ANALYTIC COMBINATORICS
—
COMPLEX ASYMPTOTICS
(Chapters IV, V, VI, VII)

PHILIPPE FLAJOLET & ROBERT SEDGEWICK

Algorithms Project
INRIA Rocquencourt
78153 Le Chesnay
France

Department of Computer Science
Princeton University
Princeton, NJ 08540
USA

Zeroth Edition, May 8, 2004
(This temporary version expires on December 31, 2004)
♡ ♡ ♡ ♡ ♡ ♡

ABSTRACT

This booklet develops in about 240 pages the basics of asymptotic enumeration through an approach that revolves around generating functions and complex analysis. Major properties of generating functions that are of interest here are singularities. The text presents the core of the theory with two chapters on complex analytic methods focusing on rational and meromorphic functions as well as two chapters on fundamentals of singularity analysis and combinatorial consequences. It is largely oriented towards applications of complex analysis to asymptotic enumeration and asymptotic properties of random discrete structures. Many examples are given that relate to words, integer compositions, paths and walks in graphs, lattice paths, constrained permutations, trees, mappings, walks, and maps.

Acknowledgements. This work was supported in part by the IST Programme of the EU under contract number IST-1999-14186 (ALCOM-FT). This booklet would be substantially different (and much less informative) without Neil Sloane's *Encyclopedia of Integer Sequences* and Steve Finch's *Mathematical Constants*, both available on the internet. Bruno Salvy and Paul Zimmermann have developed algorithms and libraries for combinatorial structures and generating functions that are based on the MAPLE system for symbolic computations and have proven to be immensely useful.

This is a set of lecture notes that are a component of a wider book project titled *Analytic Combinatorics*, which will provide a unified treatment of analytic methods in combinatorics. This text contains Chapters IV, V, VI, and VII; it is a continuation of "Analytic Combinatorics—Symbolic Methods" (by Flajolet & Sedgewick, 2002). Readers are encouraged to check Philippe Flajolet's web pages for regular updates and new developments.

PREFACE

Analytic Combinatorics aims at predicting precisely the asymptotic properties of structured combinatorial configurations, through an approach that bases itself extensively on analytic methods. Generating functions are the central objects of the theory.

Analytic combinatorics starts from an exact enumerative description of combinatorial structures by means of generating functions, which make their first appearance as purely formal algebraic objects. Next, generating functions are interpreted as analytic objects, that is, as mappings of the complex plane into itself. In this context, singularities play a key rôle in extracting a function's coefficients in asymptotic form and extremely precise estimates result for counting sequences. This chain is applicable to a large number of problems of discrete mathematics relative to words, trees, permutations, graphs, and so on. A suitable adaptation of the theory finally opens the way to the analysis of parameters of large random structures.

Analytic combinatorics can accordingly be organized based on three components:

— *Symbolic Methods* develops systematic "symbolic" relations between some of the major constructions of discrete mathematics and operations on generating functions which exactly encode counting sequences.

— *Complex Asymptotics* elaborates a collection of methods by which one can extract asymptotic counting informations from generating functions, once these are viewed as analytic transformations of the complex domain (as "analytic" also known as"holomorphic" functions). Singularities then appear to be a key determinant of asymptotic behaviour.

— *Random Structures* concerns itself with probabilistic properties of large random structures—which properties hold with "high" probability, which laws govern randomness in large objects? In the context of analytic combinatorics, this corresponds to a deformation (adding auxiliary variables) and a perturbation (examining the effect of small variations of such auxiliary variables) of the standard enumerative theory.

The approach to quantitative problems of discrete mathematics provided by analytic combinatorics can be viewed as an *operational calculus* for combinatorics. The booklets, of which this is the second installment, expose this view by means of a very large number of examples concerning classical combinatorial structures (like words, trees, permutations, and graphs). What is aimed at eventually is an effective way of quantifying "metric" properties of large random structures. Accordingly, the theory is susceptible to many applications, within combinatorics itself, but, perhaps more importantly, within other areas of science where discrete probabilistic models recurrently surface, like statistical physics, computational biology, or electrical engineering. Last but not least, the analysis of algorithms and data structures in computer science has served and still serves as an important motivation in the development of the theory.

The present booklet specifically exposes *Complex Asymptotics*, which is a unified analytic theory dedicated to the process of extractic asymptotic information from counting generating functions. As it turns out, a collection of general (and simple) theorems provide a systematic translation mechanism between generating functions and asymptotic forms of coefficients. Four chapters compose this booklet. Chapter IV serves as an *introduction to complex-analytic methods* and proceeds with the treatment of *meromorphic functions*, that is, functions whose only singularities are poles, *rational functions* being the simplest case. Chapter V develops *applications of rational and meromorphic asymptotics of generating functions*, with numerous applications related to words and languages, walks and graphs, as well as permutations. Chapter VI develops a general theory of *singularity analysis* that applies to a wide variety of singularity types like square-root or logarithmic and has applications to trees and other recursively defined combinatorial classes. Chapter VII, presents *applications of singularity analysis* to 2-regular graphs and polynomials, trees of various sorts, mappings, context-free languages, walks, and maps. It contains in particular a discussion of the analysis of coefficients of algebraic functions. (A future chapter, Chapter VIII, will explore *saddle point methods*.)

Contents

CHAPTER IV

Complex Analysis, Rational and Meromorphic Asymptotics

The shortest path between two truths in the real domain passes through the complex domain.

—JACQUES HADAMARD [1]

Generating functions are a central concept of combinatorial theory. So far, they have been treated as formal objects, that is, as formal power series. The major theme of Chapters I–III has indeed been to demonstrate how the algebraic structure of generating functions directly reflects the structure of combinatorial classes. From now on, we examine generating functions in the light of *analysis*. This means assigning *values* to the variables that appear in generating functions.

Comparatively little benefit results from assigning only real values to the variable z that figures in a univariate generating function. In contrast assigning *complex* values turns out to have serendipitous consequences. In so doing, a generating function becomes a geometric transformation of the complex plane. This transformation is very regular near the origin—one says that it is analytic or holomorphic. In other words, it only effects initially a smooth distortion of the complex plane.

Farther away from the origin, some "cracks" start appearing in the picture. These cracks—the dignified name is "singularities"—correspond to the disapperance of smoothness. What happens is that knowledge of a function's singularities provide a wealth of information regarding the function's coefficients, and especially their asymptotic rate of growth. Adopting a geometric point of view has a large pay-off.

By focussing on singularities, analytic combinatorics treads in the steps of many respectable older areas of mathematics. For instance, Euler recognized that the fact for the Riemann zeta function $\zeta(s)$ to become infinite at 1 implies the existence of infinitely many prime numbers, while Riemann, Hadamard, and de la Vallée-Poussin uncovered much deeper connections between quantitative properties of the primes and singularities of $1/\zeta(s)$.

In this chapter, we start by recalling the elementary theory of analytic functions and their singularities in a style tuned to the needs of combinatorial theory. Cauchy's integral formula expresses coefficients of analytic functions as contour integrals. Suitable uses of Cauchy's integral formula then make it possible to estimate such coefficients by suitably selecting the contour of integration. For the fairly common case of functions that have singularities at a finite distance, the *exponential growth* formula relates the *location* of the singularities closest to the origin (these are also known as "dominant" singularities) to the exponential order of growth of coefficients. The nature of these *singularities* then dictates

[1] Quoted in The Mathematical Intelligencer, v. 13, no. 1, Winter 1991.

the fine structure of the asymptotic of the function's coefficients, especially the *subexponential factors* involved. In this chapter we carry out this programme for rational functions and meromorphic functions, where the latter are defined by the fact their singularities are of the polar type.

Elementary techniques permit us to estimate asymptotically counting sequences, when these are already presented to us in closed form or as simple combinatorial sums. The methods to be exposed require no such explicit forms of counting coefficients to be available. They apply to almost any conceivable combinatorial generating function that has a decent mathematical expression—we already know from Chapters I–III that this covers a very large fragment of elementary combinatorics. In a large number of cases, complex-analytic methods can even be applied to generating functions only accessible *implicitly* from *functional equations*. This paradigm will be extensively explored in this chapter with applications found in denumerants, derangements, surjections, alignments, and several other structures introduced in Chapters I–III.

IV. 1. Generating functions as analytic objects

Generating functions, considered previously as purely *formal* objects subject to algebraic operations, are now going to be interpreted as *analytic* objects. In so doing one gains an easy access to the asymptotic form of their coefficients. This informal section offer a glimpse of themes that form the basis of this chapter and the next one.

In order to introduce the subject softly, let us start with two simple generating functions, one, $f(z)$, being the OGF of the Catalan numbers (starting at index 1), the other, $g(z)$, being the EGF of derangements:

$$f(z) = \frac{1}{2}\left(1 - \sqrt{1-4z}\right), \qquad g(z) = \frac{\exp(-z)}{1-z}. \tag{1}$$

At this stage, the forms above are merely compact descriptions of formal power series built from the elementary series

$$\begin{aligned} (1-u)^{-1} &= 1 + u + u^2 + \cdots, & (1-u)^{1/2} &= 1 - \frac{1}{2}u - \frac{1}{8}u^2 - \cdots, \\ \exp(u) &= 1 + \frac{1}{1!}u + \frac{1}{2!}u^2 + \cdots, \end{aligned}$$

by standard composition rules. Accordingly, the coefficients of both GFs are known in explicit form

$$f_n := [z^n]f(z) = \frac{1}{n}\binom{2n-2}{n-1}, \qquad g_n := [z^n]g(z) = \left(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!}\right).$$

Next, Stirling's formula and comparison with the alternating series giving $\exp(-1)$ provide respectively

$$f_n \underset{n\to\infty}{\sim} \frac{4^n}{\sqrt{\pi n^3}}, \qquad g_n = \underset{n\to\infty}{\sim} e^{-1} \doteq 0.36787. \tag{2}$$

Our purpose is to examine, heuristically for the moment, the relationship between the asymptotic forms (2) and the structure of the corresponding generating functions in (1).

Granted the growth estimates available for f_n and g_n, it is legitimate to substitute in the power series expansions of the GFs $f(z)$ and $g(z)$ any real or complex value of a small enough modulus, the upper bounds on modulus being $\rho_f = \frac{1}{4}$ (for f) and $\rho_g = 1$ (for g). Figure 1 represents the graph of the resulting functions when such real values are assigned to z. The graphs are smooth, representing functions that are differentiable any

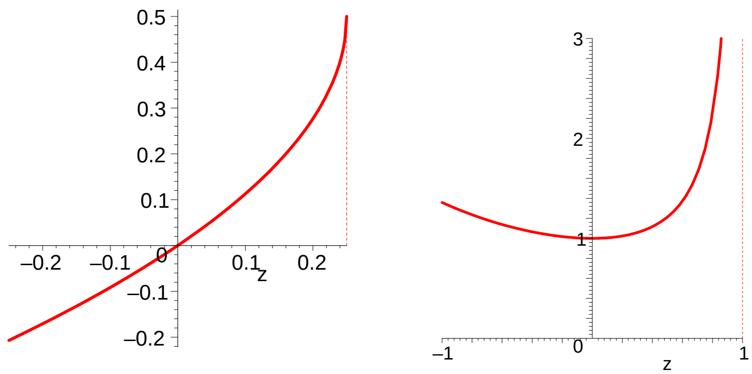


FIGURE 1. Left: the graph of the Catalan OGF, $f(z)$, for $z \in (-\frac{1}{4}, +\frac{1}{4})$; right: the graph of the derangement EGF, $g(z)$, for $z \in (-1, +1)$.

number of times for z interior to the interval $(-\rho, +\rho)$. However, at the right boundary point, smoothness stops: $g(z)$ become infinite at $z = 1$, and so it even ceases to be finitely defined; $f(z)$ does tend to the limit $\frac{1}{2}$ as $z \to (\frac{1}{4})^-$, but its derivative becomes infinite there. Such special points at which smoothness stops are called *singularities*, a term that will acquire a precise meaning in the next sections.

Observe also that, by the usual process of analysis, $f(z)$ and $g(z)$ can be *continued* in certain regions, when use is made of the global expressions (1) while $\exp$ and $\sqrt{\ }$ are assigned their usual real-analytic interpretation; for instance:

$$f(-1) = \frac{1}{2}\left(1 - \sqrt{5}\right), \qquad g(-2) = \frac{e^2}{3}.$$

Such "continuation" properties (to the complex realm) will prove essential in developing efficient methods for coefficient asymptotics.

One may proceed similarly with complex numbers, starting with numbers whose modulus is less than the radius of convergence of the series defining the GF. Figure 2 displays the images of regular grids by f and g. This illustrates the fact that a regular grid transforms into an orthogonal network of curves and more precisely that f and g preserve angles—this property corresponds to complex differentiability and is equivalent to analyticity to be introduced shortly. The singularity of f is clearly perceptible on the right of its diagram, since, at $z = \frac{1}{4}$ corresponding to $f(z) = \frac{1}{2}$, the function f folds lines and divides angles by a factor of 2.

Let us now turn to coefficient asymptotics. As is expressed by (2), the coefficients f_n and g_n each belong to a general asymptotic type,

$$A^n \theta(n),$$

corresponding to an exponential growth factor A^n modulated by a tame factor $\theta(n)$, which is subexponential; compare with (2). Here, one has $A = 4$ for f_n and $A = 1$ for g_n; also, $\theta(n) = \frac{1}{4}(\sqrt{\pi n^3})^{-1}$ for f_n and $\theta(n) = e^{-1}$ for g_n. Clearly, A should be related to the radius of convergence of the series. We shall see that, on very general grounds, the exponential rate of growth is given by $A = 1/\rho$, where ρ is the first singularity encountered along the positive real axis. In addition, under general complex-analytic conditions, it will be established that $\theta(n) = O(1)$ is systematically associated to a simple pole of the generating function, while $\theta(n) = O(n^{-3/2})$ systematically arises from a singularity that

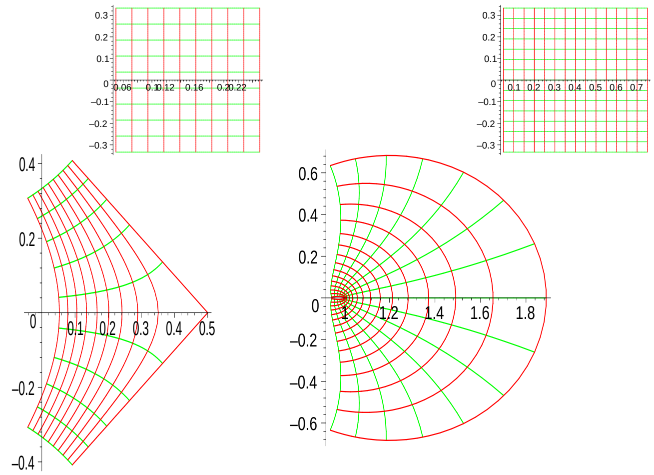


FIGURE 2. The images of regular grids by $f(z)$ (left) and $g(z)$ (right).

is of the square-root type. In summary, as this chapter and the next ones will copiously illustrate, one has:

> **Fundamental principle of complex coefficient asymptotics.** *The* location *of a function's singularities dictates the* exponential growth *of the function's coefficient, A^n, while the* nature *of the function at its singularities determines the* subexponential factor, $\theta(n)$.

Observe that the rescaling rule,

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

enables one to normalize functions so that they are singular at 1, and so "explains" the fact that the location of a function's singularities should influence the coefficients' approximation by exponential factors. Then various theorems, starting with Theorems IV.6 and IV.7, provide sufficient conditions under which the following central implication is valid,

$$h(z) \sim \sigma(z) \quad \Longrightarrow \quad [z^n]h(z) \sim [z^n]\sigma(z), \tag{3}$$

where $h(z)$ is a function singular at 1 whose Taylor coefficients are to be estimated and $\sigma(z)$ is an approximation near a singularity—usually σ is a much simpler function, typically like $(1-z)^\alpha \log^\beta(1-z)$ whose coefficients are easy to find. Under such conditions, it suffices to estimate a function locally in order to derive its coefficients asymptotically. In other words, the relation (3) provides a mapping between asymptotic scales of functions near singularities and asymptotics scales of coefficients.

$\triangleright$ **1.** *Euler, the discrete, and the continuous.* Eulers's proof the existence of infinitely many prime numbers illustrates in a striking manner the way analysis of generating functions can inform us on the discrete realm. Define, for real $s > 1$ the function

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s},$$

now known as the Riemann zeta function. The decomposition (p ranges over the prime numbers $2, 3, 5, \ldots$)

$$
\begin{aligned}
\zeta(s) &= \left(1+\frac{1}{2^s}+\frac{1}{2^{2s}}+\cdots\right)\left(1+\frac{1}{3^s}+\frac{1}{3^{2s}}+\cdots\right)\left(1+\frac{1}{5^s}+\frac{1}{5^{2s}}+\cdots\right)\cdots \\
&= \prod_p \left(1-\frac{1}{p^s}\right)^{-1}
\end{aligned}
\tag{4}
$$

expresses precisely the fact that each integer has a unique decomposition as a product of primes. Analytically, the identity (4) is easily checked to be valid for all $s > 1$. Now suppose that there were only finitely many primes. Let s tend to 1 in (4). Then, the left hand side becomes infinite, while the right hand side tends to the finite limit $\prod_p (1 - 1/p)^{-1}$. A contradiction has been reached. $\triangleleft$

$\triangleright$ **2.** *Elementary transfers.* Elementary series manipulation yield the following general result: *Let $h(z)$ be a power series with radius of convergence > 1 and assume that $h(1) \neq 0$; then one has*

$$[z^n]\frac{h(z)}{1-z} \sim h(1), \qquad [z^n]h(z)\sqrt{1-z} \sim -\frac{h(1)}{2\sqrt{\pi n^3}}, \qquad [z^n]h(z)\log\frac{1}{1-z} \sim \frac{h(1)}{n}.$$

See Bender's survey [**17**] for many similar statements. $\triangleleft$

$\triangleright$ **3.** *Asymptotics of generalized derangements.* The EGF of permutations without cycles of length 1 and 2 satisfies

$$j(z) = \frac{e^{-z-z^2/2}}{1-z} \qquad \text{with} \qquad j(z) \underset{z\to 1}{\sim} \frac{e^{-3/2}}{1-z}.$$

Analogy with derangements suggests (Note 2 can justify it) that $[z^n]j(z) \underset{n\to\infty}{\sim} e^{-3/2}$. Here is a table of exact values of $[z^n]j(z)$ (with relative error of the approximation by $e^{-3/2}$ in parentheses):

	$n = 5$	$n = 10$	$n = 20$	$n = 50$
j_n :	0.2	0.22317	0.2231301600	0.22313016014842982893328047076401 22
error :	(10^{-1})	$(2 \cdot 10^{-4})$	$(3 \cdot 10^{-10})$	(10^{-33})

The quality of the asymptotic approximation is extremely good. (Such a property is invariably attached to polar singularities.) $\triangleleft$

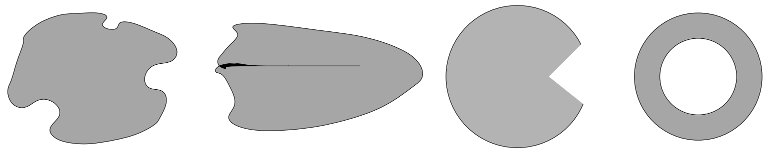
IV. 2. Analytic functions and meromorphic functions

Analytic functions are the primary mathematical concept for complex asymptotics. They can be characterized in two essentially equivalent ways (Subsection IV. 2.1): by means of convergent series expansions (à la Cauchy and Weierstraß) and by differentiability properties (à la Riemann). The first aspect is directly related to the use of generating functions for enumeration; the second one allows for a powerful abstract discussion of closure properties that usually requires little computation. *Meromorphic functions* are nothing but quotients of analytic functions.

Integral calculus with analytic or meromorphic functions (developed in Subsection IV. 2.2) assumes a shape radically different from what it is in the real domain: integrals become quintessentially independent of details of the integration contour, the residue theorem being a prime illustration of this fact. Conceptually, this makes it possible to relate properties of a function at a point (e.g., the coefficients of its expansion at 0) to its properties at another far-away point (e.g., its residue at a pole).

The presentation in this section and the next one is an informal review of basic properties of analytic functions tuned to the needs of asymptotic analysis of counting sequences. For a detailed treatment, we refer the reader to one of the many excellent treatises on the subject, like the books by Dieudonné [**46**], Henrici [**107**], Hille [**108**], Knopp [**117**], Titchmarsh [**181**], or Whittaker and Watson [**191**]. The entry in APPENDIX: *Equivalent* *definitions of analyticity*, p. 220 provides further information, in particular, a proof of the two fundamental principles stated below.

IV. 2.1. Basics. We shall consider functions defined in certain *regions* of the complex domain $\mathbb{C}$. By a region is meant an open subset Ω of the complex plane that is connected. Here are some examples:



simply connected domain   slit complex plane   indented disc   annulus

Classical treatises teach us how to extend to the complex domain the standard functions of real analysis: polynomials are immediately extended as soon as complex addition and multiplication have been defined, while the exponential is definable by means of Euler's formula, and one has for instance

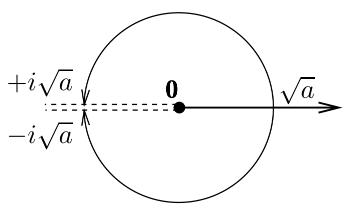
$$z^2 = (x^2 - y^2) + 2ixy, \qquad e^z = e^x \cos y + ie^x \sin y,$$

if $z = x + iy$. Both functions are consequently defined over the whole complex plane $\mathbb{C}$.

The square-root and the logarithm are conveniently described in polar coordinates by

$$\sqrt{z} = \sqrt{\rho}e^{i\theta/2}, \qquad \log z = \log \rho + i\theta, \tag{5}$$

if $z = \rho e^{i\theta}$. One can take the domain of validity of (5) to be the complex plane slit along the axis from 0 to $-\infty$, that is, restrict θ to the open interval $(-\pi, +\pi)$, in which case the definitions above specify what is known as the *principal determination*. There is no way for instance to extend by continuity the definition of $\sqrt{z}$ in any domain containing 0 in its interior since, for $a > 0$ and $z \to -a$, one has $\sqrt{z} \to i\sqrt{a}$ as $z \to -a$ from above, while $\sqrt{z} \to -i\sqrt{a}$ as $z \to -a$ from below. This situation is depicted here:



The values of $\sqrt{z}$
as z varies along $|z| = a$.

The point $z = 0$ where two determinations "meet" is accordingly known as a *branch point*.

First comes the main notion of an analytic function that arises from convergent series expansions and is closely related to the notion of generating function encountered in previous chapters.

DEFINITION IV.1. *A function $f(z)$ defined over a region Ω is* analytic *at a point $z_0 \in \Omega$ if, for z in some open disc centred at z_0 and contained in Ω, it is representable by a convergent power series expansion*

$$f(z) = \sum_{n \geq 0} c_n (z - z_0)^n. \tag{6}$$

A function is analytic in a region Ω iff it is analytic at every point of Ω.

As derives from an elementary property of power series, given a function f that is analytic at a point z_0, there exists a disc (of possibly infinite radius) with the property that the series representing $f(z)$ is convergent for z inside the disc and divergent for z

outside the disc. The disc is called the *disc of convergence* and its radius is the *radius of convergence* of $f(z)$ at $z = z_0$, which will be denoted by $\mathrm{R}_{\mathrm{conv}}(f; z_0)$. Quite elementarily, the radius of convergence of a power series conveys information regarding the rate at which its coefficients grow; see Subsection IV. 3.2 below for developments. It is also easy to prove (see APPENDIX: *Equivalent definitions of analyticity*, p. 220)) that if a function is analytic at z_0, it is thenanalytic at all points interior to its disc of convergence.

Consider for instance the function $f(z) = 1/(1-z)$ defined in the usual way (via complex division) in $\mathbb{C} \setminus \{1\}$. It is analytic at 0 by virtue of the geometric series sum,

$$\frac{1}{1-z} = \sum_{n \geq 0} 1 \cdot z^n,$$

which converges in the disc $|z| < 1$. At a point $z_0 \neq 1$, we may write

$$\begin{aligned} \frac{1}{1-z} &= \frac{1}{1 - z_0 - (z - z_0)} = \frac{1}{1-z_0} \frac{1}{1 - \frac{z-z_0}{1-z_0}} \\ &= \sum_{n \geq 0} \left(\frac{1}{1-z_0} \right)^{n+1} (z - z_0)^n. \end{aligned}$$

The last equation shows that $f(z)$ is analytic in the disc centred at 0 with radius $|1 - z_0|$, that is, the interior of the circle centred at z_0 and passing through the point 1. In particular $\mathrm{R}_{\mathrm{conv}}(f, z_0) = |1 - z_0|$ and $f(z)$ is globally analytic in the punctured plane $\mathbb{C} \setminus \{1\}$.

The last example illustrates the definition of analyticity. At the same time, it suggests that, for more complicated functions, the series rearrangement approach to analyticity might be difficult to carry out. In other words, a higher level approach to analyticity is called for—this is provided by differentiability and integrability properties developed next.

The next important notion is a geometric one based on differentiability.

DEFINITION IV.2. *A function $f(z)$ defined over a region Ω is called* complex-differentiable *(also* holomorphic*) at z_0 if the limit, for complex δ,*

$$\lim_{\delta \to 0} \frac{f(z_0 + \delta) - f(z_0)}{\delta}$$

exists. (In particular, the limit is independent of the way δ tends to 0 in $\mathbb{C}$.) This limit is denoted as usual by $f'(z_0)$ or $\frac{d}{dz} f(z)\big|_{z_0}$ A function is complex-differentiable in Ω iff it is differentiable at every $z_0 \in \Omega$.

Clearly, if $f(z)$ is complex differentiable at z_0, it acts locally as a linear transformation,

$$f(z) - f(z_0) \sim f'(z_0)(z - z_0),$$

whenever $f'(z_0) \neq 0$. Then $f(z)$ behaves in smal regions almost like a similarity transformation (composed of a translation, a rotation, and a scaling). In particular, it preserves angles[2] and infinitesimal squares get transformed into infinitesimal squares; see Figure 3 for a rendering

For instance the function $\sqrt{z}$, defined by (5) in the complex plane slit along the ray $(-\infty, 0)$, is complex-differentiable at any z of the slit plane since

$$\lim_{\delta \to 0} \frac{\sqrt{z+\delta} - \sqrt{z}}{\delta} = \lim_{\delta \to 0} \sqrt{z} \frac{\sqrt{1 + \delta/z} - 1}{\delta} = \frac{1}{2\sqrt{z}},$$

[2] A mapping that preserves angles is also called a *conformal* map.

which extends the customary rule from real analysis. Similarly, $\sqrt{1-z}$ is analytic in the complex plane slit along the ray $(1, +\infty)$. More generally, the usual real analysis proofs carry over almost verbatim to the complex realm, to the effect that

$$(f+g)' = f' + g', \quad (fg)' = f'g + fg', \quad \left(\frac{1}{f}\right)' = -\frac{f'}{f^2}, \quad (f \circ g)' = (f' \circ g)g'.$$

The notion of complex differentiability is thus much more manageable than the notion of analyticity.

It follows from a well known theorem of Riemann (see for instance [**107**, vol. 1, p 143] and APPENDIX: *Equivalent definitions of analyticity*, p. 220) that analyticity and complex differentiability are equivalent notions.

First fundamental property of analytic function theory. *A function is analytic in a region Ω if and only if it is complex-differentiable in Ω.*

The following are known facts (see again the appendix): if a function is analytic (equivalently complex-differentiable) in Ω, it admits (complex) derivatives of any order there. This property markedly differs from real analysis: complex differentiable (equivalently, analytic) functions are "smooth". Also derivatives of a function are obtained through term-by-term differentiation of the series representation of the function.

We finally introduce meromorphic[3] functions that are mild extensions of the concept of analyticity (or holomorphy) and are essential to the theory.

The quotient of two analytic functions $f(z)/g(z)$ ceases to be analytic at a point a where $g(a) = 0$. However, a simple structure for quotients of analytic functions prevails.

DEFINITION IV.3. *A function $h(z)$ is* meromorphic *at $z = z_0$ iff in a neighbourhood of $z = z_0$ with $z \neq z_0$ it is representable by an expansion of the form*

$$h(z) = \sum_{n \geq -M} h_n (z - z_0)^n. \tag{7}$$

If $h_{-M} \neq 0$, then $h(z)$ is said to have a pole of order M at $z = a$. The coefficient h_{-1} is called the residue *of $h(z)$ at $z = a$ and is written as*

$$\mathrm{Res}[h(z); z = a].$$

A function is meromorphic in a region iff it is meromorphic at any point of the region.

Equivalently, $h(z)$ is meromorphic at $z = z_0$ iff, in a neighbourhood of z_0, it can be represented as $f(z)/g(z)$, with $f(z)$ and $g(z)$ being analytic at $z = z_0$.

IV. 2.2. Integrals and residues. Integrals along curves in the complex plane are defined in the usual way from curvilinear integrals applied to the real and imaginary parts of the integrand. However integral calculus in the complex plane is of a radically different nature from what it is on the real line—in a way it is much simpler and much more powerful.

A path in a region Ω is described by its parameterization, which is a continuous function γ mapping $[0, 1]$ into Ω. Two paths γ, γ' in Ω having the same end points are said to be *homotopic* (in Ω) if one can be continuously deformed into the other while staying within Ω as in the following examples:

[3] "Holomorphic" and "meromorphic" are words coming from Greek, meaning respectively "of complete form" and "of partial form".

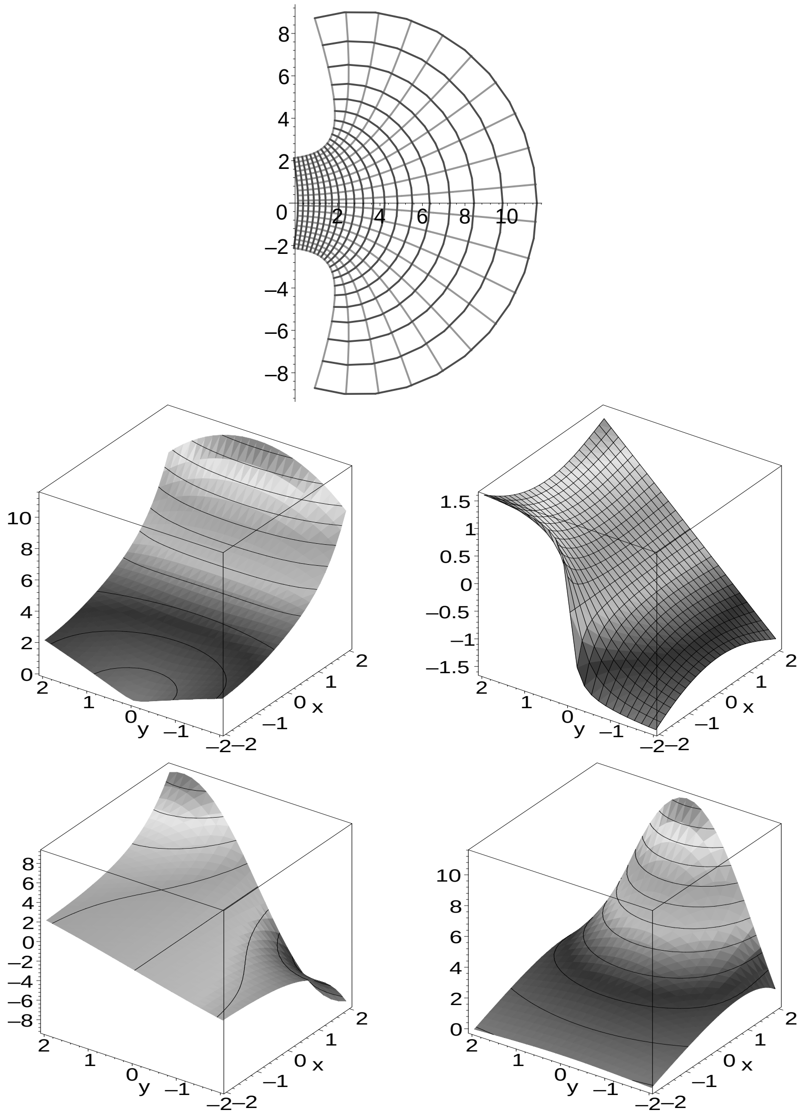
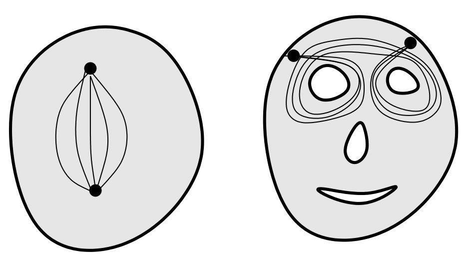


FIGURE 3. Multiple views of an analytic function. The image of the domain $\Omega = \{z \mid |\Re(z)| < 2, |\Im(z)| < 2\}$ by the function $f(z) = \exp(z) + z + 2$: (top) transformation of a square grid in Ω by f; (middle) the modulus and argument of $f(z)$; (bottom) the real and imaginary parts of $f(z)$.

homotopic paths:



A closed path is defined by the fact that its end points coincide: $\gamma(0) = \gamma(1)$, and a path is *simple* if the mapping γ is one-to-one. A closed path is said to be a *loop* of Ω if it can be continuously deformed *within* Ω to a single point; in this case one also says that the path is homotopic to 0. In what follows we implicitly restrict attention to paths that are assumed to be rectifiable. Unless otherwise stated, all integration paths will be assumed to be oriented positively.

One has:

> **Second fundamental property of analytic function theory.** *Let f be analytic in Ω and let λ be a loop of Ω. Then $\int_\lambda f = 0$.*

Equivalently, for f analytic in Ω, one has

$$\int_\gamma f = \int_{\gamma'} f, \tag{8}$$

provided γ and γ' are homotopic in Ω.

The important *Residue Theorem* due to Cauchy relates *global* properties of a meromorphic function, its integral along closed curves, to purely *local* characteristics at designated points, the residues at poles.

THEOREM IV.1 (Cauchy's residue theorem). *Let $h(z)$ be meromorphic in the region Ω and let λ be a simple loop in Ω along which the function is analytic. Then*

$$\frac{1}{2i\pi}\int_\lambda h(z)\,dz = \sum_s \operatorname{Res}[h(z); z = s],$$

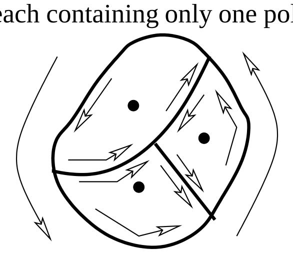
where the sum is extended to all poles s of $h(z)$ enclosed by λ.

PROOF. (Sketch) To see it in the representative case where $h(z)$ has only a pole at $z = 0$, observe by appealing to primitive functions that

$$\int_\lambda h(z)\,dz = \sum_{\substack{n \ge -M \\ n \ne -1}} h_n \left[\frac{z^{n+1}}{n+1}\right]_\lambda + h_{-1}\int_\lambda \frac{dz}{z},$$

where the bracket notation $\left[u(z)\right]_\lambda$ designates the variation of the function $u(z)$ along the contour λ. This expression reduces to its last term, itself equal to $2i\pi h_{-1}$, as is checked by using integration along a circle (set $z = re^{i\theta}$). The computation extends by translation to the case of a unique pole at $z = a$.

In the case of multiple poles, we observe that the simple loop can only enclose finitely many poles (by compactness). The proof then follows from a simple decomposition of the interior domain of λ into cells each containing only one pole. Here is an illustration



in the case of three poles. (Contributions from internal edges cancel.) $\square$

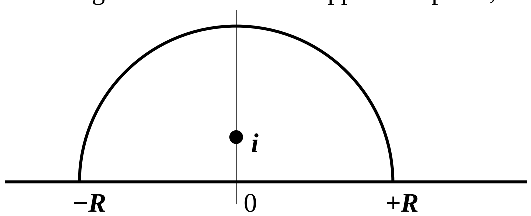
Here is a textbook example of such a reduction from global to local properties. Define the integrals

$$I_m := \int_{-\infty}^{\infty} \frac{dx}{1+x^{2m}},$$

and consider specifically I_1. Elementary calculus teaches us that $I_1 = \pi$ since the anti-derivative of the integrand is an arc tangent:

$$I_1 = \int_{-\infty}^{\infty} \frac{dx}{1+x^2} = [\arctan x]_{-\infty}^{+\infty} = \pi.$$

In the light of the residue theorem, we first consider the integral over the whole line as the limit of integrals over large intervals of the form $[-R, +R]$, then complete the contour of integration by means of a large semi-circle in the upper half-plane, as shown below:



Let γ be the contour comprised of the interval and the semi-circle. Inside γ, the integrand has a pole at $x = i$ ($i = \sqrt{-1}$), where

$$\frac{1}{1+x^2} \equiv \frac{1}{(x+i)(x-i)} = -\frac{i}{2}\frac{1}{x-i} + \frac{1}{4} + \frac{i}{8}(x-i) + \cdots,$$

so that its residue there is $-i/2$. Thus, by the residue theorem, the integral taken over γ is equal to $2\pi i$ times the residue of the integrand at i. As $R \to \infty$, the integral along the semi-circle vanishes (it is $O(R^{-1})$) while the integral along the real segment gives I_1 in the limit. There results the relation giving I_1:

$$I_1 = 2i\pi \operatorname{Res}\left(\frac{1}{1+x^2}, x = i\right) = \pi.$$

Remarkably, the evaluation of the integral in this perspective rests entirely upon the local expansion of the integrand at a special point (the point i).

$\triangleright$ **4.** *The general integral* I_m. Let $\alpha = \exp(\frac{i\pi}{2m})$ so that $\alpha^{2m} = -1$. Contour integration of the type used for I_1 yields

$$I_m = 2i\pi \sum_{j=1}^{m} \operatorname{Res}\left(\frac{1}{1+x^{2m}}; x = \alpha^{2j-1}\right),$$

while, for any $\beta = \alpha^{2j-1}$ with $1 \le j \le m$, one has

$$\frac{1}{1+x^{2m}} \underset{x\to\beta}{\sim} \frac{1}{2m\beta^{2m-1}}\frac{1}{x-\beta} \equiv -\frac{\beta}{2m}\frac{1}{x-\beta}.$$

As a consequence,

$$I_{2m} = -\frac{i\pi}{m}\left(\alpha + \alpha^3 + \cdots + \alpha^{2m-1}\right) = \frac{\pi}{m \sin\frac{\pi}{2m}}.$$

In particular, $I_2 = \pi/\sqrt{2}$, $I_3 = 2\pi/3$, $I_4 = \frac{\pi}{4}\sqrt{2}\sqrt{2+\sqrt{2}}$ as well as $\frac{1}{\pi}I_5$, $\frac{1}{\pi}I_6$ are expressible by radicals, but $\frac{1}{\pi}I_7$, $\frac{1}{\pi}I_9$ are not. The special cases $\frac{1}{\pi}I_{17}$, $\frac{1}{\pi}I_{257}$ are expressible by radicals. $\triangleleft$

▷ **5.** *Integrals of rational fractions.* Generally, all integrals of rational functions taken over the whole real line are computable by residues. In particular,

$$J_m = \int_{-\infty}^{+\infty} \frac{dx}{(1+x^2)^m}, \qquad K_m = \int_{-\infty}^{+\infty} \frac{dx}{(1^2+x^2)(2^2+x^2)\cdots(m^2+x^2)}$$

can be explicitly evaluated. ◁

Many function-theoretic consequences derive from the residue theorem. For instance, if f is analytic in Ω, $z_0 \in \Omega$ and λ is a simple loop of Ω encircling z_0, one has

$$f(z_0) = \frac{1}{2i\pi} \int_\lambda f(\zeta) \frac{d\zeta}{\zeta - z_0}. \tag{9}$$

This follows directly since

$$\text{Res}\left[f(\zeta)(\zeta - z_0)^{-1}; \zeta = z_0\right] = f(z_0).$$

Then, by differentiation with respect to z under the integral sign, one gets similarly

$$\frac{1}{k!} f^{(k)}(z_0) = \frac{1}{2i\pi} \int_\lambda f(\zeta) \frac{d\zeta}{(\zeta - z_0)^k}. \tag{10}$$

The values of a function and its derivatives at a point can thus be obtained as values of integrals of the function away from that point.

A very important application of the residue theorem concerns coefficients of analytic functions.

THEOREM IV.2 (Cauchy's Coefficient Formula). *Let $f(z)$ be analytic in a region containing 0 and let λ be a simple loop around 0 that is oriented positively. Then the coefficient $[z^n]f(z)$ admits the integral representation*

$$f_n \equiv [z^n]f(z) = \frac{1}{2i\pi} \int_\lambda f(z) \, \frac{dz}{z^{n+1}} \, .$$

PROOF. This formula follows directly from the equalities

$$\frac{1}{2i\pi} \int_\lambda f(z) \, \frac{dz}{z^{n+1}} = \text{Res}\left[f(z) z^{-n-1}; z = 0\right] = [z^n]f(z),$$

of which the first follows from the residue theorem, and the second from the identification of the residue at 0 as a coefficient. □

Analytically, the coefficient formula allows one to deduce information about the coefficients from the values of the function itself, using adequately chosen contours of integration. It thus opens the possibility of estimating the coefficients $[z^n]f(z)$ in the expansion of $f(z)$ near 0 by using information on $f(z)$ *away* from 0. The rest of this chapter will precisely illustrate this process in the case of functions whose singularities are poles, that is, rational and meromorphic functions. Note also that the residue theorem provides the simplest known proof of the Lagrange inversion theorem (see the appendices) whose rôle is *inter alia* central to tree enumerations. The supplements below explore some independent consequences of the residue theorem and the coefficient formula.

▷ **6.** *Liouville's Theorem.* If a function $f(z)$ is analytic in the whole of $\mathbb{C}$ and is of modulus bounded by an absolute constant, $|f(z)| \le B$, then it must be a constant. (By trivial bounds, upon integrating on a large circle, it is found that the Taylor coefficients at the origin of index ≥ 1 are all equal to 0.) Similarly, if $f(z)$ is of at most polynomial growth, $|f(z)| \le B\,(|z|+1)^r$ over the whole of $\mathbb{C}$, then it must be a polynomial. ◁

▷ **7.** *Lindelöf integrals.* Let $a(s)$ be analytic in $\Re(s) > \frac{1}{4}$ where it is assumed to satisfy $a(s) = O(\exp((\pi - \epsilon)|s|))$ for some $\epsilon > 0$. Then, one has for $\Re(z) > 0$,

$$\sum_{k=1}^{\infty} a(k)(-z)^k = -\frac{1}{2i\pi} \int_{1/2-i\infty}^{1/2+i\infty} a(s) z^s \, \frac{\pi}{\sin \pi s} \, ds.$$

(Close the integration contour by a large semi-circle on the right.) Such integrals, sometimes called Lindelöf integrals, provide representations for functions determined by an explicit "law" of their Taylor coefficients [**131**].

As a consequence, the generalized *polylogarithm* functions

$$\operatorname{Li}_{\alpha,k}(z) = \sum_{n \geq 1} n^{-\alpha} (\log n)^k z^n$$

are analytic in the complex plane $\mathbb{C}$ slit along $(1 + \infty)$. (More properties can be found in [**66, 86**].) For instance, one finds in this way

$$\text{“} \sum_{n=1}^{\infty} (-1)^n \log n \text{”} = -\frac{1}{8\pi} \int_{-\infty}^{+\infty} \frac{\log(\frac{1}{4} + t^2)}{\cosh(\pi t)} \, dt = 0.22579 \cdots = \log \sqrt{\frac{\pi}{2}},$$

when the divergent series on the left is interpreted as $\operatorname{Li}_{0,1}(-1) = \lim_{z \to -1^+} \operatorname{Li}_{0,1}(z)$. ◁

▷ **8.** *Magic duality.* Let ϕ be a function initially defined over the nonnegative integers but admitting a meromorphic extension over the whole of $\mathbb{C}$. Under conditions analogous to those of Note 7, the function

$$F(z) := \sum_{n \geq 1} \phi(n)(-z)^n,$$

which is analytic at the origin, is such that, near positive infinity,

$$F(z) \underset{z \to +\infty}{\sim} E(z) - \sum_{n \geq 1} \phi(-n)(-z)^{-n},$$

for some "elementary" function $E(z)$. (Starting from the representation of Note 7, close the contour of integration by a large semicircle to the left.) In such cases, the function is said to satisfy the principle of *magic duality*—its expansion at 0 and ∞ are given by one and the same "law". Functions

$$\frac{1}{1+z}, \quad \log(1+z), \quad \exp(-x), \quad \operatorname{Li}_2(-z), \quad \operatorname{Li}_3(-z)$$

satisfy magic duality. Ramanujan [**20**] made a great use of this principle, which applies to a wide class of functions including hypergeometric ones; see [**105**, Ch XI] for an insightful discussion. ◁

▷ **9.** *Euler–Maclaurin and Abel–Plana summations.* Under simple conditions on the analytic function f, one has Plana's (also known as Abel's) complex variables version of the Euler–Maclaurin summation formula:

$$\sum_{n=0}^{\infty} f(n) = \frac{1}{2} f(0) + \int_0^{\infty} f(x) \, dx + \int_0^{\infty} \frac{f(iy) - f(-iy)}{e^{2i\pi y} - 1} \, dy.$$

(See [**107**, Vol. 1, p. 274] for a proof and validity conditions.) ◁

▷ **10.** *Nörlund-Rice integrals.* Let $a(z)$ be analytic for $\Re(z) > k_0 - \frac{1}{2}$ and of at most polynomial growth in this right half plane. Then, with γ a loop around the interval $[k_0, n]$, one has

$$\sum_{k=k_0}^{n} \binom{n}{k} (-1)^{n-k} a(k) = \frac{1}{2i\pi} \int_{\gamma} a(s) \, \frac{n! \, ds}{s(s-1)(s-2)\cdots(s-n)}.$$

If $a(z)$ is meromorphic in a larger region, then the integral can be estimated by residues. For instance, with

$$S_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^k}{k}, \qquad T_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^k}{k^2+1},$$

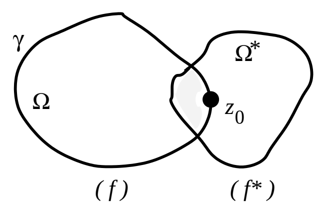
it is found that $S_n = -H_n$ (a harmonic number), while T_n oscillates boundedly as $n \to +\infty$. (This technique is a classical one in the calculus of finite differences, going back to Nörlund [**145**]. In computer science it is known as the method of "Rice's integrals" [**81**] and is used in the analysis of many algorithms and data structures including digital trees and radix sort [**121, 179**].) $\lhd$

IV. 3. Singularities and exponential growth of coefficients

For a given function, a singularity can be informally defined as a point where the function "ceases" to be analytic. Singularities are, as we have stressed repeatedly, essential to coefficient asymptotics. This section presents the bases of a discussion within the framework of analytic function theory.

IV. 3.1. Singularities. Let $f(z)$ be an analytic function defined over the interior region determined by a simple closed curve γ, and let z_0 be a point of the bounding curve γ. If there exists an analytic function $f^*(z)$ defined over some open set Ω^* containing z_0 and such that $f^*(z) = f(z)$ in $\Omega^* \cap \Omega$, one says that f is *analytically continuable* at z_0 and that $f^\star$ is an *immediate analytic continuation* of f.

Analytic continuation:



$f^*(z) = f(z)$ on $\Omega^* \cap \Omega$.

Consider for instance the function $f(z)$ defined by the power series $f(z) = \sum_{n\geq 0} z^n$ and take $z_0 = e^{i\theta}$ with $\theta \neq 0$. Simple series rearrangements yield

$$f(z) = \frac{1}{1 - z_0 - (z - z_0)} = \sum_{n\geq 0} (1 - z_0)^{-n-1}(z - z_0)^n.$$

This shows that $f(z)$ is analytically continuable in $|z - z_0| < |1 - z_0|$, that is, inside the circle of center z_0 passing through 1. Using the notion of differentiability, we also verify that the function $f(z) = \sqrt{1-z}$ is analytic at all points of $|z| = 1, z \neq 1$.

In sharp contrast to real analysis where a function admits of many smooth extensions, analytic continuation is essentially *unique*: for instance, if f^* and f^{**} continue f at z_0, then one must have $f^*(z) = f^{**}(z)$ in the vicinity of z_0. Thus, the notion of immediate analytic continuation is intrinsic. Also the process can be iterated and we say that g is an *analytic continuation*[4] of f, even if their domains of definition do not overlap, provided a finite chain of intermediate function elements connects f and g. This notion is once more intrinsic—this is known as the principle of *unicity of analytic continuation* (along paths). An analytic function is then much like a hologram: as soon as it is specified in any tiny region, it is rigidly determined in any wider region where it can be continued.

DEFINITION IV.4. *Given an f defined in the region interior to γ, a point z_0 on the boundary of the region is a* singular point *or a* singularity[5] *of f if f is* not *analytically continuable at z_0.*

Granted the intrinsic character of analytic continuation, we can usually dispense with a detailed description of the original domain Ω and the curve γ. In simple terms, a function

---

[4]The collection of all function elements continuing a given function gives rise to the notion of Riemann surface, for which manty good books exist, e.g., [**55, 173**]. We shall normally avoid appealing to this theory.

[5]For a detailed discussion, see [**46**, p. 229], [**117**, vol. 1, p. 82], or [**181**].

is singular at z_0 if it cannot be continued as an analytic function beyond z_0. A point at which a function is analytic is also called by contrast a *regular point*.

The two functions $f(z) = 1/(1-z)$ and $g(z) = \sqrt{1-z}$ may be taken as initially defined over the open unit disk by their power series representation. Then, as we already know, they can be analytically continued to larger regions, the punctured plane $\Omega = \mathbb{C}\setminus\{1\}$ for f and the complex plane slit along $(1, +\infty)$ for g. (This is achieved by the usual operations of analysis, upon taking inverses and square roots.) But both are singular at 1: for f, this results from the fact that (say) $f(z) \to \infty$ as $z \to 1$; for g this is due to the branching character of the square-root.

It is easy to check from the definitions that a converging Taylor series is analytic inside its disc of convergence. In other words, it can have no singularity inside this disc. However, it *must* have one on the boundary of the disc, as asserted by the theorem below. In addition, a classical theorem, called Pringsheim's theorem [**181**, Sec. 7.21], provides a refinement of this property in the case of functions with nonnegative coefficients.

THEOREM IV.3 (Boundary singularities). (i) *A function analytic f at the origin whose Taylor expansion at* 0 *has a finite radius of convergence R necessarily has a singularity on the boundary of its disc of convergence, $|z| = R$.*

(ii) [Pringsheim's Theorem] *If in addition f has nonnegative Taylor coefficients, then the point $z = R$ is a singularity of f.*

Define the *radius of singularity* of $f(z)$, written $\mathrm{R}_{\mathrm{sing}}(f; z_0)$ as the supremum,

$$\mathrm{R}_{\mathrm{sing}}(f; z_0) := \sup\left\{r > 0 \mid f \text{ has no singularity in the disc } D(z_0, r)\right\}.$$

In other word "radius of singularity" means the first radius at which a singularity appears. A figurative way of expressing Theorem IV.4, (i) is as follows:

The radius of convergence is identical to the radius of singularity:

$$\mathrm{R}_{\mathrm{conv}}(f; z_0) \equiv \mathrm{R}_{\mathrm{sing}}(f; z_0). \tag{11}$$

This fact together with Pringsheim's Theorem is central to asymptotic enumeration as the remainder of this section will demonstrate.

PROOF. (i) Let $f(z)$ be the function and R the radius of convergence of its Taylor series at 0, taken under the form

$$f(z) = \sum_{n\geq 0} f_n z^n. \tag{12}$$

We now that there can be no singularity of f within the disc $|z| < R$. Suppose a contrario that $f(z)$ is analytic in the whole of $|z| < \rho$ for some ρ satisfying $\rho > R$. By Cauchy's coefficient formula (theorem IV.2), upon integrating along the circle λ of radius $r = (R + \rho)/2$, it is seen that the coefficient $[z^n]f(z)$ is $O(r^{-n})$. But then, the series expansion of f would have to converge in the disc of radius $r > R$, a contradiction. (More on this theme below.)

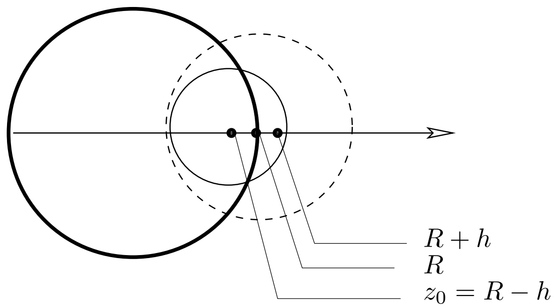
(ii) Suppose a contrario that $f(z)$ is analytic at R, implying that it is analytic in a disc of radius r centred at R. We choose a number h such that $0 < h < \frac{1}{3}r$ and consider the expansion of $f(z)$ around $z_0 = R - h$:

$$f(z) = \sum_{m\geq 0} g_m (z - z_0)^m. \tag{13}$$

By Taylor's formula and the representability of $f(z)$ together with its derivatives at z_0 by means of (12), we have

$$g_m = \sum_{n\geq 0} \binom{n}{m} f_n z_0^m,$$

and in particular, $g_m \geq 0$. By the way h was chosen, the series (13) converges at $z - z_0 = 2h$, as illustrated by the following diagram:



Consequently, one has

$$f(R+h) = \sum_{m\geq 0} \left(\sum_{n\geq 0} \binom{n}{m} f_n z_0^{m-n} \right) (2h)^m.$$

This is a converging double sum of positive terms, so that the sum can be reorganized in any way we like. In particular, one has convergence of all the series involved in

$$\begin{aligned} f(R+h) &= \sum_{m,n\geq 0} \binom{n}{m} f_n (R-h)^{m-n} (2h)^m \\ &= \sum_{n\geq 0} f_n \left[(R-h) + (2h)\right]^n \\ &= \sum_{n\geq 0} f_n (R+h)^n. \end{aligned}$$

This establishes the fact that $f_n = o((R+h)^n)$, thereby reaching a contradiction. Pringsheim's theorem is proved. $\square$

Singularities of a function analytic at 0 which are on the boundary of the disc of convergence are called *dominant singularities*. The second part of this theorem appreciably simplifies the search for dominant singularities of combinatorial generating functions since these have nonnegative coefficients.

For instance, the derangement OGF and the surjection EGF,

$$D(z) = \frac{e^{-z}}{1-z}, \qquad S(z) = (2 - e^z)^{-1}$$

are analytic except for a simple pole at $z = 1$ in the case of $D(z)$, and except for points $z_k = \log 2 + 2ik\pi$ that are simple poles in the case of $S(z)$. Thus the dominant singularities for derangements and surjections are at 1 and $\log 2$ respectively.

It is known that $\sqrt{Z}$ cannot be unambiguously defined as an analytic function in a neighbourhood of $Z = 0$. As a consequence, the function

$$C(z) = (1 - \sqrt{1-4z})/2,$$

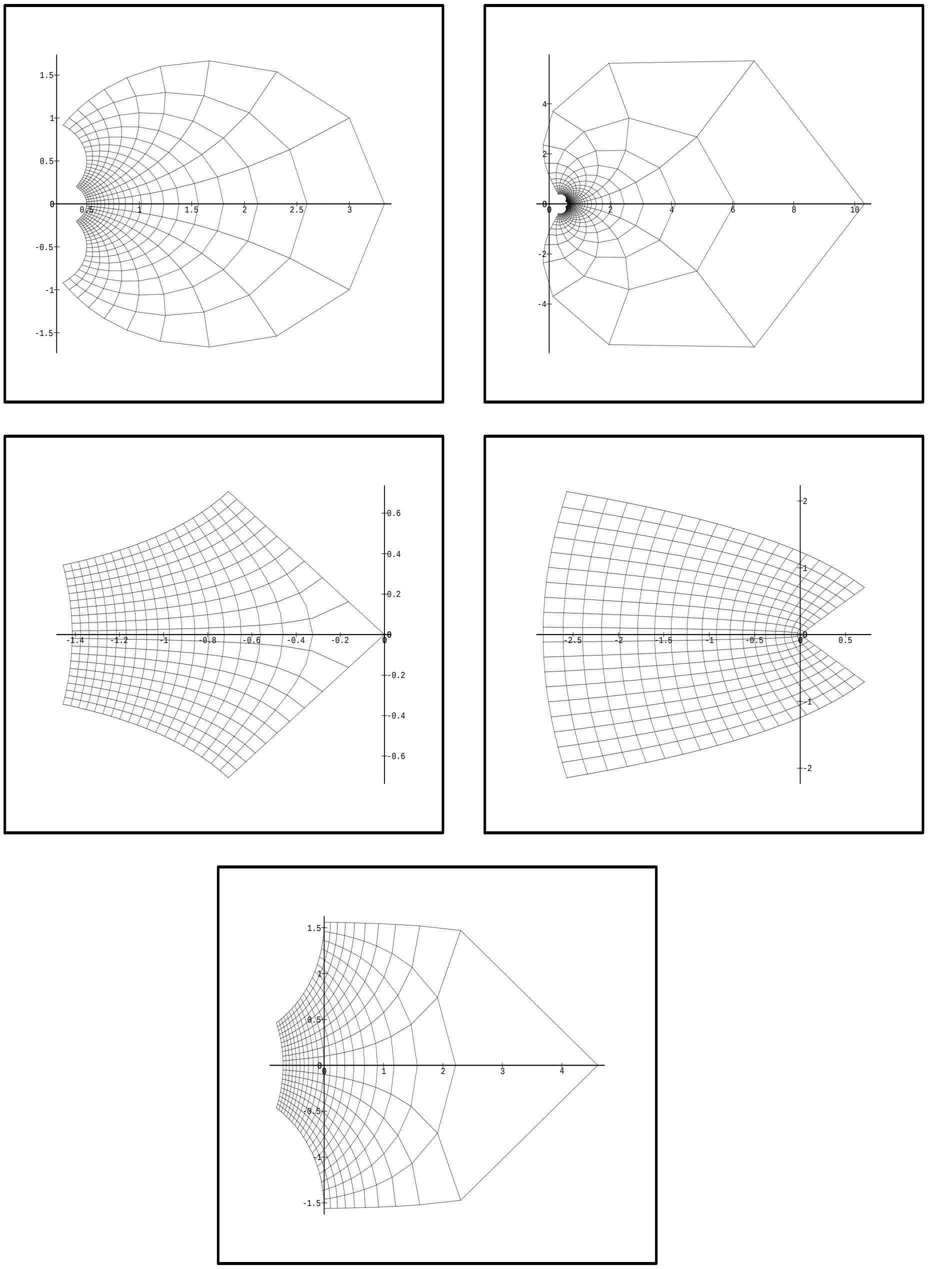


FIGURE 4. The images of a grid on the unit square (with corners $\pm 1 \pm i$) by various functions singular at $z = 1$ reflect the nature of the singularities involved. Here (from top to bottom) $f_0(z) = 1/(1-z)$, $f_1(z) = \exp(z/(1-z))$, $f_2(z) = -(1-z)^{1/2}$, $f_3(z) = -(1-z)^{3/2}$, $f_4(z) = \log(1/(1-z))$. The functions have been normalized to be increasing over the real interval $[-1, 1]$. Singularities are apparent near the right of each diagram where small grid squares get folded or unfolded in various ways. (In the case of functions f_0, f_1, f_4 that become infinite at $z = 1$, the grid has been slightly truncated to the right.)

which is the generating function of the Catalan numbers, is an analytic function in certain regions that should exclude $1/4$; for instance, one may opt to take the complex plane slit along the ray $(1/4, +\infty)$. Similarly, the function

$$L(z) = \log \frac{1}{1-z}$$

which is the EGF of cyclic permutations is analytic in the complex plane slit along $(1, +\infty)$. (An alternative way of seeing that $C(z)$ and $L(z)$ are singular at $\frac{1}{4}$ and 1 is to observe that their derivatives become infinite along rays $z \to \frac{1}{4}^-$ and $z \to 1^-$.)

A function having no singularity at a finite distance is called *entire*; its Taylor series then converges everywhere in the complex plane. The EGFs,

$$e^{z+z^2/2},\ e^{e^z-1},$$

associated to involutions and set partitions are entire.

IV. 3.2. The Exponential Growth Formula. We say that a number sequence $\{a_n\}$ is of *exponential order* K^n which we abbreviate as (the symbol $\bowtie$ is a "bowtie")

$$a_n \bowtie K^n \qquad \text{iff} \qquad \limsup |a_n|^{1/n} = K.$$

The relation $X \bowtie Y$ reads as "X *is of exponential order* Y". In other words, for any $\epsilon > 0$:

$|a_n| >_{i.o} (K - \epsilon)^n$, that is to say, $|a_n|$ exceeds $(K - \epsilon)^n$ infinitely often (for infinitely many values of n);

$|a_n| <_{a.e.} (K + \epsilon)^n$, that is to say, $|a_n|$ is dominated by $(K + \epsilon)^n$ almost everywhere (except for possibly finitely many values of n).

This relation can be rephrased as $a_n = \vartheta(n)K^n$, where ϑ is a *subexponential factor* satisfying

$$\limsup |\theta(n)|^{1/n} = 1;$$

such a factor is thus bounded from above almost everywhere by any increasing exponential (of the form $(1+\epsilon)^n$) and bounded from below infinitely often by any decaying exponential (of the form $(1-\epsilon)^n$). Typical subexponential factors are

$$1,\ n^3,\ (\log n)^2,\ \sqrt{n},\ \frac{1}{\sqrt[3]{\log n}},\ n^{-3/2},\ \log\log n.$$

Note that functions like $e^{\sqrt{n}}$ and $\exp(\log^2 n)$ must be treated as subexponential factors for the purpose of this discussion. Also the $\limsup$ definition allows in principle for factors that are infinitely often very small or 0, like $n^2 \sin n\frac{\pi}{2}$, $\log n \cos \sqrt{n}\frac{\pi}{2}$, and so on. In this and the next chapters, we shall see general methods that enable one to extract such subexponential factors from generating functions.

It is an elementary observation that the radius of convergence of the series representation of $f(z)$ at 0 is related to the exponential growth rate of the coefficients $f_n = [z^n]f(z)$. To wit, if $\mathrm{R}_{\mathrm{conv}}(f; 0) = R_0$, then we claim that

$$f_n \bowtie \left(\frac{1}{R_0}\right)^n. \tag{14}$$

This only requires the basic definition of a power series, the proof involving two simple steps.

— By definition of the radius of convergence, we have for any small $\epsilon > 0$, $f_n(R_0 - \epsilon)^n \to 0$. In particular, $|f_n|(R_0 - \epsilon)^n < 1$ for all sufficiently large n, so that $|f_n| < (R_0 - \epsilon)^{-1}$ "almost verywhere".

— In the other direction, for any $\epsilon > 0$, $|f_n|(R_0+\epsilon)^n$ cannot be a bounded sequence, since otherwise, $|f_n|(R_0 + \epsilon/2)^n$ would be a convergent series. Thus, $|f_n| > (R_0 + \epsilon)^{-1}$ "infinitely often".

For reasons already discussed, a more global approach to the determination of growth rates is desirable. This is precisely provided by the equivalence between radius of convergence and radius of singularity, as asserted in (11). As an immediate consequencne of (11) and (14), we get:

THEOREM IV.4 (Exponential Growth Formula). *If $f(z)$ is analytic at 0 and $R := \mathrm{R}_{\rm sing}(f; z_0)$ is the modulus of a singularity of $f(z)$ nearest to the origin, then the coefficient $f_n = [z^n]f(z)$ satisfies*

$$f_n \bowtie \left(\frac{1}{R}\right)^n, \qquad \textit{equivalently} \quad f_n = \left(\frac{1}{R}\right)^n \theta(n) \textit{ with } \limsup |\theta(n)|^{1/n} = 1.$$

The exponential growth formula thus directly relates the exponential order of growth of coefficients of a function to the *location* of its singularities nearest to the origin. Several direct applications to combinatorial enumeration are given below.

EXAMPLE 1. *Exponential growth and combinatorial enumeration*. Here are a few immediate applications of of exponential bounds to surjections, derangements, integer partitions, and unary binary trees.

Surjections. The function

$$R(z) = (2 - e^z)^{-1}$$

is the EGF of surjections. The denominator is an entire function, so that singularities may only arise from its zeros, to be found at the points

$$\chi_k = \log 2 + 2ik\pi, \qquad k \in \mathbb{Z}.$$

The dominant singularity of R is then at $\rho = \chi_0 = \log 2$. Thus, with $r_n = [z^n]R(z)$,

$$r_n \bowtie (\frac{1}{\log 2})^n.$$

Similarly, if "double" surjections are considered (each value in the range of the surjection is taken at least twice), the corresponding EGF is

$$R^*(z) = \frac{1}{2 - z - e^z};$$

the dominant singularity is at ρ^* defined as the positive root of equation $e^{\rho^*} - \rho^* = 2$, and the coefficient r_n^* satisfies: $r_n^* \bowtie (\frac{1}{\rho^*})^n$ Numerically, this gives

$$r_n \bowtie 1.44269^n \qquad \text{and} \qquad r_n^* \bowtie 0.87245^n,$$

with the actual figures for the corresponding logarithms being

n	$\frac{1}{n}\log r_n$	$\frac{1}{n}\log r_n^*$
10	0.33385	0.80208
20	0.35018	0.80830
50	0.35998	0.81202
100	0.36325	0.81327
∞	0.36651	0.81451
	$(\log 1/\rho)$	$(\log(1/\rho^*))$

These estimates constitutes a weak form of a more precise result to be established later in this chapter: If random surjections of size n are taken equally likely, the probability of a surjection being a double surjection is exponentially small.

Derangements. There, for $d_{1,n} = [z^n]e^{-z}(1-z)^{-1}$ and $d_{2,n} = [z^n]e^{-z-z^2/2}(1-z)^{-1}$ we have, from the poles at $z = 1$,

$$d_{1,n} \bowtie 1^n \quad \text{and} \quad d_{2,n} \bowtie 1^n.$$

The upper bound is combinatorially trivial. The lower bound expresses that the probability for a random permutation to be a derangement is *not* exponentially small. For $d_{1,n}$, we have already proved by an elementary argument the stronger result $d_{1,n} \to e^{-1}$; in the case of $d_{2,n}$, we shall establish later the precise asymptotic equivalent $d_{2,n} \to e^{-3/2}$, in accordance with what was announced in the introduction.

Unary-Binary trees. The expression

$$U(z) = \frac{1-z-\sqrt{1-2z-3z^2}}{2z} = z + z^2 + 2\,z^3 + 4\,z^4 + 9\,z^5 + \cdots,$$

represents the OGF of (plane unlabelled) unary-binary trees. From the equivalent form,

$$U(z) = \frac{1-z-\sqrt{(1-3z)(1+z)}}{2z},$$

it follows that $U(z)$ is analytic in the complex plane slit along $(\frac{1}{3}, +\infty)$ and $(-\infty, -1)$ and is singular at $z = -1$ and $z = 1/3$ where it has branch points. The closest singularity to the origin being at $\frac{1}{3}$, one has

$$U_n \bowtie 3^n.$$

In this case, the stronger upper bound $U_n \le 3^n$ results directly from the possibility of encoding such trees by words over a ternary alphabet using Łukasiewicz codes (Chapter I). A complete asymptotic expansion will be obtained in the next chapter. $\square$

$\triangleright$ **11.** *Coding theory bounds.* Let $\mathcal{C}$ be a combinatorial class. We say that it *can be encoded with $f(n)$ bits* if, for all sufficiently large values of n, elements of $\mathcal{C}_n$ can be encoded as words of $f(n)$ bits. Assume that $\mathcal{C}$ has ogf $C(z)$ with radius of singularity R satisfying $0 < R < \infty$. Then, for any ϵ, $\mathcal{C}$ can be encoded with $(1+\epsilon)\kappa n$ bits where $\kappa = -\log R$, but $\mathcal{C}$ cannot be encoded with $(1-\epsilon)\kappa n$ bits.

Similarly, if $\mathcal{C}$ has egf $\widehat{C}(z)$ with radius of singularity R satisfying $0 < R < \infty$, $\mathcal{C}$ can be encoded with $n\log(n/e) + (1+\epsilon)\kappa n$ bits where $\kappa = -\log R$, but $\mathcal{C}$ cannot be encoded with $n\log(n/e) + (1-\epsilon)\kappa n$ bits. $\triangleleft$

The exponential growth formula expressed by Theorem IV.4 can be supplemented by effective upper bounds which are very easy to derive and often turn out to be surprisingly accurate. We state:

PROPOSITION IV.1 (Saddle-Point bounds). *Let $f(z)$ be analytic in the disc $|z| < R$ with $0 < R \le \infty$. Define $M(f;r)$ for $r \in (0,R)$ by $M(f;r) := \sup_{|z|=r}|f(z)|$. Then, one has, for any r in $(0,R)$, the family of* saddle point upper bounds

$$(15) \qquad [z^n]f(z) \;\le\; \frac{M(f;r)}{r^n} \quad \textit{(any } r\textit{)}, \quad \textit{and} \quad [z^n]f(z) \;\le\; \inf_{s\in(0,R)} \frac{M(f;s)}{s^n}.$$

If in addition $f(z)$ has nonnegative coefficients at 0, then

$$(16) \qquad [z^n]f(z) \;\le\; \frac{f(r)}{r^n} \quad \textit{(any } r\textit{)}, \quad \textit{and} \quad [z^n]f(z) \;\le\; \inf_{s\in(0,R)} \frac{f(s)}{s^n}.$$

PROOF. The first bound in (15) results from trivial bounds applied to the Cauchy coefficient formula, when integration is performed along a circle:

$$[z^n] := \frac{1}{2i\pi}\int_{|z|=r} f(z)\,\frac{dz}{z^{n+1}}.$$

It is consequently valid for any r smaller than the radius of convergence of f at 0. The second inequality represents the best possible bound of this type, but the optimization problem is in general hard to solve.

The bounds (16) can be viewed as a dirceted specialization of (15). Alternatively, they can be obtained elementarily since, in the case of positive coefficients,

$$f_n \le \frac{f_0}{r^n} + \cdots + \frac{f_{n-1}}{r^{n-1}} + f_n + \frac{f_{n+1}}{r^{n+1}} + \cdots\,,$$

whenever the f_k are nonnegative. The best possible bound of thuis type is then given by the second inequality. It can be determined by cancelling a derivative,

$$s\ :\ s\frac{f'(s)}{f(s)} = n.$$

Note that because of the first inequality, any approximate solution of this last equation will in fact provide a valid upper bound. □

For reasons well explained by the saddle point method (Chapter VI), these bounds usually capture the actual asymptotic behaviour up to a polynomial factor only. A typical instance is the weak form of Stirling's formula,

$$\frac{1}{n!} \equiv [z^n]e^z \le \frac{e^n}{n^n},$$

which only overestimates the true asymptotic value by a factor of $\sqrt{2\pi n}$.

▷ **12.** *A suboptimal but easy saddle-point bound.* Let $f(z)$ be analytic in $|z| < 1$ with nonnegative coefficients. Assume that $f(x) \le (1-x)^{-\beta}$ for some $\beta \ge 0$ and all $x \in (0,1)$. Then

$$[z^n]f(z) = O(n^\beta).$$

(In fact, better bounds of the form $O(n^{\beta-1})$ are usually obtained by the method of singularity analysis exposed in Chapter VI.) ◁

EXAMPLE 2. *Combinatorial examples of saddle point bounds.* Here are applications to fragmented permutations, set partitions (Bell numbers), involutions, and integer partitions.

Fragmented permutations. Consider first the EGF of "fragmented permutations" (Chapter II) defined by $\mathcal{F} = \mathfrak{P}(\mathfrak{S}_{\ge 1}(\mathcal{Z}))$ in the labelled universe. We claim that

$$\frac{1}{n!}F_n \equiv [z^n]e^{z/(1-z)} \le e^{2\sqrt{n}-\frac{1}{2}+O(n^{-1/2})}. \tag{17}$$

Indeed, the minimizing value of r in (16) is r_0 such that

$$0 = \frac{d}{dr}\left(\frac{r}{1-r} - n\log r\right)_{r=r_0} = \frac{1}{(1-r_0)^2} - \frac{n}{r_0}.$$

The equation is solved by $r_0 = (2n+1-\sqrt{4n+1})/(2n)$. One can either use this exact value and perform asymptotic approximation of $f(r_0)/z_0^n$, or adopt the approximate value $r_1 = 1 - 1/\sqrt{n}$, which leads to simpler calculations. The estimate (17) results.

n	$\widetilde{I}_n$	I_n
100	$0.106579 \cdot 10^{85}$	$0.240533 \cdot 10^{83}$
200	$0.231809 \cdot 10^{195}$	$0.367247 \cdot 10^{193}$
300	$0.383502 \cdot 10^{316}$	$0.494575 \cdot 10^{314}$
400	$0.869362 \cdot 10^{444}$	$0.968454 \cdot 10^{442}$
500	$0.425391 \cdot 10^{578}$	$0.423108 \cdot 10^{576}$

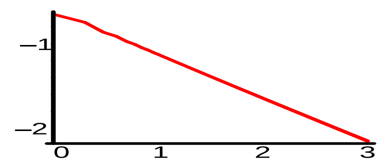


FIGURE 5. The comparison between the exact number of involutions I_n and its approximation $\widetilde{I}_n = n! e^{\sqrt{n}+n/2} n^{-n/2}$: (left) a table; (right) a plot of $\log_{10}(I_n/\widetilde{I}_n)$ against $\log_{10} n$ suggesting that the ratio is $\sim K \cdot n^{-1/2}$.

Bell numbers and set partitions. Another immediate applications is an upper bound on Bell numbers enumerating set partitions with EGF e^{e^z-1}. The best saddle point bound is

$$\frac{1}{n!} B_n \le e^{e^r - 1 - n\log r}, \qquad r \,:\, re^r = n, \tag{18}$$

with $r \sim \log n - \log\log n$.

Involutions. Regarding involutions, their EGF is $I(z) = \exp(z + \frac{1}{2}z^2)$, and one determines (see Figure 5 for numerical data)

$$\frac{1}{n!} I_n \le \frac{e^{\sqrt{n}+n/2}}{n^{n/2}}. \tag{19}$$

Similar bounds hold for permutations with all cycle lengths $\le k$ and permutations σ such that $\sigma^k = Id$.

Integer partitions. The function

$$P(z) = \prod_{k=1}^{\infty} \frac{1}{1-z^k} = \exp\left(\sum_{\ell=1}^{\infty} \frac{1}{\ell} \frac{z^\ell}{1-z^\ell}\right) \tag{20}$$

is the OGF of integer partitions, the unlabelled analogue of set partitions. Its radius of convergence is *a priori* bounded from above by 1, since the set $\mathcal{P}$ is infinite and the second form of $P(z)$ shows that it is exactly equal to 1. Therefore $P_n \bowtie 1^n$. A finer upper bound results from the estimate

$$\Lambda(t) := \log P(e^{-t}) \sim \frac{\pi^2}{6t} + \log\sqrt{\frac{t}{2\pi}} - \frac{1}{24}t + O(t^2), \tag{21}$$

which obtains from Euler–Maclaurin summation or, better, from a Mellin analysis following APPENDIX: *Mellin transform*, p. 227. Indeed, the Mellin transform of Λ is, by the harmonic sum rule,

$$\Lambda^\star(s) = \zeta(s)\zeta(s+1)\Gamma(s), \qquad s \in \langle 1, +\infty\rangle,$$

and the successive leftmost poles at $s = 1$ (simple pole), $s = 0$ (double pole), and $s = -1$ (simple pole) translate into the asymptotic expansion (21). When $z \to 1^-$, this means that

$$P(z) \sim \frac{e^{-\pi^2/12}}{\sqrt{2\pi}} \sqrt{1-z}\, \exp\left(\frac{\pi^2}{6(1-z)}\right), \tag{22}$$

from which we derive the upper bound,

$$P_n \le C n^{1/4} e^{\pi\sqrt{2n/3}}$$

(for some $C > 0$) in a way analogous to fragmented permutations above. This last bound loses only a polynomial factor, as we shall prove when studying the saddle point method in Chapter VIII. $\square$

$\triangleright$ **13.** *A natural boundary.* One has $P(re^{i\theta}) \to \infty$ as $r \to 1^-$, for any angle θ that is a rational multiple of 2π. Such points being dense on the unit circle, the function $P(z)$ admits the unit circle as a *natural boundary*, i.e., it cannot be analytically continued beyond this circle. $\triangleleft$

$\triangleright$ **14.** *Partitions into pwers and Meinardus' method.* The combination of Mellin transforms and saddle point analysis in the theory of partitions is known as Meinardus' method [**5**, Ch. 6]. Consider the set $\mathcal{R}$ of partition into rth powers ($r \geq 2$). The OGF satisfies

$$\Lambda(t) := \log R(e^{-t}) = \sum_{\ell \geq 1} \frac{1}{\ell} \frac{e^{-\ell^r t}}{1 - e^{-\ell^r t}},$$

with Mellin transform $\Lambda^\star(s) = \zeta(rs)\zeta(s+1)\Gamma(s)$ defined for $\Re(s) > r^{-1}$. From the pole of $\Lambda^\star$ at $s = 1/r$, one gets

$$R(z) = \exp\left(\frac{\xi}{(1-z)^{1/r}}\right)(1 + o(1)), \qquad \xi := \frac{1}{r}\zeta(1 + \frac{1}{r})\Gamma(\frac{1}{r}).$$

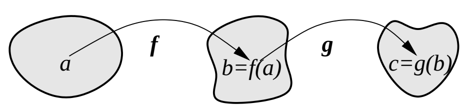
The minimizing value s_0 for saddle point bounds satisfies $1 - s_0$ $(rn/\xi)^{-r/(r+1)}$, and

$$\log R_n \leq C n^{\frac{1}{r+1}} (1 + o(1))$$

(for some $C > 0$). See Andrews' book [**5**, Ch. 6] for precise asymptotics and a general setting. $\triangleleft$

IV. 3.3. Closure properties and computable bounds. The functions analytic at a point $z = a$ are closed under sum and product, and hence form a ring. If $f(z)$ and $g(z)$ are analytic at $z = a$, then so is their quotient $f(z)/g(z)$ provided $g(a) \neq 0$. Meromorphic functions are furthermore closed under quotient and hence form a field. Such properties are proved most easily using complex-differentiability and extending the usual relations from real analysis, for instance, $(f+g)' = f' + g'$, $(fg)' = fg' = f'g$.

Analytic functions are also closed under composition: if $f(z)$ is analytic at $z = a$ and $g(w)$ is analytic at $b = f(a)$, then $g \circ f(z)$ is analytic at $z = a$. Graphically:



The proof based on complex-differentiability closely mimicks the real case. Inverse functions exist conditionally: if $f'(a) \neq 0$, then $f(z)$ is locally linear near a, hence invertible, so that there exists a g satisfying $f \circ g = g \circ f = Id$, where Id is the identity function, $Id(z) \equiv z$. The inverse function is itself locally linear, hence complex differentiable, hence analytic. In short, the inverse of an analytic function f at a place where its derivative does not vanish is an analytic function.

One way to establish closure properties, as suggested above, is to deduce analyticity criteria from complex differentiability by way of the "First Fundamental Property". An alternative approach, closer to the original notion of analyticity, can be based on a two-step process: (i) closure properties are shown to hold true for formal power series; (ii) the resulting formal power series are proved to be locally convergent by means of suitable majorizations on their coefficients. This is the basis of the classical method of *majorant series* originating with Cauchy.

▷ **15.** *The majorant series technique.* Given two power series, define $f(z) \preceq g(z)$ if $|[z^n]f(z)| \leq [z^n]g(z)$ for all $n \geq 0$. The following two conditions are equivalent: (i) $f(z)$ is analytic in the disc $|z| < \rho$; (ii) for any $r < \rho$ there exists a c such that

$$f(z) \preceq \frac{c}{1 - rz}.$$

If f, g are majorized by $c/(1 - rz), d/(1 - rz)$ respectively, then $f + g$ and $f \cdot g$ are majorized,

$$f(z) + g(z) \preceq \frac{c + d}{1 - rz}, \qquad f(z) \cdot g(z) \preceq \frac{e}{1 - sz},$$

for any $s < r$ and some e dependent on s. If f, g are majorized by $c/(1 - rz), dz/(1 - rz)$ respectively, then $f \circ g$ is majorized:

$$f \circ g(z) \preceq \frac{cz}{1 - r(1 + d)z}.$$

Constructions for $1/f$ and for the functional inverse of f can be similarly developed. See Cartan's book [**32**] and van der Hoeven's study [**186**] for a systematic treatment. ◁

For functions defined by analytic expressions, singularities can be determined inductively in an intuitively transparent manner. If $\mathrm{Sing}(f)$ and $\mathrm{Zero}(f)$ are the set of singularities and zeros of function f, then, due to closure properties of analytic functions, the following informally stated guidelines apply.

$$\begin{cases} \mathrm{Sing}(f \pm g) & \subseteq & \mathrm{Sing}(f) \cup \mathrm{Sing}(g) \\ \mathrm{Sing}(f \times g) & \subseteq & \mathrm{Sing}(f) \cup \mathrm{Sing}(g) \\ \mathrm{Sing}(f/g) & \subseteq & \mathrm{Sing}(f) \cup \mathrm{Sing}(g) \cup \mathrm{Zero}(g) \\ \mathrm{Sing}(f \circ g) & \subseteq & \mathrm{Sing}(g) \cup g^{(-1)}(\mathrm{Sing}(f)) \\ \mathrm{Sing}(\sqrt{f}) & \subseteq & \mathrm{Sing}(f) \cup \mathrm{Zero}(f) \\ \mathrm{Sing}(\log(f)) & \subseteq & \mathrm{Sing}(f) \cup \mathrm{Zero}(f) \\ \mathrm{Sing}(f^{(-1)}) & \subseteq & f(\mathrm{Sing}(f)) \cup f(\mathrm{Zero}(f')). \end{cases}$$

A mathematically rigorous treatment would require considering multivalued functions and Riemann surfaces, so that we do not state detailed validity conditions and, at this stage, keep for these formulæ the status of useful heuristics. In fact, because of Pringsheim's theorem, the search of dominant singularities of combinatorial generating function can normally avoid considering the multivalued structure of functions, since only some initial segment of the positive real half–line needs to be considered. This in turn implies a powerful and easy way of determining the exponential order of coefficients of a wide variety of generating functions, as we now explain.

As defined in Chapters I and II, a combinatorial class is *constructible* if it can be specified by a finite set of equations involving only basic constructors. A specification is *iterative* if the dependency graph of the specification is acyclic, that is, no recursion is involved and a single functional term (written with sums, products, as well as sequence, set, and cycle constructions) describes the specification. We state:

THEOREM IV.5 (Computability of growth). *Let $\mathcal{C}$ be a* constructible *unlabelled class that admits of an* iterative *specification in terms of* $(1, \mathcal{Z}; \mathfrak{S}, \mathfrak{P}, \mathfrak{M}, \mathfrak{C}; +, \times)$. *Then the radius of convergence ρ_C of the OGF $C(z)$ of $\mathcal{C}$ is a nonzero computable real number.*

Let $\mathcal{D}$ be a constructible *labelled class that admits of an* iterative *specification in terms of $(1, \mathcal{Z}; \mathfrak{S}, \mathfrak{P}, \mathfrak{C}; +, \star)$. Then the radius of convergence ρ_D of the EGF $D(z)$ of $\mathcal{D}$ is a nonzero computable real number.*

Accordingly, the exponential rate of growth of the coefficients $[z^n]C(z)$ and $[z^n]D(z)$ are computable real numbers.

A real number α is computable iff there exists a program Π_α which on input m outputs a rational number α_m that is within $\pm 10^{-m}$ of α. The theorem immediately implies that the exponential growth estimates,

$$[z^n]C(z) \equiv C_n \bowtie \left(\frac{1}{\rho_C}\right)^n, \qquad [z^n]D(z) \equiv \frac{1}{n!}D_n \bowtie \left(\frac{1}{\rho_D}\right)^n,$$

for coefficients are automatically computable from the specification itself.

PROOF. In both cases, the proof proceeds by induction on the structural specification of the class. For each class $\mathcal{F}$, with generating function $F(z)$, we associate a *signature*, which is an ordered pair $\langle \rho_F, \tau_F \rangle$, where ρ_F is the radius of convergence of F and τ_F is the value of F at ρ_F, precisely,

$$\tau_F := \lim_{x \to \rho_F^-} F(x).$$

(The value τ_F is well defined as an element of $\mathbb{R} \cup \{+\infty\}$ since F, being a counting generating function, is necessarily increasing on $(0, \rho_F)$.) We prove the assertion of the theorem together with the additional property that $\tau_F = \infty$ and as soon as one of the unary constructors $(\mathfrak{S}, \mathfrak{M}, \mathfrak{P}, \mathfrak{C})$ intervenes in the specification, that is, as soon as the class is infinite. In that case, since the OGF includes infinitely many terms of the form z^n, it must be divergent at 1, so that $\rho_F \le 1$ holds *a priori* for all infinite classes under consideration.

Consider the unlabelled case first. The signatures of the neutral class 1 and the atomic class $\mathcal{Z}$, with OGF 1 and z, are $\langle +\infty, 1 \rangle$ and $\langle +\infty, +\infty \rangle$. Any nonconstant polynomial which is the OGF of a finite set has the signature $\langle +\infty, +\infty \rangle$. The assertion is thus easily verified in these cases.

Next, let $\mathcal{F} = \mathfrak{S}(\mathcal{G})$. The OGF $G(z)$ must be nonconstant and in fact satisfy $G(0) = 0$ in order for the sequence construction to be properly defined. Thus, by the induction hypothesis, one has $0 < \rho_G \le +\infty$ and $\tau_G = +\infty$. Now, the function G being increasing and continuous along the positive axis, there must exist a value β such that $0 < \beta < \rho_G$ with $G(\beta) = 1$. For $z \in (0, \beta)$, the quasi-inverse $F(z) = (1 - G(z))^{-1}$ is well defined and analytic; as z approaches β from the left, $F(z)$ increases unboundedly. Thus, the smallest singularity of F along the positive axis is at β, and by Pringsheim's theorem, one has $\rho_F = \beta$. The argument also shows that $\tau_F = +\infty$. There only remains to check that β is computable. The coefficients of G form a computable sequence of integers, so that $G(x)$, which can be well approximated via truncated Taylor series, is an effectively computable number[6] if x is itself a positive computable number less than ρ_G. Then dichotomic search constitutes effectively an algorithm for determining β.

Next, we consider the multiset construction, $\mathcal{F} = \mathfrak{M}(\mathcal{G})$, whose translation into OGFs necessitates the "Pólya exponential":

$$F(z) = \operatorname{Exp}(G(z)) \quad \text{where} \quad \operatorname{Exp}(h(z)) := \exp\left(h(z) + \frac{1}{2}h(z^2) + \frac{1}{3}h(z^3) + \cdots\right).$$

Once more, the induction hypothesis is assumed for G. If G is polynomial, then F is a variant of the OGF of integer partitions, and in fact is expressible as a finite product of terms of the form $P(z), P(z^2), P(z^3), \ldots$ Thus, $\rho_F = 1$ and $\tau_F = \infty$ in that particular case. In the general case of $\mathcal{F} = \mathfrak{M}(\mathcal{G})$ with $\mathcal{G}$ infinite, we start by fixing arbitrarily a number r such that $0 < r < \rho_G \le 1$ and examine $F(z)$ for $z \in (0, r)$. The expression for

[6] The present argument only establishes non-constructively the *existence* of a program, based on the fact that truncated Taylor series converge geometrically fast at an interior point of their disc of convergence. Making explict this program and the involved parameters however represents a harder problem that is not touched upon here.

F rewrites as

$$\mathrm{Exp}(G(z)) = e^{G(z)} \cdot \exp\left(\frac{1}{2}G(z^2) + \frac{1}{3}G(z^3) + \cdots\right).$$

The first factor is analytic for z on $(0, \rho_G)$ since, the exponential function being entire, e^G has the singularities of G. As to the second factor, one has $G(0) = 0$ (in order for the set construction to be well-defined), while $G(x)$ is convex for $x \in [0, r]$ (since its second derivative is positive). Thus, there exists a positive constant K such that $G(x) \le Kx$ when $x \in [0, r]$. Then, the series $\frac{1}{2}G(z^2) + \frac{1}{3}G(z^3) + \cdots$ has its terms dominated by those of the convergent series

$$\frac{K}{2}r^2 + \frac{K}{3}r^3 + \cdots = K \log(1-r)^{-1} - Kr.$$

By a well known theorem of analytic function theory, a uniformly convergent sum of analytic functions is itself analytic; consequently, $\frac{1}{2}G(z^2) + \frac{1}{3}G(z^3) + \cdots$ is analytic at all z of $(0, r)$. Analyticity is then preserved by the exponential, so that $F(z)$, being analytic at $z \in (0, r)$ for any $r < \rho_G$ has a radius of convergence that satisfies $\rho_F \ge \rho_G$. On the other hand, since $F(z)$ dominates termwise $G(z)$, one has $\rho_F \le \rho_G$. Thus finally one has $\rho_F = \rho_G$. Also, $\tau_G = +\infty$ implies $\tau_F = +\infty$.

A completely parallel discussion covers the case of the powerset construction ($\mathfrak{P}$) whose associated functional $\overline{\mathrm{Exp}}$ is a minor modification of the Pólya exponential Exp. The cycle construction can be treated by similar arguments based on consideration of "Pólya's logarithm" as $\mathcal{F} = \mathfrak{C}(\mathcal{G})$ corresponds to

$$F(z) = \mathrm{Log}\,\frac{1}{1-G(z)}, \qquad \text{where} \quad \mathrm{Log}\,h(z) = \log h(z) + \frac{1}{2}\log h(z^2) + \cdots.$$

In order to conclude with the unlabelled case, there only remains to discuss the binary constructors $+$, $\times$, which give rise to $F = G + H$, $F = G \cdot H$. It is easily verified that $\rho_F = \min(\rho_G, \rho_H)$ and $\tau_F = \tau_G \circ \tau_H$ with $\circ$ being $+$ or $\times$. Computability is granted since the minimum of two computable numbers is computable.

The labelled case is covered by the same type of argument as above. The discussion is even simpler, since the ordinary exponential and logarithm replace the Pólya operators Exp and Log. It is still a fact that all the EGFs of infinite families are infinite at their dominant positive singularity, though the radii of convergence can now be of any magnitude (w.r.t. 1). $\square$

$\triangleright$ **16.** *Syntactically decidable properties.* In the unlabelled case, $\rho_F = 1$ iff the specification of $\mathcal{F}$ only involves $(1, \mathcal{Z}; \mathfrak{P}, \mathfrak{M}; +, \times)$ and at least one of $\mathfrak{P}, \mathfrak{M}$. $\triangleleft$

$\triangleright$ **17.** *Nonconstructibility of permutations and graphs.* The class $\mathcal{P}$ of all permutations cannot be specified as a constructible unlabeled class since the OGF $P(z) = \sum_n n!z^n$ has radius of convergence 0. (It is of course constructible as a labelled class.) Graphs, whether labelled or unlabelled, are too numerous to form a constructible class. $\triangleleft$

Theorem IV.5 establishes a link between analytic combinatorics, computability theory, and symbolic manipulation systems. It is based on an article of Flajolet, Salvy, and Zimmermann [**80**] devoted to such computability issues in exact and asymptotic enumeration. Recursive specifications are not discussed now since they tend to give rise to branch points, themselves amenable to singularity analysis techniques to be developed in the next chapter.

EXAMPLE 3. *Combinatorial trains.* This somewhat artificial example from [**64**] serves to

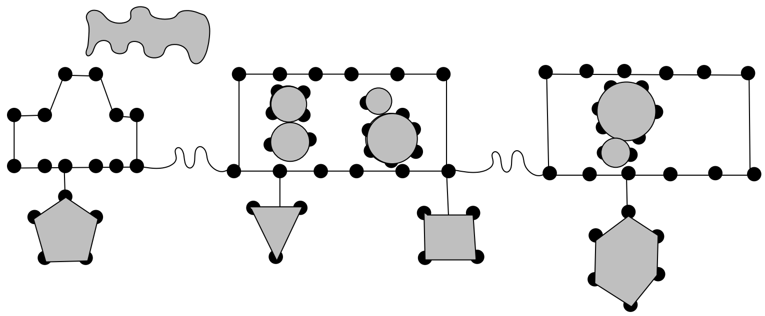


FIGURE 6. A random train.

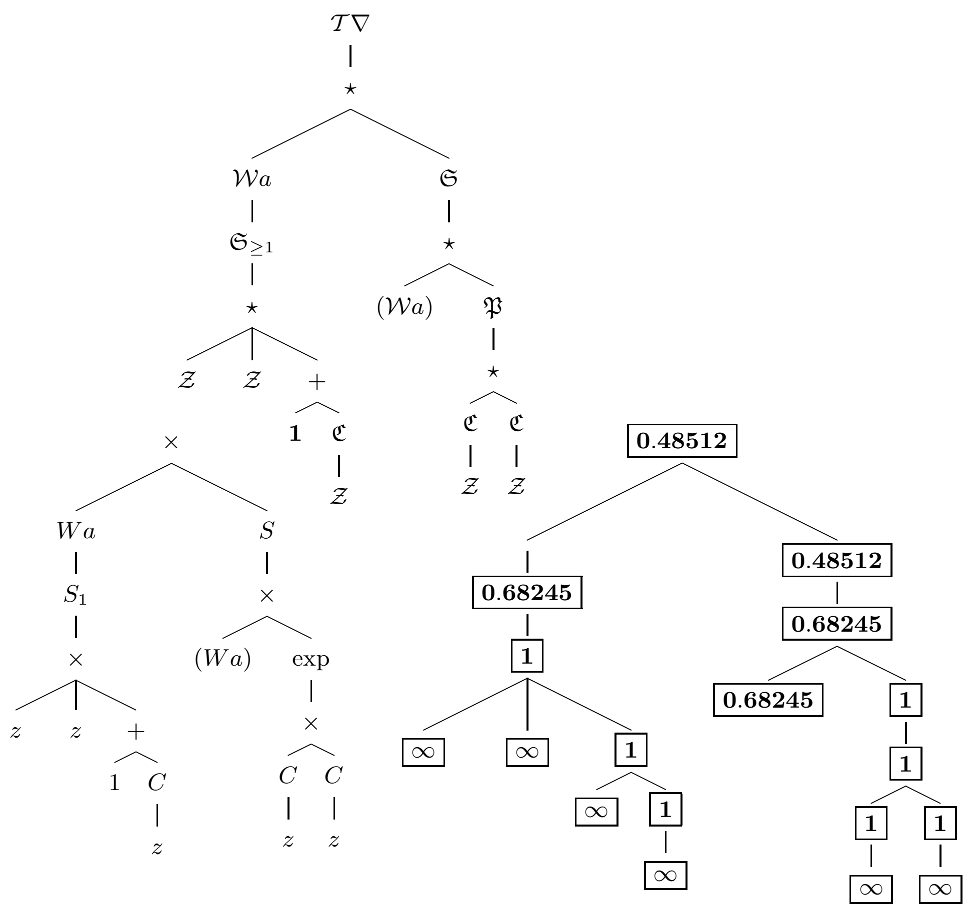


FIGURE 7. The inductive determination of the radius of convergence of the EGF of trains, $T(z)$: (top) a hierarchical view of the specification of $\mathcal{T}$; (bottom left) the corresponding expression tree of the EGF $T(z)$; (bottom right) the value of the radii for each subexpression of $T(z)$. (Notations: $L(y) = \log(1-y)^{-1}$, $S(y) = (1-y)^{-1}$, $S_1(y) = yS(y)$.)

illustrate the scope of Theorem IV.5 and demonstrate its inner mechanisms at work. Define the class of all labelled trains by the following specification,

$$\left\{\begin{array}{lcl} \mathcal{T}r & = & \mathcal{W}a \star \mathfrak{S}(\mathcal{W}a \star \mathfrak{P}(\mathcal{P}a)), \\ \mathcal{W}a & = & \mathfrak{S}_{\geq 1}(\mathcal{P}\ell), \\ \mathcal{P}\ell & = & \mathcal{Z} \star \mathcal{Z} \star (\mathbf{1} + \mathfrak{C}(\mathcal{Z})), \\ \mathcal{P}a & = & \mathfrak{C}(\mathcal{Z}) \star \mathfrak{C}(\mathcal{Z}). \end{array}\right. \tag{23}$$

In figurative terms, a train ($\mathcal{T}$) is composed of a first wagon ($\mathcal{W}a$) to which is appended a sequence of passenger wagons, each of the latter capable of containing a set of passengers ($\mathcal{P}a$). A wagon is itself composed of "planks" ($\mathcal{P}\ell$) determined by their end points ($\mathcal{Z}\star\mathcal{Z}$) and to which a circular wheel ($\mathfrak{C}(\mathcal{Z})$) may be attached. A passenger is composed of a head and a belly that are each circular arrangements of atoms (see Figure 6).

The translation into a set of EGF equations is immediate and a symbolic manipulation system readily provides the form of the EGF of trains, $T(z)$, as

$$Tr(z) = \frac{z^2\left(1+\log((1-z)^{-1})\right)}{\left(1-z^2\left(1+\log((1-z)^{-1})\right)\right)}\left(1-\frac{z^2\left(1+\log((1-z)^{-1})\right)e^{\left(\log((1-z)^{-1})\right)^2}}{1-z^2\left(1+\log((1-z)^{-1})\right)}\right)^{-1},$$

together with the expansion

$$Tr(z) = 2\,\frac{z^2}{2!} + 6\,\frac{z^3}{3!} + 60\,\frac{z^4}{4!} + 520\,\frac{z^5}{5!} + 6660\,\frac{z^6}{6!} + 93408\,\frac{z^7}{7!} + \cdots.$$

The specification (23) has a hierarchical structure, as suggested by the top representation of Figure 7, and this structure is itself directly reflected by the form of the expression tree of the GF $T(z)$. Then each node in the expression tree of $T(z)$ can be tagged with the corresponding value of the radius of convergence. This is done according to the principles of Theorem IV.5; see the bottom-right part of Figure 7. For instance, the quantity 0.68245 associated to $Wa(z)$ is given by the sequence rule and is determined as smallest positive solution to the equation

$$z^2\left(1-\log(1-z)^{-1}\right) = 1.$$

The tagging process works upwards till the root of the tree is reached; here the radius of convergence of T is determined to be $\rho \doteq 0.48512\cdots$, a quantity that happens to coincide with the ratio $[z^{49}]T(z)/[z^{50}]T(z)$ to more than 15 decimal places. □

IV. 4. Rational and meromorphic functions

The first principle that we have just discussed in great detail is:

> *The location of singularities of an analytic function determines the exponential order of growth of its Taylor coefficients.*

The second principle which refines the first one is:

> *The nature of the singularities determines the way the dominant exponential term in coefficients is modulated by a subexponential factor.*

We are now going to develop the correspondence between singular expansions and asymptotic behaviours of coefficients in the case of rational and meromorphic functions. Rational functions (fractions) are the simpler ones, and from their basic partial fraction expansion closed forms are derived for their coefficients. Next in order of difficulty comes the class of meromorphic functions; their Taylor coefficients appear to admit very accurate asymptotic expansions with error terms that are exponentially small, as results from an adequate use of the residue theore.

In the case of rational and, more generally, meromorphic functions, the net effect is summarized by the correspondence:

$$\text{Polar singularities} \rightsquigarrow \text{Subexponential factors } \theta(n) \text{ are polynomials.}$$

A distinguishing feature is the extremely good quality of the asymptotic approximations obtained; for instance 15 digits of accuracy is not uncommon in coefficients of index as low as 50.

IV. 4.1. Rational functions. A function $f(z)$ is a *rational function* iff it is of the form $f(z) = \frac{N(z)}{D(z)}$, with $N(z)$ and $D(z)$ being polynomials, which me may always assume to be relatively prime. For rational functions that are generating functions, we have $D(0) \neq 0$.

Sequences $\{f_n\}_{n\geq 0}$ that are coefficients of rational functions coincide with sequences that satisfy linear recurrence relations with constant coefficients. To see it, compute $[z^n]f(z)\cdot D(z)$, with $n > \deg(N(z))$. If $D(z) = d_0 + d_1 z + \cdots + d_m z^m$, then for $n > m$, one has:

$$\sum_{j=0}^{m} d_j f_{n-j} = 0.$$

The main theorem we prove here provides an *exact* finite expression for coefficients of $f(z)$ in terms of the poles of $f(z)$. Individual terms in corresponding expressions are sometimes called exponential polynomials.

THEOREM IV.6 (Expansion of rational functions). *If $f(z)$ is a rational function that is analytic at zero and has poles at points $\alpha_1, \alpha_2, \ldots, \alpha_m$, then there exist m polynomials $\{\Pi_j(x)\}_{j=1}^{m}$ such that:*

$$f_n \equiv [z^n]f(z) = \sum_{j=1}^{m} \Pi_j(n)\alpha_j^{-n}. \tag{24}$$

Furthermore the degree of Π_j is equal to the order of the pole of f at α_j minus one.

An expression of the form (24) is sometimes called an *exponential polynomial*.

PROOF. Since $f(z)$ is rational it admits a partial fraction expansion. Thus, assuming without loss of generality that $\deg(D) > \deg(N)$, we can decompose f into a finite sum

$$f(z) = \sum_{(\alpha,r)} \frac{c_{\alpha,r}}{(z-\alpha)^r},$$

where α ranges over the poles of $f(z)$ and r is bounded from above by the multiplicity of α as a pole of f. Coefficient extraction in this expression results from Newton's expansion,

$$[z^n]\frac{1}{(z-\alpha)^r} = \frac{(-1)^r}{\alpha^r}[z^n]\frac{1}{(1-\frac{z}{\alpha})^r} = \frac{(-1)^r}{\alpha^r}\binom{n+r-1}{r-1}\alpha^{-n}.$$

The binomial coefficient is a polynomial of degree $r-1$ in n, and collecting terms associated with a given α yields the statement of the theorem. $\square$

Notice that the expansion (24) is also an asymptotic expansion in disguise: when grouping terms according to the α's of increasing modulus, each group appears to be *exponentially smaller* than the previous one. A classical instance is the OGF of Fibonacci numbers,

$$f(z) = \frac{z}{1-z-z^2} = \frac{z}{1-z-z^2},$$

with poles at

$$\frac{-1+\sqrt{5}}{2} \doteq 0.61803, \qquad \frac{-1-\sqrt{5}}{2} \doteq -1.61803,$$

so that

$$F_n = \frac{1}{\sqrt{5}}\varphi^n - \frac{1}{\sqrt{5}}\bar{\varphi}^n = \frac{\varphi^n}{\sqrt{5}} + O(\frac{1}{\varphi^n}),$$

with $\varphi = (1+\sqrt{5})/2$ the golden ratio, and $\bar{\varphi}$ its conjugate.

The next example is certainly an artificial one. It is simply designed to demonstrate that all the details of the full decomposition are usually not required. The rational function

$$f(z) = \frac{1}{(1-z^3)^2(1-z^2)^3(1-\frac{z^2}{2})}$$

has a pole of order 5 at $z=1$, poles of order 2 at $z=\omega,\omega^2$ ($\omega = e^{2i\pi/3}$ a cubic root of unity), a pole of order 3 at $z=-1$, and simple poles at $z=\pm\sqrt{2}$. Therefore,

$$\begin{aligned} f_n = P_1(n) + P_2(n)\omega^{-n} + P_3(n)\omega^{-2n} + P_4(n)(-1)^n + \\ + P_5(n)2^{-n/2} + P_6(n)(-1)^n 2^{-n/2} \end{aligned}$$

where the degrees of $P_1,\dots,P_6$ are respectively $4,1,1,2,0,0$. For an asymptotic equivalent of f_n, only the pole at $z=1$ needs to be considered since it corresponds to the fastest exponential growth; in addition, at $z=1$, only the term of fastest growth needs to be taken into account since it gives the dominant contribution to coefficients. Thus, we have the correspondence

$$f(z) \sim \frac{1}{3^2\cdot 2^3\cdot(\frac{1}{2})}\frac{1}{(1-z)^5} \implies f_n \sim \frac{1}{3^2\cdot 2^3\cdot(\frac{1}{2})}\binom{n+4}{4} \sim \frac{n^4}{864}.$$

EXAMPLE 4. *Asymptotics of denumerants.* Denumerants are synonymous to integer partitions with summands restricted to be from a *fixed* finite set (Chapter I). We let $\mathcal{P}^{\mathcal{T}}$ be the class relative to set $\mathcal{T}$, with the known OGF,

$$P^{\mathcal{T}}(z) = \prod_{\omega\in\mathcal{T}} \frac{1}{1-z^\omega}.$$

A particular case is the one of integer partitions whose summands are in $\{1,2,\dots,r\}$,

$$P^{\{1,\dots,r\}}(z) = \prod_{m=1}^{r} \frac{1}{1-z^m}.$$

The GF has all its poles that are roots of unity. At $z=1$, the order of the pole is r, and one has

$$P^{\{1,\dots,r\}}(z) \sim \frac{1}{r!}\frac{1}{(1-z)^r},$$

as $z\to 1$. Other poles have smaller multiplicity: for instance the multiplicity of $z=-1$ is equal to the number of factors $(1-z^{2j})^{-1}$ in $P^{\{1,\dots,r\}}$, that is $\lfloor r/2\rfloor$; in general a primitive qth root of unity is found to have multiplicity $\lfloor r/q\rfloor$. There results that $z=1$ contributes a term of the form n^{r-1} to the coefficient of order n, while each of the other poles contributes a term of order at most $n^{\lfloor r/2\rfloor}$. We thus find

$$P_n^{\{1,\dots,r\}} \sim c_r n^{r-1} \qquad \text{with} \quad c_r = \frac{1}{r!(r-1)!}.$$

The same argument provides the asymptotic form of $P_n^{\mathcal{T}}$, since, to first order asymptotics, only the pole at $z=1$ counts. One then has:

PROPOSITION IV.2. *Let $\mathcal{T}$ be a finite set of integers without a common divisor ($\gcd(\mathcal{T}) = 1$). The number of partitions with summands restricted to $\mathcal{T}$ satisfies*

$$P_n^{\mathcal{T}} \sim \frac{1}{\tau}\frac{n^{r-1}}{(r-1)!}, \qquad \textit{with} \quad \tau := \prod_{n\in\mathcal{T}} n, \quad r := \operatorname{card}(\mathcal{T}).$$

For instance, in a country that would have pennies (1 cent), nickels (5 cents), dimes (10 cents) and quarters (25 cents), the number of ways to make change for a total of n cents is

$$[z^n]\frac{1}{(1-z)(1-z^5)(1-z^{10})(1-z^{25})} \sim \frac{1}{1\cdot 5\cdot 10\cdot 25}\frac{n^3}{3!} \equiv \frac{n^3}{7500},$$

asymptotically. □

IV. 4.2. Meromorphic Functions. An expansion very similar to that of Theorem IV.6 given for rational functions holds true for the larger class of coefficients of meromorphic functions.

THEOREM IV.7 (Expansion of meromorphic functions). *Let $f(z)$ be a function meromorphic for $|z| \le R$ with poles at points $\alpha_1, \alpha_2, \ldots, \alpha_m$, and analytic at all points of $|z| = R$ and at $z = 0$. Then there exist m polynomials $\{\Pi_j(x)\}_{j=1}^m$ such that:*

$$f_n \equiv [z^n]f(z) \;=\; \sum_{j=1}^{m} \Pi_j(n)\alpha_j^{-n} \;+\; \mathcal{O}(R^{-n}). \tag{25}$$

Furthermore the degree of Π_j is equal to the order of the pole of f at α_j minus one.

PROOF. We offer two different proofs, one based on subtracted singularities, the other one based on contour integration.

(i) *Subtracted singularities.* Around any pole α, $f(z)$ can be expanded locally:

$$f(z) \;=\; \sum_{k\ge -M} c_{\alpha,k}(z-\alpha)^k \tag{26}$$

$$\;=\; S_\alpha(z) + H_\alpha(z) \tag{27}$$

where the "singular part" $S_\alpha(z)$ is obtained by collecting all the terms with index in $[-M \mathinner{..} -1]$ ($S_\alpha(z) = N_\alpha(z)/(z-\alpha)^M$ with $N_\alpha(z)$ a polynomial of degree less than M) and $H_\alpha(z)$ is analytic at α. Thus setting $R(z) = \sum_j S_{\alpha_j}(z)$, we observe that $f(z) - S(z)$ is analytic for $|z| \le R$. In other words, by collecting the singular parts of the expansions and subtracting them, we have "removed" the singularities of $f(z)$, whence the name of "*method of subtracted singularities*" sometimes given to the method [**107**, vol. 2, p. 448].

Taking coefficients, we get:

$$[z^n]f(z) = [z^n]S(z) + [z^n](f(z) - S(z)).$$

The coefficient of $[z^n]$ in the rational function $S(z)$ is obtained from Theorem 1. It suffices to prove that the coefficient of z^n in $f(z) - S(z)$, a function analytic for $|z| \le R$, is $\mathcal{O}(R^{-n})$. This fact follows from trivial bounds applied to Cauchy's integral formula with the contour of integration being $\lambda = \{z \,/\, |z| = R\}$, as in the proof of Theorem IV.4:

$$|[z^n](f(z) - S(z))| = \frac{1}{2\pi}\left| \int_{|z|=R} (f(z) - S(z))\,\frac{dz}{z^{n+1}} \right| \;\le\; \frac{1}{2\pi}\frac{\mathcal{O}(1)}{R^{n+1}} 2\pi R.$$

(ii) *Contour integration.* There is another line of proof for Theorem IV.7 which we briefly sketch as it provides an insight which is useful for applications to other types of singularities treated in Chapter V. It consists in using directly Cauchy's coefficient formula

3	3
75	75
4683	4683
545835	545835
102247563	102247563
28091567595	28091567595
10641342970443	10641342970443
5315654681981355	5315654681981355
3385534663256845323	338553466325684532 ***6***
2677687796244384203115	2677687796244384203 ***088***
2574844419803190384544203	2574844419803190384544 ***450***
2958279121074145472650648875	2958279121074145472650064 ***6597***
4002225759844168492486127539083	4002225759844168492486127 5 ***55859***
6297562064950066033518373935334635	6297562064950066033518373935 ***416161***
11403568794011880483742464196184901963	11403568794011880483742464196 17 ***4527074***
2354515408573489664918449063714485547 6395	2354515408573489664918449063714 ***5314147690***

FIGURE 8. The surjection numbers pyramid: for $n = 2, 4, \dots, 32$, the exact values of the numbers R_n (left) compared to the approximation $\lceil \xi(n) \rfloor$ with discrepant digits in boldface (right).

and "pushing" the contour of integration past singularities. In other words, one computes directly the integral

$$I_n = \frac{1}{2i\pi} \int_{|z|=R} f(z) \, \frac{dz}{z^{n+1}}$$

by residues. There is a pole at $z = 0$ with residue f_n and poles at the α_j with residues corresponding to the terms in the expansion stated in Theorem IV.7; for instance, if $f(z) \sim c/(z-a)$ as $z \to a$, then

$$\operatorname{Res}(f(z) z^{-n-1}, z = a) = \operatorname{Res}(\frac{c}{(z-a)} z^{-n-1}, z = a) = \frac{c}{a^{n+1}}.$$

Finally, by the same trivial bounds as before, I_n is $\mathcal{O}(R^{-n})$. $\square$

EXAMPLE 5. *Surjections and alignments.* The surjection EGF is $R(z) = (2 - e^z)^{-1}$, and we have already determined its poles: the one of smallest modulus is at $\log 2 \doteq 0.69314$. At the dominant pole, as z tends to $\log 2$, one has $R(z) \sim -\frac{1}{2}(z - \log 2)^{-1}$. This implies an approximation for the number of surjections:

$$R_n \equiv n![z^n]R(z) \sim \xi(n), \qquad \text{with} \quad \xi(n) := \frac{n!}{2} \cdot \big(\frac{1}{\log 2}\big)^{n+1}.$$

Here is, for $n = 2, 4, \dots, 32$, a table of the values of the surjection numbers (left) compared with the asymptotic approximation rounded[7] to the nearest integer, $\lceil \xi(n) \rfloor$: It is piquant to see that $\lceil \xi(n) \rfloor$ provides the exact value of R_n for all values of $n = 1, \dots, 15$, and it starts losing one digit for $n = 17$, after which point a few "wrong" digits gradually appear, but in very limited number; see Figure 8 The explanation of such a faithful asymptotic representation owes to the fact that the error terms provided by meromorphic asymptotics are exponentially small. In effect, there is no other pole in $|z| \le 6$, the next ones being at $\log 2 \pm 2i\pi$ with modulus of about 6.32. Thus, for $r_n = [z^n]R(z)$, there holds

$$\frac{R_n}{n!} \sim \frac{1}{2} \cdot \big(\frac{1}{\log 2}\big)^{n+1} + \mathcal{O}(6^{-n}). \tag{28}$$

[7]The notation $\lceil x \rfloor$ represnets x rounded to the nearest integer: $\lceil x \rfloor := \lfloor x + \frac{1}{2} \rfloor$.

For the double surjection problem, $R^*(z) = (2 + z - e^z)$, we get similarly

$$[z^n]R^*(z) \sim \frac{1}{e^{\rho^*} - 1} (\rho^*)^{-n-1},$$

with $\rho^* = 1.14619$ the smallest positive root of $e^{\rho^*} - \rho^* = 2$.

Alignments are sequences of cycles, with EGF

$$f(z) = \frac{1}{1 - \log(1-z)^{-1}}.$$

There is a singularity when $\log(1-z)^{-1} = 1$, which is at $z = 1 - e^{-1}$ and arises before $z = 1$ where the logarithm becomes singular. Thus the computation of the asymptotic form of f_n only needs a local expansion near $(1 - e^{-1})$:

$$f(z) \sim \frac{-e^{-1}}{z - 1 + e^{-1}} \qquad \Longrightarrow \qquad [z^n]f(z) \sim \frac{e^{-1}}{(1 - e^{-1})^{n+1}}.$$

□

▷ **18.** *Some "supernecklaces".* One estimates

$$[z^n] \log\left(\frac{1}{1 - \log\frac{1}{1-z}}\right) \sim \frac{1}{n}(1 - e^{-1})^{-n},$$

where the EGF enumerates (labelled) cycles of cycles. [Hint: Take derivatives.] ◁

EXAMPLE 6. *Generalized derangements.* The probability that the shortest cycle in a random permutation of size n has length larger than k is

$$[z^n]\frac{e^{-\frac{z}{1}-\frac{z^2}{2}-\cdots-\frac{z^k}{k}}}{1-z}.$$

For any fixed k, the generating function, call it $f(z)$, is equivalent to $e^{-H_k}/(1-z)$ as $z \to 1$. Accordingly the coefficients $[z^n]f(z)$ tend to e^{-H_k} as $n \to \infty$. Thus, due to meromorphy, we have the characteristic implication

$$f(z) \sim \frac{e^{-H_k}}{1-z} \qquad \Longrightarrow \qquad [z^n]f(z) \sim e^{-H_k}.$$

Since the difference between $f(z)$ and the approximation at 1 is an entire function, the error is exponentially small:

$$[z^n]\frac{e^{-\frac{z}{1}-\frac{z^2}{2}-\cdots-\frac{z^k}{k}}}{1-z} = e^{-H_k} + O(R^{-n}), \tag{29}$$

for *fixed* k and *any* $R > 1$. The cases $k = 1, 2$ in particular justify the estimates mentioned in the introduction on p. 5.

As a side remark, the classical approximation of the harmonic numbers, $H_k \approx \log k + \gamma$ suggests $e^{-\gamma}/k$ as a further approximation to (29) that might be valid for both large n and large k in suitable regions. This can be made precise; in accordance with this heuristic argument, the expected length of the shortest cycle in a random permutation of size n is symptotic to

$$\sum_{k=1}^{n} \frac{e^{-\gamma}}{k} \sim e^{-\gamma} \log n,$$

as first proved by Shepp and Lloyd in [**169**]. □

▷ **19.** *Shortest cycles of permutations are not too long.* Let S_n be the random variable denoting the length of the shortest cycle in a random permutation of size n. Using the circle $|z| = 2$ to estimate the error in the approximation e^{-H_k} above, one finds that, for $k \le \log n$,

$$\left|\mathbb{P}(S_n > k) - e^{-H_k}\right| \le \frac{1}{2^n} e^{2^k},$$

which is exponentially small in this range of k-values. Thus, the approximation e^{-H_k} remains good when k is allowed to tend sufficiently slowly to ∞ with n. One can also explore the possibility of better bounds and larger regions of validity of the main approximation. (See Panario and Richmond's study [**154**] for a general theory of smallest components in sets.) ◁

EXAMPLE 7. *Smirnov words and Carlitz compositions.* This examples illustrates the analysis of a group of rational generating functions (Smirnov words) paralleling nicely the enumeration of a special type of integer composition (Carlitz compositions) resorting to meromorphic asymptotics.

Bernoulli trials have been discussed in Chapter III, in relation to weighted word models. Take the class $\mathcal{W}$ of all words over an r-ary alphabet, where letter j is assigned probability p_j and letters of words are drawn independently. With this weighting, the GF of all words is

$$W(z) = \frac{1}{1 - \sum p_j z} = \frac{1}{1-z}.$$

Consider the problem of determining the probability that has a random word of length n is of Smirnov type, i.e., all blocks of length 2 are formed with two distinct letters (see also [**96**, p. 69]).

By our discussion of Section III.6, the GF of Smirnov words (again with the probabilistic weighting) is

$$S(z) = \frac{1}{1 - \sum \frac{p_j z}{1 + p_j z}}.$$

This is a rational function with a unique dominant singularity at ρ such that

$$\sum_{j=1}^{r} \frac{p_j \rho}{1 + p_j \rho} = 1. \tag{30}$$

(It is easy to verify by monotonicity that this equation has a unique positive solution.) Thus, ρ is a well characterized algebraic number defined implicitly by an equation of degree r. There results that the probability for a word to be Smirnov is (not too surprisingly) exponentially small, with the precise formula being

$$[z^n]S(z) \sim C \cdot \rho^{-n}, \qquad C = \left(\rho \sum \frac{p_i \rho}{1 + p_i \rho}\right)^{-1}.$$

A similar analysis, but with bivariate generating functions shows that in a random word of length n conditioned to be Smirnov, the letter j appears with frequency asymptotic to

$$q_j = \frac{p_j \rho}{1 + p_j \rho}, \tag{31}$$

in the sense that mean number of occurrences of letter j is asymptotic to $q_j n$. All these results are seen to be consistent with the equiprobable letter case $p_j = 1/r$, for which $\rho = r/(r-1)$.

Carlitz compositions illustrate a similar situation, in which the alphabet is in a sense infinite, while letters have different sizes. Recall that a Carlitz composition of the integer n is a composition of n such that no two adjacent summands have equal values. Consider first compositions with a bound m on the largest allowable summand. The OGF of such Carlitz compositions is

$$C^{[m]}(z) = \left(1 - \sum_{j=1}^{m} \frac{z^j}{1+z^j}\right)^{-1},$$

and the OGF of all Carlitz compositions is obtained by letting m tend to infinity:

$$C^{[\infty]}(z) = \left(1 - \sum_{j=1}^{\infty} \frac{z^j}{1+z^j}\right)^{-1}. \tag{32}$$

In particular, we get *EIS* **A003242**[8]:

$$C^{[\infty]}(z) = 1 + z + z^2 + 3z^3 + 4z^4 + 7z^5 + 14z^6 + 23z^7 + 39z^8 + 71z^9 + \cdots .$$

The asymptotic form of the number of Carlitz compositions is then easily found by singularity analysis of meromorphic functions. The OGF has a simple pole at ρ which is the smallest positive root of the equation

$$\sum_{j=1}^{\infty} \frac{\rho^j}{1+\rho^j} = 1. \tag{33}$$

(Note the analogy with (30) due to commonality of the combinatorial argument.) Thus:

$$C_n^{[\infty]} \sim C \cdot \alpha^n, \quad C \doteq 0.45638, \quad \alpha \doteq 1.75024\,.$$

There, $\alpha = \rho$ with ρ as in (33). In a way analogous to Smirnov words, the asymptotic frequency of summand k appears to be $\rho^k/(1+\rho^k)$; see [**115, 135**] for further properties. $\square$

IV. 5. Localization of singularities

There are situations where a function possesses several dominant singularities, that is, several singularities are present on the boundary of its disk of convergence. We examine here the induced effect on the coefficient's coefficients and discuss ways to localize such dominant singularities.

IV. 5.1. Multiple singularities. In the presence of multiple singularities on the circle of convergence of a series, several geometric terms of the form α^n sharing the same modulus must be combined. In simpler cases, such terms induce a periodic behaviour for coefficients that is easy to describe; in the more general case, fluctuations of a somewhat "arithmetic nature" result. Finally, consideration of *all* singularities (whether dominant or not) of a meromorphic functions may lead to explicit summations expressing their coefficients.

Periodicities. When several singularities of $f(z)$ have the same modulus, they may induce complete cancellations, so that different regimes will be present in the coefficients of f. For instance

$$\frac{1}{1+z^2} = 1 - z^2 + z^4 - z^6 + z^8 - \cdots, \quad \frac{1}{1-z^3} = 1 + z^3 + z^6 + z^9 + \cdots,$$

[8] The *EIS* designates Sloane's On-Line Encyclopedia of Integer Sequences [**170**]; see [**171**] for an earlier printed version.

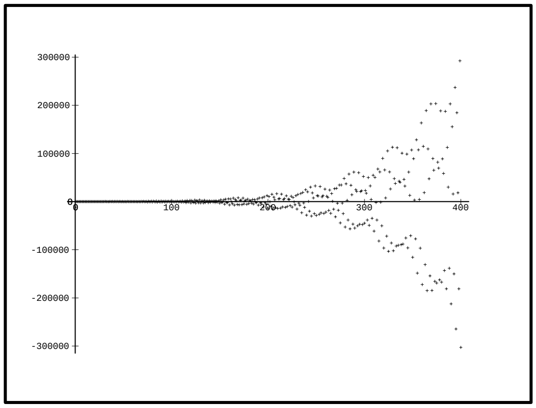


FIGURE 9. The coefficients $[z^n]f(z)$, where $f(z) = \left(1+1.02z^4\right)^{-3}\left(1-1.05z^5\right)^{-1}$ illustrate a periodic superposition of smooth behaviours that depend on the residue class of n modulo 20.

exhibit patterns of periods 4 and 3 respectively, this corresponding to roots of unity or order 4 $(\pm i)$, and 3. Accordingly,

$$\phi(z) = \frac{1}{1+z^2} + \frac{1}{1-z^3} = \frac{2-z^2+z^3+z^4+z^8+z^9-z^{10}}{1-z^{12}}$$

has a pattern of period 12, and the coefficients ϕ_n such that $n \equiv 1,5,6,7,11$ modulo 12 are zero. Consequently, if we analyze

$$[z^n]\psi(z) \qquad \text{where} \qquad \psi(z) = \phi(z) + \frac{1}{1-z/2},$$

we see that a different exponential growth manifests itself when n is taken congruent to $1,5,6,7,11 \bmod 12$. In many combinatorial applications, generating functions involving periodicities can be decomposed "at sight", and the corresponding asymptotic subproblems generated are then solved separately.

$\triangleright$ **20.** *Decidability of polynomial properties.* Given a polynomial $p(z) \in \mathbb{Q}[z]$, the following properties are decidable: (i) whether one of the zeros of p is a root of unity; (ii) whether one of the zeros of p has an argument that is commensurate with π. [One can use resultants. An algorithmic discussion of this and related issues is given in [**98**].] $\triangleleft$

Nonperiodic fluctuations. Take the polynomial $D(z) = 1 - \frac{6}{5}z + z^2$, whose roots are

$$\alpha = \frac{3}{5} + i\frac{4}{5},\ \bar\alpha = \frac{3}{5} - i\frac{4}{5},$$

both of modulus 1 (the numbers $3,4,5$ form a Pythagorean triple), with argument $\pm\theta$ where $\theta = \arctan(\frac{4}{3}) = 0.9279$. The expansion of the function $f(z) = 1/D(z)$ starts as

$$\frac{1}{1-\frac{6}{5}z+z^2} \quad = \quad 1 + \frac{6}{5}z + \frac{11}{25}z^2 - \frac{84}{125}z^3 - \frac{779}{625}z^4 - \frac{2574}{3125}z^5 + \cdots$$

the sign sequence being

$$+++---++++---+++----+++----+++---,$$

which indicates a mildly irregular oscillating behaviour, where blocks of 3 or 4 pluses follow blocks of 3 or 4 minuses.

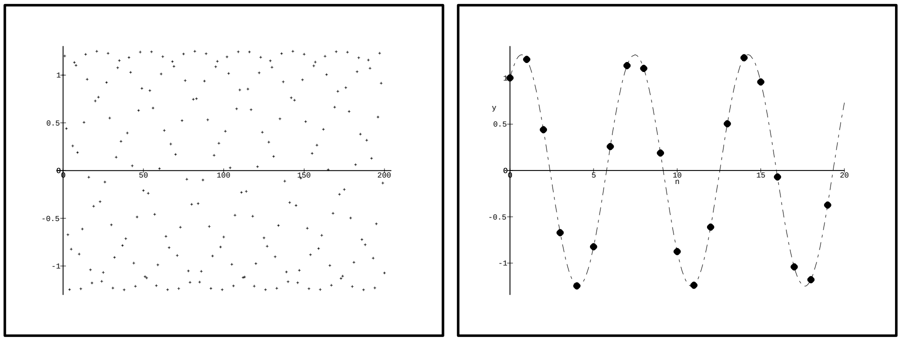


FIGURE 10. The coefficients of $f = 1/(1 - \frac{6}{5}z + z^2)$ exhibit an apparently chaotic behaviour (left) which in fact corresponds to a discrete sampling of a sine function (right), reflecting the presence of two conjugate complex poles.

The exact form of the coefficients of f results from a partial fraction expansion:

$$f(z) = \frac{a}{1 - z/\alpha} + \frac{b}{1 - z/\bar{\alpha}} \qquad \text{with } a = \frac{1}{2} + \frac{3}{8}i,\ b = \frac{1}{2} - \frac{3}{8}i.$$

Accordingly,

$$\begin{aligned} f_n &= ae^{-in\theta_0} + be^{in\theta_0} \\ &= \frac{\sin((n+1)\theta_0)}{\sin(\theta_0)}. \end{aligned}$$

This explains the sign changes observed. Since the angle θ_0 is not commensurate with π, the coefficients fluctuate but, unlike in our earlier examples, no exact periodicity is present in the sign patterns. See Figure 10 for a rendering and Figure 10 below for a meromorphic case linked to compositions into prime summands.

Complicated problems of an arithmetical nature may occur if several such singularities with non–commensurable arguments combine, and some open problem remain in the analysis of linear recurring sequences. (For instance no decision procedure is known to determine whether such a sequence ever vanishes.) Fortunately, such problems occur infrequently in combinatorial enumerations where zeros of rational functions tend to have a simple geometry.

Exact formulæ. The error terms appearing in the asymptotic expansion of coefficients of meromorphic functions are already exponentially small. By "pealing off" the singularities of a meromorphic function layer by layer, in order of increasing modulus, one is led to extremely precise expansions for the coefficients. Sometimes even, "exact" expressions may result. The latter is the case for the Bernoulli numbers B_n, the surjection numbers

R_n, the Secant numbers E_{2n} and the Tangent numbers E_{2n+1} defined by

$$\begin{array}{lcll}
\sum_{n=0}^{\infty} B_n \frac{z^n}{n!} & = & \frac{z}{e^z-1} & \text{(Bernoulli numbers)} \\
\sum_{n=0}^{\infty} R_n \frac{z^n}{n!} & = & \frac{1}{2-e^z} & \text{(Surjection numbers)} \\
\sum_{n=0}^{\infty} E_{2n} \frac{z^{2n}}{(2n)!} & = & \frac{1}{\cos(z)} & \text{(Secant numbers)} \\
\sum_{n=0}^{\infty} E_{2n+1} \frac{z^{2n+1}}{(2n+1)!} & = & \tan(z) & \text{(Tangent numbers).}
\end{array}$$

Bernoulli numbers have an EGF $z/(e^z-1)$ that has poles at the points $\chi_k = 2ik\pi$, with $k \in \mathbb{Z} \setminus \{0\}$. The residue at χ_k is equal to χ_k,

$$\frac{z}{e^z-1} \sim \frac{\chi_k}{z-\chi_k} \qquad (z \to \chi_k).$$

The expansion theorem for meromorphic functions is applicable here. To see it use the Cauchy integral formula, and proceed as in the proof of Theorem IV.7, using as external contours large circles that pass half way between poles. Along these contours, the integrand tends to 0 because the Cauchy "kernel" z^{-n-1} decreases with the radius of the integration contour while the EGF stays bounded. In the limit, corresponding to an infinitely large contour, the coefficient integral becomes equal to the sum of all residues of the meromorphic function over the whole of the complex plane.

From this argument, we thus get: $\frac{B_n}{n!} = -\sum_{k\in\mathbb{Z}\setminus\{0\}} \chi_k^{-n}$. This proves that $B_n = 0$ if n is odd. If n is even, then grouping terms two by two, we get the exact representation (which also serves as an asymptotic expansion):

$$\frac{B_{2n}}{(2n)!} = (-1)^{n-1} 2^{1-2n} \pi^{-2n} \sum_{k=0}^{\infty} \frac{1}{k^{2n}}. \tag{34}$$

Reverting the equality, we have also established that

$$\zeta(2n) = (-1)^{n-1} 2^{2n-1} \pi^{2n} \frac{B_{2n}}{(2n)!} \qquad \text{with} \quad \zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}, \quad B_n = n![z^n]\frac{z}{e^z-1},$$

a well-known identity that provides values of the Riemann zeta function ($\zeta(s)$) at even integers as rational multiples of powers of π.

In the same vein, the *surjection numbers* have as EGF $R(z) = (2-e^z)^{-1}$ with simple poles at

$$\chi_k = \log 2 + 2ik\pi \qquad \text{where} \qquad R(z) \sim \frac{1}{2}\frac{1}{\chi_k - z}.$$

Since $R(z)$ stays bounded on circles passing half way in between poles, we find the exact formula, $\frac{R_n}{n!} = \frac{1}{2}\sum_{k\in\mathbb{Z}} \chi_k^{-n-1}$. An equivalent real formulation is

$$\frac{R_n}{n!} = \frac{1}{2}\left(\frac{1}{\log 2}\right)^{n+1} + \sum_{k=1}^{\infty} \frac{\cos((n+1)\theta_k)}{(\log^2 2 + 4k^2\pi^2)^{(n+1)/2}} \text{ with } \theta_k = \arctan(\frac{2k\pi}{\log 2}), \tag{35}$$

which shows the hidden occurrence of infinitely many "harmonics" of fast decaying amplitude.

$\triangleright$ **21.** *Alternating permutations, tangent and secant numbers.* The relation (34) also provides a representation of the *tangent numbers* since $E_{2n-1} = (-1)^{n-1}B_{2n}4^n(4^n-1)/(2n)$. The secant numbers E_{2n} satisfy

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{(2k+1)^{2n+1}} = \frac{(\pi/2)^{2n+1}}{2\,(2n)!}E_{2n},$$

which can be read either as providing an asymptotic expansion of E_{2n} or as an evaluation of the sums on the left (the values of a Dirichlet L-function) in terms of π. The asymptotic number of alternating permutations (Chapter II) is consequently known to great accuracy. $\triangleleft$

$\triangleright$ **22.** *Solutions to the equation* $\tan(x) = x$. Let x_n be the nth positive root of the equation $\tan(x) = x$. For any integer $r \geq 1$, the sum $\sum_n x_n^{-2r}$ is a computable rational number. [From folklore and *The American Mathematical Monthly*.] $\triangleleft$

IV. 5.2. Localization of zeros and poles. We gather here a few results that often prove useful in determining the location of zeros of analytic functions, and hence of poles of meromorphic functions. A detailed treatment of this topic may be found in Henrici's book [**107**].

Let $f(z)$ be an analytic function in a region Ω and let γ be a simple closed curve interior to Ω, and on which f is assumed to have no zeros. We claim that the quantity

$$N(f;\gamma) = \frac{1}{2i\pi}\int_\gamma \frac{f'(z)}{f(z)}\,dz$$

exactly equals the number of zeros of f inside γ counted with multiplicity. The reason is that the function f'/f has its poles exactly at the zeros of f, and its residue at each pole is 1, so that the assertion directly results from the residue theorem

Since a primitive function of f'/f is $\log f$, the integral also represents the variation of $\log f$ along γ, which is written $[\log f]_\gamma$. The variation $[\log f]_\gamma$ reduces to i times the variation of the argument of f along γ as $\log(re^{i\theta}) = \log r + i\theta$ and the modulus r has variation equal to 0 along a closed contour, $[\log \rho]_\gamma = 0$. The quantity $[\theta]_\gamma$ is, by its definition, the number of times the transformed contour $f(\gamma)$ winds about the origin. This observation is known as the *Argument Principle*:

> **Argument Principle.** *The number of zeros of $f(z)$ (counted with multiplicities) inside γ equals the winding number of the transformed contour $f(\gamma)$ around the origin.*

By the same argument, if f is meromorphic in $\Omega \ni \gamma$, then $N(f;\gamma)$ equals the difference between the number of zeros and the number of poles of f inside γ, multiplicities being taken into account. Figure 11 exemplifies the use of the argument principle in localizing zeros of a polynomial.

By similar devices, we get Rouché's theorem:

> **Rouché's theorem.** *Let the functions $f(z)$ and $g(z)$ be analytic in a region containing in its interior the closed simple curve γ. Assume that f and g satisfy $|g(z)| < |f(z)|$* on *the curve γ. Then $f(z)$ and $f(z)+g(z)$ have the same number of zeros* inside *the interior domain delimited by γ.*

The intuition behind Rouché's theorem is that, since $|g| < |f|$, then $f(\gamma)$ and $(f+g)(\gamma)$ must have the same winding number.

$\triangleright$ **23.** *Proof of Rouché's theorem.* Under the hypothesis of Rouch´e's theorem, for $0 \leq t \leq 1$ $h(z) = (f(z) + tg(z))$ is such that $N(h;\gamma)$ is both an integer and a continuous function of t in the given range. The conclusion of the theorem follows. $\triangleleft$

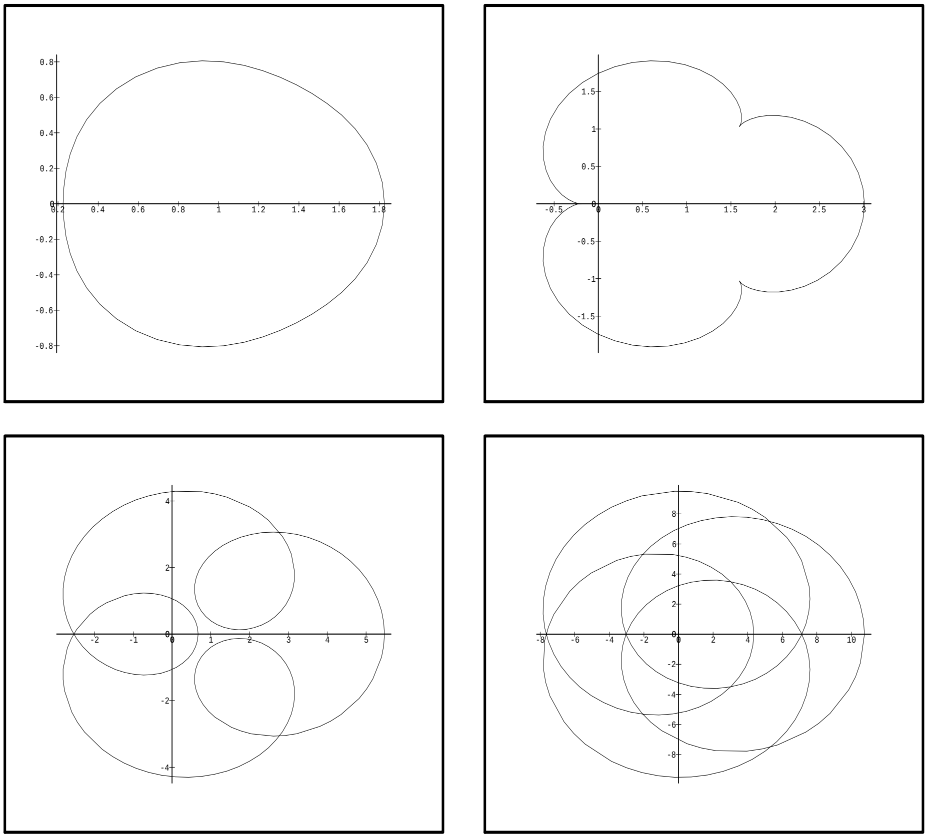


FIGURE 11. The transforms of $\gamma_j = \{|z| = \frac{4j}{10}\}$ by $P_4(z) = 1 - 2z + z^4$, for $j = 1, 2, 3, 4$, demonstrate that $P_4(z)$ has no zero inside $|z| < 0.4$, one zero inside $|z| < 0.8$, two zeros inside $|z| < 1.2$ and four zeros inside $|z| < 1.6$. The actual zeros are at $\rho_4 = 0.54368$, 1 and $1.11514 \pm 0.77184i$.

$\triangleright$ **24.** *The fundamental theorem of algebra.* Every complex polynomial $p(z)$ of degree n has exactly n roots. A proof follows by Rouch´e's theorem from the fact that, for large enough $|z| = R$, the polynomial assumed to be monic is a "perturbation" of its leading term, z^n. $\triangleleft$

These principles form the basis of numerical algorithms for locating zeros of analytic functions. For instance, one can start from an initial domain and recursively subdivide it until roots have been isolated with enough precision—the number of roots in a subdomain being at each stage determined by numerical integration; see Figure 11 and refer for instance to [**45**] for a discussion. Such algorithms can even acquire the status of full proofs if one operates with guaranteed precision routines (using, e.g., careful implementations of interval arithmetics). Examples of use of the method will appear in the next sections.

$\triangleright$ **25.** *The analytic Implicit Function Theorem from residues.* The sum of the roots of the equation $g(y) = 0$ interior to γ equals

$$\frac{1}{2i\pi} \int_\gamma \frac{g'(y)}{g(y)}\, y\, dy.$$

Let $F(z, y)$ be an analytic function in both z and y (i.e., it admits a convergent series expansion). If $F'_y(z_0, y_0) \neq 0$, then the function $y(z)$ implicitly defined by $F(z, y) = 0$ and such that $y(z_0) = y_0$ is given by

$$y(z) = \frac{1}{2i\pi} \int_\gamma \frac{F'_y(z, y)}{F(z, y)}\, y\, dy,$$

where γ is a small loop around y_0. Deduce that $y(z)$ is analytic at z_0. (Note: this requires a modicum of analytic functions of two complex variables as is to be found, e.g., in [**32**].) $\lhd$

IV. 5.3. The example of patterns in words. All patterns are not born equal. Surprisingly, in a random sequence of coin tossings, the pattern *HTT* is likely to occur much sooner (after 8 tosses on average) than the pattern *HHH* (needing 14 tosses on average); see the preliminary discussion in Chapter I. Questions of this sort are of obvious interest in the statistical analysis of genetic sequences. Say you discover that a sequence of length 100,000 on the four letters A, G, C, T contains the pattern *TACTAC* twice. Can this be assigned to chance or is this is likely to be a meaningful signal of some yet unknown structure? The difficulty here lies in quantifying precisely where the asymptotic regime starts, since, by Borges's Theorem (see the Note in Chapter I), sufficiently long texts will almost certainly contain any fixed pattern. The analysis of rational generating functions supplemented by Rouché's theorem provides definite answers to such questions.

We consider here the class $\mathcal{W}$ of words over an alphabet $\mathcal{A}$ of cardinality $m \geq 2$. A pattern $\mathfrak{p}$ of some length k is given. As seen in Chapters I and III, its autocorrelation polynomial is central to enumeration. This polynomial is defined as $c(z) = \sum_{j=0}^{k-1} c_j z^j$, where c_j is 1 if $\mathfrak{p}$ coincides with its kth shifted version and 0 otherwise. We consider here the enumeration of words containing the pattern $\mathfrak{p}$ at least once, and dually of words excluding the pattern $\mathfrak{p}$. In other words, we look at problems like: What is the probability that a random of words of length n does (or does not) contain your name as a block of consecutive letters?

The OGF of the class of words excluding $\mathfrak{p}$ is, we recall,

$$S(z) = \frac{c(z)}{z^k + (1 - mz)c(z)}. \tag{36}$$

and we shall start with the case $m = 2$ of a binary alphabet. The function $S(z)$ is simply a rational function, but the location and nature of its poles is yet unknown. We only know *a priori* that it should have a pole in the positive interval somewhere between $\frac{1}{2}$ and 1 (by Pringsheim's Theorem and since its coefficients are in the interval $[1, 2^n]$, for n large enough). Here is a small list for patterns of length $k = 3, 4$ of the pole ρ nearest to the origin:

Length (k)	Types	$c(z)$	ρ
$k = 3$	aab, abb, ...	1	0.61803
	aba, bab	$1 + z^2$	0.56984
	aaa, bbb	$1 + z + z^2$	0.54368
$k = 4$	aaab, aabb, abbb, ...	1	0.54368
	aaba, abba, abaa, ...	$1 + z^3$	0.53568
	abab, baba	$1 + z^2$	0.53101
	aaaa, bbbb	$1 + z + z^2 + z^3$	0.51879

We thus expect ρ to be close to $\frac{1}{2}$ as soon as the pattern is long enough. In order to prove this, we are going to apply Rouché's Theorem to the denominator of (36).

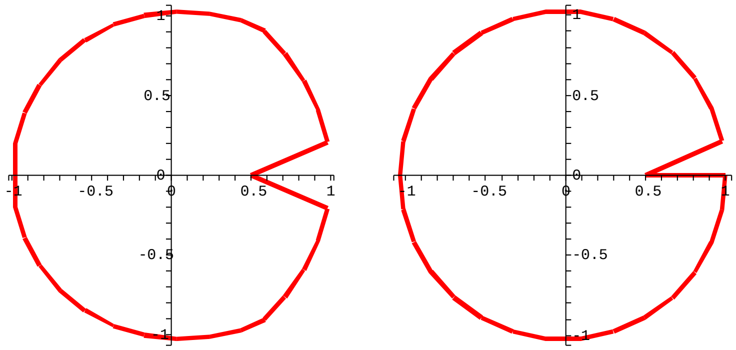


FIGURE 12. Complex zeros of $z^{31}+(1-2z)c(z)$ represented as joined by a polygonal line: (left) correlated pattern $a(ba)^{15}$; (right) uncorrelated pattern $a(ab)^{15}$.

As regards termwise domination of coefficients, the autocorrelation polynomials lies between 1 (for less correlated patterns like aaa. . . b) and $1 + z + \cdots + z^{k-1}$ (for the special case aaa. . . a). We set aside the special case of $\mathfrak{p}$ having only equal letters, i.e., a "maximal" autocorrelation polynomial—this case is discussed at length in the next chapter. Thus, in this scenario, the autocorrelation polynomial starts as $1+z^{\ell}+\cdots$ for some $\ell \geq 2$. Fix the number $A = 0.6$. On $|z| = A$, we have

$$|c(z)| \geq \left|1-(A^2+A^3+\cdots)\right| = \left|1-\frac{A^2}{1-A}\right| = \frac{1}{10}. \tag{37}$$

In addition, the quantity $(1-2z)$ ranges over the circle of diameter $[-0.2, 1.2]$ as z varies along $|z| = A$, so that $|1-2z| \geq 0.2$. All in all, we have found that, for $|z| = A$,

$$|(1-2z)c(z)| \geq 0.02.$$

On the other hand, for $k > 7$, we have $|z^k| < 0.017$ on the circle $|z| = A$. Then, amongst the two terms composing the denominator of (36), the first is strictly dominated by the second along $|z| = A$. By virtue of Rouché's Theorem, the number of roots of the denominator inside $|z| \leq A$ is then same as the number of roots of $(1-2z)c(z)$. The latter number is 1 (due to the root $\frac{1}{2}$) since $c(z)$ cannot be 0 by the argument of (37). Figure 12 exemplifies the extremely well-behaved characters of the complex zeros.

In summary, we have found that for all patterns with at least two different letters ($\ell \geq 2$) and length $k \geq 8$, the denominator has a unique root in $|z| \leq A = 0.6$. The property for lengths k satisfying $4 \leq k \leq 7$ is then easily verified directly. The case $\ell = 1$ can be subjected to an entirely similar argument (see Chapter V for details). Therefore, unicity of a simple pole ρ of $S(z)$ in the interval $(0.5, 0.6)$ is granted.

It is then a simple matter to determine the local expansion of $s(z)$ near $z = \rho$,

$$S(z) \underset{z\to\rho}{\sim} \frac{\widetilde{\Lambda}}{\rho - z}, \qquad \widetilde{\Lambda} := \frac{c(\rho)}{2c(\rho) - k\rho^{k-1}},$$

from which a precise estimate for coefficients derives by Theorems IV.6 and IV.7.

The computation finally extends almost verbatim to nonbinary alphabets, with ρ being now close to $\frac{1}{m}$. It suffices to use the disc of radius $A = 1.2/m$. The Rouché part of the argument grants us unicity of the dominant pole in the interval $(1/m, A)$ for $k \geq 5$

when $m = 3$, and for $k \geq 4$ and any $m \geq 4$. (The remaining cases are easily checked individually.)

PROPOSITION IV.3. *Consider an m-ary alphabet. Let $\mathfrak{p}$ be a pattern of length $k \geq 4$ with autocorrelation polynomial $c(z)$. Then the probability that a random word of length n does not contain $\mathfrak{p}$ as a pattern (a block of consecutive leters) satisfies*

$$\mathbb{P}_{\mathcal{W}_n}(\mathfrak{p} \textit{ does not occur}) = \Lambda_{\mathfrak{p}}(m\rho)^{-n-1} + O\left(\left(\frac{5}{6}\right)^n\right), \tag{38}$$

where $\rho \equiv \rho_{\mathfrak{p}}$ is the unique root in $(\frac{1}{m}, \frac{6}{5m})$ of the equation $z^k + (1 - mz)c(z) = 0$ and

$$\Lambda_{\mathfrak{p}} = \frac{mc(\rho)}{mc(\rho) - k\rho^{k-1}}.$$

Despite their austere appearance, these formulæ have indeed an a fairly intuitive content. First, the equation satisfied by ρ can be put under the form $mz = 1 + m^{-k}/c(z)$, and, since ρ is close to $\frac{1}{m}$, we may expect the approximation

$$m\rho \approx 1 + \frac{1}{\gamma m^k},$$

where $\gamma := c(m^{-1})$ satisfies $1 \leq \gamma < m/(m-1)$. By similar principles, the probabilities in (38) should be approximately

$$\mathbb{P}_{\mathcal{W}_n}(\mathfrak{p} \text{ does not occur}) \approx \left(1 + \frac{1}{\gamma m^k}\right)^{-n} \approx e^{-n/(\gamma m^k)}.$$

For a binary alphabet, this tells us that the occurrence of a pattern of length k starts becoming likely when n is of the order of 2^k, that is, when k is of the order of $\log_2 n$. The more precise moment when this happens must depend (via γ) on the autocorrelation of the pattern, with strongly correlated patterns having a tendency to occur a little late. (This vastly generalizes our empirical observations of Chapter I.) However, observe that the mean number of occurrences of a pattern in a text of length n does not depend on the shape of the pattern. This apparent paradox is easily resolved: correlated patterns tend to occur late, but they lend themselves to appearing in clusters. Thus, the late pattern aaa when it occurs still has probability $\frac{1}{2}$ to occur at the next position as well, and cash in another occurrence, whereas no such possibility is available to the early pattern aab whose occurrences must be somewhat spread out.

Such analyses are important as they can be used to develop a precise understanding of the behaviour of data compression algorithms (the Lempel–Ziv scheme); see Julien Fayolle's memoir (Master Thesis, Paris, 2002) for details.

$\triangleright$ **26.** *Multiple pattern occurrences.* A similar analysis applies to the generating function $S^{\langle s \rangle}(z)$ of words containing a fixed number s of occurrences of a pattern $\mathfrak{p}$. The OGF is obtained by expanding (with respect to u) the BGF $W(z,u)$ obtained in Chapter III by means of an inclusion-exclusion argument. For $s \geq 1$, one finds

$$S^{\langle s \rangle}(z) = z^k \frac{N(z)^{s-1}}{D(z)^{s+1}}, \qquad D(z) = z^k + (1 - mz)c(z), \quad N(z) = z^k + (1 - mz)(c(z) - 1)),$$

which now has a pole of multiplicity $s + 1$ at $z = \rho$. $\triangleleft$

$\triangleright$ **27.** *Patterns in Bernoulli sequences.* Similar results hold when letters are assigned nonuniform probabilities, $p_j = \mathbb{P}(a_j)$, for $a_j \in \mathcal{A}$. One only needs to define the weighted autocorrelation polynomial by its coefficient c_j being $c_j = \mathbb{P}(\mathfrak{p}_1 \cdots \mathfrak{p}_j)$, when $\mathfrak{p}$ coincides with its jth shifted version. Multiple pattern occurrences can be also analysed. $\triangleleft$

IV. 6. Singularities and functional equations

In the various combinatorial examples discussed so far in this chapter, we have been dealing with functions that are given by explicit expressions. Such situations essentially cover nonrecursive structures as well as the simplest recursive structures, like Catalan or Motzkin trees, whose generating functions are expressible in terms of radicals. In fact, as will shall see extensively in this book, complex analytic methods are instrumental in analysing coefficients of functions *implicitly* specified by functional equations. In other words: *the very nature of a functional equation can often provide clues regarding the singularities of its solution*. Chapter V will illustrate this philosophy in the case of rational functions defined by systems of positive equations; a very large number of examples will then be given in Chapters VI and VII, where singularities much more general than mere poles are treated. The purpose of this subsection is simply to offer a preliminary discussion of the way dominant singularities can be located in many cases by means means of simple iteration or inversion properties of analytic functions. Three typical functional equations are to be discussed here:

$$f(z) = ze^{f(z)}, \qquad f(z) = z + f(z^2 + z^3), \qquad f(z) = \frac{1}{1 - zf(z^2)}.$$

Inverse functions. We start with a generic problem: given a function ψ analytic at a point y_0 with $z_0 = \psi(y_0)$ what can be said about its inverse, namely the solution(s) to the equation $\psi(y) = z$ when z is near z_0 and y near y_0? Two cases occur depending on the value of $\psi'(y_0)$.

Regular case. If $\psi'(y_0) \neq 0$, then ψ admits an analytic expansion near y_0:

$$\psi(y) = \psi(y_0) + (y - y_0)\psi'(y_0) + \frac{1}{2}(y - y_0)^2\psi''(y_0) + \cdots .$$

Solving formally for y indicates a *locally linear* dependency,

$$y - y_0 \sim \frac{1}{\psi'(y_0)}(z - z_0). \tag{39}$$

A full formal expansion of $y - y_0$ in powers of $z - z_0$ is obtained by repeated substitution,

$$y - y_0 = c_1(z - z_0) + c_2(z - z_0)^2 + \cdots \tag{40}$$

and the method of majorizing series shows that the series so obtained converges locally in a sufficiently small neighbourhood of z_0. Rouché's theorem (equivalently, the analytic version of the Implicit Function Theorem, see Note 25), implies that the equation $\psi(y) = z$ admits there a unique analytic solution. In summary, *an analytic function locally admits an analytic inverse near any point where its first derivative is nonzero.*

Singular case. If to the contrary one has $\psi'(y_0) = 0$ and $\psi''(y_0) \neq 0$, then the expansion of ψ is of the form

$$\psi(y) = \psi(y_0) + \frac{1}{2}(y - y_0)^2\psi''(y_0) + \cdots . \tag{41}$$

Solving formally for y now indicates a *locally quadratic* dependency

$$(y - y_0)^2 \sim \frac{2}{\psi''(y_0)}(z - z_0),$$

and the inversion problem admits two solutions satisfying

$$y - y_0 \sim \pm\sqrt{\frac{2}{\psi''(y_0)}}(z - z_0)^{1/2}.$$

The point z_0 is thus a branch point.

A similar reasoning applies whenever the first nonzero derivative of ψ at y_0 is of order $r \geq 2$ (with a local behaviour for y then of the form $(z - z_0)^{1/r}$). Thus, the dependency between y and z cannot be analytic around (y_0, z_0). In other words, *an analytic function is not locally invertible in an analytic manner in the vicinity of any point where its first derivative is zero.*

We can now consider the problem of obtaining information on the coefficients of a function $y(z)$ defined by an implicit equation

$$y(z) = z\phi(y(z)). \tag{42}$$

For simplicity, we shall momentarily assume $\phi(u)$ to be a nonlinear entire function (possibly a polynomial of degree ≥ 2) with nonnegative coefficients. In order for the problem to be (formally) well-posed we assume that $\phi(0) \neq 0$.

The equation (42) occurs in the counting of various types of trees. For instance, $\phi(u) = e^u$ corresponds to labelled Cayley trees, $\phi(u) = (1 + u)^2$ to binary trees, and $\phi(u) = 1 + u + u^2$ to plane unary–binary trees (Motzkin trees). A full analysis of the problem was developed by Meir and Moon [**138**], themselves elaborating on earlier ideas of Pólya [**158, 159**] and Otter [**151**].

Equation (42) may be rephrased as

$$\psi(y(z)) = z \qquad \text{where} \qquad \psi(u) = \frac{u}{\phi(u)}, \tag{43}$$

so that it is a generic instance of the inversion problem for analytic functions: $y = \psi^{(-1)}$. We first observe that (42) and (43) admit unique formal power series solutions by the method of indeterminate coefficients. An application of the technique of majorizing series shows that this formal solution also represents an analytic function near the origin, with $y(0) = 0$. In addition, the coefficients of $y(z)$ are all nonnegative.

Now comes the hunt for singularities. The function $y(z)$ increases along the positive real axis. The equation $\psi'(\tau) = 0$ which is expected to create singularities for $y(z)$ is in terms of ϕ:

$$\phi(\tau) - \tau\phi'(\tau) = 0. \tag{44}$$

The function $\phi(u) = \sum_{k=0}^{\infty} \phi_k u^k$ being by assumption entire, the equation (44) is equivalent to

$$\phi_0 = \phi_2\tau^2 + 2\phi_3\tau^3 + \cdots,$$

which admits a unique positive solution.

As z increases, starting from 0 along the positive real axis, $y(z)$ increases. Let $\rho \leq \infty$ be the dominant positive singularity of $y(z)$. We are going to prove *a contrario* that $y(\rho) = \tau$ (technically, we should define $y(\rho)$ as the limit of $y(x)$ as $x \to \rho^-$). Assume that $y(\rho) < \tau$; then $y(z)$ could be analytically continued at $z = \rho$, by the discussion above of inverse functions in the regular case, since $\phi'(y(\rho)) > 0$. If on the other hand, we had $y(\rho) > \tau$, then, there would be a value $\rho^* < \rho$ such that $y(\rho^*) = \tau$; but there, we would have $\psi'(y(\rho^*)) = 0$, so that $y(z)$ would be singular at $z = \rho^*$ by the discussion on inverse functions in the singular case. Thus, in both cases, the assumption $y(\rho) \neq \tau$ leads to a contradiction. We thus obtain that $y(\rho) = \tau$, and, since y and ψ are inverse functions, this corresponds to

$$\rho = \psi(\tau) = \tau/\phi(\tau).$$

Equipped with this discussion, we state a result which covers situations more general than the case of ϕ being entire.

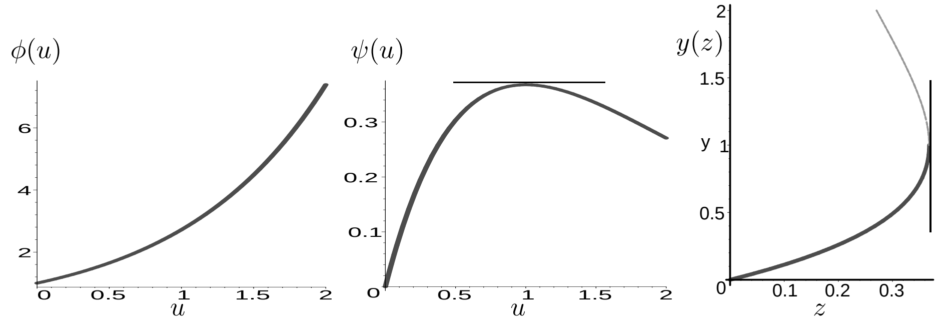


FIGURE 13. Singularities of inverse functions: $\phi(u) = e^u$ (left); $\psi(u) = u/\phi(u)$ (middle); $y = \text{Inv}(\psi)$ (right).

PROPOSITION IV.4. *Let ϕ be a nonlinear function that is analytic at 0, with nonnegative Taylor coefficients and radius of convergence $R \le +\infty$. Assume that there exists $\tau \in (0, R)$ such that*

$$\frac{\tau\phi'(\tau)}{\phi(\tau)} = 1. \tag{45}$$

Let $y(z)$ be the solution analytic at the origin of the equation $y(z) = \phi(y(z))$. Then, one has the exponential growth formula:

$$[z^n]\, y(z) \bowtie \left(\frac{1}{\rho}\right)^n \qquad \textit{where} \quad \rho = \frac{\tau}{\phi(\tau)} = \frac{1}{\phi'(\tau)}.$$

Note that, by Supplement 28 below, there can be *at most* one solution of the characteristic equation (45) in $(0, R)$, a necessary and sufficient condition for the existence of a solution in the open interval $(0, R)$ being $\lim_{x\to R^-} \frac{x\phi'(x)}{\phi(x)} > 1$. This last condition is automatically granted as soon as $\phi(R) = +\infty$.

PROOF. The discussion above applies verbatim. The function $y(z)$ is analytic around 0 (by majorizing series techniques). By the already seen argument, its value $y(\rho)$ cannot be different from τ, so that its radius of convergence must equal ρ. The form of y_n then results from general exponential bounds. $\square$

$\triangleright$ **28.** *Convexity of GFs and the Variance Lemma.* Let $\phi(z)$ be a nonlinear GF with nonnegative coefficients and a nonzero radius of convergence R. For $x \in (0, R)$ a parameter, define the *Boltzmann random variable* Ξ (of parameter x) by the property

$$\mathbb{P}(\Xi = n) = \frac{\phi_n x^n}{\phi(x)}, \qquad \text{with} \quad \mathbb{E}(s^\Xi) = \frac{\phi(sx)}{\phi(x)} \tag{46}$$

the probability generating function of Ξ. By differentiation, the first two moments of Ξ are

$$\mathbb{E}(\Xi) = \frac{x\phi'(x)}{\phi(x)}, \qquad \mathbb{E}(\Xi^2) = \frac{x^2\phi''(x)}{\phi(x)} + \frac{x\phi'(x)}{\phi(x)}.$$

There results, for any nonlinear GF $\phi(x)$, the general convexity inequality

$$\frac{d}{dx}\left(\frac{x\phi'(x)}{\phi(x)}\right) > 0,$$

since the variance of a nondegenerate random variable is always positive. Equivalently, the function $\log(\phi(e^t))$ is convex for $t \in (-\infty, \log R)$. [In statistical phsyics, a Boltzmann model (of parameter x) corresponds to a class Φ (with ogf ϕ) from which elements are drawn according to the size distribution (46).] $\triangleleft$

Take for instance general Catalan trees corresponding to

$$y = \frac{z}{1-y(z)}, \qquad \text{so that} \quad \phi(u) = \frac{1}{1-u}.$$

We have $R = 1$ and the characteristic equation reads

$$\frac{\tau}{1-\tau} = 1,$$

implying $\tau = \frac{1}{2}$, so that $\rho = \frac{1}{4}$. We obtain as anticipated $y_n \bowtie 4^n$, a weak asymptotic formula for the Catalan numbers. Similarly, for Cayley trees, we have $\phi(u) = e^u$, the characteristic equation reduces to $(\tau - 1)e^\tau = 0$, so that $\tau = 1$ and $\rho = e^{-1}$, giving a weak form of Stirling's formula:

$$[z^n]y(z) = \frac{n^{n-1}}{n!} \bowtie e^n.$$

Here is a table of a few cases of application of the method to structures already encountered in previous chapters.

Type	$\phi(u)$	(R)	τ, ρ	$y_n \bowtie \rho^{-n}$
binary tree	$(1+u)^2$	(∞)	$1, \frac{1}{4}$	$y_n \bowtie 4^n$
Motzkin tree	$1+u+u^2$	(∞)	$1, \frac{1}{3}$	$y_n \bowtie 3^n$
gen. Catalan tree	$\dfrac{1}{1-u}$	(1)	$\frac{1}{2}, \frac{1}{4}$	$y_n \bowtie 4^n$
Cayley tree	e^u	(∞)	$1, e^{-1}$	$y_n \bowtie e^n$

In fact, for all such problems, the dominant singularity is always of the square-root type as our previous discussion suggests. Accordingly, the asymptotic form of coefficients is invariably of the type

$$[z^n]\, y(z) \sim C \cdot \rho^{-n} n^{-3/2},$$

as we shall prove in Chapter VI by means of the singularity analysis method.

$\triangleright$ **29.** *A variant form of the inversion problem.* Consider the equation $y = z + \phi(y)$, where ϕ is assumed to be entire (say) and $\phi(u) = O(u^2)$ at $u = 0$. This corresponds to a simple variety of trees in which trees are counted by the number of their leaves only. For instance, we have already encountered labelled hierarchies (phylogenetic trees) in Section II.6 corresponding to $\phi(u) = e^u - 1 - u$, which is one of "Schröder's problems". Let $\widetilde{\tau}$ be the root of $\phi'(\widetilde{\tau}) = 1$ and set $\widetilde{\rho} = \widetilde{\tau} - \phi(\widetilde{\tau})$. Then $[z^n]y(z) \bowtie \widetilde{\rho}^{-n}$. For the EGF L of labelled hierarchies ($L = z + e^L - 1 - L$), this gives $L_n/n! \bowtie (2\log 2 - 1)^{-n}$. (Observe that Lagrange inversion also provides $[z^n]y(z) = \frac{1}{n}[w^{n-1}](1 - y^{-1}\phi(y))^{-n}$.) $\triangleleft$

Iteration. Consider the class $\mathcal{E}$ of *balanced 2–3 trees* defined as trees whose node degrees are restricted to the set $\{0, 2, 3\}$, with the additional property that all leaves are at the same distance from the root. Such tree trees, which are particular cases of B-trees, are a useful data structure for implementing dynamic dictionaries [**121**]. We adopt as notion of size the number of leaves (also called external nodes). The OGF of $\mathcal{E}$ satisfies the functional equation

$$E(z) = z + E(z^2 + z^3), \tag{47}$$

which reflects an inductive definition involving a substitution: given an existing tree, a new tree is obtained by substituting in all possible ways to each external node ($\square$) either a pair

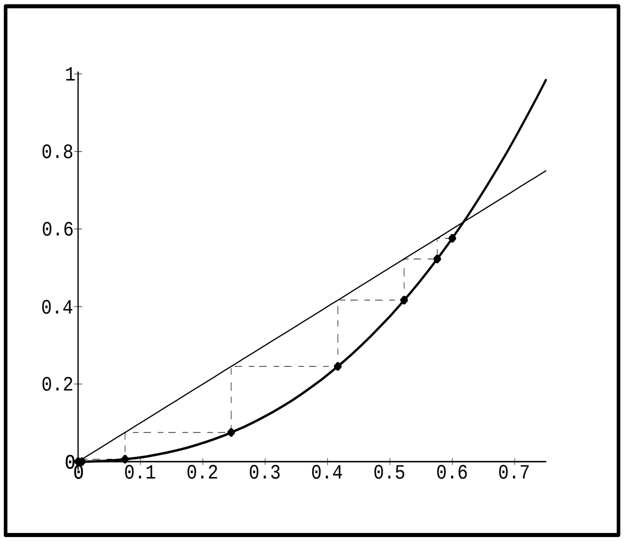


FIGURE 14. The iterates of a point $x_0 \in [0, \frac{1}{\varphi}[$ (here $x_0 = 0.6$) by $\sigma(z) = z^2 + z^3$ converge fast to 0.

$(\Box, \Box)$ or a triple $(\Box, \Box, \Box)$. On other words, we have

$$\mathcal{E}[\Box] = \Box + \mathcal{E}\left[\Box \to (\Box\Box + \Box\Box\Box)\right].$$

Equation (47) implies the seemingly innocuous recurrence

$$E_n = \sum_{k=0}^{n} \binom{k}{n-2k} E_k \qquad \text{with} \quad E_0 = 0,\ E_1 = 1,$$

but no closed-form solution is known (nor likely to exist) for E_n or $E(z)$. The expansion starts as (the coefficients are *EIS* **A014535**)

$$E(z) = z + z^2 + z^3 + z^4 + 2\,z^5 + 2\,z^6 + 3\,z^7 + 4\,z^8 + 5\,z^9 + 8\,z^{10} + \cdots.$$

We present here the first stage of an analysis due to Odlyzko [**146**] and corresponding to exponential bounds. Let $\sigma(z) = z^2 + z^3$. Equation (47) can be expanded by iteration in the ring of formal power series,

(48) $$E(z) = z + \sigma(z) + \sigma^{[2]}(z) + \sigma^{[3]}(z) + \cdots,$$

where $\sigma^{[j]}(z)$ denotes the jth iterate of the polynomial σ:

$$\sigma^{[0]}(z) = z, \qquad \sigma^{[h+1]}(z) = \sigma^{[h]}(\sigma(z)) = \sigma(\sigma^{[h]}(z)).$$

Thus, $E(z)$ is nothing but the sum of all iterates of σ. The problem is to determine the radius of convergence of $E(z)$, and by Pringsheim's theorem, the quest for dominant singularities can be limited to the positive real line.

For $z > 0$, the polynomial $\sigma(z)$ has a unique fixed point, $\rho = \sigma(\rho)$, at

$$\rho = \frac{1}{\varphi} \qquad \text{where} \qquad \varphi = \frac{1+\sqrt{5}}{2},$$

the golden ratio. Also, for any positive x satisfying $x < \rho$, the iterates $\sigma^{[j]}(x)$ must converge to 0; see Fig. 14. Furthermore, since $\sigma(z) \sim z^2$ near 0, these iterates converge

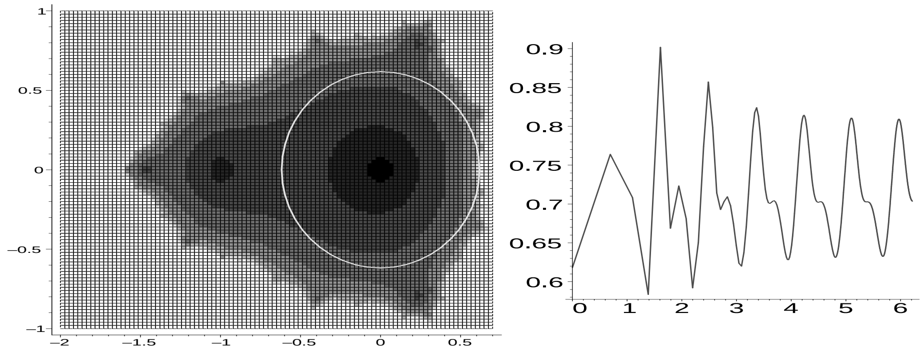


FIGURE 15. Left: the fractal domain of analyticity of $E(z)$ in gray with darker areas representing faster convergence of the sum of iterates of σ. Right: the ratio $E_n/(\varphi^n n^{-1})$ plotted against $\log n$ for $n = 1\,..\,500$ confirms that $E_n \bowtie \varphi^n$ and illustrates the periodic fluctuations expressed by Equation (51).

to 0 doubly exponentially fast. First, for $x \in [0, \frac{1}{2}]$, one has $\sigma(x) \le \frac{3}{2}x^2$ for $x \in [0, \frac{1}{2}]$, so that there

$$\sigma^{[j]}(x) \le \left(\frac{3}{2}\right)^{2^j-1} x^{2^j}. \tag{49}$$

Second, for $x \in [0, A]$, where A is any number $< \rho$, there is a number k_A such that $\sigma^{[k_A]}(x) < \frac{1}{2}$, so that, by (49), there holds:

$$\sigma^{[k]}(x) \le \frac{3}{2}\left(\frac{3}{4}\right)^{2^{k-k_A}}.$$

Thus, the series of iterates of σ is quadratically convergent for $z \in [0, A]$, any $A < \rho$.

By the triangular inequality, $|\sigma(z)| \le (\sigma(|z|)$, the sum in (48) is a normally converging sum of analytic functions, and is thus itself analytic. Consequently $E(z)$ is analytic in the whole of the open disk $|z| < \rho$.

It remains to prove that the radius of convergence of $E(z)$ is exactly equal to ρ. To that purpose it suffices to observe that $E(z)$, as given by (48), satisfies

$$E(x) \to +\infty \qquad \text{as} \qquad x \to \rho^-.$$

Let N be an arbitrarily large but fixed integer. It is possible to select a positive x_N sufficiently close to ρ with $x_N < \rho$, such that the Nth iterate $\sigma^{[N]}(x_N)$ is larger than $\frac{1}{2}$ (the function $\sigma^{[N]}(x)$ admits ρ as a fixed point and it is continuous and increasing at ρ). Given the sum expression (48), this entails the lower bound $E(x_N) > \frac{N}{2}$ for such an $x_N < \rho$ so that $E(x)$ is unbounded as $x \to \rho^-$.

The dominant positive real singularity of $E(z)$ is thus $\rho = \frac{1}{\varphi}$, and application of Cauchy bounds shows that

$$[z^n]\,E(z) \bowtie \left(\frac{1+\sqrt{5}}{2}\right)^n. \tag{50}$$

It is notable that this estimate could be established so simply by a purely qualitative examination of the basic functional equation and of a fixed point of the associated iteration scheme.

The complete asymptotic analysis of the E_n was given by Odlyzko [**146**] in a classic paper. It requires the full power of singularity analysis methods to be developed in Chapter VI. Equation (51) below states the end result, which involves periodic fluctuations; see Figure 15 (right). There is overconvergence of the representation (48), that is, convergence in certain directions beyond the disc of convergence of $E(z)$, as illustrated by Figure 15 (left). The proof techniques involve an investigation of the behaviour of iterates of σ in the complex plane, an area launched by Fatou and Julia in the first half of the past century and nowadays well-studied under the name of "complex dynamics".

▷ **30.** *The asymptotic number of 2–3 trees.* This analysis is from [**146, 147**]. The number of 2–tree trees satisfies asymptotically

$$E_n = \frac{\varphi^n}{n}\Omega(\log n) + O\left(\frac{\varphi^n}{n^2}\right), \tag{51}$$

where Ω is a periodic function with mean value $\varphi(\log(4-\varphi) \doteq 0.71208$ and period $\log(4-\phi) \doteq$. Thus oscillations are inherent in E_n. A plot of the ratio $E_n/(\phi^n/n)$ is offered in Figure 15. ◁

Complete asymptotics of a functional equation. This is Pólya's counting of certain molecules, a case where only a functional equation is known for a generating function, $M(z) = \sum_n M_n z^n$:

$$M(z) = \frac{1}{1 - zM(z^2)}. \tag{52}$$

The M_n represent the number of chemical isomeres of alcohols $C_nH_{2n+1}OH$ without asymmetric carbon atoms, and the series starts as (*EIS* **A000621**)

$$M(z) = 1 + z + z^2 + 2z^3 + 3z^4 + 5z^5 + 8z^6 + 14z^7 + 23z^8 + 39z^9 + \cdots. \tag{53}$$

By iteration of the functional equation, one finds a continued fraction representation:

$$M(z) = \cfrac{1}{1 - \cfrac{z}{1 - \cfrac{z^2}{1 - \cfrac{z^4}{\ddots}}}}.$$

Pólya [**159**] who established this functional equation in the historical paper that introduced "Pólya Theory" developed at the same time a precise asymptotic estimate for M_n.

PROPOSITION IV.5. *Let $M(z)$ be the solution analytic around 0 of the functional equation*

$$M(z) = \frac{1}{1 - zM(z^2)}.$$

Then, there exist constants K and α such that

$$M_n \sim K \cdot \alpha^n, \qquad \alpha \doteq 1.68136\,75244, \quad K \doteq 0.36071\,40971.$$

PROOF. We offer two proofs. The first one is based on direct consideration of the functional equation and is of a high degree of applicability. The second one, following Pólya, makes explicit a linear structure present in the problem and leads to more explicit results.

First proof. The first few coefficients of M are determined by the functional equation and known (53). Then, by positivity of the functional equation, $M(z)$ dominates coefficientwise any GF $(1 - zM_{<m}(z^2)^{-1}$, where $M_{<m}(z)$ is the mth truncation of $M(z)$. In particular, one has the domination relation (use $M_{<2}(z) = 1 + z$)

$$M(z) \preceq \frac{1}{1 - z - z^3}.$$

Since the rational fraction has a dominant poles at $z \doteq 0.68232$, this implies that the radius ρ of convergence of $M(z)$ satisfies $\rho < 0.69 < 1$. In the other direction, since $M(z^2) < M(z)$ for $z \in (0, \rho)$, then, one has the numerical inequality

$$M(z) \leq \frac{1}{1 - zM(z)}, \qquad 0 \leq z < \rho.$$

This can be used to show that the Catalan generating function $C(z) = (1 - \sqrt{1 - 4z})/(2z)$ is a majorant of $M(z)$ on the interval $(0, \frac{1}{4})$ and that $M(z)$ exists for $z \in (0, \frac{1}{4})$. In other words, one has $\frac{1}{4} \leq \rho < 0.69$. Altogether, the radius of convergence of M is strictly between 0 and 1.

$\triangleright$ **31.** *Alcohols and trees.* Since $M(z)$ starts as $1+z+z^2+\cdots$ while $C(z)$ starts as $1+z+2z^2+\cdots$, there is a small interval $(0, \epsilon)$ such that $M(z) \leq C(z)$. By the functional equation of $M(z)$, one has $M(z) \leq C(z)$ for z in the larger interval $(0, \sqrt{\epsilon})$. One can then bootstrap and show that $M(z) \leq C(z)$ for $z \in (0, \frac{1}{4})$. $\triangleleft$

Next, as $z \to \rho^-$, one must have $zM(z^2) \to 1$. Indeed, if this was not the case, we would have $zM(z^2) < A < 1$ for some A. But then, since $\rho^2 < \rho$, the quantity $(1 - zM(z^2))^{-1}$ would be analytic at $z = \rho$, a clear contradiction. Thus, ρ is determined implicitly by the equation

$$\rho M(\rho^2) = 1,$$

and by monotonicity, there can be only one such solution. Numerically, one can estimate ρ as the limit of quantities ρ_m satisfying

$$\sum_{n=0}^{m} M_n \rho_m^{2n+1} = 1,$$

together with $\rho \in [\frac{1}{4}, 0.069]$. In each case, only a few of the M_n are needed. One obtains in this way:

$$\rho_{10} \doteq 0.595,\ \rho_{20} \doteq 0.594756,\ \rho_{30} \doteq 0.59475397,\ \rho_{40} \doteq 0.594753964,$$

and it is not hard to verify that this provides a geometrically convergent scheme to the limit $\rho \doteq 0.59475\,39639$. (Note: Pólya determined ρ to five decimals by hand!)

The previous discussion also implies that ρ is a pole, which must be simple. Thus

$$M(z) \sim z \to \rho K \frac{1}{1 - z/\rho}, \qquad K := \frac{1}{\rho M(\rho^2) + 2\rho^3 M'(\rho^2)}. \tag{54}$$

The argument shows at the same time that $M(z)$ is meromorphic in $|z| < \sqrt{\rho} \doteq 0.77$. That $M(z)$ is a the only pole on $|z| = \rho$ can be seen from the fact that $zM(z^2) = z + z^3 + \cdots$ is unperiodic in the sense of Chapter V. (We don't detail the argument here as the property is also implied by the developments of the second proof.) The translation of the singular expansion (54) yields the statement.

Second proof. First, a sequence of formal approximants follows from (52) starting with

$$1,\ \frac{1}{1-z},\ \frac{1}{1 - \dfrac{z}{1-z^2}} = \frac{1-z^2}{1-z-z^2},\ \frac{1}{1 - \dfrac{z}{1 - \dfrac{z^2}{1-z^4}}} = \frac{1-z^2-z^4}{1-z-z^2-z^4+z^5}.$$

which permits to compute any number of terms of the series $M(z)$. Closer examination of (52) suggests to set

$$M(z) = \frac{\psi(z^2)}{\psi(z)},$$

where

$$\psi(z) = 1 - z - z^2 - z^4 + z^5 - z^8 + z^9 + z^{10} + z^{17} + z^{18} + z^{20} - z^{21} - z^{37} - \cdots$$

Back substitution into (52) yields

$$\frac{\psi(z^2)}{\psi(z)} = \frac{1}{1 - z\dfrac{\psi(z^4)}{\psi(z^2)}} \text{ or } \frac{\psi(z^2)}{\psi(z)} = \frac{\psi(z^2)}{\psi(z^2) - z\psi(z^4)},$$

which shows $\psi(z)$ to be a solution of the functional equation

$$\psi(z) = \psi(z^2) - z\psi(z^4).$$

The coefficients of ψ are all in the set $\{0, -1, +1\}$, as they satisfy the recurrence

$$\psi_{4n} = \psi_{2n},\ \psi_{4n+1} = -\psi_n,\ \psi_{4n+2} = \psi_{2n+1},\ \psi_{4n+3} = 0.$$

Thus, $M(z)$ appears as the quotient of two function, $\psi(z^2)/\psi(z)$; since $\psi(z)$ whose coefficients are bounded by 1 in absolute value, it is analytic in the unit disk, $M(z)$ is itself meromorphic in the unit disc. A numerical plot shows that that $\psi(z)$ has its smallest positive real zero at $\rho \doteq 0.59475$, which is a simple zero of $\psi(z)$ and thus a pole of $M(z)$ as $\psi(\rho^2) \neq 0$. Thus

$$M(z) \sim \frac{\psi(\rho^2)}{(z-\rho)\psi'(\rho)} \implies M_n \sim -\frac{\psi(\rho^2)}{\rho\psi'(\rho)} \left(\frac{1}{\rho}\right)^n.$$

Numerical computations then yield Pólya's estimate. Et voilà! □

The example of Pólya's alcohols is exemplary, both from a historical point of view and from a methodological perspective. It demonstrates that quite a lot of information can be pulled out of a functional equation without solving it. (A very similar situation will be discussed in Chapter V, see the enumeration of coin fountains.) In passing, we have made great use of the fact that if $f(z)$ is analytic in $|z| < r$ and some bounds imply the strict inequalities $0 < r < 1$, then one can regard functions like $f(z^2)$, $f(z^3)$, and so on, as "known" since they are analytic in the disc of convergence of f and even beyond, a situation evocative of our earlier discussion of Pólya operators in Subsection IV. 3.3. Globally, the lesson is that functional equation, even very complicated ones, can often be used to bootstrap the local singular behaviour of solutions and one can do so despite the absence of any explicit generating function solution. Then, the transition from singularities to coefficient asymptotics is a simple jump.

▷ **32.** *An arithmetic exercise* Find a characterization of $\psi_n = [z^n]\psi(z)$ based on the binary representation of n. Tabulate ψ_n for all $n \in (10^{1000}, 10^{1000} + 10^{500})$, possibly using some compressed format. Find the asymptotic proportion of the ψ_n for $n \in [1\,..\,N]$ that are nonzero. ◁

IV. 7. Notes

This chapter has been designed to serve as a refresher of basic complex analysis, with special emphasis on methods relevant for analytic combinatorics. References most useful for the discussion given in this chapter include the books of Titchmarsh [**181**] (oriented towards classical analysis), Whittaker and Watson [**191**] (stressing special functions), Dieudonné [**46**], Hille [**108**], and Knopp [**117**]. Henrici [**107**] presents complex analysis under the perspective of constructive and numerical methods, a highly valuable point of view for this book. References dealing specifically with asymptotic analysis are discussed at the end of the next chapter.

As demonstrated by the first batch of examples sprinkled over this chapter, singularities provide a royal road to coefficient asymptotics. In this regard, the two main statements of this chapter are are the theorems relative to the expansion of rational and meromorphic functions, Theorems IV.6 and IV.7. They are of course extremely classical (and easy) results. Issai Schur (1875–1941) is to be counted amongst the very first mathematicians who recognized the rôle of analytic methods in combinatorial enumerations (Example 4). This thread was developed by George Pólya in his famous paper of 1937 (see [**158, 159**]), which Read in [**159**, p. 96] describes as a "landmark in the history of combinatorial analysis". There, Pólya founded at the same time combinatorial chemistry, the enumeration of objects under group actions, and the complex-asymptotic theory of graphs and trees.

De Bruijn's classic booklet [**42**] is a wonderfully concrete introduction to effective asymptotic theory, and it contains many examples from discrete mathematics thoroughly worked out. The state of affairs in 1995 regarding analytic methods in combinatorial enumeration is superbly summarized in Odlyzko's scholarly chapter [**147**]. Wilf devotes his Chapter 5 of *Generatingfunctionoloy* [**193**] to this question. The books by Hofri [**109**] and Szpankowski [**179**] contain useful accounts in the perspective of analysis of algorithms. See also our book [**168**] for a light introduction and the chapter by Vitter and Flajolet [**188**] for more on this topic.

Paraphrasing the number theorist Hecke, we may feel confident in stating: *A function's singularities contain a wealth of asymptotic information regarding the function's coefficients; a generating function's singularities contain a wealth of information regarding quantitative properties of the corresponding combinatorial structures*. This philosophy furthermore unites analytic combinatorics and analytic number theory. It is the purpose of the next four chapters to illustrate it thoroughly by means of a great variety of combinatorial examples.

CHAPTER V

Applications of Rational and Meromorphic Asymptotics

Analytic methods are extremely powerful and when they apply,
they often yield estimates of unparalleled precision.

—ANDREW ODLYZKO [**147**]

The primary goal of this chapter is to provide combinatorial illustrations of the power of complex analytic methods, and specifically of the rational–meromorphic framework. At the same time, we shift gears and envisage counting problems at a new level of generality. Precisely, we consider combinatorial-analytic *schemas*, which, broadly speaking, are wide *families* of combinatorial types amenable to a common analytic framework and associated with a common collection of asymptotic properties.

The first schema comprises regular specifications and languages, which *a priori* leads to rational generating functions and thus systematically resort to Theorem IV.6. This is not the end of the story, however, since in general one is interested not just in a single set of combinatorial objects, but rather in a whole family of classes. The case of patterns in words at the end of the previous chapter has already exemplified this situation. Here, we extend the analysis to the determination of longest runs, corresponding to maximal sequences of good (or bad) luck in games of chance. In so doing, we develop analytical methods that apply in many cases to largest components. We then consider an important class of regular specifications, the ones that are built on nested sequences and combinatorially correspond to lattice paths. Besides providing a precise quantification of height in Dyck paths, this also leads to the determination of height in random (general) Catalan trees. The treatment is to a large extent made possible because nested sequence constructions lead naturally to nested quasi-inverses, that is, continued fractions. And continued fractions enjoy a wealth of algebraic and analytic properties.

Next, we discuss a general schema of analytic combinatorics known as the *supercritical sequence* schema, which provides a prime illustration of the power of meromorphic asymptotics while being of a very wide applicability. For instance, one can predict very precisely (and easily) the number of ways that an integer can be decomposed additively as a sum of primes (or twin primes), this even though many details of the distribution of primes are still surrounded in mystery.

Last we discuss positive linear systems of generating functions: although the resulting generating functions are once more bound to be rational, there is benefit in examining them as defined implicitly (rather than solving explicitly) and work out singularities directly. The crucial technical tool there is the Perron-Frobenus theory of nonnegative matrices, whose importance has been long recognized in the theory of finite Markov chains. A general discussion of singularities can then be conducted, leading to valuable consequences on a variety of models—paths in graphs, finite automata, and transfer matrices.

All these cases illustrate the power of rational and meromorphic asymptotics. The last example discussed treats locally constrained permutations, where rational functions even provide an entry to the world of permutations.

Universality is a term originating with statistical physics that is also nowadays increasingly used in probability theory. By universality is meant a collection of key properties that are shared by a wide family of models and are largely independent of particulars of each models. For instance, in statistical physics, random placements of pieces or random walks on a regular lattice share common properties that do not depend on the particular geometry of the lattice, whether square, triangular, or honeycomb. In probability theory, it is established that sums of random variable converge to a Gaussian limit, so that the Gaussian law is universal for sums of random variables (under suitably mild moment conditions). In this spirit, we can describe the supercritical sequence as universal accross combinatorics as it covers a large family of models simply characterized by the presence of an external sequence construction ($\mathcal{F} = \mathfrak{S}(\mathcal{G})$) accompanied with a natural analytic assumption ("supercriticality"). Alignments, compositions, and surjections for instance find themselves sheltered under a common umbrella and analytic theory tells us that they *must* share many features, like having a linear number of components in the mean and with high probability, an asymptotically predetermined proportion of components of each possible type, and so on. In a similar spirit, one can regard exponential-polynomial behaviour as universal across all problems described by regular expressions (Sections V. 1 and V. 2) or by finite state models (Section V. 4 and V. 5).

V. 1. Regular specification and languages

A combinatorial specification is said to be *regular* if it is nonrecursive ("iterative") and it involves only the constructions of Atom, Union, Product, and Sequence; see Chapter I. For convenience and without loss of analytic generality, we consider here unlabelled structures. Since the operators translating these constructions into generating functions are all of a rational nature, it follows that the corresponding OGFs are invariably rational. Then Theorem IV.6 applies directly:

THEOREM V.1 (Regular specification asymptotics). *Let $\mathcal{C}$ be an unlabelled class that is described by a regular specification. Then the coefficients of the OGF $C(z)$ satisfy an exponential-polynomial formula,*

$$C_n \equiv [z^n]C(z) = \sum_{j=1}^{m} \Pi_j(n)\alpha_j^{-n}, \tag{1}$$

for a family of algebraic numbers α_j and a family of polynomials Π_j.

General trees of bounded height, denumerants, as well as partitions and compositions into summands at most r constitute prime examples of structures admitting regular specifications.

The name "regular specification" has been chosen so as to be in agreement with the notions of regular expression and regular language from formal language theory introduced in Chapter I. We saw there that a language is called S–regular ("specification regular") if it is combinatorially isomorphic with a class $\mathcal{R}$ which admits a regular specification.

The most frequent case is that of a language specified by a *regular expression*, involving letters of the alphabet, union, catenation, and Kleene star. If the regular expression is unambiguous, i.e., every word is uniquely parsable (see APPENDIX: *Regular expressions*, p. 233), it is combinatorially isomorphic to a regular specification. In the general case, one may encounter regular expressions that are ambiguous; then, the systematic application of the translation rules amounts to counting every word with its multiplicity, that is, the number of ways in which it can be parsed.

PROPOSITION V.1 (Regular expression counting). *Given a regular expression R (assumed to be of finite ambiguity), the ordinary generating function $L_R(z)$ of the language $\mathcal{L}(R)$, counting* with multiplicity*, is given by the inductive rules:*

$$\epsilon \mapsto 1, \quad a \mapsto z, \quad \cup \mapsto +, \quad \cdot \mapsto \times, \quad \star \mapsto (1-(.))^{-1}.$$

In particular, if R is unambiguous, then the ordinary generating function satisfies $L_R(z) = L(z)$ and is given directly by the rules above. In both cases, the coefficients $[z^n]L_R(z)$ admit of an exponential-polynomial form.

Note. If R is ambiguous, it is known that one can build an unambigous R' such that $\mathcal{L}(R) = \mathcal{L}(R')$. Consequently, the conclusions of Proposition V.1 extend in principle to counting without multiplicities words in *any* regular language. One then has however to rely on an indirect automaton construction (see the appendices) which computational complexity is in general exponential.

PROOF. Formal rules associate to any proper regular expression R a specification $\mathcal{R}$:

$$\epsilon \mapsto 1 \text{ (the empty object)}, \quad a \mapsto Z_a \ (Z_a \text{ an atom}),$$
$$\cup \mapsto +, \cdot \mapsto \times, \quad \star \mapsto \mathfrak{S}\{.\}$$

It is readily recognized that this mapping is such that $\mathcal{R}$ generates exactly the collection of all parsings of words according to R. The translation rules of Chapter 1 then yield the first part of the statement. The second part follows since $L(z) = L_R(z)$ whenever R is unambiguous. ☐

EXAMPLE 1. *A potpourri of regular specifications.* We briefly recapitulate here a number of combinatorial problems already encountered in Chapters I–III that are reducible to regular specifications.

Compositions of integers (Section I.3) are specified by $\mathcal{C} = \mathfrak{S}(\mathfrak{S}_{\geq 1}(\mathcal{Z}))$, whence the OGF $(1-z)/(1-2z)$ and the closed form $C_n = 2^{n-1}$, an especially trivial exponential-polynomial form. Polar singularities are also present for compositions into k summands $(\mathfrak{S}_k(\mathfrak{S}_{\geq 1}(\mathcal{Z}))$ and for compositions whose summands are restricted to the interval $[1, r]$ $(\mathfrak{S}(\mathfrak{S}_{1\,..\,r}(\mathcal{Z}))$, with corresponding generating functions

$$\frac{z^k}{(1-z)^k}, \qquad \frac{1-z}{1-2z+z^{r+2}}.$$

In the first case, one has an explicit form for the coefficients, $\binom{n-1}{k-1}$, which is also a particular exponential-polynomial form (with the basis of the exponential being 1). The second case requires a dedicated analysis of the dominant polar singularity, a task that is undertaken in Example 2 below for the closely related problem of determining longest runs in random binary words. We shall also see later (Section V. 3 and Example 9) that a rich class of summand-restricted compositions resorts to the framework of meromorphic asymptotics.

Integer partitions involve the multiset construction. However, when summands are restricted to the interval $[1\,..\,r]$, the specification satisfies the combinatorial identity (Section I.3),

$$\mathfrak{M}(\mathfrak{S}_{1\,..\,r}(\mathcal{Z})) \simeq \mathfrak{S}(\mathcal{Z}) \times \mathfrak{S}(\mathcal{Z}^{\times 2}) \times \cdots \mathfrak{S}(\mathcal{Z}^{\times r}),$$

corresponding to the OGF

$$\prod_{j=1}^{r} \frac{1}{1-z^j}.$$

This case has already served as a leading example in our discussion of denumerants in Example IV.4, where the analysis of the pole at 1 furnishes the dominant asymptotic behaviour ($n^{k-1}/(k!(k-1)!)$) of these special partitions.

Words lead to many problems that are prototypical of the regular specification framework. In Section 1.4, we saw that one could give a regular expression describing the set of words containing the pattern abb, from which the exact and asymptotic forms of counting coefficients derive. The case of long runs of a single letter is similarly amenable to to regular specifications and is detailed below. Note however that, for a general pattern $\mathfrak{p}$, the generating functions of words constrained to include (or dually exclude) $\mathfrak{p}$ are best based on the inclusion-exclusion argument of Section III.6.4. The corresponding asymptotic analysis has already served as a pilot example in Section IV. 5.3 of the previous chapter.

Words can also be analysed under the Bernoulli model, where letter i is selected with probability p_i; cf Section III.5 for a general discussion. We saw there that one can put regular specifications to good use in order to analyse the mean number of records in a random word.

Set partitions are typically labelled objects. However, when suitably constrained, they can sometimes be encoded by words described by regular expressions; see Section I.4.3 for partitions into k classes, where the OGF has been found to be

$$S^{(k)}(z) = \frac{z^k}{(1-z)(1-2z)\cdots(1-kz)} \qquad \text{implying} \qquad S_n^{(k)} \sim \frac{k^n}{k!},$$

where the asymptotic estimate results from the dominant pole at $1/k$.

Trees have generating functions that, in all nondegenerate cases, are beyond rational functions. However, the generating function of general (Catalan) trees of height $\le h$ *is* rational; see Section III.7 relative to extremal parameters. The corresponding analysis is detailed below, Section V. 2 and Example 6, in relation to the enumeration of Dyck paths in a strip. □

▷ **1.** *Partially commutative monoids.* Let $\mathcal{W} = \mathcal{A}^\star$ be the set of all words over a finite alphabet $\mathcal{A}$ whose letters are also considered as formal indeterminates. Consider a set C of commutation rules between pairs of elements of A. For instance, if $\mathcal{A} = \{a, b, c\}$, then $\mathrm{C} = \{ab = ba, ac = ca\}$ means that a commutes with both b and c, but bc is not a commuting pair: $bc \ne cb$. Let $\mathcal{M} = \mathcal{W}/[\mathrm{C}]$ be the set of equivalent classes of words (monomials) under the rules induced by C. $\mathcal{M}$ is called a *partially commutative monoid* or a trace monoid.

If $A = \{a, b\}$, then the two possibilities for C are $\mathrm{C} = \emptyset$ and $\mathrm{C} := \{ab = ba\}$. Normal forms for $\mathcal{M}$ are given by the regular expressions $(a+b)^\star$ and $a^\star b^\star$ corresponding to the OGFs

$$\frac{1}{1-a-b}, \qquad \frac{1}{1-a-b+ab}.$$

If $\mathcal{A} = \{a, b, c\}$, the possibilities for C, the corresponding normal forms, and the OGFs M are as follows. If $C = \emptyset$, then $\mathcal{M} \simeq (a+b)^\star$ with OGF $(1-a-b-c)^{-1}$; the other cases are

$$\begin{array}{ccc} ab = ba & ab = ba,\ ac = ca & ab = ba,\ ac = ca,\ bc = cb \\ (a^\star b^\star c)^\star a^\star b^\star & a^\star (b+c)^\star & a^\star b^\star c^\star \\ \dfrac{1}{1-a-b-c+ab} & \dfrac{1}{1-a-b-c+ab+ac+bc} & \dfrac{1}{1-a-b-c+ab+ac+bc-abc} \end{array}.$$

Cartier and Foata [**33**] have proved the general result (based on extended Moebius inversion),

$$M = \left(\sum_F (-1)^{|F|} F\right)^{-1},$$

in which the sum is over all monomials F formed with distinct letters that all commute pairwise. Goldwurm and Santini [**95**] have proved that $[z^n]M(z) \sim K \cdot \alpha^n$ for some $K, \alpha > 0$. $\lhd$

EXAMPLE 2. *Longest runs in words* The analysis of longest runs in words provides an illustration of the technique of localizing dominant singularities in rational functions and of the corresponding coefficient extraction process. In Chapter I, we have determined the family of OGFs describing the length L of the longest run of consecutive a's in a binary word over the alphabet $\mathcal{W} = \{a, b\}$. The counting GF associated with the property $(L < k)$ for a fixed k is a rational function. Determining the probability distribution of L over the set of all words of length n is then equivalent to analysing the *whole* family of GFs indexed by k. The probabilistic problem is a famous one, discussed by Feller in [**58**], as it represents a basic question in the analysis of runs of good (or bad) luck in a succession of independent events. Our presentation closely follows an insightful note of Knuth [**119**] whose research was motivated by the related problem of analysing carry propagation in certain binary adders.

PROPOSITION V.2. *The longest run parameter L taken over the set of binary words of length n (endowed with the uniform distribution) satisfies, for h in any bounded set of $\mathbb{Z}$, the uniform estimate*[1]

$$\mathbb{P}_n\left(L < \lfloor \lg n \rfloor + h\right) = e^{-\alpha(n)2^{-h}} + O\left(\frac{\log n}{n}\right), \qquad \alpha(n) := 2^{\{\lg n\}}. \tag{2}$$

In particular, the mean and variance satisfy $\mathbb{E}_n(L) = \lg n + O(1)$ and $\mathbb{V}_n(L) = O(1)$, and the distribution is concentrated around its mean.

The probability distribution in (2) is called a *double exponential distribution* (Figure 1). In fact, the formula is an *asymptotic* one. It does not represent a unique *limit distribution* in the usual sense, but rather a whole family depending on the fractional part of $\lg n$, that is, on the way n is placed with respect to powers of 2. This phenomenon is further reflected by the fact that the second asymptotic term in the mean is subject to fluctuations (albeit of a tiny amplitude), see the discussion of $\Phi(x)$ below.

PROOF. The specification $\mathcal{W}^{\langle k\rangle} = a^{<k}\mathfrak{S}(ba^{<k})$ describes those words for which this length is strictly less than k. The expression of the OGF,

$$W^{\langle k\rangle}(z) = \frac{1-z^k}{1-z} \cdot \frac{1}{1 - z\frac{1-z^k}{1-z}} = \frac{1-z^k}{1-2z+z^{k+1}}, \tag{3}$$

results. Quite clearly, one should locate the dominant pole, separate it from the other poles (as this leads to constructive error terms), as well as estimate the contribution to the coefficients arising from this dominant pole.

[1] The symbol $\lg x$ denotes the binary logarithm, $\lg x = \log_2 x$.

(i) *Locating the dominant pole.* The OGF $W^{\langle k\rangle}$ has, by the first form of (3) a dominant pole ρ_k which is a root of the equation $1 = s(\rho_k)$, where $s(z) = z(1-z^k)/(1-z)$. We consider $k \geq 2$. Since $s(z)$ is an increasing polynomial and $s(0) = 0$, $s(\frac{1}{2}) < 1$, $s(1) = 1$, the root ρ_k must lie in the open interval $(\frac{1}{2}, 1)$. In fact, as one easily verifies, the condition $k \geq 2$ guarantees that $s(0.6) > 1$, hence the refined estimate

$$\frac{1}{2} < \rho_k < \frac{3}{5} \qquad (k \geq 2). \tag{4}$$

It now becomes possible to derive very precise estimates by bootstrapping. (This technique is a form of iteration for approaching a fixed point—its use in the context of asymptotic expansions is detailed in De Bruijn's book [**42**].) Writing the defining equation for ρ_k as a fixed point equation,

$$z = \frac{1}{2}(1 + z^k),$$

and making use of the rough estimates (4) yields next

$$\frac{1}{2}\left(1 + (\frac{1}{2})^k\right) < \rho_k < \frac{1}{2}\left(1 + (\frac{3}{5})^k\right). \tag{5}$$

Thus, ρ_k is exponentially close to $\frac{1}{2}$, and a further iteration from (5) shows

$$\rho_k = \frac{1}{2} + \frac{1}{2^{k+1}} + O\left(\frac{k}{2^{2k}}\right), \tag{6}$$

which constitutes a very precise estimate.

(ii) *Contribution from the dominant pole.* A straight calculation provides the value of the residue,

$$R_{n,k} := -\operatorname{Res}\left[W^{\langle k\rangle}(z) z^{-n-1}; z = \rho_k\right] = \frac{1-\rho_k^k}{2-(k+1)\rho^k}\rho_k^{-n-1}, \tag{7}$$

which is expected to provide the main approximation to the coefficients of $W^{\langle k\rangle}$ as $n \to \infty$. The meaning of (7) is better grasped if one notes that the residue resembles $2^n e^{-n/2^k}$. We shall return to such approximations shortly.

(iii) *Separation of the subdominant poles.* Consider the circle $|z| = \frac{3}{4}$ and take the second form of the denominator of $W^{\langle k\rangle}$, namely

$$1 - 2z + z^{k+1}.$$

In view of Rouché's theorem, we may regard this polynomial as the sum $f(z) + g(z)$, where $f(z) = 1 - 2z$ and $g(z) = z^{k+1}$. The term $f(z)$ has on the circle a modulus that varies between $\frac{1}{2}$ and $\frac{5}{2}$; the term $g(z)$ is at most $\frac{27}{64}$ for any $k \geq 2$. Thus, on the circle $|z| = \frac{3}{4}$, one has $|g(z)| < |f(z)|$, so that $f(z)$ and $f(z) + g(z)$ have the same number of zeros inside the circle. Since $f(z)$ admits $z = \frac{1}{2}$ as only zero there, the denominator must also have a unique root in $|z| \leq \frac{3}{4}$, and that root must coincide with ρ_k.

Similar arguments also give bounds on the error term when the number of words with longest run of length at most k is estimated by the residue (7) at the dominant pole. On the disc $|z| = \frac{3}{4}$, the denominator of $W^{\langle k\rangle}$ stays bounded away from 0 (its modulus is at least $\frac{5}{64}$ when $k \geq 2$, by previous considerations). Thus, the modulus of the remainder integral is $O((4/3)^n)$, and in fact bounded from above by $35(4/3)^n$. In summary, if we let $q_{n,k}$ represent the probability that the longest run in a random word of length n is less than k, one has available the main estimate

$$q_{n,k} := \mathbb{P}_n(L < k) = \frac{1-\rho_k^k}{1-(k+1)\rho_k^k/2}\left(\frac{1}{2\rho_k}\right)^{n+1} + O\left((\frac{2}{3})^n\right), \tag{8}$$

uniformly with respect to k. Here is table of the numerical values of the quantities appearing in the approximation of $q_{n,k}$ when written under the form $c_k \cdot (2\rho_k)^{-n}$:

k	$c_k \cdot (2\rho_k)^{-n}$
2	$1.17082 \cdot 0.80901^n$
3	$1.13745 \cdot 0.91964^n$
4	$1.09166 \cdot 0.96378^n$
5	$1.05753 \cdot 0.98297^n$
10	$1.00394 \cdot 0.99950^n$

(iv) There finally remains to transform the main estimate (8) into the limit form asserted in the statement. First, the "tail inequalities"

$$\mathbb{P}_n\left(L < \frac{3}{4}\lg n\right) = O\left(e^{-\frac{1}{2}\sqrt[4]{n}}\right), \qquad \mathbb{P}_n\left(L \geq 2\lg n\right) = O\left(\frac{1}{n}\right), \tag{9}$$

describe the tail of the probability distribution of L_n. They derive from simple bounding techniques applied to the main approximation (8) using (6). Thus, for asymptotic purposes, only a small region around $\lg n$ needs to be considered.

Regarding the central regime, for $k = \lg n + x$ and x in $[-\frac{1}{4}\lg n, \lg n]$, the approximation (6) of ρ_k and related quantities applies, and one finds

$$(2\rho_k)^{-n} = \exp\left(-\frac{n}{2^k} + O(kn2^{-2k})\right) = e^{-n/2^k}\left(1 + O(\frac{\log n}{n})\right).$$

(This results from standard expansions like $(1-a)^n = e^{-na}\exp(O(na^2))$.) At the same time, the coefficient of this quantity in (8) is

$$1 + O(k\rho_k^k) = 1 + O\left(\frac{\log n}{n^{3/4}}\right).$$

Thus a double exponential approximation holds (Figure 1) and for $k = \lg n + x$ with x in $[-\frac{1}{4}\lg n, \lg n]$, one has (uniformly)

$$q_{n,k} = e^{-n/2^k}\left(1 + O\left(\frac{\log n}{n^{3/4}}\right)\right). \tag{10}$$

In particular, upon setting $k = \lfloor \lg n \rfloor + h$, the first part of the statement follows. (The floor function takes into account the fact that k must be an integer.)

The mean and variance estimates derive from the fact that the distribution quickly decays at values away from $\lg n$ while it satisfies (10) in the central region. The mean is given by

$$\mathbb{E}_n(L) := \sum_{h\geq 0}\left[1 - \mathbb{P}_n(L < h)\right] = \Phi(n) + O\left(\frac{\log^2 n}{n}\right), \quad \Phi(x) := \sum_{h\geq 0}\left[1 - e^{-x/2^h}\right].$$

Consider the three cases $h < h_0$, $h \in [h_0, h_1]$, and $h > h_1$ with $h_0 = \lg x - \log\log x$ and $h_1 = \lg x + \log\log x$, where the general term is (respectively) close to 1, between 0 and 1, and close to 0. By summing, one finds elementarily $\Phi(x) = \lg x + O(\log\log x)$ as $x \to \infty$, and elementary ways of catching the next $O(1)$ term are discussed for instance in [**168**, p. 403].

The method of choice for precise asymptotics is to treat $\Phi(x)$ as a harmonic sum and apply Mellin transform techniques (APPENDIX: *Mellin Transform*, p. 227). The Mellin transform of $\Phi(x)$ is

$$\Phi^\star(s) := \int_0^\infty \Phi(x)x^{s-1}\,dx = \frac{\Gamma(s)}{1-2^s} \qquad \Re(s) \in (-1, 0).$$

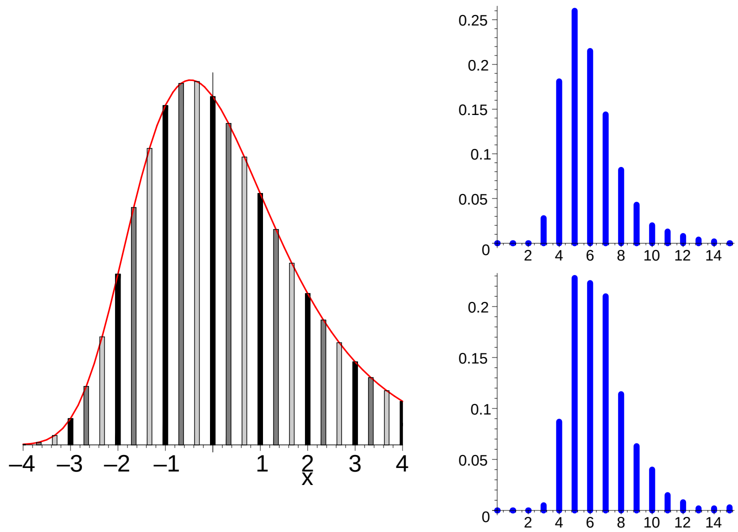


FIGURE 1. The double exponential laws: Left, histograms for n at 2^p (black), $2^{p+1/3}$ (dark gray), and $2^{p+2/3}$ (light gray), where $x = k - \lg n$. Right, empirical histograms for 1000 simulations with $n = 100$ (top) and $n = 140$ (bottom).

The double pole of $\Phi^\star$ at 0 and the simple poles at $s = \frac{2ik\pi}{\log 2}$ are reflected by the asymptotic expansion:

$$\Phi(x) = \lg x + \frac{\gamma}{\log 2} + \frac{1}{2} + P(\lg x) + O(x^{-1}), \quad \text{where} \quad P(w) = \sum_{k \in \mathbb{Z}\setminus\{0\}} \Gamma\left(\frac{2ik\pi}{\log 2}\right) e^{2ik\pi w}.$$

The oscillating function $P(w)$ has amplitude of the order of 10^{-6}. (See [**70, 119, 179**] for more on this topic.) The variance is similarly analysed. □

The analysis is closely related to the case of words excluding a patterns in Chapter IV. There, we conducted a global analysis applicable to any pattern. Here, we have specialized the discussion to patterns $\mathsf{aaa}\cdots\mathsf{a}$ and effectively extracted a whole family of limit distributions. What is striking is the existence of an *infinite family of limit laws*, which depend on the fractional part of $\lg n$. □

▷ **2.** *Longest runs in Bernoulli sequences.* Consider an alphabet $\mathcal{A}$ with letters independently chosen according to the probability distribution $\{p_j\}$. Then, the OGF of words where each letter is repeated at most k times derives from the construction of Smirnov words and is

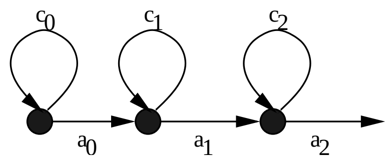
$$W^{[k]}(z) = \left(1 - \sum_i p_i z \frac{1-(p_i z)^k}{1-(p_i z)^{k+1}}\right)^{-1}.$$

Let $p_{\max}$ be the smallest of the p_j. Then the expected length of the longest run of any letter is $\log n / \log p_{\max} + O(1)$, and very precise quantitative information can be derived from the OGFs by methods akin to Example 7 (Smirnov words and Carlitz compositions) in Chapter IV, p. 34. ◁

The next batch of examples in this section develops the analysis of walks in a special type of graphs. These examples serve two purposes: they illustrate further cases of modelling by means of regular specifications, and, at the same time, provide a bridge with the analysis of lattice paths in the next section.

EXAMPLE 3. *Walks of the pure-birth type.* Consider a walk on the nonnegative integers that starts at 0 and is only allowed to either stay at the same place or progress by an increment of $+1$. Our goal is to enumerate the possible configurations that start from 0 and reach point $m-1$ in n steps. A step from j to $j+1$ will be encoded by a letter a_j; a step from state j to state j will be encoded by c_j. A diagram representing these steps is then:

(11)



(Compare with (21).) The language encoding all legal walks from state 0 to state $m-1$ can be described by a regular expression,

$$\mathcal{H}_{0,m-1} = (c_0)^\star a_0 (c_1)^\star a_1 \cdots (c_{m-2})^\star a_{m-2} (c_{m-1})^\star,$$

and the representation is certainly unambiguous. Symbolicly using letters as variables, the corresponding ordinary multivariate generating function is then

$$H_{0,m-1}(\vec{a},\vec{c}) = \frac{a_0 a_1 \cdots a_{m-2}}{(1-c_0)(1-c_1)\cdots(1-c_{m-1})}.$$

Assume that the steps are assigned weights, with α_j corresponding to a_j and γ_j to c_j. Weights of letters are extended multiplicatively to words in the usual way (cf Chapter III). If in addition, one takes $\gamma_j = 1-\alpha_j$, one obtains a probabilistic weighting: the walker starts from position 0, and, if at j, at each clock tick, she either stays at the same place with probability $1-\alpha_j$ or moves to the right with probability α_j. The OGF of such weighted walks then becomes

$$H_{0,m-1}(z) = \frac{\alpha_0\alpha_1\cdots\alpha_{m-2} z^{m-1}}{(1-(1-\alpha_0)z)(1-(1-\alpha_1)z)\cdots(1-(1-\alpha_{m-1})z)}, \tag{12}$$

and $[z^n]H$ is the probability for the walker to be found at position j at (discrete) time n. This walk process can be alternatively interpreted as a (discrete-time) pure *birth process* in the usual sense of probability theory: There is a population of individuals and, at each discrete epoch, a new birth may take place, the probability of a birth being α_j when the population is of size j.

The form (12) readily lends itself to a partial fraction decomposition. The poles of H are at the points $(1-\alpha_j)^{-1}$ and one finds as $z \to (1-\alpha_j)^{-1}$:

$$H_{0,m-1}(z) \sim \frac{r_{j,m-1}(1-\alpha_j)}{1-z(1-\alpha_j)} \quad \text{where} \quad r_{j,m-1} := \frac{\alpha_0\alpha_1\cdots\alpha_{m-2}}{\prod_{k\in[0,m-1],\ k\neq\{j\}}(\alpha_k-\alpha_j)}.$$

Thus, the probability of being in state $m-1$ at time n is

$$[z^n]H_{0,m-1}(z) = \sum_{j=0}^{m-1} r_{j,m-1}(1-\alpha_j)^{n+1}.$$

This has the form of an alternating sum that can be evaluated in each particular instance.

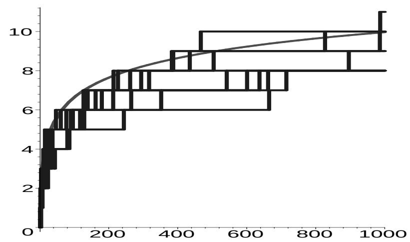


FIGURE 2. A simulation of 10 trajectories of the pure-birth process till $n = 1024$, with geometric probabilities corresponding to $q = 1/2$, compared to the curve $\log_2 x$.

An especially interesting case of the pure-birth walk is when the quantities α_k are geometric: $\alpha_k = q^{k+1}$ for some q with $0 < 1 < k$. In that case, the probability of being in state $m-1$ after n transitions becomes

$$\sum_{j=0}^{m-1} \frac{(-1)^j q^{\binom{j}{2}}}{(q)_j (q)_{m-j-1}} (1 - q^{m-j-1})^{n+1}, \quad (q)_j := (1-q)(1-q^2)\cdots(1-q^j).$$

This corresponds to a stochastic progression in a medium with exponentially increasing hardness or, equivalently, to the growth of a population where the current size of the population adversely affects fertility in an exponential manner. On intuitive grounds, we expect an evolution of the process to stay reasonably close to the curve $y = \log_{1/q} x$; see Figure 2 for a simulation confirming this fact, which can be justified by means of the analytic formulæ just described. This particular analysis is borrowed from [**63**], where it was initially developed in connection with the algorithm called "approximate counting" to be described below. □

Note. The theory of pure birth processes is discussed under a calculational and non measure-theoretic angle in the book by Bharucha-Reid [**24**]. See also the *Course* by Karlin and Taylor [**113**] for a concrete presentation.

EXAMPLE 4. *Approximate Counting.* Assume you need to keep a counter that is able to record the number of certain events (say impulses) and should have the capability of keeping counts till a certain maximal value N. A standard information-theoretic argument (with ℓ bits, one can only keep track of 2^ℓ possibilities) implies that one needs $\lceil \log_2 N+1 \rceil$ bits to perform the task—a standard binary counter will indeed do the job. However, in 1977, Robert Morris has proposed a way to maintain counters that only requires of the order of $\log\log N$ bits. What's the catch?

Morris' elegant idea consists in relaxing the constraint of exactness in the counting process and, by playing with probabilities, tolerate a small error on the counts obtained. Precisely, his solution maintains a random quantity Q which is initialized by $Q = 0$. Upon receiving an impulse, one updates Q according to the following simple procedure (with $q \in (0,1)$ a design parameter):

procedure Update(Q);
  with probability q^{Q+1} do Q := $Q+1$ (else keep Q unchanged).

When asked the number of impulses (number of times the update procedure was called) at any moment, simply use the following procedure to return an estimate:

procedure Answer(Q);
  output $\dfrac{q^{-Q}-1}{1-q}$.

Let Q_n be the value of the random quantity Q after n executions of the update procedure and X_n the corresponding estimate output by the algorithm. It is easy to verify (by recurrence or by generating functions, see Note 3 below) that

$$\mathbb{E}(q^{-Q_n}) = n(1-q)+1, \qquad \text{so that} \quad \mathbb{E}(X_n) = n. \tag{13}$$

Thus the answer provided at any instant is an *unbiased estimator* (in a mean value sense) of the actual count n. On the other hand, the analysis of the geometric pure-birth process in the previous example applies. In particular, the exponential approximation $(1-\alpha)^n \approx e^{-n\alpha}$ in conjunction with the basic formulæ show that for large n and m sufficiently near to $\log_{1/q} n$, one has (asymptotically) the *geometric-birth distribution*

$$\mathbb{P}\left(Q_n = \log_{1/q} n + x\right) = \sum_{j=0}^{\infty} \frac{(-1)^j q^{\binom{j}{2}}}{(q)_j (q)_\infty} \exp(-q^{x-j-1}) + o(1). \tag{14}$$

(We refer to [**63**] for details.) Such calculations imply that Q_n is with high probability (w.h.p.) close to $\log_{1/q} n$. Thus, if $n \le N$, the value of Q_n will be w.h.p. bounded from above by $(1+\epsilon)\log_{1/q} N$, with ϵ a small constant. But this means that the integer Q, which can itself be represented in binary, will only require

$$\log_2 \log n + O(1) \tag{15}$$

bits for storage, for fixed q.

A closer examination of the formulæ reveals that the accuracy of the estimate improves considerably when q becomes close to 1. The *standard error* is defined as $\frac{1}{n}\sqrt{\mathbb{V}(X_n)}$ and it measures (in a mean quadratic sense) the relative error to likely to be made. The variance of Q_n is, like the mean, determined by recurrence or generating functions, and one finds

$$\mathbb{V}(q^{-Q_n}) = \binom{n}{2}\frac{(1-q)^3}{q}, \qquad \frac{1}{n}\sqrt{\mathbb{V}(X_n)} \sim \sqrt{\frac{1-q}{q}}. \tag{16}$$

This means that accuracy increases as q approaches 1 and, by suitably dimensioning q, one can make it as small as desired. In summary, (13), (16), and (15) express the following property: *Approximate counting makes it possible to count till N using only about* $\log\log N$ *bits of storage, while achieving a standard error that is almost a constant and can be set to any prescribed value.* Morris' trick is now fully understood.

For instance, with $q = 2^{-1/16}$, it proves possible to count up to $2^{16} = 65536$ using only 8 bits (instead of 16), with an error likely not to exceed 20%. Naturally, there's not too much reason to appeal to the algorithm when a *single* counter needs to be managed. (Everybody can afford a few bits!) Approximate Counting turns out to be useful when a very large number of counts need to be kept *simultaneously*. It constitutes one of the early examples of a probabilistic algorithm in the management of large volumes of data, also known as *data mining*.

Functions akin to those of (14) also surface in other areas of probability theory. Guillemin, Robert, and Zwart [**101**] have detected them in processes that combine an additive increase and a multiplicative decrease (AIMD processes), in a context motivated by the adaptive transmission of "windows" of varying sizes in large communication networks (the TCP protocol of the internet). Biane, Bertoin, and Yor [**22**] encountered a function identical to (14) in their study of exponential functionals of Poisson processes. □

$\triangleright$ **3.** *Moments of* q^{-Q_n}. It is a perhaps surprising fact that any integral moment of q^{-Q_n} is a polynomial in n and q, like in (13), (16). To see it, define

$$\Phi(w) \equiv \Phi(w, \xi, q) := \sum_{m\geq 0} q^{m(m+1)/2} \frac{\xi^m w^m}{(1+\xi q)(1+\xi q^2)\cdots(1+\xi q^{m+1})}.$$

By (12), one has

$$\sum_{m\geq 0} H_{0,m} w^m = \frac{1}{1-z}\Phi\left(w; \frac{z}{1-z}, q\right).$$

On the other hand, Φ satisfies $\Phi(w) = 1 - q\xi(1-w)\Phi(qw)$, hence the q–identity,

$$\Phi(w) = \sum_{j\geq 0} (-q\xi)^j \left[(1-w)(1-qw)\cdots(1-q^{j-1}w)\right],$$

which resorts to q-calculus[2]. Thus $\Phi(q^{-r}; \xi, q)$ is a polynomial for any $r \in \mathbb{Z}_{\geq 0}$, as the expansion terminates. $\triangleleft$

Our last example makes use of regular expressions in order to estimate moments. Note that ambiguous representations are purposely used to accomplish the task.

EXAMPLE 5. *Occurrences of "hidden" patterns in texts.* Fix an alphabet $\mathcal{A} = \{a_1, \ldots, a_r\}$ of cardinality r and assume a probability distribution on $\mathcal{A}$ to be given, with p_j the probability of letter a_j. We consider the Bernoulli model on $\mathcal{W} = \mathfrak{S}(\mathcal{A})$, where the probability of a word is the product of the probabilities of its letters (cf Section III.5). A word $\mathfrak{p} = y_1 \cdots y_k$ called the pattern is fixed. The problem is to gather information on the random variable X representing the number of occurrences of $\mathfrak{p}$ in the set $\mathcal{W}_n$, where occurrences as a *"hidden pattern"*, i.e., as a *subsequence*, are counted (Section I.4.1). This is a basic example where counting with ambiguity proves useful.

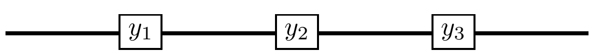
The generating function associated to $\mathcal{W}$ endowed with its probabilistic weighting is

$$W(z) = \frac{1}{1 - \sum p_j z} = \frac{1}{1-z}.$$

The regular expression

$$\mathcal{O} = \mathfrak{S}(\mathcal{A}) y_1 \mathfrak{S}(\mathcal{A}) \cdots \mathfrak{S}(\mathcal{A}) y_{k-1} \mathfrak{S}(\mathcal{A}) y_k \mathfrak{S}(\mathcal{A}) \tag{17}$$

describes all contexts of occurrences of $\mathfrak{p}$ as a subsequence in all words. Graphically, this may be rendered as follows for a pattern of length 3, $\mathfrak{p} = y_1 y_2 y_3$:

(18)


There the boxes indicate distinguished positions where letters of the pattern appear and the horizontal lines represent arbitrary separating words ($\mathfrak{S}(\mathcal{A})$). The corresponding OGF

$$O(z) = \frac{\pi(\mathfrak{p}) z^k}{(1-z)^{k+1}}, \qquad \pi(\mathfrak{p}) := p_{y_1} \cdots p_{y_{k-1}} p_{y_k} \tag{19}$$

counts elements of $\mathcal{W}$ with *ambiguity*, where the ambiguity coefficient of a word $w \in \mathcal{W}$ is precisely equal to the number of occurrences of $\mathfrak{p}$ as a subsequence in w. There results that the expected number of hidden occurrences of $\mathfrak{p}$ in a random word of length n is

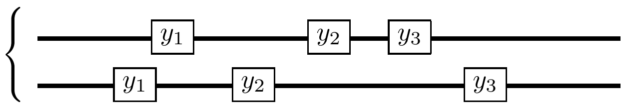
$$[z^n] O(z) = \pi(\mathfrak{p}) \binom{n}{k}, \tag{20}$$

which is consistent with what a direct probabilistic reasoning would give.

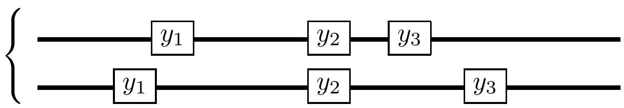
---

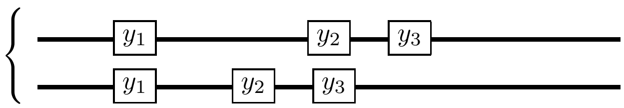
[2]By q–calculus is roughly meant the collection of special function identities relating power series of the form $\sum a_n(q) z^n$, where $a_n(q)$ is a rational fraction whose degree is quadratic in n. See [**6**, Ch. 10] for basics and [**90**] for more advanced (q–hypergeometric) material.

We next proceed to determine the variance of X over $\mathcal{W}_n$. In order to do so, we need contexts in which *pairs* of occurrences appear. Let $\mathcal{Q}$ denote the set of all words in $\mathcal{W}$ with *two* occurrences (i.e., an ordered pair of occurrences) of $\mathfrak{p}$ as a subsequence being distinguished. Then clearly $[z^n]Q(z)$ must represent $\mathbb{E}_{\mathcal{W}_n}[X^2]$. There are several cases to be considered. Graphically, a pair of occurrences may be interleaved and share no common position, like in what follows:

(21) 

But they may also have one or several overlapping positions, like in

(22) 

(23) 

(This last situation necessitates $y_2 = y_3$, typical patterns being abb and aaa.)

In the first case corresponding to (21), where there are no overlapping positions, the configurations of interest have OGF

$$Q^{[0]}(z) = \binom{2k}{k} \frac{\pi(\mathfrak{p})^2 z^{2k}}{(1-z)^{2k+1}}. \tag{24}$$

There, the binomial coefficient $\binom{2k}{k}$ counts the total number of ways of freely interleaving two copies of $\mathfrak{p}$; the quantity $\pi(\mathfrak{p})^2 z^{2k}$ takes into account the $2k$ distinct positions where the letters of the two copies appear; the factor $(1-z)^{-2k-1}$ corresponds to all the possible $2k+1$ fillings of the gaps between letters.

In the second case, let us start by considering pairs where exactly one position is overlapping, like in (22). Say this position corresponds to the rth and sth letters of $\mathfrak{p}$ (r and s may not be equal). Obviously, we need $y_r = y_s$ for this to be possible. The OGF of the configurations is now

$$\binom{r+s-2}{r-1}\binom{2m-r-s}{m-r}\frac{\pi(\mathfrak{p})^2 (p_{y_r})^{-1} z^{2k-1}}{(1-z)^{2k}}.$$

There, the first binomial coefficient $\binom{r+s-2}{r-1}$ counts the total number of ways of interleaving $y_1 \cdots y_{r-1}$ and $y_1 \cdots y_{s-1}$; the second binomial $\binom{2m-r-s}{m-r}$ is similarly associated to the interleavings of $y_{r+1} \cdots y_k$ and $y_{s+1} \cdots y_k$; the numerator takes into account the fact that $2k-1$ positions are now occupied by predetermined letters; finally the factor $(1-z)^{-2k}$ corresponds to all the $2k$ fillings of the gaps between letters. Summing over all possibilities for r, s gives the OGF of pairs with one overlapping position as

$$Q^{[1]}(z) = \left(\sum_{1 \le r,s \le k} \binom{r+s-2}{r-1}\binom{2m-r-s}{m-r} \frac{[\![y_r = y_s]\!]}{p_{y_r}} \right) \frac{\pi(\mathfrak{p})^2 z^{2k-1}}{(1-z)^{2k}}. \tag{25}$$

Similar arguments show that the OGF of pairs of occurrences with at least *two* shared positions (see, e.g., 23)) is of the form, with P a polynomial,

$$Q^{[\geq 2]}(z) = \frac{P(z)}{(1-z)^{2k-1}}, \tag{26}$$

for the essential reason that, in the finitely many remaining situations, there are at most $(2k-1)$ possible gaps.

We can now examine (24), (25), (26) in the light of singularities. The coefficient $[z^n]Q^{[0]}(z)$ is seen to cancel to first asymptotic order with the square of the mean as given in (20). The contribution of the coefficient $[z^n]Q^{[\geq 2]}(z)$ appears to be negligible as it is $O(n^{2k-2})$. The coefficient $[z^n]Q^{[1]}(z)$, which is $O(n^{2k-1})$, is seen to contribute to the asymptotic growth of the variance. In summary, after a trite calculation, we obtain:

PROPOSITION V.3. *The number X of occurrences of a hidden pattern $\mathfrak{p}$ in a random text of size n obeying a Bernoulli model satisfies*

$$\mathbb{E}_{\mathcal{W}_n}[X] = \pi(\mathfrak{p})\binom{n}{k} \sim \frac{\pi(\mathfrak{p})}{k!}n^k, \qquad \mathbb{V}_{\mathcal{W}_n}[X] = \frac{\pi(\mathfrak{p})^2\kappa(\mathfrak{p})^2}{(2k-1)!}n^{2k-1}\left(1+O(\frac{1}{n})\right),$$

where the "correlation coefficient" $\kappa(\mathfrak{p})^2$ is given by

$$\kappa(\mathfrak{p})^2 = \sum_{1\leq r,s\leq k}\binom{r+s-2}{r-1}\binom{2m-r-s}{m-r}[\![y_r = y_s]\!]\left(\frac{1}{p_{y_r}}-1\right).$$

In particular, the distribution of X is concentrated around its mean.

This example is based on an article by Flajolet, Szpankowski, and Vallée [**85**]. There the authors show further that the asymptotic behaviour of moments of higher order can be worked out. By the moment convergence theorem described in Chapter VII, this calculation entails that *the distribution of X over $\mathcal{W}_n$ is asymptotically normal*. The method also extends to a much more general notion of "hidden" pattern, e.g., distances between letters of $\mathfrak{p}$ can be constrained in various ways so as to determine a valid occurrence in the text [**85**]. It also extends to the very general framework of dynamical sources [**27**], which include Markov models as a special case. The two references [**27, 85**] thus provide a set of analyses that interpolate between the two extreme notions of pattern occurrence—as a block of consecutive symbols or as a subsequence ("hidden pattern"). Such studies demonstrate that hidden patterns are with high probability bound to occur an extremely large number of times in a long enough text—this might cast some doubts on numerological interpretations encountered in various cultures. □

▷ **4.** *Hidden patterns and shuffle relations.* To each pairs u, v of words over $\mathcal{A}$ associate the weighted-shuffle polynomial in the indeterminates $\mathcal{A}$ denoted by $\left(\!\binom{u}{v}\!\right)_t$ and defined by the properties

$$\begin{cases} \left(\!\!\binom{xu}{yv}\!\!\right)_t = x\left(\!\!\binom{u}{yv}\!\!\right)_t + y\left(\!\!\binom{xu}{v}\!\!\right)_t + t[\![x=y]\!]x\left(\!\!\binom{u}{v}\!\!\right)_t \\ \left(\!\!\binom{\mathbf{1}}{u}\!\!\right)_t = \left(\!\!\binom{u}{\mathbf{1}}\!\!\right)_t = u \end{cases}$$

where t is a parameter, x, y are elements of $\mathcal{A}$, and $\mathbf{1}$ is the empty word. Then the OGF of $Q(z)$ above is

$$Q(z) = \sigma\left[\left(\!\!\binom{\mathfrak{p}}{\mathfrak{p}}\!\!\right)_{(1-z)}\right]\frac{1}{(1-z)^{2k+1}},$$

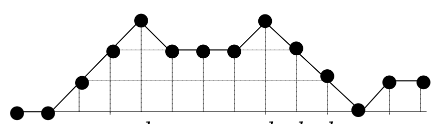
where σ is the substitution $a_j \mapsto p_j z$. ◁

V. 2. Lattice paths and walks on the line.

In this section, we consider *lattice paths* that are fundamental objects of combinatorics. Indeed, they relate to trees, permutations, and set partitions, to name a few. They also correspond to walks on the integer half-line and as such they relate to classical 1-dimensional random walks and to birth-and-death processes of probability theory. The lattice paths discussed here have steps that correspond to movements either immediately to the left or to the right. Combinatorially, such paths are the limit of paths of bounded height, themselves definable as nested sequences. As a consequence, the OGF's obtained involve a cascade of quasi-inverses, $1/(1-f)$, so that they are of the continued fraction type.

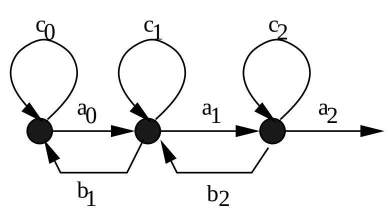
DEFINITION V.1 (Lattice path). *A (lattice)* path $\upsilon = (U_0, U_1, \ldots, U_n)$ *is a sequence of points in the lattice* $\mathbb{N}\times\mathbb{N}$ *such that if* $U_j = (x_j, y_j)$, *then* $x_j = j$ *and* $|y_{j+1} - y_j| \le 1$. *An edge* $\langle U_j, U_{j+1}\rangle$ *is called an* ascent $(\underline{a})$ *if* $y_{j+1} - y_j = +1$, *a* descent $(\underline{b})$ *if* $y_{j+1} - y_j = -1$, *and a* level step $(\underline{c})$ *if* $y_{j+1} - y_j = 0$.
The quantity n *is the* length *of the path,* $o(\upsilon) := y_0$ *is the* initial altitude, $h(\upsilon) := y_n$ *is the* final altitude. *A path is called an* excursion *if both its initial and final altitudes are zero. The extremal quantities* $\sup\{\upsilon\} := \max_j y_j$ *and* $\inf\{\upsilon\} := \min_j y_j$ *are called the* height *and* depth *of the path.*

It is assumed that paths are normalized by the condition $x_0 = 0$. With this normalization, a path of length n is encoded by a word with a, b, c representing ascents, descents, and level steps, respectively. What we call the *standard encoding* is such a word in which each step a, b, c is (redundantly) subscripted by the value of the y-coordinate of its associated point. For instance,



$$w = c_0\ a_0\ a_1\ a_2\ b_3\ c_2\ c_2\ a_2\ b_3\ b_2\ b_1\ a_0\ c_1$$

encodes a path that connects the initial point $(0, 0)$ to the point $(13, 1)$. Such a path can also be regarded as a rendering of the evolution in discrete time of a walk over the integer line:



Equivalently, lattice paths cane be read as trajectories of birth-and-death processes. (Compare with the pure-birth case in (11) above.)

Let $\mathcal{H}$ be the set of all lattice paths. Given a geometric condition (Q), it is then possible to associate to it a "language" $\mathcal{H}[Q]$ that comprises the collection of all path encodings satisfying the condition Q. This language can be viewed either as a set or as a formal sum,

$$H[Q] = \sum_{\{w \mid Q\}} w,$$

in which case it becomes the generating function in infinitely many indeterminates of the corresponding condition.

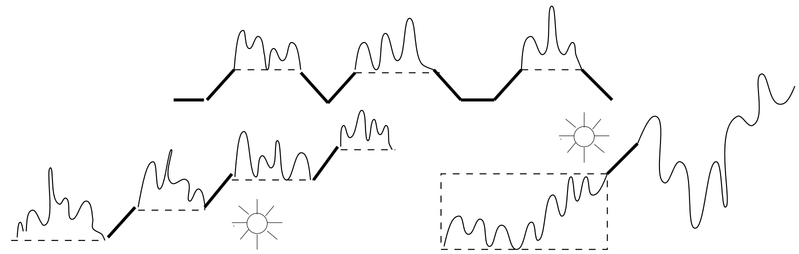


FIGURE 3. The three major decompositions of lattice paths: the arch decomposition (top), the last passages decomposition (bottom left), and the first passage decomposition (bottom right).

The general subclass of paths of interest in this subsection is defined by arbitrary combinations of flooring (m), ceiling (h), as well as fixing initial (k) and final (l) altitudes:

$$H_{k,l}^{[\geq m,<h]} \quad = \quad \{w \in H : o(w) = k,\ h(w) = l,\ \inf\{w\} \geq m,\ \sup\{w\} < h\}.$$

We also need the specializations, $H_{k,l}^{[<h]} = H_{k,l}^{[\geq 0,<h]}$, $H_{k,l}^{[\geq m]} = H_{k,l}^{[\geq m,<\infty]}$, $H_{k,l} = H_{k,l}^{[\geq 0,<\infty]}$. Three simple combinatorial decompositions of paths then suffice to derive all the basic formulæ.

Arch decomposition: An excursion from and to level 0 consists of a sequence of "arches", each made of either a c_0 or a $a_0\mathcal{H}_{1,1}^{[\geq 1]}b_1$, so that

$$\mathcal{H}_{0,0} = \left(c_0 \cup a_0\mathcal{H}_{1,1}^{[\geq 1]}b_1\right)^{\star}, \tag{27}$$

which relativizes to height $< h$.

Last passages decomposition. Recording the times at which each level $0, \ldots, k$ is last traversed gives

$$\mathcal{H}_{0,k} = \mathcal{H}_{0,0}^{[\geq 0]}a_0\mathcal{H}_{1,1}^{[\geq 1]}a_1 \cdots a_{k-1}\mathcal{H}_{k,k}^{[\geq k]} \tag{28}$$

First passage decomposition. The quantities $H_{k,l}$ with $k \leq l$ are implicitly determined by the first passage through k in a path connecting level 0 to l, so that

$$\mathcal{H}_{0,l} \quad = \quad \mathcal{H}_{0,k-1}^{[<k]}a_{k-1}\mathcal{H}_{k,l} \quad (k \leq l), \tag{29}$$

(A dual decomposition holds when $k \geq l$.)

The basic results express the generating functions in terms of a fundamental continued fraction and its associated convergent polynomials. They involve the "numerator" and "denominator" polynomials, denoted by P_h and Q_h that are defined as solutions to the second order (or "three-term") recurrence equation

$$Y_{h+1} = (1 - c_h)Y_h - a_{h-1}b_hY_{h-1},\ h \geq 1, \tag{30}$$

together with the initial conditions $(P_{-1}, Q_{-1}) = (1, 0)$, $(P_0, Q_0) = (0, 1)$, and with the convention $a_{-1}b_0 = 1$. In other words, setting $C_j = 1 - c_j$ and $A_j = a_{j-1}b_j$, we have:

$$\begin{array}{llll} P_0 = 0, & P_1 = 1, & P_2 = C_2, & P_3 = C_1C_2 - A_2 \\ Q_0 = 0, & Q_1 = C_0, & Q_2 = C_0C_1 - A_1, & Q_3 = C_0C_1C_2 - C_2A_1 - C_0A_2 \end{array}$$

These polynomials are known as continuant polynomials [**122, 190**].

THEOREM V.2 (Path continued fractions [**62**]). (i) *The generating function $H_{0,0}$ of all excursions is represented by the fundamental continued fraction:*

$$H_{0,0} = \cfrac{1}{1 - c_0 - \cfrac{a_0 b_1}{1 - c_1 - \cfrac{a_1 b_2}{1 - c_2 - \cfrac{a_2 b_3}{\ddots}}}}. \tag{31}$$

(ii) *The generating function of ceiled excursion $H_{0,0}^{[<h]}$ is given by a convergent of the fundamental fraction (with P_h, Q_h given by (30):*

$$H_{0,0}^{[<h]} = \cfrac{1}{1 - c_0 - \cfrac{a_0 b_1}{1 - c_1 - \cfrac{a_1 b_2}{\cfrac{\ddots}{1 - c_{h-1}}}}} \tag{32}$$

$$= \frac{P_h}{Q_h}. \tag{33}$$

(iii) *The generating function of floored excursions is given by the truncation of the fundamental fraction:*

$$H_{h,h}^{[\geq h]} = \cfrac{1}{1 - c_h - \cfrac{a_h b_{h+1}}{1 - c_{h+1} - \cfrac{a_{h+1} b_{h+2}}{\ddots}}} \tag{34}$$

$$= \frac{1}{a_{h-1} b_h} \frac{Q_h H_{0,0} - P_h}{Q_{h-1} H_{0,0} - P_{h-1}}, \tag{35}$$

PROOF. Repeated use of the arch decomposition (27) provides a form of $H_{0,0}^{[<h]}$ with nested quasi-inverses $(1-f)^{-1}$ that is the finite fraction representation (32), for instance,

$$\mathcal{H}_{00}^{[<1]} \cong \mathfrak{S}\{c_0\}, \quad \mathcal{H}_{00}^{[<2]} \cong \mathfrak{S}\{c_0 + a_0 \mathfrak{S}\{c_1\} b_1\},$$
$$\mathcal{H}_{00}^{[<3]} \cong \mathfrak{S}\{c_0 + a_0 \mathfrak{S}\{c_1 + a_0 \mathfrak{S}\{c_2\} b_2\} b_1\}.$$

The continued fraction representation for basic paths (namely $H_{0,0}$) is then obtained by letting $h \to \infty$ in (32). Finally, the continued fraction form (34) for ceiled excursions is nothing but the fundamental form (31), when the indices are shifted. The three continued fraction expressions (31), (32), (34) are hence established.

Finding explicit expressions for the fractions $H_{0,0}^{[<h]}$ and $H_{h,h}^{[\geq h]}$ next requires determining the polynomials that appear in the convergents of the basic fraction (31). By definition, the convergent polynomials P_h and Q_h are the numerator and denominator of the fraction $H_{0,0}^{[<h]}$. For the computation of $H_{0,0}^{[<h]}$ and P_h, Q_h, one classically introduces the linear fractional transformations

$$g_j(y) = \frac{1}{1 - c_j - a_j b_{j+1} y},$$

so that

$$H_{0,0}^{[<h]} = g_0 \circ g_1 \circ g_2 \circ \cdots \circ g_{h-1}(0) \text{ and } H_{0,0} = g_0 \circ g_1 \circ g_2 \circ \cdots \ ,. \tag{36}$$

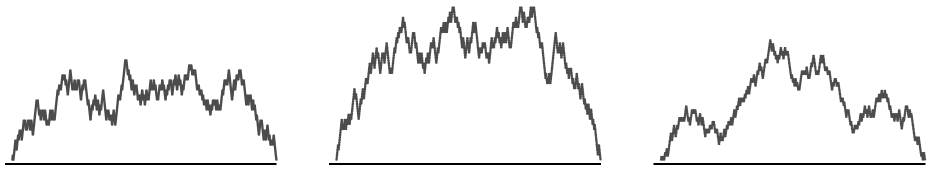


FIGURE 4. Three random Dyck paths of length $2n = 500$ have altitudes resp. 20, 31, 24: the distribution is spread, see Proposition V.4.

Now, linear fractional transformations are representable by 2×2-matrices

$$\frac{ay+b}{cy+d} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{37}$$

in such a way that composition corresponds to matrix product. By induction on the compositions that build up $H_{0,0}^{[<h]}$, there follows the equality

$$g_0 \circ g_1 \circ g_2 \circ \cdots \circ g_{h-1}(y) = \frac{P_h - P_{h-1}a_{h-1}b_h y}{Q_h - Q_{h-1}a_{h-1}b_h y}, \tag{38}$$

where P_h and Q_h are seen to satisfy the recurrence (30). Setting $y = 0$ in (38) proves (33).

Finally, $H_{h,h}^{[\geq h]}$ is determined implicitly as the root y of the equation $g_0 \circ \cdots \circ g_{h-1}(y) = H_{0,0}$, an equation that, when solved using (38), yields the form (35). □

A large number of generating functions can be derived by similar techniques. We refer to the article [**62**], where this theory was first systematically developed and to the exposition given in [**96**, Chapter 5]. Our presentation here draws upon [**72**] where the theory was put to further use in order to develop a formal algebraic theory of the general birth-and-death process in continuous time.

▷ **5.** *Transitions and crossings.* The lattice paths $\mathcal{H}_{0,l}$ corresponding to the transitions from altitude 0 to l and $\mathcal{H}_{k,0}$ (from k to 0) have OGFs

$$H_{0,l} = \frac{1}{\beta_l}\left(Q_l H_{0,0} - P_l\right), \qquad H_{k,0} = \frac{1}{\alpha_k}(Q_k H_{0,0} - P_k).$$

The crossings $\mathcal{H}_{0,h-1}^{[<h]}$ and $\mathcal{H}_{h-1,0}^{[<h]}$ have OGFs,

$$H_{0,h-1}^{[<h]} = \frac{\alpha_{h-1}}{Q_h}, \qquad H_{h-1,0}^{[<h]} = \frac{\beta_{h-1}}{Q_h},$$

obtained from the last passages decomposition. (Abbreviations used are: $\alpha_m = a_0 \cdots a_{m-1}$, $\beta_m = b_1 \cdots b_m$.) This gives combinatorial interpretations for fractions of the form $1/Q$ and results from the basic decompositions combined with Theorem V.2; see [**62, 72**] for details. ◁

We examine next a few specializations of the general formulæ provided by Theorem V.2.

EXAMPLE 6. *Height of standard lattice paths.* In order to count lattice paths, it suffices to effect one of the substitutions,

$$\sigma_M : \ a_j \mapsto z, \ b_j \mapsto z, c_j \mapsto z; \qquad \sigma_D : \ a_j \mapsto z, \ b_j \mapsto z, c_j \mapsto 0.$$

In the former case, all three step types are taken into account, giving rise to so-called "Motzkin paths"; in the latter case level steps are disallowed, and one obtains so-called "Dyck paths".

We henceforth restrict attention to the case of Dyck paths. See Figure 4 for three simulations suggesting that the distribution of height is somewhat spread. The continued fraction expressing $H_{0,0}$ is in this case purely periodic, and it represents a quadratic function:

$$H_{0,0}(z) = \cfrac{1}{1 - \cfrac{z^2}{1 - \cfrac{z^2}{1 - \ddots}}} = \frac{1}{2z^2}\left(1 - \sqrt{1-4z^2}\right),$$

since $H_{0,0}$ satisfies $y = (1 - z^2 y)^{-1}$. The families of polynomials P_h, Q_h are in this case determined by a recurrence with constant coefficients and they coincide, up to a shift of indices. Define classically the Fibonacci polynomials by the recurrence

$$F_{h+2}(z) = F_{h+1}(z) - zF_h(z), \quad F_0(z) = 0, \quad F_1(z) = 1. \tag{39}$$

One finds $Q_h = F_{h+1}(z^2)$ and $P_h = F_h(z^2)$. (The Fibonacci polynomials are essentially reciprocals of Chebyshev polynomials.) By Theorem V.2, the GF of paths of height $< h$ is then

$$H_{00}^{[<h]}(z) = \frac{F_h(z^2)}{F_{h+1}(z^2)}.$$

(We get more and, for instance, the number of ways of crossing a strip of width $h-1$ is $H_{0,h-1}^{[<h]}(z) = z^{h-1}/F_{h+1}(z^2)$.) Note that the polynomials have an explicit form,

$$F_h(z) = \sum_{k=0}^{\lfloor (h-1)/2 \rfloor} \binom{h-1-k}{k} (-z)^k,$$

as follows from the generating function expression: $\sum_h F_h(z) y^h = y/(1 - y + zy^2)$.

The equivalence between Dyck paths and (general) plane tree traversals discussed in Chapter I implies that trees of height at most h and size $n+1$ are equinumerous with Dyck paths of length $2n$ and height at most h. Set for convenience

$$G^{[h]}(z) = zH_{00}^{[<h+1]}(z^{1/2}) = z\frac{F_{h+1}(z)}{F_{h+2}(z)},$$

which is precisely the OGF of general plane trees having height $\le h$. (This is otherwise in agreement with the continued fraction form obtained directly in Chapter III). It is possible to go much further as first shown by De Bruijn, Knuth, and Rice in a beautiful paper [**43**], which also constitutes the historic application of Mellin transforms in analytic combinatorics. (We refer to this paper for and historical context and references.)

First, solving the linear recurrence (39) with z treated as a parameter yields the alternative closed form expression

$$F_h(z) = \frac{G^h - \overline{G}^h}{G - \overline{G}}, \qquad G = \frac{1 - \sqrt{1-4z}}{2}, \quad \overline{G} = \frac{1 + \sqrt{1-4z}}{2}. \tag{40}$$

There, $G(z)$ is the OGF of all trees, and an equivalent form of $G^{[h]}$ is provided by

$$G - G^{[h-2]} = \sqrt{1-4z}\,\frac{u^h(z)}{1 - u^h}, \quad \text{where} \quad u = \frac{1 - \sqrt{1-4z}}{1 + \sqrt{1-4z}} = \frac{G^2}{z}, \tag{41}$$

as is easily verified. Thus $G^{[h]}$ can be expressed in terms of $G(z)$ and z:

$$G - G^{[h-2]} = \sqrt{1-4z} \sum_{j \ge 1} z^{-jh} G(z)^{2jh}.$$

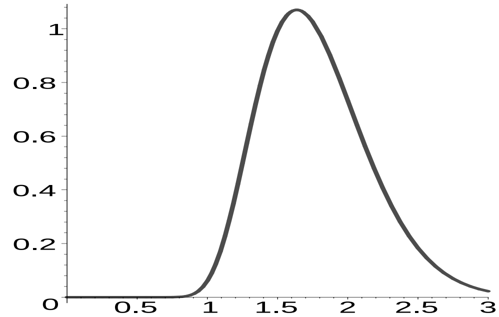


FIGURE 5. The limit density of the distribution of height $-\Theta'(x)$.

The Lagrange-Bürmann inversion theorem then gives after a simple calculation

$$G_{n+1} - G_{n+1}^{[h-2]} = \sum_{j\geq 1} \Delta^2 \binom{2n}{n-jh}, \tag{42}$$

where

$$\Delta^2 \binom{2n}{n-m} := \binom{2n}{n+1-m} - 2\binom{2n}{n-m} + \binom{2n}{n-1-m}.$$

Consequently, the number of trees of height $\geq h-1$ admits of closed form: it is a "sampled" sum by steps of h of the $2n$th line of Pascal's triangle (upon taking second order differences).

The relation (42) leads easily to the asymptotic distribution of height in random trees of size n. Stirling's formula yields the Gaussian approximation of binomial numbers: for $k = o(n^{3/4})$ and with $w = k/\sqrt{n}$, one finds

$$\frac{\binom{2n}{n-k}}{\binom{2n}{n}} \sim e^{-w^2}\left(1 - \frac{w^4+3w^2}{6n} + \frac{5w^8+6w^6-45w^4-60}{360n^2} + \cdots\right). \tag{43}$$

The use of the Gaussian approximation (43) inside the exact formula (42) then implies: *The probability that a tree of size $n+1$ has height at least $h-1$ satisfies uniformly for $h \in [\alpha\sqrt{n}, \beta\sqrt{n}]$ (with $0 < \alpha < \beta < \infty$) the estimate*

$$\frac{G_{n+1} - G_{n+1}^{[h-2]}}{G_{n+1}} = \Theta\left(\frac{h}{\sqrt{n}}\right) + O\left(\frac{1}{n}\right), \qquad \Theta(x) := \sum_{j\geq 1} e^{-j^2x^2}(4j^2x^2 - 2). \tag{44}$$

The function $\Theta(x)$ is a "theta function" which classically arises in the theory of elliptic functions [**191**]. Since binomial coefficients decay fast away from the center, simple bounds also show that the probability of height to be at least $n^{1/2+\epsilon}$ decays like $\exp(-n^{2\epsilon})$, hence is exponentially small. Note also that the probability distribution of height H itself admits of an exact expression obtained by differencing (42), which is reflected asymptotically by differentiation of the estimate of (44):

(45)

$$\mathbb{P}_{\mathcal{G}_{n+1}}\left[H = \lfloor x\sqrt{n}\rfloor\right] = -\frac{1}{\sqrt{n}}\Theta'(x) + O\left(\frac{1}{n}\right), \qquad \Theta'(x) := \sum_{j\geq 1} e^{-j^2x^2}(12j^2x - 8j^4x^3).$$

The forms (44) and (45) also give access to moments of the distribution of height. We find

$$\mathbb{E}_{\mathcal{G}_{n+1}}\left[H^r\right] \sim \frac{1}{\sqrt{n}} S_r\left(\frac{1}{\sqrt{n}}\right), \qquad \text{where} \quad S_r(y) := -\sum_{h\geq 1} h^r \Theta'(hy).$$

The quantity $y^{r+1}S_r(y)$ is a Riemann sum relative to the function $-x^r\Theta'(x)$, and the step $y = n^{-1/2}$ decreases to 0 as $n \to \infty$. Approximating the sum by the integral, one gets:

$$\mathbb{E}_{\mathcal{G}_{n+1}}\left[H^r\right] \sim n^{r/2}\mu_r \qquad \text{where} \quad \mu_r := -\int_0^\infty x^r \Theta'(x)\,dx.$$

The integral giving μ_r is a Mellin transform in disguise (set $s = r+1$) to which the treatment of harmonic sums applies. We then get upon replacing $n+1$ to n:

PROPOSITION V.4. *The expected height of a random plane rooted tree comprising n nodes is*

$$\sqrt{\pi n} - \frac{1}{2} + o(1).$$

More generally, the moment of order r of height is asymptotic to

$$\mu_r n^{r/2} \qquad \textit{where} \quad \mu_r = r(r-1)\Gamma(r/2)\zeta(r).$$

The random variable $H/\sqrt{n}$ obeys asymptotically a Theta distribution, in the sense of both the "central" estimate (44) and the "local" estimate (45). The same asymptotic estimates hold for height of Dyck paths having length $2n$.

The improved estimate of the mean is from [**43**]. The general moment forms are in fact valid for any real r (not just integers). An alternative formula for the Theta function appears in the Note below. Figure 5 plots the limit density $-\Theta'(x)$. □

▷ **6.** *Height, Fibonacci and Chebyshev polynomials.* The reciprocal polynomials $U_h(z) = z^h F_h(1/z)$ satisfy $U_h(\cos(\theta)) = \sin((h+1)\theta)/\sin(\theta)$ as is readily verified from the recurrence (39) and elementary trigonometry. Thus, the roots of $F_h(z)$ are $(4\cos^2 j\pi/h)^{-1}$ and the partial fraction expansion of $G^{[h]}(z)$ can be worked out explicitly [**43**]. There results

$$G_{n+1}^{[h-2]} = \frac{4^{n+1}}{h} \sum_{1\leq j\leq h/2} \sin^2\frac{j\pi}{h}\cos^{2n}\frac{j\pi}{h}, \tag{46}$$

which provides in particular an asymptotic form for any fixed h. (This formula can also be found directly from the sampled sum (42) by multisection of series.) Asymptotic analysis of this last expression when $h = x\sqrt{n}$ yields the alternative expression

$$\lim_{n\to\infty} \mathbb{P}_{\mathcal{G}_{n+1}}\left[H \leq x\sqrt{n}\right] = 4\pi^{5/2}x^{-3}\sum_{j\geq 0} j^2 e^{-j^2\pi^2/x^2} \quad (\equiv 1 - \Theta(x)),$$

which reflects a classical transformation formula of theta functions. See the study by Biane, Pitman, and Yor [**25**] for fascinating connections between this formula, Brownian motion, and the functional equation of the Riemann zeta function. ◁

▷ **7.** *Motzkin paths.* The OGF of Motzkin paths of height $< h$ is

$$\frac{1}{1-z} \cdot {}^{\mathrm{D}}H_{0,0}^{[h]}\left(\frac{z}{1-z}\right),$$

where ${}^{\mathrm{D}}H_{0,0}^{[h]}$ above refers to Dyck paths. Therefore, such paths of length n can be enumerated exactly by formulæderived from (42) and 46). In particular, the expected height is $\sim \sqrt{\pi n/3}$. ◁

$\triangleright$ **8.** *Height in simple varieties of trees.* Consider a simple variety of trees corresponding to the GF equation $Y(z) = z\phi(Y(z))$ (see Chapter III) and values of n such that there exists a tree of size n. Assume that there exists a positive τ strictly within the disc of convergence of ϕ such that $\tau\phi'(\tau) - \phi(\tau) = 0$. Then, the rth moment of height ($\overline{H}$) is asymptotically $\xi^{r/2} r(r-1)\Gamma(r/2)\zeta(r)n^{r/2}$. The normalized quantity $H = H/\xi$ obeys asymptotically a Theta distribution in the sense of both the central estimate (44) and the local estimate (45). [This is from [**76**] and [**68**] respectively.] For instance, $\xi = 2$ for plane binary trees and $\xi = \sqrt{2}$ for Cayley trees. $\triangleleft$

EXAMPLE 7. *Area under Dyck path and coin fountains.* Consider the case of Dyck path and the parameter equal to the area below the path. Area under a lattice path can be defined as the sum of the indices (i.e., the starting altitudes) of all the variables that enter the standard encoding of the path. Thus, the BGF $D(z,q)$ of Dyck path with z marking half-length and q marking area is obtained by the substitution

$$a_j \mapsto q^j z, \quad b_j \mapsto q^j, \quad c_j \mapsto 0$$

inside the fundamental continued fraction (31). It proves convenient to operate with the continued fraction

$$F(z,q) = \cfrac{1}{1 - \cfrac{zq}{1 - \cfrac{zq^2}{\ddots}}}, \tag{47}$$

so that $D(z,q) = F(q^{-1}z, q^2)$. Since F and D satisfy difference equations, for instance,

$$F(z,q) = \frac{1}{1 - zqF(qz,q)}, \tag{48}$$

moments of area can be determined by differentiating and setting $q = 1$ (see Chapter III for such a direct approach).

A general trick from q–calculus is effective to derive an alternative expression of F. Attempt to express the continued fraction F of (47) as a quotient $F(z,q) = A(z)/B(z)$. Then, the relation (48) implies

$$\frac{A(z)}{B(z)} = \frac{1}{1 - qz\frac{A(qz)}{B(qz)}}, \quad \text{hence } A(z) = B(qz),\ B(z) = B(qz) - qzB(q^2 z),$$

where q is treated as a parameter. The difference equation satisfied by $B(z)$ is readily solved by indeterminate coefficients: this classical technique was introduced in the theory of integer partitions by Euler. With $B(z) = \sum b_n z^n$, the coefficients satisfy the recurrence

$$b_0 = 1, \quad b_n = q^n b_n - q^{2n-1} b_{n-1}.$$

This is a first order recurrence on b_n that unwinds to give

$$b_n = (-1)^n \frac{q^{n^2}}{(1-q)(1-q^2)\cdots(1-q^n)}.$$

In other words, introducing the "q-exponential function",

$$E(z,q) = \sum_{n=0}^{\infty} \frac{(-z)^n q^{n^2}}{(q)_n}, \quad \text{where} \quad (q)_n = (1-q)(1-q^2)\cdots(1-q^n), \tag{49}$$

one finds

$$F(z,q) = \frac{E(qz,q)}{E(z,q)}. \tag{50}$$

Given the importance of the functions under discussion in various branches of mathematics, we cannot resist a quick digression. The name of the q-exponential comes form the obvious property that $E(z(q-1),q)$ reduces to e^{-z} as $q \to 1^-$. The explicit form (49) constitutes in fact the "easy half" of the proof of the celebrated Rogers-Ramanujan identities, namely,

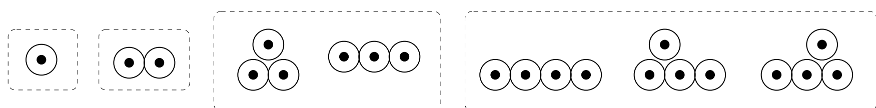
$$\begin{array}{lllll} E(-1,q) = \displaystyle\sum_{n=0}^{\infty} \frac{q^{n^2}}{(q)_n} & = & \displaystyle\prod_{n=0}^{\infty} (1-q^{5n+1})^{-1}(1-q^{5n+4})^{-1} \\ E(-q,q) = \displaystyle\sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(q)_n} & = & \displaystyle\prod_{n=0}^{\infty} (1-q^{5n+2})^{-1}(1-q^{5n+3})^{-1}, \end{array} \tag{51}$$

that relate the q-exponential to modular forms. See Andrews' book [**5**, Ch. 7] for context.

Here is finally a cute application of these ideas to asymptotic enumeration. Odlyzko and Wilf define in [**149, 147**] an (n,m) coin fountain as an arrangement of n coins in rows in such a way that there are m coins in the bottom row, and that each coin in a higher row touches exactly two coins in the next lower row. Let $C_{n,m}$ be the number of (n,m) fountains and $C(q,z)$ be the corresponding BGF with q marking n and z marking m. Set $C(q) = C(q,1)$. The question is to determine the total number of coin fountains of area n, $[q^n]C(q)$. The series starts as (this is *EIS* **A005169**)

$$C(q) = 1 + q + q^2 + 2q^3 + 3q^4 + 5q^5 + 9q^6 + 15q^7 + 26q^8 + \cdots,$$

as results from inspection of the first few cases.



The function $C(q)$ is *a priori* meromorphic in $|q| < 1$. From the bijection with Dyck paths and area, one finds

$$C(q) = \cfrac{1}{1 - \cfrac{q}{1 - \cfrac{q^2}{1 - \cfrac{q^3}{\ddots}}}}.$$

The identity (50) implies

$$C(q) = \frac{E(q,q)}{E(1,q)}.$$

An exponential lower bound of the form 1.6^n holds on $[q^n]C(q)$, since $(1-q)/(1-q-q^2)$ is dominated by $C(q)$ for $q > 0$. At the same time, the number $[q^n]C(q)$ is majorized by the number of compositions, which is 2^{n-1}. Thus, the radius of convergence of $C(q)$ has to lie somewhere between 0.5 and $0.61803\ldots$. It is then easy to check by numerical analysis the existence of a simple zero of the denominator, $E(-1,q)$, near $\rho \doteq 0.57614$. Routine computations based on Rouché's theorem then makes it possible to verify formally that ρ is the only simple pole in $|q| < 3/5$ (the process is detailed in [**147**]). Thus, singularity analysis of meromorphic functions applies:

Objects	*Weights* $(\alpha_j, \beta_j \gamma_j)$	*Counting*	*Orth. pol.*
Simple paths	$1, 1, 0$	Catalan #	Chebyshev
Permutations	$j+1, j, 2j+1$	Factorial #	Laguerre
Alternating perm.	$j+1, j, 0$	Secant #	Meixner
Involutions	$1, j, 0$	Odd factorial #	Hermite
Set partition	$1, j, j+1$	Bell #	Poisson-Charlier
Nonoverlap. set part.	$1, 1, j+1$	Bessel #	Lommel

FIGURE 6. Some special families of combinatorial objects together with corresponding weights, moments, and orthogonal polynomials.

PROPOSITION V.5. *The number of coin fountains made of n coins satisfies asymptotically*

$$[q^n]C(q) = cA^n + O((5/3)^n), \quad c \doteq 0.31236, \quad A = \rho^{-1} \doteq 1.73566.$$

This example illustrates the power of modelling by continued fractions as well as the smooth articulation with meromorphic function asymptotics. □

The systematic theory of lattice path enumerations and continued fractions was developed initially because of the need to count weighted lattice paths, notably in the context of the analysis of dynamic data structures in computer science [**67**]. In this framework, a system of multiplicative weights $\alpha_j, \beta_j, \gamma_j$ is associated with the steps a_j, b_j, c_j, each weight being an integer that represents a number of "possibilities" for the corresponding step type. A system of weighted lattice paths has counting generating functions given by an easy specialization of the corresponding multivariate expressions we have just developed, namely,

$$a_j \mapsto \alpha_j z, \quad b_j \mapsto \beta_j z, \quad c_j \mapsto \gamma_j z, \tag{52}$$

where z marks the length of paths. One can then sometimes solve an enumeration problem expressible in this way by reverse-engineering the known collection of continued fractions as found in a reference book like Wall's treatise [**190**]. Next, for general reasons, the polynomials P, Q are always elementary variants of a family of orthogonal polynomials that is determined by the weights [**37, 62, 178**]. When the multiplicities have enough structural regularity, the weighted lattice paths are likely to correspond to classical combinatorial objects and to classical families of orthogonal polynomials; see [**62, 67, 94, 96**] and Figure 6 for an outline. We illustrate this by a simple example due to Lagarias, Odlyzko, and Zagier [**127**].

EXAMPLE 8. *Interconnection networks and involutions.* The problem considered here was introduced by Lagarias, Odlyzko, and Zagier in [**127**]: *There are $2n$ points on a line, with n point-to-point connections between pairs of points. What is the probable behaviour of the* width *of such an interconnection network?* Imagine the points to be $1, \ldots, 2n$, the connections as circular arcs between points, and let a vertical line sweep from left to right; width is defined as the maximum number of edges encountered by such a line. One may freely imagine a tunnel of fixed capacity (this corresponds to the width) inside which wires can be placed to connect points pairwise. See Figure 7.

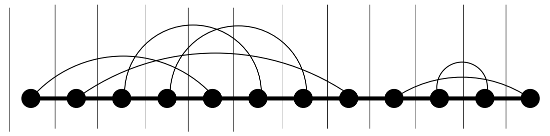


FIGURE 7. An interconnection network on $2n = 12$ points.

Let $\mathcal{I}_{2n}$ be the class of all interconnection networks on $2n$ points, which is precisely the collection of ways of grouping $2n$ elements into n pairs, or, equivalently, the class of all involutions (i.e., permutations with cycles of length 2 only). The number I_{2n} equals the "odd factorial",

$$I_{2n} = 1 \cdot 3 \cdot 5 \cdots (2n-1),$$

whose EGF is $e^{z^2/2}$ (see Chapter 2). The problem calls for determining the quantity $I_{2n}^{[h]}$ that is the number of networks corresponding to a width $\le h$.

The relation to lattice paths is as follows. First, when sweeping a vertical line across a network, define an active arc at an abscissa as one that straddles that abscissa. Then build the sequence of active arcs counts at half-integer positions $\frac{1}{2}, \frac{3}{2}, \ldots, 2n - \frac{1}{2}, 2n + \frac{1}{2}$. This constitutes a sequence of integers where each member is ± 1 the previous one, that is, a lattice path without level steps. In other words, there is an ascent in the lattice path for each element that is smaller in its cycle and a descent otherwise. One may view ascents as associated to situations where a node "opens" a new cycle, while descents correspond to "closing" a cycle.

Involutions are much more numerous than lattice paths, so that the correspondence from involutions to lattice paths is many-to-one. However, one can easily enrich lattice paths, so that the enriched objects are in one-to-one correspondence with involutions. Consider again a scanning position at a half-integer where the vertical line crosses ℓ (active) arcs. If the next node is of the closing type, there are ℓ possibilities to choose from. If the next node is of the opening type, then there is only one possibility, namely, to start a new cycle. A complete encoding of a network is obtained by recording additionally the sequence of the n possible choices corresponding to descents in the lattice path (some canonical order is fixed, for instance, oldest first). If we write these choices as superscripts, this means that the set of all enriched encodings of networks is obtained from the set of standard lattice path encodings by effecting the substitutions

$$b_j \mapsto \sum_{k=1}^{j} b_j^{(k)}.$$

The OGF of all involutions is obtained from the generic continued fraction of Theorem V.2 by the substitution

$$a_j \mapsto z, \quad b_j \mapsto j\, z,$$

where z records the number of steps in the enriched lattice path, or equivalently, the number of nodes in the network. In other words, we have obtained combinatorially a formal

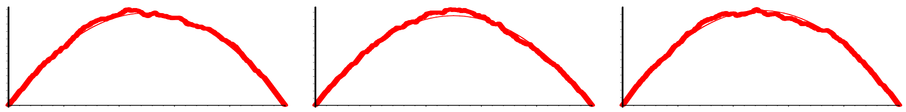


FIGURE 8. Three simulations of random networks with $2n = 1000$ illustrate the tendency of the profile to conform to a parabola with height close to $n/2 = 250$.

continued fraction representation,

$$\sum_{n=0}^{\infty}(1\cdot 3\cdots(2n-1))z^{2n} = \cfrac{1}{1-\cfrac{1\cdot z^2}{1-\cfrac{2\cdot z^2}{1-\cfrac{3\cdot z^2}{\ddots}}}},$$

which was originally discovered by Gauß [**190**]. Theorem V.2 then gives immediately the OGF of involutions of width at most h as a quotient of polynomials. Define

$$I^{[h]}(z) := \sum_{n\geq 0} I_{2n}^{[h]} z^{2n}.$$

One has

$$I^{[h]}(z) = \cfrac{1}{1-\cfrac{1\cdot z^2}{1-\cfrac{2\cdot z^2}{\cfrac{\ddots}{1-h\cdot z^2}}}} = \frac{P_h(z)}{Q_h(z)}$$

where P_h and Q_h satisfy the recurrence

$$Y_{h+1} = Y_h - hz^2 Y_{h-1}.$$

The polynomials are readily determined by their generating functions that satisfies a first-order linear differential equation reflecting the recurrence. In this way, the denominator polynomials are identified to be reciprocals of the Hermite polynomials,

$$Q_h(z) = (z/2)^h H_h(\frac{1}{2z}),$$

themselves defined classically [**2**, Ch. 22] as orthogonal with respect to the measure $e^{-x^2}\,dx$ on $(-\infty,\infty)$ and expressible via

$$H_m(x) = \sum_{m=0}^{\lfloor m/2\rfloor} \frac{(-1)^j m!}{j!(m-2j)!}(2x)^{m-2j}, \qquad \sum_{m\geq 0} H_m(x)\frac{t^m}{m!} = e^{xt-t^2}.$$

In particular, one finds

$$I^{[0]} = 1, \quad I^{[1]} = \frac{1}{1-z^2}, \quad I^{[2]} = \frac{1-2z^2}{1-3z^2}, \quad I^{[3]} = \frac{1-5z^2}{1-6z^2+3z^4}, \quad \&c.$$

The interesting analysis of the dominant poles of the rational GF's, for any fixed h, is discussed in the paper [**127**]. Furthermore, simulations strongly suggest that the width of a random interconnection network on $2n$ nodes is tightly concentrated around $n/2$; see Figure 8. Louchard [**134**] succeeded in proving this fact and a good deal more: With high probability, the profile (the profile is defined here as the number of active arcs as time evolves) of a random network conforms asymptotically to a deterministic parabola $2nx(1-x/(2*n))$ to which are superimposed random fluctuations of amplitude only $O(\sqrt{n})$ well-characterized by a Gaussian process. In particular, *the width of a random network of $2n$ nodes converges in probability to $\frac{n}{2}$.* □

V. 3. The supercritical sequence and its applications

We have seen earlier in this section that surjections and alignments with EGFs

$$\frac{1}{2-\exp(z)}, \qquad \frac{1}{1-\log(1-z)^{-1}}$$

have coefficients that satisfy simple asymptotic estimates of the form $C \cdot A^n$. A similar property holds for integer compositions, where there is even an exact counting formula, namely, 2^{n-1}. The common feature of these examples is that that they all involve a sequence construction in their specification and correspond to the schema $\mathcal{F} = \mathfrak{S}(\mathcal{G})$, in either the labelled or the unlabelled case.

We thus consider a sequence construction $\mathcal{F} = \mathfrak{S}(\mathcal{G})$, with the associated GFs (either ordinary or exponential) satisfying the usual relation

$$F(z) = \frac{1}{1-G(z)},$$

and $G(0)=0$ for well-foundedness. We shall write $f_n = [z^n]F(z)$ and $g_n = [z^n]G(z)$. We also restrict attention to the case where the radius of convergence of G is nonzero, in which case, the radius of convergence of F is also nonzero by virtue of closure properties of analytic functions. We set:

DEFINITION V.2. *Let F, G be GFs with nonnegative coefficients that are analytic at 0, with $G(0)=0$. The schema $F(z) = (1-G(z))^{-1}$ is said to be* supercritical *if $G(\rho) > 1$, where $\rho = \rho_G$ is the radius of convergence of G.*

Note that $G(\rho)$ is well defined as the limit $\lim_{x\to\rho^-} G(x)$ since $G(x)$ increases along the positive real axis. (The value $G(\rho)$ corresponds to what has been denoted earlier by τ_G when discussing "signatures" in Section IV. 3.3.) We assume that $G(z)$ is *unperiodic* in the sense that there does not exist an integer $d \geq 2$ such that $G(z) = h(z^d)$ for some h analytic at 0. (This normalization is merely a convenience that entails no loss of generality.) One has

THEOREM V.3 (Supercritical sequence asymptotics). *Let the schema $F = (1-G)^{-1}$ be supercritical and assume that G is unperiodic. Then, one has*

$$[z^n]F(z) = \frac{1}{\sigma G'(\sigma)} \cdot \sigma^{-n}\left(1 + O(A^n)\right),$$

where σ is the root in $(0, \rho_G)$ of $G(\sigma) = 1$, and A is a number less than 1. The number X of $\mathcal{G}$–components in a random $\mathcal{C}$–structure of size n has mean and variance satisfying

$$\begin{aligned} \mathbb{E}_n(X) &= \frac{1}{\rho G'(\rho)} \cdot (n+1) - 1 + \frac{G''(\rho)}{G'(\rho)^2} + O(A^n) \\ \mathbb{V}_n(X) &= \frac{\rho G''(\rho) + G'(\rho) - \rho G'(\rho)^2}{\rho^2 G'(\rho)^3} \cdot n + O(1). \end{aligned}$$

In particular, the distribution is concentrated.

PROOF [**84, 172**]. The basic observation is that G increases continuously from $G(0) = 0$ to $G(\rho_G) = \tau_G$ (with $\tau_G > 1$ by assumption) when x increases from 0 to ρ_G. Therefore, the positive number σ, which satisfies $G(\sigma) = 1$ is well defined. Then, F is analytic at all points of the interval $(0, \sigma)$. The function G being analytic at σ, satisfies, in a neighbourhood of σ

$$G(z) = 1 + G'(\sigma)(z - \sigma) + \frac{1}{2!}G''(\sigma)(z - \sigma)^2 + \cdots .$$

so that $F(z)$ has a pole at $z = \sigma$; also, this pole is simple since $G'(\rho) > 0$. Pringsheim's theorem then implies that the radius of convergence of F must coincide with σ.

There remains to show that $F(z)$ is meromorphic in a disc of some radius $R > \sigma$ with the point σ as the only singularity inside the disc. This results from the assumption that G is unperiodic. In effect, one has $G(\sigma e^{i\theta}) \leq 1$ for all θ by the triangular inequality. It suffices to verify that $G(\sigma e^{i\theta}) \neq 1$ for $\theta \in [-\pi, \pi] \setminus \{0\}$ to ensure that F is analytic at points of the circle $|z| = \sigma$, with the sole exception of σ. A contrario, $G(\sigma e^{i\theta}) = 1$ would imply, by the converse of the triangle inequality that

$$g_n \sigma^n e^{in\theta} = g_n \sigma^n,$$

for all values of n such that $g_n \neq 0$. This in turn is only possible if there is a root of unity, $\omega = e^{2i\pi/d}$, such that $\omega^n = 1$ whenever $g_n \neq 1$. This last fact is itself incompatible with the assumption that $G(z)$ is unperiodic.

In summary, $F(z)$ has a simple pole at $z = \sigma$ and is otherwise analytic at all points of $|z| = \sigma$. Thus, by compactness, there exists a disc of radius $R > \sigma$ in which F is analytic except for a unique pole at σ. Take r such that $\sigma < r < R$ and apply the main theorem of meromorphic function asymptotics to deduce the stated formula with $A = \sigma/r$.

Consider next the number of $\mathcal{G}$-components in a random $\mathcal{F}$ structure of size n. Bivariate generating functions give access to the expectation of this random variable:

$$\begin{aligned} \mathbb{E}_n(X) &= \frac{1}{f_n}[z^n] \left. \frac{\partial}{\partial u} \frac{1}{1 - uG(z)} \right|_{u=1} \\ &= \frac{1}{f_n}[z^n] \frac{G(z)}{(1 - G(z))^2}. \end{aligned}$$

The problem is now reduced to extracting coefficients in a univariate generating function with a double pole at $z = \rho$, and it suffices to expand the GF locally at ρ. The variance calculation is similar though it involves a triple pole. □

When a sequence construction is supercritical, the number of components is in the mean $\asymp n$ while its standard deviation is $\asymp \sqrt{n}$. Thus, the distribution is concentrated (see Chapter III). In fact, there results from a general theorem of Bender [**16**] that the distribution of the number of components is asymptotically Gaussian; see later chapters for details.

Direct cases of application to combinatorial generating functions are

$$a_1(z) = \frac{z}{1 - z},\ a_2(z) = e^z - 1,\ a_3(z) = \log(1 - z)^{-1},$$

corresponding respectively to integer compositions (OGF), surjections (EGF), and alignments. Thus:

- The expected number of summands in a random composition of the integer n is $\sim \frac{n+1}{2}$, with variance $\sim \frac{n}{4}$.
- The expected cardinality of the range of a random surjection whose domain has cardinality n is asymptotic to βn with $\beta = 1/(2 \log 2)$;
- The expected number of components in a random alignment of size n is asymptotic to $n/(e - 1)$.

10	16	1 ***5***
20	732	73 ***4***
30	36039	360 ***57***
40	1772207	17722 ***61***
50	87109263	871092 ***48***
60	4281550047	42815 ***49331***
70	210444532770	21044453 ***0095***
80	10343662267187	1034366226 ***5182***
90	508406414757253	5084064147 ***81706***
100	24988932929490838	24988932929 ***612479***

FIGURE 9. The pyramid relative to compositions into prime summands for $n = 10\,..\,100$: (left: exact values; right: asymptotic formula rounded).

EXAMPLE 9. *Compositions with restricted summands, compositions into primes.* Unrestricted integer compositions are well understood as regards enumeration: their number is exactly $C_n = 2^{n-1}$, their OGF is $C(z) = (1-z)/(1-z)$, and compositions with k summands are enumerated by binomial coefficients. Such simple exact formulæ disappear when restricted compositions are considered, but, as we now show, asymptotics is much more robust to changes in specifications.

Let S be a subset of the integers $\mathbb{Z}_{\geq 1}$ such that $\gcd(S) = 1$, i.e., not all members of S are multiples of a common divisor $d \geq 2$. In order to avoid trivialities, we also assume that $S \neq \{1\}$. The class $\mathcal{C}^S$ of compositions with summands constrained to the set S then satisfies:

$$\textit{Specification}: \quad \mathcal{C}^S = \mathfrak{S}(\mathfrak{S}_S(\mathcal{Z}));$$
$$\textit{OGF}: \quad D(z) = \frac{1}{1-S(z)}, \quad S(z) = \sum_{s \in S} z^s.$$

By assumption, $S(z)$ is unperiodic, so that Theorem V.3 applies directly. There is a well-defined number σ such that

$$S(\sigma) = 1, \qquad 0 < \sigma < 1,$$

and the number of S–restricted compositions satisfies

$$C_n^S := [z^n]C^S(z) = \frac{1}{\sigma S'(\sigma)} \cdot \sigma^{-n}\left(1 + O(A^n)\right). \tag{53}$$

Amongst the already discussed cases, $S = \{1, 2\}$ gives rise to Fibonacci numbers and, more generally, $S = \{1, \ldots, r\}$ corresponds to partitions with summands at most r. In this case, the OGF,

$$C^{\{1,\ldots,r\}}(z) = \frac{1}{1 - z\frac{1-z^r}{1-z}} = \frac{1-z}{1-2z+z^{r+1}}$$

is a simple variant of the OGF associated to longest runs in strings. The treatment of the latter can be copied almost verbatim to the effect that the largest component in a random composition of n is found to be $\lg n + O(1)$, both on average and with high probability.

Here is a surprising application of the general theory. Consider the case where S is taken to be the set of prime numbers, $\mathrm{Prime} = \{2, 3, 5, 7, 11, \ldots\}$, thereby defining the class of *compositions into prime summands*. The sequence starts as

$$1, 0, 1, 1, 1, 3, 2, 6, 6, 10, 16, 20, 35, 46, 72, 105,$$

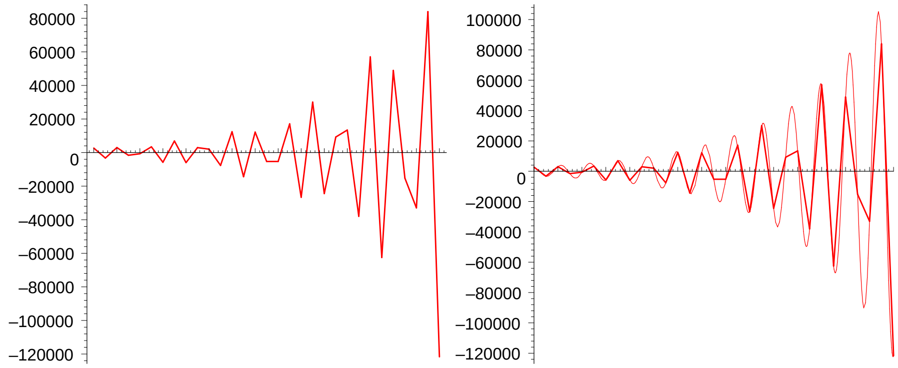


FIGURE 10. Errors in the approximation of the number of compositions into primes for $n = 70\,..\,100$: left, the values of $C_n^{\text{Prime}} - g(n)$; right, the correction $g_2(n)$ arising from the next two poles, which are complex conjugate and the continuous extrapolation of this approximation.

corresponding to the OGF $1+z^2+\cdots$, and is *EIS* **A023360** in Sloane's encyclopedia. The formula (53) applies to provide the asymptotic form of the number of such compositions. It is also well worth noting that the constants appearing in (53) are easily determined to great numerical precision, as we now explain.

By (53) and the preceding equation, the dominant singularity of the OGF of compositions into prime is the positive root $\sigma < 1$ of the characteristic equation

$$S(z) \equiv \sum_{p \ \text{Prime}} z^p = 1.$$

Fix a threshold value m_0 (for instance $m_0 = 10$ or 100) and introduce the two series

$$S^-(z) := \sum_{s\in S,\ s<m_0} z^s, \quad S^+(z) := \left(\sum_{s\in S,\ s<m_0} z^s\right) + \frac{z^{m_0}}{1-z}.$$

Clearly, for $x \in (0,1)$, one has $S^-(x) < S(x) < S^+(x)$. Define two constants σ^-, σ^+ by the conditions

$$S^-(\sigma^-) = 1, \quad S^+(\sigma^+) = 1, \qquad 0 < \sigma^-, \sigma^+ < 1.$$

These constants are algebraic numbers that are accessible to computation. At the same time, they satisfy $\sigma^+ < \sigma < \sigma^-$. As the order of truncation, m_0, increases, the values of σ^+, σ^- are expected to provide better and better approximations to σ together with an interval in which σ provably lies. For instance, $m_0 = 10$ is enough to determine that $0.66 < \sigma < 0.69$, and the choice $m_0 = 100$ gives σ to 15 guaranteed digits of accuracy, namely, $\sigma \doteq 0.67740\,17761\,30660$. Then, the asymptotic formula (53) instantiates as

(54) $$C_n^{\text{Prime}} \sim g(n), \qquad g(n) := 0.30365\,52633 \cdot 1.47622\,87836^n.$$

The constant $\sigma^{-1} \doteq 1.47622$ is akin to the family of Backhouse constants described in [**61**].

Once more, the asymptotic approximation is very good as shown by the pyramid of Figure 9. The difference between C_n^{Prime} and its approximation $g(n)$ from Eq. (54) is plotted on the left of Figure 10. The seemingly haphazard oscillations that manifest themselves are well explained by the principles discussed in the previous section. It appears that the next poles of the OGF are complex conjugate and lie near $-0.76 \pm 0.44i$, having modulus about 0.88. The corresponding residues then jointly contribute a quantity of the form

$$g_2(n) = c \cdot A^n \sin(\omega n + \omega_0), \qquad A \doteq 1.13290,$$

for some constants c, ω, ω_0. Comparing the left and right parts of Figure 10 shows this next layer of poles to explain quite well the residual error $C_n^{\mathrm{Prime}} - g(n)$. (The diagram on the right in Figure 10 also displays the values of the continuous interpolation to $g_2(n)$.)

Here is a final example that demonstrates in a striking way the scope of the method. Define the set Prime_2 of "twinned primes" as the set of primes that belong to a twin prime pair, that is, $p \in \mathrm{Prime}_2$ if one of $p-2, p+2$ is prime. The set Prime_2 starts as $3, 5, 7, 11, 13, 17, 19, 29, 31, \ldots$ (numbers like 23 or 37 are thus excluded). The asymptotic formula for the number of compositions of the integer n into summands that are twinned primes, is

$$C_n^{\mathrm{Prime}_2} \sim 0.18937 \cdot 1.29799^n.$$

It is quite remarkable that the constants involved are still computable real numbers (and of low complexity, even), this despite the fact that it is not known whether the set of twinned primes is finite or infinite. Incidentally, a sequence that starts like $C_n^{\mathrm{Prime}_2}$,

$$1, 0, 0, 1, 0, 1, 1, 1, 2, 1, 3, 4, 3, 7, 7, 8, 14, 15, 21, 28, 33, 47, 58, \ldots$$

and coincides till index 22 included (!), but not beyond, was encountered by P. A. MacMahon[3], as the authors discovered, much to their astonishment, from scanning Sloane's Encyclopedia, where it appears as *EIS* **A002124**. □

Profiles of supercritical sequences. We have seen in Chapter III that integer compositions and integer partitions, when sampled at random, tend to assume rather different aspects. Given a sequence construction, $\mathcal{F} = \mathfrak{S}(\mathcal{G})$, the profile of an element $\alpha \in \mathcal{F}$ is the vector $(X^{\langle 1\rangle}, X^{\langle 2\rangle}, \ldots)$ where $X^{\langle 1\rangle}(\alpha)$ is the number of $\mathcal{G}$–components in α that have size j. In the case of (unrestricted) integer compositions, it could be proved elementarily that, on average and for size n, the number of 1 summands is $\sim n/2$, the number of 2 summands is $\sim n/4$, and so on. Now that meromorphic asymptotic is available, such a property can be placed in a much wider perspective.

PROPOSITION V.6. *Consider a supercritical sequence construction, $\mathcal{F} = \mathfrak{S}(\mathcal{G})$, with the "unperiodic" condition. The number of $\mathcal{G}$– components of any fixed size k in a random $\mathcal{F}$– object of size n satisfies*

$$\mathbb{E}_n(X^{\langle k\rangle}) = \frac{g_k \sigma^k}{\sigma G'(\sigma)} n + O(1), \quad \mathbb{V}_n(X^{\langle k\rangle}) = O(n). \tag{55}$$

There, σ is the root in $(0, \rho_G)$ of $G(\sigma) = 1$, and $g_k = [z^k]G(z)$.

PROOF. The bivariate GF with u marking the number of $\mathcal{G}$–components of size k is

$$F(z, u) = \frac{1}{1 - (G(z) + (u-1) g_k z^k)},$$

[3]See "Properties of prime numbers deduced from the calculus of symmetric functions", *Proc. London Math. Soc.*, 23 (1923), 290-316). MacMahon's sequence corresponds to the compositions into arbitrary *odd* primes.

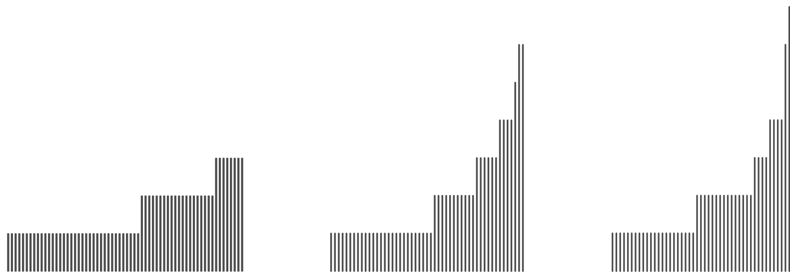


FIGURE 11. Profile of structures drawn at random represented by the sizes of their components in sorted order: (from left to right) a random surjection, alignment, and composition of size $n = 100$.

as results from the theory developed in Chapter III. The mean value is then given by a quotient,

$$\mathbb{E}_n(X^{\langle k\rangle}) = \frac{1}{f_n}\,[z^n]\,\left.\frac{\partial}{\partial u}F(z,u)\right|_{u=1} = \frac{1}{f_n}[z^n]\frac{g_k z^k}{(1-G(z))^2}.$$

The GF of cumulated values has a double pole at $z = \sigma$, and the estimate of the mean value follows. The variance is estimated similarly, after two successive differentiations and the analysis of a triple polar singularity. $\square$

The total number of components X satisfies $X = \sum X^{\langle k\rangle}$, and, by Theorem V.3, its mean is asymptotic to $n/(\sigma G'(\sigma))$. Thus, Equation (55) indicates that, at least in some average-value sense, the "proportion" of components of size k amongst all components is given by $g_k\sigma^k$. Also, since $G(\sigma) = 1$, the coefficients $g_k\sigma^k$ add up to 1.

EXAMPLE 10. *The profiles of compositions, surjections, and alignments.* Proposition V.6 immediately applies to compositions (that are sequences of sequences), surjections (sequences of sets), and alignments (sequences of cycles). The following table summarizes the conclusions:

Structures	*Specif.*	*Law* $(g_k\sigma^k)$	*Type*	σ
Compositions	$\mathfrak{S}(\mathfrak{S}_{\geq 1}(\mathcal{Z}))$	$\frac{1}{2^k}$	Geometric	$\frac{1}{2}$
Surjections	$\mathfrak{S}(\mathfrak{P}_{\geq 1}(\mathcal{Z}))$	$\frac{1}{k!}(\log 2)^k$	Poisson	$\log 2$
Alignments	$\mathfrak{S}(\mathfrak{C}(\mathcal{Z}))$	$\frac{1}{k}(1-e^{-1})^k$	Logarithmic	$1-e^{-1}$

The geometric and Poisson law are classical; the *logarithmic distribution* (also called "logarithmic-series distribution") of parameter λ is by definition the law of a random variable Y such that

$$\mathbb{P}(Y = k) = \frac{1}{\log(1-\lambda)^{-1}}\frac{\lambda^k}{k}.$$

The way the internal construction induces the law of component sizes,

Sequence $\mapsto$ Geometric; Set $\mapsto$ Poisson; Cycle $\mapsto$ Logarithmic,

stands out. Figure 11 exemplifies the phenomenon by displaying components sorted by size and represented by vertical segments of corresponding lengths for three randomly drawn objects of size $n = 100$. $\square$

▷ **9.** *Proportion of k–components and convergence in probability.* For any fixed k, the random variable $X_n^{\langle k\rangle}/X_n$ converges in probability (the notion is defined in Chapter III) to the value $g_k\sigma^k$,

$$\frac{X_n^{\langle k\rangle}}{X_n} \xrightarrow{P} g_k\sigma^k, \quad \text{i.e.,} \quad \lim_{n\to\infty} \mathbb{P}\left\{ g_k\sigma^k(1-\epsilon) \le \frac{X_n^{\langle k\rangle}}{X_n} \le g_k\sigma^k(1+\epsilon) \right\} = 1,$$

for any $\epsilon > 0$. The proof is an easy consequence of the Chebyshev inequalities (the distributions of X_n and $X_n^{\langle k\rangle}$ are both concentrated). ◁

▷ **10.** *Random generation of supercritical sequences.* Let $\mathcal{F} = \mathfrak{S}(\mathcal{G})$ be a supercritical sequence scheme. Consider a sequence of i.i.d. (independently identically distributed) random variables $Y_1, Y_2, \ldots$ each of them obeying the discrete law

$$\mathbb{P}(Y = k) = g_k\sigma^k, \qquad k \ge 1.$$

A sequence is said to be hitting n if $Y_1 + \cdots + Y_r = n$ for some $r \ge 1$. The vector $(Y_1, \ldots, Y_r)$ for a sequence conditioned to hit n has the same distribution as the sequence of the lengths of components in a random $\mathcal{F}$–object of size n.

For probabilists, this explains the shape of the formulæin Theorem V.3, which resemble renewal relations [**58**, Sec. XIII.10]. It also implies that, given a uniform random generator for $\mathcal{G}$–objects, one can generate a random $\mathcal{F}$–object of size n in $O(n)$ steps on average [**52**]. This applies to surjections, alignments, and compositions in particular. ◁

▷ **11.** *Largest components in supercritical sequences.* Let $\mathcal{F} = \mathfrak{S}(\mathcal{G})$ be a supercritical sequence. Assume that $g_k = [z^k]G(z)$ satisfies the asymptotic "smoothness" condition

$$g_k \underset{k\to\infty}{\sim} c\rho^{-k}k^\beta, \qquad c, \rho \in \mathbb{R}_{>0},\ \beta \in \mathbb{R}.$$

Then the size L of the largest $\mathcal{G}$ component in a random $\mathcal{F}$ object satisfies, for size n,

$$\mathbb{E}_{\mathcal{F}_n}(X) = \frac{1}{\log(\rho/\sigma)}\left(\log n + \beta \log\log n\right) + o(\log\log n).$$

This covers integer compositions ($\rho = 1$, $\beta = 0$) and alignments ($\rho = 1$, $\beta = -1$). [The analysis generalizes the case of longest runs in Example 2 and is based on similar principles. The GF of $\mathcal{F}$ objects with $L \le m$ is $F^{\langle m\rangle}(z) = \left(1 - \sum_{k\le m} g_k z^k\right)^{-1}$, according to Section III.7. For m large enough, this has a dominant singularity which is a simple pole at σ_m such that $\sigma_m - \sigma \sim c_1(\sigma/\rho)^m m^\beta$. There follows a double-exponential approximation

$$\mathbb{P}_{\mathcal{F}_n}(L \le m) \approx \exp\left(-c_2 n m^\beta (\sigma/\rho)^m\right)$$

in the "central" region. See Gourdon's study [**97**] for details.] ◁

V. 4. Functional equations: positive rational systems

For rational functions, positivity coupled with some simple ancillary conditions entails a host of important properties, like unicity of the dominant singularity. Such facts result from the classical Perron-Frobenius theory of nonnegative matrices that we summarize in this section. They in turn imply strong properties of large random structures.

The basic case is that of a d-dimensional column vector$\mathrm{y}(z)$ of generating functions satisfying a linear system of the form

$$\mathrm{y}(z) = \mathrm{a} + zT\,\mathrm{y}(z),$$

for some $(d \times d)$ matrix T and vector a. If T satisfies suitable positivity conditions and a is nonnegative, then any component $\mathrm{y}_j(z)$ closely resembles the extremely simple rational function,

$$\frac{1}{1-\lambda_1 z},$$

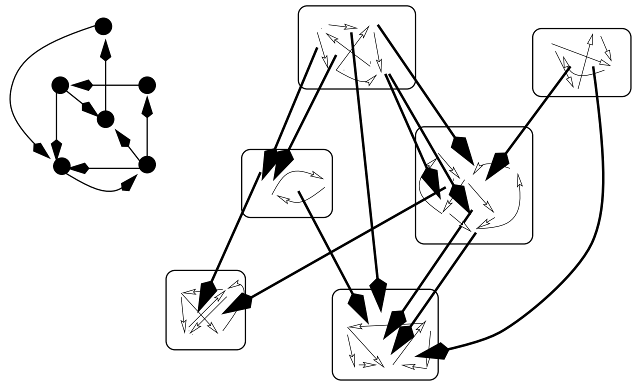


FIGURE 12. The irreducibility conditions of Perron-Frobenius theory. Left: a strongly connected digraph. Right: a weakly connected digraph that is not strongly connected is a collection of strongly connected components related by a directed acyclic graph.

where λ_1 is a well-characterized eigenvalue of T. Accordingly, the asymptotic phenomena associated with such systems are highly predictable. We propose to expose here the general theory and treat in the next section classical applications to paths in graphs and to languages recognized by finite automata.

V. 4.1. Perron-Frobenius theory of nonnegative matrices. For an arbitrary square matrix $A \in \mathbb{R}^{m\times m}$, the *spectrum* is the set of its *eigenvalues*, that is, the set of λ such that $\lambda I - A$ is not invertible (i.e., not of full rank), where I is the unit matrix with the appropriate dimension. A *dominant eigenvalue* is one of largest modulus. Finally, the spectral radius of an arbitrary matrix A is defined as

$$\sigma(A) = \max_j \{|\lambda_j|\}, \tag{56}$$

where the set $\{\lambda_j\}$ is the set of eigenvalues of A (also called spectrum). The spectral radius $\sigma(A)$ describes growth properties associated to the powers of A. Indeed, given the Jordan normal form of matrices, it is easy to see that all entries of A^n are bounded from above by a multiple of $\sigma(A)^n \cdot n^{r-1}$, where r is the maximum multiplicity of any dominant eigenvalue. When analysing a family of combinatorial models that admit a matrix formulation, it is then of obvious interest to determine the value of the spectral radius and the multiplicities attached to dominant eigenvalues.

The properties of positive and of nonnegative matrices have been superbly elicited by Perron [**157**] in 1907 and by Frobenius [**87**] in 1908–1912. The corresponding theory has far-reaching implications: it lies at the basis of the theory of finite Markov chains and it extends to positive operators in infinite-dimensional spaces [**124**].

For A a scalar matrix of dimension $m \times m$ with nonnegative entries, a crucial rôle is played by the *dependency graph*; this is the (directed) graph with vertex set $V = \{1 \mathinner{\ldotp\ldotp} m\}$ and edge set containing the directed edge $(a \to b)$ iff $A_{a,b} \neq 0$. The reason for this terminology is the following: Let A represent the linear transformation $\left\{y_i^\star = \sum_j A_{i,j} y_j\right\}_i$; then, the fact that an entry $A_{i,j}$ is nonzero means that $y_i^\star$ depends effectively on y_j and is translated by the directed edge $(i \to j)$ in the dependency graph.

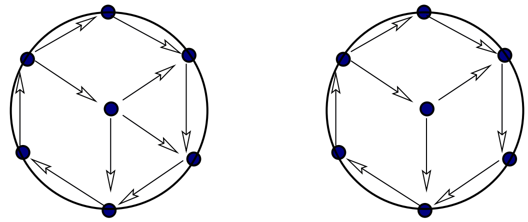


FIGURE 13. The aperiodicity conditions of Perron-Frobenius theory: an aperiodic digraph (left) and a periodic digraph (right).

From this point on, we consider matrices with nonnegative entries. Two notions are essential, irreducibility and aperiodicity (the terms are borrowed from Markov chain theory and matrix theory).

DEFINITION V.3. *The matrix A is called* irreducible *if its dependency graph is strongly connected (i.e., any two vertices are connected by a directed path). A strongly connected digraph G is* periodic *with parameter d iff all its cycles have a length that is a multiple of d. In that case, the graph decomposes into cyclically arranged layers: the vertex set V can be partitioned into d classes, $V = V_0 \cup \cdots \cup V_{d-1}$, in such a way that the edge set E satisfies*

$$E \subseteq \bigcup_{i=0}^{d-1} (V_i \times V_{(i+1) \bmod d}). \tag{57}$$

The maximal possible d is called the period. *If no decomposition exists with $d \geq 2$, so that the period has the trivial value 1, then the graph and all the matrices that admit it as their dependency graph are called* aperiodic.

By considering only simple paths, it is then seen that irreducibility is equivalent to the condition that $(I + A)^m$ has all its entries that are strictly positive. See Figure 12 for a graphical rendering of irreducibility and for the general structure of a (weakly connected) digraph. As an illustration of periodicity, a directed 10-cycle is periodic with parameter $d = 1, 2, 5, 10$ and the period is 10. See Figure 13 for representations of a periodic and an aperiodic digraph.

Periodicity also means that the existence of paths of length n between any given pair of nodes $\langle i, j\rangle$ is constrained by the congruence class $n \bmod d$. A contrario, aperiodicity entails the existence, for all n sufficiently large, of paths of length n connecting $\langle i, j\rangle$. From the definition, a matrix A with period d has, up to simultaneous permutation of its rows and columns, a cyclic block structure

$$\begin{pmatrix} 0 & \boxed{A_{0,1}} & 0 & \cdots & 0 \\ 0 & 0 & \boxed{A_{1,2}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \boxed{A_{d-2,d-1}} \\ \boxed{A_{d-1,0}} & 0 & 0 & \cdots & 0 \end{pmatrix}$$

where the blocks $A_{i,i+1}$ are reflexes of the connectivity between V_i and V_{i+1} in (57).

THEOREM V.4 (Perron-Frobenius theorem). *Let A be a matrix that is assumed to be irreducible in the sense that its dependency graph is strongly connected.*

(i) *If A has (strictly)* positive elements, *then its eigenvalues can be ordered in such a way that*

$$\lambda_1 > |\lambda_2| \geq |\lambda_3| \geq \cdots,$$

i.e., A has a unique dominant eigenvalue which is positive and simple.

(ii) *If A has* nonnegative elements, *then its eigenvalues can be ordered in such a way that*

$$\lambda_1 = |\lambda_2| = \cdots = |\lambda_d| > |\lambda_{d+1}| \geq |\lambda_{d+2}| \geq \cdots,$$

and each of the dominant eigenvalues is simple, with λ_1 being positive. Furthermore, the quantity d is precisely equal to the period of the dependency graph. If $d = 1$, in particular, then there is unicity of the dominant eigenvalue. If $d \geq 2$, the whole spectrum is invariant under the set of transformations

$$\lambda \mapsto \lambda e^{2ij\pi/d}, \qquad j = 0, 1, \ldots, d-1.$$

For proof techniques including a full proof of Part (i) of the theorem, see APPENDIX: *Perron-Frobenius theory of nonnegative matrices*, p. 233.

For short, one says that a matrix is positive (resp. nonnegative) if all its elements are positive (resp. nonnegative). Here are two useful turnkey results, Corollaries V.1 and V.2.

COROLLARY V.1. *Any one of the following conditions suffices to guarantee the existence of a unique dominant eigenvalue of a nonnegative matrix T:*

(i) *T has (strictly) positive entries;*

(ii) *T is such that, some power T^s is (strictly) positive;*

(iii) *T is irreducible and at least one diagonal element of T is nonzero;*

(iv) *T is irreducible and the dependency graph of T is such that there exist at least two paths from the same source to the same destination that are of relatively prime lengths.*

PROOF. The proof makes use of the well-known correspondence between terms in coefficients of matrix products and paths in graphs (see below Section V. 5 for more). Sufficiency of condition (i) results directly from Case (i) of Theorem V.4. Condition (ii) immediately implies irreducibility. Unicity of the dominant eigenvalue (hence aperiodicity) results from Perron-Frobenius properties of A^s, by which $\lambda_1^s > |\lambda_2|^s$. (Also, by elementary graph combinatorics, one can always take the exponent s to be at most the dimension m.) By basic combinatorics of paths in graphs, Conditions (iii) and (iv) each imply Condition (ii). $\square$

V. 4.2. Positive rational functions. The importance of Perron-Frobenius theory and of its immediate consequence, Corollary V.1, stems from the fact that uniqueness of the dominant eigenvalue is usually related to a host of analytic properties of generating functions as well as probabilistic properties of structures. In particular, as we shall see in the next section, several combinatorial problems (like automata or paths in graphs) can be reduced to the following case.

COROLLARY V.2. *Consider the matrix*

$$F(z) = (I - zT)^{-1},$$

where T, called the "transition matrix", is a scalar nonnegative matrix. It is assumed that T is irreducible. Then each entry $F_{i,j}(z)$ of $F(z)$ has a radius of convergence ρ that coincides with the smallest positive root of the determinantal equation

$$\Delta(z) := \det(I - zT) = 0.$$

Furthermore, the point ρ is a simple pole of any $F_{i,j}(z)$.

In addition, if T is aperiodic or if it satisfies any of the conditions of Corollary V.1, then all singularities other than ρ are strictly dominated in modulus by ρ.

The statement obviously applies to any positive linear combinations of entries of F and thus to solutions of any system of the form $\mathrm{y}(z) = \mathrm{a} + zT\,\mathrm{y}(z)$.

PROOF. Define first (as in the statement) $\rho = 1/\lambda_1$, where λ_1 is the eigenvalue of T of largest modulus that is guaranteed to be simple by assumption of irreducibility and by Perron-Frobenius properties. Next, the relations induced by $F = I + zTF$, namely,

$$F_{i,j}(z) = \delta_{i,j} + z \sum_k T_{i,k} F_{k,j}(z),$$

together with positivity and irreducibility entail that the $F_{i,j}(z)$ must all have the same radius of convergence r. Indeed, each F_{ij} depends positively on all the other ones (by irreducibility) so that any infinite value of an entry in the system must propagate to all the other ones.

The characteristic polynomial

$$\Delta(z) = \det(I - zT),$$

has roots that are inverses of the eigenvalues of T and $\rho = 1/\lambda_1$ is smallest in modulus. Thus, since Δ is the common denominator to all the $F_{i,j}(z)$, poles of any $F_{i,j}(z)$ can only be included in the set of zeros of this determinant, so that the inequality $r \geq \rho$ holds.

It remains to exclude the possibility $r > \rho$, which means that no "cancellations" with the numerator can occur at $z = \rho$. The argument relies on finding a positive combination of some of the $F_{i,j}$ that *must* be singular at ρ. We offer two proofs, each of interest in its own right: one (a) is conveniently based on the Jacobi trace formula, the other (b) is based on supplementary Perron–Frobenius properties.

(a) Jacobi's trace formula for matrices [**96**, p. 11],

$$\det \circ \exp = \exp \circ \mathrm{Tr} \qquad \text{or} \qquad \log \circ \det = \mathrm{Tr} \circ \log \tag{58}$$

generalizes the scalar identities[4] $e^a e^b = e^{a+b}$ and $\log ab = \log a + \log b$. Here we have (for z small enough)

$$\begin{aligned} \mathrm{Tr} \log(I - zT)^{-1} &= \sum_i \sum_{n \geq 1} T_{i,i,n} \frac{z^n}{n} \\ &= \log \det(I - zT)^{-1}, \end{aligned}$$

where the first line results from expansion of the logarithm and the second line is an instance of the trace formula. Thus, by differentiation, the sum $\sum_i M_{i,i}(z)$ is seen to be singular at $\rho = 1/\lambda_1$ and we have established that $r = \rho$.

(b) Alternatively, let v_1 be the eigenvector of T corresponding to λ_1. Perron-Frobenius theory also teaches us that, under the irreducibility and aperiodicity conditions, the vector v_1 has all its coordinates that are nonzero. Then the quantity

$$(1 - zT)^{-1} v_1 = \frac{1}{1 - z\lambda_1} v_1$$

is certainly singular at $1/\lambda_1$. But it is also a linear combination of the $F_{i,j}$'s. Thus at least one of the entries of F (hence all of them by the discussion above) must be singular at $\rho = 1/\lambda_1$. Therefore, we have again $r = \rho$.

Finally, under the additional assumption that T is aperiodic, Perron-Frobenius theory grants us that $\rho = 1/\lambda_1$ is well-separated in modulus from all other singularities F. $\square$

[4] The Jacobi trace formula is readily verified when the matrix is diagonalizable, and from there, it can be extended to all matrices by an algebraic "density" argument.

Several of these arguments will be recycled when we discuss the harder problem of analysing coefficients of positive algebraic functions in Chapter V.

We next proceed to show that properties of the Perron-Frobenius type even extend to a large class of linear systems of equations that have nonnegative polynomial coefficients. Such a case is important because of its applicability to transfer matrices; see Section V. 5 below.

Some definitions extending the ones of scalar matrices must first be set. A polynomial

$$p(z) = \sum_j c_j z^{e_j}, \qquad \text{every } c_j \neq 0,$$

is said to be primitive if the quantity $\delta = \gcd(\{e_j\})$ is equal to 1; it is imprimitive otherwise. Equivalently, $p(z)$ is imprimitive iff $p(z) = q(z^\delta)$ for some *bona fide* polynomial q and some $\delta > 1$. Thus, $z, 1+z, z^2+z^3, z+z^4+2z^8$ are primitive while $1, 1+z^2, z^3+z^6, 1+2z^8+5z^{12}$ are not.

DEFINITION V.4. *A linear system with polynomial entries,*

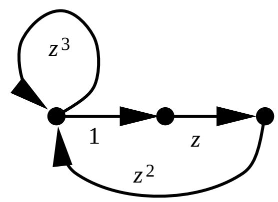
$$f(z) = v(z) + T(z) f(z) \tag{59}$$

where $T \in \mathbb{R}[z]^{r\times r}$, $v \in \mathbb{R}[z]^r$, and $f \in \mathbb{R}[z]^r$ the vector of unknowns is said to be:

- (a) rationally proper (r–proper) *if $T(0)$ is nilpotent, meaning that $T(0)^r$ is the null matrix;*
- (b) rationally nonnegative (r–nonnegative) *if each component $v_j(z)$ and each matrix entry $T_{i,j}(z)$ lies in $\mathbb{R}_{\geq 0}[z]$;*
- (c) rationally irreducible (r–irreducible) *if $(I + T(z))^r$ has all its entries that are nonzero polynomials.*
- (d) rationally aperiodic (r–periodic) *if at least one diagonal entry of some power $T(z)^e$ is a primitive polynomial.*

It is again possible to visualize these properties of matrices by drawing a directed graph whose vertices are labelled $1, 2, \ldots, r$, with the edge connecting i to j that is weighted by the entry $T_{i,j}(z)$ of matrix $T(z)$. Properness means that all sufficiently long paths (and all cycles) must involve some positive power of z— it is a condition satisfied in well-founded combinatorial problems; irreducibility means that the dependency graph is strongly connected by paths whose edges are associated with nonzero polynomials; periodicity means that all closed paths involve weights that are polynomials in some z^e for some $e > 1$.

For instance, if W is a matrix with positive entries, then zW is r–irreducible and r–aperiodic, while z^3W is r–periodic. The matrix $T = \begin{pmatrix} z & z^3 \\ 1 & 0 \end{pmatrix}$ is r–proper, r–irreducible, and r–aperiodic, since $T^2 = \begin{pmatrix} z^2+z^3 & z^4 \\ z & z^3 \end{pmatrix}$. The matrix $T = \begin{pmatrix} z^3 & 1 & 0 \\ 0 & 0 & z \\ z^2 & 0 & 0 \end{pmatrix}$ is r–proper, but it fails to be r–aperiodic since all cycles only involve powers of z^3, as is visible on the associated graph:



By abuse of language, we say that $f(z)$ is a solution of a linear system if it coincides with the first component of a solution vector, $f \equiv f_1$. The following theorem generalizes Corollary V.2.

THEOREM V.5 (Positive rational systems). (i) *Assume that a rational function $f(z)$ is a solution of a system (59) that is r– positive, r– proper, r– irreducible, and r– aperiodic. Then, $f(z)$ has a unique dominant singularity ρ that is positive, and is a simple pole; ρ is the smallest positive solution of*

$$\det(I - T(z)) = 0. \tag{60}$$

(ii) *Assume that $f(z)$ is a solution of a system that is r– positive, r– proper, and r– irreducible (but not necessarily r– aperiodic). Then, the set of dominant singularities of $f(z)$ is of the form $\{\rho_j\}_{j=0}^{d-1}$, where $\rho_0 \in \mathbb{R}_{\geq 0}$, $\rho_j/\rho_0 = \eta$ is a root of unity, and $\rho_j \eta^\ell$ is a dominant singularity for all $\ell = 0, 1, 2, \ldots$. In addition, each ρ_j is a simple pole.*

PROOF. Consider first Case (i). For any fixed $x > 0$, the matrix $T(x)$ satisfies the Perron Frobenius conditions, so that it has a maximal positive eigenvalue $\lambda_1(x)$ that is simple. More information derives from the introduction of matrix norms[5]. Spectral radius and matrix norms are intimately related since

$$\sigma(A) = \lim_{n \to +\infty} \left(||A^n||\right)^{1/n}.$$

In particular, this relation entails that the spectral radius is an increasing function of matrix entries: for nonnegative matrices, if $A \leq B$ (in the sense that $A_{i,j} \leq B_{i,j}$ for all i, j), then one has $\sigma(A) \leq \sigma(B)$; if $A < B$ (in the sense that $A_{i,j} < B_{i,j}$ for all i, j), then one has $\sigma(A) < \sigma(B)$. (To see the last inequality, note the existence of $\epsilon > 0$ such that $A \leq (1-\epsilon)B$.)

Returning to the case at hand, equation (56) and the surrounding remarks imply that the spectral radius $\sigma(T(x))$, which also equals $\lambda_1(x)$ for positive x, satisfies

$$\lambda_1(0) = 0, \qquad \lambda_1(x) \text{ strictly increasing}, \qquad \lambda_1(+\infty) = +\infty.$$

(The first condition reflects properness, the second one is a consequence of irreducibility, and the last one derives from simple majorizations.) In particular, the equation $\lambda_1(x) = 1$ admits a unique root ρ on $(0, +\infty)$. (Notice that $\lambda_1(x)$ is a real branch of the algebraic curve $\det(\lambda I - T(x)) = 0$ that dominates all other branches in absolute value for $x > 0$. There results from the general theory of algebraic functions that $\lambda_1(x)$ is analytic at every point $x > 0$.)

[5]A matrix norm $||.||$ satisfies: $||A|| = 0$ implies $A = 0$; $||cA|| = |c| \cdot ||A||$; $||A + B|| \leq ||A| + ||B||$; $||A \times B|| \leq ||A|| \cdot ||B||$.

There remains to prove that: (a) ρ is at most a simple pole of $f(z)$; (b) ρ is actually a pole; (c) there are no other singularities of modulus equal to ρ.

Fact (a) amounts to the property that ρ is a simple root of the equation $\lambda(\rho) = 1$, that is, $\lambda'(\rho) \neq 0$. (To prove $\lambda'(\rho) \neq 0$, we can argue a contrario. First derivatives $\lambda'(\rho), \lambda''(\rho)$, etc, cannot be zero till some *odd* order inclusively since this would contradict the increasing character of $\lambda(x)$ around ρ along the real line. Next, if derivatives till some *even* order ≥ 2 inclusively were zero, then we would have by the local analytic geometry of $\lambda(z)$ near ρ some complex value z_1 satisfying: $|\lambda(z_1)| = 1$ and $|z_1| < \rho$; but for such a value z_1, by irreducibility and aperiodicity, for some exponent e, the entries of $T(z_1)^e$ would be all strictly dominated in absolute value by those of $T(\rho)^e$, hence a contradiction.) Then, $\lambda'(\rho) \neq 0$ holds and by virtue of

$$\det(I - T(z)) = (1 - \lambda_1(z)) \prod_{j \neq 1} (1 - \lambda_j(z)) = (1 - \lambda_1(z)) \frac{\det(I - T(z))}{1 - \lambda_1(z)},$$

the quantity ρ is only a simple root of $\det(I - T(z))$.

Fact (b) means that no “cancellation” may occur at $z = \rho$ between the numerator and the denominator given by Cramer’s rule. It derives from an argument similar to the one employed for Corollary V.2. Fact (c) derives from aperiodicity and the Perron-Frobenius properties. $\square$

V. 5. Paths in graphs, automata, and transfer matrices.

A cluster of applications of rational functions is to problems that are naturally described as paths in digraphs, or equivalently as finite automata. In physics, the corresponding treatment is also the basis of what is called the “transfer matrix method”. We start our exposition with the enumeration of paths in graphs that constitutes the most direct introduction to the subject.

V. 5.1. Paths in graphs. Let G be a directed graph with vertex set $\{1, \ldots, m\}$, where self-loops are allowed and label each edge (a, b) by the *formal* variable $g_{a,b}$. We introduce the matrix $\mathbf{G}$ such that

$$\mathbf{G}_{a,b} = g_{a,b} \text{ if the edge } (a,b) \in G, \qquad \mathbf{G}_{a,b} = 0 \text{ otherwise,} \tag{61}$$

which is called the *formal adjacency matrix* of G. Then, from the standard definition of matrix products, the powers $\mathbf{G}^r$ have elements that are path polynomials. More precisely, one has the simple but essential relation,

$$(\mathbf{G})^r_{i,j} = \sum_{w \in \mathcal{P}(i,j;r)} w, \tag{62}$$

where $\mathcal{P}(i, j, r)$ is the set of paths in G that connect i to j and have length r, and a path w is assimilated to the monomial in indeterminates $\{g_{i,j}\}$ that represents multiplicatively the succession of its edges; for instance:

$$(\mathbf{G})^3_{i,j} = \sum_{m_1 = i, m_2, m_3, m_4 = j} g_{m_1,m_2} g_{m_2,m_3} g_{m_3,m_4},$$

In other words, powers of the matrix associated to a graph “generate” all paths in a graph. One may then treat simultaneously all lengths of paths (and all powers of matrices) by introducing the variable z to record length.

PROPOSITION V.7. (i) *Let G be a digraph and let $\mathbf{G}$ be the matrix associated to G by rules* (61) *The OGF $F^{\langle i,j\rangle}(z)$ of the set of all paths from i to j in a digraph G with z marking length and $g_{a,b}$ marking the occurrence of edge (a,b) is the entry i,j of the matrix $(I-z\mathbf{G})^{-1}$, namely*

$$F^{\langle i,j\rangle}(z) = \left.(I-z\mathbf{G})^{-1}\right|_{i,j} = \frac{\Delta^{\langle i,j\rangle}(z)}{\Delta(z)},$$

where $\Delta(z)=\det(I-z\mathbf{G})$ and $\Delta^{\langle i,j\rangle}(z)$ is the determinant of the minor of index i,j of $I-z\mathbf{G}$.

(ii) *The generating function of nonempty closed paths is given by*

$$\sum_i (F^{\langle i,i\rangle}(z)-1) = -z\frac{\Delta'(z)}{\Delta(z)}.$$

The quantity $\det(I-z\mathbf{G})$ is obviously the reciprocal polynomial of the characteristic polynomial of $\mathbf{G}$.

PROOF. Part (i) results from the discussion above which implies

$$F^{\langle i,j\rangle}(z) = \sum_{n=0}^{\infty} z^n \left(\mathbf{G}^n\right)_{i,j} = \left((I-z\mathbf{G})^{-1}\right)_{i,j},$$

and from the cofactor formula of matrix inversion. Part (ii) results from Jacobi's trace formula (58). Introduce the quantity known as the *zeta function*,

$$\begin{aligned}
\zeta(z) &:= \exp\left(\sum_i \sum_{n=1}^{\infty} F_n^{\langle i,i\rangle}\frac{z^n}{n}\right) &= \exp\left(\sum_{n=1}^{\infty}\frac{z^n}{n}\mathrm{Tr}\,\mathbf{G}^n\right)\\
&= \exp\left(\mathrm{Tr}\,\log(I-z\mathbf{G})^{-1}\right) &= \det(I-z\mathbf{G})^{-1},
\end{aligned}$$

where the last line results from the Jacobi trace formula. Thus, $\zeta(z)=\Delta(z)^{-1}$. On the other hand, differentiation combined with the definition of $\zeta(z)$ yields

$$\begin{aligned}
z\frac{\zeta'(z)}{\zeta(z)} &= -z\frac{\Delta'(z)}{\Delta(z)}\\
&= \sum_i\sum_{n=1}^{\infty} F_n^{\langle i,i\rangle} z^n,
\end{aligned}$$

and Part (ii) follows. □

The numeric substitution $\sigma:\ g_{a,b}\mapsto 1$ transforms the matrix $\mathbf{G}$ into the usual *adjacency matrix*. In particular, the number of paths of length n is obtained, under this substitution, as $[z^n](1-zG])^{-1}$. In a similar vein, it is possible to consider weighted graphs, where the $g_{a,b}$ are assigned *real-valued weights*; with the weight of a path being defined by the product of its edges weights, one finds that $[z^n](I-z\mathbf{G})^{-1}$ equals the total weight of all paths of length n. If furthermore the assignment is made in such a way that $\sum_b g_{a,b}=1$, then the matrix $\mathbf{G}$, which is called a stochastic matrix, can be interpreted as the transition matrix of a Markov chain.

▷ **12.** *Fast computation of the characteristic polynomial.* Observe that

$$z\frac{\zeta'(z)}{\zeta(z)} = \sum_{n\ge 1} z^n \mathrm{Tr}\,\mathbf{G}^n = \sum_\lambda \frac{\lambda z}{1-\lambda z},$$

(the sum is over eigenvalues). From this, one deduces an algorithm that determines the characteristic polynomial of a matrix of dimension m in $\mathcal{O}(m^4)$ arithmetic operations. [Hint: computing the quantities $\mathrm{Tr}\,G^j$ for $j=1,\ldots,m$ requires precisely m matrix multiplications.] ◁

$\triangleright$ **13.** *The matrix tree theorem.* Let G be a directed graph without loops and associated matrix $\mathbf{G}$, with $g_{a,b}$ marking edge (a,b). The Laplacian matrix $\mathbf{L}[G]$ is defined by

$$\mathbf{L}[G]_{i,j} = -g_{i,j} + [\![i=j]\!] \sum_k g_{i,k}.$$

Let $\mathbf{L}_1[G]$ be the matrix obtained by deleting the first row and first column of $\mathbf{L}[G]$. Then, the "tree polynomial"

$$T_1[G] := \det \mathbf{L}_1[G]$$

enumerates all (oriented) spanning trees of G rooted at node 1. [This classic result belongs to a circle of ideas initiated by Kirchhoff, Sylvester, Borchardt and others in the 19th century. See, e.g., the discussions by Knuth [**120**, p. 582–583] and Moon [**143**].] $\triangleleft$

Let us now assume that positive weights are assigned to the edges of G. In other words, the quantities $g_{a,b}$ in (61) have positive values. If the resulting matrix is irreducible and aperiodic, then Perron-Frobenius theory applies. There exists $\rho = 1/\lambda_1$, with $\lambda_1 > 0$ the dominant eigenvalue of $\mathbf{G}$, and the OGF of weighted paths from i to j has a simple pole at ρ. A host of probabilistic properties of paths result from there, after a certain "residue matrix" has been calculated:

LEMMA V.1 (Iteration of Perron-Frobenius matrices). *Set $M(z) = (I - z\mathbf{G})^{-1}$ where $\mathbf{G}$ has nonnegative entries, is irreducible, and is aperiodic. Let λ_1 be the dominant eigenvalue of $\mathbf{G}$. Then the "residue" matrix R such that*

$$(I - z\mathbf{G})^{-1} = \frac{R}{1 - z\lambda_1} + O(1) \qquad (z \to \lambda_1^{-1}) \tag{63}$$

has entries given by ($\langle x, y\rangle$ represents a scalar product)

$$R_{ij} = \frac{r_i \ell_j}{\langle r, \ell\rangle},$$

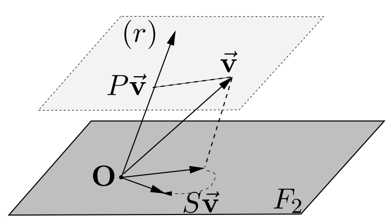
where r and ℓ are right and left eigenvectors of $\mathbf{G}$ corresponding to the eigenvalue λ_1.

PROOF. Let $\mathcal{E}$ be the ambient space. There exists a direct sum decomposition $\mathcal{E} = \mathcal{F}_1 + \mathcal{F}_2$ where $\mathcal{F}_1$ is the 1-dimensional eigenspace generated by the eigenvector (r) corresponding to eigenvalue λ_1 and $\mathcal{F}_2$ is the supplementary space which is the direct sum of characteristic spaces corresponding to the other eigenvalues $\lambda_2, \dots$. (For the purposes of the present discussion, one may freely think of the matrix as diagonalizable, with $\mathcal{F}_2$ the union of eigenspaces associated to $\lambda_2, \dots$.) Then $\mathbf{G}$ as a linear operator acting on $\mathcal{F}$ admits the decomposition

$$\mathbf{G} = \lambda_1 P + S,$$

where P is the projector on $\mathcal{F}_1$ and S acts on $\mathcal{F}_2$ with spectral radius $|\lambda_2|$, as illustrated by the diagram:

(64)



By standard properties of projections, $P^2 = P$ and $PS = SP = 0$ so that

$$\mathbf{G}^n = \lambda_1^n P + S.$$

Consequently, there holds,

$$(65)\qquad \begin{aligned}(I - z\mathbf{G})^{-1} &= \sum_{n\geq 0} z^n \lambda_1^n P + z^n S \\ &= \frac{P}{1-\lambda_1 z} + (I - zS)^{-1}.\end{aligned}$$

Thus, the residue matrix R coincides with the projector P.

Now, for any vector w, by general properties of projections, one has ($R \equiv P$):

$$Rw = c(w)r,$$

for some coefficient $c(w)$. Application of this to each of the base vectors e_j (i.e., $e_j = (\delta_{j1}, \ldots, \delta_{jd})^t$) shows that the matrix R has each of its *columns* proportional to the eigenvector r. A similar reasoning with the transpose $\mathbf{G}^t$ of $\mathbf{G}$ and the associated residue matrix R^t shows that the matrix R has each of its *rows* proportional to the eigenvector ℓ. In other words, for some constant γ, one must have

$$R_{i,j} = \gamma \ell_j r_i.$$

The normalization constant γ is itself finally determined by $\ell R r = \langle \ell, r\rangle$.

We finally observe that a full expansion can be obtained:

$$(66)\qquad (I - z\mathbf{G})^{-1}\frac{P}{1-\lambda_1 z} + \sum_{k\geq 0} R_k \left(z - \lambda_1^{-1}\right)^k, \qquad R_k := S^k (I - \lambda_1^{-1} S)^{-k-1}.$$

The proof also reveals that one needs to solve one polynomial equation for determining λ_1, and then the other quantities in (66) are all obtained by inverting matrices in the field of constants extended by the algebraic quantity λ_1. (Numerical procedures are likely to be used instead for large matrices.) $\square$

Equipped with this lemma, we can now state:

THEOREM V.6 (Random paths in digraphs). *Let* $\mathbf{G}$ *be a nonnegative matrix associated to a weighted digraph* G*, assumed to be irreducible and aperiodic. Consider the collection* $\mathcal{P}_{a,b}$ *of (weighted) paths with fixed origin* a *and final destination* b*. Then, the number of traversals of edge* (s,t) *in a random element of* $\mathcal{P}_{a,b}$ *has mean*

$$(67)\qquad \tau_{s,t} n + O(1) \qquad \textit{where} \quad \tau_{s,t} := \frac{\ell_s g_{s,t} r_t}{\lambda_1 \langle \ell, r\rangle}.$$

In other words, a long random path tends to spend asymptotically a fixed (nonzero) fraction of its time traversing any given edge. Accordingly, the number of visits of vertex s is also proportional to n and obtained by summing the expression of (67) according to all the possible values of t.

PROOF. First, the total weight ("number") of paths in $\mathcal{P}_{a,b}$ satisfies

$$(68)\qquad [z^n]\left[(I - z\mathbf{G})^{-1}\right]_{a,b} \sim \lambda_1 \frac{r_a \ell_b}{\langle \ell, r\rangle},$$

as follows from Lemma V.1. Next, introduce the modified matrix $\mathbf{H} = (h_{i,j})$ defined by

$$h_{i,j} = g_{i,j}\, u^{[\![i=s \wedge j=t]\!]}.$$

In other words, we mark each traversal of edge i,j by the variable u. Then, the quantity

$$(69)\qquad [z^n]\left[\frac{\partial}{\partial u}(I - z\mathbf{H})^{-1}\Big|_{u=1}\right]_{a,b}$$

represents the total number of traversals of edge (s,t), with weights taken into account. Simple algebra[6] shows that

$$\frac{\partial}{\partial u}(I-z\mathbf{H})^{-1}\bigg|_{u=1} = (I-z\mathbf{G})^{-1}\,(zH')\,(I-z\mathbf{G}), \tag{70}$$

where $H' := (\partial_u H)_{u=1}$ has all its entries equal to 0, except for the s,t entry whose value is $g_{s,t}$. By the calculation of the residue matrix in Lemma V.1, the coefficient of (69) is then asymptotic to

$$[z^n]\frac{R_{a,s}}{1-\lambda_1 z}g_{s,t}z\frac{R_{t,b}}{1-\lambda_1 z} \sim \upsilon n\lambda_1^{n-1}, \qquad \upsilon := \frac{r_a\ell_s g_{s,t} r_t \ell_b}{\langle \ell, r\rangle^2}. \tag{71}$$

Comparison of (71) and (68) finally yields the result since the relative error terms are $O(n^{-1})$ in each case. □

Another consequence of this last proof and Equation (68) is that the numbers of paths starting at a and ending at either b or c satisfy

$$\lim_{n\to\infty}\frac{P_{a,b,n}}{P_{a,c,n}} = \frac{\ell_b}{\ell_c}. \tag{72}$$

In other words, the quantity

$$\frac{\ell_b}{\sum_j \ell_j}$$

is the probability that a random path with origin fixed at some point a but otherwise unconstrained will end up at point b after n steps. Such properties are strongly evocative of Markov chain theory discussed below in Example 12.

▷ **14.** *Concentration of distribution for the number of passages.* Under the conditions of the theorem, the standard deviation of the number of traversals of a designated node or edge is $O(\sqrt{n})$. Thus in a random long path, the distribution of the number of such traversals is concentrated. [Compared to (70), the calculation of the second moment requires taking a further derivative, which leads to a triple pole. The second moment and the square of the mean, which are each $O(n^2)$, are then found to cancel to main asymptotic order.] ◁

EXAMPLE 11. *Walks on the interval revisited.* As a direct illustration, consider the walks associated to the graph $G(5)$ with vertex set $1,\ldots,5$ and edges being formed of all pairs (i,j) such that $|i-j|\le 1$. The matrix is

$$\mathbf{G}(5) = \begin{pmatrix} 1&1&0&0&0\\ 1&1&1&0&0\\ 0&1&1&1&0\\ 0&0&1&1&1\\ 0&0&0&1&1 \end{pmatrix}.$$

The characteristic polynomial factorizes as

$$\chi_{\mathbf{G}(5)}(z) = z(z-1)(z-2)(z^2-2z-2),$$

and its dominant root is $\lambda_1 = 1+\sqrt{3}$. From there, one finds a left eigenvector (which is also a right eigenvector since the matrix is symmetric):

$$r = \ell^t = (1,\sqrt{3},2,\sqrt{3},1).$$

[6]If A depends on u, one has $\partial_u A^{-1} = A^{-1}(\partial_u A)A^{-1}$, which is a noncommutative generalization of the usual differentiation rule for inverses.

Thus a random path (with the uniform distribution over all paths corresponding to the weights being equal to 1) visits nodes $1, \ldots, 5$ with frequencies proportional to

$$1, \quad 1.732, \quad 2, \quad 1.732, \quad 1,$$

implying that the central nodes are visited more often—such nodes have higher degrees of freedom, hence there tends to be more paths that traverse them.

In fact, this example has structure. For instance, the corresponding problem on an interval of length 11, leads to a matrix with a highly factorable characteristic polynomial

$$\chi_{\mathbf{G}(11)} = z\,(z-1)\,(z-2)\left(z^2-2\,z-2\right)\left(z^2-2\,z-1\right)\left(z^4-4\,z^3+2\,z^2+4\,z-2\right).$$

The reader may have recognized a particular case of lattice paths which resort to to the theory exposed in Section V. 2. For instance, according to Theorem V.2, the OGF of paths from vertex 1 to vertex 1 in the graph with k vertices is given by the continued fraction

$$\cfrac{1}{1-z-\cfrac{z^2}{1-\cfrac{z^2}{\cfrac{\ddots}{1-\cfrac{z^2}{1-z}}}}}.$$

(The number of fraction bars is k; the first and last quotients are $1-z$, the others being equal to 1.) From this it can be shown that the characteristic polynomial of G is an elementary variant of the Fibonacci–Chebyshev polynomial of Example 6. The analysis based on Theorem V.6 is simpler, albeit more rudimentary, as it only provides a first-order asymptotic solution to the problem. □

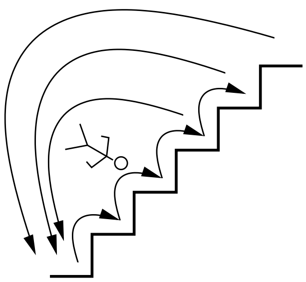
EXAMPLE 12. *Elementary theory of Markov chains.* Consider the case where the row sums of matrix $\mathbf{G}$ are all equal to 1, that is, $\sum_j g_{i,j} = 1$. Such a matrix is called a *stochastic matrix*. The quantity $g_{i,j}$ can then be interpreted as the probability of leaving state i for state j, assuming one is in state i. Assume that the matrix $\mathbf{G}$ is irreducible and aperiodic. Clearly, the matrix $\mathbf{G}$ admits the column vector $r = (1, 1, \ldots, 1)^t$ as an eigenvector corresponding to the dominant eigenvalue $\lambda_1 = 1$. The left eigenvector ℓ normalized so that its elements sum to 1 is called the (row) vector of stationary probabilities. It must be determined by linear algebra and it involves finding an element of the kernel of matrix $I - \mathbf{G}$, which can be done in a standard way.

Application of Theorem V.6 and Equation (68) shows immediately the following:

PROPOSITION V.8 (Stationary probabilities of Markov chains). *Consider a weighted graph corresponding to a stochastic matrix $\mathbf{G}$ which is irreducible and aperiodic. Let ℓ be the normalized left eigenvector corresponding to the eigenvalue 1. A random (weighted) path of length n with fixed origin and destination visits node s a mean number of times asymptotic to $\ell_s n$ and traverses edge (s,t) a mean number of times asymptotic to $\ell_s g_{s,t} n$. A random path of length n with fixed origin ends at vertex s with probability asymptotic to ℓ_s.*

This first-order asymptotic property certainly constitutes the most fundamental result in the theory of finite Markov chains. □

The next example illustrates an elementary technique often employed in calculations of eigenvalues and eigenvectors. It presupposes that the matrix to be analysed can be reduced to a sparse form and has a regular enough structure.



$$\mathbf{G} = \begin{pmatrix} 1&1&0&0&0&0\\ 1&0&1&0&0&0\\ 1&0&0&1&0&0\\ 1&0&0&0&1&0\\ 1&0&0&0&0&1\\ 1&0&0&0&0&0 \end{pmatrix} \qquad \widetilde{\mathbf{G}} = \begin{pmatrix} \frac12&\frac12&0&0&0&0\\ \frac12&0&\frac12&0&0&0\\ \frac12&0&0&\frac12&0&0\\ \frac12&0&0&0&\frac12&0\\ \frac12&0&0&0&0&\frac12\\ 1&0&0&0&0&0 \end{pmatrix}$$

FIGURE 14. The devil's staircase ($m = 6$) and the two matrices that can model it.

EXAMPLE 13. *The devil's staircase.* You live in a house that has a staircase with m steps. You come back home a bit loaded and at each second, you can either succeed in climbing a step or fall back all the way down. On the last step, you always stumble and fall back down (Figure 14). Where are you likely to be found at time n?

Precisely, two slightly different models correspond to this informally stated problem. The probabilistic model views it as a Markov chain with equally likely possibilities at each step and is reflected my matrix $\widetilde{\mathbf{G}}$ in Figure 14. The combinatorial model just assumes all possible evolutions ("histories") of the system as equally likely and it corresponds to matrix $\mathbf{G}$. We opt here for the latter, keeping in mind that the same method basically applies to both cases.

We first write down the constraints expressing the joint properties of an eigenvalue λ and its right eigenvector $x = (x_1, \ldots, x_m)^t$. The equations corresponding to $(\lambda I - \mathbf{G})x = 0$ are formed of a first batch of $m-1$ relations,

$$(\lambda - 1)x_1 - x_2 = 0, \quad -x_1 + \lambda x_2 - x_3 = 0, \quad \cdots, -x_1 + \lambda x_{m_1} - x_m = 0, \tag{73}$$

together with the additional relation (one cannot go higher than the last step):

$$-x_1 + \lambda x_m = 0. \tag{74}$$

The solution to (73) is readily found by pulling out successively $x_2, \ldots, x_m$ as functions of x_1:

$$x_2 = (\lambda - 1)x_1, \quad x_3 = (\lambda^2 - \lambda - 1)x_1, \quad \cdots, \quad x_m = (\lambda^m - \lambda^{m-1} - \cdots - 1)x_1. \tag{75}$$

Combined with the special relation (74), this last relation shows that λ must satisfy the equation

$$1 - 2\lambda^m + \lambda^{m+1}. \tag{76}$$

Let λ_1 be the largest positive root of this equation, existence and dominance beeing guaranteed by Perron-Frobenius properties. Note that the quantity $\rho := 1/\lambda_1$ satisfies the characteristic equation

$$1 - 2\rho + \rho^{m+1} = 0,$$

already encountered when discussing longest runs in words; the discussion of Example 2 then grants us the existence of an isolated ρ near $\frac12$, hence the fact that λ_1 is slightly less than 2.

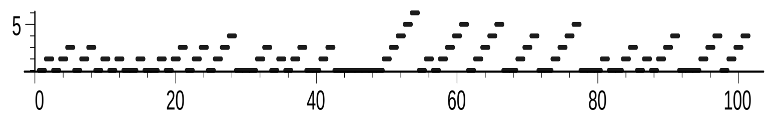
Similar devices yield the left eigenvector $y = (y_1, \ldots, y_m)$. It is found easily that y_j must be proportional to λ_1^{-j}. We thus obtain from Theorem V.6 and Equation (72): *The* *probability of being in state j (i.e., being on step j of the stair) at time n tends to the limit*

$$\varpi_j = \gamma \lambda_1^{-j}$$

where λ_1 is the root near 2 of (76) *and the normalization constant γ is determined by* $\sum_j \varpi_j = 1$. In other words, the distribution of the altitude at time n is a truncated geometric distribution with parameter $1/\lambda_1$. For instance, $m = 6$ leads to $\lambda_1 = 1.98358$, and the asymptotic probabilities of being in states $1, \dots, 6$ are

$$(77) \qquad 0.50413, \quad 0.25415, \quad 0.12812, \quad 0.06459, \quad 0.03256, \quad 0.01641,$$

exhibiting clearly a geometric decay. Here is the simulation of a random history for $n = 100$:



In this case, the frequencies observed are $0.44,\ 0.26,\ 0.17,\ 0.08,\ 0.04,\ 0.01$, pretty much in agreement with what is expected.

Finally, the similarity with the longest run problem is easily explained. Let u and d be letters representing steps upwards and downwards respectively. The set of paths from state 1 to state 1 is described by the regular expression

$$\mathcal{P}_{1,1} = \left(d + ud + \cdots + u^{m-1}d\right)^\star,$$

corresponding to the generating function

$$P_{1,1}(z) = \frac{1}{1 - z - z^2 - \cdots - z^m},$$

of variant of the OGF of words without m-runs of the letter u, which also corresponds to the enumeration of compositions with summands $\leq m$. The case of the probabilistic transition matrix $\widetilde{\mathbf{G}}$ is left as an exercise to the reader. □

This last example is typical: in many cases combinatorial problems have some amount of regularity. In such situations, all the resources of linear algebra are available, including the vast body of knowledge gathered over years on calculations of structured determinants; see for instance Krattenthaler's survey [**125**] and the book [**187**].

V. 5.2. Finite automata. Word problems corresponding to regular languages can be treated by the theory of regular specifications whenever they have enough structure and an unambiguous regular expression description is of tractable form. This was the main theme of Sections V. 1 and V. 2. The dual point of view of automata theory proves useful whenever no such direct description is in sight. Finite automata resorting essentially to the theory of paths in graphs, the results from the previous sections apply with only minor adaptation. For convenience, we start by recalling definitions already given in Chapter I.

DEFINITION V.5. *A* finite automaton A *over a finite alphabet $\mathcal{A}$ is a directed multigraph whose vertex set Q is called the set of states and whose edges are labelled by letters of the alphabet. This graph is equipped with a designated initial state $q_0 \in Q$ and a designated set of final states $Q_f \subseteq Q$.*

A word w is said to be accepted by the automaton if there exists a path π in the graph connecting the initial state q_0 to one of the final states $q \in Q_f$, so that the succession of labels of the path π corresponds to the sequence of letters composing w. The path π is then called an accepting path for w. (We can regard the finitely many states as keeping a patial memory of the past, an interpretation that proves useful in design issues.) The set of accepted words is denoted by $\mathcal{L}(A)$.

In all generality, a finite automaton may be a *nondeterministic* device: given a word w, one might not "know" *a priori* which choices to effect at vertices in order to accept it. A finite automaton is said to be *deterministic* if given any state $q \in Q$ and any letter $x \in \mathcal{A}$, there is at most one edge from vertex q that bears label x. In that case, one decides easily (in linear time) whether a word is accepted by just following edges dictated by the sequence of letters in w. All automata to be used in the examples below are deterministic.

PROPOSITION V.9 (Finite state automata counting). *Any language accepted by a deterministic finite automaton has a rational generating function obtained as follows. If the language is specified by the deterministic automaton $A = \langle Q, Q_f, q_0 \rangle$, then the corresponding ordinary generating function $L_0(z)$ is the component $L_0(z)$ of the linear system of equations*

$$\left\{ L_j(z) = \phi_j + z \sum_{a \in \mathcal{A}} L_{\tau(q_j,a)}(z) \right\},$$

where ϕ_j equals 1 if $q_j \in Q_f$ and 0 otherwise, and where $\tau(q_j, a)$ is the index of the state reachable from state q_j when the letter a is read.

As a consequence, the number of words in the language accepted by a finite-state automaton always admits of an exponential-polynomial form.

Note. The most fundamental result of the theory of regular languages is that there is complete equivalence between three descriptive models: regular expressions, deterministic finite automata, and nondeterministic finite automata. The corresponding theorems are due to Kleene (the equivalence between regular expression and nondeterministic finite automata) and to Rabin and Scott (the equivalence between nondeterministic and deterministic automata). Thus, finite automata whether deterministic or not accept ("recognize") the family of all regular languages.

PROOF. By the fundamental equivalence of models, one may freely assume the automaton to be deterministic. The quantity L_j is nothing but the OGF of the language obtained by changing the initial state of the automaton to q_j. Each equation expresses the fact that a word accepted starting from q_j may be the empty word (if q_j is final) or, else, it must consist of a letter a followed by a continuation that is itself accepted when the automaton is started from the "next" state, that is, the state of index $\tau(q_j, a)$. (Equivalently, one may reduce the proof to the enumeration of paths in graphs as detailed above.)

Existence of the exponential-polynomial form immediately results from rationality of the OGF. □

As implied by the statement of the proposition, the OGF of the language defined by a deterministic finite automaton involves a quasi-inverse $(1 - zT)^{-1}$ where the *transition matrix* T is a direct encoding of the automaton's transitions. Corollary V.2 and Lemma V.1 have been precisely custom-tailored for this situation. As is by now usual, we shall allow weights on letters of the alphabet, corresponding to a Bernoulli model on words. We say that an automaton is irreducible (resp. aperiodic) if the underlying graph and the associated matrix are irreducible (resp. aperiodic).

PROPOSITION V.10 (Random words and automata). *Let $\mathcal{L}$ be a language recognized by a deterministic finite automaton A that is irreducible and aperiodic. The number of words of $\mathcal{L}$ satisfies*

$$L_n \sim c\lambda_1^n \left(1 + O(d^{-n})\right),$$

where λ_1 is the dominant (Perron-Frobenius) eigenvalue of the transition matrix of A and c, d are positive constants with $d > 1$.

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

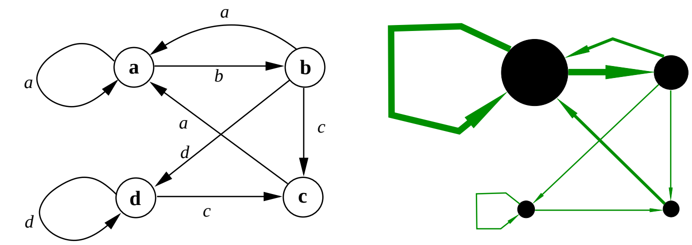


FIGURE 15. Locally constrained words: The transition matrix (T) associated to the forbidden pairs $F = \{ac, ad, bb, cb, cc, cd, da, db\}$, the corresponding automaton, and the graph with widths of vertices and edges drawn in proportion to their asymptotic frequencies.

In a random word of $\mathcal{L}_n$, the number of traversals of a designated vertex or edge has a mean that is asymptotically linear in n and is given by Theorem V.6.

▷ **15.** *Unambiguous automata.* A nondeterministic finite state automaton is said to be unambiguous if the set of accepting paths for any given words comprises at most one element. The translation into generating function as described above also applies to such automata, even though they are nondeterministic. ◁

EXAMPLE 14. *Locally constrained words.* Consider a fixed alphabet $\mathcal{A} = \{a_1, \ldots, a_m\}$ and a set $\mathcal{F} \subseteq \mathcal{A}^2$ of forbidden transitions between consecutive letters. The set of words over $\mathcal{A}$ with no such forbidden transition is denoted by $\mathcal{L}$ and is called a locally constrained language. (The particular case where exactly all pairs of equal letters are forbidden corresponds to Smirnov words and has been discussed on p. 34.)

Clearly, the words of $\mathcal{L}$ are recognized by an automaton whose state space is isomorphic to $\mathcal{A}$: state q simply memorizes the fact that the last letter read was a q. The graph of the automaton is then obtained by the collection of allowed transitions $(q, r) \mapsto a$, with $(q, r) \notin \mathcal{F}$. (In other word, the graph of the automaton is the complete graph in which all edges that correspond to forbidden transitions are deleted.) Consequently, the OGF of any locally constrained language is a rational function. Its OGF is given by

$$(1, 1, \ldots, 1)(I - zT)^{-1}(1, 1, \ldots, 1)^t,$$

where T_{ij} is 0 if $(a_i, a_j) \in \mathcal{F}$ and 1 otherwise. If each letter can follow any other letter in an accepted word, the automaton is irreducible. The graph is aperiodic except in a few degenerate cases (e.g., in the case where the allowed transitions would be $a \to b, c$, $b \to d$, $c \to d$, $d \to a$). Under irreducibility and aperiodicity, the number of words will be $\sim c\lambda_1^{-n}$ and each letter will have on average an asymptotic constant frequency. (See (30) and (31) of Chapter IV for the case of Smirnov words.)

For the example of Figure 15, the alphabet is $\mathcal{A} = \{a, b, c, d\}$. There are eight forbidden transitions and the characteristic polynomial is found to be $\lambda^3(\lambda - 2)$. Thus, one has $\lambda_1 = 2$. The right and left eigenvectors are found to be

$$r = (2, 2, 1, 1)^t, \qquad \ell = (2, 1, 1, 1).$$

Then, the matrix $(\tau_{s,t})$, where $\tau_{s,t}$ represents the asymptotic frequency of transitions from letter s to letter t is found in accordance with Theorem V.6:

$$\Gamma = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{8} & 0 & \frac{1}{16} & \frac{1}{16} \\ \frac{1}{8} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{16} & \frac{1}{16} \end{pmatrix}.$$

This means that a random path spends a proportion equal to $\frac{1}{4}$ of its time on a transition between an a and a b, but much less ($\frac{1}{16}$) on transitions between pairs of letters bc, bd, cc, ca. The letter frequencies in a random word of $\mathcal{L}$ are $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8})$, so that an a is four times more frequent than a c or a d, and so on. See Figure 15 (right) for a rendering.

Various specializations, including multivariate GF's and nonuniform letter models are readily treated by this method. Bertoni *et al.* develop in [**23**] related variance and distribution calculations in the case of the number of occurrences of a symbol in an arbitrary regular language. □

EXAMPLE 15. *De Bruijn graphs.* Two thieves want to break into a house whose entrance is protected by digital lock with an unknown four-digit code. As soon as the four digits of the code are typed consecutively, the gate opens. The first thief proposes to try in order all the four-digit sequences , resulting in as much as 40,000 key strokes in the worst-case. The second thief, who is a mathematician, says he can try *all* four-digit combinations with only 10,003 key strokes. What is the mathematician's trade secret?

Clearly certain optimizations are possible: for instance, for an alphabet of cardinality 2 and codes of 2 letters, the sequence $\mathsf{00110}$ is better than the naïve one, $\mathsf{00\,01\,10\,11}$, which is redundant; a few more attempts will lead to an optimal solution for 3–digit codes that has length 11 (rather than 24), for instance,

$$\mathsf{0001110100}.$$

The general question is then: How far can one go and how to construct such sequences?

Fix an alphabet of cardinality m. A sequence that contains as factors (contiguous blocks) *all* the k letter words is called a *de Bruijn sequence*. Clearly, its length must be at least $\delta(m, k) = m^k + k - 1$, as it must have at least m^k positions at distance at least k from the end. Such a sequence of smallest possible length $\delta(m, k)$ is called a *minimal* de Bruijn sequence. Such sequences were discovered by N. G. de Bruijn [**41**] in 1946, in response to a question coming from electrical engineering, where all possible reactions of a device presented as a black box must be tested at minimal cost. We shall expose here the case of a binary alphabet, $m = 2$, the generalization to $m > 2$ being obvious.

Let $\ell = k-1$ and consider the automaton $\mathcal{B}_\ell$ that memorizes the last block of length ℓ read when scanning the input text from left to right. A state is thus assimilated to a string of length ℓ and the total number of states is 2^ℓ. The transitions are easily calculated: let $q \in \{0,1\}^\ell$ be a state and let $\sigma(w)$ be the function that shifts all letters of a word w one position to the left, dropping the first letter of w in the process (thus σ maps $\{0,1\}^\ell$ to $\{0,1\}^{\ell-1}$); the transitions are

$$q \overset{\mathsf{0}}{\mapsto} \sigma(q)\mathsf{0}, \qquad q \overset{\mathsf{1}}{\mapsto} \sigma(q)\mathsf{1}.$$

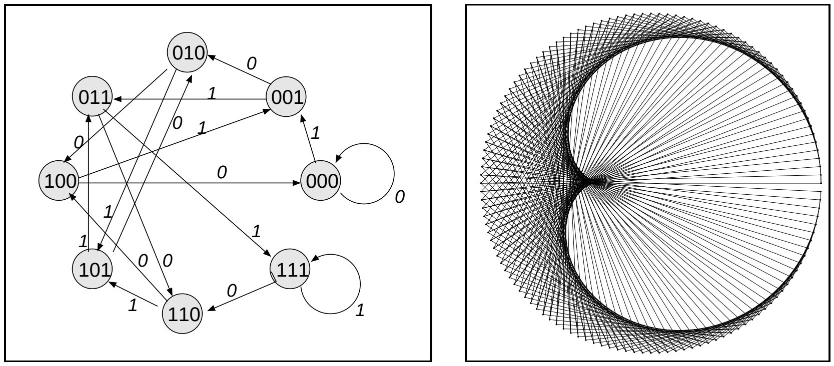


FIGURE 16. The de Bruijn graph: (left) $\ell = 3$; (right) $\ell = 7$.

If one further interprets a state q as the integer in the interval $[0\,..\,2^\ell - 1$ that it represents, then the transition matrix assumes a remarkably simple form:

$$T_{i,j} = \frac{1}{2}[\![(j \equiv 2i \bmod 2^\ell) \text{ or } (j \equiv 2i + 1 \bmod 2^\ell)]\!].$$

See Figure 16 for a rendering borrowed from [**85**].

Combinatorially, the de Bruijn graph is such that each node has indegree 2 and outdegree 2. By a well known theorem going back to Euler: *A necessary and sufficient condition for an undirected connected graph to have an Eulerian circuit (that is, a closed path that traverses each vertex exactly once) is that every node has even degree. For strongly connected digraphs, the condition is that each node should have an outdegree equal to its indegree.* This last condition is obviously satisfied here. Take an Eulerian circuit starting and ending at node 0^ℓ; its length is $2^{\ell+1} = 2^k$. Then, clearly, the sequence of edge labels encountered when prefixed with the word $0^{k-1} = 0^\ell$ constitutes a minimal de Bruijn sequence. In general, the argument gives a de Brujin sequence with minimal length $m^k + k - 1$. Et voilà! The trade secret of the thief-mathematician is exposed.

Back now to enumeration. The de Bruijn matrix is irreducible since a path labelled by sufficiently many zeros always leads any state to the state 0^ℓ, while a path ending with the letters of $w \in \{0,1\}^\ell$ leads to state w. The matrix is aperiodic since it has a loop on states 0^ℓ and 1^ℓ. Thus, by Perron Frobenius properties, it has a unique dominant eigenvalue, and it is not hard to check that its value is $\lambda_1 = 2$, corresponding to the right eigenvector $(1, 1, \ldots, 1)^t$. If one fixes a *pattern* $w \in \{0,1\}^\ell$, Theorem V.6 yields the fact that a random word contains on average $\sim \frac{n}{2^\ell}$ occurrences of pattern w. Note 14 also implies that the distribution of the number of occurrences is concentrated around the mean as the variance is $O(n)$. This gives us in a simple manner a version of what was nicknamed "Borges's Theorem" in Chapter I: *Almost every sufficiently long text contains* all *patterns of some predetermined length* ℓ. As a matter of fact, the de Bruijn graph may be used to quantify many properties of occurrences of patterns in random words, and it has been used for this purpose in several works including [**18, 74, 85**]. □

EXAMPLE 16. *Words with excluded patterns.* Fix a finite set of patterns $\Omega = \{w_1, \ldots, w_r\}$, where each w_j is a word of $\mathcal{A}^\star$. The language $\mathcal{E} \equiv \mathcal{E}^\Omega$ of words that contain no factor in Ω

is described by the extended regular expression

$$\mathcal{E} = \mathcal{A}^\star \setminus \bigcup_{j=1}^{r} (A^\star w_j A^\star),$$

which constitutes a concise but highly ambiguous description. By closure properties of regular languages, $\mathcal{E}$ is itself regular and there must exist a deterministic automaton that recognizes it.

An automaton recognizing $\mathcal{E}$ can be constructed starting from the de Bruijn automaton of index $k = -1 + \max |w_j|$ and deleting all the vertices and edges that correspond to a word of Ω. Precisely, vertex q is deleted whenever q contains a factor in Ω; the transition (edge) from q associated with letter α gets deleted whenever the word $q\alpha$ contains a factor in Ω. The pruned de Bruijn automaton, call it $\mathcal{B}_k^\circ$, accepts all words of $0^k\mathcal{E}$, when it is equiped with the initial state 0^k and all states are final. Thus, the OGF $E(z)$ is in all cases a rational function.

The matrix of $\mathcal{B}_k^\circ$ is the matrix of $\mathcal{B}_k$ with some nonzero entries replaced by 0. Assume that $\mathcal{B}_k^\circ$ is irreducible. This assumption only eliminates a few pathological cases (e.g., $\Omega = \{01\}$ on the alphabet $\{0, 1\}$). Then, the matrix of $\mathcal{B}_k^\circ$ admits a simple Perron-Frobenius eigenvalue λ_1. By domination properties ($\Omega \neq \emptyset$), we must have $\lambda_1 < m$, where m is the cardinality of the alphabet. Aperiodicity is automatically granted. We then get by a purely qualitative argument: *The number of words of length n excluding patterns from the finite set Ω is, under the assumption of irreducibility, asymptotic to $c\lambda_1^n$, for some $c > 0$ and $\lambda_1 < \|\mathcal{A}\|$.* This last result is a strong metric form of Borges' Theorem.

The construction of a pruned automaton is clearly a generalization of the case of words obeying local constraints in Example 14 above. □

▷ **16.** *Words with excluded patterns and digital trees.* Let S be a finite set of words. An automaton recognizing S, considered as a finite language, can be constructed as a tree. The tree obtained is akin to the classical *digital tree* or *trie* that serves as a data structure for maintaining dictionaries [**121**].

A modification of the construction yields an automaton of size linear in the total number of characters that appear in words of S. [Hint. The construction can be based on the Aho–Corasick automaton [**3**]). ◁

▷ **17.** *Words excluding a subsequence.* The language formed of all words that do not contain $w_1 \cdots w_k$ as a subsequence (or "hidden pattern"), except at the very end, is described by the unambiguous regular expression

$$(\mathcal{A} \setminus w_1)^\star \, w_1 \, (\mathcal{A} \setminus w_2)^\star \, w_2 \cdots w_{r-1} \, (\mathcal{A} \setminus w_k)^\star \, w_k \, .$$

Assume the alphabet is endowed with a family of weights, with p_j the weight of letter $a_j \in \mathcal{A}$. The OGF $F(z)$ of words *not* containing w as a subsequence satisfies, with $q_j := 1 - p_j$,

$$F(z) = \sum_{j=1}^{k-1} \frac{(p_1 \cdots p_j) z^j}{(1 - q_1 z) \cdots (1 - q_{j+1} z)},$$

from which an asymptotic formula for $[z^n]F(z)$ derives. E.g., in the equiprobable case ($p_i = 1/m$)

$$[z^n]F(z) \sim \frac{1}{m^n} (m-1)^{n-k+1} \frac{n^{k-1}}{(k-1)!}.$$

(This problem is closely related to the discussion of pure-birth processes on page 63.) ◁

V. 5.3. Transfer matrix methods. The transfer matrix method constitutes a variant of the modelling by deterministic automata and by paths in graphs. The very general statement of Theorem V.5 applies here with full strength. Here, we shall illustrate the situation by the width of trees following an early article by Odlyzko and Wilf [**148**] and continue with an example that draws its inspiration from the insightful exposition of domino tilings and generating functions in the book of Graham, Knuth, and Patashnik [**99**].

EXAMPLE 17. *Width of trees.* The width of a tree is defined as the maximal number of nodes that can appear on any layer at a fixed distance from the root. If a tree is drawn in the plane, then width and height can be seen as the horizontal and vertical diomensions of the bounding rectangle. Also, width is an indicator of the complexity of traversing the tree in a breadth-first search (by a queue), while height is associated to depth-first search (by a stack).

Transfer matrices are ideally suited to the problem of analysing the number of trees of fixed width. Consider a simple variety of trees corresponding to the equation $Y(z) = z\phi(Y(z))$, where the "generator" ϕ describes the formation of trees and let $Y^{[w]}(z)$ be the GF of trees of width at most w. Such trees are easily built layer by layer. Say there are k nodes at a certain level in the tree (with $1 \le k \le w$); the number of possibilities for attaching ℓ levels at the next level is the number of k-forests of depth 1 having ℓ leaves, that is, the quantity

$$t_{k,\ell} = [y^\ell]\phi(y)^k.$$

Let T be the $w \times w$ matrix with entry $T_{k,\ell} = z^\ell t_{k,\ell}$. Then, clearly, the quantity $z^i(T^h)_{i,j}$ (with $1 \le i, j \le w$) is the number of i-forests of height h, width at most w, with j nodes on level h. Thus, the GF of $\mathcal{Y}$-trees having width at most w is

$$Y^{[w]}(z) = (z, 0, 0, \ldots)(I - T)^{-1}(1, 1, 1, \ldots)^t.$$

For instance, in the case of general Catalan trees, the matrix T has the shape,

$$T = \begin{pmatrix} z\binom{1}{0} & z^2\binom{2}{0} & z^3\binom{3}{0} & z^4\binom{4}{0} \\ z\binom{2}{1} & z^2\binom{3}{1} & z^3\binom{4}{1} & z^4\binom{5}{1} \\ z\binom{3}{2} & z^2\binom{4}{2} & z^3\binom{5}{2} & z^4\binom{6}{2} \\ z\binom{4}{3} & z^2\binom{5}{3} & z^3\binom{6}{3} & z^4\binom{7}{3} \end{pmatrix},$$

for width 4. The analysis of dominant poles provides asymptotic formulae for $[z^n]Y^{[w]}(z)$:

$w=2$	$w=3$	$w=4$	$w=5$	$w=6$
$0.0085 \cdot 2.1701^n$	$0.0026 \cdot 2.8050^n$	$0.0012 \cdot 3.1638^n$	$0.0006 \cdot 3.3829^n$	$0.0004 \cdot 3.5259^n$

Additionally, the exact distribution of height in trees of size n becomes computable in polynomial time (though with a somewhat high degree polynomial).

The character of these generating functions has not been investigated in detail since the original work [**148**], so that, at the moment, analysis stops there. Fortunately, probability theory can take over the problem. Chassaing and Marckert [**35**] have shown, for Cayley trees, that the width satisfies

$$\mathbb{E}_n(W) = \sqrt{\frac{\pi}{2}} + O\left(n^{1/4}\sqrt{\log n}\right), \qquad \mathbb{P}_n(\sqrt{2}W \le x) \to 1 - \Theta(x),$$

where $\Theta(x)$ is the Theta function defined in (44). This answers very precisely an open question of Odlyzko and Wilf [**148**]. The distributional results of [**35**] extend to trees in any simple variety (under mild and natural analytic assumptions on the generator ϕ): see the paper by Chassaing, Marckert, and Yor [**36**], which builds upon earlier results of Drmota and Gittenberger [**51**]. In essence, the conclusion of these works is that the breadth first search traversal of a large tree in a simple variety gives rise to a queue whose size fluctuates asymptotically like a Brownian excursion, and is thus, in a strong sense, of a complexity comparable to depth-first search: trees don't have a preference as to the way they may be traversed. □

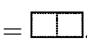
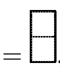
▷ **18.** *A question on width polynomials.* It is unknown whether the following assertion is true. The smallest positive root ρ_k of the denominator of $Y^{[k]}(z)$ satisfies

$$\rho_k = \rho + \frac{c}{k^2} + o(k^{-2}),$$

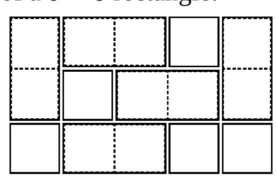
for some $c > 0$. If such an estimate holds together with suitable companion bounds, it would yield a purely analytic proof of the fact that expected width of n–trees is $\Theta(\sqrt{n})$, as well as detailed probability estimates. (The classical theory of Fredholm equations may be useful here.) ◁

EXAMPLE 18. *Monomer-dimer tilings of a rectangle.* Suppose one is given pieces that may be one of the three forms: monomers (m) that are 1×1 squares, and dimers that are dominoes, either vertically (v) oriented 1×2, or horizontally (h) oriented 2×1. In how many ways can an $n \times 3$ rectangle be covered completely and without overlap ('tiled') by such pieces?

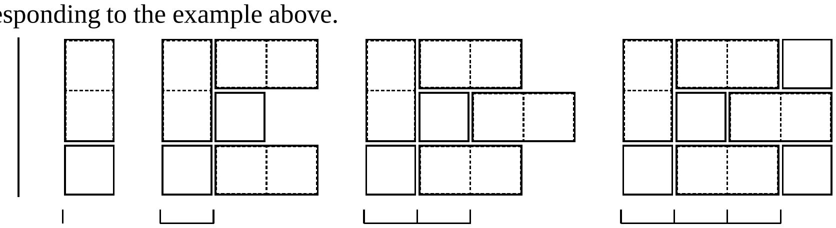
The pieces are thus of the following types,

$m =$ , $h =$ , $v =$ ,

and here is a particular tiling of a 5×3 rectangle:



In order to approach this counting problem, one defines a class $\mathcal{C}$ of combinatorial objects called configurations. A configuration relative to an $n \times k$ rectangle is a partial tiling, such that all the first $n - 1$ columns are entirely covered by dominoes while between zero and three unit cells of the last column are covered. Here are for instance, configurations corresponding to the example above.

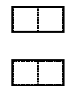


These diagrams suggest the way configurations can be built by successive addition of dominoes. Starting with the empty rectangle 0×3, one adds at each stage a collection of at most three dominoes in such a way that there is no overlap. This creates a configuration where, like in the example above, the dominoes may not be aligned in a flush-right manner. Continue to add successively dominoes whose left border is at abscissa $1, 2, 3$, etc, in a way that creates no internal "holes".

Depending on the state of filling of their last column, configuration can thus be classified into 8 classes that we may index in binary as $\mathcal{C}_{000}, \ldots, \mathcal{C}_{111}$. For instance $\mathcal{C}_{001}$

represent configurations such that the first two cells (from top to bottom, by convention) are free, while the third one is occupied. Then, a set of rules describes the new type of configuration obtained, when the sweep line is moved one position to the right and dominoes are added. For instance, we have

$$\mathcal{C}_{010} \qquad \odot$$



$$\Longrightarrow \qquad \mathcal{C}_{101}.$$

In this way, one can set up a grammar (resembling a deterministic finite automaton) that expresses all the possible constructions of longer rectangles from shorter ones according to the last layer added. The grammar comprises productions like

$$\begin{aligned} C_{000} \quad = \quad & \epsilon + \underline{mmm}C_{000} + \underline{mv}C_{000} + \underline{vm}C_{000} \\ & + \underline{\cdot mm}C_{100} + \underline{m\cdot m}C_{010} + \underline{mm\cdot}C_{001} + \underline{v\cdot}C_{001} + \underline{\cdot v}C_{100} \\ & + \underline{m\cdot\cdot}C_{011} + \underline{\cdot m\cdot}C_{101} + \underline{\cdot\cdot m}C_{110} + \underline{\cdot\cdot\cdot}C_{111}\ . \end{aligned}$$

In this grammar, a "letter" like $\underline{mv}$ represent the addition of dominoes, in top to bottom order, of types m, v respectively; the letter $\underline{m\cdot m}$ means adding two m-dominoes on the top and on the bottom, etc.

The grammar transforms into a linear system of equations with polynomial coefficients. The substitution $m \mapsto z$, $h, v \mapsto z^2$ then gives the generating functions of configurations with z marking the area covered:

$$C_{000}(z) = \frac{(1-2z^3-z^6)(1+z^3-z^6)}{(1+z^3)(1-5z^3-9z^6+9z^9+z^{12}-z^{15})}.$$

In particular, the coefficient $[z^{3n}]C_{000}(z)$ is the number of tilings of an $n \times 3$ rectangle:

$$C_{000}(z) = 1 + 3z^3 + 22z^6 + 131z^9 + 823z^{12} + 5096z^{15} + \cdots\ .$$

The sequence grows like $c\,\alpha^n$ (for $n \equiv 0 \pmod 3$) where $\alpha \doteq 1.83828$ (α is the cube root of an algebraic number of degree 5). (See [**34**] for a computer algebra session.) On average, for large n, there is a fixed proportion of monomers and the distribution of monomers in a random tiling of a large rectangle is asymptotically normally distributed, as results from the developments of Chapter IX. □

As is typical of the tiling example, one seeks to enumerate a "special" set of configurations $\mathcal{C}_f$. (In the example above, this is $\mathcal{C}_{000}$ representing complete rectangle coverings.) One determines an extended set of configurations $\mathcal{C}$ (the partial coverings, in the example) such that: (i) $\mathcal{C}$ is partitioned into finitely many classes; (ii) there is a finite set of "actions" that operate on the classes; (iii) size is affected in a well-defined additive way by the actions. The similarity with finite automata is apparent: classes play the rôle of states and actions the rôle of letters.

Often, the method of transfer matrices is used to approximate a hard combinatorial problem that is not known to decompose, the approximation being by means of a family of models of increasing "widths". For instance, the enumeration of the number T_n of tilings of an $n \times n$ square by monomers and dimers remains a famous unsolved problem of statistical physics. Here, transfer matrix methods may be used to solve the $n \times w$ version of the monomer–dimer coverings, in principle at least, for any fixed width w: the result will always be a rational function, though its degree, dicated by the dimension of the transfer matrix, will grow exponentially with w. (The "diagonal" sequence of the $n \times w$

rectangular models corresponds to the square model.) It has been at least determined by computer search that the diagonal sequence T_n starts as (this is *EIS* **A028420**):

$$1,\ 7,\ 131,\ 10012,\ 2810694,\ 2989126727,\ 11945257052321, \ldots.$$

From this and other numerical data, one estimates numerically that $(T_n)^{1/n^2} \to 1.94021\ldots$, but no expression for the constant is known to exist. The difficulty of coping with the finite-width models is that their complexity (as measured , e.g., by the number of states) blows up exponentially with w—such models are best treated by computer algebra; see [**196**]—and no law allowing to take a diagonal is visible. However, the finite width models have the merit of providing at least provable upper and lower bounds on the exponential growth rate of the hard "diagonal problem".

In contrast, for coverings by dimers only, a strong algebraic structure is available and the number of covers of an $n \times n$ square by horizontal and vertical dimers satisfies (n even) beautiful formula originally discovered by Kasteleyn:

$$U_{2n} = 2^{n^2/2} \prod_{j=1}^{n} \prod_{k=1}^{n} \left(\cos^2 \frac{j\pi}{n+1} + \cos^2 \frac{k\pi}{n+1} \right). \tag{78}$$

This sequence is *EIS* **A004003**,

$$1,\ 2,\ 36,\ 6728,\ 12988816,\ 258584046368,\ 53060477521960000,\ \ldots.$$

It is elementary to prove from (78) that

$$\lim_{n\to+\infty} \left(\widehat{U}_{2n} \right)^{1/(2n)^2} = \exp\left(\frac{1}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)^2} \right) = e^{G/\pi} \doteq 1.33851\ldots,$$

where G is Catalan's constant. This means in substance that each cell has a number of degrees of freedoms equivalent to 1.33851. See Percus' monograph [**156**] for proofs of this famous result and Finch's book [**61**, Sec. 5.23] for context and references.

$\triangleright$ **19.** *Powers of Fibonacci numbers.* Consider the OGFs

$$G(z) := \frac{1}{1-z-z^2} = \sum_{n\geq 0} F_{n+1} z^n, \qquad G^{[k]}(z) := \sum_{n \geq 0} (F_{n+1})^k z^n,$$

where F_n is a Fibonacci number. The OGF of monomer–dimer placements on a $k \times n$ board when only monomers (m) and horizontal dimers (h) are allowed is obviously $G^{[k]}(z)$. On the other hand, it is possible to set up a transfer matrix model with state i ($0 \leq i \leq k$) corresponding to i positions of the current column occupied by a previous domino. Consequently,

$$G^{[k]}(z) = \text{coeff}_{k,k}(I - zT)^{-1}, \qquad \text{where} \quad T_{i,j} = \binom{i}{i+j-k},$$

for $0 \leq i, j \leq k$. [The denominator of $G^{[k]}(z)$ is otherwise known exactly [**120**, Ex. 1.2.8.30].] $\triangleleft$

$\triangleright$ **20.** *Tours on chessboards.* The OGF of Hamiltonian tours on an $n \times w$ rectangle is rational (one is allowed to move from any cell to any other vertically or horizontally adjacent cell). The same holds for king's tours and knight's tours. $\triangleleft$

$\triangleright$ **21.** *Cover time of graphs.* Given a fixed digraph G assumed to be strongly connected, and a designated start vertex, one travels at random, moving at each time to any neighbour of the current vertex, making choices with equal likelihood. The expectation of the time to visit all the vertices is a rational number that is effectively (though perhaps not efficiently!) computable. [Hint: set up a transfer matrix, a state of which is a subset of vertices representing those vertices that have been already visited. For an interval $[0,\,..\,m]$, this can be treated by the dedicated theory of walks on the integer interval, as in Section V. 2; for the complete graph, this is equivalent to the coupon collector

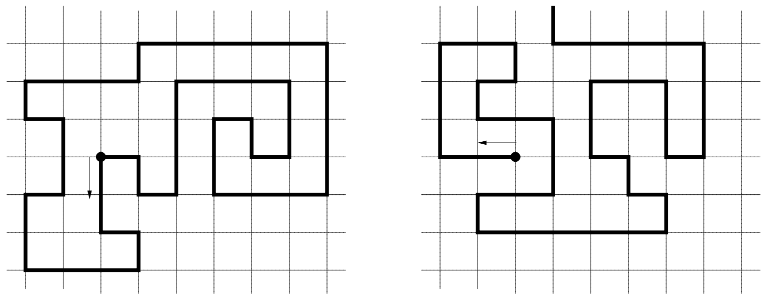
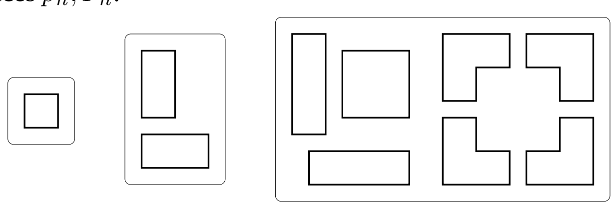


FIGURE 17. A self-avoiding polygon or SAP (left) and a self-avoiding walk or SAW (right).

problem. Most other cases are "hard" to solve analytically and one has to resort to probabilistic approximations; see Aldous and Fill's forthcoming book [**4**].] ◁

EXAMPLE 19. *Self-avoiding walks and polygons.* A long standing open problem shared by statistical physics, combinatorics, and probability theory alike is that of quantifying properties of self-avoiding configurations on the square lattice (Figure 17). Here we consider objects that, starting from the origin (the "root") follow a path, and are solely composed of horizontal and vertical steps of amplitude ± 1. The *self-avoiding walk* or *SAW* can wander but is subject to the condition that it never crosses nor touches itself. The *self-avoiding polygons* or *SAPs*, whose class is denoted by $\mathcal{P}$, are self-avoiding walks, with only an exception at the end, where the end-point must coincide with the origin. We shall focus here on polygons. It proves convenient also to consider *unrooted polygons* (also called simply-connected *polyominoes*), which are polygons where the origin is discarded, so that they plainly represent the possible shapes of SAPs up to translation. For length $2n$, the number p_n of unrooted polygons satisfies $p_n = P_n/(4n)$ since the origin ($2n$ possibilities) and the starting vertex (2 possibilities) of the corresponding SAPs are disregarded in that case. Here is a table, for small values of n, listing polyominoes and the corresponding counting sequences p_n, P_n.



n:	2	3	4	5	6	7	8	9	10
p_n (*EIS* **A002931**):	1	2	7	28	124	588	2938	15268	81826
P_n (*EIS* **A010566**):	8	24	112	560	2976	16464	94016	549648	3273040

Take the (widely open) problem of determining the number P_n of SAPs of perimeter $2n$. This (intractable) problem can be approached as a limit of the (tractable) problem[7] that consists in enumerating the collection $\mathcal{P}^{[w]}$ of SAPs of width w, for increasing values of w. The latter problem is amenable to the transfer matrix method, as first discovered by Entig in 1980; see [**54**]. Indeed, take a polygon and consider a sweepline that moves from

[7]In this version of the text, we limit ourselves to a succinct description and refer to the original papers [**54, 111**] for details.

its left to its right. Once width is fixed, there are at most 2^{2w+2} possibilities for the ways a vertical sweepline may intersect the polygon's edges at half integer abscissæ. (There are $w+1$ edges and for each of these, one should "remember" whether they connect with the upper or lower boundary.) The transitions are then themselves finitely described. In this way, it becomes possible to set up a transfer matrix for any fixed width w. For fixed n, by computing values of $P_n^{[w]}$ with increasing w, one finally determines (in principle) the exact value of any P_n.

The program suggested above has been carried out to record values by the "Melbourne School" under the impulse of Tony Guttmann. For instance, Jensen [**111**] found in 2003 that the number of unrooted polygons of perimeter 100 is

$$p_{50} = 7545649677448506970646886033356862162.$$

Attaining such record values necessitates algorithms that are much more sophisticated than the naïve approach we have just described, as well as a number of highly ingenious programming optimizations.

It is an equally open problem to estimate asymptotically the number of SAPs of perimeter n. Given the exact values till perimeter 100 or more, a battery of fitting tests for asymptotic formula can be applied, leading to highly *convincing* (though still heuristic) formulæ. Thanks to several workers in this area, we can regard the final answer as "known". From the works of Jensen and his predecessors, it results that a reliable empirical estimate is of the form

$$\begin{cases} p_n = B\mu^{2n}(2n)^{-\beta}(1+o(1)), \\ \mu \doteq 2.63815\,85303, \qquad \beta = -\dfrac{5}{2} \pm 3\cdot 10^{-7}, \qquad B \doteq 0.5623013. \end{cases}$$

Thus, the answer is almost certainly of the form $p_n \asymp \mu^{2n} n^{-5/2}$ for unrooted polygons and $P_n \asymp \mu^{2n} n^{-3/2}$ for rooted polygons. It is believed that the same connective constant μ dictates the exponential growth rate of self-avoiding walks. See Finch's book [**61**, Sec. 5.10] for a perspective and numerous references.

There is also great interest in the number $p_{m,n}$ of polyominos with perimeter $2n$ and area m, with area defined as the number of square cells composing the polyomino. Studies conducted by the Melbourne school yield numerical data that are consistent to an amazing degree (e.g., moments till order ten and small–n corrections are considered) with the following assumption: *The distribution of area in a fixed-perimeter polyomino obeys in the asymptotic limit an "Airy area distribution"*. This distribution is defined as the limit distribution of the area under Dyck paths, a problem that was briefly discussed on p. 76 and to which we propose to return in Chapter VII. See [**111, 162**] and references therein for a discussion of polyomino area. It is finally of great interest to note that the interpretation of data was strongly guided by what is already known for exactly solvable models of the type we are repeatedly considering in this book. □

V. 6. Additional constructions

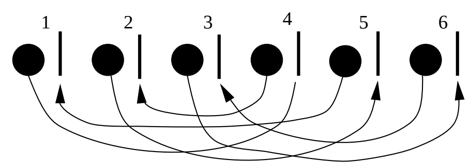
We conclude this chapter with a discussion[8] of a construction that builds on top of rational functions by means of certain transformations. Specifically, it is possible to enumerate constrained permutations by making use of the transfer matrix (or finite automaton) framework.

[8] Contents of this section are supplementary material that can be omitted on first reading.

We examine here problems whose origin lies in nineteenth century recreational mathematics. For instance, the *ménage* problem solved and popularized by Édouard Lucas in 1891, see [**39**], has the following quaint formulation: *What is the number of possible ways one can arrange n married couples (*‘ménages’*) around a table in such a way that men and women alternate, but no woman sits next to her husband?*

The ménage problem is equivalent to a permutation enumeration problem. Sit first conventionally the men at places numbered $0, \dots, n-1$, and let σ_i be the position at the right of which the ith wife is placed. Then, a ménage placement imposes the condition $\sigma_i \neq i$ and $\sigma_i \neq i+1$ for each i. We consider here a linearly arranged table (see remarks at the end for the other classical formulation that considers a round table), so that the condition $\sigma_i \neq i+1$ becomes vacuous when $i = n$. Here is a ménage placement for $n = 6$ corresponding to the permutation

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 2 & 1 & 3 \end{bmatrix}$$



Clearly, this is a generalization of the derangement problem (for which the weaker condition $\sigma_i \neq i$ is imposed), where the cycle decomposition of permutations suffices to provide a direct solution (see Chapter 2).

Given a permutation $\sigma = \sigma_1 \cdots \sigma_n$, any quantity $\sigma_i - i$ is called an *exceedance* of σ. Let Ω be a finite set of integers that we assume to be nonnegative. Then a permutation is said to be Ω-avoiding if none of its exceedances lies in Ω. The counting problem, as we now demonstrate, provides an interesting case of application of the transfer matrix method.

The set Ω being fixed, consider first for all j the class of augmented permutations $\mathcal{P}_{n,j}$ that are permutations of size n such that j of the positions are distinguished and the corresponding exceedances lie in Ω, the remaining positions having arbitrary values (but with the permutation property being satisfied!). Loosely speaking, the objects in $\mathcal{P}_{n,j}$ can be regarded as permutations with “at least” j exceedances in Ω. For instance, with $\Omega = \{1\}$ and

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 8 & 6 & 7 & 1 & 5 & 9 \end{pmatrix},$$

there are 5 exceedances that lie in Ω (at positions $1, 2, 3, 5, 6$) and with 3 of these distinguished (say by enclosing them in a box), one obtains an element counted by $\mathcal{P}_{9,3}$ like

$$2\ \boxed{3}\ \boxed{4}\ 8\ 6\ \boxed{7}\ 1\ 5\ 9.$$

Let $P_{n,j}$ be the cardinality of $\mathcal{P}_{n,j}$. We claim that the number $Q_n = Q_n^{\Omega}$ of Ω-avoiding permutations of size n satisfies

$$Q_n = \sum_{j=0}^{n} (-1)^j P_{n,j}. \tag{79}$$

Equation (79) is typically an *inclusion-exclusion* relation. To prove it formally, define the number $R_{n,k}$ of permutations that have exactly k exceedances in Ω and the generating

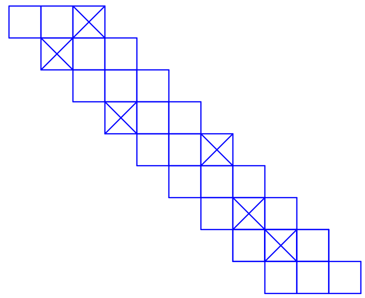


FIGURE 18. A graphical rendering of the legal template **20?02?11?** relative to $\Omega = \{0, 1, 2\}$.

polynomials

$$P_n(w) = \sum_j P_{n,j} w^j, \qquad R_n(w) = \sum_k R_{n,k} w^k.$$

The GF's are related by

$$P_n(w) = R_n(w+1) \qquad \text{or} \qquad R_n(w) = P_n(w-1)..$$

(The relation $P_n(w) = R_n(w+1)$ simply expresses symbolically the fact that each Ω-exceedance in $\mathcal{R}$ may or may not be taken in when composing an element of $\mathcal{P}$.) In particular, we have $P_n(-1) = R_n(0) = R_{n,0} = Q_n$ as was to be proved.

The preceding discussion shows that everything relies on the enumeration $P_{n,j}$ of permutations with distinguished exceedances in Ω. Introduce the alphabet $\mathcal{A} = \Omega \cup \{\text{'?'}\}$, where the symbol '?' is called the 'don't-care symbol'. A word on $\mathcal{A}$, an instance with $\Omega = \{0, 1, 2\}$ being 20?02?11?, is called a *template*. To an augmented permutation, one associates a template as follows: each exceedance that is not distinguished is represented by a don't care symbol; each distinguished exceedance (thereby an exceedance with value in Ω) is represented by its value. A template is said to be legal if it arises from an augmented permutation. For instance a template $2\,1\cdots$ cannot be legal since the corresponding constraints, namely $\sigma_1 - 1 = 2$, $\sigma_2 - 2 = 1$, are incompatible with the permutation structure (one should have $\sigma_1 = \sigma_2 = 3$). In contrast, the template 20?02?11? is seen to be legal. Figure 18 is a graphical rendering; there, letters of templates are represented by dominoes, with a cross at the position of a numeric value in Ω, and with the domino being blank in the case of a don't-care symbol.

Let $T_{n,j}$ be the set of legal templates relative to Ω that have length n and comprise j don't care symbols. Any such legal template is associated to exactly $j!$ permutations, since $n-j$ position-value pairs are fixed in the permutation, while the j remaining positions and values can be taken arbitrarily. There results that

$$P_{n,n-j} = j!\,T_{n,j} \qquad \text{and} \qquad Q_n = \sum_{j=0}^{n} (-1)^{n-j} j!\,T_{n,j}, \tag{80}$$

by (79). Thus, the enumeration of avoiding permutations rests entirely on the enumeration of legal templates.

The enumeration of legal templates is finally effected by means of a transfer matrix method, or equivalently, by a finite automaton. If a template $\tau = \tau_1 \cdots \tau_n$ is legal, then the following condition is met,

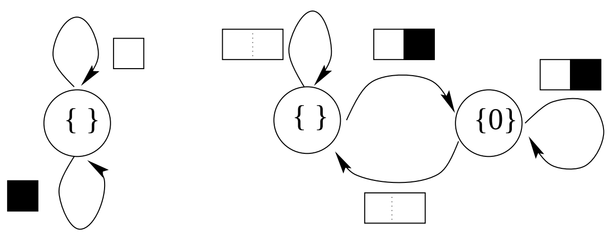
$$\tau_j + j \neq \tau_i + i, \tag{81}$$

for all pairs (i, j) such that $i < j$ and neither of τ_i, τ_j is the don't-care symbol. (There are additional conditions to characterize templates fully, but these only concern a few letters at the end of templates and we may ignore them in this discussion.) In other words, a τ_i with a numerical value preempts the value $\tau_i + i$. Figure 18 exemplifies the situation in the case $\Omega = \{0, 1, 2\}$. The dominoes are shifted one position each time (since it is the value of $\sigma - i$ that is represented) and the compatibility constraint (81) is that no two crosses should be vertically aligned. More precisely the constraints (81) are recognized by a deterministic finite automaton whose states are indexed by subsets of $\{0, \ldots, b-1\}$ where the "span" b is defined as $b = \max_{\omega\in\Omega} \omega$. The initial state is the one associated with the empty set (no constraint is present initially), the transitions are of the form

$$\begin{cases} (q_S, j) \ \mapsto \ q_{S'} & \text{where } S' = ((S-1) \cup \{j-1\}) \cap \{0, \ldots, b-1\},\ j \neq \text{`?'} \\ (q_S, ?) \ \mapsto \ q_{S'} & \text{where } S' = (S-1) \cap \{0, \ldots, b-1\}; \end{cases}$$

the final state is equal to the initial state (this translates the fact that no domino can protrude from the right, and is implied by the linear character of the ménage problem under consideration). In essence, the automaton only needs a finite memory since the dominoes slide along the diagonal and, accordingly, constraints older than the span can be forgotten. Notice that the complexity of the automaton, as measured by its number of states, is 2^b.

Here are the automata corresponding to $\Omega = \{0\}$ (derangements) and to $\Omega = \{0, 1\}$ (ménages).



For the ménage problem, there are two states depending on whether or not the currently examined value has been preempted at the preceding step.

From the automaton construction, the bivariate GF $T^{\Omega}(z, u)$ of legal templates, with u marking the position of don't care symbols, is a rational function that can be determined in an automatic fashion from Ω. For the derangement and ménage problems, one finds

$$T^{\{0\}}(z, u) = \frac{1}{1 - z(1+u)}, \qquad T^{\{0,1\}}(z, u) = \frac{1-z}{1 - z(2+u) + z^2}.$$

In general, this gives access to the OGF of the corresponding permutations. Consider the partial expansion of $T^{\Omega}(z, u)$ with respect to u, taken under the form

$$T^{\Omega}(z, u) = \sum_r \frac{c_r(z)}{1 - u u_r(z)}, \tag{82}$$

assuming for convenience only simple poles. There the sum is finite and it involves algebraic functions c_j and u_j of the variable z. Finally, the OGF of Ω-avoiding permutations is obtained from T^{Ω} by the transformation

$$z^n u^k \quad \mapsto \quad (-z)^n k!,$$

which is the transcription of (80). Define the (divergent) OGF of all permutations,

$$F(y) = \sum_{n=0}^{\infty} n!\, y^n = {}_2F_0[1,1;y],$$

in the terminology of hypergeometric functions. Then, by the remarks above and (82), we find

$$Q^{\Omega}(z) = \sum_r c_r(-z)F(-u_j(-z)).$$

In other words, *the OGF of Ω-avoiding permutations is a combination of compositions of the OGF of the factorial series with algebraic functions.*

The expressions simplify much in the case of ménages and derangements where the denominators of T are of degree 1 in u. One has

$$Q^{\{0\}}(z) = \frac{1}{1+z}F(\frac{z}{1+z}) = 1 + z^2 + 2z^3 + 9z^4 + 44z^5 + 265z^6 + 1854z^7 + \cdots,$$

for derangements, whence a new derivation of the known formula,

$$Q_n^{\{0\}} = \sum_{k=0}^{n} (-1)^k \binom{n}{k}(n-k)!.$$

Similarly, for (linear) ménage placements, one finds

$$Q^{\{0,1\}}(z) = \frac{1}{1+z}F(\frac{z}{(1+z)^2}) = 1 + z^3 + 3z^4 + 16z^5 + 96z^6 + 675z^7 + \cdots,$$

which is *EIS* **A00027** and corresponds to the formula

$$\mathbb{Q}_n^{\{0,1\}} = \sum_{k=0}^{n} (-1)^k \binom{2n-k}{k}(n-k)!.$$

Finally, the same techniques adapts to constraints that "wrap around", that is, constraints taken modulo n. (This corresponds to a round table in the ménage problem.) In that case, what should be considered is the loops in the automaton recognizing templates (see also the previous discussion of the zeta function of graphs). One finds in this way the OGF of the circular (i.e., classical) ménage problem to be *EIS* **A000179**,

$$\widehat{Q}^{\{0,1\}}(z) = \frac{1-z}{1+z}F(\frac{z}{(1+z)^2}) + 2z = 1 + z + z^3 + 2z^4 + 13z^5 + 80z^6 + 579z^7 + \cdots,$$

which yields the classical solution of the (circular) ménage problem,

$$\widehat{Q}_n^{\{0,1\}} = \sum_{k=0}^{n} (-1)^k \frac{2n}{2n-k}\binom{2n-k}{k}(n-k)!,$$

a formula that is due to Touchard; see [**39**, p. 185] for pointers to the vast classical literature on the subject. The algebraic part of the treatment above is close to the inspiring discussion offered in Stanley's book [**175**]. An application to robustness of interconnections in random graphs is presented in [**73**].

For asymptotic analysis purposes, the following general property proves useful: *Let F be the OGF of factorial numbers and assume that $y(z)$ is analytic at the origin where it satisfies $y(z) = z - \lambda z^2 + O(z^3)$; then it is true that*

$$[z^n]F(y(z)) \sim [z^n]F(z(1-\lambda z)) \sim n!e^{-\lambda}. \tag{83}$$

(The proof results from simple manipulations of divergent series in the style of [**17**].) This gives at sight the estimates

$$Q_n^{\{0\}} \sim ne^{-1}, \qquad Q_n^{\{0,1\}} \sim ne^{-2}.$$

More generally, for any set Ω containing λ elements, one has

$$Q_n^{\{\Omega\}} \sim ne^{-\lambda}.$$

Furthermore, the number $R_{n,k}^{\Omega}$ of permutations having exactly k occurrences (k fixed) of an exceedance in Ω is asymptotic to

$$Q_n^{\{\Omega\}} \sim ne^{-\lambda}\frac{\lambda^k}{k!}.$$

In other words, the rare event that an exceedance belongs to Ω obeys of Poisson distribution with $\lambda = |\Omega|$. These last two results are established by means of probabilistic techniques in the book [**14**, Sec. 4.3]. The relation (83) points to a way of arriving at such estimates by purely analytic-combinatorial techniques.

$\triangleright$ **22.** *Other constrained permutations.* Given a permutation $\sigma = \sigma_1 \cdots \sigma_n$, a *succession gap* is defined as any difference $\sigma_{i+1} - \sigma_i$. Discuss the counting of permutations whose succession gaps are constrained to lie outside of a finite set Ω. In how many ways can a kangaroo pass through all points of the integer interval $[1, n]$ starting at 1 and ending at n while making hops that belong to $\{-2, -1, 1, 2\}$? $\triangleleft$

$\triangleright$ **23.** *Shuffle products.* Let $\mathcal{L}, \mathcal{M}$ be two languages over two disjoint alphabets. Then, the shuffle product $\mathcal{S}$ of $\mathcal{L}$ and $\mathcal{M}$ is such that $\widehat{S}(z) = \widehat{L}(z) \cdot \widehat{M}(z)$, where $\widehat{S}, \widehat{L}, \widehat{M}$ are the exponential generating functions of $\mathcal{S}, \mathcal{L}, \mathcal{M}$. Accordingly, if the OGF $L(z)$ and $M(z)$ are rational then the OGF $S(z)$ is also rational. [This technique may be used to analyse generalized birthday paradox and coupon collector problems; see [**69**].] $\triangleleft$

V. 7. Notes

Applications of rational functions in discrete and continuosu mathematics are in abundance. Many examples are to be found in Goulden and Jackson's book [**96**]. Stanley [**175**] even devotes a full chapter of his book *Enumerative Combinatorics*, vol. I, to rational generating functions. These two books push the theory further than we can do here, but the corresponding asymptotic aspects which we expose lie outside of their scope. The analytic theory of positive rational functions starts with the works of Perron and Frobenius at the beginning of the twentieth century and is explained in books on matrix theory likes those of Bellman [**15**] and Gantmacher [**89**]. Its importance has been long recognized in the theory of finite Markov chains, so that the basic theory of positive matrices is well developed in many elementary treatises on probability theory. For such aspects, we refer for instance to the classic presentations by Feller [**58**] or Karlin and Taylor [**113**].

The supercritical sequence schema is the first in a list of abstract schemas that neatly exemplify the interplay between combinatorial, analytic, and probabilistic properties of large random structures. The origins of this approach are to be traced to early works of Bender [**16, 17**] followed by Soria and Flajolet [**83, 84, 172**].

Turning to more specific topics, we mention in relation to Section V. 2 the first global attempt at a combinatorial theory of continued fractions by Flajolet in [**62**] together with related works of Jackson of which an exposition is to be found in [**96**, Ch. 5] and a summary in [**72**] in relation to birth and death processes. Walks on graphs are well discussed in Godsil's book [**94**]. The discussion of local constraints in permutations based on [**73**] combines the combinatorial elements bound in Stanley's book [**175**] with the general philosophy of analytic combinatorics. Our treatment of words and languages largely draws its inspiration from the line of research started by Schützenberger in the early 1960's and on the subsequent account to be found in Lothaire's book [**132**].

There are many topics that would naturally fit into this chapter but weren't ready for the present edition. Amongst the ones that may be treated (briefly) in future editions, we mention: exactly solvable models of convex polygons, the Ehrenfest urn model, random walks on undirected graphs, shuffles and Laplace transforms, variations on cycles in graphs, digital trees and the Aho-Corasick construction, the Goulden-Jackson cluster method. Future editions will be available from Philippe Flajolet's web page.

CHAPTER VI

Singularity Analysis of Generating Functions

Es ist eine Tatsache, daß die genauere Kenntnis
des Verhaltens einer analytischen Funktion
in der Nähe ihrer singulären Stellen
eine Quelle von arithmetischen Sätzen ist.

— ERICH HECKE [**106**, Kap. VIII]

A function's singularities are reflected in the function's coefficients. For rational fractions and meromorphic functions, the local analysis of polar singularities provides contributions to coefficients in the form of exponential polynomials, that is, products of polynomials and simple exponentials. In this chapter, we present a general approach to singularity analysis of generating functions that is no longer restricted to polar singularities and extends to a very large class of functions that have moderate growth or decay at their dominant singularities. The basic principle is the existence of a *correspondence* between

the asymptotic expansion of a function near its dominant singularities
and
the asymptotic expansion of the function's coefficients.

This mapping essentially preserves orders of growth in the sense that larger functions have larger coefficients.

Precisely, the method of singularity analysis applies to "*algebraic– logarithmic*" functions whose singular expansions involve fractional powers and logarithms. It relies on two types of results: first, it is possible to set up a *catalogue* of asymptotic expansions for coefficients of standard functions occurring in such singular expansions second, *transfer theorems* allow us to extract the asymptotic order of coefficients of error terms from singular expansions and error terms.

The developments are based on Cauchy's coefficient formula used in conjunction with special contours of integration known as *Hankel contours*. The contours come very close to the singularities then steer away; by design, they have the property of capturing essential asymptotic informations contained in the functions' singularities.

An important feature of the method is to require only local asymptotic properties of the function to be analysed. In this way, it is often instrumental in the case of functions only indirectly accessible through functional equations. In particular, the method of singularity analysis allows us to treat models where singularities of the square–root type occur, which is invariably the case for simple tree types. It also applies to search trees of various kinds as well as to several searching and sorting algorithms whose analyses (see later chapters) often involve logarithmic factors.

VI. 1. Introduction

Rational and meromorphic functions have coefficients whose asymptotic form involves "exponential polynomials", that is, finite linear combinations of elements of the form

$$\left(\frac{1}{\omega}\right)^n n^{k-1},$$

with k a positive integer. This reflects the nature of polar singularities, with corresponding elements of the form

$$\frac{1}{(1-\frac{z}{\omega})^k},$$

for the function itself. We examine here a class of methods that yield a much wider range of subexponential factors. The method, called *singularity analysis*, provides asymptotic forms of coefficients which are of the type

$$\left(\frac{1}{\omega}\right)^n n^{\alpha-1}\,(\log n)^{\beta},$$

with α and β being arbitrary real (or even complex) numbers. Such forms relate to singularities of a more complicated nature than mere poles, namely, elements of the form

$$\frac{1}{(1-\frac{z}{\omega})^{\alpha}}\;\left(\log\frac{1}{1-\frac{z}{\omega}}\right)^{\beta}.$$

in the asymptotic expansion of the function at its singularity ω.

The exponential factor ω^{-n} is, as seen in Chapter IV, easily accounted for as the location of the dominant singularities always induces a multiplicative exponential factor for coefficients. If $f(z)$ is singular at $z=\omega$, then $g(z) := f(z/\omega)$ satifies, by the scaling rule of Taylor expansions

$$[z^n]\,f(z) = \omega^n [z^n]\,f(\frac{z}{\omega}) = \omega^n [z^n]\,g(z),$$

and $g(z)$ itself is singular on the unit disc. Consequently, in most of the discussion that follows, we shall examine functions $f(z)$ that are singular at $z=1$, a condition that entails no loss of generality.

Consider commonly encountered functions that are singular at 1. Here is a small sample of those admitting expansions of an elementary form:

(1)

	Function		*Coeff. (exact)*		*Coeff. (asymptotic)*
(f_1)	$[z^n]\,1-\sqrt{1-z}$	$=$	$\frac{2}{n4^n}\binom{2n-2}{n-1}$	$\sim$	$\frac{1}{2\sqrt{\pi n^3}}$
(f_2)	$[z^n]\,\frac{1}{\sqrt{1-z}}$	$=$	$\frac{1}{4^n}\binom{2n}{n}$	$\sim$	$\frac{1}{\sqrt{\pi n}}$
(f_3)	$[z^n]\,\frac{1}{1-z}$	$=$	1	$\sim$	1
(f_4)	$[z^n]\,\frac{1}{1-z}\log\frac{1}{1-z}$	$=$	H_n	$\sim$	$\log n$
(f_5)	$[z^n]\,\frac{1}{(1-z)^2}$	$=$	$n+1$	$\sim$	$n.$

Such a table obviously has structure: a logarithmic factor in the function is reflected by a similar factor in the coefficients; square-roots somehow induce square-roots; finally functions of larger growth have larger coefficients; see Figure 1.

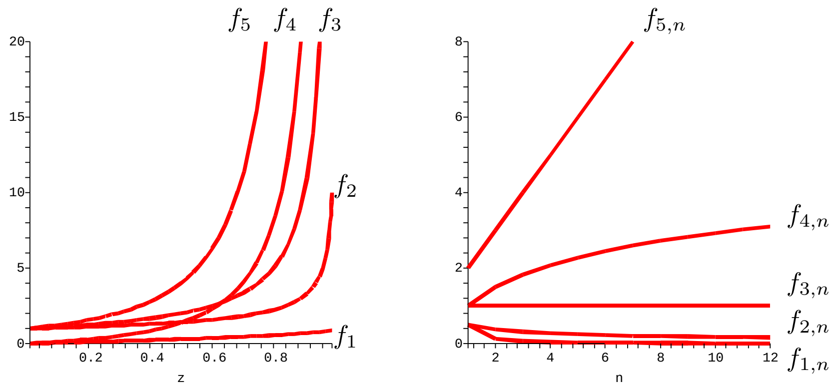


FIGURE 1. The five functions from Eq. (1) and a plot of their coefficient sequences illustrate the tendency of coefficient extraction to be consistent with orders of growth of functions.

Here is a partial explanation of such observations. First, regarding the basic functions in the scale, the Newton expansion

$$(1-z)^{-\alpha} = \sum_{n=0}^{\infty} \binom{n+\alpha-1}{n} z^n$$

when specialized to an integer k immediately gives the asymptotic form of the coefficients involved,

$$\begin{aligned} [z^n](1-z)^{-k} &\equiv \frac{(n+1)(n+2)\cdots(n+k-1)}{(k-1)!} \\ &= \frac{n^{k-1}}{(k-1)!}\left(1+\mathcal{O}(\frac{1}{n})\right). \end{aligned} \tag{2}$$

For general α, it is therefore natural to expect

$$\begin{aligned} [z^n](1-z)^{-\alpha} &\equiv \binom{n+\alpha-1}{\alpha-1} \\ &= \frac{n^{\alpha-1}}{(\alpha-1)!}\left(1+\mathcal{O}(\frac{1}{n})\right). \end{aligned} \tag{3}$$

It turns out that the asymptotic formula and even a full asymptotic expansion are valid for real or complex α, provided we interpret $(\alpha-1)!$ suitably. Indeed, one has, (see Section VI. 2 and Theorem VI.1)

$$[z^n](1-z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}\left(1+\frac{\alpha(\alpha-1)}{2n}+\cdots\right), \tag{4}$$

where $\Gamma(\alpha)$ is the *Euler Gamma function* defined as

$$\Gamma(\alpha) := \int_0^{\infty} e^{-t} t^{\alpha-1}\, dt, \tag{5}$$

for $\Re(\alpha) > 0$, which coincides with $(\alpha-1)!$ whenever α is an integer. Basic properties of this function are recalled in APPENDIX: *Gamma function*, p. 222.

We indeed observe from the pair (2)–(3) that functions of the form $(1-z)^{-\alpha}$ that are larger at the singularity $z=1$ (corresponding to larger values of α) also have larger coefficients. The correspondence that this observation suggests is very general as we are going to see repeatedly throughout this chapter.

Second, an asymptotic expansion of $f(z)$ around $z = 1$ is typically of the form

$$f(z) = \sigma(z) + \mathcal{O}(\tau(z)) \text{ where } \sigma(z) \gg \tau(z) \text{ as } z \to 1, \tag{6}$$

with σ and τ belonging to an asymptotic scale of standard functions like the collection $\{(1-z)^{-\alpha}\}_{\alpha\in\mathbb{R}}$ in simpler cases. Taking formally Taylor coefficients in the expansion (6), we arrive at

$$f_n \equiv [z^n]f(z) = [z^n]\sigma(z) + [z^n]\mathcal{O}(\tau(z)). \tag{7}$$

Therefore, in order to extract asymptotic informations on the coefficients of $f(z)$, two ingredients are needed:

(i) A catalogue of exact or asymptotic forms for coefficients of standard singular functions $\sigma(z)$;

(ii) A way of extracting coefficients of functions known only by their order of growth around the singularity.

The first aspect (i) is achieved by expansions of the type (4). The second aspect (ii) is achieved by transfer lemmas which under suitable conditions, essentially analytic continuation, guarantee that

$$[z^n]\mathcal{O}(\tau(z)) = \mathcal{O}([z^n]\tau(z)),$$

a relation which is much less trivial than its symbolic form would seem to imply (see Section VI. 3 and Theorem VI.3).

In summary, under favourable conditions which it is the goal of this chapter to elicit, we have available the implication

$$f(z) = \sigma(z) + \mathcal{O}(\tau(z)) \qquad \Longrightarrow \qquad f_n = \sigma_n + \mathcal{O}(\tau_n). \tag{8}$$

The process of singularity analysis thus parallels the analysis of coefficients of rational and meromorphic functions presented in the previous chapter. The range of singular behaviours taken into account by singularity analysis is however considerably larger. We shall allow here functions from the scale

$$\frac{1}{(1-z)^\alpha}(\log\frac{1}{1-z})^\beta \qquad (z \to 1),$$

which, for coefficients, appear to induce subexponential factors of the form

$$\theta(n) = n^{\alpha-1}(\log n)^\beta.$$

(See Theorem VI.2.) Even iterated logarithms ($\log\log$'s) and more exotic functions can be encapsulated in the method.

As an illustration of the *modus operandi*, consider the function

$$f(z) = \frac{e^{-z-z^2/2}}{\sqrt{1-z}},$$

which is the EGF of 2–regular graphs (or equivalently, "clouds", see Chapter II or [**39**]). Singularity analysis permits us to reason as follows. The function $f(z)$ is only singular at $z = 1$ where it has a branch point. Expanding the numerator around $z = 1$, we have

$$f(z) = \frac{e^{-3/4}}{\sqrt{1-z}} + \mathcal{O}((1-z)^{3/2}). \tag{9}$$

Therefore (see Theorems VI.1 and VI.3, as well as the discussion on page 134 below), upon translating formally and term–by–term, one has

$$[z^n]f(z) = e^{-3/4}\binom{n-1/2}{n} + \mathcal{O}\binom{n-3/2}{n} = \frac{e^{-3/4}}{\sqrt{\pi n}} + \mathcal{O}(n^{-3/2}), \tag{10}$$

and a full asymptotic expansion into descending powers of n can be obtained.

VI. 2. Coefficient asymptotics for the basic scale

This section and the next are essentially based on the theory developed by Flajolet and Odlyzko [**78**] and called *singularity analysis*. Technically the theory relies on a systematic use of Hankel contours in Cauchy coefficient integrals. Hankel contours classically serve to express the Gamma function: see APPENDIX: *Gamma function*, p. 222. Here they are first used to estimate coefficients of a standard scale of functions, and then to prove transfer theorems for error terms . This constitutes the basic process by which an asymptotic expansion of a function near a singularity is directly mapped to a matching asymptotic expansion of its coefficients.

Returning to the binomial expansion, we have for general α,

$$[z^n](1-z)^{-\alpha} = (-1)^n\binom{-\alpha}{n} = \binom{n+\alpha-1}{n} = \frac{\alpha(\alpha+1)\cdots(\alpha+n-1)}{n!}.$$

This quantity is expressible in terms of Gamma factors (APPENDIX: *Gamma function*, p. 222), and

$$\binom{n+\alpha-1}{n} = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\Gamma(n+1)}, \tag{11}$$

provided α is neither 0 nor a negative integer. When $\alpha \in \{0,-1,\ldots\}$, the coefficients $\binom{n+\alpha-1}{n}$ eventually vanish, so that the asymptotic problem of estimating $[z^n](1-z)^{-\alpha}$ becomes void. The asymptotic analysis of the coefficients $\binom{n+\alpha-1}{n}$ can be carried out elementarily by means of Stirling's formula or by real integral estimates: see Notes 1 and 2. However, a far more productive method consists in analysing these coefficients by means of Cauchy's coefficient formula applied to their GF. For us, this approach has two major advantages: it paves the way to the proof of transfer theorems; it readily extends to coefficients of functions involving logarithmic (and even iterated logarithmic) factors.

$\triangleright$ **1.** *Stirling's formula and asymptotics of binomial coeffi cients.* The gamma function form (11) of the binomial coefficients yields

$$[z^n](1-z)^{-\alpha} = \frac{n^{\alpha-1}}{\Gamma(\alpha)}\left(1+\mathcal{O}(\frac{1}{n})\right),$$

when Stirling's formula is applied to the gamma factors. $\triangleleft$

$\triangleright$ **2.** *Beta integrals and asymptotics of binomial coeffi cients.* The following constitutes a direct way of obtaining the general asymptotic form of $\binom{n+\alpha-1}{n}$ based on the Eulerian Beta integral (see [**191**, p.254] and APPENDIX: *Gamma function*, p. 222). Consider the quantity

$$\phi(n,\alpha) = \int_0^1 t^{\alpha-1}(1-t)^{n-1}\,dt = \frac{(n-1)!}{\alpha(\alpha+1)\cdots(\alpha+n-1)} \equiv \frac{1}{n\binom{n+\alpha-1}{n}},$$

where the second form results elementarily from successive integrations by parts. The change of variables $t = x/n$ yields

$$\phi(n,\alpha) = \frac{1}{n^\alpha}\int_0^n x^{\alpha-1}(1-x/n)^{n-1}\,dt \underset{n\to\infty}{\sim} \frac{1}{n^\alpha}\int_0^\infty x^{\alpha-1}e^{-x}\,dx \equiv \frac{\Gamma(\alpha)}{n^\alpha},$$

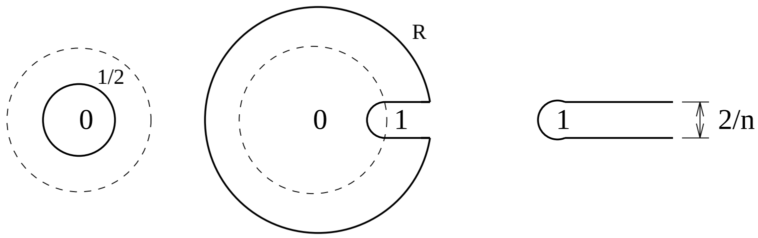


FIGURE 2. The contours $\mathcal{C}_0$, $\mathcal{C}_1$, and $\mathcal{C}_2 \equiv \mathcal{H}(n)$ used for estimating the coefficients of functions from the standard asymptotic scale.

where the asymptotic form results from the standard limit formula of the exponential: $\exp(a) = \lim_{n\to\infty}(1 + a/n)^n$. $\triangleleft$

THEOREM VI.1 (Standard function scale). *Let α be a number not belonging to the set $\{0, -1, -2, \ldots\}$. The coefficient of z^n in*

$$f(z) = (1-z)^{-\alpha}$$

admits for large n a full asymptotic expansion in descending powers of n,

$$[z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}\left(1 + \sum_{k=1}^{\infty}\frac{e_k(\alpha)}{n^k}\right),$$

where $e_k(\alpha)$ is a polynomial in α of degree $2k$.

The polynomial $e_k(\alpha)$ turns out to be divisible by $\alpha(\alpha-1)\cdots(\alpha-k)$. In particular, we have

$$f_n = [z^n](1-z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}\Bigg(1 + \frac{\alpha\,(\alpha-1)}{2\,n} + \frac{\alpha\,(\alpha-1)\,(\alpha-2)\,(3\,\alpha-1)}{24\,n^2} + \frac{\alpha^2\,(\alpha-1)^2\,(\alpha-2)\,(\alpha-3)}{48\,n^3} + \cdots\Bigg). \tag{12}$$

PROOF. First the coefficient $[z^n](1-z)^{-\alpha}$ is expressed by means of Cauchy's coefficient formula,

$$f_n = \frac{1}{2i\pi}\int_{\mathcal{C}}(1-z)^{-\alpha}\,\frac{dz}{z^{n+1}}, \tag{13}$$

with $\mathcal{C}$ a small enough contour that encircles the origin, for instance the positively oriented circle $\mathcal{C}_0 = \{z,\ |z| = \frac{1}{2}\}$. Next this contour $\mathcal{C}_0$ gets deformed into another one, $\mathcal{C}_1$, consisting of a large circle of radius $R > 1$ with a notch that comes back near and to the left of $z = 1$. (In effect, any simple closed contour around the origin that does not cross the half–line $z \geq 1$ is adequate.) Since the integrand along large circles decreases as $\mathcal{O}(R^{-n-\alpha})$, we can finally let R tend to infinity. We are then left with an integral representation for f_n where $\mathcal{C}$ is thus replaced by a contour $\mathcal{C}_2$ that starts from $-\infty$ in the lower half plane, winds around 1 clockwise and ends at $+\infty$ in the upper half plane. This is a typical case of a *Hankel contour* otherwise described in APPENDIX: *Gamma function*, p. 222.

In order to specify fully the integration path, we particularize $\mathcal{C}_2$ to be the contour $\mathcal{H}(n)$ that passes at a distance $\frac{1}{n}$ from the half line $[1, +\infty[$:

$$\mathcal{H}(n) = \mathcal{H}^-(n) + \mathcal{H}^+(n) + \mathcal{H}^\circ(n) \tag{14}$$

where

$$\begin{cases} \mathcal{H}^-(n) & = & \{z = w - \frac{i}{n},\ w \geq 1\} \\ \mathcal{H}^+(n) & = & \{z = w + \frac{i}{n},\ w \geq 1\} \\ \mathcal{H}^\circ(n) & = & \{z = 1 - \frac{e^{i\phi}}{n},\ \phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]\}. \end{cases} \tag{15}$$

Now, a change of variable

$$z = 1 + \frac{t}{n} \tag{16}$$

in the integral (13) gives the form

$$f_n = \frac{n^{\alpha-1}}{2i\pi} \int_{\mathcal{H}} (-t)^{-\alpha} \left(1 + \frac{t}{n}\right)^{-n-1} dt \tag{17}$$

where $\mathcal{H}$ is exactly the Hankel contour encountered in the proof of Theorem B.1.

We have

(18)

$$\left(1 + \frac{t}{n}\right)^{-n-1} = e^{-(n+1)\log(1+t/n)} = e^{-t}\left[1 + \frac{t^2 - 2t}{2n} + \frac{3t^4 - 20t^3 + 24t^2}{24n^2} + \cdots\right].$$

Thus, the integrand in (17) converges pointwise (as well as uniformly in any bounded domain of the t plane) to $(-t)^{-\alpha}e^{-t}$ which is precisely the "kernel" that appears in Hankel's formula for the Gamma function. Substitution of the asymptotic form

$$(1 + \frac{t}{n})^{-n-1} = e^{-t}(1 + \mathcal{O}(\frac{1}{n})),$$

as $n \to \infty$ inside the integral (17) suggests that

$$[z^n](1-z)^{-\alpha} = \frac{n^{\alpha-1}}{\Gamma(\alpha)}(1 + \mathcal{O}(\frac{1}{n})).$$

Furthermore, the full expansion (18) when plugged into the integral (17) formally leads to an expansion in descending powers of n.

To complete the argument outlined in the previous paragraph, one then proceeds as follows:

(i) Split the contour according to $|t| \leq \log^2 n$ and $|t| \geq \log^2 n$. The part corresponding to $|t| \geq \log^2 n$ is seen to be negligible in the scale of the problem; for instance, there,

$$(1 + \frac{t}{n})^{-n} = \mathcal{O}(\exp(-\log^2 n)).$$

(ii) On the remaining part of the contour, $|t| \leq \log^2 n$, the quantity $\frac{t}{n}$ is small enough, being of order $\frac{\log^2 n}{n}$, so that a terminating form of (18) may be developed to any predetermined order with uniform error terms.

These considerations justify term-by-term integration of expansion (18) within the integral of (17).

	$n=10$	$n=20$	$n=50$
$\frac{4^n}{\sqrt{\pi n^3}}\big(1$	**1** *8708*	**6** *935533866*	*2022877684829178931751713264*
$-\frac{9}{8}\,N^{-1}$	**16** *603*	**65** *45410086*	**197** *7362936920522405787299715*
$+\frac{145}{128}\,N^{-2}$	**16** *815*	**656** *5051735*	**19782** *7955337146062749074971 0*
$-\frac{1155}{1024}\,N^{-3}$	**1679** *4*	**6564** *073885*	**1978261** *3000611014266964827 32*
$+\frac{36939}{32768}\,N^{-4}$	**16796**	**656412** *2750*	**19782616** *6491988462935781 3591*
$-\frac{295911}{262144}\,N^{-5}$	**16796**	**6564120** *303*	**1978261657** *612856326190245636*
$+\frac{4735445}{4194304}\,N^{-6}$	**16796**	**656412042** *6*	**197826165775** *9023715384519184*
$-\frac{37844235}{33554432}\,N^{-7}\big)$	**16796**	**6564120420**	**19782616577561** *03402179527600*
C_n	**16796**	**6564120420**	**1978261657756160653623774456**

FIGURE 3. Improved approximations to the Catalan numbers obtained by successive terms of their asymptotic expansion.

The full expansion is then computed as follows. A term of the form $\frac{t^r}{n^s}$ in the expansion (18) induces, by Hankel's formula, a term of the form $\frac{1}{\Gamma(\alpha-r)}\,\frac{1}{n^s}$. The expansion so obtained is nondegenerate provided α differs from a negative integer or zero. Since

$$\frac{1}{\Gamma(\alpha-k)} = \frac{1}{\Gamma(\alpha)}(\alpha-1)(\alpha-2)\cdots(\alpha-k).$$

the expansion in the statement of the theorem eventually follows. $\square$

The asymptotic approximations obtained are far from being as accurate as the ones that derive from meromorphic asymptotics in Chapter IV, where exponentially small error terms could be derived. However here, when the first few terms of the asymptotic expansion are included, it is not uncommon to obtain results with about 10^{-6} accuracy, already for values of n in the range 10^1–10^2. Figure 3 examplifies this situation by displaying the approximations obtained for the Catalan numbers,

$$C_n = \frac{4^n}{n+1}[z^n](1-z)^{-1/2},$$

when C_{10}, C_{20}, C_{50} are considered and up to eight asymptotic terms are taken into account.

The basic principle underlying the method of proof of Theorem VI.1 is simple. It consists in taking a contour of integration that comes close to the singularity at $z=1$. By choosing this contour to pass at distance $\frac{1}{n}$, the kernel in Cauchy's coefficient formula transforms into an exponential, while the function can be locally expanded, with the differential coefficient only introducing a rescaling factor of $1/n$:

$$\frac{1}{z^{n+1}} \mapsto e^{-t}, \qquad (1-z)^{-\alpha} \mapsto n^{\alpha}(-t)^{-\alpha}, \qquad dz \mapsto \frac{1}{n}dt.$$

In other words, the contour $\mathcal{H}(n)$ "captures" the behaviour of the function near its singularity, thereby enabling coefficient estimation.

This principle has the further advantage of generalizing to a wide class of singular functions, most notably the ones that involve logarithmic terms, as well as leading to a whole range of transfers of $\mathcal{O}(.)$ and $o(.)$ terms, to be established in the next section.

THEOREM VI.2 (Standard coefficient scale, logarithms). *Let α be a number not in $\{0, -1, -2, \ldots\}$. The coefficient of z^n in*

$$f(z) = (1-z)^{-\alpha} \left(\frac{1}{z}\log\frac{1}{1-z}\right)^{\beta}$$

admits for large n a full asymptotic expansion in descending powers of $\log n$,

$$f_n = [z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} (\log n)^{\beta} \left[1 + \frac{C_1}{1!}\frac{\beta}{\log n} + \frac{C_2}{2!}\frac{\beta(\beta-1)}{(\log n)^2} + \cdots\right]. \tag{19}$$

There, $C_k = C_k(\alpha)$ represents

$$\Gamma(\alpha)\frac{d^k}{ds^k}\left.\frac{1}{\Gamma(s)}\right|_{s=\alpha}.$$

A coefficient of $\frac{1}{z}$ is introduced in front of the logarithm since $\log(1-z)^{-1} = z + \mathcal{O}(z^2)$. In this way, $f(z)$ is a *bona fide* power series in z, even in cases when β is not a positive integer.

PROOF. The proof is a simple variant of that of Theorem VI.1, see [**78**] for details. The basic expansion used is now

$$\begin{aligned}
f(1+\frac{t}{n})(1+\frac{t}{n})^{-n-1} &\sim e^{-t}\left(\frac{-n}{t}\right)^{\alpha}\left(\log\left(\frac{-n}{t}\right)\right)^{\beta} \\
&\sim \frac{e^{-t}(-t)^{-\alpha}}{n^{\alpha}}(\log n)^{\beta}\left(1-\frac{\log(-t)}{\log n}\right)^{\beta} \\
&\sim \frac{e^{-t}(-t)^{-\alpha}}{n^{\alpha}}(\log n)^{\beta}\left(1-\beta\frac{\log(-t)}{\log n}+\frac{\beta(\beta-1)}{2!}(\frac{\log(-t)}{\log n})^2+\cdots\right).
\end{aligned}$$

It proves again justified to employ this expansion inside the integral defining the coefficients. What comes out is a collection of Hankel integrals of the form

$$-\frac{1}{2i\pi}\int_{+\infty}^{(0)}(-t)^{-s}e^{-t}(\log(-t))^k\,dt$$

which reduce to derivatives of $\frac{1}{\Gamma(s)}$ as is seen by differentiation with respect to s under the integral sign. □

A typical example of application of Theorem VI.2 is

$$[z^n]\frac{1}{\sqrt{1-z}}\,\frac{1}{\frac{1}{z}\log\frac{1}{1-z}} = \frac{1}{\sqrt{\pi n}\,\log n}\left(1-\frac{\gamma+2\log 2}{\log n}+\mathcal{O}(\frac{1}{\log^2 n})\right).$$

(Surprising as it may seem, such singular functions do occur in combinatorics and the analysis of algorithms [**82**].)

Furthermore, a direct adaptation of the proof of our basic theorems leads to results regarding the coefficients of many functions that have a nearly polynomial growth. It is proved in [**78**] that, for a class of functions L *slowly varying* at ∞ and for $\alpha \neq 0, -1, \ldots$, one has:

$$[z^n]\frac{1}{(1-z)^{\alpha}}L(\frac{1}{1-z}) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}L(n). \tag{20}$$

	$\alpha \notin \{0,-1,-2,\ldots\}$	(Eq.)	$\alpha \in \{0,-1,-2,\ldots\}$	(Eq.)
$\beta \notin \mathbb{Z}_{\geq 0}$	$\frac{n^{\alpha-1}}{\Gamma(\alpha)}(\log n)^\beta \sum_{j=0}^{\infty} \frac{C_j}{(\log n)^j}$	(19)	$f_n \sim n^{\alpha-1}(\log n)^\beta \sum_{j=1}^{\infty} \frac{C_j^\star}{(\log n)^j}$	(21)
$\beta \in \mathbb{Z}_{\geq 0}$	$\frac{n^{\alpha-1}}{\Gamma(\alpha)} \sum_{j=0}^{\infty} \frac{E_j(\log n)}{n^j}$	(22)	$n^{\alpha-1} \sum_{j=0}^{\infty} \frac{F_j(\log n)}{n^j}$	(24)

FIGURE 4. The general and special cases of $f_n \equiv [z^n]f(z)$ when $f(z)$ is as in Theorem VI.2.

Logarithms and their powers constitute typical instances of such slowly varying functions; iterated logarithms also belong to this class and, for a general $\alpha \neq 0, -1, \ldots$, the relation (20) specializes to

$$[z^n](1-z)^{-\alpha}\left(\frac{1}{z}\log\frac{1}{1-z}\right)^\beta \left(\frac{1}{z}\log\left(\frac{1}{z}\log\frac{1}{1-z}\right)\right)^\delta \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}(\log n)^\beta(\log\log n)^\delta.$$

A full asymptotic expansion in descending powers of $\log n$ can once more be derived in this case [**78**].

Special cases. The conditions of Theorem VI.2 exclude explicity the case when α is an integer ≤ 0. The formulæ actually remain valid in this case, provided one interprets them as limit cases, making use of $0 = 1/\Gamma(0) = 1/\Gamma(-1) = \cdots$. When β is a positive integer, stronger forms are valid. Such cases are summarized in Figure 4 and discussed now.

The case of integral $\alpha \in \mathbb{Z}_{\leq 0}$. When α is an integer ≤ 0, the coefficients of $f(z) = (1-z)^{-\alpha}$ eventually reduce to zero, so that the asymptotic coefficient expansion becomes trivial. This situation is implicitly covered by the statement of Theorem VI.1 since, in that case, $1/\Gamma(\alpha) = 0$. When logarithms are present (with $\alpha \in \mathbb{Z}_{\leq 0}$ still), the expansion of Theorem VI.2 regarding

$$[z^n](1-z)^{-\alpha}(\frac{1}{z}\log\frac{1}{1-z})^\beta$$

remains valid provided we again take into account the equality $1/\Gamma(\alpha) = 0$ in formula (19) after effecting simplifications by Gamma factors: It is only the first term of (19) that vanishes,

$$(21) \qquad f_n = [z^n]f(z) \sim n^{\alpha-1}(\log n)^\beta \left[\frac{C_1^*}{1!}\frac{\beta}{\log n} + \frac{C_2^*}{2!}\frac{\beta(\beta-1)}{(\log n)^2} + \cdots\right],$$

where, $C_k^* \equiv C_k^*(\alpha)$ satisfies

$$C_k^* = \frac{d^k}{ds^k}\left.\frac{1}{\Gamma(s)}\right|_{s=\alpha}.$$

For instance, we find

$$[z^n]\frac{z}{\log(1-z)^{-1}} = -\frac{1}{n\log^2 n} + \frac{2\gamma}{n\log^3 n} + \mathcal{O}(\frac{1}{n\log^4 n}).$$

The case of integral $\beta \in \mathbb{Z}_{\geq 0}$. When β is a nonnegative integer, the error terms can be further improved with respect to the ones predicted by the general statement of

Theorem VI.2. For instance, we have (see also Note 3):

$$\begin{array}{rcl} [z^n]\dfrac{1}{1-z}\log\dfrac{1}{1-z} & = & \log n + \gamma + \dfrac{1}{2n} - \dfrac{1}{12n^2} + \mathcal{O}(\dfrac{1}{n^4}) \\ [z^n]\dfrac{1}{\sqrt{1-z}}\log\dfrac{1}{1-z} & \sim & \dfrac{1}{\sqrt{\pi n}}\left(\log n + \gamma + 2\log 2 + \mathcal{O}(\dfrac{\log n}{n})\right). \end{array}$$

(In such a case, the expansion of Theorem VI.2 terminates since only its first $(k+1)$ terms are nonzero.) In fact, in the general case of nonintegral α, there exists an expansion of the form

$$[z^n](1-z)^{-\alpha}\log^k\frac{1}{1-z} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}\left[E_0(\log n) + \frac{E_1(\log n)}{n} + \cdots\right], \tag{22}$$

where the E_j are polynomials of degree k, as can be proved by adapting the argument employed for general α. It is worth however mentioning an alternative approach due to Frobenius and Jungen [**112**], and based on the observation that

$$(1-z)^{-\alpha}\left(\log\frac{1}{1-z}\right)^k = \frac{\partial^k}{\partial\alpha^k}(1-z)^{-\alpha}.$$

If one lets the operators of differentiation ($\partial/\partial\alpha$) and coefficient extraction ($[z^n]$) commute —this can be justified by Cauchy's coefficient formula upon differentiating under the integral sign— one gets directly

$$[z^n](1-z)^{-\alpha}\left(\log\frac{1}{1-z}\right)^k = \frac{\partial^k}{\partial\alpha^k}\frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\Gamma(n+1)}. \tag{23}$$

For instance, there is an exact formula,

$$[z^n](1-z)^{-\alpha}\log\frac{1}{1-z} = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)\Gamma(n+1)}\left[\frac{1}{\alpha} + \frac{1}{\alpha+1} + \cdots + \frac{1}{n+\alpha-1}\right].$$

The joint case $\alpha \in \mathbb{Z}_{\le 0}$, $\beta \in \mathbb{Z}_{\ge 0}$. If α is an integer of $\mathbb{Z}_{\le 0}$, the coefficients appear as finite differences of coefficients of logarithmic functions. Explicit formulæ are then available elementarily from the calculus of finite differences. For instance, with $\alpha = -r$ for $r \in \mathbb{Z}_{\ge 0}$, one has

$$[z^n](1-z)^r\log\frac{1}{1-z} = (-1)^r\frac{r!}{n(n-1)\cdots(n-r)}. \tag{24}$$

The case $\alpha = -r$, $\beta = k$ is covered similarly by (23). Note that, in this case, there is a formula analogous to (22),

$$[z^n](1-z)^r\log^k\frac{1}{1-z} \sim n^{-r-1}\left[F_0(\log n) + \frac{F_1(\log n)}{n} + \cdots\right], \tag{25}$$

but now with deg $(F_j) = k-1$.

$\triangleright$ **3.** *Shifted harmonic numbers.* Define the α-shifted harmonic number by

$$h_n(\alpha) := \sum_{j=0}^{n-1}\frac{1}{j+\alpha}.$$

Set $L(z) := -\log(1-z)$. Then, one has

$$\begin{array}{rcl} [z^n](1-z)^{-\alpha}L(z) & = & \binom{n+\alpha-1}{n}h_n(\alpha) \\ [z^n](1-z)^{-\alpha}L(z)^2 & = & \binom{n+\alpha-1}{n}\left(h_n'(\alpha) + h_n(\alpha)^2\right). \end{array}$$

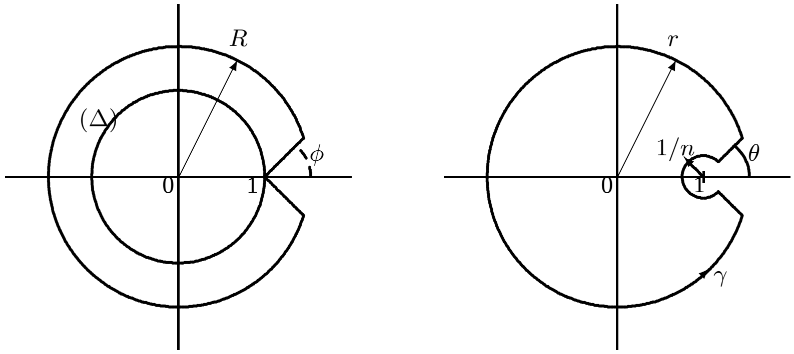


FIGURE 5. A Δ–domain and the contour used to establish Theorem 5.4.

(Note: $h_n(\alpha) = \psi(\alpha+n) - \psi(\alpha)$, where $\psi(s) := \partial_s \log \Gamma(s)$.) In particular,

$$[z^n]\frac{1}{\sqrt{1-z}} \log \frac{1}{1-z} = \frac{1}{4^n}\binom{2n}{n}[2H_{2n} - H_n],$$

where $H_n \equiv h_n(1)$ is the usual harmonic number. $\lhd$

$\rhd$ **4.** *Oscillations and complex exponents.* Fluctuations occur in the case of singular expansions involving complex exponents. From the consideration of $[z^n](1-z)^{\pm i} \asymp n^{\mp i}$, one finds

$$[z^n]\cos\left(\log\frac{1}{1-z}\right) = \frac{P(\log n)}{n} + O(\frac{1}{n^2}),$$

where $P(u)$ is a continuous and 1–periodic function. In general, oscillations are present in $[z^n](1-z)^{-\alpha}$ for any nonreal α. $\lhd$

VI. 3. Transfers

Once coefficients of a fairly extensive scale have been made explicit, there remains to show how to translate error terms in the asymptotic approximation of a function near a singularity. This task is even technically simpler as a coarser analysis suffices. It still relies on the principles of contour integration by means of Hankel-type paths.

A natural extension of the previous results is to assume the error terms valid in the complex plane slit along the real half line $[1, +\infty[$. In fact weaker conditions suffice and any domain whose boundary makes an acute angle with the half line $[1, +\infty[$ is suitable.

DEFINITION VI.1. *Given two numbers ϕ, R with $R > 1$ and $0 < \phi < \frac{\pi}{2}$, the open domain $\Delta(\phi, R)$ is defined as*

$$\Delta(\phi, R) = \{z \;\big|\; |z| < R,\ z \neq 1,\ |\operatorname{Arg}(z-1)| > \phi\}.$$

A domain is a Δ– domain if it is a $\Delta(\phi, R)$ for some R $(R > 1)$ and some ϕ $(0 < \phi < \frac{\pi}{2})$. A function is Δ– analytic if it is analytic in some Δ– domain.

Analyticity in a Δ–domain (Figure 5) is the basic condition for *transfer* to coefficients of error terms in asymptotic expansions.

THEOREM VI.3 (Transfer, Big-Oh and little-oh). (i) *Assume that $f(z)$ is Δ– analytic and that it satisfies in the intersection of a neighbourhood of 1 and of its Δ– domain the condition*

$$f(z) = \mathcal{O}\left((1-z)^{-\alpha}(\log\frac{1}{1-z})^{\beta}\right).$$

Then

$$[z^n]f(z) = \mathcal{O}(n^{\alpha-1}(\log n)^{\beta}).$$

(ii) *Assume that $f(z)$ is Δ– analytic and that it satisfies in the intersection of a neighbourhood of 1 and of the Δ– domain the condition*

$$f(z) = o\left((1-z)^{-\alpha}(\log\frac{1}{1-z})^{\beta}\right).$$

Then

$$[z^n]f(z) = o(n^{\alpha-1}(\log n)^{\beta}).$$

PROOF. The starting point is Cauchy's coefficient formula,

$$f_n \equiv [z^n]f(z) = \frac{1}{2i\pi}\int_{\gamma} f(z)\,\frac{dz}{z^{n+1}},$$

where γ is a loop around the origin which is internal to the Δ–domain of f. We choose the positively oriented contour (Figure 5) $\gamma = \gamma_1+\gamma_2+\gamma_3+\gamma_4$, with

$$\begin{cases} \gamma_1 &=& \{\, z \;\big|\; |z-1| = \frac{1}{n}, |\operatorname{Arg}(z-1)| \geq \theta]\,\} \\ \gamma_2 &=& \{\, z \;\big|\; \frac{1}{n} \leq |z-1|,\ |z| \leq r,\ \operatorname{Arg}(z-1) = \theta\,\} \\ \gamma_3 &=& \{\, z \;\big|\; |z-1| = r,\ |\operatorname{Arg}(z-1)| \geq \theta]\,\} \\ \gamma_4 &=& \{\, z \;\big|\; \frac{1}{n} \leq |z-1|,\ |z| \leq r,\ \operatorname{Arg}(z-1) = -\theta\,\}. \end{cases}$$

If the Δ domain of f is $\Delta(\phi, R)$, we assume that $1 < r < R$, and $\phi < \theta < \frac{\pi}{2}$, so that the contour γ lies entirely inside the domain of analyticity of f.

For $j = 1, 2, 3, 4$, let

$$f_n^{(j)} = \frac{1}{2i\pi}\int_{\gamma_j} f(z)\,\frac{dz}{z^{n+1}}.$$

The analysis proceeds by bounding the absolute value of the integral along each of the four parts. In order to keep notations simple, we detail the proof in the case where $\beta = 0$.

(1) *Inner circle.* From trivial bounds, the contribution there is

$$|f_n^{(1)}| = \mathcal{O}(\frac{1}{n})\cdot\mathcal{O}((\frac{1}{n})^{-\alpha}),$$

as the function is $\mathcal{O}((\frac{1}{n})^{-\alpha})$, the contour has length $\mathcal{O}(\frac{1}{n})$, and z^{-n-1} is $\mathcal{O}(1)$ there.

(2) *Rectilinear part.* Setting $\omega = e^{i\theta}$, and performing the change of variable $z = 1+\frac{\omega t}{n}$, we find

$$|f_n^{(2)}| < \frac{1}{2\pi}\int_1^{\infty} K\left(\frac{t}{n}\right)^{-\alpha}\left|1+\frac{\omega t}{n}\right|^{-n-1} dt,$$

for some constant $K > 0$ such that $|f(z)| < K(1-z)^{-\alpha}$ over the Δ–domain. From the relation

$$\left|1+\frac{\omega t}{n}\right| \geq 1 + \Re(\frac{\omega t}{n}) = 1+\frac{t}{n}\cos\theta,$$

there results

$$|f_n^{(2)}| < \frac{K}{2\pi}J_n n^{\alpha-1} \qquad \text{where} \qquad J_n = \int_1^{\infty} t^{-\alpha}\left(1+\frac{t\cos\theta}{n}\right)^{-n} dt.$$

For a given α, the integrals J_n are all bounded above by some constant since they admit a limit as n tends to infinity:

$$J_n \to \int_1^\infty t^{-\alpha} e^{-t\cos\theta}\, dt.$$

(The condition on θ that $0 < \theta < \frac{\pi}{2}$ precisely ensures convergence of the integral.) Thus, globally, on this part of the contour, we have

$$|f_n^{(2)}| = \mathcal{O}(n^{\alpha-1}),$$

and the same bound holds for γ_4 by symmetry.

(3) *Outer circle.* There, $f(z)$ is bounded while z^{-n} is of the order of r^{-n}. Thus, $f_n^{(3)}$ is exponentially small.

In summary, each of the four integrals of the split contour contributes $\mathcal{O}(n^{\alpha-1})$. The statement of Part (i) of the theorem thus follows.

(ii) An adaptation of the proof shows that $o(.)$ error terms may be translated similarly. All that is required is a further breakup of the rectilinear part in the proof of Theorem VI.3 at a distance $\log^2 n/n$ from 1, see [**78**] for details. □

An immediate corollary of Theorem VI.3 is the possibility of transferring *asymptotic equivalence* from singular forms to coefficients:

COROLLARY VI.1 (sim–transfer). *Assume f is Δ– analytic and, as $z \to 1$ in Δ,*

$$f(z) \sim (1-z)^{-\alpha},$$

with $\alpha \notin \{0, -1, -2, \cdots\}$. Then, the coefficients of f satisfy

$$[z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}.$$

PROOF. It suffices to observe that, with $g(z) = (1-z)^{-\alpha}$, one has

$$f(z) \sim g(z) \qquad \text{iff} \qquad f(z) = g(z) + o(g(z)),$$

then apply Theorem VI.1 to the first term, and Theorem VI.3 (little-oh transfer) to the remainder. □

▷ **5.** *Transfer of nearly polynomial functions.* Let $f(z)$ be Δ–singular and satisfy the singular expansion $f(z) \sim (1-z)^r$, where $r \in \mathbb{Z}_{\geq 0}$. Then, $f_n = o(n^{-r-1})$. [This is a direct consequence of the little-oh transfer.] ◁

▷ **6.** *Transfer of "large" functions.* The Δ–analyticity condition can be weakened for functions that are large at their singularity. Assume that $f(z)$ is analytic in the open disk $|z| < 1$, and that in the whole of the open disk

$$f(z) = \mathcal{O}((1-z)^{-\alpha}).$$

Then, provided that $\alpha > 1$,

$$[z^n]f(z) = O(n^{\alpha-1}).$$

[Hint. Integrate on the circle of radius $1 - \frac{1}{n}$; see also [**78**].] ◁

Function	*Coefficients*
$(1-z)^{3/2}$	$\frac{1}{\sqrt{\pi n^5}}(\frac{3}{4}+\frac{45}{32n}+\frac{1155}{512n^2}+\mathcal{O}(\frac{1}{n^3}))$
$(1-z)$	(0)
$(1-z)^{1/2}$	$-\frac{1}{\sqrt{\pi n^3}}(\frac{1}{2}+\frac{3}{16n}+\frac{25}{256n^2}+\mathcal{O}(\frac{1}{n^3}))$
$(1-z)^{1/2}\log(1-z)^{-1}$	$-\frac{1}{\sqrt{\pi n^3}}(\frac{1}{2}\log n+\frac{\gamma+2\log 2-2}{2}+\mathcal{O}(\frac{\log n}{n}))$
$(1-z)^{1/3}$	$-\frac{1}{3\Gamma(\frac{2}{3})n^{4/3}}(1+\frac{2}{9n}+\frac{7}{81n^2}+\mathcal{O}(\frac{1}{n^3}))$
$z\log^{-1}(1-z)^{-1}$	$\frac{1}{n\log^2 n}(-1+\frac{2\gamma}{\log n}+\frac{\pi^2-6\gamma^2}{2\log^2 n}+\mathcal{O}(\frac{1}{\log^3 n}))$
1	(0)
$\log(1-z)^{-1}$	$\frac{1}{n}$
$\log^2(1-z)^{-1}$	$\frac{1}{n}(2\log n+2\gamma-\frac{1}{n}-\frac{1}{6n^2}+\mathcal{O}(\frac{1}{n^4}))$
$(1-z)^{-1/3}$	$\frac{1}{\Gamma(\frac{1}{3})n^{2/3}}(1+\mathcal{O}(\frac{1}{n}))$
$(1-z)^{-1/2}$	$\frac{1}{\sqrt{\pi n}}(1-\frac{1}{8n}+\frac{1}{128n^2}+\frac{5}{1024n^3}+\mathcal{O}(\frac{1}{n^4}))$
$(1-z)^{-1/2}\log(1-z)^{-1}$	$\frac{1}{\sqrt{\pi n}}(\log n+\gamma+2\log 2-\frac{\log n+\gamma+2\log 2}{8n}+\mathcal{O}(\frac{\log n}{n^2}))$
$(1-z)^{-1}$	1
$(1-z)^{-1}\log(1-z)^{-1}$	$\log n+\gamma+\frac{1}{2n}-\frac{1}{12n^2}+\frac{1}{120n^4}+\mathcal{O}(\frac{1}{n^6}))$
$(1-z)^{-1}\log^2(1-z)^{-1}$	$\log^2 n+2\gamma\log n+\gamma^2-\frac{\pi^2}{6}+\mathcal{O}(\frac{\log n}{n})$
$(1-z)^{-3/2}$	$\sqrt{\frac{n}{\pi}}(2+\frac{3}{4n}-\frac{7}{64n^2}+\mathcal{O}(\frac{1}{n^3}))$
$(1-z)^{-3/2}\log(1-z)^{-1}$	$\sqrt{\frac{n}{\pi}}(2\log n+2\gamma+4\log 2-2+\frac{3\log n}{4n}+\mathcal{O}(\frac{1}{n}))$
$(1-z)^{-2}$	$n+1$
$(1-z)^{-2}\log(1-z)^{-1}$	$n\log n+(\gamma-1)n+\log n+\frac{1}{2}+\gamma+\mathcal{O}(\frac{1}{n})$
$(1-z)^{-2}\log^2(1-z)^{-1}$	$n(\log^2 n+2(\gamma-1)\log n+\gamma^2-2\gamma+2-\frac{\pi^2}{6}+\mathcal{O}(\frac{\log n}{n}))$
$(1-z)^{-3}$	$\frac{1}{2}n^2+\frac{3}{2}n+1$

FIGURE 6. A table of some commonly encountered functions and the asymptotic forms of their coefficients.

The theorems that we have seen justify a fairly mechanical process for translating asymptotic information on a function into information on its coefficients. The process is based on a set of simple rules. With $\alpha\notin\{0,-1,-2,\ldots\}$, we have

$$\left\{\begin{array}{lll} f(z)=(1-z)^{-\alpha} & \Longrightarrow & f_n=\frac{n^{\alpha-1}}{\Gamma(\alpha)}+\cdots \\ f(z)=\mathcal{O}((1-z)^{-\alpha}) & \Longrightarrow & f_n=\mathcal{O}(n^{\alpha-1}) \\ f(z)=o((1-z)^{-\alpha}) & \Longrightarrow & f_n=o(n^{\alpha-1}) \\ f(z)\sim(1-z)^{-\alpha} & \Longrightarrow & f_n\sim\frac{n^{\alpha-1}}{\Gamma(\alpha)} \end{array}\right.$$

together with corresponding refinements for logarithmic terms. A table that extends examples already given is displayed in Fig 6. The only requirement of the method is that the asymptotic expansion of the function should be valid in an area of the complex plane extending beyond the disk of convergence of the original series, as described by the notions of Δ–domain and Δ–analyticity. This is usually not a stringent requirement in combinatorial applications, as we shall see repeatedly in this chapter and the next one.

VI. 4. First examples of singularity analysis

The previous section has provided tools by which, starting from the expansion of a function at its singularity[1], one can justify the term-by-term transfer to coefficients,

$$\boxed{f(z) \underset{z\to 1}{=} \sigma(z) + O(\tau(z)) \quad (z \in \Delta) \qquad \Longrightarrow \qquad f_n \underset{n\to\infty}{=} \sigma_n + O(\tau_n).} \tag{26}$$

There, it is assumed that σ is a finite linear combination of standard functions of the form $(\{(1-z)^{-\alpha}\}$ or logarithmic variants, that $\tau(z)$ also lies in the scale. (Note: the case when the exponent $-\alpha$ of $\tau(z)$ lies in $\mathbb{Z}_{\leq 0}$, adjustments must be made, as already discussed.) The fundamental condition is the validity of expansion (26) in a Δ-domain in accordance with Theorem VI.3; the coefficients σ_n, τ_n are given by the basic Theorems VI.1 and VI.2.

The functions

$$\frac{1}{1-z},\ \exp(z),\ \log\frac{1}{1-z},\ \sqrt{1-z},$$

are all Δ–analytic (with $\exp(z)$ being even entire). Thus, one should expect the method of singularity analysis to be applicable to most functions that are composition of base functions, provided their singular growth is only polynomial. (For instance $\exp(z/(1-z))$ is excluded, but such fast growing functions are well covered by the saddle point method described in a later chapter.) This class includes the generating functions for many of the elementary non–recursive combinatorial structures that can be specified using sequences, sets, and cycles. Thus, singularity analysis is *a priori* broadly applicable to elementary combinatorics.

We examine here several applications of singularity analysis to such functions explicitly given by "analytic" expressions. The examples are drawn from combinatorial enumeration.

1. "Clouds" and 2– regular graphs. The function

$$C(z) = \frac{e^{-\frac{z}{2}-\frac{z^2}{4}}}{\sqrt{1-z}}$$

is the EGF of 2–regular graphs or equivalently "clouds". (Let n straight lines in the plane be given; a "cloud" is a set of n points no three of which are collinear; see Chapter II and [**39**].) Combinatorially, the generating function reflects the decomposition of 2–regular graphs as sets ($\mathfrak{P}$) of connected components that are undirected cycles ($\mathfrak{UC}$) of size ≥ 3:

$$\mathcal{C} = \mathfrak{P}(\mathfrak{UC}_{\geq 3}(\mathcal{Z})), \qquad C(z) = \exp\left(\frac{1}{2}\left(\log(1-z)^{-1} - z - \frac{z}{2}\right)\right).$$

As this is our first example, it is worth spelling out in detail the process of singularity analysis applied to this function.

The function $C(z)$ being the product of $e^{-z/2-z^2/4}$ (that is entire) and of $(1-z)^{-1/2}$ (that is analytic in the unit disk) is itself analytic in the unit disk. Furthermore, as $(1-$

[1] Such an expansion is also called a *singular expansion* for short.

$z)^{-1/2}$ is Δ–analytic (it is well-defined and analytic in the complex plane slit along $[1,+\infty[$, for instance), $C(z)$ is itself Δ–analytic, with a singularity at $z=1$.

The asymptotic expansion of $C(z)$ near $z=1$ obtains starting from the standard (analytic) expansion of $e^{-z/2-z^2/4}$ at $z=1$,

$$e^{-z/2-z^2/4} = e^{-3/4} + e^{-3/4}(1-z) + \frac{e^{-3/4}}{4}(1-z)^2 - \frac{e^{-3/4}}{12}(1-z)^3 + \cdots .$$

The factor $(1-z)^{-1/2}$ is its own asymptotic expansion, clearly valid in any Δ–domain. Forming the product yields:

$$C(z) = \frac{e^{-3/4}}{\sqrt{1-z}} + e^{-3/4}\sqrt{1-z} + \frac{e^{-3/4}}{4}(1-z)^{3/2} - \frac{e^{-3/4}}{12}(1-z)^{5/2} + \cdots .$$

By the principles of singularity analysis (Theorems VI.1 and VI.3), the asymptotic determination of the coefficients $c_n = [z^n]C(z)$ results from a direct translation which we present here in tabular form:

$C(z) =$	$c_n =$
$e^{-3/4}\dfrac{1}{\sqrt{1-z}}$	$e^{-3/4}\dbinom{n-1/2}{-1/2} \sim \dfrac{e^{-3/4}}{\sqrt{\pi n}}\left[1-\dfrac{1}{8n}+\dfrac{1}{128n^2}+\cdots\right]$
$+\, e^{3/4}\sqrt{1-z}$	$+e^{-3/4}\dbinom{n-3/2}{-3/2} \sim \dfrac{-e^{-3/4}}{2\sqrt{\pi n^3}}\left[1+\dfrac{3}{8n}+\cdots\right]$
$+\,\mathcal{O}((1-z)^{3/2})$	$+\mathcal{O}(\dfrac{1}{n^{5/2}})$.

Terms are then collected with expansions suitably truncated to the coarsest error term, so that here a 3–term expansion results.

In the sequel, we shall not detail such computations and content ourselves with putting in parallel the function's expansion and the coefficient's expansion, like

$$(27)\qquad \begin{cases} C(z) &=& \dfrac{e^{-3/4}}{\sqrt{1-z}} + e^{3/4}\sqrt{1-z} + +\mathcal{O}((1-z)^{3/2}) & (z\to 1) \\ c_n &=& \dfrac{e^{-3/4}}{\sqrt{\pi n}} - \dfrac{5e^{-3/4}}{8\sqrt{\pi n^3}} + \mathcal{O}(\dfrac{1}{n^{5/2}}) & (n\to+\infty). \end{cases}$$

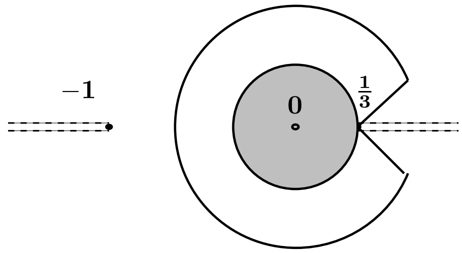
Here is a numerical check. Set $c_n^{(1)} := e^{-3/4}/\sqrt{\pi n}$ and let $c_n^{(2)}$ represent the sum of the first two terms of the expansion of c_n in (27). One finds:

n	5	50	500
$n!c_n^{(1)}$	$\mathbf{1}4.30212$	$\mathbf{1.1}462888618\cdot 10^{63}$	$\mathbf{1.45}42120372\cdot 10^{1132}$
$n!c_n^{(2)}$	$\mathbf{12}.51435$	$\mathbf{1.13196}02511\cdot 10^{63}$	$\mathbf{1.452394}2721\cdot 10^{1132}$
$n!c_n$	$\mathbf{12}$	$\mathbf{1.1319677968}\cdot 10^{63}$	$\mathbf{1.4523943224}\cdot 10^{1132}$

2. Unary–binary Trees. The function

$$U(z) = \frac{1-z-\sqrt{(1+z)(1-3z)}}{2z}$$

is the OGF of unary–binary trees enumerated by Motzkin numbers. It is singular at $z=-1$ and $z=\frac{1}{3}$, the dominant singularity being at $z=\frac{1}{3}$. By branching properties of the square-root function, $U(z)$ is analytic in a Δ–domain like the one depicted below:



Around the point $\frac{1}{3}$, a singular expansion holds, which translates into an asymptotic expression for $U_n := [z^n]U(z)$:

$$\begin{cases} U(z) &= 1 - 3^{1/2}\sqrt{1-3z} + \mathcal{O}((1-3z)^{3/2}) \\ U_n &= \sqrt{\dfrac{3}{4\pi n^3}}\,3^n + \mathcal{O}(3^n n^{-5/2}). \end{cases}$$

Further terms in the singular expansion of $U(z)$ at $z = \frac{1}{3}$ provide additional terms in the asymptotic expression of the Motzkin numbers (U_n), for instance,

$$U_n \sim \sqrt{\frac{3}{4\pi n^3}}\,3^n\left(1 - \frac{15}{16}\,n^{-1} + \frac{505}{512}\,n^{-2} - \frac{8085}{8192}\,n^{-3} + \frac{505659}{524288}\,n^{-4} + O\left(n^{-5}\right)\right).$$

3. Children's Rounds. The function

$$R(z) = \exp(z \log \frac{1}{1-z}) = (1-z)^{-z}$$

is the EGF of certain combinatorial configurations introduced by Stanley [**174**] and nicknamed by him "children's rounds". A round is a labelled set of directed cycles each of which has a center attached:

$$\mathcal{R} = \mathfrak{P}(\mathcal{Z} \star \mathfrak{C}(Z)).$$

An equivalent form of $R(z)$ is

$$R(z) = \frac{1}{1-z}e^{(1-z)\log(1-z)}$$

so that the only singularity is at $z = 1$, where

$$R(z) = \frac{1}{1-z} + \log(1-z) + \mathcal{O}((1-z)^{0.99}).$$

Thus for coefficients,

$$r_n \equiv [z^n]R(z) = 1 - \frac{1}{n} + \mathcal{O}(n^{-1.99}).$$

A more detailed analysis yields

$$r_n = 1 - \frac{1}{n} - \frac{1}{2n^2}(\log n + \gamma - 1) + \mathcal{O}(\frac{\log^2 n}{n^3}),$$

and an expansion to any order can be obtained.

▷ **7.** *The asymptotic shape of the rounds numbers.* A full asymptotic expansion of r_n is of the form

$$r_n \sim 1 - \sum_{j\geq 1}\frac{P_j(\log n)}{n^j},$$

where $P_j(x)$ is a polynomial of degree $j-1$ in x. (The coefficients of P_j are rational combinations of powers of $\gamma, \zeta(2), \ldots, \zeta(j-1)$.) This expansion can be obtained by a computer algebra program to any predetermined order. ◁

4. An elementary function. The final example is meant to show the way rather arbitrary compositions of basic functions can be treated by singularity analysis. Consider the function

$$F(z) = \frac{1}{2}\left[1 - \sqrt{1 - 4\log\frac{1}{1 - \log\frac{1}{1-z}}}\right],$$

which is built as a composition

$$F(z) = C(L(L(z))) \qquad \text{where } C(z) = \frac{1}{2}(1 - \sqrt{1-4z}),\ L(z) = \log\frac{1}{1-z}.$$

(Combinatorially, F is the EGF of trees in which nodes are replaced by cycles of cycles, a rather artificial combinatorial object!)

The problem is to locate the dominant singularity of $F(z)$ and determine its nature, which can be done inductively on the structure of $F(z)$. The dominant positive singularity ρ of $F(z)$ satisfies

$$L(L(\rho)) = \frac{1}{4} \qquad \text{so that} \qquad \rho = 1 - e^{e^{-1/4}-1},$$

since $C(z)$ is singular at $\frac{1}{4}$, $L(z)$ has positive coefficients and it assumes the value ∞ when it becomes singular.

Since $L(L(z))$ is analytic at ρ, a local expansion of $F(z)$ is obtained by composing an expansion of $C(z)$ at $\frac{1}{4}$ with the standard Taylor expansion of $L(L(z))$ at ρ. All computations done, this gives us

$$F(z) = \frac{1}{2} - C_1(\rho - z)^{1/2} + \mathcal{O}((\rho - z)^{3/2}) \text{ with } C_1 = e^{\frac{5}{8} - \frac{1}{2}e^{-1/4}}.$$

In summary, we have found

$$f_n \equiv [z^n]F(z) = \frac{C_1}{2\rho\sqrt{\pi n^3}}\left(\frac{1}{\rho}\right)^n\left[1 + \mathcal{O}(\frac{1}{n})\right],$$

with

$$\rho = 1 - e^{e^{-1/4}-1} \approx 0.198443, \qquad C_1 = e^{\frac{5}{8} - \frac{1}{2}e^{-1/4}} \approx 1.26566.$$

The method clearly applies in a large number of cases to elementary functions of combinatorial analysis that are defined explicitly by composition of exponentials, logarithms, and algebraic roots. Such functions arise systematically from elementary iterative structures studied in Section IV.3.3, and singularity analysis can often be employed in order to refine the exponential growth estimates globally obtained there.

▷ **8.** *The asymptotic number of trains.* Combinatorial trains have been introduced in Section IV.3.3 as a way to exemplify the power of complex asymptotic methods. One finds that at its dominant singularity ρ, the EGF $Tr(z)$ is of the form $Tr(z) \sim C/(1 - z/\rho)$, and, by singularity analysis,

$$[z^n]Tr(z) \sim 0.11768\,31406\,15497 \cdot 2.06131\,73279\,40138^n.$$

(This asymptotic approximation is good to 15 significant digits for $n = 50$, in accordance with the fact that the dominant singularity is a simple pole.) ◁

VI. 5. Inversion and implicitly defined functions

Recursively defined structures lead to functional equations whose solutions may, in many cases, be analysed locally near singularities. A common pattern in this context is the appearance of singularities of the square-root type, which proves to be universal for a broad class of problems involving trees and tree-like structures. Accordingly, by singularity analysis, the square-root singularity induces in coefficients subexponential terms of the form $n^{-3/2}$.

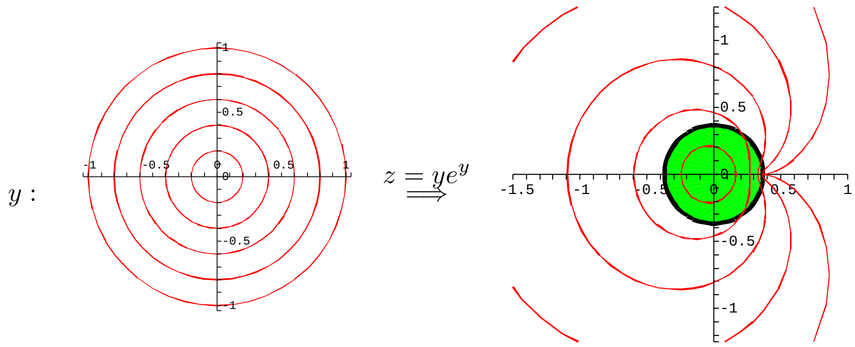


FIGURE 7. The images of concentric circles by the mapping $y \mapsto z = ye^{-y}$. It is seen that $y \mapsto z = ye^{-y}$ is injective on $|y| \leq 1$ with an image extending beyond the circle $|z| = e^{-1}$ [in grey], so that the inverse function $y(z)$ is analytically continuable in a Δ–domain around $z = e^{-1}$.

Inverse functions. We return to the analysis of the coefficients of a function defined implicitly by an equation

$$y(z) = z\phi(y(z)) \qquad \text{or equivalently} \qquad z = \frac{y(z)}{\phi(y(z))}. \tag{28}$$

Here again, it is assumed that $\phi(u)$ is a function with nonnegative coefficients and $\phi(0) \neq 0$.

The problem of solving (28) is one of functional inversion. We have seen in Chapter IV that *an analytic function admits locally an analytic inverse if and only if its first derivative is nonzero*. Set $\psi(u) = u/\phi(u)$; the equation $\psi'(u) = 0$ has, under the conditions of the problem, at most one positive solution strictly within the disc of convergence of ϕ, which satisfies

$$\phi(\tau) - \tau\phi'(\tau) = 0. \tag{29}$$

We assume from now this quantity τ to exist. For reasons already discussed (see Section 6 of Chapter IV) the radius of convergence of $y(z)$ is the corresponding positive value ρ of z such that $y(\rho) = \tau$, that is to say

$$\rho = \frac{\tau}{\phi(\tau)} = \frac{1}{\phi'(\tau)}.$$

The analysis now needs to be more precise in three respects:

(i) all the dominant singularities are to be located;
(ii) analyticity of $y(z)$ in a Δ–domain must be established;
(iii) a singular expansion needs to be determined.

These points are somewhat intertwined.

The situation corresponding to the function $\phi(u) = e^u$, so that $y(z) = ze^{y(z)}$ (the Cayley generating function), is typical of the general situation. From (29), the radius of convergence of $y(z)$ is $\rho = e^{-1}$ corresponding to $\tau = 1$. The image of a circle in the y–plane, centered at the origin and having radius $r < 1$, by the function ye^{-y} is a curve of the z–plane that properly contains the circle $|z| = re^{-r}$, see Figure 7, as $\phi(y) = e^y$ that has nonnegative coefficients satisfies

$$\phi(re^{i\theta}) \le \phi(r) \qquad \text{for all } \theta \in [-\pi, +\pi].$$

Furthermore, the inequality is strict except for $\theta \neq 0$. The image of the circle of radius 1 is a curve $\mathcal{C}$ that has a cusp at $\rho = e^{-1}$ since the first derivative of $y/\phi(y)$ vanishes there.

This geometry shows that the inverse function of $y/\phi(y)$, that is $y(z)$, is uniquely defined for z inside $\mathcal{C}$. Thus, $y(z)$ is Δ–analytic. A singular expansion for $y(z)$ is then simply derived from reversion of the power series expansion of $z = ye^{-y}$. We have

$$ye^{-y} = e^{-1} - \frac{e^{-1}}{2}(y-1)^2 + \frac{e^{-1}}{3}(y-1)^3 - \frac{e^{-1}}{8}(y-1)^4 + \cdots,$$

so that solving for y gives

$$y - 1 = \sqrt{2}(1-ez)^{1/2} + \frac{2}{3}(1-ez) + \mathcal{O}((1-ez)^{3/2}).$$

The discussion of the general case follows the same principles. The relation between z and y, in the vicinity of $(z,y) = (\rho,\tau)$, may be put under the form (see Chapter IV, Section 6),

$$\rho - z = H(y), \qquad \text{where} \quad H(y) := \left(\frac{\tau}{\phi(\tau)} - \frac{y}{\phi(y)}\right), \tag{30}$$

the function $H(y)$ in the right hand side being such that $H'(\tau) = 0$. Thus, the dependency between y and z is locally a quadratic one:

$$\rho - z = \frac{1}{2!}H''(\tau)(y-\tau)^2 + \frac{1}{3!}H'''(\tau)(y-\tau)^3 + \cdots.$$

This relation can be locally inverted: first extract square roots and derive

$$-\sqrt{\rho - z} = \sqrt{\frac{H''(\tau)}{2}}(y-\tau)\left[1 + c_1(y-\tau) + c_2(y-\tau)^2 + ...\right].$$

(The determination with a $-\sqrt{\ }$ should be chosen there as $y(z)$ increases to τ^- as $z \to \rho^-$.) This implies, by solving with respect to $y - \tau$:

$$y - \tau = -d_1(\rho - z)^{1/2} + d_2(\rho - z) + d_3(\rho - z)^{3/2} + \cdots \text{ with } d_1 = (2/H''(\tau))^{1/2}.$$

PROPOSITION VI.1. *Let ϕ be a function analytic at 0 having nonnegative Taylor coefficients with $\phi(0) = 0$, and such that there exists a positive solution τ to the characteristic equation,*

$$\phi(\tau) - \tau\phi'(\tau) = 0,$$

strictly within the disc of convergence of ϕ. Let $y(z)$ be the solution analytic at the origin of $y(z) = z\phi(y(z))$. Then $y(z)$ has a dominant singularity at

$$z = \rho \qquad \text{where} \qquad \rho = \frac{\tau}{\phi(\tau)}.$$

The singular expansion of y at ρ is of the form

$$y(z) = \tau + \sum_{j=1}^{\infty} d_j^*(1 - \frac{z}{\rho})^{j/2}.$$

for some computable constants d_j^*. In particular, one has

$$d_1^* = -\sqrt{\frac{2\phi(\tau)}{\phi''(\tau)}}.$$

▷ **9.** *Computability of the singular expansion.* Define

$$h(w) := \sqrt{\frac{\tau/\phi(\tau) - w/\phi(w)}{(\tau - w)^2}},$$

so that $y(z)$ satisfies $\sqrt{\rho - z} = (\tau - y)h(y)$. The singular expansion of y can then be deduced by Lagrange inversion from the expansion of the negative powers of $h(w)$ at $w = \tau$. This technique yields for instance explicit forms for coefficients in the singular expansion of $y = ze^y$. ◁

A simple example like $\phi(u) = 1 + u^2$ for which

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

shows that it need not however be true in all generality y has a unique dominant singularity: here there are two conjugate singularities, $-\frac{1}{2}$ and $+\frac{1}{2}$. However, the conditions for this to happen are rather simple. Let us say that a power series $h(u)$ is d–periodic if $h(u) = u^e k(u^d)$ for some power series k, with d maximal. A function is called here periodic if it is d–periodic from some $d \geq 2$ and aperiodic otherwise. An elementary argument developed in Note 10 shows that that periodicity does not occur for $y(z)$ unless $\phi(u)$ is itself periodic, in which case $y(z) = zw(z^d)$ for some analytic $w(z)$, when $\phi(u) = \psi(u^d)$ for some $d \geq 2$.

From singularity analysis, we get directly:

THEOREM VI.4 (Coefficients of tree functions). *Assume the conditions of Proposition VI.1, and, additionally, that $\phi(w)$ is aperiodic. The coefficients of the solution $y(z)$ to $y = z\phi(y)$ satisfy*

$$[z^n]y(z) \sim \frac{d_1^* \rho^{-n}}{2\sqrt{\pi n^3}} \left[1 + \sum_{k=1}^{\infty} \frac{e_k}{n^k}\right],$$

for some effectively computable coefficient sequence e_k.

In case $\phi(u)$ is d–periodic for some $d \geq 2$, the additional congruence condition, $n \equiv 1\ (d)$, must be imposed for an asymptotic expansion as the other coefficients all vanish. The dominant singularities are at $\rho e^{2ij\pi/d}$ and their contributions must be added up in accordance with the discussion of the next section (details left as an exercise).

▷ **10.** *Periodicities.* Assume that $\phi(u) = \psi(u^d)$with ψ analytic at 0. Let $y = y(z)$ be the root of $y = z\phi(y)$. Set $Z = z^d$ and let Y be the root of $Y = Z\psi(Y)^d$. One has by construction $y(z) = Y(z^d)^{1/d}$, given that $y^d = z^d\phi(y)^d$. Since $Y(Z) = Y_1 Z + Y_2 Z^2 + \cdots$, we verify that the nonzero coefficients of $y(z)$ are amongst those of index $1, 1+d, 1+2d, \ldots$.

If d is chosen maximal, then $\psi(u)^d$ is aperiodic. Thus, Theorem VI.4 applies to $Y(Z)$. The function $Y(Z)$ is Δ–regular (relative to its dominant singularity at $\rho^{1/d}$), and it has a square root singularity there and none other on $|Z| = \rho^{1/d}$. Also, since $Y = z\psi(Y)^{1/d}$, $Y(Z)$ cannot vanish on $|Z| \leq \rho^{1/d}$, $\mathbb{Z} \neq 0$. Thus, $[Z^\nu]Y(Z)^{1/d}$ is analytic in $|Z| \leq \rho^{1/d}$, except at $\rho^{1/d}$ wheer it has a $\sqrt{}$ branch point. All computations done, we find that

$$[z^n]y(z) \sim d \cdot \frac{d_1^* \rho^{-n}}{2\sqrt{\pi n^3}} \qquad \text{when} \quad n \equiv 1 \pmod{d}.$$

This is a kind of Perron-Frobenius property for periodic tree functions. ◁

Here is a table of the most basic varieties of simple trees and the corresponding asymptotic estimates found in this way:

Type	$\phi(u)$	*Sing. expansion of* $y(z)$	y_n
binary	$(1+u)^2$	$1-4\sqrt{\frac{1}{4}-z}+\cdots$	$\frac{4^n}{\sqrt{\pi n^3}}+O(n^{-5/2})$
unary-binary	$1+u+u^2$	$1-3\sqrt{\frac{1}{3}-z}+\cdots$	$\frac{3^{n+1/2}}{2\sqrt{\pi n^3}}+O(n^{-5/2})$
general	$(1-u)^{-1}$	$\frac{1}{2}-\sqrt{\frac{1}{4}-z}$	$\frac{4^{n-1}}{\sqrt{\pi n^3}}+O(n^{-5/2})$
Cayley	e^u	$1-\sqrt{2e}\sqrt{e^{-1}-z}+\cdots$	$\frac{e^n}{\sqrt{2\pi n^3}}+O(n^{-5/2})$

Combining Proposition VI.1 with methods of the previous section, we have available a method that permits us to analyse in turn $[z^n]f(y(z))$, for a wide class of implicitly defined $y(z)$. This observation will be put to good use in Chapter VII, when analysing a variety of tree parameters.

▷ **11.** *Stirling's formula via singularity analysis.* Since the solution to $Y = ze^Y$ analytic at 0 satisfies $[z^n] = n^{n-1}/n!$ (by Lagrange inversion) and, at the same time, its singularity is known from Proposition VI.1, we have:

$$\frac{n^{n-1}}{n!} \sim \frac{e^n}{\sqrt{2\pi n^3}}\left(1-\frac{1}{12}\,n^{-1}+\frac{1}{288}\,n^{-2}+\frac{139}{51840}\,n^{-3}-\cdots\right).$$

Thus Stirling's formula *also* results from singularity analysis. ◁

VI. 6. Singularity analysis and closure properties

At this stage, we have available composition rules for singular expansions of the types previously considered under operations like $\pm$, $\times$, $\div$. These are induced by corresponding rules for extended formal power series, where generalized exponents and logarithmic factors are allowed. In the previous section, we have also seen that inversion of usually gives rise to square-root singularities.

In this section we examine first the rôle of functional composition, then we show that generating functions amaneable to singularity analysis are closed under differentiation and integration.

VI. 6.1. Functional composition. Let f and g be functions analytic at the origin with nonnegative coefficients, and consider the composition

$$h = f \circ g, \qquad h(z) = f(g(z)).$$

Let ρ_f, ρ_g, ρ_h be the corresponding radii of convergence, and let $\tau_f = f(\rho_f)$, and so on. We shall asume that f and g are Δ–continuable and that they admit singular expansions in the scale of powers. There are three cases to be distinguished depending on how τ_g compares to ρ_f. Clearly one has:

- *Supercritical case*, when $\tau_g > \rho_f$. In that case, when z increases from 0, there is a value r strictly less than ρ_g such that $g(r)$ attains the value ρ_f, which triggers a singularity of $f \circ g$. In other words $r \equiv \rho_h = g^{(-1)}(\rho_f)$. Around this point, g is analytic and a singular expansion of $f \circ g$ is obtained by composing the singular expansion of f with the regular expansion of g at r. *The singularity type is that of the external function* (f).
- *Subcritical case*, when $\tau_g < \rho_f$. In this dual situation, the singularity of $f \circ g$ is driven by that of the inside function g. We have $\rho_h = \rho_g$, $\tau_h = f(\rho_g)$ and the singular expansion of $f \circ g$ is obtained by composing the regular expansion of f

with the singular expansion of g at ρ_g. *The singularity type is that of the internal function* (g).

— *Critical case*, when $\tau_g = \rho_f$. In this boundary case, there is a confluence of singularities. We have $\rho_h = \rho_g$, $\tau_h = \tau_f$, and the the singular expansion is obtained by composition rules of the singular expansions. *The singularity type is a mix of the types of the internal and external functions* (f, g).

This terminology extends the notion of supercritical sequence schema introduced in Chapter V, where we considered the case $f(z) = (1-z)^{-1}$ and discussed some of the probabilistic consequences. Rather than stating general conditions that would be unwieldy, it is better to discuss examples directly, referring to the above guidelines supplemented by the plain algebra of generalized power expansions, whenever necessary.

EXAMPLE 1. *Combinatorial sums.* Based on the discussion above, a reasonably general strategy for the asymptotic analysis of a class of combinatorial sums of the form

$$S_n = \sum_{k=1}^{n} f_k g_n^{(k)}.$$

can be developped. There f_k is a sequence of numbers, usually of a simple form and called the *weights*, while the $g_n^{(k)}$ are a triangular array of numbers, for instance Pascal's triangle.

For the weights f_k we shall consider expressly sequences such that $f(z)$ is Δ–analytic with a singular expansion involving functions of the standard scale of Theorems VI.1, VI.2, VI.3. Typical examples for $f(z)$ and (f_k) are[2]

(31)

$\frac{1}{k}$	$\frac{1}{4^k}\binom{2k}{k}$	1	H_k	k	k^2
$\log\frac{1}{1-z}$	$\frac{1}{\sqrt{1-z}}$	$\frac{1}{1-z}$	$\frac{1}{1-z}\log\frac{1}{1-z}$	$\frac{z}{(1-z)^2}$	$\frac{z+z^2}{(1-z)^3}$.

The triangular arrays $g_n^{(k)}$ discussed here are taken here to arise as coefficients of the *powers* of some fixed function,

$$g_n^{(k)} = [z^n](g(z))^k \qquad \text{where} \qquad g(z) = \sum_{n=1}^{\infty} g_n z^n,$$

with $g(z)$ an analytic function at the origin having non–negative coefficients and satisfying $g(0) = 0$. Examples are

(32)

$\frac{z}{1-z}$	ze^z	$z(1+z)$	$\frac{1-\sqrt{1-4z}}{2}$	$\frac{1-2z-\sqrt{1-4z}}{2z}$	$T(z)\ [T = ze^T]$
$\binom{n-1}{k-1}$	$\frac{k^{n-k}}{(n-k)!}$	$\binom{k}{n-k}$	$\frac{k}{n}\binom{2n-k-1}{n-1}$	$\frac{k}{n}\binom{2n}{n-k}$	$k\frac{n^{n-k-1}}{(n-k)!}$

An interesting class of such arrays arises from the Lagrange inversion theorem. Indeed, if $g(z)$ is implicily defined by $g(z) = zG(g(z))$, one has $g_{n,k} = \frac{k}{n}[w^{n-k}]G(z)^n$. (For instance, the last three cases of (32) are obtained in this way by taking $G(w)$ as $1/(1-w), (1+w)^2, e^w$.)

By design, the generating function of the S_n is simply

$$S(z) = \sum_{n=0}^{\infty} S_n z^n = f(g(z)) \qquad \text{with} \qquad f(z) = \sum_{k=0}^{\infty} f_k z^k.$$

[2]Weights like $\log k$, $\sqrt{k}$, $1/(k^2+1)$, etc, also satisfy these conditions [**86, 66**], but the proofs require advanced techniques discussed below.

Thus the asymptotic analysis of S_n can be directly based on the general discussion of composition of singularities of $f(z)$ and $g(z)$.

A. Bernoulli sums. Let ϕ be a function mapping $\mathbb{Z}_{\geq 0}$ into itself and set $f_k := \phi(k)$. Consider the sums

$$S_n := \sum_{k=0}^{n} \phi(k) \frac{1}{2^n} \binom{n}{k}.$$

If X_n is a Bernoulli random variable of $\mathrm{Bern}(n, \frac{1}{2})$, then $S_n = \mathbb{E}(\phi(X_n))$ is exactly the expectation of $\phi(X_n)$. Then, with previous notations, we find for the ogf of the sequence (S_n):

$$S(z) = \frac{2}{2-z} f\left(\frac{z}{2-z}\right).$$

Considering weights whose ogf has, like in (31) radius of convergence 1, what we have is a variant of the composition schema, with an additional prefactor. The composition scheme is of the *subcritical type* since the function $g(z) = z/(2-z)$ has radius of convergence equal to 2. The singularities of $S(z)$ are then of the same type as those of the weight gf $f(z)$ and one verifies, in all cases of (31), that, to first asymptotic order, $S_n \sim \phi(n/2)$: this is in agreement with the fact that the binomial distribution is concentrated near its mean $\frac{n}{2}$. Singularity analysis provides additionally complete asymptotic expansions, for instance,

$$\begin{aligned} \mathbb{E}(\frac{1}{X_n} \mid X_n > 0) &= \frac{2}{n} + \frac{2}{n^2} + \frac{6}{n^3} + O(n^{-4}) \\ \mathbb{E}(H_{X_n}) &= \log\frac{n}{2} + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + O(n^{-3}). \end{aligned}$$

See [**60, 66**] for more along these lines.

▷ **12.** *General Bernoulli sums.* Let $X_n \in \mathrm{Bern}(n;p)$ be a general Bernoulli random variable,

$$\mathbb{P}(X_n = k) = \binom{n}{k} p^k q^{n-k}, \qquad q = 1-p.$$

Then with $f_k = \phi(k)$, one has

$$\mathbb{E}(\phi(X_n)) = [z^n] \frac{1}{1-qz} f\left(\frac{pz}{1-qz}\right),$$

so that the analysis develops as in the case $\mathrm{Bern}(n; \frac{1}{2})$. ◁

B. Generalized Knuth–Ramanujan Q-functions. For reasons motivated by analysis of algorithms, Knuth has encountered repeatedly sums of the form

$$Q_n(\{f_k\}) = f_0 + f_1 \frac{n-1}{n} + f_2 \frac{(n-1)(n-2)}{n^2} + \cdots.$$

(See, e.g., [**123**, pp. 305–307].) There (f_k) is a sequence of coefficients (usually of at most polynomial growth). For instance, we have seen in Chapter II, Section 3 that the case $f_k \equiv 1$ yields the expected time till the first collision in the birthday paradox problem.

A closer examination shows that the analysis of such Q_n is reducible to singularity analysis. Writing

$$Q_n(\{f_k\}) = f_0 + \frac{n!}{n^{n-1}} \sum_{k \geq 1} f_k \frac{n^{n-k-1}}{(n-k)!}$$

reveals the closeness with the last column of (32). Indeed, setting

$$F(z) = \sum_{k \geq 1} \frac{f_k}{k} z^k,$$

one has ($n \geq 1$)

$$Q_n = f_0 + \frac{n!}{n^{n-1}}[z^n]S(z) \qquad \text{where} \quad S(z) = F(T(z)),$$

and $T(z)$ is the Cayley tree function ($T = ze^T$).

For weights $f_k = \phi(k)$ that are of polynomial growth, the schema is *critical*. Thus, the singular expansion $S(z)$ is obtained by composing the singular expansion of f with the expansion of T, namely, $T \sim 1 - \sqrt{2}\sqrt{1-ez}$. For instance, $\phi(k) = k^r$ for some integer $r \geq 1$ leads to $F(z)$ that has an rth order pole at $z = 1$. Then, the singularity type of $F(T(z))$ is of the form $Z^{-r/2}$ where $Z = (1-ez)$, which is reflected by $S_n \asymp e^n n^{r/2-1}$. After the final normalization, we see that $Q_n \asymp n^{(r+1)/2}$. Globally, for many weights of the form $f_k = \phi(k)$, we expect Q_n to be of the rough form $\sqrt{k}\phi(\sqrt{n})$, which agrees with the fact that the expectation of the first collision in the birthday problem is on average near $\sqrt{\pi n/2}$. □

▷ **13.** *Higher moments of the birthday problem.* Take the model where there are n days in the year and let B be the random variable representing the first birthday collision. Then $\mathbb{P}_n(B > k) = k!n^{-k}\binom{n}{k}$, and

$$\mathbb{E}_n(\Phi(B)) = \Phi(1) + Q_n(\{\Delta\Phi(k)\}), \qquad \text{where} \quad \Delta\Phi(k) := \Phi(k+1) - \Phi(k).$$

For instance $\mathbb{E}_n(B) = 1 + Q_n(\langle 1, 1, \ldots \rangle)$. We thus get moments of various functionals (here stated to two asymptotic terms):

$\Phi(x)$	x	$x^2 + x$	$x^3 + x^2$	$x^4 + x^3$
$E_n(\Phi(B))$	$\sqrt{\frac{\pi n}{2}} + \frac{2}{3}$	$2n + 2$	$3\sqrt{\frac{\pi n^3}{2}} - 2n$	$8n^2 - 7\sqrt{\frac{\pi n^3}{2}}$

via singularity analysis. ◁

▷ **14.** *How to weigh an urn? The "shake-and-paint" algorithm.* You are given an urn containing an unknown number N of identical looking balls. How to estimate this number in much fewer than $O(N)$ operations? A probabilistic solution due to Brassard and Bratley [**29**] uses a brush and paint. Shake the urn, pull out a ball, then mark it with paint and replace it into the urn. Repeat until you find an already painted ball. Let X be the number of operations. One has $\mathbb{E}(X) \sim \sqrt{\pi N/2}$. Further more the quantity $Y := X^2/2$ constitues, by the previous note, an asymptotically unbiased estimator of N, in the sense that $\mathbb{E}(Y) \sim N$. In other words, count the time till an already painted ball is first found, and return half of the square of this time. One also has $\sqrt{\mathbb{V}(Y)} = N$. By performing the experiment m times (using m different colours of paint) and by taking the arithmetic average of the m estimates, one obtains an unbiased estimator whose typical accuracy is $\sqrt{1/m}$. For instance, $m = 16$ gives an expected accuracy of 25%. (Similar principles are used in the design of data mining algorithms.) ◁

▷ **15.** *Catalan sums.* These are defined by

$$S_n := \sum_{k \geq 0} f_k \binom{2n}{n-k}, \qquad S(z) = \frac{1}{\sqrt{1-4z}} f\left(\frac{1-2z-\sqrt{1-4z}}{2z}\right).$$

The case when $\rho_f = 1$ corresponds to a critical composition and it can be discussed much in the same way as the Ramanujan sums. ◁

EXAMPLE 2. *"Supertrees"*. Let $\mathcal{G}$ be the class of general Catalan trees with ogf $G(z) = \frac{1}{2}(1 - \sqrt{1-4z})$. Its radius of convergence is $\frac{1}{4}$ and its singular value is $G(\frac{1}{4}) = \frac{1}{2}$. Consider the two generating functions:

$$H(z) = G(zG(z)), \qquad K(z) = G(2zG(z)).$$

The function $zG(z)$ is the ogf of planted trees, that is trees such that to the root is attached a stem and an extra node, corresponding to the specification $\mathcal{Z}\mathcal{G}$. Then, $H(z)$ is the ogf of the class $\mathcal{H} = \mathcal{G}[\mathcal{Z}\mathcal{G}]$ of trees such that, on each node there is grafted a planted tree (by the combinatorial substitution of Chapter I) —we shall call such objects "*supertrees*". The ogf $K(z)$ similarly corresponds to the case when the stems can be of any two colours, $\mathcal{H} = \mathcal{G}[(\mathcal{Z}+\mathcal{Z}')\mathcal{G}]$. Combinatorial sum expressions are available:

$$H_n = \sum_{k=1}^{n} \frac{1}{k}\binom{2k-2}{k-1}\binom{2n-k-1}{n-1}, \qquad K_n = \sum_{k=1}^{n} \frac{2^k}{k}\binom{2k-2}{k-1}\binom{2n-k-1}{n-1}.$$

Since $\rho_G = \frac{1}{4}$ and $\tau_G = \frac{1}{2}$, the composition scheme is subcritical in the case of $\mathcal{H}$ and critical in the case of $\mathcal{K}$. In the first case, the singularity is of square-root type and one finds easily:

$$H(z) \underset{z\to\frac{1}{4}}{\sim} \frac{2-\sqrt{2}}{4} - \frac{1}{\sqrt{8}}\sqrt{\frac{1}{4}-z}, \qquad H_n \sim \frac{4^n}{8\sqrt{2\pi n}}.$$

In the second case, one has

$$K(z) \underset{z\to\frac{1}{4}}{\sim} \frac{1}{2} - \frac{1}{\sqrt{2}}(\frac{1}{4}-z)^{1/4}, \qquad K_n \sim \frac{4^n}{8\Gamma(\frac{3}{4})n^{5/4}}.$$

The occurrence of the exponent $\frac{5}{4}$ is striking. This examples shows that asymptotic terms of the type $n^{p/q}$ with $q \neq 1, 2$ may well appear in elementary combinatorics and coefficients of simple algebraic functions. Such situations tend to be associated with nonstandard limit laws, akin to the stable distributions of probability theory [**13**]. □

EXAMPLE 3. *Supercritical cycle schema.* Consider the scheme $\mathcal{H} = \mathfrak{C}(\mathcal{G})$ which forms labelled cycles from basic components of $\mathcal{G}$. The egfs are related by

$$H(z) = \log \frac{1}{1-G(z)}.$$

Consider the case where G attains the value 1 before becoming singular, that is, $\tau_G > 1$. Then, this corresponds to a supercritical composition schema. This case can be discussed much in the same way as the supercritical sequence schema (Chapter V), with a logarithmic singularity replacing a polar singularity.

Let $\sigma := \rho_H$, which is determined by $G(\sigma) = 1$. First, one finds:

$$H(z) \underset{z\to\sigma}{\sim} \log\frac{1}{1-z/\rho} - \log(\sigma G'(\sigma)) + A(z),$$

where $A(z)$ is analytic at $z = \sigma$. Thus:

$$[z^n]H(z) \sim \frac{\sigma^{-n}}{n}.$$

(The error term implicit in this estimate is exponentially small).

The bgf $H(z,u) = \log(1-uG(z))^{-1}$ has the variable u marking the number of components in $\mathcal{H}$-objects. In particular, the mean number of components in a random $\mathcal{H}$-object of size is $\sim \lambda n$, where $\lambda = 1/(\sigma G'(\sigma))$, and the distribution is concentrated around its mean. Similarly, the mean number of components with size k in a random $\mathcal{H}_n$ object is found to be asymptotic to $\lambda g_k \sigma^k$, where $g_k = [z^k]G(z)$, □

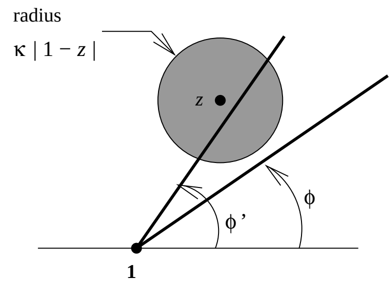


FIGURE 8. The geometry of the contour $\gamma(z)$ used in the proof of the differentiation theorem.

VI. 6.2. Differentiation and integration. Functions amenable to singularity analysis are closed under differentiation[3]. This is once more in sharp contrast with real analysis. The following statement is a version tuned to our needs of well-known differentiability properties of complex asymptotic expansions (see, e.g., Olver's book [**150**, p. 9]). For simplicity, we restrict attention to functions whose singular expansion is of the form

$$f(z) = \sum_{j=0}^{J} c_j (1-z)^{\alpha_j} + O((1-z)^A). \tag{33}$$

THEOREM VI.5 (Singular differentiation). *If $f(z)$ is Δ-regular and admits a singular expansion near its singularity in the sense of* (33), *then for each integer $r > 0$, $\frac{d^r}{dz^r} f(z)$ is also Δ-regular and admits an expansion obtained through term-by-term differentiation:*

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

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

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

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

Let $g(z)$ be a function that is regular in a domain $\Delta(\phi, \eta)$ where it is assumed to satisfy $g(z) = O((1-z)^A)$ for $z \in \Delta$. Choose a subdomain $\Delta' := \Delta(\phi', \eta')$, where $\phi < \phi' < \frac{\pi}{2}$ and $0 < \eta' < \eta$. By elementary geometry, for any sufficiently small $\kappa > 0$, the disc of radius $\kappa(z-1)$ centered at a value $z \in \Delta'$ lies entirely in Δ; see Figure 8. We fix such a small value κ and let $\gamma(z)$ represent the boundary of that disc oriented positively.

The starting point is Cauchy's integral formula

$$g'(z) = \frac{1}{2\pi i} \int_C g(w) \frac{dw}{(w-z)^2}, \tag{34}$$

a direct consequence of the residue theorem. Here C should encircle z while lying inside the domain of regularity of g, and we opt for the choice $C \equiv \gamma(z)$. Then trivial bounds applied to (34) give:

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

[3]The presentation of this section is borrowed from [**60**].

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

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

THEOREM VI.6 (Singular integration). *Let $f(z)$ be Δ-regular and admit a Δ-expansion near its singularity in the sense of* (33)*. Then $\int_0^z f(t)\,dt$ is also Δ-regular. Assume that none of the quantities α_j and A equals -1.*

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

$$\int_0^z f(t)\,dt = -\sum_{j=0}^{J} \frac{c_j}{\alpha_j+1}(1-z)^{\alpha_j+1} + O\left((1-z)^{A+1}\right). \tag{35}$$

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

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

where the "integration constant" L_0 has the value

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

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

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

that are consistent with elementary integration, and similar rules are easily derived for powers of logarithms. Furthermore, the corresponding O–transfers hold true. (The proofs are simple modifications of the one given below for the basic case.)

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

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

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

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

(i) *Case* $A < -1$. Straight-line integration between 0 and z, provides (35), as soon as it has been established that

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

By Cauchy's integral formula, we can choose any path of integration that stays within the region of analyticity of r. We choose the contour $\gamma := \gamma_1 \cup \gamma_2$, shown in Figure 9. Then,

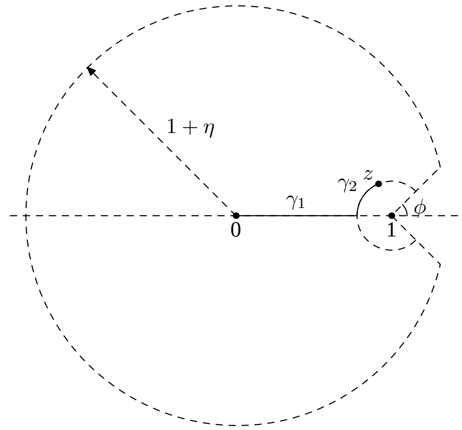


FIGURE 9. The contour used in the proof of the integration theorem.

one has

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

where the symbol $|dt|$ designates the differential line element (often denoted by ds) in the corresponding curvilinear integral. Both integrals are $O(|1-z|^{A+1})$: for the integral along γ_1, this results from explicitly carrying out the integration; for the integral along γ_2, this results from the trivial bound $O(\|\gamma_2\|(1-z)^A)$.

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

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

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

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

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

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

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

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

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

VI. 6.3. Polylogarithms. The generalized *polylogarithm* $\operatorname{Li}_{\alpha,r}$, where α is an arbitrary complex number and r a nonnegative integer is defined for $|z|<1$ by

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

and the notation Li_α abbreviates $\operatorname{Li}_{\alpha,0}$. In particular, one has $\operatorname{Li}_{1,0}(z)=\operatorname{Li}_1(z)$, the usual logarithm. The singular expansion of the polylogarithm, taken from [**66**], involves the Riemann ζ function:

THEOREM VI.7 (Singularities of polylogarithms). *The function* $\operatorname{Li}_{\alpha,r}(z)$ *is* Δ*–continuable and, for* $\alpha\notin\{1,2,\dots\}$*, it satisfies the singular expansion*

$$\operatorname{Li}_{\alpha,0}(z) \sim \Gamma(1-\alpha)w^{\alpha-1}+\sum_{j\ge 0}\frac{(-1)^j}{j!}\zeta(\alpha-j)w^j, \quad w=-\log z=\sum_{\ell=1}^{\infty}\frac{(1-z)^\ell}{\ell}. \tag{36}$$

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

$$\operatorname{Li}_{\alpha,r}(z)=(-1)^r\frac{\partial^r}{\partial\alpha^r}\operatorname{Li}_{\alpha,0}(z), \qquad L(z):=\log\frac{1}{1-z},$$

and corresponding termwise differentiation of (36) *with respect to* α*.*

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

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

Similar expansions hold when α is a positive integer; see [**66**] for details.

PROOF (SKETCH). The proof when $z\to 1^-$ along the real line is a simple consequence of Mellin transform techniques (see APPENDIX: *Mellin transform,* p. 227). Take $r=0$ and set $\Lambda(w)=\operatorname{Li}_\alpha(e^{-w})$. The Mellin transform of Λ is

$$\Lambda^\star(s)=\zeta(s+\alpha)\Gamma(s),$$

since Λ is a harmonic sum. There are poles at $s=0,-1,-2,\dots$ due to the Gamma factor and a pole at $s=1-\alpha$ due to the zeta function. A standard Mellin analysis then yields the estimate (36) when $w\to 0$, at least when z is real.

In order to extend the estimate beyond the disc of convergence $|z|=1$, one starts from a Lindel¨of integral representation of the polylogarithm (Chapter IV),

$$\operatorname{Li}_\alpha(-z)=-\frac{1}{2i\pi}\int_{1/2-i\infty}^{1/2+i\infty}\frac{z^s}{s}\,\frac{\pi}{\sin\pi s}\,ds.$$

Setting $z=-e^{i(w-\pi)}$ and $s=1/2+it$, the integral can then be analysed as a "harmonic integral" (a continuous analogue of harmonic sums) by means of Mellin transforms; see [**66**] for details. $\square$

$\triangleright$ **16.** *Stirling's formula from polylogarithms.* One has $\log n! = [z^n](1-z)^{-1}\operatorname{Li}_{0,1}(z)$, to which singularity analysis is applicable. Theorem VI.7 yields the singular expansion

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

from which Stirling's formula reads off:

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

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

$$1^1 2^2 \cdots n^n \sim A n^{\frac{1}{2}n^2+\frac{1}{2}n+\frac{1}{12}} e^{-\frac{1}{4}n^2}, \quad A := \exp\left(\tfrac{1}{12}-\zeta'(-1)\right) = \exp\left(-\tfrac{\zeta'(2)}{2\pi^2}+\tfrac{\log(2\pi)+\gamma}{12}\right).$$

The constant A is the Glaisher-Kinkelin constant [**61**, p. 135]. ◁

VI. 6.4. Hadamard Products. The *Hadamard product* of two functions $f(z), g(z)$ analytic at the origin is defined as their term-by-term product,

$$f(z) \odot g(z) = \sum_{n\geq 0} f_n g_n z^n, \qquad \text{where} \quad f(z) = \sum_{n\geq 0} f_n z^n, \quad g(z) = \sum_{n\geq 0} g_n z^n. \tag{37}$$

As we are going to see following an article of Fill, Flajolet, and Kapur [**60**], the class of functions amenable to singularity analysis is closed under Hadamard products. Establishing sucha closure property requires methods for composing functions from the basic scale, namely $(1-z)^a$, as well as error terms of the form $O((1-z)^A)$. We address these problems in turn.

The expansion around the origin,

$$(1-z)^a = 1 + \frac{-a}{1}z + \frac{(-a)(-a+1)}{2!}z^2 + \cdots, \tag{38}$$

gives through term-by-term multiplication

$$(1-z)^a \odot (1-z)^b = {}_2F_1[-a,-b;1;z]. \tag{39}$$

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

$${}_2F_1[\alpha,\beta;\gamma;z] = 1 + \frac{\alpha\beta}{\gamma}\frac{z}{1!} + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}\frac{z^2}{2!} + \cdots. \tag{40}$$

From the transformation theory of hypergeometrics, see e.g. [**191**, Ch XIV], we know that, in general, hypergeometric functions can be expanded in the vicinity of $z=1$ by means of the $z \mapsto 1-z$ transformation. Instantiation of this transformation with $\gamma = 1$ yields

$$\begin{aligned} {}_2F_1[\alpha,\beta;1;z] = {} & \frac{\Gamma(1-\alpha-\beta)}{\Gamma(1-\alpha)\Gamma(1-\beta)}\, {}_2F_1[\alpha,\beta;\alpha+\beta;1-z] \\ & + \frac{\Gamma(\alpha+\beta-1)}{\Gamma(\alpha)\Gamma(\beta)}(1-z)^{-\alpha-\beta+1}\, {}_2F_1[1-\alpha,1-\beta;2-\alpha-\beta;1-z]. \end{aligned} \tag{41}$$

From theer, we state:

PROPOSITION VI.2 (Singularity analysis of Hadamard products). (i) *When a, b, and $a+b$ are not integers, the Hadamard product $(1-z)^a \odot (1-z)^b$ has an infinite Δ-expansion with exponent scale $\{0,1,2,\ldots\} \cup \{a+b+1, a+b+2, \ldots\}$, namely,*

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

where the coefficients λ and μ are given by

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

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

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

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

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

$$(f \odot g)(z) = O((1-z)^{a+b+1}). \tag{42}$$

Part (ii) is proved by means of contour integration techniquesin [**60**]. Globally, Theorem VI.2 establishes the closure under Hadamard products of functions amenable to singularity analysis in the sense of (33). The treatment of boundary cases and of logarithmic factors is discussed in [**60**].

EXAMPLE 4. *Pólya's drunkard problem.* (This example is taken from [**60**].) In the d–dimensional lattice $\mathbb{Z}^d$ of points with integer coordinates, the drunkard performs a random walk starting from the origin with steps in $\{-1,+1\}^d$, each taken with equal likelihood. The probability that the drunkard is back at the origin after $2n$ steps is

$$q_n^{(d)} = \left(\frac{1}{2^{2n}}\binom{2n}{n}\right)^d, \tag{43}$$

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

$$\left(1-\sum_{n=1}^{\infty} p_n^{(d)} z^n\right)^{-1} = \sum_{n=0}^{\infty} q_n^{(d)} z^n, \tag{44}$$

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

The asymptotic analysis of the q_n's is straightforward; the one of the p_n's is more involved and is of interest in connection with recurrence and transience of the random walk; see, e.g., [**49, 130**]. The Hadamard closure theorem provides a direct access to this problem. Define

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

Then, Equations (43) and (44) imply:

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

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

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

$$P(z) = 1-\sqrt{1-z}, \qquad \text{implying} \quad p_n^{(1)} = \frac{1}{n2^{2n-1}}\binom{2n-2}{n-1} \sim \frac{1}{2\sqrt{\pi n^3}}.$$

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

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

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

and reconstruct the only singular expansion that is compatible, namely

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

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

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

so that, by Theorems VI.2 and VI.3, one has

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

(See the study by Louchard *et al.* [**136**, Sec. 4] for somewhat similar calculations.)

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

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

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

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

leading to

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

By singularity analysis, the last expansion gives

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

A complete asymptotic expansion in powers $n^{-3/2}, n^{-5/2}, \ldots$ can be obtained by the same devices. In particular this improves the error term above to $O(n^{-5/2})$. The explicit form of $Q(1)$ results from its expression as the generalized hypergeometric ${}_3F_2[\frac{1}{2}, \frac{1}{2}, \frac{1}{2}; 1, 1; 1]$, which evaluates by Clausen's theorem and Kummer's identity to the square of a complete elliptic integral. (See the papers by Larry Glasser for context, for instance [**93**]; nowadays, several computer algebra systems even provide this value automatically.)

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

VI. 7. Multiple singularities

The basic principle for a function with multiple dominant singularities parallels the situation of rational and meromorphic functions: *the contributions from each singularity must be added up.*

Let us demonstrate the *modus operandi* before stating the general theorem. Take for instance the function

$$g(z) = \frac{e^z}{\sqrt{1-z^2}}. \tag{45}$$

There are two singularities at $z = +1$ and $z = -1$, with

$$g(z) \sim \frac{e}{\sqrt{2}\sqrt{1-z}} \quad z \to +1 \qquad \text{and} \qquad g(z) \sim \frac{e^{-1}}{\sqrt{2}\sqrt{1+z}} \quad z \to -1.$$

We have

$$[z^n]\frac{e}{\sqrt{2}\sqrt{1-z}} \sim \frac{e}{\sqrt{2\pi n}} \qquad \text{and} \qquad [z^n]\frac{e^{-1}}{\sqrt{2}\sqrt{1+z}} \sim \frac{e^{-1}(-1)^n}{\sqrt{2\pi n}}.$$

To get the coefficient $[z^n]g(z)$, it proves justified to add up these two contributions (by Theorem VI.8 below), so that

$$[z^n]g(z) \sim \frac{1}{\sqrt{2\pi n}}[e + (-1)^n e^{-1}].$$

The justification for this process is provided by the technical result below.

THEOREM VI.8 (Multiple singularities). *Let $f(z)$ have a finite number of singularities on the unit circle at points $\zeta_j = e^{i\theta_j}$, for $j = 1\,..\,r$.*

(H_1). *Let Δ_0 be a Δ– domain. Assume that $f(z)$ is analytic in the indented disk*

$$\mathcal{D} = \bigcap_{j=1}^{r} (\zeta_j \cdot \Delta_0),$$

with $\zeta \cdot \Delta_0$ the image of Δ_0 by the rotation $z \mapsto \zeta z$.

(H_2). *Assume that there are r functions $h_1(z), \ldots, h_r(z)$ with $h_j(z)$ analytic in a circular neighbourhood of ζ_j such that as $z \to \zeta_j$ in $\mathcal{D}$, one has*

$$f(z) = h_j(z) + \mathcal{O}\left((1 - \frac{z}{\zeta_j})^{-\alpha}\right).$$

Under these assumptions, the coefficients of $f(z)$ satisfy

$$[z^n]f(z) = \mathcal{O}(n^{\alpha-1}).$$

A function analytic in a domain like $\mathcal{D}$ is sometimes said to be star–continuable, a notion that is the natural generalization of Δ–analyticity for functions with several dominant singularities.

PROOF. Like in the case of a single variable, the proof bases itself on Cauchy's coefficient formula. A composite contour γ like the one depicted on Figure 10 is used. Estimates on each fragment of the contour obey the same principles as in the proof of Theorem VI.3. $\square$

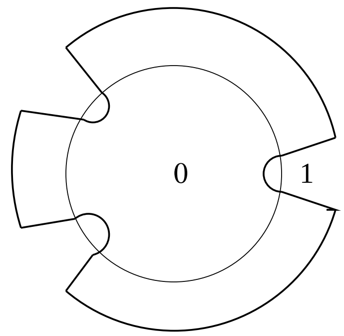


FIGURE 10. A contour used to prove the transfer theorem in the case of multiple singularities (here $r = 3$).

In other words, each dominant singularity can be analysed independently, the singular expansions are then each transferred to coefficients and the corresponding asymptotic contributions are finally collected.

This theorem applies to a function like $g(z)$ defined in Eq. (45) as follows. Define

$$f(z) = g(z) - \frac{e}{\sqrt{2}\sqrt{1-z}} - \frac{e^{-1}}{\sqrt{2}\sqrt{1+z}}.$$

Let $\zeta_1 = 1$ and $\zeta_2 = -1$ be the two dominant singularities. Then, $f(z)$ satisfies the conditions of Theorem VI.8 with $\alpha = \frac{1}{2}$, and

$$h_1(z) = -\frac{e^{-1}}{\sqrt{2}\sqrt{1+z}},\ h_2(z) = -\frac{e}{\sqrt{2}\sqrt{1-z}}.$$

Thus,

$$[z^n]f(z) = \mathcal{O}(n^{-3/2}).$$

The coefficient $[z^n]g(z)$ is then recovered by

$$[z^n]g(z) = [z^n]\frac{e}{\sqrt{2}\sqrt{1-z}} + [z^n]\frac{e^{-1}}{\sqrt{2}\sqrt{1+z}} + \mathcal{O}(n^{-3/2}).$$

The process is simple though its justification is slightly complicated by the fact that one must deal with multiple expansions valid at different points (this is the rôle played by the "regular" parts h_j).

As yet another example, consider the problem of estimating the coefficients of

$$f(z) = \sqrt{\frac{1+z}{1-z}} = \exp\left(\frac{1}{2}\log\frac{1+z}{1-z}\right),$$

which is the EGF of permutations having only cycles of odd length. We have

$$\begin{aligned} f(z) &= \frac{2^{1/2}}{\sqrt{1-z}} - 2^{-3/2}\sqrt{1-z} + \mathcal{O}((1-z)^{3/2}) \quad (z \to 1) \\ f(z) &= 2^{-1/2}\sqrt{1+z} + \mathcal{O}((1+z)^{3/2}) \quad (z \to -1). \end{aligned}$$

Therefore,

$$[z^n]f = \frac{2^{1/2}}{\sqrt{\pi n}} + \frac{-2^{-1/2} + (-1)^n 2^{1/2}}{\sqrt{\pi n^3}} + \mathcal{O}(n^{-5/2}).$$

This last example illustrates the occurrence of singular parts with different weights.

The situation of multiple dominant singularities ties with the corresponding discussion of Chapter IV. In the periodic case where the dominant singularities are at all roots of unity, different regimes manifest themselves cyclically depending on modular properties of the index n, like in the two examples above. In the (seldom occurring) case where dominant singularities have arguments incommensurable to π, aperiodic fluctuations may appear, the discussion being similar to the corresponding case for rational functions.

VI. 8. Tauberian theory and Darboux's method

There are several alternative approaches to the analysis of coefficients of generating functions with moderate growth. *All of them naturally provide estimates compatible with singularity analysis methods* (Theorems VI.1, VI.2, and VI.3). Each one requires some sort of "*regularity condition*" either on the part of the function or on the part of the coefficient sequence, the regularity condition of singularity analysis being in essence analytic continuation.

The methods briefly surveyed here fall into three broad categories:

(i) Elementary real analytic methods;
(ii) Tauberian theorems;
(iii) Darboux's method.

Elementary real analytic methods assume some *a priori* smoothness conditions on the coefficient sequence; they are included here for the sake of completeness, though properly speaking they do not belong to the galaxy of complex asymptotic methods. Their scope is mostly limited to the analysis of products while the other methods permit to approach more general functional composition patterns.

Tauberian theorems belong to the category of advanced real analysis methods; they also needs some *a priori* regularity on the coefficients, typically positivity or monotonicity.

Darboux's method requires some smoothness of the function on the closed unit disk, and, by its techniques and scope, it is the closest to singularity analysis.

We content ourselves with a brief discussion of the main results. For more information, the reader is referred to Odlyzko's excellent survey [**147**].

Elementary real analytic methods. An asymptotic equivalent of the coefficients of a function can sometimes be worked out elementarily from simple properties of the component functions. The regularity conditions are a smooth asymptotic behaviour of the coefficients of one of the two factors in a product of generating functions. A good source for these techniques is Bender's survey [**17**].

THEOREM VI.9 (Bender's method). *Let $a(z) = \sum a_n z^n$ and $b(z) = \sum b_n z^n$ be two power series with radii of convergence $\alpha > \beta \geq 0$ respectively. Assume that $b(z)$ satisfies the ratio test,*

$$\frac{b_{n-1}}{b_n} \to \beta \qquad \text{as} \qquad n \to \infty.$$

Then the coefficients of the product $f(z) = a(z) \cdot b(z)$ satisfy, provided $a(\beta) \neq 0$,

$$[z^n]f(z) \sim a(\beta) b_n \qquad \text{as} \qquad n \to \infty.$$

PROOF. (Sketch) The basis of the proof is the following chain:

$$\begin{aligned} f_n &= a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0) \\ &= b_n \left(a_0 + a_1 \frac{b_{n-1}}{b_n} + a_2 \frac{b_{n-2}}{b_n} + \cdots + a_n \frac{b_0}{b_n} \right) \\ &= b_n \left(a_0 + a_1 (\frac{b_{n-1}}{b_n}) + a_2 (\frac{b_{n-2}}{b_{n-1}})(\frac{b_{n-1}}{b_n}) + \cdots \right) \\ &\sim b_n (a_0 + a_1 \beta + a_2 \beta^2 + \cdots). \end{aligned}$$

There, only the last line requires a little elementary analysis that is left as an exercise to the reader. □

This theorem applies for instance to the EGF of 2–regular graphs:

$$f(z) = a(z) \cdot b(z) \text{ with } a(z) = e^{-z/2 - z^2/4},\ b(z) = \frac{1}{\sqrt{1-z}}$$

$$\Longrightarrow \qquad f_n \sim e^{-3/4} \binom{n - 1/2}{n} \sim \frac{e^{-3/4}}{\sqrt{\pi n}}.$$

Clearly, a whole collection of lemmas could be given in the same vein. Singularity analysis usually provides more complete expansions, though Theorem VI.9 does apply to a few situations not covered by it.

▷ **17.** Estimate asymptotically the coefficients of z^n in

$$\left(\sum \frac{z^n}{n!} \right) \left(\sum n! z^n \right), \left(\sum 2^n z^n \right) \left(\sum n z^{2^n} \right).$$

◁

Tauberian theory. Tauberian methods apply to functions whose growth is known along the positive real line. The regularity conditions are in the form of additional assumptions on the coefficients (positivity or monotonicity) known under the name of Tauberian "side conditions". An insightful introduction to the subject may be found in Titchmarsh's book [**181**], and a detailed exposition in Postnikov's monograph [**160**]. We cite the most famous of all Tauberian theorems due to Hardy, Littlewood, and Karamata. Here, a function is said to be *slowly varying* at infinity iff, for any $c > 0$, one has $L(cx)/L(x) \to 1$ as $x \to +\infty$; examples of slowly varying functions are provided by powers of logarithms or iterated logarithms.

THEOREM VI.10 (The HLK Tauberian theorem). *Let $f(z)$ be a power series with radius of convergence equal to 1, satisfying*

$$f(z) \sim \frac{1}{(1-z)^\alpha} L(\frac{1}{1-z}), \tag{46}$$

for some $\alpha \geq 0$ with L a slowly varying function. Assume that the coefficients $f_n = [z^n] f(z)$ are all non– negative (this is the "side condition"). Then

$$\sum_{k=0}^{n} f_k \sim \frac{n^\alpha}{\Gamma(\alpha+1)} L(n). \tag{47}$$

The conclusion (47) is consistent with what singularity analysis gives: Under the conditions, and if in addition analytic continuation is assumed, then

$$f_n \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} L(n), \tag{48}$$

which by summation yields the estimate (47).

It must be noted that a Tauberian theorem requires very little on the part of the function. However, it also gives less since the result it provides is valid in the more restrictive sense of mean values, or Cesàro averages. (However, if further regularity conditions on the f_n are injected, for instance monotonicity, then the conclusion of (48) can be deduced from (47) by purely elementary real analysis.) The method applies only to functions that are large enough at their singularity, and despite numerous efforts to improve the conclusions, it is the case that Tauberian theorems have little concrete to offer in terms of error estimates.

Appeal to a Tauberian theorem is justified when a function has, apart from the positive half line, a very irregular behaviour near its circle of convergence, for instance when each point of the unit circle is a singularity. (The function is then said to admit the unit circle as a natural boundary.) An interesting example of this situation is discussed by Greene and Knuth [**100**] who consider the function

$$f(z) = \prod_{k=1}^{\infty}\left(1 + \frac{z^k}{k}\right) \tag{49}$$

that is the EGF of permutations having cycles all of different lengths. A little computation shows that

$$\begin{aligned} \log \prod_{k=1}^{\infty}\left(1 + \frac{z^k}{k}\right) &= \sum_{k=1}^{\infty} \frac{z^k}{k} - \frac{1}{2}\sum_{k=1}^{\infty}\frac{z^{2k}}{k^2} + \frac{1}{3}\sum_{k=1}^{\infty}\frac{z^{3k}}{k^3} - \cdots \\ &\sim \log\frac{1}{1-z} - \gamma + o(1). \end{aligned}$$

(Only the last line requires some care, see [**100**].)

Thus, we have

$$f(z) \sim \frac{e^{-\gamma}}{1-z} \qquad \Longrightarrow \qquad \frac{1}{n}(f_0 + f_1 + \cdots + f_n) \sim e^{-\gamma},$$

by virtue of Theorem VI.9. In fact, Greene and Knuth were able to supplement this argument by a "bootstrapping" technique and show a stronger result, namely

$$f_n \to e^{-\gamma}.$$

▷ **18.** Find estimates for

$$[z^n]\prod_{k=1}^{\infty}\left(1 + \frac{z^k}{\sqrt{k}}\right).$$

◁

Darboux's method. The method of Darboux requires, as regularity condition, that functions be smooth enough —i.e., sufficiently differentiable— on their circle of convergence. What lies at the heart of this many–facetted method is a simple relation between the smoothness of a function and the corresponding decrease of its Taylor coefficients.

THEOREM VI.11 (Darboux's method). *Assume that $f(z)$ is continuous in the closed disk $|z| \le 1$, and is in addition k times continuously differentiable $(k \ge 0)$ on $|z| = 1$. Then*

$$[z^n]f(z) = o\left(\frac{1}{n^k}\right). \tag{50}$$

PROOF. Start from Cauchy's coefficient formula

$$f_n = \frac{1}{2i\pi} \int_{\mathcal{C}} f(z) \frac{dz}{z^{n+1}}.$$

Because of the continuity assumption, one may take as integration contour $\mathcal{C}$ the unit circle. Setting $z = e^{i\theta}$ yields the Fourier version of Cauchy's coefficient formula,

$$f_n = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) e^{-ni\theta} \, d\theta. \tag{51}$$

The integrand in (51) is strongly oscillating and the Riemann–Lebesgue lemma of classical analysis (see [**181**, p. 403]) shows that the integral giving f_n tends to 0 as $n \to \infty$.

This argument covers the case $k = 0$. The case of a general k is then derived through successive integrations by parts, as

$$[z^n]f(z) = \frac{1}{2\pi(in)^k} \int_0^{2\pi} f^{(k)}(e^{i\theta}) e^{-ni\theta} \, d\theta.$$

□

Various consequences of Theorem VI.11 are given in reference texts also under the name of Darboux's method. See for instance [**39, 100, 107, 193**]. We shall only illustrate the mechanism by rederiving in this framework the analysis of the EGF of 2–regular graphs. Clearly, we have

$$\begin{aligned} f(z) &= \frac{e^{-z/2-z^2/4}}{\sqrt{1-z}} \\ &= \frac{e^{-3/4}}{\sqrt{1-z}} + e^{-3/4}\sqrt{1-z} + R(z). \end{aligned} \tag{52}$$

There $R(z)$ is the product of $(1-z)^{3/2}$ with a function analytic at $z = 1$ that is a rest in the Taylor expansion of $e^{-z/2-z^2/4}$. Thus, $R(z)$ is of class $\mathbf{C}^1$, *i.e.*, continuously differentiable once. By Theorem VI.11, we have

$$[z^n]R(z) = o\left(\frac{1}{n}\right),$$

so that

$$[z^n]f(z) = e^{-3/4}\binom{n-1/2}{n} + e^{-3/4}\binom{n-3/2}{n} + o\left(\frac{1}{n}\right) = \frac{e^{-3/4}}{\sqrt{\pi n}} + o\left(\frac{1}{n}\right). \tag{53}$$

Darboux's method bears some resemblance to singularity analysis in that the estimates derive from translating error terms in expansions. Smoothness conditions, rather than plain order of growth information, are required by it. It is often applied in situations like in (52)–(53) to functions that are products of the type $h(z)(1-z)^\alpha$ with $h(z)$ analytic at 1, or combinations thereof. In such particular cases, Darboux's method is however subsumed by singularity analysis.

It is inherent to Darboux's method that it cannot be applied to functions whose expansions only involve terms that become infinite, while singularity analysis can. A clear example arises in the analysis of the common subexpression problem [**82**] where there occurs a function with a singular expansion of the form

$$\frac{1}{\sqrt{1-z}} \frac{1}{\sqrt{\log \frac{1}{1-z}}} \left[1 + \frac{c_1}{\log \frac{1}{1-z}} + \cdots\right].$$

▷ **19.** This exercise gives an instance where Darboux's method applies but not singularity analysis. Let

$$F_r(z) = \sum_{n=0}^{\infty} \frac{z^{2^n}}{(2^n)^r}.$$

Show that $F_0(z)$, and hence each F_r, is singular at every point of the unit circle. [Hint: examine the growth of F_0 near 2^nth roots of unity.] Use Darboux's method to analyse asymptotically

$$[z^n]e^z F_5(z).$$

◁

VI. 9. Notes

General surveys of asymptotic methods in enumeration have been given by Bender [**17**] and more recently Odlyzko [**147**]. A general reference to asymptotic analysis that has a remarkably concrete approach is De Bruijn's book [**42**]. Comtet's book [**39**] and Wilf's book [**193**] each devote a chapter to these questions.

This chapter is almost entirely based on the theory developed by Flajolet and Odlyzko in [**78**], where the term "singularity analysis" originates from. That theory itself draws its inspiration from classical analytic number theory, for instance the prime number theorem where similar constours are used (see the discussion in [**78**] for sources). Another area where Hankel contours are used is the inversion theory of integral transforms [**47**], in particular in the case of algebraic and logarithmic singularities.

As seen repeatedly in this chapter, singularity analysis applies to functions with a moderate (at most polynomial) rate of growth near their dominant singularities. The conditions are simply analytic continuation and validity of the singular expansions in the corresponding domains. Such conditions are automatically satisfied by a large number of functions given by explicit "analytic" expressions, most notably a large subset of the generating functions of elementary combinatorial structures defined by the basic constructions of sequence, set, and cycle. Thus singularity analysis ties well with the symbolic enumeration methods of Chapters 1–3.

The cases of molecules and of simple trees show that singularity analysis also applies to many functions defined by functional equations corresponding to recursively defined combinatorial structures. A detailed study of *singular functional equations* does permit in cases like these to analyse coefficients of generating functions that are only defined *implicitly*. Many of these problems belong to an area of on-going research. Prototypes are to be found in Pólya's paper that contains numerous examples of nonplane trees, Odlyzko's analysis of balanced 2-3 trees (we gave earlier a first approximation of the analysis), as well as in two problems to be examined later: the analysis of the height of binary trees in (a singular iteration problem), and the analysis of multidimensional search in k–d–trees and quad-trees (a singular differential system).

The application of the method of singularity analysis is rather mechanical since it corresponds to a direct term by term translation, once general analyticity conditions are recognized to hold. Salvy [**164**] has indeed succeeded in automating the analysis of a large class of generating functions in this way; related decision procedures in this context are also discussed in [**80**].

Darboux's method can often be employed as an alternative to singularity analysis. It is still by far the most widely used technique in the literature, though the direct mapping of asymptotic scales afforded by singularity analysis appears to us much more transparent. Darboux's method is well explained in the books by Comtet [**39**], Henrici [**107**], Olver [**150**], and Wilf [**193**]. Tauberian theory is treated in detail in Postnikov's monograph [**160**], with an excellent introduction to be found in Titchmarsh's book [**181**].

Finally, another range of asymptotic behaviour —that of very rapidly growing functions and entire functions— can be covered by the use of the *saddle point method* examined in Chapter VIII. The *Mellin transform* is the basis of another range of complex-asymptotic techniques applying to functions of an "arithmetical nature": in this book, it is only discussed briefly in APPENDIX: *Mellin transform*, p. 227.

CHAPTER VII

Applications of Singularity Analysis

So their combinations with themselves and with each other give rise to endless complexities, which anyone who is to give a likely account of reality must survey.
— PLATO (speaking of Platonic solids), The Timaeus

Mathematics is being lazy. Mathematics is letting the principles do the work for you so that you do not have to do the work for yourself.
— GEORGE PÓLYA[1]

I wish to God these calculations had been executed by steam.
— CHARLES BABBAGE (1792-1871)

Singularity analysis paves the way to the analysis of a large variety of generating functions. In accordance with Pólya's aphorism, it enables us to "be lazy" and "let the principles work for you". In this chapter we illustrate this situation with singularity analysis developed in Chapter VI being put to use in order to analyse whole classes of generating functions lavishly provided by the symbolic methods of Chapters I–III.

The exp–log schema (Section VII. 1) is a general schema of analytic combinatorics that covers the set construction, either labelled or unlabelled, applied to generators whose singularity is of logarithmic type. This schema parallels in generality the supercritical schema of Chapter V. It applies to permutations, derangements, 2–regular graphs, mappings, and functional graphs. It is even the case that properties relative to the factorization of polynomials with coefficients over a finite field can be attached to it. In particular, one can obtain in a transparent manner a prime number theorem for such polynomials as well as several other characteristics of random polynomials.

The next sections deal with recursively defined structures. In that case, generating functions are accessible by means of an equation or a system that implicitly defines them. A distinctive feature of many such combinatorial generating functions is yet another type of *universality*: square-root singularities are universal, a fact that translates into universality of the exponent $-\frac{3}{2}$ in corresponding asymptotic estimates of coefficients.

Trees are the prototypical recursively defined combinatorial type. For simple varieties, equations merely involve properties of inverses of analytic functions. This applies to simple varieties of trees determined by degree constraints. Universality of square-root singularities entails that several quantities assume the same behaviour across quite different looking tree varieties: the subexponential growth factor of tree counts has a universal $-\frac{3}{2}$ exponent, nodes tend to be found at level about $\sqrt{n}$ and height is of that $\sqrt{n}$ order with high probability, path length grows like $n\sqrt{n}$, and so on. Such results hold for classical tree types (e.g., Catalan, unary-binary, Cayley) and the methods extend to many tree-like classes of combinatorics including functional graphs, mappings, and hierarchies

[1]Quoted in M Walter, T O'Brien, Memories of George P´olya, Mathematics Teaching 116 (1986)

(Section VII. 2). Essentially the same methods apply to functions defined by a single implicit equation: in that case failure cases of the Implicit Function Theorem replace failure of invertiblity in analytic functions, but square-root singularity is still universal. Consequences are found in the general enumeration of nonplane unlabelled, secondary structures of molecular biology, nonplane unlabelled rees, as well as isomers of alkanes in theoretical chemistry (Section VII. 3).

A number of generating functions of combinatorics are algebraic functions (Section VII. 4, meaning that they satisfy either one polynomial equation or are components of a polynomial system. At this level of generality, a whole family of singular behaviours is possible, though singular expansions can only involve fractional exponents. Singularity analysis is invariably applicable. The investigation of algebraic functions requires viewing them as plane algebraic curves and making use of the famous Newton-Puiseux theorem of elementary algebraic geometry, which strongly constrains the allowable types of singularities. For functions given by positive polynomial systems, the general results specialize and one encounters once more square-root singularities (Subsection VII. 4.2).

Algebraic functions manifest themselves first and foremost when dealing with context-free specifications and languages. In that case, under a technical condition of irreducibility, the theory of positive polynomial systems applies. As an example we discuss geometric configurations in the plane satisfying a non-crossing constraint. Algebraic functions may also surface as solutions of various types of functional equations: this is in particular the case for many types of walks generalizing Dyck and Motzkin paths (via the kernel method) and for many types of random maps (via the quadratic method). In all these cases, singular exponents of various forms are bound to occur.

VII. 1. The "exp–log" schema

In this section, we examine a schema that is of a level of generality comparable to the supercritical sequence schema of Chapter V but whose "physics" is rather different. This schema extends what is encountered when constructing of permutations ($\mathcal{P}$) as labelled sets of cycles ($\mathcal{K}$):

$$P(z) = \exp\left(K(z)\right), \qquad K(z) = \log\frac{1}{1-z}.$$

The distinctive feature here is the fact that a logarithmic singularity for the $\mathcal{K}$–components gets composed with an exponential, to the effect that the generating function of the composed $\mathcal{P}$–objects exhibits a singularity of polynomial growth (here $P(z) = (1-z)^{-1}$). In the case of permutations, everything is explicit and we have seen in Chapter III chapters that the distribution of the number of components is concentrated around its mean $H_n = \log n + \gamma + o(1)$. We also know from Chapter IV that with a positive probability a large $\mathcal{P}$-structure contains no $\mathcal{K}$–component of size 1: see the discussion of derangements in Chapter IV.

Very similar properties hold true under very general conditions. We start with the definition of a logarithmic function:

DEFINITION VII.1. *A function $G(z)$ with radius of convergence satisfying $0 < \rho < \infty$ is said to be of* logarithmic type *if the following conditions hold:*

(i) *the number ρ is the unique singularity of $G(z)$ on $|z| = \rho$;*
(ii) *the function $G(z)$ is continuable to a Δ– domain;*
(iii) *as $z \to \rho$ in Δ, function $G(z)$ satisfies*

$$G(z) = \kappa \log\frac{1}{1-z/\rho} + \lambda + O\left(\frac{1}{(\log(1-z/\rho))^2}\right),$$

$\mathcal{F}$	κ	$n=100$	$n=272$	$n=739$
Permutations	1	5.18737	6.18485	7.18319
Derangements	1	4.19732	5.18852	6.18454
2–regular	$\frac{1}{2}$	2.53439	3.03466	3.53440
Mappings	$\frac{1}{2}$	2.97898	3.46320	3.95312

FIGURE 1. Some exp–log structures ($\mathcal{F}$) and the mean number of $\mathcal{G}$–components for $n = 100, 272 \equiv \lceil 100 \cdot e \rfloor, 739 \equiv \lceil 100 \cdot e^2 \rfloor$.

for some $\kappa > 0$ and $\lambda \in \mathbb{R}$.

An exp–log schema is then defined as follows:

DEFINITION VII.2. *Let $\mathcal{F} = \mathfrak{P}(\mathcal{G})$ be a labelled set construction. The schema is said to be of the* exp–log type *if the egf $G(z)$ is of logarithmic type.*

Let $\mathcal{F} = \mathfrak{K}(\mathcal{G})$ be an unlabelled set construction: $\mathfrak{K} = \mathfrak{M}$ (multiset) or $\mathfrak{P}$ (powerset). The schema is said to be of the exp–log type *if the ogf $G(z)$ is of logarithmic type and its radius of convergence satisfies $0 < \rho < 1$.*

As we shall see below, beyond permutations, this schema covers mappings, unlabelled functional graphs, polynomials over finite fields, 2–regular graphs, as well as generalized derangements. Singularity analysis gives precise information on the decomposition of large $\mathcal{F}$ objects into $\mathcal{G}$ components.

THEOREM VII.1 (Exp–log schema). *Given a schema $\mathcal{F} = \mathfrak{P}(\mathcal{G})$ of the exp– log type, one has*

$$\begin{aligned} [z^n]F(z) &= \frac{e^{\lambda+r_0}}{\Gamma(\kappa)} n^{\kappa-1}\rho^{-n}\left(1+O\left((\log n)^{-2}\right)\right), \\ [z^n]G(z) &= \frac{\kappa}{n}\rho^{-n}\left(1+O\left((\log n)^{-2}\right)\right), \end{aligned}$$

where $r_0 = 0$ in the labelled case and r_0 is given by (1) *in the case of unlabelled multisets.*

Let X be the number of $\mathcal{G}$– components in a random $\mathcal{F}$– object. Then:

$$\begin{aligned} \mathbb{E}_{\mathcal{F}_n}(X) &= \kappa(\log n - \psi(\kappa)) + \lambda + r_1 + O\left((\log n)^{-2}\right) \quad (\psi(s) \equiv \tfrac{d}{ds}\Gamma(s)) \\ \mathbb{V}_{\mathcal{F}_N}(X) &= (\kappa + r_2)\log n, \end{aligned}$$

where $r_1 = r_2 = 0$ in the labelled case and $r_1 > 0$ is given by (2) *in the case of unlabelled multisets. In particular the distribution of X in a large $\mathcal{F}_n$ object is concentrated around its mean.*

This result is from an article by Flajolet and Soria [**83**] (with a correction to the logarithmic type condition given by Jennie Hansen [**102**]). We shall show in a later chapter that, in addition, the asymptotic distribution of X is invariably Gaussian under such exp-log conditions.

PROOF. We first discuss the labelled case where $F(z) = \exp G(z)$. The estimate for $[z^n]G(z)$ follows directly from singularity analysis with the transfer of error terms of type $\log^{-2}$. For $F(z)$, it is based on

$$F(z) \sim \frac{e^{\lambda}}{(1 - z/\rho)^{\kappa}},$$

(with a $\log^{-2}$ relative error term), to which singularity analysis applies.

The BGF of $\mathcal{F}$ with u marking the number of $\mathcal{G}$–components is $F(z,u) = \exp(uG(z))$, so that the function

$$f_1(z) := \frac{\partial}{\partial u} F(z,u)\Big|_{u=1} = F(z)G(z),$$

which is an EGF of cumulated values, satisfies near ρ

$$f_1(z) \sim \frac{e^\lambda}{(1-z/\rho)^\kappa}\left(\kappa \log \frac{1}{1-z/\rho} + \lambda\right)$$

to the effect that

$$[z^n]f_1(z) \equiv ([z^n]F(z)) \cdot \mathbb{E}_{\mathcal{F}_n}(X) = \frac{e^\lambda}{\Gamma(\kappa)}\rho^{-n}\left(\kappa \log n - \kappa\psi(\kappa) + \lambda + O\left((\log n)^{-2}\right)\right).$$

The variance analysis is conducted in the same way, but using a second derivative.

For the unlabelled case, let us consider the multiset construction. The analysis of $[z^n]G(z)$ obeys the same principles as in the labelled case as it only depends on the logarithmic assumption. The usual translation of multisets can be put under the form

$$F(z) = \exp\left(G(z) + R(z)\right), \qquad R(z) := \sum_{j=2}^{\infty} \frac{G(z^j)}{j},$$

and $R(z)$ involves terms $R(z^2), \ldots$ that are each analytic in $|z| < \rho^{1/2}$. Thus, $R(z)$ is itself analytic (as a uniformly convergent sum of analytic functions) in $|z| < \rho^{1/2}$, which properly contains the disc $|z| < \rho$ since by assumption $\rho < 1$. One thus has Δ-analyticity of F and, as $z \to \rho$,

$$F(z) \sim \frac{e^{\lambda + r_0}}{(1-z/\rho)^\kappa}, \qquad r_0 := \sum_{j=2}^{\infty} \frac{G(\rho^j)}{j}. \tag{1}$$

The asymptotic expansion of $[z^n]G(z)$ follows.

Regarding the BGF $F(z,u)$ of of $\mathcal{F}$ with u marking the number of $\mathcal{G}$–components, one has

$$F(z,u) = \exp\left(\frac{uG(z)}{1} + \frac{u^2G(z^2)}{2} + \cdots\right).$$

Consequently,

$$f_1(z) := \frac{\partial}{\partial u} F(z,u)\Big|_{u=1} = F(z)\left(G(z) + R_1(z)\right), \qquad R_1(z) = \sum_{j=2}^{\infty} G(z^j).$$

Again, the singularity type is that of $F(z)$ multiplied by a logarithmic term:

$$f_1(z) \sim F(z)(G(z) + r_1), \qquad r_1 := \sum_{j=2}^{\infty} G(\rho^j) \tag{2}$$

and the statement for the mean value of X follows.

The variance analysis proceeds along similar lines. For instance, in the labelled case, one has

$$([z^n]F(z))\mathbb{E}_n(X^2) = [z^n](f_1(z) + f_2(z)), \qquad f_2(z) = \frac{\partial^2}{\partial u^2} F(z,u)\Big|_{u=1} = F(z)G(z)^2,$$

which involves at its singularity a log-squared term. □

EXAMPLE 1. *Direct instances of the exp– log schema.* The case of *permutations* corresponds to $\kappa = 1$, $\lambda = 0$, and it is easily seen to be in agreement with the statement of Theorem VII.1. Let Ω be a finite set of the integers and consider next permutations *without* any cycle of length in Ω. This includes *derangements* ($\Omega = \{1\}$) and their generalizations. Then,

$$G(z) = \log \frac{1}{1-z} - \sum_{\omega \in \Omega} \frac{z^\omega}{\omega}.$$

The theorem applies with $\kappa = 1$ but with now $\lambda := -\sum_{\omega\in\Omega} \omega^{-1}$.

The class of *2– regular graphs* is obtained by the set construction applied to undirected cycles of length ≥ 3. In this case

$$F(z) = \exp(G(z)), \qquad G(z) = \frac{1}{2}\log\frac{1}{1-z} - \frac{z}{2} - \frac{z^2}{4}.$$

This is an exp–log scheme with $\kappa = \frac{1}{2}$ and $\lambda = -\frac{3}{4}$. In particular the mean number of cycles is asymptotic to $\frac{1}{2}\log n$.

In Chapter II, we have encountered the class $\mathcal{F}$ of *mappings* (functions from a finite set to itself) as labelled sets of connected components ($\mathcal{K}$), themselves (directed) cycles of trees ($\mathcal{T}$). The class of all mappings has an EGF given by

$$F(z) = \exp(K(z)), \qquad K(z) = \log \frac{1}{1-T(z)}, \qquad T(z) = ze^{T(z)},$$

with T the Cayley tree function. The analysis of the previous chapter shows that $T(z)$ is singular at $z = e^{-1}$ where it admits the singular expansion $T(z) \sim 1 - \sqrt{2}\sqrt{1-ez}$. Thus $G(z)$ is logarithmic with $\kappa = \frac{1}{2}$ and $\lambda = \log\sqrt{2}$. In particular, the number of connected mappings satisfies

$$K_n \equiv n![z^n]K(z) = n^n\sqrt{\frac{\pi}{2n}}\left(1 + O(n^{-1/2})\right).$$

In other words: *the probability for a random mapping of size n to consist of a single component is* $\sim \sqrt{\frac{\pi}{2n}}$. Also, the mean number of components in a random mapping of size n is

$$\frac{1}{2}\log n + \log\sqrt{2e^\gamma} + O(n^{-1/2}).$$

Similar properties hold for mappings without fixed points that are analogous to derangements and were discussed in Chapter II. □

We shall see below on page 181 that unlabelled functional graphs, which are the counterpart of (labelled) mappings, also resort to the exp-log schema.

EXAMPLE 2. *Polynomials over finite fields.* Given its importance in various areas of mathematics and in applications to coding theory, symbolic computation, and cryptography, we devote a special item to this factorization properties of random polynomials over finite fields. A preliminary discussion has already been given at the end of Chapter I.

Let $\mathbb{F}_p$ be the finite field with p elements and $\mathcal{P} = \mathbb{F}_p[X]$ the set of monic polynomials with coefficients in the field. We view these polynomials as (unlabelled) combinatorial objects with size identified to degree. Since a polynomial is specified by the sequence of its coefficients, there are p^n monic polynomials of degree n, and the OGF is $P(z) = (1-pz)^{-1}$.

Polynomials are a unique factorization domain, since they can be subjected in the usual way to Euclidean division. A nonconstant polynomial that has no proper nonconstant

$$(X+1)\left(X^{10}+X^9+X^8+X^6+X^4+X^3+1\right)\left(X^{14}+X^{11}+X^{10}+X^3+1\right)$$
$$X^3(X+1)\left(X^2+X+1\right)^2\left(X^{17}+X^{16}+X^{15}+X^{11}+X^9+X^6+X^2+X+1\right)$$
$$X^5(X+1)\left(X^5+X^3+X^2+X+1\right)\left(X^{12}+X^8+X^7+X^6+X^5+X^3+X^2+X+1\right)\left(X^2+X+1\right)$$
$$X^2\left(X^2+X+1\right)^2\left(X^3+X^2+1\right)\left(X^8+X^7+X^6+X^4+X^2+X+1\right)\left(X^8+X^7+X^5+X^4+1\right)$$
$$\left(X^7+X^6+X^5+X^3+X^2+X+1\right)\left(X^{18}+X^{17}+X^{13}+X^9+X^8+X^7+X^6+X^4+1\right)$$

FIGURE 2. The factorizations of five random polynomials of degree 25 over $\mathbb{F}_2$. One out of five polynomials in this sample has no root in the base field (the asymptotic probability is $\frac{1}{4}$ by Note 3).

divisor is termed *irreducible* —irreducibles are the analogues of the primes in the integer realm. The unique factorization property implies that the collection $\mathcal{I}$ of monic irreducible polynomials is implicitly determined by $\mathcal{P} \cong \mathfrak{M}(\mathcal{I})$, which is reflected by a functional relation determining $I(z)$:

$$\log \frac{1}{1-pz} = I(z) + \frac{1}{2}I(z^2) + \frac{1}{3}I(z^3) + \cdots .$$

As seen in Chapter I, one can solve explicitly for $I(z)$ using Moebius inversion, to the effect that

$$\begin{aligned} I(z) &= \sum_{k\geq 1} \frac{\mu(k)}{k} \log \frac{1}{1-pz^k} \\ &= \log \frac{1}{1-pz} + R(z), \end{aligned}$$

where $R(z)$ is analytic in $|z| < p^{-1/2}$. Thus $I(z)$ is logarithmic. There results that $I_n \sim p^n/n$, which constitutes a "Prime Number Theorem" for polynomials over finite fields: *A fraction asymptotic to* $\frac{1}{n}$ *of the polynomials in* $\mathbb{F}_p[X]$ *are irreducible.* This says that a polynomial of degree n is roughly comparable to a number written in base p having n digits as the proportion of prime numbers amongst such numbers is about $1/(n \log p)$, by virtue of the classical Prime Number Theorem.

In addition, Theorem VII.1 yields: *The mean number of factors of a random polynomial of degree* n *is* $\sim \log n$ *and the distribution is concentrated.* This and similar developments lead to a complete analysis of some of the basic algorithms known for factoring polynomials over finite fields; see [**71**]. □

▷ **1.** *The divisor function for polynomials.* Let $\delta(\varpi)$ for $\varpi \in \mathcal{P}$ be the total number of monic polynomials (not necessarily irreducible) dividing ϖ: if $\varpi = \iota_1^{e_1} \cdots \iota_k^{e_k}$, where the ι_j are distinct irreducibles, then $\delta(\varpi) = (e_1+1)\cdots(e_k+1)$. One has

$$\mathbb{E}_{\mathcal{P}_n}(\delta) = \frac{[z^n]\prod_{j\geq 1}(1+2z^j+3z^{2j}+\cdots)}{[z^n]\prod_{j\geq 1}(1+z^j+z^{2j}+\cdots)} = \frac{[z^n]P(z)^2}{[z^n]P(z)},$$

so that the mean value of δ over $\mathcal{P}_n$ is exactly $(n+1)$. This evaluation is relevant to polynomial factorization over $\mathbb{Z}$ since it gives an upper bound on the number of ireducible factor combinations that need to be considered in order to lift a factorization from $\mathbb{F}_p(X)$ to $\mathbb{Z}(X)$; see [**189, 122**]. ◁

▷ **2.** *The cost of fi nding irreducible polynomials.* Assume that it takes expected time $t(n)$ to *test* a random polynomial of degree n for irreducibility. Then it takes expected time $\sim nt(n)$ to *fi nd* a random polynomial of degree n: simply draw a polynomial at random and test it for irreducibility. Testing fro ireducibility can be achieved by developing a polynomial factorization algorithm which is stopped as soon as a nontrivial factor is found. See works by Panario *et al.* for details [**152, 153**]. ◁

Under the exp–log conditions, it is also possible (and easy) to analyse the *profile* of structures, that is the number of components of size r for each fixed r. We recall here that the Poisson distribution of parameter ν is the law of a discrete random variable Y such that

$$E(u^Y) = e^{-\nu(1-u)}, \qquad \mathbb{P}(Y = k) = e^{-\nu}\frac{\nu^k}{k!}.$$

A variable Y is said to be a *negative binomial* of parameter (m, α) if its probability generating function and its individual probabilities satisfy:

$$E(u^Y) = \left(\frac{1-\alpha}{1-\alpha u}\right)^m, \qquad \mathbb{P}(Y = k) = \binom{m+k-1}{k}\alpha^k(1-\alpha)^m.$$

(The quantity $\mathbb{P}(Y = k)$ is the probability that the mth success in a sequence of independent trials with individual success probability α occurs at time $m + k$; see [**59**, p. 165].)

PROPOSITION VII.1 (Profiles of exp–log structures). *Assume the conditions of Theorem VII.1 and let the number of $X^{(r)}$ of components of size r in a random $\mathcal{F}_n$ object of size n. In the labelled case, $X^{(r)}$ admits a limit distribution of the Poisson type in the sense that, for any fixed k,*

$$\lim_{n\to\infty} \mathbb{P}_{\mathcal{F}_n}(X^{(r)} = k) = e^{-\nu}\frac{\nu^k}{k!}, \quad \nu = g_r\rho^r, \quad g_r \equiv [z^r]G(z). \tag{3}$$

In the unlabelled case, $X^{(r)}$ admits a limit distribution of the negative binomial type in the sense that, for any fixed k,

$$\lim_{n\to\infty} \mathbb{P}_{\mathcal{F}_n}(X^{(r)} = k) = \binom{G_r+k-1}{k}\alpha^k(1-\alpha)^{G_r}, \quad \alpha = \rho^r,\ G_r \equiv [z^r]G(z). \tag{4}$$

PROOF. In the labelled case, the BGF of $\mathcal{F}$ with u marking the number $X^{(r)}$ of r–components is

$$F(z, u) = \exp\left((u-1)g_r z^r\right) F(z).$$

Extracting the coefficient of u^k leads to

$$[u^k]F(z, u) = \exp\left(-g_r z^r\right)\frac{z^{kr}}{k!}F(z),$$

to which singularity analysis applies directly. Observe that the factor of $F(z)$ contributes the probability of (3) as $z \to \rho$ while the singularity type of $F(z)$ remains unaffected.

In the unlabelled case, the starting BGF equation is

$$F(z, u) = \left(\frac{1-z^r}{1-uz^r}\right)^{G_r} F(z),$$

and the analytic reasoning is similar to the labelled case. $\square$

The unlabelled version covers in particular polynomials over finite fields; see [**71, 116**] for related results.

$\triangleright$ **3.** *Mean profi les.* The mean value of $X^{(r)}$ satisfies

$$\mathbb{E}_{\mathcal{F}_n}(X^{(r)}) \sim g_r\rho^r, \qquad \mathbb{E}_{\mathcal{F}_n}(X^{(r)}) \sim G_r\frac{\rho^r}{1-\rho^r},$$

in the labelled and unlabelled (multiset) case respectively. In particular, the mean number of roots of a random polynomial over $\mathbb{F}_p$ that lie in the base field $\mathbb{F}_p$ is asymptotic to $\frac{p}{p-1}$; the asymptotic probability that a polynomial has no root in the base field is $(1 - 1/p)^p$. $\triangleleft$

▷ **4.** *Profi les of powersets.* In case of unlabelled powersets $\mathcal{F} = \mathfrak{P}(\mathcal{G})$ (no repetitions of elements allowed), the distribution of $X^{(r)}$ satisfies

$$\lim_{n\to\infty} \mathbb{P}_{\mathcal{F}_n}(X^{(r)} = k) = \binom{G_r}{k}\alpha^k(1-\alpha)^{G_r-k}, \quad \alpha = \frac{\rho^r}{1+\rho^r},$$

i.e., the limit is a binomial law of parameters $(G_r, \rho^r/(1+\rho^r))$. ◁

VII. 2. Simple varieties of trees

A *simple variety of trees* $\mathcal{V}$ is a class of trees determined by a subset Ω of the integers, so that all node degrees of a tree in $\mathcal{V}$ are constrained to belong to $\Omega \ni 0$. Such simple varieties exist in four versions: plane or nonplane, unlabelled or labelled. In three of the four cases, the generating function of $\mathcal{V}$ satisfies an equation of the form

$$y(z) = z\phi(y(z)), \tag{5}$$

corresponding to the fact that a tree is recursively formed as a root to which is appended a collection (either a sequence or a set) of subtrees. The algebraic situation is then summarized by the following table:

(6)

	Plane	*Non-plane*
Unlabelled (OGF)	$\mathcal{V} = \mathcal{Z} \times \mathfrak{S}_\Omega(\mathcal{V})$	$\mathcal{V} = \mathcal{Z} \times \mathfrak{M}_\Omega(\mathcal{V})$
	$V(z) = z\phi(V(z))$	$V(z) = z\Phi(V(z)))$
	$\phi(w) := \sum_{\omega\in\Omega} u^\omega$	(Φ a Pólya operator)
Labelled (EGF)	$\mathcal{V} = \mathcal{Z} \star \mathfrak{S}_\Omega(\mathcal{V})$	$\mathcal{V} = \mathcal{Z} \star \mathfrak{P}_\Omega(\mathcal{V})$
	$\widehat{V}(z) = z\phi(\widehat{V}(z))$	$\widehat{V}(z) = z\phi(\widehat{V}(z))$
	$\phi(w) := \sum_{\omega\in\Omega} u^\omega$	$\phi(w) := \sum_{\omega\in\Omega} \frac{u^\omega}{\omega!}$

The generating functions are ordinary (V) in the unlabelled case, exponential ($\widehat{V}$) otherwise. The nonplane unlabelled trees further involve a Pólya operator Φ, which is a sum of monomials in the quantities $V(z^2), V(z^3), \ldots$.

The relation $y = z\phi(y)$, which prevails in the first three cases of (6) have been treated in Section VI.5. In essence, y is defined by inversion of the relation $z = y/\phi(y)$ and inversion “fails” when the first derivative of the function to be inverted vanishes. At this point, the dependency $y \mapsto z$ becomes quadratic, so that its inverse $z \mapsto y$ gives rise to a square-root singularity. We are going to explore the effect of this situation on the probabilistic behaviour of tree parameters.

VII. 2.1. Basic analyses. In the three cases resorting to Equation (5), the main quantities of interest involve a characteristic equation and a condition on the basic constructor function ϕ of (6), which is invariably assumed to be analytic at the origin. The condition is the existence of the characteristic quantity $\tau > 0$ satisfying

$$\tau\phi'(\tau) - \phi(\tau) = 0, \qquad 0 < \tau < R_{\mathrm{conv}}(\phi). \tag{7}$$

We recall also that $\phi(w)$ is said to be *unperiodic* if a decomposition $\psi(w) = w^a h(w^d)$ with h analytic at 0 implies $d = 1$. Paraphrasing the results of Section VI.5, we state:

THEOREM VII.2 (Enumeration of simple trees). *Let $y(z)$ be defined by associated to a simple variety according to* (5)*. Assume that ϕ is unperiodic and such that the characteristic condition* (7) *is satisfied. Then the coefficients of $y(z)$ admit a complete asymptotic expansion*

$$[z^n]y(z) \sim \frac{\gamma\rho^{-n}}{2\sqrt{\pi n^3}}\left[1+\sum_{k=1}^{\infty}\frac{e_k}{n^k}\right], \qquad \rho := \frac{\tau}{\phi(\tau)}, \quad d_1^\star = \sqrt{\frac{2\phi(\tau)}{\phi''(\tau)}}.$$

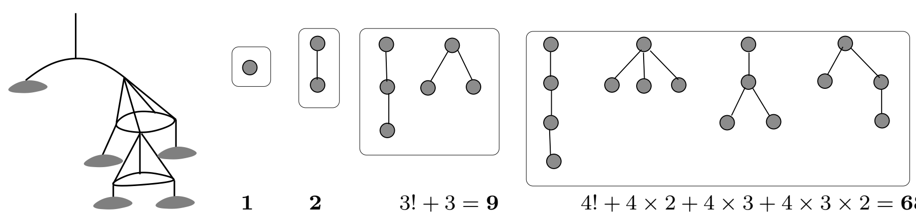
The heart of the matter is, under the conditions of the theorem, the singular expansion of $y(z)$ at $z=\rho$,

$$y(z) = \tau + \sum_{j=1}^{\infty}(-1)^j d_j^\star \left(1-\frac{z}{\rho}\right)^{j/2}, \tag{8}$$

that will prove essential in the analysis of many tree parameters. The developments that follow all make use of the assumptions of Theorem VII.2. We set for convenience of notations:

$$\gamma := d_1^\star = \sqrt{\frac{2\phi(\tau)}{\phi''(\tau)}}.$$

▷ **5.** *Mobiles.* A (labelled) *mobile*, as defined by Bergeron, Labelle, and Leroux [**19**, p. 240], is a (labelled) tree in which subtrees dangling from the root are taken up to cyclic shift:



$\mathbf{1}$ $\qquad$ $\mathbf{2}$ $\qquad$ $3!+3=\mathbf{9}$ $\qquad$ $4!+4\times 2+4\times 3+4\times 3\times 2=\mathbf{68}$

(Think of Calder's creations.) The EGF satisfies (*EIS* **A038037**)

$$\begin{aligned} M(z) &= z+2\frac{z^2}{2!}+9\frac{z^3}{3!}+68\frac{z^4}{4!}+730\frac{z^5}{5!}+\cdots \\ &= z\left(1+\log\frac{1}{1-M(z)}\right). \end{aligned}$$

The asymptotic formula for the number of of mobiles is $\frac{1}{n!}M_n \sim C\cdot A^n n^{-3/2}$, where $C \doteq 0.46563$, $A \doteq 1.15741$; see [**19**, p. 261]. ◁

EXAMPLE 3. *Root degrees in simple varieties.* As an immediate application, we discuss the probability distribution of *root-degree* in simple varieties of trees under the conditions of Theorem VII.2. Let $\mathcal{V}^{[k]}$ be the subset of $\mathcal{V}$ composed of all trees whose root has degree equal to k. The quantity $V_n^{[k]}/V_n$ is the probability that the root of a random tree of size n has degree k. Since a tree in $\mathcal{V}^{[k]}$ is formed by appending a root to a collection of k trees, one has

$$V^{[k]}(z) = \phi_k z V(z)^k, \qquad \phi_k := [w^k]\phi(w).$$

For any *fixed* k, a singular expansion results from raising both members of (8) to the kth power, so that, in particular,

$$V^{[k]}(z) = \tau^k - k\gamma\tau^{k-1}\sqrt{1-\frac{z}{\rho}} + O\left(1-\frac{z}{\rho}\right). \tag{9}$$

This is to be compared to the basic estimate (8): the ratio $V_n^{[k]}/V_n$ is then asymptotic to the ratio of the coefficients of $\sqrt{1-z/\rho}$ in the corresponding generating functions, $V^{[k]}(z)$ and $V(z) \equiv y(z)$. Thus, for any fixed k, we have found that

$$\frac{V_n^{[k]}}{V_n} = \rho k \phi_k \tau^{k-1} + O(n^{-1/2}). \tag{10}$$

(The error term is in fact of the form $O(n^{-1})$, as seen when pushing the expansion one step further.) Since $\rho = 1/\phi'(\tau)$, one can rephrase (10) as follows: *The random variable* Δ *representing the root-degree admits a discrete limit distribution given by*

$$\lim_{n\to\infty} \mathbb{P}_{\mathcal{V}_n}(\Delta) = \frac{k\phi_k \tau^{k-1}}{\phi'(\tau)}. \tag{11}$$

By general principles (Note 6), the convergence is uniform and a stronger form can be developed from singularity analysis techniques, see (12). The probability generating function of the limit law admits the simple form $u\phi'(\tau u)/\phi'(\tau)$.

For our four pilot examples, this gives:

Tree	$\phi(w)$	τ, ρ	*Probability distr.*	(*type*)
binary	$(1+w)^2$	$1, \frac{1}{4}$	PGF: $\frac{1}{2}u + \frac{1}{2}u^2$	(Bernoulli)
unary-binary	$1+w+w^2$	$1, \frac{1}{3}$	PGF: $\frac{1}{3}u + \frac{2}{3}u^2$	(Bernoulli)
general	$(1-w)^{-1}$	$\frac{1}{2}, \frac{1}{4}$	PGF: $u/(2-u)$	(Sum of two geometric)
Cayley	e^w	$1, e^{-1}$	PGF: ue^{u-1}	(shifted Poisson)

The probability distribution is thus characterized by the fact that its probability generating function is a rescaled version of the *derivative* of the basic tree constructor $\phi(w)$. □

▷ **6.** *Simple convergence implies uniform convergence.* Given two *probability distributions* f, g supported at the integers and given by the sequences $(f_k), (g_k)$ define their *local distance* and their *total variation distance* by

$$D(f,g) := \sup_k |f_k - g_k|, \qquad D_{TV}(f,g) := \sum_{k\geq 0} |f_k^{(n)} - g_k|,$$

so that $D(f,g) \leq D_{TV}(f,g)$. Let $f_k^{(n)}$ be a family of discrete probability distributions (indexed by n) and assume that pointwise for each *fi xed* value of k:

$$\lim_{n\to\infty} f_k^{(n)} = g_k.$$

Then $D(f^{(n)}, g) \to 0$ as $n \to \infty$. [Hint: for some k_0, one has $\sum_{k\geq k_0} g_k < \epsilon$. For $k = 0, \ldots, k_0 - 1$, one can make each of the $|f_k^{(n)} - g_k|$ less than ϵ/k_0 by taking n large enough. The triangle inequality implies next that $\sum_{k\geq k_0} f_k^{(n)} < 2\epsilon$, for n large enough. Thus $D_{TV}(f^{(n)}, g) < 4\epsilon$.] ◁

▷ **7.** *Bivariate GFs and limit PGFs.* The BGF of trees with u marking root degree is

$$F(z,u) = z\phi(uy(z)).$$

Choose u_0 such that $u_0 > 1$ and $u_0\tau < R_{\mathrm{conv}}(\phi)$.. Then, one has, for any $|u| < u_0$ and as $z \to \rho$:

$$F(z,u) = \rho\phi(u\tau) - \rho\gamma\phi'(u\tau)\sqrt{1-\frac{z}{\rho}} + O\left(1-\frac{z}{\rho}\right),$$

and the estimate is uniform for $|u| < u_0$. Thus, the PGF of root degree converges to the PGF $u\phi'(\tau u)/\phi'(\tau)$ with spead $O(n^{-1/2})$ [by uniformity of the singularity analysis process]. This implies

$$\sup_{n\to\infty} \left| \mathbb{P}_{\mathcal{V}_n}(\Delta) = \frac{k\phi_k\tau^{k-1}}{\phi'(\tau)} \right| = O\left(\frac{1}{\sqrt{n}}\right), \tag{12}$$

a stronger form of the convergence in distribution asserted in (11). ◁

▷ **8.** *Uniformly exponential tails.* With notations as in Note 7, one finds $\mathbb{E}_{\mathcal{V}_n}(u_1^{\Delta}) = O(1)$ with $u_1 = \frac{1}{2}(1 + u_0) > 0$. By Markov's inequality, there exists an absolute constants K_1, K_2 such that $\mathbb{P}_{\mathcal{V}_n}(\Delta > k) < K_1 K_2^{-n}$ for all large enough n. ◁

EXAMPLE 4. *Mean level profile in simple varieties.* The question we address here is that of determining the mean number of nodes at level k (i.e., at distance k from the root) in a random tree of some large size n. An explicit expression for the joint distribution of nodes at all levels has been developed in Section 5 of Chapter III, but this exact multivariate representation is somewhat hard to interpret in concrete terms.

Let $\xi_j(t)$ be the number of nodes at level k in tree t. Define the generating function of cumulated values,

$$X_k(z) := \sum_{t \in \mathcal{V}} \xi_k(t) z^{|t|}.$$

Clearly, $X_0(z) \equiv y(z)$ since each tree has a unique root. Then, since the parameter ξ_k is the sum over subtrees of parameter ξ_{k-1}, we are in the situation of inheritance, as discussed in Chapter III. We find for a tree with root subtrees $t_1, \ldots, t_{\deg(t)}$,

$$\begin{aligned} X_k(z) &= \sum_{t \in \mathcal{V}} \left(\sum_{j=0}^{\deg(t)} \xi_k(t_j) \right) z^{|t|} \\ &= z \sum_r r \phi_r y(z)^{r-1} X_{k-1}(z) \quad = \quad z\phi'(y(z)) X_{k-1}(z), \end{aligned}$$

so that by recurrence:

$$X_k(z) = (z\phi'(y(z)))^k \, y(z). \tag{13}$$

Making use of the (analytic) expansion of ϕ' at τ, namely, $\phi'(y) \sim \phi'(\tau) + \phi''(\tau)(y - \tau)$ and of $\rho\phi'(\tau) = 1$, one gets for any fixed k

$$X_k(z) \sim \left(1 - k\gamma\rho\phi''(\tau)\sqrt{1 - \frac{z}{\rho}}\right) \left(\tau - \gamma\sqrt{1 - \frac{z}{\rho}}\right) \sim \tau - \gamma(\tau\rho\phi''(\tau)k + 1)\sqrt{1 - \frac{z}{\rho}}.$$

Thus comparing the singular part of $X_k(z)$ to that of $y(z)$, we find: *For* fixed k*, the mean number of nodes at level* k *in a tree is of the asymptotic form*

$$\mathbb{E}_{\mathcal{V}_n}[\xi_k] \sim Ak + 1, \qquad A := \tau\rho\phi''(\tau).$$

This result was first given by Meir and Moon in an important paper of 1978, which started the general theory of simple families of trees [**138**]. The striking fact is that, although the number of nodes at level k can at least double at each level (in the case of the most shrubby trees), the growth is only linear on average. In figurative terms, the immediate vicinity of the root starts like a "cone". and trees of simple varieties tend to be rather skinny near their base.

When used in conjunction with saddle point bounds, the exact GF expression of (13) additionally provides a probabilistic upper bound on the height of trees of the form $O(n^{1/2+\delta})$ for any $\delta > 0$. Indeed restrict z to the interval $(0, \rho)$ and assume that $k = n^{1/2+\delta}$. Let χ be the height parameter. First, we have

$$\mathbb{P}_{\mathcal{V}_n}(\chi \geq k) \equiv \mathbb{E}_{\mathcal{V}_n}([\![\xi_k \geq 1]\!]) \leq \mathbb{E}_{\mathcal{V}_n}(\xi_k). \tag{14}$$

Next by saddle point bounds, for any legal positive x ($0 < x < R_{\mathrm{conv}}(\phi)$),

$$\mathbb{E}_{\mathcal{V}_n}(\xi_k) \leq (x\phi'(y(x)))^k \, y(x) x^{-n} \leq \tau \, (x\phi'(y(x)))^k \, x^{-n}. \tag{15}$$

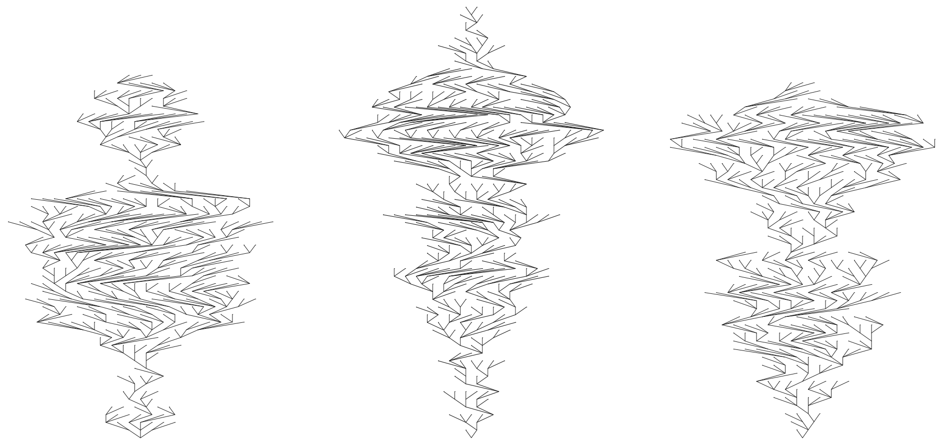


FIGURE 3. Three random 2–3 trees ($\Omega = \{0, 2, 3\}$) of size $n = 500$ have height respectively 48, 57, 47. The skinny aspect of the base, albeit subject to wide variation, is in accordance with the analysis developed in the text.

Fix now

$$x = \rho - \frac{n^{\delta}}{n}.$$

Then, local expansions show that

$$\log\left((x\phi'(y(x)))^k \, x^{-n} \right) \sim -\gamma_1 n^{3\delta/2} + \gamma_2 n^{\delta}, \qquad (16)$$

for some positive constants γ_1, γ_2. Thus, by (14) and (15): *In a simple variety of trees, the probability of height exceeding $n^{1/2+\delta}$ is exponentially small, being of the rough form $\exp(-n^{3\delta/2})$. Accordingly, the mean height is $O(n^{1/2+\delta})$ for any $\delta > 0$.* Flajolet and Odlyzko [**76**] have characterized the moments of height, thenmean being in particular asymptotic to $\lambda\sqrt{n}$ and the limit distribution being of the Theta type already encountered in Chapter V in relation to the height of general Catalan trees. Further local limit and large deviation estimates appear in [**68**]. Figure 3 displays three random trees of size $n = 500$. $\square$

$\triangleright$ **9.** *The variance of level profi les.* The BGF of trees with u marking nodes at level k has an explicit expression, in accordance with the developments of Chapter III. For instance for $k = 3$, this is $z\phi(z\phi(z\phi(uy(z))))$. Double differentiation followed by singularity analysis shows that

$$\mathbb{V}_{\mathcal{V}_n}[\xi_k] \sim \frac{1}{2}A^2k^2 - \frac{1}{2}A(3-4A)k + \tau A - 1,$$

another result of Meir and Moon [**138**]. The precise analysis of the mean and variance in the interesting regime where $k \asymp \sqrt{n}$ is also given in [**138**], but it requires the saddle point method of Chapter VIII. $\triangleleft$

VII. 2.2. Additive functionals. We consider next an important class of recursive parameters of trees, which generalize path length considered in Section 4 of Chapter III. A tree parameter ξ is said to be an *additive functional* if it is defined in terms of a simpler tree parameter η by a recursion of the type:

$$\xi(t) = \eta(t) + \sum_{j=1}^{\deg(t)} \xi(t_j), \tag{17}$$

where $\deg(t)$ is the degree of the root of t and the t_j are the root subtrees of t (whose number is $\deg(t)$). Unwinding the recursion shows that

$$\xi(t) := \sum_{s \preceq t} \eta(s), \tag{18}$$

where the sum is extended to all subtrees s of t (written $s \preceq t$). What is needed is access to moments of ξ viewed as a random variable over the subclass $\mathcal{V}_n$ of all trees of size n.

Expectation of a recursive parameter over trees of a fixed size is of prime relevance to the analysis of algorithms operating on trees [**168**]. For ease of notations, take the case of a simple variety of the *unlabelled plane type*. The generating function of cumulated values are defined in the usual way as

$$X(z) := \sum_{t \in \mathcal{V}} \xi(t) z^{|t|}, \qquad H(z) := \sum_{t \in \mathcal{V}} \eta(t) z^{|t|},$$

and the goal is to determine their relationship. We have

$$X(z) = H(z) + \widetilde{X}(z), \qquad \widetilde{X}(z) := \sum_{t \in \mathcal{V}} \left(z^{|t|} \sum_{j=1}^{\deg(t)} \xi(t_j) \right).$$

Spitting the expression of $\widetilde{X}(z)$ according to the values r of degree, we get

$$\begin{aligned} \widetilde{X}(z) &= \sum_{r \geq 0} \left(\phi_r z^{1+|t_1|+\cdots+|t_r|} \sum_{j=1}^{r} \xi(t_j) \right) \\ &= \sum_{r \geq 0} \left(\phi_r z^{1+|t_1|+\cdots+|t_r|} r \xi(t_1) \right) \\ &= zX(z) \sum_{r \geq 0} \left(r \phi_r y(z)^{r-1} \right). \end{aligned}$$

In summary, this leads to a linear equation satisfied by X,

$$X(z) = H(z) + z\phi'(y(z)) X(z),$$

which solves to

$$X(z) = \frac{H(z)}{1 - z\phi'(y(z))} = \frac{zy'(z)}{y(z)} H(z). \tag{19}$$

This is our main equation. To get its second form, note that differentiating the fundamental relation $y = z\phi(y))$ yields the identity

$$y'(1 - z\phi'(y)) = \phi(y) = \frac{y}{z}, \qquad \text{i.e.,} \qquad 1 - z\phi'(y) = \frac{y}{zy'}.$$

$\triangleright$ **10.** *A combinatorial interpretation.* Equation (18) suggest to view $X(z)$ as the gf of trees with one subtree marked to which is attached a weight of η. Then (19) can be read as follows: point to an arbitrary node at a tree in $\mathcal{V}$ (the gf is $zy'(z)$), “subtract” the tree attached to this node (a factor of $y(z)^{-1}$), and replace it by the same tree but now weighted by η (the gf is $H(z)$). $\triangleleft$

▷ **11.** *Labelled varieties.* Formula (19) holds verbatim for labelled trees (either of the plane or nonplane type), provided we interpret $y(z), X(z), H(z)$ as egf's: $X(z) := \sum_{t \in \mathcal{V}} \xi(t) z^{|t|}/|t|!$, and so on. ◁

Given Equation (19), it is an easy task to churn out a number of mean value estimates for many tree parameters of interest.

EXAMPLE 5. *Mean degree profile.* Let $\xi(t) \equiv \xi_k(t)$ be the number of nodes of degree k in random tree of some variety $\mathcal{V}$. The analysis extends that of the root degree seen earlier. The parameter ξ is an additive functional induced by the basic parameter $\eta(t) \equiv \eta_k(t)$ defined by $\eta_k(t) := [\![\deg(t) = k]\!]$. By the analysis of root degree, we have for the GF of cumulated values associated to η

$$H(z) = \phi_k z y(z)^k, \qquad \phi_k := [w^k]\phi(w),$$

so that, by the fundamental formula (19),

$$X(z) = \phi_k z y(z)^k \frac{zy'(z)}{y(z)} = z^2 \phi_k y(z)^{k-1} y'(z).$$

The singular expansion of $zy'(z)$ results from that of $y(z)$ by differentiation (Chapter VI),

$$zy'(z) = \frac{1}{2}\gamma \frac{1}{\sqrt{1 - z/\rho}} + O(1),$$

and the corresponding coefficient is $[z^n](zy') = ny_n$. This gives immediately the singularity type of X, which is of the form of an inverse square root. Thus,

$$X(z) \sim \rho \phi_k \tau^{k-1}(zy'(z))$$

implying ($\rho = \tau/\phi(\tau)$)

$$\frac{X_n}{ny_n} \sim \frac{\phi_k \tau^k}{\phi(\tau)}.$$

Consequently, one has: *In a simple variety, the mean number of nodes of degree k is asymptotic to $\lambda_k n$, where $\lambda_k := \phi_k \tau^k/\phi(\tau)$; in other words, the probability distribution of the degree Δ' of a random node in a random tree of size n satisfies*

$$\lim_{n\to\infty} \mathbb{P}_n(\Delta) = \frac{\phi_k \tau^k}{\phi(\tau)}, \qquad \text{with PGF :} \quad \sum_k \lambda_k u^k = \frac{\phi(u\tau)}{\phi(\tau)}.$$

▷ **12.** *Variances.* The variance of the number of k-ary nodes is $\sim \nu n$, so that the distribution of the number of nodes of this type is concentrated, for each fixed k. The starting point is the BGF defined implicitly by

$$Y(z,u) = z\left(\phi(Y(z,u)) + \phi_k(u-1)Y(z,u)^k\right),$$

upon taking a double derivative with respect to u, setting $u = 1$, and finally performing singularity analysis on the resulting GF of cumulated values. ◁

For the usual tree varieties this gives:

Tree	$\phi(w)$	τ, ρ	*Probability distr.*	(*type*)
binary	$(1+w)^2$	$1, \frac{1}{4}$	PGF: $\frac{1}{4} + \frac{1}{2}u + \frac{1}{4}u^2$	(Bernoulli)
unary-binary	$1 + w + w^2$	$1, \frac{1}{3}$	PGF: $\frac{1}{3} + \frac{1}{3}u + \frac{1}{3}u^2$	(Bernoulli)
general	$(1-w)^{-1}$	$\frac{1}{2}, \frac{1}{4}$	PGF: $1/(2-u)$	(Geometric)
Cayley	e^w	$1, e^{-1}$	PGF: e^{u-1}	(Poisson)

For instance, asymptotically, a general Catalan tree has on average $\frac{n}{2}$ leaves, $\frac{n}{4}$ nodes of degre 1 $\frac{n}{8}$ of degree 2, and so on; a Cayley tree has $\sim ne^{-1}/k!$ nodes of degree k; for binary (Catalan) trees, the four possible types of nodes each appear each with asymptotic frequency $\frac{1}{4}$. (These data are in agreement with the fact that a random tree under $\mathcal{V}_n$ is distributed like a branching process tree determined by the PGF $\phi(u\tau)/\phi(\tau)$; see the corresponding remarks in Chapter III.) $\square$

$\triangleright$ **13.** *The father of a random node.* The discrepancy in distributions between the root degree and the the degree of a random node deserves an explanation. Pick up a node distinct from the root at random in a tree and look at the degree of its mother. The PGF of the law is in the limit $u\phi'(u\tau)/\phi'(\tau)$. Thus the degree of the root is asymptotically the same as that of the mother of any non-root node.

More generally, let X have distribution $p_k := \mathbb{P}(X = k)$. Construct a random variable Y such that the probability $q_k := \mathbb{P}(Y = k)$ is proportional both to k and p_k. Then for the associated PGFs, the relation $q(u) = p'(u)/p'(1)$ holds. The law of Y is said to be the *size-biased* version of the law of X. Here, a mother is picked up with an importance proportional to its degree. $\triangleleft$

EXAMPLE 6. *Path length.* Path length in trees can be analysed starting from the bivariate generating function given (see Chapter III) by a functional equation of the difference type that involves the transformation $z \mapsto uz$. This is useful for the computation of higher moments, but for mean values, one may as well proceed directly from the additive functional scheme. Path length is definable inductively by

$$\xi(t) = |t| - 1 + \sum_{j=1}^{\deg(t)} \xi(t_j).$$

In this case, we have $\eta(t) = |t| - 1$ corresponding to the GF of cumulated values

$$H(z) = zy'(z) - y(z).$$

Thus, by the fundamental relation (19), one finds:

$$X(z) = (zy'(z) - y(z))\frac{zy'(z)}{y(z)} = \frac{z^2y'(z)^2}{y(z)} - zy'(z).$$

Since the type of $y'(z)$ at its singularity is $Z^{-1/2}$ where $Z := (1 - z/\rho)$, and the formula for $X(z)$ involves the square of y', the singularity of $X(z)$ is of type $Z^{-1/2}$, that is, it resembles a simple pole. This means that $X_n = [z^n]X(z)$ grows like ρ^{-n}, so that the mean value of ξ over $\mathcal{V}_n$ has growth $n^{3/2}$. Determining the dominant terms is done by combining local singular expansions as usual:

$$X(z) + zy'(z) \sim \frac{\gamma^2}{4\tau}\frac{1}{Z} + O(Z^{-1/2}).$$

As a result: *In a random tree of size n in a simple variety, the expectation of path satisfies*

$$\mathbb{E}_{\mathcal{V}_n}(\xi) = \lambda\sqrt{\pi n^3} + O(n), \qquad \lambda := \sqrt{\frac{\phi(\tau)}{2\tau^2\phi''(\tau)}}. \tag{20}$$

For our classical varieties, the main terms of (20) are then:

Binary	Unary-binary	General	Cayley
$\sim \sqrt{\pi n^3}$	$\sim \frac{1}{2}\sqrt{3\pi n^3}$	$\sim \frac{1}{2}\sqrt{\pi n^3}$	$\sim \sqrt{\frac{1}{2}\pi n^3}$

Observe that the quantity $\frac{1}{n}\mathbb{E}_{\mathcal{V}_n}(\xi)$ represents the expected depth of a random node in random tree (the probability model is then $[1\,..\,n] \times \mathcal{V}_n$), which is thus $\sim \lambda\sqrt{n}$. This result is consistent with the previously noted fact that height of a tree is with high probability not much larger than $O(n^{1/2})$. □

▷ **14.** *Variance of path length.* The variance of path length is asymptotic to $\lambda_2 n^{3/2}$ for some computable constant $\lambda_2 > 0$. Hence the distribution is "spread". Louchard [**133**] and Tak´acs [**180**] have additionally worked out the asymptotic form of all moments, leading to a characterization of the limit law of path length that can be described in terms of the Airy function and coincides with the Brownian excursion area. ◁

▷ **15.** *Generalizations of path length.* Take the elementary cost $\eta(t)$ to be $|t|^\alpha$ for some exponent $\alpha > 0$. Then the results of Chapter VI make it possible to analyse the relevant singularity of the GF $H(z)$ (via Hadamard products and singularities of polylogarithms). This entails for the mean values of ξ precise estimates generalizing those of path length (which corresponds to $\alpha = 1$). There is a quantitative difference depending on whether $\alpha < 1\frac{1}{2}$, $\alpha = \frac{1}{2}$, $\alpha > \frac{1}{2}$. For instance for binary trees, one finds (K_α and K_0' are explicit constants):

$\eta(t)$ *in the case* $\lvert t\rvert = n$		$\mathbb{E}_n(\xi)$	
n^α	$(\frac{3}{2} < \alpha)$	$\dfrac{\Gamma(\alpha-\frac{1}{2})}{\Gamma(\alpha)} n^{\alpha+\frac{1}{2}}$	$+O(n^{\alpha-\frac{1}{2}})$
$n^{3/2}$		$\dfrac{1}{\Gamma(3/2)} n^2$	$+O(n\log n)$
n^α	$(\frac{1}{2} < \alpha < \frac{3}{2})$	$\dfrac{\Gamma(\alpha-\frac{1}{2})}{\Gamma(\alpha)} n^{\alpha+\frac{1}{2}}$	$+O(n)$
$n^{1/2}$		$\dfrac{1}{\sqrt{\pi}} n\log n$	$+O(n)$
n^α	$(0 < \alpha < \frac{1}{2})$	$K_\alpha n$	$+O(1)$
$\log n$		$K_0' n$	$-2\sqrt{\pi}n^{1/2} + O(1)$.

This is based on an article by Fill, Flajolet, and Kapur [**60**]. ◁

In summary, the developments of this section thus provide precise information on the shape of large random trees in simple varieties. The root degree has a special distribution, a random node has degree obeying a distribution (with PGF $\phi(u\tau)/\phi(\tau)$) which directly reflects the nature of the tree constructor ϕ. A typical node is at depth about $\sqrt{n}$, the height of the tree being itself in all likelihood not much larger and path length being $\sim \lambda n^{3/2}$ on average.

VII. 2.3. Enumeration of some non-plane unlabelled trees. We shall discuss here the enumeration of two classes of non-plane trees following Pólya [**158, 159**] and Otter [**151**], which are very important sources for the asymptotic theory of tree enumeration (see also the brief account in [**104**]). These authors used the more traditional method of Darboux instead of singularity analysis, but this distinction is immaterial here as calculations develop under completely parallel lines under both theories. The two classes under consideration are those of "general" and binary non-plane unlabelled trees. Pólya operators are then central to the enumeration, and their treatment in terms of radii of convergence and singularities is typical of the asymptotic theory of unlabelled objects obeying symmetries, i.e., involving the unlabelled $\mathfrak{M}$, $\mathfrak{P}$, $\mathfrak{C}$ constructions.

We prove here:

PROPOSITION VII.2 (Special non-plane unlabelled trees). *Consider the two classes of* non-plane unlabelled trees

$$\mathcal{H} = \mathcal{Z} \times \mathfrak{M}(\mathcal{H}), \quad \mathcal{W} = \mathcal{Z} \times \mathfrak{M}_{\{0,2\}}(\mathcal{W})$$

respectively of the general and binary type. Then, with constants C_H, A_H *and* C_W, A_W *given in* (25) *and* (26)*, one has*

$$H_n \sim \frac{C_H}{2\sqrt{\pi n^3}} A_H^n, \qquad W_{2n-1} \sim \frac{C_W}{2\sqrt{\pi n^3}} A_W^n.$$

PROOF. (i) *General case.* The OGF of nonplane unlabelled trees is the analytic solution to the functional equation

$$H(z) = z \exp\left(\frac{H(z)}{1} + \frac{H(z^2)}{2} + \cdots\right). \tag{21}$$

Let T be the solution to

$$T(z) = z e^{T(z)}, \tag{22}$$

that is to say, the Cayley function. The function $H(z)$ has a radius of convergence ρ strictly less than 1 as its coefficients dominate those of $T(z)$, the radius of convergence of the latter being exactly $e^{-1} \doteq 0.367$. The radius ρ cannot be 0 since the number of trees is bounded from above by the number of plane trees whose OGF has radius $\frac{1}{4}$. Thus, one has $\frac{1}{4} \le \rho \le e^{-1}$.

Rewriting the defining equation of $H(z)$ as

$$H(z) = \zeta e^{H(z)} \qquad \text{with} \qquad \zeta := z \exp\left(\frac{H(z^2)}{2} + \frac{H(z^3)}{3} + \cdots\right),$$

we observe that $\zeta = \zeta(z)$ is analytic for $|z| < \rho^{1/2}$, that is to say in a disk that properly contains the disk of convergence of $H(z)$. We may thus rewrite $H(z)$ as

$$H(z) = T(\zeta(z)).$$

Since $\zeta(z)$ is analytic at $z = \rho$, a singular expansion of $H(z)$ near $z = \rho$ results from composing the singular expansion of T at e^{-1} with the analytic expansion of ζ at ρ. In this way, we get:

$$H(z) = 1 - C(1 - \frac{1}{\rho})^{1/2} + \mathcal{O}((1 - \frac{z}{\rho})), C = \sqrt{2e\rho\zeta'(\rho)}. \tag{23}$$

Thus,

$$[z^n]H(z) \sim \frac{C}{2\sqrt{\pi n^3}} \rho^{-n}.$$

(ii) *Binary case.* Consider the functional equation

$$f(z) = z + \frac{1}{2} f(z)^2 + \frac{1}{2} f(z^2). \tag{24}$$

This enumerates non-plane binary trees with size defined as the number of leaves, so that $W(z) = \frac{1}{z} f(z^2)$. Thus, it suffices to analyse $[z^n]f(z)$, which avoids dealing with periodicity phenomena.

The OGF $f(z)$ has a radius of convergence ρ that is at least $\frac{1}{4}$ (since there are fewer non-plane trees than plane ones). It is also at most $\frac{1}{2}$ as results from comparison of f with the solution to the equation $g = z + \frac{1}{2} g^2$. We may then proceed as before: treat the term $\frac{1}{2} f(z^2)$ as a function in $|z| < \rho^{1/2}$, as though it were known, then solve.

To this effect, set

$$\zeta(z) := z + \frac{1}{2} f(z^2),$$

which exists in $|z| < \rho^{1/2}$. Then, the equation (24) becomes a plain quadratic equation, $f = \zeta + \frac{1}{2}f^2$, with solution

$$f(z) = 1 - \sqrt{1 - 2\zeta(z)}.$$

The singularity ρ is the smallest positive solution of $\zeta(\rho) = \frac{1}{2}$. The singular expansion of f is obtained by composing the analytic expansion of ζ at ρ with $\sqrt{1-2\zeta}$. The usual square-root singularity results:

$$f(z) \sim 1 - C\sqrt{1 - z/\rho}, \qquad C := \sqrt{2\rho\zeta'(\rho)}.$$

This induces the $\rho^{-n}n^{-3/2}$ form for the coefficients $[z^n]f(z) \equiv [z^{2n-1}]W(z)$. □

▷ **16.** *Full asymptotic expansions for* H_n, W_{2n-1}. These exist since the OGFs admit complete asymptotic expansions in powers of $\sqrt{1 - z/\rho}$. ◁

The argument used in the proof of the proposition may seem nonconstructive. However, numerically, the values of ρ and C may be determined to great accuracy. See the notes below as well as Finch's section on "Otter's tree enumeration constants" [**61**, Sec. 5.6].

▷ **17.** *Numerical evaluation of constants I.* Here is an unoptimized procedure controlled by a parameter $m \geq 0$ for general non-plane unlabelled trees.

Procedure Get_value_of_ρ(m : integer);
1. Set up a procedure to compute and memorize the H_n on demand;
(this can be based on recurrence relations implied by $H'(z)$; see [**144**])
2. Define $f^{[m]}(z) := \sum_{j=1}^{m} H_n z^n$;
3. Define $\zeta^{[m]}(z) := z\exp\left(\sum_{k=2}^{m} \frac{1}{k} f^{[m]}(z^k)\right)$;
4. Solve numerically $\zeta^{[m]}(x) = e^{-1}$ for $x \in (0,1)$ to $\max(m, 10)$ digits of accuracy;
5. Return x as an approximation to ρ.

For instance, here is a conservative estimate the accuracy attained for $m = 0, 10, \ldots, 50$ (in a few billion machine instructions):

$m=0$	$m=10$	$m=20$	$m=30$	$m=40$	$m=50$
$3\cdot 10^{-2}$	10^{-6}	10^{-11}	10^{-16}	10^{-21}	10^{-26}

Empirically, accuracy appears to be a little better than $10^{-m/2}$. This yields to 25D:

$$\begin{aligned} &\rho \doteq 0.3383218568992076951961126, \ A_H \equiv \rho^{-1} \doteq 2.9557652856519949747148 18\\ &C_H \doteq 1.559490020374640885542206. \end{aligned} \tag{25}$$

The formula of the Proposition correctly estimates H_{100} with a relative error of 10^{-3}. ◁

▷ **18.** *Numerical evaluation of constants II.* The procedure of the previous note adapts easily to give:

$$\begin{aligned} &\rho \doteq 0.4026975036714412909690453, \ A_W \equiv \rho^{-1} \doteq 2.4832535361726368585622 89\\ &C_W \doteq 1.130033716398972007144137. \end{aligned} \tag{26}$$

The formula of the Proposition correctly estimates $[z^{100}]f(z)$ with a relative error of $7\cdot 10^{-3}$. ◁

The two results, general and binary, are thus obtained by a modification of the method used for simple varieties of trees, upon treating the Pólya operator part as a "perturbation" of the corresponding equations of simple varieties of trees. A more general theory is possible for any simple variety of unlabelled non-plane trees $\mathcal{T} = \mathcal{Z}\mathfrak{M}_\Omega(\mathcal{T})$ (for some $\Omega \subset \mathbb{Z}_{\geq 0}$, but it requires more advanced methods to be discussed below.

VII. 2.4. Tree like structures. It is possible to treat many combinatorial structures related to trees by techniques exposed earlier in this section. A square root singularity is usually the norm, as is its corollary, the $\rho^{-n}n^{-3/2}$ asymptotic form for coefficients. As the routine is now well established, we shall content ourselves with a brief discussion of some examples encountered earlier in this book.

1. Trees by leaves and hierarchies. Given a nonempty set $\Omega \subset \mathbb{Z}_{\geq 0}$ that does not contain 0,1, it makes sense to consider the class of labelled trees,

$$\mathcal{C} = Z + \mathfrak{K}_{\Omega}(\mathcal{C}), \qquad \mathfrak{K} = \mathfrak{S} \text{ or } \mathfrak{P}.$$

These are trees (plane or non-plane) with size counted as the number of leaves and with degrees constrained to lie in Ω. The EGF is then of the form

$$C(z) = z + \phi(C(z)).$$

(The fact that $0, 1 \notin \Omega$ ensures well foundedness of the definition.)

The discussion given for simple families of trees adapts easily. Assume for simplicity ϕ here to be entire. Then $C(z)$ is a solution to $y = z + \phi(y)$ that remains analytic, when z increases from 0 along the positive axis, as long as the function $y - \phi(y)$ has a nonzero derivative. Thus, the smallest positive singularity ρ and the corresponding singular value $\tau := y(\rho)$ satisfy the system of two equations:

$$\tau = \rho + \phi(\tau), \qquad 1 = \phi'(\tau). \tag{27}$$

In other words, τ is the smallest root of $\phi'(\tau) = 1$ and ρ is then determined by $\rho = \tau - \phi(\tau)$. Near (ρ, τ), the dependence between z and y is locally a quadratic one:

$$\rho - z \sim \frac{1}{2}\phi''(\tau)(\tau - y),$$

that is,

$$y \sim \tau - \gamma\left(1 - \frac{z}{\rho}\right)^{1/2}, \qquad \gamma := \sqrt{\frac{1}{2}\rho\phi''(\tau)}.$$

(A full expansion can be obtained.) Thus, for coefficients, one has:

$$[z^n]y \sim \frac{\gamma}{2\sqrt{\pi n^3}}\rho^{-n}, \qquad \gamma = \sqrt{\frac{1}{2}\rho\phi''(\tau)}. \tag{28}$$

In Chapter II, we have considered the class $\mathcal{H}$ of labelled hierarchies defined by the choice $\Omega = \{2, 3, \ldots\}$, i.e., $\mathcal{H} = \mathcal{Z} + \mathfrak{P}_{\geq 2}(\mathcal{H})$. These occur in statistical classification theory: given a collection of n distinguished items, H_n is the number of ways of superimposing a nontrivial classification (cf Figure 4). Such abstract classifications usually no planar structure so that $\phi(w) = \sum_{\omega \in \Omega} w^{\omega}/\omega!$ is here $e^w - 1 - w$. Thus, we find mechanically $\tau = \log 2$, $\rho = 2\log 2 - 1$, and

$$\frac{1}{n!}H_n \sim \frac{1}{2\sqrt{\pi n^3}}\left(2\log 2 - 1\right)^{-n+1/2}.$$

For the unlabelled version, $\widetilde{\mathcal{H}}$, a calculation combines the analysis above with the principles employed for nonplane trees to the effect that

$$\widetilde{H}_n \sim \frac{\gamma}{2\sqrt{\pi n^3}}\rho^{-n}, \qquad \rho \doteq 0.29224.$$

2. Mappings. The basic decomposition of mappings,

$$F = \exp(K), \qquad K = \log\frac{1}{1-T}, \qquad T = ze^T,$$

```
|===|===|===|German
|   |   |===|English
|   |   |===|Dutch
|   |
|   |===|===|Swedish
|       |===|Danish
|===|Greek
|
|===|===|===|Portuguese
|   |   |===|Spanish
|   |   |===|Italian
|   |   |===|French
|   |
|   |===|Romanian
|
|===|Armenian
```

FIGURE 4. A hierarchy placed on some of the modern Indoeuropean languages.

# components	$\sim \frac{1}{2}\log n$
# cyclic nodes	$\sim \sqrt{\pi n/2}$
# terminal nodes	$\sim ne^{-1}$

tail length (λ)	$\sim \sqrt{\pi n/8}$
cycle length (μ)	$\sim \sqrt{\pi n/8}$
tree size	$\sim n/3$
component size	$\sim 2n/3$

FIGURE 5. Expectations of the main additive parameters of random mappings of size n.

lends itself to a number of multivariate extensions. For instance, the parameter $\chi(\phi)$ equal to the number of cyclic points gives rise to the BGF

$$F(z,u) = \exp\left(\log\frac{1}{1-T}\right) = (1-uT)^{-1}.$$

The mean number of a cyclic points in a random mapping of size n is accordingly

$$\mu_n = \frac{n!}{n^n}[z^n]\left(\frac{\partial}{\partial u}F(z,u)\Big|_{u=1}\right) = \frac{n!}{n^n}[z^n]\frac{T}{(1-T)^2}.$$

Singularity analysis is immediate as

$$\frac{T}{(1-T)^2} \underset{z\to e^{-1}}{\sim} \frac{1}{2}\frac{1}{1-ez} \qquad \text{implying} \qquad [z^n]\frac{T}{(1-T)^2} \underset{n\to\infty}{\sim} \frac{1}{2}e^n.$$

The mean number of cyclic points in a random n–mapping is asymptotic to $\sqrt{\pi n/2}$. A large number of parameters can be analysed in this way systematically as shown in the survey [**77**]: see Figure 5 for a summary of results whose proof we leave as an exercise to the reader. The leftmost table describes global parameters of mappings; the rightmost table is relative to properties of random point in random n-mapping: λ is the distance to its cycle of a random point, μ the length of the cycle to which the point leads, tree size and component size are respectively the size of the largest tree containing the point and the size of its (weakly) connected component. In particular, a random mapping of size n has relatively few components, some of which are expected to be of a fairly large size.

3. Simple varieties of mappings. Let Ω be a subset of the integers and consider mappings $\phi \in \mathcal{F}$ such that the number of preimages of any point is constrained to lie in Ω. Such special mappings may serve to model the behaviour of special classes of functions under iteration. For instance the quadratic functions $\phi(x) = x^2 + a$ over $\mathbb{F}_p$ have the property that each element y has either zero, one, or two preimages (depending on whether $y - a$ is a quadratic nonresidue, 0, or a quadratic residue). Such constrained mappings are of interest in various areas of computational number theory and cryptography.

The basic decomposition of general mappings needs to be amended in this case. Start with the family of trees $\mathcal{T}$ that are the simple variety corresponding to Ω:

$$T = z\phi(T), \qquad \phi(w) := \sum_{\omega\in\Omega} \frac{u^\omega}{\omega 1}.$$

At any point of a cycle, one must graft r trees with the constraint that $r + 1 \in \Omega$ (since one arrow "comes from" the cycle itself). Such legal tuples with a root appended are represented by

$$U = z\phi'(T),$$

since ϕ is an exponential generating function and shift corresponds to differentiation. Then connected components and components are formed in the usual way by

$$K = \log\frac{1}{1-U}, \qquad F = \exp(K) = \frac{1}{1-U}.$$

We assume that ϕ (i.e., Ω) satisfies the general conditions of Theorem VII.2, with τ the characteristic value. Then $T(z)$ has a square-root singularity at $\rho = \tau/\phi(\tau)$. The same holds for U which satisfies the singular expansion

$$U(z) \sim 1 - \rho\phi''(\tau)\gamma\sqrt{1-\frac{z}{\rho}}. \tag{29}$$

Thus, eventually,

$$F(z) \sim \frac{\lambda}{\sqrt{1-\frac{z}{\rho}}}.$$

There results the universality of the exponent $-1/2$ in such constrained mappings:

$$[z^n]F(z) \sim \frac{\lambda}{\sqrt{\pi n}}\rho^{-n},$$

which nicely extends what is known to hold for unrestricted mappings. The analysis of additive functionals can then proceed on lines very similar to the case of standard mappings, to the effect that the estimates of Figure 5 hold, albeit with different multiplicative constants. The programme just sketched has been carried out in a thorough way by Arney and Bender in [**7**] to which we refer for a detailed treatment.

4. Unlabelled functional graphs. Unlabelled functional graphs or mapping patterns ($\mathcal{F}$) are unlabelled digraphs in which each vertex has outdegree equal to 1. Equivalently, they can be regarded as multisets of components ($\mathcal{L}$) that are cycles of nonplane unlabelled trees ($\mathcal{H}$),

$$\mathcal{F} = \mathfrak{M}(\mathcal{L}); \quad \mathcal{L} = \mathfrak{C}(\mathcal{U}); \quad \mathcal{H} = \mathcal{Z} \times \mathfrak{M}(\mathcal{H}),$$

a specification that entirely parallels that of mappings.

The OGF $H(z)$ has a square-root singularity by virtue of (23) above, with additionally $H(\rho) = 1$. The translation of the unlabelled cycle construction,

$$L(z) = \sum_{j\geq 1} \frac{\varphi(j)}{j} \log\frac{1}{1-H(z^j)},$$

implies that $L(z)$ is logarithmic, and $F(z)$ has a singularity of type $1/\sqrt{Z}$ where $Z := 1 - z/\rho$. Thus, *unlabelled functional graphs constitute an exp-log structure* with $\kappa = \frac{1}{2}$. The number of unlabelled functional graphs thus grows like $C\rho^{-n}n^{-1/2}$ and the mean number of components in a random functional graph is $\sim \frac{1}{2}\log n$, like for the labelled mapping counterpart. See [**139**] for more on this topic. Unlabelled functional graphs are sometimes called "mapping patterns" in the literature.

▷ **19.** *Alternative form for $F(z)$.* Arithmetical simplifications associated with the Euler totient function yield:

$$F(z) = \prod_{k=1}^{\infty} \left(1 - H(z^k)\right)^{-1}.$$

A similar form applies generally to multisets of unlabelled cycles. ◁

5. Unrooted trees. All the trees considered so far have been rooted and this version is the one most useful in applications. An *unrooted tree*[2] is by definition a connected acyclic (undirected) graph. In that case, the tree is clearly non-plane and no special root node is distinguished.

The counting of the class $\mathcal{U}$ of *unrooted labelled trees* is easy: there are plainly $U_n = n^{n-2}$ of these since all nodes are distinguished by a label, which entails that $nU_n = T_n$ with $T_n = n^{n-1}$ by Cayley's formula. Incidentally, the EGF $U(z)$ satisfies

$$U(z) = \int_0^z T(y)\,\frac{dy}{y} = T(z) - \frac{1}{2}T(z)^2, \tag{30}$$

as already seen when we discussed labelled graphs in Chapter II.

For *unrooted unlabelled trees*, symmetries are in the way and an tree can be rooted in a number of ways that depends on its shape. For instance of star graph leads to a number of different rooted trees that equals 2 (choose either the center or one of the peripheral nodes), while a line graph gives rise to $\lceil n/2 \rceil$ structurally different rooted trees. With $\mathcal{H}$ the class of rooted unlabelled trees and $\mathcal{I}$ the class of unrooted trees, we have at this stage only minor refinements of the general inequality

$$I_n \le H_n \le nI_n.$$

A table of values of the ratio H_n/I_n suggests that the answer is closer to the upper bound:

(31)

n	10	20	30	40	50	60
H_n/I_n	6.78	15.58	23.89	32.15	40.39	48.62

The solution is provided by a famous exact formula due to Otter, namely

$$I(z) = H(z) - \frac{1}{2}\left(H(z)^2 - H(z^2)\right). \tag{32}$$

It gives in particular (*EIS* **A000055**)

$$I(z) = z + z^2 + z^3 + 2\,z^4 + 3\,z^5 + 6\,z^6 + 11\,z^7 + 23\,z^8 + 47\,z^9 + 106\,z^{10} + \cdots.$$

Given (32), it is child's play to determine the singular expansion of I knowing that of H. The radius of convergence of I is the same as that of H as the term $H(z^2)$ only introduces exponentially small coefficients. Thus, it suffices to analyse $H - \frac{1}{2}H^2$:

$$H(z) - \frac{1}{2}H(z)^2 \sim \frac{1}{2} - \delta_2\left(1 - \frac{z}{\rho}\right) + \delta_3\left(1 - \frac{z}{\rho}\right)^{3/2} + O\left(\left(1 - \frac{z}{\rho}\right)^2\right).$$

[2] Unrooted trees are also called sometimes *free trees*.

What is noticeable is the cancellation in coefficients for the term $Z^{1/2}$ (since $1-x-\frac{1}{2}(1-x)^2 = \frac{1}{2}+O(x^2)$), so that $Z^{3/2}$ is the actual singularity type of I. Clearly, the constant δ_3 is computable from the first four terms in the singular expansion of H at ρ. Then singularity analysis yields: *The number of unrooted trees of size n satisfies the formula*

$$I_n \sim \frac{3\delta_3}{4\sqrt{\pi n^5}}\rho^{-n}, \qquad I_n \sim 0.5349496061 \cdot 2.9955765856^n n^{-5/2}. \tag{33}$$

The numerical values are from [**61**] and the result is Otter's original [**151**]. The formula gives an error slighly under 10^{-2} for $n = 100$. An unrooted tree of size n gives rise to about different $0.8n$ rooted trees on average, which agrees well with the observations of (31).

▷ **20.** *Dissimilarity property of trees.* Fix an unrooted tree. Two vertices u, v are *similar* there exists an automorphism of the tree that exchanges u and v. Two edges e, f are similar if there is an automorphism of the tree that exchanges e and f. Let p^* and q^* be the number of equivalence classes of vertices and edges under similarity. Then the following identity, called the dissimilarity theorem, holds,

$$p^* - q^* = 1 - s,$$

where s is 1 if the tree admits a "central symmetric" edge and 0 otherwise. Summing over all unrooted trees gives the basic formula (32). [This is based on Read's account in [**159**, p. 107] and on [**104**, p. 56].] ◁

VII. 3. Positive implicit functions

Our goal here is to show that square root singularity holds in the wide context of generating functions satisfying a single functional equation,

$$y(z) = G(z, y(z)), \qquad G(0,0) = 0,\ G'_w(0,0) = 0, \tag{34}$$

under suitable conditions on G. (By trivial algebra, the initial conditions on G can be further weakened to $G(0,0) > 0$ and $0 \le G'_w(0,0) < 1$). For many combinatorial applications, it is legitimate to assume that G is *bivariate analytic* around $(0,0)$ and that it has nonnegative Taylor coefficients there, an assumption that is made throughout this section. Thus, we postulate that the representation

$$G(z,w) = \sum_{m,n\ge 0} g_{m,n} z^m w^n, \qquad \text{with} \quad g_{m,n} \ge 0,\ g_{0,0} = 0,\ g_{0,1} = 0, \tag{35}$$

holds in a domain $|z| < R$, $|y| < S$. Next, in order to avoid trivialities, we consider only $G's$ that are *nonlinear*, that is, one at least of the coefficients $g_{m,n}$ for $n \ge 2$ is strictly positive. (Else, one could plainly solve the linear equation for $y(z)$.) Accordingly, we also postulate that

$$g_{m,n} > 0 \quad \text{for some } m \text{ and for some } n \ge 2. \tag{36}$$

Finally, for reasons of the same nature as in the discussion of simple families of trees, singularities may only occur from a failure of the implicit function theorem. As we shall see in the proof of the main theorem below, this necessitates the existence of two numbers $r, s > 0$ such that

$$G(r,s) = s, \quad G_w(r,s) = 1, \qquad \text{with} \quad r < R, \quad s < S. \tag{37}$$

The system (37) is called the *characteristic system*. Finally, we recall that a function $f(z) = \sum_{n\ge 0} f_n z^n$ is *aperiodic* if cannot be put under the form $f(z) = z^a h(z^d)$ with

$d > 2$ and h analytic at 0. (Put otherwise, there exist three indices n_1, n_2, n_3 such that $f_{n_1} f_{n_2} f_{n_3} \neq 0$ and $\gcd(n_2 - n_1, n_3 - n_1) = 1$.) The main result here [**17, 140**] is[3]

THEOREM VII.3 (Positive implicit functions). *Let $G(z, w)$ be a positive bivariate analytic function satisfying* (35) *and let the equation* (34) *admit a solution $y(z)$ that is analytic at 0, has nonnegative coefficients, and is aperiodic. Assume further the conditions* (36) *and* (37)*. Then, $y(z)$ converges at $z = r$ where it has a square-root singularity:*

$$y(z) \underset{z\to r}{=} s - \gamma\sqrt{1 - z/r} + O(1 - z/r), \qquad \gamma := \sqrt{\frac{2rG'_z(r,s)}{G''_{ww}(r,s)}}.$$

The expansion is valid in a Δ-domain, so that

$$[z^n]y(z) \underset{n\to\infty}{=} \frac{\gamma}{2\sqrt{\pi n^3}} r^{-n} \left(1 + O(n^{-1})\right).$$

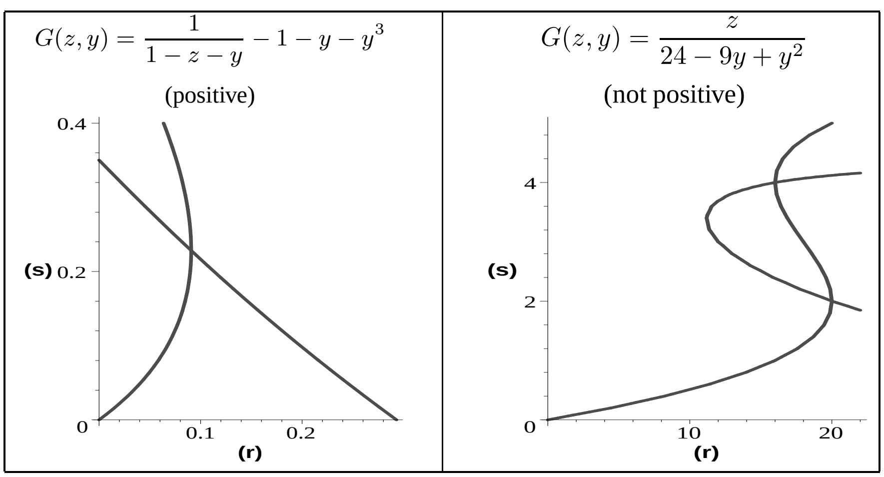
Note that statement tacitly implies the existence of at most one root of the characteristic system within the analyticity domain of G.

PROOF. By assumption $y(z)$ is analytic at the origin with nonnegative Taylor coefficients. Let ρ be its radius of convergence, and $\tau = y(\rho)$. By Pringsheim's theorem, ρ is a singularity of y. Mair and Moon [**140**] provide an argument to the effect that $\rho = r$ and $\tau = s$. The square-root singularity then follows from the usual argument based on the failure of the implicit function theorem. $\square$

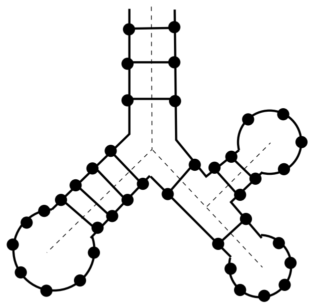
The solutions to the characteristic system (37) can be regarded as the intersection points of two curves, namely,

$$G(r, s) - s = 0, \qquad G'_w(r, s) = 1.$$

Here are plots in the case of two functions G: the first one has nonnegative coefficients while the second one (corresponding to a counterexample of Canfield [**31**]) involves negative coefficients. Positivity of coefficients implies convexity properties that avoid pathological situations.



[3]This theorem has an interesting history. A version of it was first stated by Bender in 1974 (Theorem 5 of [**17**]). Canfield [**31**] then pointed out the fact that Bender's conditions were not quite sufficient. A corrected statement was given by Meir and Moon in [**140**] with a further (minor) erratum in [**141**]. We follow heer the form given in Theorem 10.13 by Odlyzko [**147**] with the correction of another minor misprint (regarding $g_{0,1}$ which should read $g_{0,1} \neq 1$ or, better, $g_{0,1} \in [0, 1[$).



A fragment of RNA is, in first approximation, a tree-like structure with edges corresponding to bases pairs and "loops" corresponding to leaves. There are constraints on the sizes of leaves (taken here between 4 and 7) and length of edges (here between 1 and 4 base pairs). Such a RNA fragment can then be viewed as a planted tree P attached to a binary tree (Y) with equations:

$$\begin{cases} P = AY, \quad Y = AY^2 + B, \\ A = z^2 + z^4 + z^6 + z^8, \ B = z^4 + z^5 + z^6 + z^7. \end{cases}$$

FIGURE 6. A simplified combinatorial model of RNA structures as considered by Waterman *et al.* [**110, 167, 177**].

EXAMPLE 7. *Trees with variable edges lengths and node sizes.* Consider unlabelled plane trees in which nodes can be of different sizes: what is given is a set $\widehat{\Omega}$ of ordered pairs (ω, σ), where a value (ω, σ) means that a node of degree ω and size σ is allowed. Simple varieties in their basic form correspond to $\sigma \equiv 1$; trees enumerated by leaves (including hierarchies) correspond to $\sigma \in \{0, 1\}$ with $\sigma = 1$ iff $\omega = 0$. Figure 6 indicates the way such trees can model the self-bonding of single stranded nucleic acids like RNA, according to Waterman *et al.* [**110, 167, 177**]. Clearly an extremely large number of variations are possible.

The fundamental equation in the case of a finite $\widehat{\Omega}$ is

$$Y(z) = P(z, Y(z)), \qquad P(z, w) := \sum_{(\omega,\sigma)\in\widehat{\Omega}} z^\sigma w^\omega.$$

In the aperiodic case, we shall invariably have a formula of the form

$$Y_n \sim C \cdot A^n n^{3/2},$$

with the universal square-root singularity and the universal count exponent $-\frac{3}{2}$. □

▷ **21.** *Schröder numbers.* Consider the class $\mathcal{Y}$ of unary-binary trees where unary nodes have size 2, while leaves and binary nodes have the usual size 1. The GF satisfies $Y = z + z^2 Y + zY^2$, so that

$$Y(z) = zD(z^2), \qquad D(z) = \frac{1 - z - \sqrt{1 - 6z + z^2}}{2z}.$$

We have $D(z) = 1 + 2\,z + 6\,z^2 + 22\,z^3 + 90\,z^4 + 394\,z^5 + \cdots$, which is *EIS* **A006318** ("Large Schröder numbers"). By the bijective correspondence between trees and lattice paths, $\mathcal{Y}_{2n+1}$ is in bijective correspondence with excursions of length n made of steps $(1, 1)$, $(2, 0)$, $(1, -1)$. Upon tilting by 45°, this is equivalent to paths connecting the lower left corner to the upper right corner of an $(n \times n)$ square that are made of horizontal, vertical, and diagonal steps, and never going under the diagonal. The series $S = \frac{z}{2}(1 + D)$ enumerates Schröder's generalized parenthesis systems (Chapter I): $S := z + S^2/(1 - S)$, and the asymptotic formula

$$S_n \sim \frac{1}{2} D_{n-1} \sim \frac{1}{4\sqrt{\pi n^3}} \left(3 - 2\sqrt{2}\right)^{-n+1/2}$$

follows straightforwardly. ◁

EXAMPLE 8. *Nonplane trees and alkanes.* In chemistry, carbon atoms (C) are known to have valency 4 while hydrogen (H) has valency 1. *Alkanes*, also known as paraffins (Figure 7), are are acyclic molecules formed of carbon and hydrogen atoms according to this rule and without multiple bonds; they are thus of the type $C_n H_{2n+2}$. In combinatorial

```
   H         H  H          H  H  H          H  OH H
   |         |  |          |  |  |          |  |  |
   |         |  |          |  |  |          |  |  |
H--C--H   H--C--C--H    H--C--C--C--H    H--C--C--C--H
   |         |  |          |  |  |          |  |  |
   |         |  |          |  |  |          |  |  |
   H         H  H          H  H  H          H  H  H

 Methane    Ethane         Propane          Propanol
```

FIGURE 7. A few examples of alkanes (CH_4, C_2H_6, C_3H_8) and an alcohol.

terms, we are talking of unrooted trees with (total) node degrees in $\{1, 4\}$. The rooted version of these trees are determined by the fact that a root is chosen and (out)degrees of nodes lie in the set $\Omega = \{0, 3\}$; these are rooted ternary trees and they correspond to alcohols (with the OH group marking one of the carbon atoms). The problem of enumerating isomers of alkanes and alcohols has been at the origin of Pólya paper [**158, 159**].

Alcohols ($\mathcal{A}$) are the simplest to enumerate as they are rooted trees. The OGF starts as (*EIS* **A000598**)

$$A(z) = 1 + z + z^2 + z^3 + 2\,z^4 + 4\,z^5 + 8\,z^6 + 17\,z^7 + 39\,z^8 + 89\,z^9 + \cdots ,$$

with size being taken here as the number of internal nodes. The specification is

$$\mathcal{A} = \{\epsilon\} + \mathcal{Z}\mathfrak{M}_3(\mathcal{A}).$$

(Equivalently $\mathcal{A}^+ := \mathcal{A} \setminus \{\epsilon\}$ satisfies $\mathcal{A}^+ = \mathcal{Z}\mathfrak{M}0, 1, 2, 3(\mathcal{A}^+)$.) This implies that $A(z)$ satisfies the functional equation:

$$A(z) = 1 + z\left(\frac{1}{3}A(z^3) + \frac{1}{2}A(z)A(z^2) + \frac{1}{6}A(z)^3\right).$$

In order to apply Theorem VII.3, introduce the function

$$G(z, w) = 1 + z\left(\frac{1}{3}A(z^3) + \frac{1}{2}A(z^2)w + \frac{1}{6}w^3\right),$$

which exists in $|z| < |\rho|^{1/2}$ and $|w| < \infty$, with ρ the (yet unknown) radius of convergence of A. Like before, the Pólya terms $A(z^2), A(z^3)$ are teated as known functions. By methods similar to those used in the analysis of binary and general trees (Subsection VII. 2.3), we find that the characteristic system admits a solution,

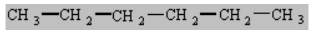
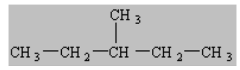
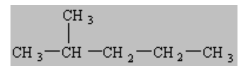
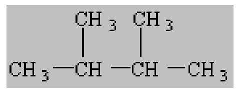
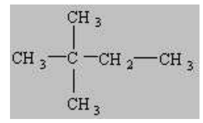
$$r \doteq 0.35518174231437739 28, \quad s \doteq 2.11742070095363102 25,$$

so that $\rho = r$ and $y(\rho) = s$. Thus the growth of the number of alcohols is of the form $C \cdot 2.81546^n n^{-3/2}$.

Let $B(z)$ be the OGF of alkanes (*EIS* **A000602**):

$$\begin{aligned} B(z) \quad &= \quad 1 + z + z^2 + 2\,z^3 + 3\,z^4 + 5\,z^5 + 9\,z^6 + 18\,z^7 \\ &\qquad + 35\,z^8 + 75\,z^9 + 159\,z^{10} + 355\,z^{11} + 802\,z^{12} + 1858\,z^{13} + \cdots \end{aligned}$$

For instance, $B_6 = 5$ because there are 5 isomers of hexane, C_6H_{14}, for which chemists had to develop a nomenclature system, interestingly enough based on a diameter of the tree:

Hexane	3-Methylpentane	2-Methylpentane
		
2,3-Dimethylbutane	2,2-Dimethylbutane	
		

The number of structurally different alkanes can then be found an adaptation of the dissimilarity formula (32) for which we refer to [**19**, p.290]. This problem has served as a powerful motivation for the enumeration of graphical trees and its has a fascinating history which goes back to Cayley. (See Rains and Sloane's article [**161**] and [**159**]). The asymptotic formula of alkanes is of the usual form with an $n^{-5/2}$ term, as these are unrooted molecules. □

The pattern of analysis should by now be clear, and we state:

THEOREM VII.4 (Nonplane unlabelled trees). *Let $\Omega \ni 0$ be a finite subset of $\mathbb{Z}_{\geq 0}$ and consider the variety $\mathcal{V}$ of (rooted) nonplane unlabelled trees. Assume aperiodicity ($\gcd(\Omega) = 1$) and Ω contains at least one element larger than 1. Then the number of trees of size n in $\mathcal{V}$ satisfies an asymptotic formula:*

$$V_n \sim C \cdot A^n n^{-3/2}.$$

PROOF. The argument given for alcohols is transposed verbatim. Only the existence of a root of the characteristic system needs to be established.

The radius of convergence of $V(z)$ is a priori ≤ 1. The fact that ρ is strictly less than 1 is established by means of a lower bound, $V_n > B^n$ for some $B > 1$ and infinitely many large enough values of n. To obtain this exponential variability, first choose an n_0 such that $V_{n_0} > 1$, then build a perfect d-ary tree (for some $d \in \Omega$, $d \neq 0, 1$) tree of height h, and finally graft freely subtrees of size n_0 at $n/(4n_0)$ of the leaves of the perfect tree. Choosing d such that $d^h > n/(4n_0)$ yields the lower bound. That the radius of convergence is nonzero results from the upper bound provided by corresponding plane trees whose growth is at leat exponential. Thus, one has $0 < \rho < 1$.

By the translation of multisets of bounded cardinality, the function G is polynomial in finitely many of the quantities $\{V(z), V(z^2), \ldots\}$. Thus the function $G(z, w)$ constructed like in the case of alcohols converges in $|z| < \rho^{1/2}$, $|w| < \infty$. As $z \to \rho^{-1}$, we must have $\tau := V(\rho)$ finite, since otherwise, there would be a contradiction in orders of growth in the nonlinear equation $V(z) = \cdots + \cdots V(z)^d \cdots$ as $z \to \rho$. Thus (ρ, τ) satisfies $\tau = G(\rho, \tau)$. For the derivative, one must have $G'_w(\rho, \tau) = 1$ since': (i) a smaller value would mean that V is analytic at ρ (by the Implicit Function Theorem); (ii) a larger value would mean that a singularity has been encountered earlier (by the usual argument on failure of the Implicit Function Theorem). Thus, Theorem VII.3 on positive implicit functions is applicable. □

A large number of variations are clearly possible as evidenced by the title of an article [**103**] published by Harary, Robinson, and Schwenk in 1975, namely, "Twenty-step algorithm for determining the asymptotic number of trees of various species"

VII. 4. The analysis of algebraic functions

Algebraic series and algebraic functions are simply defined as solutions of a polynomial equation (Definition VII.3). It is a nontrivial fact established by elimination theory (which can itself be implemented by way of resultants or Groebner bases) that they are equivalently defined as components of solutions of polynomial systems .

The starting point is the following definition of an algebraic function.

DEFINITION VII.3. *A function $f(z)$ analytic in a neighbourhod $\mathcal{V}$ of a point z_0 is said to be* algebraic *if there exists a (nonzero) polynomial $P(z,y) \in \mathbb{C}[z,y]$, such that*

$$P(z, f(z)) = 0, \qquad z \in \mathcal{V}.$$

A power series $f \in \mathbb{C}[\![z]\!]$ is said to be an algebraic power series if it coincides with the expansion of an algebraic function at 0.

$\triangleright$ **22.** *Algebraic definition of algebraic series.* It is customary to define f to be an algebraic series if it satisfies $P(z,f) = 0$ in the sense of formal power series, without *a priori* consideration of convergence issues. Then the technique of majorizing series may be used to prove that the coefficients of f grow at most exponentially. Thus, the new definition is equivalent to Definition —refalg-def. $\triangleleft$

The *degree* of an algebraic series f is by definition the minimal value of $\deg_y P(z,y) = 1$ over all polynomials that are cancelled by f (so that rational series are algebraic of degree 1). Note that one can always assume P to be irreducible (that is $P = QR$ implies that one of Q or R is a scalar) and of minimal degree.

An algebraic function may also be defined by starting with a polynomial system of the form

$$\left\{\begin{array}{lcl} P_1(z, y_1, \dots, y_m) & = & 0 \\ \vdots & & \\ P_m(z, y_1, \dots, y_m) & = & 0, \end{array}\right. \tag{38}$$

where each P_j is a polynomial. A solution of (38) is by definition an m-tuple $(f_1, \dots, f_m)$ that cancels each P_j, that is, $P_j(z, f_1, \dots, f_m) = 0$. Any of the f_j is called a component solution. A basic result of *elimination theory* is that any component solution of a non-degenerate polynomial system is an algebraic series (APPENDIX: *Algebraic elimination*, p. 219). In other words, one can eliminate the auxiliary variables $y_2, \dots, y_m$ and construct a single bivariate polynomial Q such that $Q(z, y_1) = 0$.

Algebraic functions have singularities constrained to be *branch points*. By this is meant that the local expansion at such a singularity is a fractional power series known as a Newton–Puiseux expansion. Singularity analysis is systematically applicable to algebraic functions—hence the characteristic form of asymptotic expansions that involve terms of the form $\omega^n n^{p/q}$ (for some algebraic number ω and some rational exponent p/q). In this section, we develop such basic structural results (Subsection VII. 4.1). However, coming up with effective solutions (i.e., decision procedures) is not obvious in the algebraic case. Hence, a number of nontrivial algorithms are also described in order to locate and analyse singularities (Newton's polygon method), and eventually determine the asymptotic form of coefficients. In particular, the multivalued character of algebraic functions creates a need to solve "connection problems". Finally, like in the rational case, positive systems (Subsection VII. 4.2) enjoy special properties that further constrain what can be observed as regards asymptotic behaviours, including a return of the square-root singularity. Our presentation of positive systems is based on an essential result of the theory, the Drmota–Lalley–Woods theorem, that plays for algebraic functions a rôle quite similar to that of Perron-Frobenius theory for rational functions.

VII. 4.1. General algebraic functions. Let $P(z,y)$ be an irreducible polynomial of $\mathbb{C}[z,y]$,

$$P(z,y) = p_0(z) y^d + p_1(z) y^{d-1} + \cdots + p_d(z).$$

The solutions of the polynomial equation $P(z, y) = 0$ define a locus of points (z, y) in $\mathbb{C} \times \mathbb{C}$ that is known as a *complex algebraic curve*. Let d be the y-degree of P. Then, for each z there are at most d possible values of y. In fact, there exist d values of y "almost always", that is except for a finite number of cases:

— If z_0 is such that $p_0(z_0) = 0$, then there is a reduction in the degree in y and hence a reduction in the number of y-solutions for the particular value of $z = z_0$. One can conveniently regard the points that disappear as "points at infinity".

— If z_0 is such that $P(z_0, y)$ has a multiple root, then some of the values of y will coalesce.

Define the *exceptional set* of P as the set ($\mathbf{R}$ is the resultant):

$$\Xi[P] := \{z \mid R(z) = 0\}, \; R(z) := \mathbf{R}(P(z, y), \partial_y P(z, y), y). \tag{39}$$

(The quantity $R(z)$ is also known as the discriminant of $P(z, y)$ taken as a function of y.) If $z \notin \Xi[p]$, then we have a guarantee that there exist d distinct solutions to $P(z, y) = 0$, since $p_0(z) \neq 0$ and $\partial_y P(z, y) \neq 0$. Then, by the implicit function theorem, each of the solutions y_j lifts into a locally analytic function $y_j(z)$. What we call a *branch* of the algebraic curve $P(z, y) = 0$ is the choice of such an $y_j(z)$ together with a connected region of the complex plane throughout which this particular $y_j(z)$ is analytic.

Singularities of an algebraic function can thus only occur if z lies in the exceptional set $\Xi[P]$. At a point z_0 such that $p_0(z_0) = 0$, some of the branches escape to infinity, thereby ceasing to be analytic. At a point z_0 where the resultant polynomial $R(z)$ vanishes but $p_0(z) \neq 0$, then two or more branches collide. This can be either a multiple point (two or more branches happen to assume the same value, but each one exists as an analytic function around z_0) or a branch point (some of the branches actually cease to be analytic). An example of an exceptional point that is not a branch point is provided by the classical lemniscate of Bernoulli: at the origin, two branches meet while each one is analytic there (see Figure 8).

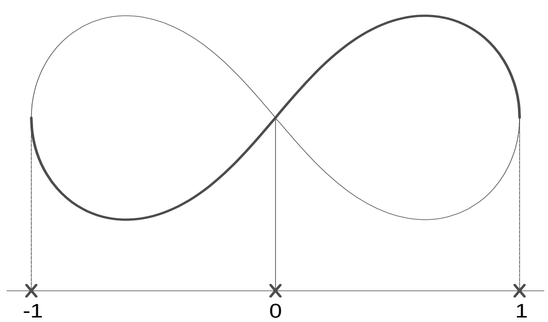


FIGURE 8. The lemniscate of Bernoulli defined by $P(z, y) = (z^2 + y^2)^2 - (z^2 - y^2) = 0$: the origin is a double point where two analytic branches meet.

A partial knowledge of the topology of a complex algebraic curve may be gotten by first looking at its restriction to the reals. Consider for instance the polynomial equation $P(z, y) = 0$, where

$$P(z, y) = y - 1 - zy^2,$$

which defines the OGF of the Catalan numbers. A rendering of the real part of the curve is given in Figure 9. The complex aspect of the curve as given by $\Im(y)$ as a function of z is also displayed there. In accordance with earlier observations, there are normally two

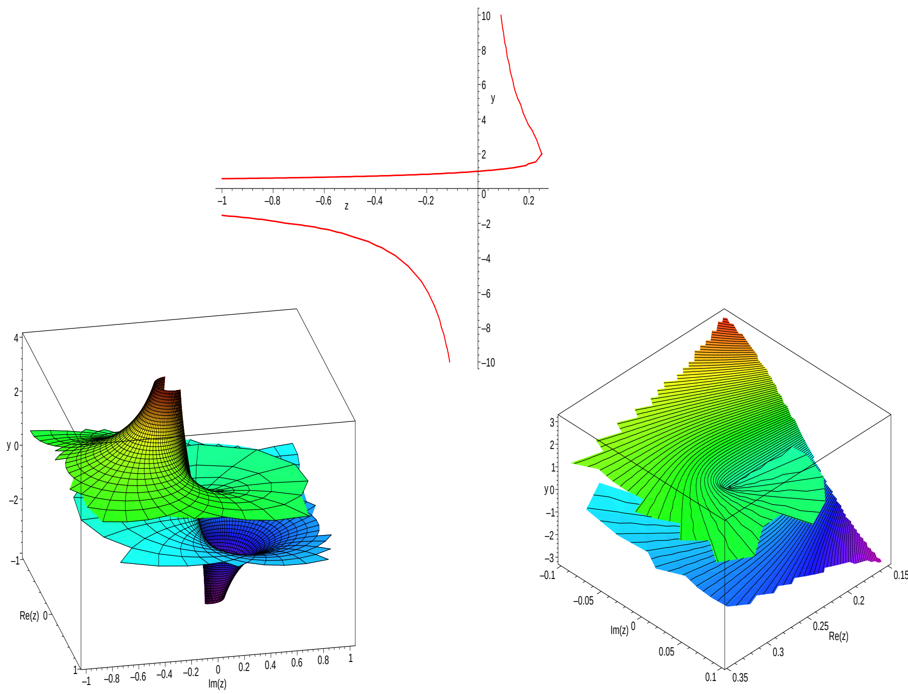


FIGURE 9. The real section of the Catalan curve (top). The complex Catalan curve with a plot of $\Im(y)$ as a function of $z = (\Re(z), \Im(z))$ (bottom left); a blowup of $\Im(y)$ near the branch point at $z = \frac{1}{4}$ (bottom right).

sheets (branches) above each each point. The exceptional set is given by the roots of the discriminant,

$$\mathcal{R} = z(1 - 4z).$$

For $z = 0$, one of the branches escapes at infinity, while for $z = 1/4$, the two branches meet and there is a branch point; see Figure 9.

In summary the exceptional set provides a set of *possible candidates* for the singularities of an algebraic function. This discussion is summarized by the slightly more general lemma that follows.

LEMMA VII.1 (Location of algebraic singularities). *Let $Y(z)$, analytic at the origin, satisfy a polynomial equation $P(z, Y) = 0$. Then, $Y(z)$ can be analytically continued along any half-line emanating from the origin that does not cross any point of the exceptional set* (39)*.*

Nature of singularities. We start the discussion with an exceptional point that is placed at the origin (by a translation $z \mapsto z + z_0$) and assume that the equation $P(0, y) = 0$ has k equal roots $y_1, \ldots, y_k$ where $y = 0$ is this common value (by a translation $y \mapsto y + y_0$ or an inversion $y \mapsto 1/y$, if points at infinity are considered). Consider a punctured disk $|z| < r$ that does not include any other exceptional point relative to P. In the argument that follows, we let $y_1, (z), \ldots, y_k(z)$ be analytic determinations of the root that tend to 0 as $z \to 0$.

Start at at some arbitrary value interior to the real interval $(0, r)$, where the quantity $y_1(z)$ is locally an analytic function of z. By the implicit function theorem, $y_1(z)$ can be continued analytically along a circuit that starts from z and returns to z while simply encircling the origin (and staying within the punctured disk). Then, by permanence of analytic relations, $y_1(z)$ will be taken into another root, say, $y_1^{(1)}(z)$. By repeating the process, we see that after a certain number of times κ with $1 \leq \kappa \leq k$, we will have obtained a collection of roots $y_1(z) = y_1^{(0)}(z), \ldots, y_1^{(\kappa)}(z) = y_1(z)$ that form a set of κ distinct values. Such roots are said to form a *cycle*. In this case, $y_1(t^\kappa)$ is an analytic function of t except possibly at 0 where it is continuous and has value 0. Thus, by general principles (regarding removable singularities), it is in fact analytic at 0. This in turn implies the existence of a convergent expansion near 0:

$$y_1(t^\kappa) = \sum_{n=1}^{\infty} c_n t^n. \tag{40}$$

The parameter t is often called the *local uniformizing parameter*, as it reduces a multivalued function to a single value one. This translates back into the world of z: each determination of $z^{1/\kappa}$ yields one of the branches of the multivalued analytic function as

$$y_1(z) = \sum_{n=1}^{\infty} c_n z^{n/\kappa}. \tag{41}$$

Alternatively, with $\omega = e^{2i\pi/\kappa}$ a root of unity, the κ determinations are obtained as

$$y_1^{(j)} = \sum_{n=1}^{\infty} c_n \omega^n z^{n/\kappa},$$

each being valid in a sector of opening $< 2\pi$. (The case $\kappa = 1$ corresponds to an analytic branch.)

If $r = k$, then the cycle accounts for all the roots which tend to 0. Otherwise, we repeat the process with another root and, in this fashion, eventually exhaust all roots. Thus, all the k roots that have value 0 at $z = 0$ are grouped into cycles of size $\kappa_1, \ldots, \kappa_\ell$. Finally, values of y at infinity are brought to zero by means of the change of variables $y = 1/u$, then leading to negative exponents in the expansion of y.

THEOREM VII.5 (Newton–Puiseux expansions at a singularity). *Let $f(z)$ be a branch of an algebraic function $P(z, f(z)) = 0$. In a circular neighbourhood of a singularity ζ slit along a ray emanating from ζ, $f(z)$ admits a fractional series expansion (Puiseux expansion) that is locally convergent and of the form*

$$f(z) = \sum_{k \geq k_0} c_k (z - \zeta)^{k/\kappa},$$

for a fixed determination of $(z - \zeta)^{1/\kappa}$, where $k_0 \in \mathbb{Z}$ and κ is an integer ≥ 2, called the "branching type".

Newton (1643-1727) discovered the algebraic form of Theorem VII.5, published it in his famous treatise *De Methodis Serierum et Fluxionum* (completed in 1671). This method was subsequently developed by Victor Puiseux (1820–1883) so that the name of Puiseux series is customarily attached to fractional series expansions. The argument given above is taken from the neat exposition offered by Hille in [**108**, Ch. 12, vol. II]. It is known as a "monodromy argument", meaning that it consists in following the course of values of an analytic function along paths in the complex plane till it returns to its original value.

Newton polygon. Newton also described a constructive approach to the determination of branching types near a point (z_0, y_0), that by means of the previous discussion can always be taken to be $(0, 0)$. In order to introduce the discussion, let us examine the Catalan generating function near $z_0 = 1/4$. Elementary algebra gives the explicit form of the two branches

$$y_1(z) = \frac{1}{2z}\left(1 - \sqrt{1-4z}\right), \quad y_2(z) = \frac{1}{2z}\left(1 + \sqrt{1-4z}\right),$$

whose forms are consistent with what Theorem VII.5 predicts. If however one starts directly with the equation,

$$P(z, y) \equiv y - 1 - zy^2 = 0$$

then, the translation $z = 1/4 - Z$ (the minus sign is a mere notational convenience), $y = 2 + Y$ yields

$$Q(Z, Y) \equiv -\frac{1}{4}Y^2 + 4Z + 4ZY + ZY^2. \tag{42}$$

Look for solutions of the form $Y = cZ^\alpha(1 + o(1))$ with $c \neq 0$ (the existence is granted *a priori* by the Newton-Puiseux Theorem). Each of the monomials in (42) gives rise to a term of a well determined asymptotic order. respectively $Z^{2\alpha}, Z^1, Z^{\alpha+1}, Z^{2\alpha+1}$. If the equation is to be identically satisfied, then the main asymptotic order of $Q(Z, Y)$ should be 0. Since $c \neq 0$, this can only happen if two or more of the exponents in the sequence $(2\alpha, 1, \alpha + 1, 2\alpha + 1)$ coincide *and* the coefficients of the corresponding monomial in $P(Z, Y)$ is zero, a condition that is an algebraic constraint on the constant c. Furthermore, exponents of all the remaining monomials have to be larger since by assumption they represent terms of lower asymptotic order.

Examination of all the possible combinations of exponents leads one to discover that the only possible combination arises from the cancellation of the first two terms of Q, namely $-\frac{1}{4}Y^2 + 4Z$, which corresponds to the set of constraints

$$2\alpha = 1, \quad -\frac{1}{4}c^2 + 4 = 0,$$

with the supplementary conditions $\alpha + 1 > 1$ and $2\alpha + 1 > 1$ being satisfied by this choice $\alpha = \frac{1}{2}$. We have thus discovered that $Q(Z, Y) = 0$ is consistent asymptotically with

$$Y \sim 4Z^{1/2}, \quad Y \sim -4Z^{1/2}.$$

The process can be iterated upon subtracting dominant terms. It invariably gives rise to complete formal asymptotic expansions that satisfy $Q(Z, Y) = 0$ (in the Catalan example, these are series in $\pm Z^{1/2}$). Furthermore, elementary majorizations establish that such formal asymptotic solutions represent indeed convergent series. Thus, local expansions of branches have indeed been determined.

An algorithmic refinement (also due to Newton) can be superimposed on the previous discussion and is known as the method of *Newton polygons*. Consider a general polynomial

$$Q(Z, Y) = \sum_{j \in J} Z^{a_j} Y^{b_j},$$

and associate to it the finite set of points (a_j, b_j) in $\mathbb{N} \times \mathbb{N}$, which is called the Newton diagram. It is easily verified that the only asymptotic solutions of the form $Y \propto Z^t$ correspond to values of t that are inverse slopes (i.e., $\Delta x/\Delta y$) of lines connecting two or more points of the Newton diagram (this expresses the cancellation condition between two monomials of Q) *and* such that all other points of the diagram are on this line or to the right of it. In other words:

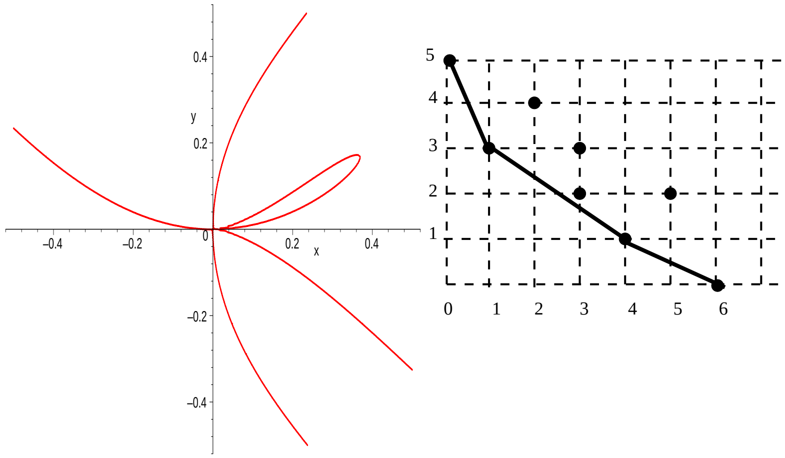


FIGURE 10. The real curve defined by the equation $P = (y-x^2)(y^2-x)(y^2-x^3) - x^3y^3$ near $(0,0)$ (left) and the corresponding Newton diagram (right).

Newton's polygon method. *Any possible exponents τ such that $Y \sim cZ^\tau$ is a solution to a polynomial equation corresponds to one of the inverse slopes of the leftmost convex envelope of the Newton diagram. For each viable τ, a polynomial equation constrains the possible values of the corresponding coefficient c. Complete expansions are obtained by repeating the process, which means deflating Y from its main term by way of the substitution $Y \mapsto Y - cZ^\tau$.*

Figure 10 illustrates what goes on in the case of the curve $p = 0$ where

$$\begin{aligned} P(z,y) &= (y-z^2)(y^2-z)(y^2-z^3) - z^3y^3 \\ &= y^5 - y^3z - y^4z^2 + y^2z^3 - 2\,z^3y^3 + z^4y + z^5y^2 - z^6, \end{aligned}$$

considered near the origin. As the partly factored form suggests, we expect the curve to resemble the union of two orthogonal parabolas and of a curve $y = \pm z^{3/2}$ having a cusp, i.e., the union of

$$y = z^2, \quad y = \pm\sqrt{z}, \quad y = \pm z^{3/2},$$

respectively. It is visible on the Newton diagram of the expanded form that the possible exponents $y \propto z^t$ at the origin are the inverse slopes of the segments composing the envelope, that is,

$$\tau = 2, \quad \tau = \frac{1}{2}, \quad \tau = \frac{3}{2}.$$

For computational purposes, once determined the branching type κ, the value of k_0 that dictates where the expansion starts, and the first coefficient, the full expansion can be recovered by deflating the function from its first term and repeating the Newton diagram construction. In fact, after a few initial stages of iteration, the method of indeterminate coefficients can always be eventually applied[4]. Computer algebra systems usually have this routine included as one of the standard packages; see [**165**].

[4]Bruno Salvy, private communication, August 2000

Asymptotic form of coefficients. The Newton–Puiseux theorem describes precisely the local singular structure of an algebraic function. The expansions are valid around a singularity and, in particular, they hold in indented disks of the type required in order to apply the formal translation mechanisms of singularity analysis (Chapter VI).

THEOREM VII.6 (Algebraic asymptotics). *Let $f(z) = \sum_n f_n z^n$ be an algebraic series. Assume that the branch defined by the series at the origin has a unique dominant singularity at $z = \alpha_1$ on its circle of convergence. Then, the coefficient f_n satisfies the asymptotic expansion,*

$$f_n \sim \alpha_1^{-n} \left(\sum_{k \geq k_0} d_k n^{-1-k/\kappa} \right),$$

where $k_0 \in \mathbb{Z}$ and κ is an integer ≥ 2.

If $f(z)$ has several dominant singularities $|\alpha_1| = |\alpha_2| = \cdots = |\alpha_r|$, then there exists an asymptotic decomposition (where ϵ is some small fixed number, $\epsilon > 0$)

$$f_n = \sum_{j=1}^{r} \phi^{(j)}(n) + O((|\alpha_1| + \epsilon))^{-n},$$

where

$$\phi^{(j)}(n) \sim \alpha_j^{-n} \left(\sum_{k \geq k_0^{(j)}} d_k^{(j)} n^{-1-k/\kappa_j} \right),$$

each $k_0^{(j)}$ is in $\mathbb{Z}$, and each κ_j is an integer ≥ 2.

PROOF. The directional expansions granted by Theorem VII.5 are of the exact type required by singularity analysis (Chapter VI). Composite contours are to be used in the case of multiple singularities, in which case each $\phi^{(j)}(n)$ is the contribution obtained by transfer of a local singular element. □

In the case of multiple singularities, arithmetic cancellations may occur: consider for instance the case of

$$\frac{1}{\sqrt{1 - \frac{6}{5}z + z^2}} = 1 + 0.60z + 0.04z^2 - 0.36z^3 - 0.408z^4 - \cdots,$$

and refer to the corresponding discussion of rational coefficients. Fortunately, such delicate situations tend not to arise in combinatorial situations.

EXAMPLE 9. *Branches of unary-binary trees.* the generating function of unary binary trees is defined by $P(z, f) = 0$ where

$$P(z, y) = y - z - zy - zy^2,$$

so that

$$f(z) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z} = \frac{1 - z - \sqrt{(1+z)(1-3z)}}{2z}.$$

There exist only two branches: f and its conjugate $\bar{f}$ that form a 2–cycle at $\frac{1}{3}$. The singularities of all branches are at $0, -1, \frac{1}{3}$ as is apparent from the explicit form of f or from the defining equation. The branch representing $f(z)$ at the origin is analytic there (by a general argument or by the combinatorial origin of the problem). Thus, the dominant singularity of $f(z)$ is at $\frac{1}{3}$ and it is unique in its modulus class. The "easy" case of Theorem VII.7 then applies once $f(z)$ has been expanded ear $\frac{1}{3}$. As a rule, the organization of computations

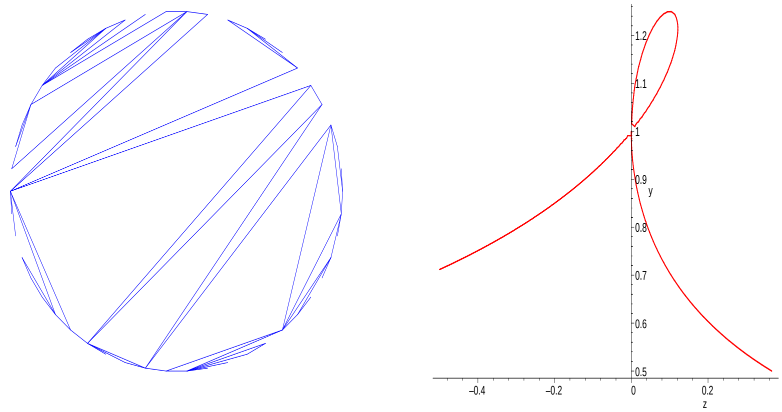


FIGURE 11. Non-crossing graphs: (a) a random connected graph of size 50; (b) the real algebraic curve corresponding to non-crossing forests.

is simpler if one makes use of the local uniformizing parameter with a choice of sign in accordance to the direction along which the singularity is approached. In this case, we set $z = \frac{1}{3} - \delta^2$ and find

$$f(z) = 1 - 3\,\delta + \frac{9}{2}\delta^2 - \frac{63}{8}\delta^3 + \frac{27}{2}\delta^4 - \frac{2997}{128}\delta^5 + \cdots, \qquad \delta = (\frac{1}{3} - z)^{1/2}.$$

This translates immediately into

$$f_n \equiv [z^n]f(z) \sim \frac{3^{n+1/2}}{2\sqrt{\pi n^3}}\left(1 - \frac{15}{16n} + \frac{505}{512n^2} - \frac{8085}{8192n^3} + \cdots\right).$$

The approximation provided by the first three terms is quite good: for $n = 10$ already, it estimates $f_{10} = 835.$ with an error less than 1. □

EXAMPLE 10. *Branches of non-crossing forests.* Consider the polynomial equation $P(z, y) = 0$, where

$$P(z, y) = y^3 + (z^2 - z - 3)y^2 + (z + 3)y - 1,$$

and the combinatorial GF satisfying $P(z, F) = 0$ determined by the initial conditions (*EIS* **A054727**:

$$F(z) = 1 + 2z + 7z^2 + 33z^3 + 181z^4 + 1083z^5 + \cdots.$$

(Combinatorial aspects are developed below in Section VII. 5.1, and $F_n = [z^n]F(z)$ is the number of non-crossing graphs of size n that are forests.)

The exceptional set is mechanically computed: its elements roots of the discriminant

$$R = -z^3(5z^3 - 8z^2 - 32z + 4).$$

Newton's algorithm shows that two of the branches at 0, say y_0 and y_2, form a cycle of length 2 with $y_0 = 1 - \sqrt{z} + O(z)$, $y_2 = 1 + \sqrt{z} + O(z)$ while it is the "middle branch" $y_1 = 1 + z + O(z^2)$ that corresponds to the combinatorial GF $F(z)$.

The nonzero exceptional points are the roots of the cubic factor of $\mathcal{R}$, namely

$$\Omega = \{-1.93028,\ 0.12158,\ 3.40869\}.$$

Let $\xi \doteq 0.1258$ be the root in $(0, 1)$. By Pringsheim's theorem and the fact that the OGF of an infinite combinatorial class must have a positive dominant singularity in $[0, 1]$, the only possibility for the dominant singularity of $y_1(z)$ is ξ. (For a more general argument, see below.)

For z near ξ, the three branches of the cubic give rise to one branch that is analytic with value approximately 0.67816 and a cycle of two conjugate branches with value near 1.21429 at $z = \xi$. The expansion of the two conjugate branches is of the singular type,

$$\alpha \pm \beta\sqrt{1 - z/\xi},$$

where

$$\alpha = \frac{43}{37} + \frac{18}{37}\xi - \frac{35}{74}\xi^2 \doteq 1.21429, \quad \beta = \frac{1}{37}\sqrt{228 - 981\xi - 5290\xi^2} \doteq 0.14931.$$

The determination with a minus sign must be adopted for representing the combinatorial GF when $z \to \xi^-$ since otherwise one would get negative asymptotic estimates for the nonnegative coefficients. Alternatively, one may examine the way the three real branches along $(0, \xi)$ match with one another at 0 and at ξ^-, then conclude accordingly.

Collecting partial results, we finally get by singularity analysis the estimate

$$F_n = \frac{\beta}{2\sqrt{\pi n^3}}\omega^n \left(1 + O(\frac{1}{n})\right), \qquad \omega = \frac{1}{\xi} \doteq 8.22469$$

where the cubic algebraic number ξ and the sextic β are as above. $\square$

The example above illustrates several important points in the analysis of coefficients of algebraic functions when there are no simple explicit radical forms. First of all a given combinatorial problem determines a unique branch of an algebraic curve at the origin. Next, the dominant singularity has to be identified by "connecting" the combinatorial branch with the branches at every possible singularity of the curve. Finally, computations tend to take place over algebraic numbers and not simply rational numbers.

So far, examples have illustrated the common situation where the exponent at the dominant singularity is $\frac{1}{2}$, which is reflected by a factor of $n^{-3/2}$ in the asymptotic form of coefficients. Our last example shows a case where the exponent assumes a different value, namely $\frac{1}{4}$.

EXAMPLE 11. *Branches of "supertrees".* Consider the quartic equation

$$y^4 - 2\,y^3 + (1 + 2\,z)\,y^2 - 2\,yz + 4\,z^3 = 0$$

and let K be the branch analytic at 0 determined by the initial conditions:

$$K(z) = 2\,z^2 + 2\,z^3 + 8\,z^4 + 18\,z^5 + +64\,z^6 + 188\,z^7 + \cdots .$$

(This ogf in fact corresponds to bicoloured "supertrees" already studied in Chapter VI, Section 6.)

The discriminant is found to be

$$\mathcal{R} = 16\,z^4\left(16\,z^2 + 4\,z - 1\right)(-1 + 4\,z)^3,$$

with roots at $\frac{1}{4}$ and $(-1 \pm \sqrt{5})/8$. The dominant singularity of the branch of combinatorial interest turns out to be at $z = \frac{1}{4}$ where $K(\frac{1}{4}) = \frac{1}{2}$. The translation $z = \frac{1}{4} + Z, y = \frac{1}{2} + Y$ then transforms the basic equation into

$$4\,Y^4 + 8\,ZY^2 + 16\,Z^3 + 12\,Z^2 + Z = 0.$$

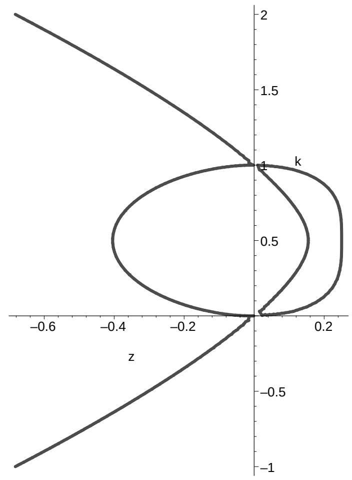


FIGURE 12. The algebraic curve associated to the generating function of supertrees of type K.

According to Newton's polygon, the main cancellation arises from $\frac{1}{4}Y^4 - Z = 0$: this corresponds to a segment of inverse slope $1/4$ in the Newton diagram and accordingly to a cycle formed with 4 conjugate branches, i.e., a fourth-root singularity. Thus, one has,

$$K(z) \underset{z\to\frac{1}{4}}{\sim} 1/2 - \frac{1}{\sqrt{2}}\left(\frac{1}{4}-z\right)^{1/4} - \frac{1}{\sqrt{2}}\left(\frac{1}{4}-z\right)^{3/4} + \cdots, \qquad [z^n]K(z) \underset{n\to\infty}{\sim} \frac{4^n}{8\Gamma(\frac{3}{4})n^{5/4}},$$

which is consistent with earlier found values. Observe that we have started here with the raw algebraic equation satisfies by K. □

Computable coefficient asymptotics. The previous discussion contains the germ of a complete algorithm for deriving an asymptotic expansion of coefficients of any algebraic function. We sketch here the main principles leaving some of the details to the reader. Observe that the problem is a *connection problems*: the "shapes" of the various sheets around each point (including the exceptional points) are known, but it remains to connect them together and see which ones are encountered first when starting from a given branch at the origin.

Algorithm ACA: *Algebraic Coeffi cient Asymptotics*.
Input: A polynomial $P(z, y)$ with $d = \deg_y P(z, y)$; a series $Y(z)$ such that $P(z, Y) = 0$ and assumed to be specified by sufficiently many initial terms so as to be distinguished from all other branches.
Output: The asymptotic expansion of $[z^n]Y(z)$ whose existence is granted by Theorem VII.6.
The algorithm consists of three main steps: *Preparation*, *Dominant singularities*, and *Translation*.

I. Preparation: Define the discriminant $R(z) = \mathbf{R}(P, P'_y, y)$.

(P_1) Compute the exceptional set $\Xi = \{z \mid R(z) = 0\}$ and the points of infinity $\Xi_0 = \{z \mid p_0(z) = 0\}$, where $p_0(z)$ is the leading coefficient of $P(z, y)$ considered as a function of y.

(P_2) Determine the Puiseux expansions of all the d branches at each of the points of $\Xi \cup \{0\}$ (by Newton diagrams and/or indeterminate coefficients). This includes the expansion of analytic branches as well. Let $\{y_{\alpha,j}(z)\}_{j=1}^{d}$ be the collection of all such expansions at some $\alpha \in \Xi \cup \{0\}$.

(P_3) Identify the branch at 0 that corresponds to $Y(z)$.

II. Dominant singularities (Controlled approximate matching of branches). Let $\Xi_1, \Xi_2, \ldots$ be a partition of the elements of $\Xi \cup \{0\}$ sorted according to the increasing values of their modulus: it is assumed that the numbering is such that if $\alpha \in \Xi_i$ and $\beta \in \Xi_j$, then $|\alpha| < |\beta|$ is equivalent to $i < j$. Geometrically, the elements of Ξ have been grouped in concentric circles. First, a preparation step is needed.

(D_1) Determine a nonzero lower bound δ on the radius of convergence of any local Puiseux expansion of any branch at any point of Ξ. Such a bound can be constructed from the minimal distance between elements of Ξ and from the degree d of the equation.

The sets Ξ_j are to be examined in sequence until it is detected that one of them contains a singularity. At step j, let $\sigma_1, \sigma_2, \ldots, \sigma_s$ be an arbitrary listing of the elements of Ξ_j. The problem is to determine whether any σ_k is a singularity and, in that event, to find the right branch to which it is associated. This part of the algorithm proceeds by controlled numerical approximations of branches and constructive bounds on the minimum separation distance between distinct branches.

(D_2) For each candidate singularity σ_k, with $k \geq 2$, set $\zeta_k = \sigma_k(1 - \delta/2)$. By assumption, each ζ_k is in the domain of convergence of $Y(z)$ and of any $y_{\sigma_k,j}$.

(D_3) Compute a nonzero lower bound η_k on the minimum distance between two roots of $P(\zeta_k, y) = 0$. This separation bound can be obtained from resultant computations.

(D_4) Estimate $Y(\zeta_k)$ and each $y_{\sigma_k,j}(\zeta_k)$ to an accuracy better than $\eta_k/4$. If two elements, $Y(z)$ and $y_{\sigma_k,j}(z)$ are (numerically) found to be at a distance less than η_k for $z = \zeta_k$, then they are matched: σ_k *is* a singularity and the corresponding $y_{\sigma_k,j}$ is the corresponding singular element. Otherwise, σ_k is declared to be a regular point for $Y(z)$ and discarded as candidate singularity.

The main loop on j is repeated until a singularity has been detected., when $j = j_0$, say. The radius of convergence ρ is then equal to the common modulus of elements of Ξ_{j_0}; the corresponding singular elements are retained.

III. Coefficient expansion. Collect the singular elements at all the points σ determined to be a dominant singularity at Phase III. Translate termwise using the singularity analysis rule,

$$(\sigma - z)^{p/\kappa} \mapsto \sigma^{p/\kappa - n} \frac{\Gamma(-p/\kappa + n)}{\Gamma(-p/\kappa)\Gamma(n+1)},$$

and reorganize into descending powers of n, if needed.

This algorithm vindicates the following assertion:

PROPOSITION VII.3 (Decidability of algebraic connections.). *The dominant singularities of a branch of an algebraic function can be determined by the algorithm* ACA *in a finite number of operations.*

VII. 4.2. Positive algebraic systems. The discussion of algebraic singularities specializes nicely to the case of positive functions. We first indicate a procedure that determines the radius of convergence of any algebraic series with *positive* coefficients. The procedure takes advantage of Pringsheim's theorem that allows us to restrict attention to candidate singularities on the positive half-line. It represents a shortcut that is often suitable for human calculation and, in fact, the it systematizes some of the techniques already used implicitly in earlier examples.

Algorithm ROCPAF: *Radius of Convergence of Positive Algebraic Functions.*
Input: A polynomial $P(z, y)$ with $d = \deg_y P(z, y)$; a series $Y(z)$ such that $P(z, Y) = 0$ that is known to have only nonnegative coefficients ($[z^n]Y(z) \geq 0$) and is assumed to be specified by sufficiently many initial terms.
Output: The radius of convergence ρ of $Y(z)$.

Plane-sweep. Let Ξ^+ be the subset of those elements of the exceptional set Ξ which are positive real.

(R_1) Sort the subset of those branches $\{y_{0,j}\}$ at 0^+ that have totally real coefficients. This is essentially a lexicographic sort that only needs the initial parts of each expansion. Set initially $\xi_0 = 0$ and $U(z) = Y(z)$.

(R_2) Sweep over all $\xi \in \Xi^+$ in increasing order. To detect whether a candidate ξ is the dominant positive singularity, proceed as follows:

- — Sort the branches $\{y_{\xi,j}\}$ at ξ^- that have totally real coefficients.
- — using the orders at ξ_0^+ and ξ^-, match the branch $U(z)$ with its corresponding branch at ξ^-, say $V(z)$; this makes use of the total ordering between real branches at ξ_0^+ and ξ^-. If the branch $V(z)$ is singular, then return $\rho = \xi$ as the radius of convergence of $Y(z)$ and use $V(z)$ as the singular element of $Y(z)$ at $z = \rho$; otherwise continue with the next value of $\xi \in \Xi^+$ while replacing $U(z)$ by $V(z)$ and ξ_0 by ξ.

This algorithm is a plane-sweep that takes advantage of the fact that the real branches near a point can be totally ordered; finding the ordering only requires inspection of a finite number of coefficients. The plane-sweep algorithm enables us to trace at each stage the original branch and keep a record of its order amongst all branches. The method works since no two real branches can cross at a point other than a multiple point, such a point being covered as an element of Ξ^+.

We now turn to positive systems. Most of the combinatorial classes known to admit algebraic generating functions involve singular exponents that are multiples of $\frac{1}{2}$. This empirical observation is supported by the fact, to be proved below, that a wide class of positive systems have solutions with a square-root singularity. Interestingly enough, the corresponding theorem is due to independent research by several authors: Drmota [**50**] developed a version of the theorem in the course of studies relative to limit laws in various families of trees defined by context-free grammars; Woods [**195**], motivated by questions of Boolean complexity and finite model theory, gave a form expressed in terms of colouring rules for trees; finally, Lalley [**128**] came across a similarly general result when quantifying return probabilities for random walks on groups. The statement that follows is a fundamental result in the analysis of algebraic systems arising from combinatorics and is (rightly) called the "Drmota-Lalley-Woods" theorem. Notice that the authors of [**50, 128, 195**] prove more: Drmota and Lalley show how to pull out limit Gaussian laws for simple parameters (e.g., as in [**50**] by a perturbative analysis; see Chapter 9); Woods shows how to deduce estimates of coefficients even in some periodic or non-irreducible cases (see definitions below).

In the treatment that follows we start from a polynomial system of equations,

$$\{y_j = \Phi_j(z, y_1, \dots, y_m)\}, \quad j = 1, \dots, m.$$

We shall discuss in the next section a class of combinatorial specifications, the "context-free" specifications, that leads systematically to such fixed-point systems. The case of linear systems has been already dealt with, so that we limit ourselves here to *nonlinear systems* defined by the fact that at least one polynomial Φ_j is nonlinear in some of the indeterminates $y_1, \dots, y_m$.

First, for combinatorial reasons, we define several possible attributes of a polynomial system.

- — *Algebraic positivity* (or a-positivity). A polynomial system is said to be *a-positive* if all the component polynomials Φ_j have nonnegative coefficients.

Next, we want to restrict consideration to systems that determine a unique solution vector $(y_1, \dots, y_m) \in (\mathbb{C}[[z]])^m$. (This discussion is related to 0-dimensionality in the sense alluded to earlier.) Define the *z-valuation* $\mathrm{v}al(\vec{y})$ of a vector $\vec{y} \in \mathbb{C}[\![z]\!]^m$ as the minimum over all j's of the individual valuations[5] $\mathrm{v}al(y_j)$. The distance between two vectors is defined as usual by $d(\vec{y}, \vec{y}') = 2^{-\mathrm{v}al(\vec{y}-\vec{y}')}$. Then, one has:

— *Algebraic properness* (or a-properness). A polynomial system is said to be *a-proper* if it satisfies a Lipschitz condition

$$d(\Phi(\vec{y}), \Phi(\vec{y}\,')) < K d(\vec{y}, \vec{y}\,') \quad \text{for some } K < 1.$$

In that case, the transformation Φ is a contraction on the complete metric space of formal power series and, by the general fixed point theorem, the equation $y = \Phi(y)$ admits a unique solution. In passing, this solution may be obtained by the iterative scheme,

$$\vec{y}^{(0)} = (0, \ldots, 0)^t, \quad \vec{y}^{(h+1)} = \Phi(y^{(h)}), \quad y = \lim_{h \to \infty} y^{(h)}.$$

The key notion is irreducibility. To a polynomial system, $\vec{y} = \Phi(\vec{y})$, associate its *dependency graph* defined as a graph whose vertices are the numbers $1, \ldots, m$ and the edges ending at a vertex j are $k \to j$, if y_j figures in a monomial of $\Phi_k(j)$. (This notion is reminiscent of the one already introduced for linear system on page V.4.)

— *Algebraic irreducibility* (or a-irreducibility). A polynomial system is said to be *a-irreducible* if its dependency graph is strongly connected.

Finally, one needs a technical notion of periodicity to dispose of cases like

$$y(z) = \frac{1}{2z}\left(1 - \sqrt{1-4z}\right) = z + z^3 + 2z^5 + \cdots,$$

(the OGF of complete binary trees) where coefficients are only nonzero for certain residue classes of their index.

— *Algebraic aperiodicity* (or a-aperiodicity). A power series is said to be *aperiodic* if it contains three monomials (with nonzero coefficients), $z^{e_1}, z^{e_2}, z^{e_3}$, such that $e_2 - e_1$ and $e_3 - e_1$ are relatively prime. A proper polynomial system is said to be aperiodic if each of its component solutions y_j is aperiodic.

THEOREM VII.7 (Positive polynomial systems). *Consider a nonlinear polynomial system $\vec{y} = \Phi(\vec{y})$ that is a-proper, a-positive, and a-irreducible. In that case, all component solutions y_j have the same radius of convergence $\rho < \infty$. Then, there exist functions h_j analytic at the origin such that*

$$y_j = h_j\left(\sqrt{1 - z/\rho}\right) \qquad (z \to \rho^-). \tag{43}$$

In addition, all other dominant singularities are of the form $\rho\omega$ with ω a root of unity. If furthermore the system is a-aperiodic, all y_j have ρ as unique dominant singularity. In that case, the coefficients admit a complete asymptotic expansion of the form

$$[z^n]y_j(z) \sim \rho^{-n}\left(\sum_{k\ge1} d_k n^{-1-k/2}\right). \tag{44}$$

PROOF. The proof consists in gathering by stages consequences of the assumptions. It is essentially based on close examination of "failures" of the implicit function theorem and the way these lead to singularities.

(a) As a preliminary observation, we note that each component solution y_j is an algebraic function that has a nonzero radius of convergence. In particular, singularities are constrained to be of the algebraic type with local expansions in accordance with the Newton-Puiseux theorem (Theorem VII.5).

[5] Let $f = \sum_{n=\beta}^{\infty} f_n z^n$ with $f_\beta \neq 0$; the valuation of f is by definition $\mathrm{v}al(f) = \beta$.

(b) Properness together with the positivity of the system implies that each $y_j(z)$ has nonnegative coefficients in its expansion at 0, since it is a formal limit of approximants that have nonnegative coefficients. In particular, each power series y_j has a certain nonzero radius of convergence ρ_j. Also, by positivity, ρ_j is a singularity of y_j (by virtue of Pringsheim's theorem). From the nature of singularities of algebraic functions, there exists some order $R \geq 0$ such that each Rth derivative $\partial_z^R y_j(z)$ becomes infinite as $z \to \rho_j^-$.

We establish now that $\rho_1 = \cdots = \rho_m$. In effect, differentiation of the equations composing the system implies that a derivative of arbitrary order r, $\partial_z^r y_j(z)$, is a linear form in other derivatives $\partial_z^r y_j(z)$ of the same order (and a polynomial form in lower order derivatives); also the linear combination and the polynomial form have nonnegative coefficients. Assume a contrario that the radii were not all equal, say $\rho_1 = \cdots = \rho_s$, with the other radii $\rho_{s+1}, \ldots$ being strictly greater. Consider the system differentiated a sufficiently large number of times, R. Then, as $z \to \rho_1$, we must have $\partial_z^R y_j$ tending to infinity for $j \leq s$. On the other hand, the quantities y_{s+1}, etc., being analytic, their Rth derivatives that are analytic as well must tend to finite limits. In other words, because of the irreducibility assumption (and again positivity), infinity *has to* propagate and we have reached a contradiction. Thus, all the y_j have the same radius of convergence and we let ρ denote this common value.

(c_1) The key step consists in establishing the existence of a square-root singularity at the common singularity ρ. Consider first the scalar case, that is

$$y - \phi(z, y) = 0, \tag{45}$$

where ϕ is assumed to depend nonlinearly on y and have nonnegative coefficients. The requirement of properness means that z is a factor of all monomials, except the constant term $\phi(0, 0)$.

Let $y(z)$ be the unique branch of the algebraic function that is analytic at 0. Comparison of the asymptotic orders in y inside the equality $y = \phi(z, y)$ shows that (by nonlinearity) we cannot have $y \to \infty$ when z tends to a finite limit. Let now ρ be the radius of convergence of $y(z)$. This argument shows that $y(z)$ is necessarily finite at its singularity ρ. We set $\tau = y(\rho)$ and note that, by continuity $\tau - \phi(\rho, \tau) = 0$.

By the implicit function theorem, a solution (z_0, y_0) of (45) can be continued analytically as $(z, y_0(z))$ in the vicinity of z_0 as long as the derivative with respect to y,

$$J(z_0, y_0) := 1 - \phi_y'(z_0, y_0)$$

remains nonzero. The quantity ρ being a singularity, we must have $J(\rho, \tau) = 0$. (In passing, the system

$$\tau - \phi(\rho, \tau) = 0, \qquad J(\rho, \tau) = 0,$$

determines only finitely many candidates for ρ.) On the other hand, the second derivative $-\phi_{yy}''$ is nonzero at (ρ, τ) (by positivity, since no cancellation can occur); there results by the classical argument on local failures of the implicit function theorem that $y(z)$ has a singularity of the square-root type (see also Chapters 4 and 5). More precisely, the local expansion of the defining equation (45) at (ρ, τ) binds (z, y) locally by

$$-(z - \rho)\phi_z'(\rho, \tau) - \frac{1}{2}(y - \tau)^2 \phi_{yy}''(\rho, \tau) + \cdots = 0,$$

where the subsequent terms are negligible by Newton's polygon method. Thus, we have

$$y - \tau = -\sqrt{\frac{\phi_z(\rho, \tau)}{\phi_{yy}''(\rho, \tau)}}(\rho - z)^{1/2} + \cdots,$$

the negative determination of the square-root being chosen to comply with the fact that $y(z)$ increases as $z \to \rho^-$. This proves the first part of the assertion in the scalar case.

(c_2) In the multivariate case, we graft an ingenious argument [**128**] that is based on a linearized version of the system to which Perron-Frobenius theory is applicable. First, irreducibility implies that any component solution y_j depends nonlinearly on itself (by possibly iterating Φ), so that a discrepancy in asymptotic behaviours would result for the implicitly defined y_j in the event that some y_j tends to infinity.

Now, the multivariate version of the implicit function theorem grants locally the analytic continuation of any solution $y_1, y_2, \dots, y_m$ at z_0 provided there is no vanishing of the Jacobian determinant

$$J(z_0, y_1, \dots, y_m) := \det\left(\delta_{i,j} - \frac{\partial}{\partial y_j}\Phi_i(z_0, y_1, \dots, y_m)\right),$$

where $\delta_{i,j}$ is Kronecker's symbol. Thus, we must have

$$J(\rho, \tau_1, \dots, \tau_m) = 0 \qquad \text{where} \quad \tau_j := y_j(\rho).$$

The next argument (we follow Lalley [**128**]) uses Perron-Frobenius theory and linear algebra. Consider the modified Jacobian matrix

$$K(z_0, y_1, \dots, y_m) := \left(\frac{\partial}{\partial y_j}\Phi_i(z_0, y_1, \dots, y_m)\right),$$

which represents the "linear part" of Φ. For $z, y_1, \dots, y_m$ all nonnegative, the matrix K has positive entries (by positivity of Φ) so that it is amenable to Perron-Frobenius theory. In particular it has a positive eigenvalue $\lambda(z, y_1, \dots, y_m)$ that dominates all the other in modulus. The quantity

$$\widehat{\lambda}(z) = \lambda(y_1(z), \dots, y_m(z))$$

is increasing as it is an increasing function of the matrix entries that themselves increase with z for $z \geq 0$.

We propose to prove that $\widehat{\lambda}(\rho) = 1$, In effect, $\widehat{\lambda}(\rho) < 1$ is excluded since otherwise $(I - K)$ would be invertible at $z = \rho$ and this would imply $J \neq 0$, thereby contradicting the singular character of the $y_j(z)$ at ρ. Assume *a contrario* $\widehat{\lambda}(\rho) > 1$ in order to exclude the other case. Then, by the increasing property, there would exists $\rho_1 < \rho$ such that $\widehat{\lambda}(\rho_1) = 1$. Let v_1 be a left eigenvector of $K(\rho_1, y_1(\rho_1), \dots, y_m(\rho_1))$ corresponding to the eigenvalue $\widehat{\lambda}(\rho_1)$. Perron-Frobenius theory grants that such a vector v_1 has all its coefficients that are positive. Then, upon multiplying on the left by v_1 the column vectors corresponding to y and $\Phi(y)$ (which are equal), one gets an identity; this derived identity upon expanding near ρ_1 gives

$$A(z - \rho_1) = -\sum_{i,j} B_{i,j}(y_i(z) - y_i(\rho_1))(y_j(z) - y_j(\rho_1)) + \cdots, \tag{46}$$

where $\cdots$ hides lower order terms and the coefficients $A, B_{i,j}$ are nonnegative with $A > 0$. There is a contradiction in the orders of growth if each y_i is assumed to be analytic at ρ_1 since the left side of (46) is of exact order $(z - \rho_1)$ while the right side is at least as small as to $(z - \rho_1)^2$. Thus, we must have $\widehat{\lambda}(\rho) = 1$ and $\widehat{\lambda}(x) < 1$ for $x \in (0, \rho)$.

A calculation similar to (46) but with ρ_1 replaced by ρ shows finally that, if

$$y_i(z) - y_i(\rho) \sim \gamma_i(\rho - z)^\alpha,$$

then consistency of asymptotic expansions implies $2\alpha = 1$, that is $\alpha = \frac{1}{2}$. (The argument here is similar to the first stage of a Newton polygon construction.) We have thus proved that the component solutions $y_j(z)$ have a square-root singularity. (The existence of a complete expansion in powers of $(\rho - z)^{1/2}$ results from examination of the Newton diagram.) The proof of the general case (43) is at last completed.

(d) In the aperiodic case, we first observe that each $y_j(z)$ cannot assume an infinite value on its circle of convergence $|z| = \rho$, since this would contradict the boundedness of $|y_j(z)|$ in the open disk $|z| < \rho$ (where $y_j(\rho)$ serves as an upperbound). Consequently, by singularity analysis, the Taylor coefficients of any $y_j(z)$ are $O(n^{-1-\eta})$ for some $\eta > 1$ and the series representing y_j at the origin converges on $|z| = \rho$.

For the rest of the argument, we observe that if $y = \Phi(z, \vec{y})$, then $y = \Phi^{\langle m \rangle}(z, \vec{y})$ where the superscript denotes iteration of the transformation Φ in the variables $\vec{y} = (y_1, \ldots, y_m)$. By irreducibility, $\Phi^{\langle m \rangle}$ is such that *each* of its component polynomials involves *all* the variables.

Assume that there would exists a singularity ρ^* of some $y_j(z)$ on $|z| = \rho$. The triangle inequality yields $|y_j(\rho^*)| < y_j(\rho)$ where strictness is related to the general aperiodicity argument encountered at several other places in this book. But then, the modified Jacobian matrix $K^{\langle m \rangle}$ of $\Phi^{\langle m \rangle}$ taken at the $y_j(\rho^*)$ has entries dominated strictly by the entries of $K^{\langle m \rangle}$ taken at the $y_j(\rho)$. There results (see page 93) that the dominant eigenvalue of $K^{\langle m \rangle}(z, \vec{y}_j(\rho^*))$ must be strictly less than 1. But this would imply that $I - K^{\langle m \rangle}(z, \vec{y}_j(\rho^*))$ is intervertible so that the $y_j(z)$ would be analytic at ρ^*. A contradiction has been reached: ρ is the sole dominant singularity of each y_j and this concludes the argument. $\square$

We observe that the dominant singularity is obtained amongst the positive solutions of the system

$$\vec{\tau} = \Phi(\rho, \vec{\tau}), \quad J(\rho, \vec{\tau}) = 0.$$

For the Catalan GF, this yields for instance

$$\tau - 1 - \rho\tau^2 = 0, \quad 1 - 2\rho\tau = 0,$$

giving back as expected: $\rho = \frac{1}{4}$, $\tau = \frac{1}{2}$.

VII. 5. Combinatorial applications of algebraic functions

In this section, we first present context-free specifications that admit a direct translation into polynomial systems (Section VII. 5.1). When particularized to formal languages, this gives rise to context-free languages that, provided an unambiguity condition is met, lead to algebraic generating functions.

The next two subsections introduce objects whose constructions still lead to algebraic functions, but in a non-obvious way. This includes: walks with a finite number of allowed basic jumps (Section VII. 5.2) and planar maps (Section VII. 5.3). In that case, bivariate functional equations are induced by the combinatorial decompositions. The common form is

$$\Phi(z, u, F(z, u), h_1(z), \ldots, h_r(z)) = 0, \tag{47}$$

where Φ is a known polynomial and the unknowns are F and $h_1, \ldots, h_r$. Specific methods are needed in order to attain solutions to such functional equations that would seem at first glance to be grossly underdetermined. Random walks lead to a linear version of (47) that is treated by the so-called "kernel method". Maps lead to nonlinear versions that are solved by means of Tutte's "quadratic method". In both cases, the strategy consists in binding z and u by forcing them to lie on an algebraic curve (suitably chosen in order to eliminate the dependency on $F(z, u)$), and then pulling out the algebraic consequences of such a specialization.

VII. 5.1. Context-free specifications and languages. A *context-free system* is a collection of combinatorial equations,

$$\left\{\begin{array}{lcl} \mathcal{C}_1 & = & \Phi_1(\vec{a}, \mathcal{C}_1, \ldots, \mathcal{C}_m) \\ \vdots & & \vdots \\ \mathcal{C}_m & = & \Phi_m(\vec{a}, \mathcal{C}_1, \ldots, \mathcal{C}_m), \end{array}\right. \tag{48}$$

where $\vec{a} = (a_1, \ldots)$ is a vector of atoms and each of the Φ_j only involves the combinatorial constructions of disjoint union and cartesian product. A combinatorial class $\mathcal{C}$ is said to be context-free if it is definable as the first component ($\mathcal{C} = \mathcal{C}_1$) of a well-founded context-free system. The terminology comes from linguistics and it stresses the fact that objects can be "freely" generated by the rules in (48), this without any constraints imposed by an outside context[6].

For instance the class of plane binary trees defined by

$$\mathcal{B} = e + (i \times \mathcal{B} \times \mathcal{B}) \qquad (e, i \text{ atoms})$$

is a context-free class. The class of general plane trees defined by

$$\mathcal{G} = o \times \texttt{sequence}(\mathcal{G}) \qquad (o \text{ an atom})$$

is definable by the system

$$\mathcal{G} = o \times \mathcal{F}, \quad \mathcal{F} = \mathbf{1} + (\mathcal{F} \times \mathcal{G}),$$

with $\mathcal{F}$ defining forests, and so it is also context-free. (This example shows more generally that sequences can always be reduced to polynomial form.)

Context-free specifications may be used to describe all sorts of combinatorial objects. For instance, the class $\mathcal{T}$ of triangulations of convex polygons is specified symbolically by

$$\mathcal{T} = \nabla + (\nabla \times \mathcal{T}) + (\mathcal{T} \times \nabla) + (\mathcal{T} \times \nabla \times \mathcal{T}), \tag{49}$$

where ∇ represents a generic triangle.

The general symbolic rules given in Chapter 1 apply in all such cases. Therefore the Drmota-Lalley-Woods theorem (Theorem VII.7) provides the asymptotic solution to an important category of problems.

THEOREM VII.8 (Context-free specifications). *A context-free class $\mathcal{C}$ admits an OGF that satisfies a polynomial system obtained from the specification by the translation rules:*

$$\mathcal{A} + \mathcal{B} \mapsto A + B, \quad \mathcal{A} \times \mathcal{B} \mapsto A \cdot B.$$

The OGF $C(z)$ is an algebraic function to which algebraic asymptotics applies. In particular, a context-free class $\mathcal{C}$ that gives rise to an algebraically aperiodic irreducible system has an enumeration sequence satisfying

$$C_n \sim \frac{\gamma}{\sqrt{\pi n^3}} \omega^n,$$

where γ, ω are computable algebraic numbers.

This last result explains the frequently encountered estimates involving a factor of $n^{-3/2}$ (corresponding to a square-root singularity of the OGF) that can be found throughout analytic combinatorics.

[6]Formal language theory also defines context-sensitive grammars where each rule (called a production) is applied only if it is enabled by some external context. Context-sensitive grammars have greater expressive power than context-free ones, but they depart significantly from decomposability since they are surrounded by strong undecidability properties; accordingly context-sensitive grammars cannot be associated with any global generating function formalism.

▷ **23.** *Extended context-free specifications.* If $\mathcal{A}, \mathcal{B}$ are context-free specifications then: (i) the sequence class $\mathcal{C} = \texttt{sequence}(\mathcal{A})$ is context-free; (ii) the substitution class $\mathcal{D} = \mathcal{A}[b \mapsto \mathcal{B}]$ is context-free. ◁

We detail below an example from combinatorial geometry.

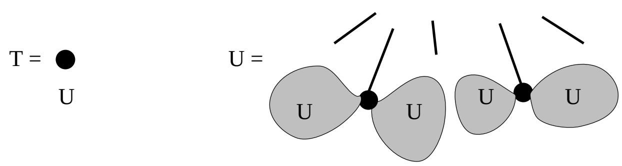
EXAMPLE 12. *Planar non-crossing configurations.* The enumeration of non-crossing planar configurations is discussed here at some level of generality. (An analytic problem in this orbit has been already treated in Example 10.) The purpose is to illustrate the fact that context-free descriptions can model naturally very diverse sorts of objects including particular topological-geometric configurations. The problems considered have their origin in combinatorial musings of the Rev. T.P. Kirkman in 1857 and were revisited in 1974 by Domb and Barett [**48**] for the purpose of investigating certain perturbative expansions of statistical physics. Our presentation follows closely the synthesis offered in [**75**].

Consider for each value of n the regular n-gon built from vertices taken for convenience to be the n complex roots of unity and numbered $0, \dots, n-1$. A non-crossing graph is a graph on this set of vertices such that no two of its edges cross. From there, one defines non-crossing connected graphs, non-crossing forests (that are acyclic), and non-crossing trees (that are acyclic and connected); see Figure 13. Note that there is a well-defined orientation of the complex plane and also that the various graphs considered can always be rooted in some canonical way (e.g., on the vertex of smallest index) since the placement of vertices is rigidly fixed.

Trees. A non-crossing tree is rooted at 0. To the root vertex, is attached an ordered collection of vertices, each of which has an end-node ν that is the common root of two non-crossing trees, one on the left of the edge $(0, \nu)$ the other on the right of $(0, \nu)$. Let $\mathcal{T}$ denote the class of trees and $\mathcal{U}$ denote the class of trees whose root has been severed. With o denoting a generic node, we then have

$$\mathcal{T} = o \times \mathcal{U}, \quad \mathcal{U} = \texttt{sequence}(\mathcal{U} \times o \times \mathcal{U}),$$

which corresponds graphically to the "butterfly decomposition":



In terms of OGF, this gives the system

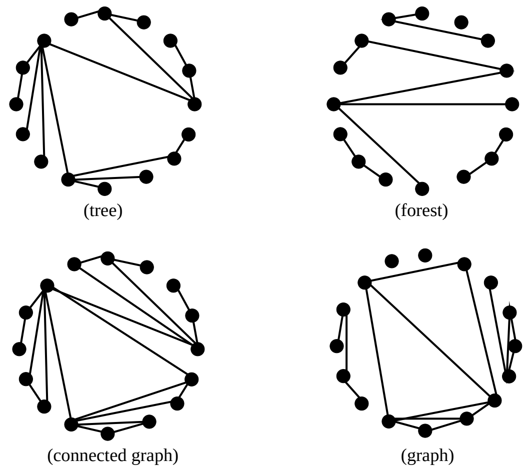
$$\{T = zU,\ U = (1 - zU^2)^{-1}\} \quad \Longleftrightarrow \quad \{T = zU,\ U = 1 + UV,\ V = zU^2\}, \tag{50}$$

where the latter form corresponds to the expansion of the sequence operator. Consequently, T satisfies $T = T^3 - zT + z^2$, which by Lagrange inversion gives $T_n = \frac{1}{2n-1}\binom{3n-3}{n-1}$.

Forests. A (non-crossing) forest is a non-crossing graph that is acyclic. In the present context, it is not possible to express forests simply as sequences as trees, because of the geometry of the problem.

Starting conventionally from the root vertex 0 and following all connected edges defines a "backbone" tree. To the left of every vertex of the tree, a forest may be placed. There results the decomposition (expressed directly in terms of OGF's),

$$F = 1 + T[z \mapsto zF], \tag{51}$$



Configuration / OGF	Coefficients (exact / asymptotic)
Trees (*EIS*: **A001764**)	$z + z^2 + 3z^3 + 12z^4 + 55z^5 + \cdots$
$T^3 - zT + z^2 = 0$	$\frac{1}{2n-1}\binom{3n-3}{n-1}$
	$\sim \frac{\sqrt{3}}{27\sqrt{\pi n^3}}(\frac{27}{4})^n$
Forests (*EIS*: **A054727**)	$1 + z + 2z^2 + 7z^3 + 33z^4 + 181z^5 \cdots$
$F^3 + (z^2 - z - 3)F^2 + (z+3)F - 1 = 0$	$\sum_{j=1}^{n} \frac{1}{2n-j}\binom{n}{j-1}\binom{3n-2j-1}{n-j}$
	$\sim \frac{0.07465}{\sqrt{\pi n^3}}\,(8.22469)^n$
Connected graphs (*EIS*: **A007297**)	$z + z^2 + 4z^3 + 23z^4 + 156z^5 + \cdots$
$C^3 + C^2 - 3zC + 2z^2 = 0$	$\frac{1}{n-1}\sum_{j=n-1}^{2n-3}\binom{3n-3}{n+j}\binom{j-1}{j-n+1}$
	$\sim \frac{2\sqrt{6}-3\sqrt{2}}{18\sqrt{\pi n^3}}\left(6\sqrt{3}\right)^n$
Graphs (*EIS*: **A054726**)	$1 + z + 2z^2 + 8z^3 + 48z^4 + 352z^5 + \cdots$
$G^2 + (2z^2 - 3z - 2)G + 3z + 1 = 0$	$\frac{1}{n}\sum_{j=0}^{n-1}(-1)^j\binom{n}{j}\binom{2n-2-j}{n-1-j}2^{n-1-j}$
	$\sim \frac{\sqrt{140-99\sqrt{2}}}{4\sqrt{\pi n^3}}\left(6+4\sqrt{2}\right)^n$

FIGURE 13. (Top) Non-crossing graphs: a tree, a forest, a connected graph, and a general graph. (Bottom) The enumeration of non-crossing configurations by algebraic functions.

where T is the OGF of trees and F is the OGF of forests. In (51), the term $T[z \mapsto zF]$ denotes a functional composition. A context-free specification in standard form results mechanically from (50) upon replacing z by zF, namely

$$F = 1 + T,\ T = zFU,\ U = 1 + UV,\ V = zFU^2. \tag{52}$$

This system is irreducible and aperiodic, so that the asymptotic shape of F_n is of the form $\gamma\omega_n n^{-3/2}$, as predicted by Theorem VII.8. This agrees with the precise formula determined in Example 10.

Graphs. Similar constructions (see [**75**]) give the OGF's of connected graphs and general graphs. The results are summarized in Figure 13. Note the common shape of the asymptotic estimates and also the fact that simple binomial terms or sums are available in each case. □

Note on "tree-like" structures. A context-free specification can always be regarded as defining a class of trees. Indeed, if the jth term in the construction Φ_j is "coloured" with the pair (i, j), it is seen that a context-free system yields a class of trees whose nodes are tagged by pairs (i, j) in a way that is consistent with the system's rules (48). However, despite this correspondence, it is often convenient to preserve the possibility of operating directly with objects[7] when the tree aspect is unnatural. By a terminology borrowed from the theory of syntax analysis in computer science, such trees are referred to as "parse trees" or "syntax trees".

Let $\mathcal{A}$ be a fixed finite alphabet whose elements are called letters. A *grammar* G is a collection of equations

$$G \quad : \quad \begin{cases} \mathcal{L}_1 & = & \Psi_1(\vec{a}, \mathcal{L}_1, \ldots, \mathcal{L}_m) \\ \vdots & & \vdots \\ \mathcal{L}_m & = & \Psi_m(\vec{a}, \mathcal{L}_1, \ldots, \mathcal{L}_m), \end{cases} \tag{53}$$

where each Ψ_j involves only the operations of union ($\cup$) and catenation product ($\cdot$) with $\vec{a}$ the vector of letters in $\mathcal{A}$. For instance,

$$\Psi_1(\vec{a}, \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3) = a_2 \cdot \mathcal{L}_2 \cdot \mathcal{L}_3 \cup a_3 \cup \mathcal{L}_3 \cdot a_2 \cdot \mathcal{L}_1.$$

A solution to (53) is an m-tuple of languages over the alphabet $\mathcal{A}$ that satisfies the system. By convention, one declares that the grammar G defines the first component, $\mathcal{L}_1$.

To each grammar (53), one can associate a context-free specification (48) by transforming unions into disjoint union, '$\cup$' $\mapsto$ '$+$', and catenation into cartesian products, '$\cdot$' $\mapsto$ '$\times$'. Let $\widehat{G}$ be the specification associated in this way to the grammar G. The objects described by $\widehat{G}$ appear in this perspective to be trees (see the discussion above regarding parse trees). Let h be the transformation from trees of $\widehat{G}$ to languages of G that lists letters in infix (i.e., left-to-right) order: we call such an h the erasing transformation since it "forgets" all the structural information contained in the parse tree and only preserves the succession of letters. Clearly, application of h to the combinatorial specifications determined by $\widehat{G}$ yields languages that obey the grammar G. For a grammar G and a word $w \in \mathcal{A}^\star$, the number of parse trees $t \in \widehat{G}$ such that $h(t) = w$ is called the *ambiguity coefficient* of w with respect to the grammar G; this quantity is denoted by $\kappa_G(w)$.

[7] Some authors have even developed a notion of "object grammars"; see for instance [**53**] itself inspired by techniques of polyomino surgery in [**44**].

A grammar G is unambiguous if all the corresponding ambiguity coefficients are either 0 or 1. This means that there is a bijection between parse trees of $\widehat{G}$ and words of the language described by G: each word generated is uniquely "parsable" according to the grammar. From Theorem VII.8, we have immediately:

PROPOSITION VII.4 (Contex-free languages). *Given a context-free grammar G, the ordinary generating function of the language $L_G(z)$, counting words* with *multiplicity, is an algebraic function. In particular, a context-free language that admits an unambiguous grammar specification has an ordinary generating function $L(z)$ that is an algebraic function.*

This theorem originates from early works of Chomsky and Sch¨utzenberger [**38**] which have exerted a strong influence on the philosophy of the present book.

For example consider the Łukasiewicz language

$$\mathcal{L} = (a \cdot \mathcal{L} \cdot \mathcal{L} \cdot \mathcal{L}) \cup b.$$

This can be interpreted as the set of functional terms built from the ternary symbol a and the nullary symbol b:

$$\begin{aligned} \mathcal{L} &= \{b,\ abbb,\ aabbbbb,\ ababbbb, \ldots\} \\ &\simeq \{b,\ a(b,b,b),\ a(a(b,b,b),b,b),\ a(b,a(b,b,b),b), \ldots\}, \end{aligned}$$

where $\simeq$ denotes combinatorial isomorphism. It is easily seen that the terms are in bijective correspondence with their parse trees, themselves isomorphic to ternary trees. Thus the grammar is unambiguous, so that the OGF equation translates directly from the grammar,

$$L(z) = zL(z)^3 + z. \tag{54}$$

As another example, we revisit Dyck paths that are definable by the grammar,

$$D = 1 \cup (a \cdot D \cdot b \cdot D)\,, \tag{55}$$

where a denotes ascents and b denotes descents. Each word in the language must start with a letter a that has a unique matching letter b and thus it is uniquely parsable according to the grammar (55). Since the grammar is unambiguous, the OGF reads off:

$$D(z) = z + z^2 D(z)^2.$$

VII. 5.2. Walks and the kernel method. Start with a set Ω that is a finite subset of $\mathbb{Z}$ and is called the set of *jumps*. A *walk* (relative to Ω) is a sequence $w = (w_0, w_1, \ldots, w_n)$ such that $w_0 = 0$ and $w_{i+1} - w_i \in \Omega$, for all i, $0 \le i < n$. A *nonnegative walk* satisfies $w_i \ge 0$ and an *excursion* is a nonnegative walk such that, additionally, $w_n = 0$. The quantity n is called the length of the walk or the excursion. For instance, Dyck paths and Motzkin paths analysed in Section V. 2 are excursions that correspond to $\Omega = \{-1, +1\}$ and $\Omega = \{-1, 0, +1\}$ respectively. (Walks and excursions can be viewed as particular cases of paths in a graph in the sense of Section V. 5, with the graph taken to be the infinite set $\mathbb{Z}_{>0}$ of integers.)

We propose to determine f_n, the number of excursions of length n and type Ω, via the corresponding OGF

$$F(z) = \sum_{n=0}^{\infty} f_n z^n.$$

In fact, we shall determine the more general BGF

$$F(z,u) := \sum_{n,k} f_{n,k} u^k z^n,$$

where $f_{n,k}$ is the number of walks of length n and final altitude k (*i.e.*, the value of w_n in the definition of a walk is constrained to equal k). In particular, one has $F(z) = F(z, 0)$.

We let $-c$ denote the smallest (negative) value of a jump, and d denote the largest (positive) jump. A fundamental rôle is played in this discussion by the "characteristic polynomial" of the walk,

$$S(y) := \sum_{\omega \in \Omega} y^{\omega} = \sum_{j=-c}^{d} S_j y^j$$

that is a Laurent polynomial[8]. Observe that the bivariate generating function of generalized walks where intermediate values are allowed to be negative, with z marking the length and u marking the final altitude, is rational:

$$G(z, u) = \frac{1}{1 - zS(u)}. \tag{56}$$

Returning to nonnegative walks, the main result to be proved below is the following: *For each finite set $\Omega \in \mathbb{Z}$, the generating function of excursions is an algebraic function that is explicitly computable from Ω.* There are many ways to view this result. The problem is usually treated within probability theory by means of Wiener-Hopf factorizations [**163**]. In contrast, Labelle and Yeh [**126**] show that an unambiguous context-free specification can be systematically constructed, a fact that is sufficient to ensure the algebraicity of the GF $F(z)$. (Their approach is based implicitly on the construction of a finite pushdown automaton itself equivalent, by general principles, to a context-free grammar.) The Labelle-Yeh construction reduces the problem to a large, but somewhat "blind", combinatorial preprocessing, and, for analysts it has the disadvantage of not extracting a simpler (and noncombinatorial) structure inherent in the problem. The method described below is often known as the "kernel" method. It takes its inspiration from exercises in the 1968 edition of Knuth's book [**118**] (Ex. 2.2.1.4 and 2.2.1.11) where a new approach was proposed to the enumeration of Catalan and Schroeder objects. The technique has since been extended and systematized by several authors; see for instance [**10, 11, 28, 56, 57**].

Let $f_n(u) = [z^n]F(z, u)$ be the generating function of walks of length n with u recording the final altitude. There is a simple recurrence relating $f_{n+1}(u)$ to $f_n(u)$, namely,

$$f_{n+1}(u) = S(u) \cdot f_n(u) - r_n(u), \tag{57}$$

where $r_n(u)$ is a Laurent polynomial consisting of the sum of all the monomials of $S(u)f_n(u)$ that involve negative powers[9] of u:

$$r_n(u) := \sum_{j=-c}^{-1} u^j \left([u^j]\, S(u) f_n(u)\right) = \{u^{<0}\} S(u) f_n(u). \tag{58}$$

The idea behind the formula is to subtract the effect of those steps that would take the walk below the horizontal axis. For instance, one has

$$\begin{array}{lcl} S(u) = \dfrac{S_{-1}}{u} + O(1) & : & r_n(u) = \dfrac{S_{-1}}{u} f_n(0) \\ S(u) = \dfrac{S_{-2}}{u^2} + \dfrac{S_{-1}}{u} + O(1) & : & r_n(u) = \left(\dfrac{S_{-2}}{u^2} + \dfrac{S_{-1}}{u}\right) f_n(0) + \dfrac{S_{-2}}{u} f_n'(0) \end{array}$$

[8] If Ω is a set, then the coefficients of S lie in $\{0, 1\}$. The treatment above applies in all generality to cases where the coefficients are arbitrary positive real numbers. This accounts for probabilistic situations as well as multisets of jump values.

[9] The convenient notation $\{u^{<0}\}$ denotes the singular part of a Laurent expansion: $\{u^{<0}\}f(z) := \sum_{j<0} \left([u^j]f(u)\right) \cdot u^j$.

and generally:

$$\lambda_j(u) = \frac{1}{j!}\{u^{<0}\}u^j S(u). \tag{59}$$

Thus, from (57) and (58) (multiply by z^{n+1} and sum), the generating function $F(z,u)$ satisfies the fundamental functional equation

$$F(z,u) = 1 + zS(u)F(z,u) - z\{u^{<0}\}\left(S(u)F(z,u)\right). \tag{60}$$

Explicitly, one has

$$F(z,u) = 1 + zS(u)F(z,u) - z\sum_{j=0}^{c-1} \lambda_j(u)\left[\frac{\partial^j}{\partial u^j}F(z,u)\right]_{u=0}, \tag{61}$$

for Laurent polynomials $\lambda_j(u)$ that depend on $S(u)$ in an effective way by (59).

The main equations (60) and (61) involve one unknown bivariate GF, $F(z,u)$ and c univariate GF's, the partial derivatives of F specialized at $u=0$. It is true, but not at all obvious, that the single functional equation (61) fully determines the $c+1$ unknowns. The basic technique is known as "cancelling the kernel" and it relies on strong analyticity properties; see the book by Fayolle *et al.* [**57**] for deep ramifications. The form of (61) to be employed for this purpose starts by grouping on one side the terms involving $F(z,u)$,

$$F(z,u)(1-zS(u)) = 1 - z\sum_{j=0}^{c-1}\lambda_j(u)G_j(z), \qquad G_j(z) := \left[\frac{\partial^j}{\partial u^j}F(z,u)\right]. \tag{62}$$

If the right side was not present, then the solution would reduce to (56). In the case at hand, from the combinatorial origin of the problem and implied bounds, the quantity $F(z,u)$ is bivariate analytic at $(z,u)=(0,0)$ (by elementary exponential majorizations on the coefficients). The main principle of the kernel method consists in *coupling* the values of z and u in such a way that $1-zS(u)=0$, so that $F(z,u)$ disappears from the picture. A condition is that both z and u should remain small (so that F remains analytic). Relations between the partial derivatives are then obtained from such a specializations, $(z,u)\mapsto(z,u(z))$, which happen to be just in the right quantity.

Consequently, we consider the "kernel equation",

$$1 - zS(u) = 0, \tag{63}$$

which is rewritten as

$$u^c = z\cdot(u^c S(u)).$$

Under this form, it is clear that the kernel equation (63) defines $c+d$ branches of an algebraic function. A local analysis (Newton's polygon method) shows that, amongst these $c+d$ branches, there are c branches that tend to 0 as $z\to 0$ while the other d tend to infinity as $z\to 0$. Let $u_0(z),\ldots,u_{c-1}(z)$ be the c branches that tend to 0, that we call "small" branches. In addition, we single out $u_0(z)$, the "principal" solution, by the reality condition

$$u_0(z)\sim \gamma z^{1/c}, \quad \gamma := (S_c)^{1/c}\in\mathbb{R}_{>0} \qquad (z\to 0^+).$$

By local uniformization (40), the conjugate branches are given locally by

$$u_\ell(z) = u_0(\mathrm{e}^{2i\ell\pi}z) \qquad (z\to 0^+).$$

Coupling z and u by $u=u_\ell(z)$ produces interesting specializations of Equation (62). In that case, (z,u) is close to $(0,0)$ where F is bivariate analytic so that the substitution is admissible. By substitution, we get

$$1 - z\sum_{j=0}^{c-1}\lambda_j(u_\ell(z))\left[\frac{\partial^j}{\partial u^j}F(z,u)\right]_{u=0}, \qquad \ell = 0\,..\,c-1. \tag{64}$$

This is now a linear system of c equation in c unknowns (the partial derivatives) with algebraic coefficients that, in principle, determines $F(z,0)$.

A convenient approach to the solution of (64) is due to Mireille Bousquet-Mélou. The argument goes as follows. The quantity

$$M(u) := u^c - zu^c\sum_{j=0}^{c-1}\lambda_j(u)\frac{\partial^j}{\partial u^j}F(z,0) \tag{65}$$

can be regarded as a polynomial in u. It is monic while it vanishes by construction at the c small branches $u_0,\dots,u_{c-1}$. Consequently, one has the factorization,

$$M(u) = \prod_{\ell=0}^{c-1}(u - u_\ell(z)). \tag{66}$$

Now, the constant term of $M(u)$ is otherwise known to equal $-zS_{-c}F(z,0)$, by the definition (65) of $M(u)$ and by Equation (59) specialized to $\lambda_0(u)$. Thus, the comparison of constant terms between (65) and (66) provides us with an explicit form of the OGF of excursions:

$$F(z,0) = \frac{(-1)^{c-1}}{S_{-c}z}\prod_{\ell=0}^{c-1}u_\ell(z).$$

One can then finally return to the original functional equation and pull the BGF $F(z,u)$. We can thus state:

PROPOSITION VII.5 (Kernel method for walks). *Let Ω be a finite step of jumps and let $S(u)$ be the characteristic polynomial of Ω. In terms of the c small branches of the "kernel" equation,*

$$1 - zS(u) = 0,$$

denoted by $u_0(z),\dots,u_{c-1}(z)$, the generating function of excursions is expressible as

$$F(z) = \frac{(-1)^{c-1}}{zS_{-c}}\prod_{\ell=0}^{c-1}u_\ell(z) \quad \textit{where } S_{-c} = [u^{-c}]S(u)$$

is the multiplicity (or weight) of the smallest element $-c\in\Omega$.

More generally the bivariate generating function of nonnegative walks is bivariate algebraic and given by

$$F(z,u) = \frac{1}{u^c - z\,(u^cS(u))}\prod_{\ell=0}^{c-1}(u - u_\ell(z))\,.$$

We give next a few examples illustrating this kernel technique.

Trees and Łukasiewicz codes. A particular class of walks is of special interest; it corresponds to cases where $c = 1$, that is, the largest jump in the negative direction has amplitude 1. Consequently, $\Omega + 1 = \{0, s_1, s_2, \dots, s_d\}$. In that situation, combinatorial theory teaches us the existence of fundamental isomorphisms between walks defined by steps Ω and trees whose degrees are constrained to lie in $1+\Omega$. The correspondence is by way of

Łukasiewicz codes[10] (‘also known as ‘Polish” prefix codes, “Polish” prefix notation), and from it we expect to find tree GF’s in such cases.

As regards generating functions, there now exists only *one* small branch, namely the solution $u_0(z)$ to $u_0(z) = z\phi(u_0(z))$ (where $\phi(u) = uS(u)$) that is analytic at the origin. One then has $F(z) = F(z,0) = \frac{1}{z}u_0(z)$, so that the walk GF is determined by

$$F(z,0) = \frac{1}{z}u_0(z), \quad u_0(z) = z\phi(u_0(z)), \quad \phi(u) := uS(u).$$

This form is consistent with what is already known regarding the enumeration of simple families of trees. In addition, one finds

$$F(z,u) = \frac{1 - u^{-1}u_0(z)}{1 - zS(u)} = \frac{u - u_0(z)}{u - z\phi(u)}.$$

Classical specializations are rederived in this way:

— the Catalan walk (Dyck path), defined by $\Omega = \{-1, +1\}$ and $\phi(u) = 1 + u^2$, has
$$u_0(z) = \frac{1}{2z}\left(1 - \sqrt{1 - 4z^2}\right);$$

— the Motzkin walk, defined by $\Omega = \{-1, 0, +1\}$ and $\phi(u) = 1 + u + u^2$ has
$$u_0(z) = \frac{1}{2z}\left(1 - z - \sqrt{1 - 2z - 3z^2}\right);$$

— the modified Catalan walk, defined by $\Omega = \{-1, 0, 0+1\}$ (with two steps of type 0) and $\phi(u) = 1 + 2u + u^2$, has
$$u_0(z) = \frac{1}{2z}\left(1 - 2z - \sqrt{1 - 4z}\right);$$

— the d-ary tree walk (the excursions encode d-ary trees) defined by $\Omega = \{-1, d - 1\}$, has $u_0(z)$ that is defined implicitly by
$$u_0(z) = z(1 + u_0(z)^d).$$

Examples of the general case. Take now $\Omega = \{-2, -1, 1, 2\}$ so that

$$S(u) = u^{-2} + u^{-1} + u + u^2.$$

Then, $u_0(z), u_1(z)$ are the two branches that vanish as $z \to 0$ of the curve

$$y^2 = z(1 + y + y^3 + y^4).$$

The linear system that determines $F(z,0)$ and $F'(z,0)$ is

$$\begin{cases} 1 - \left(\dfrac{z}{u_0(z)^2} + \dfrac{z}{u_0(z)}\right) F(z,0) - \dfrac{z}{u_0(z)} F'(z,0) & = & 0 \\ 1 - \left(\dfrac{z}{u_1(z)^2} + \dfrac{z}{u_1(z)}\right) F(z,0) - \dfrac{z}{u_1(z)} F'(z,0) & = & 0 \end{cases}$$

(derivatives are taken with respect to the second argument) and one finds

$$F(z,0) = -\frac{1}{z}u_0(z)u_1(z), \qquad F'(z,0) = \frac{1}{z}(u_0(z) + u_1(z) + u_0(z)u_1(z)).$$

This gives the number of walks, through a combination of series expansions,

$$F(z) = 1 + 2z^2 + 2z^3 + 11z^4 + 24z^5 + 93z^6 + 272z^7 + 971z^8 + 3194z^9 + \cdots.$$

---

[10] Such a code [**132**] is obtained by a preorder traversal of the tree, recording a jump of $r - 1$ when a node of outdegree r is encountered. The sequence of jumps gives rise to an excursion followed by an extra -1 jump.

A single algebraic equation for $F(z) = F(z, 0)$ is then obtained by elimination (*e.g.*, via Groebner bases) from the system:

$$\begin{cases} u_0^2 - z(1 + u_0 + u_0^3 + u_0^4) &= 0 \\ u_1^2 - z(1 + u_1 + u_1^3 + u_1^4) &= 0 \\ zF + u_0 u_1 &= 0 \end{cases}$$

Elimination shows that $F(z)$ is a root of the equation

$$z^4y^4 - z^2(1 + 2z)y^3 + z(2 + 3z)y^2 - (1 + 2z)y + 1 = 0.$$

For walks corresponding to $\Omega = \{-2, -1, 0, 1, 2\}$, we find similarly $F(z) = -\frac{1}{z}u_0(z)u_1(z)$, where u_0, u_1 are the small branches of $y^2 = z(1 + y + y^2 + y^3 + y^4)$, the expansion starts as

$$F(z) = 1 + z + 3z^2 + 9z^3 + 32z^4 + 120z^5 + 473z^6 + 1925z^7 + 8034z^8 + \cdots,$$

and $F(z)$ is a root of the equation

$$z^4y^4 - z^2(1 + z)y^3 + z(2 + z)y^2 - (1 + z)y + 1 = 0.$$

VII. 5.3. Maps and the quadratic method. A (planar) map is a connected planar graph together with an embedding into the plane. In all, generality, loops and multiple edges are allowed. A planar map therefore separates the plane into regions called faces (Figure 14). The maps considered here are in addition rooted, meaning that a face, an incident edge, and an incident vertex are distinguished. In this section, only rooted maps are considered[11]. When representing rooted maps, we shall agree to draw the root edge with an arrow pointing away from the root node, and to take the root face as that face lying to the left of the directed edge (represented in grey on Figure 14).

Tutte launched in the 1960's a large census of planar maps, with the intention of attacking the four-colour problem by enumerative techniques[12]; see [**30, 182, 183, 184, 185**]. There exists in fact an entire zoo of maps defined by various degree or connectivity constraints. In this chapter, we shall limit ourselves to conveying a flavour of this vast theory, with the goal of showing how algebraic functions arise. The presentation takes its inspiration from the book of Goulden and Jackson [**96**, Sec. 2.9]

Let $\mathcal{M}$ be the class of all maps where size is taken to be the number of edges. Let $M(z, u)$ be the BGF of maps with u marking the number of edges on the outside face. The basic surgery performed on maps distinguishes two cases based upon the nature of the root edge. A rooted map will be declared to be isthmic if the root edge r of map μ is an "isthmus" whose deletion would disconnect the graph. Clearly, one has,

$$\mathcal{M} = o + \mathcal{M}^{(i)} + \mathcal{M}^{(n)}, \tag{67}$$

where $\mathcal{M}^{(i)}$ (resp. $\mathcal{M}^{(n)}$) represent the class of isthmic (resp. non-isthmic) maps and 'o' is the graph consisting of a single vertex and no edge. There are accordingly two ways to build maps from smaller ones by adding a new edge.

---

[11]Nothing is lost regarding asymptotic properties of random structures when a rooting is imposed. The reason is that a map has, with probability exponentially close to 1, a trivial automorphism group; consequently, almost all maps of m edges can be rooted in $2m$ ways (by choosing an edge, and an orientation of this edge), and there is an almost uniform $2m$-to-1 correspondence between unrooted maps and rooted ones.

[12]The four-colour theorem to the effect that every planar graph can be coloured using only four colours was eventually proved by Appel and Haken in 1976, using structural graph theory methods supplemented by extensive computer search.

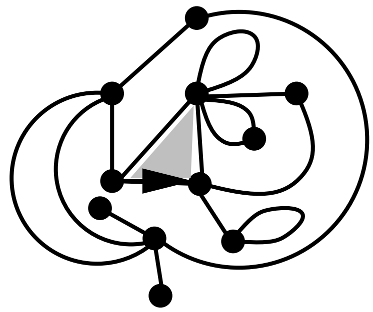
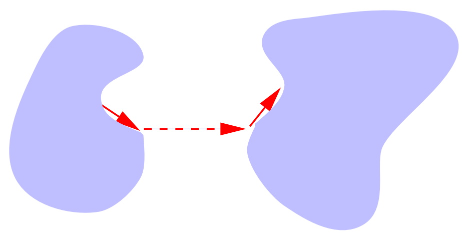


FIGURE 14. A planar map.

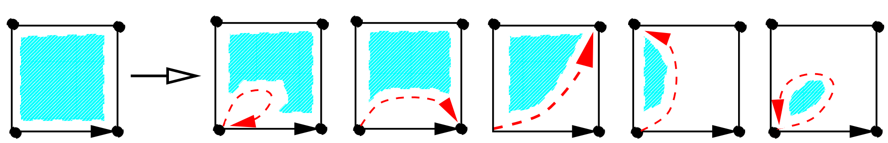
(i) The class of all isthmic maps is constructed by taking two arbitrary maps and joining them together by a new root edge, as shown below:



The effect is to increase the number of edges by 1 (the new root edge) and have the root face degree become 2 (the two sides of the new root edge) plus the sum of the root face degrees of the component maps. The construction is clearly revertible. In other words, the BGF of $\mathcal{M}^{(i)}$ is

$$M^{(i)}(z,u) = zu^2 M(z,u)^2. \tag{68}$$

(ii) The class of non-isthmic maps is obtained by taking an already existing map and adding an edge that preserves its root node and "cuts across" its root face in some unambiguous fashion (so that the construction should be revertible). This operation will therefore result in a new map with an essentially smaller root-face degree. For instance, there are 5 ways to cut across a root face of degree 4, namely,



This corresponds to the linear transformation

$$u^4 \quad \mapsto \quad zu^5 \quad + \quad zu^4 \quad + \quad zu^3 \quad + \quad zu^2 \quad + \quad zu^1.$$

In general the effect on a map with root face of degree k is described by the transformation $u^k \mapsto z(1-u^{k+1})/(1-u)$; equivalently, each monomial $g(u) = u^k$ is transformed into $u(g(1) - ug(u))/(1-u)$. Thus, the OGF of $\mathcal{M}^{(n)}$ involves a discrete difference operator:

$$M^{(n)}(z,u) = zu\frac{M(z,1) - uM(z,u)}{1-u}. \tag{69}$$

Collecting the contributions from (68) and (69) in (67) then yields the basic functional equation,

$$M(z,u) = 1 + u^2 z M(z,u)^2 + uz\frac{M(z,1) - uM(z,u)}{1-z}. \tag{70}$$

The functional equation (70) binds two unknown functions, $M(z,u)$ and $M(z,1)$. Much like in the case of walks, it would seem to be underdetermined. Now, a method due to Tutte and known as the quadratic method provides solutions. Following Tutte and the account in [**96**, p. 138], we consider momentarily the more general equation

$$\left(g_1 F(z,u) + g_2\right)^2 = g_3, \tag{71}$$

where $g_j = G_j(z,u,h(z))$ and the G_j are explicit functions—here the unknown functions are $F(z,u)$ and $h(z)$ (cf. $M(z,u)$ and $M(z,1)$ in (70)). Bind u and z in such a way that the left side of (71) vanishes, that is, substitute $u = u(z)$ (a yet unknown function) so that $g_1 F + g_2 = 0$. Since the left-hand side of (71) now has a double root in u, so must the right-hand side, which implies

$$g_3 = 0, \qquad \left.\frac{\partial g_3}{\partial u}\right|_{u=u(z)}. \tag{72}$$

The original equation has become a system of two equations in two unknowns that determines implicitly $h(z)$ and $u(z)$. From there, elimination provides individual equations for $u(z)$ and for $h(z)$. (If needed, $F(z,u)$ can then be recovered by solving a quadratic equation.) It will be recognized that, if the quantities q_1, g_2, g_3 are polynomials, then the process invariably yields solutions that are algebraic functions.

We now carry out this programme in the case of maps and Equation (70). First, isolate $M(z,u)$ by completing the square, giving

$$\left(M(z,u) - \frac{1}{2}\frac{1-u+u^2z}{u^2z(1-u)}\right)^2 = Q(z,u) + \frac{M(z,1)}{u(1-u)}, \tag{73}$$

where

$$Q(z,u) = \frac{z^2u^4 - 2zu^2(u-1)(2u-1) + (1-u^2)}{4u^4z^2(1-u)^2}.$$

Next, the condition expressing the existence of a double root is

$$Q(z,u) + \frac{1}{u(1-u)}M(z,1) = 0, \quad Q'_u(z,u) + \frac{2u-1}{u^2(1-u)^2}M(z,1) = 0.$$

It is now easy to eliminate $M(z,1)$, since the dependency in M is linear, and a straightforward calculation shows that $u = u(z)$ should satisfy

$$\big(u^2z + (u-1)\big)\big(u^2z + (u-1)(2u-3)\big) = 0.$$

The first parameterization would lead to $M(z,1) = 1/z$ which is not admissible. Thus, $u(z)$ is to be taken as the root of the second factor, with $M(z,1)$ being defined parametrically by

$$z = \frac{(1-u)(2u-3)}{u^2}, \quad M(z,1) = -u\frac{3u-4}{(2u-3)^2}.$$

The change of parameter $u = 1 - 1/w$ reduces this further to the "Lagrangean form",

$$z = \frac{w}{1-3w}, \quad M(z,1) = \frac{1-4w}{(1-3w)^2}. \tag{74}$$

To this the Lagrange inversion theorem can be applied. The number of maps with n edges, $M_n = [z^n]M(z,1)$ is then determined as

$$M_n = 2\frac{(2n)!3^n}{n!(n+2)!},$$

and one obtains Sequence **A000168** of the *EIS*:

$$M(z,1) = 1 + 2z + 9z^2 + 54z^3 + 378z^4 + 2916z^5 + 24057z^6 + 208494z^7 + \cdots .$$

We refer to [**96**, Sec. 2.9] for detailed calculations (that are nowadays routinely performed with assistance of a computer algebra system). Currently, there exist many applications of the method to maps satisfying all sorts of combinatorial constraints (*e.g.*, multiconnectivity); see [**166**] for a recent panorama.

The derivation above has purposely stressed a parameterized approach as this constitutes a widely applicable approach in many situations. In a simple case like this, we may also eliminate u and solve explicitly for $M(z,1)$, to wit,

$$M(z) \equiv M(z,1) = -\frac{1}{54\,z^2}\left(1 - 18z - (1-12z)^{3/2}\right).$$

It is interesting to note that the singular exponent here is $\frac{3}{2}$, a fact further reflected by the somewhat atypical factor of $n^{-5/2}$ in the asymptotic form of coefficients:

$$M_n \sim \frac{2}{\sqrt{\pi n^5}} 12^n \qquad (n \to \infty).$$

Accordingly, randomness properties of maps are appreciably different from what is observed in trees and many commonly encountered context-free objects.

VII. 6. Notes

The exp-log schema, like its companion, the supercritical-sequence schema, illustrates the level of generality that can be attained by singularity analysis techniques. Refinements of the results we have given can be found in the book by Arratia, Barbour, and Tavaré [**9**], which develops a stochastic process approach to these questions; see also [**8**] by the same authors for an accessible introduction.

The rest of the chapter deals in an essential manner with recursively defined structures. As noted repeatedly in the course of this chapter, this is very often conducive to square-root singularity and universal behaviours that are quantified by exponents of the form $\frac{1}{2}, \frac{3}{2}, \ldots$. Simple varieties of trees have been introduced in an important paper of Meir and Moon [**138**], that bases itself on methods developed earlier by Pólya [**158, 159**] and Otter [**151**]. One of the merits of [**138**] is to demonstrate that a high level of generality is attainable when discussing properties of trees. A similar treatment can be inflicted more generally to recursively defined structures when their generating function satisfies an implicit equation. In this way, nonplane unlabelled trees are shown to exhibit properties very similar to their plane counterparts. It is of interest to note that such of the enumerative questions in this area had been initially motivated by problems of theoretical chemistry: see the colourful account of Cayley and Sylvester's works in [**26**], the reference books by Harary–Palmer [**104**] and Finch [**61**], as well as PÓlya's original studies [**158, 159**].

Algebraic functions are the modern counterpart of the study of curves by classical Greek mathematicians. They are either approached by algebraic methods (this is the core of algebraic geometry) or by transcendental methods. For our purposes, however, only rudiments of the theory of curves are needed. For this, there exist several excellent introductory books, of which we recommend the ones by Abhyankar [**1**], Fulton [**88**], and

Kirwan [**114**]. On the algebraic side, we have striven to provide an introduction to algebraic functions that requires minimal apparatus. At the same time the emphasis has been put somewhat on algorithmic aspects, since most algebraic models are nowadays likely to be treated with the help of computer algebra. As regards symbolic computational aspects, we recommend the treatise by von zur Gathen and J¨urgen [**189**] for background, while polynomial systems are excellently reviewed in the book by Cox, Little, and O'Shea [**40**].

In the combinatorial domain, algebraic functions have been used early: in Euler and Segner's enumeration of triangulations (1753) as well as in Schr¨oder's famous "*vier combinatorische Probleme*" described in [**176**, p. 177]. A major advance was the realization by Chomsky and Sch¨utzenberger that algebraic functions are the "exact" counterpart of context-free grammars and languages (see again the historic paper [**38**]). A masterful summary of the early theory appears in the proceedings edited by Berstel [**21**] while a modern and precise exposition forms the subject of Chapter 6 of Stanley's book [**176**]. On the analytic-asymptotic side, many researchers have long been aware of the power of Puiseux expansions in conjunction with some version of singularity analysis (often in the form of the Darboux–Pólya method: see [**159**] based on Pólya's classic paper [**158**] of 1937). However, there appeared to be difficulties in coping with the fully general problem of algebraic coefficient asymptotics [**31, 142**]. We believe that Section VII. 4.1 sketches the first complete theory. In the case of positive systems, the "Drmota-Lalley-Woods" theorem is the key to most problems encountered in practice—its importance should be clear from the developments of Section VII. 4.2.

The applications of algebraic functions to context-free languages have been known for some time (e.g., [**65**]). Our presentation of 1-dimensional walks of a general type follows a recent article by Banderier and Flajolet [**12**], whch can be regarded as the analytic pendant of algebraic studies by Gessel [**92, 91**]. The kernel method to which we apply has its origins in problems of queueing theory and random walks [**56, 57**] and is further explored in an article by Bousquiet-Mélou and Petkovšek [**28**]. The algebraic treatment of random maps by the quadratic method is due to brilliant studies of Tutte in the 1960's: see for instance his census [**182**] and the account in the book by Jackson and Goulden [**96**]. A combinatorial-analytic treatment of multiconnectivity issues is given in [**13**], where the possibility of treating in a unified manner about a dozen families of maps appears clearly.

APPENDIX B

Basic Complex Analysis

This appendix contains entries on the following topics:

Algebraic elimination; Equivalent definitions of analyticity; Gamma function; Implicit Function Theorem; Mellin transform; Perron-Frobenius theory of nonnegative matrices; Regular expressions.

1. Algebraic elimination. Auxiliary quantities can be eliminated from systems of polynomial equations. In essence, elimination is achieved by suitable combinations of the equations themselves. One of the best strategies is based on Gröbner bases and is presented in the excellent book [**40**] of Cox, Little, and O'Shea. This entry develops a more elementary approach based on *resultants*.

Consider a field of coefficients $\mathbb{K}$ which may be specialized as $\mathbb{Q}, \mathbb{C}, \mathbb{C}(z), \ldots$, as the need arises. A polynomial of degree d in $\mathbb{K}[x]$ has at most d roots in $\mathbb{K}$ and exactly d roots in the algebraic closure $\bar{\mathbb{K}}$ of $\mathbb{K}$. Given two polynomials,

$$P(x) = \sum_{j=0}^{\ell} a_j x^{\ell-j}, \qquad Q(x) = \sum_{k=0}^{m} b_{m-k} x^k,$$

their *resultant* (with respect to the variable x) is the determinant of order $(\ell+m)$,

$$\mathbf{R}(P,Q,x) = \det \begin{vmatrix} a_0 & a_1 & a_2 & \cdots & 0 & 0 \\ 0 & a_0 & a_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{\ell-1} & a_\ell \\ b_0 & b_1 & b_2 & \cdots & 0 & 0 \\ 0 & b_0 & b_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{m-1} & b_m \end{vmatrix}, \tag{1}$$

also called the Sylvester determinant. By its definition, the resultant is a polynomial form in the coefficients of P and Q. The main property of resultants is the following: (i) *If $P(x), Q(x) \in \mathbb{K}[x]$ have a common root in the algebraic closure $\bar{\mathbb{K}}$ of $\mathbb{K}$, then $\mathbf{R}(P(x), Q(x), x) = 0$.* (ii) *Conversely, if $\mathbf{R}(P(x), Q(x), x) = 0$ holds, then either $a_0 = 0$ or $b_0 = 0$ or else $P(x), Q(x)$ have a common root in $\bar{\mathbb{K}}$.* (The idea is as follows. Let S be the matrix in (1). Then the linear system $Sw = 0$ admits a solution $w = (\xi^{\ell+m-1}, \ldots, \xi^2, \xi, 1)$, if ξ is a commont root of P and Q. We refer to [**129**, V§10] for a presentation of resultants.)

$\triangleright$ **1.** *Root separation and discriminant.* Let $f(x)$ be a polynomial of degree ℓ with leading coefficient a_0 and m distinct roots $\alpha_1, \ldots, \alpha_\ell$. Then the *discriminant* defined as $\mathbf{R}(f(x), f'(x), x)$ satisfies

$$\mathbf{R}(f(x), f'(x), x) = a_0^\ell \prod_{i \neq j} (\alpha_i - \alpha_j).$$

There results a computable lower bound on the distance between any two roots of f (assumed to have distinct roots) as $a_0^{2/(\ell-1)}\,\mathbf{R}(f,f')^{1/\binom{\ell}{2}}$. ◁

The resultant thus provides a *sufficient* condition for the existence of common roots, but not always a necessary one. This has implications in situations where the coefficients a_j, b_k depend on one or several parameters. In that case, the condition $\mathbf{R}(P,Q,x) = 0$ will certainly capture all the situations where P and Q have a common root, but it may also include some situations where there is a reduction in degree, although the polynomials have no common root. For instance, taking $P = tx - 1, Q = tx^2 - 1$ (with t a parameter), the resultant with respect to x is found to be

$$\mathcal{R} = t(1-t).$$

Indeed, the condition $\mathcal{R} = 0$ corresponds to either a common root ($t = 1$ implying $P(1) = Q(1) = 0$) or to some degeneracy in degree ($t = 0$).

Given a system

$$\{P_j(z, y_1, y_2, \ldots, y_m) = 0\}, \qquad j = 1\,..\,m, \tag{2}$$

defining an algebraic curve, we can then proceed as follows in order to extract a single equation satisfied by one of the indeterminates. By taking resultants with P_m, eliminate all occurrences of the variable y_m from the first $m-1$ equations, thereby obtaining a new system of $m-1$ equations in $m-1$ variables (and z kept as a parameter, so that the base field is $\mathbb{C}(z)$). Repeat the process and successively eliminate $y_{m-1}, \ldots, y_2$. The strategy (in the simpler case where variables are eliminated in succession exactly one at a time) is summarized in the skeletton procedure Eliminate:

procedure Eliminate $(P_1, \ldots, P_m, y_1, y_2, \ldots y_m)$;
$\{$*Elimination of* $y_2, \ldots, y_m$ *by resultants*$\}$
$(A_1, \ldots, A_m) \; := \; (P_1, \ldots, P_m)$;
for j **from** m **by** -1 to 2 **do**
for k **from** $j-1$ **by** -1 to 1 **do**
$A_k \; := \; \mathbf{R}(A_k, A_j, y_{k+1})$;
return(A_1).

The polynomials obtained need not be minimal, in which case, one should appeal to multivariate polynomial factorization in order to select the relevant factors at each stage. (Groebner bases provide a neater alternative to these questions, see [**40**].)

Computer algebra systems usually provide implementations of both resultants and Groebner bases. The complexity of elimination is however exponential in the worst-case. For instance, degrees essentially multiply, which is somewhat intrinsic as y_0 in the system

$$y_0 - z - y_k = 0,\; y_k - y_{k-1}^2 = 0,\; \ldots\; , y_1 - y_0^2 = 0,$$

defines with k equations the OGF of regular trees of degree 2^k, while it represents an algebraic function of degree 2^k and no less.

2. Equivalent definitions of analyticity. Three different notions are introduced at the beginning of Chapter IV: analyticity (defined by power series expansions), holomorphy (defined as complex differentiability), and the property of having null integrals along loops. As is known from any textbook on complex analysis, these three notions are equivalent.

This appendix entry sketches a proof of the equivalence, which is summarized by the following diagram:

$$\begin{array}{ccc}
\text{Analyticity} & \overset{[A]}{\Longrightarrow} & \mathbb{C}\text{-differentiability} \\
\Uparrow\ [C] & & \Downarrow\ [B] \\
\text{(Integral representation)} & \overset{[C]}{\Longleftarrow} & \text{Null integrals}
\end{array}$$

A. Analyticity implies complex-differentiability. Asume that $f(z)$ is analytic in the disc $D(z_0; R)$. We may assume without loss of generality that $z_0 = 0$ and $R = 1$ (else effect a linear transformation on the argument z). According to the definition of analyticity, the series representation

$$f(z) = \sum_{n=0}^{\infty} f_n z^n,$$

converges for all z with $|z| < 1$. We are first going to prove that $f(z)$ given by this representation is analytic at any z_1 interior to $D(0; 1)$ by means of simple series rearrangements. First, formally, the binomial theorem provides

$$\begin{aligned}
f(z) &= \sum_{n\geq 0} f_n z^n \quad = \quad \sum_{n\geq 0} f_n (z_1 + z - z_1)^n \\
&= \sum_{n\geq 0} \sum_{k=0}^{n} \binom{n}{k} f_n z_1^k (z - z_1)^{n-k} \\
&= \sum_{m\geq 0} c_m (z - z_1)^m, \qquad c_m := \sum_{k\geq 0} \binom{m+k}{k} f_{m+k} z_1^k.
\end{aligned} \tag{3}$$

Let r_1 be any number smaller than $1 - |z_1|$. We observe that (3) makes analytic sense. Indeed, one has the bound $|f_n| \leq CA^n$, valid for any $A > 1$ and some $C > 0$. Thus, the terms in (3) are dominated in absolute value by those of the double series

$$\sum_{n\geq 0} \sum_{k=0}^{n} \binom{n}{k} CA^n |z_1|^k r_1^{n-k} = C \sum_{n\geq 0} A^n (|z_1| + r_1)^n = \frac{C}{1 - A(|z_1| + r_1)}, \tag{4}$$

which is absolutely convergent as soon as A is chosen such that $A < (|z_1| + r_1)^{-1}$.

Complex differentiability of at any $z_1 \in D(0; 1)$ derives from the calculation, valid for small enough δ,

$$\begin{aligned}
\frac{1}{\delta}(f(z_1 + \delta) - f(z_1))) &= \sum_{n\geq 0} n f_n z_1^{n-1} + \delta \sum_{n\geq 0} \sum_{k=2}^{n} \binom{n}{k} f_n z_1^k (\delta)^{n-k-2} \\
&= \sum_{n\geq 0} n f_n z_1^{n-1} + O(\delta),
\end{aligned} \tag{5}$$

where boundedness of the coefficient of δ results from an argument analogous to the one used in (4).

This argument also shows that the derivative of f at z_1 is obtained by differentiating termwise the series representing f. More generally derivatives of all orders exist. In view of this fact, the equalities of (3) can also be interpreted as the *Taylor expansion* (by grouping terms according to values of k first):

$$f(z_1 + \delta) = f(z_1) + \delta f'(z_1) + \frac{\delta^2}{2!} f''(z_1) + \cdots . \tag{6}$$

B. Complex differentiability implies the "Null Integral" Property. This starts from the *Cauchy– Riemann equations*. Let $P(x,y) = \Re f(x+iy)$ and $Q(x,y) = \Im f(x+iy)$. By adopting successively in the definition of complex differentiability $\delta = h$ and $\delta = ih$, one finds $P'_x + iQ'_x = Q'_y - iP'_y$, implying

$$\frac{\partial P}{\partial x} = \frac{\partial Q}{\partial y} \qquad \text{and} \qquad \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \tag{7}$$

known as the Cauchy–Riemann equations. (The functions P and Q satisfy the partial differential equations $\Delta f = 0$, where Δ is the 2-dimensional *Laplacian* $\Delta := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$; such functions are known as *harmonic functions*.)

The Proof of the Null Integral Property (i.e., $\int_\lambda f = 0$ for a loop λ) from differentiability results from the Cauchy–Riemann equations, taking into account Greene's formula,

$$\int_{\partial K} A dx + B dy = \int\int_K \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\right) dx\, dy,$$

which is valid for any (compact) domain K enclosed by a simple curve ∂K.

C. "Null Integrals" implies analyticity. The starting point is the formula

$$f(a) = \frac{1}{2i\pi}\int_\gamma \frac{f(z)}{z-a}\, dz, \tag{8}$$

knowing only differentiability of f and its consequence, the Null Integral Property (but *not* postulating the existence of an analytic expansion). There γ is a simple positive loop encircling a inside a region where f is analytic.

The proof of (8) is obtained by decomposing $f(z)$ in the original integral as $f(z) = f(z) - f(a) + f(a)$. Define accordingly

$$g(z) = \begin{cases} \frac{f(z)-f(a)}{z-a} & \text{for} \quad z \neq a \\ f'(a) & \text{for} \quad z \neq a. \end{cases}$$

By the differentiability assumption, g is continuous and holomorphic (differentiable) at any point other than a. Its integral is thus 0 along γ. On the other hand, we have

$$\int_\gamma \frac{1}{z-a}\, dz = 2i\pi,$$

by a simple computation: deform γ to a small circle along a and evaluate the integral directly by setting $z - a = re^{i\theta}$.

Once (8) is granted, it suffices to write, e.g., for an expansion at 0,

$$\begin{aligned} f(z) &= \frac{1}{2i\pi}\int_\gamma f(t)\,\frac{dt}{t-z} \\ &= \frac{1}{2i\pi}\int_\gamma f(t)\left(1 + \frac{z}{t} + \frac{z^2}{t^2} + \cdots\right)\frac{dt}{t} \\ &= \sum_{n\geq 0} f_n z^n, \qquad f_n := \frac{1}{2i\pi}\int_\gamma f(t)\,\frac{dt}{t^{n+1}}. \end{aligned}$$

(Exchanges of integration and summation are justified by normal convergence.) Analyticity is thus proved.

3. Gamma function. The formulæ of singularity analysis involve the *Gamma function* in an essential manner. The Gamma function extends to nonintegral arguments the factorial function and we collect in this appendix a few classical facts regarding it. Proofs may be found in classic treatises like [**107**] or [**191**]. We first list the basic function–theoretic properties. Next we prove the Hankel contour representation that illustrates a technique fundamental to singularity analysis. Last, we conclude with a few classical expansions that are of use in Chapter VI.

Euler introduced the Gamma function as

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt, \tag{9}$$

where the integral converges provided $\Re(s) > 0$. Through integration by parts, one immediately derives the basic functional equation of the Gamma function,

$$\Gamma(s+1) = s\Gamma(s). \tag{10}$$

Since $\Gamma(1) = 1$, one has $\Gamma(n+1) = n!$, so that the Gamma function serves to extend the factorial function for nonintegral arguments. For combinatorial purposes, the special value,

$$\Gamma\left(\frac{1}{2}\right) := \int_0^\infty e^{-t}\, \frac{dt}{\sqrt{t}} = 2\int_0^\infty e^{-x^2}\, dx = \sqrt{\pi},$$

proves to be quite important. It implies in turn $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$.

$\triangleright$ **2.** *Evaluation of the Gaussian integral.* Define $J := \int_0^\infty e^{-x^2}\, dx$. The idea is to evaluate J^2:

$$J^2 = \int_0^\infty \int_0^\infty e^{-(x^2+y^2)}\, dxdy.$$

Going to polar coordinates, $(x^2+y^2)^{1/2} = \rho$, $x = \rho\cos\theta$, $y = \rho\sin\theta$ yields, via the standard change of variables formula:

$$J^2 = \int_0^\infty \int_0^{\frac{\pi}{2}} e^{-\rho^2}\, \rho d\rho d\theta,$$

so that $J^2 = \pi/4$. $\triangleleft$

From (10) that the Gamma function can be analytically continued to the whole of $\mathbb{C}$ with the exception of poles at $0, -1, -2, \ldots$. As $s \to -m$, with m an integer, the functional equation used backwards yields

$$\Gamma(s) \sim \frac{(-1)^m}{m!}\frac{1}{s+m}\ ,$$

so that the residue of $\Gamma(s)$ at $s = -m$ is $(-1)^m/m!$. Figure 1 depicts the graph of $\Gamma(s)$ for real values of s.

Hankel contour representation. Euler's integral representation of $\Gamma(s)$ used in conjunction with the functional equation permits us to continue $\Gamma(s)$ to the whole of the complex plane. A direct approach due to Hankel provides an alternative integral representation valid for all values of s.

THEOREM B.1 (Hankel's contour integral). *Let $\int_{+\infty}^{(0)}$ denote an integral taken along a contour starting at $+\infty$ in the upper plane, winding counterclockwise around the origin, and proceeding towards $+\infty$ in the lower half plane. Then, for all $s \in \mathbb{C}$,*

$$\frac{1}{\Gamma(s)} = -\frac{1}{2i\pi}\int_{+\infty}^{(0)} (-t)^{-s} e^{-t}\, dt. \tag{11}$$

In (11), $(-t)^{-s}$ is assumed to have its principal determination when t is negative real, and this determination is then extended uniquely by continuity throughout the contour. The left hand side of (11) can also be rewritten as $\frac{1}{\pi}\sin(\pi s)\Gamma(1-s)$, by virtue of the complement formula (this is absurd!).

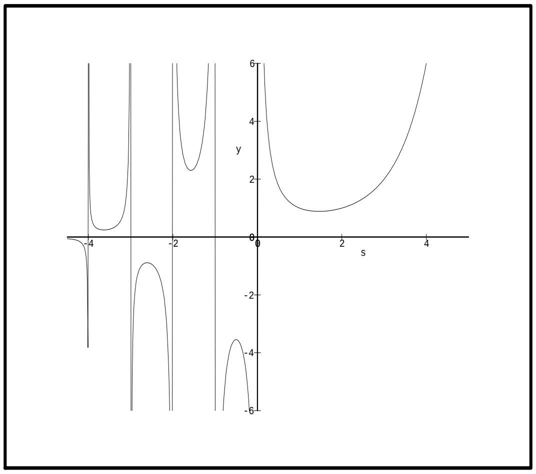


FIGURE 1. A plot of $\Gamma(s)$ for real s.

PROOF. We refer to volume 2 of Henrici's book [**107**, p. 35] or Whittaker and Watson's treatise [**191**, p. 245] for a detailed proof.

A contour of integration that fulfills the conditions of the theorem is typically the contour $\mathcal{H}$ that is at distance 1 of the positive real axis comprising three parts: a line parallel to the positive real axis in the upper half–plane; a connecting semi–circle centered at the origin; a line parallel to the positive real axis in the lower half–plane. More precisely, $\mathcal{H} = \mathcal{H}^- \cup \mathcal{H}^+ \cup \mathcal{H}^\circ$, where

$$\begin{cases} \mathcal{H}^- & = & \{z = w - i, \ w \geq 0\} \\ \mathcal{H}^+ & = & \{z = w + i, \ w \geq 0\} \\ \mathcal{H}^\circ & = & \{z = -e^{i\phi}, \ \phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]\}. \end{cases} \tag{12}$$

Let ϵ be a small positive real number, and denote by $\epsilon \cdot \mathcal{H}$ the image of $\mathcal{H}$ by the transformation $z \mapsto \epsilon z$. By analyticity, for the integral representation, we can equally well adopt as integration path the contour $\epsilon \cdot \mathcal{H}$, for any $\epsilon > 0$. The main idea is then to let ϵ tend to 0.

Assume momentarily that $s < 0$. (The extension to arbitrary s then follows by analytic continuation.) The integral along $\epsilon \cdot \mathcal{H}$ decomposes into three parts:

The integral along the semi–circle is 0 if we take the circle of a vanishing small radius, since $-s > 0$.

The contributions from the upper and lower lines give, as $\epsilon \to 0$

$$\int_{+\infty}^{(0)} (-t)^{-s} e^{-t} \, dt = (-U + L) \int_0^\infty t^{-s} e^{-t} \, dt$$

where U and L denote the determinations of $(-1)^{-s}$ on the half-lines lying in the upper and lower half planes respectively.

By continuity of determinations, $U = (e^{-i\pi})^{-s}$ and $L = (e^{+i\pi})^{-s}$. Therefore, the right hand side of (11) is equal to

$$-\frac{(-e^{i\pi s} + e^{-i\pi s})}{2i\pi} \Gamma(1-s) = -\frac{\sin(\pi s)}{\pi} \Gamma(1-s)$$

which reduces to $1/\Gamma(s)$ by the complement formula for the Gamma function. $\square$

Expansions. The Gamma function has poles at the nonpositive integers but has no zeros. Accordingly, $1/\Gamma(s)$ is an entire function with zeros at $0, -1, \ldots$, and the position of the zeros is reflected by the product decomposition,

$$\frac{1}{\Gamma(s)} = se^{\gamma s} \prod_{m=1}^{\infty} \left[(1 + \frac{s}{n}) e^{-s/n} \right] \tag{13}$$

(of the so–called Weierstraß type). There $\gamma = 0.57721$ denotes Euler's constant

$$\gamma = \lim_{n \to \infty} (H_n - \log n) \equiv \sum_{n=1}^{\infty} \left[\frac{1}{n} - \log(1 + \frac{1}{n}) \right].$$

The logarithmic derivative of the Gamma function is classically known as the psi function and is denoted by $\psi(s)$:

$$\psi(s) = \frac{d}{ds} \log \Gamma(s) = \frac{\Gamma'(s)}{\Gamma(s)}.$$

In accordance with (13), $\psi(s)$ admits a partial fraction decomposition

$$\psi(s+1) = -\gamma - \sum_{n=1}^{\infty} \left[\frac{1}{n+s} - \frac{1}{n} \right]. \tag{14}$$

From (14), there results that the Taylor expansion of $\psi(s+1)$, hence of $\Gamma(s+1)$, involves values of the Riemann zeta function,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

at the positive integers: for $|s| < 1$,

$$\psi(s+1) = -\gamma + \sum_{n=2}^{\infty} (-1)^n \zeta(n) s^{n-1}.$$

so that the coefficients in the expansion of $\Gamma(s)$ around any integer are polynomially expressible in terms of Euler's constant γ and values of the zeta function at the integers. For instance,

$$\Gamma(s+1) = 1 - \gamma\, s + \left(\frac{\pi^2}{12} + \frac{\gamma^2}{2} \right) s^2 + \left(-\frac{\zeta(3)}{3} - \frac{\pi^2 \gamma}{12} - \frac{\gamma^3}{6} \right) s^3 + \mathcal{O}(s^4).$$

Another direct consequence of the infinite product formulæ for $\Gamma(s)$ and $\sin \pi s$ is the complement formula for the Gamma function,

$$\Gamma(s)\Gamma(-s) = -\frac{\pi}{s \sin \pi s}, \tag{15}$$

which directly results from the factorization of the sine function (due to Euler),

$$\sin s = s \prod_{n=1}^{\infty} \left(1 - \frac{s^2}{n^2 \pi^2} \right).$$

In particular, the complement formula entails the special value

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}.$$

▷ **3.** *The duplication formula.* This is

$$2^{2s-1} \Gamma(s) \Gamma(s + \frac{1}{2}) = \pi^{1/2} \Gamma(2s),$$

which provides the expansion of Γ near $1/2$:

$$\Gamma(s+\frac{1}{2}) = \pi^{1/2} - (\gamma + 2\,\log 2)\,\pi^{1/2}s + \left(\frac{\pi^{5/2}}{4} + \frac{(\gamma+2\,\log 2)^2\,\pi^{1/2}}{2}\right)s^2 + \mathcal{O}(s^3).$$

$\lhd$

Finally, a famous asymptotic formula is Stirling's approximation, familiarly known as "Stirling's formula":

$$\Gamma(s+1) = s\Gamma(s) \sim s^s e^{-s}\sqrt{2\pi s}\left[1 + \frac{1}{12s} + \frac{1}{288s^2} - \frac{134}{51840s^4} + \cdots\right].$$

It is valid for (large) real $s > 0$, and more generally for all $s \to \infty$ in $|\operatorname{Arg}(s)| < \pi - \delta$ (any $\delta > 0$). For the purpose of obtaining effective bounds, the following quantitative relation [**191**, p. 253] often proves useful

$$\Gamma(s+1) = s^s e^{-s}(2\pi s)^{1/2} e^{\theta/(12s)} \text{ where } 0 < \theta < 1,$$

an equality that holds now for all $s \geq 1$.

$\rhd$ **4.** *Stirling's formula via the method of Laplace.* Stirling's formula for large s can be derived by applying Laplace's method to the integral

$$\int_0^\infty e^{-t}t^s\,dt \equiv \int_0^\infty e^{-t+s\log t}\,dt,$$

and by expanding near the maximum of the integrand, namely, $t = s$. [See [**39**, p. 267] for an explicit form of the full expansion related to derangement numbers.] $\lhd$

$\rhd$ **5.** *Stirling's formula via Euler–Maclaurin summation.* Stirling's formula can be derived from Euler–Maclaurin summation applied to $\log\Gamma(s)$. $\lhd$

$\rhd$ **6.** *The Eulerian Beta function.* It is defined for $\Re(p), \Re(q) > 0$ by any of the following integrals:

$$B(p,q) := \int_0^1 x^{p-1}(1-x)^{q-1}\,dx = \int_0^\infty \frac{y^{p-1}}{(1+y)^{p+q}}\,dy = 2\int_0^{\frac{\pi}{2}} \cos^{2p-1}\theta\,\sin^{2p-1}\theta\,d\theta.$$

It satisfies:

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

[See [**191**, p. 254] for a proof generalizing that of Note 2.] $\lhd$

4. Implicit Function Theorem. In its real variable version, the implicit function theorem asserts that, for a sufficiently smooth function $F(z,w)$ of two variables, a solution to the the equation $F(z,w) = 0$ exists in the vicinity of a solution point (z_0, w_0) (therefore satisfying $F(z_0, w_0) = 0$) provided the partial derivative satisfies $F'_w(z_0, w_0) \neq 0$. This theorem admits a complex-analytic extension.

Without loss of generality, one restrict attention to $(z_0, w_0) = (0, 0)$. Let $F(z,w)$ be an analytic function of two complex variables in the sense that it admits a convergent representation valid in a polydisc,

$$F(z,w) = \sum_{m,n\geq 0} f_{m,n} z^m w^n, \qquad |z| < R, \quad |w| < S. \tag{16}$$

for some $R, S > 0$.

THEOREM B.2 (Analytic Implicit Functions). *Let F be bivariate analytic in the sense of* (16)*. Assume that $F(0,0) \equiv f_{0,0} = 0$ and $F'_w(0,0) \equiv f_{0,1} \neq 0$. Then, there exists a unique function $f(z)$ analytic in a neighbourhood $|z| < \rho$ of 0 such that $f(0) = 0$ and*

$$F(z, f(z)) = 0, \qquad |z| < \rho.$$

PROOF (SKETCH). We refer to the discussion in Hille's book [**108**] for details. Three different proofs can be given.

(i) Proof by residues. Make use of the principle of the argument and Rouché's Theorem to see that the equation $F(z,w)$ has a unique solution near 0 for $|z|$ small enough. Appeal to the related result of Chapter IV (based on the residue theorem) that expresses the sum of the solutions to an equation as a contour integral. Here, this expresses the solution as (C a small enough contour around 0)

$$f(z) = \frac{1}{2i\pi}\int_C w\frac{F'_w(z,w)}{F(z,w)}\,dw,$$

which is checked to represent an analytic function.

(ii) Proof by majorant series. Set $G(z,w) := w - f_{0,1}^{-1}F(z,w)$. The equation $F(z,w) = 0$ becomes the fixed-point equation $w = G(z,w)$. The bivarite series G has its coefficients dominated termwise by those of

$$\widehat{G}(z,w) = \frac{A}{(1-z/R)(1-w/S)} - A - A\frac{w}{S}.$$

The equation $w = \widehat{G}(z,w)$ is quadratic. It admits a solution $\widehat{f}(z)$ analytic at 0,

$$\widehat{f}(z) = A\frac{z}{R} + \frac{A(A^2+AS+S^2)}{S^2}\frac{z^2}{R^2} + \cdots,$$

whose coefficients dominate termwise those of f.

(iii) Proof by Picard's method of successive approximants. With G like before, define the sequence of functions

$$\phi_0(z) := 0; \quad \phi_{j+1}(z) = G(z,\phi_j(z)),$$

all analytic in a small neighbourhood of 0. Then $f(z)$ can be obtained as

$$f(z) = \lim_{j\to\infty}\phi_j(z) \equiv \phi_0(z) - \sum_{j=0}^{\infty}\left(\phi_j(z) - \phi_{j+1}(z)\right),$$

which is itself checked to be analytic near 0. □

5. Mellin transform. The Mellin transform of a function f defined over $\mathbb{R}_{>0}$ is the complex variable function $f^\star(s)$ defined by the integral

$$f^\star(s) := \int_0^\infty f(x)x^{s-1}\,dx. \tag{17}$$

This transform is also occasionally denoted by $\mathcal{M}[f]$ or $\mathcal{M}[f(x);s]$. Its importance devolves from two properties: (i) it *maps* asymptotic expansions of a function at 0 and $+\infty$ to singularities of the transform; (ii) it *factorizes* harmonic sums (defined below). The conjunction of the mapping property and the harmonic sum property makes it possible to analyse asymptotically rather complicated sums arising from a linear superposition of models taken at different scales. Major properties are summarized in Figure 2

It is assumed that f is locally integrable. Then, the two conditions,

$$f(x) \underset{x\to 0^+}{=} O(x^u), \qquad f(x) \underset{x\to+\infty}{=} O(x^v),$$

guarantee that f^* exists for s in a *strip*,

$$s \in \langle -u,-v\rangle, \qquad \text{i.e.,} \qquad -u < \Re(s) < -v.$$

Thus existence of the transform is granted provided $v < u$. The prototypical Mellin transform is the Gamma function discussed elsewhere in this appendix:

$$\Gamma(s) := \int_0^\infty e^{-x} x^{s-1}\, dx = \mathcal{M}[e^{-x}; s], \qquad 0 < \Re(s) < \infty.$$

Similarly $f(x) = (1+x)^{-1}$ is $O(x^0)$ at 0 and $O(x^{-1})$ at infinity, and hence its transform exists in the strip $\langle -1, 0\rangle$; it is in fact $\pi/\sin\pi s$. The Heaviside function $H(x) := [\![0 \le x < 1]\!]$ exists in $\langle 0 < +\infty\rangle$ and has transform $1/s$.

Harmonic sum propery. The Mellin transform is a linear transform. In addition, it satisfies the important rescaling rule

$$f(x) \overset{\mathcal{M}}{\mapsto} f^\star(s) \qquad \text{implies} \qquad f(\mu x) \overset{\mathcal{M}}{\mapsto} \mu^{-s} f^\star(s),$$

for any $\mu > 0$. Linearity then implies the derived rule

$$\sum_k \lambda_k f(\mu_k x) \overset{\mathcal{M}}{\mapsto} \left(\lambda_k \mu_k^{-s}\right) \cdot f^\star(s), \tag{18}$$

valid a priori for any finite set of pairs (λ_k, μ_k) and extending to infinite sums whenever the interchange of $\int$ and $\sum$ is permissible. A sum of the form (18) is called a *harmonic sum*, the function f is the "base function", the λ's the "amplitudes" and the μ's the "frequencies". Equation (18) then yields the "harmonic sum rule": *The Mellin transform of a harmonic sum* factorizes *as the product of the transform of the base function and a generalized Dirichlet series associated to amplitudes and frequencies*. Harmonic sums surface recurrently in the context of analytic combinatorics and Mellin transforms are a method of choice for coping with them.

Here are a few examples of application of the rule (18):

$$\begin{array}{llll}
\displaystyle\sum_{k\ge1} e^{-k^2x^2} & \mapsto & \frac{1}{2}\Gamma(s/2)\zeta(s) & \Re(s) > 1 \\
\displaystyle\sum_{k\ge0} e^{-x2^k} & \mapsto & \dfrac{\Gamma(s)}{1-2^{-s}} & \Re(s) > 0 \\
\displaystyle\sum_{k\ge0} (\log k) e^{-\sqrt{k}x} & \mapsto & -\frac{1}{2}\zeta'(s/2)\Gamma(s) & \Re(s) > 2 \\
\displaystyle\sum_{k\ge1} \frac{1}{k(k+x)} & \mapsto & \zeta(2-s)\dfrac{\pi}{\sin\pi s} & 0 < \Re(s) < 1
\end{array}$$

$\triangleright$ **7.** *Connection between power series and Dirichlet series.* Let (f_n) be a sequence of numbers with at most polynomial growth, $f_n = O(n^r)$, and with OGF $f(z)$. Then, one has

$$\sum_{n\ge1} \frac{f_n}{n^s} = \frac{1}{\Gamma(s)} \int_0^\infty f\left(e^{-x}\right) x^{s-1}\, dx, \qquad \Re(s) > r+1.$$

For instance, one obtains the Mellin pairs

$$\frac{e^{-x}}{1-e^{-x}} \overset{\mathcal{M}}{\mapsto} \zeta(s)\Gamma(s) \quad (\Re(s) > 1), \qquad \log\frac{1}{1-e^{x}} \overset{\mathcal{M}}{\mapsto} \zeta(s+1)\Gamma(s) \quad (\Re(s) > 0).$$

These may be used to analyse sums or, conversely, deduce analytic properties of Dirichlet series. $\triangleleft$

Mapping properties. Mellin transfoms map asymptotic terms in the expansions of a function f at 0 and $+\infty$ onto singular terms of the transform $f^\star$. This property stems from the

Function ($f(x)$)	Mellin transform ($f^\star(s)$)	
$f(x)$	$\int_0^\infty f(x)x^{s-1}\,dx$	**definition**, $s \in \langle -u, -v\rangle$
$\frac{1}{2i\pi}\int_{c-i\infty}^{c+i\infty} f^\star(s)x^{-s}\,ds$	$f^\star(s)$	**inversion th.**, $-u < c < -v$
$\sum_i \lambda_i f_i(x)$	$\sum_i \lambda_i f_i^\star(s)$	linearity
$f(\mu x)$	$\mu^{-s} f^\star(s)$	scaling rule ($\mu > 0$)
$x^\rho f(x^\theta)$	$\frac{1}{\theta} f^\star\left(\frac{s+\rho}{\theta}\right)$	power rule
$\sum_i \lambda_i f(\mu_i x)$	$\left(\sum_i \lambda_i \mu_i^{-s}\right) \cdot f^\star(s)$	**harmonic sum rule** ($\mu_i > 0$)
$\int_0^\infty \lambda(t) f(tx)\,dt$	$\int_0^\infty \lambda(t)t^{-s}\,dt \cdot f^\star(s)$	harmonic integral rule
$f(x)\log^k x$	$\partial_s^k f^\star(s)$	diff. I, $k \in \mathbb{Z}_{\geq 0}$, $\partial_s := \frac{d}{ds}$
$\partial_x^k f(x)$	$\frac{(-1)^k \Gamma(s)}{\Gamma(s-k)} f^\star(s-k)$	diff. II, $k \in \mathbb{Z}_{\geq 0}$, $\partial_x := \frac{d}{dx}$
$\underset{x\to 0}{\sim} x^\alpha (\log x)^k$	$\underset{s\to -\alpha}{\sim} \frac{(-1)^k k!}{(s+\alpha)^{k+1}}$	**mapping**: $x \to 0$, left poles
$\underset{x\to +\infty}{\sim} x^\beta (\log x)^k$	$\underset{s\to -\beta}{\sim} \frac{(-1)^{k-1} k!}{(s+\beta)^{k+1}}$	**mapping**: $x \to \infty$, right poles

FIGURE 2. Major properties of Mellin transforms.

basic identities

$$H(x)x^\alpha \overset{\mathcal{M}}{\mapsto} \frac{1}{s+\alpha}\ (s \in \langle -\alpha, +\infty\rangle), \qquad (1 - H(x))x^\beta \overset{\mathcal{M}}{\mapsto} \frac{1}{s+\beta}\ (s \in \langle -\infty, -\beta\rangle),$$

as well as what one obtains by differentiation with respect to α, β.

The converse mapping property also holds. Like for other integral transforms, there is an inversion formula: if f is continuous in an interval containing x, then

$$f(x) = \frac{1}{2i\pi}\int_{c-i\infty}^{c+i\infty} f^\star(s)x^{-s}\,ds, \tag{19}$$

where the abscissa c should be chosen in the "fundamental strip" of f; for instance any c satisfying $-u < c < -v$ with u, v as above is suitable.

In many cases of practical interest, $f^\star$ is continuable into a meromorphic function to the whole of $\mathbb{C}$. If the continuation of $f^\star$ does not grow too fast along vertical lines, then one can estimate the inverse Mellin integral of (19) by residues. This corresponds to shifting the line of integration to some $d \neq c$ and taking poles into account by the residue theorem. Since the residue at a pole s_0 of $f^\star$ involves a factor of x^{-s_0}, the contribution of s_0 will give useful information on $f(x)$ as $x \to \infty$ if s_0 lies to the right of c, and on $f(x)$ as $x \to 0$ if s_0 lies to the left. Higher order poles introduce additional logarithmic factors. The "dictionary" is simply

$$\frac{1}{(s-s_0)^k} \overset{\mathcal{M}^{-1}}{\mapsto} \pm\frac{(-1)^k}{k!} x^{-s_0}(\log x)^k, \tag{20}$$

where the sign is '$+$' for a pole on the left of the fundamental strip and '$-$' for a pole on the right.

Mellin asymptotic summation. The combination of mapping properties and the harmonic sum property constitutes a powerful tool of asymptotic analysis. As an example, let us first investigate the pair

$$F(x) := \sum_{k\geq 1} \frac{1}{1+k^2x^2}, \qquad F^\star(s) = \frac{1}{2}\frac{\pi}{\sin\frac{1}{2}\pi s}\zeta(s),$$

where $F^\star$ results from the harmonic sum rule and is is originally defined in the strip $\langle 1, 2\rangle$. The function is meromorphically continuable to the whole of $\mathbb{C}$ with poles at the points $0, 1, 2$ and $4, 6, 8, \ldots$. The transform $F^\star$ is small towards infinity, so that application of the dictionary (20) is justified. One then finds mechanically:

$$F(x) \underset{x\to 0^+}{\sim} \frac{\pi}{2x} - \frac{1}{2} + O(x^M), \qquad F(x) \underset{x\to+\infty}{\sim} \frac{\pi^2}{6x^2} - \frac{\pi^4}{90x^4} + \cdots,$$

for any $M > 0$.

A particularly important quantity in analytic combinatorics is the harmonic sum

$$\Phi(x) := \sum_{k=0}^{\infty}\left(1 - e^{-x/2^k}\right).$$

It occurs for instance in the analysis of longest runs on page 61. By the harmonic sum rule, one finds

$$\Phi^\star(s) = -\frac{\Gamma(s)}{1-2^s}, \qquad s \in \langle -1, 0\rangle$$

(The transform of $e^{-x} - 1$ is also $\Gamma(s)$, but in the shifted strip $\langle -1, 0\rangle$.) The singularities of $\Phi^\star$ are at $s = 0$, where there is a double pole, at $s = -1, -2, \ldots$ which are simple poles, but also at the complex points

$$\chi_k = \frac{2ik\pi}{\log 2}.$$

The Mellin dictionary (20) can still be applied provided one integrates along a long rectangular contour that passes in-between poles. The salient feature is here the presence of fluctuations induced by the imaginary poles, since

$$x^{-\chi_k} = \exp\left(2ik\pi \log_2 x\right),$$

and each pole induces a Fourier element. All in all, one finds (any $M > 0$):

$$\text{(21)} \qquad \begin{cases} \Phi(x) & \underset{x\to+\infty}{\sim} \quad \log_2 x + \dfrac{\gamma}{\log 2} + \dfrac{1}{2} + P(x) + O(x^M) \\ & \qquad P(x) \ := \ \dfrac{1}{\log 2}\displaystyle\sum_{k\in\mathbb{Z}\setminus\{0\}} \Gamma\left(\frac{2ik\pi}{\log 2}\right) e^{2i\pi\log_2 x}. \end{cases}$$

The analysis for $x \to 0$ is also possible: in this particular case, it yields

$$\Phi(x) \underset{x\to 0}{\sim} \sum_{n\geq 1} \frac{(-1)^{n-1}}{1-2^{-n}}\,\frac{x^n}{n!},$$

which is what would result from expanding the exponential in $\Phi(x)$ and reorganizing the terms, and consequently constitutes an exact formula.

▷ **8.** *Mellin-type derivation of Stirling's formula.* One has the Mellin pair

$$L(x) = \sum_{k\geq 1} \log\left(1+\frac{x}{k}\right) - \frac{x}{k}, \qquad L^\star(s) = -\frac{\pi}{s\sin\pi s}\zeta(1-s), \quad s \in \langle -2,-1\rangle$$

Note that $L(x) = \log(e^{-\gamma x}/\Gamma(1+x))$. Mellin asymptotics provides

$$L(x) \underset{x\to+\infty}{\sim} -x\log x - (\gamma-1)x - \frac{1}{2}\log x - \log\sqrt{2\pi} - \frac{1}{12x} + \frac{1}{360}x^3 - \frac{1}{1260x^5} + \cdots,$$

where one recognizes Stirling's expansion of $x!$,

$$\log x! \underset{x\to+\infty}{\sim} \log\left(x^x e^{-x}\sqrt{2\pi x}\right) + \sum_{n\geq 1} \frac{B_{2n}}{2n(2n-1)} x^{1-2n},$$

with B_n the Bernoulli numbers. ◁

▷ **9.** *Mellin-type analysis of the harmonic numbers.* For a parameter $\alpha > 0$, one has the Mellin pair:

$$K_\alpha(x) = \sum_{k\geq 1}\left(\frac{1}{k^\alpha} - \frac{1}{(k+x)^\alpha}\right), \qquad K_\alpha^\star(s) = -\zeta(\alpha-s)\frac{\Gamma(s)\Gamma(\alpha-s)}{\Gamma(\alpha)}.$$

This serves to estimate harmonic numbers and their generalisations, for instance

$$H_n \underset{n\to\infty}{\sim} \log n + \gamma + \frac{1}{2n} + \sum_{k\geq 2}\frac{B_k}{k}n^{-k} \sim \log n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \cdots,$$

since $K_1(n) = H_n$. ◁

EXAMPLE 1. *Euler-Maclaurin summation via Mellin analysis.* Let f be continuous on $(0,+\infty)$ and satisfy $f(x) =_{x\to+\infty} O(x^{-1-\delta})$, for some $\delta > 0$, and

$$f(x) \underset{x\to 0^+}{\sim} \sum_{k=0}^{\infty} f_k x^k.$$

The summatory function $F(x)$ satisfies

$$F(x) := \sum_{n\geq 1} f(nx), \qquad F^\star(s) = \zeta(s) f^\star(s),$$

by the harmonic sum rule. By the mapping property, the collection of singular expansions of $f^\star$ at $s = 0, -1, -2, \ldots$ is summarized by the formal sum

$$f^\star(s) \asymp \left(\frac{f_0}{s}\right)_{s=0} + \left(\frac{f_1}{s+1}\right)_{s=1} + \left(\frac{f_2}{s+2}\right)_{s=1} + \cdots.$$

Thus, provided $F^\star(s)$ is small towards $\pm i\infty$ in finite strips, one has

$$F(x) \underset{x\to 0}{\sim} \frac{1}{x}\int_0^\infty f(t)\,dt + \sum_{j=0}^{\infty} f_j \zeta(-j) x^j,$$

where the main term is associated to the singularity of $F^\star$ at 1 and arises from the pole of $\zeta(s)$, with $F^\star(1)$ giving the integral of f. The interest of this approach is that it is very versatile and allows for various forms of asymptotic expansions of f at 0 as well as multipliers like $(-1)^k$, $\log k$, and so on; see [**70**] for details. □

Good references on Mellin transforms are the books by Doetsch [**47**] and Widder [**192**]. The term "harmonic sum" and some of the corresponding technology originates with the abstract [**79**]. This brief presentation is based on the survey article [**70**] to which we refer for a detailed treatment. Mellin analysis of "harmonic integrals" is a classical topic of applied mathematics for which we refer to the books by Wong [**194**] and Paris–Kaminski [**155**]. Good treatments of properties of use in discrete mathematics and analysis of algorithms now appear in the books by Hofri [**109**], Mahmoud [**137**], and Szpankowski [**179**].

6. Perron-Frobenius theory of nonnegative matrices. Perron-Frobenius theory gives access to growth properties associated to nonnegative matrices and hence to the dominant singularities of generating functions that satisfy linear systems of equations with nonnegative coefficients. Applications to rational asymptotics, paths, graphs, and automata are detailed in Chapter IV. The purpose here is only to sketch the main techniques from elementary matrix analysis that intervene in this theory. Excellent treatments are to be found in the books of Bellman [**15**, Ch. 16], Gantmacher [**89**, Ch. 13], as well as Karlin and Taylor [**113**, p. 536–551].

THEOREM B.3. *Let A be a matrix whose entries are all positive. Then, A has a unique eigenvalue $\lambda(A)$ which has greatest modulus. This eigenvalue is positive and simple.*

PROOF. The main idea consists in investigating the set of possible "expansion factors"

$$S := \left\{ \lambda \;\middle|\; \exists\, \mathrm{v} \geq 0, \quad A\,\mathrm{v} \geq \lambda\,\mathrm{v} \right\}. \tag{22}$$

(There $\mathrm{v} \geq 0$ means that all components of v are nonnegative and $\mathrm{v} \geq \mathrm{w}$ means that the entries of v are at least as large as the corresponding entries of w.) The largest of the expansion factors,

$$\mu := \sup(S),$$

plays a vital rôle in the argument. The proof relies on establishing that it coincides with the dominant eigenvalue $\lambda(A)$. We set $d = \dim(A)$.

Simple inequalities show that S contains at least the interval $[0, d\min_{i,j} a_{i,j}]$. Inequalities relative to the norm $\|\cdot\|_1$ show that $S \subseteq [0, \sum_{i,j} a_{i,j}]$. Thus, μ is finite and nonzero. That the supremum value μ is actually attained (i.e., $\mu \in S$) results from a simple topological argument detailed in [**15**]: take a bounded family $\mathrm{v}^{(j)}$ corresponding to a sequence $\lambda^{(j)}$ tending to μ and extract a convergent subsequence tending to a vector $\mathrm{v}^{(\infty)}$ which must then satisfy $A\,\mathrm{v}^{\infty)} \geq \mu\,\mathrm{v}^{(\infty)}$. We let w be such a vector of $\mathbb{R}^d_{\geq 0}$ satisfying $A\,\mathrm{w} \geq \mu\,\mathrm{w}$.

Next, one has $A\,\mathrm{w} = \mu\,\mathrm{w}$. Indeed, suppose a contrario that this is not the case and that, without loss of generality,

$$\sum_j A_{1,j} w_j - \mu\,\mathrm{w}_1 = \eta, \qquad \sum_j A_{i,j} w_j - \mu\,\mathrm{w}_1 \geq 0 \quad (i = 2, \ldots, d), \tag{23}$$

for $\eta > 0$. Then, given the slack afforded by η, one could construct a small perturbation $\mathrm{w}^\star$ of w (by $\mathrm{w}^\star_j = \mathrm{w}_j$ for $j = 2, \ldots, d$ and $\mathrm{w}^\star_1 = w_1 + e_1/(2\mu)$) as well as a value $\mu^\star$ such that $A\,\mathrm{w}^\star \geq \mu^\star\,\mathrm{w}$ with $\mu^\star > \mu$, a contradiction. Thus, μ is an eigenvalue of A and w is an eigenvector corresponding to this eigenvalue.

Furthermore, all eigenvalues are dominated in modulus by μ. Let indeed ν and x be such that $A\,\mathrm{x} = \nu\,\mathrm{x}$. One has $A|\,\mathrm{x}\,| \geq |\nu|\,|\,\mathrm{x}\,|$, where $|\,\mathrm{x}\,|$ designates the vector whose entries are the absolute values of the corresponding entries of x. Thus, by the maximality property defining μ, one must have $|\nu| \leq \mu$. If $|\nu| = \mu$ and x is a corresponding eigenvector, then $A|\,\mathrm{x}\,| \geq \mu|\,\mathrm{x}\,|$, and by the same argument as in (23), one must have $A|\,\mathrm{x}\,| = \mu|x|$. Thus $|x|$ is also an eigenvector corresponding to μ. Then, by the triangular inequality, one has $|A\,\mathrm{x}\,| \geq A|\,\mathrm{x}\,|$, so that in fact $A|\,\mathrm{x}\,| = |A\,\mathrm{x}\,|$, which by the converse triangular inequality implies that $\mathrm{x} = \omega\,\mathrm{y}$, where $\omega \in \mathbb{C}$ and y has nonnegative entries. From this observation and the fact that $A\,\mathrm{y} = \nu\,\mathrm{y}$, it results that ν is positive real, so that $\nu = \mu$. Unicity of the dominant eigenvalue is therefore established.

Finally, simplicity of the eigenvalue μ results from a specific argument based on submatrices. If B_k is obtained from A by deleting the kth row and the kth column, then, on general grounds, one has $\lambda(A) > \lambda(B_k)$. From there, through the equality

$$-\frac{d}{d\lambda}|A - \lambda I| = |B_1 - \lambda I| + \cdots + |B_d - \lambda I|$$

(here $|A| = \det(A)$), it can be verified that the derivative of the characteristic polynomial of A at μ is strictly negative, and in particular nonzero; hence simplicity of the eigenvalue μ. See [**15**] for details. $\square$

7. Regular expressions[1]. A *language* is a set of words over some fixed alphabet $\mathcal{A}$. The structurally simplest (yet nontrivial) languages are the *regular languages* that can be defined in a variety of ways: by regular expressions and by finite automata, either deterministic or nondeterministic.

DEFINITION B.1. *The category* RegExp *of* regular expressions *is defined by the property that it contains all the letters of the alphabet* $(a \in \mathcal{A})$ *as well as the empty symbol* ϵ*, and is such that, if* $R_1, R_2 \in$ RegExp*, then the formal expressions* $R_1 \cup R_2$*,* $R_1 \cdot R_2$ *and* $R_1^\star$ *are regular expressions.*

Regular expressions are meant to specify *languages*. The language $\mathcal{L}(R)$ denoted by a regular expression R is defined inductively by the rules: (i) $\mathcal{L}(R) = \{a\}$ if R is the letter $a \in \mathcal{A}$ and $\mathcal{L}(R) = \{\epsilon\}$ (with ϵ the empty word) if R is the symbol ϵ; (ii) $\mathcal{L}(R_1 \cup R_2) = \mathcal{L}(R_1) \cup \mathcal{L}(R_2)$ (with $\cup$ the set-theoretic union); (iii) $\mathcal{L}(R_1 \cdot R_2) = \mathcal{L}(R_1) \cdot \mathcal{L}(R_2)$ (with $\cdot$ the concatenation of words extended to sets); (iv) $\mathcal{L}(R_1^\star) = \{\epsilon\} + \mathcal{L}(R_1) + \mathcal{L}(R_1) \cdot \mathcal{L}(R_1) + \cdots$. A language is said to be a *regular language* if it is specified by a regular expression.

A language is a *set* of words, but a word $w \in \mathcal{L}(R)$ may be parsable in several ways according to R. More precisely, one defines the *ambiguity coefficient* (or *multiplicity*) of w with respect to the regular expression R as the number of parsings, written $\kappa(w) = \kappa_R(w)$. In symbols, we have

$$\kappa_{R_1 \cup R_2}(w) = \kappa_{R_1}(w) + \kappa_{R_2}(w), \quad \kappa_{R_1 \cdot R_2}(w) = \sum_{u \cdot v = w} \kappa_{R_1}(u)\kappa_{R_2}(v),$$

with natural initial conditions ($\kappa_a(b) = \delta_{a,b}$, $\kappa_\epsilon(w) = \delta_{\epsilon,w}$), and with the definition of $\kappa_{R^\star(w)}$ taken as induced by the definition of $R^\star$ via unions and products, namely,

$$\kappa_{R^\star}(w) = \delta_{\epsilon,w} + \sum_{j=1}^{\infty} \kappa_{R^j}(w).$$

As such, $\kappa(w)$ lies in the completed set $\mathbb{N} \cup \{+\infty\}$. We shall only consider here regular expressions R that are *proper*, in the sense that $\kappa_R(w) < +\infty$. It can be checked that this condition is equivalent to requiring that no $S^\star$ with $\epsilon \in \mathcal{L}(S)$ enters in the inductive definition of the regular expression R. (This condition is substantially equivalent to the notion of well-founded specification in Chapter 1.) A regular expression R is said to be *unambiguous* iff for all w, we have $\kappa_R(w) \in \{0, 1\}$; it is said to be *ambiguous*, otherwise.

Given a language $L = \mathcal{L}(R)$, we are interested in two enumerating sequences

$$L_{R,n} = \sum_{|w|=n} \kappa_R(w), \qquad L_n = \sum_{|w|=n} \mathbf{1}_{w \in L},$$

corresponding to the counting of words in the language, respectively, with and without multiplicities.

[1] This entry supplements Appendix A: *Regular languages*.

Bibliography

1. S.-S. Abhyankar, *Algebraic geometry for scientists and engineers*, American Mathematical Society, 1990.
2. Milton Abramowitz and Irene A. Stegun, *Handbook of mathematical functions*, Dover, 1973, A reprint of the tenth National Bureau of Standards edition, 1964.
3. Alfred V. Aho and Margaret J. Corasick, *Efficient string matching: an aid to bibliographic search*, Communications of the ACM **18** (1975), 333–340.
4. David Aldous and James A. Fill, *Reversible Markov chains and random walks on graphs*, 2003, Book in preparation; mansucript available electronically.
5. George E. Andrews, *The theory of partitions*, Encyclopedia of Mathematics and its Applications, vol. 2, Addison–Wesley, 1976.
6. George E. Andrews, Richard Askey, and Ranjan Roy, *Special functions*, Encyclopedia of Mathematics and its Applications, vol. 71, Cambridge University Press, 1999.
7. J. Arney and E. D. Bender, *Random mappings with constraints on coalescence and number of origins*, Pacific Journal of Mathematics **103** (1982), 269–294.
8. Richard Arratia, A. D. Barbour, and Simon Tavaré, *Random combinatorial structures and prime factorizations*, Notices of the American Mathematical Society **44** (1997), no. 8, 903–910.
9. ______, *Logarithmic combinatorial structures: a probabilistic approach*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2003. MR 2 032 426
10. Cyril Banderier, *Combinatoire analytique des chemins et des cartes*, Ph.D. thesis, Université Paris VI, June 2001.
11. Cyril Banderier, Mireille Bousquet-Mélou, Alain Denise, Philippe Flajolet, Danièle Gardy, and Dominique Gouyou-Beauchamps, *On generating functions of generating trees*, Formal Power Series and Algebraic Combinatorics (C. Martínez, M. Noy, and O. Serra, eds.), Universitat Politècnica de Catalunya, June 1999, (Proceedings of FPSAC'99, Barcelona. Also available as INRIA Res. Rep. 3661, April 1999.), pp. 40–52.
12. Cyril Banderier and Philippe Flajolet, *Basic analytic combinatorics of directed lattice paths*, Theoretical Computer Science **281** (2002), no. 1-2, 37–80.
13. Cyril Banderier, Philippe Flajolet, Gilles Schaeffer, and Michèle Soria, *Random maps, coalescing saddles, singularity analysis, and Airy phenomena*, Random Structures & Algorithms **19** (2001), no. 3/4, 194–246.
14. A. D. Barbour, Lars Holst, and Svante Janson, *Poisson approximation*, The Clarendon Press Oxford University Press, New York, 1992, Oxford Science Publications.
15. Richard Bellman, *Matrix analysis*, S.I.A.M. Press, Philadelphia, Pa, 1997, A reprint of the second edition, first published by McGraw-Hill, New York, 1970.
16. Edward A. Bender, *Central and local limit theorems applied to asymptotic enumeration*, Journal of Combinatorial Theory **15** (1973), 91–111.
17. ______, *Asymptotic methods in enumeration*, SIAM Review **16** (1974), no. 4, 485–515.
18. Edward A. Bender and Fred Kochman, *The distribution of subword counts is usually normal*, European Journal of Combinatorics **14** (1993), 265–275.
19. F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial species and tree-like structures*, Cambridge University Press, Cambridge, 1998.
20. Bruce C. Berndt, *Ramanujan's notebooks, part I*, Springer Verlag, 1985.
21. Jean Berstel (ed.), *Séries formelles*, LITP, University of Paris, 1978, (Proceedings of a School, Vieux–Boucau, France, 1977).
22. Jean Bertoin, Philippe Biane, and Marc Yor, *Poissonian exponential functionals, q-series, q-integrals, and the moment problem for log-normal distributions*, Tech. Report PMA-705, Laboratoire de Probabilités et Modèles Aléatoires, Université Paris VI, 2002.
23. Alberto Bertoni, Christian Choffrut, Massimiliano Goldwurm Goldwurm, and Violetta Lonati, *On the number of occurrences of a symbol in words of regular languages*, Theoretical Computer Science **302** (2003), no. 1–3, 431–456.
24. A. T. Bharucha-Reid, *Elements of the theory of Markov processes and their applications*, Dover, 1997, A reprint of the original McGraw-Hill edition, 1960.
25. Philippe Biane, Jim Pitman, and Marc Yor, *Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions*, Bulletin of the American Mathematical Society (N.S.) **38** (2001), no. 4, 435–465.

26. Norman Biggs, E. Keith Lloyd, and Robin Wilson, *Graph theory, 1736–1936*, Oxford University Press, 1974.
27. Jérémie Bourdon and Brigitte Vallée, *Generalized pattern matching statistics*, Mathematics and computer science, II (Versailles, 2002) (B. Chauvin *et al.*, ed.), Trends Math., Birkhäuser, Basel, 2002, pp. 249–265.
28. Mireille Bousquet-Mélou and Marko Petkovšek, *Linear recurrences with constant coefficients: the multivariate case*, Discrete Mathematics **225** (2000), no. 1-3, 51–75.
29. Gilles Brassard and Paul Bratley, *Algorithmique: conception et analyse*, Masson, Paris, 1987.
30. W. G. Brown and W. T. Tutte, *On the enumeration of rooted non-separable planar maps*, Canadian Journal of Mathematics **16** (1964), 572–577.
31. E. Rodney Canfield, *Remarks on an asymptotic method in combinatorics*, Journal of Combinatorial Theory, Series A **37** (1984), 348–352.
32. Henri Cartan, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Hermann, 1961.
33. Pierre Cartier and Dominique Foata, *Problèmes combinatoires de commutation et réarrangements*, Lecture Notes in Mathematics, vol. 85, Springer Verlag, 1969.
34. Frédéric Cazals, *Monomer-dimer tilings*, Studies in Automatic Combinatorics **2** (1997), Electronic publication `http://algo.inria.fr/libraries/autocomb/autocomb.html`.
35. Philippe Chassaing and Jean-François Marckert, *Parking functions, empirical processes, and the width of rooted labeled trees*, Electronic Journal of Combinatorics **8** (2001), no. 1, Research Paper 14, 19 pp. (electronic).
36. Philippe Chassaing, Jean-François Marckert, and Marc Yor, *The height and width of simple trees*, Mathematics and computer science (Versailles, 2000), Trends Math., Birkhäuser, Basel, 2000, pp. 17–30.
37. T. S. Chihara, *An introduction to orthogonal polynomials*, Gordon and Breach, New York, 1978.
38. Noam Chomsky and Marcel Paul Schützenberger, *The algebraic theory of context–free languages*, Computer Programing and Formal Languages (P. Braffort and D. Hirschberg, eds.), North Holland, 1963, pp. 118–161.
39. Louis Comtet, *Advanced combinatorics*, Reidel, Dordrecht, 1974.
40. David Cox, John Little, and Donal O'Shea, *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, 2nd ed., Springer, 1997.
41. N. G. de Bruijn, *A combinatorial problem*, Nederl. Akad. Wetensch., Proc. **49** (1946), 758–764, Also in *Indagationes Math.* **8**, 461–467 (1946).
42. ———, *Asymptotic methods in analysis*, Dover, 1981, A reprint of the third North Holland edition, 1970 (first edition, 1958).
43. N. G. de Bruijn, D. E. Knuth, and S. O. Rice, *The average height of planted plane trees*, Graph Theory and Computing (R. C. Read, ed.), Academic Press, 1972, pp. 15–22.
44. Marie-Pierre Delest and Gérard Viennot, *Algebraic languages and polyominoes enumeration*, Theoretical Computer Science **34** (1984), 169–206.
45. Michael Dellnitz, Oliver Schütze, and Qinghua Zheng, *Locating all the zeros of an analytic function in one complex variable*, J. Comput. Appl. Math. **138** (2002), no. 2, 325–333.
46. Jean Dieudonné, *Calcul infinitésimal*, Hermann, Paris, 1968.
47. G. Doetsch, *Handbuch der Laplace Transformation*, vol. 1–3, Birkhäuser Verlag, Basel, 1955.
48. C. Domb and A.J. Barrett, *Enumeration of ladder graphs*, Discrete Mathematics **9** (1974), 341–358.
49. Peter G. Doyle and J. Laurie Snell, *Random walks and electric networks*, Mathematical Association of America, Washington, DC, 1984.
50. Michael Drmota, *Systems of functional equations*, Random Structures & Algorithms **10** (1997), no. 1–2, 103–124.
51. Michael Drmota and Bernhard Gittenberger, *On the profile of random trees*, Random Structures & Algorithms **10** (1997), no. 4, 421–451.
52. Philippe Duchon, Philippe Flajolet, Guy Louchard, and Gilles Schaeffer, *Boltzmann samplers for the random generation of combinatorial structures*, Preprint, 2003, 44 pages. To appear in *Combinatorics, Probability and Computing*.
53. Isabelle Dutour and Jean-Marc Fédou, *Object grammars and random generation*, Discrete Mathematics and Theoretical Computer Science **2** (1998), 47–61.
54. I. G. Entig, *Generating functions for enumerating self-avoiding rings on the square lattice*, Journal of Physics A: Mathematical and General **18** (1980), 3713–3722.
55. H. M. Farkas and I. Kra, *Riemann surfaces*, second ed., Graduate Texts in Mathematics, vol. 71, Springer-Verlag, New York, 1992.
56. Guy Fayolle and Roudolf Iasnogorodski, *Two coupled processors: the reduction to a Riemann-Hilbert problem*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **47** (1979), no. 3, 325–351.
57. Guy Fayolle, Roudolf Iasnogorodski, and Vadim Malyshev, *Random walks in the quarter-plane*, Springer-Verlag, Berlin, 1999.
58. W. Feller, *An introduction to probability theory and its applications*, third ed., vol. 1, John Wiley, 1968.
59. ———, *An introduction to probability theory and its applications*, vol. 2, John Wiley, 1971.

60. James A. Fill, Philippe Flajolet, and Nevin Kapur, *Singularity analysis, Hadamard products, and tree recurrences*, Preprint, June 2003, 47 pages. Accepted for publication in *Journal of Computational and Applied Mathematics*.
61. Steven Finch, *Mathematical constants*, Cambridge University Press, New-York, 2003.
62. Philippe Flajolet, *Combinatorial aspects of continued fractions*, Discrete Mathematics **32** (1980), 125–161.
63. ———, *Approximate counting: A detailed analysis*, BIT **25** (1985), 113–134.
64. ———, *Elements of a general theory of combinatorial structures*, Fundamentals of Computation Theory (Lothar Budach, ed.), Lecture Notes in Computer Science, vol. 199, Springer Verlag, 1985, Proceedings of FCT'85, Cottbus, GDR, September 1985 (Invited Lecture), pp. 112–127.
65. ———, *Analytic models and ambiguity of context–free languages*, Theoretical Computer Science **49** (1987), 283–309.
66. ———, *Singularity analysis and asymptotics of Bernoulli sums*, Theoretical Computer Science **215** (1999), no. 1-2, 371–381.
67. Philippe Flajolet, Jean Françon, and Jean Vuillemin, *Sequence of operations analysis for dynamic data structures*, Journal of Algorithms **1** (1980), 111–141.
68. Philippe Flajolet, Zhicheng Gao, Andrew Odlyzko, and Bruce Richmond, *The distribution of heights of binary trees and other simple trees*, Combinatorics, Probability and Computing **2** (1993), 145–156.
69. Philippe Flajolet, Dani`ele Gardy, and Lo¨ys Thimonier, *Birthday paradox, coupon collectors, caching algorithms, and self–organizing search*, Discrete Applied Mathematics **39** (1992), 207–229.
70. Philippe Flajolet, Xavier Gourdon, and Philippe Dumas, *Mellin transforms and asymptotics: Harmonic sums*, Theoretical Computer Science **144** (1995), no. 1–2, 3–58.
71. Philippe Flajolet, Xavier Gourdon, and Daniel Panario, *The complete analysis of a polynomial factorization algorithm over finite fields*, Journal of Algorithms **40** (2001), no. 1, 37–81.
72. Philippe Flajolet and Fabrice Guillemin, *The formal theory of birth-and-death processes, lattice path combinatorics, and continued fractions*, Advances in Applied Probability **32** (2000), 750–778.
73. Philippe Flajolet, Kostas Hatzis, Sotiris Nikoletseas, and Paul Spirakis, *On the robustness of interconnections in random graphs: A symbolic approach*, Theoretical Computer Science **287** (2002), no. 2, 513–534.
74. Philippe Flajolet, Peter Kirschenhofer, and Robert F. Tichy, *Deviations from uniformity in random strings*, Probability Theory and Related Fields **80** (1988), 139–150.
75. Philippe Flajolet and Marc Noy, *Analytic combinatorics of non-crossing configurations*, Discrete Mathematics **204** (1999), no. 1-3, 203–229, (Selected papers in honor of Henry W. Gould).
76. Philippe Flajolet and Andrew M. Odlyzko, *The average height of binary trees and other simple trees*, Journal of Computer and System Sciences **25** (1982), 171–213.
77. ———, *Random mapping statistics*, Advances in Cryptology (J-J. Quisquater and J. Vandewalle, eds.), Lecture Notes in Computer Science, vol. 434, Springer Verlag, 1990, Proceedings of EUROCRYPT'89, Houtalen, Belgium, April 1989, pp. 329–354.
78. ———, *Singularity analysis of generating functions*, SIAM Journal on Algebraic and Discrete Methods **3** (1990), no. 2, 216–240.
79. Philippe Flajolet, Mireille R´egnier, and Robert Sedgewick, *Some uses of the Mellin integral transform in the analysis of algorithms*, Combinatorial Algorithms on Words (A. Apostolico and Z. Galil, eds.), NATO Advance Science Institute Series. Series F: Computer and Systems Sciences, vol. 12, Springer Verlag, 1985, (Invited Lecture), pp. 241–254.
80. Philippe Flajolet, Bruno Salvy, and Paul Zimmermann, *Automatic average–case analysis of algorithms*, Theoretical Computer Science **79** (1991), no. 1, 37–109.
81. Philippe Flajolet and Robert Sedgewick, *Mellin transforms and asymptotics: finite differences and Rice's integrals*, Theoretical Computer Science **144** (1995), no. 1–2, 101–124.
82. Philippe Flajolet, Paolo Sipala, and Jean-Marc Steyaert, *Analytic variations on the common subexpression problem*, Automata, Languages, and Programming (M. S. Paterson, ed.), Lecture Notes in Computer Science, vol. 443, 1990, Proceedings of the 17th ICALP Conference, Warwick, July 1990, pp. 220–234.
83. Philippe Flajolet and Mich`ele Soria, *Gaussian limiting distributions for the number of components in combinatorial structures*, Journal of Combinatorial Theory, Series A **53** (1990), 165–182.
84. ———, *General combinatorial schemas: Gaussian limit distributions and exponential tails*, Discrete Mathematics **114** (1993), 159–180.
85. Philippe Flajolet, Wojciech Szpankowski, and Brigitte Vall´ee, *Hidden word statistics*, Preprint, 2002, 34 pages; submitted to *J. of the ACM*.
86. W. B. Ford, *Studies on divergent series and summability and the asymptotic developments of functions defined by Maclaurin series*, 3rd ed., Chelsea Publishing Company, New York, 1960, (From two books originally published in 1916 and 1936.).
87. G. Frobenius, *Über Matrizen aus nicht negativen Elementen*, Sitz.-Ber. Akad. Wiss., Phys-Math Klasse, Berlin (1912), 456–477.
88. William Fulton, *Algebraic curves*, W.A. Benjamin, Inc., New York, Amsterdam, 1969.
89. F. R. Gantmacher, *Matrizentheorie*, Deutscher Verlag der Wissenschaften, Berlin, 1986, A translation of the Russian original Teoria Matriz, Nauka, Moscow, 1966.

90. George Gasper and Mizan Rahman, *Basic hypergeometric series*, Encyclopedia of Mathematics and its Applications, vol. 35, Cambridge University Press, 1990.
91. Ira M. Gessel, *A factorization for formal Laurent series and lattice path enumeration*, J. Combin. Theory Ser. A **28** (1980), no. 3, 321–337.
92. ———, *A noncommutative generalization and $q-$analog of the Lagrange inversion formula*, Transactions of the American Mathematical Society **257** (1980), no. 2, 455–482.
93. M. L. Glasser, *A Watson sum for a cubic lattice*, Journal of Mathematical Physics **13** (1972), 1145–1146.
94. C. D. Godsil, *Algebraic combinatorics*, Chapman and Hall, 1993.
95. Massimiliano Goldwurm and Massimo Santini, *Clique polynomials have a unique root of smallest modulus*, Information Processing Letters **75** (2000), no. 3, 127–132.
96. Ian P. Goulden and David M. Jackson, *Combinatorial enumeration*, John Wiley, New York, 1983.
97. Xavier Gourdon, *Largest component in random combinatorial structures*, Discrete Mathematics **180** (1998), no. 1-3, 185–209.
98. Xavier Gourdon and Bruno Salvy, *Asymptotics of linear recurrences with rational coefficients*, Tech. Report 1887, INRIA, March 1993, To appear in Proceedings FPACS'93.
99. Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete mathematics*, Addison Wesley, 1989.
100. D. H. Greene and D. E. Knuth, *Mathematics for the analysis of algorithms*, Birkhäuser, Boston, 1981.
101. Fabrice Guillemin, Philippe Robert, and Bert Zwart, *AIMD algorithms and exponential functionals*, In press, 2003, To appear in *Annals of Applied Probability*. 20 pages.
102. Jennie C. Hansen, *A functional central limit theorem for random mappings*, Annals of Probability **17** (1989), no. 1.
103. F. Harary, R. W. Robinson, and A. J. Schwenk, *Twenty-step algorithm for determining the asymptotic number of trees of various species*, Journal of the Australian Mathematical Society (Series A) **20** (1975), 483–503.
104. Frank Harary and Edgar M. Palmer, *Graphical enumeration*, Academic Press, 1973.
105. G. H. Hardy, *Ramanujan: Twelve lectures on subjects suggested by his life and work*, third ed., Chelsea Publishing Company, New-York, 1978, Reprinted and Corrected from the First Edition, Cambridge, 1940.
106. Erich Hecke, *Vorlesungen über die Theorie der algebraischen Zahlen*, Akademische Verlagsgesellschaft, Leipzig, 1923.
107. Peter Henrici, *Applied and computational complex analysis*, vol. 2, John Wiley, New York, 1974.
108. E. Hille, *Analytic function theory*, Blaisdell Publishing Company, Waltham, 1962, 2 Volumes.
109. Micha Hofri, *Analysis of algorithms: Computational methods and mathematical tools*, Oxford University Press, 1995.
110. J. A. Howell, T. F. Smith, and M. S. Waterman, *Computation of generating functions for biological molecules*, SIAM Journal on Applied Mathematics **39** (1980), no. 1, 119–133.
111. Iwan jensen, *A parallel algorithm for the enumeration of self-avoiding polygons on the square lattice*, Journal of Physics A: Mathematical and General **36** (2003), 5731–5745.
112. R. Jungen, *Sur les séries de Taylor n'ayant que des singularités algébrico-logarithmiques sur leur cercle de convergence*, Commentarii Mathematici Helvetici **3** (1931), 266–306.
113. Samuel Karlin and Howard Taylor, *A first course in stochastic processes*, second ed., Academic Press, 1975.
114. Frances Kirwan, *Complex algebraic curves*, London Mathematical Society Student Texts, no. 23, Cambridge University Press, 1992.
115. Arnold Knopfmacher and Helmut Prodinger, *On Carlitz compositions*, European Journal of Combinatorics **19** (1998), no. 5, 579–589.
116. John Knopfmacher and Arnold Knopfmacher, *Counting irreducible factors of polynomials over a finite field*, Discrete Mathematics **112** (1993), 103–118.
117. K. Knopp, *Theory of functions*, Dover Publications, New York, 1945.
118. Donald E. Knuth, *The art of computer programming*, vol. 1: Fundamental Algorithms, Addison-Wesley, 1968, Second edition, 1973.
119. ———, *The average time for carry propagation*, Indagationes Mathematicae **40** (1978), 238–242.
120. ———, *The art of computer programming*, 3rd ed., vol. 1: Fundamental Algorithms, Addison-Wesley, 1997.
121. ———, *The art of computer programming*, 2nd ed., vol. 3: Sorting and Searching, Addison-Wesley, 1998.
122. ———, *The art of computer programming*, 3rd ed., vol. 2: Seminumerical Algorithms, Addison-Wesley, 1998.
123. Donald E. Knuth and Arnold Schönhage, *The expected linearity of a simple equivalence algorithm*, Theoretical Computer Science **6** (1978), 281–315.
124. M. Krasnoselskii, *Positive solutions of operator equations*, P. Noordhoff, Groningen, 1964.
125. Christian Krattenthaler, *Advanced determinant calculus*, Seminaire Lotharingien de Combinatoire **42** (1999), Paper B42q, 66 pp.
126. Jacques Labelle and Yeong Nan Yeh, *Generalized Dyck paths*, Discrete Mathematics **82** (1990), 1–6.
127. J. C. Lagarias, A. M. Odlyzko, and D. B. Zagier, *On the capacity of disjointly shared networks*, Computer Netwoks **10** (1985), no. 5, 275–285.

128. Steven P. Lalley, *Finite range random walk on free groups and homogeneous trees*, Ann. Probab. **21** (1993), no. 4, 2087–2130.
129. Serge Lang, *Linear algebra*, Addison-Wesley, Reading, Mass., 1966.
130. Gregory F. Lawler, *Intersections of random walks*, Birkhäuser Boston Inc., Boston, MA, 1991.
131. Ernst Lindelöf, *Le calcul des résidus et ses applications à la théorie des fonctions*, Collection de monographies sur la théorie des fonctions, publiée sous la direction de M. Émile Borel, Gauthier-Villars, Paris, 1905, Reprinted by Gabay, Paris, 1989.
132. M. Lothaire, *Combinatorics on words*, Encyclopedia of Mathematics and its Applications, vol. 17, Addison–Wesley, 1983.
133. Guy Louchard, *The Brownian excursion: a numerical analysis*, Computers and Mathematics with Applications **10** (1984), no. 6, 413–417.
134. ———, *Random walks, Gaussian processes and list structures*, Theoretical Computer Science **53** (1987), no. 1, 99–124.
135. Guy Louchard and Helmut Prodinger, *Probabilistic analysis of Carlitz compositions*, Discrete Mathematics & Theoretical Computer Science **5** (2002), no. 1, 71–96.
136. Guy Louchard, R. Schott, M. Tolley, and P. Zimmermann, *Random walks, heat equations and distributed algorithms*, Journal of Computational and Applied Mathematics **53** (1994), 243–274.
137. Hosam M. Mahmoud, *Evolution of random search trees*, John Wiley, New York, 1992.
138. A. Meir and J. W. Moon, *On the altitude of nodes in random trees*, Canadian Journal of Mathematics **30** (1978), 997–1015.
139. ———, *On random mapping patterns*, Combinatorica **4** (1984), 61–70.
140. ———, *Erratum: "On an asymptotic method in enumeration"*, J. Combin. Theory Ser. A **52** (1989), no. 1, 163.
141. ———, *On an asymptotic method in enumeration*, J. Combin. Theory Ser. A **51** (1989), no. 1, 77–89.
142. ———, *The asymptotic behaviour of coefficients of powers of certain generating functions*, European Journal of Combinatorics **11** (1990), 581–587.
143. J. W. Moon, *Counting labelled trees*, Canadian Mathematical Monographs N.1, William Clowes and Sons, 1970.
144. Albert Nijenhuis and Herbert S. Wilf, *Combinatorial algorithms*, second ed., Academic Press, 1978.
145. Niels Erik Nörlund, *Vorlesungen über Differenzenrechnung*, Chelsea Publishing Company, New York, 1954.
146. A. M. Odlyzko, *Periodic oscillations of coefficients of power series that satisfy functional equations*, Advances in Mathematics **44** (1982), 180–205.
147. ———, *Asymptotic enumeration methods*, Handbook of Combinatorics (R. Graham, M. Grötschel, and L. Lovász, eds.), vol. II, Elsevier, Amsterdam, 1995, pp. 1063–1229.
148. A. M. Odlyzko and H. S. Wilf, *Bandwidths and profiles of trees*, Journal of Combinatorial Theory, Series B **42** (1987), 348–370.
149. ———, *The editor's corner: n coins in a fountain*, American Matematical Monthly **95** (1988), 840–843.
150. F. W. J. Olver, *Asymptotics and special functions*, Academic Press, 1974.
151. Richard Otter, *The number of trees*, Annals of Mathematics **49** (1948), no. 3, 583–599.
152. Daniel Panario, Boris Pittel, Bruce Richmond, and Alfredo Viola, *Analysis of Rabin's irreducibility test for polynomials over finite fields*, Random Structures Algorithms **19** (2001), no. 3-4, 525–551, Analysis of algorithms (Krynica Morska, 2000).
153. Daniel Panario and Bruce Richmond, *Analysis of Ben-Or's polynomial irreducibility test*, Proceedings of the Eighth International Conference "Random Structures and Algorithms" (Poznan, 1997), vol. 13, 1998, pp. 439–456.
154. ———, *Exact largest and smallest size of components*, Algorithmica **31** (2001), no. 3, 413–432.
155. R. B. Paris and D. Kaminski, *Asymptotics and Mellin-Barnes integrals*, Encyclopedia of Mathematics and its Applications, vol. 85, Cambridge University Press, Cambridge, 2001.
156. J. K. Percus, *Combinatorial Methods*, Applied Mathematical Sciences, vol. 4, Springer-Verlag, 1971.
157. Oskar Perron, *Über Matrizen*, Mathematische Annalen **64** (1907), 248–263.
158. G. Pólya, *Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen*, Acta Mathematica **68** (1937), 145–254.
159. G. Pólya and R. C. Read, *Combinatorial enumeration of groups, graphs and chemical compounds*, Springer Verlag, New York, 1987.
160. A. G. Postnikov, *Tauberian theory and its applications*, Proceedings of the Steklov Institute of Mathematics, vol. 144, American Mathematical Society, 1980.
161. E. M. Rains and N. J. A. Sloane, *On Cayley's enumeration of alkanes (or 4-valent trees)*, Journal of Integer Sequences **2** (1999), Article 99.1.1; avilable electronically.
162. Christoph Richard, *Scaling behaviour of two-dimensional polygon models*, Journal of Statistical Physics **108** (2002), no. 3/4, 459–493.
163. Philippe Robert, *Réseaux et files d'attente: méthodes probabilistes*, Mathématiques & Applications, vol. 35, Springer, Paris, 2000.
164. Bruno Salvy, *Asymptotique automatique et fonctions génératrices*, Ph. D. thesis, École Polytechnique, 1991.

165. Bruno Salvy and Paul Zimmermann, *GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable*, ACM Transactions on Mathematical Software **20** (1994), no. 2, 163–167.
166. Gilles Schaeffer, *Conjugaison d'arbres et cartes combinatoires aléatoires*, Ph.d. thesis, Université de Bordeaux I, December 1998.
167. William R. Schmitt and Michael S. Waterman, *Linear trees and RNA secondary structure*, Discrete Applied Mathematics. Combinatorial Algorithms, Optimization and Computer Science **51** (1994), no. 3, 317–323.
168. Robert Sedgewick and Philippe Flajolet, *An introduction to the analysis of algorithms*, Addison-Wesley Publishing Company, 1996.
169. L. A. Shepp and S. P. Lloyd, *Ordered cycle lengths in a random permutation*, Transactions of the American Mathematical Society **121** (1966), 340–357.
170. N. J. A. Sloane, *The on-line encyclopedia of integer sequences*, 2000, Published electronically at `http://www.research.att.com/~njas/sequences/`.
171. N. J. A. Sloane and Simon Plouffe, *The encyclopedia of integer sequences*, Academic Press, 1995.
172. Michèle Soria-Cousineau, *Méthodes d'analyse pour les constructions combinatoires et les algorithmes*, Doctorat ès sciences, Université de Paris–Sud, Orsay, July 1990.
173. George Springer, *Introduction to Riemann surfaces*, Addison-Wesley Publishing Company, Inc., Reading, Mass., 1957, Reprinted by Chelsea, New York.
174. Richard P. Stanley, *Generating functions*, Studies in Combinatorics, M.A.A. Studies in Mathematics, Vol. 17. (G-C. Rota, ed.), The Mathematical Association of America, 1978, pp. 100–141.
175. ______, *Enumerative combinatorics*, vol. I, Wadsworth & Brooks/Cole, 1986.
176. ______, *Enumerative combinatorics*, vol. II, Cambridge University Press, 1998.
177. P. R. Stein and M. S. Waterman, *On some new sequences generalizing the Catalan and Motzkin numbers*, Discrete Mathematics **26** (1979), no. 3, 261–272.
178. Gabor Szegő, *Orthogonal polynomials*, American Mathematical Society Colloquium Publications, vol. XXIII, A.M.S, Providence, 1989.
179. Wojciech Szpankowski, *Average-case analysis of algorithms on sequences*, John Wiley, New York, 2001.
180. Lajos Takács, *Conditional limit theorems for branching processes*, Journal of Applied Mathematics and Stochastic Analysis **4** (1991), no. 4, 263–292.
181. E. C. Titchmarsh, *The theory of functions*, second ed., Oxford University Press, 1939.
182. W. T. Tutte, *A census of planar maps*, Canadian Journal of Mathematics **15** (1963), 249–271.
183. ______, *On the enumeration of planar maps*, Bull. Amer. Math. Soc. **74** (1968), 64–74.
184. ______, *On the enumeration of four-colored maps*, SIAM Journal on Applied Mathematics **17** (1969), 454–460.
185. ______, *Planar enumeration*, Graph theory and combinatorics (Cambridge, 1983), Academic Press, London, 1984, pp. 315–319.
186. Joris van der Hoeven, *Majorants for formal power series*, Preprint, 2003, 29 pages. Available from author's webpage.
187. Robert Vein and Paul Dale, *Determinants and their applications in mathematical physics*, Applied Mathematical Sciences, vol. 134, Springer-Verlag, 1998.
188. Jeffrey Scott Vitter and Philippe Flajolet, *Analysis of algorithms and data structures*, Handbook of Theoretical Computer Science (J. van Leeuwen, ed.), vol. A: Algorithms and Complexity, North Holland, 1990, pp. 431–524.
189. Joachim von zur Gathen and Jürgen Gerhard, *Modern computer algebra*, Cambridge University Press, New York, 1999.
190. H. S. Wall, *Analytic theory of continued fractions*, Chelsea Publishing Company, 1948.
191. E. T. Whittaker and G. N. Watson, *A course of modern analysis*, fourth ed., Cambridge University Press, 1927, Reprinted 1973.
192. David Vernon Widder, *The Laplace transform*, Princeton University Press, 1941.
193. Herbert S. Wilf, *Generatingfunctionology*, Academic Press, 1990.
194. Roderick Wong, *Asymptotic approximations of integrals*, Academic Press, 1989.
195. Alan R. Woods, *Coloring rules for finite trees, and probabilities of monadic second order sentences*, Random Structures Algorithms **10** (1997), no. 4, 453–485.
196. Doron Zeilberger, *Symbol-crunching with the transfer-matrix method in order to count skinny physical creatures*, Integers **0** (2000), Paper A9, Published electronically at `http://www.integers-ejcnt.org/vol0.html`.

Index