# THE AVERAGE CASE ANALYSIS OF ALGORITHMS

PHILIPPE FLAJOLET AND ROBERT SEDGEWICK
Rocquencourt & Princeton

*Counting and Generating Functions*

March 25, 1993

# Foreword

This report is part of a projected series whose aim is to present in a synthetic way the major methods and models in the average–case analysis of algorithms. The following items are to be treated in the series. First, there will be a collection of reports on *Methods*:

I. Counting and generating functions.

II. Asymptotics from generating functions.

III. Saddle point and transform methods.

Next, a collection of reports on *Models*:

IV. Strings, languages, and grammars.

V. Term trees and expression trees.

VI. Permutations, searching, and sorting.

VII. Digital Structures.

VIII. Mappings, occupancy, and hashing.

The present work (*I. Counting and Generating Functions*) introduces a symbolic approach through formal specifications to the analysis of basic combinatorial structures. It consists of three chapters: 1. Symbolic enumeration and ordinary generating functions; 2. Labelled structures and exponential generating functions; 3. Parameters and multivariate generating functions.

# Chapter 1

# Symbolic Enumeration and Ordinary Generating Functions

*Laplace discovered the remarkable correspondence between*
*set theoretic operations and operations on formal power series*
*and put it to great use to solve a variety of combinatorial problems.*
— GIAN–CARLO ROTA [13]

At an elementary level, combinatorial enumeration problems are often expressed by recurrences, with generating functions being then introduced as a privileged solution tool. We examine a framework that explains the surprising efficiency of generating functions for combinatorial enumeration problems. In fact, in combinatorial analysis and the average case analysis of algorithms, *generating functions are the central mathematical object*, rather than a mere tool.

This chapter introduces a *symbolic* approach to combinatorial enumerations. Its principle is that many general set–theoretic constructions have direct translations over generating functions. We give a catalogue based on a core of important constructions which includes the operations of union, cartesian product, sequence, set, multiset, and cycle. Supplementary operations like pointing and substitution can be also be treated by similar devices.

In this way, a specification language for elementary combinatorial objects is defined. The problem of enumerating a class of combinatorial structures then simply reduces to finding a proper specification, a sort of formal "grammar", for the class in terms of the basic constructions.

We show how to describe in such a context integer partitions and compositions, as well as several elementary string and tree enumeration problems. A parallel approach, to be developed in the next chapter, applies to labelled structures and exponential generating functions, and in comparison the plain structures considered in this chapter are often called *unlabelled*.

This methodology is susceptible to multivariate extensions with which many characteristic parameters of combinatorial objects can also be analyzed in a unified manner. It also has the merit of connecting nicely with complex asymptotic methods that can be applied directly to generating functions. These topics will be explored in later chapters.

## 1.1 Symbolic enumeration methods

In the framework to be described, classes of combinatorial structures are defined, either iteratively or recursively, in terms of simpler classes by means of a collection of elementary combinatorial constructions. The approach followed resembles the description of formal languages by means of context–free grammars, as well as the construction of structured data types in classical programming languages.

A *class of combinatorial structures* often called simply a *class* is a finite or denumerable set on which a *size* function is defined, the size of an element being a non negative integer. If $\mathcal{A}$ is a class, the size of an element $\alpha \in \mathcal{A}$ is denoted by $|\alpha|$, or $|\alpha|_{\mathcal{A}}$ in the few cases where the underlying class needs to be made explicit. Given a class $\mathcal{A}$, we consistently let $\mathcal{A}_n$ be the set of objects in $\mathcal{A}$ that have size $n$ and use the same group of letters for the counts $A_n = \mathrm{card}(\mathcal{A}_n)$ (alternatively, also $a_n = \mathrm{card}(\mathcal{A}_n)$). We further assume that the $\mathcal{A}_n$ are all finite. A more axiomatic presentation is then as follows: a class of combinatorial structures is a pair $(\mathcal{A}, |\,|)$ where $\mathcal{A}$ is at most denumerable and the mapping $|\,| \in \mathcal{A} \mapsto \mathbb{N}$ is such that the inverse image of any integer is finite.

The *counting sequence* of $\mathcal{A}$ is the sequence of integers $\{A_n\}_{n\geq 0}$. For instance, binary sequences $\mathcal{S}$, permutations $\mathcal{P}$, general plane trees (or ordered rooted trees) $\mathcal{G}$ constitute classes of combinatorial structures, with the usual conventions that the size of a word is its length, the size of a permutation is the number of its elements, the size of a tree is the number of its nodes. The corresponding counting sequences are then given by

$$S_n = 2^n, \quad P_n = n\,!, \quad G_n = \frac{1}{n}\binom{2n-2}{n-1}. \tag{1.1}$$

Average case analysis of algorithms typically reduces to counting problems for combinatorial structures like these.

A path often taken in the literature consists in decomposing the structures to be enumerated into smaller structures either of the same type or of simpler types, and then in extracting from such a decomposition *recurrence relations* satisfied by the $\{A_n\}$. In this context, the recurrence relations are either solved directly —whenever they are simple enough— or by means of generating function techniques.

The approach developed here is direct, more "symbolic", as it relies on a specification language for combinatorial structures. It is based on so–called *admissible constructions* that have the important feature of admitting direct translations into generating functions. In this chapter, we specifically examine constructions whose natural translation is in terms of ordinary generating functions.

The *ordinary generating function* (OGF) of a sequence $\{A_n\}$ is, we recall,

$$A(z) = \sum_{n=0}^{\infty} A_n z^n. \tag{1.2}$$

An alternative and more combinatorial form of counting generating functions is often useful; it is

$$A(z) = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|}, \tag{1.3}$$

as results from observing that the term $z^n$ occurs as many times as there are objects in $\mathcal{A}$ having size $n$.

We adhere consistently to the convention of representing classes, their counting sequences, and their generating functions, by the same groups of letters: for instance, $\mathcal{A}$ for a class, $\{A_n\}$ (or $\{a_n\}$) for the counting sequence, and $A(z)$ (or $a(z)$) for its OGF. The OGF's corresponding to Eq. (1.2) are then

$$S(z) = \frac{1}{1-2z}, \quad P(z) = \sum_{n=0}^{\infty} n!\, z^n, \quad G(z) = \frac{1-\sqrt{1-4z}}{2}.$$

The OGF's $S(z)$ and $G(z)$ exist as standard analytic objects since the series converge in a neighbourhood of 0 (they represent functions analytic at the origin), while the OGF $P(z)$ is a purely formal power series (its radius of convergence is 0) that can nonetheless be subjected to the algebraic treatment of power series.

**Definition 1.1** *Assume that $\Phi$ is a binary construction that associates to two classes $\mathcal{B}$ and $\mathcal{C}$ a new class*

$$\mathcal{A} = \Phi\{\mathcal{B}, \mathcal{C}\},$$

*in a finitary way (each $\mathcal{A}_n$ depends on finitely many of the $\mathcal{B}_n$ and $\mathcal{C}_n$). Then $\Phi$ is admissible iff the counting sequence $\{A_n\}$ of $\mathcal{A}$ is a function of the counting sequences $\{B_n\}$ and $\{C_n\}$ of $\mathcal{B}$ and $\mathcal{C}$ only:*

$$\{A_n\} = \Xi[\{B_n\}, \{C_n\}].$$

In that case, there exists a well defined operator $\Psi$ relating the corresponding ordinary generating functions

$$A(z) = \Psi[B(z), C(z)].$$

The definition given for a binary construction is readily extended to constructions of arbitrary sorts.

As an introductory example, take the construction of cartesian product that corresponds to forming "records" in classical programming languages. Assume that $\mathcal{A}$ is the cartesian product of $\mathcal{B}$ and $\mathcal{C}$,

$$\mathcal{A} = \mathcal{B} \times \mathcal{C}, \qquad (a)$$

the size of a pair $\alpha = (\beta, \gamma)$ being defined by $|\alpha|_{\mathcal{A}} = |\beta|_{\mathcal{B}} + |\gamma|_{\mathcal{C}}$. Then, considering all possibilities, the counting sequences corresponding to $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are related by the convolution relation

$$A_n = \sum_{k=0}^{n} B_k C_{n-k}. \qquad (b)$$

We recognize here a product of ordinary generating functions. Therefore, with $A(z) = \sum_{n\geq 0} A_n z^n$ etc, we find

$$A(z) = B(z) \cdot C(z). \qquad (c)$$

Thus in our terminology, the cartesian product is admissible and it directly translates into ordinary generating functions as a product.

The correspondence between Eq. $(a)$ and $(c)$ is clearly a very general one. Its merit is that it can be stated as a lemma that needs to be established only once. Once the problem of counting elements of a cartesian product is recognized, it becomes possible to dispense altogether with the intermediate stage of writing recurrences $(b)$. This is the spirit of the symbolic method for combinatorial enumerations. Its interest lies in the fact that several powerful set-theoretic constructions are amenable to this treatment.

EXERCISE 1. [For formalists only!] The valuation of a formal power series $f = \sum_n a_n z^n$ is defined as the smallest $r$ such that $a_r \neq 0$ and is denoted by $\mathrm{val}(f)$. Given two power series $f$ and $g$, their distance $d(f, g)$ is defined as $2^{-\mathrm{val}(f-g)}$.

Show that the space of all formal power series is an ultrametric space, i.e.,

$$d(f, g) \leq \max(d(f, h), d(h, g)),$$

for any $h$. Show that the operator $\Psi$ associated with an admissible construction is a continuous function over that space.

Define a corresponding ultrametrics over classes and prove that any admissible construction $\Phi$ is a continuous mapping.

## 1.2 Admissible constructions for OGFs

The main goal of this section is to introduce the admissible constructions that form the core of a specification language for combinatorial structures.

The *cartesian product* has, as we just saw, its usual set–theoretic meaning; in addition, the size of a pair is the sum of the sizes of its components.

We next consider the *disjoint union* also called the *sum* of classes, the intent being to capture the union of disjoint sets. We take the sum (as a combinatorial construction) of two classes $\mathcal{B}$ and $\mathcal{C}$ to represent the union (in the standard set–theoretic sense) of two disjoint copies, $\mathcal{B}^\circ$ and $\mathcal{C}^\circ$, of $\mathcal{B}$ and $\mathcal{C}$. One way of formalizing this notion is to introduce two distinct "markers" $\epsilon_1$ and $\epsilon_2$, each of size zero, and define the (disjoint) union $\mathcal{B} + \mathcal{C}$ of $\mathcal{B}, \mathcal{C}$ by

$$\mathcal{B} + \mathcal{C} = (\{\epsilon_1\} \times \mathcal{B}) \cup (\{\epsilon_2\} \times \mathcal{C}).$$

Disjoint union in the above sense is thus equivalent to a standard union whenever it is applied to disjoint sets. The size of an object in a disjoint union $\mathcal{A} = \mathcal{B} + \mathcal{C}$ is by definition inherited from its size in its class of of origin.

The reason for this definition of disjoint union is as follows. It would be inconvenient to have a construction that translates into generating functions under some external condition —disjointness— of a logical nature that would need to be proved separately. With our definition, the sum of two classes is always defined.

EXERCISE 2. With the above definition of disjoint union, show that the counting sequence of $\mathcal{B} = \mathcal{A} + \mathcal{A}$ is $b_n = 2a_n$. Find the generating function of $(\mathcal{A} + \mathcal{A}) \times (\mathcal{A} + \mathcal{A})$.

**1.** The main constructions of union, product, sequence, set, multiset, and cycle and their translation into generating functions.

| Construction | | OGF |
|---|---|---|
| Union | $\mathcal{A} = \mathcal{B} + \mathcal{C}$ | $A(z) = B(z) + C(z)$ |
| Product | $\mathcal{A} = \mathcal{B} \times \mathcal{C}$ | $A(z) = B(z) \cdot C(z)$ |
| Sequence | $\mathcal{A} = \mathfrak{S}\{\mathcal{B}\}$ | $A(z) = \dfrac{1}{1 - B(z)}$ |
| Set | $\mathcal{A} = \mathfrak{P}\{\mathcal{B}\}$ | $A(z) = \exp(B(z) - \frac{1}{2}B(z^2) + \cdots)$ |
| Multiset | $\mathcal{A} = \mathfrak{M}\{\mathcal{B}\}$ | $A(z) = \exp(B(z) + \frac{1}{2}B(z^2) + \cdots)$ |
| Cycle | $\mathcal{A} = \mathfrak{C}\{\mathcal{B}\}$ | $A(z) = \log \dfrac{1}{1 - B(z)} + \cdots$ |

**2.** The translation for sets, multisets, and cycles of small cardinality.

$$\begin{array}{ll}
\mathcal{A} = \mathfrak{P}\{\mathcal{B}, \text{card} = 2\} & \frac{B(z)^2}{2} - \frac{B(z^2)}{2} \\
\mathcal{A} = \mathfrak{M}\{\mathcal{B}, \text{card} = 2\} & \frac{B(z)^2}{2} + \frac{B(z^2)}{2} \\
\mathcal{A} = \mathfrak{C}\{\mathcal{B}, \text{card} = 2\} & \frac{B(z)^2}{2} + \frac{B(z^2)}{2} \\
\mathcal{A} = \mathfrak{P}\{\mathcal{B}, \text{card} = 3\} & \frac{B(z)^3}{6} - \frac{B(z)B(z^2)}{2} + \frac{B(z^3)}{3} \\
\mathcal{A} = \mathfrak{M}\{\mathcal{B}, \text{card} = 3\} & \frac{B(z)^3}{6} + \frac{B(z)B(z^2)}{2} + \frac{B(z^3)}{3} \\
\mathcal{A} = \mathfrak{C}\{\mathcal{B}, \text{card} = 3\} & \frac{B(z)^3}{3} + \frac{2B(z^3)}{3} \\
\mathcal{A} = \mathfrak{P}\{\mathcal{B}, \text{card} = 4\} & \frac{B(z)^4}{24} - \frac{B(z)^2B(z^2)}{4} + \frac{B(z)B(z^3)}{3} + \frac{B(z^2)^2}{8} - \frac{B(z^4)}{4} \\
\mathcal{A} = \mathfrak{M}\{\mathcal{B}, \text{card} = 4\} & \frac{B(z)^4}{24} + \frac{B(z)^2B(z^2)}{4} + \frac{B(z)B(z^3)}{3} + \frac{B(z^2)^2}{8} + \frac{B(z^4)}{4} \\
\mathcal{A} = \mathfrak{M}\{\mathcal{B}, \text{card} = 4\} & \frac{B(z)^4}{4} + \frac{B(z^2)^2}{2} + \frac{B(z^4)}{2}.
\end{array}$$

**3.** The additional constructions of pointing and substitution.

| Construction | | OGF |
|---|---|---|
| Pointing | $\mathcal{A} = \Theta\mathcal{B}$ | $A(z) = z\frac{d}{dz}B(z)$ |
| Substitution | $\mathcal{A} = \mathcal{B} \circ \mathcal{C}$ | $A(z) = B(C(z))$ |

Figure 1.1: A summary of the various constructions for unlabelled structures together with their translation into ordinary generating functions.

If $\mathcal{C}$ is a class then the *sequence* class $\mathfrak{S}\{\mathcal{C}\}$ is defined as the infinite sum

$$\mathfrak{S}\{\mathcal{C}\} = \{\epsilon\} + \mathcal{C} + (\mathcal{C} \times \mathcal{C}) + (\mathcal{C} \times \mathcal{C} \times \mathcal{C}) + \cdots$$

with $\epsilon$ being a "null" structure, meaning a structure of size 0. The null structure plays a rôle similar to that of the empty word in formal language theory, and the sequence construction is analogous to the Kleene star operation ($\mathcal{C}^*$). Notice however the fact that the various components in a sequence are individuated (graphically, their boundaries are separated by commas while word boundaries in a catenation product get erased).

It can then be checked that the construction $\mathcal{A} = \mathfrak{S}\{\mathcal{C}\}$ defines a proper class satisfying the finiteness condition for sizes if and only if $\mathcal{C}$ contains no object of size 0. From the definition of size for sums and products, there results that the size of a sequence is the sum of the sizes of its components:

$$\gamma = (\alpha_1, \ldots, \alpha_\ell) \qquad \Longrightarrow \qquad |\gamma| = |\alpha_1| + \cdots + |\alpha_\ell|.$$

The *powerset* class (or set class) $\mathcal{A} = \mathfrak{P}\{\mathcal{B}\}$ is defined as the class consisting of all *finite* subsets of class $\mathcal{B}$. *Multisets* are like sets except that repetitions of elements are allowed, the notation being $\mathcal{A} = \mathfrak{M}\{\mathcal{B}\}$. Another way of defining $\mathfrak{M}\{\mathcal{B}\}$ is as the quotient $\mathfrak{S}\{\mathcal{B}\}/\mathbf{R}$ with $\mathbf{R}$ the equivalence relation between sequences defined by $(\alpha_1, \ldots, \alpha_r)\,\mathbf{R}\,(\beta_1, \ldots, \beta_r)$ iff, there exists some permutation $\sigma$ of $[1\,..\,n]$ such that for all $j$, $\beta_j = \alpha_{\sigma(j)}$.

Directed *cycles* are just sequences defined up to cyclic permutations, the notation being $\mathfrak{C}\{\mathcal{B}\}$. Thus, $\mathfrak{C}\{\mathcal{B}\} = \mathfrak{S}\{\mathcal{B}\}/\mathbf{S}$ with $\mathbf{S}$ the equivalence relation between sequences defined by $(\alpha_1, \ldots, \alpha_r)\,\mathbf{S}\,(\beta_1, \ldots, \beta_r)$ iff there exists some cyclic permutation $\sigma$ of $[1\,..\,n]$ such that for all $j$, $\beta_j = \alpha_{\sigma(j)}$; in other words, for some $d$, $\beta_j = \alpha_{1+(j+d) \bmod n}$. We do not make much use of the cycle construction in the context of OGFs, in these notes; however, its counterpart plays a rather important rôle in the context of labelled structures and exponential generating functions.

We again need to make explicit the way the size function is defined when such constructions are performed. Like for products and sequences, the size of a composite object —set, multiset, or cycle— is defined as the sum of the sizes of its components.

In this framework, the class of all binary sequences is described as an *iterative* (i.e., non–recursive) class by

$$\mathcal{S} = \mathfrak{S}\{\mathcal{B}\} \qquad \text{where} \qquad \mathcal{B} = \{a, b\},$$

the ground alphabet, comprises two elements of size 1. The size of a binary sequence then coincides with its length (the number of letters it contains). In contrast, trees require *recursive* descriptions.

A *specification* for an $r+1$–tuple $(\mathcal{A}^{(0)}, \ldots, \mathcal{A}^{(r)})$ of classes is a collection of $r+1$ equations,

$$\begin{cases} \mathcal{A}^{(0)} & = & \Xi_0\{\mathcal{A}^{(0)}, \ldots, \mathcal{A}^{(r)}\} \\ \mathcal{A}^{(1)} & = & \Xi_1\{\mathcal{A}^{(0)}, \ldots, \mathcal{A}^{(r)}\} \\ & \cdots & \\ \mathcal{A}^{(r)} & = & \Xi_r\{\mathcal{A}^{(0)}, \ldots, \mathcal{A}^{(r)}\} \end{cases}$$

where $\Xi_j$ denotes a term built from the $\mathcal{A}$'s using the constructions of disjoint union, cartesian product, sequence, set, multiset, and cycle, as well as the "initial structures" $\mathcal{E}$, the class consisting of the empty structure only, and $\mathcal{N}$, the class consisting of a single object (node, letter) of size 1. We also say that the system is a specification of $\mathcal{A}^{(0)}$. A specification for a class of combinatorial structures is thus a sort of formal grammar defining that class.

For instance, the class $\mathcal{B}$ of pure plane binary trees (all nodes have degree 0 or 2) admits the one–line *recursive* specification

$$\mathcal{B} = \mathcal{N} + (\mathcal{N} \times \mathcal{B} \times \mathcal{B}),$$

where $\mathcal{N}$ is a class consisting of a single object of size 1, the generic node. The class $\mathcal{F}$ of ordered forests of binary trees is defined by

$$\begin{cases} \mathcal{B} & = & \mathcal{N} + (\mathcal{N} \times \mathcal{B} \times \mathcal{B}) \\ \mathcal{F} & = & \mathfrak{S}\{\mathcal{B}\} \end{cases}$$

Similarly, $\mathcal{G} = \mathfrak{M}\{\mathcal{B}\}$ represents the class of unordered forests of binary trees.

Two classes $\mathcal{A}$ and $\mathcal{B}$ are said to be *isomorphic* iff their counting sequences are identical. This is equivalent to saying that there exists a bijection from $\mathcal{A}$ to $\mathcal{B}$ that preserves size. Since we are interested only in counting problems, it is convenient to identify isomorphic classes and plainly consider them as identical. We then write simply $\mathcal{A} = \mathcal{B}$ in lieu of the more explicit$\mathcal{A} \simeq \mathcal{B}$. With this convention, for instance, the cartesian product becomes an associative operation,

$$\mathcal{A} \times (\mathcal{B} \times \mathcal{C}) = (\mathcal{A} \times \mathcal{B}) \times \mathcal{C}.$$

EXERCISE 3. Describe the objects of size $n$ in

$$\mathcal{B}' = \{\epsilon\} + (\mathcal{N} \times \mathcal{B}' \times \mathcal{B}'), \quad \mathcal{B}'' = \mathcal{N} + (\{\epsilon\} \times \mathcal{B}'' \times \mathcal{B}'').$$

Prove that $\mathcal{A}$ and $\mathcal{B}$ are isomorphic, where

$$\mathcal{A} = \mathcal{N} \times \mathfrak{S}\{\mathcal{A}\}, \quad \mathcal{B} = \mathcal{N} + (\mathcal{B} \times \mathcal{B}).$$

[Hint: the rotation correspondence between binary and general trees.]

EXERCISE 4. Establish the basic isomorphisms

$$(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C}), \qquad (\mathcal{A} + \mathcal{B}) \times \mathcal{C} = \mathcal{A} \times \mathcal{C} + \mathcal{B} \times \mathcal{C}.$$

**Definition 1.2** *A class of combinatorial structures is said to be* constructible *iff it admits a specification in terms of sum, product, sequence, set, multiset, and cycle constructions.*

EXERCISE 5. Show that the following classes are constructible: binary sequences, plane binary trees, general plane trees, integer partitions, integer compositions, general non–plane trees, binary sequences containing the pattern $bba$, binary sequences excluding the pattern $bba$.

It is easily recognized that this framework encompasses the usual one of context–free languages which, in a way, may be seen as the restricted case where only union, products, and sequences are allowed in specifications. There is however a subtle difference with formal language theory where, as already mentioned, the "boundaries" between words in a catenation product get erased. The exercises below (for language theorists) explore some of these differences.

EXERCISE 6. [For language theorists] Let $\mathcal{B} = \{a, b\}$ be a two letter alphabet. Prove the fundamental isomorphism

$$\mathfrak{S}\{\mathcal{B}\} = \mathfrak{S}\{b\} \times \mathfrak{S}\{a\} \times \mathfrak{S}\{b\},$$

by providing a direct combinatorial interpretation of both members of the equality.

Show that $\mathfrak{S}\{\mathfrak{S}\{\mathcal{B}\}\}$ does not define a class of combinatorial structures as the subset of elements of any given size is infinite, thus

$$\mathfrak{S}\{\mathfrak{S}\{\mathcal{B}\}\} \neq \mathfrak{S}\{\mathcal{B}\}.$$

Similarly, $\mathfrak{S}\{a, b\} \neq \mathfrak{S}\{\mathfrak{S}\{a\} \times \mathfrak{S}\{b\}\}$.

Thus the sequence construction $\mathfrak{S}\{\cdot\}$ differs from the language theoretic Kleene star in several aspects.

EXERCISE 7.[For formalists!] A specification is said to be *proper* iff it defines a class of combinatorial structures and thus satisfies the finiteness condition for size. State syntactical conditions and an algorithm to test specifications for properness.

[Hint: See [6], and model the treatment on the corresponding question for context–free languages.]

## 1.3 The admissibility theorem for OGFs

This section shows that any specification of a constructible class translates directly into generating function equations.

**Theorem 1.1** *The constructions of union, cartesian product, sequence, set, multiset, and cycle are all admissible. The associated operators are*

$$\begin{array}{llcl}
\textit{Union:} & \mathcal{A} = \mathcal{B} + \mathcal{C} & \Longrightarrow & A(z) = B(z) + C(z) \\
\textit{Product:} & \mathcal{A} = \mathcal{B} \times \mathcal{C} & \Longrightarrow & A(z) = B(z) \cdot C(z) \\
\textit{Sequence:} & \mathcal{A} = \mathfrak{S}\{\mathcal{B}\} & \Longrightarrow & A(z) = \dfrac{1}{1 - B(z)} \\
\textit{Set:} & \mathcal{A} = \mathfrak{P}\{\mathcal{B}\} & \Longrightarrow & A(z) = \exp(\displaystyle\sum_{k=1}^{\infty} (-1)^{k-1} \frac{B(z^k)}{k}) \\
\textit{Multiset:} & \mathcal{A} = \mathfrak{M}\{\mathcal{B}\} & \Longrightarrow & A(z) = \exp(\displaystyle\sum_{k=1}^{\infty} \frac{B(z^k)}{k}) \\
\textit{Cycle:} & \mathcal{A} = \mathfrak{C}\{\mathcal{B}\} & \Longrightarrow & A(z) = \displaystyle\sum_{k=1}^{\infty} \frac{\varphi(k)}{k} \log \frac{1}{1 - B(z^k)}
\end{array}$$

*where $\varphi(k)$ is the Euler totient function.*

The class $\mathcal{E} = \{\epsilon\}$ consisting of the empty structure only, and the class $\mathcal{N}$ consisting of a single "atomic" object (node, letter) of size 1 have OGFs

$$E(z) = 1 \qquad \text{and} \qquad N(z) = z.$$

**Proof.** *Union:* If $\mathcal{A} = \mathcal{B} + \mathcal{C}$, since the union is *disjoint*, and the size of an $\mathcal{A}$–element coincides with its size in $\mathcal{B}$ or $\mathcal{C}$, we have clearly $A_n = B_n + C_n$ and

$$A(z) = B(z) + C(z).$$

*Cartesian Product:* We have already seen the admissibility result for $\mathcal{A} = \mathcal{B} \times \mathcal{C}$. It follows from $A_n = \sum_{k=0}^{n} B_k C_{n-k}$ that

$$A(z) = B(z) \times C(z).$$

Notice also the alternative simple derivation based on the combinatorial form of GF's,

$$\sum_{\alpha \in \mathcal{A}} z^{|\alpha|} = \sum_{(\beta,\gamma) \in (\mathcal{B} \times \mathcal{C})} z^{|\beta|+|\gamma|} = \sum_{\beta \in \mathcal{B}} z^{|\beta|} \times \sum_{\gamma \in \mathcal{C}} z^{|\gamma|},$$

as follows from distributing products over sums. The result readily extends to an arbitrary number of factors.

*Sequence:* Admissibility for $\mathcal{A} = \mathfrak{S}\{\mathcal{B}\}$ (with $B_0 = 0$) follows from the union and product relations. Since

$$\mathcal{A} = \{\epsilon\} + \mathcal{B} + (\mathcal{B} \times \mathcal{B}) + (\mathcal{B} \times \mathcal{B} \times \mathcal{B}) + \cdots$$

we find

$$A(z) = 1 + B(z) + B^2(z) + B^3(z) + \cdots = \frac{1}{1 - B(z)},$$

where the geometric sum converges in the sense of formal power series since $[z^0]B(z) = 0$.

*Set (or powerset) construction:* Let $\mathcal{A} = \mathfrak{P}\{\mathcal{B}\}$ and first take $\mathcal{B}$ to be finite. Then, $\mathcal{A}$ is isomorphic to the the product

$$\prod_{\beta \in \mathcal{B}} (\{\epsilon\} + \{\beta\})$$

with $\epsilon$ a null structure of size 0, since distributing the product forms all possible combinations, i.e., sets of elements of $\mathcal{B}$. Thus, either directly from the combinatorial form (1.3) of OGFs' or from the sum and product rules, we have

$$B(z) = \prod_{\beta \in \mathcal{B}} (1 + z^{|\beta|}) = \prod_{n} (1 + z^n)^{b_n}.$$

Using the *exp-log* "trick", $A(z) = \exp(\log A(z))$, we find

$$\begin{aligned} A(z) &= \exp\big(\sum_{n=1}^{\infty} B_n \log(1 + z^n)\big) \\ &= \exp\big(\sum_{n=1}^{\infty} B_n \cdot \sum_{k=1}^{\infty} \frac{z^{nk}}{k}\big) \\ &= \exp\left(\frac{B(z)}{1} - \frac{B(z^2)}{2} + \frac{B(z^3)}{3} - \cdots\right), \end{aligned}$$

where the second line results from expanding the logarithm, and the third line results from exchanging summations.

The proof extends to the case of $\mathcal{B}$ infinite by noting that each $\mathcal{A}_n$ depends only on the $\mathcal{B}_j$ for $j \le n$, to which the relations given above for the finite case apply. More precisely, let $\mathcal{A}^{(\le n)} = \sum_{k=1}^{n} \mathcal{A}_j$ and $\mathcal{B}^{(\le n)} = \mathfrak{P}\{\mathcal{A}^{(\le n)}\}$. Then,

$$A(z) = A^{(\le n)}(z) + O(z^{n+1}) \qquad \text{and} \qquad B(z) = B^{(\le n)}(z) + O(z^{n+1})$$

while $A^{(\le n)}(z)$ and $B^{(\le n)}(z)$ are connected by the fundamental exponential relation, since $\mathcal{A}^{(\le n)}$ is finite. Letting $n$ tend to infinity, there follows

$$A(z) = \exp(\frac{B(z)}{1} - \frac{B(z^2)}{2} + \frac{B(z^3)}{3} - \cdots).$$

The necessary condition for convergence is again that $[z^0]B(z) = 0$, a restriction that also applies to multisets and cycles.

*Multiset:* First for finite $\mathcal{B}$ (with $B_0 = 0$), the multiset class $\mathcal{A} = \mathfrak{M}\{\mathcal{B}\}$ is definable as

$$\mathfrak{M}\{\mathcal{B}\} = \prod_{\beta \in \mathcal{B}} \mathfrak{S}\{\beta\},$$

since distributing the product forms all combinations with repetitions, i.e., multisets of elements of $\mathcal{B}$. This relation translates into generating functions by the product and sequence rules,

$$A(z) = \prod_{\beta \in \mathcal{B}} (1 - z^{|\beta|})^{-1} = \prod_{n=1}^{\infty} (1 - z)^{-B_n}$$

$$= \exp(\frac{B(z)}{1} + \frac{B(z^2)}{2} + \frac{B(z^3)}{3} + \ldots),$$

by the same type of computation as for sets. The case of an infinite class $\mathcal{B}$ follows also by a similar continuity argument.

*Cycle:* The translation of the cycle relation $\mathcal{A} = \mathfrak{C}\{\mathcal{B}\}$ is

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

where $\varphi(k)$ is the Euler totient function: $\varphi(k)$ equals the number of integers in $[1, k]$ that are relatively prime to $k$, with $\varphi(1) = 1$. The first terms, with $L_k = \log(1 - A(z^k))^{-1}$ are

$$A(z) = \frac{1}{1}L_1 + \frac{1}{2}L_2 + \frac{2}{3}L_3 + \frac{2}{4}L_4 + \frac{4}{5}L_5 + \frac{2}{6}L_6 + \frac{6}{7}L_7 + \cdots.$$

This result will not be proven here (and not used either). It was first established by Read within the framework of Pólya's theory of counting [12], an elementary combinatorial derivation appearing in [7]. □

The results for sets, multisets, and cycles are actually particular cases of the well known *Pólya theory* that deals more generally with the enumeration of objects under symmetry constraints [12].

At this stage, we have therefore defined a specification language for combinatorial structures which is some fragment of set theory. Each constructible class has by virtue of Theorem 1.1 an ordinary generating function that can be determined systematically, given a specification. In fact, it is even possible to use computer algebra systems in order to compute it *automatically*! See [6] for the description of such a system.

From Theorem 1.1, specifications translate into generating functions, which may be summarized by the following informally stated theorem.

**Theorem 1.2** *The generating function of a constructible class is a component of a system of generating function equations whose terms are built from* 1 *and* $z$ *using the operations*

$$+\ ,\quad \times\ ,\quad Q(f)=\frac{1}{1-f},\quad \Phi_{\mathrm{set}}[f],\quad \Phi_{\mathrm{mul}}[f],\quad \Phi_{\mathrm{cyc}}[z],$$

*where*

$$\Phi_{\mathrm{set}}[f]=\exp(\sum_{k=1}^{\infty}(-1)^{k-1}\frac{f(z^k)}{k}),\ \ \Phi_{\mathrm{mul}}[f]=\exp(\sum_{k=1}^{\infty}\frac{f(z^k)}{k}),$$

$$\Phi_{\mathrm{cyc}}[f]=\sum_{k=1}^{\infty}\log\frac{1}{1-f(z^k)}.$$

It can be further recognized that iterative (*i.e.*, non–recursive) structures have explicit generating functions involving the above operators only, while recursive structures have OGF's that are *a priori* accessible indirectly via systems of equations.

## 1.4 Compositions and Partitions

This section and the next one provide illustrations of the symbolic method of counting via specifications. In this framework, generating functions obtain with hardly any computation. The most direct applications described here relate to the additive decomposition of integers into summands with the classical combinatorial structures of partitions and compositions.

**Definition 1.3** *A composition of an integer $n$ is a sequence $(x_1, x_2, \ldots, x_k)$ of integers (for some $k \geq 1$) such that*

$$n = x_1 + x_2 + \cdots + x_k, \qquad 1 \leq x_j.$$

*A partition of an integer $n$ is a sequence $(x_1, x_2, \ldots, x_k)$ of integers ($k \geq 1$) such that*

$$n = x_1 + x_2 + \cdots + x_k \qquad \text{and} \qquad 1 \leq x_1 \leq x_2 \leq \cdots \leq x_k.$$

*In both cases, the $x_i$'s are called the summands or the parts.*

It is convenient to extend these definitions by regarding 0 as obtained by the empty sequence of summands ($k = 0$).

Compositions and partitions of integers are specified simply by means of the sequence and set constructions. From this, a whole collection of related enumeration results are derived in a uniform manner.

**Compositions.** First, from the definition, the class $\mathcal{C}$ of all compositions is described by an the equation between combinatorial classes,

$$\mathcal{C} = \mathfrak{S}\{\mathcal{I}\}, \tag{1.4}$$

where $\mathcal{I} = \{1, 2, \ldots\}$ is the class of summands that are integers at least 1 and the size of an integer $\sigma \in \mathcal{I}$ is taken to be $\sigma$ itself; the induced size of an element $\gamma \in \mathcal{C}$, $\gamma = (\sigma_1, \ldots, \sigma_k)$, is then the sum of the $\sigma_i$'s. Thus, counting the number of compositions of an integer $n$ is the same as enumerating $\mathcal{C}_n$.

The specification $\mathcal{C} = \mathfrak{S}\{\mathcal{I}\}$ has a direct translation into OGFs,

$$C(z) = \frac{z}{1 - I(z)}, \tag{1.5}$$

while the generating functions of summands derives directly from the definition,

$$I(z) = z + z^2 + z^3 + \cdots = \frac{z}{1-z}. \tag{1.6}$$

The collection of equations (1.5), (1.6) determines $C(z)$,

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

The counting problem for compositions is thus solved by a direct expansion,

$$C_n = 2^{n-1}, \quad n \geq 1; \qquad C_0 = 1.$$

When considering variations of this scheme, a number of counting results follow immediately by inspection. For instance, in order to enumerate the class $\mathcal{C}^{\{1,2\}}$ of compositions of $n$ whose parts are only allowed to be taken from the set $\{1, 2\}$, simply write

$$\mathcal{C}^{\{1,2\}} = \mathfrak{S}\{\mathcal{I}^{\{1,2\}}\} \qquad \text{with } \mathcal{I}^{\{1,2\}} = \{1, 2\}.$$

Thus, in terms of generating functions, the relation

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

holds (see Eq. (1.5)), with

$$I^{\{1,2\}}(z) = z + z^2.$$

Then,

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

and the number of compositions of $n$ in this class is a Fibonacci number,

$$C_n^{\{1,2\}} = F_{n+1} \quad \text{where} \quad F_n = \frac{1}{\sqrt{5}} \left[ (\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n \right].$$

Similarly, the compositions such that all their summands lie in the set $\{1, 2, \ldots, r\}$ have generating function

$$C^{\{1,\ldots,r\}}(z) = \frac{1}{1 - z - z^2 - \cdots z^r} = \frac{1}{1 - z\frac{1-z^r}{1-z}} = \frac{1-z}{1 - 2z + z^{r+1}},$$

and the corresponding counts are given by generalized Fibonacci numbers.

This fairly general process is encapsulated in the following proposition.

**Proposition 1.1** *Let $\mathcal{T}$ be a subset of the positive integers. The OGF of the class $\mathcal{D}$ of compositions with summands constrained to $\mathcal{T}$ is given by*

$$D(z) = \frac{1}{1 - T(z)} \qquad \textit{where} \qquad T(z) = \sum_{n \in \mathcal{T}} z^n.$$

EXERCISE 8. Give a combinatorial sum expressing $[z^n]C^{\{1,\dots,r\}}(z)$.
Show that, amongst all compositions of $n$, the proportion of those that have all their summands $\leq 10$ is exponentially small.

EXERCISE 9. Show that the OGFs of compositions with all summands even and all summands odd are respectively

$$\frac{1}{1-\frac{z^2}{1-z^2}} \quad \text{and} \quad \frac{1}{1-\frac{z}{1-z^2}}.$$

Express the corresponding counts in terms of known quantities and find direct combinatorial explanations.

Let now $\mathcal{C}^{(k)}$ denote the class of compositions made of $k$ summands, $k$ a fixed integer $\geq 1$. We have

$$\mathcal{C}^{(k)} = \mathcal{I} \times \mathcal{I} \times \cdots \times \mathcal{I},$$

where the number of terms in the cartesian product is $k$, and $\mathcal{I}$ still represents the summands, i.e., the class of positive integers. From there, the corresponding generating function is found to be

$$C^{(k)} = (I(z))^k \qquad \text{with} \qquad I(z) = \frac{z}{1-z}.$$

The number of compositions of $n$ having $k$ parts is thus

$$C_n^{(k)} = [z^n]\frac{z^k}{(1-z)^k} = \binom{n-1}{k-1},$$

a result which constitutes a combinatorial refinement of $C_n = 2^{n-1}$.

EXERCISE 10. Find a direct combinatorial proof of $C_n^{(k)} = \binom{n-1}{k-1}$.
Find a combinatorial sum for the number of compositions of $n$ that have $k$ summands, each at most $r$.

EXERCISE 11. Show that the number of lattice points with integer coordinates that belong to the closed ball of radius $n$ in $d$–dimensional space is

$$[z^{n^2}]\frac{1}{1-z}(\Theta(z))^d \qquad \text{where} \qquad \Theta(z) = 1 + 2\sum_{n=1}^{\infty} z^{n^2}.$$

(Such OGF's are useful in cryptography.)

| $n$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $+\infty$ |
|---|---|---|---|---|---|---|---|---|---|---|
| *Compositions* | 1 | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | $2^{n-1}$ |
| *Partitions* | 1 | 1 | 2 | 3 | 5 | 7 | 11 | 15 | 22 | $\sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$ |

Figure 1.2: The number of partitions and compositions for $n = 0\,..\,8$ and the corresponding asymptotic forms.

**Partitions.** Partitions have been defined as nondecreasing sequences of integers. They can be equivalently specified as multisets of summands,

$$\mathcal{P} = \mathfrak{M}\{I\},$$

since a multiset can always be presented in sorted order. Thus, the difference between compositions and partitions (the order of summands *does* or *does not* count) is reflected by the use of a sequence against a multiset construction.

The form of the partition generating function then derives from Theorem 1.1; the general translation mechanism provides the relation

$$P(z) = \exp\left(I(z) + \frac{1}{2}I(z^2) + \frac{1}{3}I(z^3) + \cdots\right) \quad \text{with} \quad I(z) = \frac{z}{1-z}. \tag{1.7}$$

In a special case like this, it is just as easy, however, to appeal directly to the product representation and get the more familiar form

$$P(z) = \prod_{m=1}^{\infty} \frac{1}{1-z^m}. \tag{1.8}$$

Contrary to compositions that are counted by the explicit formula $2^{n-1}$, so simple form exists for $p_n$. Asymptotic analysis based of the OGF (1.7) on the saddle point shows that $p_n = e^{O(\sqrt{n})}$. In fact a very famous theorem of Hardy and Ramanujan later improved by Rademacher, see [1], provides a full expansion of which the dominant term is

$$P_n \sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right).$$

There are consequently much fewer partitions than compositions.

EXERCISE 12. Show that

$$z\frac{P'(z)}{P(z)} = \sum_{n=1}^{\infty} \frac{nz^n}{(1-z^n)^2}.$$

Deduce a recurrence satisfied by the $P_n = [z^n]P(z)$ that involves a number theoretic function related to the divisor function. Show that $P_n$ can be computed in time polynomial in $n$.

A similar argument gives the generating function of partitions whose summands lie in the set $\{1, 2, \ldots, r\}$,

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

In general an analogue to Proposition 1.1 holds.

**Proposition 1.2** *Let $\mathcal{T}$ be a subset of the positive integers. The OGF of the class $\mathcal{Q}$ of partitions with summands constrained to $\mathcal{T}$ is given by*

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

EXERCISE 13. Represent partitions as collections of points ("Ferrers graphs") in the $\mathbb{N}\times\mathbb{N}$ lattice. Using a geometric symmetry that exchanges number of summands and value of largest summand, show that the OGF of partitions with $k$ summands is

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

Find the OGF of partitions into $k$ summand, each summand being $\leq r$.

Whenever summands are restricted to a finite set, the special partitions that result are called denumerants. A popular denumerant problem consists in finding the number of ways of giving change of 99 cents using coins that are pennies (1 ¢), nickels (5 ¢), dimes (10 ¢) and quarters (25 ¢). (The order in which the coins are taken does not matter and repetitions are allowed.)

For the case of a finite $\mathcal{T}$, we predict from Proposition 2 that $Q(z)$ is always a *rational* function with poles that are roots of unity and the $Q_n$ satisfy a linear recurrence related to the structure of $\mathcal{T}$.

The solution to the original coin change problem is found to be

$$[z^{99}]\,\frac{1}{(1-z)(1-z^5)(1-z^{10})(1-z^{25})}=242.$$

In the same vein, one proves [4, p. 108] that

$$P_n^{\{1,2\}}=\lceil\frac{2n+3}{4}\rfloor\qquad P_n^{\{1,2,3\}}=\lceil\frac{(n+3)^2}{12}\rfloor.$$

There $\lceil x\rfloor\equiv\lfloor x+\frac{1}{2}\rfloor$ denotes the integer closest to real number $x$. Such results are typically obtained by the two step process: (i) decompose the rational generating function into simple fractions; (ii) compute the coefficients of each simple fraction and combine them to get the final result.

EXERCISE 14. Let $k$ be a fixed integer; show that, with $\mathcal{T}=\{1,\ldots,k\}$,

$$Q_n\sim c_k n^{k-1}\quad\text{with}\quad c_k=\frac{1}{k!(k-1)!}.$$

Obtain the asymptotic form of $Q_n$ when $\mathcal{T}$ is an arbitrary finite set:

$$Q_n\sim\frac{1}{\tau}\frac{n^{k-1}}{(k-1)!}\quad\text{with}\quad\tau=\prod_{n\in\mathcal{T}}n.$$

(This is due to Schur.)

## 1.5 Languages and specifications

The purpose of this section is to illustrate the way complex specifications of combinatorial structures translate into systems of equations over generating functions. We adopt as leading examples some classical formal languages (a language is a set of strings).

First, we consider words or strings over a binary alphabet of letters, $\mathcal{A}=\{a,b\}$. Then, the class of all strings (words) over $\mathcal{A}$ is

$$\mathcal{W}=\mathfrak{S}\{\mathcal{A}\}.$$

From the obvious fact that $A(z)=2z$, we get

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

and the number of binary words of length $n$ is found to be $W_n = 2^n$, as expected!

There is however a less obvious way to construct binary strings. It is based on the observation that (with a minor adjustment at the beginning) a string decomposes into a succession of "blocks" each formed with a single $b$ followed by an arbitrary (possibly empty) sequence of $a$'s. For instance *aaabaababaabbabbaaa* decomposes as

$$aaa \mid\mid baa \mid ba \mid baa \mid b \mid ba \mid b \mid baaa.$$

Omitting redundant $\{\,\}$'s, we have the alternative decomposition:

$$\mathcal{W} = \mathfrak{S}\{a\} \cdot \mathfrak{S}\{b \cdot \mathfrak{S}\{a\}\}. \tag{1.9}$$

A check is provided by computing again the OGF corresponding to this new specification,

$$W(z) = \frac{1}{1-z}\frac{1}{1-z\frac{1}{1-z}}, \tag{1.10}$$

which reduces to $(1-2z)^{-1}$.

**Longest runs.** The interest of the decomposition that we have just seen for strings is to take into account various other interesting properties, for example the longest runs. Define $a^{\leq k}$ to be the collection of all words formed with the letter $a$ only, whose length is between 0 and $k-1$; the corresponding OGF is

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

The collection $\mathcal{W}^{\langle k\rangle}$ of words which do not have $k+1$ consecutive $a$'s is described by an amended form of (1.9), namely

$$\mathcal{W}^{\langle k\rangle} = a^{\leq k} \cdot \mathfrak{S}\{ba^{\leq k}\}. \tag{1.11}$$

The corresponding OGF obtains immediately from (1.11)

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

This is therefore the generating functions of words whose longest run of $a$'s is of length at most $k$.

From this computation and some asymptotic analysis, it can be deduced that the longest run of $a$'s in a random binary string of length $n$ is about $\log_2 n$. This will be further explored in later chapters.

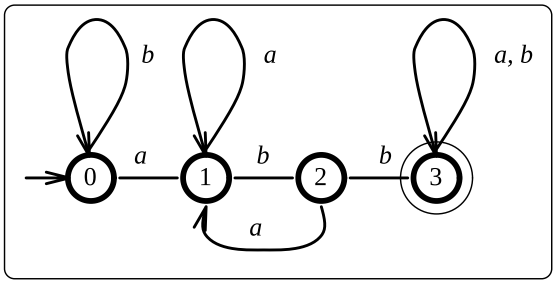



Figure 1.3: Words that contain the pattern *abb* are recognized by a 4–state automaton with initial state $q_0$, final state $q_3$.

---

EXERCISE 15. From the isomorphism

$$\mathfrak{S}\{a, b\} = \mathfrak{S}\{a\}\mathfrak{S}\{b\}\mathfrak{S}\{a\mathfrak{S}\{a\}\,b\mathfrak{S}\{b\}\},$$

deduce that the OGF of words with no more than $k$ consecutive $a$'s and no more than $k$ consecutive $b$'s is

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

**Patterns.** The previous example has shown some uses of the sequence and product constructions in order to describe the occurrence of special patterns in strings, namely runs. We now examine an instance of a general pattern occurrence problem, where the pattern need no longer have a simple structure.

Consider the class $\mathcal{L}$ of all words that contain the pattern *abb* as a factor (the letters of the pattern should appear contiguously). Such words are recognized by a finite automaton with 4 states, $q_0, q_1, q_2, q_3$. The construction is classical: state $q_j$ is interpreted as meaning "*the first j characters of the pattern have just been scanned*", and the corresponding automaton appears in Figure 1.5. The initial state is $q_0$, and there is a unique final state $q_3$.

The language accepted by the automaton is by definition the collection of words formed with the edge labels read along any path in the graph that starts with the initial state ($q_0$) and end at the final state (here $q_3$).

For $j \in \{0, 1, 2, 3\}$, introduce the class (language) $\mathcal{L}_j$ of all words such that the automaton, when started in state $q_j$, terminates in the final state

($q_3$). Since the initial state is $q_0$, we have $\mathcal{L} \equiv \mathcal{L}_0$. The languages are connected by the system of equations

$$\begin{aligned}
\mathcal{L}_0 &= a\mathcal{L}_1 + b\mathcal{L}_0 \\
\mathcal{L}_1 &= a\mathcal{L}_1 + b\mathcal{L}_2 \\
\mathcal{L}_2 &= a\mathcal{L}_1 + b\mathcal{L}_3 \\
\mathcal{L}_3 &= a\mathcal{L}_3 + b\mathcal{L}_3 + \epsilon,
\end{aligned}$$

which directly reflects the graph structure of the automaton. This gives rise to a set of equations for the associated OGFs

$$\begin{aligned}
L_0 &= zL_1 + zL_0 \\
L_1 &= zL_1 + zL_2 \\
L_2 &= zL_1 + zL_3 \\
L_3 &= zL_3 + zL_3 + 1.
\end{aligned}$$

Solving the system, we find the OGF of all words containing the pattern $abb$: it is $L_0(z)$ since the initial state of the automaton is $q_0$, and

$$L_0(z) = \frac{z^3}{(1-z)(1-2z)(1-z-z^2)}.$$

The partial fraction decomposition is

$$L_0(z) = \frac{1}{1-2z} - \frac{2+z}{1-z-z^2} + \frac{1}{1-z},$$

so that

$$L_{0,n} = 2^n - 2F_{n+1} - F_n + 1.$$

In particular $2^n - L_{0,n}$ behaves like $O(\varphi^n)$ with $\varphi$ the golden ratio. In other words, all but an exponentially vanishing proportion of the strings of length $n$ contain the given pattern $abb$, a fact that was to be expected on probabilistic grounds.

The process developed here for counting patterns in strings is clearly a general one. This fact was first discovered in the late 1950's by Chomsky and Schützenberger.

**Theorem 1.3** *Any regular language (a language is regular if it is recognized by a deterministic finite automaton) admits a generating function that is rational.*

EXERCISE 16. Given any fixed pattern string $\pi$, show constructively that the OGF of words that contain $\pi$ is a rational function. Write a computer algebra programme that computes the OGF given $\pi$.

[Guibas and Odlyzko] Relate the OGF to the overlap structure of the pattern.

EXERCISE 17. Using the fact that any regular language $L$ is recognised by a deterministic finite automaton, show that the ogf $l(z)$ of $L$ is a component of a linear system of equations with polynomial coefficients. Deduce an expression of $l(z)$ as a quotient of two sparse determinants.

EXERCISE 18. [Chomsky–Schützenberger] The OGF of an unambiguous context free (and the OGF of an arbitrary context free language with words counted with their ambiguity multiplicity) satisfies a system of polynomials equations and is thus an algebraic functions.

## 1.6 Trees and recursive specifications

This section shows some basic tree enumerations. There, recursive specifications normally lead to nonlinear equations (and systems of equations) over generating functions.

**Plane trees.** Plane trees are also called sometimes ordered trees, since the subtrees dangling from a node are ordered between themselves. Alternatively, they may be viewed as abstract graph structures accompanied by an embedding into the plane. They are precisely described in terms of unions, cartesian products, and sequence constructions.

First, the family $\mathcal{G}$ of "general" plane trees (where all node degrees are allowed) satisfies

$$\mathcal{G} = \mathcal{N} \times \mathfrak{S}\{\mathcal{G}\},$$

where $\mathcal{N}$ represents the class comprising a single node of size 1. For corresponding OGFs, we have $N(z) = z$, and $G(z)$ is determined by the equation

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

The equation reduces to a quadratic equation, $G(1 - G) = z$, whose solution is

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

As a result, the number of general trees of size $n$ is

$$G_n = \frac{1}{n}\binom{2n-2}{n-1} = \frac{1}{2n-1}\binom{2n-1}{n} = \frac{(2n-2)!}{n!\,(n-1)!},$$

a Catalan number. An asymptotic equivalent results from Stirling's formula:

$$G_n \sim \frac{4^{n-1}}{\sqrt{\pi n^3}}.$$

Next, the family $\mathcal{B}$ of binary trees is specified by

$$\mathcal{B} = \mathcal{E} + (\mathcal{N} \times \mathcal{B} \times \mathcal{B}), \tag{1.12}$$

where $\mathcal{E}$ will momentarily represent the type of an external node. The transition to OGF'ss depends on what is fixed as the notion of size. Usually, one considers the size of a tree to be the number of its internal nodes. In that case, $\mathcal{E}$ is taken to be the class $\{\epsilon\}$ which may be viewed as representing the class $\{\epsilon\}$ of the null pointer (of size 0), so that we have $E(z) = 1$, and Eq. (1.12) gives

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

Thus,

$$B(z) = \frac{1 - \sqrt{1 - 4z}}{2z},$$

so that

$$B_n = G_{n+1} = \frac{1}{n+1}\binom{2n}{n}.$$

**Note.** Two other generating functions arise naturally: the first one, $\beta_1(z)$, is relative to size measured by number of external nodes; the second one, $\beta_2(z)$, is relative to size measured by the total number of nodes, internal and external nodes all contributing. The equations are then

$$\beta_1(z) = z + (\beta_1(z))^2; \qquad \beta_2(z) = z + z(\beta_2(z))^2,$$

from which we find

$$\beta_1(z) = \frac{1-\sqrt{1-4z}}{2}; \qquad \beta_2(z) = \frac{1-\sqrt{1-4z^2}}{2z}.$$

The relations

$$\beta_1(z) = zB(z); \qquad \beta_2(z) = zB(z^2)$$

imply

$$\beta_{1,n} = B_{n-1} \qquad \text{and} \qquad \beta_{2,2n+1} = B_n.$$

This reflects obvious linear relations between the different notions of size in binary trees: a tree with $n$ internal binary nodes is composed of $n+1$ external nodes and has $2n+1$ nodes in total.

In general, it is possible to determine generating functions for any family of plane trees defined by restrictions on allowed node degrees.

**Proposition 1.3** *Let $\Omega$ be a subset of the natural integers. The generating function for the class $\mathcal{T}$ of trees all of which nodes have outdegree belonging to $\Omega$ is determined implicitly by the equation*

$$T(z) = z\,\omega(T(z)) \qquad \textit{where} \qquad \omega(u) = \sum_{k\in\Omega} u^k.$$

We shall see in the next chapter that such generating functions may be systematically expanded by means of an inversion formula due to Lagrange.

Exercise 19. Let $M(z)$ be the generating function for trees where nodes have degrees $0, 1, 2$ only. The coefficients $M_n = [z^n]M(z)$ are sometimes referred to as Motzkin numbers. Prove

$$M(z) = z(1 + M(z) + M^2(z)) \quad M(z) = \frac{1-z-\sqrt{1-2z-3z^2}}{2z}.$$

By expanding the OGF, find two different expressions for $M_n$.

| $n$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $n$ | $+\infty$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Plane general* | 0 | 1 | 1 | 2 | 5 | 14 | 42 | 132 | 429 | $\frac{1}{n}\binom{2n-2}{n-1}$ | $\sim \frac{4^{n-1}}{\sqrt{\pi n^3}}$ |
| *Plane binary* | 1 | 1 | 2 | 5 | 14 | 42 | 132 | 429 | 1430 | $\frac{1}{n+1}\binom{2n}{n}$ | $\sim \frac{4^n}{\sqrt{\pi n^3}}$ |
| *Unord. general* | 0 | 1 | 1 | 2 | 4 | 9 | 20 | 48 | 115 | – | $C \cdot \frac{A^n}{n^{3/2}}$ |
| *Unord. binary* | 1 | 1 | 1 | 2 | 3 | 6 | 11 | 23 | 46 | – | $C_2 \cdot \frac{A_2^n}{n^{3/2}}$ |

Figure 1.4: The number of trees of various sorts for $n = 0\,..\,8$ and the corresponding asymptotic forms ($C = 0.43992$, $A = 2.95576$; $C_2 = 0.79160$, $A_2 = 2.48325$).

**Unordered trees** An *unordered tree* (sometimes called non–plane tree) is a tree in the general graph–theoretic sense so that there is no order distinction between subtrees dangling from a common root. The unordered trees considered here are furthermore rooted. Accordingly, an unordered tree is a root node linked to a multiset of trees, in the language of constructible structures. Thus, the class $\mathcal{U}$ of all unordered trees, admits the recursive specification

$$\mathcal{U} = \mathcal{N} \times \mathfrak{M}\{\mathcal{U}\},$$

which translates into the functional equation

$$U(z) = z \exp\left(U(z) + \frac{1}{2}U(z^2) + \frac{1}{3}U(z^3) + \cdots\right).$$

This equation does not admit a closed form solution. However, the local analysis of its singularities yields a *bona fide* asymptotic expansion for $U_n$, a fact first discovered by Polya [12] who proved that

$$U_n \sim C \cdot \frac{A^n}{n^{3/2}}, \qquad (1.13)$$

for some positive constants $C$ and $A$.

The example of unordered trees is instructive: Asymptotic expansions to be examined later can often proceeed directly from functional equations even in cases where no clear algebraic structure is available to reach explicit forms for coefficients.

EXERCISE 20. Show that

$$\sum_{n=0}^{\infty} U_{n+1} z^n = \prod_{n=1}^{\infty} (1 - z^n)^{-U_n}.$$

By logarithmic differentiation, derive for the $U_n$ a recurrence that permits to compute $U_n$ in time polynomial in $n$.

Regarding unordered binary trees counted according to the number of binary nodes, we have the equation

$$\mathcal{V} = \mathcal{E} + \mathcal{N} \times \mathfrak{M}\{\mathcal{V}, \text{card} = 2\}, \tag{1.14}$$

where $\mathfrak{M}\{\mathcal{X}, \text{card} = 2\}$ represents the class of all unordered pairs with repetition allowed (i.e., multisets of "cardinality" equal to 2) of elements of $\mathcal{X}$. From there, the associated functional equation is found to be

$$V(z) = 1 + \frac{z}{2}(V(z)^2 + V(z^2)), \tag{1.15}$$

a fact that results from the developments below concerning sets of fixed cardinality. The corresponding asymptotic analysis was carried out by Otter around 1948 who found a result of the same shape as (1.13).

EXERCISE 21. Prove equation (1.15) by first establishing a recurrence for the numbers $V_n$.

Find the corresponding equations for ternary trees and quaternary trees. Can you describe a general pattern?

EXERCISE 22. Define the class of hierarchies to be trees without nodes of outdegree 1. Show that the corresponding OGFs, $H(z)$ and $\widehat{H}(z)$, in the planar and non–planar cases respectively, satisfy

$$H = z + \frac{zH^2}{1-H} \qquad \text{and} \qquad \widehat{H}(z) = z\left[\exp(\widehat{H}(z) + \frac{\widehat{H}(z^2)}{2} + \cdots) - \widehat{H}(z)\right].$$

Find an explicit form for $H_n$. (Such objects arise in statistical classification theory.)

The enumeration of the class of trees defined by an arbitrary set $\Omega$ of nodes degree results from the treatment of sets of fixed cardinality given below. We shall leave aside the discussion of unrooted trees which is treated extensively in Harary and Palmer's book [9].

## 1.7 Bijections and specifications

The point to be illustrated here is that some amount of "combinatorial pre-processing" is sometimes necessary in order to bring combinatorial structures into the framework of symbolic methods. We present an example where string encodings provide the solution to a counting problem relative to set partitions.

Let $\mathcal{S}_n^{(k)}$ denote the collection of all partitions of the set $[1\,..\,n]$ into $k$ non–empty blocks and $S_n^{(k)} = ||\mathcal{S}_n^{(k)}||$ the corresponding cardinality. The basic object under consideration here is a *set partition*, not to be confused with the integer partitions considered earlier.

It is possible to find an encoding of partitions in $\mathcal{S}_n^{(k)}$ of an $n$–set into $k$ blocks by words over a $k$ letter alphabet, $\mathcal{B} = \{b_1, b_2, \ldots, b_k\}$ as follows.

Consider a fixed partition $\varpi$ formed of $k$ blocks. Identify each block by its smallest element called the block *leader*; then sort the block leaders in increasing order. Define the index of a block as the rank of its leader amongst all the $k$ leaders, with ranks conventionally starting at 1. Each element of $[1\,..\,n]$, can also be associated to an index in $[1\,..\,k]$ that is the index of the block to which the element belongs.

Scan the elements 1 to $n$ in order and produce sequentially $n$ letters from the alphabet $\mathcal{B}$. When $j \in [1\,..\,n]$ is scanned, if $r$ is its index, then the letter $b_r$ is produced. In this way, a partition is encoded as a word of length $n$ over $\mathcal{B}$ with the additional properties that: (i) all $k$ letters occur; (ii) the first occurrence of $b_1$ precedes the first occurrence of $b_2$ which itself precedes the first occurrence of $b_3$, etc. For instance to $n = 5$, $k = 3$, the set partition

$$\varpi = \{\{4\}, \{5, 1, 2\}, \{3\}\},$$

is reorganized by putting leaders in first position of the blocks, and then sorting the blocks by increasing order of their leaders:

$$\varpi = \{\{1, 2, 5\}, \{3\}, \{4\}\}.$$

The table that associates indices to elements and the string encoding are thus

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 2 & 3 & 1 \end{pmatrix} \quad \text{and} \quad b_1 b_1 b_2 b_3 b_1.$$

In this fashion, we have mapped $\mathcal{S}_n^{(k)}$ into words of length $n$ in the language

$$b_1\,\mathfrak{S}\{b_1\}\cdot b_2\,\mathfrak{S}\{b_1+b_2\}\cdot b_3\,\mathfrak{S}\{b_1+b_2+b_3\}\ \cdots\ b_k\,\mathfrak{S}\{b_1+b_2+\cdots+b_k\}. \tag{1.16}$$

The encoding is clearly revertible.

By the sum, product, and sequence rules, the language specification (1.16) translates into OGF

$$\begin{aligned} S^{(k)}(z) &= z\frac{1}{1-z}\cdot z\frac{1}{1-2z}\cdot z\frac{1}{1-3z}\ \cdot\ z\frac{1}{1-kz} \\ &= \sum_{n=0}^{\infty} S_n^{(k)} z^n = \frac{z^k}{(1-z)(1-2z)\cdots(1-kz)}. \end{aligned}$$

The partial fraction expansion of $S^{(k)}(z)$ is readily computed,

$$S^{(k)}(z) = \frac{1}{k!}\sum_{j=0}^{k}\binom{k}{j}\frac{(-1)^{k-j}}{1-jz},$$

and by a direct expansion,

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

In particular, one has

$$S_n^{(1)} = 1;\quad S_n^{(2)} = \frac{1}{2!}(2^n-2);\quad S_n^{(3)} = \frac{1}{3!}(3^n - 3\cdot 2^n + 3).$$

These numbers are known as the Stirling numbers of the second kind, or better, as the Stirling partition numbers, and the $S_n^{(k)}$ are often denoted by $\left\{ {n \atop k} \right\}$. They will be examined further in the next chapter.

Eventually, the counting of set partitions could be done successfully thanks an encoding into strings, the corresponding language forming a constructible class of combinatorial structures (actually a regular language). In the next chapter, we shall examine another approach to the counting of set partitions that is based on labelled structures and exponential generating functions.

Exercise 23. Find a direct recurrence on the $S_n^{(k)}$ and deduce the exponential generating function:

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

Also derive this formula directly from the OGF. [Hint: See Comtet's book [4, Ch. V].]

EXERCISE 24. Find the number of words of length $n$ containing $k$ distinct letters, and such that no two consecutive letters are identical [4, p. 219].

## 1.8 Additional constructions

Constructions with a fixed number of components, like sets with a fixed number of elements, are admissible as explained here. In addition, two less frequently used constructions, pointing and composition, also translate agreeably into generating functions.

**Fixed cardinality.** An immediate formula for OGF's is that of the *diagonal* later of a cartesian product defined as $\mathcal{B} = \{(\alpha, \alpha) \mid \alpha \in \mathcal{A}\}\}$. Then, clearly $B_{2n} = A_n$ so that

$$B(z) = A(z^2).$$

With the diagonal, we can construct the unordered pairs of (distinct) elements of $\mathcal{A}$, which we write as $\mathcal{B} = \mathfrak{P}\{\mathcal{A}, \text{card} = 2\}$. An informal argument then runs as follows: the unordered pair $\{\alpha, \beta\}$ is associated to the two ordered pairs $(\alpha, \beta)$ and $(\beta, \alpha)$ except when $\alpha = \beta$. Thus, pairs of distinct elements are obtained by taking all ordered pairs to which a multiplicity of $\frac{1}{2}$ is assigned and subtracting the diagonal with a multiplicity of $\frac{1}{2}$. The resulting OGF schema is thus

$$\mathcal{B} = \mathfrak{P}\{\mathcal{A}, \text{card} = 2\} \qquad \Longrightarrow \qquad B(z) = \frac{1}{2}A(z)^2 - \frac{1}{2}A(z^2).$$

Similarly, for multisets, we find

$$\mathcal{C} = \mathfrak{M}\{\mathcal{A}, \text{card} = 2\} \qquad \Longrightarrow \qquad C(z) = \frac{1}{2}A(z)^2 + \frac{1}{2}A(z^2).$$

This type of direct reasoning could be extended to treat triples, and so on, but the computations (if not the reasoning) tend to grow out of control. An approach based on multivariate generating functions is preferable.

More generally, we consider refinements of the set and multiset constructions. We let $\mathfrak{P}\{\mathcal{A}, \text{card} = k\}$, $\mathfrak{M}\{\mathcal{A}, \text{card} = k\}$, $\mathfrak{C}\{\mathcal{A}, \text{card} = k\}$ represent the collection of sets, multisets, and cycles built out of $\mathcal{A}$ and formed with $k$ components.

**Theorem 1.4** *The OGF of sets, $\mathcal{B} = \mathfrak{P}\{\mathcal{A}, \text{card} = k\}$, is a polynomial in the quantities $A(z), A(z^2), A(z^3)$, etc,*

$$B(z) = [u^k] \exp(\frac{u}{1}A(z) - \frac{u^2}{2}A(z^2) + \frac{u^3}{3}A(z^3) - \cdots).$$

*The OGF of multisets, $\mathcal{B} = \mathfrak{M}\{\mathcal{A}, \text{card} = k\}$, is*

$$B(z) = [u^k] \exp(\frac{u}{1}A(z) + \frac{u^2}{2}A(z^2) + \frac{u^3}{3}A(z^3) + \cdots).$$

*The OGF of cycles, $\mathcal{B} = \mathfrak{C}\{\mathcal{A}, \text{card} = k\}$, is*

$$B(z) = [u^k] \sum_{\ell=1}^{\infty} \log \frac{1}{1 - u^k A(z^k)}.$$

The explicit forms for small values of $k$ are summarized in Figure 1.1. The proof based on bivariate generating functions will be discussed in the chapter devoted to parameters and multivariate generating functions.

Exercise 25. Write a computer algebra programme for the general cases of Theorem 1.3. How many monomials do the involved polynomials contain?

Exercise 26. Let $\mathcal{A}$ be the class of the finite sets of elements from $\mathcal{B}$, with the additional constraint that no two elements in a set have the same size. Prove that

$$A(z) = \prod_{n=1}^{\infty}(1 + B_n z^n).$$

**Pointing and substitution.** The idea underlying the definitions of pointing and substitution is that combinatorial structures are often formed of "atoms" (words are formed with letters, graphs with nodes...) which determine their sizes. In this context, the notions of pointing and substitution can be interpreted in a simple manner: Pointing means "pointing a distinguished atom"; substitution $\mathcal{B}[\mathcal{C}]$ means "substitute elements of $\mathcal{C}$ for atoms of $\mathcal{B}$".

*Pointing.* Formally, the pointing of a class $\mathcal{B}$, noted $\mathcal{A} = \Theta\mathcal{B}$, is defined by

$$\Theta\mathcal{B} = \sum_{n\geq 0} \mathcal{B}_n \times [1\,..\,n]\,.$$

Then clearly

$$A_n = n \cdot B_n \qquad \text{and} \qquad A(z) = z\frac{d}{dz}B(z).$$

(The operator $z\frac{d}{dz}$ itself is occasionally denoted by $\Theta$.)

If $B_n$ is the number of $\mathcal{B}$ structures of size $n$, then $nB_n$ can be interpreted as counting pointed structures where one of the $n$ atoms has been distinguished (for instance by a special pointer of size 0 attached to it).

Though we do not make a heavy use of this construction, it sometimes appears handy as it may account for the occurrence of derivatives in various contexts.

*Substitutions.* Formally, if $\mathcal{B}_n$ is the set of objects in $\mathcal{B}$ having size $n$, the substitution of $\mathcal{C}$ into $\mathcal{B}$ (also known as composition of $\mathcal{B}$ and $\mathcal{C}$), noted $\mathcal{B}\circ\mathcal{C}$ or $\mathcal{B}[\mathcal{C}]$, is defined as

$$\mathcal{B}\circ\mathcal{C} \equiv \mathcal{B}[\mathcal{C}] = \sum_{k\geq 0} \mathcal{B}_k \times (\mathcal{C}\times\mathcal{C}\times\cdots\times\mathcal{C}) \qquad (k \text{ factors } \mathcal{C}).$$

Elements of $\mathcal{B}\circ\mathcal{C}$ may thus be viewed as obtained by selecting in all possible ways an element $\beta\in\mathcal{B}$ and substituting for each of its atom an arbitrary element of $\mathcal{C}$.

From the definition, if $\mathcal{A} = \mathcal{B}[\mathcal{C}]$, we have by the sum and product rules

$$\begin{aligned} A(z) &= \sum_{n\geq 0} B_n \cdot (C(z))^k \\ &= B(C(z)). \end{aligned}$$

For instance the class $\mathcal{U}$ of balanced 2-3 trees, a familiar data structure [11], is defined as trees whose internal nodes have degree 2 or 3 such that all leaves are at the same distance from the root. Since only leaves contain informations, it is customary to adopt as a measure of size of such trees the number of their leaves. Let $\mathcal{N}$ denote a generic leaf of size 1. By growing trees at the leaves, we find the equation

$$\mathcal{U} = \mathcal{N} + \mathcal{U}[(\mathcal{N} \times \mathcal{N}) + (\mathcal{N} \times \mathcal{N} \times \mathcal{N})]$$

This specification obtains as a consequence of the substitution construction used recursively. Accordingly, the generating function of 2-3 trees is characterized by the functional equation

$$U(z) = z + U(z^2 + z^3).$$

From this functional equation, Odlyzko showed that $U_n$ grows like $\phi^n/n$ where $\phi = (1+\sqrt{5})/2$ is the golden ratio, the precise asymptotics involving intricate periodic fluctuations that are captured by complex asymptotic methods (singularity analysis and iteration theory).

Exercise 27. The equation for 2–3 trees alternatively results from the relations

$$\mathcal{U} = \sum_{h=0}^{\infty} \mathcal{U}^{[h]} \quad \text{with} \quad \mathcal{U}^{[0]} = \mathcal{N}, \quad \mathcal{U}^{[h+1]} = \mathcal{U}^{[h]}[(\mathcal{N} \times \mathcal{N}) + (\mathcal{N} \times \mathcal{N} \times \mathcal{N})],$$

where $\mathcal{U}^{[h]}$ is the class of trees of height $h$.

We can summarize the results of this section by a statement.

**Theorem 1.5** *The constructions of pointing and substitution are admissible:*

$$\begin{aligned} \mathcal{A} = \Theta\mathcal{B} \quad &\Longrightarrow \quad A(z) = z\tfrac{d}{dz}B(z) \\ \mathcal{A} = \mathcal{B}[\mathcal{C}] \quad &\Longrightarrow \quad A(z) = B(C(z)). \end{aligned}$$

## 1.9 Notes

There are several lessons to be lernt from the uses that we have surveyed of the symbolic method.

First, for a given class of problems, one is often lead to a unified treatment that reveals a natural class of functions in which generating functions lie. Thus denumerants with a finite set of coin denominations always lead to rational generating functions with poles on the unit circle. Such an observation is useful since then a common strategy for coefficient extraction can be applied, in such a case, based on partial fraction expansion.

In the same style, the run statistics constitue a particular case of the general theorem of Chomsky and Schützenberger: *the generating function of a regular language (definable by regular expressions or finite automata) is always a rational function.* Theorems of this sort establish a bridge between combinatorial specifications and special functions.

The example of counting set partitions shows that application of the symbolic method may require finding an adequate presentation of the combinatorial structures to be counted (there under the equivalent form of a regular language).

Second, our initial examples were classes of combinatorial structures with *explicit* "iterative" definitions, a fact leading in turn to explicit generating function expressions. The tree examples have introduced *recursively defined* structures. In that case, the recursive definition translates into a *functional equation* that only determines the generating function implicitly. In simpler situations (like binary or general trees), the equation can be solved and explicit counting results still follow. In other cases (like non-planar trees) one can often proceed with complex asymptotic analysis directly from the functional equation obtained, as we shall see later.

**Bibliography** Modern presentations of combinatorial analysis appear in the books of Comtet [4] (a superb book largely example driven), Stanley [15] (algebra and order structures) and Wilf [17] (generating functions oriented). An encyclopedic reference is the book of Jackson & Goulden [8] whose descriptive approach is very much parallel to ours. The framework presented here was used already in some surveys [5, 16] that deal more specifically with the analysis of algorithms.

The sources of the modern approaches to combinatorial analysis are hard to trace since these approaches are usually based on earlier traditions and informally stated mechanisms that were well mastered by practicing combinatorial analysts. One source in recent times is the Chomsky–Schützenberger theory of formal languages and enumerations [3]. Rota [13] and Stanley [14] developed an approach which is based on partially ordered sets. Bender and Goldman developed a theory of "prefabs" [2] whose purposes are similar to the theory developed here. Joyal [10] proposed an extensive framework based on category theory that addresses foundational issues in combinatorial enumerations.

The symbolic ideas exposed here are applied to the analysis of algorithms in surveys [5, 16]. Flajolet, Salvy, and Zimmermann [6] have shown how to use them in order to automate the analysis of some well characterized classes of algorithms and data structures by means of computer algebra systems.

## Problems and Exercises

Integer partitions have a rich combinatorial structure. Their OGF's are associated with special types of identities, some connected with elliptic functions, others with basic hypergeometric functions and so–called "$q$–analogues", see [1, 4]. The oldest identities go back to Euler.

EXERCISE 28. A sequence $\gamma_1, ..., \gamma_k$ is said to be superincreasing if for all $j$: $\gamma_1 + \gamma_2 + ... + \gamma_{j-1} < \gamma_j$. Show that, if $\mathcal{T}$ is superincreasing, then the number $Q_n = D_n[\mathcal{C}]$ of partitions with summands in $\mathcal{T}$ satisfies (for $n$ large enough) a recurrence of the type:

$$Q_n = \sum_{i=1}^{\ell} \epsilon_i D_{n-i}$$

where $\epsilon_i = 0, -1, +1$. Determine $\ell$ and give a simple characterisation of the $\epsilon_i$.

EXERCISE 29. The generating function of partitions, all of whose summands are odd is

$$A(z) = \prod_{n=0}^{\infty} \frac{1}{1 - z^{2n+1}}.$$

The generating function of partitions into distinct summands is

$$B(z) = \prod_{n=1}^{\infty} (1 + z^n).$$

Show that $A(z) \equiv B(z)$ and find a combinatorial bijection.

EXERCISE 30. [Euler] Define

$$e(u, z) = \prod_{m=1}^{\infty} (1 + uz^m).$$

Show that $e(u, z)$ can be expanded with respect to $u$ and use identification of coefficients in the relation $e(u, z) = (1 + u)e(uz, z)$ in order to determine the coefficients $a_k(z)$ in the expansion

$$e(z, u) = \sum_{k=0}^{\infty} a_k(z)u^k.$$

Give an interpretation of the coefficients of $e(z,u)$, and obtain a combinatorial proof of the formula

$$e(u,z) = 1 + \sum_{k=1}^{\infty} \frac{z^{k(k-1)/2} u^k}{(1-z)(1-z^2)\cdots(1-z^k)}.$$

Consider similarly

$$E(u,z) = \prod_{m=1}^{\infty} \frac{1}{(1-uz^m)} = \sum_{k=0}^{\infty} A_k(z) u^k.$$

Deduce the identity

$$1+\sum_{n=1}^{\infty} \frac{z^n}{(1-z)(1-z^2)\cdots(1-z^n)} = \left(1 + \sum_{n=1}^{\infty} (-1)^n \frac{z^{n(n-1)/2}}{(1-z)(1-z^2)\cdots(1-z^n)}\right)^{-1}.$$

EXERCISE 31. Consider partitions into summands that are distinct and powers of 2; get the OGF $E(z) = \prod_{j=0}^{\infty}(1+z^{2^j})$. However, since each integer has a unique binary representation, we have $E_n = 1$, hence derive (one of) Euler's classical identities:

$$\frac{1}{1-z} = \prod_{j=0}^{\infty}(1+z^{2^j}).$$

Objects apparently "non–combinatorial" can sometimes be enumerated by symbolic methods. Here is an indirect decomposition for irreducible ("prime") polynomials over finite fields.

EXERCISE 32. [Gauss] A polynomial is monic if its leading coefficient equals 1. Fix a prime number $p$ and consider integer coefficient polynomials whose coefficients are reduced modulo $p$ (i.e., polynomials over the Galois field $GF(p)$). Let $\mathcal{I}$ be the collection of all monic polynomials that are irreducible (they do not factor!). Show that

$$\log \frac{1}{1-pz} = I(z) + \frac{1}{2} I(z^2) + \frac{1}{3} I(z^3) + \cdots.$$

Find an explicit sum involving the Moebius function $\mu(n)$ for $I_n$. Deduce that $I_n \sim p^n/n$.

EXERCISE 33. Find a meaning for

$$[z^n]\prod_{j=0}^{\infty}(1-z^{2^j}), \qquad [z^n]\prod_{j=0}^{\infty}\frac{1}{1-z^{2^j}}.$$

Let $\nu(n)$ be the number of ones in the binary representation of $n$. Express $\sum_{n=0}^{\infty}\nu(n)z^n$.

Patterns in strings lead to rational generating functions with some interesting probabilistic and combinatorial implications.

EXERCISE 34. Let $L \subset \{a,b\}^*$ be a regular language and $S = \{a,b\}^{\infty}$ be the set of infinite strings with the product probability induced by $\Pr(a) = \Pr(b) = \frac{1}{2}$. Show that the probability that a random string $\omega \in S$ starts with a word of $L$ is a rational number (i.e. of the form $\frac{p}{q}$ with $p, q$ integers).

[Hint: Show that the sought probability is $l^*(1/2)$, where $l^*(z)$ is the OGF of a regular language closely related to $L$, the "prefix language" of $L$.]

Show that the pattern $\pi_1 = ab$ tends to occur "sooner" than $\pi_2 = aa$.

Show that a random $\omega \in S$ contains the block $\beta = abaab$ with probability 1. What is the expectation of the smallest $j$ such that $\beta$ occurs in $\omega$ at starting position $j$. Compare with the result for $\beta' = aabbb$.

The cycle construction is not treated here. However, it has some interesting connections with the combinatorics of strings.

EXERCISE 35. Use basic combinatorial principles and elementary properties of arithmetic functions in order to prove the cycle construction result of Theorems 1.1 and 1.4. [See Flajolet & Soria [7].]

Generating functions are sometimes only formal power series with radius of convergence equal to 0. They can nonetheless be useful for solving counting problems and deriving asymptotic estimates.

EXERCISE 36. Let $P_n = n!$ and $P(z)$ be the corresponding OGF. Thus $P(z)$ is the OGF of permutations. Show that the (purely divergent!) series $P(z)$ satisfies the differential equation

$$1 + z\frac{d}{dz}(zP(z)) = P(z),$$

and interpret it combinatorially. [Hint: One can go from the permutations of $[1..n]$ to those of $[1..n+1]$ by pointing at all places where $n+1$ can be inserted.

EXERCISE 37. With $P(z) = \sum_{n=0}^{\infty} n!z^n$, we have

$$\int_0^{+\infty} \frac{e^{-t}}{1+zt}\,dt \sim P(-z) \qquad (\text{as } z \to 0^+).$$

Thus, the OGF of permutations appears as the (formal, divergent) asymptotic series of an integral.

EXERCISE 38. A permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n$ of $[1..n]$ is said to be decomposable if, for some $k < n$, $\sigma_1\sigma_2\cdots\sigma_k$ is a permutation of $[1..k]$. Let $I(z)$ be the OGF of indecomposable permutations. Show that

$$I(z) = 1 - \frac{1}{1+Q(z)},$$

where $Q(z) = \sum_{n\geq 1} n!z^n$.

Deduce that $I_n \sim n!$, so that almost all permutations are indecomposable. [See [4, p. 262]]

Combinatorial analysis is concerned with the interactions between certain formal operations of analysis and combinatorics. A number of classical identities can be interpreted and even proved using combinatorics.

EXERCISE 39. Explain the combinatorial "meaning" of a derivation, $g(z) = \mathrm{D}f(z)$ with $\mathrm{D} = d/dz$, by assuming that $f(z)$ is the OGF of a class $\mathcal{F}$.

Give a combinatorial interpretation of the chain rule,

$$\mathrm{D}[f(g(z))] = \mathrm{D}[f](g(z)) \cdot \mathrm{D}g(z).$$

Interpret combinatorially $\mathrm{D}^j$. Deduce a combinatorial proof of Taylor's formula:

$$f(z+h) = \sum_{j=0}^{\infty} \frac{1}{j!}\mathrm{D}^j f(z) \cdot h^j.$$

Describe the form of $\mathrm{D}^j[f(g(z))]$. (Faà di Bruno's formula).

## Bibliography

[1] ANDREWS, G. E. *The Theory of Partitions*, vol. 2 of *Encyclopedia of Mathematics and its Applications*. Addison–Wesley, 1976.

[2] BENDER, E. A., AND GOLDMAN, J. R. Enumerative uses of generating functions. *Indiana University Mathematical Journal* (1971), 753–765.

[3] CHOMSKY, N., AND SCHÜTZENBERGER, M. P. The algebraic theory of context–free languages. In *Computer Programing and Formal Languages* (1963), P. Braffort and D. Hirschberg, Eds., North Holland, pp. 118–161.

[4] COMTET, L. *Advanced Combinatorics*. Reidel, Dordrecht, 1974.

[5] FLAJOLET, P. Mathematical methods in the analysis of algorithms and data structures. In *Trends in Theoretical Computer Science*, E. Börger, Ed. Computer Science Press, Rockville, Maryland, 1988, ch. 6, pp. 225–304. (Lecture Notes for *A Graduate Course in Computation Theory*, Udine, 1984).

[6] FLAJOLET, P., SALVY, B., AND ZIMMERMANN, P. Automatic average–case analysis of algorithms. *Theoretical Computer Science, Series A 79*, 1 (Feb. 1991), 37–109.

[7] FLAJOLET, P., AND SORIA, M. The cycle construction. *SIAM Journal on Discrete Mathematics 4*, 1 (Feb. 1991), 58–60.

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

[9] HARARY, F., AND PALMER, E. M. *Graphical Enumeration*. Academic Press, 1973.

[10] JOYAL, A. Une théorie combinatoire des séries formelles. *Advances in Mathematics 42*, 1 (1981), 1–82.

[11] KNUTH, D. E. *The Art of Computer Programming*, vol. 3: Sorting and Searching. Addison-Wesley, 1973.

[12] PÓLYA, G., AND READ, R. C. *Combinatorial Enumeration of Groups, Graphs and Chemical Compounds*. Springer Verlag, New York, 1987.

[13] ROTA, G.-C. *Finite Operator Calculus*. Academic Press, 1975.

[14] STANLEY, R. P. Generating functions. In *Studies in Combinatorics,* M.A.A. Studies in Mathematics, Vol. 17. (1978), G.-C. Rota, Ed., The Mathematical Association of America, pp. 100–141.

[15] STANLEY, R. P. *Enumerative Combinatorics*, vol. I. Wadsworth & Brooks/Cole, 1986.

[16] VITTER, J. S., AND FLAJOLET, P. Analysis of algorithms and data structures. In *Handbook of Theoretical Computer Science*, J. van Leeuwen, Ed., vol. A: Algorithms and Complexity. North Holland, 1990, ch. 9, pp. 431–524.

[17] WILF, H. S. *Generatingfunctionology*. Academic Press, 1990.

# Chapter 2

# Labelled Structures and Exponential Generating Functions

*Cette approche évacue pratiquement tous les calculs.*
— Dominique Foata &
Marcel P. Schützenberger [10]

Many objects of classical combinatorics present themselves naturally as labelled structures where each "atom" of a structure (typically nodes in a graph or a tree) bears a distinctive integer label. For instance the cycle decomposition of a permutation represents the permutation as an unordered collection of cyclic graphs whose nodes are labelled by integers.

Commonly encountered classes of labelled structures are permutations, set partitions, labelled graphs and labelled trees, functional graphs that are associated to mappings of a finite set into itself, as well as structures related to occupancy problems and hash tables.

Operations on labelled structures are based on a special product, the labelled (or partitional) product that distributes labels between components. This operation is a natural analogue of the cartesian product for plain unlabelled structures. The labelled product in turn leads to labelled analogues of the sequence, set, and cycle constructions.

The labelled constructions all translate over exponential generating functions. The translation schemes are analytically simpler than in the unlabelled case considered in the previous chapter.

Labelled constructions allows to take into account constructions that are in many ways richer combinatorially, in particular regarding order properties of combinatorial structures. They thus constitute a facet with powerful descriptive powers of the symbolic method for combinatorial enumerations.

## 2.1 Introduction

In this chapter, we consider classes of combinatorial structures in the sense previously introduced. The structures to be considered here are in addition labelled structures where, roughly speaking, a structure of size $n$ bears $n$ distinct labels of the interval $[1\,..\,n]$ at designated places (typically nodes). The counting of labelled structures is normally achieved by means of exponential generating functions.

The *exponential generating function* (EGF) of a sequence $\{a_n\}$ is, we recall,

$$a(z) = \sum_{n=0}^{\infty} a_n \frac{z^n}{n!}. \tag{2.1}$$

The coefficient $a_n$ in such an exponential generating function is then recovered by

$$a_n = n! \cdot [z^n]\, a(z).$$

Like in the previous chapter, we adhere to a systematic naming convention for generating functions of combinatorial structures. A class $\mathcal{A}$, its counting sequence $\{A_n\}$ (or $a_n$) and its exponential generating function $A(z)$ (or $a(z)$) will all be denoted by the same group of letters[1]. The EGF of a class $\mathcal{A}$ admits the "combinatorial form"

$$A(z) = \sum_{\alpha\in\mathcal{A}} \frac{z^{|\alpha|}}{|\alpha|!}.$$

If $a(z), b(z), c(z)$ are EGF's, with $a(z) = \sum_n a_n z^n/n!$ and so on, we have

$$a(z) = b(z)\cdot c(z) \qquad \Longrightarrow \qquad a_n = \sum_{k=0}^{n} \binom{n}{k} b_k c_{n-k},$$

so that the product of EGF's leads to a binomial convolution formula for coefficients. In the same vein,

$$a(z) = a^{(1)}(z)\, a^{(2)}(z) \cdots a^{(r)}(z) \quad \Longrightarrow$$

[1] In the few contexts where we need to discuss simultaneously ordinary and exponential generating functions, we use $\widehat{a}(z)$ to distinguish the EGF from the OGF.

$$a_n = \sum_{n_1+n_2+\cdots+n_r=n} \binom{n}{n_1, n_2, \ldots, n_r} a_n^{(1)} a_n^{(2)} \cdots a_n^{(r)}. \tag{2.2}$$

In Eq. (2.2) there occurs the multinomial coefficient

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}.$$

This multinomial coefficient also counts the number of ways of splitting $n$ elements into $r$ classes, a fact which is at the origin of many applications of binomial convolutions and EGF's.

In the sequel, we start the discussion with two groups of examples, surjections and set partitions first, alignments and permutations next. There, the convolution relations (2.2) are employed to attain enumeration results, and at the same time smoothly introduce the techniques needed for the more abstract treatment of labelled products in later sections.

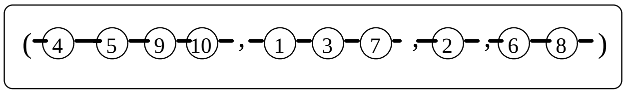
## 2.2 Surjections and set partitions.

Fix some integer $r \geq 1$. Let $\mathcal{R}_n^{(r)}$ denote the class of all surjections (or *onto mappings*) from the set $[1\,..\,n]$ *onto* $[1\,..\,r]$; let $\mathcal{S}_n^{(r)}$ denote the number of ways of partitioning the set $[1\,..\,n]$ into $r$ (disjoint, nonempty) equivalence classes. In accordance with previous notations, we set $\mathcal{R}^{(r)} = \bigcup_n \mathcal{R}_n^{(r)}$ and $\mathcal{S}^{(r)} = \bigcup_n \mathcal{S}_n^{(r)}$; the corresponding structures are called *surjections* and *set partitions* (the latter not to be confused with integer partitions that we considered in the previous chapter).

The number $R_n^{(r)}$ of $r$–surjections is computable as follows: an $r$–surjection $\phi \in \mathcal{R}_n^{(r)}$ is determined by the ordered $r$–tuple formed with the preimages, $(\phi^{-1}(1), \phi^{-1}(2), \ldots, \phi^{-1}(r))$. Each preimage is a nonempty set. We thus have

$$R_n^{(r)} = \sum_{(n_1, n_2, \ldots, n_r)} \binom{n}{n_1, n_2, \ldots, n_r}, \tag{2.3}$$

the sum being taken over $n_j \geq 1$, $n_1 + n_2 + \cdots + n_r = n$. In this formula the indices $n_j$ vary over all allowable cardinalities of preimages, and the multinomial coefficient counts the number of ways of distributing the elements of $[1\,..\,n]$ amongst the $r$ preimages.



A surjection, here the mapping from $[1\,..\,10]$ onto $[1\,..\,4]$ given by the table

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 3 & 2 & 1 & 1 & 4 & 2 & 4 & 1 & 1 \end{pmatrix},$$

may be viewed as an ordered tuple of linear sorted graphs by looking at the collection of preimages of $1, 2, 3$, etc.

Figure 2.1: The decomposition of surjections.

A simple way to compute the EGF of the $R_n^{(r)}$ is to introduce the numbers $V_n$ by

$$V_0 = 0, \qquad V_n = 1 \ \text{ if } n \geq 1.$$

The formula (2.3) then takes the simpler form

$$R_n^{(r)} \equiv \sum_{n_1, n_2, \ldots, n_r} \binom{n}{n_1, n_2, \ldots, n_r} V_{n_1} V_{n_2} \cdots V_{n_r}, \tag{2.4}$$

where the summation now extends to *all* tuples $(n_1, n_2, \ldots, n_r)$. The EGF of the $V_n$ is

$$V(z) = \sum_{n=0}^{\infty} V_n \frac{z^n}{n!} = e^z - 1. \tag{2.5}$$

Thus the convolution relations (2.3) and (2.4) lead to

$$R^{(r)}(z) = (e^z - 1)^r. \tag{2.6}$$

Equation (2.6) does solve the counting problem for surjections. For small $r$, we have

$$R^{(2)}(z) = e^{2z} - 2e^z + 1, \quad R^{(3)}(z) = e^{3z} - 3e^{2z} + 3e^z - 1,$$

whence, by expanding,

$$R_n^{(2)} = 2^n - 2, \quad R_n^{(3)} = 3^n - 3 \cdot 2^n + 3\,.$$

The general formula follows similarly from expanding the $r$th power in (2.6) by the binomial theorem, and then extracting coefficients:

$$\begin{aligned} R_n^{(r)} &= n!\,[z^n] \sum_{j=0}^{r} \binom{r}{j} (-1)^j e^{r-j} z \\ &= \sum_{j=0}^{r} \binom{r}{j} (-1)^j (r-j)^n. \end{aligned} \tag{2.7}$$

The enumeration problem for set partitions is closely related to that of surjections, the basic formula connecting the two counting sequences being

$$S_n^{(r)} = \frac{1}{r!} R_n^{(r)}. \tag{2.8}$$

The rationale for (2.8) is that an $r$–partition is associated with a group of exactly $r!$ distinct $r$–surjections, two surjections belonging to the same group iff one obtains from the other by permuting the range values, $[1\,..\,r]$. (Incidentally, the discussion conducted here resembles that of distinguishable versus undistinguishable boxes in classical occupancy problems of probability theory.)

The numbers $S_n^{(r)} = \frac{1}{r!} R_n^{(r)}$ are known as the Stirling numbers of the second kind, or better, the Stirling "partition" numbers. They were briefly encountered in the previous chapter. Knuth advocated for the $S_n^{(r)}$ the notation $\left\{ n \atop r \right\}$. From (2.7), an explicit form also exists:

$$\left\{ n \atop r \right\} = \frac{1}{r!} \sum_{j=0}^{r} \binom{r}{j} (-1)^j (r-j)^n.$$

The books by Graham, Knuth, and Patashnik [12] and Comtet [3] contain a thorough discussion of these numbers.

Returning to generating functions, we have obtained

$$R^{(r)}(z) = (V(z))^r, \quad S^{(r)}(z) = \frac{1}{r!}(V(z))^r \quad \text{where } V(z) = (e^z - 1). \tag{2.9}$$

Define now the collection of all surjections and all set partitions by

$$\mathcal{R} = \bigcup_r \mathcal{R}^{(r)}, \qquad \mathcal{S} = \bigcup_r \mathcal{S}^{(r)}.$$

Thus $\mathcal{R}_n$ is the class of all surjections of $[1\,..\,n]$ onto *any* initial segment of the integers, and $\mathcal{S}_n$ is the class of all partitions of the set $[1\,..\,n]$ into *any* number of parts. Then, from (2.9), a summation over values of $r$ provides

$$R(z) = \frac{1}{1 - V(z)} = \frac{1}{2 - e^z}, \qquad S(z) = e^{V(z)} = e^{e^z - 1}. \tag{2.10}$$

The numbers $R_n = n!\,[z^n]R(z)$ and $S_n = n![z^n]S(z)$ are called the *surjection numbers* and the *Bell numbers* respectively.

Explicit expressions as finite double sums result from summing the $R_n^{(r)}$ of Eq. (2.7). Alternative single (though infinite) sums result from the expansions

$$\left\{ \begin{array}{rcl} R(z) & = & \dfrac{1}{2}\dfrac{1}{1 - \frac{1}{2}e^z} \\ & = & \displaystyle\sum_{k=0}^{\infty} \frac{1}{2^{k+1}} e^{kz} \end{array} \right. \qquad \text{and} \qquad \left\{ \begin{array}{rcl} S(z) & = & e^{e^z - 1} = \dfrac{1}{e} e^{e^z} \\ & = & \displaystyle\frac{1}{e}\sum_{k=0}^{\infty} e^{kz}, \end{array} \right.$$

with which coefficient extraction yields

$$R_n \;=\; \frac{1}{2}\sum_{k=0}^{\infty} \frac{k^n}{2^k} \qquad \text{and} \qquad S_n \;=\; \frac{1}{e}\sum_{k=0}^{\infty} \frac{k^n}{k!}.$$

The formula for the Bell numbers was found by Dobinski in 1877.

The asymptotic analysis of the surjection numbers $(R_n)$ will be done in a later chapter by means of singularity analysis of the meromorphic function $R(z)$; that of Bell's partition numbers $(S_n)$ is best done by means of the saddle point method. The asymptotic forms found are

$$R_n \sim \frac{n!}{2}\frac{1}{(\log 2)^{n+1}} \qquad \text{and} \qquad S_n \sim n!\,\frac{e^{e^r - 1}}{r^{n+1}\sqrt{2\pi \exp(r)}},$$

where $r \equiv r(n)$ is the positive root of $re^r = n$ (one has $r \approx \log n - \log\log n$). Elementary derivations of these asymptotic forms are explored in the exercises.

EXERCISE 1. Find the index $k_0$ near which the terms in Dobinski's formula are maximal. Approximate the general term of index $k = k_0 \pm h$. By comparing the sum of the approximations to the Riemann sum of the Gaussian integral, derive the asymptotic form stated for $S_n$.

[Note: This is an instance of the Laplace method for sums. See De Bruijn's book [5] for details.]

EXERCISE 2. [Comtet] Treat similarly the asymptotics of $R_n$ by the Laplace method for sums.

EXERCISE 3. The function

$$R(z) - \frac{1}{2}\frac{1}{\log 2 - z}$$

is analytic for $|z| \leq 6$. hence, by Cauchy's bounds for analytic functions,

$$R_n = \frac{n!}{2}\left(\frac{1}{(\log 2)^{n+1}} + O(\frac{1}{6^n})\right).$$

Relations (2.10) obey the scheme

$$R(z) = \frac{1}{1 - V(z)}, \qquad S(z) = e^{V(z)}, \tag{2.11}$$

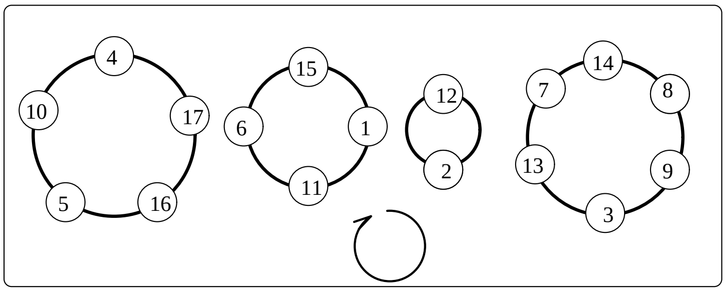
which can be interpreted in a graph–theoretic way as follows. Define a linear sorted graph as a line graph that is directed and has at least $n \geq 1$ nodes labelled by distinct integers that go in increasing order along directed edges. Let $\mathcal{V}$ be the subclass of the linear sorted graphs whose integer labels form an initial segment ($[1\,..\,n]$ for a graph of size $n$) of the integers. The EGF of $\mathcal{V}$ is $V(z) = e^z - 1$.

One can then view a surjection in $\mathcal{R}_n$ as a *sequence* of linear sorted graphs with its labels of $[1\,..\,n]$ distributed between components, and a partition in $\mathcal{S}$ as a *set* of linear sorted graphs again with a distribution of labels. Relations between EGF's of the form (2.10) are characteristic of such labelled sequence and labelled set constructions. They constitute a major paradigm in combinatorial enumerations.

## 2.3 Alignments and permutations.

As a further illustration of labelled constructions, consider the problem of enumerating the class $\mathcal{P}_n^{(r)}$ of permutations of $[1\,..\,n]$ whose cycle decomposition involves $r$ cycles.

First of all, let $\mathcal{K}_n$ represent the subclass of circular permutations of $[1\,..\,n]$. The corresponding counts satisfy $K_n = (n-1)!$. To see it, organize the $n!$ permutations into $(n-1)!$ groups each of cardinality $n$, two permutations being in the same group if one can be deduced from the other modulo




A permutation may be viewed as a *set* of cycles that are labelled cyclic digraphs. The diagram shows the decomposition of the permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 \\ 11 & 12 & 13 & 17 & 10 & 15 & 14 & 9 & 3 & 4 & 6 & 2 & 7 & 8 & 1 & 5 & 16 \end{pmatrix}.$$

(Cycles read clockwise and $i$ is connected to $\sigma_i$ in the graph.)

Figure 2.2: The cycle decomposition of permutations.

a circular shift; each group corresponds to a unique circular permutation, so that $K_n = \frac{1}{n}\, n!$. The corresponding EGF results:

$$K(z) = \sum_n (n-1)! \frac{z^n}{n!} = \log \frac{1}{1-z}.$$

A permutation composed of $r$ cycles is equivalent to an unordered collection of $r$ labelled (directed) cycle graphs. For the present discussion, it proves convenient to introduce first a companion structure: an *alignment* is an ordered sequence of such cycle graphs, with $\mathcal{O}^{(r)}$ being the class of alignments comprising $r$ cycles. Thus, an alignment resembles a permutation safe that cycles are ordered between themselves. Clearly, we have —compare with Eq. (2.8):

$$P_n^{(r)} = \frac{1}{r!} O_n^{(r)}. \tag{2.12}$$

Next by expressing the fact that an alignment decomposes into an $r$–tuple of cycles together with a splitting of the labels $[1\,..\,n]$ into $r$ classes we get the recurrence:

$$O_n^{(r)} = \sum_{n_1, n_2, \ldots, n_r} \binom{n}{n_1, n_2, \ldots, n_r} K_{n_1} K_{n_2} \cdots K_{n_r}. \tag{2.13}$$

Thus going over to EGF's:

$$O^{(r)}(z) = (K(z))^r, \quad P^{(r)}(z) = \frac{1}{r!}(K(z))^r \qquad \text{where } K(z) = \left(\log \frac{1}{1-z}\right).$$

For the classes $\mathcal{O}$ and $\mathcal{P}$ of all alignments and all permutations, we next find by summation

$$O(z) = \frac{1}{1 - K(z)}, \qquad P(z) = e^{K(z)},$$

that is to say

$$O(z) = \frac{1}{1 - \log(1-z)^{-1}}, \quad P(z) = \exp(\log(1-z)^{-1}) = \frac{1}{1-z}. \tag{2.14}$$

The latter result for $P(z)$ corresponds to $P_n = n!$, as was of course to be expected.

In addition, we have obtained an explicit generating function for permutations having $r$ cycles, and

$$P_n^{(r)} = \frac{n!}{r!} [z^n] \left(\log \frac{1}{1-z}\right)^r.$$

These numbers are fundamental quantities of combinatorial analysis. They are known as the Stirling numbers of the first kind, or better, according to a proposal of Knuth, the *Stirling "cycle" numbers.* Together with the Stirling partition numbers, the properties of the Stirling cycle numbers are explored in the book by Graham, Knuth, and Patashnik [12] where they are denoted by $\left[{n \atop r}\right]$.

EXERCISE 4. The differential relation

$$(1-z)\frac{d}{dz} P^{(r)}(z) = P^{(r-1)}(z),$$

implies the recurrence

$$\begin{bmatrix} n \\ r \end{bmatrix} = (n-1) \begin{bmatrix} n-1 \\ r \end{bmatrix} + \begin{bmatrix} n-1 \\ r-1 \end{bmatrix}.$$

A complicated but explicit form for the Stirling cycle numbers was obtained by Schlömilch in 1852 [3, p. 216]:

$$\left[{n \atop k}\right] = \sum_{0\le j\le h\le n-k} (-1)^{j+h}\binom{h}{j}\binom{n-1+h}{n-k+h}\binom{2n-k}{n-k-h}\frac{(h-j)^{n-k+h}}{h!}.$$

This alternating double sum is not too useful, however. (A proof is suggested in the exercises.) A more important relation is that of the generating polynomials of the $\left[{n \atop r}\right]$ for fixed $n$,

$$P_n(u) \equiv \sum_{r=1}^{n} P_n^{(r)}u^r = u\cdot(u+1)\cdot(u+2)\cdots(u+n-1).$$

From there, it can be easily proved that the expected number of cycles in a random permutation of size $n$ is the harmonic number,

$$H_n = 1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n},$$

and is thus asymptotic to $\log n$.

Exercise 5. Introduce the bivariate generating function

$$P(u,z) := \sum_{r=0}^{\infty} P^{(r)}(z)u^r$$

and observe that

$$\begin{aligned} P(u,z) &= \sum_{r=0}^{\infty}\frac{u^r}{r!}\left(\log\frac{1}{1-z}\right)^r = \exp\left(u\log\frac{1}{1-z}\right) \\ &= (1-z)^{-u}. \end{aligned}$$

Deduce the form of the Stirling cycle polynomials from the equality

$$[z^n](1-z)^{-u} = (-1)^n\binom{-u}{n},$$

[Such bivariate generating function techniques will be examined systematically in a later chapter.]

## 2.4 Labelled constructions

A *labelled structure* is formed of "atoms", nodes in the case of graphs or trees, that are labelled in some way by distinct integers. The size of a structure is the number of its atoms, which is thus also the number of its labels. A labelled structure is said to be *well–labelled* (or canonically labelled) if its labels form an initial segment of $\mathbb{N}^*$. It is again convenient to introduce the empty (null) structure $\epsilon$ of size 0 that bears no label, and to consider it as a special case of a labelled structure.

A labelled structure may be relabelled. *We only consider relabellings that preserve the order relations between labels.*

- Reduction: For a non–canonically 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 note $\rho_0(\alpha)$ the reduction of the structure $\alpha$.
- Expansion: This operation is defined by means of a relabelling function $e \in [1\,..\,n] \mapsto \mathbb{N}^*$ assumed to be strictly increasing. We note $e(\alpha)$ the result of relabelling $\alpha$ by $e$.

We next define a product called the *labelled product*, or simply *product* (originally this was named *partitional product* by Foata who proposed an early formalization in [9])).

Given two labelled structures $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$, the product $\alpha \star \beta$ is a finite collection of structures that are ordered pairs $(\alpha', \beta')$ of relabelled copies of $(\alpha, \beta)$,

$$\alpha\star\beta = \{\, (\alpha',\beta') \ \mid\ \rho_0(\alpha') = \alpha,\ \rho_0(\beta') = \beta,\ (\alpha',\beta')\ \text{is well–labelled}\,\}, \quad (2.15)$$

the relabellings preserving the order structure present in $\alpha$ and $\beta$. An equivalent form is via expansion of labels:

$$\alpha \star \beta = \{\, (e(\alpha), f(\beta) \ \mid\ \mathrm{Im}(e) \cap \mathrm{Im}(f) = \emptyset,\ \mathrm{Im}(e) \cup \mathrm{Im}(f) = [1\,..\,|\alpha| + |\beta|]\,\}, \quad (2.16)$$

where $e, f$ are again assumed to be *increasing* with ranges $\mathrm{Im}(e), \mathrm{Im}(f)$.

If $\mathcal{A}$ and $\mathcal{B}$ are two classes of combinatorial structures, the partitional product $\mathcal{C} = \mathcal{A} \star \mathcal{B}$ is defined by the usual extension of operations to sets:

$$\mathcal{A} \star \mathcal{B} = \bigcup_{\substack{\alpha\in\mathcal{A}\\ \beta\in\mathcal{B}}} (\alpha \star \beta). \quad (2.17)$$

**1.** The main constructions of union, and product, sequence, set, and cycle for labelled structures together with their translation into exponential generating functions.

| Construction | | EGF |
|---|---|---|
| Union | $\mathcal{A} = \mathcal{B} + \mathcal{C}$ | $A(z) = B(z) + C(z)$ |
| Product | $\mathcal{A} = \mathcal{B} \star \mathcal{C}$ | $A(z) = B(z) \cdot C(z)$ |
| Sequence | $\mathcal{A} = \mathfrak{S}\{\mathcal{B}\}$ | $A(z) = \dfrac{1}{1 - B(z)}$ |
| Set | $\mathcal{A} = \mathfrak{P}\{\mathcal{B}\}$ | $A(z) = \exp(B(z))$ |
| Cycle | $\mathcal{A} = \mathfrak{C}\{\mathcal{B}\}$ | $A(z) = \log \dfrac{1}{1 - B(z)}$ |

**2.** The translation for sets, multisets, and cycles of fixed cardinality.

| Construction | | EGF |
|---|---|---|
| Sequence | $\mathcal{A} = \mathfrak{S}\{\mathcal{B}, \text{card} = k\}$ | $A(z) = (A(z))^k$ |
| Set | $\mathcal{A} = \mathfrak{P}\{\mathcal{B}, \text{card} = k\}$ | $A(z) = \dfrac{1}{k!}(A(z))^k$ |
| Cycle | $\mathcal{A} = \mathfrak{C}\{\mathcal{B}, \text{card} = k\}$ | $A(z) = \dfrac{1}{k}(A(z))^k$ |

**3.** The additional constructions of pointing and substitution.

| Construction | | EGF |
|---|---|---|
| Pointing | $\mathcal{A} = \Theta\mathcal{B}$ | $A(z) = z\frac{d}{dz}B(z)$ |
| Substitution | $\mathcal{A} = \mathcal{B} \circ \mathcal{C}$ | $A(z) = B(C(z))$ |

**4.** The "boxed" product.

$$\mathcal{A} = (\mathcal{B}^{\square} \star \mathcal{C}) \quad \Longrightarrow \quad A(z) = \int_0^z \left(\frac{d}{dt}B(t)\right) \cdot C(t)\,dt.$$

Figure 2.3: The various constructions over *labelled* structures considered in this chapter together with their translation into *exponential* generating functions (EGF's). The first constructions are counterparts of the unlabelled constructions of the previous chapter (the multiset construction is not meaningful here). The translation for composite constructions of bounded cardinality is simpler. Finally, the boxed product, is specific to labelled structures.

**Definition 2.1** *The product of $\mathcal{A}$ and $\mathcal{B}$ is obtained by forming ordered pairs from $\mathcal{A} \times \mathcal{B}$ and performing all possible order consistent relabellings, ensuring that the resulting pairs are well–labelled, as described by (2.15), (2.16), (2.17).*

From the same line of reasoning as in the previous section, we find for corresponding EGF's the relation,

$$A_n = \sum_{n_1+n_2=n} \binom{n}{n_1, n_2} B_{n_1} C_{n_2} \qquad \text{hence} \qquad a(z) = b(z) \cdot c(z).$$

There the binomial arises since the cardinality of $(\alpha \star \beta)$ is $\binom{|\alpha|+|\beta|}{|\alpha|,|\beta|}$, and the product formula results from tracing all possibilities for relabellings and for the $\mathcal{A}$ and $\mathcal{B}$ components. Thus, the labelled product simply corresponds to the product operation on exponential generating functions.

The $k$–th (labelled) *power* of $\mathcal{A}$ is defined as $(\mathcal{A} \star \mathcal{A} \cdots \mathcal{A})$, with $k$ factors equal to $\mathcal{A}$. It is denoted $\mathfrak{S}\{\mathcal{A}, \text{card} = k\}$.

This corresponds to forming $k$–sequences and performing all consistent relabellings. The (labelled) *sequence* class of $\mathcal{A}$ is denoted by $\mathfrak{S}\{\mathcal{A}\}$ and is defined by

$$\mathfrak{S}\{\mathcal{A}\} \stackrel{\text{def}}{=} \{\epsilon\} + \mathcal{A} + (\mathcal{A} * \mathcal{A}) + (\mathcal{A} * \mathcal{A} * \mathcal{A}) + \cdots = \bigcup_{k \geq 0} \mathfrak{S}\{\mathcal{A}, \text{card} = k\}.$$

The product relation for EGF's clearly extends to arbitrary products, so that

$$\mathcal{B} = \mathfrak{S}\{\mathcal{A}, \text{card} = k\} \quad \Longrightarrow \quad B(z) = (A(z))^k,$$

and

$$\mathcal{C} = \mathfrak{S}\{\mathcal{A}\} \quad \Longrightarrow \quad C(z) = \sum_{k=0}^{\infty} (A(z))^k = \frac{1}{1 - A(z)}.$$

We denote by $\mathfrak{P}\{\mathcal{A}, \text{card} = k\}$ the class of $k$–sets formed from $\mathcal{A}$. This can defined formally like in the unlabelled case as the quotient of the sequence class $\mathfrak{S}\{\mathcal{A}, \text{card} = k\}$ by an equivalence relation based on arbitrary permutations of the elements in the sequence: a "set" is a sequence regarded up to arbitrary permutation of its elements. The (labelled) *powerset* class of $\mathcal{A}$, denoted $\mathfrak{P}\{\mathcal{A}\}$, is defined by

$$\mathfrak{P}\{\mathcal{A}\} \stackrel{\text{def}}{=} \{\epsilon\} + \mathcal{A} + \mathfrak{P}\{\mathcal{A}, \text{card} = 2\} + \cdots = \bigcup_{k \geq 0} \mathfrak{P}\{\mathcal{A}, \text{card} = k\}.$$

A labelled $k$–set is associated with exactly $k!$ different sequences. (There is here is subtle difference with the the unlabelled case where formulæ are more complex as an unlabelled sequence may contain repeated elements while components of a labelled sequence are all distinguished by their labels.) Thus in terms of EGF's

$$\mathcal{B} = \mathfrak{P}\{\mathcal{A}, \mathrm{card} = k\} \implies B(z) = \frac{(A(z))^k}{k!},$$

and

$$\mathcal{C} = \mathfrak{P}\{\mathcal{A}\} \implies C(z) = \sum_{k=0}^{\infty} \frac{(A(z))^k}{k!} = \exp(A(z)).$$

We also define $k$–cycles, $\mathfrak{C}\{\mathcal{A}, \mathrm{card} = k\}$ and the cycle class, $\mathfrak{C}\{\mathcal{A}\}$ of a labelled class $\mathcal{A}$ in the obvious manner as sequences taken up to circular shifts of their elements. In terms of EGF's, we have

$$\mathcal{B} = \mathfrak{C}\{\mathcal{A}, \mathrm{card} = k\} \implies B(z) = \frac{(A(z))^k}{k},$$

and

$$\mathcal{C} = \mathfrak{C}\{\mathcal{A}\} \implies C(z) = \sum_{k=0}^{\infty} \frac{(A(z))^k}{k} = \log \frac{1}{1 - A(z)},$$

since each cycle admits exactly $k$ representations as a sequence.

We sometimes write $\mathcal{A}^k$ when no confusion arises with the (unlabelled) cartesian product. Symbolically, the class of $k$–sets can also be written in the more suggestive way as $\frac{1}{k!}\mathcal{A}^k$, and $k$–cycles can be represented as $\frac{1}{k}\mathcal{A}^k$.

**Theorem 2.1** *The constructions of labelled product, $k$–th power, and sequence class,*

$$\mathcal{A} = \mathcal{B} \star \mathcal{C}, \qquad \mathcal{A} = \mathfrak{S}\{\mathcal{B}, \mathrm{card} = k\}, \qquad \mathcal{A} = \mathfrak{S}\{\mathcal{B}\}$$

*are admissible:*

$$A(z) = B(z) \cdot C(z), \qquad A(z) = (B(z))^k, \qquad A(z) = \frac{1}{1 - B(z)}.$$

*The constructions of $k$–set and powerset class*

$$\mathcal{A} = \mathfrak{P}\{\mathcal{B}, \mathrm{card} = k\}, \quad \mathcal{A} = \mathfrak{P}\{\mathcal{B}\},$$

*are admissible*

$$A(z) = \frac{1}{k!}\left(B(z)\right)^k, \qquad A(z) = \exp(B(z)).$$

*The constructions of $k$–cycle and cycle class,*

$$\mathcal{A} = \mathfrak{C}\{\mathcal{B}, \mathrm{card} = k\}, \quad \mathcal{A} = \mathfrak{C}\{\mathcal{B}\},$$

*are admissible*

$$A(z) = \frac{1}{k}\left(B(z)\right)^k, \qquad A(z) = \log\frac{1}{1 - B(z)}.$$

**Constructible classes.** Like in the previous chapter, we say that a class of labelled objects is constructible if it admits a specification in terms of sums (disjoint unions), the labelled constructions of product, sequence, set, cycle, and the initial classes defined by the empty structure of size 0 and the atomic node $\mathcal{N} = \{1\}$ of size 1.

Set partitions, surjections, permutations, and alignments are thus particular constructible classes. An immediate consequence of Theorem 2.1 is:

**Theorem 2.2** *The exponential generating function of a constructible class of labelled objects is a component of a system of generating function equations whose terms are built from* 1 *and* $z$ *using the operators*

$$+\,,\ \times\,,\ Q(f) = \frac{1}{1-f}\,,\ E(f) = e^f\,,\ L(f) = \log\frac{1}{1-f}.$$

If we further allow cardinality restrictions in composite constructions, the operators $f^k$, $\frac{1}{k!}f^k$ and $\frac{1}{k}f^k$ are to be added to the list of operators.

## 2.5 Labelled graphs.

This section presents a first immediate application of the admissibility theorem to the counting of connected graphs.

In the context of graphical enumerations, the labelled set construction takes the form of an enumerative formula relating a class of graphs $\mathcal{G}$ and the subclass of its connected graphs $\mathcal{K} \subset \mathcal{G}$:

$$\mathcal{G} = \mathfrak{P}\{\mathcal{K}\} \implies G(z) = e^{K(z)}.$$

This is well known in graph theory [15] as the *exponential formula*. It has been encountered already when dealing with the cycle decomposition of permutations and it will surface again in the analysis of functional graphs.

Consider the class $\mathcal{G}$ of all (undirected) labelled graphs, the size of a graph being the number of its nodes. Since a graph is determined by the choice of its set of edges; there are $\binom{n}{2}$ potential edges each of which may be taken in or out, so that

$$G_n = 2^{\binom{n}{2}}.$$

Let $\mathcal{K} \subset \mathcal{G}$ be the subclass of all connected graphs. The exponential formula determines $K(z)$ implicitly,

$$K(z) = \log\left(1 + \sum_{n\geq 1} 2^{\binom{n}{2}} \frac{z^n}{n!}\right),$$

from which a complicated convolution expression follows for $K_n$ when one expands $\log(1+u)$:

$$\begin{aligned} K_n &= 2^{\binom{n}{2}} - \frac{1}{2} \sum_{\substack{n_1+n_2=n \\ n_1,n_2\geq 1}} \binom{n}{n_1, n_2} 2^{\binom{n_1}{2}+\binom{n_2}{2}} \\ &\quad + \frac{1}{3} \sum_{\substack{n_1+n_2+n_3=n \\ n_1,n_2,n_3\geq 1}} \binom{n}{n_1, n_2, n_3} 2^{\binom{n_1}{2}+\binom{n_2}{2}+\binom{n_3}{2}} - \cdots . \end{aligned}$$

With the very fast increase of $G_n$ with $n$, for instance

$$2^{\binom{n+1}{2}} = 2^n \, 2^{\binom{n}{2}},$$

a detailed analysis of the various terms of the expression of $K_n$ shows predominance of the first term, and

$$K_n = 2^{\binom{n}{2}}(1 - O(2^{-n})).$$

Thus, almost all labelled graphs of size $n$ are connected, a fact first found by Korshunov. (Bollobás's book [2] contains an extensive account of probabilistic and combinatorial properties of random graphs.) Notice that here use could be made of a purely divergent generating function for asymptotic enumeration purposes.

*Note.* The class of all graphs is not constructible in the sense that it does not admit a complete construction from single atoms involving only sums, products, sets and cycles. (This assertion can be established rigorously by complex analysis since EGF's of constructible classes all have a nonzero radius of convergence.) In contrast, the structures encountered in the rest of this chapter are all constructible.

## 2.6 Cycles in permutations

A relation between classes of combinatorial structures, like the (labelled) set or sequence relation, induces a large number of related counting results. We propose to examine here results that accompany the decomposition

$$\mathcal{P} = \mathfrak{P}\{\mathcal{K}\},$$

with $\mathcal{P}$ the class of permutations and $\mathcal{K}$ the class of cyclic permutations.

*A. Involutions.* A permutation $\sigma$ is an *involution* if $\sigma^2 = Id$ with $Id$ the identity permutation. Quite clearly, an involution can have only cycles of sizes 1 and 2. The class $\mathcal{I}$ of all involutions thus satisfies

$$\mathcal{I} = \mathfrak{P}\{\mathcal{K}_1 + \mathcal{K}_2\}. \tag{2.18}$$

The translation of (2.18) is immediate:

$$I(z) \equiv \sum_n I_n \frac{z^n}{n!} = \exp\left(z + \frac{z^2}{2}\right). \tag{2.19}$$

This last equation then provides the formula

$$I_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n!}{(n-2k)!2^k k!},$$

which resolves the counting problem explicitly.

A *pairing* is an involution without fixed point. The EGF $J(z)$ of all pairings is

$$J(z) = e^{z^2/2} \text{ so that } J_{2n} = 1 \cdot 3 \cdot 5 \cdots (2n-1)$$

as was to be anticipated from a direct reasoning.

*B. Constrained cycle lengths.* The foregoing discussion readily generalizes. Fix some integer $r$ and let $\mathcal{A}^{(r)}$ be the class of permutations $\sigma$ satisfying

$\sigma^r = Id$; let $\mathcal{B}^{(r)}$ be the class of permutations, all of whose cycles have length at most equal to $r$. Clearly,

$$A^{(r)}(z) = \exp\left(\sum_{j|r} \frac{z^j}{j}\right), \quad B^{(r)}(z) = \exp\left(\sum_{j=1}^{r} \frac{z^j}{j}\right). \tag{2.20}$$

(Both formulæ coincide when $r = 1, 2$.)

Exercise 6. The numbers $B_n^{(r)}$ satisfy the recurrence

$$B_{n+1}^{(r)} = (n+1)B_n^{(r)} - n(n-1)\cdots(n-r+1)B_{n-r}^{(r)}.$$

*C. Derangements.* Classically, a derangement is defined as a permutation without fixed points, i.e., $\sigma_i \neq i$ for all $i$. Given an integer $r$, an $r$–derangement is a permutation all of whose cycles have length larger than $r$. Let $\mathcal{D}^{(r)}$ be the class of all $r$–derangements. A specification is

$$\mathcal{D}^{(r)} = \mathfrak{P}\{\mathcal{K} \setminus \bigcup_{j=1}^{r} \mathcal{K}_j\}, \tag{2.21}$$

the corresponding EGF being then

$$D^{(r)}(z) = \exp\left(\log\frac{1}{1-z} - \sum_{j=1}^{r} \frac{z^j}{j}\right) = \frac{\exp(-\sum_{j=1}^{r} \frac{z^j}{j})}{1-z}. \tag{2.22}$$

For instance, when $r = 1$, we find by a direct expansion that

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

which is a truncated series of $\exp(-1)$ and thus converges fast to $\frac{1}{e}$. Phrased differently, this is a famous combinatorial problem from the nineteenth century. A number $n$ of people go to opera, leave their hats on hook in the cloakroom and grab them at random when leaving; the probability that nobody gets back his own hat is

$$\frac{1}{e} + O(\frac{1}{(n+1)!}).$$

*D. Even and odd size cycles.* In the same vein, the EGF's of permutations having only even size cycles ($E(z)$) and odd size cycles ($O(z)$) are

$$\begin{aligned} E(z) &= \exp\left(\frac{1}{2}\log\frac{1}{1-z^2}\right) = \frac{1}{\sqrt{1-z^2}}, \\ O(z) &= \exp\left(\frac{1}{2}\log\frac{1+z}{1-z}\right) = \sqrt{\frac{1+z}{1-z}}. \end{aligned} \tag{2.23}$$

This formula results from retaining only the even and odd terms of $K(z)$,

$$\begin{aligned} K^{\text{even}}(z) &= \textstyle\sum_{n\equiv 0\ (2)} \frac{z^n}{n} = \frac{1}{2}\log\frac{1}{1-z^2}, \\ K^{\text{odd}}(z) &= \textstyle\sum_{n\equiv 1\ (2)} \frac{z^n}{n} = \frac{1}{2}\log\frac{1+z}{1-z}. \end{aligned}$$

Exercise 7. Prove that

$$E_{2n} = (1\cdot 3\cdot 5\cdots(2n-1))^2.$$

Express $O_{2n}, O_{2n+1}$ in terms of $E_{2n}$. Find direct combinatorial explanations for these formulæ.

*E. Even and odd number of cycles.* The EGF's of permutations having an even number of cycles ($E^*(z)$) and an odd number of cycles ($O^*(z)$) are

$$\begin{aligned} E^*(z) &= \cosh(\log\frac{1}{1-z}) = \frac{1}{2}\frac{1}{1-z} + \frac{1}{2} - \frac{z}{2}, \\ O^*(z) &= \sinh(\log\frac{1}{1-z}) = \frac{1}{2}\frac{1}{1-z} - \frac{1}{2} + \frac{z}{2}. \end{aligned}$$

This time the formula results from keeping the terms of even or odd rank in the exponential formula that corresponds to the set construction:

$$\cosh(u) = \sum_{k\equiv 0\ (2)} \frac{u^k}{k!}, \qquad \sinh(u) = \sum_{k\equiv 1\ (2)} \frac{u^k}{k!}.$$

The pattern behing these various examples is expressed by the following proposition.

**Proposition 2.1** *Let $\mathcal{P}^{(A,B)}$ be the class of permutations with cycle sizes in $A$ and with a number of cycles that belongs to $B$. The corresponding EGF is*

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

Exercise 8. Find explicit EGF's and formulæ for permutations having an odd/even number of cycles each of odd/even length.

## 2.7 Words, occupancy problems, and partitions

This section deals with some enumerative problems that present themselves in hashing, random allocations, and statistics on letters in words.

**Words.** Fix an alphabet

$$\mathcal{X} = \{a_1, a_2, \ldots, a_r\}$$

of cardinality $r$, and let $\mathcal{W}$ be the class of all words on the alphabet $\mathcal{X}$, the size of a word being its length. A word of length $n$, $w \in \mathcal{W}_n$, can be viewed as an unconstrained function from $[1\,..\,n]$ to $[1\,..\,r]$, the function associating to each position the value of the corresponding letter in the word (numbered from 1 to $r$).

As was done for surjections, a function $\phi \in [1\,..\,n] \mapsto [1, ..\, r]$ is determined by the collection of its preimages $(\phi^{(-1)}(1), \phi^{(-1)}(2), \ldots, \phi^{(-1)}(r))$. Let $\mathcal{U} = \mathcal{V} \cup \{\epsilon\}$ denote the class of all linear sorted graph augmented by the addition of the empty graph. Then, we have

$$\mathcal{W} = \mathfrak{S}\{\mathcal{U}, \mathrm{card} = r\} = \mathcal{U}^r. \tag{2.24}$$

The EGF of $\mathcal{U}$ is $U(z) = e^z$. Thus, the EGF of $\mathcal{W}$ is

$$W(z) = (e^z)^r = e^{rz}, \tag{2.25}$$

which yields $W_n = r^n$, as was to be expected.

For the situation where restrictions are imposed on the number of occurrences of letters, the decomposition (2.24) generalizes as follows.

**Proposition 2.2** *Let $\mathcal{W}^{(A)}$ denote the family of words such that the number of occurrences of each letter lies in a set $A$. Then*

$$W^{(A)}(z) = (\alpha(z))^r \qquad \textit{where} \qquad \alpha(z) = \sum_{a \in A} \frac{z^a}{a!}. \tag{2.26}$$

The truncated exponential function is classically defined by

$$e_b(z) = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^b}{b!}. \tag{2.27}$$

From Proposition 2.2, the EGF of words containing *at most* $b$ times each letter, and that of words containing *more* than $b$ times each letter are

$$(e_b(z))^r \qquad \text{and} \qquad (e^z - e_b(z))^r\,, \tag{2.28}$$

respectively. This gives the number of $r$–arrangements and surjections (take $b = 1$ in the second formula). Also, the number of doubly surjective functions (each value in the range is taken at least twice) from $[1 \,..\, n]$ to $[1 \,..\, r]$, as results from taking $b = 2$, is

$$n!\,[z^n](e^z - 1 - z)^r.$$

Exercise 9. Find an explicit formula for double surjection numbers.
Estimate asymptotically the number of words where each letter occurs at least $k$ times ($k, r$ fixed and $n \to \infty$). Show that, for almost all words, the number of occurrences of the least frequent letter stays unbounded as $n \to \infty$. (This is a very weak form of the *weak law of large numbers.*)

**Set partitions.** We return to the classes $\mathcal{R}^{(k)}$ of $k$–surjections, and $\mathcal{S}^{(k)}$ of set partitions with $k$ classes. We have the isomorphism

$$\mathcal{R}^{(k)} \cong \mathfrak{S}\{\mathcal{V}, \mathrm{card} = k\} = \mathcal{V}^k, \qquad \mathcal{S}^{(k)} \cong \mathfrak{P}\{\mathcal{V}, \mathrm{card} = k\} = \frac{1}{k!}\mathcal{V}^k$$

with $\mathcal{V}$ the class of nonempty urns, or what amounts to the same the class of all linear sorted graphs of size $\geq 1$. This gives the corresponding EGF's directly:

$$R^{(k)}(z) = (e^z - 1)^k, \quad S^{(k)}(z) = \frac{1}{k!}(e^z - 1)^k.$$

The line of reasoning already adopted for permutations and words provides a more general result.

**Proposition 2.3** *Let $\mathcal{R}^{(A,B)}$ be the class of surjections where the cardinalities of the preimages lie in $A$ and the cardinality of the range belongs to $B$. The corresponding EGF is*

$$R^{(A,B)}(z) = \beta(\alpha(z)) \quad \textit{where} \quad \alpha(z) = \sum_{a \in A} \frac{z^a}{a!}, \quad \beta(z) = \sum_{b \in B} z^b.$$

*Let $\mathcal{S}^{(A,B)}$ be the class of set partitions with part sizes in $A$ and with a number of parts that belongs to $B$. The corresponding EGF is*

$$S^{(A,B)}(z) = \beta(\alpha(z)) \quad \textit{where} \quad \alpha(z) = \sum_{a \in A} \frac{z^a}{a!}, \quad \beta(z) = \sum_{b \in B} \frac{z^b}{b!}.$$

This gives the EGF's of special classes of partitions,

$$f_1 = \exp(e_b(z)-1),\ \ f_2 = \exp(e^z - e_b(z)),\ \ f_3 = \exp(\sinh(z)),\ \ f_4 = \sinh(e^z - 1),$$

corresponding to: (i) largest part $\le b$; (ii) smallest part $> b$; (iii) only odd sized parts; (iv) an odd number of parts.

Exercise 10. The EGF of partitions without singleton parts is

$$e^{e^z-1-z}.$$

Find the EGF of partitions with an odd/even number of parts each of odd/even size.

What distinguishes a labelled structure from an unlabelled one? There is nothing intrinsic there, and everything is in the eyes of the beholder! Take the class of words $\mathcal{W}$ over an alphabet of cardinality $r$. The two generating functions

$$W(z) \equiv \sum_n W_n z^n = \frac{1}{1-rz} \qquad \text{and} \qquad \widehat{W}(z) \equiv \sum_n W_n \frac{z^n}{n!} = e^{rz},$$

leading in both cases to $W_n = r^n$, correspond to two different ways of constructing words, the first one directly as an unlabelled sequence, the other one as a labelled power of letter positions. A similar situation arises for $r$–partitions, for which we found

$$S^{(r)}(z) = \frac{z^r}{(1-z)(1-2z)\cdots(1-rz)} \qquad \text{and} \qquad \widehat{S}^{(r)}(z) = \frac{(e^z-1)^r}{r!},$$

by viewing them as either unlabelled structures (an encoding via words of a regular language) or directly as labelled structures (a labelled set).

## 2.8 Labelled trees

The trees considered here are labelled structures, meaning as usual that their nodes bear integer labels. We restrict attention to the rooted variety of unordered (nonplane) trees. We denote by $\mathcal{T}$ this class.

The class $\mathcal{T}$ is definable by the symbolic equation

$$\mathcal{T} = \mathcal{N} \star \mathfrak{P}\{\mathcal{T}\}, \tag{2.29}$$

where $\mathcal{N}$ represents the class consisting of a single labelled node: $\mathcal{N} = \{1\}$. From the specification (2.29), we obtain that the EGF $T(z)$ is defined implicitly by the "functional equation"

$$T(z) = ze^{T(z)}. \tag{2.30}$$

The first few values are easily found:

$$T(z) = z + 2^1\frac{z^2}{2!} + 3^2\frac{z^3}{3!} + 4^3\frac{z^4}{4!} + 5^4\frac{z^5}{5!} + \cdots .$$

This leads to expect that $T_n = n^{n-1}$, a fact proved by the *Lagrange Inversion Theorem.*

**Theorem 2.3 (Lagrange Inversion)** *Let $\phi(u) = \sum_{j=0}^{\infty} \phi_j u^j$ be a formal power series with $\phi_0 \neq 0$, and let $Y(z)$ be the unique formal power series solution of the equation $Y = z\phi(Y)$. The coefficients of $Y$, $Y^k$, and $\psi(Y)$ (for an arbitrary series $\psi$) are given by*

$$\begin{aligned}
[z^n]Y(z) &= \frac{1}{n}[u^{n-1}](\phi(u))^n \\
[z^n]Y^k(z) &= \frac{k}{n}[u^{n-k}](\phi(u))^n \\
[z^n]\psi(Y(z)) &= \frac{1}{n}[u^{n-1}](\phi(u))^n\psi'(u).
\end{aligned} \tag{2.31}$$

**Proof.** [Sketch] By Cauchy's formula, we have

$$[z^n]Y(z) = \frac{1}{2i\pi}\oint Y(z)\,\frac{dz}{z^{n+1}} \qquad \text{where} \qquad Y = z\phi(Y).$$

Take $y = Y(z)$ as the independent variable. We have

$$z = \frac{y}{\phi(y)}, \quad dz = \frac{\phi(y) - y\phi'(y)}{\phi^2(y)}\,dy.$$

Thus

$$\begin{aligned}
[z^n]Y(z) &= \frac{1}{2i\pi}\oint\left[\frac{\phi^n(y)}{y^n} - \frac{\phi^{n-1}(y)\phi'(y)}{y^{n-1}}\right]dy \\
&= [y^{n-1}](\phi(y))^n - [y^{n-2}](\phi(y))^{n-1}\phi'(y) \\
&= \frac{1}{n}[y^{n-1}](\phi(y))^n.
\end{aligned}$$

The second line follows from Cauchy's coefficient formula applied backwards; the third line follows from the observation that

$$[y^{n-2}](\phi(y))^{n-1}\phi'(y) = \frac{n-1}{n}[y^{n-1}](\phi(y))^n,$$

by standard rules on coefficients of derivatives.

The developments for $Y^k$ and $\psi(Y)$ are entirely similar. Furthermore, the result is purely algebraic and its proof can be freed of analytic conditions by a continuity argument. □

As an immediate consequence, one gets the number of trees

$$T_n = n![z^n]T(z) = n! \cdot \frac{1}{n}[u^{n-1}]\,(e^u)^n = n^{n-1},$$

a result originally due to Arthur Cayley in 1889, whence the name of Cayley trees sometimes given to unordered labelled trees. The number of unordered forests ($k$–sets of trees) also follows from Lagrange inversion:

$$T_n^{(k)} = n![z^n]\frac{(T(z))^k}{k!} = \frac{(n-1)!}{(k-1)!}[u^{n-k}](e^u)^n = \binom{n-1}{k-1}n^{n-k}.$$

A similar process gives the number of trees where all (out)degrees of nodes are restricted to lie in a set $\Omega$.

**Proposition 2.4** *The number of trees $T_n$ and the number of $k$–forests $T_n^{(k)}$ are given by*

$$T_n = n^{n-1}, \qquad T_n^{(k)} = \binom{n-1}{k-1}n^{n-k}.$$

*The number of trees and $k$–forests, where all nodes have degree in $\Omega$, is*

$$T_n^{(\Omega)} = (n-1)![u^{n-1}](\omega(u))^n, \quad T_n^{(\Omega,k)} = \frac{(n-1)!}{(k-1)!}[u^{n-k}](\omega(u))^n,$$

*where $\omega(u) = \sum_{d\in\Omega}\frac{u^d}{d!}$.*

Exercise 11. The number of ordered (plane) rooted labelled trees with all node degrees allowed is

$$n! \cdot \frac{1}{n}\binom{2n-2}{n-1} = \frac{(2n-2)!}{(n-1)!}.$$

This follows either from an EGF computation, or from a reduction to the unlabelled case via a canonical traversal order.

Find the EGF for ordered and rooted labelled trees, all of whose node degrees are in a set $\Omega$.

EXERCISE 12. The number of labelled (rooted unordered) binary trees of size $2n+1$, where each node has outdegree either 0 or 2, involves the Catalan numbers. Find a direct combinatorial explanation for the result by relating it to unlabelled plane (ordered) trees. [Hint: Use symmetries and a canonical traversal order for the labelling.]

**Functional graphs and finite mappings.** Let $\mathcal{F}$ be the class of mappings from $[1\,..\,n]$ to itself. A mapping $f \in [1\,..\,n] \mapsto [1\,..\,n]$ can be represented by a directed graph over the set of vertices $[1\,..\,n]$ with an edge connecting $x$ to $f(x)$, for all $x \in [1\,..\,n]$. The graphs so obtained are called *functional graphs* and they have the characteristic property that the outdegree of each vertex is exactly equal to 1.

Starting from any point $x_0$, the succession of (directed) edges in the graph traverses the iterates of the function, $x_0, f(x_0), f(f(x_0)), \ldots$, and since the domain is finite, each such sequence must eventually loop on itself. When the operation is repeated, the elements group themselves into components. This leads to another characterization of functional graphs:

> A functional graph is a set of connected functional graphs. A connected functional graph is a collection of rooted trees arranged in a cycle.

Thus, with $\mathcal{T}$ still being the class of all Cayley trees, and with $\mathcal{K}$ denoting the class of all connected functional graphs, we have the specification:

$$\left\{\begin{array}{rcl} \mathcal{F} & = & \mathfrak{P}\{\mathcal{K}\} \\ \mathcal{K} & = & \mathfrak{C}\{\mathcal{T}\} \\ \mathcal{T} & = & \{1\} \star \mathfrak{P}\{\mathcal{T}\}. \end{array}\right. \tag{2.32}$$

This translates at sight into a set of equations for EGF's

$$\left\{\begin{array}{rcl} F(z) & = & e^{K(z)} \\ K(z) & = & \log \dfrac{1}{1-T(z)} \\ T(z) & = & e^{T(z)}. \end{array}\right. \tag{2.33}$$

Eventually, the EGF $F$ satisfies $F = (1-T)^{-1}$. It can be checked from there, by Lagrange inversion once more, that we have

$$F_n = n^n,$$

as was to be expected (!).

The constructions (2.32) lead to a large number of counting results. For instance, the functions without fixed points, $f(x) \neq x$, and without 2–cycles, $f(f(x)) \neq x$, have EGFs

$$\frac{e^{-T(z)}}{1-T(z)}, \qquad \frac{e^{-T(z)-T^2(z)/2}}{1-T(z)}.$$

The first equation is consistent with what a direct count yields, namely $(n-1)^n$.

Several analyses of this type are of relevance to cryptography and the study of random number generators. For instance, the fact that a random mapping over $[1\,..\,n]$ tends to reach a cycle in $O(\sqrt{n})$ steps led Pollard to design a Monte Carlo integer factorization algorithm, see [17, p. 371]. The algorithm suitably optimised in turn led to the factorization of the Fermat number $F_8 = 2^{2^8}+1$ by Brent in 1980.

Exercise 13. The class $\mathcal{BF}$ of binary mappings (where each point has either 0 or 2 preimages) is enumerated by

$$BF(z) = \frac{1}{\sqrt{1-2z^2}} \qquad BF_{2n} = \frac{((2n)!)^2}{2^n (n!)^2}.$$

Also find a direct combinatorial argument.

## 2.9 Combinatorial models and interpretations.

We briefly pause at this stage and discuss a reverse problem, that of going from a given generating function to a class of structures enumerated by the function. The exercise of "interpreting" a generating function by finding a suitable "combinatorial model" has some merit in that it sheds some more light on the symbolic approach exposed here.

Consider the four "initial" generating functions,

$$z, \quad \frac{1}{1-z}, \quad \log\frac{1}{1-z}, \quad e^z.$$

The EGF $f_0(z) = z$ is the EGF of class that contains a unique labelled structure of size one, that may be freely taken to be a single node labelled by 1 and called the *labelled atom*. Then $f_1(z) = (1-z)^{-1}$ is, by the sequence construction, a labelled sequence of atoms: this corresponds exactly to the representation of a permutation as a sequence of distinct integers. The interpretation of $f_2(z) = \log(1-z)^{-1}$ is similar, and we get the cyclic permutations from the cycle construction applied to atoms.

The case of $f_3(z) = e^z$ is interesting. It corresponds to having for each $n$ a unique labelled structure of size $n$. There are several possible models for this.

- The graph where the labels $[1\,..\,n]$ are presented on a line in increasing order (technically the empty graph should be allowed for $n = 0$). This corresponds to the *linear sorted graphs* already encountered.
- The totally disconnected graph on $[1\,..\,n]$. This corresponds to an *urn*, with the empty urn allowed for $n = 0$. The corresponding class is denoted by $\mathcal{U}$.
- Any other graph structure that has only 1 structure for each $n$. For instance, if combinatorially meaningful for a given application, we could take the complete graph over $[1\,..\,n]$.

The model of linear sorted graphs has already proved useful in enumerating surjections and set partitions. That of urns is closely related to hashing and occupancy problems.

An *arrangement* of $[1\,..\,n]$ is a permutation of a subset of $[1\,..\,n]$: in simpler terms, an arrangement specifies a selection of some elements of $[1\,..\,n]$ in some order. Let $\mathcal{A}_n$ be the class of arrangements of $[1\,..\,n]$. We have

$$A_n = 1 + \sum_{k=1}^{n} (n)_k \qquad \text{where} \quad (n)_k = n(n-1)\cdots(n-k+1),$$

as follows from a straight counting of possibilities. The EGF is then readily computed,

$$A(z) = \frac{e^z}{1-z}.$$

We can now look back at this formula and interpret it. It corresponds to the isomorphism

$$\mathcal{A} \cong \mathcal{U} \star \mathcal{P},$$

whose meaning is *a posteriori* clear: an arrangement decomposes into an urn of elements (the ones that we do not take in the arrangement), the rest consisting of the selected elements in the order specified by the permutation.

More generally, it is that any suitably positive "exp–log" function as provided by Theorem 2.2 can receive a combinatorial interpretation. This interplay between combinatorial form and algebraic analysis is one of the corner stones of combinatorial analysis. Such an approach often proves to be an instructive way of establishing identities that special functions of analysis satsify by interpreting the identities as reflecting simple combinatorial bijections); it is also a means of getting deeper insight into the class of generating functions associated to a given type of combinatorial problem.

Exercise 14. Find combinatorial models (i.e., interpret combinatorially)

$$e^{z^2}, e^{z^2/2}, e^{z^3}, e^{z^3/3}, e^{z^3/6}, \frac{1}{1-z^3}, \frac{1}{1-z^3/3}, \frac{1}{1-z^3/6}.$$

Exercise 15. Find combinatorial models for

$$e^{z/(1-z)}, \quad e^{ze^z}, \quad \log\frac{1}{2-e^z}, \quad e^{e^{e^z-1}-1}, \quad \frac{1}{1-ze^z}, \quad e^{ze^z/(1-z)}.$$

## 2.10 Additional constructions

The *pointing* of a class $\mathcal{A}$ is defined by

$$\mathcal{B} = \Theta\mathcal{A} \ \text{ iff } \ \mathcal{B}_n = [1\,..\,n] \times \mathcal{A}_n.$$

In other words, in order to generate an element of $\mathcal{B}$, select one of the $n$ labels and point at it. Clearly

$$B_n = n \cdot A_n \quad \Longrightarrow \quad B(z) = z\frac{d}{dz}A(z).$$

The *composition* or *substitution* can be defined so that it corresponds *a priori* to composition of generating functions. It is formally defined as

$$\mathcal{B} \circ \mathcal{C} = \sum_{k=0}^{\infty} \mathcal{B}_k \times \mathfrak{P}\{\mathcal{C}, \mathrm{card} = k\},$$

so that its EGF is

$$\sum_{k=0}^{\infty} B_k \frac{(C(z))^k}{k!} = B(C(z)).$$

A combinatorial way of realizing this definition and form $\mathcal{B} \circ \mathcal{C}$, is as follows: select some element of $\mathcal{B}$ of some size $k$, then a $k$–set of $\mathcal{C}^k$; the elements of the $k$–set are naturally ordered by value of their "leader" (the leader of an object being the value of its smallest label); the element with leader of rank $r$ is then substituted to the labelled node of value $r$ in $\mathcal{B}$.

**Theorem 2.4** *The combinatorial constructions of pointing and substitution are admissible.*

$$\mathcal{B} = \Theta\mathcal{A} \implies B(z) = z\frac{d}{dz}A(z)$$

$$\mathcal{C} = \mathcal{A}[\mathcal{B}] \implies C(z) = A(B(z)).$$

Exercise 16. The EGF of relabelled pairings of elements of $\mathcal{A}$ is

$$e^{A+A^2/2}.$$

Exercise 17. The sequence class of $\mathcal{A}$ may be defined by composition as $\mathcal{P} \circ \mathcal{A}$ where $\mathcal{P}$ is the set of all permutations. The powerset class of $\mathcal{A}$ may be defined as $\mathcal{U} \circ \mathcal{A}$ where $\mathcal{U}$ is the class of all urns. Thus,

$$\mathfrak{S}\{\mathcal{A}\} \cong \mathcal{P} \circ \mathcal{A}, \quad \mathfrak{P}\{\mathcal{A}\} \cong \mathcal{U} \circ \mathcal{A}.$$

[Hint. See Joyal's work [16] for an extensive use of such ideas.]

Exercise 18. The EGF's of permutations with cycles of distinct lengths and of set partitions with parts of distinct sizes are

$$\prod_{n=1}^{\infty}(1+\frac{z^n}{n}), \quad \prod_{n=1}^{\infty}(1+\frac{z^n}{n!}).$$

The probability that a permutation of $\mathcal{P}_n$ has distinct cycle sizes tends to $e^{-\gamma}$, see [14] for a Tauberian argument.

## 2.11 Order constraints

A construction well suited to taking into account many order properties of combinatorial structures is that of the modified labelled product,

$$\mathcal{A} = (\mathcal{B}^{\Box} \star \mathcal{C}).$$

This denotes the subset of the product $\mathcal{B} \star \mathcal{C}$ formed with elements such that the smallest label is constrained to lie in the $\mathcal{B}$ component. (To make this definition consistent, it is assumed that $B_0 = 0$.) We call this binary operation on structures the *boxed* product.

**Theorem 2.5** *The boxed product is admissible.*

$$\mathcal{A} = (\mathcal{B}^{\Box} \star \mathcal{C}) \quad \Longrightarrow \quad A(z) = \int_0^z \left(\frac{d}{dt}B(t)\right) \cdot C(t)\,dt.$$

**Proof.** From the definition,

$$A_n = \sum_{k=1}^{n} \binom{n-1}{k-1} B_k C_{n-k}.$$

The binomial coefficient that appears in the standard labelled product is now modified since only $n-1$ labels need to be distributed between the two components, $k-1$ going to the $\mathcal{B}$ component (that is constrained to contain the label 1 already) and $n-k$ to the $\mathcal{C}$ component. From the form

$$\frac{A_n}{n!} = \frac{1}{n}\sum_{k=0}^{n} \binom{n}{k} (kB_k)\, C_{n-k},$$

the result follows by taking EGF's. $\Box$

A useful special case is the min–rooting operation,

$$\mathcal{A} = \{1\}^{\Box} \star \mathcal{C},$$

for which a variant definition goes as follows. Take in all possible ways elements $\gamma \in \mathcal{C}$, prepend an atom with a label smaller than the labels of $\gamma$, for instance 0, and relabel in the canonical way over $[1\,..\,(n+1)]$. Clearly $A_{n+1} = C_n$ which yields

$$A(z) = \int_0^z C(t)\,dt,$$

a result also consistent with the general formula of boxed products. In passing, we have a way of combinatorially interpreting integration.

EXERCISE 19. Find a one line combinatorial proof of *integration by parts*,

$$\int_0^z A'(t) \cdot B(t)\, dt = A(z) \cdot B(z) - \int_0^z A(t) \cdot B'(t)\, dt.$$

**Increasing trees.** To each permutation, one can associate bijectively a binary tree of a special type[2] called an *increasing binary trees* and sometimes a heap–ordered tree or a tournament tree. This is a plane rooted binary tree in which internal nodes bear labels in the usual way.

The correspondence is as follows: Given a permutation of some sort written as a word, $\sigma = \sigma_1 \sigma_2 \ldots \sigma_n$, factor it in the form $\sigma = \sigma_L \cdot \min(\sigma) \cdot \sigma_R$, with $\min(\sigma)$ the smallest label value in the permutation, and $\sigma_L, \sigma_R$ the factors left and right of $\min(\sigma)$. Then the binary tree $\beta(\sigma)$ is defined recursively by

$$\beta(\sigma) = \langle \min(\sigma), \beta(\sigma_L), \beta(\sigma_R) \rangle, \qquad \beta(\epsilon) = \epsilon.$$

The empty tree (consisting of a unique external node of size 0) goes with the empty permutation $\epsilon$. Conversely, reading the labels of the tree in symmetric (infix) order gives back the originla permutation.

Thus, the family $\mathcal{I}$ of binary increasing trees satisfies the recursive definition

$$\mathcal{I} = \{\epsilon\} + (\{1\}^{\square} \star \mathcal{I} \star \mathcal{I}),$$

which implies the nonlinear integral equation for the EGF

$$I(z) = 1 + \int_0^z I^2(t)\, dt.$$

This equation reduces to $I'(z) = I^2(z)$ and, under the initial condition $I(0) = 1$, it admits the solution $I(z) = (1 - z)^{-1}$. Thus $I_n = n!$, which is consistent with the fact that there are as many increasing trees as there are permutations.

The construction of increasing trees associated with permutation is instrumental in deriving EGF's relative to various local order patterns in permutations, like the number of ascents and descents, rises, falls, peaks and troughs, etc. We illustrate its use here by counting the number of *up-and-down* permutations, also known as *alternating* permutations. The result

[2] Such trees are closely related to data structures like heaps and binomial queues [4, 19].

was first derived by Désiré André in 1881 by means of a direct recurrence argument.

A permutation $\sigma = \sigma_1\sigma_2 \ldots \sigma_{2n+1}$ is an alternating permutation if

$$\sigma_1 > \sigma_2 < \sigma_3 > \cdots < \sigma_{2n+1}.$$

It can be checked that the corresponding increasing trees have no one–way branching nodes. Thus, the corresponding specification is

$$\mathcal{J} = \{1\} + (\{1\}^{\Box} \star \mathcal{I} \star \mathcal{I}),$$

so that

$$J(z) = z + \int_0^z J^2(t)\,dt \quad \text{or} \quad \frac{d}{dz}J(z) = 1 + J^2(z),$$

which implies (with $J(0) = 0$)

$$J(z) = \tan(z) = z + 2\frac{z^3}{3!} + 16\frac{z^5}{5!} + 272\frac{z^7}{7!} + \cdots.$$

The $J_{2n+1}$ are known as the *tangent numbers* or the *Euler numbers* of odd index. Alternating permutations are thus counted by the tangent numbers.

Analyses of increasing trees also model performance issues in binary search trees, quicksort, and heap–like priority queue structures.

EXERCISE 20. The number of alternating permutations of $[1\,..\,2n]$ is

$$(2n)!\,[z^{2n}]\frac{1}{\cos(z)}.$$

**Labelled constructions.** It is possible to base a fair part of the theory of labelled constructions on the box operation. Consider the three relations

$$\begin{array}{llll}
\mathcal{F} = \mathfrak{S}\{\mathcal{G}\} & \Longrightarrow & f(z) = \frac{1}{1-g(z)}, & f = 1 + gf \\
\mathcal{F} = \mathfrak{P}\{\mathcal{G}\} & \Longrightarrow & f(z) = e^{g(z)}, & f = \int g'f \\
\mathcal{F} = \mathfrak{C}\{\mathcal{G}\} & \Longrightarrow & f(z) = \log\frac{1}{1-g(z)}, & f = \int g'\frac{1}{1-g}
\end{array}$$

The last column is easily checked to provide an alternative form of the standard operator corresponding to sequences, powersets, and cycles. Each case is then itself deduced directly from Theorem 2.5 and the labelled product rule:

Sequences: they obey the recursive definition

$$\mathcal{F} = \mathfrak{S}\{\mathcal{G}\} \qquad \Longrightarrow \qquad \mathcal{F} \cong \{\epsilon\} + (\mathcal{G} \star \mathcal{F}).$$

Sets: we have

$$\mathcal{F} = \mathfrak{P}\{\mathcal{G}\} \qquad \Longrightarrow \qquad \mathcal{F} \cong \{\epsilon\} + (\mathcal{G}^{\Box} \star \mathcal{F}),$$

which means that a set can always be presented in an ordered way by having the component with the smallest label first. In other words, the elements of a set can be canonically arranged by increasing values of their smallest labels (the "leaders" that have surfaced already on several occasions).

Cycles: The element of a cycle that contains the smallest label can be taken canonically as the cycle "leader", which is then followed by a sequence of elements when the cycle is traversed in order. Thus

$$\mathcal{F} = \mathfrak{C}\{\mathcal{G}\} \qquad \Longrightarrow \qquad \mathcal{F} \cong (\mathcal{G}^{\Box} \times \mathfrak{S}\{\mathcal{G}\}).$$

Greene [13] has developed a complete framework of labelled grammars based on standard and boxed labelled products. In its basic form, its expressive power is essentially equivalent to ours, because of the above relations. More complicated order constraints, dealing simultaneously with a collection of larger and smaller elements, can be furthermore taken into account by Greene's framework.

Exercise 21. Give a generating function translation for the subclass of the product $\mathcal{A} = (\mathcal{B} \star \mathcal{C})$ in which the smallest label is in $\mathcal{B}$ and the largest one in $\mathcal{C}$. Same question with the two smallest elements in $\mathcal{B}$. [See [13].]

Exercise 22. Interpret combinatorially

$$\tan \frac{z}{1-z}, \quad \tan(\tan(z)), \quad \tan(e^z - 1).$$

Exercise 23. [Foata's fundamental correspondence] Show combinatorially that the number of cycles and the number of left to right maxima in permutations of $\mathcal{P}_n$ have the same distribution.

[Hint: Organize the cycle decomposition by cycle leaders and then read the elements sequentially.]

## 2.12 Notes

Labelled constructions are a frequently used paradigm of combinatorial analysis with applications to order statistics and graphical enumerations for instance. See the books by Comtet [3], Wilf [22], Stanley [21], or Goulden and Jackson [11] for many examples.

The labelled set construction and the exponential formula were recognized early by researchers working in the area of graphical enumerations [15]. Foata [9] proposed a detailed formalization in 1974 of labelled constructions, especially sequences and sets, under the names of partitional complex; a brief account is also given by Stanley in his survey [20]. This is parallel to the concept of "prefab" due to Bender and Goldman [1].

Greene developed a general framework of "labelled grammars" largely based on the boxed product with implications for the random generation of combinatorial structures. Joyal's framework [16], already mentioned in the previous chapter, is based on category theory and it encompasses both cases of unlabelled and labelled objects.

Flajolet, Salvy, and Zimmermann have developed a specification language closely related to the system exposed here. They show in [8] how to compile automatically specifications into generating functions; this is complemented by an automatic computation of asymptotic forms which is operational in many cases of interest, using a computer algebra system.

## Problems and exercises

Stirling numbers and their cognates satisfy a host of algebraic relations.

EXERCISE 24. From the differential relation

$$\frac{d}{dz}S^{(k)}(z) = e^z\, S^{(k-1)}(z),$$

the $S_n^{(k)} = \left\{ {n \atop k} \right\}$ satisfy the recurrence

$$\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k\left\{ {n-1 \atop k} \right\}.$$

Expand $(1+(e^z-1))^u$ and deduce the identity

$$u^n = \sum_k \left\{ {n \atop k} \right\} u(u-1)(u-2)\cdots(u-k+1).$$

From the recurrence of from the EGF, show that the ordinary generating function of the $\left\{ {n \atop k} \right\}_{n\geq 0}$ is

$$\sum_{n\geq 0} \left\{ {n \atop k} \right\} z^n = \frac{z^k}{(1-z)(1-2z)\cdots(1-kz)}.$$

EXERCISE 25. By forming a differential equation satisfied by $S(z)$ show that the $S_n$ obey the "full history" recurrence

$$S_{n+1} = \sum_{k=0}^{n} \binom{n}{k} S_k.$$

Find a direct combinatorial interpretation of this recurrence.

Find a full history linear recurrence satisfied by the surjection number.

From these recurrences, show that the $R_n$ and $S_n$ can be computed in time polynomial in $n$. Deduce an algorithm that draws at random a surjection or a set partition of size $n$. [Hint: see the book [18] for a general treatment of such random generation problems.]

Symbolic methods also provide valuable intuitions regarding the probabilistic analysis of certain allocation processes, and they can sometimes be adapted to nonuniform probability distributions [6].

EXERCISE 26. [An urn model] You have $m$ distinguishable balls and tow urns $A$ and $B$. At any time $t = 0, 1, 2, \ldots$, one of the balls changes urns. The EGF of the number of moves of duration $n$ that start with urn $A$ full and end with urn $A$ again full is

$$(\cosh(z))^m.$$

Find explicit expressions. Generalize to the case when $A$ contains initially $k$ balls and finally $\ell$ balls.

Show that in a long span of time each urn contains approximately $m/2$ balls on average.

EXERCISE 27. The coupon collector problem. A company issues coupons of $m$ different types, type $j$ being issued with probability $p_j$. Let $W$ (the "wait" time) be the random variable representing the number of coupons that one needs to gather until a full collection with $m$ different coupons is obtained. Show that

$$\Pr\{W \geq n\} = n!\,[z^n] \prod_{j=1}^{m} (e^{p_j z} - 1).$$

Using Laplace transforms, show that the expectation of $W$ is

$$\mathrm{E}\{W\} = \int_0^\infty \left[1 - \prod_{j=1}^{m} (1 - e^{-p_i x})\right] dx.$$

In the uniform case where $p_j = \frac{1}{m}$, show that

$$\mathrm{E}\{W\} = \sum_{k \geq 1} \binom{m}{k} \frac{(-1)^{k-1}}{k}$$

and that the last expression is equal to the harmonic number $H_m$. Find a direct argument for the uniform case.

Interesting applications of symbolic methods to graphical enumerations are provided by bipartite, acyclic, and 2–regular graphs.

EXERCISE 28. A plane bipartite graph is a pair $(G, \omega)$ where $G$ is labelled graph, $\omega = (\omega_W, \omega_E)$ is a bipartition of the nodes (into *West* and *East* categories), and the edges are such that they only connect nodes from $\omega_W$ to nodes of $\omega_E$. A direct count shows that the EGF of bipartite graphs is

$$\Gamma(z) = \sum_n \gamma_n \frac{z^n}{n!} \quad \text{with} \quad \gamma_n = \sum_k \binom{n}{k} 2^{k(n-k)}.$$

The EGF of plane bipartite graphs that are connected is $\log \Gamma(z)$.

A bipartite graph is a labelled graph whose nodes can be partitioned into two groups so that edges only connect nodes of different groups. The EGF of bipartite graphs is

$$\exp\left(\frac{1}{2}\log \Gamma(z)\right) = \sqrt{\Gamma(z)}.$$

[Hint. The EGF of a connected bipartite graph is $\frac{1}{2}\log\Gamma(z)$ as a factor of $\frac{1}{2}$ kills the East–West orientation present in a connected plane bipartite graph. See Wilf's book [22, p. 78] for details.]

Exercise 29. The number of unrooted labelled unordered trees is $n^{n-2}$ with EGF

$$U(z) = T(z) - \frac{1}{2}(T(z))^2 \quad \text{where} \quad T(z) = ze^{T(z)}$$

is the Cayley function. The GF of acyclic labelled graphs is

$$e^{T(z) - \frac{1}{2}T^2(z)}.$$

Find the EGF of unicyclic graphs, *i.e.e*, graphs containing a single cycle [7]. Derive explicit counting formulæ.

Exercise 30. A 2–regular graph is an undirected graph in which each vertex has degree exactly 2. Show that the connected 2–regular graphs are undirected cycles of length $n \geq 3$. Deduce that the EGF of 2–regular graphs is

$$H(z) = \frac{e^{-z/2 - z^2/4}}{\sqrt{1-z}}.$$

Find a linear recurrence with polynomial coefficients satisfied by $H_n = n![z^n]H(z)$. [Hint: See [3].]

With $I(z)$ the EGF of involutions, justify combinatorially the equality

$$I(z) \cdot H^2(z) = \frac{1}{1-z}.$$

Given $n$ straight lines in general position, a cloud is defined to be a set of $n$ intersection points no three being collinear. Show that clouds and 2–regular graphs are equinumerous. [Hint: Use duality.]

The EGF of nonnegative integer matrices whose row sum and column sums are all equal to 2 is

$$\frac{e^{z/2 + z^2/4}}{\sqrt{1-z}}.$$

[Hint: Convert to a multigraph problem.]

Computations akin to the Lagrange inversion theorem may be used to express the Stirling cycle numbers and prove the Abel identities.

Exercise 31. Prove Schlömilch's formula starting from

$$\frac{k!}{n!}\begin{bmatrix} n \\ k \end{bmatrix} = \frac{1}{2i\pi}\oint \log^k \frac{1}{1-z}\,\frac{dz}{z^{n+1}},$$

and performing the change of variable *a la* Lagrange: $z = 1 - e^{-t}$. [3, p.216].

Exercise 32. By computing in two different ways the coefficient

$$[z^n]e^{(\alpha+\beta)y} = [z^n]e^{\alpha y} \cdot e^{\beta y},$$

where $y = ze^y$ is the Cayley function, derive the *Abel identity*

$$(\alpha+\beta)(n+\alpha+\beta)^{n-1} = \alpha\beta\sum_{k=1}^{n-1}\binom{n}{k}(k+\alpha)^{k-1}(n-k+\beta)^{n-k-1}.$$

Exercise 33. The number of trees with maximum outdegree of node $\leq b$ is

$$(n-1)![u^{n-1}](e_b(u))^n.$$

Give the formulæ corresponding to $b = 1, 2, 3$. Relate to a word counting problem by interpreting back the formula.

Exercise 34.[Prüfer] Find a bijective proof that $T_n = n^{n-1}$. [Hint: See [3], page 63.]

# Bibliography


[1] Bender, E. A., and Goldman, J. R. Enumerative uses of generating functions. *Indiana University Mathematical Journal* (1971), 753–765.

[2] Bollobás, B. *Random Graphs.* Academic Press, 1985.

[3] Comtet, L. *Advanced Combinatorics.* Reidel, Dordrecht, 1974.

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

[5] De Bruijn, N. G. *Asymptotic Methods in Analysis.* Dover, 1981. A reprint of the third North Holland edition, 1970 (first edition, 1958).

[6] Flajolet, P., Gardy, D., and Thimonier, L. Birthday paradox, coupon collectors, caching algorithms, and self–organizing search. *Discrete Applied Mathematics 39* (1992), 207–229.

[7] Flajolet, P., Knuth, D. E., and Pittel, B. The first cycles in an evolving graph. *Discrete Mathematics 75* (1989), 167–215.

[8] Flajolet, P., Salvy, B., and Zimmermann, P. Automatic average–case analysis of algorithms. *Theoretical Computer Science, Series A 79*, 1 (Feb. 1991), 37–109.

[9] Foata, D. *La série génératrice exponentielle dans les problèmes d'énumération.* S.M.S. Montreal University Press, 1974.

[10] Foata, D., and Schützenberger, M.-P. *Théorie Géométrique des Polynômes Euleriens*, vol. 138 of *Lecture Notes in Mathematics.* Springer Verlag, 1970.

[11] Goulden, I. P., and Jackson, D. M. *Combinatorial Enumeration.* John Wiley, New York, 1983.

[12] Graham, R., Knuth, D., and Patashnik, O. *Concrete Mathematics.* Addison Wesley, 1989.

[13] Greene, D. H. *Labelled formal languages and their uses.* PhD thesis, Stanford University, June 1983. Available as Report No. STAN-CS-83-982.

[14] Greene, D. H., and Knuth, D. E. *Mathematics for the analysis of algorithms.* Birkhauser, Boston, 1981.

[15] Harary, F., and Palmer, E. M. *Graphical Enumeration.* Academic Press, 1973.

[16] Joyal, A. Une théorie combinatoire des séries formelles. *Advances in Mathematics 42*, 1 (1981), 1–82.

[17] Knuth, D. E. *The Art of Computer Programming*, vol. 3: Sorting and Searching. Addison-Wesley, 1973.

[18] Nijenhuis, A., and Wilf, H. S. *Combinatorial Algorithms*, second ed. Academic Press, 1978.

[19] Sedgewick, R. *Algorithms*, second ed. Addison–Wesley, Reading, Mass., 1988.

[20] Stanley, R. P. Generating functions. In *Studies in Combinatorics,* M.A.A. Studies in Mathematics, Vol. 17. (1978), G.-C. Rota, Ed., The Mathematical Association of America, pp. 100–141.

[21] Stanley, R. P. *Enumerative Combinatorics*, vol. I. Wadsworth & Brooks/Cole, 1986.

[22] Wilf, H. S. *Generatingfunctionology.* Academic Press, 1990.

# Chapter 3

# Parameters and Multivariate Generating Functions

*Generating functions find averages, etc.*
— HERBERT WILF [17]

Combinatorics and the analysis of algorithms often demand informations on probabilistic properties of parameters of combinatorial objects. For sorting purposes, it may be of interest to know how many inversions a permutation is likely to have, on average, or —if feasible— in distribution, either exactly or asymptotically.

It turns out that the symbolic methods initially developed for counting both unlabelled and labelled structures adapt nicely to the analysis of various sorts of parameters of such structures.

Bivariate generating functions —ordinary or exponential— can keep track of the number of components in a composite construction, a sequence, a (multi)set, or a cycle. From such generating functions, there results either explicit probability distributions or, at least, mean and variance expressions. Typical applications include the number of summands in a composition, the number of blocks in a set partition, the number of cycles in a permutation, or the root degree of a tree.

More generally, bivariate generating functions give access to additive parameters defined inductively over combinatorial objects. An especially important case is the analysis of the number of designated "basic" configurations in an object of a given size, for example, the number of fixed point in a permutation, the number of singleton blocks in a set partition, the number of leaves in trees of various sorts.

The translation schemes from combinatorial constructions to bivariate generating functions appear to be natural refinements of the corresponding ones for the basic constructions that have been previously examined.

## 3.1 Multivariate generating functions

Consider a class $\mathcal{A}$ of combinatorial structures. For us, a *parameter* $\xi$ on the class is a function from $\mathcal{A}$ to the natural numbers $\mathbb{N}$. For instance one may take for $\mathcal{A}$ the class $\mathcal{P}$ of all permutations, and for $\xi$ the parameter that associates to a permutation $\sigma \in \mathcal{P}$ the number of its cycles. Natural questions are then: How many permutations of size $n$ have $k$ cycles? What is the expected number of cycles in a random permutation? Does this parameter have a distribution that can be made explicit, either exactly or asymptotically? What are the characteristics of this distribution in terms of shape, concentration, or limit law?

Let $A_{n,k}$ denote the number of objects in $\mathcal{A}_n$ whose value of the $\xi$–parameter equals $k$:

$$A_{n,k} = \mathrm{card}\{\, \alpha \in \mathcal{A} \mid |\alpha| = n \text{ and } \xi[\alpha] = k\,\}.$$

The quantities $A_{n,k}$ are proportional to the probabilities of $\xi$ over $\mathcal{A}_n$ since

$$\Pr\{\xi = k \mid \mathcal{A}_n\} = \frac{A_{n,k}}{A_n} = \frac{A_{n,k}}{\sum_k A_{n,k}}.$$

An analysis giving the values of the $A_{n,k}$ is, for that reason, often called an analysis in *distribution*.

The *bivariate generating function* (BGF) of $\mathcal{A}$ with respect to $\xi$ is defined as[1]

$$\begin{aligned} A(u,z) &= \sum_{n,k} A_{n,k} u^k z^n \quad \text{(ordinary GF, unlabelled case)} \\ A(u,z) &= \sum_{n,k} A_{n,k} u^k \frac{z^n}{n!} \quad \text{(exponential GF, labelled case).} \end{aligned}$$

In the first case, the bivariate generating function is said to be *ordinary*, in the second case, it is said to be *exponential*.

[1] One often says that $A(u,z)$ is the (bivariate) generating function of $\mathcal{A}$ with the variable $u$ "*marking*" parameter $\xi$.

The counting of $\mathcal{A}$–structures of size $n$ and $\xi$–parameter equal to $k$ is recovered by applying the coefficient operator twice,

$$A_{n,k} = \omega_n \cdot [u^k z^n] A(z, u),$$

where

$$\omega_n = 1 \;\; \text{(OGF, unlabelled case)}, \qquad \omega_n = n! \;\; \text{(EGF, labelled case)}.$$

From the definition, we also see that $A(1, z)$ coincides with the plain generating function of $\mathcal{A}$ (either ordinary or exponential). We observe that a bivariate generating function admits a "combinatorial" form, which for OGF's and EGF's reads

$$A(u, z) = \sum_{\alpha \in \mathcal{A}} u^{\xi[\alpha]} z^{|\alpha|} \qquad \text{and} \qquad A(u, z) = \sum_{\alpha \in \mathcal{A}} u^{\xi[\alpha]} \frac{z^{|\alpha|}}{|\alpha|!}.$$

The *probability generating function* of $\xi$ over $\mathcal{A}_n$ is

$$\mathrm{E}\{u^\xi \mid \mathcal{A}_n\} = \frac{[z^n]A(u, z)}{[z^n]A(1, z)},$$

and with the notation

$$A_n(u) = \omega_n \cdot [z^n] A(u, z)$$

the probability generating function is expressed as

$$\frac{A_n(u)}{A_n(1)}.$$

The *mean* or *expectation* of $\xi$ over $\mathcal{A}_n$ is simply obtained by a differentiation,

$$\mathrm{E}\{\xi \mid \mathcal{A}_n\} = \frac{A_n'(1)}{A_n(1)} \;\; \text{where} \;\; A_n'(1) = \omega_n \cdot [z^n] \left. \frac{\partial}{\partial u} A(u, z) \right|_{u=1},$$

as results from standard computations with probability generating functions.

In combinatorial enumerations, it is usually the unnormalized quantity appearing in the numerator that plays an important rôle:

$$A_n'(1) \equiv \sum_k k A_{n,k} \equiv \sum_{\alpha \in \mathcal{A}_n} \xi[\alpha].$$

This quantity represents the "total" value or "cumulated" value of $\xi$ over $\mathcal{A}_n$. It determines the expectation after normalization so that we call it the *unnormalized mean*. From what we have seen, its GF is

$$\frac{\partial}{\partial u} A(u,z)\Big|_{u=1},$$

which we call the *generating function of (unnormalized) means*.

Estimates of *variance* are useful in order to get an idea of the dispersion of a parameter. For computations, one starts with the moment of order 2 which involves a double differentiation

$$\mathrm{E}\{\xi^2 \mid \mathcal{A}_n\} = \frac{A_n''(1)}{A_n(1)} + \frac{A_n'(1)}{A_n(1)}.$$

From there, the standard deviation ($\sigma_n$) and the variance ($\sigma_n^2$) are recovered by a classical formula:

$$\sigma_n^2 \equiv \mathrm{E}\{(\xi - \mu_n)^2 \mid \mathcal{A}_n\} = \frac{A_n''(1)}{A_n(1)} + \frac{A_n'(1)}{A_n(1)} - \left(\frac{A_n'(1)}{A_n(1)}\right)^2.$$

**Cycles in permutations.** Let us return to the example of cycles in permutations which is of interest in connection with certain sorting algorithms like bubble sort or insertion sort, maximum finding, and *in situ* rearrangement.

We are dealing with labelled objects, hence exponential generating functions. As seen in the previous chapter, the EGF of permutations having $k$ cycles is expressed by

$$\frac{1}{k!}(K(z))^k,$$

where

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

is the EGF of cyclic permutations. From there, the (exponential) bivariate generating function of permutations with respect to number of cycles is computed as follows:

$$\begin{aligned} P(u,z) &= \sum_{k=0}^{\infty} u^k \cdot \frac{1}{k!}K(z))^k \\ &= \exp(uK(z)) \\ &= (1-z)^{-u}. \end{aligned}$$

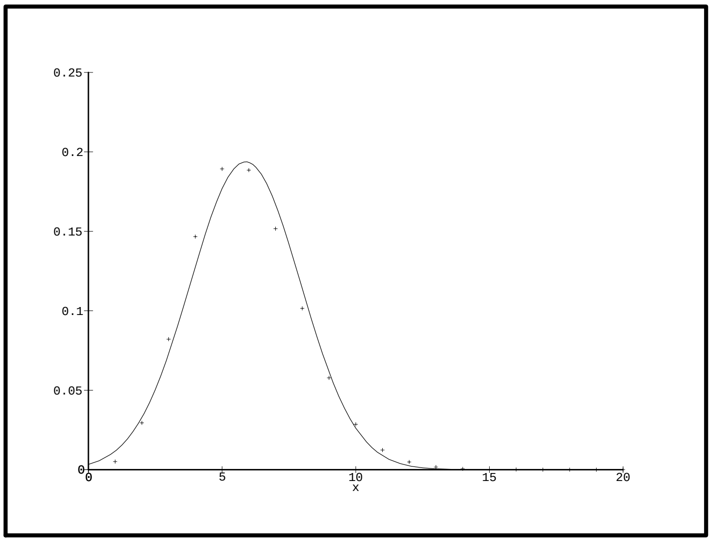



Figure 3.1: The histogram of the distribution of the number of cycles in permutations of size $n = 200$ and a Gaussian curve with the same mean and variance.

The probability generating function for fixed $n$ assumes a simple form in this case, since $[z^n](1-z)^{-u}$ is given by the binomial theorem:

$$\mathrm{E}\{u^\xi \mid \mathcal{P}_n\} = \binom{-u}{n} = \frac{u(u+1)\cdots(u+n-1)}{n!}.$$

This is the generating polynomial for the Stirling cycle numbers, $\left[{n \atop k}\right]$. In a way, these developments completely solve the distribution problem for cycles in permutations, since the $\left[{n \atop k}\right]$ have an explicit form given by Schlömilch's formula:

$$\left[{n \atop k}\right] = \sum_{0 \le j \le h \le n-k} (-1)^{j+h} \binom{h}{j} \binom{n-1+h}{n-k+h} \binom{2n-k}{n-k-h} \frac{(h-j)^{n-k+h}}{h!}.$$

However this formula is somewhat complicated, and asymptotic results are wanted since they should prove easier to interpret.

By differentiation of the bivariate GF with respect to $u$, then setting $u = 1$, we next get the expected number of cycles in a random permutation of size $n$ as a Taylor coefficient

$$\mu_n = [z^n] \frac{1}{1-z} \log \frac{1}{1-z},$$

a quantity which is equal to the $n$th harmonic number $H_n$, a remarkably simple expression indeed when compared to Schlömilch's formula. Thus, on average, a random permutation of size $n$ has about $\log n + \gamma$ cycles, a well known fact of discrete probability. Observe that the normalization in order to get the mean was especially simple here since the number of permutations of size $n$ is $n!$, and this normalization happens to coincide with the factor of $1/n!$ present in all exponential generating functions.

For the variance, a further differentiation of the bivariate EGF shows that

$$\sigma_n^2 = \left(\sum_{k=1}^{n} \frac{1}{k}\right)^2 - \left(\sum_{k=1}^{n} \frac{1}{k^2}\right).$$

Thus, asymptotically,

$$\sigma_n \sim \sqrt{\log n}.$$

The standard deviation is of smaller order than the mean, and therefore deviations from the mean have an asymptotically negligible probability of occurrence. (This results from the familiar Markov–Chebychev inequalities of probability theory [2, p. 74]). Furthermore, the distribution was proved to be asymptotically Gaussian by Gončarov, near 1944,

The situation encountered with cycles in permutations is typical of iterative (non–recursive) structures. In many cases, especially when dealing with recursive structures, the bivariate GF may satisfy complicated functional equations in two variables (see the example of path length below), and explicit expressions for the distribution are not always available. At best, only asymptotic laws can be found. Nonetheless, in all cases, the BGF's are the central tool in obtaining mean and variance estimates, since their derivatives instantiated at $u = 1$ tend to satisfy much simpler relations than the BGF's themselves.

Exercise 1. With $\epsilon$ being a fixed but arbitrary small positive number, the probability that the number of cycles in a random permutation of size $n$ lies outside of the interval

$$[(1-\epsilon)\log n, (1+\epsilon)\log n]$$

tends to 0 as $n \to \infty$.

Exercise 2. Prove that the moment of order $k$ of the distribution of cycles satisfies asymptotically

$$\mathrm{E}\{\xi^k \mid \mathcal{A}_n\} \sim (\log n)^k,$$

and express it exactly in terms of generalized harmonic numbers.

## 3.2 Components in composite structures

General schemas that we detail now allow to build bivariate generating functions for number of components corresponding to each of the composite constructions of sequence, set, or multiset both in the unlabelled and in the labelled case.

**Theorem 3.1** *Consider* unlabelled *structures and a construction* $\mathcal{B} = \Phi\{\mathcal{A}\}$*, where* $\Phi$ *is of the type sequence, set, multiset or cycle. The ordinary bivariate generating function for* $\mathcal{B}$ *with* $u$ *marking the number of components is*

$$\begin{array}{ll} \textit{Sequence:} & \dfrac{1}{1-uA(z)} \\ \textit{Set:} & \exp\left(\displaystyle\sum_{k=1}^{\infty}(-1)^{k-1}\frac{u^k}{k}A(z^k)\right) \\ \textit{Multiset:} & \exp\left(\displaystyle\sum_{k=1}^{\infty}\frac{u^k}{k}A(z^k)\right) \\ \textit{Cycle:} & \displaystyle\sum_{k=1}^{\infty}\frac{\varphi(k)}{k}\log\frac{1}{1-u^kA(z^k)}. \end{array}$$

*For* labelled *structures, and* $\mathcal{B} = \Phi\{\mathcal{A}\}$*, the exponential bivariate generating functions are*

$$\begin{array}{ll} \textit{Sequence:} & \dfrac{1}{1-uA(z)} \\ \textit{Set:} & \exp(uA(z)) \\ \textit{Cycle:} & \log\dfrac{1}{1-uA(z)}. \end{array}$$

**Proof.** *Unlabelled case.* For sequences, the OGF of sequences formed with $k$ components is $(A(z))^k$. Thus the bivariate OGF follows by summation,

$$B(u,z) = \sum_{k=0}^{\infty} u^k (A(z))^k = \frac{1}{1-uA(z)}.$$

For sets, let $\xi$ be the parameter on $\mathcal{B}$ that represents the number of components. Then, by definition,

$$B(u,z) = \sum_{\beta\in\mathcal{B}} u^{\xi[\beta]} z^{|\beta|}.$$

By the same reasoning as in the univariate case, we have

$$B(u,z) = \prod_{\alpha \in \mathcal{A}} (1 + uz^{|\alpha|}),$$

as can be seen by distributing the product. Application of the exp–log technique shows that

$$\begin{aligned}
\log B(u,z) &= \sum_{\alpha \in \mathcal{A}} \log(1 + uz^{|\alpha|}) \\
&= \sum_{\alpha \in \mathcal{A}} \sum_{k=1}^{\infty} (-1)^{k-1} \frac{u^k}{k} z^{k|\alpha|} \\
&= \sum_{k=1}^{\infty} (-1)^{k-1} \frac{u^k}{k} \sum_{\alpha \in \mathcal{A}} z^{k|\alpha|} \\
&= \sum_{k=1}^{\infty} (-1)^{k-1} \frac{u^k}{k} A(z^k).
\end{aligned}$$

The result thus follows for sets, the proof for multisets being entirely similar. (The proof for cycles is omitted.)

*Labelled case.* There, everything is straightforward as the EGF of sequences formed with $k$ components is $(A(z))^k$, the corresponding EGF's for sets and cycles being $\frac{1}{k!}(A(z))^k$ and $\frac{1}{k}(A(z))^k$. Thus, for the three cases, a summation gives

$$\begin{array}{ll}
\text{Sequence:} & B(u,z) = \sum_{k=0}^{\infty} u^k (A(z))^k \\
\text{Set:} & B(u,z) = \sum_{k=0}^{\infty} \frac{u^k}{k!} (A(z))^k \\
\text{Cycle:} & B(u,z) = \sum_{k=1}^{\infty} \frac{u^k}{k} (A(z))^k.
\end{array}$$

The result then follows. □

An immediate application is to the generating function for the mean number $\xi$ of components in a composite structure $\mathcal{B} = \Phi\{\mathcal{A}\}$. The GFof unnormalized means, $\Xi(z)$, is univariate, and the mean number of components in a random structure of size $n$ is given by

$$\mathrm{E}\{\xi \mid \mathcal{B}_n\} = \frac{[z^n]\Xi(z)}{[z^n]A(z)}.$$

From the developments of the introduction, such a GF is known to derive from the bivariate GF's of Theorem 3.1 by

$$\Xi(z) = \frac{\partial}{\partial u} B(u,z)\bigg|_{u=1}.$$

The computations are best encapsulated once and for all in a general statement.

**Theorem 3.2** *The ordinary generating functions of the unnormalized mean value of number of components in the unlabelled case are*

$$\begin{array}{ll}
\textit{Sequence:} & \Xi(z) = \dfrac{A(z)}{(1-A(z))^2} \\
\textit{Set:} & \Xi(z) = B(z) \cdot \displaystyle\sum_{k=1}^{\infty} (-1)^{k-1} A(z^k) \\
\textit{Multiset:} & \Xi(z) = B(z) \cdot \displaystyle\sum_{k=1}^{\infty} A(z^k) \\
\textit{Cycle:} & \Xi(z) = \displaystyle\sum_{k=1}^{\infty} \varphi(k) \frac{A(z^k)}{1-A(z^k)}.
\end{array}$$

*In the labelled case, the EGF's of unnormalized mean values are*

$$\begin{array}{ll}
\textit{Sequence:} & \Xi(z) = \dfrac{A(z)}{(1-A(z))^2} \\
\textit{Set:} & \Xi(z) = A(z)\, e^{A(z)} \\
\textit{Cycle:} & \Xi(z) = \dfrac{A(z)}{1-A(z)}.
\end{array}$$

Exercise 3. In the labelled case $\mathcal{B} = \mathfrak{P}\{\mathcal{A}\}$, the mean number of components in a $\mathcal{B}$–structure is

$$\frac{\sum_{k=1}^{n} \binom{n}{k} A_k B_{n-k}}{B_n}.$$

We now examine how several counting problems can be solved using these techniques.

**Compositions.** Consider the problem of determining the expected number of summands in a random composition of integer $n$. For instance, $(2,3,5,1,2,5,4)$ is a composition of integer 22 with 7 summands. As seen earlier, the set of compositions is $\mathcal{C} = \mathcal{I}^*$ where $\mathcal{I} = \{1,2,\ldots\}$ and the size of integer $i \in \mathcal{I}$ is $|i| = i$. Thus,

$$C(z) = \frac{1}{1-I(z)} \qquad \text{with} \qquad I(z) = \frac{z}{1-z}.$$

By direct expansion, we found in this way that $C_n = 2^{n-1}$ for $n \geq 1$ and $C_0 = 1$.

The bivariate OGF with $u$ marking the number of summands results directly from Theorem 3.1; it is

$$C(u,z) = \frac{1}{1-uI(z)}.$$

The OGF of mean values, $\Xi(z)$, obtains by differentiation, or by applying directly Theorem 3.2:

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

A direct expansion gives the expected number of summands in a random composition as

$$\frac{\Xi_n}{C_n} = \frac{n+1}{2}.$$

Thus a composition of $n$ has about $n/2$ summands on average: a random composition tends to be made of a large number of small summands.

**Set partitions.** Set partitions $\mathcal{S}$ are built of blocks, a construction that is reflected by the EGF equation

$$S(z) = e^{V(z)} \qquad \text{with} \qquad V(z) = e^z - 1.$$

The bivariate EGF with $u$ marking the number of blocks is then

$$S(u,z) = e^{uV(z)} = e^{u(e^z-1)}.$$

The EGF of mean values, $\Xi(z)$ is thus

$$\Xi(z) = V(z)e^{V(z)} = (e^z-1)e^{e^z-1}.$$

Due to the simple shape of $V(z)$, this is almost a derivative of $S(z)$:

$$\Xi(z) = \frac{d}{dz}S(z) - S(z).$$

Thus, the mean number of blocks in a random partition of size $n$ is

$$\frac{\Xi_n}{S_n} = \frac{S_{n+1}}{S_n} - 1,$$

a quantity directly expressible in terms of Bell numbers. A computation based on the asymptotics of the Bell numbers reveals this expected value to be asymptotic to

$$\frac{n}{\log n}.$$

**Trees.** Consider the parameter $\xi$ equal to the degree of the root in a tree. Take first the class $\mathcal{G}$ of all plane unlabelled trees. A plane tree is a root to which is appended a sequence of trees,

$$\mathcal{G} = \mathcal{N} \times \mathfrak{S}\{\mathcal{G}\},$$

where $\mathcal{N}$ is the class formed of a single node, so that

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

The bivariate GF with $u$ marking $\xi$ is then

$$G(u, z) = \frac{z}{1 - uG(z)},$$

from which the cumulative GF is found,

$$\Xi(z) = \frac{zG(z)}{(1 - G(z))^2}.$$

The recursive relation satisfied by $G$ entails a further simplification,

$$\Xi(z) = \frac{1}{z}G^3(z).$$

A closed form for the coefficient results (for instance by the Lagrange–Bürmann inversion theorem),

$$\Xi_n = \frac{3}{n+1}\binom{2n-3}{n-2},$$

and the mean root degree is found to be asymptotic to $\frac{3}{2}$. A random plane tree is thus usually composed of a small number of root subtrees, at least one of which should be fairly large.

For the class $\mathcal{T}$ of non–plane labelled trees (Cayley trees) the basic EGF equation is

$$T(z) = z\, e^{T(z)},$$

since non–planarity is taken into account by a set construction. In that case, the bivariate EGF satisfies $T(u,z) = ze^{uT(z)}$, and we find

$$\Xi(z) = zT(z)e^{T(z)} = (T(z))^2,$$

so that the mean root degree is, by Lagrange inversion,

$$2(1-\frac{1}{n}) \sim 2.$$

A probabilistic phenomenon similar to that of plane trees is observed here.

## 3.3 Additive parameters

The foregoing treatment of the number of components in structures is susceptible to extensions that apply to any additively defined parameter. The proof techniques are simple and they completely parallel what has already been done.

To start the discussion, consider an unlabelled cartesian product of the type $\mathcal{B} = (\mathcal{A})^2$ and an *additive parameter* $\psi$ defined structurally on $\mathcal{B}$ by

$$\psi[\beta] = \theta(\alpha_1) + \theta(\alpha_2) \qquad \text{when} \quad \beta = (\alpha_1, \alpha_2).$$

There, enumerative properties of $\mathcal{A}$ with respect to $\theta$ are assumed to be known. The problem is to derive informations on the distribution of $\psi$ over $\mathcal{B}$.

Lest $A(u,z)$ and $B(u,z)$ be the bivariate OGF's of the pairs $(\mathcal{A},\theta)$ and $(\mathcal{B},\psi)$. Since $\psi$ is an additive parameter, the quantity $u^{\psi[\beta]}$ is determined multiplicatively from the values of $u^{\theta[\alpha]}$ over components. Using the combinatorial form of generating functions, we have

$$\begin{aligned}
B(u,z) &= \sum_{\beta\in\mathcal{B}} u^{\psi(\beta)} z^{|\beta|} \\
&= \sum_{(\alpha_1,\alpha_2)\in\mathcal{B}} u^{\theta(\alpha_1)+\theta(\alpha_2)} z^{|\alpha_1|+|\alpha_2|} \\
&= \left(\sum_{\alpha\in\mathcal{A}} u^{\theta(\alpha)} z^{|\alpha|}\right)^2 \\
&= A(u,z)^2.
\end{aligned}$$

Thus a cartesian square can be treated by the product rule exactly as in the case of univariate generating functions. The same principle extends readily to all products, and in turn to sequences. For sets, multisets, and cycles, "derivatives" of Pólya operators introduce themselves naturally, as summarized below.

For labelled constructions, the developments are similar with binomial convolutions leading to bivariate EGF's that are exponential.

**Theorem 3.3** *Consider an unlabelled construction $\mathcal{B} = \Phi\{\mathcal{A}\}$ and two parameters, $\theta$ defined on $\mathcal{A}$ and $\psi$ defined on $\mathcal{B}$, that are additively related,*

$$\psi[(\alpha_1, \alpha_2, \ldots \alpha_\ell)] = \theta[\alpha_1] + \theta[\alpha_2] + \cdots + \theta[\alpha_\ell],$$

*when $\beta \in \mathcal{B}$ comprises $\ell$ components The ordinary bivariate generating functions $A(u,z)$ and $B(u,z)$ are connected by the relations*

$$\begin{array}{ll}
\textit{Sequence:} & B(u,z) = \dfrac{1}{1-A(u,z)} \\
\textit{Set:} & B(u,z) = \exp\left(\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} A(u^k, z^k)\right) \\
\textit{Multiset:} & B(u,z) = \exp\left(\displaystyle\sum_{k=1}^{\infty} \frac{1}{k} A(u^k, z^k)\right) \\
\textit{Cycle:} & B(u,z) = \displaystyle\sum_{k=1}^{\infty} \frac{\varphi(k)}{k} \log \frac{1}{1-A(u^k,z^k)}.
\end{array}$$

*In the labelled case and for* exponential *bivariate generating functions,* *one has*

$$\begin{array}{ll} \textit{Sequence:} & B(u,z) = \dfrac{1}{1-A(u,z)} \\ \textit{Set:} & B(u,z) = \exp(A(u,z)) \\ \textit{Cycle:} & B(u,z) = \log \dfrac{1}{1-A(u,z)}. \end{array}$$

**Proof.** For sequences, and for labelled constructions, the proof is a direct application of the combinatorial form of generating functions discussed above. For unlabelled sets and multisets, the exp–log technique does the rest. $\square$

The generating functions of mean values are deduced immediately by differentiation. Several examples follow.

**Compositions.** Consider again compositions, $\mathcal{C} = \mathcal{I}^*$ with $\mathcal{I}$ representing the integer summands. Let first $\psi$ be the number of summands equal to 1 in a composition. Parameter $\psi$ is defined additively from a parameter $\theta$ defined on integers whose value satisfies $\theta[1] = 1$ and $\theta(i) = 0$ when $i \neq 1$. The bivariate GF of $\mathcal{I}$ with $u$ marking $\theta$ is

$$I(u,z) = uz + z^2 + z^3 + \cdots = uz + \frac{z^2}{1-z} = (u-1)z + \frac{z}{1-z}.$$

Thus, the bivariate GF of $\mathcal{C}$ with $u$ marking $\psi$ is $C(u,z) = 1/(1 - I(u,z))$. The cumulative GF of $\theta$ obtains by differentiation:

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

From the partial fraction expansion, a random composition of $n$ has on average $\sim n/8$ summands equal to 1 (out of a total of $\sim n/2$), a fair amount!

More generally, for a fixed integer $k$, let $\psi$ be the number of summands having value $k$. We find

$$C(u,z) = \frac{1}{1-I(u,z)} \quad \text{where} \quad I(u,z) = \frac{z}{1-z} + (u-1)z^k.$$

Thus,

$$\Theta(z) = \frac{z^k}{(1-\frac{z}{1-z})^2}.$$

There results, by a detailed analysis of the partial fraction decomposition, that the average number of summands equal to $k$ is asymptotically (for fixed $k$ and $n \to +\infty$)

$$\sim \frac{n}{2^{k+2}}.$$

**Permutations.** In a similar way, let $\mathcal{P}$ be the class of all permutations constructed as sets of cycles ($\mathcal{K}$). Let $\psi$ represent the number of cycles of a fixed size $k$ in a permutation; $\psi$ is defined additively from a characteristic parameter $\theta$ of cycles whose value equals 1 only on cycles of size $k$. The GF of $\mathcal{K}$ with $u$ marking $\theta$ is

$$K(u, z) = K(z) - \frac{z^k}{k} + u\frac{z^k}{k} = K(z) + (u-1)\frac{z^k}{k}$$

(take all terms of $K$ except those corresponding to size $k$ for which a factor of $u$ is inserted).

From Theorem 3.3, the GF of $\mathcal{P}$ with $u$ marking $k$–cycles is $P(u, z) = e^{K(u,z)}$, which using the form of $K(z)$ simplifies to

$$P(u, z) = \frac{e^{(u-1)z^k/k}}{1-z}.$$

The cumulative GF of $\psi$ over $\mathcal{P}$ is thus plainly

$$\Psi(z) = \frac{1}{k}\frac{z^k}{1-z}.$$

In other words, *a random permutation of size $n \geq k$ has on average $\frac{1}{k}$ cycles of length $k$*. This result constitutes a refinement of the property that the mean number of cycles is the harmonic number $H_n = \sum_{k=1}^{n} \frac{1}{k}$, so that each term in the harmonic sum receives a direct interpretation.

**Set partitions.** A general pattern emerges from the previous two examples, and after a while it is no longer necessary to specify parameters fully, as bivariate generating functions may be written down "at sight". For instance the GF of set partitions with $u$ marking the number of blocks of size $k$ is

$$\exp\left(e^z - 1 + (u-1)\frac{z^k}{k!}\right),$$

with corresponding cumulative GF

$$\frac{z^k}{k!} e^{e^z - 1}.$$

Thus, the mean number of $k$–blocks is

$$\binom{n}{k} \frac{S_{n-k}}{S_n},$$

with $S_n$ the Bell number. Asymptotically (for fixed $k$ and $n \to +\infty$), this is found to be

$$\sim \frac{1}{k!} \left( \log \frac{n}{\log n} \right)^k.$$

Exercise 4. Throw $n$ balls into $m$ urns at random ($m$ fixed). The bivariate EGF for the number of urns containing $k$ components is

$$\left( e^z + (u-1) \frac{z^k}{k!} \right)^m.$$

Let $m$ and $n$ tend to infinity in such a way that $\frac{n}{m} = \alpha$, a fixed constant. The proportion of urns containing $k$ elements approaches the Poisson law with distribution

$$e^{-\alpha} \frac{\alpha^k}{k!}.$$

Exercise 5. Find formulæ for the expected cardinality of the image of a random surjection in $\mathcal{R}_n$, and the number of values that are taken $k$ times.

**Trees.** This example illustrates the use of multivariate generating functions for a recursive parameter over a recursively defined structures.

A leaf of a tree is a node without descendents. Consider again the class $\mathcal{G}$ of plane unlabelled trees, $\mathcal{G} = \mathcal{N} \times \mathfrak{S}\{\mathcal{G}\}$. The number $G_{n,k}$ of trees with $n$ nodes and $k$ leaves is to be determined.

Let $\theta$ be the parameter "number of leaves" and $G(u, z)$ the associated bivariate OGF. In order to individuate leaves, rewrite the original specification of plane trees as

$$\mathcal{G} = \mathcal{N} + (\mathcal{N} \times \mathcal{G} \times \mathfrak{S}\{\mathcal{G}\}).$$

The parameter $\theta$ is additive, so that to the defining relation, there corresponds termwise

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

In passing, the induced quadratic equation can be solved explicitly

$$G(u,z) = \frac{1}{2}\left(1+(u-1)z - \sqrt{1-2(u+1)z+(u-1)^2z^2}\right).$$

It is however simpler to expand using the Lagrange inversion theorem which provides

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

The mean number of leaves derives from the cumulative GF,

$$\frac{1}{2}z + \frac{1}{2}\frac{z}{\sqrt{1-4z}},$$

so that the mean is $n/2$ exactly for $n \geq 2$.

In the same vein, for Cayley trees, the bivariate EGF is the solution to

$$Y(u,z) = uz + z(e^{Y(u,z)} - 1).$$

EXERCISE 6. The mean number of nodes of degree $k$ in a tree of $\mathcal{G}_n$ is asymptotic to

$$\frac{n}{2^{k+1}},$$

which is proportional to a geometric law of parameter $1/2$.

EXERCISE 7. Determine the statistics of the number of leaves in Cayley trees and relate it to Stirling partition numbers Show that the mean number of leaves is asymptotic to $e^{-1}n$.

Show that the mean number of nodes of degree $k$ is asymptotic to

$$n \cdot e^{-1}\frac{1}{k!}.$$

The number of node of degree $k$ in a large Cayley tree is proportional to a Poisson law of rate 1.

## 3.4 Path length in trees.

This example is a little less direct than the preceding ones. It represents a standard paradigm in the analysis of algorithms, though. Instances are the evaluation of quicksort, of path length in binary search trees or digital trees.

Path length on trees is classically defined by the recursion

$$\lambda[t] = \begin{cases} 1 & \text{for } |t| = 1 \\ |t| + \lambda[t_1] + \cdots + \lambda[t_\ell] & \text{for } |t| > 1 \text{ and } t_j\text{'s the root subtrees of } t. \end{cases}$$

The parameter $\lambda[t]$ also equals the sum of distances of all nodes to the root of $t$ (with distance measured by the number of edges on a connecting branch).

Consider first in all generality two parameters on elements of $\mathcal{G}$ $\lambda[t]$ and $\mu[t]$ that are connected by a relation

$$\lambda[t] = |t| + \mu[t] \qquad \text{so that} \qquad u^{\lambda[t]} = u^{|t|} \cdot u^{\mu[t]}. \tag{3.1}$$

Let $G(u,z)$ and $H(u,z)$ be the bivariates GF's of $\mathcal{G}$ with respect to $\lambda$ and $\mu$. By the combinatorial form of bivariate GF's, such a general relation implies

$$G(u,z) = H(u,uz), \tag{3.2}$$

since an extra factor of $u^n$ appears as a multiplier of $z^n$ when forming $G(u,z)$.

Now specialize $\mu$ to be

$$\mu[t] = \lambda[t_1] + \cdots + \lambda[t_\ell] \text{for } |t| > 1 \text{ and } t_j\text{'s the root subtrees of } t,$$

with $\lambda[t]$ the path length parameter. From the additive scheme on sequences, we have

$$H(u,z) = \frac{z}{1 - G(u,z)}. \tag{3.3}$$

The comparison of (3.2) and (3.3) shows that $G(u,z)$, which is the GF of interest, satisfies the nonlinear "difference equation"

$$G(u,z) = \frac{zu}{1 - G(u,uz)}. \tag{3.4}$$

The generating function $\Lambda(z)$ of mean values of $\lambda$ then obtains by differentiation with respect to $u$ and setting $u = 1$. We have

$$\frac{\partial}{\partial u} G(u,z) = \frac{z}{(1 - G(u,z))^2} \left[ \left.\frac{\partial G(u,z)}{\partial u}\right|_{(u,uz)} + z \left.\frac{\partial G(u,z)}{\partial z}\right|_{(u,uz)} \right],$$

which by specialization to $u = 1$ gives

$$\Lambda(z) = \frac{z}{(1 - G(z))^2} \left[\Lambda(z) + zG'(z)\right].$$

Solving the linear equation for $\Lambda(z)$ yields after simplifications

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

The total path length of all trees of size $n$ is thus

$$\Lambda_n = \frac{1}{2}4^n + \frac{1}{2}\binom{2n-2}{n-1},$$

which yields a normalized mean asymptotic to $\frac{1}{2}\sqrt{\pi n}$.

The intricacy of this problem arises from a combination of a recursive rule and a shift by $|t|$ in the definition (3.1) of path length $\lambda$. In an abstract sense the shift by $|t| = n$ is reflected by the transformation $z \mapsto uz$ in the bivariate GF.

EXERCISE 8. The bivariate EGF of path length in Cayley trees satisfies the functional equation

$$T(u, z) = zue^{T(u,uz)}.$$

Find the mean value of path length in such trees.

## 3.5 Combinatorial schemas

The discussion of additive parameters and its accompanying examples permits us to develop a highler level view of the counting of parameters in constructible classes.

**Marks.** Additive parameters can be viewed a parameters that are "*inherited*" through constructions. In most cases they record the occurrence of certain designated configurations, like components or substructures of a given size. The general result of Theorem 3.3 gives us a way of computing the bivariate GF as follows:

> *A variable $u$ is to be introduced in generating function equations at the place where configurations are to be recorded. Two common schemes for a class with GF $f(z)$ are*
>
> $$uf(z) \quad \text{and} \quad f(z) + (u-1)f_k z^k \qquad (f_k = [z^k]f(z)),$$
>
> *corresponding to the occurrence of $\mathcal{F}$ structures and to $\mathcal{F}_k$ structures respectively. Then the BGF is determined by applying "upwards" the general translation mechanisms granted by the symbolic method.*

This provides a simple intuition regarding the placement of the auxiliary variable $u$ which thus appears as a *formal marker* of designated configurations in specifications.

For instance, reconsider the class $\mathcal{F}$ of finite mappings discussed in the chapter on EGF's:

$$\mathcal{F} = \mathfrak{P}\{\mathcal{K}\}, \quad \mathcal{K} = \mathfrak{C}\{\mathcal{T}\}, \quad \mathcal{T} = \{1\} \star \mathfrak{P}\{\mathcal{T}\}.$$

The translation on EGF's is

$$F(z) = e^{K(z)}, \quad K(z) = \log \frac{1}{1-T(z)}, \quad T(z) = e^{T(z)}.$$

Here are bivariate EGF's for $(i)$ the number of components, $(ii)$ the number of trees, $(iii)$ the number of leaves:

$$(i)\ e^{uK(z)}, \qquad (ii)\ \exp(\log(1-uT(z))^{-1} = \frac{1}{1-uT(z)},$$

$$(iii)\ \exp(\log(1-T(u,z))^{-1} = \frac{1}{1-T(u,z)} \ \text{with}\ T(u,z) = (u-1)z + ze^{T(u,z)}.$$

**Multivariate generating functions.** It is also possible to keep track *simultaneously* of several parameters of combinatorial structures. Formally, this corresponds to defining additive parameters over $\mathbb{N}^r$ instead of $\mathbb{N}$. In that case, one should introduce a collection of $r$ different variables $\mathbf{u} = (u_1, u_2, \ldots, u_r)$. The previous theory of additive parameters immediately generalizes as does the intuition about the interpretation of auxiliary variables as markers of designated configurations. It would be too heavy (and uninspiring!) to develop a complete theory, and a few example should suffice in order to understand what goes on.

For instance the trivariate EGF $F(u_1, u_2, z)$ of functional graphs with $u_1$ marking components and $u_2$ marking trees is

$$F(u_1, u_2, z) = \exp(u_1 \log(1 - u_2 T(z))^{-1}) = \frac{1}{(1 - u_2 T(z))^{u_1}}.$$

EXERCISE 9. Find an explicit expression for the coefficients of the trivariate $F$ involving Stirling cycle numbers.

Some sorts of "universal" generating functions may then be formed that simultaneously record a large number of characteristics of combinatorial objects. For instance, the EGF of permutations with $u_j$ marking the number of cycles of size $j$ is

$$\exp\left(u_1 \frac{z}{1} + u_2 \frac{z^2}{2} + u_3 \frac{z^3}{3} + \cdots\right);$$

the EGF of set partitions with $u_j$ marking the number of blocks cycles of size $j$ is

$$\exp\left(u_1 \frac{z}{1!} + u_2 \frac{z^2}{2!} + u_3 \frac{z^3}{3!} + \cdots\right);$$

the OGF of compositions with $u_j$ marking the number of summands of value $j$ is

$$\frac{1}{(1 - u_1 z - u_2 z^2 - u_3 z^3 - \cdots)}.$$

(In these cases, we even have formal generating functions in infinitely variables.)

Generating functions like this are sometimes simpler to expand. For instance, the number of permutations with $n_1$ cycles of size 1, $n_2$ of size 2, etc, is

$$\frac{n!}{c_1!\, c_2! \cdots c_n!\; 1^{c_1}\, 2^{c_2} \cdots n^{c_n}},$$

provided $\sum j c_j = n$. This is a result originally due to Cauchy.

EXERCISE 10. Find explicit expressions for the number of set partitions with $n_1$ blocks of size 1, $n_2$ of size 2, etc. Proceed similarly with compositions.

Interpret the EGF

$$(1 - \sum_{n=1}^{\infty} x_n \frac{z^n}{n!})^{-1}$$

in terms of surjections, and

$$(1 - \sum_{n=1}^{\infty} x_n n^n \frac{z^n}{n!})^{-1},$$

in terms of finite mappings.

Several examples of such "universal" generating functions are presented in Comtet's book, see [3], pages 225 and 233.

## 3.6 Extremal parameters

Apart from additive parameters already examined at length, another important category is that of parameters defined by a maximum rule. Two major cases are the largest component in a combinatorial structure (for instance, the largest cycle of a permutation) and the maximum degree of nesting of certain constructions in a recursive structure (typically, the height parameter of a tree). In this case, bivariate generating functions are of no help, and the standard technique consists in introducing a collection of univariate generating functions that are defined by imposing a bound on the parameter of interest. Such GF's can then be constructed by the symbolic method.

**Largest components.** Consider a construction $\mathcal{B} = \Phi\{\mathcal{A}\}$, where $\Phi$ may involve an arbitary combination of basic constructions, and assume here for simplicity that the construction for $\mathcal{B}$ is a non–recursive one. This corresponds to a relation between generating functions

$$B(z) = \Psi[A(z)],$$

where $\Psi$ is the functional that is the "image" of the combinatorial construction $\Phi$. Elements of $\mathcal{A}$ thus appear as components in an object $\beta \in \mathcal{B}$. Let $\mathcal{B}^{\langle b \rangle}$ denote the subclass of $\mathcal{B}$ formed with objects whose $\mathcal{A}$–components all have a size at most $b$. The GF of $\mathcal{B}^{\langle b \rangle}$ is obtained by the same process as that of $\mathcal{B}$ itself, safe that $A(z)$ should be replaced by the GF of elements of size at most $b$. Thus,

$$B^{\langle b \rangle}(z) = \Psi[\mathbf{T}_b A(z)],$$

where the *truncation operator* is defined on series by

$$\mathbf{T}_b f(z) = \sum_{n=0}^{b} f_n z^n \qquad (f(z) = \sum_{n=0}^{\infty} f_n z^n).$$

Some cases of this situation have already been encountered in passing in earlier chapters.

In this way, the cycle decomposition of permutations,

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

gives the EGF of permutations with largest cycle $\leq b$:

$$P^{\langle b \rangle}(z) = \exp\left(\frac{z}{1} + \frac{z^2}{2} + \cdots + \frac{z^b}{b}\right),$$

which involves the truncated logarithm.

Similarly, the EGF of words over an $m$–ary alphabet

$$W(z) = (e^z)^m$$

leads to the EGF of words such that each letter occurs at most $b$ times:

$$W^{\langle b \rangle}(z) = \left(1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^b}{b!}\right)^m,$$

which now involves the truncated exponential. One finds similarly the EGF of set partitions with largest block of size at most $b$,

$$S^{\langle b \rangle}(z) = \exp\left(\frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^b}{b!}\right).$$

A slightly less direct example is that of the longest run in a sequence of binary drawings. The collection $\mathcal{W}$ of binary strings over the alphabet $\{x, y\}$ admits the decomposition

$$\mathcal{W} = \mathfrak{S}\{y\} \cdot \mathfrak{S}\{x\,\mathfrak{S}\{y\}\},$$

corresponding to a "scansion " dictated by the occurrences of the letter $x$. The corresponding OGF then appears under the form

$$W(z) = Y(z) \cdot \frac{1}{1 - zY(z)} \qquad \text{where } Y(z) = \frac{1}{1-z}$$

corresponds to $\mathcal{Y} = \mathfrak{S}\{y\}$. Thus, the OGF of strings with at most $b$ consecutive occurrences of the letter $y$ obtains upon replacing $Y(z)$ by its truncation:

$$W^{\langle b \rangle}(z) = Y^{\langle b \rangle}(z) \frac{1}{1 - zY^{\langle b \rangle}(z)} \quad \text{where} \quad Y^{\langle b \rangle}(z) = 1 + z + z^2 + \cdots + z^b,$$

so that

$$W^{\langle b\rangle}(z)=\frac{1-z^{b+1}}{1-2z+z^{b+2}}.$$

Such generating functions are thus easy to derive. The asymptotic analysis of their coefficients is however often harder, as few general complex analytic properties of the truncation operator seem to have been found (or even investigated!) yet.

EXERCISE 11. The EGF of permutations with smallest cycle of size $> b$ is

$$\frac{\exp(-\frac{z}{1}-\frac{z^2}{2}-\frac{z^b}{b})}{1-z}.$$

Develop a theory of *smallest* components in combinatorial structures.

**Heights.** The degree of nesting of a recursive construction is a generalization of the notion of height in the simpler case of trees. Consider for instance a recursively defined class

$$\mathcal{B}=\Phi\{\mathcal{B}\},$$

where $\Phi$ is a construction. Let $\mathcal{B}^{[h]}$ denote the subclass of $\mathcal{B}$ composed solely of elements whose construction involves at most $h$ applications of $\Phi$. We have by definition

$$\mathcal{B}^{[h+1]}=\Phi\{\mathcal{B}^{[h]}\}.$$

Thus, with $\Psi$ the image functional of construction $\Phi$, the corresponding GF's are defined by a *recurrence*,

$$B^{[h+1]}=\Psi[B^{[h]}].$$

It is usually convenient to start the recurrence with the initial condition $B^{[-1]}(z)=0$.

Consider for instance general plane trees defined by

$$\mathcal{G}=\mathcal{N}\times\mathfrak{S}\{\mathcal{G}\}\qquad\text{so that}\qquad G(z)=\frac{z}{1-z}.$$

Define the height of a tree as the number of nodes on its longest branch. Then the set of trees of height $\leq h$ satisfies the recurrence

$$\mathcal{G}^{[0]}=\mathcal{N},\quad \mathcal{G}^{[h+1]}=\mathcal{N}\times\mathfrak{S}\{\mathcal{G}^{[h]}\}.$$

Accordingly, the OGF of trees of bounded height satisfies

$$G^{[-1]}(z) = 0,\ G^{[0]}(z) = z,\ G^{[h+1]}(z) = \frac{z}{1 - G^{[h]}(z)}.$$

The recurrence unwinds and one finds

$$G^{[h]}(z) = \cfrac{z}{1 - \cfrac{z}{1 - \cfrac{z}{\ddots \over 1 - z}}},$$

where the number of stages in the fraction equals $b$. This is the finite form (technically known as a "convergent") of a *continued fraction* expansion. From implied linear recurrences and an analysis based on Mellin transforms, De Bruijn, Knuth, and Rice have worked out that the average height of a general plane tree is $\sim \sqrt{\pi n}$.

For plane binary trees defined by

$$\mathcal{B} = \mathcal{N} + \mathcal{B} \times \mathcal{B} \qquad \text{so that} \qquad B(z) = z + (B(z))^2,$$

the recurrence is

$$B^{[0]}(z) = z, \quad B^{[h+1]}(z) = z + (B^{[h]}(z))^2.$$

In this case, the $B^{[h]}$ are related to a "continuous quadratic form",

$$B^{[h]}(z) = z + (z + (z + (\cdots)^2)^2)^2.$$

These are polynomials of degree $2^h$ for which no closed form expression is known, nor even likely to exist. However, using complex asymptotic methods and singularity analysis, Flajolet and Odlyzko have shown that the average height of a binary plane tree is $\sim 2\sqrt{\pi n}$.

For Cayley trees, finally, the defining equation is

$$\mathcal{T} = \{1\} \star \mathfrak{P}\{\mathcal{T}\} \qquad \text{so that} \qquad T(z) = ze^{T(z)}.$$

The EGF of trees of bounded height satisfy the recurrence

$$T^{[0]}(z) = z, \quad T^{[h+1]}(z) = ze^{T^{[h]}(z)}.$$

We are now confronted with a "continuous exponential",

$$T^{[h]}(z) = ze^{ze^{ze^{\cdot^{\cdot^{\cdot} ze^{z}}}}}.$$

The average height was found by Renyi and Szekeres who appealed again to complex asymptotics and found it to be $\sim \sqrt{2\pi n}$.

These examples show that height statistics are closely related to iteration theory. Except in a few cases like general plane trees, normally no algebra is available and the only resort is to complex analytic methods.

**Averages.** For extremal parameters, the GF of mean values obey a general pattern. Let $\mathcal{F}$ be some combinatorial class with GF $f(z)$. Consider for instance height with $f^{[h]}(z)$ the GF of objects of height *at most* $h$. The GF of objects of height *exactly* $h$ obtained by differencing is

$$f^{[h]}(z) - f^{[h-1]}(z).$$

These differences describe the probability distribution of height over $\mathcal{F}$.

The generating function of unnormalized mean values is then

$$\begin{aligned} \Xi(z) &= \sum_{h=0}^{\infty} h\left[f^{[h]}(z) - f^{[h-1]}(z)\right] \\ &= \sum_{h=0}^{\infty}\left[f(z) - f^{[h]}(z)\right], \end{aligned}$$

as is readily checked by rearranging the second sum (or equivalently using summation by parts). The same computation applies to largest components as well.

For height, this involves in all generality the differences between the fixed point of a functional $\Phi$ (the GF $f(z)$) and the iterative approximations $(f^{[h]}(z)$ to the fixed point. This is a common scheme in extremal statistics.

## 3.7 Notes

Multivariate generating functions are a common tool from classical combinatorial analysis. Comtet's book [3] is again an excellent source of examples. A systematization of multivariate generating functions for additive parameters is given in the book by Jackson and Goulden [11].

In contrast generating functions for averages seemed to have received relatively little attention before the advent of digital computers and the analysis of algorithms. Many important techniques are implicit in Knuth's books, especially [12, 13]. Wilf discusses related issues in his book [17] and in [16]. Early systems specialized to tree algorithms have been proposed by

Flajolet and Steyaert [6, 15, 9]; see also Berstel and Reutenauer's work [1]. Some of the ideas developed there (viewing generating functions of averages as images of combinatorial structures with multiplicities attached) took their inspiration from the well established treatment of formal power series in non-commutative indeterminates (that can be seen as words with multiplicities attached), see Eilenberg's book [5] or [14].

The global framework of constructible structures affords a clear discussion of the various categories of structural parameters on combinatorial objects —additive parameters, marks, largest component, and height. This becomes even clearer when examined in the light of asymptotic properties of structures.

The ideas developed here correspond to an extension of the symbolic framework in which well characterized classes of parameters can also be systematically determined. The additive parameters include time complexity measures for a closed class of programmes. Several such computations can then even be automated with the help of computer algebra systems. A unified approach is developed by Flajolet, Salvy and Zimmermann in [8] and further explored in [18].

## Problems and exercises

In graph theory purely divergent bivariate generating functions tend to show up.

Exercise 12. The EGF of connected labelled graphs with $u$ marking the number of edges is

$$\log\left(1+\sum_{n=1}^{\infty}(1+u)^{n(n-1)/2}\frac{z^n}{n!}\right).$$

Various parameters may be analyzed by suitably introducing marks in bivariate GF's.

Exercise 13. Take the number of *distinct* block sizes and cycle sizes in partitions and permutations. The bivariate EGF's are

$$\prod_{n=1}^{\infty}\left(1-u+ue^{z^n/n!}\right),\quad \prod_{n=1}^{\infty}\left(1-u+ue^{z^n/n}\right).$$

Find a comparable OGF for the number of distinct summands in an integer partition.

The OGF for compositions all of whose summands are distinct is

$$\int_0^{\infty} e^{-x}\left[\prod_{n=1}^{\infty}(1+xz^n)\right]dx.$$

Symmetric functions may be manipulated by mechanisms that are often reminiscent of the set and multiset construction. They appear in many areas of combinatorial enumeration.

Exercise 14. Let $X=\{x_i\}_{i=1}^{r}$ be a collection of formal variables. Define the symmetric functions

$$\begin{aligned}
\prod_i(1+x_iz) &= \sum_n a_nz^n\\
\prod_i\frac{1}{1-x_iz} &= \sum_n b_nz^n\\
\sum_i\frac{x_iz}{1-x_iz} &= \sum_n c_nz^n.
\end{aligned}$$

Show that the $a_n, b_n, c_n$ are symmetric functions,

$$\begin{aligned} a_n &= \sum_{i_1<i_2<\cdots<i_r} x_{i_1}x_{i_2}\cdots x_{i_r} \\ b_n &= \sum_{i_1\le i_2\le\cdots\le i_r} x_{i_1}x_{i_2}\cdots x_{i_r} \\ c_n &= \sum_{i=1}^{r} x_i^r. \end{aligned}$$

Establish the relations

$$B(z) = \frac{1}{A(-z)}, \quad A(z) = \frac{1}{B(-z)},$$

$$C(z) = z\frac{d}{dz}\log B(z), \quad B(z) = \exp\int_0^z C(t)\,\frac{dt}{t}.$$

Deduce that each of $a_n, b_n, c_n$ is polynomially expressible in terms of any of the other. Make explicit the connection coefficients in each case.

Exercise 15. A graph is $r$–regular iff each node has degree exactly equal to $r$. The number of $r$–regular graphs of size $n$ is

$$[x_1^r x_2^r \cdots x_n^r] \prod_{1\le i<j\le n} (1 + x_i x_j).$$

[Note: Gessel [10] showed how to extract explicit expressions from such huge symmetric functions.]

Balanced structures lead to counting GF's close to the ones obtained for height statistics.

Exercise 16. The OGF of balanced 2-3 trees of height $h$ counted by the number of leaves satisfies the recurrence

$$Z^{[h+1]}(z) = Z^{[h]}(z^2 + z^3) = (Z^{[h]}(z))^2 + (Z^{[h]}(z))^3.$$

Express it in terms of the iterates of $\sigma(z) = z^2 + z^3$.

Find the OGF of mean values of the number of internal nodes in such trees.

EXERCISE 17. [Hierarchical partitions] Let $\varepsilon(z) = e^z - 1$. Find a combinatorial interpretation for

$$\varepsilon(\varepsilon(\cdots(\varepsilon(z)))) \qquad (h \text{ times}).$$

(Such structures show up in statistical classification theory.)

A universal continued fraction arises as the universal GF of trees counted according to the level of nodes.

EXERCISE 18. The GF of general plane trees with $u_j$ marking the number of nodes at level $j$ is

$$G(\mathbf{u}, z) = \cfrac{u_0 z}{1 - \cfrac{u_1 z}{1 - \cfrac{u_2 z}{1 - \cfrac{u_3 z}{\ddots}}}}.$$

Find explicit expressions for the coefficients.

Do the same for binary trees and Cayley trees.

The model of height in general plane trees can be solved algebraically.

EXERCISE 19. Show that the OGF $G^{[h]}(z)$ of trees of height $\leq h$ is of the form

$$z\frac{F_{h+1}(z)}{F_{h+2}(z)},$$

where the $F$'s are the Fibonacci polynomials

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

Express the $F_h$ in terms of $G(z)$ itself. Find explicit forms for the distribution of height in trees of $\mathcal{G}_n$ by means of Lagrange inversion.

[Hint: this is due to De Bruijn, Knuth, and Rice [4].]

Extremal statistics can also be obtained for components of complex structures.

EXERCISE 20. Find the EGF's for the largest cycle, longest branch, and diameter of functional graphs. Do the same for the largest tree, largest component. [Hint: see [7] for details.]

EXERCISE 21. Find the GF of mean values of the number of nodes at maximal depth in a general plane tree and in a Cayley tree.

# Bibliography


[1] Berstel, J., and Reutenauer, C. Recognizable formal power series on trees. *Theoretical Computer Science 18* (1982), 115–148.

[2] Billingsley, P. *Probability and Measure*, 2nd ed. John Wiley & Sons, 1986.

[3] Comtet, L. *Advanced Combinatorics.* Reidel, Dordrecht, 1974.

[4] De Bruijn, N. G., Knuth, D. E., and Rice, S. O. The average height of planted plane trees. In *Graph Theory and Computing* (1972), R. C. Read, Ed., Academic Press, pp. 15–22.

[5] Eilenberg, S. *Automata, Languages, and Machines,* Volume A. Academic Press, 1974.

[6] Flajolet, P. *Analyse d'algorithmes de manipulation d'arbres et de fichiers*, vol. 34–35 of *Cahiers du Bureau Universitaire de Recherche Opérationnelle.* Université Pierre et Marie Curie, Paris, 1981. 209 pages.

[7] Flajolet, P., and Odlyzko, A. M. Random mapping statistics. In *Advances in Cryptology* (1990), J.-J. Quisquater and J. Vandewalle, Eds., vol. 434 of *Lecture Notes in Computer Science*, Springer Verlag, pp. 329–354. Proceedings of Eurocrypt'89, Houtalen, Belgium, April 1989.

[8] Flajolet, P., Salvy, B., and Zimmermann, P. Automatic average–case analysis of algorithms. *Theoretical Computer Science, Series A 79*, 1 (Feb. 1991), 37–109.

[9] Flajolet, P., and Steyaert, J.-M. A complexity calculus for recursive tree algorithms. *Mathematical Systems Theory 19* (1987), 301–331.

[10] Gessel, I. M. Symmetric functions and $P$–recursiveness. *Journal of Combinatorial Theory,* Series A *53* (1990), 257–285.

[11] Goulden, I. P., and Jackson, D. M. *Combinatorial Enumeration.* John Wiley, New York, 1983.

[12] Knuth, D. E. *The Art of Computer Programming*, vol. 1: Fundamental Algorithms. Addison-Wesley, 1968. Second edition, 1973.

[13] Knuth, D. E. *The Art of Computer Programming*, vol. 3: Sorting and Searching. Addison-Wesley, 1973.

[14] Salomaa, A., and Soittola, M. *Automata-Theoretic Aspects of Formal Power Series.* Springer, Berlin, 1978.

[15] Steyaert, J.-M. *Structure et complexité des algorithmes.* Doctorat d'état, Université Paris VII, Apr. 1984.

[16] Wilf, H. S. Some examples of combinatorial averaging. *American Mathematical Monthly 92* (Apr. 1985), 250–261.

[17] Wilf, H. S. *Generatingfunctionology.* Academic Press, 1990.

[18] Zimmermann, P. *Séries génératrices et analyse automatique d'algorithmes.* Ph. d. thesis, École Polytechnique, 1991.

# Contents