COMBINATORIAL COUNTING

MARTIN KLAZAR

Contents

Preface

These notes are inspired by the course *Combinatorial Counting* (*Kombinatorické počítání*, NDMI015) which I have been teaching on Faculty of Mathematics and Physics of Charles University in Prague. Their main theme is how to count finite things, precisely or, with less emphasize, asymptotically. Following topics are covered. Chapter 1 revolves around the Catalan numbers. Chapter 2 presents several analytic deductions of the Stirling (asymptotic) formula for factorial.

February 2018 Martin Klazar

Chapter 1

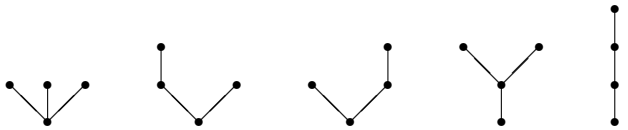
Catalan numbers

We begin this text with the Catalan numbers c_n and results they inspire. In Section 1.1 we define them and find for them a recurrence relation, determine their parity and bound them exponentially both from below and above. Section 1.2 discusses the concept of PIO formulas and PIO algorithms and illustrates it on c_n and the parity-modified Catalan numbers c'_n. In Section 1.3 we derive more recurrences and several explicit formulas and from the Stirling formula we deduce asymptotics of c_n, which we then prove in a weaker form $cn^{-3/2}4^n$. In Section 1.4 we present a bijective proof of the formula $c_n = \frac{1}{n}\binom{2n-2}{n-1}$ and show that c_n count each of the six families of permutations avoiding a fixed pattern of length three. Can c_n satisfy a linear recurrence with constant coefficients? Four arguments why they cannot are given in Section 1.5. A refinement $c_{n,k}$ is treated in Section 1.6. In Section 1.7 we quote an excerpt from the list of R. P. Stanley of problems in enumeration solved by the Catalan numbers, and prove Valtr's theorem: n random points in a square form a convex chain, assuming that they already form a convex polygon, with probability $\frac{1}{c_n}$. Concluding Comments and references contain some comments and references.

1.1 Definition, recurrence, parity, exponential growth

We define the Catalan numbers c_n as numbers of certain trees with n vertices. All trees considered in this chapter are finite. A *rooted tree* $T = (r, V, E)$ is a tree (V, E) with a distinguished vertex $r \in V$, called a *root (of T)*. Recall that a *tree* is a graph (V, E) $(E \subset \binom{V}{2})$ such that every two vertices in V are joined by a unique path. For every edge $e \in E$ in a rooted tree (r, V, E) there is a unique path P whose one end is r and the other e. If $u \in e$ is the vertex of e that is the end of P different from r and $v \in e$ is the other vertex (if $v \neq r$ then v is an inner vertex of P), we write $v \to u$; the presence of root determines this orientation of edges in E away from r. A *rooted plane tree*, abbreviated *rp tree*, is a rooted tree $T = (r, V, E, L)$ enriched with a list L of linear orders on the sets

$\{u \in V \mid v \to u\}$, $v \in V$. We call such set a *set of children (of v)*. Two rp trees are *isomorphic* if there is a bijection between their vertex sets that preserves roots, edges and linear orders of sets of children. Two rp trees are regarded as distinct only if they are non-isomorphic. Here we list all five distinct rp trees with four vertices:



We draw an rp tree in the plane with the root at the lowest position, with edges as upgoing straight segments (arrows), and with the linear orders on sets of children given by the left-to-right order. It is because of these orders that the second and third tree are distinct because as rooted trees they are isomorphic.

Thus there is one rp tree with one vertex, one with two vertices, two with three vertices and five with four vertices. *How many rp trees with n vertices exist?* This is a text on enumerative combinatorics and not a philosophical essay, and we are primarily concerned with the first two words and not so much with the final verb. For $n \in \mathbb{N} = \{1, 2, \dots\}$ we define

the n-th Catalan number c_n = the number of distinct rp trees with n vertices

—recall that "distinct" means "non-isomorphic". Let $\mathcal{T}$ be the set of all distinct rp trees. (Precisely, what are the elements of $\mathcal{T}$? We return to this question at the end of the section.) For $T \in \mathcal{T}$ let $|T|$ denote the number of vertices in T and for $n \in \mathbb{N}$ let $\mathcal{T}_n = \{T \in \mathcal{T} \mid |T| = n\}$. In general we denote by $|X|$ and $\#X$ the cardinality of a set X. We call $|T|$ also the *size of an rp tree T*. Let

$$C = C(x) := \sum_{T \in \mathcal{T}} x^{|T|} = \sum_{n=1}^{\infty} |\mathcal{T}_n| = \sum_{n=1}^{\infty} c_n x^n = x + x^2 + 2x^3 + 5x^4 + \dots$$

be the *generating function* (GF) of c_n. More precisely, this is an *ordinary generating function* (OGF) of the numbers c_n; later we will meet generating functions of other kinds. We want to have an effective formula for the numbers c_n—see the next section for precise definition of this notion—and we will find several nice ones with the help of the GF $C(x)$. We explain why c_n are "Catalan" in the final Comments and references.

The term "generating function" is a bit of a misnomer because primarily GF is a formal (algebraic) object, an infinite sequence like $(c_1, c_2, \dots)$ or a formal power series like $c_1 x + c_2 x^2 + \dots$, which can be operated with without associating to it any function. Of course, handling a GF as a function is often very useful, and for determinig asymptotic behaviour of its coefficients practically indispensable, whence originated the term. It is not apriori clear why replacing a sequence of numbers by a GF should help in determining a formula for them. The reason for the success of GFs lies in mirroring of combinatorial decompositions of structures in algebraic relations for GFs—this text contains many examples.

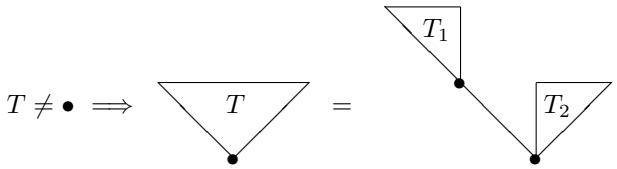
Every non-singleton rp tree decomposes uniquely in two parts as follows.

Proposition 1.1.1. *There is a bijection*

$$f:\ \mathcal{T}\backslash\mathcal{T}_1 \to \mathcal{T}\times\mathcal{T},\ T\mapsto (T_1,T_2)\ ,$$

such that always $|T| = |T_1| + |T_2|$.

Proof. We define f by setting T_1 to be the rp subtree of T rooted in the first child of T's root, and T_2 to be the rest of T:



It is easy to check that f has the stated properties. □

Exercise 1.1.2. *Check it.*

This decomposition of rp trees is the key to their enumeration. Restricting f to $\mathcal{T}_n$ we get for every $n \geq 2$ the bijection

$$\mathcal{T}_n \longleftrightarrow \bigcup_{k=1}^{n-1} \mathcal{T}_k \times \mathcal{T}_{n-k}\ ,$$

which translates to the basic recurrence relation

$$c_1 = 1,\ c_n = \sum_{k=1}^{n-1} c_k c_{n-k} \text{ for } n \geq 2\ . \tag{1.1}$$

For a similar but different recurrence for c_n see Proposition 1.4.11 and Exercise 1.4.12. Recurrence (1.1) readily gives as many Catalan numbers as we wish:

$$(c_n) = (c_n)_{n\geq 1} = (1,\ 1,\ 2,\ 5,\ 14,\ 42,\ 132,\ 429,\ 1430,\ 4862,\ 16796,\ 58786,\ \dots)\ .$$

Without reading further, can you guess when c_n is odd?

Proposition 1.1.3. *The Catalan number c_n is odd if and only if n is a power of 2, $n = 2^m$ for $m \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$.*

Proof. Induction using recurrence (1.1). For $n = 1$ indeed $c_1 = 1$ is odd. If $n > 1$ is even then $c_n = \sum_{k=1}^{n-1} c_k c_{n-k} = c_{n/2}^2 + 2\sum_{k=1}^{n/2-1} c_k c_{n-k}$, and c_n has the same parity as $c_{n/2}$. If $n > 1$ is odd then $c_n = \sum_{k=1}^{n-1} c_k c_{n-k} = 2\sum_{k=1}^{(n-1)/2} c_k c_{n-k}$, and c_n is always even. Thus c_n is odd iff n can be completely divided by 2, i.e., iff $n = 2^m$. □

Symmetry of the sum in recurrence (1.1) reveals parity of c_n easily. It is harder to deduce it from simpler recurrences and formulas for c_n that we derive later.

Exercise 1.1.4. *How do c_n behave modulo 3? (You may use Google.)*

The problem of effective computation of modular reductions of (c_n) and similar sequences of numbers will be treated in Section 4.5.

How fast do Catalan numbers grow? By recurrence (1.1), for $n \geq 3$ we have $c_n \geq 2c_1c_{n-1} = 2c_{n-1}$. Since $c_7 = 132 > 2^7 = 128$, induction gives

$$c_n > 2^n \quad \text{for every} \quad n \geq 7 \,.$$

It is a little harder to obtain an exponential *upper* bound for c_n directly from recurrence (1.1). We need for it the next lemma.

Lemma 1.1.5. *For every $n = 2, 3, \ldots,$*

$$\sum_{k=1}^{n-1} \frac{1}{k^2(n-k)^2} < \frac{8}{n^2} \,.$$

Proof. As $(k(n-k))^{-1} = n^{-1}(k^{-1} + (n-k)^{-1})$, we have

$$\sum_{k=1}^{n-1} \frac{1}{k^2(n-k)^2} = \frac{2}{n^2} \sum_{k=1}^{n-1} \frac{1}{k^2} + \frac{4}{n^3} \sum_{k=1}^{n-1} \frac{1}{k} < \frac{4}{n^2} + \frac{4}{n^2} = \frac{8}{n^2}$$

because $\sum_{k=1}^{n-1} \frac{1}{k^2} < 1 + \sum_{k=2}^{\infty} \frac{1}{k(k-1)} = 1 + \sum_{k=2}^{\infty} \left(\frac{1}{k-1} - \frac{1}{k} \right) = 2.$ □

Thus c_n grow only exponentially:

Proposition 1.1.6. *For every $n = 1, 2, \ldots,$*

$$c_n \leq \frac{8^{n-1}}{n^2} < 8^n \,.$$

Proof. For $n \in \mathbb{N}$ we seek an upper bound $c_n \leq c\frac{\alpha^n}{n^2}$ where $c > 0$ and $\alpha > 1$ are constants, which are to be determined. Suppose it holds for c_k for every $k < n$, then recurrence (1.1) and the previous lemma for $n \geq 2$ give

$$c_n = \sum_{k=1}^{n-1} c_k c_{n-k} \leq \sum_{k=1}^{n-1} \frac{c\alpha^k}{k^2} \cdot \frac{c\alpha^{n-k}}{(n-k)^2} < \frac{8c^2\alpha^n}{n^2} \,.$$

For this to be $\leq c\frac{\alpha^n}{n^2}$, we set $c = \frac{1}{8}$. To start induction, we need that $c_n \leq c\frac{\alpha^n}{n^2} = \frac{\alpha^n}{8n^2}$ holds for $n = 1$ and thus set $\alpha = 8$. So $c_n \leq \frac{\alpha^n}{8n^2} = \frac{8^{n-1}}{n^2}$ for every $n \in \mathbb{N}$. □

Using minimalistic tools (just induction and recurrence (1.1)), we derived the bounds

$$2^n < c_n < 8^n \tag{1.2}$$

where the first inequality holds for $n \geq 7$ and the second for every $n \geq 1$. We give better bounds in inequalities (1.9), and precise asymptotics in Corollaries 1.3.8 and 1.3.10. We extend the bound of Proposition 1.1.6 to more general recurrences in Proposition 4.6.1 in Chapter 4.

Exercise 1.1.7. *Can you decrease the constant 8? Or increase the 2? By how much?*

What objects exactly the numbers c_n count? What exactly are the elements of the set $\mathcal{T}$ of all distinct rp trees? We intuitively view them as blocks of mutually isomorphic rp trees, but are not these blocks as sets unnecessarily big? Could not we take from each some nice representative? This is what we do now. For $n \in \mathbb{N}$ and $[n] = \{1, 2, \ldots, n\}$ we call a rooted tree

$$(1, [n], E),\ E \subset \binom{[n]}{2},$$

a *natural tree* if

- $v \to u$ implies $v < u$ for every edge $\{u, v\} \in E$ and
- each set of children $\{u \in [n] \mid v \to u\}$, $v \in [n]$, is an interval in $[n]$.

Every natural tree $(1, [n], E)$ is obviously also an rp tree: the sets of children are ordered by the natural order $<$ of $\mathbb{N}$. For example, if $n = 4$ we can cast the five rp trees of page 2, respectively, as natural:

$$\begin{aligned}
&(1,\ [4],\ \{\{1,2\},\ \{1,3\},\ \{1,4\}\})\ ,\\
&(1,\ [4],\ \{\{1,2\},\ \{1,3\},\ \{2,4\}\})\ ,\\
&(1,\ [4],\ \{\{1,2\},\ \{1,3\},\ \{3,4\}\})\ ,\\
&(1,\ [4],\ \{\{1,2\},\ \{2,3\},\ \{2,4\}\})\ \text{ and}\\
&(1,\ [4],\ \{\{1,2\},\ \{2,3\},\ \{3,4\}\})\ .
\end{aligned}$$

We revisit the definition of c_n and give it a simpler and concrete form:

$$\textit{the } n\textit{-th Catalan number } c_n = \textit{the number of natural trees } (1, [n], E)\ .$$

Of course, we have to prove equivalence of both definitions of Catalan numbers.

Proposition 1.1.8. *Every rp tree with n vertices is isomorphic to exactly one natural tree $(1, [n], E)$.*

Proof. Suppose $U = (r, V, E, L)$ is an rp tree. We merge the linear orders in L in the single linear order $(V, \prec)$ of the breadth-first search in U. If $V_i \subset V$, $i = 0, 1, \ldots$, are vertices with distance i from r, so $V_0 = \{r\}$, V_1 are the children of r and so on, we put $V_0 \prec V_1 \prec V_2 \prec \ldots$ and for distinct $u, v \in V_i$ we put $u \prec v$ (resp. $v \prec u$) if the path joining u to r branches to the left (resp. to the right), according to L, from that for v. We label the vertices in V by $1, 2, \ldots, |V|$ according to $\prec$. The resulting tree $T = (1, [|V|], E)$ is clearly a natural tree that is isomorphic, as an rp tree, to U. It remains to show that there are no two different natural trees that are isomorphic as rp trees. Suppose that $(1, [n], E)$ and $(1, [n], F)$ are two natural trees that are isomorphic as rp trees. □

1.2 What is an effective formula? For example, for $n \mapsto c_n$?

In this section we give concrete and general answer. What we mean by an *effective formula* (we mentioned this term in the previous section) for the Catalan numbers $n \mapsto c_n$? Or, more generally, for a *counting function*?

A counting function is any computable function $f : \mathbb{N} \to \mathbb{Z}$.

We think of the values of f as giving cardinalities $f(n) = |S_n|$ of finite sets $(S_n)_{n\geq 1}$ coming up in an enumerative problem, or giving solutions to some other problem in enumeration or number theory. For example, we may have $f(n) = |\mathcal{T}_n| = c_n$. Usually, but not always, already the very statement of the problem gives at once an algorithm, effective or not, computing $f(n)$. It is natural to require f be computable; for non-computable f there is nothing to talk about, at least for enumerative combinatorialist. To continue the example, how the very first algorithm for $n \mapsto c_n$, coming to one's mind immediately after learning the definition $c_n := |\mathcal{T}_n|$, could look like? (Recurrence (1.1) or formula (1.7), of course, but try to imagine that you are not yet indoctrinated, never heard about the Catalan numbers and rooted plane trees, and are learning about them only now.) It might work as follows. For the set

$$S = \{E \mid E \subset [n] \times [n]\}$$

with 2^{n^2} elements, the algorithm first determines the subset $S_0 \subset S$ of the edge sets E of rp trees with vertex set $[n]$ and root 1 (the sets of children are ordered by the natural order on $[n]$). For example, if $n = 4$ then S_0 contains $\{(1,2), (1,3), (2,4)\}$ and $\{(1,3), (1,4), (3,2)\}$, both an isomorphic copy of the second rp tree on page 2, but $\{(1,2), (3,4)\} \notin S_0$. Then the algorithm determines (by checking $n!$ bijections from $[n]$ to itself) which rp trees $\{E, E'\} \subset S_0$, $E \neq E'$, are isomorphic. Finally, the algorithm finds the maximum subset $S_1 \subset S_0$ of pairwise non-isomorphic rp trees and outputs $c_n = |S_1|$. Thus $n \mapsto c_n$ is a computable function and so a counting function. The algorithm is not very effective, though, as it does more than 2^{n^2} steps. The interested reader certainly knows a much more effective algorithm and soon we mention it.

But not always a number theoretic or enumerative problem leads to a computable function. We mention two examples, the second one is a big embarassment at the heart of enumerative combinatorics (or rather the theory of algebraic numbers).

Definition 1.2.1 (PIO formula, PIO algorithm). *A counting function*

$$f : \mathbb{N} \to \mathbb{Z}$$

has a PIO formula if there is an algorithm $\mathcal{A}$, called a PIO algorithm, that for every input $n \in \mathbb{N}$ computes the output $f(n)$

$$\textit{in time polynomial in } m(n) = m_f(n) := \log(1+n) + \log(2+|f(n)|)\ .$$

That is to say, for some constants $c, d \in \mathbb{N}$, for every $n \in \mathbb{N}$ the algorithm $\mathcal{A}$ computes $f(n)$ in at most $c \cdot m(n)^d$ steps. Similarly, if f is defined only on a subset $X \subset \mathbb{N}$.

1.3 Recurrences, explicit formulae, asymptotics

We employ the GF $C(x)$ to derive for c_n a recurrence relation simpler than (1.1), and then two explicit formulae. Recurrence (1.1) translates in an algebraic relation for $C = C(x)$:

$$C - x = C^2, \quad \text{or} \quad C^2 - C + x = 0 \,. \tag{1.3}$$

Thus $C \in \mathbb{C}[[x]]$ is algebraic over the field of rational functions $\mathbb{C}(x)$ because it satisfies a quadratic equation with coefficients in $\mathbb{C}[x]$. Note that one can bypass basic recurrence (1.1) and derive equation (1.3) directly from the combinatorial decomposition of rp trees: since

$$C = \sum_{T \in \mathcal{T}} x^{|T|} \,,$$

$C - x = C^2$ is a direct translation of the bijection in Proposition 1.1.1.

Exercise 1.3.1. *What algebraic relation for $C(x)$ you get by using another decomposition of rp trees: $T \mapsto (T_1, T_2, \ldots, T_k)$ where T_i is the rp subtree of T rooted in the i-th child of T's root.*

Writting equation (1.3) as $C = \frac{x}{1-C}$, we expand C in a continued fraction:

$$C = \frac{x}{1-C} = \frac{x}{1-\dfrac{x}{1-C}} = \cdots = \cfrac{x}{1-\cfrac{x}{1-\cfrac{x}{1-\cfrac{x}{\ddots}}}} \,.$$

We deduce from the algebraic equation for C a differential equation. Differentiation of equation (1.3) yields

$$2CC' - C' + 1 = 0, \text{ or } C' = \frac{1}{1-2C} \,. \tag{1.4}$$

Exercise 1.3.2. *Can you solve this differential equation?*

Since C satisfies a quadratic equation, by multiplying both the denominator and the numerator in the last fraction with a linear polynomial in C we get rid of C in the denominator. We get for C a linear differential equation of first order:

$$C' = \frac{1}{1-2C} = \frac{-\frac{C}{2} + \frac{1}{4}}{(1-2C)(-\frac{C}{2} + \frac{1}{4})} = \frac{-\frac{C}{2} + \frac{1}{4}}{C^2 - C + \frac{1}{4}} = \frac{-\frac{C}{2} + \frac{1}{4}}{\frac{1}{4} - x} \,,$$

or

$$(1-4x)C' + 2C - 1 = 0\ . \tag{1.5}$$

In terms of the coefficients in $C = \sum_{n\geq 1} c_n x^n$ we get for $n \geq 2$ by collecting coefficients of x^{n-1} the relation

$$nc_n - 4(n-1)c_{n-1} + 2c_{n-1} = 0, \text{ or } c_n = \frac{4n-6}{n} c_{n-1}\ . \tag{1.6}$$

This is certainly a recurrence simpler than (1.1), and we can solve it:

$$\begin{aligned} c_n &= \prod_{i=2}^{n} \frac{4i-6}{i} \cdot c_1 = \frac{2\cdot 6\cdot 10\cdot \ \ldots\ \cdot (4n-6)}{n!} \\ &= \frac{(n-1)!\cdot 2^{n-1}\cdot 1\cdot 3\cdot 5\cdot \ \ldots\ \cdot (2n-3)}{(n-1)!\cdot n!} = \frac{(2n-2)!}{(n-1)!\cdot n!} \\ &= \frac{1}{n}\binom{2n-2}{n-1} \end{aligned} \tag{1.7}$$

which is the classical explicit formula for the Catalan numbers. Of course, it is "explicit" only so far as we regard factorial $n! = 1\cdot 2\cdot \ldots \cdot n$ as explicit. But we can solve the recurrence also in another way:

$$\begin{aligned} c_n &= \prod_{i=2}^{n} \frac{4i-6}{i} \cdot c_1 = -\frac{1}{2} \prod_{i=0}^{n-1} \frac{4i-2}{i+1} \\ &= \frac{(-1)^{n+1}4^n}{2} \cdot \frac{\frac{1}{2}(\frac{1}{2}-1)\ldots(\frac{1}{2}-n+1)}{n!} \\ &= \frac{(-1)^{n+1}}{2}\binom{\frac{1}{2}}{n} 4^n \end{aligned} \tag{1.8}$$

which is the less known explicit formula for the Catalan numbers. Hence

$$c_n = \frac{1}{n}\binom{2n-2}{n-1} = \frac{(-1)^{n+1}}{2}\binom{\frac{1}{2}}{n} 4^n\ .$$

The advantage of formula (1.8) over formula (1.7) is that it reveals at once the asymptotic order of c_n (however, see Exercise 1.3.4). Indeed, from

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} = \prod_{i=0}^{n-1} \frac{\alpha-i}{i+1}$$

it follows that if $\alpha \in \mathbb{C}$ with $|\alpha| \leq 1$ then $\left|\binom{\alpha}{n}\right| \leq 1$ for every $n \in \mathbb{N}_0$. For $\alpha = \frac{1}{2}$ we can say more precisely that for every $n \in \mathbb{N}$ (empty products are set to 1),

$$\frac{1}{4n^2} \leq \frac{1}{4}\cdot \frac{1\cdot 2\cdot \ldots \cdot (n-2)}{n!} \leq \left|\binom{\frac{1}{2}}{n}\right| \leq \frac{1}{2}\cdot \frac{1\cdot 2\cdot \ldots \cdot (n-1)}{n!} = \frac{1}{2n}\ .$$

Thus by formula (1.8) for every $n \in \mathbb{N}$,

$$\frac{4^n}{8n^2} < c_n \leq \frac{4^n}{4n} \tag{1.9}$$

which is a considerable tightening of bounds (1.2).

Exercise 1.3.3. *How do we exactly get the strict lower bound $<$ from composing the two non-strict $\leq$s above it? Should not one of these be $<$ instead?*

Exercise 1.3.4. *Deduce similar bounds on c_n, with multiplicative gap roughly n between the lower and upper bound, from the formula (1.7).*

The route from the algebraic equation (1.3), via the differential equations (1.4) and (1.5) and the recurrence (1.6), to the formulae (1.7) and (1.8) is a little runaround. By it we wanted to demonstrate that one can deduce from an algebraic equation for GF a recurrence like (1.6) without actually solving the equation; we return to this method later. But we can solve equation (1.3), so let us rederive formula for c_n also this way. By the quadratic formula,

$$C(x) = \frac{1 - \sqrt{1-4x}}{2} \tag{1.10}$$

(we selected $-\sqrt{\dots}$ in $\pm\sqrt{\dots}$ because $C(0) = 0$). But what is $\sqrt{1-4x}$? By Newton's binomial theorem, it is the formal power series

$$(1-4x)^{\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-4)^n x^n \ ,$$

and formula (1.8) follows. It was easy to solve the quadratic equation (1.3) but had we for $C(x)$ a cubic or some higher degree equation, things would be less easy (but not hopeless, as we will see in Section 4.4).

Exercise 1.3.5. *Devise a variant of rp trees, with GF satisfying a cubic or a higher degree algebraic equation.*

Suppose you guessed the formula $c_n = \frac{1}{n}\binom{2n-2}{n-1}$ and want to verify it by the basic recurrence (1.1). This can be done more easily using the alternative form $c_n = \frac{(-1)^{n+1}}{2}\binom{\frac{1}{2}}{n}4^n$. To check it we apply the Vandermonde convolution identity

$$\sum_{k=0}^{a} \binom{\alpha}{k}\binom{\beta}{a-k} = \binom{\alpha+\beta}{a}$$

where α, β are variables and $a \in \mathbb{N}_0$. By combinatorics, it holds for all numbers $\alpha, \beta \in \mathbb{N}$ because then the right side counts a-element subsets $X \subset [\alpha+\beta]$ (for $n \in \mathbb{N}_0$ we set $[n] = \{1, 2, \dots, n\}$ with $[0] = \emptyset$) while the left side just pigeonholes X according to the size k of the intersection $X \cap [\alpha]$. Since we have equality for infinitely many numeric values of α and β, the identity holds formally as an equality between two bivariate polynomials with degree a.

Exercise 1.3.6. *Is this argument correct? In*

$$\alpha^2 + \beta^2 - 1 = 0$$

we also have equality for infinitely many numeric values $\alpha, \beta \in \mathbb{R}$, for coordinates of the points on the unit circle, but the polynomial on the left side is far from being identically zero.

Anyhow, the Vandermonde identity holds and using it for $\alpha = \beta = \frac{1}{2}$ we easily check by the basic recurrence (1.1) and induction on n the explicit formula $c_n = \frac{(-1)^{n+1}}{2}\binom{\frac{1}{2}}{n}4^n$: for $n = 1$ it holds, $1 = c_1 = \frac{(-1)^{1+1}}{2}\binom{\frac{1}{2}}{1}4^1$, and for $n \geq 2$ we indeed have (the Vandermonde identity is used on the third line)

$$\begin{aligned} c_n = \sum_{k=1}^{n-1} c_k c_{n-k} &= \sum_{k=1}^{n-1} \frac{(-1)^{k+1}}{2}\binom{\frac{1}{2}}{k}4^k \cdot \frac{(-1)^{n-k+1}}{2}\binom{\frac{1}{2}}{n-k}4^{n-k} \\ &= \frac{(-1)^n 4^n}{4}\left(\sum_{k=0}^{n}\binom{\frac{1}{2}}{k}\binom{\frac{1}{2}}{n-k} - 2\binom{\frac{1}{2}}{n}\right) \\ &= \frac{(-1)^n 4^n}{4}\left(\binom{1}{n} - 2\binom{\frac{1}{2}}{n}\right) \\ &= \frac{(-1)^{n+1}}{2}\binom{\frac{1}{2}}{n}4^n \end{aligned}$$

as $\binom{1}{n} = 0$ for $n > 1$.

We saw that the GF $C(x)$ of the Catalan numbers satisfies algebraic equation (1.3) and two differential equations (1.4) and (1.5). Does it satisfy some functional equation, equation involving the operation of composition? If we substitute $x - x^2$ for x in $C(x)$, due to the identity

$$1 - 4(x - x^2) = (1 - 2x)^2$$

the right side of formula (1.10) becomes x. So $C(x)$ satisfies the functional equation

$$C(x - x^2) = x\,. \tag{1.11}$$

Interestingly, substitution of $C(x)$ for x in the same polynomial $x - x^2$ gives by equation (1.3) x again:

$$C(x) - C(x)^2 = x\,. \tag{1.12}$$

As we explain in Section 4.3, the pair of identities (1.11) and (1.12) is an instance of a general phenomenon rooted in associativity of composition of formal power series.

What identity for c_n does equation (1.11) imply? Equating for $n \geq 2$ the coefficient of x^n in $\sum_{k\geq 1} c_k(x - x^2)^k = \sum_{k\geq 1} c_k x^k(1 - x)^k$ to zero, we obtain third recurrence for c_n:

$$c_1 = 1,\ c_n = \sum_{k\geq 1}(-1)^{k+1}\binom{n-k}{k}c_{n-k} \quad \text{for} \quad n \geq 2\,. \tag{1.13}$$

Summation here goes efectively from $k = 1$ to $k = \lfloor n/2 \rfloor$. With formula (1.7) this yields the binomial identity

$$\frac{1}{n}\binom{2n-2}{n-1} = \sum_{k=1}^{\lfloor n/2 \rfloor} (-1)^{k+1} \frac{1}{n-k}\binom{2n-2k-2}{n-k-1}\binom{n-k}{k}, \ n \geq 2 \, .$$

As we know, equation (1.12) is equivalent with basic recurrence (1.1).

With the help of Stirling's asymptotic formula for $n!$ we deduce from formula (1.7) precise asymptotics for c_n.

Theorem 1.3.7 (Stirling's formula). *For $n \to \infty$ we have the asymptotic relation*

$$n! = 1 \cdot 2 \cdot 3 \cdot \ \ldots \ \cdot n \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$$

where $\pi = 3.14159\ldots$ and $e = 2.71828\ldots$ are the well known constants.

We prove Stirling's formula in three ways in Chapter 2.

Corollary 1.3.8. *The Catalan numbers for $n \to \infty$ satisfy*

$$c_n = \frac{1}{n}\binom{2n-2}{n-1} \sim \frac{1}{4\sqrt{\pi}} \cdot \frac{4^n}{n^{3/2}} \, .$$

Proof. By Stirling's formula, for $n \to \infty$

$$c_n = \frac{n}{2n(2n-1)} \cdot \frac{(2n)!}{n! \cdot n!} \sim \frac{1}{4n} \cdot \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)^2}$$

which simplifies to the stated expression. □

However, unless we estimate rate of convergence in $\sim$, this says nothing about the size of c_n for a particular $n \in \mathbb{N}$. Fortunately, explicit versions of Stirling's formula are known and we mention them in Section 1.7.

In conclusion we derive in a self-contained way the asymptotics for c_n in weaker form without the factor $\frac{1}{4\sqrt{\pi}}$. We use the following more generally applicable result. Recall that sequences $(a_n) \subset \mathbb{C}$ defined by

$$a_n = \prod_{j=1}^{n} r(j)$$

for a rational function $r \in \mathbb{C}(x)$ (i.e., $r(x)$ is a ratio of two complex polynomials) are called *hypergeometric*.

Theorem 1.3.9 (asymptotics of hypergeometric sequences). *Suppose that $(a_n) \subset (0, +\infty)$ is a sequence of positive real numbers such that for every $n > n_0$ we have*

$$\frac{a_{n+1}}{a_n} = A \cdot \frac{n^k + \alpha_1 n^{k-1} + \cdots + \alpha_k}{n^k + \beta_1 n^{k-1} + \cdots + \beta_k}$$

where $A, \alpha_i, \beta_i \in \mathbb{R}$ are constants such that $A > 0$, $k \geq 1$ and $\alpha_1 \neq \beta_1$. Then there is a real constant $c > 0$ such that for $n \to \infty$ we have the asymptotic relation

$$a_n \sim c n^{\alpha_1 - \beta_1} A^n \ .$$

Proof.

□

Catalan numbers form a hypergeometric sequence and we can apply to them Theorem 1.3.9.

Corollary 1.3.10. *For some constant $c > 0$ and $n \to \infty$,*

$$c_n = \frac{1}{n}\binom{2n-2}{n-1} \sim c \cdot \frac{4^n}{n^{3/2}} \ .$$

Proof. By recurrence (1.6), for every $n \in \mathbb{N}$ one has

$$\frac{c_{n+1}}{c_n} = \frac{4n-2}{n+1} = 4 \cdot \frac{n - \frac{1}{2}}{n+1} \ .$$

For the sequence (c_n) the constants therefore are $A = 4$, $\alpha_1 = -\frac{1}{2}$ and $\beta_1 = 1$. The asymptotics follows from Theorem 1.3.9. □

Exercise 1.3.11. *Which is larger for $n \to +\infty$,*

$$\prod_{i=n+1}^{2n} i \quad \textit{or} \quad \prod_{i=0}^{n-1} (3i+1) \ ?$$

Determine asymptotics of the ratio of the first product to the second.

1.4 Dyck words, good bracketings and pattern-avoiding permutations

We derive the classic formula $c_n = \frac{1}{n}\binom{2n-2}{n-1}$ for Catalan numbers combinatorially, without generating functions. We find a bijection between rp trees and simpler objects, which we count directly by a clever transformation. Then we present other families of structures counted by c_n: good bracketings and permutations avoiding a fixed three-term subpermutation.

The objects are *Dyck words*. A Dyck word

$$D = (d_1, \ldots, d_{2n}) \in \{-1, 1\}^{2n}$$

of *size* $n \in \mathbb{N}_0$ is a $2n$-tuple of 1s and -1s such that

$$d_1 + d_2 + \cdots + d_j \geq 0, \ 1 \leq j \leq 2n, \ d_1 + d_2 + \cdots + d_{2n} = 0 \ .$$

So D has equal number, n, of 1s and -1s and never more -1s than 1s in any initial segment. Let $\mathcal{D}_n$ be the set of Dyck words with size n. We establish a bijection between them and rp trees.

Proposition 1.4.1. *For every $n \in \mathbb{N}$ there is a bijection $f : \mathcal{T}_n \to \mathcal{D}_{n-1}$. Thus $c_n = |\mathcal{D}_{n-1}|$.*

Proof. We define f both recursively and explicitly. For $n = 1$ and the only rp tree T with $|T| = 1$ we set $f(T) = \emptyset$. For $n > 1$ and $T \in \mathcal{T}_n$,

$$f(T) := (1,\, f(T_1),\, -1,\, 1,\, f(T_2),\, -1,\, \ldots,\, 1,\, f(T_k),\, -1)$$

where $T_1, T_2, \ldots, T_k$ are the rp subtrees of T rooted in the children of the root, listed in the order of the children. Explicitly, we walk around T in clockwise direction, starting and finishing at the root, and write for each edge 1 if we go upwards and -1 if we go downwards. Clearly, this is the same as the recursive definition, and f is a bijection between $\mathcal{T}_n$ and $\mathcal{D}_{n-1}$. The walk in fact does the depth-first search in T. □

Exercise 1.4.2. *Check in details that f is a bijection between $\mathcal{T}_n$ and $\mathcal{D}_{n-1}$.*

For example, the five rp trees of $\mathcal{T}_4$ (see p. 2) are sent by f to the respective Dyck words (we write 0 for -1)

$$\mathcal{D}_3 = \{101010,\ 110010,\ 101100,\ 110100,\ 111000\}\ .$$

To count $\mathcal{D}_n$, we consider the set $\mathcal{E}_n$ of all $2n$-tuples of n 1s and n -1s, and the set $\mathcal{F}_n$ of all $2n$-tuples of $n+1$ 1s and $n-1$ -1s. Clearly,

$$\mathcal{D}_n \subset \mathcal{E}_n,\ |\mathcal{E}_n| = \binom{2n}{n} \quad \text{and} \quad |\mathcal{F}_n| = \binom{2n}{n+1} = \binom{2n}{n-1}\ .$$

Proposition 1.4.3. *For every $n \in \mathbb{N}$ there is a bijection $f : \mathcal{E}_n \backslash \mathcal{D}_n \to \mathcal{F}_n$. Thus $|\mathcal{E}_n| - |\mathcal{D}_n| = |\mathcal{F}_n|$.*

Proof. For $D = (d_1, \ldots, d_{2n}) \in \mathcal{E}_n$, $D \notin \mathcal{D}_n \iff d_1 + d_2 + \cdots + d_j = -1$ for some $j \in [2n]$. For $D \in \mathcal{E}_n \backslash \mathcal{D}_n$ we take minimum such j and set

$$f(D) := (-d_1, -d_2, \ldots, -d_j, d_{j+1}, \ldots, d_{2n})\ .$$

Since the initial part of D with length j has one more -1 than 1s, and the rest of D contrarywise, $f(D)$ has two more 1s than -1s and belongs to $\mathcal{F}_n$. For $D \in \mathcal{F}_n$ take the shortest initial part with sum 1 (which exists) and change signs in it. This gives a mapping $g : \mathcal{F}_n \to \mathcal{E}_n \backslash \mathcal{D}_n$, inverse to f. So f is a bijection. □

Since

$$|\mathcal{E}_n| - |\mathcal{F}_n| = \binom{2n}{n} - \binom{2n}{n-1} = \frac{(2n)!(1 - \frac{n}{n+1})}{n!n!} = \frac{1}{n+1}\binom{2n}{n}\ ,$$

the previous two propositions prove that $c_n = |\mathcal{D}_{n-1}| = |\mathcal{E}_{n-1}| - |\mathcal{F}_{n-1}| = \frac{1}{n}\binom{2n-2}{n-1}$. For reference we record the count of $\mathcal{D}_n$ as a proposition.

Proposition 1.4.4 (counting Dyck words). *The number of Dyck words with size $n \in \mathbb{N}_0$ is*

$$\#\mathcal{D}_n = c_{n+1} = \frac{1}{n+1}\binom{2n}{n} .$$

Exercise 1.4.5. *Do we really have to translate rp trees to Dyck words? Could not the argument be carried on just with rp trees?*

Let us have a look at so called good bracketings, counted by the Catalan numbers too. They relate both to rp trees and to Dyck words. What is a good bracketing? A basic syntactic combinatorial concept in logic (even though many textbooks on mathematical logic do not recognize it as such, but some do) without which we would stay helpless over a formula and could not parse it correctly.

> *A good bracketing is a bijection $f : A \to B$ between two disjoint subsets $A, B \subset [n]$ ($n \in \mathbb{N}_0$), such that $i < f(i)$ for every $i \in A$ and, for every $i, j \in A$, $i < j < f(i) \Rightarrow f(j) < f(i)$.*

Obviously, we may assume in addition that $A \cup B = [2n]$ and $|A| = |B| = n$ and will do so; we call such partitions *bracketings*, the elements of A and B record, respectively, the positions of left and right brackets "(" and ")". In a good bracketing f pairs each opening bracket with its closing partner. By the next proposition, f is unique; we can recover these pairs and parse each well bracketed formula unambiguously. For example, consider the bracketings

$$(\,(\,(\,)\,(\,)\,)\,)\,(\,)\,(\,) \quad \text{and} \quad (\,(\,(\,)\,(\,)\,)\,)\,(\,(\,)\,)\,(\,) .$$

The former is not good but the latter is, with $f(1) = 8$, $f(2) = 7$, $f(3) = 4$, $f(5) = 6$, $f(9) = 12$, $f(10) = 11$ and $f(13) = 14$.

Proposition 1.4.6 (uniqueness of good bracketings). *For every $n \in \mathbb{N}_0$ and every partition $[2n] = A \cup B$ with two n-element blocks, there exists at most one good bracketing $f : A \to B$.*

Proof. Let $f : A \to B$ be a good bracketing where A and B are as given. We consider labeling $\ell : [2n] \to \{-1, 1\}$, $\ell(A) = \{1\}$ and $\ell(B) = \{-1\}$, and claim that for every $i \in A$ and $j \in [2n]$ with $i \le j < f(i)$,

$$\ell(i) + \ell(i+1) + \cdots + \ell(j) > 0 \quad \text{and} \quad \ell(i) + \ell(i+1) + \cdots + \ell(f(i)) = 0 .$$

Indeed, by the definition of f each element $e \in B \cap [i, j]$ pairs with the element $f^{-1}(e) \in A \cap [i, e]$ but i remains unpaired, and for the whole interval $[i, f(i)]$ the elements $f^{-1}(e)$ exhaust $A \cap [i, f(i)]$ and $|B \cap [i, f(i)]| = |A \cap [i, f(i)]|$. So

if f exists, it can be recovered just from A and B by setting for $i \in A$ the value $f(i)$ to be the first $j \in (i, 2n]$ with $\ell(i) + \ell(i+1) + \cdots + \ell(j) = 0$. □

Thus we identify good bracketings with those words from $\{-1, 1\}^{2n}$ with n -1s and n 1s, for which there is a bijection f as above between the positions of 1s and -1s. We denote the set of good bracketings in this format by $\mathcal{B}_n$. They look quite similar to Dyck words. Indeed, they are the same.

Proposition 1.4.7 (good bracketings are Dyck words). *For any $n \in \mathbb{N}_0$,*

$$\mathcal{B}_n = \mathcal{D}_n \ .$$

Hence $|\mathcal{B}_n| = c_{n+1} = \frac{1}{n+1}\binom{2n}{n}$.

Proof. Let $D = (d_1, d_2, \ldots, d_{2n}) \in \{-1, 1\}^{2n}$ have n -1s and n 1s. Assume that $D \in \mathcal{B}_n$ and take the bijection $f : A \to B$ (A are the positions of 1s, and B of -1s) proving it. Consider an initial segment $(d_1, \ldots, d_j)$, $1 \leq j \leq 2n$, of D. We cover $[1, j]$ by the intervals $[1, f(1)]$, $[f(1)+1, f(f(1)+1)]$, $\ldots$, $[i, f(i)]$ where $i \leq j \leq f(i)$. By the displayed inequalities in the proof of Proposition 1.4.6, $\sum_{k \in I} d_k = 0$ if I is any on these intervals and is ≥ 0 if $I = [i, j]$. Thus $\sum_{k=1}^{j} d_k \geq 0$ and $D \in \mathcal{D}_n$.

Assume that $D \in \mathcal{D}_n$. We prove by induction on n that $D \in \mathcal{D}_n$. For $n = 0$ it is true as then $D = ()$ and the bijection f is an empty mapping. Suppose that $n > 0$, consider the minimum j with $d_1 + d_2 + \cdots + d_j = 0$ and set $D_1 = (d_2, \ldots, d_{j-1})$ and $D_2 = (d_{j+1}, \ldots, d_{2n})$. It is easy to see that D_1 and D_2 are Dyck words, shorter than D. By induction, they are good bracketings and we have for them the corresponding bijections f_1 and f_2. Thus D is a good bracketing in $\mathcal{D}_n$, witnessed by the bijection $f = f_1 \cup f_2 \cup \{(1, j)\}$.

The enumeration result follows from Proposition 1.4.4. □

We conclude that of the two equivalent decriptions of the same object, as a Dyck word or as a good bracketing, the former is simpler and the latter explicitly (and thus better) captures the inherent recursive structure.

Exercise 1.4.8. *Devise another combinatorial proof of the previous result by partitioning the words in $\{-1, 1\}^{2n}$ with n -1s and n 1s into $n+1$ blocks with equal sizes, one of them being the good bracketings.*

We switch to permutations with forbidden patterns. They add to rp trees, Dyck words and good bracketings many more families counted by the Catalan numbers. Here we present six. For $n \in \mathbb{N}$ symbol $\mathcal{S}_n$ denotes the set of $n!$ permutations of $[n] = \{1, 2, \ldots, n\}$, i.e. the set of all n-tuples $a_1 a_2 \ldots a_n$ with $\{a_1, a_2, \ldots, a_n\} = \{1, 2, \ldots, n\}$, and $\mathcal{S} = \bigcup_{n \geq 1} \mathcal{S}_n$ the set of all permutations. For $\pi \in \mathcal{S}_m$ we write $|\pi| = m$. A permutation $\pi = a_1 a_2 \ldots a_m \in \mathcal{S}_m$ *is contained* in a permutation $\rho = b_1 b_2 \ldots b_n \in \mathcal{S}_n$, written

$$\pi \preceq \rho \ ,$$

if there is a subsequence $1 \le i_1 < i_2 < \cdots < i_m \le n$ of $1, 2, \ldots, n$ (in particular, $m \le n$) such that for every $j, k \in [m]$ we have

$$a_j < a_k \iff b_{i_j} < b_{j_k} .$$

For example, $12 \not\preceq \rho$ if and only if $\rho = n(n-1)\ldots 21$. In more visual terms, $\pi \preceq \rho$ if and only if in any graph (in the sense of graphs of real function) R of ρ,

$$R = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \subset \mathbb{R}^2$$

where $x_1 < x_2 < \cdots < x_n$ and $y_i < y_j \iff b_i < b_j$, contains as a subset a graph (in this sense) of π. In yet another way we express the containment by the *normalization* $n : \mathbb{N}^{**} \to \mathcal{S}$, where $\mathbb{N}^{**}$ are all finite injective sequences of the natural numbers,

$$n(x_1 x_2 \ldots x_k) = a_1 a_2 \ldots a_k \in \mathcal{S}_k \text{ such that } x_i < x_j \iff a_i < a_j .$$

Then $\pi \preceq \rho$ iff ρ has a subsequence whose normalization is π. For any $k \in \mathbb{N}$ we define a mapping from $\mathcal{S}$ to $\mathbb{N}^{**}$ by

$$a_1 a_2 \ldots a_m \mapsto a_1 a_2 \ldots a_m + k := (a_1 + k)(a_2 + k)\ldots(a_m + k) .$$

For any (possibly infinite) set of permutations $X \subset \mathcal{S}$ and $n \in \mathbb{N}$ we define

$$\mathcal{S}(X) = \{\rho \in \mathcal{S} \mid \forall\, \pi \in X : \ \pi \not\preceq \rho\} \text{ and } s_n(X) = |\mathcal{S}(X) \cap \mathcal{S}_n| \in \mathbb{N}_0 .$$

Thus $s_n(X)$ counts the permutations of $1, 2, \ldots, n$ avoiding (not containing) every permutation in X. If $\#X = 1$ and $X = \{\pi\}$, we omit the curly brackets and write just $\mathcal{S}(\pi)$ and $s_n(\pi)$. For example, $s_n(12) = s_n(21) = 1$ for every $n \in \mathbb{N}$. We determine the counting functions $s_n(\pi)$ for all six $\pi \in \mathcal{S}_3$. They turn out to be all equal and, as the reader already rightly guesses, are equal to the sequence of Catalan numbers. First we reduce by a symmetry argument the six cases to just two.

Proposition 1.4.9 (not six but two cases). *For*

$$\pi \in \{132, 231, 312, 213\} \text{ and } \pi \in \{123, 321\} ,$$

the four, respectively two, counting functions $s_n(\pi)$ coincide.

Proof. Consider the *reversal* and *complement* $r, c : \mathcal{S} \to \mathcal{S}$,

$$r(a_1 a_2 \ldots a_n) = a_n a_{n-1} \ldots a_1$$

and

$$c(a_1 a_2 \ldots a_n) = (n - a_1 + 1)(n - a_2 + 1)\ldots(n - a_n + 1) .$$

From the definition of the containment we have that $s_n(\pi) = s_n(r(\pi))$ and $s_n(\pi) = s_n(c(\pi))$ for every $\pi \in \mathcal{S}$ and $n \in \mathbb{N}$. Since $r(132) = 231$, $c(132) = 312$, $r(c(132)) = 213$ and $r(123) = 321$, the result follows. □

Exercise 1.4.10. *In more details, why is it true that $s_n(\pi) = s_n(r(\pi))$ and $s_n(\pi) = s_n(c(\pi))$? What group do r and c generate?*

Proposition 1.4.11 (132-avoiding permutations). *For every $n \in \mathbb{N}$,*

$$s_n(132) = \#\{\rho \in \mathcal{S}_n \mid \rho \not\geq 132\} = c_{n+1} = \frac{1}{n+1}\binom{2n}{n} .$$

Proof. Let $s_n = s_n(132)$. We set $s_0 = 1$. Then for every $n \in \mathbb{N}$ we have the recurrence

$$s_n = \sum_{i=0}^{n-1} s_i s_{n-1-i} . \tag{1.14}$$

It mirrors the bijection (we denote $\mathcal{S}^0(132) = \mathcal{S}(132) \cup \{\emptyset\}$)

$$f : \mathcal{S}(132) \to \mathcal{S}^0(132) \times \mathcal{S}^0(132),\ \rho \mapsto (\sigma_1, \sigma_2) = (n(\rho_1), n(\rho_2)) ,$$

satisfying $|\rho| = |\sigma_1| + |\sigma_2| + 1$. Here for $\rho \in \mathcal{S}(132) \cap \mathcal{S}_n$ the (possibly empty) ρ_i are given by $\rho = \rho_1 n \rho_2$. We check that f has the stated properties. Crucial observation is that, since $132 \not\preceq \rho$, $\rho_1 > \rho_2$. So we recover uniquely ρ from (σ_1, σ_2) and f is injective. In the other way, for any pair (σ_1, σ_2) in $\mathcal{S}^0(132) \times \mathcal{S}^0(132)$ with $|\sigma_1| = k$ and $|\sigma_2| = l$, the permutation

$$(\sigma_1 + l)\,(k + l + 1)\,\sigma_2 \in \mathcal{S}(132) \cap \mathcal{S}_{k+l+1}$$

and maps by f to (σ_1, σ_2), which shows that f is onto. The relation $|\rho| = |\sigma_1| + |\sigma_2| + 1$ is obvious.

This decomposition and recurrence (1.14) are very similar but not identical to Proposition 1.1.1 and recurrence (1.1) (see Exercise 1.4.12). Let $S = S(x) = \sum_{n\geq 0} s_n x^n$. Recurrence (1.14) says that

$$S - 1 = xS^2, \quad \text{or} \quad xS^2 - S + 1 = 0 , \tag{1.15}$$

which should be compared with equation (1.3). Multiplying with x we get $(xS)^2 - xS + x = 0$ and see that $xS = C$, which gives the result $s_n = c_{n+1} = \frac{1}{n+1}\binom{2n}{n}$. □

Exercise 1.4.12. *We revisit recurrences (1.1) and (1.14). Let $a_n, b_n \in \mathbb{N}$ be given by the recurrences $a_1 = 1$, $a_n = \sum_{i=1}^{n-1} a_i a_{n-i}$ for $n \geq 2$, $b_0 = 1$ and $b_n = \sum_{i=0}^{n-1} b_i b_{n-i-1}$ for $n \geq 1$. Prove directly, without resorting to generating functions, that $b_n = a_{n+1}$.*

Proposition 1.4.13 (123-avoiding permutations). *For every $n \in \mathbb{N}$,*

$$s_n(123) = \#\{\rho \in \mathcal{S}_n \mid \rho \not\geq 123\} = c_{n+1} = \frac{1}{n+1}\binom{2n}{n} .$$

Proof.

□

Corollary 1.4.14 (Catalan numbers everywhere). *For all six $\pi \in \mathcal{S}_3$ and $n \in \mathbb{N}$,*

$$s_n(\pi) = \#\{\rho \in \mathcal{S}_n \mid \rho \not\supset \pi\} = c_{n+1} = \frac{1}{n+1}\binom{2n}{n}.$$

Proof.

□

1.5 Can (c_n) satisfy a linear recurrence with constant coefficients?

From recurrence (1.1) we derived simpler recurrence (1.6). Perhaps recurrence (1.1) could be simplified even more and the Catalan numbers could be shown to satisfy a linear recurrence of the form

$$c_n = \alpha_1 c_{n-1} + \alpha_2 c_{n-2} + \cdots + \alpha_k c_{n-k},\ n > n_0 \geq k\ ? \tag{1.16}$$

Here the fixed $k \in \mathbb{N}$ is the order of the recurrence and all coefficients $\alpha_i \in \mathbb{C}$ with $\alpha_k \neq 0$ are constants. Note that recurrence (1.6) has order $k = 1$ but the coefficient $\alpha_1 = \alpha_1(n) = \frac{4n-6}{n}$ depends on n. Sequences satisfying recurrences like (1.16) and their GFs are discussed in Chapter 5. We show that (c_n) does not belong to them.

Proposition 1.5.1. *The Catalan numbers c_n satisfy no linear recurrence with constant coefficients, that is, no recurrence of the form (1.16).*

Certainly, this is not an unexpected result. Its beauty lies in the various ways one can prove it. We present four proofs, all of which but the last use properties of rational generating functions that will be established in Chapter 5.

First proof, by 2-adic valuation

Proposition 1.1.3 says that odd values of c_n are highly lacunary. Linear recurrence sequences with constant coefficients cannot behave thus. By Proposition 5.1.2 in Chapter 5 we may assume that the coefficients α_i in equation (1.16) all lie in the field of fractions $\mathbb{Q}$. For $\alpha \in \mathbb{Q}\backslash\{0\}$ we denote by $\mathrm{ord}_2(\alpha) \in \mathbb{Z}$ the exponent of 2 in the prime factorization of α, the unique $m \in \mathbb{Z}$ such that $\alpha = 2^m\beta$ where $\beta \in \mathbb{Q}$ can be expressed with both numerator and denominator odd, and set $\mathrm{ord}_2(0) = +\infty$. For example, $\mathrm{ord}_2(\frac{3}{28}) = -2$. Recall two basic properties of the function $\mathrm{ord}_2 : \mathbb{Q} \to \mathbb{Z} \cup \{+\infty\}$ which it shares with all other p-adic valuations ord_p for prime numbers p:

Exercise 1.5.2. *Let $\alpha, \beta \in \mathbb{Q}$. Prove the following.*

1. *We have* $\mathrm{ord}_2(\alpha\beta) = \mathrm{ord}_2(\alpha) + \mathrm{ord}_2(\beta)$.
2. *Also,* $\mathrm{ord}_2(\alpha + \beta) \geq \min(\mathrm{ord}_2(\alpha), \mathrm{ord}_2(\beta))$, *with equality if* $\mathrm{ord}_2(\alpha) \neq \mathrm{ord}_2(\beta)$.

To prove Proposition 1.5.1 we assume for contradicion that equality (1.16) holds for every $n > n_0$ for some $k \in \mathbb{N}$ constant coefficients $\alpha_i \in \mathbb{Q}$, $\alpha_k \neq 0$. We write it symmetricly as

$$\alpha_0 c_n + \alpha_1 c_{n-1} + \cdots + \alpha_k c_{n-k} = 0, \ n > n_0 \ ,$$

where $\alpha_0 := -1$. Let

$$m = \min_{0 \leq i \leq k} \mathrm{ord}_2(\alpha_i), \quad \text{thus } m \leq \mathrm{ord}_2(\alpha_0) = 0 \ .$$

We fix a j, $0 \leq j \leq k$, such that $\mathrm{ord}_2(\alpha_j) = m$. Now we can certainly take an $N \in \mathbb{N}$ larger that n_0 and $2k$ and such that $N - j$ is a power of 2. Then none of the other numbers $N - i$, $0 \leq i \leq k$ and $i \neq j$, is a power of 2. By Proposition 1.1.3, c_{N-j} is odd but all other Catalan numbers c_{N-i} in equality (1.16) with $n = N$ are even. Hence the minimum

$$\min_{0 \leq i \leq k} \mathrm{ord}_2(\alpha_i c_{N-i}) = \min_{0 \leq i \leq k} (\mathrm{ord}_2(\alpha_i) + \mathrm{ord}_2(c_{N-i}))$$

is attained for the unique index $i = j$ and, by Exercise 1.5.2,

$$+\infty = \mathrm{ord}_2(0) = \mathrm{ord}_2(\textstyle\sum_{i=0}^{k} \alpha_i c_{N-i}) = \mathrm{ord}_2(\alpha_j c_{N-j}) = \mathrm{ord}_2(\alpha_j) = m \leq 0 \ ,$$

a contradiction. □

Second proof, by irrationality of $C(x)$

By the results in Section 5.1, every linear recurrence sequence with constant coefficients has a rational GF, that is, its GF is a ratio of two polynomials. Thus equality (1.16) for (c_n) would imply that $C(x)$ is rational and therefore, by formula (1.10),

$$\sqrt{1 - 4x} = \frac{p(x)}{q(x)}, \quad \text{for some } p, q \in \mathbb{C}[x] \backslash \{0\} \ .$$

Thus

$$(1 - 4x) q(x)^2 = p(x)^2 \ .$$

This is impossible: the degree of the left side is odd but the degree of the right side is even. □

This resembles the well known proof shown in calculus classes of irrationality of the real number $\sqrt{2}$.

Exercise 1.5.3. *What are, in algebraic terms, the common features of the above proof that* $\sqrt{1-4x} \notin \mathbb{C}(x)$ *and the proof that* $\sqrt{2} \notin \mathbb{Q}$*?*

We may as well get contradiction from equation (1.11): if $C(x) = \frac{p(x)}{q(x)}$ for two (coprime) polynomials $p, q \in \mathbb{C}[x]$ then substituting $x - x^2$ for x we get

$$x = C(x-x^2) = \frac{p(x-x^2)}{q(x-x^2)}, \quad \text{or} \quad x \cdot q(x-x^2) = p(x-x^2)\ .$$

Again, the left degree is odd but the right degree is even, a contradiction.

Polynomial substitution in fact preserves coprimality of two polynomials, which once more shows that the above displayed equality is impossible:

Exercise 1.5.4. *Prove that if* $p, q \in \mathbb{C}[x]$ *are coprime polynomials and* $r \in \mathbb{C}[x]$ *is arbitrary but non-constant then the polynomials* $p(r(x))$ *and* $q(r(x))$ *remain coprime.*

Exercise 1.5.5. *By Proposition 1.1.3 the sequence* $(c_n \bmod 2)$ *is not eventually periodic: no* $a, b \in \mathbb{N}$ *exist such that* $c_n \equiv c_{n+a}$ *modulo* 2 *whenever* $n > b$*. Extend this from* $p = 2$ *to any prime modulus* p*.*

Third proof, by asymptotic analysis

The first idea coming to the mind of an enumerative combinatorialist when you ask her or him about a proof of Proposition 1.5.1 might be that the asymptotics $c_n \sim cn^{-3/2}4^n$ (Corollaries 1.3.8 and 1.3.10) is clearly incompatible with power sums, the general explicit form for terms of linear recurrence sequences with constant coefficients, and that's it. However, to show it without handwaving takes some effort.

What are power sums? We prove in Section 5.1 that if (c_n) satisfies recurrence (1.16) then for every $n > n_0$ we have equality

$$c_n = \sum_{j=1}^{r} p_j(n)\gamma_j^n \tag{1.17}$$

where $r \in \mathbb{N}$, $\gamma_j \in \mathbb{C}\backslash\{0\}$ are distinct numbers, and $p_j \in \mathbb{C}[x]$ are nonzero polynomials. It seems clear that $c_n \sim cn^{-3/2}4^n$ and expression (1.17) are for $n \to \infty$ incompatible, because the leading term in expression (1.17) is $cn^k\gamma^n$ with $c, \gamma \neq 0$ and $k \in \mathbb{N}_0$. However, this really only seems; it holds only if the maximum modulus $|\gamma_j|$ is attained for one index j. If it is attained more often, situation gets complicated. The correct leading term in expression (1.17) is, in general,

$$c_n = (d_1 z_1^n + d_2 z_2^n + \cdots + d_t z_t^n)n^k\gamma^n + O(n^{k-1}\gamma^n) \tag{1.18}$$

where $d_j \in \mathbb{C}\backslash\{0\}$, $t \in \mathbb{N}$, $k \in \mathbb{N}_0$, $z_j \in \mathbb{C}$ are distinct numbers with $|z_j| = 1$, and $\gamma > 0$ is the maximum modulus $|\gamma_j|$. In Lemma 1.5.7 below we show that the coefficient $(\cdots)$ of $n^k\gamma^n$ satisfies $\limsup_{n\to\infty} |(\cdots)| > 0$. Trivially,

$|(\cdots)| \leq |d_1|+\cdots+|d_t| = O(1)$ for every n. Thus intuition did not betray us and expression (1.18) and hence expression (1.17) is incompatible with asymptotics $c_n \sim cn^{-3/2}4^n$ of Corollary 1.3.10. □

But it remains to prove the promised lemma. Examples like

$$1^n + (-1)^n \ ,$$

which is 2 for even $n \in \mathbb{N}_0$ but 0 for odd $n \in \mathbb{N}_0$, show that the coefficient $(\cdots)$ may vanish on whole infinite arithmetic progressions. But we prove that on the other hand it is bounded away from 0 for infinitely many n. We need the next determinantal formula where we met again Alexandre-Théophile Vandermonde.

Exercise 1.5.6 (Vandermonde determinant). *Suppose $x_1, x_2, \ldots, x_n$ are variables. Then*

$$\det\left(x_i^{j-1}\right)_{i,j=1}^{n} = \prod_{1\leq i<j\leq n} (x_j - x_i) \ .$$

Lemma 1.5.7. *Let $t \in \mathbb{N}$, $d_j \in \mathbb{C}\backslash\{0\}$ and $z_j \in \mathbb{C}$ with $|z_j| = 1$ (z_j lie on the unit complex circle) for $j = 1, 2, \ldots, t$, and let the t numbers d_j be mutually distinct. Then*

$$\limsup_{n\to\infty} |d_1 z_1^n + d_2 z_2^n + \cdots + d_t z_t^n| > 0 \ .$$

Proof. Suppose (for contradiction) that the limsup is 0,

$$\lim_{n\to\infty} (d_1 z_1^n + d_2 z_2^n + \cdots + d_t z_t^n) = 0 \ .$$

So, denoting the linear combination $(\cdots)$ by $v(n)$, for every $k \in \mathbb{N}$ there is an $n_k \in \mathbb{N}$ with $|v(n_k + m)| \leq \frac{1}{k}$ for $m = 1, 2, \ldots, t$. We express $d_1, \ldots, d_t$ in terms of $v(n_k + 1), \ldots, v(n_k + t)$ by Cramer's rule:

$$d_j = \frac{\det M(j)}{\det(z_l^{n_k+m})_{m,l=1}^{t}}, \ j = 1, 2, \ldots, t \ ,$$

where the $t \times t$ matrix $M(j)$ in the numerator arises from the denominator matrix by replacing in it the j-th column with the column $(v(n_k + 1), v(n_k + 2), \ldots, v(n_k + t))^T$. By the definition of determinant, the triangle inequality, and the assumption on $v(n_k + m)$, we have the bound

$$|\det M(j)| \leq \frac{t!}{k} \ .$$

If we take out $z_l^{n_k+1}$ from the l-th column of the denominator matrix, it becomes a Vandermonde matrix: in the l-th column there remain powers z_l^{m-1}, $m = 1, 2, \ldots, t$. By Exercise 1.5.6 and since the numbers z_l are all distinct,

$$|\det(z_l^{n_k+m})_{m,l=1}^{t}| = \prod_{1\leq l<l'\leq t} |z_{l'} - z_l| =: d > 0 \ ,$$

a nonzero constant independent of k. Thus, for every $j = 1, 2, \ldots, t$,

$$|d_j| \leq \frac{t!}{dk} \to 0, \ k \to \infty \ .$$

So $d_j = 0$ (which contradicts the assumption that every d_j is nonzero). □

Exercise 1.5.8. *Recall Cramer's rule from linear algebra and its proof.*

Exercise 1.5.9. *What does the lemma say geometricly? What can be said about the density of the $n \in \mathbb{N}$ with $v(n)$ bounded away from 0?*

Fourth proof, by algebra

I found this proof as the last one. With hindsight it seems to be the simplest and most natural of the four. We assume that relation (1.16) holds and substitute in it explicit expressions $c_{n-i} = \frac{1}{n-i}\binom{2n-2i-2}{n-i-1} = \frac{1}{n-i}\frac{(2n-2i-2)!}{(n-i-1)!^2}$ (formula (1.7)). We multiply it by $\frac{n(n-1)\ldots(n-k)\cdot(n-1)!^2}{(2n-2k-2)!}$, to cancel common factors and denominators, and rearrange it. Using the *Pochhammer symbol*

$$(x)_k := x(x-1)(x-2)\ldots(x-k+1), \ k \in \mathbb{N}_0, \ (x)_0 := 1 \ ,$$

and setting $\alpha_0 := -1$, we obtain the resulting equation

$$\sum_{i=0}^{k} \alpha_i \cdot \widehat{(n)_{k+1}} \cdot (2n-2-2i)_{2k-2i} \cdot (n-1)_i^2 = 0$$

in which $\widehat{\cdots}$ signifies omission of the factor $n-i$. We regard n as an indeterminate and set $n = 0$. Then all summands vanish except for $i = 0$ that is $-(-1)_k(-2)_{2k} \neq 0$. Hence the sum is a nonzero polynomial in n, with degree at most $3k$. It vanishes for every $n > n_0$, which is impossible. □

Exercise 1.5.10. *Why is it impossible? (Cf. Exercise 1.3.6.)*

1.6 Refining c_n — the Narayana numbers

One can refine the count of rp trees, $c_n = \frac{1}{n}\binom{2n-2}{n-1}$, by many statistics. One of the basic is the number of leaves where a leaf is a childless vertex. Thus we define for $k, n \in \mathbb{N}$ and $1 \leq k \leq n$,

> *the n,k-th Narayana number $c_{n,k}$ = the number of rp tress with n vertices and k leaves .*

We explain why $c_{n,k}$ are "Narayana" in Section 1.7. Clearly, $c_n = \sum_{k=1}^{n} c_{n,k}$ and $c_{1,1} = 1$ but $c_{n,n} = 0$ for $n > 1$. The earlier picture on p. 2 shows that $c_{4,1} = c_{4,3} = 1$ and $c_{4,2} = 3$. Is there a nice formula for $c_{n,k}$? We reveal the answer here, but prove it later.

Proposition 1.6.1 (Narayana numbers). *Let $n \geq 2$ and $1 \leq k \leq n-1$. There are*

$$c_{n,k} = \frac{1}{n-1}\binom{n-1}{k}\binom{n-1}{k-1}$$

rp trees with n vertices and k leaves.

It is easy to generalize the argument of Section 1.3 and to deduce for the bivariate GF counting rp trees by vertices and leaves a formula extending (1.10). Compared to the univariate case, it is harder to extract from it an explicit formula for coefficients. We defer it to Section 4.4.

We enrich the decomposition of Proposition 1.1.1 with the information about leaves. For an rp tree T, the number of vertices is $|T|$ and the number of leaves is denoted by $\|T\|$.

Proposition 1.6.2. *There is a bijection*

$$f:\ \mathcal{T}\backslash\mathcal{T}_1 \to \mathcal{T}\times\mathcal{T},\ T \mapsto (T_1, T_2)\ ,$$

such that always $|T| = |T_1| + |T_2|$ and $\|T\| = \|T_1\| + \|T_2\|$, except for $|T_2| = 1$ when $\|T\| = \|T_1\| + \|T_2\| - 1 = \|T_1\|$.

Proof. We use the decomposition of Proposition 1.1.1. One-vertex tree has just one leaf. The number of leaves in T is the sum of their numbers in T_1 and in T_2, except when T_2 has just one vertex and does not contribute to $\|T\|$. □

We consider the bivariate GF $C = C(x,y) \in \mathbb{C}[y][[x]]$,

$$C = C(x,y) = \sum_{T\in\mathcal{T}} x^{|T|}y^{\|T\|} = \sum_{n,k\geq 1} c_{n,k}x^n y^k = xy + x^2y + x^3(y+y^2) + \dots\ .$$

By Proposition 1.6.2,

$$C - xy = C\cdot(C - xy + x), \quad \text{or} \quad C^2 + (x - xy - 1)C + xy = 0\ .$$

The quadratic formula gives

$$C(x,y) = \frac{1 + xy - x - \sqrt{(1+xy-x)^2 - 4xy}}{2} \tag{1.19}$$

—this extends the formula (1.10) and specializes to it when $y = 1$. Compared to $C(x)$, the difficulty is now that we have more than two terms under $\sqrt{\dots}$ and cannot use the binomial theorem easily. Still, there is a way to extract the nice formula of Proposition 1.6.1, and we explain it in Section 4.4 when we learn more about algebraic manipulations with GFs.

We are used to the complementarity of binomial coefficients:

$$\binom{n}{k} = \binom{n}{n-k}$$

because the k-element subsets of $[n]$ are in bijection with their complements, the $(n-k)$-element ones. Interestingly, the numbers $c_{n,k}$ enjoy the same symmetry. We prove it by their GF.

Proposition 1.6.3. *Let $n \geq 2$ and $1 \leq k \leq n-1$. Then*

$$c_{n,k} = c_{n,n-k} \ ,$$

that is, there are as many rp trees with k leaves and $n-k$ non-leaves as rp trees with $n-k$ leaves and k non-leaves.

Proof. To prove that $c_{n,k} = c_{n,n-k}$ via $C(x,y)$ means to show that $C(x,y)$ is invariant upon the substitution $x := xy, y := y^{-1}$, with the exception of the first monomial xy that turns in x. Applying it to the formula (1.19), we get

$$D(x,y) := C(xy, y^{-1}) = \frac{1 - xy + x - \sqrt{(1-xy+x)^2 - 4x}}{2} \ .$$

It is easy to check the identity

$$(1 + xy - x)^2 - 4xy = (1 - xy + x)^2 - 4x$$

(either side equals $1 + x^2y^2 + x^2 - 2xy - 2x - 2x^2y$), and thus indeed $D(x,y) = C(x,y) - xy + x$. □

Exercise 1.6.4. *Find bijective proof for the identity $c_{n,k} = c_{n,n-k}$, based on the combinatorial definition of $c_{n,k}$.*

By the combinatorial definitions, $c_n = \sum_{k=1}^{n-1} c_{n,k}$. So we have the binomial identity

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

1.7 Stanley's list and Valtr's theorem

We call the list after its author, R. P. Stanley. In Section 1.4 we saw that the Catalan numbers count besides rp trees also Dyck words. In fact, c_n count very many other families of structures. Stanley in [73, Exercise 6.19] collected 66 such Catalan problems. In [74] he has extended the list of Exercise 6.19 to the (current) total of 207 items. Below we quote verbatim from Stanley's list [73, Exercise 6.19] and [74] and include every ($c_5 = 14$)-th Catalan problem from it; we put our skippings or remarks in square brackets, like [...].

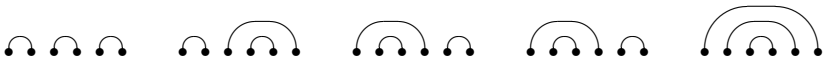
"

6.19. [1]–[3$^+$] Show that the Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ count the number of elements of the 66 sets S_i, (a) $\leq i \leq$ (nnn), given below. We illustrate the elements of each S_i for $n = 3$, hoping that these illustrations will make any undefined terminology clear. (The terms used in (vv)–(yy)) are defined in Chapter 7.) Ideally S_i and S_j should be proved to have the same cardinality by exhibiting a simple, elegant bijection $\phi_{ij} : S_i \to S_j$ (so 4290 bijections in all). In some cases the sets S_i and S_j will actually coincide but their description will differ.

a. Triangulations of a convex $(n+2)$-gon into n triangles by $n-1$ diagonals that do not intersect in their interiors:

[. . .]

o. Ways of connecting $2n$ points in the plane lying on a horizontal line by n nonintersecting arcs, each arc connecting two of the points and lying above the points:



[. . .]

cc. Permutations $a_1 a_2 \dots a_{2n}$ of the multiset $\{1^2, 2^2, \dots, n^2\}$ such that: (i) the first occurrences of $1, 2, \dots, n$ appear in increasing order, and (ii) there is no subsequence of the form $\alpha\beta\alpha\beta$:

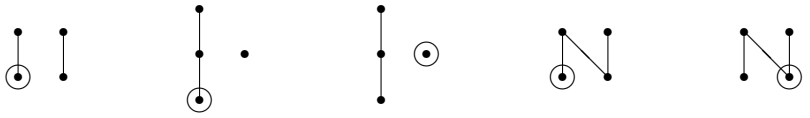
112233 112332 122331 123321 122133

[. . .]

qq. Partitions $\{B_1, \dots, B_k\}$ of $[n]$ such that if the numbers $1, 2 \dots, n$ are arranged in order around a circle, then the convex hulls of the blocks $B_1, \dots, B_k$ are pairwise disjoint:

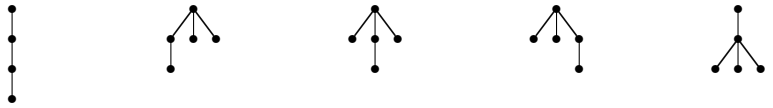
[. . .]

eee. Nonisomorphic $(n+1)$-element posets that are a union of two chains and that are not a (nontrivial) ordinal sum, rooted at a minimal element:



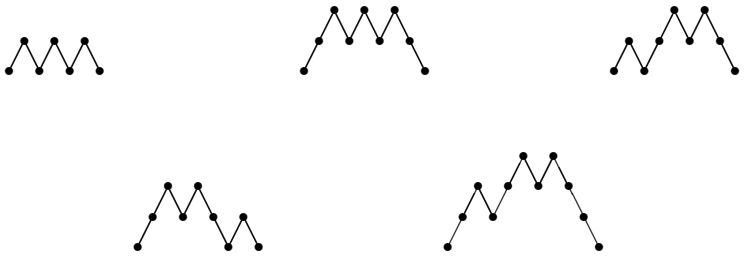
[. . . , we move to [74]]

(sss) Plane trees for which every vertex has $0, 1$, or 3 children, with a total of $n + 1$ vertices with 0 or 1 child



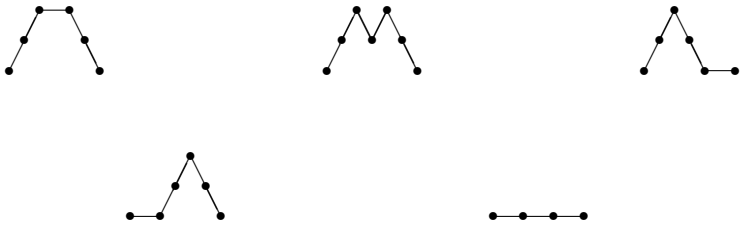
[. . .]

(g^4) Dyck paths with n peaks such that there are no factors (consecutive steps) UUU and $UUDD$



[. . .]

(u^4) Schröder paths as in Exercise 6.39(t) from $(0, 0)$ to $(2n, 0)$ with neither peak nor level step at odd height



[. . .]

(i^5) Ways of connecting $n + 1$ points in the plane lying on a horizontal line by noncrossing arcs above the line such that no arc connects adjacent points and the right endpoints of the arcs are all distinct

[...]

(w[5]) Sequences $1 \le a_1 < a_2 < \cdots < a_n \le 2n$ such that $a_i \le 2i-1$

123 124 125 134 135

[added, the illustration is missing in [74], ...]

(k[6]) Sequences $(a_1, \ldots, a_n) \in \mathbb{N}_n$ for which there exists a distributive lattice of rank n with a_i join-irreducibles of rank i, $1 \le i \le n$

300 210 120 201 111

[...]

(y[6]) Number of distinct terms (monomials) appearing in the expansion of $\prod_{i=1}^{n}(x_1 + x_2 + \cdots + x_i)$

$$x(x+y)(x+y+z) = x^3 + 2x^2y + xy^2 + x^2z + xyz$$

[...]

(m[7]) 321-avoiding permutations $w \in \mathfrak{S}_{2n+1}$ such that i is an excedance of w (i.e., $w(i) > i$) if and only if $i \ne 2n+1$ and $w(i) - 1$ is not an excedance of w (so that w has exactly n excedances)

4512736 3167245 3152746 4617235 5671234

[...]

(a[8]) Arrays

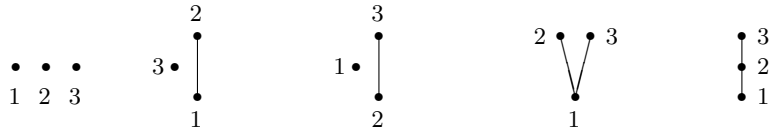
$$\begin{pmatrix} a_1 & a_2 & \ldots & a_{r-1} & a_r \\ b_1 & b_2 & \ldots & b_{r-1} & \end{pmatrix}$$

of integers, for some $r \ge 1$, such that $a_i > 0$, $b_i \ge 0$, $\sum a_i = n$, and $b_i < a_i + b_{i-1}$ for $1 \le i \le r-1$ (setting $b_0 = 0$)

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 0 & \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 1 & \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 0 & \end{pmatrix} \quad \begin{pmatrix} 3 \\ \end{pmatrix}$$

[...]

(o^8) Natural partial orderings $<_P$ of $[n]$ such that if $i <_P k$ and $i <_{\mathbb{Z}} j <_{\mathbb{Z}} k$, then $i <_P j$



”

Exercise 1.7.1. *Establish some of the* 4290 *bijections hinted to above.*

Finitely many points in the plane form a *convex chain* if they all lie on the graph of a convex function. In other words, the points $p_1 = (x_1, y_1), p_2 = (x_2, y_2), \ldots, p_n = (x_n, y_n)$ with $x_1 < x_2 < \cdots < x_n$ form a convex chain if the vectors $p_2 - p_1, p_3 - p_2, \ldots, p_n - p_{n-1}$ are ordered counter-clockwisely. Convex chains is a particular case of convex n-gons. The following remarkable theorem, due to P. Valtr, places for the first time (as far as we know) the Catalan numbers in the context of random discrete geometry. By $\Pr(A \,|\, B)$ we denote the probability of an event A conditioned by another event B, that is,

$$\Pr(A \,|\, B) = \frac{\Pr(A \cap B)}{\Pr(B)} \quad \text{if } \Pr(B) > 0 \,.$$

If $A \subset B$, as will be the case here, then in fact $\Pr(A \,|\, B) = \frac{\Pr(A)}{\Pr(B)}$.

Theorem 1.7.2 (Valtr, 1995). *Let $p_1, p_2, \ldots, p_n$ be n random and mutually independent points in the unit square $[0,1] \times [0,1]$. Then*

$$\begin{aligned} &\Pr(\textit{the points } p_i \textit{ form a convex chain} \,|\, \textit{they already form a convex } n\textit{-gon}) \\ &= \frac{1}{c_n} = \frac{1}{\frac{1}{n}\binom{2n-2}{n-1}} \,. \end{aligned}$$

Exercise 1.7.3. *Prove Valtr's theorem for $n = 3$.*

But in the spirit of Stanley's list we should provide an illustration of the case $c_4 = 5$. Here it is.

Proposition 1.7.4. *Let p_1, p_2, p_3, and p_4 be four points, picked at random and mutually independently in the unit square $[0,1] \times [0,1]$. Then*

$$\Pr(\textit{they form a convex chain} \,|\, \textit{they form a convex quadrangle}) = \frac{1}{c_4} = \frac{1}{5} \,.$$

Proof.

□

Comments and references

The Catalan numbers were named after the Belgian and French mathematician Eugène Ch. Catalan (1814–1894). However, L. Euler investigated them already in 1751 and they appeared in 1730s in the work of a Chinese scientist and mathematician Ming Antu—see Pak [62] for more information on history of c_n. For more information on modular behaviour of the Catalan (and related) numbers consult Chen and Jiang [20], Eu, Liu and Yeh [31] and Kauers, Krattenthaler and Müller [44] (and further references mentioned in these articles). An argument similar to but more complicated than that in the proof of Proposition 1.1.6 was given by Villarino [80] to prove somewhat stronger bound that has 6 instead of our 8. This is an example of the *symbolic method* when combinatorial relations for counted structures are directly mirrored by relations for generating functions. See Flajolet and Sedgewick [36], Goulden and Jackson [39] and Bergeron, Labelle and Leroux [11] for more information. (named after the French musician, mathematician and chemist Alexandre-Théophile Vandermonde (1735–1796)) (named after the Scottish mathematician James Stirling (1692–1770)) taken from Lando [48, Chapter 4.2]. Asymptotic relations of the form

$$a_n \sim cn^{-3/2}A^n, \ n \to \infty \ ,$$

occur surprisingly often, in enumerations of trees of various kinds but also for many other objects. Bell, Burris and Yeats [10] explain this ubiquity. These words are named after the German mathematician Walther F. A. von Dyck (1856–1934). (named after the Swiss mathematician Gabriel Cramer (1704–1752)) These are so called *Narayana numbers* (named after the Indian mathematician Tadepali Venkata Narayana (1930–1987)). Valtr [78, 79] He has published the material as a monograph on Catalan problems in [75]. , the paths are named after the German mathematician Ernst Schröder (1841–1902)

Chapter 2

Analytic intermezzo I. Stirling's formula for factorial

The second chapter presents three proofs for the fundamental result in combinatorial asymptotics, *Stirling's formula for factorial.* Let $n! = \prod_{i=1}^{n} i$ for $n \in \mathbb{N}$ and $0! := 1$. Stirling's formula reads:

$$n! = \left(\sqrt{2\pi} + o(1)\right)\sqrt{n}\left(\frac{n}{e}\right)^n, \ n \to \infty \tag{2.1}$$

(Theorem 1.3.7). The proofs are based on the following integral representations. For every $n \in \mathbb{N}_0$ one has

$$\log(n!) = \int_{\frac{1}{2}}^{n+\frac{1}{2}} \log x \, dx + c_1 + O(n^{-1}) \tag{2.2}$$

$$n! = \int_0^{+\infty} e^{-x} x^n \, dx \tag{2.3}$$

$$\frac{1}{n!} = \frac{1}{2\pi i} \oint \frac{e^z \, dz}{z^{n+1}} \, . \tag{2.4}$$

These are remarkable formulas. Only the second represents $n!$ directly, other two have simple expressions in $n!$. The first is an asymptotic equality with an unspecified constant $c_1 > 0$. In the first integral the integrand does not depend on the parameter n which appears in the integration path, and in the other two integrals it is the other way around. In the third complex integral we integrate over any positively oriented circle centered at the origin. In the next three sections we deduce relation (2.1) from the three identities (2.2), (2.3) and (2.4) in this order. We always start by proving the identity. In the last section we give references and commments.

2.1 Approximating sums with integrals

We prove identity (2.2). For $n \in \mathbb{N}_0$,

$$\log(n!) = \textstyle\sum_{m=1}^{n} \log m \ .$$

Also, for $m \in \mathbb{N}$,

$$\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} \log x \ dx = \log m + O(m^{-2}) \tag{2.5}$$

because

$$\begin{aligned}
(x \log x - x)\big|_{m-\frac{1}{2}}^{m+\frac{1}{2}} &= m \log\left(\frac{m+\frac{1}{2}}{m-\frac{1}{2}}\right) + \frac{\log(m^2-\frac{1}{4})}{2} - 1 \\
&= m \log\left(1 + \frac{1}{m-\frac{1}{2}}\right) + \frac{\log\left(1-\frac{1}{4m^2}\right)}{2} + \log m - 1 \\
&= \frac{m}{m-\frac{1}{2}} - \frac{m}{2\left(m-\frac{1}{2}\right)^2} - 1 + O(m^{-2}) + \log m \\
&= \frac{-\frac{1}{2}}{2\left(m-\frac{1}{2}\right)^2} + O(m^{-2}) + \log m = \log m + O(m^{-2})
\end{aligned}$$

where we used the Taylor expansion $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$ $(|x| < 1)$ in the form $\log(1+x) = x - \frac{x^2}{2} + O(x^3)$ $(|x| \le \frac{1}{2}$, say). Summing equation (2.5) over $m = 1, 2, \ldots, n$ we get by Exercise 2.1.1 identity (2.2):

$$\log(n!) = \sum_{m=1}^{n} \left(\int_{m-\frac{1}{2}}^{m+\frac{1}{2}} \log x \ dx + O(m^{-2}) \right) = \int_{\frac{1}{2}}^{n+\frac{1}{2}} \log x \ dx + c_1 + O(n^{-1}) \ .$$

Exercise 2.1.1. *Justify that* $\sum_{m=1}^{n} O(m^{-2}) = c_1 + O(n^{-1})$.

Evaluating the integral once more we have (again using the above Taylor expansion of logarithm)

$$\begin{aligned}
\log(n!) &= \int_{\frac{1}{2}}^{n+\frac{1}{2}} \log x \ dx + c_1 + O(n^{-1}) \\
&= (n+\tfrac{1}{2}) \log(n+\tfrac{1}{2}) - (n+\tfrac{1}{2}) + c_2 + O(n^{-1}) \\
&= n \log n - n + \tfrac{1}{2} \log n + c_3 + O(n^{-1})
\end{aligned}$$

for some constants c_i. So, by Exercise 2.1.2,

$$n! = \exp(\log n!) = \sqrt{n}\left(\frac{n}{e}\right)^n \exp\left(c_3 + O(n^{-1})\right) = \left(c + O(n^{-1})\right) \sqrt{n}\left(\frac{n}{e}\right)^n$$

where $c = \exp(c_3) > 0$. This is formula (2.1), even with the error term $o(1)$ improved to $O(n^{-1})$, but with yet undetermined constant c.

Exercise 2.1.2. *Justify that* $\exp\left(c + O(n^{-1})\right) = \exp(c) + O(n^{-1})$ *as* $n \to \infty$.

To show $c = \sqrt{2\pi}$ we derive a recurrence for the quantity

$$W_n := \int_0^{\pi/2} (\cos x)^n dx, \ n = 0, 1, 2, \dots \ .$$

We have $W_0 = \frac{\pi}{2}$ and $W_1 = 1$. Integration by parts gives

$$\begin{aligned} W_n &= \sin x \cdot (\cos x)^{n-1}|_0^{\pi/2} + (n-1)\int_0^{\pi/2} \sin^2 x \cdot (\cos x)^{n-2} \cdot dx \ . \\ &= 0 + (n-1)(W_{n-2} - W_n) \end{aligned}$$

(due to $\sin^2 x = 1 - \cos^2 x$). So, for $n \geq 2$,

$$W_n = \frac{n-1}{n} \cdot W_{n-2} \ .$$

Thus

$$\begin{aligned} W_{2n} &= \frac{(2n-1)(2n-3)\dots 1}{2n(2n-2)\dots 2} \cdot \frac{\pi}{2} = \frac{(2n)!}{(2^n n!)^2} \cdot \frac{\pi}{2} \\ W_{2n+1} &= \frac{2n(2n-2)\dots 2}{(2n+1)(2n-1)\dots 1} \cdot 1 = \frac{(2^n n!)^2}{(2n+1)!} \end{aligned}$$

and the ratio is

$$\frac{W_{2n}}{W_{2n+1}} = \frac{(2n)!^2(2n+1)}{(2^n n!)^4} \cdot \frac{\pi}{2} \ .$$

Replacing the factorials with incomplete Stirling's formulas we get that

$$\frac{W_{2n}}{W_{2n+1}} \sim \frac{2\pi}{c^2} \ .$$

It follows from the definition of W_n that $W_n < W_{n-1} < W_{n-2}$. By the recurrence,

$$1 < \frac{W_{n-1}}{W_n} < \frac{W_{n-2}}{W_n} = 1 + \frac{1}{n-1}$$

and $\frac{W_{n-1}}{W_n} \to 1$ as $n \to \infty$. Thus $\frac{2\pi}{c^2} \sim 1$ and $c = \sqrt{2\pi}$. □

2.2 The gamma function

We prove the identity (2.3). Denoting the integral in it by I_n, we have $I_0 = -e^{-x}|_{x=0}^{+\infty} = 0 - (-1) = 1$. For $n \geq 1$ integration by parts gives

$$I_n = (-e^{-x})x^n|_{x=0}^{+\infty} - \int_0^{+\infty} (-e^{-x})nx^{n-1}, \ \text{ so } \ I_n = nI_{n-1} \ .$$

By induction, $I_n = n!$ for every $n \in \mathbb{N}_0$.

We proceed to deduce Stirling's formula. We make the substitution $x = n(1+y)$,

$$n! = \int_0^{+\infty} e^{-x} x^n \, dx = e^{-n} n^{n+1} \int_{-1}^{+\infty} \left(e^{-y}(1+y)\right)^n \, dy \, .$$

This shifts the peak of the integrand from $x = n$ to $y = 0$: it is easy to check that $e^{-y}(1+y)$ on $[-1, 0]$ increases from 0 to 1, and on $[0, +\infty)$ it decreases exponentially fast from 1 to 0^+. Also,

$$\begin{aligned} e^{-y}(1+y) &= e^{-y\log(1+y)} = e^{-y^2/2+y^3/3-y^4/4+\dots} = e^{-y^2/2} e^{y^3/3-y^4/4+\dots} \\ &= e^{-y^2/2}(1+O(y^3)), \ |y| < \tfrac{1}{2} \, . \end{aligned}$$

We split the integration interval by a small $\delta = \delta(n) > 0$, which we determine later, into three subintervals ($\delta \to 0$ fast enough so that $n\delta^3 \to 0$):

$$\begin{aligned} \int_{-1}^{+\infty} \left(e^{-y}(1+y)\right)^n \, dy &= \int_{-1}^{-\delta} \dots + \int_{-\delta}^{\delta} \dots + \int_{\delta}^{+\infty} \dots \\ &= J_1 + \int_{-\delta}^{\delta} e^{-ny^2/2}(1+O(ny^3)) \, dy + J_2 \\ &= \left(1+O(n\delta^3)\right) \int_{-\delta}^{\delta} e^{-ny^2/2} \, dy + J_1 + J_2 \\ &= \left(1+O(n\delta^3)\right)(J_3 - 2J_4) + J_1 + J_2 \, , \end{aligned}$$

with

$$J_3 = \int_{-\infty}^{+\infty} e^{-ny^2/2} \, dy = \sqrt{\frac{2}{n}} \int_{-\infty}^{+\infty} e^{-t^2} \, dt = \sqrt{\frac{2\pi}{n}}, \ J_4 = \int_{\delta}^{+\infty} e^{-ny^2/2} \, dy \, .$$

We used the classic evaluation of the Gauss integral $\int_{-\infty}^{+\infty} e^{-t^2} \, dt = \sqrt{\pi}$, which we prove in the lemma below. We assume that $n\delta^3 \to 0$ but $n\delta^2 \gg 1$ (by this we mean that for a constant $c > 0$, $n\delta^2 > c$ for every $n \in \mathbb{N}$). It follows from the above observations on $e^{-y}(1+y)$ that then

$$|J_1|, \ |J_2| = O\left(e^{-n\delta^2/2}\right) \, ,$$

and this estimate holds for $|J_4|$ as well.

Exercise 2.2.1 *Prove in detail this estimate of J_1, J_2 and J_4.*

Hence, with $\delta = n^{-1/2+\varepsilon/3}$ for small $\varepsilon > 0$,

$$\int_{-1}^{+\infty} \dots = \sqrt{\frac{2\pi}{n}} + O\left(\sqrt{n}\delta^3\right) + O\left(e^{-n\delta^2/2}\right) = \sqrt{\frac{2\pi}{n}} + O\left(n^{-1+\varepsilon}\right) \, .$$

Alltogether,

$$n! = e^{-n} n^{n+1} \left(\sqrt{2\pi/n} + O\left(n^{-1+\varepsilon}\right) \right) = \left(1 + O(n^{-1/2+\varepsilon})\right) \sqrt{2\pi n} \left(\frac{n}{e}\right)^n .$$

□

Lemma 2.2.2 *It is true that*

$$\int_{-\infty}^{+\infty} e^{-t^2} \, dt = \sqrt{\pi} \, .$$

Proof. Since the integrand is a nonnegative even function, the result is equivalent with $I := \int_0^{+\infty} e^{-t^2} \, dt = \sqrt{\pi}/2$ and $I^2 = \pi/4$. By the Fubini theorem and the substitution $t = vu$ we have

$$\begin{aligned} I^2 &= \int_0^{+\infty} e^{-t^2} \, dt \int_0^{+\infty} e^{-u^2} \, du = \int_0^{+\infty} \left(\int_0^{+\infty} e^{-t^2-u^2} \, dt \right) \, du \\ &= \int_0^{+\infty} \left(\int_0^{+\infty} u e^{-u^2(1+v^2)} \, dv \right) \, du \\ &= \int_0^{+\infty} \left(\int_0^{+\infty} u e^{-u^2(1+v^2)} \, du \right) \, dv = \int_0^{+\infty} \frac{dv}{2(1+v^2)} \\ &= (\arctan(+\infty) - \arctan(0))/2 = \pi/4 \, . \end{aligned}$$

□

Guido Fubini (1879–1943) was an Italian mathematician.

Exercise 2.2.3 *What does the Fubini theorem say? Where and how we exactly used it in the calculation?*

2.3 Cauchy's formula

We prove the identity (2.4). We take Cauchy's formula (for Cauchy's personalia see later)

$$a_n = [z^n] f(z) = \frac{1}{2\pi i} \oint f(z) z^{-n-1} \, dz$$

for the coefficient of z^n in the expansion $f(z) = \sum_{n \geq 0} a_n z^n$ of a function analytic in a neighborhood of 0 in terms of the complex integral over a simple contour encircling the origin, and apply it to the exponential function

$$f(z) = e^z = \sum_{n \geq 0} \frac{z^n}{n!}$$

which is entire and has the coefficient of z^n equal to $1/n!$.

Exercise 2.3.1 *If these words make you feel uncomfortable, it is a good idea to familiarize yourself with complex analysis and its use in asymptotics: Flajolet and Sedgewick [36]. It is a wonderful world!*

We take for the contour the circle $z = re^{i\theta}$ with radius $r > 0$ and angle $\theta \in [-\pi, \pi)$:

$$\frac{1}{n!} = \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{e^{re^{i\theta}}}{(re^{i\theta})^{n+1}} \frac{d\, re^{i\theta}}{d\theta}\, d\theta = \frac{1}{2\pi} \frac{e^r}{r^n} \int_{-\pi}^{\pi} e^{r(e^{i\theta}-1)-ni\theta}\, d\theta\ .$$

The exponent has expansion $r(i\theta - \theta^2/2 + O(\theta^3)) - ni\theta$ for θ close to 0, which suggests to set $r = n$ (thus the integration path will depend on n, after all). We denote

$$f(n,\theta) = n(e^{i\theta} - 1) - ni\theta = -\tfrac{n\theta^2}{2} + \tfrac{n(i\theta)^3}{6} + \cdots = -\tfrac{n\theta^2}{2} + O(n\theta^3),\ |\theta| < \tfrac{1}{2}\ .$$

We split the integral very similarly to the previous proof. Let $\varphi = \varphi(n) \in (0, \pi)$ be a small angle such that $n\varphi^3 \to 0$ but $n\varphi^2 \gg 1$. Then

$$\begin{aligned}
\int_{-\pi}^{\pi} e^{n(e^{i\theta}-1)-ni\theta}\, d\theta &= \int_{-\varphi}^{\varphi} e^{f(n,\theta)}\, d\theta + \int_{\varphi}^{\pi} \left(e^{f(n,\theta)} + e^{f(n,-\theta)}\right)\, d\theta \\
&= \int_{-\varphi}^{\varphi} e^{-n\theta^2/2}(1 + O(n\theta^3)\, d\theta + I \\
&= \big(1 + O(n\varphi^3)\big) \int_{-\varphi}^{\varphi} e^{-n\theta^2/2}\, d\theta + I \\
&= \big(1 + O(n\varphi^3)\big) \left(\sqrt{2\pi/n} + O\left(e^{-n\varphi^2/2}\right)\right) + I\ ,
\end{aligned}$$

where we calculated the Gauss integral and estimated its tail exactly as in the previous proof. As for I, if $\theta \in [\varphi, \pi]$ then

$$\left|e^{f(n,\theta)}\right| = \left|e^{f(n,-\theta)}\right| = e^{n(\cos\theta - 1)} \le e^{n(\cos\varphi - 1)} \le e^{-n\varphi^2/2}$$

(by the Taylor expansion of $\cos\varphi$ for small $\varphi > 0$). Thus for $|I|$ we have the same bound as for the Gauss tail. Setting $\varphi = n^{-1/2+\varepsilon/3}$ as before we get by a very similar calculation the same estimate

$$\frac{1}{n!} = \frac{1}{2\pi} \frac{e^n}{n^n} \left(\sqrt{\frac{2\pi}{n}} + O(n^{-1+\varepsilon})\right) = \big(1 + O(n^{-1/2+\varepsilon})\big) \frac{1}{\sqrt{2\pi n}} \left(\frac{e}{n}\right)^n\ .$$

□

2.4 Comments and references

Our exposition of the three proofs is based on de Bruijn [15] and Flajolet and Sedgewick [36]. One can find in the literature and on the Internet many articles proving Stirling's formula. For example, just the *American Mathematical*

Monthly contains the articles by de Angelis [6], Blyth and Patak [12], Coleman [21], Diaconis and Freedman [25], Dutkay et al. [27], Feller [34], Impens [43], Khan [45], Lou [50], Maria [53], Marsaglia and Marsaglia [54], Michel [56] and [57], Namias [58], Nanjundiah [59], Neuschel [61], Patin [63], Pinsky [65], Robbins [67], and Romik [68]. The book [36] contains five proofs of Stirling's formula. Three of them make use of the formulas (2.2), (2.3) and (2.4). What are the representations of $n!$ in the other two proofs?

Chapter 3

$0-1$ laws and the Blatter–Specker theorem

Section 3.1 presents the general result (Theorem 3.1.1) of B. Bollobás and A. Thomason that for every increasing (closed to supersets) property of subsets of an n-element set X there exists a so called threshold function $t = t(n)$: as n goes to ∞, $m = m(n)$-element subsets of X for m much smaller than t almost surely do not have the property, but those for m much larger than t almost surely do. This applies to properties of graphs (Corollary 3.1.3). Key role in the proof plays the Kruskal–Katona theorem (Theorem 3.1.8) from extremal theory of set systems.

Section 3.2 deals with enumerative $0-1$ laws for relational structures. These assert that asymptotically (when the size of the universe goes to ∞) either almost all such structures have a certain property or almost all do not. R. Fagin derived a simple asymptotic relation (Theorem 3.2.2) between labeled and unlabeled count of relational structures. Their properties are best defined in terms of formulae of mathematical logic, and we introduce and explain this approach here. One of the earliest $0-1$ laws of this type, from 1950, is due to R. Carnap who states it at the end of his tract [18]. It is Theorem 3.2.10, a first-order $0-1$ law for unlabeled count of relational structures with only unary predicates. More famous is the result in Theorem 3.2.11, associated with the names of Ju. V. Glebskij, D. I. Kogan, M. I. Liogon'kij and V. A. Talanov and the name of R. Fagin: a first-order $0-1$ law for numbers of general relational structures. To simplify its exposition we prove in detail only the version for undirected simple graphs (Theorem 3.2.12). The remarkable proof is an enumerative application of the completeness theorem for the first-order logic (here without functions), Theorem 3.2.13, which is due to K. Gödel. We follow the exposition in the book [71] of J. Spencer, who by his own words was much inspired by R. Fagin's original article [33]. Unfortunately, both sources omit the crucial and most interesting part of the proof, namely the proof of the completeness theorem itself. We fill this lacuna and offer to the interested and patient reader a proof of

Gödel's Theorem 3.2.13.
Section 3.3

3.1 The Bollobás–Thomason theorem: thresholds exist

For $n \in \mathbb{N}$ we will work with the sets $[n] = \{1, 2, \dots, n\}$, $[n]_0 = \{0, 1, \dots, n\}$ and $\mathcal{P}([n]) = \{X \mid X \subset [n]\}$. A *property Q (of subsets of $[n]$)* is any subset

$$Q \subset \mathcal{P}([n]) .$$

One could call Q also a hypergraph with vertices $[n]$. We say that Q is *non-trivial* if $Q \neq \emptyset, \mathcal{P}([n])$, and that it is *increasing* (resp. *decreasing* or an *ideal*) if for every $A, B \in \mathcal{P}([n])$ with $A \subset B$, $A \in Q$ implies $B \in Q$ (resp. $B \in Q$ implies $A \in Q$). For any property Q and $m \in [n]_0$ we define $Q_m = Q \cap \binom{[n]}{m}$ where $\binom{[n]}{m}$ denotes the set of m-element subsets of $[n]$. The ratio

$$P_m(Q) := \frac{|Q_m|}{\left|\binom{[n]}{m}\right|} = \frac{|Q_m|}{\binom{n}{m}}$$

is called the *probability (of Q_m)* (it is the probability that a random element of $\binom{[n]}{m}$ has property Q). We allow that Q depends on n and consider sequences $Q = (Q^n) = (Q^1, Q^2, \dots)$ of properties $Q^n \subset \mathcal{P}([n])$. We say that $t(n) \in [n]$ is a *threshold function* for (Q^n) if for any function $m(n) \in [n]_0$ asymptotically larger (resp. smaller) than $t(n)$, the probability $P_{m(n)}(Q^n)$ goes to 1 (resp. goes to 0). The general result on existence of threshold functions is as follows.

Theorem 3.1.1 (Bollobás–Thomason, 1986). *Let $Q = (Q^n)$ where every property $Q^n \subset \mathcal{P}([n])$, $n = 1, 2, \dots$, is non-trivial and increasing. Then there exists a function $t : \mathbb{N} \to \mathbb{N}$, $t(n) \in [n]$ for every $n \in \mathbb{N}$, such that for every function $m : \mathbb{N} \to \mathbb{N}_0$, $m(n) \in [n]_0$ for every $n \in \mathbb{N}$,*

$$\lim_{n\to\infty} \frac{m(n)}{t(n)} = +\infty \Rightarrow \lim_{n\to\infty} P_{m(n)}(Q^n) = 1$$

and

$$\lim_{n\to\infty} \frac{m(n)}{t(n)} = 0 \Rightarrow \lim_{n\to\infty} P_{m(n)}(Q^n) = 0$$

— $t(n)$ is the threshold function for property Q.

Exercise 3.1.2. *Statement of the theorem excludes the trivial increasing properties $\emptyset$ and $\mathcal{P}([n])$. Could one include them?*

We apply the theorem to graphs, which are systems of two-element sets (edges). For $m \in \mathbb{N}_0$ and $n \in \mathbb{N}$ we let

$$G(m, n) = \binom{\binom{[n]}{2}}{m}$$

denote the set of all graphs on $[n]$ with m edges. A property $Q \subset \mathcal{P}(\binom{[n]}{2})$ of graphs on $[n]$ (a *graph G (on $[n]$)* is any subset $G \subset \binom{[n]}{2}$) is non-trivial if it is non-empty and some graph does not lie in it, and it is increasing if it is closed to supergraphs (if G, H are graphs on $[n]$ with $G \subset H$ and $G \in Q$ then $H \in Q$). For example, connected graphs — graphs G on $[n]$ such that $[n]$ cannot be partitioned in two blocks so that every edge in G lies completely in one block — form an increasing and (for $n > 1$) non-trivial property.

Corollary 3.1.3. *Every sequence $Q = (Q^n)$, $n \in \mathbb{N}$, of non-trivial and increasing graph properties $Q^n \subset \mathcal{P}(\binom{[n]}{2})$ has a threshold function, a function $t(n)$, $1 \le t(n) \le \binom{n}{2}$, such that for every function $m(n)$, $0 \le m(n) \le \binom{n}{2}$,*

$$\lim_{n\to\infty} \frac{m(n)}{t(n)} = +\infty \text{ (resp. } = 0) \Rightarrow \lim_{n\to\infty} \frac{|Q^n \cap G(m(n), n)|}{|G(m(n), n)|} = 1 \text{ (resp. } = 0) .$$

Proof. For $n \in \mathbb{N}$ let $N = N(n) = \binom{n}{2}$. After fixing a bijection between $[N]$ and $\binom{[n]}{2}$ we also have a bijection between $\mathcal{P}([N])$ and graphs on $[n]$. It follows that the property $R^N \subset \mathcal{P}([N])$ of subsets of $[N]$ corresponding to Q^n is non-trivial and increasing and, for any $m(n)$ with $0 \le m(n) \le \binom{n}{2}$, $P_{m(n)}(R^N)$ equals to the above displayed ratio. The result therefore follows from Theorem 3.1.1 (in (R^N) for N not of the form $N = \binom{n}{2}$ we let R^N be any non-trivial and increasing property of subsets of $[N]$). □

Exercise 3.1.4. *Show that the property of graphs on $[n]$ "to have exactly n edges" has no threshold function. Find a property of graphs with threshold function $t(n) = n$.*

We begin the proof of Theorem 3.1.1. Clearly, if $Q \subset \mathcal{P}([n])$ is increasing then Q is non-trivial if and only if $\emptyset \notin Q$ and $[n] \in Q$. Thus for non-trivial and increasing Q we have $P_0(Q) = 0$ and $P_n(Q) = 1$. We show that then $P_m(Q)$ is a non-decreasing function of $m \in [n]_0$.

Proposition 3.1.5. *Let $n \in \mathbb{N}$ and $m \in [n]$. If $Q \subset \mathcal{P}([n])$ is increasing then*

$$P_{m-1}(Q) \le P_m(Q) .$$

Proof. Double counting! We have, first grouping by A and then by B ($B \subset [n]$),

$$|Q_{m-1}|(n - m + 1) = |\{(A, B) \mid A \in Q_{m-1}, A \subset B, |B\backslash A| = 1\}| \le m|Q_m|$$

(note that every B lies in Q_m since Q is increasing). If $|Q_m| = 0$ then also $|Q_{m-1}| = 0$ and the inequality holds as both probabilities are 0. If $|Q_m| > 0$, we get $\frac{|Q_{m-1}|}{|Q_m|} \le \frac{m}{n-m+1}$ and

$$\frac{P_{m-1}(Q)}{P_m(Q)} = \frac{|Q_{m-1}|/\binom{n}{m-1}}{|Q_m|/\binom{n}{m}} = \frac{|Q_{m-1}|}{|Q_m|} \cdot \frac{n - m + 1}{m} \le 1 .$$

□

We need the folowing representation of natural numbers.

Proposition 3.1.6 (cascade form). *For every $m, k \in \mathbb{N}$ there exists a unique sequence of integers $n_0 > n_1 > \cdots > n_l \geq 1$, $l \in \mathbb{N}_0$ and $k - l \geq 1$, such that*

$$m = \binom{n_0}{k} + \binom{n_1}{k-1} + \cdots + \binom{n_l}{k-l} .$$

We call it the cascade form of m.

Proof. We initialize the variables i, M, K as $i := 0, M := m, K := k$, set n_i to the largest $n \in \mathbb{N}$ such that $\binom{n}{K} \leq M$, terminate if $M - \binom{n_i}{K} = 0$, and else actualize the variables by $M := M - \binom{n_i}{K}$, $i := i+1$, $K := K-1$ and repeat. For $K = 1$ the zero difference is always achieved and thus the algorithm terminates. It is also clear that we get representation $m = \sum_{i=0}^{l} \binom{n_i}{k-i}$, $k - l \geq 1$. We leave the proof of the rest (the n_i decrease and the cascade form is unique) as Exercise 3.1.7. □

Exercise 3.1.7. *Prove that the n_i generated by the above greedy algorithm satisfy $n_0 > n_1 > \cdots > n_l \geq 1$ and that the cascade form of m is unique.*

For every $k \in \mathbb{N}$ we then can define the function $f_k : \mathbb{N} \to \mathbb{N}$ by

$$f_k(m) = \binom{n_0}{k-1} + \binom{n_1}{k-2} + \cdots + \binom{n_l}{k-l-1}$$

where $m = \sum_{i=0}^{l} \binom{n_i}{k-i}$ is the cascade form of m.

To prove Theorem 3.1.1 we need the following nice result from extremal theory of set systems. For $Q \subset \mathcal{P}([n])$ we define $\Delta(Q) \subset \mathcal{P}([n])$ by

$$\Delta(Q) = \{A \backslash \{a\} \mid a \in A \in Q\} .$$

If $Q = \{F\}$ and this is known from the context, we simplify $\Delta(\{F\})$ to $\Delta(F)$.

Theorem 3.1.8 (Kruskal, 1963; Katona, 1968). *Let $n \in \mathbb{N}$, $k \in [n]$ and $\emptyset \neq Q \subset \binom{[n]}{k}$. Then*

$$|\Delta(Q)| \geq f_k(|Q|) .$$

Proof. (Frankl, 1984.) In the first part of the proof we formulate and prove a lemma on exchange of two elements in edges of set systems. For $n \in \mathbb{N}$ with $n > 1$, two distinct $a, b \in [n]$ and $A \subset [n]$ we define $A_a^b := (A \backslash \{a\}) \cup \{b\}$ if $a \in A \,\&\, b \notin A$, and else leave A_a^b undefined. For $j \in [n]$ with $j > 1$ and $Q \subset \mathcal{P}([n])$ we define

$$S_j(Q) = \{S_j(A) = S_{j,Q}(A) \mid A \in Q\}$$

where $S_j(A) = A_j^1$ if A_j^1 is defined and $A_j^1 \notin Q$, and $S_j(A) = A$ else. Thus for all $A \in Q$ simultaneously we exchange j for 1 if it is possible and yields a new

set, and else leave A as it is. Note that $S_j(A)$ also depends on the set system Q the set A belongs to. The lemma asserts that always

$$\Delta(S_j(Q)) \subset S_j(\Delta(Q)) \ .$$

We prove it by showing for every $F \in Q$ the inclusion

$$\Delta(S_j(F)) \subset S_j(\Delta(Q)) \ .$$

On the left sides we have in fact $S_{j,Q}$ and on the right sides $S_{j,\Delta(Q)}$ but we write just S_j and hope it will be always clear which of the two is meant. We distinguish four cases. The first one is when $j \notin F$. Then $S_j(F) = F$ and for every $E \in \Delta(F)$ one has $E \in \Delta(F) = S_j(\Delta(F)) \subset S_j(\Delta(Q))$. The second case is when $j \in F \,\&\, 1 \in F$. Then again $S_j(F) = F$ and every $E \in \Delta(F)$ with $1 \in E$ again lies in $S_j(\Delta(F)) \subset S_j(\Delta(Q))$, since $E = S_j(E)$. If $E \in \Delta(F)$ with $E = F\backslash\{1\}$ then E lies in $S_j(\Delta(Q))$ too as $S_j(F\backslash\{1\}) = F\backslash\{1\} = E$ because $(F\backslash\{1\})^1_j = F\backslash\{j\} \in \Delta(F) \subset \Delta(Q)$. The third case is when $j \in F \,\&\, 1 \notin F$ but still $S_j(F) = F$ because $F^1_j \in Q$. If $E \in \Delta(F)$ with $E = F\backslash\{j\}$ then $E = S_j(E)$ and E lies in $S_j(\Delta(F)) \subset S_j(\Delta(Q))$. If $E \in \Delta(F)$ with $j \in E$ then $E = S_j(E)$ as $E^1_j \in \Delta(F^1_j) \subset \Delta(Q)$ and so $E \in S_j(\Delta(F)) \subset S_j(\Delta(Q))$. The final fourth case is when $j \in F \,\&\, 1 \notin F$ and $S_j(F) = F^1_j$ because $F^1_j \notin Q$. For $E \in \Delta(S_j(F)) = \Delta(F^1_j)$ with $1 \in E$ we take $G = E^j_1 \in \Delta(F)$. Thus $E = G^1_j$. If $E \in \Delta(Q)$ then E belongs to $S_j(\Delta(Q))$ since $E = S_j(E)$. If $E \notin \Delta(Q)$ then $E = S_j(G)$ and again E belongs to $S_j(\Delta(F)) \subset S_j(\Delta(Q))$. If $1 \notin E$, that is $E = F\backslash\{j\}$, then $E = S_j(E)$ and again E belongs to $S_j(\Delta(F)) \subset S_j(\Delta(Q))$. The lemma is proven.

Note that $|S_j(Q)| = |Q|$ and $|S_j(\Delta(Q))| = |\Delta(Q)|$ because the mapping $A \mapsto S_j(A)$ is injective. This and the lemma implies that $|\Delta(Q)| \geq |\Delta(S_j(Q))|$. Also, replacement of Q with $S_j(Q)$ does not change sizes of edges but increases (if $S_j(Q) \neq Q$) the number of edges containig 1. Hence sufficiently many replacements transform the given set system $Q \subset \mathcal{P}([n])$ in a set system $Q' \subset \mathcal{P}([n])$ whose edges have the same sizes and satisfies $|Q| = |Q'|$, $|\Delta(Q)| \geq |\Delta(Q')|$ and $S_j(Q') = Q'$ for every $j \in [n]$ with $j > 1$. The last property of Q' is equivalent with

$$E \in \Delta(F),\ F \in Q',\ 1 \notin F \Rightarrow E \cup \{1\} \in Q' \ .$$

Let $k, n \in \mathbb{N}$ and $Q \subset \binom{[n]}{k}$ be nonempty. In the second part of the proof we establish the claimed inequality by induction on $|Q|$ and k. For $k = 1$ and every $|Q|$ it holds: $|\Delta(Q)| = 1$ and $f_1(|Q|) = 1$. We assume that $k \geq 2$, $|Q| = \sum_{i=0}^{l} \binom{n_i}{k-i}$ is the cascade form of $|Q|$ and that the inequality holds for all smaller k and any $|Q|$. By the above reduction of Q to Q' we may assume that

$$E \in \Delta(F),\ F \in Q,\ 1 \notin F \Rightarrow E \cup \{1\} \in Q \ .$$

Let $Q(1) = \{F\backslash\{1\} \mid 1 \in F \in Q\}$. We have

$$|\Delta(Q)| \geq |Q(1)| + |\Delta(Q(1))|$$

(Exercise 3.1.9). □

Exercise 3.1.9. *Prove that for any system Q of k-element subsets of $[n]$, $k \geq 2$, and with $Q(1) = \{F\backslash\{1\} \mid 1 \in F \in Q\}$ we have $|\Delta(Q)| \geq |Q(1)| + |\Delta(Q(1))|$. Is it true for $k = 1$? For which other systems $Q \subset \mathcal{P}([n])$ this bound holds?*

Proposition 3.1.10. *Let $n, k \in \mathbb{N}$, $n \geq k+1$, $A_n = \binom{n}{k}$ and $B_n = \binom{n}{k-1}$. Then for $x \in \mathbb{Z}$,*

$$0 \leq x \leq A_n - A_{n-1} = B_{n-1}$$
$$\Rightarrow f_k(A_{n-1} + x) \geq B_{n-1} + \frac{B_n - B_{n-1}}{A_n - A_{n-1}} x = B_{n-1} + \frac{k-1}{n-k+1} x \; .$$

Proof.

□

3.2 The GKLT–F theorem: enumerative first-order $0-1$ laws

A *type*, or a *signature*, is any tuple $S = (S_1, \ldots, S_k) \in \mathbb{N}^k$. A *relational structure R of type S on a set X*, briefly an *S-structure on X*, is any k-tuple

$$R = (X_1, \ldots, X_k) \;\text{ where }\; X_i \subset X^{S_i} = X \times X \times \cdots \times X \;\; (S_i \text{ components})$$

is an S_i-ary relation on X. So (as before, $\mathcal{P}(A) = \{B \mid B \subset A\}$)

$$\{R \mid R \text{ is an } S\text{-structure on } X\} = \prod_{i=1}^{k} \mathcal{P}(X^{S_i}) \; .$$

For example, an undirected simple graph with two-colored vertices $1, 2, \ldots, n$ is a $(2,1)$-structure on $[n]$ such that the binary relation is irreflexive and symmetric. Clearly, the number $l(|X|, S)$ of S-structures on a finite set X is given by

$$l(|X|, S) = 2^{\sum_{i \geq 1} u_i |X|^i}, \;\; u_i = \#\{j \in [k] \mid S_j = i\} \in \mathbb{N}_0 \; .$$

Two S-structures, $R = (X_1, \ldots, X_k)$ on X and $R' = (X'_1, \ldots, X'_k)$ on X', are *isomorphic*, written $R \cong R'$, if a bijection $F : \; X \to X'$ satisfies

$$(a_1, a_2, \ldots, a_{S_j}) \in X_j \iff (F(a_1), F(a_2), \ldots, F(a_{S_j})) \in X'_j, \;\; j = 1, 2, \ldots, k \; .$$

The binary relation $\cong$ of isomorphism is an equivalence relation on the set of S-structures on X. *Unlabeled S-structure on X* refers to an equivalence class of $\cong$ while *labeled S-structure on X* refers to an individual S-structure on X, regardless of isomorphism. We let $u(|X|, S)$ denote the number of unlabeled S-structures on a finite set X. We show that for $|X| = n$ almost all equivalence classes of $\cong$ have size close to $n!$ and $l(n, S) \sim n! \cdot u(n, S)$, except for $S = (1, 1, \ldots, 1)$.

Exercise 3.2.1. *Find a formula for $u(n,(1,1,\dots,1))$ (k 1s).*

Theorem 3.2.2 (Fagin, 1977). *Let $S=(S_1,\dots,S_k)$ be a type different from $(1,1,\dots,1)$. Then*

$$\lim_{n\to\infty}\frac{n!\cdot u(n,S)}{l(n,S)}=\lim_{n\to\infty}\frac{n!\cdot\left|\prod_{i=1}^{k}\mathcal{P}([n]^{S_i})/\cong\right|}{\left|\prod_{i=1}^{k}\mathcal{P}([n]^{S_i})\right|}=1$$

—for $n\to\infty$ the number of labeled S-structures on $[n]$ is asymptotic to the number of unlabeled ones multiplied by $n!$.

The proof uses the next popular lemma.

Lemma 3.2.3 (Burnside, 1897; Frobenius, 1887). *Let X be a finite set, G be a finite group of bijections $g:\ X\to X$, equivalence $\sim$ on X be defined by $a\sim b$ iff $a=g(b)$ for some $g\in G$, and for any $g\in G$ let $F(g)=|\{a\in X\mid g(a)=a\}|$. Then*

$$|X/\!\sim|=\frac{1}{|G|}\sum_{g\in G}F(g)\ .$$

Proof. It is clear that $\sim$ is an equivalence relation (Exercise 3.2.4). For a fixed $x\in X$ with $x\in B_x\in X/\!\sim$, double counting gives

$$|G|=|\{(g,y)\in G\times X\mid g(x)=y\}|=|G_x|\cdot|B_x|$$

where $G_x=\{g\in G\mid g(x)=x\}$. The second displayed equality follows from the fact that for any $y\in B_x$ the sets $M_y=\{g\in G\mid g(x)=y\}$ and G_x are in bijection, given by $g\mapsto h^{-1}g$ where $h\in M_y$ is arbitrary but fixed. Thus $|G_x|=|G|/|B_x|$ and another double counting gives

$$\begin{aligned}\sum_{g\in G}F(g) &= |\{(g,x)\in G\times X\mid g(x)=x\}|=\sum_{x\in X}|G_x|=|G|\sum_{x\in X}\frac{1}{|B_x|}\\ &= |G|\cdot|X/\!\sim|\ .\end{aligned}$$

□

Exercise 3.2.4. *Check that the relation $\sim$ defined in the statement of the lemma is an equivalence.*

Exercise 3.2.5. *For a random permutation $\pi:\ [n]\to[n]$, what is the expected number of fixed points? Note for fans of the probabilistic method: it is forbidden to use linearity of expectation.*

We prove Theorem 3.2.2. We may assume that $S_1\geq 2$. For $i,n\in\mathbb{N}$, each permutation $\pi:\ [n]\to[n]$ induces a permutation $\pi_i:\ [n]^i\to[n]^i$ via $\pi_i((a_1,\dots,a_i))=(\pi(a_1),\dots,\pi(a_i))$. We denote by $C_i(\pi)\in\mathbb{N}$ the number of

cycles of π_i. Since for an (i)-structure on $[n]$ the permutation π is its automorphism iff it is a union of some cycles of π_i, we see that for a given permutation $\pi : [n] \to [n]$ the number of labeled S-structures on $[n]$ for which π is an automorphism equals

$$N(\pi) := 2^{\sum_{i\geq 1} u_i C_i(\pi)}$$

where as before u_i counts components i in S. By Lemma 3.2.3 (applied to $G = \{(\pi_{S_1}, \ldots, \pi_{S_k}) \mid \pi : [n] \to [n] \text{ a perm.}\}$ and $X = \prod_{j=1}^{k} [n]^{S_j}$), for $n \in \mathbb{N}$ we have a formula for the number of unlabeled structures:

$$u(n, S) = \frac{1}{n!} \sum_{\pi} N(\pi) = \frac{1}{n!} \sum_{\pi} 2^{\sum_{i\geq 1} u_i C_i(\pi)}$$

and we need to prove that for $n \to \infty$,

$$\frac{n! \cdot u(n, S)}{l(n, S)} = \sum_{\pi} 2^{\sum_{i\geq 1} u_i (C_i(\pi) - n^i)} \to 1 \, .$$

For $\pi = \mathrm{id}_n$, the identical permutation, $C_i(\pi) = n^i$ for every i and the summand equals 1. We show that for any $\pi \neq \mathrm{id}_n$ the sumand is negligible. □

Let $S = (S_1, \ldots, S_k) \in \mathbb{N}^k$ be a signature and let $S_0 = 2$. We introduce a formal but quite general way of precise definitions for many classes of S-structures on finite or infinite sets. The classes will consist of models of first-order closed formulas. The formulas are special words (i.e., finite sequences of symbols) over a countable alphabet $\mathcal{A} = \mathcal{A}(S)$. The alphabet consists of three auxiliary and seven logical symbols, $k{+}1$ predicate symbols $V = \{V_0, V_1, \ldots, V_k\}$ (V_i has arity S_i), where we write 'V_0' also as '=', and countably many symbols $N = \{x_i \mid i \in \mathbb{N}\}$ for variables:

$$\begin{aligned} \mathcal{A} &= \{\ ,\ (\ )\ \neg\ \vee\ \wedge\ \to\ \leftrightarrow\ \exists\ \forall\ =\ V_1\ \ldots\ V_k\ x_1\ x_2\ \ldots\ \} \\ &= \{\ ,\ (\ )\ \neg\ \vee\ \wedge\ \to\ \leftrightarrow\ \exists\ \forall\ \} \cup V \cup N \end{aligned}$$

(since the comma ',' belongs to $\mathcal{A}$, we separate elements of $\mathcal{A}$ only by spaces, and the two elipses '...' of course are not elements of $\mathcal{A}$). The first ten symbols are, respectively, comma, left bracket, right bracket, negation, disjunction, conjuction, implication, equivalence, existential quantifier and general quantifier. An *atomic formula (of the language $L(S)$)* is any word over $\mathcal{A}$ of the form

$$V_i(y_1, y_2, \ldots, y_{S_i}) \ \text{ where } i \in [k]_0 \text{ and so } V_i \in V, \text{ and } y_1, \ldots, y_{S_i} \in N \, .$$

This word has $2S_i + 2$ symbols from $\mathcal{A}$. For binary predicates, when $S_i = 2$, we use the synonymous notation $y_1 V_i y_2$ that saves three symbols. If $S = (2, 3, 1)$ then

$$x_5 = x_{24} \ \ V_3(x_2) \ \ V_2(x_3, x_3, x_1) \ \ V_1(x_4, x_4) \ \ x_6 V_1 x_5 \ \ V_0(x_1, x_2)$$

are examples of six atomic formulas from $L(S)$. From now on we return to using comma as a separator and hope that the reader will have no troubles to

distinguish its usages belonging to the language $L(S)$ over $\mathcal{A}$ from those that are part of the metalanguage.

Formulas (of type S) or the *language* $L(S) \subset \mathcal{A}^*$, where $\mathcal{A}^*$ denotes the set of all words over the alphabet $\mathcal{A}$, is the smallest (to $\subset$) subset $L(S)$ of $\mathcal{A}^*$ that contains all atomic formulas and is closed to the production rules

$$x \in N,\, F,\, G \in L(S) \quad \Rightarrow \quad (\neg\, F),\ (F \vee G),\ (F \wedge G),\ (F \to G),\ (F \leftrightarrow G),\\ (\exists\, x\, F),\ (\forall\, x\, F) \in L(S)$$

(on the right side of $\Rightarrow$ all commas belong to the metalanguage but all brackets to $L(S)$).

Exercise 3.2.6. *Prove that such (unique) smallest set $L(S) \subset \mathcal{A}^*$ of formulas exists.*

It is clear that every formula is atomic or uniquely decomposes in one or two simpler (that is, shorter) formulas, but one should prove it because for other production rules generating other languages such decomposition need not hold.

Exercise 3.2.7. *Prove that for every formula $F \in L(S)$ there exist formulas $G, H \in L(S)$, variable $x \in N$, connective*

$$\circ \in \{\wedge, \vee, \to, \leftrightarrow\}$$

and quantifier $Q \in \{\forall, \exists\}$ such that F is atomic or

$$F \text{ is } (\neg\, G) \text{ or } (G \circ H) \text{ or } (Q\, x\, G)$$

and this decomposition of F is unique (including the case of atomic F).

For example, if $S = (2)$ and we write '$\sim$' for 'V_1' then the word $\varphi_{usg} \in \mathcal{A}(S)^*$ with 30 symbols given as

$$((\forall\, x_1\, (\neg\, x_1 \sim x_1)) \wedge (\forall\, x_1\, (\forall\, x_2\, (x_1 \sim x_2 \to x_2 \sim x_1))))$$

is a formula in $L(S)$. The formula φ_{usg} is built by connectives and quantifiers from three atomic formulas and decomposes uniquely as $(G \wedge H)$. It asserts (via the standard intepretation which we discuss later) irreflexivity and symmetry of the binary predicate V_1. An S-structure with property φ_{usg} is an undirected simple graph. But before turning to interpretation we need to deal with two syntactic points.

We simplify brackets in a formula by omitting the outermost pair and applying two conventions. One orders the connectives and quantifiers according to decreasing strength of bond: Q, $\neg$, $\wedge$, $\vee$, $\to$ and $\leftrightarrow$; we parse formulas with missing brackets according to it. By another convention, without brackets we understand connectives of the same strength or quantifiers as bracketed from the right. For example, if $H, G, F \in L(S)$ and $x \in N$ then

$$\exists\, x \neg\, F \vee G \ \text{ means } \ ((\exists\, x\, (\neg\, F)) \vee G)$$

and not, for example, $(\exists\, x\, (\neg\, (F \vee G)))$, and

$$F \to G \to H \;\text{ means }\; (F \to (G \to H))\ .$$

The formula φ_{usg} thus simplifies to

$$\forall\, x_1\, \neg\, x_1 \sim x_1 \wedge \forall\, x_1\, \forall\, x_2\, (x_1 \sim x_2 \to x_2 \sim x_1)$$

and the remaining brackets cannot be omitted without altering the intended sense of φ_{usg}.

The second and more important point concerns occurrences of variables in formulas. Let $F \in L(S)$ and $x \in N$. We say that x *has an occurrence in* F if some symbol of F (as a word in $\mathcal{A}^*$) equals x; this is then an *occurrence of* x *in* F. Clearly, only finitely many variables occur in any formula. For example, x_2 has three occurrences in φ_{usg}, x_1 six and x_n for $n > 2$ none. We distinguish *free* and *bound* occurrences of x in F. One variable may have both. We define them by induction on the decomposition of a formula (Exercise 3.2.7). If F is atomic and equals $V_i(y_1, \ldots, y_{S_i})$ with $V_i \in V$ and $y_1, \ldots, y_{S_i} \in N$, then the set of variables having an occurrence in F is $\{y_1, \ldots, y_{S_i}\}$ and all their occurrences in F (S_i in total) are free. Similarly, if F is $\neg\, G$ or $G \circ H$ then all variables having an occurrence in F are those having it in G or in H and the type of an occurrence in G and H remains the same in F. Critically, if F is $Q\, x\, G$ then the variables occurring in F are those occurring in G plus x, and we define all occurrence of x in F as bound (regardless that before some might have been free in G). A formula is *closed* if it has only bound occurrences of any variable. For example, φ_{usg} is a closed formula but the formula $(\exists\, x\, x = x) \leftrightarrow x = x$ (here $x \in N$), which is the same as $\exists\, x\, x = x \leftrightarrow x = x$, is not closed because the last two occurrences of x are free.

We introduce interpretation of formulas and the symbols $\models_f$ and $\models$. Let $S \in \mathbb{N}^k$ be a type, $\varphi \in L(S)$ be a formula, $R = (X_1, \ldots, X_k)$ be an S-structure on a set X, let $X_0 = \{(a, a) \mid a \in X\}$, and $f :\ N \to X$ be a mapping assigning to any variable in $\mathcal{A}(S)$ an element in the universe X of R. We define the relation between S-structures and formulas

$$R \models_f \varphi$$

(its refutation is written as $\not\models_f$), in words R *is a model for* φ *relative to* f, by induction on the decomposition of φ. If φ is atomic, equal to $V_i(y_1, \ldots, y_{S_i})$, we set

$$R \models_f \varphi \iff (f(y_1), \ldots, f(y_{S_i})) \in X_i\ .$$

If φ is $\neg\, G$ then $R \models_f \varphi$ iff $R \not\models_f G$. If φ is $G \vee H$ then $R \models_f \varphi$ iff $R \models_f G$ or $R \models_f H$. If φ is $G \wedge H$ then $R \models_f \varphi$ iff $R \models_f G$ and $R \models_f H$. If φ is $G \to H$ then $R \not\models_f \varphi$ iff $R \models_f G$ and $R \not\models_f H$. If φ is $G \leftrightarrow H$ then $R \models_f \varphi$ iff $R \models_f G$ if and only if $R \models_f H$. This of course agrees with the standard interpretation of the propositional connectives. Critically, if φ is $\exists\, x\, G$ then we define

$$R \models_f \varphi \iff R \models_g G \;\text{ for some } g :\ N \to X \text{ with } g(y) = f(y) \text{ if } y \neq x$$

and for φ equal to $\forall\, x\, G$ we define

$$R \models_f \varphi \iff R \models_g G \text{ for every } g:\ N \to X \text{ with } g(y) = f(y) \text{ if } y \neq x\ .$$

The model relation in fact depends only on the free occurrences in φ.

Exercise 3.2.8. *Prove that if $f, g:\ N \to X$ are such that $f(x) = g(x)$ for every variable $x \in N$ with a free occurrence in φ then*

$$R \models_f \varphi \iff R \models_g \varphi\ .$$

Thus for closed φ the mapping f is completely irrelevant and we may define

$$R \models \varphi \iff R \models_f \varphi \text{ for any } f\ .$$

We say then that R *is a model of (the closed formula)* φ. Model relation cannot distinguish isomorphic structures.

Exercise 3.2.9. *Prove that if R and R' are two isomorphic S-structures and $\varphi \in L(S)$ is a closed formula then*

$$R \models \varphi \iff R' \models \varphi\ .$$

Finally, if $S \in \mathbb{N}^k$ is a type, $\varphi \in L(S)$ is a closed formula and $n \in \mathbb{N}$ is a number, we define notation (the second definition is unambiguous by Exercise 3.2.9)

$$\begin{aligned} l(n, S, \varphi) &= \#(\text{labeled } S\text{-structures on } [n] \text{ such that } R \models \varphi) \\ u(n, S, \varphi) &= \#(\text{unlabeled } S\text{-structures on } [n] \text{ such that } R \models \varphi)\ . \end{aligned}$$

This completes all definitions needed to understand statements of the three Theorems 3.2.10, 3.2.11 and 3.2.12 below, but we still have not introduced all tools needed for their proofs.

We give an example for the numbers $l(n, S, \varphi)$ and $u(n, S, \varphi)$. Let $S = (2)$ and as before we write '$\sim$' for 'V_1'. Consider the closed formula $\varphi_{mat} \in L(S)$ given as (we abbreviate $\neg\, x = y$ by $x \neq y$)

$$\neg\, \exists\, x_1\, \exists\, x_2\, \exists\, x_3\, (x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_2 \neq x_3 \wedge x_1 \sim x_2 \wedge x_1 \sim x_3)\ .$$

We leave it to the interested reader to check as an exercise that

$$\begin{aligned} l(n, S, \varphi_{usg} \wedge \varphi_{mat}) &= \sum_{k \geq 0} \binom{n}{2k} \cdot (2k-1)!! \\ u(n, S, \varphi_{usg} \wedge \varphi_{mat}) &= 1 + \lfloor n/2 \rfloor \end{aligned}$$

where $(-1)!! = 1$ and $(2k-1)!! = \prod_{i=1}^{k}(2i-1)$ for $k \in \mathbb{N}$ is the *odd factorial*.

Theorem 3.2.10 (Carnap, 1950). *Let $S = (1, 1, \ldots, 1)$ (k unary predicates) and $\varphi \in L(S)$ be a closed formula. Then*

$$\lim_{n\to\infty} \frac{u(n, S, \varphi)}{u(n, S)} = \lim_{n\to\infty} \frac{|\{R \in \mathcal{P}([n])^k \mid R \models \varphi\}/\cong|}{|\mathcal{P}([n])^k/\cong|} = 0 \quad or \quad 1$$

(here $\mathcal{P}([n])^k = \mathcal{P}([n]) \times \cdots \times \mathcal{P}([n])$ with k components). In other words, for $n \to \infty$ the proportion of unlabeled S-structures on $[n]$ with property φ goes either to 0 or to 1.

Theorem 3.2.11 (Glebskij, Kogan, Liogon'kij and Talanov, 1969; Fagin, 1976). *Let $S \in \mathbb{N}^k$ be a type and $\varphi \in L(S)$ be a closed formula. Then*

$$\lim_{n\to\infty} \frac{l(n, S, \varphi)}{l(n, S)} = \lim_{n\to\infty} \frac{\left|\{R \in \prod_{i=1}^{k} \mathcal{P}([n]^{S_i}) \mid R \models \varphi\}\right|}{\left|\prod_{i=1}^{k} \mathcal{P}([n]^{S_i})\right|} = 0 \quad or \quad 1$$

—for $n \to \infty$ the proportion of S-structures on $[n]$ with property φ goes either to 0 or to 1.

In particular, there is no way to define by a first-order formula a property of S-structures holding asymptotically with probability, say, $\frac{1}{3}$.

Theorem 3.2.12 (version for graphs). *Let φ be a closed formula of the language $L(S)$ where $S = (2)$. Then (recall the closed formula $\varphi_{usg} \in L(S)$ defining undirected simple graphs)*

$$\lim_{n\to\infty} \frac{l(n, S, \varphi \wedge \varphi_{usg})}{l(n, S, \varphi_{usg})} = \lim_{n\to\infty} \frac{|\{G \subset \binom{[n]}{2} \mid G \models \varphi\}|}{2^{\binom{n}{2}}} = 0 \quad or \quad 1$$

—for $n \to \infty$ the proportion of graphs on $[n]$ with property φ goes either to 0 or to 1.

We think the reader understands our abused notation in the middle displayed fraction (we defined $\models$ only for relational structures).

Theorem 3.2.13 (Gödel, 1930). *Let $S \in \mathbb{N}^k$ be a type and let $\mathcal{T} \subset L(S)$ be a possibly infinite set of closed formulas. The theory $\mathcal{T}$ is consistent, free of contradictions, if and only if there exists an S-structure R on a possibly infinite set such that $R \models \mathcal{T}$, that is, $\mathcal{T}$ has a model.*

3.3 The Friedgut–Kalai theorem: thresholds are sharp

3.4 The Shelah–Spencer theorem: irrational exponents are not first order

3.5 The Blatter–Specker theorem: second-order binary structures are periodic

Chapter 4

Algebra of generating functions

4.1 The ring of formal power series

The reader is certainly familiar with complex polynomials, finite linear combinations $\mathbb{C}[[x]] = \{f = f(x) = \sum_{n=0}^{d} a_n x^n \mid d \in \mathbb{N}_0, a_n \in \mathbb{C}\}$. They form a commutative ring with 1. The *degree function* $\deg : \mathbb{C}[[x]] \to \mathbb{N}_0 \cup \{-\infty\}$, $\deg(f) = \max n$ with $a_n \neq 0$, $\deg(0) = -\infty$, has the properties that (i) $\deg(fg) = \deg f + \deg g$ and (ii) $\deg(f+g) \leq \max(\deg f, \deg g)$, with equality if $\deg f \neq \deg g$. Thus, by (i), $\mathbb{C}[x]$ has no zero divisors (but this is of course already subsumed in the proof of (i)).

By $\mathbb{C}[[x]]$ we denote the set of all *formal infinite* linear combinations with coefficients in $\mathbb{C}$ of the powers $1 = x^0, x = x^1, x^2, \ldots$ of the variable x,

$$\mathbb{C}[[x]] = \{f = f(x) = a_0 + a_1 x + a_2 x^2 + \cdots = \sum_{n=0}^{\infty} a_n x^n \mid a_n \in \mathbb{C}\}$$

—we call them *formal power series, fps.* Two usual binary operations of addition and multiplication on $\mathbb{C}[[x]]$ are

$$\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

and

$$\sum_{n=0}^{\infty} a_n x^n \cdot \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i}\right) x^n \ .$$

This product of formal power series is sometimes called the *Cauchy product*; it is named after the French mathematician Augustin-Louis Cauchy (1789–1857). The *Hadamard product*, named after his compatriot Jacques Hadamard (1865–1963), is defined by

$$\sum_{n=0}^{\infty} a_n x^n \odot \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} a_n b_n x^n \ .$$

Exercise 4.1.1 *Verify that* $\mathbb{C}[[x]] = (\mathbb{C}[[x]], +, \cdot)$ *is a commutative ring with* 1.

It is clear that $\mathbb{C}[x]$ is a subring of $\mathbb{C}[[x]]$.

For $f = \sum_{n\geq 0} a_n x^n \in \mathbb{C}[[x]]$ and $n \in \mathbb{N}_0$ we denote by $[x^n]f = a_n$ the coefficient of x^n, and for $f \neq 0$ we denote by $\operatorname{ord}(f)$ the smallest $n \in \mathbb{N}_0$ with $[x^n]f \neq 0$; we set $\operatorname{ord}(0) = +\infty$. It is easy to see that for every $f, g \in \mathbb{C}[[x]]$,

$$\operatorname{ord}(fg) = \operatorname{ord}(f) + \operatorname{ord}(g) \quad \text{and} \quad \operatorname{ord}(f+g) \geq \min(\operatorname{ord}(f), \operatorname{ord}(g)) \ ,$$

with equality if $\operatorname{ord}(f) \neq \operatorname{ord}(g)$. Thus, by the first equation, $\mathbb{C}[[x]]$ has no zero divisors.

Proposition 4.1.2 *$f \in \mathbb{C}[[x]]$ is a unit (i.e., $fg = 1$ for some $g \in \mathbb{C}[[x]]$) if and only if $\operatorname{ord}(f) = 0$.*

Proof. If $\operatorname{ord}(f) > 0$ then $\operatorname{ord}(fg) = \operatorname{ord}(f) + \operatorname{ord}(g) > 0$ for every fps g, so f is not a unit as $\operatorname{ord}(1) = 0$. If $\operatorname{ord}(f) = 0$ then $f(x) = a_0 + a_1 x + a_2 x^2 + \dots$ with $a_0 \neq 0$. We seek a $g(x) = b_0 + b_1 x + b_2 x^2 + \dots$ such that, for $n = 0, 1, 2, \dots$,

$$[x^n]fg = a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0 = \delta_{n,0} \ (= 1 \text{ if } n = 0 \text{ and } = 0 \text{ else}) \ .$$

This infinite system of equations in the unknowns $b_0, b_1, \dots$ has a unique solution: $b_0 = a_0^{-1}$, and if we know $b_0, b_1, \dots, b_{n-1}$ for $n > 0$ already then b_n is uniquely determined from the n-th equation by

$$b_n = -a_0^{-1}(a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0) \ .$$

□

Later we express the (multiplicative) inverse $f^{-1} = 1/f$ to f as a formal sum of a geometric series.

Example 4.1.3 Since $(1-x)(1+x+x^2+\dots) = 1$ and $(1-x-x^2)(1+x+2x^2+3x^3+5x^3+\dots) = (1-x-x^2)\sum_{n\geq 0} f_n x^n = 1$, where the *Fibonacci numbers* f_n follow the recurrence $f_0 = f_1 = 1$, $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$, we have inverses

$$\frac{1}{1-x} = \sum_{n\geq 0} x^n \quad \text{and} \quad \frac{1}{1-x-x^2} = \sum_{n\geq 0} f_n x^n \ .$$

Leonardo Pisano Bigollo (c. 1170–c. 1250), known as Fibonacci, was an Italian mathematician. □

We cannot divide by a fps f if (and only if) it has zero constant term: $a_0 = [x^0]f = 0$. After embedding the ring $\mathbb{C}[[x]]$ in the field $\mathbb{C}((x))$ of the *formal Laurent series (briefly fps)*, we can divide by every nonzero element;

$$\mathbb{C}((x)) = \{\textstyle\sum_{n=k}^{+\infty} a_n x^n \mid k \in \mathbb{Z}, a_n \in \mathbb{C}\} \ .$$

(Pierre A. Laurent (1813–1854) was a French mathematician.) Compared to $\mathbb{C}[[x]]$, in a fps f we allow terms with negative exponents but only finitely many.

The arithmetic operations $+$ and $\cdot$ and the order and coefficient functions $\mathrm{ord}(\cdot)$ and $[x^n]$ extend from $\mathbb{C}[[x]]$ to $\mathbb{C}((x))$ in the obvious way. It is easy to check that $(\mathbb{C}((x)), +, \cdot)$ is a commutative ring with 1. Moreover, every nonzero $f \in \mathbb{C}((x))$ has a unique expression as $f = x^k g$ where $k \in \mathbb{Z}$ and $\mathrm{ord}(g) = 0$ (so $g \in \mathbb{C}[[x]]$), thus f is a unit and has the inverse $f^{-1} = x^{-k} g^{-1}$. Therefore $(\mathbb{C}((x)), +, \cdot)$ is a field and we have the following.

Proposition 4.1.4 *$\mathbb{C}((x))$ is the field of fractions of $\mathbb{C}[[x]]$.*

Recall that each commutative ring R with no zero divisors has a unique *field of fractions*, a superfield $F \supset R$ such that each $a \in F$ is of the form $a = b/c$ with $b, c \in R$.

In Chapter 1 we saw that the GF $C(x) = \sum_{n \geq 1} c_n x^n$ of the Catalan numbers satisfies the quadratic equation $y^2 - y + x = 0$. Many GFs in enumerative combinatorics are algebraic, i.e., are solutions of polynomial equations. Thus solvability of such equations over $\mathbb{C}[[x]]$ or $\mathbb{C}((x))$ comes up naturally in enumeration, and also in algebraic geometry, algebra and number theory. The field $\mathbb{C}((x))$ is not algebraicly closed; the simplest polynomial equation unsolvable over it is $y^2 - x = 0$. Indeed, $\mathrm{ord}(f^2)$ is an even integer or $+\infty$ for every $f \in \mathbb{C}((x))$ but $\mathrm{ord}(x) = 1$. To be able to solve polynomial equations over $\mathbb{C}((x))$, we have to allow fractional exponents in the powers of x. We define

$$\begin{aligned} \mathbb{C}\langle x \rangle &= \{\textstyle\sum_{n=k}^{+\infty} a_n x^{n/m} \mid k \in \mathbb{Z}, m \in \mathbb{N}, a_n \in \mathbb{C}\} \\ &= \{f(x^{1/m}) \mid m \in \mathbb{N}, f(x) \in \mathbb{C}((x))\} \ . \end{aligned}$$

These formal infinite linear combinations are called *formal Puiseux series*; they are named after the French mathematician and astronomer Victor A. Puiseux (1820–1883) but already Isaac Newton (1642–1727) was using them. The operations $+$, $\cdot$ and functions $\mathrm{ord}(\cdot)$, $[x^k]$ extend to them naturally and $(\mathbb{C}\langle x \rangle, +, \cdot)$ is a field. Thus we have the extension of rings and fields

$$\mathbb{C}[x] \subset \mathbb{C}[[x]] \subset \mathbb{C}((x)) \subset \mathbb{C}\langle x \rangle \ .$$

The equation $y^2 - x = 0$ is now solvable in $\mathbb{C}\langle x \rangle$: $y = \pm x^{1/2}$. The proof of the following result is nontrivial and we refer for it to Fischer [35, Chapter 7].

Theorem 4.1.5 (Puiseux, 1850) *$\mathbb{C}\langle x \rangle$ is the algebraic closure of $\mathbb{C}((x))$.*

Recall that each field F has a unique *algebraic closure*, a superfield $\overline{F} \supset F$ such that (i) each element $a \in \overline{F}$ solves a polynomial equation with coefficients in F and (ii) each polynomial equation with coefficients in $\overline{F}$ has a solution in $\overline{F}$.

Exercise 4.1.6 *Prove that each equation $y^2 - f(x) = 0$, $f(x) \in \mathbb{C}\langle x \rangle$, has a solution in $\mathbb{C}\langle x \rangle$.*

4.2 Formal convergence in $\mathbb{C}[[x]]$ and $\mathbb{C}((x))$

We say that a sequence $(f_1, f_2, \dots) \subset \mathbb{C}[[x]]$ of fps *formally converges* to a fps $f \in \mathbb{C}[[x]]$ if

$$\lim_{n\to\infty} \operatorname{ord}(f - f_n) = +\infty \ .$$

In other words, for each $k \in \mathbb{N}_0$ the sequence of coefficients

$$([x^k]f_1, [x^k]f_2, \dots) \subset \mathbb{C}$$

is eventually constant and equal to the coefficient $[x^k]f$. We write

$$\lim f_n = \lim_{n\to\infty} f_n = f \ .$$

Hopefully no confusion arises by using the same notation for the ordinary limit of a sequence of real or complex numbers. The same definition of formal convergence works in $\mathbb{C}((x))$ and also all results on it we give below extend without problems from $\mathbb{C}[[x]]$ to $\mathbb{C}((x))$.

It is useful to express formal convergence in the analytic language of norms. We define $\|\cdot\| : \mathbb{C}((x)) \to [0, +\infty)$ by

$$\|f\| = 2^{-\operatorname{ord} f}$$

(with the convention that $2^{-\infty} = 0$).

Exercise 4.2.1 *Prove that for every $f, g \in \mathbb{C}((x))$, (i) $\|f\| \geq 0$, with equality exactly for $f = 0$, (ii) $\|fg\| = \|f\| \cdot \|g\|$ and (iii) $\|f + g\| \leq \max(\|f\|, \|g\|)$, with equality if $\|f\| \neq \|g\|$.*

Thus $(\mathbb{C}((x)), +, \cdot, \|\cdot\|)$ is a non-Archimedean normed field. ("non-Archimedean" refers to the strengthening of the triangle inequality $\|f + g\| \leq \|f\| + \|g\|$.) Clearly, $\lim f_n = f$ means that $\|f - f_n\| \to 0$ as $n \to \infty$, that is, $\lim f_n = f \iff \lim \|f_n - f\| = 0$.

Exercise 4.2.2 *Prove that a sequence $(f_n) \subset \mathbb{C}[[x]]$ formally converges (to some fps f) if and only if it is Cauchy: for every $\varepsilon > 0$ there is an $n_0 \in \mathbb{N}$ such that*

$$m, n > n_0 \Rightarrow \|f_m - f_n\| < \varepsilon \ .$$

Does it remain true for $f_m - f_n$ replaced by $f_{n+1} - f_n$?

Formal limits commute with arithmetic operations:

Exercise 4.2.3 *Verify that if $f = \lim f_n$ and $g = \lim g_n$ (where $f, g, f_n, g_n \in \mathbb{C}[[x]]$) then $f + g = \lim(f_n + g_n)$ and $fg = \lim(f_n g_n)$.*

Hence the same holds for inverses: if the fps f is a unit and $f = \lim f_n$ then for $n > n_0$ every f_n is a unit and $1/f = \lim(1/f_n)$. Using formal convergence we define infinite sums and infinite products of sequences of formal power series.

Definition 4.2.4 *Let $(f_n) \subset \mathbb{C}[[x]]$ be a sequence of formal power series. We define*

$$\sum_{n=1}^{\infty} f_n = \lim_{n\to\infty} (f_1 + f_2 + \cdots + f_n)$$

and similarly, assuming in adddition that $\operatorname{ord}(f_n) > 0$ *for each* n*,*

$$\prod_{n=1}^{\infty} (1 + f_n) = \lim_{n\to\infty} (1 + f_1)(1 + f_2) \dots (1 + f_n) \ ,$$

if the formal limits exist. If they exist we say that the infinite series, respective product, formally converges.

The notational ambiguity $f = \sum_{n=0}^{\infty} a_n x^n$ causes no problem because the series $\sum_{n=0}^{\infty} f_n$ where $f_n = a_n x^n$ formally converges to f. In plane terms, formal convergence of $f = \sum_{n=1}^{\infty} f_n$ and $f = \prod_{n=1}^{\infty}(1+f_n)$ means that for each $k \in \mathbb{N}_0$ the coefficient $[x^k]f$ is given by an expression in the coefficients in the f_ns containing only *finitely many* terms, so the fps f_n with $n > n_0 = n_0(k)$ are irrelevant.

Exercise 4.2.5 *Explain why instead of considering $\prod_{n=1}^{\infty} f_n$ for general fps f_n we may restrict to the case when each f_n has constant term* 1*.*

Proposition 4.2.6 *Let $(f_n) \subset \mathbb{C}[[x]]$. Then*

$$\textstyle\sum_{n=1}^{\infty} f_n \text{ formally converges} \iff \lim f_n = 0 \ ,$$

which is the same as $\operatorname{ord}(f_n) \to +\infty$ *or as* $\|f_n\| \to 0$*. The same holds for $\prod_{n=1}^{\infty}(1 + f_n)$ (each f_n has zero constant term).*

Proof. Let $k \in \mathbb{N}_0$ and $\operatorname{ord}(f_n) \to +\infty$. Then $[x^l]f_n = 0$ whenever $n > n_0$ and $l \le k$, hence the sequence $([x^k]\sum_{i=1}^{n} f_i = \sum_{i=1}^{n}[x^k]f_i)_{n\ge1}$ is constant for $n > n_0$ and partial sums formally converge. The same holds for partial products because for $n > n_0$ we have $[x^k]\prod_{i=1}^{n}(1+f_i) = [x^k]\prod_{i=1}^{n-1}(1+f_i)+[x^k]f_n\prod_{i=1}^{n-1}(1+f_i) = [x^k]\prod_{i=1}^{n-1}(1 + f_i)$.

If $\operatorname{ord}(f_n) \not\to +\infty$ then there is a $k \in \mathbb{N}_0$ such that $\operatorname{ord}(f_n) = k$ for infinitely many n. Thus $[x^k]\sum_{i=1}^{n-1} f_i \ne [x^k]\sum_{i=1}^{n} f_i$ and $[x^k]\prod_{i=1}^{n-1}(1 + f_i) \ne [x^k]\prod_{i=1}^{n}(1 + f_i)$ for (the same) infinitely many n, and the sequences of partial sums and partial products do not formally converge. □

It is well known that $\sum_{n=1}^{\infty}(-1)^{n+1}/n = \log 2$ but that this sum can be changed arbitrarily by permuting the sequence of summands $(1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \dots)$; the series is not absolutely convergent. This does not happen for formal convergence which corresponds to absolute numeric convergence.

Proposition 4.2.7 *Let $f_n \in \mathbb{C}[[x]]$ for $n \in \mathbb{N}$ and $\pi : \mathbb{N} \to \mathbb{N}$ be a permutation of $\mathbb{N}$. Then*

$$\sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} f_{\pi(n)}$$

whenever one of the series formally converges. The same holds for $\prod_{n=1}^{\infty}(1+f_n)$ (each f_n has zero constant term).

Proof. We prove it for products; the argument for sums is similar and easier. Suppose that $k \in \mathbb{N}$ and $\prod_{n=1}^{\infty}(1+f_n)$ formally converges. Thus, by the previous proposition, $\operatorname{ord}(f_n) > k$ if $n > n_0$. We take n_1 large enough so that $\{\pi(1), \pi(2), \ldots, \pi(n_1)\} \supset \{1, 2, \ldots, n_0\}$ (clearly $n_1 \geq n_0$). For $n > n_1$ we then have

$$\operatorname{ord}\big(\textstyle\prod_{i=1}^{n}(1+f_i) - \prod_{i=1}^{n}(1+f_{\pi(i)})\big) = \operatorname{ord}\big(\textstyle\prod_{i=1}^{n_0}(1+f_i)\cdot f\big) > k$$

because the fps f is a sum of nonconstant monomials in f_is with $i > n_0$. Thus $\prod_{n=1}^{\infty}(1+f_{\pi(n)})$ formally converges to the same limit as $\prod_{n=1}^{\infty}(1+f_n)$. □

Exercise 4.2.8 *We have shown in effect that infinite sums and products of fps satisfy the commutative law. Show that they are also associative and distributive.*

We present several examples of infinite series and products of fps.

Example 4.2.9 If the fps f has zero constant term then

$$\frac{1}{1-f} = \sum_{n=0}^{\infty} f^n \,.$$

The series formally converges because $\operatorname{ord}(f^n) \geq n$, and the equality follows by taking formal limit of the identity $(1-f)(1+f+f^2+\cdots+f^{n-1}) = 1-f^n$. In particular, if $f = a_0 + a_1x + a_2x^2 + \ldots$ with $a_0 \neq 0$ then

$$\frac{1}{f} = \frac{1}{a_0}\sum_{n=0}^{\infty}\left(-a_0^{-1}a_1x - a_0^{-1}a_2x^2 - \ldots\right)^n \,.$$

This is the promised expression of a multiplicative inverse as a sum of formal geometric series. We can expres the same inverse element also as an infinite product: if the fps f has zero constant term then

$$\frac{1}{1-f} = \prod_{n=0}^{\infty}\left(1+f^{2^n}\right) \,.$$

The product formally converges because $\operatorname{ord}(f^{2^n}) \geq 2^n$, and the equality follows from the previous identity as $(1+f)(1+f^2)(1+f^4)\ldots(1+f^{2^n}) = \sum_{k=0}^{2^n} f^k + g$ with $\operatorname{ord}(g) > 2^n$ (each $k \in \mathbb{N}_0$ has exactly one expression as a sum of distinct

powers of 2). Alternatively, since $(1-f)(1+f) = 1-f^2$, $(1-f^2)(1+f^2) = 1-f^4$, $(1-f^4)(1+f^4) = 1-f^8$ and so on, multiplying the infinite product, which we denote P, by $1-f$ we get the identity $(1-f)P = (1-f^{2^k})\prod_{n\geq k}(1+f^{2^n})$. For $k \to \infty$ the right side goes formally to 1, thus $(1-f)P = 1$. □

A interesting fact is that the relation $g = 1/(1-f)$ forces regular growth of the coefficients in the fps g.

Proposition 4.2.10 *Let $a_1, a_2, \ldots$ be nonnegative real numbers and the coefficients b_n be defined by*

$$1 + \sum_{n=1}^{\infty} b_n x^n = \frac{1}{1 - \sum_{n=1}^{\infty} a_n x^n} \, .$$

Then the limit $\lim_{n\to\infty} b_n^{1/n}$ exists (it may be $+\infty$).

Proof. A *composition* of $n \in \mathbb{N}$ is any tuple $c = (m_1, \ldots, m_k) \in \mathbb{N}^k$ with $m_1 + \cdots + m_k = n$; let $C(n)$ be all compositions of n. We weight c by the product $w(c) = a_{m_1} \ldots a_{m_k}$. Expanding the right side of the definition of b_n into a geometric series we see that

$$b_n = \sum_{c \in C(n)} w(c) \, .$$

If $n = n_1 + n_2$ and c_i is a composition of n_i then $c = (c_1, c_2)$ (abusing notation slightly) gives a composition of n. This correspondence is in fact injective and $w(c) = w(c_1)w(c_2)$. Hence

$$b_n = \sum_{c \in C(n)} w(c) \geq \sum_{c_1 \in C(n_1)} w(c_1) \sum_{c_2 \in C(n_2)} w(c_2) = b_{n_1} b_{n_2} \, .$$

So $b_n \geq b_{n_1} b_{n_2}$ whenever $n = n_1 + n_2$—the sequence (b_n) is *supermultiplicative*. By Fekete's lemma below, such sequences always posses the stated limit. □

Exercise 4.2.11 *Prove the next lemma and deduce from it the above result for supermultiplicative sequences.*

Lemma 4.2.12 (Fekete, 1927) *If $a_1, a_2, \ldots$ are nonnegative real numbers such that $a_{n_1+n_2} \geq a_{n_1} + a_{n_2}$ whenever $n = n_1 + n_2$ (one says that the sequence (a_n) is* superadditive*) then the limit $\lim_{n\to\infty} a_n/n$ exists and may be $+\infty$. The same holds, with finite limits only, for subadditive sequences.*

The result is due to the Hungarian mathematician Michael Fekete (1886–1957).

Exercise 4.2.13 *Prove bijectively and by GF that the number of compositions of n is*

$$|C(n)| = 2^{n-1} \, .$$

Example 4.2.14 By Example 4.2.9,

$$\prod_{n=0}^{\infty}\left(1+x^{2^n}\right)=\sum_{n=0}^{\infty}x^n\ .$$

This is one of the very many identities equaling an infinite product with an infinite series. Possibly the most famous such identity (but we will discuss a rival shortly) is *Euler's partition identity*

$$\prod_{n=1}^{\infty}\frac{1}{1-x^n}=\sum_{n=0}^{\infty}p(n)x^n\ .$$

Here $p(0)=1$ and for $n\in\mathbb{N}$ the *partition function* $p(n)\in\mathbb{N}$ counts so called *(integer) partitions of* n, the expressions $n=a_1+a_2+\cdots+a_k$ where $a_i\in\mathbb{N}$ and $a_1\geq a_2\geq\cdots\geq a_k$. So $p(1)=1, p(2)=2, p(3)=3, p(4)=5, p(5)=7$ and so on. The identity is due to the Swiss mathematician Leonhard Euler (1707–1783).

From the miriads of partition $\prod=\sum$ identities we quote just one more, *Euler's pentagonal identity*

$$\prod_{n=1}^{\infty}(1-x^n)=\sum_{n=-\infty}^{\infty}(-1)^n x^{n(3n+1)/2}=1-x-x^2+x^5+x^7-x^{12}-x^{15}+\dots$$

We give a proof based on manipulations with fps in the proposition below. The numbers $1,2,5,7,12,15,\dots$ are so called *pentagonal numbers*; they can be obtained as numbers of dots in certain pentagonal arrays, in the same way as $1,3,6,10,\dots$ are triangular and $1,4,9,16,\dots$ square numbers. □

Exercise 4.2.15 *Prove Euler's partition identity.*

Exercise 4.2.16 *Deduce from Euler's pentagonal identity the recurrence for partition numbers*

$$p(n)=p(n-1)+p(n-2)-p(n-5)-p(n-7)+p(n-12)+p(n-15)-\dots$$

where $p(m)=0$ *if* $m<0$.

Exercise 4.2.17 *Restate Euler's pentagonal identity in terms of partitions.*

Proposition 4.2.18 *We prove Euler's pentagonal identity*

$$\prod_{n=1}^{\infty}(1-x^n)=1+\sum_{n=1}^{\infty}(-1)^n(x^{n(3n+1)/2}+x^{n(3n-1)/2})\ .$$

Proof.

□

Example 4.2.19 We give two more examples of $\prod = \sum$ type identities, both in fact outside the scope of formal convergence in $\mathbb{C}[[x]]$. One of the greatest discoveries of Euler says that

$$\prod_p \left(1 - \frac{1}{p^s}\right)^{-1} = \sum_{n=0}^{\infty} \frac{1}{n^s} \, .$$

Here $p = 2, 3, 5, 7, 11, 13, \ldots$ runs through the prime numbers and s is a *formal* variable. This famous identity holds on the formal level because multiplying out the factors $(1 - 1/p^s)^{-1} = 1 + 1/p^2 + 1/p^{2s} + \ldots$, the result is well defined and gives $\sum_{n\geq 1} a_n/n^s$ with a_n counting the prime factorizations of n (expressions of n as a product of powers of distinct primes), and always $a_n = 1$ due to uniqueness of prime factorization. The identity is based on formal convergence but in a ring different from $\mathbb{C}[[x]]$, the ring of formal Dirichlet series which we will not introduce here. (Peter L. Dirichlet (1805–1859) was a German mathematician, also a brother in law of the composer and musician Felix Mendelssohn Bartholdy (1809–1847).)

Another Euler's discovery is that

$$\prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n (\pi x)^{2n}}{(2n+1)!} \, . \tag{4.1}$$

Here $\pi = 3.14159\ldots$ is the well known constant, defined for example by the relation $\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$ (also due to Euler). The identity relates elements of $\mathbb{C}[[x]]$ (the right side with the fps $1 - x^2/n^2$, $n = 1, 2, \ldots$) but it cannot use formal convergence because the infinite product does not formally converge. However, we can interpret the left side *semiformally*: after expanding it the coefficients of $x^2, x^4, x^6, \ldots$ are given by expressions with infinitely many terms—so formal convegence is out of question—but these expressions are absolutely convergent (numeric) series, which can be summed regardless of order (and conveniently manipulated in other ways). For example, the coefficients of x^2 and x^4 come out as the sums of the absolutely convergent infinite series

$$s_2 = -\sum_{n\geq 1} \frac{1}{n^2} \quad \text{and} \quad s_4 = \sum_{m>n\geq 1} \frac{1}{m^2 n^2} \, ,$$

respectively. The identity thus claims, for example, that $(6s_2)^2 = 120 s_4$. We will not develop semiformal convegence here. But we have to say that the more standard interpretation of Euler's identity (4.1) is as an equality between functions, the righ side being of course $\sin(\pi x)/\pi x$. □

Exercise 4.2.20 *If we change the $-$ signs on the left side of (4.1) to $+$, after multiplying out formally we see that on the right side it results in changing exactly the signs of the terms with odd n. Thus we get another identity (actually equivalent with the original one)*

$$\prod_{n=1}^{\infty} \left(1 + \frac{x^2}{n^2}\right) = \sum_{n=0}^{\infty} \frac{(\pi x)^{2n}}{(2n+1)!} \, .$$

Is this a valid argument? Expansion of what function do we have on the right side?

We put formal convergence in $\mathbb{C}((x))$ in a wider context. In Section 4.1 we mentioned two ways of completing a domain by adding solutions to hitherto unsolvable equations, the field of fractions and the algebraic closure. Now we encounter another construction of this sort; we show that the normed field $\mathbb{C}((x))$ is the metrical completion of the normed field of rational functions $\mathbb{C}(x)$. This is analogous to obtaining the field of real numbers $\mathbb{R}$ by completing the field of fractions $\mathbb{Q}$. In fact, more apt analogy is with another completion of $\mathbb{Q}$, the field of p-adic numbers $\mathbb{Q}_p$.

The field of rational functions $\mathbb{C}(x)$ is the field of fractions of the ring of complex polynomials $\mathbb{C}[x]$. Thus

$$\mathbb{C}(x) = \{p/q \mid p, q \in \mathbb{C}[x], q \neq 0\} ,$$

factored by the equivalence relation

$$p_1/q_1 \sim p_2/q_2 \iff p_1q_2 - p_2q_1 = 0 .$$

There is a natural injective embedding of $\mathbb{C}(x)$ in $\mathbb{C}((x))$, and we regard $\mathbb{C}(x)$ as a subfield of $\mathbb{C}((x))$. Due to the division algorithm, $\mathbb{C}[x]$ is a UFD (unique factorization domain). The irreducible elements are exactly the linear polynomials $x-\alpha$, $\alpha \in \mathbb{C}$, and their associates, and the units in $\mathbb{C}[x]$ are exactly the nonzero constant polynomials. So every nonzero $r \in \mathbb{C}(x)$ has the unique expression

$$r = r(x) = \gamma \prod_{\alpha \in \mathbb{C}} (x-\alpha)^{k_\alpha},\ \gamma \in \mathbb{C}^*,\ k_\alpha \in \mathbb{Z} ,$$

with only finitely many exponents $k_\alpha \neq 0$. We define the *order of* $r(x)$ *at* α by

$$\mathrm{ord}_\alpha(r) = k_\alpha$$

and set $\mathrm{ord}_\alpha(0) = +\infty$. It is useful to define also the order at infinity:

$$\mathrm{ord}_\infty(r(x)) := \mathrm{ord}_0(r(1/x)) .$$

The order function on $\mathbb{C}(x)$ has the same two basic properties as the order function on $\mathbb{C}((x))$:

$$\mathrm{ord}_\alpha(rs) = \mathrm{ord}_\alpha(r) + \mathrm{ord}_\alpha(s) \ \text{ and } \ \mathrm{ord}_\alpha(r+s) \geq \min(\mathrm{ord}_\alpha(r), \mathrm{ord}_\alpha(s)) ,$$

with equality if $\mathrm{ord}_\alpha(r) \neq \mathrm{ord}_\alpha(s)$. In fact, for every $r \in \mathbb{C}(x)$ we have

$$\mathrm{ord}_0(r) = \mathrm{ord}(r)$$

where the ord on the right side is the order function on $\mathbb{C}((x))$.

Exercise 4.2.21 *Show that for every nonzero* $r \in \mathbb{C}(x)$,

$$\sum_{\alpha \in \mathbb{C} \cup \{\infty\}} \mathrm{ord}_\alpha(r) = 0 .$$

Exercise 4.2.22 *Prove Lüroth's theorem, or look it up in the literature: if $\mathbb{C} \subset K \subset \mathbb{C}(x)$ is an intermediate field then either $K = \mathbb{C}$ or K is isomorphic to the field $\mathbb{C}(x)$. The theorem is named after the German mathematician Jacob Lüroth (1844–1910). For that matter, how do we know that the fields $\mathbb{C}$ and $\mathbb{C}(x)$ are not isomorphic?*

As with ord, we turn the order at α function into the norm $\|\cdot\|_\alpha : \mathbb{C}(x) \to [0, +\infty)$ by

$$\|r\|_\alpha := 2^{-\mathrm{ord}_\alpha(r)}$$

(again $\|0\|_\alpha = 0$). We omit the proof of the next proposition.

Proposition 4.2.23 *For every $\alpha \in \mathbb{C} \cup \{\infty\}$, $\mathbb{C}(x) = (\mathbb{C}(x), +, \cdot, \|\cdot\|_\alpha)$ is a non-Archimedean normed field: for every $r, s \in \mathbb{C}(x)$,*

1. *$\|r\|_\alpha \geq 0$ and equality holds if only if $r = 0$,*
2. *(multiplicativity) $\|rs\|_\alpha = \|r\|_\alpha \cdot \|s\|_\alpha$ and*
3. *(strong triangle inequality) $\|r + s\|_\alpha \leq \max(\|r\|_\alpha, \|s\|_\alpha)$, with equality if $\|r\|_\alpha \neq \|s\|_\alpha$.*

We all know that the field of fractions $\mathbb{Q}$ is not complete to the usual norm $|p/q|$: some Cauchy sequences in it do not have limits. An example is the sequence $(a_n) \subset \mathbb{Q}$ where $a_1 = 1$, $a_2 = 2$ and the next a_n are constructed by repeatedly halving the initial interval $[1, 2]$ so that always 2 lies between the squares of the endpoints. Then (a_n) is Cauchy and its limit $a = \lim a_n$ should satisfy $a^2 = 2$ but there is no such element in $\mathbb{Q}$.

We show that, similarly, $\mathbb{C}(x)$ is not complete with respect to the norm $\|\cdot\|_\alpha$. For simplicity we restrict to the case $\alpha = 0$ and use notation $\|\cdot\|_0 = \|\cdot\|$. As we know,

$$\|r\| = 2^{-\mathrm{ord}(r)},\ r \in \mathbb{C}(x) \subset \mathbb{C}((x))\ .$$

Proposition 4.2.24 *Suppose that $0 \leq n_1 < n_2 < \dots$ is an infinite sequence of integers such that the length of the gaps $n_{i+1} - n_i$ is unbounded. Then*

$$(x^{n_1},\ x^{n_1} + x^{n_2},\ x^{n_1} + x^{n_2} + x^{n_3}, \dots)$$

is a Cauchy sequence of polynomials that has no limit in $\mathbb{C}(x)$ (with respect to the norm $\|\cdot\|$).

Proof. Let $p_i = x^{n_1} + x^{n_2} + \cdots + x^{n_i}$. The sequence $(p_1, p_2, \dots)$ is Cauchy because $\|p_i - p_j\| < 2^{-n_i}$ if $j > i$. Note that if $p, q \in \mathbb{C}[x]$ and $\deg(p) < \mathrm{ord}(q)$ then $\mathrm{ord}(q - p) \leq \mathrm{ord}(q)$. Now let $p/q \in \mathbb{C}(x)$ be given. By the assumption, there is an i such that $n_i > \deg(p)$ and $n_{i+1} - n_i > \deg(q)$. Then for $j > i$ we have

$$\begin{aligned}
\left\| \frac{p}{q} - p_j \right\| &= \left\| \frac{p - p_i q - (p_j - p_i)q}{q} \right\| = \|1/q\| \cdot \|(p_j - p_i)q - (p - p_i q)\| \\
&= 2^{\mathrm{ord}(q)} 2^{-\mathrm{ord}((p_j - p_i)q - (p - p_i q))} \geq 2^{\mathrm{ord}(q)} 2^{-(n_{i+1} + \mathrm{ord}(q))} \\
&= 2^{-n_{i+1}}
\end{aligned}$$

because $\operatorname{ord}((p_j - p_i)q) = n_{i+1} + \operatorname{ord}(q) \geq n_{i+1} > n_i + \deg(q) = \deg(p - p_i q)$. Thus $\|p/q - p_j\| \geq 2^{-n_{i+1}}$ for every j with $j > i$ and the sequence (p_i) does not converge to p/q. □

The combinatorial idea behind the previous example is that no sequence $(a_n) \subset \mathbb{C}$ with infinitely many nonzero terms and arbitrarily long intervals of zeros can satisfy for $n > n_0$ a linear recurrence with constant coefficients; hence the GF $\sum_{n\geq 0} a_n x^n$ is not rational. This is exactly the case with $\sum_{i\geq 0} x^{n_i} = \lim p_i(x)$— the limit is in $\mathbb{C}((x))$ but not in $\mathbb{C}(x)$.

Exercise 4.2.25 *What about the Cauchy sequences (in $(\mathbb{C}(x), \|\cdot\|)$)*

$$(\textstyle\sum_{i=1}^{n} x^i)_{n\geq 1},\ (\sum_{i=1}^{n} c_i x^i)_{n\geq 1} \quad \textit{and} \quad (\sum_{i=1}^{n} f_i x^i)_{n\geq 1}$$

where (c_i) are the Catalan and (f_i) the Fibonacci numbers. Do they have limits in $\mathbb{C}(x)$? If yes, what are they?

Each metric space (M, d) has a unique *metrical completion*, a metric superspace $N \supset M$ such that N is complete (every Cauch sequence in it converges) and M is dense in N (every point in N is a limit of points in M). If M is a field and the metric comes from a norm then N is a normed field as well. This is the situation for $\mathbb{C}((x)) \supset \mathbb{C}(x)$ with respect to the norm $\|\cdot\| = \|\cdot\|_0$. We 'proved' $\mathbb{C}((x))$ to be complete in Exercise 4.2.2 and it is clear that already the rational functions $p(x)/x^k$, $p \in \mathbb{C}[x]$ and $k \in \mathbb{N}_0$, are dense in $\mathbb{C}((x))$. Thus we have the following.

Proposition 4.2.26 *The normed field of formal Laurent series $\mathbb{C}((x))$ is the metrical completion of the field of rational functions $\mathbb{C}(x)$ with respect to the norm $\|\cdot\| = \|\cdot\|_0$.*

At the close of the section we show how formal Laurent series produce another, important, complete normed field.

Proposition 4.2.27 *For any prime number p, the normed field*

$$\mathbb{Z}((x))/(x - p)$$

is the metrical completion of the field of fractions $\mathbb{Q}$ with respect to the p-adic norm $\|\cdot\|_p$.

4.3 Differentiation, composition, exp and log

For $f = f(x) = \sum_{n\geq k} a_n x^n \in \mathbb{C}((x))$, the *(formal) derivative of f* is the fls

$$f'(x) = \tfrac{df}{dx} = \textstyle\sum_{n\geq k} n a_n x^{n-1}\ .$$

Formal differentiation has the familiar properties, which we leave for the interested reader to derive.

Exercise 4.3.1 *Let $f, g \in \mathbb{C}((x))$ and $\alpha, \beta \in \mathbb{C}$. Show that*

$$(\alpha f+\beta g)' = \alpha f'+\beta g',\ (fg)' = f'g+fg' \ \text{ and } \ (f/g)' = (f'g-fg')/g^2\ (g \neq 0)\ .$$

Hence $(f^k)' = kf^{k-1}f'$ for every $f \in \mathbb{C}((x))$ and $k \in \mathbb{Z}$, and $f' = 0$ iff f is constant, i.e., $[x^k]f = 0$ if $k \neq 0$. Differentiation commutes with formal limits.

Exercise 4.3.2 *Let $f_n, f \in \mathbb{C}((x))$ and $f = \lim f_n$. Show that*

$$f' = \lim f_n'\ .$$

Thus we may always differentiate infinite series of fls term by term: $(\sum_{n\geq 1} f_n)' = \sum_{n\geq 1} f_n'$, if one of the series formally converges. The same holds for infinite products of fps: in the convergent case, $(\prod_{n=1}^{\infty}(1+f_n))' = \lim((1+f_1)(1+f_2)\dots(1+f_n))'$ $(\mathrm{ord}(f_n) > 0)$.

Iterating the above *Leibniz rule* (the second formula in Exercise 4.3.1) we get for every nonzero $f_1, f_2, \dots, f_k \in \mathbb{C}((x))$ the *logarithmic derivative identity*

$$(f_1f_2\dots f_k)' = \textstyle\sum_{n=1}^{k} f_1\dots f_{n-1}f_n'f_{n+1}\dots f_k = \sum_{n=1}^{k} f_n'/f_n \cdot f_1f_2\dots f_k\ ,$$

or

$$\frac{(f_1f_2\dots f_k)'}{f_1f_2\dots f_k} = \sum_{n=1}^{k} \frac{f_n'}{f_n}\ .$$

(Gottfried W. von Leibniz (1646–1716) was a German mathematician and philosopher.) It has the following infinite version.

Proposition 4.3.3 *Let $(f_n) \subset \mathbb{C}[[x]]$ be a sequence of fps such that each f_n has constant term 1 and $\|f_n - 1\| \to 0$ as $n \to \infty$. Then the next infinite product and series formally converge and*

$$\Big(\prod_{n=1}^{\infty} f_n\Big)' = \sum_{n=1}^{\infty} \frac{f_n'}{f_n} \cdot \prod_{n=1}^{\infty} f_n\ .$$

Proof. Take formal limit for $k \to \infty$ of the finite form of the identity. □

The identity takes its name from the alternative derivation by differentiating the logarithm of the product. We introduce formal logarithm later.

We present two applications of the logarithmic derivative identity, first for the finite and then for the infinite version. We introduce for any nonzero polynomial $f \in \mathbb{C}[x]$ the function

$$\mathrm{rad}(f) = |\{\alpha \in \mathbb{C} \mid f(\alpha) = 0\}|\ ,$$

which counts distinct roots of f (without multiplicity). It should be clear that $\mathrm{rad} f \leq \deg f$ and $\mathrm{rad}(gf^n) = \mathrm{rad}(gf)$ for every $n \in \mathbb{N}$ and $g, f \in \mathbb{C}[x]$

Theorem 4.3.4 (the Stothers–Mason theorem) *If the polynomials $a, b, c \in \mathbb{C}[x]$ are pairwise coprime, not all constant and satisfy relation $a + b = c$, then*

$$\max(\deg a, \deg b, \deg c) \leq \operatorname{rad}(abc) - 1 \ .$$

Proof. Dividing $a+b=c$ by c ($c \neq 0$ by coprimality), setting $f = a/c$, $g = b/c$ and differentiating, we get the equations

$$f + g = 1 \ \text{ and } \ f' + g' = f \cdot \frac{f'}{f} + g \cdot \frac{g'}{g} = 0 \ ,$$

which gives

$$-\frac{f'/f}{g'/g} = \frac{b}{a} \ .$$

We factorize a, b, c:

$$f = \frac{a}{c} = \frac{\alpha \prod (x-\alpha_i)^{m_i}}{\gamma \prod (x-\gamma_i)^{o_i}} \ \text{ and } \ g = \frac{b}{c} = \frac{\beta \prod (x-\beta_i)^{n_i}}{\gamma \prod (x-\gamma_i)^{o_i}} \ ,$$

where $\alpha, \beta, \gamma \in \mathbb{C}$ are nonzero, the α_i are distinct roots of a with multiplicities $m_i \in \mathbb{N}$, and similarly for the β_i and γ_i. Expressing f'/f and g'/g by the logarithmic derivative identity mentioned above, we get

$$\frac{b}{a} = -\frac{\sum m_i/(x-\alpha_i) - \sum o_i/(x-\gamma_i)}{\sum n_i/(x-\beta_i) - \sum o_i/(x-\gamma_i)} \ .$$

If we multiply the denominator and the numerator on the right side by

$$N = \prod (x-\alpha_i) \cdot \prod (x-\beta_i) \cdot \prod (x-\gamma_i) \ ,$$

we get

$$\frac{b}{a} = -\frac{N\left(\sum m_i/(x-\alpha_i) - \sum o_i/(x-\gamma_i)\right)}{N\left(\sum n_i/(x-\beta_i) - \sum o_i/(x-\gamma_i)\right)} = \frac{Q}{P} \ ,$$

where the polynomials $P, Q \in \mathbb{C}[x]$ have degrees at most $\deg N - 1 = \operatorname{rad}(abc) - 1$. Since a, b are coprime, $\deg a \leq \deg P \leq \operatorname{rad}(abc) - 1$ and $\deg b \leq \deg Q \leq \operatorname{rad}(abc) - 1$. From $a + b = c$ we deduce that $\deg c \leq \max(\deg a, \deg b) \leq \operatorname{rad}(abc) - 1$ too. □

The theorem is due, independently, to Stothers [77] and Mason [55]. Our proof is taken from Lang [49, p. 194] who writes that it is due to Mason. The S.–M. theorem has many applications, of which we mention here just the proof of the FLT (*Fermat's last theorem*, Pierre de Fermat (1601 or 1607–1665) was a French lawyer and an amateur mathematician) for polynomials.

Corollary 4.3.5 *If the polynomials $a, b, c \in \mathbb{C}[x]$, not all constant, satisfy relation $a^n + b^n = c^n$, $n \in \mathbb{N}$, then $n \leq 2$.*

Proof. We may assume that the polynomials a, b, c are coprime. Let $d \geq 1$ be their maximum degree. The S.–M. theorem: $nd = \max(\deg a^n, \deg b^n, \deg b^n) \leq \mathrm{rad}(a^n b^n c^n) - 1 = \mathrm{rad}(abc) - 1 \leq \deg(abc) - 1 \leq 3d - 1$. So $n \leq 2$. □

For the exponent $n = 2$ we have infinitely many solutions, coming from the bivariate identity $(x^2 - y^2)^2 + (2xy)^2 = (x^2 + y^2)^2$.

For the second application of the logarithmic derivative identity we define for $n \in \mathbb{N}$ the *sum-of-divisors function*

$$\sigma(n) = \sum_{d \,|\, n} d \,.$$

For example, $\sigma(7) = 1 + 7 = 8$ and $\sigma(20) = 1 + 2 + 4 + 5 + 10 + 20 = 42$.

Exercise 4.3.6 *Prove the identity*

$$\sum_{n=1}^{\infty} \frac{x^n}{1 - x^n} = \sum_{n=1}^{\infty} \sigma(n) x^n \,.$$

The series on the left side (and similar expressions) is called *Lambert series*, after Johann Ch. Lambert (1728–1777), a Swiss mathematician. You can guess who discovered the next beautiful identity.

Proposition 4.3.7 *For every $n \in \mathbb{N}$ we have the recurrence*

$$\sigma(n) = \sigma(n-1) + \sigma(n-2) - \sigma(n-5) - \sigma(n-7) + \sigma(n-12) + \sigma(n-15) - \dots \,.$$

Here we subtract from n the pentagonal numbers (defined in Example 2 in the previous section), $\sigma(m) = 0$ if $m < 0$ and if $\sigma(0)$ appears on the right side (i.e., n is pentagonal) then we set $\sigma(0) = \sigma(n - n) = n$.

Proof. We denote by $P \subset \mathbb{N}$ the pentagonal numbers and apply the logarithmic derivative identity on Euler's pentagonal identity $\prod_{n \geq 1}(1 - x^n) = 1 + \sum_{n \in P} s(n) x^n$, where $s(n) \in \{-1, 1\}$ is the pentagonal sign of n. We get

$$\begin{aligned} x\Big(1 + \sum_{n \in P} s(n) x^n\Big)' &= x \sum_{n \geq 1} \frac{(1 - x^n)'}{1 - x^n} \Big(1 + \sum_{n \in P} s(n) x^n\Big) \,, \\ -\sum_{n \in P} s(n) n x^n &= \sum_{n \geq 1} \frac{n x^n}{1 - x^n} \Big(1 + \sum_{n \in P} s(n) x^n\Big) \\ &= \sum_{n \geq 1} \sigma(n) x^n \Big(1 + \sum_{n \in P} s(n) x^n\Big) \,. \end{aligned}$$

Applying $[x^n]$ to both sides, we get the recurrence in the equivalent form $-s(n)\sigma(n - n) = \sigma(n) - \sigma(n-1) - \sigma(n-2) + \sigma(n-5) + \sigma(n-7) - \dots$ (in this equality, on the right side $\sigma(m) = 0$ if $m \leq 0$). □

After differentiation we introduce another important operation on $\mathbb{C}((x))$, the *(formal) composition of fls*. It should be the formal counterpart of composition of functions. It is easy to see that for $f, g \in \mathbb{C}((x))$ the infinite series

$$\sum_{n\in\mathbb{Z}}([x^n]f)g^n$$

formally converges if and only if

(i) $\operatorname{ord}(f) \geq 0$ and $g = 0$ or

(ii) f is arbitrary and $0 < \|g\| < 1$ (i.e., $0 \neq g \in x\mathbb{C}[[x]]$) or

(iii) $[x^n]f \neq 0$ for only finitely many n (i.e., f is a Laurent polynomial) and $g \neq 0$ is arbitrary.

Definition 4.3.8 *If $f, g \in \mathbb{C}((x))$ and (i) or (ii) holds then we say that g* is substituable in f *and define the* composition of f and g *as the formal sum*

$$f(g(x)) = f \circ g = \sum_{n\in\mathbb{Z}}([x^n]f)g(x)^n \ .$$

Example 4.3.9 As we saw earlier, if $\operatorname{ord}(f) > 0$ then

$$(1 + x + x^2 + \dots) \circ f = 1/(1 - f) \ .$$

□

It is clear that $x \circ f = f \circ x = f$ for every $f \in \mathbb{C}((x))$, thus x is the neutral element for $\circ$. The composition case (i) may seem trivial but the next exercise shows its importance. The case (iii), which we omitted from the definition of $f \circ g$, is in fact used often in applications of field operation to fls, for example, g^2 for $g \in \mathbb{C}((x))$ may be interpreted as $x^2 \circ g$. The problem with including case (iii) in the definition of $f \circ g$ is that such extended composition ceases to be associative, see Example 4.3.17 below.

Exercise 4.3.10 *Let $f \in \mathbb{C}[[x]]$ and $n \in \mathbb{N}_0$. Show that*

$$n![x^n]f = ((d/dx)^n f)(0)$$

where $(d/dx)^n f$ is the n-th formal derivative of f.

Exercise 4.3.11 *Show that composition from the right is an endomorphism of the field $\mathbb{C}((x))$: if $f, g, h \in \mathbb{C}((x))$ and h is substituable in the fls on its left then $(f \pm g) \circ h = f \circ h \pm g \circ h$, $1 \circ h = 1$ and $(fg) \circ h = (f \circ h)(g \circ h)$.*

Exercise 4.3.12 *Composition is continuous (i.e., commutes with formal limits): if $f_n, g_n, f, g \in \mathbb{C}((x))$ with $f = \lim f_n$, $g = \lim g_n$ and $f \circ g$ is defined, then $f_n \circ g_n$ is defined for $n > n_0$ and*

$$\lim(f_n \circ g_n) = f \circ g \ .$$

Exercise 4.3.13 *Prove the formal chain rule: for $f, g \in \mathbb{C}((x))$, $(f \circ g)' = (f' \circ g)g'$ if the compositions are defined.*

We prove the expected property of formal composition, associativity. We expect it on the base of our experience with composing functions and mappings, where associativity holds trivially.

Proposition 4.3.14 *Composition is associative, for $f, g, h \in \mathbb{C}((x))$ one has*

$$(f \circ g) \circ h = f \circ (g \circ h)$$

whenever the involved compositions are defined.

Proof. Since every $f \in \mathbb{C}((x))$ is a limit $f = \lim f_n$ of some Laurent polynomials f_n, by Exercise 4.3.12 it suffices to prove associativity of formal composition just for them. Each Laurent polynomial f defines a mapping $f^* : \mathbb{C}\backslash\{0\} \to \mathbb{C}$, and the correspondence $f \mapsto f^*$ is injective in the sense that if $f^* = g^*$ on infinitely many points $x \in \mathbb{C}$ then $f = g$. It is easy to see that formal and functional composition commute: for every two Laurent polynomials we have $(f \circ g)^* = f^* \circ g^* : \mathbb{C}\backslash A \to \mathbb{C}$ where A is finite and on the right side we compose functions (if the formal composition is defined). Since composition of functions is trivially associative, it follows that formal composition of Laurent polynomials is associative. □

Exercise 4.3.15 *Give a purely formal proof that formal composition is associative, without relying on composition of functions.*

We establish the formal counterpart of inverse function.

Proposition 4.3.16 *Let $f \in \mathbb{C}((x))$. Then a $g \in \mathbb{C}((x))$ exists with*

$$f \circ g = x$$

if and only if $\operatorname{ord}(f) = 1$. *The fls g is then uniquely determined,* $\operatorname{ord}(g) = 1$ *as well and $f \circ g = g \circ f = x$; we call g the* compositional inverse of f *and denote it by*

$$g(x) = f(x)^{\langle -1 \rangle} .$$

Proof. It is easy to check that $\operatorname{ord}(f \circ g) = \operatorname{ord}(f) \cdot \operatorname{ord}(g)$ if $g \neq 0$ and the composition is defined. Thus if $\operatorname{ord}(f) \neq 1$ then $\operatorname{ord}(f \circ g) \neq 1$, and $\operatorname{ord}(f) = 1$ is a necessary condition for existence of a compositional inverse (we do not accept $x^{-1} \circ x^{-1} = x$, more on this later). We show that it is also a sufficient condition; let $f(x) = a_1 x + a_2 x^2 + \dots$ be a fls with $a_1 \neq 0$. We look for a $g \in \mathbb{C}((x))$ with $\operatorname{ord}(g) = 1$, so $g(x) = b_1 x + b_2 x^2 + \dots$ with $b_1 \neq 0$, such that

$$f \circ g = \sum_{n=1}^{\infty} a_n (b_1 x + b_2 x^2 + \dots)^n = x .$$

Expanding and comparing the coefficients of x^n on both sides, we look for solution in the unknown b_n of the infinite system of equations

$$a_1 b_n + p_n(a_1, \dots, a_n, b_1, \dots, b_{n-1}) = \delta_{1,n},\ n = 1, 2, \dots$$

where $p_n(\dots)$ are some integral polynomials in the stated variables; p_1 is zero. Again (as in the proof of Proposition 4.1.2) the system has a unique solution: $b_1 = a_1^{-1}$ from the first equation and if $b_1, b_2, \dots, b_{n-1}$ are already determined, the n-th equation determines uniquely b_n.

It remains to show that the right compositional inverse is also the left inverse. This follows by associativity. If g is the right inverse of f and h that of g then

$$g \circ f = (g \circ f) \circ (g \circ h) = g \circ ((f \circ g) \circ h) = x\ ,$$

and g is also the left inverse of f. □

The left inverse is unique too: $g \circ f = h \circ f \Rightarrow g = h$, upon composing $f^{\langle -1 \rangle}$ from the right. The GF $C(x) = \sum_{n \geq 1} c_n x^n$ of Catalan numbers satisfies the quadratic equation $x = C - C^2 = (x - x^2) \circ C$. It follows that $C(x)^{\langle -1 \rangle} = x - x^2$ and

$$x = C(x) - C(x)^2 = C(x - x^2) = \sum_{n \geq 1} c_n (x - x^2)^n = \sum_{n \geq 1} c_n x^n (1 - x)^n\ ,$$

which is equivalent with the identity

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

Example 4.3.17 We consider the extended composition $f \circ g$ of $f, g \in \mathbb{C}((x))$, defined if one of the cases (i), (ii) or (iii) occurs, that is, exactly if the series obtained by replacing x^n in the expansion of $f(x)$ with $g(x)^n$ formally converges. This operation is *not* associative. Indeed, let

$$f(x) = C(x) = \textstyle\sum_{n \geq 1} n^{-1} \binom{2n-2}{n-1} x^n,\ g(x) = C(x)^{\langle -1 \rangle} = x - x^2 \ \text{ and } \ h(x) = 1\ .$$

Then all extended compositions below are defined but

$$(f \circ g) \circ h = x \circ 1 = 1 \ \neq \ f \circ (g \circ h) = f \circ 0 = 0\ .$$

□

Exercise 4.3.18 *In the example, the three fps define functions analytic in a neighborhood of* 0*: f in the disc with radius $\frac{1}{4}$ and g, h are even entire. How come that their composition is not associative? Aren't we composing functions?*

Exercise 4.3.19 *Where does the proof of Proposition 4.3.14 fail for the fps $f = C(x), g = C(x)^{\langle -1 \rangle}, h = 1$?*

In view of Propositions 4.3.14 and 4.3.16, we have the following.

Proposition 4.3.20 *Let* $P_1 = \{f \in \mathbb{C}((x)) \mid \mathrm{ord}(f) = 1\}$. *Then* $(P_1, \circ)$ *is a non-commutative group.*

4.4 Composition of GF: exponential GF and the Lagrange inversion formula

4.5 Effective computation of modular reductions

4.6 More on algebra in $\mathbb{C}[[x]]$ and $\mathbb{C}((x))$

For a finite tuple of natural numbers $t = (t_1, \ldots, t_l) \in \bigcup_{i \geq 0} \mathbb{N}^i$ we use notation $|t| = t_1 + t_2 + \cdots + t_l$. The empty tuple $t = \emptyset$ for $l = 0$ has $|t| = 0$. We say that a sequence $(a_n) \subset K$ in a field K satisfies a *catalanian recurrence* if there is a finite set $S \subset K$ such that for every index $n = 1, 2, \ldots$ we have equality (empty product here is 1)

$$a_n = \sum_{0 \leq |t| \leq n,\ t \neq (n)} c_{n,t} a_{t_1} a_{t_2} \ldots a_{t_l},\ c_{n,t} \in S\ .$$

Thus $a_1 = c_{1,\emptyset}$, $a_2 = c_{2,\emptyset} + c_{2,(1)} a_1 + c_{2,(1,1)} a_1^2$, and so on.

Proposition 4.6.1 *Let a sequence* $(a_n) \subset K$ *satisfy a catalanian recurrence, with coefficients in a finite set* $S \subset K$.

1. *If* $R \subset K$ *is a subring such that* K *is a field of fractions of* R *then there exists an element* $e \in R$ *such that* $e^n a_n \in R$ *for every* $n \in \mathbb{N}$.

2. *If* $K = (K, |\cdot|)$ *is a normed field then there exists a constant* $c > 1$ *such that* $|a_n| \leq c^n$ *for every* $n \in \mathbb{N}$.

Proof. 1. Induction on n. For every $x \in S$ there is an $r \in R$ with $rx \in R$. The product $d \in R$ of all these r (now we use finiteness of S) has the property that $dx \in R$ for every $x \in S$. It follows from the catalanian recurrence satisfied by (a_n) that if $d^{2i-1} a_i \in R$ for $i = 1, 2, \ldots, n-1$ then $d^{2n-1} a_n \in R$ as well. Thus $e = d^2$ works.

2. We bound the growth of a_n by a simpler majorant catalanian sequence, which we resolve by its GF which satisfies a quadratic equation. Let $d = \max |e|, e \in S$. We take the sequence $(b_n) \subset \mathbb{R}_{\geq 0}$ defined by the catalanian recurrence

$$b_n = \sum_{0 \leq |t| \leq n,\ t \neq (n)} d b_{t_1} b_{t_2} \ldots b_{t_l}$$

$(b_1 = d)$. It follows by induction on n that $|a_n| \le b_n$ for every $n \in \mathbb{N}$. Let $B = B(x) = \sum_{n\ge1} b_n x^n$. The recurrence translates to the equation

$$B = \frac{d}{(1-x)(1-B)} - dB - d, \quad \text{or} \quad (1+d)B^2 - B + \frac{dx}{1-x} = 0\ .$$

So, since $B(0) = 0$,

$$B(x) = \frac{1-(1-x)^{-1/2}(1-(1+2d)^2x)^{1/2}}{2(1+d)}\ .$$

As we noted earlier, $|\binom{\pm1/2}{n}| \le 1$ for every $n \in \mathbb{N}_0$. Newton's binomial theorem gives that

$$|a_n| \le b_n = [x^n]B(x) < \frac{(n+1)(1+2d)^{2n}}{2(1+d)} < (2(1+2d)^2)^n\ .$$

Thus $c = 2(1+2d)^2$ works, d being the maximum norm of an element in S. □

Proposition 4.6.2 *Let K be a field and $f(x) = \sum_{n\ge0} a_n x^n \in K[[x]]$ be a fps that is algebraic over $K(x)$. Then there exists an $m \in \mathbb{N}_0$ such that the shifted sequence of coefficients*

$$(b_n) = (a_{n+m}),\ n = 1, 2, \dots\ ,$$

satisfies a catalanian recurrence.

Proof. There are polynomials $p_0, p_1, \dots, p_k \in K[x]$, $k \in \mathbb{N}$ and $p_k \ne 0$, such that $\sum_{i=0}^k p_i(x)f(x)^i = 0$. We assume that k is minimum and reexpand $P(x,y) = \sum_{i=0}^k p_i(x)y^i$ by the binomial theorem as

$$P(x, y+z) = \sum_{i=0}^{k} q_i(x,y)z^i,\ q_i \in K[x,y]\ .$$

Since $q_1(x,y) = \sum_{i=1}^k ip_i(x)y^{i-1}$, q_1 is nonzero and has y-degree $k-1$, which implies that $q_1(x, f(x)) \ne 0$. Thus we may set $m = \operatorname{ord}(q_1(x, f(x))) \in \mathbb{N}_0$. We split f as $f(x) = u(x) + x^m v(x)$ where $u(x) = a_0 + a_1x + \cdots + a_mx^m$ and $v(x) = \sum_{n\ge m+1} a_n x^{n-m} = \sum_{n\ge1} b_n x^n$. In

$$0 = P(x,f) = P(x, u(x) + x^m v(x)) = \sum_{i=0}^{k} q_i(x, u(x))x^{im}v(x)^i$$

the polynomial $q_1(x, u(x))$ has order exactly m because $\operatorname{ord}(u(x) - f(x)) \ge m+1$ and hence $\operatorname{ord}(q_1(x,u(x)) - q_1(x,f(x))) \ge m+1$. Therefore in the last sum

every summand with $i \geq 1$ has order at least $2m$, and so must have $q_0(x, u(x))$. Dividing by x^{2m} we get the identity

$$0 = \sum_{i=0}^{k} r_i(x) v(x)^i, \ r_i \in K[x], \ r_1(0) \neq 0 \ .$$

Equating the coefficient of x^n for $n = 1, 2, \ldots$ on the right side to zero, we get a catalanian recurrence for the sequence $(b_n) = (a_{n+m})$ of coefficients in $v(x)$. Indeed, that coefficient equals

$$[x^n] r_0(x) + \sum_{i=1}^{k} \sum_{j \geq 0} [x^j] r_i(x) \sum_{t_l \in \mathbb{N}, \ t_1 + \cdots + t_i = n - j} b_{t_1} b_{t_2} \ldots b_{t_i} \ ,$$

where always $t_l \leq n$, the only term involving $t_l = n$ is $[x^0] r_1(x) \cdot b_n$, and the coefficients in the recurrence attain only finitely many values as the $r_i(x)$ are polynomials. □

Exercise 4.6.3 *What is the point of the manipulation producing the shift m? Could not we obtain some recurrence for $(a_n = [x^n] f(x))$ directly from the initial equation $P(x, f(x)) = 0$?*

Theorem 4.6.4 (convergence of algebraic fps) *Suppose that $K = (K, |\cdot|)$ is a normed field and $f(x) = \sum_{n \geq 0} a_n x^n \in K[[x]]$ is a fps that is algebraic over $K(x)$. Then $f(x)$ is convergent: there is a constant $c > 1$ such that*

$$|a_n| \leq c^n, \ n \in \mathbb{N} \ .$$

Proof. This follows by combining part 2 of Proposition 4.6.1 and Proposition 4.6.2. Indeed, if for some $m \in \mathbb{N}_0$ and $c > 1$ one has $|a_{n+m}| \leq c^n$ for every $n \in \mathbb{N}$, then there is a $d > 1$ such that $|a_n| < d^n$ for every $n \in \mathbb{N}$. Namely, $d = \max\{|a_1|, |a_2|, \ldots, |a_m|, c\}$. □

Theorem 4.6.5 (Eisenstein, 1851) *Let K be a field, $R \subset K$ be a subring such that K is a field of fractions of R, and $f(x) = \sum_{n \geq 0} a_n x^n \in K[[x]]$ be a fps that is algebraic over $K(x)$. Then there exists an element $e \in R$ such that $e^n a_n \in R$ for every $n \in \mathbb{N}$. In other words,*

$$f(ex) - a_0 \in R[[x]] \ .$$

Proof. This follows by combining part 1 of Proposition 4.6.1 and Proposition 4.6.2. Indeed, if for some $m \in \mathbb{N}_0$ and $e \in R$ one has $e^n a_{n+m} \in R$ for every $n \in \mathbb{N}$, then there is a $d \in R$ such that $d^n a_n \in R$ for every $n \in \mathbb{N}$. Namely, d is the product of the m denominators for $a_1, a_2, \ldots, a_m$, and e. □

The theorem was announced in [30] by the German (Prussian) mathematician Gotthold Eisenstein (1823–1852). He stated it for the case $R = \mathbb{Z}$ and $K = \mathbb{Q}$ in the following single sentence; his note contains no proof.

(. . .) In jeder Reihen-Entwicklung dieser Art, wenn sie nur aus einer *algebraischen Funktion* stammt, mag dieselbe übrigens explicite oder implicite gegeben sein, kommen in sämmtlichen Coëfficienten, so fern dieselben rational sind, als nothwendige Nenner, d. h. als solche, die sich nich weiter gegen Faktoren des Zählers fortheben lassen, stets nur eine *endliche Anzahl ganz bestimmter Primfaktoren* und deren Potenzen vor; es sind dieser Primzahlen zugleich die Divisoren einer aus der algebraischen Gleichung, der die Funktion Genüge leistet, leicht zu bildenden characteristischen Zahl, nämlich ihrer dem speciellen Werthe $x = 0$ entsprechenden von Gauß so genannten Determinante, welche bekanntlich nicht verschwiden darf, wenn die Reihen-Entwicklung überhaupt möglich sein soll; endlich kann statt x immer ein solches Vielfache von x gesetzt werden, daß alle Coëfficienten der Reihe in ganze Zahlen übergehen. (. . .)

The first published proof is due to Eduard H. Heine (1821–1881), a German mathematician, two years later in [42]; by then Eisenstein, not yet 30, succumbed to tuberculosis. (We have learned about Heine's work from Allouche [5].) Our proof is a reworked proof from Cassels [19, Chapter ?].

4.7 Comments and references

Chapter 5

Rational generating functions

5.1 Generalities

We have to prove that wlog $\alpha_i \in \mathbb{Q}$. This is based on the useful fact that the coefficients in the recurrence may be taken from the field containing the sequence; we apply the lemma for the extension $\mathbb{Q} \subset \mathbb{C}$. The lemma follows from the next result that every system of linear equations with more unknowns than equations has a nontrivial solution.

Exercise 5.1.1. *Let K be a field and $\sum_{i=1}^{n} a_{i,j}x_i = 0$, $j = 1, 2, \ldots, m$ and $a_{i,j} \in K$, be a system of m (homogeneous and linear) equations with n unknowns, where $n > m$. Prove that the system has a solution $x_i \in K$ with not all x_i zero.*

Proposition 5.1.2. *Let $K \subset L$ be an extension of fields, $(a_n) \subset K$ be a sequence, and $\alpha_1, \ldots, \alpha_k \in L$ be k coefficients, not all zero, such that*

$$\sum_{i=1}^{k} \alpha_i a_{n+i} = 0, \text{ for every } n > n_0 \, .$$

Then some k coefficients $\beta_1, \ldots, \beta_k \in K$ exist, not all zero, such that

$$\sum_{i=1}^{k} \beta_i a_{n+i} = 0, \text{ for every } n > n_0 \, .$$

Proof. Let $U = \{x \in K^k \mid \sum_{i=1}^{k} \alpha_i x_i = 0\}$. Clearly, U is a vector subspace of the k-dimensional vector space K^k over K, and $U \neq K^k$. Thus if $b_j = (b_{j,1}, \ldots, b_{j,k}) \in U$, $j = 1, 2, \ldots, l$, form the basis of U, then $l < k$. By the previous exercise there is a $(\beta_1, \ldots, \beta_k) \in K^k$ with not all $\beta_i = 0$ such

that $\sum_{i=1}^{k} \beta_i b_{j,i} = 0$ for every $j = 1, 2, \ldots, l$. Thus $\sum_{i=1}^{k} \beta_i b_i = 0$ for every $(b_1, \ldots, b_k) \in U$. Since every k-tuple $(a_{n+1}, \ldots, a_{n+k})$ with $n > n_0$ lies in U, $\beta_1, \ldots, \beta_k$ are the desired coefficients. □

5.2 The transfer matrix method

5.3 Counting lattice points in polytopes

A *polyhedron* $M \subset \mathbb{R}^d$ is a set defined by finitely many linear inequalities,

$$M = \{x \in \mathbb{R}^d \mid \sum_{j=1}^{d} a_{i,j} x_j \geq b_i,\ i = 1, 2, \ldots, k\} ,$$

where $a_i = (a_{i,1}, a_{i,2}, \ldots, a_{i,d}) \in \mathbb{R}^d$ are k given vectors and $b_i \in \mathbb{R}$ are k given numbers. If we allow in the definition of M also inequalities $\leq$ and equalities $=$, we get again polyhedra.

Exercise 5.3.1 *Why? What about strict inequalities $>$?*

Geometricly, M is an intersection of k closed halfspaces. A *polytope* is a bounded polyhedron. A well known fact is that polytopes are exactly convex hulls of finite sets:

$$\begin{aligned} P \subset \mathbb{R}^d \text{ is a polytope } \iff P &= \operatorname{conv}(V) \\ &:= \{\textstyle\sum_{i=1}^{l} \lambda_i v_i \mid \lambda_i \geq 0 \ \&\ \sum_{i=1}^{l} \lambda_i = 1\} , \end{aligned}$$

where $V = \{v_1, \ldots, v_l\} \subset \mathbb{R}^d$ are l given points. The above expression $\sum_{i=1}^{l} \lambda_i v_i$ (plus the restriction $\lambda_i \geq 0$ & $\sum_{i=1}^{l} \lambda_i = 1$) for a point of P is the *convex combination* of the points $v_1, \ldots, v_l$. If a point v_j is a convex combination of other points in V, we may omit it from V without diminishing $\operatorname{conv}(V)$.

Exercise 5.3.2 *Why?*

The unique minimum subset $V' \subset V$ such that $P = \operatorname{conv}(V) = \operatorname{conv}(V')$ consists of the *vertices of P*.

Exercise 5.3.3 *Prove that $v \in P \subset \mathbb{R}^d$ is a vertex of the polytope P if and only if there is a closed halfspace $H \subset \mathbb{R}^d$ such that $H \cap P = \{v\}$.*

We say that the polytope P is a *lattice polytope*, respectively a *rational polytope*, if the vertices of P lie in $\mathbb{Z}^d$, respectively in $\mathbb{Q}^d$. For $n \in \mathbb{N}_0$ and a set $M \subset \mathbb{R}^d$ we define the n-th blow-up of M as

$$nM := \{nx = (nx_1, nx_2, \ldots, nx_d) \mid x \in M\} .$$

The main goal of the section is to deduce the next theorem.

Theorem 5.3.4 (Ehrhardt, 1960) *Let $P \subset \mathbb{R}^d$ be a rational polytope and $m > 0$ be an integer such that mP is a lattice polytope. Then the counting function*

$$f_P(n) := |\mathbb{Z}^d \cap nP|, \ n \in \mathbb{N}_0,$$

is a rational quasipolynomial in n with period m. In particular, if P is a lattice polytope then $f_p(n)$ is a rational polynomial in n

For example, the denumerant $p_A(n)$, $A = \{1, 2\}$, which counts partitions of n into parts $1, 2$, is clearly given by $p_A(n) = \lfloor n/2 \rfloor + 1$. But also $p_A(n) = f_P(n)$ for the polytope $P \subset \mathbb{R}^2$,

$$P = \{x \in \mathbb{R}^2 \mid x_1 + 2x_2 = 1, \ x_i \geq 0\} = \text{conv}(\{(1, 0), (0, 1/2)\}) \ .$$

Indeed, $p_A(n) = f_P(n) = n/2 + 1$ for even n and $n/2 + 1/2$ for odd n, a rational quasipolynomial with period 2.

5.4 Comments and references

Chapter 6

Analytic intermezzo II. Asymptotics via complex analysis

6.1 Comments and references

Chapter 7

Lattice walks

7.1 Random (?) walks in $\mathbb{Z}^d$ and other graphs

A *(simple undirected) graph*

$$G = (V, E)$$

has a finite or infinite set V of *vertices* and a set $E \subset \binom{V}{2} = \{e \subset V \mid |e| = 2\}$ of *edges*. A *walk* w in G is a finite or infinite sequence of vertices ($n \in \mathbb{N}_0$)

$$w = (v_0, v_1, \dots, v_n) \subset V \quad \text{or} \quad w = (v_0, v_1, \dots) \subset V$$

such that $\{v_{i-1}, v_i\} \in E$ for every $i \in [n]$ or for every $i \in \mathbb{N}$. The *length* of w is n in the former case and ∞ in the latter. For example, every vertex $v \in V$ is a walk in G with length 0. Injective walk with $i \neq j \Rightarrow v_i \neq v_j$ is called a *path*. The vertices v_0 and v_n are, respectively, the *starting* and the *final* vertex of w. For $v \in V$ and $n \in \mathbb{N}_0$ we denote by

$$d_n = d_n(v, G) \in \mathbb{N}_0 \quad \text{and} \quad a_n = a_n(v, G) \in \mathbb{N}_0 \,,$$

respectively, the number of walks w in G with length $n \in \mathbb{N}_0$ and $v_0 = v$, and the number of those of them with $v_0 = v_j = v$ for some $j > 0$ (i.e., revisitting its starting vertex v). The *degree* $\deg v = \deg_G v \in \mathbb{N}_0 \cup \{\infty\}$ of a vertex $v \in V$ in G is the number of neighbours of v in G and equals to the cardinality

$$\deg_G v = |\{\{v, u\} \mid u \in V\} \cap E| \,.$$

If all degrees in G are finite and the same, equal to an $r \in \mathbb{N}_0$, we call graph G *r-regular*.

Basic graphs for this section are, for $d \in \mathbb{N}$, the countable $2d$-regular graphs $\mathbb{Z}^d$ of nearest neighbours among the lattice points in $\mathbb{R}^d$,

$$\mathbb{Z}^d = (\mathbb{Z}^d, E) \quad \text{where} \quad \{a, b\} \in E \iff \sum_{i=1}^{d} |a_i - b_i| = 1 \,.$$

That is, $\{a, b\}$ is an edge in $\mathbb{Z}^d$ if and only if

$$a - b \in \{(\pm 1, 0, 0, \ldots, 0),\ (0, \pm 1, 0, 0, \ldots, 0),\ \ldots,\ (0, 0, \ldots, 0, \pm 1)\}$$

(we can move from a to b and back by a unit step in the direction of one of the d coordinated axes). For every vertex $v \in \mathbb{Z}^d$, every $d \in \mathbb{N}$ and every $n \in \mathbb{N}_0$ we clearly have

$$d_n(v, \mathbb{Z}^d) = (2d)^n \ .$$

In 1921, the Hungarian mathematician G. Pólya in [66] posed and answered the following original and intriguing question, which was the starting point of the new probabilistic discipline of random walks. A man aimlessly and randomly wanders ("Irrfahrt") in a rectangular net of streets ("Straßennetz"), as is the one in the city New York. From the point of view of probability calculus ("Wahrscheinlichkeitsrechnung"), how likely is it that he returns to the starting point (crossroad)? We model the net of streets and crossroads with the graph $\mathbb{Z}^d$ where New York has $d = 2$. The answer G. Pólya found is that for $d = 1, 2$ the wanderer always, with probability 1, sooner or later returns to the start, but in dimensions $d \geq 3$ with a positive probability he never returns to the start. Now we state these results precisely and prove them. We cast them enumeratively, without invoking (and defining) probability.

Proposition 7.1.1 (wandering in $\mathbb{Z}^2$). *Consider any starting vertex in $\mathbb{Z}^2$, say the origin $\overline{0} = (0, 0)$. Then*

$$\lim_{n\to\infty} \frac{a_n(\overline{0}, \mathbb{Z}^2)}{d_n(\overline{0}, \mathbb{Z}^2)} = \lim_{n\to\infty} \frac{a_n}{d_n} = \lim_{n\to\infty} \frac{a_n}{4^n} = 1 \ .$$

Vaguely restated, random walk in $\mathbb{Z}^2$ returns to the start with probability 1.

Proof. Let $w = (v_0, v_1, \ldots, v_n)$ be a walk in $\mathbb{Z}^2$ with $v_0 = \overline{0}$ and length $n \in \mathbb{N}_0$. We define two more quantities counting walks w starting at $\overline{0}$: b_n is the number of w with $v_n = \overline{0}$ and c_n (no Catalan numbers now) is the number of w with $v_n = \overline{0}$ and $v_j \neq \overline{0}$ for $0 < j < n$. We set $c_0 = 0$. Clearly, for $n \in \mathbb{N}$ we have $a_n \leq d_n$, $c_n \leq b_n \leq d_n$ and $d_n = 4^n$. Pigeonholing the walks counted by a_n by their first revisit of $\overline{0}$ at step j, we see that

$$a_n = \sum_{j=0}^{n} c_j d_{n-j}, \quad \text{or} \quad \frac{a_n}{4^n} = \sum_{j=0}^{n} \frac{c_j}{4^j} \leq 1$$

for every $n \in \mathbb{N}_0$ ($a_n \leq 4^n$). Thus it suffices to show that

$$\sum_{j=0}^{\infty} \frac{c_j}{4^j} = 1 \ .$$

The GFs $B(x) := \sum_{n\geq 0} \frac{b_n}{4^n} x^n = 1 + \ldots$ and $C(x) := \sum_{n\geq 0} \frac{c_n}{4^n} x^n = \frac{x^2}{4} + \ldots$ satisfy the relation

$$B(x) = \frac{1}{1 - C(x)} = \sum_{k\geq 0} C(x)^k$$

—split a walk counted by b_n by its returns to $\overline{0}$ into k segments. This relation also holds for the real functions $B(x)$ and $C(x)$ if $x \in [0,1)$ because both power series have radii of convergence ≥ 1 ($b_n, c_n \leq 4^n$). Now it suffices to show that

$$\lim_{x \to 1^-} B(x) = +\infty \, .$$

Indeed, then by the above relation $\lim_{x\to 1^-} C(x) = 1$. But by easy Abel's Theorem 7.1.5 below, $\lim_{x\to 1^-} C(x) = C(1)$. Thus $C(1) = 1$, which is the above required infinite series evaluation. To show that $\lim_{x\to 1^-} B(x) = +\infty$, it suffices to prove, again by Theorem 7.1.5, that $B(1) = +\infty$. We show it by computing b_n. Clearly, $b_n = 0$ for odd n (Exercise 7.1.2). For even lengths,

$$b_{2n} = \sum_{j=0}^{n} \frac{(2n)!}{j!\, j!\, (n-j)!\, (n-j)!} = \binom{2n}{n} \sum_{j=0}^{n} \binom{n}{j}^2 = \binom{2n}{n}^2 \, .$$

The first equality follows by considering all positions of the j steps of w to the right, which force the same number j of steps to the left and the same number $n-j$ for steps up and for steps down, thus the multinomial coefficient $\binom{2n}{j,\, j,\, n-j,\, n-j}$. The last equality follows from the well known binomial identity $\sum_{j=0}^{n} \binom{n}{j}^2 = \binom{2n}{n}$ (Exercise 7.1.3). We know from Corollary 1.3.10, say, that $\binom{2n}{n} \sim cn^{-1/2} 4^n$ with a constant $c > 0$. Hence the n-th summand in $B(1)$ for even n is $\sim cn^{-1}$ and

$$B(1) = \sum_{n=0}^{\infty} \frac{b_n}{4^n} = \sum_{n=0}^{\infty} \binom{2n}{n}^2 4^{-2n} = +\infty$$

because $\sum n^{-1} = +\infty$. □

Exercise 7.1.2. *Why does every walk in $\mathbb{Z}^2$ starting and finishing at the same vertex have an even length?*

Exercise 7.1.3. *Prove that for every $n \in \mathbb{N}_0$,*

$$\sum_{j=0}^{n} \binom{n}{j}^2 = \binom{2n}{n} \, .$$

Exercise 7.1.4. *Prove Proposition 7.1.1 in dimension one,*

$$\lim_{n\to\infty} \frac{a_n(0, \mathbb{Z}^1)}{d_n(0, \mathbb{Z}^1)} = \lim_{n\to\infty} \frac{a_n(0, \mathbb{Z}^1)}{2^n} = 1 \, .$$

Theorem 7.1.5 (easy Abel's theorem). *If a power series*

$$U(x) = \sum_{n \geq 0} u_n x^n \in \mathbb{R}[[x]]$$

converges for every $x \in [0, R)$, where $0 < R < +\infty$, and has all $u_n \geq 0$ then always

$$\lim_{x \to R^-} U(x) = \sum_{n=0}^{\infty} u_n R^n \ (=: U(R))$$

—no matter whether the limit and the sum are finite or $+\infty$.

Proof. For every $N \in \mathbb{N}$,

$$\sum_{n=0}^{N} u_n R^n \leq \lim_{x \to R^-} U(x) = \lim_{x \to R^-} \sum_{n=0}^{\infty} u_n x^n \leq \sum_{n=0}^{\infty} u_n R^n \ .$$

The first inequality follows from nonnegativity of u_n and the fact that for every $n \in \mathbb{N}_0$, $\lim_{x \to R^-} \frac{x^n}{R^n} = 1$. The second inequality follows just from nonnegativity of u_n. Sending $N \to +\infty$ gives the theorem. □

We turn to wandering on streets of a city in dimension $d = 3$. This might be one of the cities that Marco Polo describes to Kublai Khan in the book [16], perhaps the city Hypatia.

Proposition 7.1.6 (wandering in $\mathbb{Z}^3$). *We again start wandering at the origin $\overline{0} = (0, 0, 0)$. Then*

$$\lim_{n \to \infty} \frac{a_n(\overline{0}, \mathbb{Z}^3)}{d_n(\overline{0}, \mathbb{Z}^3)} = \lim_{n \to \infty} \frac{a_n}{6^n} < 1 \ .$$

Vaguely restated, random walk in $\mathbb{Z}^3$ returns to the start with probability smaller than 1 (and with positive probability disappears in infinity without return).

Proof. We define the quantities b_n and c_n and the generating functions $B(x)$ and $C(x)$ like before in the proof of Proposition 7.1.1, so $B(x) := \sum_{n \geq 0} \frac{b_n}{6^n} x^n$ and $C(x) := \sum_{n \geq 0} \frac{c_n}{6^n} x^n$. The argument does not change, but now we have to show that $B(1) < +\infty$, i.e. that the series $\sum_{n \geq 0} \frac{b_n}{6^n}$ converges. Then since as before $B(x) = \frac{1}{1 - C(x)}$ and $B(1) = \lim_{x \to 1^-} B(x)$ and $C(1) = \lim_{x \to 1^-} C(x)$ by Theorem 7.1.5, we get $C(1) = \lim_{x \to 1^-} C(x) < 1$. By this we are done as, like before, $C(1)$ equals to the above displayed limit.

We prove convergence of $\sum_{n \geq 0} \frac{b_n}{6^n}$. For odd n again $b_n = 0$. We upperbound $b_{2n}/6^{2n}$. Like in the proof of Proposition 7.1.1,

$$\begin{aligned} \frac{b_{2n}}{6^{2n}} &= \frac{1}{6^{2n}} \sum_{j,k \in \mathbb{N}_0,\, j+k \leq n} \frac{(2n)!}{j!\, j!\, k!\, k!\, (n-j-k)!\, (n-j-k)!} \\ &= \binom{2n}{n} 4^{-n} \sum_{\cdots} \left[\frac{1}{3^n} \binom{n}{j,\, k,\, n-j-k} \right]^2 \ . \end{aligned}$$

Here j is the number of steps of the walk to the right, k the number of steps up and $n - j - k$ the number of steps back. The numbers in $[\cdots]$s sum up to 1, by

the multinomial expansion $3^n = (1+1+1)^n = \sum_{x+y+z=n} \binom{n}{x,\,y,\,z}$, $x, y, z \in \mathbb{N}_0$. Thus by Exercises 7.1.7 and 7.1.8,

$$\sum_{\dots} [\cdots]^2 \le \max_{x,y,z \in \mathbb{N}_0,\, x+y+z=n} \frac{1}{3^n} \binom{n}{x,\,y,\,z} = \frac{1}{3^n} \binom{n}{x_0,\,y_0,\,z_0}$$

where $(m \in \mathbb{N})$ (x_0, y_0, z_0) equals (m, m, m) if $n = 3m$, $(m+1, m, m)$ if $n = 3m+1$ and $(m+1, m+1, m)$ if $n = 3m+2$. By Exercise 7.1.9, always

$$\binom{n}{x_0,\,y_0,\,z_0} \ll \frac{3^n}{n} \,.$$

Since $\binom{2n}{n} 4^{-n} \sim cn^{-1/2}$, $c > 0$, we get the bound

$$\frac{b_{2n}}{6^{2n}} \ll n^{-1/2} n^{-1} = n^{-3/2} \,.$$

Hence

$$B(1) = \sum_{n \ge 0} \frac{b_n}{6^n} = \sum_{n \ge 0} \frac{b_{2n}}{6^{2n}} \ll \sum_{n \ge 1} \frac{1}{n^{3/2}} < +\infty \,.$$

□

Exercise 7.1.7. *If $\alpha_1, \dots, \alpha_k \ge 0$ are real numbers with $\sum_{i=1}^k \alpha_i = 1$ then*

$$\sum_{i=1}^k \alpha_i^2 \le \max_{1 \le i \le k} \alpha_i \,.$$

Exercise 7.1.8. *If $a > b \ge 0$ are integers with $a \ge b+2$ then*

$$a!\, b! > (a-1)!\, (b+1)! \,.$$

Exercise 7.1.9. *Show that $(m \in \mathbb{N})$ for $n = 3m$, $n = 3m+1$ and $n = 3m+2$, respectively,*

$$\frac{n!}{m!^3} \ll \frac{3^n}{n}, \quad \frac{n!}{(m+1)!\, m!^2} \ll \frac{3^n}{n} \quad \textit{and} \quad \frac{n!}{(m+1)!^2\, m!} \ll \frac{3^n}{n} \,.$$

Exercise 7.1.10. *Prove Proposition 7.1.6 for any dimension $d \ge 3$,*

$$\lim_{n \to \infty} \frac{a_n(\overline{0}, \mathbb{Z}^d)}{d_n(\overline{0}, \mathbb{Z}^d)} = \lim_{n \to \infty} \frac{a_n(\overline{0}, \mathbb{Z}^d)}{(2d)^n} < 1 \,.$$

7.2 Selfwavoiding walks in the honeycomb lattice

7.3 Algebra of power series and lattice paths

Comments and references

N. H. Abel proved his Theorem **??** in [1, Lehrsatz IV] and inequality (Exercise **??**) in [1, Lehrsatz III]. His paper was republished as a booklet in 1890's, is today accessible on-line and worth reading.

Hints and solutions to some exercises

Exercise 1.1.4. Deutsch and Sagan [24] proved that if $A \subset \mathbb{N}$ is the set of numbers that may have in its ternary expansion digit 2 only on the rightmost place, and if for $n \in A$ the number of 1s in this expansion, the rightmost place not counted, is denoted $r(n)$, then $c_n \equiv (-1)^{r(n)}$ modulo 3 for $n \in A$ and $c_n \equiv 0$ else.

Exercise 1.3.1. $C - x = x \sum_{k \geq 1} C^k$. What quadratic equation this gives?

Exercise 1.3.2. This is an ODE with separated variables.

Exercise 1.3.3. The first $\leq$ is equality only for $n = 1$ and the second $\leq$ only for $n = 2$, so for every $n \in \mathbb{N}$ at least one of them is strict and compose to $<$.

Exercise 1.3.4. Bound the binomial coefficient $\binom{2n-2}{n-1}$ both from above and below by expanding $(1+1)^{2n-2}$.

Exercise 1.3.5. For example, the GF $A = A(x)$ of rp trees such that every vertex has 0 or 5 children satisfies the equation $A - x = xA^5$.

Exercise 1.3.6. The argument is correct because the set of specializations for α, β is not just infinite but has the grid form $\mathbb{N} \times \mathbb{N}$, an infinite set times an infinite set.

Exercise 1.3.11. By Theorem 1.3.9 the ratio has asymptotics $cn^{1/6} \left(\frac{4}{3}\right)^n$.

Exercise 1.5.4. Use Bachet's identity: $\alpha p + \beta q = 1$ for some $\alpha, \beta \in \mathbb{C}[x]$.

Exercise 1.5.5. The eventual periodicity of c_n modulo p would mean that, in the ring $(\mathbb{Z}/p\mathbb{Z})[[x]]$, $C(x) \bmod p = a(x)/b(x)$ where $a, b \in (\mathbb{Z}/p\mathbb{Z})[x]$. Bring it to contradiction by tools of this proof.

Exercise 1.7.3. Turn the square upside down.

Exercise 2.1.1. Bound the remainder $\sum_{m \geq n+1} m^{-2}$ by $\int_n^{+\infty} x^{-2}\, dx$.

Exercise 2.1.2. Use the expansion $\exp x = \sum_{i=0}^{\infty} x^i / i!$.

Exercise 3.1.2. Empty properties $Q^n = \emptyset$ are fine, for such n just set $t(n) = n$, say. But if $Q^n = \mathcal{P}([n])$ for infinitely many n then (Q^n) does not have threshold function (consider $m(n) = 0$).

Bibliography

[1] N. H. Abel, Untersuchungen über die Reihe: $1 + \frac{m}{1}x + \frac{m\cdot(m-1)}{1\cdot 2}\cdot x^2 + \frac{m\cdot(m-1)\cdot(m-2)}{1\cdot 2\cdot 3}\cdot x^3 + \ldots$, *Journal für die reine und angewandte Mathematik* **1** (1826), 311–339.

[2] M. Aigner, *A Course in Enumeration*, Springer, Berlin, 2007.

[3] M. Aigner and G. M. Ziegler, *Proofs from THE BOOK*, Springer, Berlin, 2001.

[4] G. L. Alexanderson, *The Random Walks of George Pólya*, MAA, Washington, DC, 2000.

[5] J.-P. Allouche, Transcendence of formal power series with rational coefficients, *Theor. Comput. Sci.* **218** (1999), 143–160.

[6] V. de Angelis, Stirling's series revisited, *Amer. Math. Monthly* **116** (2009), 839–843.

[7] A. Barvinok, *Integer Points in Polyedra*, European Mathematical Society, Zurich, 2008.

[8] M. Beck and S. Robins, *Computing the Continuous Discretely. Integer-point Enumeration in Polyhedra*, Springer, New York, 2007.

[9] M. Beck and F. Sottile, Irrational proofs for three theorems of Stanley, *European J. Combin.* **28** (2007), 403–409.

[10] J. P. Bell, S. N. Burris and K. A. Yeats, Counting rooted trees: the universal law $t(n) \sim C\rho^{-n}n^{-3/2}$, *Electron. J. Combin.*, **13** (2006), Research Paper 63, 64 pp.

[11] F. Bergeron, G. Labelle and P. Leroux, *Combinatorial Species and Tree-like Structures*, Cambridge University Press, Cambridge, UK, 1998. [Translated from the 1994 French edition.]

[12] C. R. Blyth and P. K. Pathak, A note on easy proofs of Stirling's theorem, *Amer. Math. Monthly* **93** (1986), 376–379.

[13] M. Bóna, *Combinatorics of Permutations*, CRC Press, Boca Raton, FL, 2012.

[14] B. Bollobás and A. Thomason, Threshold functions, *Combinatorica* **7** (1986), 35–38.

[15] N. G. de Bruijn, *Asymptotic Methods in Analysis*, North-Holland, Amsterdam, 1957.

[16] I. Calvino, *Invisible Cities*

[17] P. J. Cameron, *Notes on Counting: An Introduction to Enumerative Combinatorics*, preprint, 2010, 217 pp. [Available from the author's homepage.]

[18] R. Carnap, *Logical Foundations of Probability*, The University of Chicago Press, Chicago, Ill., 1950.

[19] J. W. S. Cassels, *Local Fields*, Cambridge University Press, Cambridge, UK, 1986.

[20] Y.-G. Chen and W. Jiang, Catalan numbers modulo a prime power, *Integers* **13** (2013), Paper No. A36, 4 pp.

[21] A. J. Coleman, A simple proof of Stirling's formula, *Amer. Math. Monthly* **58** (1951), 334–336.

[22] L. Comtet, *Advanced Combinatorics*, D. Reidel Publishing Co., Dordrecht, 1974.

[23] A. De, P. Kurur, Ch. Saha and R. Saptharishi, Fast integer multiplication using modular arithmetic, *SIAM J. Comput.* **42** (2013), 685–699.

[24] E. Deutsch and B. E. Sagan, Congruences for Catalan and Motzkin numbers and related sequences, *J. Number Theory* **117** (2006), 191–215.

[25] P. Diaconis and D. Freedman, An elementary proof of Stirling's formula, *Amer. Math. Monthly* **93** (1986), 123–125.

[26] Peter G. Doyle and J. Laurie Snell, *Random walks and electric networks*, MAA, Washington, DC, 1984.

[27] D. E. Dutkay, C. P. Niculescu and F. Popovici, Stirling's formula and its extension for the gamma function, *Amer. Math. Monthly* **120** (2013), 737–740.

[28] H. M. Edwards, Roots of solvable polynomials of prime degree, *Expo. Math.* **32** (2014), 79–91.

[29] G. Eisenstein, Allgemeine Auflösung der Gleichungen von den ersten vier Graden, *J. Reine Angew. Math.* **27** (1844), 81–83.

[30] G. Eisenstein, Über eine allgemeine Eigenschaft der Reihen-Entwicklungen aller algebraischen Funktionen, *Bericht Königl. Preuss. Akad. d. Wiss. zu Berlin* (1851), 441–443.

[31] S.-P. Eu, S.-Ch. Liu and Y.-N. Yeh, Catalan and Motzkin numbers modulo 4 and 8, *European J. Combin.* **29** (2008), 1449–1466.

[32] G. Everest, A. van der Poorten, I. Shparlinski and T. Ward, *Recurrence Sequences*, American Mathematical Society, Providence, RI, 2003.

[33] R. Fagin, The number of finite relational structures, *Discrete Math.* **19** (1977), 17–21.

[34] W. Feller, A direct proof of Stirling's formula, *Amer. Math. Monthly* **74** (1967), 1223–1225.

[35] G. Fischer, *Plane Algebraic Curves*, AMS, Providence, RH, 2001.

[36] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, UK, 2010. [Preliminary version available on the Internet from the web page of the late P. F.]

[37] P. Frankl, A new short proof for the Kruskal–Katona theorem, *Discrete Math.* **48** (1984), 327–329.

[38] M. Fürer, Faster integer multiplication, *SIAM J. Comput.* **39** (2009), 979–1005.

[39] I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, John Wiley & Sons, Inc., New York, 1983.

[40] Y. Gurevich and S. Shelah, Time polynomial in input or output, *J. Symbolic Logic* **54** (1989), 1083–1088.

[41] F. Harary, Note on Carnap's relational asymptotic relative frequencies, *J. Symb. Logic* **23** (1958), 257–260.

[42] E. Heine, Der Eisensteinsche Satz über Reihen-Entwicklung algebraischer Functionen, *J. Reine Angew. Math.* **45** (1853), 285–302.

[43] Ch. Impens, Stirling's series made easy, *Amer. Math. Monthly* **110** (2003), 730–735.

[44] M. Kauers, C. Krattenthaler and T. Müller, A method for determining the mod-$2k$ behaviour of recursive sequences, with applications to subgroup counting, *Electron. J. Combin.* **18** (2011), Paper 37, 83 pp.

[45] R. A. Khan, A probabilistic proof of Stirling's formula, *Amer. Math. Monthly* **81** (1974), 366–369.

[46] M. Klazar, Some general results in combinatorial enumeratio, *Permutation patterns*, 3–40, London Math. Soc. Lecture Note Ser., 376, Cambridge Univ. Press, Cambridge, 2010.

[47] M. Klazar, So what is an answer?—a few remarks and ideas on PIO algorithms in combinatorial enumeration, in preparation.

[48] S. K. Lando, *Lectures on Generating Functions*, The American Mathematical Society, Providence, RI, USA, 2003 (translated from the Russian original by the author).

[49] S. Lang, *Algebra*, Springer, 2002 (revised 3rd edition).

[50] H. Lou, A short proof of Stirling's formula, *Amer. Math. Monthly* **121** (2014), 15–157.

[51] R. Lyons and Y. Peres, *Probability on Trees and Networks*, Cambridge University Press, New York, 2016.

[52] Yu. I. Manin, *A course in mathematical logic for mathematicians. Second edition. Chapters I–VIII translated from the Russian by Neal Koblitz. With new chapters by Boris Zilber and the author*, Springer, New York, 2010.

[53] A. J. Maria, A remark on Stirling's formula, *Amer. Math. Monthly* **72** (1965), 1096–1098.

[54] G. Marsaglia and J. C. W. Marsaglia, A new derivation of Stirling's approximation to $n!$, *Amer. Math. Monthly* **97** (1990), 826–829.

[55] R. C. Mason, *Diophantine Equations over Function Fields*, London Mathematical Society Lecture Note Series, 96, Cambridge University Press, 1984.

[56] R. Michel, On Stirling's formula, *Amer. Math. Monthly* **109** (2002), 388–390.

[57] R. Michel, The $(n+1)$th proof of Stirling's formula, *Amer. Math. Monthly* **115** (2008), 844–845.

[58] V. Namias, A simple derivation of Stirling's asymptotic series, *Amer. Math. Monthly* **93** (1986), 25–29.

[59] T. S. Nanjundiah, Note on Stirling's formula, *Amer. Math. Monthly* **66** (1959), 701–703.

[60] O. Nash, On Klein's icosahedral solution of the quintic, *Expo. Math.* **32** (2014), 99–120.

[61] T. Neuschel, A new proof of Stirling's formula, *Amer. Math. Monthly* **121** (2014), 350–352.

[62] I. Pak, History of Catalan numbers, arXiv:1408.5711v1, 2014.

[63] J. M. Patin, A very short proof of Stirling's formula, *Amer. Math. Monthly* **96** (1989), 41–42.

[64] S. J. Patterson, Eisenstein and the quintic equation, *Historia Math.* **17** (1990), 132–140.

[65] M. A. Pinsky, Stirling's formula via the Poisson distribution, *Amer. Math. Monthly* **114** (2007), 256–258.

[66] G. Pólya, Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz, *Math. Ann.* **84** (1921), 149–160.

[67] H. Robbins, A remark on Stirling's formula, *Amer. Math. Monthly* **62** (1955), 26–29.

[68] D. Romik, Stirling's approximation for $n!$: the ultimate short proof?, *Amer. Math. Monthly* **107** (2000), 556–557.

[69] S. Shelah and J. Spencer, Zero-one laws for sparse random graphs, *J. Amer. Math. Soc.* **1** (1988), 97–115.

[70] J. R. Shoenfield, *Mathematical logic*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont. 1967.

[71] J. Spencer, *The Strange Logic of Random Graphs*, Springer, Berlin, 2001.

[72] R. P. Stanley, *Enumerative Combinatorics. Volume 1*, Cambridge University Press, Cambridge, UK, 2012 (second edition, the fist edition appeared in 1986).

[73] R. P. Stanley, *Enumerative Combinatorics. Volume 2*, Cambridge University Press, Cambridge, UK, 1999.

[74] R. P. Stanley, *Catalan Addendum*, 96 pp., `http://www-math.mit.edu/~rstan/ec/catadd.pdf` (accessed May 30, 2014).

[75] R. P. Stanley, *Catalan Numbers*, Cambridge University Press, Cambridge, UK, 2015.

[76] J. Stillwell, Eisenstein's footnote, *Math. Intelligencer* **17** (1995), 58–62.

[77] W. W. Stothers, Polynomial identities and hauptmoduln, *Quarterly J. Math. Oxford*, **2** (1981), 349–370.

[78] P. Valtr, Probability that n random points are in convex position, *Discrete Comput. Geom.* **13** (1995), 637–643.

[79] P. Valtr, Catalan numbers via random planar point sets. In: *Intuitive geometry (Budapest, 1995)*, 441–443, Bolyai Soc. Math. Stud., 6, János Bolyai Math. Soc., Budapest, 1997.

[80] M. B. Villarino, The convergence of the Catalan number generating function, arXiv:1511.08555v2, 2016.

[81] H. S. Wilf, What is an answer?, *Amer. Math. Monthly* **89** (1982), 289–292.

[82] H. S. Wilf, *generatingfunctionology*, Academic Press, Inc., San Diego, Ca, USA, 1994 (second edition, available on the Internet).

[83] W. Woess, *Random Walks on Infinite Graphs and Groups*, Cambridge University Press, Cambridge, UK, 2000.

[84] K. Woods, The unreasonable ubiquitousness of quasi-polynomials, *Electron. J. Combin.* **21** (2014), Paper 1.44, 23 pp.

[85] M. Yannakakis, Algorithms for acyclic database schemas, 82–94. In: W. Chu, G. Gardarin, S. Ohsuga and Y. Kambayashi (Eds.), *Proc. 7th International Conference on Very Large Data Bases*, Morgan Kaufmann, Cannes, 1981.

[86] Wikipedia, `http://en.wikipedia.org/wiki/Main_Page`

[87] Complexity Zoo,
`https://complexityzoo.uwaterloo.ca/Complexity_Zoo` (Zookeeper: S. Aaronson).

Index