define $\Omega_1(\vartheta) := \Lambda_r(\vartheta)/(\vartheta - \lambda_1)$. Then by Lemma 1,

$$\Omega_1(\vartheta) f(z) = \left(\Omega_1(\vartheta) f\right)(0)(1-z)^{-\lambda_1} + I_{\lambda_1}[(1-z)^r g^{(r)}](z). \tag{19}$$

By applying the elementary identity

$$\frac{\vartheta^{\overline{j}}}{\vartheta - \lambda_k} = \lambda_k^{\overline{j}} \sum_{0 \le s < j} \frac{\vartheta^{\overline{s}}}{\lambda_k^{\overline{s+1}}} + \frac{\lambda_k^{\overline{j}}}{\vartheta - \lambda_k} \qquad (k = 1, \dots, r;\, j = 0, 1, \dots),$$

we have

$$\Omega_1(\vartheta) = \sum_{0 \le k < r} \frac{\vartheta^{\overline{k}}}{\lambda_1^{\overline{k+1}}} \left(\lambda_1^{\overline{r}} - \sum_{k < j < r} c_j \lambda_1^{\overline{j}} \right).$$

Now applying this operator to f, using the relations

$$\vartheta^{\overline{j}} f(z)\big|_{z=0} = f^{(j)}(0) = j! b_j \qquad (0 \le j < r),$$

we obtain

$$C_1 := (\Omega_1(\vartheta)f)(0) = \sum_{0\le k<r} \frac{k!b_k}{\lambda_1^{\overline{k+1}}} \left(\lambda_1^{\overline{r}} - \sum_{k<j<r} c_j \lambda_1^{\overline{j}} \right). \tag{20}$$

Thus

$$\Omega_1(\vartheta)f(z) = C_1(1-z)^{-\lambda_1} + I_{\lambda_1}[(1-z)^r g^{(r)}](z). \tag{21}$$

Contribution from the second zero. By definition $\Omega_1(\vartheta) = \prod_{2\le j\le r}(\vartheta - \lambda_j)$. Define $\Omega_2(\vartheta) := \Omega_1(\vartheta)/(\vartheta - \lambda_2)$ (whether λ_2 is simple or not). Applying Lemma 1 to (21) using (9), we obtain

$$\Omega_2(\vartheta)f(z) = \frac{C_1}{\lambda_1 - \lambda_2}(1-z)^{-\lambda_1} + I_{\lambda_2}[I_{\lambda_1}[(1-z)^r g^{(r)}]](z) + Q_1(z),$$

where $Q_1(z) \in \mathbf{Q}_{\nu-1}$.

Contribution from the remaining zeros. Since the preceding analysis used only Lemma 1 and $\Re(\lambda_1) > \Re(\lambda_2)$, it can be further applied to other zeros no matter they are simple or not. By repeating the same argument, we obtain

$$f(z) = \frac{C_1(1-z)^{-\lambda_1}}{(\lambda_1-\lambda_2)\cdots(\lambda_1-\lambda_r)} + I_{\lambda_r}\left[\cdots I_{\lambda_1}[(1-z)^r g^{(r)}]\cdots\right](z) + Q_2(z), \tag{22}$$

where $Q_2(z) \in \mathbf{Q}_{\nu-1}$.

It remains to simplify the compositional term

$$U_r(\lambda_1,\dots,\lambda_r)(z) := I_{\lambda_r}\left[\cdots I_{\lambda_1}[(1-z)^r g^{(r)}]\cdots\right](z) \qquad (r \ge 1).$$

We prove by induction that

$$U_r(\lambda_1,\dots,\lambda_r)(z) = \frac{\lambda_1^{\overline{r}} I_{\lambda_1}[g](z)}{(\lambda_1-\lambda_2)\cdots(\lambda_1-\lambda_r)} - \lambda_1^{\overline{r}} \sum_{0\le k<r} \frac{k!b_k}{\lambda_1^{\overline{k+1}}}(1-z)^{-\lambda_1} + Q_3(z), \tag{23}$$

where $Q_3(z) \in \mathbf{Q}_{\nu-1}$.

First induction. Before proving (23), we first show that

$$U_k(\lambda_{i_1},\dots,\lambda_{i_k})(z) \in \mathbf{Q}_{\nu-1}, \tag{24}$$

for any subset $\{i_1,\dots,i_k\} \subset \{2,\dots,r\}$ and for any $1 \le k < r$.

First if $k=1$ then by integration by parts

$$\begin{aligned} U_1(\lambda_{i_1})(z) &= I_{\lambda_{i_1}}[(1-z)g'](z) \\ &= g(z) - g(0)(1-z)^{-\lambda_{i_1}} + \lambda_{i_1} I_{\lambda_{i_1}}[g](z) \\ &\in \mathbf{Q}_{\nu-1}, \end{aligned}$$

where we used again Lemma 3.

For $k \ge 2$, again by integration by parts,

$$\begin{aligned} U_k(\lambda_{i_1},\dots,\lambda_{i_k})(z) &= U_{k-1}(\lambda_{i_2},\dots,\lambda_{i_k})(z) - g^{(k-1)}(0) I_{\lambda_{i_k}}\left[\cdots I_{\lambda_{i_2}}[(1-z)^{-\lambda_{i_1}}]\cdots\right] \\ &\quad + (\lambda_{i_1}+k-1) I_{\lambda_{i_k}}\left[U_{k-1}(\lambda_{i_1},\dots,\lambda_{i_{k-1}})\right](z). \end{aligned}$$

The second term on the right-hand side is easily seen to be in $\mathbf{Q}_{\nu-1}$ by (9). For the other two terms, we have, by induction, $U_{k-1}(\lambda_{i_2},\dots,\lambda_{i_k})(z) \in \mathbf{Q}_{\nu-1}$ and $U_{k-1}(\lambda_{i_1},\dots,\lambda_{i_{k-1}})(z) \in \mathbf{Q}_{\nu-1}$; so that the last term is also in $\mathbf{Q}_{\nu-1}$ by Lemma 3. This proves (24).

Proof of (23) by induction. If $r = 1$, then, similarly as above,

$$\begin{aligned} U_1(\lambda_1)(z) &= g(z) - g(0)(1-z)^{-\lambda_1} + \lambda_1 I_{\lambda_1}[g](z) \\ &= \lambda_1 I_{\lambda_1}[g](z) - b_0(1-z)^{-\lambda_1} + Q_4(z), \end{aligned}$$

where $Q_4(z) \in \mathbf{Q}_{\nu-1}$. Thus (23) holds for $r = 1$.

For $r \geq 2$, by induction, (9) and (24),

$$\begin{aligned} U_r(\lambda_1, \dots, \lambda_r)(z) &= U_{r-1}(\lambda_2, \dots, \lambda_r)(z) - g^{(r-1)}(0) I_{\lambda_r}\left[\cdots I_{\lambda_2}[(1-z)^{-\lambda_1}]\cdots\right] \\ &\quad + (\lambda_1 + r - 1) I_{\lambda_r}\left[U_{r-1}(\lambda_1, \dots, \lambda_{r-1})\right](z) \qquad (25) \\ &= -\frac{g^{(r-1)}(0)(1-z)^{-\lambda_1}}{(\lambda_1 - \lambda_2)\cdots(\lambda_1 - \lambda_{r-1})} + \frac{(\lambda_1 + r - 1)\lambda_1^{\overline{r-1}}}{(\lambda_1 - \lambda_2)\cdots(\lambda_1 - \lambda_{r-1})} I_{\lambda_r}[I_{\lambda_1}[g]](z) \\ &\quad - \frac{(\lambda_1 + r - 1)\lambda_1^{\overline{r-1}}}{(\lambda_1 - \lambda_2)\cdots(\lambda_1 - \lambda_{r-1})} \sum_{0 \leq k < r-1} \frac{k! b_k}{\lambda_1^{\overline{k+1}}} I_{\lambda_r}[(1-z)^{-\lambda_1}](z) + Q_5(z), \end{aligned}$$

where $Q_5(z) \in \mathbf{Q}_{\nu-1}$. Simplifying the last expression yields (23).

Collecting all contributions. Noting that

$$(\lambda_1 - \lambda_2)\cdots(\lambda_1 - \lambda_r) = \Lambda'(\lambda_1),$$

we thus obtain

$$f(z) = \frac{\lambda_1^{\overline{r}}}{\Lambda_r'(\lambda_1)} \left(I_{\lambda_1}[g](z) - (1-z)^{-\lambda_1} \sum_{0 \leq k < r-1} \frac{k! b_k}{\lambda_1^{\overline{r}} \cdot \lambda_1^{\overline{k+1}}} \sum_{k < j < r} c_j \lambda_1^{\overline{j}} \right) + Q_6(z), \qquad (26)$$

where $Q_6(z) \in \mathbf{Q}_{\nu-1}$.

Now by (7), we have

$$\begin{aligned} [z^n] I_{\lambda_1}[g](z) &= \sum_{0 \leq k < n} \frac{b_k}{k+1} \prod_{k+2 \leq j \leq n} \left(1 + \frac{\lambda_1 - 1}{j}\right) \\ &= \frac{\Gamma(\lambda_1 + n)}{\Gamma(\lambda_1)\, n!} \sum_{0 \leq k < n} \frac{k!\, b_k}{\lambda_1^{\overline{k+1}}} \\ &= \frac{n^{\lambda_1 - 1}}{\Gamma(\lambda_1)} \sum_{k \geq 0} \frac{k!\, b_k}{\lambda_1^{\overline{k+1}}} + o(n^{\nu-1}), \qquad (27) \end{aligned}$$

since we assume that the series $\sum_k b_k k!/\lambda_1^{\overline{k+1}}$ converges. Thus (15) follows and this proves the sufficiency part of (*i*).

Necessary part of (*i*)**.** Observe that the result (26) relies only on the fact that $b_n = o(n^{\nu-1})$ and the condition that K_0 converges was used only in deriving the asymptotics of $[z^n]I_{\lambda_1}[g](z)$. We assume that $\sum_{0 \leq j < r} c_j \Gamma(\lambda_1 + j)/\Gamma(\lambda_1 + r) = 1$ and $a_n \sim K' n^{\lambda_1 - 1}$ for some K'. Then by the recurrence (13), we have

$$\begin{aligned} b_n &= a_n - \sum_{0 \leq j < r} c_j \frac{j!}{r!} \sum_{0 \leq k < n} \frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}} a_k \\ &= K n^{\lambda_1 - 1} + o(n^{\nu-1}) \\ &\quad - K \sum_{0 \leq j < r} c_j \frac{j!}{r!} \sum_{0 \leq k < n} \frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}} \left(\Gamma(\lambda_1) \binom{k + \lambda_1 - 1}{k} + o((k+1)^{\nu-1}) \right). \end{aligned}$$

Now

$$\begin{aligned}\sum_{0\le k<n}\binom{k}{j}\binom{\lambda_1+k-1}{k}\binom{n-1-k}{r-1-j} &= \frac{1}{j!}[z^{n-1}]\frac{\mathrm{d}^j}{\mathrm{d}z^j}(1-z)^{-\lambda_1}\cdot\frac{z^{r-1}}{(1-z)^{r-j}}\\ &\sim \frac{\lambda_1^{\overline{j}}}{j!\Gamma(\lambda_1+r)}\,n^{\lambda_1+r-1};\end{aligned}$$

so that

$$\Gamma(\lambda_1)\sum_{0\le k<n}\frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}}\binom{k+\lambda_1-1}{k}\sim\frac{r!\Gamma(\lambda_1+j)}{j!\Gamma(\lambda_1+r)}\,n^{\lambda_1-1}.$$

It remains to estimate the error term, which satisfies

$$\begin{aligned}o(1)\sum_{0\le k<n}\frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}}(k+1)^{\nu-1} &= o(1)n^{-r}\sum_{0\le k<n}k^{j+\nu-1}(n-k)^{r-1-j}\\ &= o(n^{\nu-1}),\end{aligned}$$

implying that $b_n=o(n^{\nu-1})$ by $\sum_{0\le j<r}c_j\Gamma(\lambda_1+j)/\Gamma(\lambda_1+r)=1$.

From this estimate, we obtain, by (26) and (27), $K'=K_1$ and $|K_0|<\infty$.

Proof of Theorem 1. Case ***(ii)*****.**

The proof is almost the same as that for case (*i*); we sketch the main differences. First, denote by $w(z):=\sum_{n\ge0}w_jz^j$; then

$$(1-z)^rg^{(r)}(z)=K_2\lambda_1^{\overline{r}}(1-z)^{-\lambda_1}+(1-z)^rw^{(r)}(z).$$

From Lemma 1, it follows that

$$\Omega_1(\vartheta)f(z)=K_2\lambda_1^{\overline{r}}(1-z)^{-\lambda_1}\log\frac{1}{1-z}+\left(\Omega_1(\vartheta)f\right)(0)(1-z)^{-\lambda_1}+I_{\lambda_1}\left[(1-z)^rw^{(r)}\right](z).$$

By the same argument as in Case (*i*), the cumulative contribution of the last two terms on the right-hand side (from the remaining zeros) to $f(z)$ is given by

$$\frac{(1-z)^{-\lambda_1}}{\Lambda_r'(\lambda_1)}\left(\sum_{0\le k<r}\frac{k!b_k}{\lambda_1^{\overline{k+1}}}\sum_{0\le j\le k}c_j\lambda_1^{\overline{j}}-\lambda_1^{\overline{r}}\sum_{0\le j<r}\frac{j!w_j}{\lambda_1^{\overline{j+1}}}\right)+\lambda_1^{\overline{r}}I_{\lambda_1}[w](z)+Q_7(z),$$

where $Q_7(z)\in\mathbf{Q}_{\nu-1}$.

By arguing similarly as above, we deduce that the contribution of the term $-K_2\lambda_1^{\overline{r}}(1-z)^{-\lambda_1}\log(1-z)$ to $f(z)$ is given by

$$\frac{K_2\lambda_1^{\overline{r}}}{\Lambda_r'(\lambda_1)}(1-z)^{-\lambda_1}\log\frac{1}{1-z}-\frac{K_2\lambda_1^{\overline{r}}(1-z)^{-\lambda_1}}{\Lambda_r'(\lambda_1)}\sum_{2\le j\le r}\frac{1}{\lambda_1-\lambda_j}+Q_8(z),$$

where $Q_8(z)\in\mathbf{Q}_{\nu-1}$ and we used the identity

$$(1-z)^{-\rho}\int_0^z(1-x)^{\rho-\lambda_1-1}\log\frac{1}{1-x}\,\mathrm{d}x=\frac{(1-z)^{-\lambda_1}}{\lambda_1-\rho}\log\frac{1}{1-z}-\frac{(1-z)^{-\lambda_1}}{(\lambda_1-\rho)^2}+\frac{(1-z)^{-\rho}}{(\lambda_1-\rho)^2},$$

for $\Re(\rho) < \Re(\lambda_1)$.

Adding both contributions leads to

$$f(z) = \frac{K_2 \lambda_1^{\overline{r}}}{\Lambda_r'(\lambda_1)} (1-z)^{-\lambda_1} \log \frac{1}{1-z} + \frac{K_7 \lambda_1^{\overline{r}}}{\Lambda_r'(\lambda_1)} (1-z)^{-\lambda_1} + \frac{\lambda_1^{\overline{r}}}{\Lambda_r'(\lambda_1)} I_{\lambda_1}[w](z) + Q_9(z),$$

where $Q_9(z) \in \mathbf{Q}_{\nu-1}$ and

$$K_7 \quad = \quad \frac{1}{\lambda_1^{\overline{r}}} \sum_{0\le k<r} \frac{k! b_k}{\lambda_1^{\overline{k+1}}} \sum_{0\le j\le k} c_j \lambda_1^{\overline{j}} - \sum_{0\le j<r} \frac{j! w_j}{\lambda_1^{\overline{j+1}}} - K_2 \sum_{2\le j\le r} \frac{1}{\lambda_1 - \lambda_j}.$$

From the identity

$$[z^n](1-z)^{-\lambda_1} \log \frac{1}{1-z} = \frac{\Gamma(\lambda_1 + n)}{\Gamma(\lambda_1)\Gamma(n+1)} \sum_{0\le j<n} \frac{1}{\lambda_1 + j},$$

(16) follows.

Proof of Theorem 1. Case (*iii*).

The proof is almost the same as in Case (*i*). We start from (22) and we need only to prove that

$$[z^n] U_r(\lambda_1, \dots, \lambda_r)(z) \sim \frac{K_4 \Gamma(\upsilon + r + 1)}{(\upsilon + 1 - \lambda_1) \cdots (\upsilon + 1 - \lambda_r) \Gamma(\upsilon + 1)} n^{\upsilon} \log^{\beta} n. \tag{28}$$

First if $r = 1$, then, by Lemma 4,

$$\begin{aligned} [z^n] U_1(\lambda_1)(z) &= b_n - b_0 (1-z)^{-\lambda_1} + \lambda_1 [z^n] I_{\lambda_1}[g](z) \\ &\sim K_4 \left(1 + \frac{\lambda_1}{\upsilon + 1 - \lambda_1}\right) n^{\upsilon} \log^{\beta} n. \end{aligned}$$

So (28) holds for $r = 1$.

For $r \ge 2$, (28) follows by induction and (25).

This proves Case (*iii*) and completes the proof of Theorem 1.

A special case. In the special case when $\lambda_1 = 2$, the conditions (14) for asymptotic linearity of a_n become

$$b_n = o(n), \quad \text{and} \quad \sum_{k\ge 0} \frac{b_k}{(k+1)(k+2)} \quad \text{converges.}$$

Interestingly, in addition to [20, 39, 62], such conditions also appeared in other problems related to divide-and-conquer; see [28, 51, 52, 63].

Further refinements for small "toll functions". The asymptotic estimate $a_n \sim K_1 n^{\lambda_1 - 1}$ in Theorem 1 can be further improved if more asymptotics of b_n is known. For example, if $b_n = O(n^{\tau})$, where $\tau < \nu - 1$, then (assuming that λ_1 is simple)

$$a_n - K_1 n^{\lambda_1 - 1} = \begin{cases} O(n^{\tau}), & \text{if } \tau > \alpha - 1; \\ O(n^{\alpha-1} (\log n)^{m_0}), & \text{if } \tau = \alpha - 1; \\ O(n^{\alpha-1} (\log n)^{m_0 - 1}), & \text{if } \tau < \alpha - 1, \end{cases}$$

where m_0 denotes the largest multiplicity of zeros in $\{\lambda_j : \Re(\lambda_j) = \alpha\}$. Further refinements in the second and third cases are also possible without additional conditions on b_n; such results will be needed in Section 3.

Analytic heuristics. The proof of the theorem can be largely simplified if we assume that g can be analytically continued outside the unit disk and then apply the singularity analysis of Flajolet and Odlyzko [43]. We briefly consider Case (*iii*). So, by assumptions, we have (see [43])

$$g(z) \sim K_4 \Gamma(\upsilon+1)(1-z)^{-\upsilon-1} \log^{\beta} \frac{1}{1-z},$$

as $z \sim 1$. Then, by the assumption of analytic continuation,

$$(1-z)^r g^{(r)}(z) \sim K_4 \Gamma(\upsilon+r+1)(1-z)^{-\upsilon-1} \log^{\beta} \frac{1}{1-z},$$

and it follows, by singularity analysis, that

$$[z^n](1-z)^r g^{(r)}(z) \sim K_4 \frac{\Gamma(\upsilon+r+1)}{\Gamma(\upsilon+1)} n^{\upsilon} \log^{\beta} n.$$

Then (17) is obtained by applying successively Lemma 4.

Such an analytic approach based on singularity analysis is very efficient in computing the constants and the price is that we need stronger assumptions that may not be easy to be justified. On the other hand, our elementary approach imposes no analytic continuation, although it requires more analysis to handle the cancellations of coefficients (caused by $(1-z)^r$). Also our approach is easily applicable to huge "toll functions" such as $b_n \sim n!/\sqrt{n}$ for which the analytic approach fails.

2.4 Non-simple dominant zero

We strengthen Theorem 1 by considering the case when λ_1 is not simple.

Theorem 1'. *Assume that* $\lambda_1 = \cdots = \lambda_m$ *and* $\Re(\lambda_m) > \Re(\lambda_j)$ *for all* $m+1 \le j \le r$. *Let* $\nu = \Re(\lambda_1) > 0$.

(*i*) *If* b_n *satisfies*

$$b_n = o(n^{\nu-1}(\log n)^{m-1}), \quad \text{and} \quad \sum_{k\ge 0} \frac{k! b_k}{\lambda_1^{\overline{k+1}}} \text{ converges}, \tag{29}$$

then $a_n \sim K_1' n^{\lambda_1 - 1} \log^{m-1} n$, *where*

$$K_1' := \frac{m}{\Lambda_r^{(m)}(\lambda_1)\Gamma(\lambda_1)} \left(\lambda_1^{\overline{r}} K_0 - \sum_{0\le k<r-1} \frac{k! b_k}{\lambda_1^{\overline{k+1}}} \sum_{k<j<r} c_j \lambda_1^{\overline{j}} \right); \tag{30}$$

(*ii*) *If* $b_n \sim K_2' n^{\lambda_1 - 1} \log^{\beta} n$, *where* $\beta > -1$, *then*

$$a_n \sim \frac{K_2' m! \Gamma(\lambda_1 + r)}{\Lambda_r^{(m)}(\lambda_1)\Gamma(\lambda_1)(\beta+1)\cdots(\beta+m)} n^{\lambda_1 - 1} \log^{\beta+m} n.$$

(*iii*) *If* $b_n \sim K_4 n^{\upsilon} \log^{\beta} n$, *where* $\Re(\upsilon) > \nu - 1$ *and* $\beta \in \mathbb{R}$, *then* a_n *satisfies (17).*

Proof. The proof is similar to that of Theorem 1 with suitable modifications. We start from (21)

$$\Omega_1(\vartheta) f(z) = C_1 (1-z)^{-\lambda_1} + I_{\lambda_1}[(1-z)^r g^{(r)}](z),$$

where the constant C_1 is given in (20).

Applying Lemma 1 yields

$$\Omega_2(\vartheta)f(z) = C_2(1-z)^{-\lambda_1} + C_1(1-z)^{-\lambda_1}\log\frac{1}{1-z} + I_{\lambda_1}^{[2]}[(1-z)^r g^{(r)}](z),$$

where C_2 can be computed if desired.

Iterating this argument for another $m-1$ times, we obtain

$$\begin{aligned}\Omega_m(\vartheta)f(z) &= \sum_{0\le j\le m-2}\frac{C_{m-j}}{j!}(1-z)^{-\lambda_1}\left(\log\frac{1}{1-z}\right)^j \\ &\quad + \frac{C_1}{(m-1)!}(1-z)^{-\lambda_1}\left(\log\frac{1}{1-z}\right)^{m-1} + I_{\lambda_1}^{[m]}[(1-z)^r g^{(r)}](z),\end{aligned}$$

where $\Omega_m(\vartheta) := \Lambda_r(\vartheta)/(\vartheta-\lambda_1)^m$ and C_j, $0\le j\le m-2$, are constants.

Then the remaining zeros, no matter simple or not, will only change the constants but not the dominant order, and add additional negligible terms. Let

$$\mathbf{Q}_{\tau,\beta} := \{Q(z) \,:\, [z^n]Q(z) = o(n^\tau(\log n)^\beta)\} \qquad (\tau,\beta\in\mathbb{R}).$$

The following lemma is easily extended from Lemma 3.

Lemma 5. *If* $\tau > \Re(\rho)-1$ *and* $Q(z)\in\mathbf{Q}_{\tau,\beta}$, *where* $\tau,\beta\in\mathbb{R}$, *then* $I_\rho[Q](z)\in\mathbf{Q}_{\tau,\beta}$.

Proof. Omitted. ▮

We then obtain

$$\begin{aligned}f(z) &= \frac{C_1}{(m-1)!(\lambda_1-\lambda_{m+1})\cdots(\lambda_1-\lambda_r)}(1-z)^{-\lambda_1}\left(\log\frac{1}{1-z}\right)^{m-1} \\ &\quad + I_{\lambda_r}[\cdots[I_{\lambda_1}^{[m]}[(1-z)^r g^{(r)}]\cdots](z) + Q_{10}(z), \qquad (31)\end{aligned}$$

where $Q_{10}(z)\in\mathbf{Q}_{\nu-1,m-1}$ by Lemma 5. Now

$$(\lambda_1-\lambda_{m+1})\cdots(\lambda_1-\lambda_r) = \frac{\Lambda_r^{(m)}(\lambda_1)}{m!},$$

so we are left, again as in Case (i), Theorem 1, with the compositional term

$$I_{\lambda_r}[\cdots[I_{\lambda_1}^{[m]}[(1-z)^r g^{(r)}]\cdots](z).$$

This term is treated similarly as before. First we have, by (10),

$$\begin{aligned}I_{\lambda_1}^{[m]}[(1-z)^r g^{(r)}](z) &= (\lambda_1+r-1)I_{\lambda_1}^{[m]}[(1-z)^{r-1}g^{(r-1)}](z) + I_{\lambda_1}^{[m-1]}[(1-z)^{r-1}g^{(r-1)}](z) \\ &\quad - g^{(r-1)}(0)I_{\lambda_1}^{[m-1]}[(1-z)^{-\lambda_1}](z) \\ &= (\lambda_1+r-1)I_{\lambda_1}^{[m]}[(1-z)^{r-1}g^{(r-1)}](z) + I_{\lambda_1}^{[m-1]}[(1-z)^{r-1}g^{(r-1)}](z) \\ &\quad - \frac{g^{(r-1)}(0)}{(m-1)!}(1-z)^{-\lambda_1}\left(\log\frac{1}{1-z}\right)^{m-1} \\ &= \lambda_1^{\overline{r}} I_{\lambda_1}^{[m]}[g](z) + \sum_{0\le j<r}\frac{\lambda_1^{\overline{r}}}{\lambda_1^{\overline{j+1}}} I_{\lambda_1}^{[m-1]}[(1-z)^j g^{(j)}](z) \\ &\quad - \left(\sum_{0\le j<r}\frac{\lambda_1^{\overline{r}}}{\lambda_1^{\overline{j+1}}}g^{(j)}(0)\right)\frac{(1-z)^{-\lambda_1}}{(m-1)!}\left(\log\frac{1}{1-z}\right)^{m-1},\end{aligned}$$

for $m \geq 1$, where $I_{\lambda_1}^{[0]}[f](z) = f(z)$.

Thus

$$\begin{aligned} I_{\lambda_r}[\cdots[I_{\lambda_1}^{[m]}[(1-z)^r g^{(r)}]\cdots](z) &= \lambda_1^{\bar{r}} I_{\lambda_r}[\cdots[I_{\lambda_1}^{[m]}[g]\cdots] + Q_{11}(z) \\ &- \left(\sum_{0\le j<r} \frac{\lambda_1^{\bar{r}} j! b_j}{\lambda_1^{j+1}(m-1)!}\right) I_{\lambda_r}\left[\cdots I_{\lambda_{m+1}}\left[(1-z)^{-\lambda_1}\left(\log\frac{1}{1-z}\right)^{m-1}\right]\cdots\right](z), \end{aligned} \tag{32}$$

where

$$Q_{11}(z) := \sum_{0\le j<r} \frac{\lambda_1^{\bar{r}}}{\lambda_1^{j+1}} I_{\lambda_r}\left[\cdots\left[I_{\lambda_1}^{[m-1]}[(1-z)^j g^{(j)}]\right]\cdots\right](z).$$

By arguing similarly as in the proof of Case (i), Theorem 1, we deduce that

$$Q_{11}(z) \in \mathbf{Q}_{\nu-1,m-1}.$$

On the other hand, by applying successively (9),

$$\begin{aligned} I_{\lambda_r}&\left[\cdots I_{\lambda_{m+1}}\left[(1-z)^{-\lambda_1}\left(\log\frac{1}{1-z}\right)^{m-1}\right]\cdots\right](z) \\ &= \frac{(1-z)^{-\lambda_1}}{(\lambda_1-\lambda_{m+1})\cdots(\lambda_1-\lambda_r)}\left(\log\frac{1}{1-z}\right)^{m-1} + Q_{12}(z), \end{aligned} \tag{33}$$

where $Q_{12}(z) \in \mathbf{Q}_{\nu-1,m-1}$.

It remains to derive an asymptotic approximation for the coefficient

$$[z^n]I_{\lambda_r}[\cdots[I_{\lambda_1}^{[m]}[g]\cdots](z).$$

By (14), $|K_0| < \infty$; so that

$$[z^n]I_{\lambda_1}[g](z) \sim \frac{K_0}{\Gamma(\lambda_1)}\, n^{\lambda_1},$$

Applying I_{λ_1} once again and (18), we have

$$[z^n]I_{\lambda_1}^{[2]}[g](z) \sim \frac{K_0}{\Gamma(\lambda_1)}\, n^{\lambda_1-1}\log n.$$

And in general

$$[z^n]I_{\lambda_1}^{[m]}[g](z) \sim \frac{K_0}{\Gamma(\lambda_1)(m-1)!}\, n^{\lambda_1-1}(\log n)^{m-1}.$$

The remaining zeros change the coefficients as above, and we obtain

$$[z^n]I_{\lambda_r}[\cdots[I_{\lambda_1}^{[m]}[g]\cdots](z) \sim \frac{mK_0}{\Gamma(\lambda_1)\Lambda_r^{(m)}(\lambda_1)}\, n^{\lambda_1-1}(\log n)^{m-1}. \tag{34}$$

Collecting the contributions from (31), (32), (33) and (34), we complete the proof of Case (i). The remaining proof is similar and is omitted. ▌

3 Method of moments and phase changes

The method of moments for recursively defined random variables can be summarized as follows. Assume that $P_n(y)$ denotes the probability (or moment) generating function of the random variable, say X_n in question.

- Assume that
$$P_n(y) = \Phi[P_0, \dots, P_{n-1}; Q_0, \dots, Q_n](y),$$
for some functional Φ and for given $Q_0(y), \dots, Q_n(y)$.
- Derive first the asymptotics of the mean and/or the variance.
- Consider the probability generating function of a suitably normalized random variable (by exact mean or asymptotic mean), which satisfies the same type of recurrence. From this derive the recurrence of higher centralized moments.
- Find the asymptotics of higher centralized moments and check if the limiting sequence uniquely characterizes a distribution.
- Conclude the convergence in distribution by the moment convergence theorem of Frechet and Shohat [47].

A feature of this method of moments is that, since the recurrences of moments (centered or non-centered) are of the same form, all asymptotic analysis required for solving the recurrence can be encapsulated into the same framework, which we call "asymptotic transfers" that bridges the asymptotics of the "toll sequence" ("inhomogeneous part" of the recurrence) to that of the recurrence.

For other approaches to recursively defined random variables, see [29, 30, 67, 89, 92, 97].

3.1 A class of quicksort.

Recurrence. We consider in this section the following simple recurrence:

$$X_n \stackrel{d}{=} X_{I_n} + X^*_{n-1-I_n} + b_n \qquad (n \ge r), \tag{35}$$

with X_n suitably defined for $n < r$, where the X_n's and the X_n^*'s are independent and have the same distribution,

$$P(I_n = k) = \pi_{n,k} := \sum_{0 \le j < r} p_j \frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}}, \qquad \sum_{0 \le j < r} p_j = 1 \qquad (0 \le k < n),$$

and the b_n's are known "toll functions". The random variable X_n corresponds to the cost of quicksort described in Introduction when sorting a random permutation.

For simplicity, we consider in details only the case

$$\begin{cases} X_n \stackrel{d}{=} X_{I_n} + X^*_{n-1-I_n} + 1, & \text{if } n \ge r; \\ X_n = 0, & \text{if } n < r, \end{cases} \tag{36}$$

focusing on the change of the limit laws. Such a recurrence corresponds to the number of times the r-sample-partitioning is used throughout the sorting process. By definition, it is easy to see that $X_n \equiv n$ iff $r = 1$. So we assume throughout this section that $r \ge 2$.

Motivations. Although from an algorithmic point of view, this cost measure is less important than the major cost measures like the number of comparisons and the number of exchanges, we study this variable for several reasons. First, the derivation of the limit law of the major cost measures is more or less the same as the linear cost measures once we have available the asymptotic transfers. Second, the more interesting phase change phenomenon for X_n does not occur for the major cost measures; for linear cost measures such as X_n, we can easily produce different phase changes by simply varying the underlying distribution $\{p_j\}$, these phase changes suggesting further questions. Third, the new ingredient of our analysis is that we do not impose any condition on the second largest zero(s), while previous applications of the method of moments for similar problems necessitate that the second largest zero(s) (in real parts) are simple. Fourth, this investigation was motivated by the analysis of quicksort using ninther or remedians, the underlying splitting probabilities being more complicated than previous variants of quicksort. The reason to introduce more general framework is to capture more regularities for more classes of quicksort.

A description in terms of growing BSTs. The above quicksort can be described alternatively in terms of a tree-growing process. So we are inserting keys one after another into an initially empty tree. If the number of keys is less than r, then the tree consists of a node with all keys in it. When the r-th key is inserted, the tree is split into a root and two subtrees: the root holds the j-th smallest key, say x, in this r-sample with probability p_{j-1}, the left (right) subtree then is a node containing all keys less (greater) than x, respectively. The $(r+1)$-st key is inserted into the left subtree with probability $j/(r+1)$ and into the right subtree with probability $(r+1-j)/(r+1)$. And the process continues. Each time the number of keys in a node equals r, a split as above is performed. See [7, 78, 94] for similar cases. We will give another description in terms of urn models in the next section.

Moment generating function and differential equation. In terms of moment generating functions $P_n(y) := E(e^{X_n y})$, we have $P_n(y) = 1$ for $0 \le n < r$, and

$$P_n(y) = e^y \sum_{0 \le k < n} \pi_{n,k} P_k(y) P_{n-1-k}(y) \qquad (n \ge r).$$

The expected cost $E(X_n)$ thus satisfies the recurrence $a_n = 0$ for $n < r$ and

$$a_n = \sum_{0 \le k < n} (\pi_{n,k} + \pi_{n,n-1-k}) a_k + 1 \qquad (n \ge r),$$

so that the corresponding generating functions satisfies

$$f^{(r)}(z) = r! \sum_{0 \le j < r} \frac{p_j + p_{r-1-j}}{j!} (1-z)^{-r+j} f^{(j)}(z) + r!(1-z)^{-r-1}, \tag{37}$$

with the initial conditions $f^{(j)}(0) = 0$ for $0 \le j < r$. Also the indicial equation is given by

$$\Lambda_r(\vartheta) = \vartheta^{\overline{r}} - r! \sum_{0 \le j < r} (p_j + p_{r-1-j}) \frac{\vartheta^{\overline{j}}}{j!} = 0.$$

Let again the r zeros of $\Lambda_r(\vartheta)$ be arranged according to their real part:

$$\Re(\lambda_1) \ge \Re(\lambda_2) \ge \Re(\lambda_3) \ge \cdots \ge \Re(\lambda_r).$$

Obviously, $\Lambda(2) = 0$, so that $\lambda_1 = 2$. We will show that all other zeros are less than 2 in real parts.

Let $\alpha = \Re(\lambda_2)$ denote the real part of the second largest zero(s) (in real part). Define

$$\mathscr{Z}_1 := \{\lambda_j : \Re(\lambda_j) < 2 \text{ and } \Re(\lambda_j) \geq \Re(\lambda_k) \text{ for } k = 2, \ldots, r\},$$

to be the multiset of all λ_j's whose real part is α.

To state our main result, we also need to define two integers: m is the multiplicity of α if $\Lambda_r(\alpha) = 0$, and $m = 0$ otherwise; q is the largest multiplicity of the complex zeros in $\mathscr{Z}_1$. Let $N(0,1)$ denote a normal random variable with zero mean and unit variance.

3.2 Phase changes of the limit laws.

Theorem 2. *(i) If $\alpha < 3/2$, then*

$$\frac{X_n - \mu n}{\sigma_1 \sqrt{n}} \xrightarrow{d} N(0,1),$$

where the symbol $\xrightarrow{d}$ denotes convergence in distribution,

$$\mu = \frac{1}{(r+1)H_{r+1} - \sum_{0 \leq j < r} p_j \left((j+1)H_{j+1} + (r-j)H_{r-j}\right)}, \tag{38}$$

and σ_1 is a positive constant defined in (44);

(ii) if $\alpha = 3/2$ and $m \geq q$, then

$$\frac{X_n - \mu n}{\sigma_2 \sqrt{n} (\log n)^{m-1/2}} \xrightarrow{d} N(0,1),$$

where $\sigma_2 > 0$ is a constant given in (45); if $\alpha = 3/2$ and $q > m$, then the limiting distribution of $(X_n - \mu n)/(\sqrt{n}(\log n)^{q-1})$ does not exist;

(iii) if $\alpha > 3/2$ and $m > q$, then

$$\frac{X_n - \mu n}{\sigma_3 n^{\alpha-1} (\log n)^{m-1}} \xrightarrow{d} X,$$

where σ_3 is given in (46), X is uniquely characterized by its moments and $E(X^h)$ is given in (47); if $\alpha > 3/2$ and $q \geq m$, then $(X_n - \mu n)/(n^{\alpha-1}(\log n)^{q-1})$ does not exist.

In each case, $E(X_n) \sim \mu n$ and

$$\mathrm{Var}(X_n) \sim \begin{cases} \sigma_1^2 n, & \text{if } \alpha < 3/2; \\ \sigma_2^2 n (\log n)^{2m-1}, & \text{if } \alpha = 3/2 \text{ and } m \geq q; \\ \Psi_1(n) n (\log n)^{2q-2}, & \text{if } \alpha = 3/2 \text{ and } q > m; \\ \sigma_3^2 n^{2\alpha-2} (\log n)^{2m-2}, & \text{if } \alpha > 3/2 \text{ and } m > q; \\ \Psi_2(n) n^{2\alpha-2} (\log n)^{2q-2}, & \text{if } \alpha > 3/2 \text{ and } q \geq m, \end{cases} \tag{39}$$

where the σ_j's are positive constants and $\Psi_1(n), \Psi_2(n)$ are bounded periodic functions.

3.3 Proof of Theorem 2

Properties of the zeros. First $\Lambda_r(2) = 0$ since

$$(r+1)! = r! \sum_{0 \leq j < r} (p_j + p_{r-1-j})(j+1) \qquad (r \geq 1),$$

for any probability distribution $\{p_j\}$.

Next we show that $\vartheta = 2$ is a simple zero, namely, $\Lambda_r'(2) \neq 0$. For,

$$\begin{aligned}
\Lambda_r'(2) &= (r+1)! \sum_{2 \le \ell \le r+1} \frac{1}{\ell} - r! \sum_{0 \le j < r} (p_j + p_{r-1-j})(j+1) \sum_{2 \le \ell \le j+1} \frac{1}{\ell} \\
&= r! \sum_{0 \le j < r} (p_j + p_{r-1-j})(j+1) \sum_{j+2 \le \ell \le r+1} \frac{1}{\ell} \\
&> 0,
\end{aligned}$$

for $r \ge 1$.

Finally, we prove that all other zeros have real part less than two. Assume that $\vartheta = 2 + \delta + iv$ is a zero, where $\delta^2 + v^2 > 0$ and $\delta \ge 0$. Then

$$\begin{aligned}
\left|\vartheta^{\overline{r}}\right| &= (r+1)! \prod_{0 \le \ell < r} \left| 1 + \frac{\delta + iv}{2 + \ell} \right| \\
&= r! \sum_{0 \le j < r} (p_j + p_{r-1-j})(j+1) \prod_{0 \le \ell < r} \left| 1 + \frac{\delta + iv}{2 + \ell} \right| \\
&> r! \sum_{0 \le j < r} (p_j + p_{r-1-j})(j+1) \prod_{0 \le \ell < j} \left| 1 + \frac{\delta + iv}{2 + \ell} \right| \\
&= r! \sum_{0 \le j < r} (p_j + p_{r-1-j}) \frac{|\vartheta^{\overline{j}}|}{j!} \\
&\ge r! \left| \sum_{0 \le j < r} (p_j + p_{r-1-j}) \frac{\vartheta^{\overline{j}}}{j!} \right|,
\end{aligned}$$

for $r \ge 2$, which is a contradiction. Thus $\Re(\lambda_j) < 2$ for $j = 2, \ldots, r$.

Special cases. If $r = 1$, then $\Lambda_r(\vartheta) = \vartheta - 2$. If $r = 2$, then $\Lambda_r(\vartheta) = \vartheta^2 - \vartheta - 2$. If $r = 3$, then

$$\Lambda_r(\vartheta) = \vartheta^3 + 3p_1\vartheta^2 - (1 + 9p_1)\vartheta - 6(1 - p_1).$$

Thus if $p_1 < (2\sqrt{6} - 4)/3$, then the second largest zeros are complex conjugates; if $p_1 = (2\sqrt{6} - 4)/3$, then $1 - \sqrt{6}$ is a double zero; if $p_1 > (2\sqrt{6} - 4)/3$, then the three zeros are all real. The pattern of zeros becomes more complicated for larger values of r.

Mean. Let $f(z) := \sum_n E(X_n) z^n$. Then f satisfies (37). Applying Theorem 1 with $\lambda_1 = 2$, $c_j = r!(p_j + p_{r-1-j})/j!$, and $b_j = 0$ for $j < r$, we obtain

$$E(X_n) = [z^n] f(z) \sim \mu n,$$

where $\mu = r!/\Lambda_r'(2)$ is, after simplification, given by (38).

This quick estimate is insufficient for our uses. We need the more precise expansions

$$E(X_n) = \mu n + \begin{cases} O\left(n^{\alpha-1} (\log n)^{\max\{m,q\}-1} + 1\right), & \text{if } \alpha < 3/2; \\ (T_1 + o(1))\, n^{\alpha-1} (\log n)^{m-1}, & \text{if } 3/2 \le \alpha < 2 \text{ and } m > q; \\ (\Pi_1(n) + o(1))\, n^{\alpha-1} (\log n)^{q-1}, & \text{if } 3/2 \le \alpha < 2 \text{ and } q \ge m, \end{cases} \tag{40}$$

where

$$T_1 = \frac{mr!}{(\alpha-1)\Lambda_r^{(m)}(\alpha)\Gamma(\alpha)}, \tag{41}$$

and $\Pi_1(n)$ is a periodic function of the form

$$\Pi_1(n) = \sum_{\substack{\lambda_j\in\mathscr{Z}_1\\ \mathrm{mul}(\lambda_j)=q}} \frac{qr!}{(\lambda_j-1)\Lambda_r^{(q)}(\lambda_j)\Gamma(\lambda_j)}\, n^{i\Im(\lambda_j)},$$

with $\mathrm{mul}(\lambda)$ denoting the multiplicity of the zero λ.

The expansions do not follow directly from Theorem 1 or 1' but the same method of proof applies. We sketch the proof for T_1. By the initial conditions $f^{(j)}(0) = 0$ for $j < r$, we deduce as above that

$$\Omega_1(\vartheta)(f)(z) = r!(1-z)^{-2} - r!(1-z)^{-1}.$$

Then, defining $\Omega_2(\vartheta) := \Omega_1(\vartheta)/(\vartheta-\alpha)^m$, we derive, by Corollary 1

$$\begin{aligned}\Omega_2(\vartheta)(f)(z) &= r!I_\alpha^{[m]}[(1-z)^{-2}] - r!I_\alpha[(1-z)^{-1}]\\ &= r!\frac{(1-z)^{-2}}{(2-\alpha)^m} + r!\frac{(1-z)^{-\alpha}}{(\alpha-2)(\alpha-1)(m-1)!}\log^{m-1}\frac{1}{1-z} + Q_{13}(z),\end{aligned}$$

where $Q_{13}(z) \in \mathbf{Q}_{\alpha-1,m-1}$ by (9) and (11). Then (41) follows from applying successively Lemma 4.

The proof for $\Pi_1(n)$ is similar.

For later uses, we write $\Pi_1(n)$ in the form

$$\Pi_1(n) = \sum_{1\le j\le u} \varpi_j n^{i\beta_j} \qquad (\varpi_j \in \mathbb{C}),$$

for some integer u and $\beta_j \in \{\Im(\lambda_j) : \lambda_j \in \mathscr{Z}_1 \text{ and } \mathrm{mul}(\lambda_j) = q\}$.

Shifting the mean. To compute higher central moments, it is more convenient to shift first the asymptotic mean by considering

$$\overline{P}_n(y) := E(e^{(X_n-\mu n-\mu)y}) = P_n(y)e^{-\mu(n+1)y},$$

which satisfies the recurrence $\overline{P}_n(y) = e^{-\mu(n+1)y}$ for $n < r$ and

$$\overline{P}_n(y) = e^y \sum_{0\le k<n} \pi_{n,k}\overline{P}_k(y)\overline{P}_{n-1-k}(y) \qquad (n \ge r). \tag{42}$$

Recurrence of higher moments. Let $P_{n,h} := \overline{P}_n^{(h)}(0) = E(X_n-\mu n-\mu)^h$. Then, by (42), $P_{n,h} = (-\mu)^h(n+1)^h$ for $n < r$ and

$$P_{n,h} = \sum_{0\le k<n} (\pi_{n,k} + \pi_{n,n-1-k})\, P_{k,h} + R_{n,h} \qquad (n \ge r), \tag{43}$$

where

$$R_{n,h} = \sum_{\substack{a+b+c=h\\ a,b<h}} \binom{h}{a,b,c} \sum_{0\le k<n} \pi_{n,k}P_{k,a}P_{n-1-k,b} \qquad (h \ge 1).$$

Variance. The order of the variance is pivotal for the change of the limit laws and it depends heavily on the second order term in the asymptotic expansion of the mean, namely, $P_{n,1}$.

By (43),

$$R_{n,2} = 1 + 2 \sum_{0\le k<n} (\pi_{n,k} + \pi_{n,n-1-k})\, P_{k,1} + 2 \sum_{0\le k<n} \pi_{n,k} P_{k,1} P_{n-1-k,1}.$$

By applying the estimates (40), we deduce that

$$R_{n,2} = \begin{cases} O\left(n^{2\alpha-2}(\log n)^{2\max\{m,q\}-2} + 1\right), & \text{if } \alpha < 3/2;\\ (C_2 + o(1))n^{2\alpha-2}(\log n)^{2m-2}, & \text{if } \alpha \ge 3/2 \text{ and } m > q;\\ (\tilde{\Pi}_1(n) + o(1))n^{2\alpha-2}(\log n)^{2q-2}, & \text{if } \alpha \ge 3/2 \text{ and } q \ge m, \end{cases}$$

where

$$C_2 := 2T_1^2 r \sum_{0\le j<r} p_j \binom{r-1}{j} \frac{\Gamma(j+\alpha)\Gamma(r-1-j+\alpha)}{\Gamma(r-1+2\alpha)},$$

and

$$\tilde{\Pi}_1(n) = 2r \sum_{0\le j<r} p_j \binom{r-1}{j} \sum_{1\le s,t\le u} \varpi_s \varpi_t \frac{\Gamma(j+\alpha+i\beta_s)\Gamma(r-1-j+\alpha+i\beta_t)}{\Gamma(r-1+2\alpha+i\beta_s+i\beta_t)}\, n^{i\beta_s+i\beta_t}.$$

From Theorems 1 and 1', it follows that

$$P_{n,2} = \begin{cases} \sigma_1^2 n + o(n), & \text{if } \alpha < 3/2;\\ \dfrac{C_2(r+1)! + o(1)}{(2m-1)\Lambda_r'(2)}\, n(\log n)^{2m-1}, & \text{if } \alpha = 3/2 \text{ and } m \ge q;\\ (\Pi_2(n) + o(1))n(\log n)^{2q-2}, & \text{if } \alpha = 3/2 \text{ and } q > m;\\ (T_2 + o(1))n^{2\alpha-2}(\log n)^{2m-2}, & \text{if } \alpha > 3/2 \text{ and } m > q;\\ (\Pi_2(n) + o(1))n^{2\alpha-2}(\log n)^{2q-2}, & \text{if } \alpha > 3/2 \text{ and } q \ge m, \end{cases}$$

where

$$\begin{aligned} \sigma_1^2 &:= \mu \sum_{0\le j<r} (p_j + p_{r-1-j})(j+1) \sum_{k\ge j} \frac{R_{k,2}}{(k+1)(k+2)}, && (44)\\ T_2 &:= 2C_2^2 r \sum_{0\le j<r} p_j \binom{r-1}{j} \frac{\Gamma(j+\alpha)\Gamma(r-1-j+\alpha)}{\Lambda_r(2\alpha-1)\Gamma(2\alpha-1)},\\ \Pi_2(n) &= 2r \sum_{0\le j<r} p_j \binom{r-1}{j} \sum_{1\le s,t\le u} \frac{\varpi_s \varpi_t \Gamma(j+\alpha+i\beta_s)\Gamma(r-1-j+\alpha+i\beta_t)}{\Lambda_r(2\alpha-1+i\beta_s+i\beta_t)\Gamma(2\alpha-1+i\beta_s+i\beta_t)}\, n^{i\beta_s+i\beta_t}. \end{aligned}$$

Consequently, by definition, $\mathrm{Var}(X_n) = P_{n,2} - P_{n,1}^2$, and (39) follows with

$$\begin{aligned} \sigma_2^2 &:= \frac{C_2(r+1)!}{\Lambda_r'(2)(2m-1)} = \frac{C_2(r+1)}{2m-1}\mu, && (45)\\ \Psi_1(n) &= \Pi_2(n) - \Pi_1(n)^2,\\ \sigma_3^2 &:= T_2 - T_1^2. && (46) \end{aligned}$$

Asymptotic normality: $\alpha < 3/2$. We prove by induction that for $h \geq 1$

$$\begin{cases} P_{n,2h} \sim \dfrac{(2h)!}{h!2^h}\sigma_1^{2h}\, n^h \\ P_{n,2h-1} = o(n^{h-1/2}). \end{cases}$$

This will imply that

$$\begin{cases} E\left(\dfrac{X_n - \mu n}{\sigma_1\sqrt{n}}\right)^{2h} \sim \dfrac{(2h)!}{h!2^h}; \\ E\left(\dfrac{X_n - \mu n}{\sigma_1\sqrt{n}}\right)^{2h-1} = o(1), \end{cases}$$

from this the asymptotic normality of $(X_n - \mu n)/(\sigma_1\sqrt{n})$ will follow.

The case $h = 1$ is proved by (39) and (40). Assume that $h \geq 2$. By (43) and induction,

$$\begin{aligned} R_{n,2h} &\sim \sum_{1\leq a<h} \binom{2h}{2a} \frac{(2a)!(2h-2a)!}{a!2^a(h-a)!2^{h-a}}\, \sigma_1^{2h} \sum_{0\leq k<n} \pi_{n,k} k^a (n-1-k)^{h-a} \\ &= \frac{(2h)!}{h!2^h}\, \sigma_1^{2h} \sum_{1\leq a<h} \binom{h}{a} \sum_{0\leq j<r} p_j \sum_{0\leq k<n} \frac{\binom{k}{j}\binom{n-1-k}{r-1-j}}{\binom{n}{r}}\, k^a (n-1-k)^{h-a} \\ &\sim \frac{(2h)!}{h!2^h}\, \sigma_1^{2h}\, n^h \sum_{0\leq j<r} p_j r \binom{r-1}{j} \sum_{1\leq a<h} \binom{h}{a} \int_0^1 x^{a+j}(1-x)^{r-1-j+h-a}\, \mathrm{d}x \\ &= \frac{(2h)!}{h!2^h}\, \sigma_1^{2h}\, n^h \left(1 - \frac{r!}{(r+h)!} \sum_{0\leq j<r} (p_j + p_{r-1-j}) \frac{(j+h)!}{j!}\right). \end{aligned}$$

So that by applying Case (*iii*) of Theorem 1 with $\upsilon = h$ and $\nu = 2$, we get

$$P_{n,2h} \sim \frac{(2h)!(r+h)!}{h!2^h h!\Lambda_r(h+1)}\, \sigma_1^{2h}\, n^h \left(1 - \frac{r!}{(r+h)!} \sum_{0\leq j<r} (p_j + p_{r-1-j}) \frac{(j+h)!}{j!}\right).$$

Observe that

$$\begin{aligned} \frac{(r+h)!}{h!\Lambda_r(h+1)} &= \frac{(r+h)!}{(r+h)! - r!\sum_{0\leq j<r}(p_j + p_{r-1-j})(j+h)!/j!} \\ &= \frac{1}{1 - \frac{r!}{(r+h)!}\sum_{0\leq j<r}(p_j + p_{r-1-j})\frac{(j+h)!}{j!}}. \end{aligned}$$

Accordingly,

$$P_{n,2h} \sim \frac{(2h)!}{h!2^h}\, \sigma_1^{2h}\, n^h.$$

Similarly, we deduce that

$$P_{n,2h-1} = o(n^{h-1/2}).$$

Asymptotic normality: $\alpha = 3/2$ and $m \geq q$. The proof is similar and is omitted.

Non-existence of limit laws: $\alpha = 3/2$ and $q > m$. Our approach can be applied to prove that

$$P_{n,h} \sim \Pi_h(n) n^{h/2} (\log n)^{h(q-1)},$$

so that

$$E\left(\frac{X_n - \mu n}{\sqrt{n}(\log n)^{q-1}}\right)^h \sim \Pi_h(n),$$

where the $\Pi_h(n)$'s are periodic functions that can be computed recursively; see [20]. In particular, $\Pi_1(n)$ and $\Pi_2(n)$ are given as above. The non-existence of the limit laws follows from the Frechet-Shohat moment convergence theorem [47]; see [20] for similar details.

Non-normal limit laws: $\alpha > 3/2$ and $m > q$. By induction and the asymptotic transfers, we can show that

$$P_{n,h} \sim T_h n^{h(\alpha-1)} (\log n)^{h(m-1)} \qquad (h \ge 1),$$

where $T_0 = 1$, T_1 and T_2 are defined as above, and

$$\begin{aligned} T_h &= \frac{\Gamma(h(\alpha-1)+r+1)}{\Lambda_r(h(\alpha-1)+1)\Gamma(h(\alpha-1)+1)} \sum_{0 \le j < r} p_j r \binom{r-1}{j} \\ &\qquad \times \sum_{1 \le k < h} \binom{h}{k} T_k T_{h-k} \int_0^1 x^{k(\alpha-1)+j} (1-x)^{(h-k)(\alpha-1)+r-1-j} \, \mathrm{d}x \qquad (h \ge 2). \end{aligned} \tag{47}$$

Thus

$$E\left(\frac{X_n - \mu n}{n^{\alpha-1}(\log n)^{q-1}}\right)^h \to T_h \qquad (h \ge 1).$$

We now show that the sequence $\{T_h\}_h$ uniquely determines a distribution by proving that $|T_h| \le K^h h!$ for a sufficiently large $K > 0$. Then the convergence in distribution of $(X_n - \mu n)/(n^{\alpha-1}(\log n)^{m-1})$ will follow from the Frechet-Shohat theorem (see [47, 76]).

The proof for $|T_h| \le K^h h!$ is as follows. First we have, by induction and by using the inequality $p_j \le 1$,

$$|T_h| \le h! K^h W_h,$$

where

$$W_h := \frac{r\Gamma(h(\alpha-1)+r+1)}{\Lambda_r(h(\alpha-1)+1)\Gamma(h(\alpha-1)+1)} \sum_{1 \le k < h} \int_0^1 x^{k(\alpha-1)} (1-x)^{(h-k)(\alpha-1)} \, \mathrm{d}x.$$

Note that

$$\frac{r\Gamma(h(\alpha-1)+r+1)}{\Lambda_r(h(\alpha-1)+1)\Gamma(h(\alpha-1)+1)} = \frac{r(h(\alpha-1)+1)^{\overline{r}}}{\Lambda_r(h(\alpha-1)+1)} \sim r,$$

as $h \to \infty$. On the other hand,

$$\begin{aligned} \sum_{1 \le k < h} \int_0^1 x^{k(\alpha-1)} (1-x)^{(h-k)(\alpha-1)} \, \mathrm{d}x &= 2 \int_0^{1/2} \frac{x^{\alpha-1}(1-x)^{h(\alpha-1)} - (1-x)^{\alpha-1} x^{h(\alpha-1)}}{(1-x)^{\alpha-1} - x^{\alpha-1}} \, \mathrm{d}x \\ &\sim 2 \int_0^{\infty} x^{\alpha-1} e^{-h(\alpha-1)x} \, \mathrm{d}x \\ &= O\left(h^{-\alpha}\right). \end{aligned}$$

Thus $W_h < 1$ for $h \ge h_0 > 1$. It follows that $|T_h| \le h! K^h$ for $h \ge 0$ if K was chosen so large that $|T_h| \le h! K^h$ for all $h < h_0$.

Non-existence of limit laws: $\alpha > 3/2$ and $q \geq m$. This case is similar to the case when $\alpha = 3/2$ and $q > m$. The proof is omitted.

Variability condition. The constants or periodic functions in the dominant terms of the variance (39) are *a priori* positive, but, with the exception of σ_2, it is not obvious to see the positivity of these constants and periodic functions. Here is a general approach for achieving this.

For convenience, let $v_n = \mathrm{Var}(X_n)$. Then $v_n = 0$ for $n < r$ and satisfies the recurrence

$$v_n = \sum_{0\le k<n} (\pi_{n,k} + \pi_{n,n-1-k})v_k + u_n \qquad (n \geq r),$$

where

$$u_n := \sum_{0\le k<n} \pi_{n,k} \left(\mu_k + \mu_{n-1-k} - \mu_n + 1\right)^2 \qquad (\mu_n := E(X_n)).$$

Thus when $\alpha < 3/2$, $u_n \geq 0$ satisfies the conditions (14) and it follows that σ_1^2 has an alternative expression:

$$\sigma_1^2 = (r+1)\mu \sum_{k\ge r} \frac{u_k}{(k+1)(k+2)},$$

which is positive since $u_n \not\equiv 0$.

The same approach applies to σ_3 since $T_1 \neq 0$, and to the periodic functions Ψ_1, Ψ_2 since $\Pi_1(n) \not\equiv 0$.

3.4 Limit laws for the number of comparisons

Let X_n be the number of comparisons used by the quicksort described above. Then $P_n(y) := E(e^{X_n y})$ satisfies

$$P_n(y) = e^{\tau_n y} \sum_{0\le k<n} \pi_{n,k} P_k(y) P_{n-1-k}(y) \qquad (n \geq r),$$

with suitable initial conditions, where we assume that τ_n is a function of n such that $\tau_n - n$ satisfies (14) with $\lambda_1 = 2$. Then

$$E(X_n) \sim \mu n \log n + Cn,$$

where $\mu = (r+1)!/\Lambda_r'(2)$ and C is a constant depending on initial conditions. Also the same method of moments and asymptotic transfers can be applied to show that (see [62])

$$\frac{X_n - \mu n \log n - Cn}{n} \xrightarrow{d} X,$$

with convergence of all moments, where $V_h := E(X^h)$ satisfies $V_0 = 1$, $V_1 = 0$ and

$$V_h = \sum_{a+b+c=h} \binom{h}{a,b,c} V_a V_b \sum_{0\le j<r} p_j r \binom{r-1}{j} \int_0^1 x^{a+j}(1-x)^{b+r-1-j} \Delta(x)^c \,\mathrm{d}x, \tag{48}$$

for $h \geq 2$, where $\Delta(x) := 1 + \mu x \log x + \mu(1-x)\log(1-x)$.

4 Applications

We consider applications of the tools developed in Sections 2 and 3 and briefly discuss the major variants of quicksort and the associated BSTs falling in our framework, with an emphasis on phase changes of limit laws.

4.1 Quicksort proper and BSTs

Quicksort, in its original primitive version, uses no sampling and chooses a random element as the pivot. The general cost measures satisfy the form (35) with $r = 1$ and $\pi_{n,k} = 1/n$ (or $p_0 = 1$). If X_n satisfies (36), then $X_n \equiv n$ for $n \geq 1$.

The relationship between quicksort proper and BSTs has already been described in Introduction. For a systematic study on cost measures of quicksort and BSTs, as well as rather complete references, see [62]. See also Devroye [30] for an alternative approach. In particular, the main result in [62] says that if b_n is small (roughly, $b_n \ll n^{1/2}$) then X_n is asymptotically normal, and if b_n is large (roughly, $b_n \gg n^{1/2}$), then the limit law exists and is non-normal, with a normal limit law in between ($b_n \sim n^{1/2}$). Thus the "phase change" occurs at $n^{1/2}$. See [61] for an approach to Berry-Esseen bounds, which gives convergence rates to the normal limit law.

The limiting distribution of the number of comparisons used by quicksort proper has recently received much attention in the literature; see [31, 40, 71] and the references therein. Most of these problems are highly challenging and it seems that more new tools are required.

4.2 Quicksort with median-of-$(2t + 1)$ and locally balanced BSTs

Quicksort proper, although simple, suffers from the drawback that the sizes of the two subfiles after each partition may have large difference, rendering a quadratic performance for the algorithm. Thus it is tempting to choose the pivot more carefully to avoid the worst case.

It is well known that the median is robust to outliers. Thus a simple idea to avoid the potentially worst-case of quicksort proper is to take first a random sample of odd size > 1 and then choose the median of this sample as the pivot. So the general cost measures satisfy the recurrence (35) with $r = 2t+1$ and $p_t = 1$. Quadratic worst-case is still possible but less likely; see [71, 84] for more recise results on large deviations.

In terms of BSTs, this version of quicksort corresponds to picking the median-of-$(2t+1)$ elements as the root once the size of the tree exceeds $2t + 1$.

The number of partitions used to find the median-of-$(2t + 1)$ satisfies (36). It is known (see [20]) that $X_n = X_n(t)$ is asymptotically normally distributed if $1 \leq t \leq 58$, and the limiting distribution does not exist for $t > 58$. For other probabilistic results, see [7, 26, 29, 50, 53, 54, 78, 90, 93, 94, 102].

4.3 Quicksort using ninther

The ninther was first introduced by J. W. Tukey [112] as a quick, robust summary of a sample (indeed he used the median of many ninthers). When used in quicksort, we choose first a ninther and then partition the input in the usual way and sort the subfiles recursively; see Bentley and McIlroy [11] for a detailed description and implementation. Quicksort using ninther has been widely used in practical implementations; see for example [4, 74, 85].

The general cost measures for this quicksort satisfy (35) with $r = 9$ and

$$\{p_j\}_{0\leq j<9} = \left\{0, 0, 0, \frac{3}{14}, \frac{4}{7}, \frac{3}{14}, 0, 0, 0\right\}.$$

Our theory applies. In particular, the number of comparisons used satisfies (cf. [36])

$$E(X_n) \sim \frac{12600}{8027} nH_n + Cn,$$

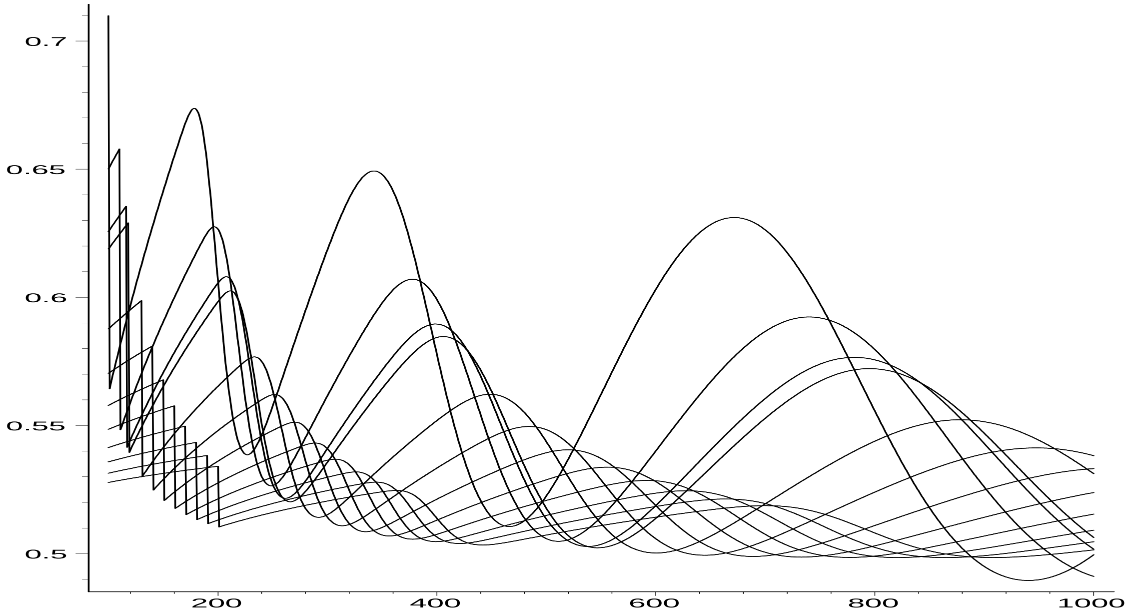

Figure 1: *Periodic fluctuations of the normalized sequence* $-(E(X_n)-\mu n)/n^{\alpha-1}$ *(median-of-*$(2t+1)$*), for* n *from 100 to 1000 and for* $t \in \{50, 55, 58, 59, 65, 70, 75, 80, 85, 90, 95, 100\}$.

where C depends on initial conditions, and $(X_n - \frac{12600}{8027} n \log n - Cn)/n$ converges in distribution to some random variable X; see (48). And the variance is asymptotic to

$$\mathrm{Var}(X_n) \sim \left(\frac{96701354802804451177 7}{623634635288174590080} - \frac{1050}{8027}\pi^2 \right) n^2 \approx 0.25958 n^2.$$

4.4 Quicksort using remedian-of-3^d

Beyond the median of three medians-of-three, we can consider the remedian of 3^d elements, which is defined recursively as the median of three remedians of 3^{d-1} elements for $d \geq 2$, where the remedian of three elements is simply the median. It is known that the remedian is a consistent estimator of the sample median; see [16, 98].

The general cost measures of quicksort using remedian as the pivot at each partitioning step satisfy (35) with $r = 3^d$. The probability distribution $\{p_j\}$ is now more complicated.

We show that for $d \geq 1$ the probability that the remedian of 3^d elements is the $(j+1)$-st order statistics is given recursively by

$$p_j = p_j^{(d)} = 3! \sum_{\substack{a+b+c=j \\ a,b,c \geq 0}} \varrho(a,b,c) \frac{\binom{j}{a,b,c}\binom{3^d-1-j}{3^{d-1}-1-a,3^{d-1}-b,3^{d-1}-c}}{\binom{3^d}{3^{d-1},3^{d-1},3^{d-1}}}, \tag{49}$$

where

$$\varrho(a,b,c) = p_a^{(d-1)} \left(p_0^{(d-1)} + \cdots + p_{b-1}^{(d-1)} \right) \left(p_c^{(d-1)} + \cdots + p_{3^{d-1}-1}^{(d-1)} \right),$$

with $p_0^{(0)} = 1$.

To see this, we first randomly partition the set $\{1, \ldots, 3^d\}$ into three subsets $\mathcal{A}, \mathcal{B}, \mathcal{C}$ each of which contains 3^{d-1} elements. Then the ratio of the binomial coefficients in (49) is the probability that

$j+1 \in \mathcal{A}$ and that there are exactly a, b, c elements that are less than $j+1$ in $\mathcal{A}, \mathcal{B}, \mathcal{C}$, respectively; and $\varrho(a,b,c)$ is the probability, conditioned on $j+1 \in \mathcal{A}$ and on a, b, c, that the remedian of $\mathcal{A}$ is $j+1$, the remedian of $\mathcal{B}$ is less than $j+1$ and the remedian of $\mathcal{C}$ is greater than $j+1$. Thus the expression for $\varrho(a,b,c)$ follows.

The second largest zero(s) (in real parts) of the indicial polynomials $\Lambda_{3^d}(\vartheta)$ for d from 1 to 6 are given in Table 2.

d	1	2	3	4	5	6
α	0	0	0	0.52691	1.33805	1.70469

Table 2: *Approximate values of* α *for* $2 \leq d \leq 6$.

Thus if we consider the number X_n of partitioning stages used by quicksort with remedians, then the limit law is normal for $d \leq 5$ and does not exist for $d > 5$.

4.5 Quicksort using (s,t)-median

Another way of extending the ninther is to consider the median of $2s+1$ elements each of which is the median of $2t+1$ elements; so that $r = (2s+1)(2t+1)$. Such a median is called (s,t)-median.

The probability p_j that the (s,t)-median is the $(j+1)$-st order statistics satisfy

$$p_j = (2s+1)\binom{2s}{s}\frac{\binom{j}{t}\binom{r-1-j}{t}}{\binom{r}{2t+1,\ldots,2t+1}}\,\varrho(s, j-t),$$

where $\varrho(s,k)$ can be computed recursively as

$$\varrho(s,k) = \sum_{0\leq a\leq t}\sum_{t<b\leq 2t+1}\binom{k}{a,b}\binom{2s(2t+1)-k}{2t+1-a,\, 2t+1-b}\varrho(s-1, k-a-b),$$

with the initial condition $\varrho(0,k) = 1$.

To see this, we partition the set $\{1, \ldots, r\}$ into $2s+1$ subsets $\{\mathcal{A}_1, \ldots, \mathcal{A}_{2s+1}\}$ each containing $2t+1$ elements. The desired (s,t)-median is the median of the medians of $\{\mathcal{A}_1, \ldots, \mathcal{A}_{2s+1}\}$. First, the binomial coefficients $\binom{j}{t}\binom{r-1-j}{t}$ is the number of combinations that the median of $\mathcal{A}_{2s+1}$ is exactly $j+1$, and $\varrho(s, j-t)$ enumerates the number of arrangements that exactly s medians of $\mathcal{A}_\ell$ are less than $j+1$, and the remaining s medians of $\mathcal{A}_\ell$ are all greater than $j+1$. Thus the ratio $\binom{j}{t}\binom{r-1-j}{t}\varrho(s, j-t)/\binom{r}{2t+1,\ldots,2t+1}$ is the probability that the median of $\mathcal{A}_{2s+1}$ is exactly $j+1$, the s medians of $\mathcal{A}_1, \ldots, \mathcal{A}_s$ are all less than $j+1$, and the s medians of $\mathcal{A}_{s+1}, \ldots, \mathcal{A}_{2s}$ are all greater than $j+1$. Note that the total number of combinations of arranging the $\mathcal{A}_j$'s so that they have the above configuration is given by $(2s+1)\binom{2s}{s}$. The recurrence for $\varrho(s,k)$ is obvious.

All pairs of (s,t) leading to a phase change of the limit law are given in Table 1.

4.6 Quicksort using minimaxer

Yet a different idea of choosing the pivot is to take the *minimaxer*[1] defined as the minimum of a set of elements each of which is the maximum of some smaller set of elements. Such a scheme has a few features: the probability distribution $\{p_j\}$ is in general asymmetric, and the computational cost

---

[1] By definition, a minimaxer is a player who manipulates the character creation rules to produce extremely powerful characters.

for finding the minimaxer is usually less than that for finding the medians or approximate medians. The idea of minimaxer has been widely used in artificial intelligence and game theory; see [91, 104].

For simplicity, consider the minimum of t elements each of which is the maximum of s elements. Then $r = st$ and the p_j's are given by

$$p_j = \frac{t\binom{j}{s-1}\varrho_s(t-1, st-j-1)}{\binom{st}{s,\ldots,s}},$$

where $\varrho_s(t,k)$ denotes the number of ways of putting k different black balls and $st-k$ different white balls into t different boxes such that each box contains exactly s balls in which at least one is black, so that

$$\varrho_s(t,k) = \sum_{1\le \ell \le k-t+1} \binom{k}{\ell}\binom{st-k}{s-\ell}\varrho_s(t-1, k-\ell).$$

Computing the second largest zero(s) for the indicial equation associated with this problem is much more involved due to the skewness of the distribution, resulting in very slow convergence and large power of polynomials. For $s = 2$, the phase change for the number of partitioning stages occurs at $t = 143$. Note that the distributions are moving towards the left as r increases. We have not computed higher values of s, which is too time-consuming.

4.7 Theoretical quicksorts

Our framework in Section 3 makes it relatively easy to produce phase changes of limit laws, although the associated probability distributions may be hard to be realized in practice. We briefly consider some interesting cases here. Our main motivation of considering these cases was to investigate the fundamental question: *what distributions cause phase changes, is it because the distributions are more and more concentrated as r gets large or because they are more and more biased?* We still have no answer for this question.

The simplest case is the uniform distribution $p_j = 1/r$ for $j = 0, \ldots, r-1$. Obviously, this is a trivial case that brings the pivot-choosing process back to the quicksort proper, and there is no phase change.

In the case of binomial distributions where $p_j = 2^{1-r}\binom{r-1}{j}$ for $j = 0, \ldots, r-1$, the distribution of X_n (defined by (35)) is asymptotically normal for $r \le 236$, and the limit law does not exist for $r > 236$.

Another interesting case is when $p_0 = 1$, which means that we take the smallest of a random sample of r elements as the pivot. Then the phase change occurs at $r = 26$, exactly the same as the storage requirement of m-ary search trees. Taking the second smallest element in the r-sample yields a phase change at $r = 33$, the third smallest at $r = 41$, the fourth smallest at $r = 48$, the fifth smallest at $r = 55$, etc.

If $p_j = \mathbf{s}(n, j+1)/n!$, where the $\mathbf{s}(n,j)$'s denote the signless Stirling numbers of the first kind (see [22]), then the limit laws change nature at $r = 69$. Another phase change occurs at $r = 158$ when taking Eulerian numbers (see [22]) as the underlying distribution.

All these are type I phase changes, as mentioned in the Introduction, it is open to produce type II phase changes by constructing suitable probability distributions $\{p_j\}$.

4.8 Quicksort with multiple pivots and m-ary search trees

We discussed up to now only quicksort with single pivot at each stage. It is natural to use $m-1$ pivots ($m \ge 2$) and to partition the input into m subfiles according to the pivots chosen; the pivots are in their final position after partition and the elements falling between two pivots are sorted recursively. The search tree version of such a sorting scheme is called m-ary search tree. See [53, 54, 68, 77, 109, 111] for more information.

In its simplest version, the cost measures X_n of quicksort using $m-1$ pivots (when given a random permutation) have moment generating functions of the form

$$\begin{aligned} P_n(y) &:= E(e^{X_n y}) \\ &= \frac{Q_n(y)}{\binom{n}{m-1}} \sum_{n_1+\cdots+n_m=n-m+1} P_{n_1}(y)\cdots P_{n_m}(y) \qquad (n \geq m), \end{aligned}$$

with suitable initial conditions for $P_n(y)$, $n < m$, where $Q_n(y)$ denotes the moment generating functions of the "toll function"—cost used to partition the input.

Such a recurrence, although different from (35), produces also differential equations (for the generating functions of moments) that are of CE type in nature. Random variables of this sort exhibit a type I phase change at $m = 26$ when $Q_n(y) = e^y$; see [20]. This corresponds to the number of nodes used by a random m-ary search tree of n keys. For other results on m-ary search trees, see [6, 39, 50, 54, 75, 80].

4.9 Generalized quicksort of Hennequin and generalized m-ary search trees

In this sorting algorithm, a random sample of r elements are selected from the input (a random permutation of n elements), and then $m-1$ pivots are chosen so that the probability that the m subfiles have sizes $n_1, \dots, n_m$ is of the form

$$\pi_{n;n_1,\dots,n_m} = \sum_{j_1+\cdots+j_m=r-m+1} p_{j_1,\dots,j_m} \frac{\binom{n_1}{j_1}\cdots\binom{n_m}{j_m}}{\binom{n}{r}},$$

for some probability distribution

$$\sum_{j_1+\cdots+j_m=r-m+1} p_{j_1,\dots,j_m} = 1,$$

so that the cost measures have MGF of the form

$$P_n(y) = Q_n(y) \sum_{n_1+\cdots+n_m=n-m+1} \pi_{n;n_1,\dots,n_m} P_{n_1}(y)\cdots P_{n_m}(y) \qquad (n \geq r).$$

Due to the hypergeometric factor, such problems are also manageable by our CE theory, and more phase changes can be further produced. For example, if

$$\pi_{n;n_1,\dots,n_m} = \frac{\binom{n_1}{t}\cdots\binom{n_m}{t}}{\binom{n}{m(t+1)-1}},$$

then the phase changes are described in Table 3 (see also [20]).

4.10 Levelwise analysis for quicksort and profiles of search trees

Our discussions have been up to now restricted to the case when $\lambda_1 = 2$. We consider here cases of moving λ_1.

t	0	1	2	3	4	5,6	7,...,10	11,...,19	20,..., 58
$2 \le m \le ?$	26	13	9	7	6	5	4	3	2

Table 3: *The upper bound of m (for each fixed t) for which the asymptotic normality holds. The special case $(m,t) = (2,0)$ has to be discarded ($X_n \equiv n$ in this case).*

For simplicity, consider the quantity $X_{n,k}$, the number of external nodes at level k (the root being at level 0) in a random m-ary search tree of n keys (constructed from a random permutation). Let $M_k(z) := \sum_n E(X_{n,k})z^n$. Then by definition (see Mahmoud [77, Chapter 3])

$$M_{k+1}^{(m-1)}(z) = \frac{m!}{(1-z)^{m-1}} M_k(z) \qquad (k \ge 0),$$

with the initial conditions $M_0(z) \equiv 1$ and $M_{k+1}^{(j)}(0) = (j+1)!\delta_{k0}$, $0 \le j \le m-2$.

Now consider $M(w,z) := \sum_k M_k(z)w^k$. Then

$$\frac{\partial^{m-1}}{\partial z^{m-1}} M(w,z) = \frac{m!w}{(1-z)^{m-1}} M(w,z),$$

with the initial conditions $M(w,z) = 1 + 2wz + \cdots + (m-1)wz^{m-2} + \cdots$. This is a CE equation with indicial equation

$$\vartheta^{\overline{m-1}} - m!w = 0.$$

Each zero of this equation defines an analytic function of w and the dominant zero $\lambda_1(w)$ varies in a neighborhood of 2 when w is near 1. Also we can find a small constant $\varepsilon >$ such that $\Re(\lambda_1(w)) > \Re(\lambda_j(w))$ for all $j = 2, \ldots, m-1$, uniformly for $|w-1| \le \varepsilon$; see Hille [55, §9.4] for the required theory. On the other hand, it is straightforward to prove that all zeros are simple. It follows, by Theorem 1, that

$$[z^n]M(w,z) = K(w)n^{\lambda_1(w)-1}(1+o(1)),$$

uniformly for $|w-1| \le \varepsilon$, where $K(w)$ is analytic for $|w-1| \le \varepsilon$. Then the techniques used in [19] can be further applied to derive the asymptotics of $E(X_{n,k}) = [w^k z^n]M(w,z)$.

See also [19] for an analysis of the levelwise improvements of quicksort and other related problems, and [54, 77] for more examples of the same type.

4.11 Introspective quicksort and balanced BSTs

Introspective sort (see [87, 113]) is a variant of quicksort in which subfiles whose recursion depths (number of nodes between the root and the node itself in terms of BST) exceed a given threshold, say $c\lfloor \log_2 n \rfloor$, are sorted by a heapsort. The idea can, of course, be combined with other schemes of finding the median of an odd sample or an approximate median. The standard STL (Standard Template Library) uses Introspective sort (with median-of-three) as the standard sort function.

Interestingly, exactly the same idea was also proposed for balancing BSTs in [3]. So no balancing scheme is used unless the depth of a node exceeds a given threshold and the total path length can be controlled to be within $O(n \log n)$.

We can analyze the cost measures of such a strategy (stopping a path if it becomes too high) by techniques similar to the one used in [19]. In particular, the dominant zero of the indicial equation also depends on another parameter as in the profile of m-ary search trees.

4.12 Increasing trees

A labeled tree is called increasing if the labels of each path from the root to a node are increasing. We mentioned in Introduction that the binary increasing trees are another isomorphic model to BSTs. Recursive trees are non-plane increasing trees with all node degrees allowed. Such trees are another isomorphic model to BSTs by the well-known transformation from multiway trees to binary trees (see Cormen et al. [24]); see also [13]. Thus parameters in such trees also lead to CE equations of first order.

An extension of recursive trees to multiple branching factors was proposed by Mahmoud and Smythe [81]. The underlying differential equation for the generating functions of moments is of the form

$$(1-z)^m f^{(m)}(z) = m! f(z) + \phi(z).$$

This form is manageable by our approach. The phase change for the number of nodes used to store n keys occurs, as already observed in [81], at $m = 26$ (asymptotic normality for $m \leq 26$ is proved in [81] using Smythe's central limit theorems in [105]).

Another interesting variety of increasing trees is studied by Prodinger [95], where differential equations of the form

$$f'(z) = \frac{c}{1-4z} f(z) + \phi(z),$$

appear. Such an equation is, after a change of variables, of CE type. The associated recurrences are surprisingly very complicated.

4.13 Urn models

There is a close relationship between urn models and the divide-and-conquer sorting or searching schemes discussed in this paper; see [79, 106].

We consider briefly the matrix structures of the associated urn models for the quicksort studied in Section 3; see Mahmoud [79] for more details of related structures. This model is based on the bottom-up description (in Section 3) in terms of growing BSTs instead of the top-down description of quicksort. The urn model consists of balls of $r-1+\lceil r/2 \rceil$ different colors. Color i corresponds to the node holding $i-1$ keys for $1 \leq i \leq r-1$. For nodes holding $r-1$ keys in the BST, we further split the possible future configurations (according to the sizes of two subtrees) into $\lceil r/2 \rceil$ cases, giving rise to further colors; so color $r+j$ corresponds to the node with $r-1$ keys, which, when a new key is inserted, is further split into subtrees of sizes j and $r-1-j$, for $0 \leq j < \lceil r/2 \rceil$. The complete urn scheme is as follows:

$$\begin{pmatrix}
-1 & 2 & 0 & 0 & 0 & \cdots & & & & & & & \cdots & 0 & 0 \\
0 & -2 & 3 & 0 & 0 & \cdots & & & & & & & \cdots & 0 & 0 \\
0 & 0 & -3 & 4 & 0 & \cdots & & & & & & & \cdots & 0 & 0 \\
0 & 0 & 0 & -4 & 5 & \cdots & & & & & & & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & & & & & & & & & \vdots & \vdots \\[2ex]
0 & 0 & 0 & 0 & \cdots & & \cdots & 3-r & r-2 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & & \cdots & 0 & 2-r & r-1 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & & \cdots & 0 & 0 & 1-r & q_0 & q_1 & \cdots & q_{h-1} & q_h \\
1 & 0 & 0 & 0 & \cdots & & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 2 & 0 & 0 & & & & 0 & 0 & r-1 & -q_0 & -q_1 & \cdots & -q_{h-1} & -q_h \\
0 & 0 & 3 & 0 & & & & 0 & r-2 & 0 & -q_0 & -q_1 & & -q_{h-1} & -q_h \\
0 & 0 & 0 & 4 & & & & r-3 & 0 & 0 & -q_0 & -q_1 & & -q_{h-1} & -q_h \\
\vdots & & & & \ddots & & & & & \vdots & \vdots & & & & \vdots \\
0 & 0 & 0 & 0 & \cdots & 2\left\lceil \frac{r}{2} \right\rceil & \cdots & 0 & 0 & 0 & -q_0 & -q_1 & \cdots & -q_{h-1} & -q_h
\end{pmatrix},$$

for odd r, where $h = \lceil r/2 \rceil - 1$, $q_i = r(p_i + p_{r-1-i})$ for $0 \le i < h$ and $q_h = rp_h$. If r is even, then $q_h = r(p_h + p_{h+1})$ and the last row in the matrix is the same as above except replacing properly $2\lceil r/2 \rceil$ by the two numbers $r/2, r/2+1$.

In particular, if $p_{(r+1)/2} = 1$, where $r \ge 3$ is odd, and $p_j = 0$ for $j \ne (r+1)/2$, then the matrix can be simplified; this corresponds to the matrix associated with the paged BSTs in [79]. If $p_0 = 1$ and $p_j = 0$ for $j \ne 0$, then the above matrix can be simplified so that it corresponds to the matrix associated with recursive bucket trees in [79].

Urn schemes of the above form can be treated by our CE theory as discussed in Section 3. The scheme can be further extended to Hennequin's generalized quicksort. An interesting question is: which class of urn models is manageable via our CE theory? But this is beyond the scope of this paper.

4.14 Other applications

For other problems in the analysis of algorithms where CE equations appear, see [62, 63, 70] for quickselect and its variants (for which $\lambda_1 = 1$), [49, 62, 103] for sorting on a broadcast communication model, and [36, Chapter 3].

References

[1] W. J. Adams, *The Life and Times of the Central Limit Theorem*, Kaedmon Publishing Co., New York, 1974.

[2] E. A. Albacea, Leapfrogging samplesort, *Lecture Notes in Computer Science*, **1023** (1995), 1-9.

[3] A. Andersson, General balanced trees, *Journal of Algorithms*, **30** (1999), 1-28.

[4] C. Ashcraft and R. Grimes, SPOOLES: an object-oriented sparse matrix library, in *Proceedings of the 9th SIAM Conference on Parallel Processing for Scientific Computing 1999*, (San Antonio, TX), 10 pp., SIAM, Philadelphia, PA, 1999.

[5] J.-L. Baer and Y.-B. Lin, Improving quicksort performance with a codeword data structure, *IEEE Transactions on Software Engineering*, **15** (1989), 622-631.

[6] R. A. Baeza-Yates, Some average measures in m-ary search trees, *Information Processing Letters*, **25** (1987), 375–381.

[7] R. A. Baeza-Yates, Height balance distribution of search trees, *Information Processing Letters*, **39** (1991), 317–324.

[8] Z.-D. Bai, H.-K. Hwang, W.-Q. Liang and T.-H. Tsai, Limit theorems for the number of maxima of random samples from planar regions, *Electronic Journal of Probability*, **6**, paper 3 (2001), 41 pages; available via `http://www.math.washington.edu/~ejpecp/EjpVol6/paper3.pdf`.

[9] S. Battiato, D. Cantone, D. Catalano, G. Cincotti and M. Hofri, An efficient algorithm for the approximate median selection problem, in *Proceedings of the Fourth Italian Conference, CIAC 2000, Rome*, Ed. by G. Bongiovanni, G. Gambosi, R. Petreschi, Lecture Notes in Computer Science, 1767, pp. 226–238.

[10] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, Springer, NY, 1999.

[11] J. L. Bentley and M. D. McIlroy, Engineering a sort function, *Software–Practice and Experience*, **23** (1993), 1249–1265.

[12] J. L. Bentley and R. Sedgewick, Fast algorithms for sorting and searching strings, in *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, New Orleans Louisiana, January 5–7, 1997, pp. 360-369, ACM Press, 1997.

[13] F. Bergeron, P. Flajolet and B. Salvy, Varieties of increasing trees, in *CAAP '92* (Rennes, 1992), Lecture Notes in Computer Science, **581**, pp. 24–48, Springer, Berlin, 1992; available via `ftp://ftp.inria.fr/INRIA/publication/publi-ps-gz/RR/RR-1583.ps.gz`.

[14] P. Billingsley, *Probability and Measure*, Third Edition, John Wiley & Sons, NY, 1995.

[15] W. H. Burge, Analysis of a tree sorting method and some properties of a set of trees, in *First USA-JAPAN Computer Conference Proceedings*, AFIPS and IPSJ, pp. 372–378, 1972.

[16] M. T. Chao and G. D. Lin, The asymptotic distributions of the remedians, *Journal of Statistical Planning and Inference*, **37** (1993), 1-11.

[17] P. L. Chebyshev, *Oeuvres de P. L. Tchebychev, Tome II*, Edited by A. Markov and N. Sonin, Chelsea, NY, 1962.

[18] W.-M. Chen and H.-K. Hwang, Analysis of two randomized algorithms for finding the maximum in a broadcast communication model, submitted (2001); available via `http://algo.stat.sinica.edu.tw`.

[19] H.-H. Chern and H.-K. Hwang, Transitional behaviors of the average cost of quicksort with median-of-$(2t+1)$, *Algorithmica*, **29** (2001), 44-69.

[20] H.-H. Chern and H.-K. Hwang, Phase changes in random m-ary search trees and generalized quicksort, *Random Structures and Algorithms*, **19** (2001), 316-358.

[21] H.-H. Chern, H.-K. Hwang and Y.-N. Yeh, Distribution of the number of consecutive records, *Random Structures and Algorithms*, **17** (2000), 169-196.

[22] L. Comtet, *Advanced Combinatorics. The Art of Finite and Infinite Expansions*, Revised and enlarged edition, D. Reidel, Dordrecht, 1974.

[23] C. R. Cook and D. J. Kim, Best sorting algorithms for nearly sorted lists, *Communications of the ACM*, **23** (1980), 620-624.

[24] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*, MIT Press, 1990.

[25] N. J. Cox, The ninther and its relatives: summary measures useful in geomorphology, *Zeitschrift für Geomorphologie*, **36** (1992), 491–495.

[26] W. Cunto and J. L. Gascon, Improving time and space efficiency in generalized binary search trees, *Acta Informatica*, **24** (1987), 583–594.

[27] H. Delange, On the use of the method of moments for the study of additive functions, *Journal of Number Theory*, **39** (1991), 144–161.

[28] L. Devroye, Moment inequalities for random variables in computational geometry, *Computing*, **30** (1983), 111–119.

[29] L. Devroye, Universal limit laws for depths in random trees, *SIAM Journal on Computing*, **28** (1999), 409–432.

[30] L. Devroye, Limit laws for sums of functions of subtrees of random binary search trees, preprint, (2001).

[31] L. Devroye, J. A. Fill and R. Neininger, Perfect simulation from the quicksort limit distribution, *Electronic Communications in Probability*, **5** (2000), 95-99 (electronic); available via `http://www.math.washington.edu/~ejpecp/EcpVol5/paper12.pdf`.

[32] P. Diaconis, Application of the method of moments in probability and statistics, in "Moments in mathematics" (San Antonio, Texas, 1987), pp. 125–142, AMS, Providence, RI (1987).

[33] F. Dionisi, B. Hug, J. M. Aeschlimann and A. Houllemar, Supercritical CO_2 extraction for total fat analysis of food products, *Journal of Food Science*, **64** (1999), 612–615.

[34] R. G. Dromey, Exploiting partial order with quicksort, *Software—Practice and Experience*, **14** (1984), 509-518.

[35] O. Dubois, R. Monasson, B. Selman and R. Zecchina (Editors), *Phase Transitions in Combinatorial Problems*, Papers from the *Conference on NP-Hardness and Phase Transitions* held in Trieste, September 1999, *Theoretical Computer Science*, **265**, 2001.

[36] M. Durand, *Holonomie et applications en l'analyse d'algorithmes et combinatoire*, DEA Mémoire, LIX, Ecole Polytechnique, 2000; available via `http://algo.inria.fr/durand`.

[37] B. Durian, Quicksort without a stack, in *Proceedings of the 12th Symposium on Mathematical Foundations of Computer Science*, Lecture Notes in Computer Science, Vol. 233, pp. 283-289, Springer, 1986.

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

[39] J. A. Fill, Transfer theorems and asymptotic distributional results for m-ary search trees, manuscript, 2001.

[40] J. A. Fill and S. Janson, Quicksort asymptotics, preprint (2001); available via `http://www.mts.jhu.edu/~fill`.

[41] P. Flajolet, G. Gonnet, C. Puech and J. M. Robson, Analytic variations on quadtrees, *Algorithmica*, **10** (1993), 473-500.

[42] P. Flajolet and A. Odlyzko, The average height of binary trees and other simple trees, *Journal of Computer System and Sciences*, **25** (1982) 171-213.

[43] P. Flajolet and A. Odlyzko, Singularity analysis of generating functions, *SIAM Journal on Discrete Mathematics*, **3** (1990), 216–240.

[44] P. Flajolet, P. Poblete and A. Viola, On the analysis of linear probing hashing, *Algorithmica*, **22** (1998), 490–515.

[45] P. Flajolet and C. Puech, Partial match retrieval of multidimensional data, *Journal of the ACM*, **33** (1986), 371–407.

[46] J. Françon, Arbres binaires de recherche: propriétés combinatoires et application, *RAIRO Informatique Théorique*, **10** (1976), 35–50.

[47] M. Fréchet and J. Shohat, A proof of the generalized second-limit theorem in the theory of probability, *Transactions of the American Mathematical Society*, **33** (1931), 533-543.

[48] G. H. Gonnet and R. Baeza-Yates, *Handbook of Algorithms and Data Structures*, Addison-Wesley, Workingham, 1991.

[49] P. Grabner and H. Prodinger, Sorting algorithms for broadcast communications: Mathematical analysis, *Theoretical Computer Science*, to appear (2001).

[50] D. H. Greene, *Labelled Formal Languages and Their Uses*, Ph. D. Thesis, Stanford University, 1983.

[51] J. M. Hammersley, Generalization of the fundamental theorem on sub-additive functions, *Proceedings of the Cambridge Philosophical Society*, **58** (1962), 235–238.

[52] J. M. Hammersley and G. R. Grimmett, Maximal solutions of the generalized subadditive inequality, in "Stochastic Geometry," edited by E. F. Harding and D. G. Kendall, John Wiley and Sons, London, pp. 270–284, 1974.

[53] P. Hennequin, Combinatorial analysis of quicksort algorithm, *RAIRO Informatique Théorique et Applications*, **23** (1989), 317–333.

[54] P. Hennequin, *Analyse en moyenne d'algorithme, tri rapide et arbres de recherche*, Ph.D. Thesis, Ecole Polytechnique, 1991.

[55] E. Hille, *Analytic Function Theory, Volume I*, Gin and Company, Boston, MA, 1959.

[56] C. A. R. Hoare, Algorithm 64: Quicksort, *Communications of the ACM*, **4** (1961), 321.

[57] C. A. R. Hoare, Quicksort, *Computer Journal*, **5** (1962), 10–15.

[58] B.-C. Huang and D. E. Knuth, A one-way, stackless quicksort algorithm, *BIT*, **26** (1986), 127-130.

[59] F. Hubalek, H.-K. Hwang, W. Lew, H. M. Mahmoud and H. Prodinger, A multivariate view of random bucket digital search trees, submitted (2001); available via `http://algo.stat.sinica.edu.tw`.

[60] C. Hurley and R. Modarres, Low-Storage quantile estimation, *Computational Statistics*, **10** (1995), 311–325.

[61] H.-K. Hwang, Second phase changes in random m-ary search trees and generalized quicksort, submitted (2001).

[62] H.-K. Hwang and R. Neininger, Phase changes of the limit laws in the quicksort recurrence under varying toll functions, submitted (2001); available via `http://algo.stat.sinica.edu.tw`.

[63] H.-K. Hwang and T.-H. Tsai, Quickselect and Dickman function, *Combinatorics, Probability and Computing*, accepted for publication (2001); available via `http://algo.stat.sinica.edu.tw`.

[64] H.-K. Hwang and T.-H. Tsai, An asymptotic theory for recurrence relations based on minimization and maximization, submitted for publication (2001); available via `http://algo.stat.sinica.edu.tw`.

[65] E. L. Ince, *Ordinary Differential Equations*, Dover, New York, 1926.

[66] J. JaJa, A perspective on quicksort, *Computing in Science & Engineering*, January/February, 43–49, 2000.

[67] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, New York, 2000.

[68] M. Kaykobad, Md. M. Islam, M. E. Amyeen and M. M. Murshed, 3 is a more promising algorithmic parameter than 2, *Computers and Mathematics with Applications*, **36** (1998), 19–24.

[69] P. R. Kersten, Fuzzy order statistics and their applications to fuzzy clustering, *IEEE Transactions on Fuzzy Systems*, **7** (1999), 708–712.

[70] P. Kirschenhofer, H. Prodinger and C. Martínez, Analysis of Hoare's FIND algorithm with median-of-three partition, *Random Structures and Algorithms*, **10** (1997), 143–156.

[71] C. Knessl and W. Szpankowski, Quicksort algorithm again revisited, *Discrete Mathematics and Theoretical Computer Science (Electronic Journal)*, **3** (1999), 43-64; available via `http://dmtcs.loria.fr/volumes/abstracts/dm030202.abs.html`.

[72] D. E. Knuth, *The Art of Computer Programming, Volume III, Sorting and Searching*, Second Edition, Addison Wesley, Reading, MA, 1998.

[73] S. Kundu, Sorting tree, nestling tree and inverse permutation, *Information Processing Letters*, **6** (1977), 94-96.

[74] J. N. Larsson and K. Sadakane, Faster suffix sorting, Technical Report, LU-CS-TR:99-214, LUNDFD6/(NFCS-3140)/1–20/(1999), Department of Computer Science, Lund University; available via `http://www.cs.lth.se/~jesper/ssrev-tr.pdf`.

[75] W. Lew and H. M. Mahmoud, The joint distribution of elastic buckets in multiway search trees, *SIAM Journal on Computing*, **23** (1994), 1050–1074.

[76] M. Loève, *Probability Theory. I*, Fourth edition, Graduate Texts in Mathematics, Vol. 45, Springer-Verlag, New York, 1977.

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

[78] H. M. Mahmoud, On rotations in fringe-balanced binary trees, *Information Processing Letters*, **65** (1998), 41–46.

[79] H. M. Mahmoud, The size of random bucket trees via urn models, manuscript (2001).

[80] H. M. Mahmoud and B. Pittel, Analysis of the space of search trees under the random insertion algorithm, *Journal of Algorithms*, **10** (1989), 52–75.

[81] H. M. Mahmoud and R. T. Smythe, Probabilistic analysis of bucket recursive trees, *Theoretical Computer Science*, **144** (1995), 221–249.

[82] G. S. Manku, S. Rajagopalan and B. G. Lindsay, Approximate medians and other quantiles in one pass and with limited memory, in *Proceedings, 1998 ACM SIGMOD*, Vol. 27 (1998), 426-435.

[83] C. Martínez and S. Roura, Optimal sampling strategies in quicksort and quickselect, *SIAM Journal on Computing*, **31** (2001), 683-705.

[84] C. J. H. McDiarmid and R. B. Hayward, Large deviations for quicksort, *Journal of Algorithms*, **21** (1996), 476-507.

[85] A. Moffat and J. Katajainen, In-place calculation of minimum-redundancy codes, in *Proceedings of the Workshop on Algorithms and Data Structures*, Kingston, Ontario, 1995, pp. 393-402.

[86] D. Motzkin, A stable quicksort, *Software–Practice and Experience*, **11** (1981), 607-611.

[87] D. R. Musser, Introspective sorting and selection algorithms, *Software—Practice and Experience*, **27** (1997), 983-993.

[88] R. K. Nagle, E. B. Saff and A. D. Snider, *Fundamentals of Differential Equations and Boundary Value Problems*, Third edition, Addison-Wesley, Reading, MA, 2000.

[89] R. Neininger and L. Rüschendorf, A general contraction theorem and asymptotic normality in combinatorial structures, preprint (2001); available via `http://www.stochastik.uni-freiburg.de/homepages/neininger`.

[90] A. Panholzer and H. Prodinger, An analytic approach for the analysis of rotations in fringe-balanced binary search trees, *Annals of Combinatorics*, **2** (1998), 173–184.

[91] J. Pearl, A space-efficient on-line method of computing quantile estimates, *Journal of Algorithms*, **2** (1981), 164–177.

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

[93] P. V. Poblete, The analysis of heuristics for search trees, *Acta Informatica*, **30** (1993), 233-248.

[94] P. V. Poblete and J. I. Munro, The analysis of a fringe heuristic for binary search trees, *Journal of Algorithms*, **6** (1985), 336-350.

[95] H. Prodinger, Depth and path length of heap ordered trees, *International Journal of Foundations of Computer Science*, **7** (1996), 293-299.

[96] J. Röhrich, A hybrid of quicksort with $O(n \log n)$ worst case complexity, *Information Processing Letters*, **14** (1982), 119-123.

[97] U. Rösler and L. Rüschendorf, The contraction method for recursive algorithms, *Algorithmica*, **29** (2001), 3–33.

[98] P. J. Rousseeuw and G. W. Bassett, The remedian: A robust averaging method for large data sets, *Journal of the American Statistical Association*, **85** (1990), 97-104.

[99] S. M. Sarwar, S. A. Sarwar, M. H. A. Jaragh and J. Brandeburg, Engineering quicksort, *Computer Languages*, **22** (1996), 39-47.

[100] W. Schachinger, Limiting distribution of the costs of partial match retrievals in multidimensional tries, *Random Structures and Algorithms*, **17**, 428–459 (2000).

[101] R. Sedgewick, Quicksort with equal keys, *SIAM Journal on Computing*, **6** (1977), 240–269.

[102] R. Sedgewick, *Quicksort*, Garland, NY, 1980.

[103] S.-H. Shiau and C.-B. Yang, A fast sorting algorithm on broadcast communications, In *Proceedings of the Second International Conference on Parallel Computing Systems (PCS99)*, pp. 87-92, Ensenada, Mexico, August 16-20, 1999.

[104] W. D. Smith, Fixed point for negamaxing probability distributions on regular trees, preprint (1995), available via `http://citeseer.nj.nec.com/67965.html`.

[105] R. T. Smythe, Central limit theorems for urn models, *Stochastic Processes and Applications*, **65** (1996), 115–137.

[106] R. T. Smythe and H. M. Mahmoud, A survey of recursive trees, *Theory of Probability and Mathematical Statistics*, **51** (1996), 1–27.

[107] R. P. Stanley, *Enumerative Combinatorics*, Volume 1, Corrected reprint of the 1986 original, Cambridge University Press, Cambridge, 1997.

[108] C. J. Stephenson, A method of constructing binary search trees by making insertions at the root, *International Journal of Computer and Information Sciences*, **9** (1980), 15–29.

[109] J. L. Szwarcfiter and L. B. Wilson, Some properties of ternary trees, *Computer Journal*, **21** (1978), 66-72.

[110] L. Takács, On the total heights of random rooted binary trees, *Journal of Combinatorial Theory, Series B*, **61** (1994), 155–166.

[111] K.-H. Tan, *An Asymptotic Analysis of the Number of Comparisons in Multipartition Quicksort*, Ph. D. Dissertation, Carnegie-Mellon University, 1993.

[112] J. W. Tukey, The ninther, a technique for low-effort robust (resistant) location in large samples, in *Contributions to Survey Sampling and Applied Statistics in Honor of H. O. Hartley*, Ed. by H. A. David, pp. 251-258, Academic Press, New York, 1978.

[113] J. D. Valois, Introspective sorting and selection revisited, *Software–Practice and Experience*, **30** (2000), 617-638.

[114] J. Vuillemin, A unifying look at data structures, *Communications of the ACM*, **23** (1980), 229–239.

[115] R. L. Wainwright, A class of sorting algorithms based on quicksort, *Communications of the ACM*, **28** (1984), 396-403.

[116] L. M. Wegner, Quicksort for equal keys, *IEEE Transactions on Computers*, **34** (1985), 362-367.

[117] L. M. Wegner, A generalized, one-way, stackless quicksort, *BIT*, **27** (1987), 44-48.

[118] B. W. Weide, *Statistical Methods in Algorithm Design and Analysis*, Ph. D. Dissertation, Carneige-Mellon University, 1978.