ANALYTIC METHODS

HELMUT PRODINGER

Contents

1. Introduction

This chapter should have really been written by Philippe Flajolet (1948–2011), who could not make it this time. *He* coined the name *analytic combinatorics*. The present author knew Flajolet since 1979 and followed all the developments closely since then, also being a coauthor on various occasions. Flajolet and his followers started out in *analysis of algorithms*, a subject founded by Knuth in his series of books *The art of computer programming*; it became clear over the years that many techniques from classical mathematics had to be unearthed and many new ones had to be discovered. Flajolet was a pioneer in this direction; Doron Zeilberger called him a combinatorialist who became an analyst, and it was his understanding that combinatorics should have an analytic component, like number theory has analytic number theory. Flajolet, apart from being an exceptional problem solver, had a strong desire to be clear and systematic. Eventually, with coauthor Robert Sedgewick, after many years of preparations, the book *Analytic Combinatorics* [24] was published. It has 810 pages, and only a fraction of it can be represented here.

When analyzing algorithms, there is often an algebraic (combinatorial) part, followed by an asymptotic (analytic) part. And, indeed, analytic combinatorics, as understood by Flajolet, follows the same pattern. The central objects are generating functions. First, through combinatorial constructions (a bit reminiscent of grammars for formal languages), symbolic equations for the objects are obtained. There are rules how to translate the symbolic equations into equations for the associated generating functions. From here, there are a few typical scenarios. Ideally, one can write explicit expressions for the generating functions, and then get from it explicit expressions for the coefficients, which are the numbers of interest. But often, these expressions are involved, and one needs asymptotic techniques. It is often better to derive asymptotic equivalents directly from the generating functions. Sometimes, one cannot solve equations and thus has no explicit form for the generating functions. But even then there might still be hope. Asymptotic methods (in combinatorics) have been known for decades, and we might cite De Bruijn's book [5] and Odlyzko's treatise [40]. One of Flajolet's favorite methods was the Mellin transform. He wrote a series of survey papers about it, and a draft of some 100+ pages about it existed, but eventually did not make it into the book [24].

We follow these guidelines and include material here that – in one way or another – can be traced back to [24]. Whatever we cite from this great book we do with due respect. Even in places where it is not explicitly said, the concept and notations are probably taken from this authoritative text, since we firmly believe that it is vain to try to improve on the masters themselves.

The choice is perhaps a bit personal; during some 35 years as a working mathematician, I came accross many things that are useful and also necessary to know. Whatever is included here, was interesting to myself over the years. I hope that the selection presented here will be useful for the readers of this handbook as well.

Plan of this chapter. We start with combinatorial constructions, both, for unlabelled and labelled objects. On the way, we discuss various classical combinatorial objects, like compositions, partitions, trees, set partitions etc. Then we elaborate on techniques related to generating functions. Before we move to asymptotic considerations, we need a few preparations, like Gamma function, zeta function, harmonic numbers, etc. Then we move to the Mellin transform. Here, we start with so-called harmonic sums, move to digital sums and eventually to divide-and-conquer type recursions. Important here is the Mellin inversion formula and the Mellin-Perron summation formula. Then we discuss Rice's method, which is based on the Cauchy integral formula, and a running example that is both instructive and important: approximate counting. Then we sketch singularity analysis of generating functions, which is a toolkit allowing us to translate from the local behaviour of the generating functions around their (dominant) singularities to the asymptotic behaviour of its (Taylor-)coefficients. Another interesting running example is the one of longest runs in strings consisting of two symbols. Another important asymptotic technique, the saddle point method, is only sketched. The last section deals with Gaussian limiting distributions and how it can be obtained in an important family of special instances; these developments are due to H.-K. Hwang. This section must be seen as an appetizer; it will increase the desire to read more comprehensive texts about limiting distributions in combinatorial analysis.

Let us collect a few conventions and notations and facts that will be encountered later:

Iverson's notation. We write $[P]=1$ when condition P is true, $[P]=0$ otherwise. This notation is more flexible, say, than Kronecker's $\delta_{i,j}=[i=j]$.

Probability generating functions and moments. Assume that the power series $f(z)$ has non-negative coefficients and $f(1)=1$ (a probability generating function). Then expectation and variance (when they exist) are given by

$$\mathbb{E}=f'(1), \qquad \mathbb{V}=f''(1)+f'(1)-\big(f'(1)\big)^2.$$

Coefficient extraction of power series. If $f(z)=\sum_{n\ge0}a_nz^n$, then $[z^n]f(z)=a_n$.

2. Combinatorial constructions and associated ordinary generating functions

Let $\mathcal{A}$ be a denumerable set and a *size* associated to each of its elements, which is a non-negative number, often written as $|a|$ for $a \in \mathcal{A}$. Furthermore, we assume that, for each n, there is only a finite number of elements of size n in $\mathcal{A}$; we call it a_n. Then we define the associated (ordinary) generating function

$$A(z) = \sum_{n \ge 0} a_n z^n.$$

For the time being, this is just a formal construction, but later on, we will interpret $A(z)$ as a function of a complex variable z. This will be particularly useful when we discuss how to get asymptotic equivalents for the numbers a_n. At the moment, we are just interested in combinatorial constructions, and how they are reflected by their associated generating functions. Such constructions allow to create, starting from basic objects, more complicated ones. Another, but equivalent point of view is decomposition, where a certain combinatorial class of objects is (uniquely) decomposed into simpler ingredients. Combinatorial constructions are also relevant in symbolic algebra systems.

The most common constructions are now discussed.

Union. Let $\mathcal{B}$ and $\mathcal{C}$ be such combinatorial classes with associated generating functions $B(z)$ and $C(z)$ and assume that the classes are mutually disjoint. Then the (disjoint) union $\mathcal{A} = \mathcal{B} + \mathcal{C}$, also written as $\mathcal{A} = \mathcal{B} \cup \mathcal{C}$, has associated generating function $A(z) = B(z) + C(z)$. This follows from the elementary $a_n = b_n + c_n$.

Product. Let $\mathcal{B}$ and $\mathcal{C}$ be such combinatorial classes with associated generating functions $B(z)$ and $C(z)$. Then we form the (cartesian) product $\mathcal{A} = \mathcal{B} \times \mathcal{C}$. The size of an object (b, c) is defined to be $|b| + |c|$. Then the associated generating function is $A(z) = B(z)C(z)$. This follows from the fact that

$$a_n = \sum_{k=0}^{n} b_k c_{n-k},$$

and this is just the Cauchy product of two series:

$$B(z)C(z) = \sum_{m \ge 0} b_m z^m \cdot \sum_{n \ge 0} c_n z^n = \sum_{n \ge 0} \left(\sum_{k=0}^{n} b_k c_{n-k} \right) z^n = A(z).$$

This idea immediately extends to several factors, not just two. In particular, for a fixed number k, we can consider k-tuples $(x_1, \dots, x_k)$, where all $x_i \in \mathcal{A}$, and $|(x_1, \dots, x_k)| = |x_1| + \cdots + |x_k|$. Then the generating function associated to $\mathcal{A}^k$ is $A^k(z)$.

SEQUENCE. Let $\mathcal{B}$ be a combinatorial class, that does not contain elements of size 0. Then we form

$$\mathcal{A} = \mathcal{B}^0 + \mathcal{B}^1 + \mathcal{B}^2 + \cdots ,$$

which describes sequences of elements of $\mathcal{B}$. The (unique) sequence of zero elements is traditionally written as ε. We write $\mathcal{A} = \textsc{Seq}(\mathcal{B})$, and the associated generating function is

$$A(z) = 1 + B(z) + B^2(z) + \cdots = \frac{1}{1 - B(z)}.$$

Especially when dealing with languages (sets of words), the notion $\mathcal{B}^*$ instead of $\textsc{Seq}(\mathcal{B})$ is common; then also $\mathcal{B}^+ = \mathcal{B}^1 + \mathcal{B}^2 + \cdots$.

POWER SET. For a given $\mathcal{B}$, we form finite sets of elements taken from $\mathcal{B}$; the result is $\mathcal{A} = \textsc{PSet}(\mathcal{B})$, and the size of such a set is defined to be the sum of the sizes of its elements. We must assume that $\mathcal{B}$ does not contain an element of size 0. We have

$$\mathcal{A} \equiv (\varepsilon + \{\beta_1\}) \times (\varepsilon + \{\beta_2\}) \times \cdots$$

for an enumeration $(\beta_1, \beta_2, \dots)$ of the class $\mathcal{B}$. Now let

$$B(z) = \sum_{n\ge1} B_n z^n,$$

then we can compute

$$\begin{aligned} A(z) &= \prod_{\beta\in\mathcal{B}} (1 + z^{|\beta|}) = \prod_{n\ge1} (1 + z^n)^{B_n} \\ &= \exp\Big(\sum_{n\ge1} B_n \log(1 + z^n)\Big) = \exp\Big(\sum_{n\ge1} B_n \sum_{k\ge1} \frac{(-1)^{k-1} z^{nk}}{k}\Big) \\ &= \exp\Big(\frac{B(z)}{1} - \frac{B(z^2)}{2} + \frac{B(z^3)}{3} - \cdots\Big). \end{aligned}$$

The operations with infinite series are justified, since, for given n, only a finite number of B_j's contribute to A_n.

MULTISET. This is very similar to forming sets, but now repeated elements are allowed, leading to multisets. The computation is similar:

$$\begin{aligned} A(z) &= \prod_{n\ge1} (1 + z^n + z^{2n} + z^{3n} + \cdots)^{B_n} \\ &= \exp\Big(\sum_{n\ge1} B_n \log\frac{1}{1 - z^n}\Big) = \exp\Big(\sum_{n\ge1} B_n \sum_{k\ge1} \frac{z^{nk}}{k}\Big) \\ &= \exp\Big(\frac{B(z)}{1} + \frac{B(z^2)}{2} + \frac{B(z^3)}{3} + \cdots\Big). \end{aligned}$$

CYCLES. For a given $\mathcal{B}$, we form cycles of elements taken from $\mathcal{B}$, where again $B_0 = 0$. A cycle is $(b_1, \dots, b_n)$ with $b_i \in \mathcal{B}$. It is identified with all cyclic rotations. So, for example, $(a,b,a,b) = (b,a,b,a)$, but (a,a,b,b) is a different cycle. Again, the size of a cycle is the sum of the sizes of its elements. Then, for the associated generating functions,

$$A(z) = \sum_{k\ge1} \frac{\phi(k)}{k} \log \frac{1}{1-B(z^k)},$$

where $\phi(k)$ is Euler's totient function, in other words, the number of i's less than k which are relatively prime to k.

The proof of this relation will not be given; it is usually done using Polya's enumeration theory (enumeration under group action, here just the cyclic group).

Constructions under restrictions are also considered, for example $\mathrm{SEQ}_k(\mathcal{B})$, $\mathrm{SEQ}_{\ge k}(\mathcal{B})$ of sequences with exactly k or $\ge k$ elements, and various others.

Further constructions will be introduced in this text when they occur.

Now we turn to a few examples.

COMPOSITIONS. A composition of a positive integer n is a representation $n = i_1 + \cdots + i_k$ with positive integers i_j; the number k is referred to as the number of parts. We can interpret the integers $\mathcal{I} = \{1, 2, \dots\} = \{\bullet, \bullet\bullet, \bullet\bullet\bullet, \dots\} = \mathrm{SEQ}_{\ge1}\{\bullet\}$. The size of integer $i \cong \bullet^i$ is just i, and so

$$I(z) = \frac{z}{1-z}.$$

Further, compositions are described by $\mathcal{C} = \mathrm{SEQ}_{\ge1}(\mathcal{I})$, whence

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

and I_n, the number of compositions of n, is given by $I_n = 2^{n-1}$, which is also easy to see directly.

PARTITIONS. They are defined like compositions, except that the order of the terms ("parts") is irrelevant. They can be seen as multisets of $\mathcal{I}$; the multiset construction then gives us

$$P(z) = \exp\Big(I(z) + \frac{I(z^2)}{2} + \frac{I(z^3)}{3} + \cdots\Big).$$

This form is, however, not very useful; the more natural way to write this is

$$P(z) = \prod_{n\ge1} \frac{1}{1-z^n};$$

on expanding the product, a typical term is $z^{i_1+2i_2+3i_3+\cdots}$, which just describes a partition (i_1 ones, i_2 twos, i_3 threes, etc.) The size is the sum of the parts. There will be a separate section providing the very basic elements of the extremely rich and useful theory of partitions.

Some families of trees. The class of binary trees $\mathcal{B}$ is either an external node or a root (an internal node) followed by a left and a right subtree, both again binary trees. This recursive definition can be stated as $\mathcal{B} = \Box + \circ \cdot \mathcal{B} \cdot \mathcal{B}$. External nodes are not counted when one speaks about size, and sometimes not drawn. The equation

$$B(z) = 1 + zB^2(z)$$

is immediate, leading to

$$B(z) = \frac{1-\sqrt{1-4z}}{2z} = \sum_{n\ge 0} \frac{1}{n+1}\binom{2n}{n} z^n,$$

so that binary trees with n (internal) nodes are enumerated by Catalan numbers $\frac{1}{n+1}\binom{2n}{n}$.

The generalization to t-ary trees is immediate: the root has an ordered list of t subtrees, and $t = 2$ means binary trees.

Another important family is $\mathcal{P}$, the family of planar trees. They are known under several names: planted plane trees, plane trees, ordered trees, ... There is a root node and a sequence of r planar subtrees ($r \ge 0$). Thus $\mathcal{P} = \circ + \circ \cdot \mathcal{P} + \circ \cdot \mathcal{P} \cdot \mathcal{P} + \circ \cdot \mathcal{P} \cdot \mathcal{P} \cdot \mathcal{P} + \cdots$ Consequently

$$\begin{aligned} P(z) &= zP(z) + zP^2(z) + zP^3(z) + \cdots = \frac{z}{1-P(z)} \\ &= \frac{1-\sqrt{1-4z}}{2} = \sum_{n\ge 1} \frac{1}{n}\binom{2n-2}{n-1} z^n. \end{aligned}$$

There are two standard bijections: Between planar trees and binary trees the rotation correspondence. Take a planar tree, let only the leftmost edge survive, connect siblings instead, cut off the root and turn the tree by 45° to obtain a binary tree with one node less. This is reversible. This construction is described in more detail in many textbooks. Here is an example:

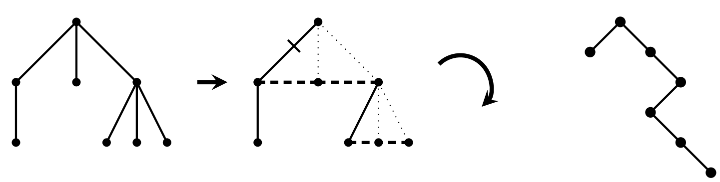


Figure 1. A planar tree with 8 nodes (=7 edges) and the corresponding binary tree with 7 internal nodes.

Planar trees are also in bijection with non-negative lattice path (Dyck paths) which are described a little later in this text. Here is already an example.

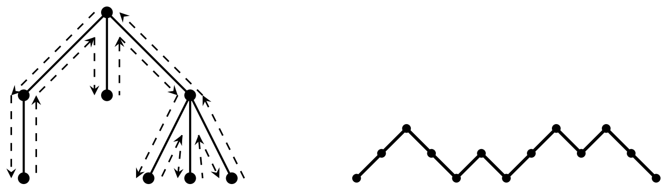


FIGURE 2. A planar tree with 8 nodes (=7 edges) and the corresponding Dyck path of length 12 (=semi-length 6)

As this example shows, one just walks around the tree and records the steps (up or down) in a diagramm.

Much more about these and other families of trees can be found in this handbook in various chapters that specialize on trees.

SET PARTITIONS. A partition of the set $\{1, 2, \dots, n\}$ into k blocks consists of k nonempty subsets which are mutally disjoint, and their union is the full set. The number of them is denoted by $\left\{{n \atop k}\right\}$, a Stirling subset number (in the older literature Stirling numbers of the second kind). In order to describe these set partitions symbolically, we call the smallest element in each block the leader. Then we order the blocks according to the block leaders in ascending order. In this way, it makes sense to talk about block 1, block 2, ..., block k. And now we write a string $b_1 b_2 \dots b_n$, where b_i is the number of the block in which i lies (its address). Such a string has the properties that $b_1 = 1$, and if a new number appears for the first time (scanning the string from left to right), it is one higher than the previous highest number, and altogether all the numbers $1, \dots, k$ appear. Example. The string 112122313241 codes the set partition $\{1, 2, 4, 8, 12\}, \{3, 5, 6, 10\}, \{7, 9\}, \{11\}$. The set of admissible strings admits the following representation:

$$1\,\textsc{Seq}(1)\,2\,\textsc{Seq}(1+2)\,3\,\textsc{Seq}(1+2+3)\dots k\,\textsc{Seq}(1+\dots+k).$$

This translates into the generating function

$$S^{(k)}(z) = \frac{z}{1-z}\frac{z}{1-2z}\cdots\frac{z}{1-kz};$$

partial fraction decomposition gives

$$S^{(k)}(z) = \frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}\frac{1}{1-jz}, \text{ so } \left\{{n \atop k}\right\} = \frac{1}{k!}\sum_{j=1}^{k}\binom{k}{j}(-1)^{k-j}j^n.$$

3. Combinatorial constructions and associated exponential generating functions

We are now discussing labelled classes. The idea is as follows. In the previous section, the function “size” was abstractly introduced. Now, if an object has size n, we assume that there are n atoms present, and we label them. Here is a precise definition.

Definition 3.1. *A weakly labelled object of size n is a graph whose set of vertices is a subset of the integers. Equivalently, we say that the vertices bear labels, with the implied condition that labels are distinct integers from $\mathbb{Z}$. An object of size n is said to be well-labelled, or, simply, labelled, if it is weakly labelled and, in addition, its collection of labels is the complete integer interval $[1..n]$. A labelled class is a combinatorial class comprised of well-labelled objects.*

We will also use reduction of labels: For a weakly labelled structure of size n, this operation reduces its labels to the standard interval $[1..n]$ while preserving the relative order of labels. For instance, the sequence $\langle 7,3,9,2\rangle$ reduces to $\langle 3,2,4,1\rangle$. We use $\rho(\alpha)$ to denote this canonical reduction of the structure α.

In order to count labelled objects, we appeal to exponential generating functions. The exponential generating function of a sequence (A_n) is the formal power series

$$A(z) = \sum_{n\geq 0} A_n \frac{z^n}{n!}.$$

The exponential generating function of a class $\mathcal{A}$ is the exponential generating function of the numbers $A_n = \mathrm{card}(\mathcal{A}_n)$. It is also said that the variable z marks the size in the generating function. With the standard notation for coefficients of series, the coefficient A_n in an exponential generating function is then recovered by

$$A_n = n![z^n]A(z),$$

since $[z^n]A(z) = A_n/n!$.

Neutral and atomic classes. It proves useful to introduce a neutral (empty, null) object ε that has size 0 and bears no label at all, and consider it as a special labelled object; a neutral class $\mathcal{E}$ is then by definition $\mathcal{E} = \varepsilon$ and is also denoted by $\mathbf{1}$. The (labelled) atomic class $\mathcal{Z} = ①$ is formed of a unique object of size 1 that, being well-labelled, bears the integer label ①. The exponential generating functions of the neutral class and the atomic class are, respectively, $E(z) = 1$, $Z(z) = z$.

Labelled product. The labelled product of $\mathcal{B}$ and $\mathcal{C}$, denoted $\mathcal{B}\star\mathcal{C}$, is obtained by forming ordered pairs from $\mathcal{B}\times\mathcal{C}$ and performing all possible order-consistent relabellings. When $\mathcal{A} = \mathcal{B}\star\mathcal{C}$, the corresponding counting sequences satisfy

$$A_n = \sum_{k=0}^{n} \binom{n}{k} B_k C_{n-k};$$

the binomial coefficients count the relabellings. The new object has size $k + (n - k) = n$; k of the numbers $\{1, \dots, n\}$ are selected for the labels of the first component, $n - k$ for the second. But this is just the way exponential generating functions are multiplied:

$$\sum_{n\ge0} A_n \frac{z^n}{n!} = \sum_{n\ge0} B_n \frac{z^n}{n!} \cdot \sum_{n\ge0} C_n \frac{z^n}{n!}.$$

Sequences. General sequences and sequences with k factors can be formed as before. Here and in the following we assume again that $\varepsilon \notin \mathcal{B}$.

$$\textsc{Seq}(\mathcal{B}) = \{\varepsilon\} + \mathcal{B} + \mathcal{B} \star \mathcal{B} + \mathcal{B} \star \mathcal{B} \star \mathcal{B} + \cdots = \bigcup_{k\ge0} \textsc{Seq}_k(\mathcal{B}).$$

The translations into exponential generating functions are

$$A(z) = B(z)^k \quad \text{and} \quad A(z) = \frac{1}{1 - B(z)},$$

respectively.

Sets. We denote by $\textsc{Set}_k(\mathcal{B})$ the class of k-sets formed from $\mathcal{B}$. The set class is defined formally as the quotient $\textsc{Set}_k(\mathcal{B}) := \textsc{Seq}_k(\mathcal{B})/\boldsymbol{R}$, where the equivalence relation $\boldsymbol{R}$ identifies two sequences when the components of one are a permutation of the components of the other. A "set" is like a sequence, but the order between components is immaterial. The (labelled) set construction applied to $\mathcal{B}$, denoted $\textsc{Set}(\mathcal{B})$, is then defined by

$$\textsc{Set}(\mathcal{B}) = \bigcup_{k\ge0} \textsc{Set}_k(\mathcal{B}).$$

The translations into exponential generating functions are

$$A(z) = \frac{B(z)^k}{k!} \quad \text{and} \quad A(z) = \sum_{k\ge0} \frac{B(z)^k}{k!} = \exp(B(z)),$$

respectively.

Cycles. We start with k-cycles. The class of k-cycles, $\textsc{Cyc}_k(\mathcal{B})$ is formally defined to be the quotient $\textsc{Cyc}_k(\mathcal{B}) := \textsc{Set}_k(\mathcal{B})/\boldsymbol{S}$, where the equivalence relation $\boldsymbol{S}$ identifies two sequences when the components are one cyclic permutation of the components of each other. A cycle is like a sequence whose components can be cyclically shifted, so that there is now a uniform k-to-one correspondence between k-sequences and k-cycles. We assume that $\mathcal{B} \neq \emptyset$ and $k \geq 1$. Then

$$\begin{aligned} \mathcal{A} = \textsc{Cyc}_k(\mathcal{B}) \quad &\Longrightarrow \quad A(z) = \frac{1}{k} B(z)^k, \\ \mathcal{A} = \textsc{Cyc}(\mathcal{B}) \quad &\Longrightarrow \quad A(z) = \sum_{k \geq 1} \frac{1}{k} B(z)^k = \log \frac{1}{1 - B(z)}. \end{aligned}$$

In the sequel we describe a few important combinatorial objects as families of labelled objects.

Surjections. Fix some integer $r \geq 1$ and let $\mathcal{R}_n^{(r)}$ denote the class of all surjections from the set $\{1, \ldots, n\}$ onto $\{1, \ldots, r\}$ whose elements are called r-surjections. We set $\mathcal{R}^{(r)} = \bigcup_{n \geq 1} \mathcal{R}_n^{(r)}$ and compute the corresponding exponential generating function, $R^{(r)}(z)$. We observe that an r-surjection $\phi \in \mathcal{R}_n^{(r)}$ is determined by the ordered r-tuple formed from the collection of all preimage sets, $(\phi^{-1}(1), \phi^{-1}(2), \ldots, \phi^{-1}(r))$; they are disjoint non-empty sets of integers that cover the interval $[1..n]$. One has the combinatorial specification

$$\mathcal{R}^{(r)} = \textsc{Seq}_r(\mathcal{V}), \quad \mathcal{V} = \textsc{Set}_{\geq 1}(\mathcal{Z}),$$

(a surjection is a sequence of non-empty sets), from which we conclude $R^{(r)}(z) = (e^z - 1)^r$. From this, we find also

$$\begin{aligned} \mathcal{R}_n^{(r)} = n![z^n](e^z - 1)^r &= n![z^n] \sum_{j=0}^{r} \binom{r}{j} (-1)^{r-j} e^{jz} \\ &= \sum_{j=0}^{r} \binom{r}{j} (-1)^{r-j} j^n = r! \left\{ {n \atop r} \right\}. \end{aligned}$$

Set partitions into r blocks. Let $\mathcal{S}_n^{(r)}$ denote the number of ways of partitioning the set $\{1, \ldots, n\}$ into r disjoint and non-empty equivalence classes (blocks). We set $\mathcal{S}^{(r)} = \bigcup_{n \geq 1} \mathcal{S}_n^{(r)}$; the corresponding objects are called set partitions, as defined already earlier. The enumeration problem for set partitions is closely related to that of surjections:

$$\mathcal{S}^{(r)} = \textsc{Set}_r(\mathcal{V}), \quad \mathcal{V} = \textsc{Set}_{\geq 1}(\mathcal{Z}) \quad \Longrightarrow \quad S^{(r)}(z) = \frac{(e^z - 1)^r}{r!}.$$

Thus we find again the formula for the Stirling set partition numbers:

$$\left\{ {n \atop r} \right\} = \frac{1}{r!} \sum_{j=1}^{r} \binom{r}{j} (-1)^{r-j} j^n.$$

Talking about *all* surjections resp. set partitions just means to sum over r. This leads to

$$R(z)=\sum_{r\ge0}(e^z-1)^r=\frac{1}{2-e^z}$$

and

$$S(z)=\sum_{r\ge0}\frac{(e^z-1)^r}{r!}=e^{e^z-1}.$$

The numbers $R_n=n![z^n]R(z)$ and $S_n=n![z^n]S(z)$ are called surjection numbers resp. Bell numbers. Clearly,

$$R_n=\sum_{r\ge0}r!\left\{{n \atop r}\right\}\quad\text{and}\quad S_n=\sum_{r\ge0}\left\{{n \atop r}\right\}.$$

We have

$$R(z)=\frac12\frac{1}{1-\frac{e^z}{2}}=\sum_{l\ge0}\frac{e^{lz}}{2^{l+1}},$$

therefore

$$R_n=\sum_{l\ge0}\frac{l^n}{2^{l+1}}.$$

Similarly,

$$S(z)=\frac1e e^{e^z}=\frac1e\sum_{j\ge0}\frac{e^{jz}}{j!},$$

whence

$$S_n=n![z^n]\frac1e\sum_{j\ge0}\frac{e^{jz}}{j!}=\frac1e\sum_{j\ge0}\frac{j^n}{j!}.$$

This is known as Dobinski's formula.

The present approach is also flexible with respect to restrictions. For example, $\exp(e_b(z)-1)$, with the truncated exponential series

$$e_b(z):=1+z+\frac{z^2}{2!}+\cdots+\frac{z^b}{b!}$$

corresponds to partitions with all blocks of size $\le b$, e^{e^z-1-z} corresponds to partitions with no singletons, and $\cosh(e^z-1)$ to partitions with an even number of blocks.

Restricted words and random allocation. Consider an alphabet with r letters, say, $\{a_1,\dots,a_r\}$. For a word of length n, the sequence $\{$set of indices of letter $a_1\}$, …, $\{$set of indices of letter $a_r\}$ is forming an "ordered" partition of the sets of labels $\{1,\dots,n\}$; without restrictions, this yields the exponential generating function $(e^z)^r$.

Now we want to determine the exponential generating function of all words where all letters appear at least b times. For that, we use again the truncated exponential series $e_b(z)$. Then we get $(e^z - e_{b-1}(z))^r$ as an answer. Observe that this is clear if the alphabet has just one letter, and the concept of exponential generating functions takes automatically care of the mixing of letters, whence the rth power. Variations of this also work, like: all letters appear at most b times leads to $e_b(z)^r$, and all kinds of restrictions can be handled.

Now we consider a balls-in-bins model. Throw at random n distinguishable balls into m distinguishable bins. We might think of the balls numbered from 1 to n. Each bin corresponds to one of m letters, and each realization of the experiment is coded by a word of length n. Let Min and Max represent the size of the least filled and most filled bins, respectively. Then

$$\mathbb{P}(\textsc{Max} \leq b) = \frac{n!}{m^n}[z^n]e_b(z)^m = n![z^n]e_b(\tfrac{z}{m})^m$$

and

$$\mathbb{P}(\textsc{Min} > b) = n![z^n]\big(e^{z/m} - e_b(\tfrac{z}{m})\big)^m.$$

Birthday paradox. This is a classical example: Assume that there is a line of persons entering a large room one by one. Each person is let in and declares her birthday upon entering the room. How many people must enter in order to find two that have the same birthday? The birthday paradox is the counterintuitive fact that on average a birthday collision is likely to take place as early as at time $n \approx 24$. Let B be the time of the first collision, which is a random variable ranging between 2 and $r+1$ (where the upper bound is derived from the pigeonhole principle; we assume that the year has r days; $\mathcal{X}$ denotes an alphabet with r letters). A collision has not yet occurred at time n, if the sequence of birthdays $\beta_1, \ldots, \beta_n$ has no repetition. In other words, the function β from $[1..n]$ to $\mathcal{X}$ must be injective; equivalently, $\beta_1, \ldots, \beta_n$ is an n-arrangement of r objects (=r ordered objects). Thus, we have the fundamental relation

$$\mathbb{P}(B > n) = \frac{r(r-1)\ldots(r-n+1)}{r^n} = \frac{n!}{r^n}[z^n](1+z)^r = n![z^n]\Big(1+\frac{z}{r}\Big)^r.$$

The expectation of the random variable B is

$$\mathbb{E}(B) = \sum_{n \geq 0} \mathbb{P}(B > n) = 1 + \sum_{n=1}^{r} \frac{r(r-1)\ldots(r-n+1)}{r^n}.$$

An alternative form of the expectation is now derived, which easily leads to generalizations. Let $f(z) = \sum_n f_n z^n$ be an entire function with nonnegative coefficients. Then

$$\sum_n f_n n! = \sum_n f_n \int_0^\infty e^{-t} t^n dt = \int_0^\infty f(t) e^{-t} dt$$

(a Laplace transform). Therefore

$$\mathbb{E}(B) = \int_0^\infty e^{-t}\Big(1+\frac{t}{r}\Big)^r dt.$$

Exactly the same reasoning leads to the following: the expected time necessary for the first occurrence of the event "b persons have the same birthday" has expectation given by the integral

$$\int_0^\infty e^{-t}\Big(e_{b-1}\Big(\frac{t}{r}\Big)\Big)^r dt,$$

where the classical case means $b=2$.

COUPON COLLECTOR. This problem is dual to the birthday paradox. We ask for the first time C when $\beta_1,\dots,\beta_C$ contains all the elements of $\mathcal{X}$; that is, all the possible birthdays have been "collected." In other words, the event $\{C \le n\}$ means the equality between sets, $\{\beta_1,\dots,\beta_n\} = \mathcal{X}$. Thus, the probabilities satisfy

$$\mathbb{P}\{C \le n\} = \frac{r!\left\{ n \atop r \right\}}{r^n} = \frac{n!}{r^n}[z^n](e^z-1)^r = n![z^n](e^{z/r}-1)^r.$$

The complementary probabilities are then

$$\mathbb{P}\{C > n\} = n![z^n]\Big(e^z - (e^{z/r}-1)^r\Big).$$

In the same style as before we get

$$\mathbb{E}(C) = \int_0^\infty \Big(1-(1-e^{-t/r})^r\Big)dt = r\sum_{j=1}^r \binom{r}{j}\frac{(-1)^{j-1}}{j}.$$

Alternatively, we might substitute $v = 1-e^{-t/r}$, then expand and integrate termwise; this process provides the answer in the form rH_r, with harmonic numbers $H_n = 1+\frac12+\dots+\frac1n$. More on these numbers will appear later. This answer can also be obtained in an elementary fashion: To get the first copy, you need on average 1 drawing, to get a second one needs $r/(r-1)$, a third one needs $r/(r-2)$, and so on.

The symbolic approach (leading to an integral) has the advantage of straight-forward generalizations. For instance, the expected time till each coupon is obtained b times is

$$\int_0^\infty \Big(1-(1-e_{b-1}(\tfrac{t}{r})e^{-t/r})^r\Big)dt.$$

PERMUTATIONS AND CYCLES. It is known that a permutation admits a unique decomposition into cycles: Let σ be a permutation. Start with any element, say 1, and draw a directed edge from 1 to $\sigma(1)$, then continue connecting to $\sigma^2(1)$, $\sigma^3(1)$, and so on; a cycle containing 1 is obtained after at most n steps. If one repeats the construction, taking at each stage an element not yet connected to earlier ones, the cycle decomposition of the permutation σ is obtained. This argument shows that the class of sets-of-cycles is isomorphic to the class of permutations:

$$\mathcal{P} \cong \textsc{Set}(\textsc{Cyc}(\mathcal{Z})) \cong \textsc{Seq}(\mathcal{Z}).$$

This combinatorial isomorphism is reflected by the obvious series identity

$$P(z) = \exp\Big(\log\frac{1}{1-z}\Big) = \frac{1}{1-z}.$$

The advantage of it is that restrictions are handled in an almost automatic fashion:

The class $\mathcal{P}^{(A,B)}$ of permutations with cycle lengths in $A \subseteq \mathbb{N}$ and with cycle numbers that belongs to $B \subseteq \mathbb{N}_0$ has exponential generating function

$$P^{(A,B)}(z) = \beta(\alpha(z)) \quad \text{with} \quad \alpha(z) = \sum_{a\in A}\frac{z^a}{a}, \quad \beta(z) = \sum_{b\in B}\frac{z^b}{b!}.$$

A popular instance is derangements (fix-point free permutations). The restriction is that no cycles of length one are allowed, therefore

$$\alpha(z) = \sum_{a\ge 2}\frac{z^a}{a} = \log\frac{1}{1-z} - z, \quad \beta(z) = \sum_{b\ge 0}\frac{z^b}{b!} = e^z,$$

leading to

$$\exp\Big(\log\frac{1}{1-z} - z\Big) = \frac{e^{-z}}{1-z};$$

this produces the number of derangements as

$$n![z^n]\frac{e^{-z}}{1-z} = n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}.$$

Notice that the probability that a random permutation of n elements has no fix-points is very close (for large n) to $\frac{1}{e}$, which is a popular result.

Stirling cycle numbers. The class $\mathcal{P}^{(r)}$ of permutations that decompose into r cycles, can be represented as

$$\mathcal{P}^{(r)} = \textsc{Set}_r(\textsc{Cyc}(\mathcal{Z})),$$

which leads to

$$P^{(r)}(z) = \frac{1}{r!}\Big(\log\frac{1}{1-z}\Big)^r.$$

Therefore we get

$$P_n^{(r)} = \begin{bmatrix} n \\ r \end{bmatrix} = \frac{n!}{r!}[z^n]\Big(\log\frac{1}{1-z}\Big)^r.$$

These numbers are called Stirling cycle numbers; in the older literature, they are often called sign-less Stirling numbers of the first kind. This is a somewhat strange name, so the new name should be favoured as it makes much more sense.

4. Partitions and q-series

Partitions have already appeared briefly before. Here, we want to describe them in more detail.

A partition of a positive integer n is a representation $n = i_1 + i_2 + \cdots + i_k$ with integers $1 \le i_1 \le i_2 \le \cdots$. The i_j's are called parts and k is the number of parts. So partitions can be described by the formal expression

$$\mathbf{1}^*\mathbf{2}^*\mathbf{3}^*\dots,$$

since $n = 1 \cdot j_1 + 2 \cdot j_2 + \cdots$ with integers $j_s \ge 0$. If we denote $p(n)$ the number of partitions of n and

$$P(q) = 1 + \sum_{n \ge 1} p(n) q^n$$

the (ordinary) generating function of partitions, then we get immediately from the formal expression that

$$P(q) = \prod_{i \ge 1} \frac{1}{1 - q^i}.$$

Recall that the star '$*$' is a handy alternative for the construction 'Seq' introduced earlier. Note [1] that in the context of partitions it is customary to use the variable q instead of z in generating functions. Andrews' encyclopedic book contains all of this, and much more. We also introduce $p(0) = 1$ to have smoother expressions. Let us fix some notation (we assume $|q| < 1$):

$$(q)_n = (q;q)_n = (1-q)(1-q^2)\dots(1-q^n),$$
$$(q)_\infty = (q;q)_\infty = (1-q)(1-q^2)\dots,$$

and, more generally,

$$(x)_n = (x;q)_n = (1-x)(1-xq)\dots(1-xq^{n-1}),$$
$$(x)_\infty = (x;q)_\infty = (1-x)(1-xq)\dots.$$

So we have $P(q) = 1/(q)_\infty$, and also the number of partitions of n into k parts is given as

$$[q^n t^k] \prod_{i \ge 1} \frac{1}{1 - tq^i} = [q^n t^k] \frac{1}{(tq;q)_\infty}.$$

There is a graphical representation of a partition, called a Ferrers diagram. One simply codes a part k as a row of k unit squares and arranges them in decreasing order. Upon reflecting the diagram at the diagonal, we get another partition, called conjugate partition.

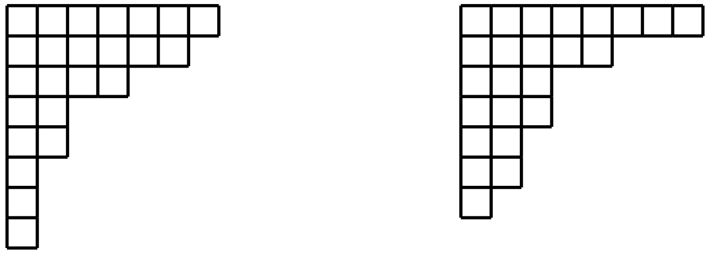


FIGURE 3. The partition $(7, 6, 4, 2, 2, 1, 1, 1)$ and its conjugate $(8, 5, 3, 3, 2, 2, 1)$

The number of partitions of n where the parts are of size $\leq k$ is

$$[q^n]\frac{1}{(q)_k},$$

and using the concept of conjugate partitions, this number also equals the number of partitions of n where the number of parts are $\leq k$.

The following theorem is very basic for the manipulation of "q-series."

Theorem 4.1. *We have*

$$F(t) = \frac{(at)_\infty}{(t)_\infty} = \sum_{n\geq 0} \frac{(a)_n}{(q)_n} t^n.$$

This theorem is attributed to Cauchy in [1] and often called the q-binomial theorem.

Proof. Splitting off the first factors in the product, we get

$$F(t) = \frac{1-at}{1-t}F(qt), \quad \text{or} \quad (1-t)F(t) = (1-at)F(qt).$$

Writing $F(t) = \sum_{n\geq 0} A_n t^n$ and comparing coefficients of t^n, we get

$$A_n - A_{n-1} = q^n A_n - aq^{n-1}A_{n-1}, \quad \text{or} \quad A_n = \frac{1-aq^{n-1}}{1-q^n}A_{n-1}.$$

Since $A_0 = F(0) = 1$, iteration of this recursion results in

$$A_n = \frac{(a)_n}{(q)_n},$$

as claimed. □

The first important special case arises from setting $a = 0$:

$$\frac{1}{(t)_\infty} = \prod_{n\geq 0}\frac{1}{1-tq^n} = \sum_{n\geq 0}\frac{t^n}{(q)_n}. \tag{1}$$

It is called Euler's partition identity, because replacing t by tq results in

$$\frac{1}{(qt)_\infty} = \sum_{n\ge 0} \frac{q^n t^n}{(q)_n},$$

which is the generating function of partitions of n where the number of parts is labelled by t. Thus, comparing coefficients, we find that

$$\frac{q^k}{(q)_k}$$

is the generating function of partitions with k parts. The other important special case is obtained by replacing a by a/b and t by bt. Then

$$(a/b)_n(bt)^n = (b-a)(b-aq)\dots(b-aq^{n-1})t^n.$$

Setting now $a=-1$ and $b=0$ results in

$$(-t)_\infty = \prod_{n\ge 0}(1+tq^n) = \sum_{n\ge 0}\frac{q^{\binom{n}{2}}}{(q)_n}t^n. \tag{2}$$

This is also called Euler's partition identity. For this we consider the number of partitions into distinct parts $p_{\mathcal{D}}(n)$. Since each number i can be a part either 0 or 1 times, the formal equation

$$(\boldsymbol{\varepsilon}+\mathbf{1})(\boldsymbol{\varepsilon}+\mathbf{2})(\boldsymbol{\varepsilon}+\mathbf{3})\dots$$

describes these objects. Thus

$$\sum_{n\ge 0} p_{\mathcal{D}}(n)q^n = \prod_{k\ge 1}(1+q^k) = (-q)_\infty,$$

and $(-qt)_\infty$ is the generating function of partitions of n into distinct parts and k parts. Now we replace t by tq in Euler's partition identity:

$$(-tq)_\infty = \prod_{n\ge 1}(1+tq^n) = \sum_{n\ge 0}\frac{q^{\binom{n}{2}+n}}{(q)_n}t^n.$$

Comparing coefficients of t^k, we find that the generating function of partitions into distinct parts and k parts is given by

$$\frac{q^{\binom{k+1}{2}}}{(q)_k}.$$

Euler's partition identities appear frequently in Analytic Combinatorics and Analysis of Algorithms.

Now we want to compute the generating function of partitions where the number of parts is bounded by M and the parts are bounded by N. There is only a finite number of possibilities, whence this generating function is actually a polynomial, call it $G(M,N)$. Now $G(M,N) - G(M-1,N)$ is the generating function where the number of parts is exactly equal to M; removing one from each part shows that this equals $q^M G(M, N-1)$. Together with $G(0,N) = G(M,0) = 1$, the solution of

$$G(M,N) - G(M-1,N) = q^M G(M,N-1)$$

is given by

$$G(M,N) = \frac{(q)_{M+N}}{(q)_M (q)_N},$$

which is called a Gaussian q-binomial coefficient $\left[{M+N \atop M}\right]$.

Once again, we get

$$G(M,\infty) = G(\infty, M) = \frac{1}{(q)_M}.$$

Here is another technique of interest, nicknamed "adding a new slice" by Flajolet. Define

$$F(q,u) = \sum_{n\geq 1}\sum_{i\geq 1}[\text{number of partitions of } n \text{ with last part } i]q^n u^i.$$

Now, to create a new slice, i.e., a new part $j \geq i$, means to replace u^i by

$$\sum_{j\geq i}(qu)^j = \frac{(qu)^i}{1-qu}.$$

So, taking the partitions with just one part separately into account,

$$F(q,u) = \frac{qu}{1-qu} + \frac{1}{1-qu}F(q,qu).$$

This recursion can be iterated:

$$\begin{aligned} F(q,u) &= \frac{qu}{1-qu} + \frac{1}{1-qu}\left[\frac{q^2u}{1-q^2u} + \frac{1}{1-q^2u}\left[\frac{q^3u}{1-q^3u} + \frac{1}{1-q^3u}\left[\dots\right.\right.\right. \\ &= \sum_{k\geq 1}\frac{uq^k}{(qu)_k}. \end{aligned}$$

Forgetting what the last part is means setting $u = 1$; adding 1 for the empty partition results in

$$\sum_{k\geq 0}\frac{q^k}{(q)_k} = \frac{1}{(q)_\infty},$$

which follows directly from Euler's partition identity by setting $t = 1$. This describes the generating function of all partitions in terms of those with exactly k parts, as discussed earlier.

In various applications (see for instance [22]) it is important to expand

$$Q(x) = (qx)_\infty = \prod_{k\geq 1}(1 - xq^k)$$

around $x = 1$, viz.

$$Q(x) = Q(1) + Q'(1)(x-1) + \frac{Q''(1)}{2}(x-1)^2 + \cdots .$$

We have

$$Q'(x) = Q(x)\sum_{k\ge1}\frac{-q^k}{1-xq^k}, \quad Q''(x) = 2Q(x)\sum_{1\le j<k}\frac{q^{j+k}}{(1-xq^j)(1-xq^k)}$$

and therefore

$$\frac{Q'(1)}{Q(1)} = -\sum_{k\ge1}\frac{1}{q^{-k}-1}, \quad \frac{Q''(1)}{Q(1)} = \left(\sum_{k\ge1}\frac{1}{q^{-k}-1}\right)^2 - \sum_{k\ge1}\left(\frac{1}{q^{-k}-1}\right)^2.$$

5. Some applications of the adding a slice technique

A restricted composition of a natural number n in the sense of Carlitz [6], (Carlitz composition) is defined to be a composition

$$n = a_1 + a_2 + \cdots + a_k \quad \text{such that} \quad a_i \ne a_{i+1} \text{ for } i = 1, \dots, k-1.$$

We refer to n as the size and to k as the number of parts of the composition.

Observe that there are 2^{n-1} unrestricted compositions of the integer n with generating function $z/(1-2z)$.

Let $c(n)$ denote the number of Carlitz compositions of n. In [6], Carlitz found the generating function

$$C(z) := \sum_{n\ge0} c(n)z^n.$$

We will rederive this here with the method called "adding a new slice." This appears in [30] and is also described in the book [24].

We proceed from a Carlitz composition with k parts to one with $k+1$ parts by allowing a_{k+1} to be any number and then subtracting the forbidden case $a_{k+1} = a_k$. In terms of generating functions this reads as follows. Let $f_k(z,u)$ be the generating function of those Carlitz compositions with k parts where the coefficient of $z^n u^j$ refers to size n and last part $a_k = j$. Then

$$f_{k+1}(z,u) = f_k(z,1)\frac{zu}{1-zu} - f_k(z,zu) + [k=0] \quad \text{for } k \ge 0, \;\; f_0(z,u) = 1.$$

The first term means that we forget the labelling of the last part ($u := 1$) and add any term, together with a labelling by u, and the second one means that we subtract the forbidden term, which is a repetition of the previous last part. Introducing $F(z,u) := \sum_{k\ge1} f_k(z,u)$ and summing on $k \ge 0$, we get

$$F(z,u) = F(z,1)\frac{zu}{1-zu} + \frac{zu}{1-zu} - F(z,zu).$$

This functional equation can now be iterated:

$$F(z,u) = \big(1+F(z,1)\big)\frac{zu}{1-zu} - \big(1+F(z,1)\big)\frac{z^2u}{1-z^2u} + \big(1+F(z,1)\big)\frac{z^3u}{1-z^3u} - + \cdots ;$$

now setting $u := 1$ and abbreviating

$$\sigma(z) = \sum_{j\ge 1} \frac{z^j(-1)^{j-1}}{1-z^j},$$

we get

$$F(z,1) = \sigma(z) + F(z,1)\,\sigma(z).$$

Since $C(z) = 1 + F(z,1)$, we find the formula of Carlitz,

$$C(z) = \frac{1}{1-\sigma(z)}.$$

The next example is about LEVEL NUMBER SEQUENCES OF TREES [20]. However, the enumeration of these is equivalent to many other objects; the paper [12] describes the somewhat erratic history. The objects are sequences $(a_1, a_2, \dots, a_k)$ with positive integers a_i such that always $1 \le a_{i+1} \le 2a_i$. The number k is arbitrary, and $n := a_1 + \cdots + a_k$. The starting value is $a_1 = 1$. The interest is in the number H_n of sequences satisfying these conditions. Let $H_{n,j}^{[k]}$ be the number of such sequences relative to a fixed number k, and let the last element be fixed as well: $a_k = j$. Furthermore, set

$$H^{[k]}(q,u) = \sum_{n,j\ge 1} H_{n,j}^{[k]} q^n u^j \quad \text{and}$$

$$H(q,u) = \sum_{k\ge 1} H^{[k]}(q,u).$$

Now we describe how $H^{[k+1]}(q,u)$ can be obtained from $H^{[k]}(q,u)$. This is done using the substitution

$$u^j \to (uq) + (uq)^2 + \cdots + (uq)^{2j} = \frac{uq}{1-uq}\big(1-(uq)^{2j}\big).$$

This means

$$H^{[k+1]}(q,u) = \frac{uq}{1-uq}\big[H^{[k]}(q,1) - H^{[k]}(q,u^2q^2)\big]$$

and by summing on k (noticing that $H^{[1]}(q,u) = uq$),

$$H(q,u) = qu + \frac{uq}{1-uq}\big[H(q,1) - H(q,u^2q^2)\big].$$

This can again be iterated: set $G(u)=qu+\frac{uq}{1-uq}H(q,1)$, then

$$H(q,u)=G(u)-\frac{uq}{1-uq}\Big[G(u^2q^2)-\frac{u^2q^3}{1-u^2q^3}\Big[G(u^4q^6)\\ -\frac{u^4q^7}{1-u^4q^7}\Big[G(u^8q^{14})-\cdots;$$

now it is possible to set $u=1$, which means that we do not care about the value of the last element in the sequence anymore, and get

$$H(q,1)=G(1)-\frac{q}{1-q}\Big[G(q^2)-\frac{q^3}{1-q^3}\Big[G(q^6)-\frac{q^7}{1-q^7}\Big[G(q^{14})-\cdots.$$

Now this equation can be solved for $H(q,1)$, with the answer

$$H(q,1)=\frac{\displaystyle\sum_{j\ge1}\frac{(-1)^{j-1}q^{2^{j+1}-j-2}}{(1-q)(1-q^3)\dots(1-q^{2^{j-1}-1})}}{\displaystyle\sum_{j\ge0}\frac{(-1)^{j}q^{2^{j+1}-j-2}}{(1-q)(1-q^3)\dots(1-q^{2^{j}-1})}}.$$

The last example is about WORDS $a_1a_2\dots a_{2n+1}$ of odd length where a letter $k\in\mathbb{N}$ is weighted by a geometric probability pq^{k-1} $(p+q=1)$, i. e., $\Pr\{a_j=k\}=pq^{k-1}$, $k\ge1$, and the letters *obey the pattern* $a_1\ge a_2\le a_3\ge a_4\le\cdots$. Let $T_{2n+1}(u)$ be the generating function such that the coefficient of u^i in it is the mass of correct words and last letter i. Then we have for $n\ge1$

$$T_{2n+1}(u)=\frac{p^2u}{(1-qu)(1-q^2u)}T_{2n-1}(1)-\frac{p^2u}{(1-qu)(1-q^2u)}T_{2n-1}(q^2u),\\ T_1(u)=\frac{pu}{1-qu}.$$

Adding a new slice means adding a pair (k,j) with $1\le k\le i$, $j\ge k$, replacing u^i by 1 and providing the factor u^j. But

$$\sum_{k=1}^{i}pq^{k-1}\sum_{j\ge k}pq^{j-1}u^j=\frac{p^2u}{(1-qu)(1-q^2u)}-\frac{p^2u}{(1-qu)(1-q^2u)}\big(q^2u\big)^i,$$

which explains the recursion. The starting value is just

$$\sum_{j\ge1}pq^{j-1}u^j=\frac{pu}{1-qu}.$$

We introduce the generating functions

$$F(z,u)=\sum_{n\ge0}T_{2n+1}(u)z^{2n+1}\quad\text{and}\quad f(z)=F(z,1).$$

Summing up we find

$$F(z,u)=\frac{puz}{1-qu}+\frac{p^2uz^2}{(1-qu)(1-q^2u)}F(z,1)-\frac{p^2uz^2}{(1-qu)(1-q^2u)}F(z,q^2u).$$

Iterating that we find

$$f(z)=\frac{pz}{1-q}+\frac{p^2z^2}{(1-q)(1-q^2)}f(z)-\frac{p^2q^2z^3}{(1-q)(1-q^2)(1-q^3)}$$
$$-\frac{p^4q^2z^4}{(1-q)(1-q^2)(1-q^3)(1-q^4)}f(z)+\cdots$$

and eventually

$$f(z)=\sum_{n\ge0}\frac{(-1)^n(pz)^{2n+1}}{(q)_{2n+1}}q^{n(n+1)}\Big/\sum_{n\ge0}\frac{(-1)^n(pz)^{2n}}{(q)_{2n}}q^{n(n-1)}.$$

This example was taken from [44]; in this paper it is also explained why the limit for $q\to1$ of $f(z)$ is the tangent function $\tan z$. It is a classical result that $\tan z$ is the exponential generating function of up-down (or down-up) alternating permutations of odd length.

6. Lagrange inversion formula

Let $y=x\Phi(y)$, where $\Phi(y)$ is a power series such that $\Phi(0)\neq0$. It is obvious to expand x as a power series in y, but we want just the opposite. This is the celebrated inversion formula of Lagrange. We give three versions of it.

$$[x^n]y=\frac{1}{n}[y^{n-1}]\big(\Phi(y)\big)^n,\qquad (n\ge1).$$

Slightly more general (for $n\ge1$, $p\ge0$):

$$[x^n]y^p=\frac{p}{n}[y^{n-p}]\big(\Phi(y)\big)^n.$$

Even more general (for $n\ge1$ and a power series $g(y)$):

$$[x^n]g(y)=\frac{1}{n}[y^{n-1}]g'(y)\big(\Phi(y)\big)^n.$$

Proof. We use Cauchy's integral formula. In case this is not justified analytically, it can be done on a purely formal level as explained in [25]. The present approach is also easy to remember. Observe that

$$dx=dy\frac{\Phi(y)-y\Phi'(y)}{\Phi^2(y)}.$$

Now

$$[x^n]y^p=\frac{1}{2\pi i}\oint\frac{dx}{x^{n+1}}y^p$$

$$\begin{aligned}
&= \frac{1}{2\pi i}\oint dy \frac{\Phi(y) - y\Phi'(y)}{\Phi^2(y)} \frac{\Phi^{n+1}(y)}{y^{n+1}} y^p \\
&= \frac{1}{2\pi i}\oint dy \Big(\Phi(y) - y\Phi'(y)\Big) \frac{\Phi^{n-1}(y)}{y^{n+1-p}} \\
&= [y^{n-p}]\Big(\Phi(y) - y\Phi'(y)\Big)\Phi^{n-1}(y) \\
&= [y^{n-p}]\Phi^n(y) - [y^{n-p-1}]\Phi'(y)\Phi^{n-1}(y) \\
&= [y^{n-p}]\Phi^n(y) - \frac{1}{n}[y^{n-p-1}]\big(\Phi^n(y)\big)' \\
&= [y^{n-p}]\Phi^n(y) - \frac{n-p}{n}[y^{n-p}]\Phi^n(y) \\
&= \frac{p}{n}[y^{n-p}]\Phi^n(y).
\end{aligned}$$

Now we set

$$g(y) = \sum_{p\ge1} c_p y^p$$

and get

$$\begin{aligned}
[x^n]g(y) &= \sum_{p\ge1} c_p [x^n] y^p = \sum_{p\ge1} c_p \frac{p}{n}[y^{n-p}]\big(\Phi(y)\big)^n \\
&= \frac{1}{n}[y^{n-1}]\sum_{p\ge1} p c_p y^{p-1}\big(\Phi(y)\big)^n = \frac{1}{n}[y^{n-1}]g'(y)\big(\Phi(y)\big)^n.
\end{aligned}$$

Applications:

(The following tree structures have been introduced in Section 2.)

t-ary trees. These objects are recursively built from a root and t successors, which are themselves t-ary trees. A tree might be empty as well. For $t = 2$, we get the important special case of binary trees. The equation $B = 1 + xB^t$ for the generating function, counting trees according to the number of vertices, is immediate. Set $B = 1 + y$, then $y = x(1+y)^t$, in order to make the Lagrange inversion formula applicable. Then $\Phi(y) = (1+y)^t$, and thus

$$b_n = [x^n]B(x) = [x^n]y(x) = \frac{1}{n}[y^{n-1}]\big(1+y\big)^{tn} = \frac{1}{n}\binom{tn}{n-1}.$$

Number of leaves in planar trees. (*Narayana* numbers.) The recursive description of planar trees immediately translates into a bivariate generating function:

$$G(z,u) = zu + \frac{zG(z,u)}{1 - G(z,u)};$$

the variable u counts leaves, z nodes (in planar trees). Note that for $u = 1$ and $y = G(z, 1)$, we have $y = z\Phi(y)$ with $\Phi(y) = \frac{1}{1-y}$.

$$\begin{aligned} G_{n,k} = [u^k][z^n]G(z,u) &= [u^k]\frac{1}{n}[y^{n-1}]\Big(u + \frac{y}{1-y}\Big)^n \\ &= \frac{1}{n}[y^{n-1}]\binom{n}{k}\frac{y^{n-k}}{(1-y)^{n-k}} \\ &= \frac{1}{n}\binom{n}{k}[y^{k-1}]\frac{1}{(1-y)^{n-k}} \\ &= \frac{1}{n}\binom{n}{k}\binom{n-2}{k-1}. \end{aligned}$$

These numbers are sometimes called Narayana numbers.

Planar trees according to degree of the root. Let $P_{n,k}$ be the number of planar trees with n nodes and root degree k. Then, again, $y = z\Phi(y)$ with $\Phi(y) = y/(1-y)$ enumerates planar trees. Let us assume that $p \geq 1$ and $n \geq 2$, since $P_{n,0} = [n = 1]$. We find

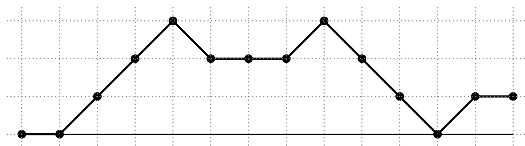
$$P_{n,k} = [z^{n-1}]y^k = \frac{k}{n-1}[y^{n-1-k}](1-y)^{-n+1} = \frac{k}{n-1}\binom{2n-3-k}{n-2}.$$

7. Lattice path enumeration: the continued fraction theorem

We follow here [24] and [13].

Definition 7.1 (Lattice path)**.** *A Motzkin path $\upsilon = (U_0, U_1, \ldots, U_n)$ is a sequence of points in the discrete quarter-plane $\mathbb{N}_0 \times \mathbb{N}_0$, such that $U_j = (j, y_j)$ and the jump condition $|y_{j+1} - y_j| \leq 1$ is satisfied. An edge $\langle U_j, U_{j+1}\rangle$ is called an ascent if $y_{j+1} - y_j = 1$, a descent if $y_{j+1} - y_j = -1$, and a level step if $y_{j+1} - y_j = 0$. A path that has no level steps is called a Dyck path. The quantity n is the length of the path, ini$(\upsilon) := y_0$ is the initial altitude, fin$(\upsilon) := y_n$ 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.*

A path can always be 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 initial point. For instance,



$c_0\ a_0\ a_1\ a_2\ b_3\ c_2\ c_2\ a_2\ b_3\ b_2\ b_1\ a_0\ c_1$

encodes a path that connects the initial point $(0,0)$ to the point $(13,1)$.

Let us examine the description of the class written $\mathcal{H}_{0,0}^{[<1]}$ of Motzkin excursions of height <1. We have

$$\mathcal{H}_{0,0}^{[<1]} \cong (c_0)^* \quad \Longrightarrow \quad H_{0,0}^{[<1]} = \frac{1}{1-c_0}.$$

The class of excursions of height <2 is obtained from here by a substitution

$$c_0 \mapsto c_0 + a_0(c_1)^* b_1,$$

whence

$$H_{0,0}^{[<2]} = \cfrac{1}{1-c_0-\cfrac{a_0b_1}{1-c_1}}.$$

Iteration of this simple mechanism yields already the finite version of the continued fraction theorem of Flajolet [13].

Theorem 7.2 (Continued fraction theorem, finite version)**.**

$$H_{0,0}^{[<h]} = \cfrac{1}{1-c_0-\cfrac{a_0b_1}{1-c_1-\cfrac{a_1b_2}{\cfrac{\ddots}{1-c_{h-1}}}}} = \frac{P_h}{Q_h}.$$

The unrestricted version leads to an infinite continued fraction.

Theorem 7.3 (Continued fraction theorem, infinite version)**.**

$$H_{0,0} = \cfrac{1}{1-c_0-\cfrac{a_0b_1}{1-c_1-\cfrac{a_1b_2}{\ddots}}}.$$

Generating functions written in this way are nothing but a concise description of usual counting generating functions: for instance if individual weights α_j, β_j, γ_j are assigned to the letters a_j, b_j, c_j, respectively, then the ordinary generating function of multiplicatively weighted paths with z marking length is obtained by setting $a_j = \alpha_j z$, $b_j = \beta_j z$, $c_j = \gamma_j z$.

The "numerator" and "denominator" polynomials, denoted by P_h and Q_h are defined as solutions to the second-order (or "three-term") linear recurrence equation

$$Y_{h+1} = (1 - c_h)Y_h - a_{h-1}b_h Y_{h-1}, \quad h \geq 0,$$

together with the initial conditions $(P_{-1}, Q_{-1}) = (-1, 0)$, $(P_0, Q_0) = (0, 1)$, and with the convention $a_{-1} = b_0 = 1$. These recursions are easy to obtain by replacing $1 - c_{n-1}$ by $1 - c_{n-1} - \frac{a_{n-1}b_n}{1-c_n}$ and comparing numerators and denominators separately.

These polynomials are also known as continuant polynomials [26, 48]. 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

$$H_{0,0}^{[<h]} = g_0 \circ g_1 \circ \cdots \circ g_{h-1}(0).$$

Linear fractional transformations are representable by 2×2 matrices

$$\frac{ay + b}{cy + d} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

in such a way that the composition corresponds to the matrix product. By induction on the compositions that build up $H_{0,0}^{[<h]}$, there follows the equality

$$g_0 \circ g_1 \circ \cdots \circ g_{h-1}(y) = \frac{P_h - P_{h-1}a_{h-1}b_h y}{Q_h - Q_{h-1}a_{h-1}b_h y}.$$

Eventually, one sets $y := 0$. The polynomials P_h and Q_h both satisfy the recursion for Y_h as just given.

Now we come to applications. In order to count Dyck paths, it is sufficient to substitute

$$a_j \mapsto z, \qquad b_j \mapsto z, \qquad c_j \mapsto 0.$$

Because of the natural bijection (sometimes calles the glove bijection), described earlier, a Dyck path of length $2n$ and height h translates into a planar tree with $n + 1$ nodes and height $h + 1$, so that the results translate directly; compare [10]. In order to avoid misunderstandings, we state explicitly that the height of a planar tree is the length of the longest path from the root to a leaf; the length of a path is counted in terms of the number of nodes on it. This is the original definition used in [10]; sometimes people count the number of edges, which is then one less than what is considered here. The height of a Dyck path is of course the maximal vertical level reached.

The families of polynomials P_h, Q_h are in this case determined by a recurrence with constant coefficients. Define the Fibonacci polynomials by the recurrence

$$F_{h+2}(z) = F_{h+1}(z) - zF_h(z), \quad F_0(z) = 0,\ F_1(z) = 1,$$

then it turns out that $Q_h(z) = F_{h+1}(z^2)$ and $P_h(z) = F_h(z^2)$. The Fibonacci polynomials admit an explicit form

$$F_h(z) = \frac{1}{\sqrt{1-4z}}\left[\left(\frac{1+\sqrt{1-4z}}{2}\right)^h - \left(\frac{1-\sqrt{1-4z}}{2}\right)^h\right].$$

If we take the limit $h \to \infty$ in

$$\frac{F_h(z)}{F_{h+1}(z)},$$

then we get

$$D(z) = \frac{1-\sqrt{1-4z}}{2z} = \sum_{n\ge0} \frac{1}{n+1}\binom{2n}{n} z^n,$$

the generating function of Dyck paths of halflength n; the coefficients are Catalan numbers. The limit is taken in the sense of the "discrete topology," but it works as well as the "pointwise" limit for analytic functions. As demonstrated in [10], all our expressions become easier with the substitution

$$z = \frac{u}{(1+u)^2} \quad \Longrightarrow \quad dz = \frac{1-u}{(1+u)^3}du.$$

Then, by solving the second order recursion and substituting,

$$F_h = \frac{1}{1-u}\frac{1-u^h}{(1+u)^{h-1}}, \qquad \frac{F_h}{F_{h+1}} = (1+u)\frac{1-u^h}{1-u^{h+1}}.$$

The limit of the last expression for $h \to \infty$ is $1+u$. The function $D(z) - F_h(z)/F_{h+1}(z) = \frac{1-u^2}{u}\frac{u^{h+1}}{1-u^{h+1}}$ is easier and describes the Dyck paths with height $\ge h$, according to halflength; call it $H^{[\ge h]}(z)$.

The following method to extract coefficients is in [10]; it is the Lagrange inversion formula in disguise and uses the Cauchy integral formula:

$$\begin{aligned}
[z^n]H^{[\geq h-1]}(z) &= \frac{1}{2\pi i}\oint \frac{dz}{z^{n+1}}\frac{1-u^2}{u}\frac{u^h}{1-u^h} \\
&= \frac{1}{2\pi i}\oint \frac{du(1+u)^{2n-1}(1-u)}{u^{n+1}}\frac{1-u^2}{u}\frac{u^h}{1-u^h} \\
&= [u^{n+1}](1-u)^2(1+u)^{2n}\sum_{k\geq 1}u^{hk} \\
&= \sum_{k\geq 1}[u^{n+1-hk}](1-2u+u^2)(1+u)^{2n} \\
&= \sum_{k\geq 1}\left[\binom{2n}{n+1-hk} - 2\binom{2n}{n-hk} + \binom{2n}{n-1-hk}\right].
\end{aligned} \tag{3}$$

We would like to introduce an alternative method to compute the generating function of Dyck paths of height $< h$; it appears for instance in [41]. Define $\varphi_i(z)$ the generating function (according to length) of non-negative lattice paths starting at $(0,0)$, ending at (n,i), and height $< h$, for $i = 0,\dots,h-1$. Then

$$\begin{gathered}
\varphi_0(z) = 1 + z\varphi_1(z), \quad \varphi_{h-1}(z) = z\varphi_{h-2}(z), \\
\varphi_i(z) = z\varphi_{i-1}(z) + z\varphi_{i+1}(z) \quad \text{for } 1 \leq i \leq h-2.
\end{gathered}$$

This system is best written as a matrix equation:

$$\begin{pmatrix}\varphi_0\\ \varphi_1\\ \vdots\\ \varphi_{h-1}\end{pmatrix} = \begin{pmatrix}1 & -z & 0 & \dots\\ -z & 1 & -z & \dots\\ & & \ddots & \\ & & -z & 1\end{pmatrix}\begin{pmatrix}\varphi_0\\ \varphi_1\\ \vdots\\ \varphi_{h-1}\end{pmatrix} + \begin{pmatrix}1\\ 0\\ \vdots\\ 0\end{pmatrix}$$

It can be solved using Cramer's rule. Denote by a_h the determinant of the matrix

$$\begin{pmatrix}1 & -z & 0 & \dots\\ -z & 1 & -z & \dots\\ & & \ddots & \\ & & -z & 1\end{pmatrix}$$

with h rows and columns. Expanding with respect to the first row yields the recursion

$$a_h = a_{h-1} - z^2 a_{h-2}, \quad \text{for } h \geq 2,\ a_0 = a_1 = 1.$$

The solution is

$$a_h = \frac{1}{\sqrt{1-4z^2}}\biggl[\Bigl(\frac{1+\sqrt{1-4z^2}}{2}\Bigr)^{h+1} - \Bigl(\frac{1-\sqrt{1-4z^2}}{2}\Bigr)^{h+1}\biggr],$$

and thus

$$\varphi_i(z) = \frac{z^i a_{h-1-i}}{a_h}.$$

The good substitution in this case is $z = \frac{u}{1+u^2}$, because then

$$a_h = \frac{1}{1-u^2}\frac{1-u^{2h+2}}{(1+u^2)^h}.$$

This approach works also in the Motzkin case (level steps allowed), see [42]; the equation is then

$$\begin{pmatrix}\varphi_0\\ \varphi_1\\ \vdots\\ \varphi_{h-1}\end{pmatrix} = \begin{pmatrix}1-z & -z & 0 & \dots\\ -z & 1-z & -z & \dots\\ & & \ddots & \\ & & -z & 1-z\end{pmatrix}\begin{pmatrix}\varphi_0\\ \varphi_1\\ \vdots\\ \varphi_{h-1}\end{pmatrix} + \begin{pmatrix}1\\ 0\\ \vdots\\ 0\end{pmatrix};$$

the recursion for the determinants is

$$a_h = (1-z)a_{h-1} - z^2 a_{h-2},$$

and

$$a_h = \frac{1}{\sqrt{1-2z-3z^2}}\biggl[\Bigl(\frac{1-z+\sqrt{1-2z-3z^2}}{2}\Bigr)^{h+1} - \Bigl(\frac{1-z-\sqrt{1-2z-3z^2}}{2}\Bigr)^{h+1}\biggr],$$

and the good substitution is $z = \frac{v}{1+v+v^2}$. Explicit expressions are also possible; they involve trinomial coefficients $\binom{n,3}{k} = [v^k](1+v+v^2)^n$ (notation from [9]).

It is worthwhile to notice that Motzkin paths can be obtained from Dyck paths by squeezing in arbitrary sequences of level steps between consecutive steps. That amounts to replace z by $z(1+z+\cdots) = \frac{z}{1-z}$, if we assume that an up/down-step is followed by a sequence of level steps. We must provide for an arbitrary sequence of level steps in the beginning. So we get

$$\frac{1}{1-z}D\biggl(\Bigl(\frac{z}{1-z}\Bigr)^2\biggr) = \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2},$$

as expected. Since this procedure does not affect the height, functions like $H_{0,0}^{[<h]}$ also translate.

8. Lattice path enumeration: the kernel method

This is taken from the survey paper [45]; I am confident that the present Handbook will have much more about the subject in different chapters.

In the present author's view, the *kernel method* originated in Knuth's book [31], where it was presented as an innocent exercise 2.2.1.-4. Later, it was turned into a method; see [4] and the literature cited therein. It was probably rediscovered independently by many people; I recommend to follow the references in [4].

I feel that I cannot do anything better as an introduction than to reproduce Knuth's original exercise. One starts at the origin, and can advance from (n, i) to both $(n+1, i \pm 1)$, except in the case when $i = 0$, when one can only go to $(n+1, 1)$. In this way, one models non-negative lattice paths (or random walks). The Dyck paths of the previous section are the case where one ends at level 0 after n steps. We want to know how many paths lead from the origin to $(n, 0)$, and, more generally, to (n, i). (Clearly, this is a very classical subject, but the derivation that Knuth presented is the subject of this presentation.) One uses generating functions $f_i(z)$, describing walks leading to (n, i); the coefficient of z^n is the number of walks from the origin to (n, i). The following recursions are immediate:

$$\begin{aligned} f_i(z) &= zf_{i-1}(z) + zf_{i+1}(z), \qquad i \geq 1, \\ f_0(z) &= 1 + zf_1(z). \end{aligned}$$

Now one introduces $F(z, x) = \sum_{n \geq 0} f_n(z)x^n$, multiplies the recursion by x^i and sums:

$$F(z, x) - f_0(z) = zxF(z, x) + \frac{z}{x}\big[F(z, x) - f_0(z) - xf_1(z)\big],$$

or

$$F(z, x) = zxF(z, x) + \frac{z}{x}\big[F(z, x) - F(z, 0)\big] + 1,$$

whence

$$F(z, x) = \frac{zF(z, 0) - x}{zx^2 - x + z}.$$

Plugging in $x = 0$ leads to nothing, but the denominator factors as $z(x - r_1(z))(x - r_2(z))$, with

$$r_{1,2}(z) = \frac{1 \mp \sqrt{1 - 4z^2}}{2z}.$$

Note that $x - r_1(z) \sim x - z$ as $x, z \to 0$. Therefore the factor $1/(x - r_1(z))$ has no power series expansion around $(0, 0)$, but $F(z, x)$ *has,* so this "bad" factor must actually disappear, i.e., $(x - r_1(z))$ must be a factor of the *numerator* as well, which leads to the equation $zF(z,0) = r_1(z)$, from which $F(z,0)$ can be computed. Consequently, $F(z,x)$ is then also explicitly computed, and the factor $(x - r_1(z))$ can be cancelled from both, numerator and denominator.

From this, one finds for instance that $[z^{2n}]F(z,0) = \frac{1}{n+1}\binom{2n}{n}$, a well-known *Catalan number*, and similar expressions for $[z^n x^i]F(z,x)$, for $n \equiv i \bmod 2$.

The next example revisits the *toilet paper problem*, a popular subject introduced by Knuth [34]. He considers two rolls of tissues, with m resp. n units, and random users, who are with probability p *big-choosers* (taking one unit from the larger roll) resp. with probability $q = 1 - p$ *little-choosers* (taking one unit from the smaller roll). The parameter of interest is the (average) number of units remaining on the larger roll, when the smaller one became empty.

Let m be the number of units on the larger, and n on the smaller roll; $M_{m,n}$ is the expected number of units left on the larger roll, when the smaller one becomes empty.

The recursions are

$$\begin{aligned}
M_{m,0} &= m, \\
M_{m,m} &= M_{m,m-1}, \qquad m \ge 1, \\
M_{m,n} &= pM_{m-1,n} + qM_{m,n-1}, \qquad m > n > 0.
\end{aligned}$$

Define

$$F_0(z) = \sum_{m\ge 0} M_{m,m} z^m, \qquad F_1(z) = \sum_{m\ge 1} M_{m,m-1} z^m.$$

Note that

$$F_0(z) = \sum_{m\ge 0} M_{m,m} z^m = \sum_{m\ge 1} M_{m,m-1} z^m = F_1(z).$$

Define

$$F(z,x) = \sum_{m\ge n\ge 0} M_{m,n} z^m x^{m-n}.$$

Then, by summing up,

$$\begin{aligned}
F(z,x) &= \sum_{m>n>0} M_{m,n} z^m x^{m-n} + \sum_{m>0} M_{m,0} z^m x^m + \sum_{m\ge 0} M_{m,m} z^m \\
&= \sum_{m>n>0} [pM_{m-1,n} + qM_{m,n-1}] z^m x^{m-n} + \frac{zx}{(1-zx)^2} + F_0(z) \\
&= pzx \sum_{m-1\ge n>0} M_{m-1,n} z^{m-1} x^{(m-1)-n}
\end{aligned}$$

$$
\begin{aligned}
&\quad + \frac{q}{x} \sum_{m>n>0} M_{m,n-1} z^m x^{m-(n-1)} + \frac{zx}{(1-zx)^2} + F_0(z) \\
&= pzx\Big[F(z,x) - \sum_{m\geq 0} M_{m,0} z^m x^m\Big] \\
&\quad + \frac{q}{x} \sum_{m\gg n\geq 0} M_{m,n} z^m x^{m-n} + \frac{zx}{(1-zx)^2} + F_0(z) \\
&= pzx\Big[F(z,x) - \frac{zx}{(1-zx)^2}\Big] \\
&\quad + \frac{q}{x}\Big[F(z,x) - x\sum_{n\geq 0} M_{n+1,n} z^{n+1} - \sum_{n\geq 0} M_{n,n} z^n\Big] \\
&\quad + \frac{zx}{(1-zx)^2} + F_0(z) \\
&= pzxF(z,x) + \frac{q}{x}\Big[F(z,x) - xF_1(z) - F_1(z)\Big] \\
&\quad + \frac{zx(1-pzx)}{(1-zx)^2} + F_1(z).
\end{aligned}
$$

(Here, we used the ad hoc notation $a \gg b :\Leftrightarrow a - b \geq 2$.) Solving,

$$
\begin{aligned}
F(z,x) &= \frac{\frac{zx(1-pzx)}{(1-zx)^2} + F_1(z)[1-q-q/x]}{1-pzx-q/x} = \frac{F_1(z)[q-px] - \frac{zx^2(1-pzx)}{(1-zx)^2}}{pzx^2 - x + q} \\
&= \frac{F_1(z)[q-px] - \frac{zx^2(1-pzx)}{(1-zx)^2}}{pz(x-r_1(z))(x-r_2(z))},
\end{aligned}
$$

with

$$r_{1,2}(z) = \frac{1 \mp \sqrt{1-4pqz}}{2pz}.$$

Therefore, for $x = r_1(z)$, the numerator must vanish, yielding

$$F_1(z)[q - pr_1(z)] - \frac{zr_1^2(z)(1-pzr_1(z))}{(1-zr_1(z))^2} = 0,$$

or

$$F_1(z) = \frac{zr_1^2(z)(1-pzr_1(z))}{(q-pr_1(z))(1-zr_1(z))^2} = \frac{z}{q(1-z)^2}\big(q - C(pqz)\big),$$

with

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

Note that $r_1(z) = \frac{C(pqz)}{pz}$ and that $1/r_2(z) = r_1(z)pz/q$. The argument that $1/(x - r_1(z))$ has no power series expansion around $(x,z) \sim (0,0)$ must be replaced here by something else, for instance, that $F(z,x)/z \sim 1$ for $(x,z) \sim (0,0)$, which follows from the combinatorial description of the problem.

The expression for $F(z,x)$ is ugly, but we can extend Knuth's asymptotic analysis to $M_{m,m-n}$ for $m\to\infty$ and fixed n; the instance $n=0$ was given in [34]. This asymptotic analysis is perhaps best understood after consulting the Section 17 first. However, we decided to put it in here, so that we avoid having to repeat the description of the problem later.

For $q<p$, Knuth has shown that the local expansion of $C(pqz)$ around $z=1$ starts like

$$q+\tfrac{pq}{p-q}(z-1)+\tfrac{(pq)^2}{(p-q)^3}(z-1)^2+\cdots.$$

Hence the local expansion of $F(z,x)$ around $z=1$ is given by

$$\frac{1}{1-z}\cdot\frac{p}{(2p-1)(1-x)}+Q(z),$$

where $Q(z)$ has a radius of convergence >1. So $M_{m,m-n}=p/(2p-1)+O(r^m)$, (for a suitable $0<r<1$), and the n plays no role here. This is intuitive, the big-choosers dominate, so it does not really make a difference whether the second roll is slightly smaller. Now let us assume that $p<q$. Then $F(z,x)$ starts like

$$\frac{1}{(1-z)^2}\cdot\frac{2p-1}{(p-1)(1-x)}\\
+\frac{1}{1-z}\biggl[-\frac{1}{1-x}+\frac{p}{q(1-x)^2}-\frac{p(1-p)}{(2p-1)(q-px)}\biggr]+\cdots.$$

The coefficient of z^m is asymptotic to

$$(m+1)\cdot\frac{1-2p}{(1-p)(1-x)}+\biggl[-\frac{1}{1-x}+\frac{p}{q(1-x)^2}-\frac{p(1-p)}{(2p-1)(q-px)}\biggr].$$

And the coefficient of x^n (n fixed) in this is

$$(m+1)\cdot\frac{1-2p}{1-p}+\biggl[-1+\frac{p}{q}(n+1)-\frac{p}{2p-1}\Bigl(\frac{p}{q}\Bigr)^n\biggr],$$

or

$$m\cdot\frac{1-2p}{q}+\frac{pn}{q}+\frac{p}{1-2p}\Bigl(\frac{p}{q}\Bigr)^n.$$

For $n=0$, we find again Knuth's value $\frac{q-p}{q}m+\frac{p}{q-p}$. Perhaps it is not very intuitive at the first glance why this *grows* with n. However, for larger n, the process tends to be over more quickly, and so more will be left on the large roll.

Now let us discuss the case $p=q$. Then

$$C(pqz)=\frac12-\frac12\sqrt{1-z},$$

and[1]

$$F(z,x) \sim (1-z)^{-3/2} \cdot \frac{1}{1-x} - (1-z)^{-1/2} \cdot \frac{1-2x}{(1-x)^3},$$

and the coefficient of z^m behaves like

$$\left(2\sqrt{\frac{m}{\pi}} + \frac{3}{4\sqrt{\pi m}}\right) \cdot \frac{1}{1-x} - \frac{1}{\sqrt{\pi m}} \cdot \frac{1-2x}{(1-x)^3}.$$

Furthermore the coefficient of x^n (n fixed) in this is

$$2\sqrt{\frac{m}{\pi}} + \frac{3}{4\sqrt{\pi m}} + \frac{1}{\sqrt{\pi m}} \cdot \frac{(n+1)(n-2)}{2}.$$

For $n = 0$ we find again

$$2\sqrt{\frac{m}{\pi}} - \frac{1}{4\sqrt{\pi m}}.$$

Again, since everybody takes at random, the process tends to be over more quickly, leaving more on the larger roll.

9. Gamma and zeta function

These two special functions appear in many contexts; therefore we only collect here a few basic facts. General references are [49] and [24].

Euler defined

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \quad \text{for } \Re z > 0.$$

It is easy to show via integration by parts that $\Gamma(z+1) = z\Gamma(z)$ and $\Gamma(1) = 1$, whence $\Gamma(n+1) = n!$ for positive integers n. There is another definition due to Gauss,

$$\Gamma(z) = \lim_{n\to\infty} \frac{n! n^z}{z(z+1)\dots(z+n)}.$$

The meaning of n^z is here $e^{z \log n}$, with $\log n$ being real and positive. Now

$$\begin{aligned} \frac{z(z+1)\dots(z+n)}{n! e^{z\log n}} &= e^{-z\log n} z\Big(1+\frac{z}{1}\Big)\Big(1+\frac{z}{2}\Big)\dots\Big(1+\frac{z}{n}\Big) \\ &= e^{zH_n - z\log n} z \prod_{k=1}^{n} \Big(1+\frac{z}{k}\Big) e^{-z/k}. \end{aligned}$$

[1]We use the symbol '$\sim$' in the sense that a full *asymptotic series* in powers $1-z$ would be available, at least in principle, to which the method of *singularity analysis of generating functions (transfer theorems)* [19] is applicable, see also Section 17.

Here, $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ is a *harmonic number.* Thus, the limit can be made explicit by writing

$$\frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{n\ge1} \Big[\Big(1+\frac{z}{n}\Big)e^{-z/n}\Big],$$

with Euler's constant

$$\gamma = \lim_{n\to\infty}(H_n - \log n) = \sum_{n\ge1}\Big[\frac{1}{n} - \log\Big(1+\frac{1}{n}\Big)\Big] = 0.5772156649.$$

Gauss' definition is more general, but coincides with Euler's when both make sense. From it, it is not hard to derive Legendre's *duplication* formula

$$\Gamma(z)\Gamma(z+\tfrac{1}{2}) = \Gamma(2z)\sqrt{\pi}\,2^{1-2z}$$

as well as the *reflection* formula

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}.$$

From this, the special value $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ follows.

The logarithmic derivate of the Gamma function is

$$\psi(z) := \frac{\Gamma'(z)}{\Gamma(z)};$$

using the product form it can be written as

$$\psi(z+1) = -\gamma - \sum_{n\ge1}\Big(\frac{1}{n+z} - \frac{1}{n}\Big) = -\gamma - \sum_{n\ge2}(-1)^n\zeta(n)z^{n-1},$$

with the Riemann zeta function $\zeta(s) = \sum_{n\ge1} n^{-s}$ for $\Re s > 1$. This also expresses harmonic numbers: $H_n = \psi(n+1) + \gamma$. From the last expansion of $\psi(z+1)$, the expansion of $\Gamma(z)$ around any integer can be reconstructed; for example $\Gamma(z+1) \sim 1-\gamma z$ for $z \to 0$; the Gamma function has simple poles at $z = -k$, with residue $(-1)^k/k!$, $k = 0, 1, 2, \ldots$. This can be seen directly from Gauss' definition as well.z‘

For purely imaginary values, we have

$$|\Gamma(iz)| = \sqrt{\frac{\pi}{z\sinh(\pi z)}}\ ,$$

this also shows the rapid decay when z becomes large. □

The Riemann zeta function is for $\Re s > 1$ defined by

$$\zeta(s) = \sum_{n\ge1}\frac{1}{n^s}.$$

It can be continued to the whole of $\mathbb{C}$ and has only one simple pole at $s=1$ with the Laurent expansion

$$\zeta(s) = \frac{1}{s-1} + \gamma + \cdots .$$

The following special values are useful:

$$\zeta(2k) = \frac{2^{2k-1} B_{2k} (-1)^{k-1} \pi^{2k}}{(2k)!} \quad \text{for } k \in \mathbb{N},$$

$$\zeta(-2k+1) = -\frac{B_{2k}}{2k}, \quad \zeta(-2k) = 0,$$

$$\zeta(s) \sim -\frac{1}{2} - \log\sqrt{2\pi} \cdot s + \cdots , \qquad (s \to 0),$$

with B_n the Bernoulli numbers.

The following functional equation due to Riemann is very famous:

$$\Gamma\Big(\frac{s}{2}\Big)\zeta(s) = \pi^{s-\frac{1}{2}} \Gamma\Big(\frac{1-s}{2}\Big)\zeta(1-s).$$

Some properties of the Bernoulli numbers (taken from [26]) are now collected: They are defined by their exponential generating function

$$\frac{z}{e^z - 1} = \sum_{n \geq 0} B_n \frac{z^n}{n!};$$

from this the recursion

$$\sum_{k=0}^{n} \binom{n+1}{k} B_k = [n=0] \quad \text{for all } n \geq 0$$

follows. Bernoulli numbers with odd indices, except $B_1 = -\frac{1}{2}$, are zero. The coefficients of the tangent may be expressed by them:

$$\tan z = \sum_{n \geq 0} (-1)^{n-1} 4^n (4^n - 1) B_{2n} \frac{z^{2n-1}}{(2n)!}.$$

And the celebrated sum of the first mth powers:

$$1^m + 2^m + \cdots + (n-1)^m = \frac{1}{m+1}(B_{m+1}(n) - B_{m+1}),$$

here we have Bernoulli polynomials defined by

$$B_m(x) = \sum_k \binom{m}{k} B_k x^{m-k}.$$

They satisfy $B_m(0) = B_m$ and have the exponential generating function

$$\sum_{m \geq 0} B_m(x) \frac{z^m}{m!} = \frac{z e^{xz}}{e^z - 1}.$$

10. Harmonic numbers and their generating functions

We start with

$$[z^n]\frac{1}{1-z}\log\frac{1}{1-z}=\sum_{k=1}^{n}[z^k]\log\frac{1}{1-z}=\sum_{k=1}^{n}\frac{1}{k}=H_n.$$

For many applications it is necessary to know the coefficients of

$$\frac{1}{(1-z)^{m+1}}\log^k\frac{1}{1-z}$$

for integers $m,k\ge 0$. The following derivation is based on [50]. We start from a bivariate generating function

$$\begin{aligned}
\sum_{k\ge0}\frac{1}{(1-z)^{m+1}}\log^k\frac{1}{1-z}\frac{t^k}{k!}&=\frac{1}{(1-z)^{m+1}}\exp\Big(t\log\frac{1}{1-z}\Big)\\
&=\frac{1}{(1-z)^{m+1+t}}=\sum_{n\ge0}\binom{m+t+n}{n}z^n\\
&=\sum_{n\ge0}\binom{m+n}{n}z^n\prod_{j=1}^{n}\Big(1+\frac{t}{m+j}\Big)\\
&=\sum_{n\ge0}\binom{m+n}{n}z^n\exp\bigg(\sum_{j=1}^{n}\log\Big(1+\frac{t}{m+j}\Big)\bigg)\\
&=\sum_{n\ge0}\binom{m+n}{n}z^n\exp\bigg(\sum_{i\ge1}\frac{(-1)^{i-1}t^i}{i}\sum_{j=1}^{n}\frac{1}{(m+j)^i}\bigg)\\
&=\sum_{n\ge0}\binom{m+n}{n}z^n\exp\bigg(\sum_{i\ge1}\frac{(-1)^{i-1}t^i}{i}\Big(H^{(i)}_{m+n}-H^{(i)}_{m}\Big)\bigg),
\end{aligned}$$

with harmonic numbers of order i,

$$H^{(i)}_n=\sum_{k=1}^{n}\frac{1}{k^i}.$$

Now we can read off coefficients:

$$\begin{aligned}
&k![z^nt^k]\frac{1}{(1-z)^{m+1}}\log^k\frac{1}{1-z}\frac{t^k}{k!}=\binom{m+n}{n}k!(-1)^k\\
&\times\sum_{1j_1+2j_2+\cdots=k}\frac{(-1)^{j_1+j_2+\cdots}\big(H^{(1)}_{m+n}-H^{(1)}_{m}\big)^{j_1}\big(H^{(2)}_{m+n}-H^{(2)}_{m}\big)^{j_2}\cdots}{1^{j_1}2^{j_2}\cdots j_1!j_2!\cdots}.
\end{aligned}$$

For $k=1$ and $k=2$ we get the important special cases

$$\frac{1}{(1-z)^{m+1}}\log\frac{1}{1-z}=\sum_{n\ge 0}\binom{m+n}{n}\big(H_{m+n}-H_m\big),$$

$$\frac{1}{(1-z)^{m+1}}\log^2\frac{1}{1-z}=\sum_{n\ge 0}\binom{m+n}{n}\Big[\big(H_{m+n}-H_m\big)^2-\big(H^{(2)}_{m+n}-H^{(2)}_m\big)\Big].$$

The general expression can be written using Bell polynomials [50].

11. Approximation of binomial coefficients

One often needs to approximate binomial coefficients $\binom{2n}{n+k}$ in a central region (e.g. $|k|\le\sqrt{n}\log n$). We obtain, for instance by Stirling's formula (6),

$$\frac{\binom{2n}{n+k}}{\binom{2n}{n}}\sim e^{-k^2/n}\cdot\Big(1+\frac{k^2}{2n^2}-\frac{k^4}{6n^3}+\cdots\Big). \tag{4}$$

A derivation of Stirling's formula will be given later (6). Quite often r-th differences of binomial coefficients appear, like (for $k=2$)

$$\frac{\binom{2n}{n+k+1}-2\binom{2n}{n+k}+\binom{2n}{n+k-1}}{\binom{2n}{n}}.$$

Let us recall the difference operator Δ (operating on k):

$$\Delta^r\binom{2n}{n+k}=\sum_{l=0}^{r}\binom{r}{l}(-1)^{r-l}\binom{2n}{n+k+l}.$$

Let us consider $\Delta^r e^{-k^2/n}$. It is not difficult to see that one can apply this difference operator term by term to the expansion (4).

For this, we need the Hermite polynomials. Here are the necessary properties [2]:

$$e^{2zt-t^2}=\sum_{n\ge 0}H_n(z)\frac{t^n}{n!},$$
$$H_0=1,\quad H_1=2z,\quad H_{n+1}=2zH_n-2nH_{n-1},$$
$$H_n(z)=\sum_{k\ge 0}\frac{n!}{k!(n-2k)!}(-1)^k(2z)^{n-2k},$$
$$H_n(-z)=(-1)^nH_n(z).$$

Thus

$$\Delta^r e^{-k^2/n}=\sum_{l=0}^{r}(-1)^{r-l}\binom{r}{l}e^{-(k+l)^2/n}$$

$$= e^{-k^2/n}\sum_{l=0}^{r}(-1)^{r-l}\binom{r}{l}e^{-\frac{2kl}{n}-\frac{l^2}{n}}$$

$$= e^{-k^2/n}\sum_{l=0}^{r}(-1)^{r-l}\binom{r}{l}\sum_{m\ge 0}H_m\Big(-\frac{k}{\sqrt{n}}\Big)\Big(\frac{l}{\sqrt{n}}\Big)^m\frac{1}{m!}.$$

We know (with Stirling subset numbers) that

$$\sum_{l=0}^{r}(-1)^{r-l}\binom{r}{l}l^m = r!\left\{{m \atop r}\right\}.$$

Hence

$$(-1)^r\Delta^r e^{-k^2/n}\sim e^{-k^2/n}\left[H_r\Big(\frac{k}{\sqrt{n}}\Big)\frac{1}{n^{r/2}}-H_{r+1}\Big(\frac{k}{\sqrt{n}}\Big)\frac{1}{n^{(r+1)/2}}\frac{r}{2}+\cdots\right].$$

Therefore we have

$$\sum_{l=0}^{r}(-1)^l\binom{r}{l}\frac{\binom{2n}{n+k+r}}{\binom{2n}{n}}\sim e^{-k^2/n}H_r\Big(\frac{k}{\sqrt{n}}\Big)\frac{1}{n^{r/2}},$$

which was announced in [21].

To obtain more terms we have to consider $\Delta^r k^t e^{-k^2/n}$. For this, we can use the general formula (of Leibniz type):

$$\Delta^r\big(f(k)g(k)\big)=\sum_{l=0}^{r}\binom{r}{l}\big(\Delta^{r-l}f(k+l)\big)\big(\Delta^l g(k)\big)$$

with $f(k)=k^t$ and $g(k)=e^{-k^2/n}$.

A very general answer that includes the approximation of binomial coefficients as a special case is presented in [27], which we sketch here; compare the original text for some technical conditions.

Let

$$g(z)=\sum_{k\ge 0}p_k z^k$$

be a probability generating function, and its Thiele expansion given as

$$g(e^t)=\exp\Big(\mu t+\frac{\sigma^2t^2}{2!}+\frac{\kappa_3t^3}{3!}+\frac{\kappa_4t^4}{4!}+\cdots\Big),$$

where $\mu=g'(1)$ and $\sigma^2=g''(1)+g'(1)-(g'(1))^2$ are expectation and variance. The constants κ_n are called *cumulants*. Further, let

$$A_{n,k}=[z^{\mu n+k}]g(z)^n,$$

where k is chosen to make the exponent an integer, which is "not too far away" from the expected value μn.

For our application, we choose $g(z) = \frac{1+z}{2}$, so that $\mu = \frac{1}{2}$ and $\sigma^2 = \frac{1}{4}$. Then

$$A_{2n,k} = \frac{\binom{2n}{n+k}}{2^{2n}};$$

it does not matter whether we normalize by 2^{2n} or $\binom{2n}{n}$, since, by Stirling's formula (6),

$$\binom{2n}{n} 2^{-2n} = \frac{1}{\sqrt{\pi n}} - \frac{1}{8\sqrt{\pi} n^{3/2}} + O(n^{-5/2}),$$

where any number of terms would be available.

The answer in [27] is:

$$A_{n,k} = \frac{1}{\sigma\sqrt{2\pi n}} \exp\Big(-\frac{k^2}{2\sigma^2 n}\Big)\Big[1 - \frac{\kappa_3}{2\sigma^4}\frac{k}{n} + \frac{\kappa_3}{6\sigma^6}\frac{k^3}{n^2}\Big] + O(n^{-3/2});$$

there is also a general version given, including terms of the form k^K/n^N, which we do not reproduce here.

12. Mellin transform and asymptotics of harmonic sums

This section is based on the survey [16]. For details and proofs we refer to this classical paper.

The Mellin transform associates to a function $f(x)$ defined over the positive reals the complex function $f^*(s)$ where

$$f^*(s) = \int_0^\infty f(x) x^{s-1} dx;$$

it is a close relative to the Laplace and Fourier transform.

The major use of the Mellin transform examined here is for the asymptotic analysis of sums obeying the general pattern

$$F(x) = \sum_k \lambda_k f(\mu_k x),$$

either as $x \to 0$ or as $x \to \infty$. Sums of this type are called harmonic sums; $f(x)$ is called the base function.

Harmonic sums surface at many places in combinatorial mathematics as well as in the analysis of algorithms and data structures. De Bruijn and Knuth are responsible in an essential way for introducing the Mellin transform in this range of problems, as attested by Knuth's account in [32] and the classic paper [10] which have been the basis of many later combinatorial applications.

By a simple change of variable, we find

$$F^*(s) = \sum_k \lambda_k \mu_k^{-s} \cdot f^*(s).$$

It is this factorization property that makes the Mellin transform useful for harmonic sums. The Mellin transform exists in a fundamental strip $\langle -u, -v\rangle := \{z \in \mathbb{C} \mid -u < \Re z < -v\}$, if

$$f(x) = O(x^u) \quad \text{as } x \to 0, \qquad f(x) = O(x^v) \quad \text{as } x \to \infty.$$

As the integral defining $f^*(s)$ depends analytically on the complex parameter s, a Mellin transform is in addition analytic in its fundamental strip.

For instance, the function $f(x) = (1+x)^{-1}$ is $O(x^0)$ at 0 and $O(x^{-1})$ at infinity, hence a guaranteed existence strip for $f^*(s)$ is $(0, 1)$, which here coincides with the fundamental strip. In this case, the Mellin transform may be found from the classical Beta integral to be $f^*(s) = \frac{\pi}{\sin \pi s}$ which is analytic in $\langle 0, 1\rangle$.

The function $f(x) = e^{-x}$ satisfies

$$e^{-x} \sim 1 \quad \text{as } x \to 0, \qquad e^{-x} = O(x^{-b}) \quad \text{for any } b \text{ as } x \to \infty.$$

Therefore the Mellin transform

$$f^*(s) = \int_0^\infty e^{-x} x^{s-1} dx = \Gamma(s)$$

exists in $\langle 0, \infty\rangle$ and is analytic there.

Let $H(x)$ be the step function defined by $H(x) = [0 \le x < 1]$. Then $H^*(s) = \frac{1}{s}$ in $\langle 0, \infty\rangle$.

The following list covers the essential functional properties.

$f(x)$	$f^*(s)$	$\langle \alpha, \beta\rangle$
$x^\nu f(x)$	$f^*(s+\nu)$	$\langle \alpha-\nu, \beta-\nu\rangle$
$f(x^\rho)$	$\frac{1}{\rho} f^*(\frac{s}{\rho})$	$\langle \alpha\rho, \beta\rho\rangle \quad (\rho > 0)$
$f(\frac{1}{x})$	$-f^*(-s)$	$\langle -\beta, -\alpha\rangle$
$f(\mu x)$	$\mu^{-s} f^*(s)$	$\langle \alpha, \beta\rangle, \ \ (\mu > 0)$
$\sum_k \lambda_k f(\mu_k x)$	$\sum_k \lambda_k \mu_k^{-s} \cdot f^*(s)$	
$f(x) \log x$	$\frac{d}{ds} f^*(s)$	$\langle \alpha, \beta\rangle$
$x \frac{d}{dx} f(x)$	$-s f^*(s)$	
$\frac{d}{dx} f(x)$	$-(s-1) f^*(s-1)$	
$\int_0^\infty f(t)dt$	$-\frac{1}{s} f^*(s+1)$	

The empty ranges depend on the situation. They are usually the intersection of the strips of the ingredients, like the Dirichlet-like series and the strip for the base function. For instance, let

$$F(x) = \frac{1}{e^x - 1} = e^{-x} + e^{-2x} + e^{-3x} + \cdots,$$

then

$$F^*(s) = (1 + 2^{-s} + 3^{-s} + \cdots)\Gamma(s) = \zeta(s)\Gamma(s).$$

It is valid in the intersection of $\langle 1, \infty \rangle$ (ζ-function) and $\langle 0, \infty \rangle$ (Γ-function), so it is $\langle 1, \infty \rangle$.

Here is a little list of some common Mellin transforms:

e^{-x}	$\Gamma(s)$	$\langle 0, \infty \rangle$
$e^{-x} - 1$	$\Gamma(s)$	$\langle -1, 0 \rangle$
$e^{-x} - 1 + x$	$\Gamma(s)$	$\langle -2, -1 \rangle$
e^{-x^2}	$\frac{1}{2}\Gamma(\frac{s}{2})$	$\langle 0, \infty \rangle$
$\frac{1}{1+x}$	$\frac{\pi}{\sin \pi s}$	$\langle 0, 1 \rangle$
$\log(1 + x)$	$\frac{\pi}{s \sin \pi s}$	$\langle -1, 0 \rangle$
$H(x) \equiv [0 \leq x < 1]$	$\frac{1}{s}$	$\langle 0, \infty \rangle$
$x^\alpha (\log x)^k H(x)$	$\frac{(-1)^k k!}{(s+\alpha)^{k+1}}$	$\langle -\alpha, \infty \rangle$

Theorem 12.1 (Mellin's inversion theorem)**.** *(i) Let $f(x)$ be integrable with fundamental strip $\langle \alpha, \beta \rangle$. If c is such that $\alpha < c < \beta$ and $f^*(c+it)$ is integrable, then the equality*

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f^*(s) x^{-s} ds = f(x)$$

holds almost everywhere. Moreover, if $f(x)$ is continuous, then equality holds everywhere on $(0, \infty)$.

(ii) Let $f(x)$ be locally integrable with fundamental strip $\langle \alpha, \beta \rangle$ and be of bounded variation in a neighborhood of x_0. Then, for any c in the interval (α, β),

$$\lim_{T \to \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} f^*(s) x^{-s} ds = \frac{f(x_0^+) + f(x_0^-)}{2}.$$

Recall that a function of bounded variation is a real-valued function whose total variation is bounded; the total variation of a real-valued function $f(x)$, defined on an interval $[a, b]$ is the quantity

$$\sup_P \sum_{0 \leq i < n_P} |f(x_{i+1}) - f(x_i)|,$$

taken over all partitions $P = (x_0, \ldots, x_{n_P})$ of the interval $[a, b]$. $\square$

We need the notion of a singular expansion. Let $\phi(s)$ be meromorphic in Ω, and let $\mathcal{S}$ include all the poles of $\phi(s)$ in Ω. A singular expansion of $\phi(s)$ in Ω is a formal sum of singular elements of $\phi(s)$ at all points of $\mathcal{S}$. When E is a singular expansion of $\phi(s)$ in Ω, we write $\phi(s) \asymp E$, $s \in \Omega$. This formal expansion is a concise way of combining information contained in the Laurent expansions of the function $\phi(s)$ at various points. Example:

$$\Gamma(s) \asymp \sum_{k \ge 0} \frac{(-1)^k}{k!} \frac{1}{s+k}$$

is the singular expansion in the whole of $\mathbb{C}$.

Theorem 12.2 (Direct mapping)**.** *Let $f(x)$ have a transform $f^*(s)$ with nonempty fundamental strip $\langle \alpha, \beta \rangle$.*

(i) Assume that $f(x)$ admits as $x \to 0^+$ a finite asymptotic expansion of the form

$$f(z) = \sum_{(\xi,k) \in A} c_{\xi,k} x^{\xi} (\log x)^k + O(x^{\gamma}), \tag{5}$$

where the ξ satisfy $-\gamma < -\xi \le \alpha$ and the k are non-negative. Then $f^*(s)$ is continuable to a meromorphic function in the strip $\langle -\gamma, \beta \rangle$ where it admits the singular expansion

$$f(z) \asymp \sum_{(\xi,k) \in A} c_{\xi,k} \frac{(-1)^k k!}{(s+\xi)^{k+1}} + O(x^{\gamma}), \qquad s \in \langle -\gamma, \beta \rangle.$$

(ii) Similarly, assume that $f(x)$ admits as $x \to \infty$ a finite asymptotic expansion of the form (5) where now $\beta \le -\xi < -\gamma$. Then $f^*(s)$ is continuable to a meromorphic function in the strip $\langle \alpha, -\gamma \rangle$ where

$$f(z) \asymp -\sum_{(\xi,k) \in A} c_{\xi,k} \frac{(-1)^k k!}{(s+\xi)^{k+1}} + O(x^{\gamma}), \qquad s \in \langle \alpha, -\gamma \rangle.$$

Thus terms in the asymptotic expansion of $f(x)$ at 0 induce poles of $f^*(s)$ in a strip to the left of the fundamental strip; terms in the expansion at ∞ induce poles in a strip to the right.

This principle is general: subtracting from a function a truncated form of its asymptotic expansion at either 0 or ∞ does not alter its Mellin transform and only shifts the fundamental strip. An instance is provided by the functions e^{-x}, $e^{-x}-1$, $e^{-x}-1+x$, all having the Mellin transform $\Gamma(s)$, in different strips.

Under a set of mild conditions, a converse to the Direct Mapping theorem also holds: The singularities of a Mellin transform which is small enough towards $i\infty$ encode the asymptotic properties of the original function. See [16] for a precise statement.

The Mellin summation formula is

$$\sum_k \lambda_k f(\mu_k x) \sim \pm \sum_{s\in H} \operatorname{Res}(f^*(s)\Lambda(s)x^{-s}),$$

with $\Lambda(s) = \sum_k \lambda_k \mu_k^{-s}$.

For an expansion near 0, the sum is over the set H of poles to the left of the fundamental strip, and the sign is $+$.

For an expansion near ∞, the sum is over the set H of poles to the right of the fundamental strip, and the sign is $-$.

Example. Harmonic numbers. We refer here to Section 9 about properties of the zeta function and the Bernoulli numbers. We write

$$h(x) = \sum_{k\geq 1}\Big(\frac{1}{k} - \frac{1}{k+x}\Big) = \sum_{k\geq 1}\frac{1}{k}\frac{x/k}{1+x/k}$$

and notice that $H_n = h(n)$ and that $h(x)$ is a harmonic sum, with $\lambda_k = \mu_k = \frac{1}{k}$. We have

$$h^*(s) = -\frac{\pi}{\sin \pi s}\zeta(s),$$

and the fundamental strip is $\langle -1, 0\rangle$. The singular expansion to the right of this fundamental strip is

$$h^*(s) \asymp \Big(\frac{1}{s^2} - \frac{\gamma}{s}\Big) - \sum_{k\geq 1}(-1)^k\frac{\zeta(1-k)}{s-k}.$$

Hence

$$H_n \sim \log n + \gamma + \frac{1}{2n} + \sum_{k\geq 2}\frac{(-1)^k B_k}{k}\frac{1}{n^k}.$$

The dominant terms come from the expansion at 0. We have to take the residue (with a negative sign) of

$$-\frac{\pi}{\sin \pi s}\zeta(s)x^{-s},$$

which is

$$[s^{-1}]\Big(-s^{-2} + (\log x + \gamma)s^{-1} + \cdots\Big) = \log x + \gamma.$$

The full expansion for the harmonic numbers is

$$H_n \sim \log n + \gamma + \frac{1}{2n} - \sum_{k\geq 1}\frac{B_{2k}}{2kn^{2k}}.$$

Example. Stirling's formula for the Γ-function. From the product decomposition of the Gamma function, one has

$$l(x) = \Gamma(x+1) - \gamma x = \sum_{n\geq 1}\Big[\frac{x}{n} - \log\Big(1+\frac{x}{n}\Big)\Big],$$

One computes

$$l^*(s) = -\zeta(-s)\frac{\pi}{s\sin(\pi s)},$$

with fundamental strip $\langle -2, -1\rangle$. There are double poles at $s = -1$, $s = 0$ and simple poles at the positive integers. The main contribution is

$$[(s+1)^{-1}]\zeta(-s)\frac{\pi}{s\sin(\pi s)}x^{-s} = x\log x + x(\gamma - 1).$$

The full expansion is

$$\log(x!) \sim \log\left(x^x e^{-x}\sqrt{2\pi x}\right) + \sum_{n\ge 1}\frac{B_{2n}}{2n(2n-1)}\frac{1}{x^{2n-1}}. \tag{6}$$

Example. A divisor sum. Consider

$$D(x) = \sum_{k\ge 1} d(k)e^{-kx},$$

where $d(k)$ is the number of (positive) divisors of k. Since

$$\sum_{k\ge 1}\frac{d(k)}{k^s} = \zeta^2(s),$$

we find

$$D^*(s) = \Gamma(s)\zeta^2(s),$$

and the main contribution of the asymptotic expansion at $x = 0$ is given by

$$[(s-1)^{-1}]\Gamma(s)\zeta^2(s)x^{-s} = \frac{-\log x + \gamma}{x}.$$

Example. Height of planar trees

We have seen (3) that the probability that a Dyck path of length $2n$ has height $\ge h-1$ is given by

$$\sum_{k\ge 1}\frac{\binom{2n}{n+1-hk} - 2\binom{2n}{n-hk} + \binom{2n}{n-1-hk}}{\frac{1}{n+1}\binom{2n}{n}}.$$

Because of the fundamental correspondence between Dyck paths and planar trees this enumerates as well the family of planar trees with $n+1$ nodes and height $\ge h$. In order to compute the expectation, one has to sum this on $h \ge 1$, which leads to

$$E_{n+1} = (n+1)\sum_{k\ge 1} d(k)\frac{\binom{2n}{n+1-k} - 2\binom{2n}{n-k} + \binom{2n}{n-1-k}}{\binom{2n}{n}}.$$

Here, again, $d(k)$ is the number of (positive) divisors of the integer k. Approximation of the second difference of binomial coefficients as described earlier (Section 11) leads to the approximation

$$E_{n+1} \sim n \sum_{k \geq 1} d(k) e^{-k^2/n} \Big(\frac{4k^2}{n} - 2 \Big).$$

The asymptotic evaluation of this series is a typical application of the Mellin transform. With $x = 1/\sqrt{n}$, we need to estimate

$$\sum_{k \geq 1} d(k) e^{-k^2 x^2} (4k^2 x^2 - 2),$$

which is a harmonic sum, and its Mellin transform is

$$\sum_{k \geq 1} d(k) k^{-s} \cdot \int_0^\infty e^{-x^2} (4x^2 - 2) x^{s-1} dx = \zeta^2(s)(s-1)\Gamma\Big(\frac{s}{2}\Big).$$

Using the inversion formula, this sum can be written as

$$\frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \zeta^2(s)(s-1)\Gamma\Big(\frac{s}{2}\Big) x^{-s} ds.$$

The residue at $s = 1$ (after shifting the line of integration to the left) produces $\frac{\sqrt{\pi}}{x}$, which is $\sqrt{\pi n}$, which is the asymptotic equivalent of the average height of a planar trees with n nodes. For more details, compare [24] and the literature cited therein.

Example. A doubly exponential sum and periodicities. The prototype of harmonic sums with a fluctuating behavior is the function

$$F(x) = \sum_{k \geq 0} e^{-x 2^k},$$

whose behavior is sought as $x \to 0$. The Mellin transform is

$$\sum_{k \geq 0} 2^{-ks} \Gamma(s) = \frac{1}{1 - 2^{-s}} \Gamma(s),$$

and the fundamental strip is $\langle 0, \infty \rangle$. The poles to the left of the strip are at $s = 0$ (double) and at $s = \chi_k := \frac{2\pi i k}{\log 2}$ (simple). Now, as $s \sim 0$,

$$\frac{1}{1 - 2^{-s}} \Gamma(s) x^{-s} \sim \frac{1}{(\log 2)^2 s^2} + \frac{1}{s} \Big(\frac{1}{2} - \log_2 x - \frac{\gamma}{\log 2} \Big) + \cdots,$$

and thus, as $x \to 0$,

$$F(x) \sim \frac{1}{2} - \log_2 x - \frac{\gamma}{\log 2} + \frac{1}{\log 2} \sum_{k \neq 0} \Gamma(-\chi_k) e^{2\pi i k \log_2 x}.$$

The series is a Fourier series and represents a periodic function. Since the Gamma-function is small along the imaginary axis [49], as mentioned earlier in Section 9, the amplitude of this function is small.

13. The Mellin-Perron formula

The following treatment is borrowed from [18]. We start from the Mellin inversion formula

$$\sum_k \lambda_k f(\mu_k x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\Big(\sum_k \lambda_k \mu^{-s}\Big) f^*(s) x^{-s} ds, \tag{7}$$

where c is in the fundamental strip. Introduce the step function

$$H_0(x) = \begin{cases} 1 & \text{if } x \in [0,1], \\ 0 & \text{otherwise,} \end{cases}$$

also written as $H_0(x) = [0 \le x \le 1]$, together with $H_m(x) = H_0(x)(1-x)^m$. Then we get

Theorem 13.1. *Let $c > 0$ lie in the half-plane of absolute convergence of $\sum_k \lambda_k \mu_k^{-s}$. Then, for any $m \ge 1$, we have*

$$\frac{1}{m!}\sum_{1\le k<n} \lambda_k \Big(1-\frac{k}{n}\Big)^m = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_k \frac{\lambda_k}{\mu_k^s}\cdot n^s \frac{ds}{s(s+1)\dots(s+m)}.$$

For $m = 0$,

$$\sum_{1\le k<n} \lambda_k + \frac{\lambda_n}{2} = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \sum_k \frac{\lambda_k}{\mu_k^s}\cdot n^s \frac{ds}{s}.$$

This formula is obtained by setting $x := 1/n$, $f(x) := H_m(x)$ in the Mellin inversion formula (7) and noticing that

$$H_m^*(s) = \frac{m!}{s(s+1)\dots(s+m)}.$$

For $m = 0$, the formula has to be modified slightly by taking a principal value, since $H_0(x)$ is discontinuous at $x = 1$. See [3] for a direct proof of this instance.

For example, for $\lambda_k \equiv 1$, $\mu_k \equiv k$ and $m = 1$, we get

$$\sum_{1\le k<n}\Big(1-\frac{k}{n}\Big) = \frac{n-1}{2} = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \zeta(s) n^s \frac{ds}{s(s+1)}.$$

Shifting the line of integration to the left and taking the poles at $s = 1$ and $s = 0$ into account (note that $\zeta(0) = -\frac{1}{2}$), we get

$$0 = \frac{1}{2\pi i}\int_{-\frac14-i\infty}^{-\frac14+i\infty} \zeta(s) n^s \frac{ds}{s(s+1)}. \tag{8}$$

This formula is the basis of some *exact* formulæ.

We apply the Mellin-Perron machinery now to the binary sum of digits function. Let $\nu(k)$ be the number of digits 1 in the binary expansion of the integer k. Furthermore, let $v_2(k)$ be the exponent of 2 in the prime decomposition of k (so, if $k = 2^i(2j+1)$, then $v_2(k) = i$). Then we notice that the binary representation looks like $***10^i$ (for k), and $***01^i$ (for $k-1$). Taking differences, we see that $\nu(k) - \nu(k-1) = 1 - v_2(k)$. Summing this on k, we see that

$$\nu(k) = k - \sum_{j \le k} v_2(j).$$

We will study the summatory function

$$S(n) := \sum_{k<n} \nu(k) = \frac{n(n-1)}{2} - \sum_{j \le k < n} v_2(j) = \binom{n}{2} - \sum_{j \le n} (n-j) v_2(j).$$

Now we compute

$$V(s) = \sum_{k \ge 1} \frac{v_2(k)}{k^s} = \sum_{k \ge 0} \frac{v_2(2k+1)}{(2k+1)^s} + \sum_{k \ge 1} \frac{v_2(2k)}{(2k)^s} = \frac{1}{2^s} \sum_{k \ge 1} \frac{1 + v_2(k)}{k^s},$$

or

$$V(s) = \frac{1}{2^s}\zeta(s) + \frac{1}{2^s} V(s) \quad \Longrightarrow \quad V(s) = \frac{\zeta(s)}{2^s - 1}.$$

Thus

$$S(n) = \frac{n(n-1)}{2} - \frac{n}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{\zeta(s)}{2^s - 1} n^s \frac{ds}{s(s+1)}.$$

The integrand has a simple pole at $s = 1$, a double pole at $s = 0$ and simple poles at $s = \chi_k := \frac{2\pi i k}{\log 2}$. Shifting the line of integration[2] to $\Re(s) = -\frac{1}{4}$ and taking residues into account, we get

$$S(n) = \frac{1}{2} n \log_2 n + n F_0(\log_2 n) - n R(n),$$

where the Fourier series akin to F_0 occurs as the sum of residues of the integrand at the imaginary poles $s = \chi_k$. The remainder term is

$$R(n) = \frac{1}{2\pi i} \int_{-\frac{1}{4}-i\infty}^{-\frac{1}{4}+i\infty} \frac{\zeta(s)}{2^s - 1} n^s \frac{ds}{s(s+1)}.$$

We are going to prove that $R(n) \equiv 0$ whenever n is an integer. The integral converges since $|\zeta(-\frac{1}{4} + it)| \ll |t|^{3/4}$ (cf. [49]). Using the expansion

$$\frac{1}{2^s - 1} = -1 - 2^s - 2^{2s} - 2^{3s} - \cdots$$

---

[2]Technically, we integrate along a rectangle with upper and lower sides passing through $(2N+1)\pi i/\log 2$, respectively, and let $N \to \infty$. Because of growth properties of the zeta function, the contribution along the horizontal segments vanishes. This also proves directly that the sum of residues at the complex points (which gives the Fourier series) converges.

in the integral, which is legitimate since now $\Re(s) < 0$, we find that $R(n)$ is a sum of terms of the form

$$\frac{1}{2\pi i}\int_{-\frac14-i\infty}^{-\frac14+i\infty}\zeta(s)(2^k n)^s\frac{ds}{s(s+1)},$$

and each of these terms is 0 by virtue of (8).

This proves a result originally due to Delange [11].

Theorem 13.2. *The sum-of-digits function $S(n)$ satisfies*

$$S(n)=\frac12 n\log_2 n+nF_0(\log_2 n),$$

where $F_0(x)$ is representable as a Fourier series $F_0(x)=\sum_{k\in\mathbb{Z}} f_k e^{2\pi ikx}$ and

$$f_0=\frac{\log_2\pi}{2}-\frac{1}{2\log 2}-\frac34,$$
$$f_k=-\frac{1}{\log 2}\frac{\zeta(\chi_k)}{\chi_k(\chi_k+1)}\quad \textit{for}\quad \chi_k=\frac{2\pi ik}{\log 2},\ k\neq 0.$$

It is interesting to compare this derivation with the elementary and pretty arguments provided by Delange [11]: We use the notation $x=\lfloor x\rfloor+\{x\}$ with integer part $\in\mathbb{Z}$ and fractional part $0\le\{x\}<1$ and write m in binary as $(\dots a_2a_1a_0)_2$; it is not hard to see that

$$a_k=\Big\lfloor\frac{m}{2^k}\Big\rfloor-2\Big\lfloor\frac{m}{2^{k+1}}\Big\rfloor=\int_m^{m+1}\bigg(\Big\lfloor\frac{t}{2^k}\Big\rfloor-2\Big\lfloor\frac{t}{2^{k+1}}\Big\rfloor\bigg)dt.$$

Therefore, with $l=\lfloor\log_2 n\rfloor$,

$$S(n)=\sum_{k=0}^{l}\int_0^n\bigg(\Big\lfloor\frac{t}{2^k}\Big\rfloor-2\Big\lfloor\frac{t}{2^{k+1}}\Big\rfloor\bigg)dt=\sum_{k=0}^{l}2^{k+1}\int_0^{n/2^{k+1}}\big(\lfloor 2u\rfloor-2\lfloor u\rfloor\big)du.$$

Introducing $g(u)=\lfloor 2u\rfloor-2\lfloor u\rfloor-\frac12$, then

$$\begin{aligned}S(n)&=\frac{n(l+1)}{2}+\sum_{k=0}^{l}2^{k+1}g\Big(\frac{n}{2^{k+1}}\Big)\\&=\frac{n\log_2 n}{2}+n\frac{1-\{\log_2 n\}}{2}+n2^{1-\{\log_2 n\}}\sum_{k\ge0}2^{-k}g(2^{\{\log_2 n\}-1}\cdot 2^k);\end{aligned}$$

the last step was the change $k:=l-k$ and noticing that $g(x)=0$ whenever x is an integer. Introducing

$$h(x)=\sum_{k\ge0}2^{-k}g(x\cdot 2^k),$$

this reads as

$$S(n) = \frac{n \log_2 n}{2} + n\frac{1 - \{\log_2 n\}}{2} + n2^{1-\{\log_2 n\}}h(2^{\{\log_2 n\}-1}).$$

The final step is to introduce

$$F_0(x) = \frac{1 - \{x\}}{2} + 2^{1-\{x\}}h(2^{\{x\}-1}),$$

then

$$S(n) = \frac{n \log_2 n}{2} + nF_0(\log_2 n).$$

It can be shown that $F_0(x)$ is periodic and continuous, and the Fourier coefficients can be computed as well.

We want to finish the discussion of digital properties by considering the Gray code and the sum-of-digits function of it, using again the Mellin-Perron technique.

The Gray code representation of the integers starts like

$$0, 1, 11, 10, 110, 111, 101, 100, 1100, 1101, \ldots;$$

its characteristic is that the representations of n and $n+1$ differ in exactly one binary position, and it is constructed in a simple manner by reflections based on powers of two. Let $\gamma(k)$ be the number of 1-digits in the Gray code representation of k, and $\delta_k = \gamma(k) - \gamma(k-1)$. It is easy to see that $\delta_{2k} = \delta_k$, and the pattern for odd values is $\delta_{2k+1} = (-1)^k$. Thus the Dirichlet series $\delta(s)$ relative to the sequence δ_k is given by

$$\delta(s) = \sum_{k \geq 1} \frac{\delta_k}{k^s} = \frac{2^s}{2^s - 1} \sum_{k \geq 0} \frac{(-1)^k}{(2k+1)^s}.$$

So the summatory function $G(n) = \sum_{k<n} \gamma(k)$ of the sum-of-digits function of the Gray code representation can be expressed via

$$\frac{n}{2\pi i} \int_{2-i\infty}^{2+i\infty} \delta(s) n^s \frac{ds}{s(s+1)};$$

and the rest of the analysis is very similar to before, and leads to the explicit result

$$G(n) = \frac{1}{2} n \log_2 n + nF_1(\log_2 n),$$

with explicit Fourier coefficients of $F(x)$. We refer for details to the fundamental paper [18].

We would like to mention the Hurwitz zeta function [49]

$$\zeta(s, a) := \sum_{n \geq 0} \frac{1}{(n+a)^s} \qquad \text{for } 0 < a \leq 1 \text{ and } \Re s > 1.$$

Then

$$\sum_{k\ge 0}\frac{(-1)^k}{(2k+1)^s}=\sum_{k\ge 0}\frac{1}{(4k+1)^s}-\sum_{k\ge 0}\frac{1}{(4k+3)^s}=\frac{1}{4^s}\zeta(s,\tfrac14)-\frac{1}{4^s}\zeta(s,\tfrac34).$$

This shows that $\delta(s)$ can be expressed via $\zeta(z,a)$.

14. Mellin-Perron formula: divide-and-conquer recursions

The aim of this section, as described in [15], is to solve the divide-and-conquer recursion

$$f_n=f_{\lfloor n/2\rfloor}+f_{\lceil n/2\rceil}+e_n,$$

where e_n is a given sequence. The initial conditions are chosen to be $e_0=e_1=f_0=0$. The tool will be a special case ($m=2$) of the Mellin-Perron formula as discussed before:

Lemma 14.1. *Let w_n be a sequence and $W(s)=\sum_{n\ge1}w_n/n^s$ its generating Dirichlet series. Assume that $W(s)$ converges absolutely for $\Re(s)>2$. Then*

$$\frac{n}{2\pi i}\int_{3-i\infty}^{3+i\infty}W(s)n^s\frac{ds}{s(s+1)}=\sum_{k=1}^{n}(n-k)w_k.$$

The recursion is now written for even and odd indices separately:

$$\begin{aligned}f_{2m}&=2f_m+e_{2m},\\ f_{2m+1}&=f_m+f_{m+1}+e_{2m+1},\end{aligned}$$

which holds for $m\ge1$. Taking backward differences with $\nabla f_n=f_n-f_{n-1}$ and $\nabla e_n=e_n-e_{n-1}$ yields

$$\begin{aligned}\nabla f_{2m}&=\nabla f_m+\nabla e_{2m},\\ \nabla f_{2m+1}&=\nabla f_{m+1}+\nabla e_{2m+1}\end{aligned}$$

for $m\ge1$. Now we take forward differences. Note that $\Delta\nabla f_m=\Delta(f_m-f_{m-1})=f_{m+1}-2f_m+f_{m-1}$. So

$$\begin{aligned}\Delta\nabla f_{2m}&=\Delta\nabla f_m+\Delta\nabla e_{2m},\\ \Delta\nabla f_{2m+1}&=\Delta\nabla e_{2m+1}\end{aligned}$$

for $m\ge1$, with $\Delta\nabla f_1=f_2-2f_1=e_2=\Delta\nabla e_1$. Now set $w_n=\Delta\nabla f_n$ and its Dirichlet generating function $W(s)=\sum_{n\ge1}w_nn^{-s}$. From the recursion we get by summing

$$W(s)=\sum_{m\ge1}\frac{\Delta\nabla f_m}{(2m)^s}+\Delta\nabla f_1+\sum_{m\ge2}\frac{\Delta\nabla e_m}{m^s}=\frac{W(s)}{2^s}+\sum_{m\ge1}\frac{\Delta\nabla e_m}{m^s},$$

or

$$W(s) = \frac{1}{1-2^{-s}} \sum_{m\ge 1} \frac{\Delta\nabla e_m}{m^s}.$$

It is easy to check that

$$\sum_{k=1}^{n}(n-k)\Delta\nabla f_k = f_n - nf_1.$$

We assume that $e_n = O(n)$. The Mellin-Perron formula thus gives us

$$f_n = nf_1 + \frac{n}{2\pi i}\int_{3-i\infty}^{3+i\infty} \frac{\Xi(s)n^s}{1-2^{-s}} \frac{ds}{s(s+1)},$$

where

$$\Xi(s) = \sum_{n\ge 1} \frac{\Delta\nabla e_n}{n^s}.$$

The growth condition on e_n ensures that this Dirichlet series converges for $\Re(s) > 2$, as required.

This is an exact formula. Asymptotics can be derived from it as in previous instances, by shifting the line of integration to the left, and taking residues into account.

Example. We study the “worst case of the number of comparisons in mergesort,” which, according to [15] and the literature cited therein is given by our recursion, for $e_n = n-1$ and $f_1 = 0$, which implies $\Delta\nabla f_1 = e_2 = 1$ and $\Delta\nabla e_n = 0$ for $n \ge 2$. Then

$$\frac{f_n}{n} = \frac{1}{2\pi i}\int_{3-i\infty}^{3+i\infty} \frac{n^s}{1-2^{-s}} \frac{ds}{s(s+1)}.$$

The residue calculations involve a double pole at $s = 0$ and simple poles at $s = \chi_k = 2\pi ik/\log 2$. Then one gets

$$f_n = n\log_2 n + nA(\log_2 n) + O(\sqrt{n}\,),$$

with

$$A(x) = \sum_{k\in\mathbb{Z}} a_k e^{2\pi ikx}$$

and

$$a_0 = \frac{1}{2} - \frac{1}{\log 2} \quad \text{and} \quad a_k = \frac{1}{\log 2}\frac{1}{\chi_k(\chi_k+1)} \text{ for } k \ne 0.$$

For more details and more examples we refer to [15].

15. Rice's method

Rice's method made its first appearence as exercise 5.2.2–54 in [32]: It allows to write the alternating sum

$$\sum_{k=0}^{n}\binom{n}{k}(-1)^k f_k$$

as a contour integral

$$\frac{1}{2\pi i}\int_{\mathcal{C}}\frac{(-1)^n n!}{z(z-1)\dots(z-n)}f(z)dz,$$

where $f(z)$ is an analytic function that extrapolates the sequence f_k ($f(k)=f_k$), and $\mathcal{C}$ encircles the interval $[0..n]$. As was pointed out in [23], such integrals can be traced back to Nörlund [39]. The advantage of this representation is that, by extending the contour of integration, asymptotic equivalents can be derived, usually by collecting additional residues. In many cases, the integral disappears when the contour goes to infinity, leading to identities. Furthermore, the alternating sum usually involves heavy cancellations, and is not always easy to analyze in a direct way.

Consider the difference operator Δ, defined by $\Delta f_k = f_{k+1}-f_k$. Then it is a standard exercise by induction to prove that

$$\Delta^n f_0 = \sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}f_k.$$

The transformation of sequences

$$f_n \mapsto g_n = \sum_{k=0}^{n}\binom{n}{k}(-1)^k f_k$$

is an involution, called Euler transform; if $F(z)$ and $G(z)$ are the ordinary generating functions of the sequences f_n and g_n, then

$$\begin{aligned}
G(z) &= \sum_{0\le k\le n} z^n\binom{n}{k}(-1)^k f_k = \sum_{0\le k}(-1)^k f_k\sum_{n\ge k}\binom{n}{k}z^n\\
&= \sum_{0\le k}(-1)^k f_k\frac{z^k}{(1-z)^{k+1}} = \frac{1}{1-z}F\Big(\frac{z}{z-1}\Big).
\end{aligned}$$

If $f(z)$ and $g(z)$ are the exponential generating functions, then

$$g(z)=\sum_{0\le k\le n}\frac{z^n}{n!}\binom{n}{k}(-1)^k f_k = \sum_{k\ge 0}\frac{(-z)^k f_k}{k!}\sum_{n-k\ge 0}\frac{z^{n-k}}{(n-k)!} = e^z f(-z).$$

We state the integral formula as a lemma.

Lemma 15.1. *Lef $f(z)$ be analytic in a domain that contains the half-line $[n_0,\infty)$. Then, the differences of the sequence (f_n) admit the integral representation*

$$\sum_{k=n_0}^{n}(-1)^k\binom{n}{k}f_k=\frac{(-1)^n}{2\pi i}\int_{\mathcal{C}}\frac{n!}{z(z-1)\dots(z-n)}f(z)dz, \tag{9}$$

where $\mathcal{C}$ is a positively oriented curve that lies in the domain of analyticity of $f(z)$, encircles $[n_0,n]$, and does not include any of the integers $0,1,\dots,n_0-1$.

Proof. This is a direct application of residue calculus:

$$\begin{aligned}\mathrm{Res}_{z=k}\frac{n!}{z(z-1)\dots(z-n)}f(z)&=\frac{n!}{k(k-1)\dots1(-1)(-2)\dots(k-n)}f_k\\&=(-1)^{n-k}\binom{n}{k}f_k.\quad\square\end{aligned}$$

The kernel in (9) can be expressed by Gamma functions:

$$\frac{n!}{z(z-1)\dots(z-n)}=\frac{\Gamma(n+1)\Gamma(z-n)}{\Gamma(z+1)}=\frac{(-1)^{n+1}\Gamma(n+1)\Gamma(-z)}{\Gamma(n+1-z)}.$$

We will deal now with rational and meromorphic functions $f(z)$. It turns out that the differences of the sequence (f_n) can be expressed via collecting residues.

Theorem 15.1 (Rational functions)**.** *Let $f(z)$ be a rational function analytic on $[n_0,\infty)$ Then, except for a finite number of values of n, one has*

$$\sum_{k=n_0}^{n}(-1)^k\binom{n}{k}f_k=(-1)^{n-1}\mathrm{Res}_z\frac{n!f(z)}{z(z-1)\dots(z-n)}$$

where the sum is extended to all poles z of $f(z)/(z(z-1)\dots(z-n))$ not on $[n_0,\infty)$.

Proof. The idea is to integrate along a large circle with radius R and use trivial bounds. The details are in [23]. $\square$

The first example that we treat is

$$S_n(m)=\sum_{k=1}^{n}\binom{n}{k}\frac{(-1)^k}{k^m}$$

for m a positive integer. Note that for negative m, this is basically a Stirling subset number. In this example, $f(z)=z^{-m}$ is a rational function, and there is only one additional pole at $z=0$. So, we get (compare again [23])

$$S_n(m)=(-1)^{n-1}\mathrm{Res}_{z=0}\frac{n!}{z^{m+1}(z-1)\dots(z-n)}$$

$$
\begin{aligned}
&= (-1)^{n-1}[z^{-1}]\frac{n!}{z^{m+1}(z-1)\dots(z-n)}\\
&= (-1)^{n-1}[z^{m}]\frac{n!}{(z-1)\dots(z-n)} = -[z^m]\frac{1}{(1-\frac{z}{1})\dots(1-\frac{z}{n})}\\
&= -[z^m]\exp\left\{\sum_{k=1}^{n}\log\left(\frac{1}{1-\frac{z}{k}}\right)\right\} = -[z^m]\exp\left\{\sum_{k=1}^{n}\sum_{j\ge1}\frac{1}{j}\left(\frac{z}{k}\right)^j\right\}\\
&= -[z^m]\exp\left\{\sum_{j\ge1}\frac{1}{j}z^jH_n^{(j)}\right\} = -[z^m]\prod_{j\ge1}\exp\left\{\frac{1}{j}z^jH_n^{(j)}\right\}\\
&= -\sum_{1l_1+2l_2+3l_3+\dots=m}\frac{1}{l_1!l_2!l_3!\dots}\left(\frac{H_n^{(1)}}{1}\right)^{l_1}\left(\frac{H_n^{(2)}}{2}\right)^{l_2}\left(\frac{H_n^{(3)}}{3}\right)^{l_3}\dots
\end{aligned}
$$

The quantities that arise here are harmonic numbers:

$$H_n^{(m)} = \sum_{k=1}^{n}\frac{1}{k^m}.$$

In particular, we get

$$\sum_{k=1}^{n}\binom{n}{k}\frac{(-1)^{k-1}}{k} = H_n,$$

which is a classic formula.

Theorem 15.2 (Meromorphic functions). *Let $f(z)$ be a function that is analytic on $[n_0,\infty)$.*

(i) Assume that $f(z)$ is meromorphic in the whole of $\mathbb{C}$ and analytic on $\Omega = \bigcup_{j=1}^{\infty}\gamma_j$ where the γ_j are positively oriented concentric circles whose radii tend to infinity. Let $f(z)$ be of polynomial growth on Ω. Then, for n large enough,

$$\sum_{k=n_0}^{n}(-1)^k\binom{n}{k}f_k = (-1)^{n-1}\mathrm{Res}_z\frac{n!f(z)}{z(z-1)\dots(z-n)}$$

where the sum is extended to all poles z not on $[n_0,\infty)$.

(ii) Assume that $f(z)$ is meromorphic in the half-plane Ω defined by $\Re z \ge d$ for some $d < n_0$. Let $f(z)$ be of polynomial growth in the complement in Ω of some compact set. Then, for n large enough,

$$\sum_{k=n_0}^{n}(-1)^k\binom{n}{k}f_k = (-1)^{n-1}\mathrm{Res}_z\frac{n!f(z)}{z(z-1)\dots(z-n)} + O(n^d)$$

where the sum is extended to all poles z in $\Re z > d$ and not on $[n_0,\infty)$.

Proof. See [23]. □

Now we study the next example, which is about trie sums; see [32, 22, 23]. They originate from the solution of the divide-and-conquer recursion

$$f_n = a_n + 2^{-n} \sum_{k=0}^{n} \binom{n}{k} (f_k + f_{n-k}),$$

for a given (toll) sequence a_n. A prototype is the sequence

$$U_n = \sum_{k=2}^{n} \binom{n}{k} \frac{(-1)^k}{2^{k-1} - 1}.$$

It arises when we set $a_n = n - 1$ for $n \geq 2$ and $f_0 = f_1 = 0$: Translating the recursion into an equation for the exponential generating function $F(z) = \sum_{n \geq 0} f_n z^n / n!$, we get

$$F(z) = (z-1)e^z + 1 + 2e^{z/2} F\Big(\frac{z}{2}\Big).$$

Setting $G(z) = e^{-z} F(z) = \sum_{n \geq 0} g_n z^n / n!$, this means

$$G(z) = z - 1 + e^{-z} + 2G\Big(\frac{z}{2}\Big),$$

and so

$$g_n = \frac{(-1)^n}{1 - 2^{1-n}}, \quad n \geq 2, \quad g_0 = g_1 = 0.$$

Therefore, with $F(z) = e^z G(z)$,

$$f_n = \sum_{k=2}^{n} \binom{n}{k} \frac{(-1)^k}{1 - 2^{k-1}} = \sum_{k=2}^{n} \binom{n}{k} \frac{(-1)^k}{2^{k-1} - 1} + U_n = n - 1 + U_n.$$

The analysis of U_n is a direct application of Theorem 15.2 when taking as integration contours large circles that go in between the poles of the function $(2^{s-1} - 1)^{-1}$. The poles are at

$$\chi_k = 1 + \frac{2\pi i k}{\log 2};$$

each of these induces a contribution of the form

$$n^{\chi_k} = n e^{2\pi i \cdot \log_2 n}.$$

The asymptotic formula follows now from the residues at $s = 0$ (double pole) and $s = \chi_k$, $k \neq 0$ (simple poles):

$$\begin{aligned} U_n &= \frac{n}{\log 2}(H_{n-1} - 1) - \frac{n}{2} + 2 + \frac{1}{\log 2} \sum_{k \neq 0} \frac{\Gamma(n+1)\Gamma(-1+\chi_k)}{\Gamma(n+\chi_k)} \\ &= n \log_2 n + n\Big(\frac{\gamma - 1}{\log 2} - \frac{1}{2}\Big) + \frac{n}{\log 2} \sum_{k \neq 0} \Gamma(-\chi_k) e^{2\pi i k \cdot \log_2 n} + O(\sqrt{n}\,). \end{aligned}$$

The error term $O(\sqrt{n}\,)$ is rather arbitrary here and could be replaced by $O(n^a)$ for any $0 < a < 1$.

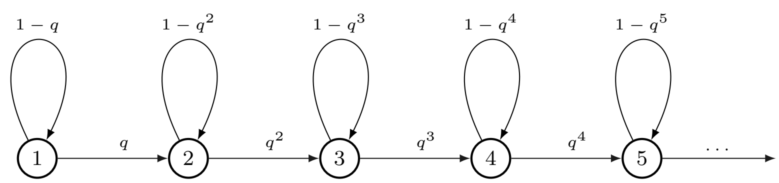
Here, we used the approximation

$$\frac{\Gamma(n+1)}{\Gamma(n+1-\chi_k)} = n^{\chi_k}\Big(1+O\Big(\frac{|\chi_k|^2}{n}\Big)\Big),$$

which is uniform in k and follows from Stirling's formula. $\square$

16. Approximate counting

Consider the following graph.



Assume that a random walk starts at state 1. Random steps are done as indicated: if we are in state i, the probability to advance to state $i+1$ is given by q^i, and with the complementary probability, we stay in state i. The question is where are we after n random steps? In [24], this is called a *walk of the pure-birth type*, but in the older literature [14], it is called *approximate counting*, and that is where the motivation comes from. The states represent a counter, and the final state k is considered to be an approximate count for — not n of course, but $\log_Q n$, with $Q = \frac{1}{q}$. We are not discussing the applications for that idea, but since the example is quite instructive and appears often, also in disguised form, and was also rediscovered several times, we decided to include it here. A relatively new application is in [8, 46].

Let $p(n,k)$ be the probability that, starting in state 1, we end up in state k after n random steps.

Flajolet [14] proves

Theorem 16.1.

$$p(n,k) = \sum_{t=0}^{k-1} \frac{(-1)^t q^{\binom{t}{2}}}{(q)_t (q)_{k-1-t}} (1-q^{k-t})^n.$$

The following derivation is from [36]. We obtain by a direct translation from the graph the recursion

$$p(n,k) = q^{k-1}p(n-1,k-1) + \big(1-q^k\big)p(n-1,k), \quad p(0,1)=1.$$

We will use a bivariate *generating function.* If we set

$$F(z,u) := \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} z^n u^k p(n,k),$$

we derive

$$F(z,u) - u = zuF(z,qu) + zF(z,u) - zF(z,qu),$$

or

$$F(z,u) = \frac{u}{1-z} + \frac{z(u-1)}{1-z} F(z,qu).$$

Iterating, this gives

$$\begin{aligned} F(z,u) &= \frac{u}{1-z} + \frac{z(u-1)}{1-z}\frac{uq}{1-z} + \frac{z(u-1)}{1-z}\frac{z(qu-1)}{(1-z)^2}uq^2 \\ &\quad + \frac{z(u-1)}{1-z}\frac{z(qu-1)}{1-z}\frac{z(q^2u-1)}{(1-z)^2}uq^3 + \cdots \\ &= \sum_{j=0}^{\infty} \frac{(-1)^j z^j (u;q)_j uq^j}{(1-z)^{j+1}}. \end{aligned} \tag{10}$$

This expression was derived in [43], using a transformation formula due to Heine. Now we have several ways of computing $[z^n u^k]F(z,u)$. We write

$$(u;q)_j = \frac{(u;q)_\infty}{(uq^j;q)_\infty},$$

and with Euler's partition identity (see Section 4), we have

$$\begin{aligned} p(n,k) &= \sum_{t=0}^{k-1} \frac{(-1)^t q^{\binom{t}{2}}}{(q)_t (q)_{k-1-t}} \sum_{j=0}^{n} (-1)^j q^{(k-1-t)j} q^j [z^{n-j}](1-z)^{-(j+1)} \\ &= \sum_{t=0}^{k-1} \frac{(-1)^t q^{\binom{t}{2}}}{(q)_t (q)_{k-1-t}} (1-q^{k-t})^n, \end{aligned}$$

which is exactly Flajolet's formula.

There is a second expression (given by Charalambides [7]),

$$p_C(n,k) = \frac{q^{\binom{k}{2}}}{(q)_k} \sum_{j=k}^{n} (-1)^{j-k} \frac{(q)_j}{(q)_{j-k}} \binom{n}{j}, \quad p_C(0,0) = 1.$$

This is equivalent to Flajolet's formula for $p(n-1,k)$; we give an independent proof of this fact.

$$p(n-1,k) = \sum_{j=0}^{k-1} \frac{(-1)^j q^{\binom{j}{2}}}{(q)_j (q)_{k-1-j}} (1-q^{j-k})^{n-1} = \sum_{j=0}^{k} \frac{(-1)^j q^{\binom{j}{2}}}{(q)_j (q)_{k-j}} (1-q^{k-j})^n.$$

Let us consider the generating function

$$S = \sum_{k\ge 0} x^k p(n-1,k)$$

$$
\begin{aligned}
&= \sum_{k\ge0} x^k \sum_{j=0}^{k} \frac{(-1)^j q^{\binom{j}{2}}}{(q)_j(q)_{k-j}}(1-q^{k-j})^n\\
&= \sum_{l=0}^{n}\binom{n}{l}(-1)^l \sum_{k\ge0} x^k \sum_{j=0}^{k} \frac{(-1)^j q^{\binom{j}{2}}}{(q)_j(q)_{k-j}} q^{(k-j)l}\\
&= \sum_{l=0}^{n}\binom{n}{l}(-1)^l \sum_{j\ge0} \frac{(-1)^j q^{\binom{j}{2}}}{(q)_j} \sum_{k\ge j} x^k \frac{1}{(q)_{k-j}} q^{(k-j)l}\\
&= \sum_{l=0}^{n}\binom{n}{l}(-1)^l \sum_{j\ge0} \frac{(-1)^j x^j q^{\binom{j}{2}}}{(q)_j} \sum_{k\ge0} (xq^l)^k \frac{1}{(q)_k}\\
&= \sum_{l=0}^{n}\binom{n}{l}(-1)^l \frac{1}{(xq^l;q)_\infty} \sum_{j\ge0} \frac{(-1)^j x^j q^{\binom{j}{2}}}{(q)_j}\\
&= \sum_{l=0}^{n}\binom{n}{l}(-1)^l \frac{1}{(xq^l;q)_\infty}(x;q)_\infty\\
&= \sum_{l=0}^{n}\binom{n}{l}(-1)^l (x;q)_l.
\end{aligned}
$$

On the other hand, let us consider the generating function

$$
\begin{aligned}
T = \sum_{k\ge0} x^k p_C(n,k) &= \sum_{k\ge0} x^k q^{\binom{k}{2}} \sum_{j=k}^{n} (-1)^{j-k} \begin{bmatrix} j\\ k\end{bmatrix}_q \binom{n}{j}\\
&= \sum_{j=0}^{n}\binom{n}{j}(-1)^j \sum_{k=0}^{j} x^k q^{\binom{k}{2}} (-1)^k \begin{bmatrix} j\\ k\end{bmatrix}_q\\
&= \sum_{j=0}^{n}\binom{n}{j}(-1)^j \prod_{l=0}^{j-1}(1-q^l x)\\
&= \sum_{j=0}^{n}\binom{n}{j}(-1)^j (x;q)_j.
\end{aligned}
$$

So $S=T$, that ends the proof.

A third expression for Flajolet's formula consists in using a q-binomial in (10) to extract $[u^{k-1}]$. First,

$$
[u^k]F(z,u) = \sum_{j\ge0} \frac{(-1)^j z^j q^j}{(1-z)^{j+1}} [u^{k-1}](u;q)_j
$$

$$= \sum_{j\ge 0} \frac{(-1)^j z^j q^j}{(1-z)^{j+1}} q^{\binom{k-1}{2}} \begin{bmatrix} j \\ k-1 \end{bmatrix}_q (-1)^{k-1},$$

and consequently:

$$\begin{aligned} p(n,k) &= [z^n] \sum_{j\ge 0} \frac{(-1)^j z^j q^j}{(1-z)^{j+1}} q^{\binom{k-1}{2}} \begin{bmatrix} j \\ k-1 \end{bmatrix}_q (-1)^{k-1} \\ &= q^{\binom{k-1}{2}} (-1)^{k-1} \sum_{j=k-1}^{n} (-1)^j \begin{bmatrix} j \\ k-1 \end{bmatrix}_q \binom{n}{j} q^j. \end{aligned}$$

For the asymptotics of the expected value C_n and the variance, Flajolet [14] first approximated the probabilities $p(n,k)$ using real analysis, and then continued with the approximate values, using the Mellin transform. This provides additional information about the probability distribution in the limit. To compute C_n, there is, however a more direct way, starting from the generating function $F(z,u)$, differentiating it w.r.t. u, followed by $u := 1$. It should be noted that this operation, applied to $(u;q)_j$ for $j \ge 1$, simply results in zero. Therefore

$$\begin{aligned} C_n &= [z^n] \frac{\partial}{\partial u} F(z,u)\Big|_{u=1} \\ &= -[z^n] \sum_{j\ge 1} \frac{(-1)^j z^j (q)_{j-1} q^j}{(1-z)^{j+1}} + [z^n] \frac{1}{1-z} \\ &= -\sum_{j\ge 1} (-1)^j \binom{n}{j} (q)_{j-1} q^j + 1. \end{aligned}$$

The asymptotic evaluation of the sum is a typical application of Rice's method, see Section 15. We extend $(q)_{j-1} q^j$ by

$$\frac{(q)_\infty}{(1-q^z)(q^z)_\infty} q^z.$$

Then we have to compute residues of

$$\frac{(-1)^n n!}{z(z-1)\dots(z-n)} \frac{(q)_\infty}{(1-q^z)(q^z)_\infty} q^z$$

at $z = 0$ (double pole) and at $z = \chi_k$ (simple poles for $k \ne 0$); we write $\chi_k = 2\pi i k/L$, with $L = \log Q$ and $Q = 1/q$. The double pole requires more work, since all the factors have to be expanded to two terms. We collect the expansions:

$$\frac{(-1)^n n!}{z(z-1)\dots(z-n)} \sim \frac{1}{z}(1+H_n), \qquad \frac{1}{1-q^z} \sim \frac{1}{Lz}\Big(1+\frac{Lz}{2}\Big),$$
$$\frac{(q)_\infty}{(q^z)_\infty} \sim 1-\alpha L z, \qquad q^z \sim 1 - Lz,$$

with $\alpha = \sum_{k\ge 1} \frac{1}{Q^k-1}$. Using the asymptotics for H_n and adding the extra term 1, we get $\log_Q n + \frac{\gamma}{L} + \frac{1}{2} - \alpha$. The residue at χ_k is simpler:

$$-\frac{1}{L}\frac{\Gamma(n+1)\Gamma(-\chi_k)}{\Gamma(n+1-\chi_k)} \sim \frac{1}{L} n^{\chi_k}\Gamma(-\chi_k).$$

Altogether we have a theorem, first proved by Flajolet [14].

Theorem 16.2.

$$C_n = \log_Q n + \frac{\gamma}{L} + \frac{1}{2} - \alpha + \delta(\log_Q n) + O(n^{-1/2}),$$

where the (small) periodic function $\delta(x)$ is given by its Fourier expansion

$$\delta(x) = -\frac{1}{L}\sum_{k\ne 0}\Gamma(-\chi_k)e^{2\pi i k x}.$$

For more results on the subject we refer to the original papers and [24].

17. Singularity analysis of generating functions

Our goal is, to say something about $[z^n]f(z)$, using information about the singularities of $f(z)$. Let us start with the simple case of a rational function $f(z)$. Using partial fraction decomposition, it can be written as a polynomial plus a finite number of terms of the form

$$\frac{A}{(1-z/\rho)^k},$$

for positive integers k. Now, the polynomial part can only influence a finite number of terms, is thus irrelevant for asymptotics, and

$$[z^n]\frac{A}{(1-z/\rho)^k} = A\rho^{-n}\binom{n+k-1}{n} \sim A\rho^{-n}\frac{n^{k-1}}{(k-1)!}.$$

We see that singularities ρ, which are just poles in this instance, which are closest to the origin, produce the largest exponential growth ρ^{-n}. For example, let $f(z) = \frac{8}{(1-z^2)(1-z/3)}$. Then there are two poles at ± 1 of smallest modulus 1, and one at 3, which will produce an exponentially small error term. We find the local expansions

$$f(z) \sim \frac{6}{1-z} \quad (z\to 1), \qquad f(z) \sim \frac{3}{1+z} \quad (z \to -1),$$

whence the asymptotic formula

$$[z^n]f(z) = 6 + 3(-1)^n + O(3^{-n}).$$

Now this instance was extremely simple, but we can deduce an important principle: The location of the singularity is responsible for the exponential growth, and that a pole of order k produces a (leading) term of order n^{k-1}.

Now let us consider Catalan numbers, given by

$$f(z) = \frac{1-\sqrt{1-4z}}{2z} = \sum_{n\geq 0} \frac{1}{n+1}\binom{2n}{n} z^n.$$

Stirling's formula gives us

$$\frac{1}{n+1}\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi}\, n^{3/2}}.$$

Can we see this directly from the generating function $f(z)$? Well, the singularity is at $z = \frac{1}{4}$, and the local expansion is

$$f(z) = 2 - 2\sqrt{1-4z} + \cdots,$$

and

$$[z^n]\big(-2\sqrt{1-4z}\,\big) \sim \frac{-2\, 4^n}{\Gamma(-\frac{1}{2}) n^{3/2}} = \frac{4^n}{\sqrt{\pi} n^{3/2}}.$$

The rest of this section is devoted to the study of why and how this works. The main references are the highly cited article [19] and of course [24].

For a precise notion of exponential growth, we write

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

(read a_n is of exponential growth K^n).

Theorem 17.1 (Exponential growth formula)**.** *If $f(z)$ is analytic at 0 and R is the modulus of the singularity nearest to the origin, then the coefficient $f_n = [z^n]f(z)$ satisfies*

$$f_n \bowtie \Big(\frac{1}{R}\Big)^n.$$

This allows us already to deal with meromorphic functions (the only singularities are poles). We can consider all poles nearest to the origin; there can only be a finite number of them, usually it is just one pole. Then one can subtract these poles (the principal parts); the resulting function has a larger radius of convergence, and thus what remains is exponentially small compared to the contribution from the dominant poles. Let us consider the example

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

which is the exponential generating function of surjections. The poles are the solutions of $e^z = 2$, or the points $\chi_k = \log 2 + 2\pi i k$, $k \in \mathbb{Z}$. The closest pole to the origin is at $\log 2$, so

$$[z^n]R(z) \bowtie \Big(\frac{1}{\log 2}\Big)^n.$$

But we can write

$$R(z) = -\frac{1}{2}\frac{1}{z - \log 2} + S(z),$$

and $S(z)$ has radius of convergence $|\log 2 + 2\pi i| = 6.321302922$. Consequently

$$[z^n]R(z) = \frac{1}{2}\frac{1}{(\log 2)^{n+1}} + O(6^{-n}).$$

In general, if ρ is a dominant pole of order k, there will be a contribution $p(n)\rho^{-n}$, where the polynomial $p(n)$ has degree $k-1$.

The process can be iterated, by including more poles, and obtaining a smaller (exponential) error term. In a variety of cases, the formal process that includes all the poles will lead to a series expansion that is exact and asymptotic. In general, there is a theorem due to Mittag-Leffler which explains how a series expansion can be obtained. For asymptotic purposes, collecting the contributions from dominant poles will be sufficient.

The process called singularity analysis of generating functions is about transfer theorems, from information of the generating function around the singularity to the asymptotics of the coefficients. Sufficient conditions will be provided for the implications

$$\begin{aligned}
&(T1) \qquad & f(z) = O(g(z)) \qquad &\Longrightarrow \qquad & f_n = O(n),\\
&(T2) \qquad & f(z) = o(g(z)) \qquad &\Longrightarrow \qquad & f_n = o(n),\\
&(T3) \qquad & f(z) \sim g(z) \qquad &\Longrightarrow \qquad & f_n \sim g_n.
\end{aligned}$$

We will also speak about O-transfers, o-transfers, and $\sim$-transfers. O-transfers are the most basic; refinements usually lead to o-transfers, and $\sim$-transfers follow from these, since

$$f(z) \sim g(z) \quad \text{is equivalent to} \quad f(z) = g(z) + o(g(z)).$$

Basic transfer. We know from basic principles that

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

for $\alpha \notin \{0, -1, -2, \dots\}$; from Stirling's formula this give us

$$[z^n](1-z)^{-\alpha} \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)}\Big[1 + \frac{\alpha(\alpha-1)}{2n} + \frac{\alpha(\alpha-1)(\alpha-2)(3\alpha-1)}{24n^2} + \cdots\Big].$$

The full expansion is given in [19].

Without loss of generality, when analyzing a singularity at ρ, we can assume that it is as 1, because of the simple transformation $g(z) := f(z/\rho)$, and

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

From a technical point of view, the following domains are important:

$$\Delta(\phi, \eta) = \{z \mid |z| \leq 1 + \eta, \ |\arg(z-1)| \geq \phi\},$$

where we take $\eta > 0$ and $0 < \phi < \pi/2$. This domain has the form of an indented disk.

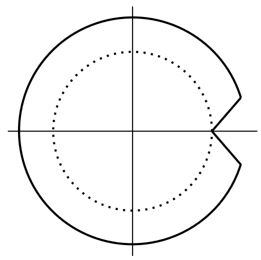


FIGURE 4. Domain $\Delta(\phi, \eta)$.

Extraction of coefficients is then done using Cauchy's integral formula, with a contour that stays as closely to the boundary of the domain $\Delta(\phi, \eta)$ as possible. We must refer to the original paper for details.

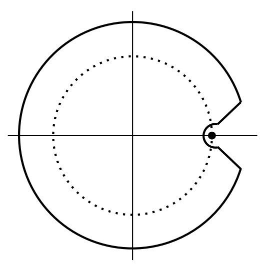


FIGURE 5. A typical path of integration

Theorem 17.2. *Assume that, with the sole exception of the singularity $z = 1$, $f(z)$ is analytic in the domain $\Delta = \Delta(\phi, \eta)$, where $\eta > 0$ and $0 < \phi < \pi/2$. Assume further that as z tends to 1 in Δ,*

$$f(z) = O(|1-z|^{-\alpha}),$$

for some real number α. Then the n-th Taylor coefficient of $f(z)$ satisfies

$$f_n = [z^n]f(z) = O(n^{\alpha-1}).$$

Corollary 17.1. *Assume that $f(z)$ is analytic in $\Delta \setminus \{1\}$, and that as z tends to 1 in Δ,*

$$f(z) = o(|1-z|^{-\alpha}),$$

Then, as $n \to \infty$,

$$f_n = o(n^{\alpha-1}).$$

Corollary 17.2. *Assume that $f(z)$ is analytic in $\Delta \setminus \{1\}$, and that as z tends to 1 in Δ,*

$$f(z) \sim K(1-z)^{-\alpha},$$

for $\alpha \notin \{0, -1, -2, \dots\}$. Then, as $n \to \infty$,

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

This settles already the instance about the Catalan numbers from before. Since $\sqrt{1-4z}$ can be extended to the complex plane except for a cut from $\frac{1}{4}$ to infinity on the real axis, this includes a fortiori also a Δ-domain.

Now we also allow logarithmic factors, in other words we consider

$$f(z) = (1-z)^{-\alpha} \left(\frac{1}{z} \log\left(\frac{1}{1-z} \right) \right)^{\beta}.$$

The factor $1/z$ is introduced for convenience: since $\log 1/(1-z) = z + O(z^2)$, dividing by z lets this expansion start with 1, and then raising it to the power β results in a power series expansion. The expansion is given in [24]:

$$[z^n]f(z) = \frac{n^{\alpha-1}}{\Gamma(\alpha)} (\log n)^{\beta} \Big[1 + \frac{C_1}{\log n} + \frac{C_2}{\log^2 n} + \cdots \Big],$$

with

$$C_k = \binom{\beta}{k} \Gamma(\alpha) \frac{d^k}{ds^k} \Gamma(s) \Big|_{s=\alpha}.$$

Note that this expansion comprises terms that have the factor $n^{\alpha-1}$. Terms including $n^{\alpha-2}$ go into the remainder term.

Here is an example:

$$[z^n] \frac{1}{\sqrt{1-z}} \frac{1}{\frac{1}{z}\log\frac{1}{1-z}} = \frac{1}{\sqrt{\pi n}\,\log n} \left(1 - \frac{\gamma + 2\log 2}{\log n} + O\Big(\frac{1}{\log^2 n} \Big) \right).$$

It is also possible to introduce an additional factor of the form $(\log \log n)^{\delta}$; we are not citing the details here.

The instance α being a negative integer had to be excluded before, since it resulted in a polynomial. With the presence of logarithmic factors, however, this still makes sense. Here is one particular expansion:

$(n > m \geq 0)$

$$[z^n](1-z)^m \log \frac{1}{1-z} = \sum_{k=0}^{m} \binom{m}{k} (-1)^k \frac{1}{n-k} = \frac{m!(-1)^m}{n(n-1)\dots(n-m)}.$$

The following proof is quite instructive: Consider

$$f(z) = \frac{m!}{z(z-1)\dots(z-m)} \frac{1}{z-n}$$

and perform partial fraction decomposition:

$$f(z) = \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} \frac{1}{k-n} \frac{1}{z-k} + \frac{m!}{n(n-1)\dots(n-m)} \frac{1}{z-n}.$$

Now multiply this by z and let $z \to \infty$. The result is

$$0 = \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} \frac{1}{k-n} + \frac{m!}{n(n-1)\dots(n-m)}.$$

18. Longest runs in words

We consider a binary alphabet $\{0, 1\}$, and are interested in the length of the longest run of ones in a random word of length n. Much more general scenarios are discussed in [24]. For a given parameter k, we want to find the generating functions of words where all runs of ones have length $< k$; call it $W^{<k}(z)$. There is a natural decomposition of the words with this property:

$$\mathbf{1}^{<k}\big(\mathbf{0}\mathbf{1}^{<k}\big)^*,$$

where $\mathbf{1}^{<k} = \{\varepsilon, \mathbf{1}, \mathbf{11}, \dots, \mathbf{1}^{k-1}\}$. From this,

$$W^{<k}(z) = \frac{1-z^k}{1-z} \frac{1}{1 - z\frac{1-z^k}{1-z}} = \frac{1-z^k}{1-2z+z^{k+1}}.$$

The first step is the location of the dominant pole ρ_k, which we expect to be close to $\frac{1}{2}$. Set $Q_k(z) = 1 - 2z + z^{k+1}$. As z traverses the circle $|z| = 1$ in the complex plane, the value of $Q_k(z)$ winds around the origin exactly once, hence the polynomial Q_k has exactly one root in $|z| < 1$. We call this root $\rho_k = \frac{1}{2} + \varepsilon_k$. It satisfies the equation

$$z = \frac{1}{2} + \frac{1}{2} z^{k+1}.$$

One can start with a crude bound $\varepsilon_k = O(\frac{1}{k})$ and plug this in:

$$\frac{1}{2} + \varepsilon_k = \frac{1}{2} + \frac{1}{2}\Big(\frac{1}{2} + O\Big(\frac{1}{k}\Big)\Big)^{k+1}.$$

This leads to the approximation

$$\rho_k = \frac{1}{2} + \frac{1}{2^{k+2}} + O\Big(\frac{k}{2^{2k}}\Big),$$

which is enough for practical purposes. We can do better, however, using the Lagrange inversion formula. We write $y = z - \frac{1}{2}$ and $x = \frac{1}{2}$, then

$$y = x\Big(\frac{1}{2} + y\Big)^{k+1} = x\Phi(y).$$

Hence

$$[x^n]y = \frac{1}{n}[y^{n-1}]\Big(\frac{1}{2} + y\Big)^{n(k+1)} = \frac{1}{n}\binom{n(k+1)}{n-1}\frac{1}{2^{nk+1}},$$

and

$$y = \sum_{n\ge 1}\frac{1}{n}\binom{n(k+1)}{n-1}\frac{1}{2^{nk+1}}x^n.$$

Plugging in the special value for x, we get

$$\rho_k = \frac{1}{2} + \sum_{n\ge 1}\frac{1}{n}\binom{n(k+1)}{n-1}\frac{1}{2^{nk+n+1}} = \frac{1}{2} + \frac{1}{2^{k+2}} + \cdots .$$

The iterative process to get better and better approximations for ρ_k is called bootstrapping and appears already in [33]. The next step is to expand $W^{<k}(z)$ around the dominant pole. We have

$$W^{<k}(z) \sim \frac{A}{1 - z/\rho_k},$$

and

$$A = \lim_{z\to\rho_k}(1 - z/\rho_k)\frac{1 - z^k}{Q_k(z)} = \frac{1 - \rho_k^k}{-\rho_k Q'(\rho_k)} = \frac{1 - \rho_k^k}{\rho_k(2 - (k+1)\rho_k^k)}.$$

Therefore

$$[z^n]W^{<k}(z) = \frac{1 - \rho_k^k}{2 - (k+1)\rho_k^k}\rho_k^{-n-1} + O(1).$$

The error comes from the fact that the remaining poles are larger than 1 in absolute value. This derivation, as it stands, works for k fixed and $n \to \infty$, but it is not too hard to extend the range of validity of it. In particular, one replaces A by the easier $\frac{1}{2}$ and

$$[z^n]W^{<k}(z) \sim \frac{1}{2}\Big(\frac{1}{2} + \frac{1}{2^{k+2}}\Big)^{-n-1} \sim 2^n\Big(1 - \frac{1}{2^{k+1}}\Big)^{n+1} \sim 2^n e^{-n/2^{k+1}},$$

when k is close to $\log_2 n$ where the main contribution comes from. We must refer to [24] for this series of approximations.

It is interesting to sketch the enumeration problem that Knuth [33] encountered in his carry propagation problem. Words of length n over the alphabet $\{0, 1, 2\}$ are studied where letter 1 appears with probability

$\frac{1}{2}$, and 0 and 2 with probability $\frac{1}{4}$ each. For given k, the (contiguous) substring 1^k2 is forbidden. The allowed words may be described as

$$\Big(\big(\mathbf{1}^{<k}\mathbf{2}\big)^*\mathbf{1}^*\mathbf{0}\Big)^*\big(\mathbf{1}^{<k}\mathbf{2}\big)^*\mathbf{1}^*.$$

Translating that into a generating function ($\mathbf{0}\mapsto z/4$, $\mathbf{1}\mapsto z/2$, $\mathbf{2}\mapsto z/4$), we get

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

The rest of the analysis is very similar to the previous discussion; the dominant pole is here close to 1.

19. Inversions in permutations and pumping moments

We consider permutations $\pi = p_1p_2\dots p_n$ of $\{1,2,\dots,n\}$ (written in one-line notation) and assume that all $n!$ of them are equally likely. An inversion in π is a pair $i<j$ with $p_i>p_j$, and $I(\pi)$ is the total number of inversions of π. There is a convenient way to study the statistics of this parameter, namely the inversion table $b_1b_2\dots b_n$. The meaning is that b_i counts the number of elements larger than i that stand to the left of i. Clearly, $I(\pi)=b_1+\dots+b_n$. We have the natural restrictions $0\le b_i\le n-i$, since there are only $n-i$ elements larger than i altogether. The cute observation is that, given an inversion table, the permutation itself can be reconstructed: The value b_1 tells us that number 1 will be in position b_1+1, then b_2 tells us the position of number 2 (b_2 slots must be left open for later numbers), and so on. There is thus a bijection between permutations and inversion tables; the latter ones are easier to handle when it comes to generating functions, because of independence of the entries. If we consider the product

$$1(1+q)(1+q+q^2)\cdots(1+q+\dots+q^{n-1})=\prod_{j=1}^{n}\frac{1-q^j}{1-q},$$

then the coefficient of q^k in it is the number of permutations of n elements with k inversions. Now let

$$F_n(q)=\prod_{j=1}^{n}\frac{1-q^j}{j(1-q)}=F_{n-1}(q)\frac{1-q^n}{n(1-q)}$$

be the probability generating function of the parameter $I(\pi)$. Our goal is to get information about the s-th factorial moment, $\frac{d^s}{dq^s}F_n(q)\big|_{q=1}$. We form a bivariate generating function

$$H(z,q)=\sum_{n\ge0}z^nF_n(q).$$

The recursion translates then into

$$(1-q)\frac{\partial}{\partial z}H(z,q) = H(z,q) - qH(zq,q), \qquad H(z,1) = \frac{1}{1-z}.$$

It is impossible to solve this functional differential equation explicitly, but it is possible to derive information about

$$g_s(z) = \sum_{n\ge 0} z^n \frac{d^s}{dq^s}F_n(q)\big|_{q=1} = \frac{d^s}{dq^s}H(z,q)\big|_{q=1},$$

by differentiation the functional equation several times, and express $g_s(z)$ by already computed $g_i(z)$ with $i < s$. This procedure is nicknamed "pumping moments" in [24]; the following computations are taken from [29].

With a bit of patience, this program results in

$$g_s'(z) - \frac{1}{1-z}g_s(z) = h_s(z), \qquad g_0(z) = \frac{1}{1-z},$$

where

$$h_s(z) = \frac{1}{1-z}\sum_{j=1}^{s}\binom{s}{j}g_{s-j}^{j}z^j + \frac{1}{(s+1)(1-z)}\sum_{j=2}^{s+1}\binom{s+1}{j}g_{s+1-j}^{(j)}z^j.$$

Note that $h_s(z)$ is a "known" function since it involes only g_i's and its derivatives that were already computed. This differential equation is easy to solve:

$$g_s(z) = \frac{1}{1-z}\int_0^z h_s(t)(1-t)dt.$$

By inspection, one "sees" (and then proves by induction) that

$$h_s(z) = \frac{(2s)!}{4^s(1-z)^{s+1}} + \frac{c_s(2s-1)!}{(1-z)^{2s}} + \textsf{lower order terms}.$$

The constants c_s satisfy

$$c_s = c_{s-1}\frac{s}{2(2s-1)} + \frac{s(1-4s)}{3(2s-1)4^{s-1}}, \quad \text{with} \quad c_0 = 0.$$

It is easy to prove that

$$c_s = -\frac{s(4s+5)}{9\cdot 4^{s-1}}.$$

This leads by direct translation to the formula

$$\frac{d^s}{dq^s}F_n(q)\big|_{q=1} = \frac{1}{4^s}n^{2s} + \frac{s(2s-11)}{9\cdot 4^s}n^{2s-1} + O(n^{2s-2}).$$

For $s = 1$, we get the expected value $\frac{n(n-1)}{4}$, which is exact.

If one applies the pumping moments method to a suitably shifted random variable, one can derive limiting distribution results, see [24].

Area under Dyck paths.

In this example, we consider Dyck paths (counted by Catalan numbers) and its area; if the path is $x_0x_1\dots x_{2n}$ with $x_0 = x_{2n} = 0$, $x_i \ge 0$, $x_i - x_{i+1} = \pm 1$, then the area is defined to be $x_0 + x_1 + \cdots + x_{2n}$. Let $F(z, q)$ be the generating function according to half-length and area. Mapping $a_j \mapsto q^j z$, $b_j \mapsto q^j$, $c_j \mapsto 0$, the continued fraction theorem (Theorem 7.3, Section 7) leads directly to

$$F(z,q) = \cfrac{1}{1 - \cfrac{zq}{1 - \cfrac{zq^2}{\ddots}}} = \frac{1}{1 - zqF(zq, q)}.$$

It is natural to set $F(z, q) = A(z)/B(z)$, then

$$\frac{A(z)}{B(z)} = \frac{1}{1 - zq\frac{A(zq)}{B(zq)}} = \frac{B(zq)}{B(zq) - zqA(zq)}.$$

Comparing numerator and denominator,

$$A(z) = B(zq), \qquad B(z) = 1 - zqA(zq) = 1 - zqB(zq^2).$$

Setting $B(z) = \sum_n b_n z^n$, this leads to

$$b_n = q^n b_n - q^{2n-1} b_{n-1}, \qquad b_0 = 1.$$

This recursion can be iterated, with the result

$$b_n = \frac{(-1)^n q^{n^2}}{(1-q)(1-q^2)\dots(1-q^n)}.$$

Defining

$$E(z,q) = \sum_{n\ge 0} \frac{(-z)^n q^{n^2}}{(q)_n},$$

we may express the sought generating function as

$$F(z,q) = \frac{E(zq,q)}{E(z,q)}.$$

Now we look at the generating functions $\mu_r(z) = \frac{\partial^r}{\partial q^r} F(z,q)\big|_{q=1}$; they are (apart from normalization) the generating functions of the r-th factorial moments. Pumping the moments can now be done as follows: First we rewrite the functional equation as $F(z, q) = 1 + zF(z, q)F(zq, q)$. Then this will be differentiated r times according to the Leibniz rule, resulting in

$$\mu_r(z) = z \sum_{j=0}^{r} \binom{r}{j} \mu_{r-j}(z) \sum_{k=0}^{j} z^k \mu_{j-k}^{(k)}(z).$$

One sees by inspection that

$$\mu_r(z) = \frac{K_r}{(1-4z)^{(3r-1)/2}} + O((1-4z)^{-(3r-2)/2}).$$

The constants K_r follow a recursion, which cannot be solved explicitly, but this information is sufficient to characterize its probability distribution, see [24, 47]; this is also true for our previous (simpler) example of inversions, which leads to the normal distribution.

There exist many similar problems in the literature where such an approach works.

20. Tree function

The objects we consider are labelled rooted non-planar trees. There is a symbolic description of this family:

$$\mathcal{T} = \mathcal{Z} \star \textsc{Set}(\mathcal{T}).$$

This is a recursive description, which used the labelled product which conveniently does the relabelling for us once we use exponential generating functions:

$$T(z) = ze^{T(z)}.$$

Reading off coefficients is a typical application of the Lagrange inversion formula:

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

Therefore the number of labelled rooted non-planar trees with n nodes is given by $n![z^n]T = n^{n-1}$.

This is a famous formula, and there exist also many direct proofs for it. We refer to [24, 31] for pointers to the literature.

For comparison, let us also study the unlabelled counterparts, defined by

$$\mathcal{A} = \mathcal{Z} \cdot \textsc{Mset}(\mathcal{A}),$$

which leads to

$$A(z) = z\exp\Big(A(z) + \frac{1}{2}A(z^2) + \frac{1}{3}A(z^3) + \cdots\Big);$$

this equation for the ordinary generating function $A(z) = \sum_{n\ge1} a_n z^n$ was first derived by Polya. There is an equivalent formula,

$$A(z) = \frac{z}{(1-z)^{a_1}(1-z^2)^{a_2}(1-z^3)^{a_3}\dots},$$

which is also implicit, but allows to compute the numbers a_n in a recursive fashion. The formula follows directly from the combinatorial description: What follows the root is a multiset of already existing trees of all possible sizes.

The implicitly defined function $y = ze^y$ is known as the tree function. This is an important function that, although implicitly defined, we should augment to our arsenal of known functions. Whenever one sees quantities like n^n, one should think about this tree function. We know that

$$y = \sum_{n \geq 1} n^{n-1} \frac{z^n}{n!},$$

and then

$$\sum_{n \geq 0} n^n \frac{z^n}{n!} = 1 + zy'(z) = 1 + \frac{y}{1-y} = \frac{1}{1-y}.$$

Many similar quantities may be expressed via the tree function. A celebrated example is Ramanujan's Q-function [17]:

$$Q(n) = 1 + \frac{n-1}{n} + \frac{(n-1)(n-2)}{n^2} + \cdots .$$

It follows immediately that

$$\frac{n^n}{n!} Q(n) = \frac{n^{n-1}}{(n-1)!} + \frac{n^{n-2}}{(n-2)!} + \cdots + 1.$$

Now consider $\log \frac{1}{1-y(z)}$ and its coefficients. This is a typical application of the third version of the Lagrange inversion formula, with $g(y) = \log \frac{1}{1-y}$. We compute

$$[z^n] \log \frac{1}{1-y} = \frac{1}{n}[y^{n-1}] g'(y) e^{yn} = \frac{1}{n}[y^{n-1}] \frac{1}{1-y} e^{yn} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{n^k}{k!},$$

so we see that $\log \frac{1}{1-y}$ is essentially the generating function of Ramanujan's Q-function:

$$\log \frac{1}{1-y(z)} = \sum_{n \geq 1} \frac{n^{n-1}}{n!} Q(n) z^n.$$

Next, we look again at the Lagrange inversion scenario, $y = z\Phi(y)$. In combinatorial contexts, it is natural to assume that $\Phi(y)$ is given as a power series with non-negative coefficients $\Phi(y) = \sum_k \phi_k y^k$ and that $\phi_0 > 0$. To avoid trivialities, we exclude the linear function $\Phi(z) = \phi_0 + \phi_1 y$. This describes planar trees, and the number ϕ_k can be interpreted as a weight when branching with k successors occurs; in particular, if $\phi_k = 0$, k-way branching is not allowed. The family of such trees is often called *simply generated family of trees*, introduced by Meir and Moon [38]. The tree function $y = ze^y$ is the instance $\Phi(y) = e^y$. The radius of convergence of $y(z)$ (and thus the exponential growth of the coefficients) can be determined by a theorem that we cite from [24].

Theorem 20.1. *Let Φ be a function analytic at 0, having non-negative Taylor coefficients, and such that $\Phi(0) \neq 0$. Let $R \leq \infty$ be the radius of convergence of the series representing Φ at 0. Under the condition*

$$\lim_{z \to R^-} \frac{z\Phi'(z)}{\Phi(z)} > 1,$$

there exists a unique solution $\tau \in (0, R)$ of the characteristic equation

$$\frac{\tau\Phi'(\tau)}{\Phi(\tau)} = 1.$$

Then, the formal solution $y(z)$ of the equation $y(z) = z\Phi(y(z))$ is analytic at 0 and its coefficients satisfy the exponential growth formula:

$$[z^n]y(z) \bowtie \rho^{-n} \quad \textit{where} \quad \rho = \frac{\tau}{\Phi(\tau)} = \frac{1}{\Phi'(\tau)}. \quad \square$$

We just give the remark that the equation $\frac{\tau\Phi'(\tau)}{\Phi(\tau)} = 1$ follows from the implicit function theorem.

Now we apply this to the tree function. It is plain to see that $\tau = 1$ and hence $\rho = \frac{1}{e}$. We want to expand $y(z)$ around its singularity $\rho = \frac{1}{e}$. In order to do so, it is easier to look at $z = z(y) = ye^{-y}$, expand this around $\tau = 1$, and then invert the expansion. We easily get

$$ez = 1 - \frac{1}{2}(1-y)^2 - \frac{1}{3}(1-y)^3 + \cdots .$$

Note that the linear term is not present, which will lead us to a singularity of the square-root type. We get

$$(1-y)^2 = 2(1-ez) + \cdots \Longrightarrow y = 1 - \sqrt{2}\sqrt{1-ez} + \cdots .$$

The conditions of singularity analysis are satisfied, since the function $y(z)$ is analytic except for a cut from $\frac{1}{e}$ to ∞ on the real axis. Therefore

$$[z^n]y(z) \sim -\sqrt{2}\frac{1}{\Gamma(-\frac{1}{2})n^{3/2}}e^n = \frac{e^n}{\sqrt{2\pi}\, n^{3/2}}.$$

But we know the exact answer $[z^n]y(z) = \frac{n^{n-1}}{n!}$. Hence

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Note that in principle as many terms as desired could be obtained. Singularity analysis has proved Stirling's formula via the tree function!

21. The saddle point method

Our main source here is Odlyzko's chapter [40].

The saddle point method is the most useful method for obtaining asymptotic information about rapidly growing functions. It is based on the freedom to shift contours of integration when estimating integrals of analytic functions. We assume that $f(z)$ is analytic in $|z| < R \leq \infty$. We will also make the assumption that for some R_0, if $R_0 < r < R$, then

$$\max_{|z|=r} |f(z)| = f(r).$$

This assumption is clearly satisfied by all functions with real non-negative coefficients, which are the most common ones in combinatorial enumeration. We will also suppose that $|z| = r$ is the unique point with $|z| = r$ where the maximum value is assumed. The first step in estimating $[z^n]f(z)$ by the saddle point method is to find the saddle point. Under our assumptions, that will be a point $r \in (R_0, R)$ which minimizes $r^{-n}f(r)$. The minimizing $r = r_0$ will usually be unique, at least for large n. Cauchy's integral formula is then applied with the contour $|z| = r_0$. The reason for this choice is that for many functions, on this contour the integrand is large only near $z = r_0$, the contributions from the region near $z = r_0$ do not cancel each other, and remaining regions contribute little. By Cauchy's theorem, any simple closed contour enclosing the origin gives the correct answer. However, on most of them the integrand is large, and there is so much cancellation that it is hard to derive any estimates. The circle going through the saddle point, on the other hand, yields an integral that can be controlled well.

Example. Stirling's formula. We compute the reciprocal of $n!$ as $[z^n]e^z$. The saddle point is that real r that minimizes $r^{-n}e^r$, which is $r = n$. Consider the contour $|z| = n$, and set $z = n\exp(i\theta)$, $-\pi \leq \theta \leq \pi$. Then

$$[z^n]e^z = \frac{1}{2\pi i}\int_{|z|=n} e^z \frac{dz}{z^{n+1}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} n^{-n}\exp(ne^{i\theta} - ni\theta)\, d\theta.$$

Since $|\exp(z)| = \exp(\Re(z))$, the absolute value of the integrand is $n^{-n}\exp(n\cos\theta)$, which is maximized for $\theta = 0$. Now

$$e^{i\theta} = \cos\theta + i\sin\theta = 1 - \frac{\theta^2}{2} + i\theta + O(|\theta|^3),$$

so for any $\theta_0 \in (0;\pi)$,

$$\int_{-\theta_0}^{\theta_0} n^{-n}\exp(ne^{i\theta} - ni\theta)\, d\theta = \int_{-\theta_0}^{\theta_0} n^{-n}\exp\Big(n - \frac{n\theta^2}{2} + O(n|\theta|^3)\Big)\, d\theta.$$

Note the cancellation of the $ni\theta$ term. We select $\theta_0 = n^{-2/5}$, so that $n|\theta|^3 \le n^{-1/5}$, and therefore

$$\exp(n - \frac{n\theta^2}{2} + O(n|\theta|^3)) = \exp\big(n - \frac{n\theta^2}{2}\big)(1 + O(n^{-1/5})).$$

Hence

$$\int_{-\theta_0}^{\theta_0} n^{-n} \exp(ne^{i\theta} - ni\theta)\, d\theta = (1 + O(n^{-1/5}))n^{-n}e^n \int_{-\theta_0}^{\theta_0} \exp(-\frac{n\theta^2}{2})\, d\theta.$$

But

$$\begin{aligned}\int_{-\theta_0}^{\theta_0} \exp(-\frac{n\theta^2}{2})\, d\theta &= \int_{-\infty}^{\infty} \exp(-\frac{n\theta^2}{2})\, d\theta - 2\int_{\theta_0}^{\infty} \exp(-\frac{n\theta^2}{2})\, d\theta \\ &= (2\pi/n)^{1/2} + O(\exp(-n^{1/5}/2)),\end{aligned}$$

so

$$\int_{-\theta_0}^{\theta_0} n^{-n} \exp(ne^{i\theta} - ni\theta)\, d\theta = (1 + O(n^{-1/5}))n^{-n}e^n(2\pi/n)^{1/2}.$$

On the other hand, for $\theta_0 < |\theta| \le \pi$,

$$\cos\theta \le \cos\theta_0 = 1 - \frac{\theta_0^2}{2} + O(\theta_0^4)$$

so

$$n\cos\theta \le n - n^{1/5}/2 + O(n^{3/5})$$

and therefore for large n

$$\Big|\int_{\theta_0}^{\pi} n^{-n} \exp(ne^{i\theta} - ni\theta)\, d\theta\Big| \le n^{-n}\exp(n - n^{1/5}/2).$$

Combining all these estimates we find that

$$\frac{1}{n!} = [z^n]e^z = (1 + O(n^{-1/5}))(2\pi/n)^{1/2}n^{-n}e^n,$$

which is a weak form of Stirling's approximation.

Let us summarize: We decompose the contour $\mathcal{C} = \mathcal{C}^{(0)} \cup \mathcal{C}^{(1)}$, where $\mathcal{C}^{(0)}$ (the "central part") contains the saddle point (or passes very near to it) and $\mathcal{C}^{(0)}$ is formed of the two remaining "tails." This splitting has to be determined in each case in accordance with the growth of the integrand. The basic principle rests on two major conditions: the contributions of the two tails should be asymptotically negligible; in the central region, the quantity $f(z)$ in the integrand (which is written as $\exp(f(z))$) should be asymptotically well approximated by a quadratic function. Under these conditions, the integral is asymptotically equivalent to an incomplete Gaussian integral. It then suffices to verify that tails can be completed back, introducing only negligible error terms. By this sequence of steps, the original integral is asymptotically reduced to a complete Gaussian integral, which evaluates in closed form.

Let us consider another example where the outcome is less predictable. It goes back to van Lint [35], but compare with [37] for further developments.

Representations of 0 as a weighted sum. Let $A(N)$ be the number of solutions of the equation

$$\sum_{k=-N}^{N} \varepsilon_k k = 0, \quad \text{where} \quad \varepsilon_k \in \{0,1\}.$$

Now

$$A(N) = [z^0] \prod_{k=-N}^{N} (1+z^k) = \frac{1}{2\pi i} \oint \prod_{k=-N}^{N} (1+z^k) \frac{dz}{z}.$$

In order to find the saddle point, it is convenient to take the logarithm of the integrand, differentiate, and solve the equation

$$1 = \sum_{k=1}^{N} \frac{k(z^k-1)}{1+z^k}.$$

The solution (approximate saddle point) is close to one; if one wants a closer approximation, one can use bootstrapping, as explained earlier. We thus choose the unit circle as the path of integration:

$$A(N) = \frac{2^{2N+2}}{\pi} \int_0^{\frac{\pi}{2}} \prod_{k=1}^{N} \cos^2 kx \, dx.$$

Now we sketch how to choose appropriate ranges and approximations. For the range $\frac{\pi}{2N} \le x \le \frac{\pi}{2}$, the integrand is exponentially small, and can be ignored. For $0 \le x < N^{-4/3}$, the Gaussian approximation is valid:

$$\prod_{k=1}^{N} \cos^2 kx = \exp\Big(-x^2 \frac{N(N+1)(2N+1)}{6} + O(N^{-1/3}) \Big);$$

the integral over the remaining range $N^{-4/3} \le x < \frac{\pi}{2N}$ is not neglible, but smaller than the contribution from the central range. All error terms can be made explicit, and since

$$\int_0^{N^{-4/3}} \exp\Big(-x^2 \frac{N(N+1)(2N+1)}{6} \Big) \, dx \sim \frac{1}{2} (3\pi)^{1/2} N^{-3/2},$$

there is the result

$$A(N) \sim \Big(\frac{3}{\pi}\Big)^{1/2} 2^{2N+1} N^{-3/2}.$$

22. Hwang's quasi-power theorem

Here, we want to discuss a method, which turns out to be very useful in combinatorial contexts to prove Gaussian limit laws. This section is by no means a full description of the rich interplay between combinatorics and probability theory. The source of our short treatment is once again [24].

A real random variable Y is specified by its distribution function,

$$\mathbb{P}\{Y \le x\} = F(x).$$

It is said to be continuous if $F(x)$ is continuous. In that case, $F(x)$ has no jump, and there is no single value in the range of Y that carries a non-zero probability mass. If in addition $F(x)$ is differentiable, the random variable Y is said to have a density, $g(x) = F'(x)$, so that

$$\mathbb{P}\{Y \le x\} = \int_{\infty}^{x} g(x)dx.$$

A particularly important case is the standard Gaussian or normal $\mathcal{N}(0,1)$ distribution function,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{x} e^{-w^2/2} dw;$$

the corresponding density is

$$\Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Let Y be a continuous random variable with distribution function $F_Y(x)$. A sequence of random variables Y_n with distribution functions $F_{Y_n}(x)$ is said to converge *in distribution* to Y if, pointwise, for each x,

$$\lim_{n\to\infty} F_{Y_n}(x) = F_Y(x).$$

In that case, one writes $Y_n \Rightarrow Y$ and $F_{Y_n} \Rightarrow F_Y$.

For the readers' convenience, we cite a classic theorem:

Theorem 22.1 (Basic Central Limit Theorem)**.** *Let T_j be independent random variables supported by $\mathbb{R}$ with a common distribution of (finite) mean μ and (finite) standard deviation σ. Let $S_n := T_1 + \cdots + T_n$. Then the standardized sum S_n^* converges in distribution to the standard normal distribution,*

$$S_n^* = \frac{S_n - \mu n}{\sigma\sqrt{n}} \Rightarrow \mathcal{N}(0,1).$$

Short proofs use the concept of characteristic functions, which we do not discuss here.

A particularly simple application is when $T \equiv T_j$ takes the values 0 and 1, both with probability $\frac{1}{2}$. Then we can consider the probability generating function

$$p_n(u) = \left(\frac{1}{2} + \frac{u}{2}\right)^n,$$

so that $[u^k]p_n(u) = 2^{-n}\binom{n}{k}$ is the probability that $S_n = k$. We see here that the probability generating function is a large power of a fixed function, here $\frac{1+u}{2}$.

We will see that it suffices that the probability generating function of a combinatorial parameter behaves nearly like a large power of a fixed function to ensure convergence to a Gaussian limit – this is the quasi-powers framework, a concept that is largely due to Hwang [28].

Before we state a theorem, we discuss the Stirling cycle distribution. Assume that we have a random permutation of n elements, and the random variable X_n counts the number of cycles. Since a permutation can be seen as a set of cycles, it is easy to write down the relevant generating functions:

$$\mathcal{P} = \textsc{Set}(\textsc{Cyc}(\mathcal{Z})) \quad \Longrightarrow \quad P(z,u) = \exp\left(u \log \frac{1}{1-z}\right) = (1-z)^{-u}.$$

So, $n![z^n u^k]P(z,u)$ is the probability that a random permutation of n elements has k cycles. Extracting the coefficient of z^n, we find the probability generating function $p_n(u)$ related to X_n:

$$p_n(u) = \binom{n+u-1}{n} = \frac{\Gamma(u+n)}{\Gamma(u)\Gamma(n+1)};$$

near $u = 1$:

$$p_n(u) = \frac{n^{u-1}}{\Gamma(u)}\Big(1 + O\Big(\frac{1}{n}\Big)\Big) = \frac{\left(e^{u-1}\right)^{\log n}}{\Gamma(u)}\Big(1 + O\Big(\frac{1}{n}\Big)\Big).$$

Thus, as $n \to \infty$, the probability generating function $p_n(u)$ approximately equals a large power of e^{u-1}, taken with exponent $\log n$ and multiplied by the fixed function, $\Gamma(u)^{-1}$. This is enough to ensure a Gaussian limit law, as we will see.

The following notations will be convenient: given a function $f(u)$ analytic at $u = 1$ and assumed to satisfy $f(1) \neq 0$, we set

$$\mathsf{mean}(f) = \frac{f'(1)}{f(1)}, \qquad \mathsf{var}(f) = \frac{f''(1)}{f(1)} + \frac{f'(1)}{f(1)} - \left(\frac{f'(1)}{f(1)}\right)^2.$$

Theorem 22.2 (Hwang's quasi-power theorem)**.** *Let the X_n be non-negative discrete random variables (supported by $\mathbb{N}_0$), with probability generating functions $p_n(u)$. Assume that, uniformly in a fixed complex* *neighbourhood of $u=1$, for sequences $\beta_n, \kappa_n \to \infty$, there holds*

$$p_n(u) = A(u)B(u)^{\beta_n}\left(1 + O\Big(\frac{1}{\kappa_n}\Big)\right),$$

where $A(u)$, $B(u)$ are analytic at $u=1$ and $A(1)=B(1)=1$. Assume finally that $B(u)$ satisfies the so-called "variability condition,"

$$\mathsf{var}(B(u)) = B''(1) + B'(1) - (B'(1))^2 \neq 0.$$

Under these conditions, the mean and variance of X_n satisfy

$$\mu_n = \mathbb{E}X_n = \beta_n \mathsf{mean}(B(u)) + \mathsf{mean}(A(u)) + O(\kappa_n^{-1}),$$
$$\sigma_n^2 = \mathbb{V}X_n = \beta_n \mathsf{var}(B(u)) + \mathsf{var}(A(u)) + O(\kappa_n^{-1}).$$

The distribution of X_n is, after standardization, asymptotically Gaussian, and the speed of convergence to the Gaussian limit is $O(\kappa_n^{-1} + \beta_n^{-1/2})$:

$$\mathbb{P}\left\{\frac{X_n - \mathbb{E}X_n}{\sqrt{\mathbb{V}X_n}} \le x\right\} = \Phi(x) + O\Big(\frac{1}{\kappa_n} + \frac{1}{\sqrt{\beta_n}}\Big).$$

This theorem is a direct application of a lemma, also due to Hwang, that applies more generally to arbitrary discrete or continuous distributions.

References

[1] G. Andrews. *The Theory of Partitions*, volume 2 of *Encyclopedia of Mathematics and its Applications*. Addison–Wesley, 1976.
[2] G. Andrews, R. Askey, and R. Roy. *Special Functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1999.
[3] T. Apostol. *Introduction to Analytic Number Theory*. Springer, 1976.
[4] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy, and D. Gouyou-Beauchamps. Generating functions of generating trees. *Discrete Mathematics*, 246:29–55, 2002.
[5] N. G. De Bruijn. *Asymptotic Methods in Analysis*. Cover, 1970.
[6] L. Carlitz. Restricted compositions. *FQ*, 14:254–264, 1976.
[7] Ch. A. Charalambides. Moments of a class of discrete q-distributions. *Journal of Statistical Planning and Inference*, 135:64–76, 2005.
[8] J. Cichoń and W. Macyna. Approximate counters for flash memory. *17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA), Toyama, Japan*, 2011.
[9] L. Comtet. *Advanced Combinatorics*. Reidel, Dordrecht, 1974.
[10] N. G. De Bruijn, D. E. Knuth, and S. O. Rice. The average height of planted plane trees. In R. C. Read, editor, *Graph Theory and Computing*, pages 15–22. Academic Press, 1972.
[11] H. Delange. Sur la fonction sommatoire de la fonction somme des chiffres. *Enseignement Mathématique*, 21:31–47, 1975.
[12] C. Elsholtz, C. Heuberger, and H. Prodinger. The number of huffman codes, compact trees, and sums of unit fractions. *IEEE Transactions of Information Theory*, 59:1065–1075, 2013.

[13] P. Flajolet. Combinatorial aspects of continued fractions. *Discrete Mathematics*, 32:125–161, 1980.
[14] P. Flajolet. Approximate counting: a detailed analysis. *BIT*, 25:113–134, 1985.
[15] P. Flajolet and M. Golin. Mellin transforms and asymptotics. The mergesort recurrence. *Acta Inform.*, 31(7):673–696, 1994.
[16] P. Flajolet, X. Gourdon, and P. Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144:3–58, 1995.
[17] P. Flajolet, P. Grabner, P. Kirschenhofer, and H. Prodinger. On Ramanujan's $Q(n)$–function. *Journal of Computational and Applied Mathematics*, 58:103–116, 1995.
[18] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, and R. Tichy. Mellin transforms and asymptotics: Digital sums. *TCS*, 123(2):291–314, 1994.
[19] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, 3:216–240, 1990.
[20] P. Flajolet and H. Prodinger. Level number sequences for trees. *Discrete Mathematics*, 65:149–156, 1987.
[21] P. Flajolet, J.-C. Raoult, and J. Vuillemin. The number of registers required for evaluating arithmetic expressions. *TCS*, 9(1):99–125, 1979.
[22] P. Flajolet and R. Sedgewick. Digital search trees revisited. *SIAM Journal on Computing*, 15:748–767, 1986.
[23] P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: Finite differences and Rice's integrals. *Theoretical Computer Science*, 144:101–124, 1995.
[24] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, Cambridge, 2009.
[25] I. Goulden and D. Jackson. *Combinatorial Enumeration*. John Wiley, 1983.
[26] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics (Second Edition)*. Addison Wesley, 1994.
[27] D. H. Greene and D. E. Knuth. *Mathematics for the analysis of algorithms*. Birkhäuser, Boston, second edition, 1982.
[28] H.-K. Hwang. On convergence rates in the central limit theorems for combinatorial structures. *EJC*, 19:329–343, 1998.
[29] P. Kirschenhofer, H. Prodinger, and R. F. Tichy. A contribution to the analysis of in situ permutation. *Glasnik Mathematicki*, 22(42):269–278, 1987.
[30] A. Knopfmacher and H. Prodinger. On Carlitz compositions. *European Journal on Combinatorics*, 19:579–589, 1998.
[31] D. E. Knuth. *The Art of Computer Programming*, volume 1: Fundamental Algorithms. Addison-Wesley, 1973. Third edition, 1997.
[32] D. E. Knuth. *The Art of Computer Programming*, volume 3: Sorting and Searching. Addison-Wesley, 1973.
[33] D. E. Knuth. The average time for carry propagation. *Indagationes mathematicae*, 40:238–242, 1978.
[34] D. E. Knuth. The toilet paper problem. *Amer. Math. Monthly*, 91:465–470, 1984.
[35] J. H. Van Lint. Representation of 0 as $\sum_{k=-N}^{n} \varepsilon_k k$. *Proc. Amer. Math. Soc.*, 18:182–184, 1967.
[36] G. Louchard and H. Prodinger. Generalized approximate counting revisited. *Theoretical Computer Science*, 391:109–125, 2008.
[37] G. Louchard and H. Prodinger. Representations of numbers as $\sum_{k=-n}^{n} \varepsilon_k k$: A saddle point approach. In M. Archibald, V. Brattka, V. F. Goranko, and B. Löwe, editors, *LNCS 5489 (Infinity in Logic and Computation)*, pages 87–96. Springer Verlag, 2010.

[38] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 30:997–1015, 1978.

[39] N. F. Nörlund. *Vorlesungen über Differenzenrechnung*. Chelsea, New York, 1954.

[40] A. Odlyzko. *Asymptotic Enumeration Methods (in: Handbook of Combinatorics)*. Eds. R. Graham, M. Grötschel, L. Lovász. Elsevier Science, 1995.

[41] W. Panny and H. Prodinger. The expected height of paths for several notions of height. *Studia Scientiarum Mathematicarum Hungarica*, 20:119–132, 1985.

[42] H. Prodinger. The average height of a stack where three operations are allowed and some related problems. *Journal of Combinatorics, Information and System Sciences*, 5:287–304, 1980.

[43] H. Prodinger. Approximate counting via Euler transform. *Mathematica Slovaka*, 44:569–574, 1994.

[44] H. Prodinger. Combinatorics of geometrically distributed random variables: New q–tangent and q–secant numbers. *International Journal of Mathematical Sciences*, 24:825–838, 2000.

[45] H. Prodinger. The kernel method: a collection of examples. *Séminaire Lotharingien de Combinatoire*, B50f:19 pages, 2004.

[46] H. Prodinger. Approximate counting with m counters: a detailed analysis. *Theoretical Computer Science*, 439:58–68, 2012.

[47] L. Takács. Conditional limit theorems for branching processes. *Journal of Applied Mathematics and Stochastic Analysis*, 4:263–292, 1991.

[48] H. S. Wall. *Analytic Theory of Continued Fractions*. Chelsea Publishing Company, 1948.

[49] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, fourth edition, 1927. Reprinted 1973.

[50] D. A. Zave. A series expansion involving the harmonic numbers. *Inf. Process. Lett.*, 5(3):75–77, 1976.

Index

Department of Mathematics, University of Stellenbosch, Stellenbosch, 7602, South Africa
E-mail address: hproding@sun.ac.za