Algebraic combinatorics

RWTH Aachen, WS 2021
RWTH Aachen, SS 2017, WS 2017
Leibniz-Universität Hannover, SS 2013
Universität Duisburg-Essen, SS 2011

Jürgen Müller

Contents

0 What is combinatorics?

• *'Combinatorial theory' is the name now given to the subject formerly called 'combinatorial analysis' or 'combinatorics', though these terms are still used by many people. Like many branches of mathematics, its boundaries are not clearly defined, but the central problem may be considered that of arranging objects according to specified rules and finding out in how many ways this may be done.*

If the specified rules are very simple, then the chief emphasis is on the enumeration of the number of ways in which the arrangement may be made. If the rules are subtle or complicated, the chief problem is whether or not such arrangements exist, and to find methods for constructing the arrangements. An intermediate area is the relationship between related choices, and a typical theorem will assert that the maximum for one kind of choice is equal to the minimum for another kind. *[6, p.ix]*

• *The basic problem of enumerative combinatorics is that of counting the number of elements of a finite set S. This definition, as it stands, tells us little about the subject since virtually any mathematical problem can be cast in these terms. In a genuine enumerative problem, the elements of S will usually have a rather simple combinatorial definition and very little additional structure. It will be clear that S has many elements, and to main issue will be to count or estimate them all. ...*

There has been an explosive growth in combinatorics in recent years. ... One important reason for this growth has been the fundamental role that combinatorics plays as a tool in computer science and related areas. A further reason has been the prodigious effort ... to bring coherence and unity to the discipline of combinatorics. ... Enumerative combinatorics has been greatly elucidated by this effort, as has its role in such areas of mathematics as finite group theory, representation theory, commutative algebra, algebraic geometry, and algebraic topology. *[11, p.1]*

(0.1) Example: Fibonacci numbers. a) The following problem was posed in the medieval book 'Liber abbaci' [**Leonardo da Pisa 'Fibonacci', 1202**]: Any female rabbit gives birth to a couple of rabbits monthly, from its second month of life on. If there is a single couple in the first month, how many are there in month $n \in \mathbb{N}$?

Hence let $[F_n \in \mathbb{N}_0; n \in \mathbb{N}_0]$ be the **linear recurrent sequence** of **degree** 2 given by $F_0 := 0$ and $F_1 := 1$, and $F_{n+2} := F_n + F_{n+1}$ for $n \in \mathbb{N}_0$. Thus we obtain the sequence of **Fibonacci numbers**, see also Table 1:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
F_n	0	1	1	2	3	5	8	13	21	34	55	89	144	233	377	610

To find a closed formula for F_n, and to determine the growth behavior of F_n, we proceed as follows; see also (13.1) for a different treatment:

Table 1: Fibonacci numbers.

n	F_n
1	1
2	1
4	3
8	21
16	987
32	2178309
64	10610209857723
128	251728825683549488150424261
256	141693817714056513234709965875411919657707794958199867

Letting $A := \begin{bmatrix} . & 1 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ we have $A \cdot [F_n, F_{n+1}]^{\text{tr}} = [F_{n+1}, F_{n+2}]^{\text{tr}}$, thus $[F_n, F_{n+1}]^{\text{tr}} = A^n \cdot [F_0, F_1]^{\text{tr}}$ for $n \in \mathbb{N}_0$. We have $\chi_A = \det(XE_2 - A) = X^2 - X - 1 = (X - \rho_+)(X - \rho_-) \in \mathbb{R}[X]$, where $\rho_\pm := \frac{1}{2}(1 \pm \sqrt{5}) \in \mathbb{R}$. From $\ker(A - \rho_\pm E_2) = \langle [1, \rho_\pm]^{\text{tr}} \rangle_{\mathbb{R}}$ we get the diagonalising matrix $P := \begin{bmatrix} 1 & 1 \\ \rho_+ & \rho_- \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$, where $P^{-1} := \frac{1}{\rho_- - \rho_+} \cdot \begin{bmatrix} \rho_- & -1 \\ -\rho_+ & 1 \end{bmatrix}$. Thus we have $P^{-1}A^nP = (P^{-1}AP)^n = (\text{diag}[\rho_+, \rho_-])^n = \text{diag}[\rho_+^n, \rho_-^n]$, implying

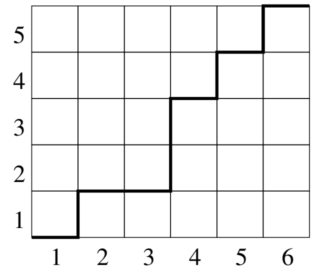
$$A^n = P \cdot \text{diag}[\rho_+^n, \rho_-^n] \cdot P^{-1} = \frac{1}{\rho_- - \rho_+} \cdot \begin{bmatrix} \rho_+^n \rho_- - \rho_+ \rho_-^n & \rho_-^n - \rho_+^n \\ \rho_+^{n+1} \rho_- - \rho_+ \rho_-^{n+1} & \rho_-^{n+1} - \rho_+^{n+1} \end{bmatrix}.$$

Hence we have the **Moivre-Binet Formula** [1718, 1843] ([Bernoulli, 1728]) $F_n = \frac{\rho_-^n - \rho_+^n}{\rho_- - \rho_+} = \frac{1}{\sqrt{5}}(\rho_+^n - \rho_-^n)$. Since $|\rho_+| > 1$ and $|\rho_-| < 1$ this yields $\lim_{n\to\infty} \frac{F_n \cdot \sqrt{5}}{\rho_+^n} = 1$, in particular the F_n grow exponentially.

b) The number $\rho_+ := \frac{1}{2}(1+\sqrt{5}) \in \mathbb{R}$ is called the **golden ratio**, featuring in the following classical problem: How has a line segment to be cut into two pieces, such that length ratio between the full segment and the longer piece coincides with the length ratio between the longer and the shorter piece? Assume that the line segment has length 1, and letting $\frac{1}{2} < x < 1$ be the length of the longer piece, we thus have $\frac{1}{x} = \frac{x}{1-x}$, or equivalently $x^2 + x - 1 = 0$, which yields $x = \frac{1}{2}(-1 + \sqrt{5}) \in \mathbb{R}$ as the unique positive solution. Thus the above ratio indeed equals $\frac{x}{1-x} = \frac{1}{x} = \frac{2}{-1+\sqrt{5}} = \frac{1}{2}(1 + \sqrt{5}) = \rho_+ \sim 1.61803$.

From the above we infer $\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \rho_+$, saying that the Fibonacci numbers yield (good) rational approximations of the golden ratio: For example, we have

Table 2: Paths in a square grid.



1	6	21	56	126	252
1	5	15	35	70	126
1	4	10	20	35	56
1	3	6	10	15	21
1	2	3	4	5	6
	1	1	1	1	1

$\frac{F_8}{F_7} = \frac{21}{13} \sim 1.61538$, $\frac{F_9}{F_8} = \frac{34}{21} \sim 1.61905$, $\frac{F_{10}}{F_9} = \frac{55}{34} \sim 1.61765$, $\frac{F_{11}}{F_{10}} = \frac{89}{55} \sim 1.61818$, $\frac{F_{12}}{F_{11}} = \frac{144}{89} \sim 1.61798$, $\frac{F_{13}}{F_{12}} = \frac{233}{144} \sim 1.61806$, $\frac{F_{14}}{F_{13}} = \frac{377}{233} \sim 1.61803$.

(0.2) Example: Shortest paths. Given an $(m \times n)$-square grid, where $m, n \in \mathbb{N}_0$, we want to determine the number of shortest paths from the bottom leftmost vertex to the upper rightmost vertex, see Table 2 for the case $m = 6$ and $n = 5$. Hence any of these paths is of length $m + n$, consisting of a series of m right moves and n up moves, which is uniquely determined by the m-subset indicating the right moves, or alternatively by the n-subset indicating the up moves.

Using coordinates, we count the number $f(i, j) \in \mathbb{N}$, where $i \in \{0, \ldots, m\}$ and $j \in \{0, \ldots, n\}$, of ways of getting from the bottom leftmost vertex $[0, 0]$ to vertex $[i, j]$. Hence we have $f(0, j) = f(i, 0) = 1$, and $f(i, j) = f(i - 1, j) + f(i, j - 1)$ for $i, j \geq 1$, yielding the pattern as shown in Table 2 for $m = n = 5$. We now derive a concise description of $f(m, n)$; see also (2.6):

The paths in question have length $m+n$, and are uniquely described by the m-set of positions where the right moves are made, or equivalently are uniquely described by the (complementary) n-set of positions where the up moves are made. Hence we infer that $f(m, n) = \binom{m+n}{m} = \binom{m+n}{n}$ is a **binomial coefficient**, see (2.1). The above recursion translates into $\binom{m+n}{m} = \binom{m+n-1}{m-1} + \binom{m+n-1}{n-1} = \binom{m+n-1}{m-1} + \binom{m+n-1}{m}$, for $m, n \geq 1$.

(0.3) Example: Parity of binomial coefficients. From the above considerations we derive **palindromicity** $\binom{n}{k} = \binom{n}{n-k}$, for $n \in \mathbb{N}_0$ and $k \in \{0, \ldots, n\}$, and the **triangle identity** $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$, for $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$. This shows by way of example how the 'path model' can be used to prove identities for binomial coefficients.

The triangle identity gives rise to the **Pascal triangle** shown in Table 3, allowing to compute binomial coefficients by using additions alone, but no multiplications. The Pascal triangle has a rich structure, and actually binomial

Table 3: The Pascal triangle.

$n\backslash k$	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

coefficients fulfill lots of (miraculous) identities. For example, we may ask for the parity of $\binom{n}{k}$, and in particular how many of the entries in row n of the Pascal triangle are odd; see also (2.7) for a more general treatment:

i) The positions of the odd entries in the Pascal triangle yield the following repeating pattern, see Table 4: If the 2^i rows $0, \ldots, 2^i - 1$ are given, where $i \in \mathbb{N}_0$, then the next rows $2^i, \ldots, 2^{i+1} - 1$ are obtained by copying down the given rows by 2^i rows, and then copying the new part to the right by 2^i columns:

By the binomial formula we have $(X+1)^n = \sum_{k=0}^{n} \binom{n}{k} X^k \in \mathbb{Z}[X]$, so that row n of the Pascal triangle can be interpreted as the polynomial $(X+1)^n \in \mathbb{Z}[X]$. Hence by reducing modulo 2, the parity pattern of the binomial coefficients in row n can be interpreted as the polynomial $(X+1)^n \in \mathbb{F}_2[X]$. Recall that we have $(X+1)^2 = X^2 + 1 \in \mathbb{F}_2[X]$.

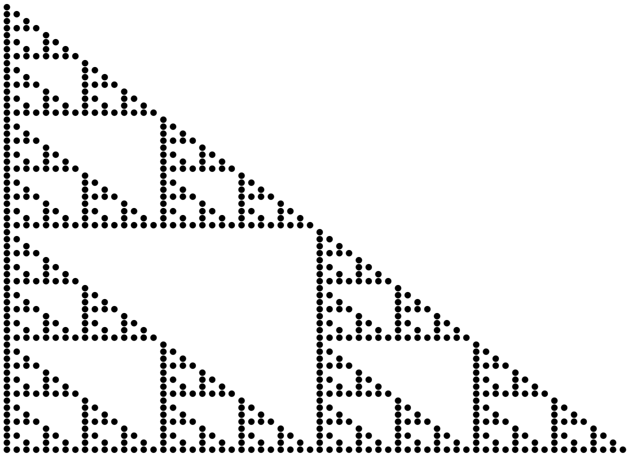
Now let $n = \sum_{i=0}^{l} a_i \cdot 2^i$ be the 2-adic representation of n, where $l = l(n) = \lfloor \log_2(n) \rfloor \in \mathbb{N}_0$ and $a_i = a_i(n) \in \{0, 1\}$; thus $a_l = 1$. Letting $n' := n - 2^l$ we have $n' \in \{0, \ldots, 2^l - 1\}$, and we get $(X+1)^n = (X+1)^{2^l + n'} = (X+1)^{2^l} (X+1)^{n'} = (X^{2^l} + 1)(X+1)^{n'}$. Hence the positions of the odd entries in row n are found as those in row n', occupying positions $\{0, \ldots, 2^l - 1\}$ (this is the piece copied down), as well as those in row n' shifted by 2^l steps to the right, occupying positions $\{2^l, \ldots, 2^{l+1} - 1\}$ (this is the piece copied to the right). $\sharp$

In particular, we observe that the Pascal triangle with 2^i rows, where $i \in \mathbb{N}_0$, has a total of 3^i odd entries.

ii) For the number of odd entries in row n of the Pascal triangle we observe:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
odd	1	2	2	4	2	4	4	8	2	4	4	8	4	8	8	16	2	4

Table 4: Parity in the Pascal triangle.



Letting again $n = \sum_{i=0}^{l} a_i \cdot 2^i$ be the 2-adic representation of n, the number of odd entries in row n is given as 2^w, where $w = w(n) := \sum_{i=0}^{l} a_i \in \mathbb{N}_0$ is the number of binary digits 1 in the 2-adic representation of n:

We proceed by induction, the assertion being clear for $n = 0$. By the repeating pattern described above, the number of odd entries in row $n \geq 1$ is twice the number of odd entries in row $n' := n - 2^l = \sum_{i=0}^{l-1} a_i \cdot 2^i$. The latter is given as $2^{w(n')}$, where $w(n') = w(n) - 1$, so that for row n we get $2 \cdot 2^{w(n')} = 2^w$. ♯

(0.4) The Sierpinski Sieve [1915]. Considering the Pascal triangles with 2^i rows, for $i \in \mathbb{N}_0$, and the repeating pattern of odd entries reveals a remarkable relationship to the following construction in fractal geometry. Actually, the construction of the sets $\mathcal{T}_i$ below shows that there is a natural bijection between the set of triangles in $\mathcal{T}_i$ and the odd entries in the Pascal triangle with 2^i rows:

We consider the open rectangular triangle $\mathcal{T}_0$ in the Euclidean plane with vertices $[0, 0]$, $[1, 0]$ and $[0, 1]$. We cut out the closed triangle with vertices $[\frac{1}{2}, \frac{1}{2}]$, $[\frac{1}{2}, 0]$ and $[0, \frac{1}{2}]$, leaving a union $\mathcal{T}_1$ of three open triangles of halved edge lengths, each being a rescaled copy of $\mathcal{T}_0$; thus $\mathcal{T}_1$ covers $\frac{3}{4}$ of the area of $\mathcal{T}_0$. For each of the smaller triangles obtained we iterate this process, yielding an infinite descending chain $\mathcal{T}_0 \supset \mathcal{T}_1 \supset \mathcal{T}_2 \supset \cdots$, where $\mathcal{T}_i$ consists of a total of 3^i triangles covering an area of $\frac{3^i}{2^{2i+1}}$, for $i \in \mathbb{N}_0$. The limit $\mathcal{T}_\infty := \bigcap_{i \geq 0} \mathcal{T}_i$ approached is called the **fractal Sierpinski Sieve**.

By construction, $\mathcal{T}_\infty$ is a measurable set of zero area, and we have $[y, x] \in \mathcal{T}_\infty$ whenever $[x, y] \in \mathcal{T}_\infty$. Moreover, letting $\mathcal{T}' \subseteq [0, 1]^2$ be the set of points $[x, y]$ such that either of $x, y, x + y$ equals $z \cdot 2^{-i}$, for some $i \in \mathbb{N}_0$ and $z \in \mathbb{N}_0$, we have $\mathcal{T}_\infty \cap \mathcal{T}' = \emptyset$. (Starting with a closed triangle and removing open triangles instead, yields a larger, but less interesting fractal set next to $\mathcal{T}_\infty$ encompassing the points $[x,y] \in \mathcal{T}'$ such that $x+y \leq 1$.) We derive a numerical description of the elements of $\mathcal{T}_\infty$, in particular showing that $\mathcal{T}_\infty$ is an infinite set:

Letting $\mathcal{T}_i' := \mathcal{T}_i \cap \mathcal{T}'$ for $i \in \mathbb{N}_0$, we consider the map $\tau \colon \mathcal{T}_0' \to [0,1]^2$ defined by

$$\tau \colon [x,y] \mapsto \begin{cases} [2x, 2y], & \text{if } x < \frac{1}{2} \text{ and } y < \frac{1}{2}, \\ [2x-1, 2y], & \text{if } x > \frac{1}{2}, \\ [2x, 2y-1], & \text{if } y > \frac{1}{2}. \end{cases}$$

Note that $\tau(\mathcal{T}_0') \cap \mathcal{T}' = \emptyset$; and recall that $[x,y] \in \mathcal{T}_0$ if and only if $x > 0$, $y > 0$ and $x + y < 1$. Then, for $[x,y] \in \mathcal{T}_0'$, we have $[x,y] \in \mathcal{T}_1'$ if and only if $\tau([x,y]) \in \mathcal{T}_0'$, and similarly $[x,y] \in \mathcal{T}_i'$ if and only if $\tau([x,y]) \in \mathcal{T}_{i-1}'$, for $i \in \mathbb{N}$, so that $[x,y] \in \mathcal{T}_i'$ if and only if $\tau^i([x,y]) \in \mathcal{T}_0'$. This implies that $[x,y] \in \mathcal{T}_\infty$ if and only if $\tau^i([x,y]) \in \mathcal{T}_0'$ for all $i \in \mathbb{N}_0$.

Now, for $[x,y] \in \mathcal{T}_0'$, the coordinates x and y have infinite 2-adic representation $x = \sum_{i>0} a_i \cdot 2^{-i}$ and $y = \sum_{i>0} b_i \cdot 2^{-i}$, respectively, where $a_i, b_i \in \{0,1\}$; note that the coordinates having finite 2-adic representations are captured in $\mathcal{T}'$. Then in terms of 2-adic representations τ is given by shifting each of the given representations by one step and ignoring the first binary digit, that is $[x', y'] := \tau([x,y])$ is given as $x' = \sum_{i>0} a_{i+1} \cdot 2^{-i}$ and $y' = \sum_{i>0} b_{i+1} \cdot 2^{-i}$.

Hence if $[x,y] \in \mathcal{T}_\infty'$ then, applying all powers of τ in turn, the condition $x+y < 1$ implies that the case $a_i = b_i = 1$ is excluded for all $i \in \mathbb{N}$, so that we have $a_i + b_i \leq 1$ for all $i \in \mathbb{N}$. Moreover, assume that $a_i + b_i = 1$ for almost all $i \in \mathbb{N}$, then after applying a suitable power of τ we may assume that $a_i + b_i = 1$ for all $i \in \mathbb{N}$, so that $x + y = 1$, a contradiction. Hence we conclude that we additionally have $a_i = b_i = 0$ for infinitely many $i \in \mathbb{N}$.

Conversely, let $[x,y] \in [0,1]^2 \setminus \mathcal{T}'$ be such that $a_i + b_i \leq 1$ for all $i \in \mathbb{N}$, and $a_i = b_i = 0$ for infinitely many $i \in \mathbb{N}$. Then, after applying any power of τ, there is $k \in \mathbb{N}$ such that $a_i + b_i = 1$ for $i \in \{1, \ldots, k-1\}$, and $a_k = b_k = 0$. This indeed implies that $x + y < 1$; note that in view of the conditions imposed it is not necessary to require that $x + y$ has an infinite 2-adic representation. $\sharp$

I Counting

1 Sets

(1.1) Sets. a) A **set** is a collection of well-defined distinct objects forming a new entity **[Cantor 1895]**: *Eine* **Menge** *ist eine gedankliche Zusammenfassung von bestimmten, wohlunterschiedenen Objekten der Anschauung oder des Denkens zu einem Ganzen.* Hence we stick to **naive set theory**, but Russell's antinomy below shows that this generality leads to a contradiction.

The objects collected are called the **elements** of the set. For any set M and any object x either $x \in M$ or $x \notin M$ holds. Moreover, any set is uniquely determined by its elements, hence contains a particular element only once, and we disregard the order of the elements. The **empty set** $\emptyset = \{\}$ is the set without elements.

For example, there are the **positive integers** $\mathbb{N} := \{1, 2, 3, \ldots\}$, the **non-negative integers** $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$, the **integers** $\mathbb{Z} := \{0, 1, -1, 2, -2, \ldots\}$, the **rational numbers** $\mathbb{Q}$, the **real numbers** $\mathbb{R}$, the **complex numbers** $\mathbb{C}$.

Sets can be given by explicit enumeration or by description. The latter means that from a given set M a new set N is formed by giving a logical formula saying which of the elements of M are elements of N and which are not; for example, we have $\{x \in \mathbb{Z}; x^2 = 1\} = \{1, -1\}$.

b) Russell's antinomy [1901]. By the generality of naive set theory, there is the set of all sets. Hence let $\mathcal{M} := \{M \text{ set}; M \notin M\}$ be the set of all sets, which do not contain themselves as one of their elements, thus either $\mathcal{M} \in \mathcal{M}$ or $\mathcal{M} \notin \mathcal{M}$. Assume we have $\mathcal{M} \in \mathcal{M}$, then $\mathcal{M}$ does not contain $\mathcal{M}$ as one of its elements, hence $\mathcal{M} \notin \mathcal{M}$, a contradiction. Assume we have $\mathcal{M} \notin \mathcal{M}$, then $\mathcal{M}$ does contain $\mathcal{M}$ as one of its elements, hence $\mathcal{M} \in \mathcal{M}$, again a contradiction.

Hence the set of all sets cannot possibly exist. Thus we indeed have to impose restrictions on which objects we may collect to form a set.

(1.2) Elementary constructions. a) Let M and N be sets. If for all $x \in M$ we have $x \in N$ then M is called a **subset** of N, and N is called a **superset** of M; we write $M \subseteq N$. If $M \subseteq N$ and $M \neq N$ then M is called a **proper** subset of N; we write $M \subset N$. We have $M = N$ if and only if $M \subseteq N$ and $N \subseteq M$, and we have $\emptyset \subseteq M$ and $M \subseteq M$. For example, we have $\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

b) The sets $M \cap N := \{x; x \in M \text{ and } x \in N\} = \{x \in M; x \in N\} = \{x \in N; x \in M\}$ and $M \setminus N := \{x \in M; x \notin N\}$ are called **intersection** and **difference** of M and N, respectively. If $M \cap N = \emptyset$ then M and N are called **disjoint**. If $\mathcal{I} \neq \emptyset$ is a set, and M_i is a set for $i \in \mathcal{I}$, then their intersection is defined as $\bigcap_{i \in \mathcal{I}} M_i := \{x; x \in M_i \text{ for all } i \in \mathcal{I}\}$.

Moreover, $M \cup N := \{x; x \in M \text{ or } x \in N\}$ is called the **union** of M and N; if additionally $M \cap N = \emptyset$ then the union of M and N is called **disjoint**, written as $M \,\dot{\cup}\, N$. If $\mathcal{I}$ is a set, and M_i is a set for $i \in \mathcal{I}$, then their union is defined as $\bigcup_{i \in \mathcal{I}} M_i := \{x; x \in M_i \text{ for some } i \in \mathcal{I}\}$.

c) Let $\mathcal{P}(M) := \{L; L \subseteq M\}$ be the **power set** of M; we have $\mathcal{P}(\emptyset) = \{\emptyset\}$.

d) For $x \in M$ and $y \in N$ let $[x, y] := \{x, \{x, y\}\}$ be the associated **ordered pair** with first and second **components** x and y, respectively; hence for $x \neq y$ we have $[x, y] \neq [y, x]$. Let $M \times N := \{[x, y]; x \in M, y \in N\}$ be the **Cartesian product** of M and N; hence $M \times N \neq \emptyset$ if and only if both $M \neq \emptyset$ and $N \neq \emptyset$.

(1.3) Maps. Let M and N be sets. A **relation** between M and N is a subset $R \subseteq M \times N$, and $x \in M$ and $y \in N$ such that $[x, y] \in R$ are said to be **related** by R; we also write xRy. A relation $f \subseteq M \times N$ such that for all $x \in M$ there is a unique $y \in N$ satisfying $[x, y] \in f$ is called a **map** from M to N; we write $f\colon M \to N\colon x \mapsto y = f(x)$, and also $[f(x) \in N; x \in M]$, where the element y is called the **image** of x, while x is called a **preimage** of y.

The sets M and N are called the **source** and the **domain** of f, respectively. The set $\mathrm{im}(f) := \{y \in N; y = f(x) \text{ for some } x \in M\}$ is called the **image** of f. For a subset $N' \subseteq N$, the set $f^{-1}(N') := \{x \in M; f(x) \in N'\}$ is called the **preimage** of N', with respect to f. For a subset $M' \subseteq M$, the **restriction** of f to M' is defined as $f|_{M'}\colon M' \to N\colon x \mapsto f(x)$.

The map $f\colon M \to N$ is called **surjective** if $\mathrm{im}(f) = N$, that is for all $y \in N$ there is some $x \in M$ such that $y = f(x)$. Moreover, f is called **injective** if for all $y \in N$ the preimage $f^{-1}(y)$ has at most one element, that is for all $y \in N$ there is at most one element $x \in M$ such that $y = f(x)$, or equivalently we have $f(x) \neq f(x') \in N$ whenever $x \neq x' \in M$. Finally, f is called **bijective** if it is both surjective and injective, that is $f^{-1}(y)$ is a singleton set for all $y \in N$, or equivalently for all $y \in N$ there is a unique $x \in M$ such that $y = f(x)$.

The map $\mathrm{id}_M\colon M \to M\colon x \mapsto x$ is called the **identity map**. The **composition** of maps $f\colon M \to N$ and $g\colon N \to U$, where U is a set, is defined as $gf = g \cdot f = g \circ f\colon M \to U\colon x \mapsto g(f(x))$. We have $f \cdot \mathrm{id}_M = f$ and $\mathrm{id}_N \cdot f = f$; moreover gf is surjective whenever f and g are surjective, and gf is injective whenever f and g are injective.

If $f\colon M \to N$ is bijective, the relation $f^{-1} := \{[y, x] \in N \times M; [x, y] \in f\}$ is a map as well, $f^{-1}\colon N \to M$ is called the **inverse map** of f. Hence we have $f(f^{-1}(y)) = y$ for all $y \in N$, thus $ff^{-1} = \mathrm{id}_N$, and $f^{-1}(f(x)) = x$ for all $x \in M$, thus $f^{-1}f = \mathrm{id}_M$. Moreover, f^{-1} is bijective such that $(f^{-1})^{-1} = f$.

Let $\mathrm{Maps}(M, N) := \{f \subseteq M \times N; f \text{ map}\}$. Moreover, let $\mathrm{Inj}(M, N) := \{f \in \mathrm{Maps}(M, N); f \text{ injective}\}$ and $\mathrm{Surj}(M, N) := \{f \in \mathrm{Maps}(M, N); f \text{ surjective}\}$, as well as $\mathrm{Bij}(M, N) := \mathrm{Inj}(M, N) \cap \mathrm{Surj}(M, N)$.

(1.4) Dedekind-Peano axioms. a) A set N fulfilling the following conditions is called a **set of positive integers**: There is an element $1 \in N$ and an injective **successor map** $N \to N \setminus \{1\}\colon n \mapsto n'$, such that the **principle of induction** holds: For any subset $M \subseteq N$ such that $1 \in M$, and such that for any $n \in M$ we also have $n' \in M$, we already have $M = N$.

The successor map is surjective as well, hence is bijective: Let $M := \{1\} \,\dot\cup\, \mathrm{im}(')$, then by induction we have $M = N$.

b) The set $\mathbb{N} := \{1, 2, \ldots\}$ of **positive integers** together with the successor map $\mathbb{N} \to \mathbb{N} \setminus \{1\}\colon n \mapsto n+1$ fulfills the above conditions; we take the existence of $\mathbb{N}$ and its arithmetic properties for granted. The set $\mathbb{N}$ is the unique model of a set of positive integers N, that is there is a unique map $f\colon \mathbb{N} \to N$ fulfilling $f(1) = 1$ and $f(n+1) = f(n)'$ for all $n \in \mathbb{N}$, and f is bijective:

Let $M \subseteq \mathbb{N}$ be the set of all $n \in \mathbb{N}$ such that there is a unique $f_n\colon \{1, \ldots, n\} \to N$ fulfilling $f_n(1) = 1$ and $f_n(m+1) = f_n(m)'$ for all $m < n$; hence we have

$1 \in M$, and for $1 \neq n \in M$ we have $f_n|_{\{1,\ldots,n-1\}} = f_{n-1}$. The surjectivity of the successor map of $\mathbb{N}$ implies that f_n is uniquely extendible to an admissible map $f_{n+1}\colon \{1,\ldots,n+1\} \to N$, thus by induction for $\mathbb{N}$ we have $M = \mathbb{N}$.

We have $1 \in \mathrm{im}(f)$, and for $f(n) \in \mathrm{im}(f)$ we have $f(n)' = f(n+1) \in \mathrm{im}(f)$, thus induction for N yields $\mathrm{im}(f) = N$, that is f is surjective. Let $M \subseteq \mathbb{N}$ be the set of all $n \in \mathbb{N}$ such that $f^{-1}(f(n)) = \{n\}$, then $f(\mathbb{N} \setminus \{1\}) \subseteq N \setminus \{1\}$ implies $f^{-1}(1) = \{1\}$, and we have $f^{-1}(f(n+1)) = f^{-1}(f(n)') = f^{-1}(f(n)) + 1 = \{n+1\}$ for $n \in \mathbb{N}$, thus induction for $\mathbb{N}$ yields $M = \mathbb{N}$, that is f is injective. ♯

(1.5) Cardinality. a) Let N be a set. If there is $n \in \mathbb{N}_0$ such that there is a bijection $f\colon \{1,\ldots,n\} \to N$, where we let $\{1,\ldots,0\} := \emptyset$, then N is called **finite** of **cardinality** $|N| := n$; otherwise N is called **infinite**, and we write $|N| = \infty$. For a finite set N the cardinality is uniquely determined:

We may assume that $N = \{1,\ldots,n\}$, for some $n \in \mathbb{N}$, and there is a bijection $f\colon \{1,\ldots,m\} \to \{1,\ldots,n\}$ for some $m \in \mathbb{N}$ such that $m \leq n$. We proceed by induction on $n \in \mathbb{N}$: If $n = 1$, then $m = 1$ as well. If $n \geq 2$, then letting $k := f(m)$ we get a bijection $f\colon \{1,\ldots,m-1\} \to \{1,\ldots,k-1,k+1,\ldots,n\}$. Using the bijection $\{1,\ldots,n-1\} \to \{1,\ldots,k-1,k+1,\ldots,n\}$ defined by $i \mapsto i$ for $i \leq k-1$, and $i \mapsto i+1$ for $i \geq k$, we may assume that $f\colon \{1,\ldots,m-1\} \to \{1,\ldots,n-1\}$ is a bijection as well, and by induction we have $n-1 = m-1$. ♯

If N is a finite set and $M \subseteq N$, then we may assume that $M = \{1,\ldots,m\} \subseteq \{1,\ldots,n\} = N$, where $n = |N|$, hence M is finite as well and we have $|M| \leq |N|$, where $|M| = |N|$ if and only if $M = N$.

Moreover, a map $f\colon N \to N$ is injective if and only if f is surjective: There is a subset $N' \subseteq N$ such that $f|_{N'}\colon N' \to \mathrm{im}(f)$ is a bijection; we have $N' = N$ if and only if f is injective, and from $|N'| = |\mathrm{im}(f)| \leq |N|$ we conclude that $|N'| = |N|$ if and only if f is surjective. In particular, since $\mathbb{N}$ possesses an injective but not surjective successor map, we conclude that $\mathbb{N}$ is infinite.

b) Any set N is infinite if and only if there is an injective map $\mathbb{N} \to N$: If there is such a map, then we may assume that $\mathbb{N} \subseteq N$, and since $\mathbb{N}$ is infinite we infer that N is infinite as well. Conversely, let now N be infinite. Then naively we might want to proceed as follows: We choose $x_1 \in N$, and then successively $x_{n+1} \in N \setminus \{x_1,\ldots,x_n\}$, since by assumption we have $N \setminus \{x_1,\ldots,x_n\} \neq \emptyset$ for all $n \in \mathbb{N}$, yielding the injective map $f\colon \mathbb{N} \to N\colon n \mapsto x_n$. But this is not justified, neither by the induction principle nor directly by the choice principle! Instead, we have to argue by **transfinite induction** as follows:

Assume to the contrary that there is no injective map $\mathbb{N} \to N$. We consider the set $\mathcal{F} := \bigcup_{n\in\mathbb{N}_0} \mathrm{Inj}(\{1,\ldots,n\},N)$; note that $\mathcal{F} \neq \emptyset$ since $\mathrm{Inj}(\emptyset,N) \neq \emptyset$. Denoting the upper bound of the source of $f \in \mathcal{F}$ by $n(f) \in \mathbb{N}_0$, the set $\mathcal{F}$ is partially ordered by letting $f \leq f'$ whenever $n(f) \leq n(f')$ and $f'|_{\{1,\ldots,n(f)\}} = f$. If $f_1 \leq f_2 \leq \cdots \leq f_k \leq \cdots$ is a chain in $\mathcal{F}$, then $f_\infty\colon M_\infty := \bigcup_{k\in\mathbb{N}}\{1,\ldots,n(f_k)\} \to N\colon i \mapsto f_k(i)$, whenever $i \leq n(f_k)$, is a well-defined map, and since for any $i,j \in M_\infty$ there is $k \in \mathbb{N}$ such that $i,j \leq n(f_k)$ we infer that f_∞ is injective. By induction we have $M_\infty = \mathbb{N}$ or $M_\infty = \{1, \ldots, n(f_\infty)\}$ is finite. By assumption, the former case cannot occur, thus $f_\infty \in \mathcal{F}$ is an **upper bound** of the given chain. Thus by **Zorn's Lemma**, see (5.2), there is a **maximal element** $f_0 \in \mathcal{F}$. Since $f_0 \colon \{1, \ldots, n(f_0)\} \to N$ is not surjective, choosing $x_0 \in N \setminus \operatorname{im}(f_0)$, we may extend f_0 to a map $\widetilde{f_0} \in \mathcal{F}$ by letting $\widetilde{f_0}(n(f_0)+1) := x_0$, contradicting the maximality of $f_0 \in \mathcal{F}$. ♯

c) Sets M and N are called **equicardinal** if there is a bijection $M \to N$. A set N is called **Dedekind infinite** [1888] if there is a proper subset $M \subset N$ such that M and N are equicardinal, that is there is an injective map $f \colon N \to N$ which is not surjective. Hence any Dedekind infinite set is infinite; conversely, any infinite set N is Dedekind infinite: We may assume $\mathbb{N} \subseteq N$, then $f \colon N \to N$ defined by $f(x) = x$ for $x \in N \setminus \mathbb{N}$, and $f(n) := n+1$ for $n \in \mathbb{N} \subseteq N$, is injective but not surjective.

(1.6) Basic counting principles. Let M and N be finite sets. Then we have the **sum principle**, saying that if $M \cap N = \emptyset$ then $|M \,\dot\cup\, N| = |M| + |N|$: Let $f \colon \{1, \ldots, m\} \to M = \{x_1, \ldots, x_m\} \colon i \mapsto x_i$ and $g \colon \{1, \ldots, n\} \to N = \{y_1, \ldots, y_n\} \colon j \mapsto y_j$ be bijective, where $m := |M| \in \mathbb{N}_0$ and $n := |N| \in \mathbb{N}_0$. Then the map $h \colon \{1, \ldots, m+n\} \to M \,\dot\cup\, N$ defined by $h(k) = x_k$ for $k \leq m$, and $h(k) = y_{k-m}$ for $k \geq m+1$, is a bijection.

Moreover, we have the **product principle** $|M \times N| = |M| \cdot |N|$: The map $\{1, \ldots, mn\} \to M \times N \colon (i-1)n + j \mapsto [x_i, y_j]$ is a bijection; in other words we have m and n **possibilities** for the first and second components, respectively.

We have $|\operatorname{Maps}(M, N)| = |N|^{|M|}$: For $m \in \mathbb{N}$ let $N^m := N \times \cdots \times N$ be the **m-fold Cartesian power** of N, where N^1 can be identified with N; we let $N^0 := \{[]\}$, a singleton set. Then $\operatorname{Maps}(M, N) \to N^m \colon \alpha \mapsto [\alpha(x_1), \ldots, \alpha(x_m)]$ is a bijection, hence by induction on m we get $|\operatorname{Maps}(M, N)| = n^m$.

In particular, we have $|\mathcal{P}(N)| = 2^{|N|}$: The map $\operatorname{Maps}(N, \{0,1\}) \to \mathcal{P}(N) \colon f \mapsto f^{-1}(1)$, is a bijection, implying $|\mathcal{P}(N)| = |\operatorname{Maps}(N, \{0,1\})| = 2^{|N|}$; the elements of $\operatorname{Maps}(N, \{0,1\})$ are called the **indicator maps** of N.

2 Selections

(2.1) Selections without repetitions. Let N be a finite set of cardinality $n := |N| \in \mathbb{N}_0$, hence we may assume that $N = \{1, \ldots, n\}$, and let $k \in \mathbb{N}_0$. We discuss various ways of selecting from N, see Table 5:

a) A **k-arrangement** or **k-permutation** of N is an **ordered selection without repetitions**, that is an injective map $f \colon K := \{1, \ldots, k\} \to N$. Let $\mathcal{S}_k(n) := \operatorname{Inj}(K, N)$; hence we have $\mathcal{S}_k(n) = \emptyset$ for $k > n$. For $k = n$ we get $\mathcal{S}_n := \mathcal{S}_n(n) = \operatorname{Bij}(N, N)$, being called the **permutations** of N.

We show that for $k \leq n$ we have $|\mathcal{S}_k(n)| = n_{(k)} := \prod_{i=0}^{k-1}(n-i) \in \mathbb{N}$, called a **falling factorial**, where we define the empty product as $n_{(0)} := 1$: We proceed by induction on $k \in \mathbb{N}_0$, the case $k = 0$ being trivial. For $k \geq 1$ we have n possibilities to choose $f(k) \in N$, for any of these there are $|\mathrm{Inj}(\{1, \ldots, k-1\}, N \setminus \{f(k)\})| = |\mathcal{S}_{k-1}(n-1)|$ possibilities for f left, hence we get $|\mathcal{S}_k(n)| = n \cdot |\mathcal{S}_{k-1}(n-1)| = n \cdot \prod_{i=0}^{k-2}(n-1-i) = n \cdot \prod_{i=1}^{k-1}(n-i) = \prod_{i=0}^{k-1}(n-i)$.

Since $n_{(k)} = 0$ for $k > n$, we have $|\mathcal{S}_k(n)| = n_{(k)} \in \mathbb{N}_0$ for all $k \in \mathbb{N}_0$. Moreover, we have $|\mathcal{S}_n| = n! := n_{(n)} \in \mathbb{N}$, called n**-factorial**; in particular we have $0! := 1$. Using this, for $k \leq n$ we get $|\mathcal{S}_k(n)| = n_{(k)} = \frac{n!}{(n-k)!} \in \mathbb{N}$.

For example, if any of $k = 4$ men marries one of $n = 6$ women, then this can be done in $\frac{6!}{2!} = 6 \cdot 5 \cdot 4 \cdot 3 = 360$ ways.

b) A k**-subset** or k**-combination** of N is an **unordered selection without repetitions**, that is a subset $M \subseteq N$ such that $|M| = k$, which hence can be identified with the associated indicator map $N \to \{0, 1\}$. Let $\mathcal{P}_k(n) := \{M \subseteq N; |M| = k\}$; hence we have $\mathcal{P}_k(n) = \emptyset$ for $k > n$.

We determine $|\mathcal{P}_k(n)| \in \mathbb{N}$ for $k \leq n$: Any k-subset $M \subseteq N$ gives rise to $|\mathcal{S}_k|$ arrangements, that is injective maps $f\colon \{1, \ldots, k\} \to N$ such that $\mathrm{im}(f) = M$, hence we have $|\mathcal{S}_k(n)| = |\mathcal{S}_k| \cdot |\mathcal{P}_k(n)|$, yielding $|\mathcal{P}_k(n)| = \frac{n_{(k)}}{k!} =: \binom{n}{k} \in \mathbb{N}$, being called a **binomial coefficient**; note that it is not obvious at all that this is an integer, here it follows from its interpretation as a cardinality. Since $\frac{n_{(k)}}{k!} = 0$ for $k > n$, we have $|\mathcal{P}_k(n)| = \binom{n}{k} \in \mathbb{N}_0$ for all $k \in \mathbb{N}_0$. Moreover, for $k \leq n$ we have $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k} \in \mathbb{N}$.

For example, in the lottery '6-from-49', there are $k = 6$ balls drawn from an urn containing $n = 49$ distinguishable balls, hence there are $\binom{49}{6} = \frac{49}{6 \cdot 43} = 13\,983\,816 \sim 1.4 \cdot 10^7$ distinct draws.

(2.2) Selections with repetitions. We discuss further ways of selecting from $N = \{1, \ldots, n\}$, where $n \in \mathbb{N}_0$, and let $k \in \mathbb{N}_0$, see Table 5:

a) A k**-tuple** of N is an **ordered selection with repetitions**, that is a map $K := \{1, \ldots, k\} \to N$, which hence can be identified with an element of the k-fold Cartesian power N^k of N; we have $|\mathrm{Maps}(K, N)| = |N|^{|K|} = n^k \in \mathbb{N}$.

For example, in the football pool '13er-Wette', the outcome of $k = 13$ soccer matches is encoded into a map $\{1, \ldots, 13\} \to \{0, 1, 2\}$, by identifying 'home team wins' with 1, 'guest team wins' with 2, and 'draw' with 0; hence we have $n = 3$ and there are $3^{13} = 1\,594\,323 \sim 1.6 \cdot 10^6$ distinct outcomes.

b) A k**-multiset** on N is an **unordered selection with repetitions**, that is a map $\mu\colon N \to \mathbb{N}_0\colon i \mapsto \mu_i$, or equivalently an n-tuple $\mu = [\mu_1, \ldots, \mu_n] \in \mathbb{N}_0^n$, such that $\sum_{i=1}^n \mu_i = k$, where $\mu_i \in \mathbb{N}_0$ is called the associated **multiplicity**, and $|\mu| := k$ is called the **cardinality** of μ; we also write the multiset associated with μ as $1^{\mu_1} \cdots n^{\mu_n}$. In particular, the **multiplicity-free** multisets, that is we have $\mu_i \leq 1$ for all $i \in N$, can be identified with the k-subsets of N. Let $\mathcal{M}_k(n) := \{[\mu_1, \ldots, \mu_n] \in \mathbb{N}_0^n; \sum_{i=1}^n \mu_i = k\}$; hence we have $\mathcal{M}_0(0) = \{[]\}$ and $\mathcal{M}_k(0) = \emptyset$ for $k \geq 1$.

Table 5: Selections of an n-set.

k-selections	without repetitions	with repetitions	of k-multisets
ordered	$n_{(k)}$ arrangements	n^k tuples	$\binom{k}{\mu_1,\ldots,\mu_n}$
unordered	$\binom{n}{k}$ subsets	$\binom{n+k-1}{k}$ multisets	1

We determine $|\mathcal{M}_k(n)| \in \mathbb{N}$ for $n \geq 1$: Given $\mu = [\mu_1, \ldots, \mu_n] \in \mathcal{M}_k(n)$, writing the associated multiset as $[\bullet\cdots\bullet \mid \bullet\cdots\bullet \mid \ldots \mid \bullet\cdots\bullet]$, with μ_j entries '$\bullet$' in the j-th slot, for $j \in \{1, \ldots, n\}$, and counting the entries '$\mid$', we get a tuple of length $(n-1)+\sum_{j=1}^n \mu_j = n+k-1$, uniquely determining μ by the $(n-1)$-subset of the positions of the '$\mid$'.

Formally, letting $\sigma(\mu) := \{\sum_{j=1}^i (\mu_j + 1) \in \mathbb{N}; i \in \{1, \ldots, n-1\}\}$, an $(n-1)$-subset of $\{1, \ldots, n+k-1\}$, yields an injective map $\sigma \colon \mathcal{M}_k(n) \to \mathcal{P}_{n-1}(n+k-1)$. Conversely, given a subset $\{s_1, \ldots, s_{n-1}\} \subseteq \{1, \ldots, n+k-1\}$, letting $\mu_1 := s_1 - 1 \in \mathbb{N}_0$ as well as $\mu_i := s_i - s_{i-1} - 1 \in \mathbb{N}_0$ for $i \in \{2, \ldots, n-1\}$, and $\mu_n := k - \sum_{i=1}^{n-1} \mu_i = k - s_{n-1} + (n-1) \geq (n+k-1) - (n+k-1) = 0$, we have $\mu := [\mu_1, \ldots, \mu_n] \in \mathcal{M}_k(n)$, and from $\sum_{j=1}^i (\mu_j + 1) = s_i$, for $i \in \{1, \ldots, n-1\}$, we get $\sigma(\mu) = \{s_1, \ldots, s_{n-1}\}$. Hence σ is surjective, thus is a bijection.

Either picture yields $|\mathcal{M}_k(n)| = |\mathcal{P}_{n-1}(n+k-1)| = \binom{n+k-1}{n-1} = \binom{n+k-1}{k}$. ♯

Note that the former formula is not defined for $n = 0$, but the latter does hold for $n = 0$ as well. For example, $\mathsf{I}^4\mathsf{MP}^2\mathsf{S}^4$ is a multiset on the Latin alphabet $\{\mathsf{A}, \ldots, \mathsf{Z}\}$, hence we have $n = 26$, and $[4_\mathsf{I}, 1_\mathsf{M}, 2_\mathsf{P}, 4_\mathsf{S}]$ has cardinality $k = 4 + 1 + 2 + 4 = 11$; the number of 11-multisets on the Latin alphabet is given as $\binom{36}{11} = 600\,805\,296 \sim 6 \cdot 10^8$.

(2.3) Selections from multisets. We generalize the notion of selection without repetition to multisets. To this end, let $N := \{1, \ldots, n\}$, where $n \in \mathbb{N}_0$, let $\mu = [\mu_1, \ldots, \mu_n] \in \mathcal{M}_k(n)$, where $k \in \mathbb{N}_0$, and let $l \in \mathbb{N}_0$, see Table 5:

a) An **l-arrangement** or **l-permutation** of μ is an ordered selection of μ, that is a map $f \colon L := \{1, \ldots, l\} \to N$ such that $|f^{-1}(i)| \leq \mu_i$, for all $i \in N$; in particular, if μ is multiplicity-free describing the subset $M := \{i \in N; \mu_i = 1\} \subseteq N$, we recover $\operatorname{Inj}(L, M)$. Hence we have $l \leq k$, and for $k = l$ we have $|f^{-1}(i)| = \mu_i$, for all $i \in N$; in the latter case f is called a **permutation** of μ.

We show that the number of permutations of μ is given by the **multinomial coefficient** $\binom{k}{\mu_1,\ldots,\mu_n} := \frac{k!}{\prod_{i=1}^n \mu_i!} \in \mathbb{N}$: We proceed by induction on $n \in \mathbb{N}_0$. For $n = 0$ we have $k = 0$ as well, thus there is a unique map $f \colon \emptyset \to \emptyset$, where $\frac{0!}{1} = 1$. For $n \geq 1$ we have $\binom{k}{\mu_n}$ possibilities to choose $f^{-1}(n) \subseteq K := \{1, \ldots, k\}$, for any of these there are $\binom{k-\mu_n}{\mu_1,\ldots,\mu_{n-1}}$ possibilities for $f \colon K \setminus f^{-1}(n) \to N \setminus \{n\}$ left, hence there are $\frac{k!}{\mu_n!(k-\mu_n)!} \cdot \frac{(k-\mu_n)!}{\prod_{i=1}^{n-1} \mu_i!} = \frac{k!}{\prod_{i=1}^n \mu_i!}$ possibilities for f.

b) An **l-submultiset** of μ is an unordered selection of μ, that is a map $\lambda = [\lambda_1, \ldots, \lambda_n] \in \mathcal{M}_l(n)$ such that $\lambda_i \leq \mu_i$, for all $i \in N$. Hence we have $l \leq k$, and in the case $k = l$ we have $\lambda = \mu$. Moreover, summing over $l \in \{0, \ldots, k\}$ there are $\prod_{i=1}^n (\mu_i + 1) \in \mathbb{N}$ submultisets of μ.

For example, the multiset $\mathsf{I}^4\mathsf{MP}^2\mathsf{S}^4$ has $\binom{11}{4,1,2,4} = \frac{11!}{4!\cdot 1!\cdot 2!\cdot 4!} = 34\,650$ permutations, for example $\mathsf{IIIIMPPSSSS}$ and $\mathsf{MISSISSIPPI}$; moreover, IMS^2 is the 4-submultiset $[1_{\mathsf{I}}, 1_{\mathsf{M}}, 2_{\mathsf{S}}]$ of $[4_{\mathsf{I}}, 1_{\mathsf{M}}, 2_{\mathsf{P}}, 4_{\mathsf{S}}]$, allowing for the permutation MISS.

(2.4) Binomial coefficients. a) We generalize the notion of falling factorials and binomial coefficients. To this end, let $k \in \mathbb{N}_0$ and let X be an indeterminate. Then the polynomials $X_{(k)} := \prod_{i=0}^{k-1} (X - i) \in \mathbb{Z}[X]$ and $X^{(k)} := \prod_{i=0}^{k-1} (X + i) \in \mathbb{Z}[X]$ are called **falling** and **rising factorials**, respectively, where again $X_{(0)} = X^{(0)} := 1 \in \mathbb{Z}[X]$. Hence both $X_{(k)}$ and $X^{(k)}$ are monic of degree k, and we have $X^{(k)} = (X + k - 1)_{(k)}$ as well as the **reciprocity** $(-X)_{(k)} = \prod_{i=0}^{k-1} (-X - i) = (-1)^k \cdot \prod_{i=0}^{k-1} (X + i) = (-1)^k X^{(k)}$. Moreover, for any $z \in \mathbb{C}$ by evaluating we get the complex numbers $z_{(k)} \in \mathbb{C}$ and $z^{(k)} \in \mathbb{C}$.

Still having $k! = k_{(k)} \in \mathbb{N}$, we let $\binom{X}{k} := \frac{X_{(k)}}{k!} \in \mathbb{Q}[X]$, being called the associated **binomial coefficient**, thus $\binom{X}{k}$ has leading coefficient $\frac{1}{k!}$ and degree k, and we have the **negation** $\binom{-X}{k} = \frac{(-X)_{(k)}}{k!} = (-1)^k \cdot \frac{(X+k-1)_{(k)}}{k!} = (-1)^k \cdot \binom{X+k-1}{k}$. For $k \neq 0$ we let $\binom{X}{-k} := 0 \in \mathbb{Z}[X]$; note that $(-k)!$ is not defined.

For any $z \in \mathbb{C}$ we get $\binom{z}{k} = \frac{z_{(k)}}{k!} \in \mathbb{C}$, while for $k \neq 0$ we have $\binom{z}{-k} = 0 \in \mathbb{C}$. In particular we have the **combinatorial reciprocity** $|\mathcal{M}_k(n)| = \binom{n+k-1}{n-1} = \binom{n+k-1}{k} = (-1)^k \cdot \binom{-n}{k}$, for all $n \in \mathbb{N}_0$, relating the number of k-multisets of $N := \{1, \ldots, n\}$ to the number of k-subsets of N; this is elucidated in (12.3).

b) To explain the name 'binomial coefficient', let X and Y be indeterminates. Then $(X + Y)^n = \sum_{\mu \in \mathcal{M}_n(2)} \binom{n}{\mu_1, \mu_2} X^{\mu_1} Y^{\mu_2} = \sum_{k=0}^n \binom{n}{k} X^k Y^{n-k} \in \mathbb{Z}[X, Y]$ holds in the bivariate polynomial ring $\mathbb{Z}[X, Y]$, where $\binom{n}{\mu_1, \mu_2}$ is the number of permutations of the multiset $\mu \in \mathcal{M}_n(2)$. Hence $\binom{n}{k}$ appears in the expansion of the n-th power, where $n \in \mathbb{N}_0$, of the **binomial** $X + Y \in \mathbb{Z}[X, Y]$.

Hence evaluating at $y := 1$ yields $(X + 1)^n = \sum_{k=0}^n \binom{n}{k} X^k \in \mathbb{Z}[X]$, thus evaluating further at $x := 1$ yields $\sum_{k=0}^n \binom{n}{k} = 2^n \in \mathbb{C}$, while evaluating further at $x := -1$ for $n \neq 0$ yields $\sum_{k=0}^n (-1)^k \binom{n}{k} = 0 \in \mathbb{C}$.

To explain the name 'multinomial coefficient', let $X_1, \ldots, X_n$ be indeterminates, using the **multinomial** $\sum_{i=1}^n X_i \in \mathbb{Z}[X_1, \ldots, X_n]$ in the multivariate polynomial ring $\mathbb{Z}[X_1, \ldots, X_n]$, we get $(\sum_{i=1}^n X_i)^k = \sum_{\mu \in \mathcal{M}_k(n)} (\binom{k}{\mu_1, \ldots, \mu_n} \cdot \prod_{i=1}^n X_i^{\mu_i})$.

(2.5) Identities for binomial coefficients. a) We show that for all $k \in \mathbb{Z}$ we have the **triangle identity** $\binom{X}{k} = \binom{X-1}{k-1} + \binom{X-1}{k} \in \mathbb{Q}[X]$: If $k < 0$ then both sides vanish, if $k = 0$ then both sides are equal to the constant polynomial 1. Hence we may assume that $k \geq 1$. Then both sides have degree k, hence their difference either vanishes or has at most k complex zeroes. Thus it suffices to show that the above identity holds for all $x := n \in \mathbb{N}$:

Let $N := \{1, \ldots, n\}$. Then we have $\mathcal{P}_k(n) = \{M \subseteq N; |M| = k\} = \{M \subseteq N; |M| = k, n \in M\} \,\dot\cup\, \{M \subseteq N; |M| = k, n \not\in M\}$, where $\{M \subseteq N; |M| = k, n \in M\} \to \mathcal{P}_{k-1}(n-1)\colon M \mapsto (M \setminus \{n\})$ and $\{M \subseteq N; |M| = k, n \not\in M\} \to \mathcal{P}_k(n-1)\colon M \mapsto M$ are bijections, thus we get $\binom{n}{k} = |\mathcal{P}_k(n)| = |\mathcal{P}_{k-1}(n-1)| + |\mathcal{P}_k(n-1)| = \binom{n-1}{k-1} + \binom{n}{k-1}$. ♯

b) We show that for all $k \in \mathbb{Z}$ we have the **Vandermonde identity** $\binom{X+Y}{k} = \sum_{i=0}^{k} \binom{X}{i}\binom{Y}{k-i} \in \mathbb{Q}[X, Y]$: If $k < 0$ then the left hand side vanishes, while the right hand side is the empty sum. Hence we may assume that $k \geq 0$, then both sides have total degree k. If their difference does not vanish, then viewing it as an element of $\mathbb{Q}[X][Y]$, there are at most $k(k+1)$ complex roots of any of its at most $k+1$ non-vanishing coefficient polynomials in $\mathbb{Q}[X]$, which each have degree at most k, and evaluating at a complex number not in this set there are at most k complex roots of the resulting non-vanishing polynomial in $\mathbb{Q}[Y]$. Thus it suffices to show that the above identity holds for all $[x, y] := [m, n] \in \mathbb{N}_0^2$:

Let M and N be disjoint sets such that $|M| = m$ and $|N| = n$, then we have $\binom{m+n}{k} = |\{L \subseteq M \,\dot\cup\, N; |L| = k\}| = \coprod_{i=0}^{k} \{L \subseteq M \,\dot\cup\, N; |L| = k, |L \cap M| = i\}$, where the condition $|L \cap M| = i$ is equivalent to saying $|L \cap N| = k - i$. Thus we get $|\{L \subseteq M \,\dot\cup\, N; |L| = k, |L \cap M| = i\}| = \binom{m}{i}\binom{n}{k-i}$, implying $|\{L \subseteq M \,\dot\cup\, N; |L| = k\}| = \sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i}$. ♯

(2.6) The Pascal triangle. Letting $k \in \mathbb{N}_0$, evaluating at $n \in \mathbb{N}_0$ the recursion $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ gives rise to the **Pascal triangle** shown in Table 3, allowing to compute binomial coefficients by using additions alone, but no multiplications. The Pascal triangle has a rich structure, where we point out only a few of its properties; note that these typically do not generalize to the polynomial case $\binom{X}{k}$:

i) We have **palindromicity** $\binom{n}{k} = \binom{n}{n-k}$: Since $\binom{n}{k} = 0$ for $k > n$, and $\binom{n}{-k} = 0$ for $k \neq 0$, we may assume $k \leq n$, where we observe $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$.

Alternatively, for $k \leq n$ there is the bijection $\mathcal{P}_k(n) \to \mathcal{P}_{n-k}(n)\colon M \mapsto (N \setminus M)$.

ii) From $\mathcal{P}(n) = \coprod_{k=0}^{n} \mathcal{P}_k(n)$ we recover the **row sum** formula $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

iii) For $n \geq 1$, the **alternating row sum** formula $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$ is equivalent to $|\mathcal{P}_{\text{even}}(n)| = \sum_{0 \leq k \leq n \text{ even}} \binom{n}{k} = \sum_{0 \leq k \leq n \text{ odd}} \binom{n}{k} = |\mathcal{P}_{\text{odd}}(n)|$, where $\mathcal{P}_{\text{even}}(n) := \coprod_{0 \leq k \leq n \text{ even}} \mathcal{P}_k(n)$ and $\mathcal{P}_{\text{odd}}(n) := \coprod_{0 \leq k \leq n \text{ odd}} \mathcal{P}_k(n)$. This in turn in view of $|\mathcal{P}(n)| = 2^n$ is equivalent to $|\mathcal{P}_{\text{even}}(n)| = 2^{n-1} = |\mathcal{P}_{\text{odd}}(n)|$.

Alternatively, the latter can be shown by induction on $n \in \mathbb{N}$ as follows: For $n = 1$ we have $\mathcal{P}_{\text{even}}(1) = \{\emptyset\}$ and $\mathcal{P}_{\text{odd}}(1) = \{\{1\}\}$, and for $n \geq 2$ letting $N := \{1, \ldots, n\}$ we have $\mathcal{P}_{\text{even}}(n) = \{M \subseteq N; |M| \text{ even}, n \in M\} \,\dot\cup\, \{M \subseteq N; |M| \text{ even}, n \not\in M\}$, where $\{M \subseteq N; |M| \text{ even}, n \in M\} \to \mathcal{P}_{\text{odd}}(n-1)\colon M \mapsto$

$(M \setminus \{n\})$ and $\{M \subseteq N; |M| \text{ even}, n \not\in M\} \to \mathcal{P}_{\text{even}}(n-1)\colon M \mapsto M$ are bijections. Hence we have $|\mathcal{P}_{\text{even}}(n)| = |\mathcal{P}_{\text{odd}}(n-1)| + |\mathcal{P}_{\text{even}}(n-1)| = 2^n$.

iv) We have the **partial column sums** $\sum_{i=0}^{n} \binom{i}{k} = \binom{n+1}{k+1}$: We proceed by induction on $n \in \mathbb{N}_0$, where for $n = 0$ we have $\binom{0}{k} = \binom{1}{k+1}$, and for $n \geq 1$ we have $\sum_{i=0}^{n} \binom{i}{k} = \binom{n}{k} + \sum_{i=0}^{n-1} \binom{i}{k} = \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.

Alternatively, we have $\mathcal{P}_{k+1}(n+1) = \coprod_{i=1}^{n+1} \{M \subseteq \{1, \ldots, n+1\}; |M| = k+1, \max(M) = i\}$, where $\{M \subseteq \{1, \ldots, n+1\}; |M| = k+1, \max(M) = i\} \to \mathcal{P}_k(i-1)\colon M \mapsto (M \setminus \{i\})$ is a bijection, hence $\binom{n+1}{k+1} = \sum_{i=1}^{n+1} \binom{i-1}{k} = \sum_{i=0}^{n} \binom{i}{k}$.

v) For **partial diagonal sums** we have $\sum_{i=0}^{k} \binom{X+i}{i} = \binom{X+k+1}{k} \in \mathbb{Q}[X]$: We proceed by induction on $k \in \mathbb{N}_0$, where for $k = 0$ we have $\binom{X+0}{0} = 1 = \binom{X+1}{0} \in \mathbb{Z}[X]$, while for $k \geq 1$ we get $\sum_{i=0}^{k} \binom{X+i}{i} = \binom{X+k}{k} + (\sum_{i=0}^{k-1} \binom{X+i}{i}) = \binom{X+k}{k} + \binom{X+k}{k-1} = \binom{X+k+1}{k} \in \mathbb{Q}[X]$.

For **partial alternating row sums** using negation this yields $\sum_{i=0}^{k} (-1)^i \binom{X}{i} = \sum_{i=0}^{k} \binom{-X+i-1}{i} = \binom{(-X-1)+k+1}{k} = \binom{-(X-1)+k-1}{k} = (-1)^k \binom{X-1}{k} \in \mathbb{Q}[X]$. Evaluating at $x := n$ we get $\sum_{i=0}^{k} (-1)^{k-i} \binom{n}{i} = \binom{n-1}{k}$; for $n = k$ we recover $\sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} = \binom{n-1}{n} = \delta_{0,n}$, using the **Kronecker symbol** $\delta \in \{0, 1\}$.

(2.7) Congruences for binomial coefficients. a) Let $k, n \in \mathbb{N}_0$, and let $p \in \mathbb{Z}$ be a prime. We consider the question of when $p \mid \binom{n}{k}$: To this end, let $n = \sum_{i \geq 0} a_i p^i$ and $k = \sum_{i \geq 0} b_i p^i$ be the p-adic representations of n and k, respectively, where $a_i, b_i \in \{0, \ldots, p-1\}$. Then we have the **Lucas congruence** [1878] saying that $\binom{n}{k} \equiv \prod_{i \geq 0} \binom{a_i}{b_i} \pmod{p}$:

Since for $i \in \{0, \ldots, p\}$ we have $p \not| \binom{p}{i} = \frac{p!}{i!(p-i)!}$ if and only if $i \in \{0, p\}$, we have $(X+Y)^p = \sum_{i=0}^{p} \binom{p}{i} X^i Y^{p-i} = X^p + Y^p \in \mathbb{F}_p[X, Y]$, where $\mathbb{F}_p \cong \mathbb{Z}/p\mathbb{Z}$ is the finite field of order p, and where we identify an integer in $\mathbb{Z}$ with its natural image in $\mathbb{F}_p$. This yields $\sum_{l=0}^{n} \binom{n}{l} X^l = (X+1)^n = \prod_{i \geq 0} (X+1)^{a_i p^i} = \prod_{i \geq 0} (X^{p^i} + 1)^{a_i} = \prod_{i \geq 0} (\sum_{j=0}^{p-1} \binom{a_i}{j} X^{j p^i}) \in \mathbb{F}_p[X]$. The coefficient of $X^k = \prod_{i \geq 0} X^{b_i p^i}$ in the latter polynomial equals $\prod_{i \geq 0} \binom{a_i}{b_i}$. $\sharp$

Hence we have $p \not| \binom{n}{k}$ if and only if $p \not| \binom{a_i}{b_i}$ for all $i \geq 0$. Since $a_i \leq p-1$ we have $p \not| \binom{a_i}{b_i} = \frac{(a_i)_{(b_i)}}{b_i!}$ if and only if $b_i \leq a_i$. This implies that $p \not| \binom{n}{k}$ if and only if $b_i \leq a_i$ for all $i \geq 0$, that is the p-adic expansion of k is **contained** in the p-adic expansion of n.

In particular, for $p = 2$, letting again $w(n) := \sum_{i \geq 0} a_i \in \mathbb{N}_0$ be the number of binary digits 1 in the 2-adic representation of n, then there are $2^{w(n)}$ possible 2-adic representations contained in the 2-adic representation of n. Thus we recover the number of odd entries in row n of the Pascal triangle as given in (0.3). Moreover, we have obtained a number theoretic description of where the odd entries are actually located.

b) It follows from the Lucas congruence, that $\binom{pn}{pk} \equiv \binom{n}{k} \pmod{p}$. We show combinatorially that actually $\binom{pn}{pk} \equiv \binom{n}{k} \pmod{p^2}$:

To this end, we consider a $(p \times n)$-square grid, in which pk squares out of the pn ones are chosen; this yields $\binom{pn}{pk}$ possibilities. We let the group C_p^n act on these grids by cyclically permuting the rows independently, and we consider the associated orbits. Considering a fixed row of a choice, since p is a prime, it is either fixed or yields an orbit of length p, where the former case occurs if and only if it is fully chosen or consists of unchosen squares only. Thus the orbit length of a choice is given as p^i, where i is the number of mixed rows in the above sense. Now, there are $\binom{n}{k}$ choices entirely consisting of fully chosen rows, while for the other ones there are at least 2 mixed rows. Hence there are $\binom{pn}{pk} - \binom{n}{k}$ of the latter, all of which have orbit length divisible by p^2. ♯

In the same vein, we show algebraically that $\binom{pn}{pk} \equiv \binom{n}{k} \pmod{p^3}$ for $p \geq 5$:

We proceed keeping the above picture. If a choice contains at least 3 mixed rows, then its orbit has length divisible by p^3. Hence we only have to consider choices with precisely 2 mixed rows. Since in this case we have $k-1$ fully chosen rows, keeping the mixed rows fixed, this amounts to $\binom{2p}{p} - 2$ possibilities (all choices excluding the cases where either of the rows considered is fully chosen), so that we have to show that $\binom{2p}{p} \equiv 2 \pmod{p^3}$; note that this is just the case $n = 2$ and $k = 1$ of the claim:

The Vandermode identity yields $\binom{2p}{p} - 2 = -2 + \sum_{i=0}^{p} \binom{p}{i}\binom{p}{p-i} = \sum_{i=1}^{p-1} \binom{p}{i}^2 = p^2 \cdot \sum_{i=1}^{p-1} (\frac{(p-1)_{(i-1)}}{i!})^2$. Hence we show that $\sum_{i=1}^{p-1} (\frac{(p-1)_{(i-1)}}{i!})^2 \equiv 0 \pmod{p}$:

We have $\sum_{i=1}^{p-1} (\frac{(p-1)_{(i-1)}}{i!})^2 \equiv \sum_{i=1}^{p-1} (\frac{(i-1)!}{i!})^2 \equiv \sum_{i=1}^{p-1} (\frac{1}{i})^2 \equiv \sum_{i=1}^{p-1} i^2 \pmod{p}$; recall that $\mathbb{Z}_p^* = \{1, \ldots, p-1\}$, so that this is well-defined indeed. Finally, we use the well-known identity $\sum_{i=1}^{n} i^2 = \frac{1}{6} n(n+1)(2n+1)$, which is immediately proved by induction; see also (4.3). From this, for $p \geq 5$ we indeed get $\sum_{i=1}^{p-1} i^2 \equiv \frac{1}{6} p(p-1)(2p-1) \equiv 0 \pmod{p}$. ♯

The statement does not hold for $p \leq 3$: For $p = 2$ we get $\binom{4}{2} \equiv 6 \not\equiv 2 \pmod{8}$ and for $p = 3$ we get $\binom{6}{3} \equiv 20 \not\equiv 2 \pmod{27}$.

3 Partitions and permutations

(3.1) Partitions of sets. a) We discuss ways of forming partitions of $N := \{1, \ldots, n\}$, where $n \in \mathbb{N}_0$; hence let $k \in \mathbb{N}_0$: A k**-composition** of N is an ordered decomposition $[N_1, \ldots, N_k]$ of $N = \coprod_{i=1}^{k} N_i$ into k pairwise disjoint **blocks** $N_i \neq \emptyset$. A k**-partition** of N is an unordered decomposition of $N = \coprod_{i=1}^{k} N_i$ into k pairwise disjoint blocks $N_i \neq \emptyset$. For example, for $n = 3$ the 2-partitions of N are given as $N = \{1,2\} \,\dot\cup\, \{3\} = \{1,3\} \,\dot\cup\, \{2\} = \{2,3\} \,\dot\cup\, \{1\}$, hence the 2-compositions of N are given as

$$[\{1,2\},\{3\}], [\{3\},\{1,2\}], [\{1,3\},\{2\}], [\{2\},\{1,3\}], [\{2,3\},\{1\}], [\{1\},\{2,3\}].$$

Table 6: The Stirling triangle of the second kind.

$n\backslash k$	0	1	2	3	4	5	6	7
0	1							
1	0	1						
2	0	1	1					
3	0	1	3	1				
4	0	1	7	6	1			
5	0	1	15	25	10	1		
6	0	1	31	90	65	15	1	
7	0	1	63	301	350	140	21	1

b) The number $S_{n,k} \in \mathbb{N}_0$ of k-partitions of N is called the associated **Stirling number of the second kind**. Since any k-partition of N gives rise to $k!$ compositions of N, the number of k-compositions of N is given as $k! \cdot S_{n,k}$.

We have $S_{n,k} = 0$ for $k > n$, and $S_{0,0} = 1$. For $n \geq 1$ we have $S_{n,0} = 0$ and $S_{n,1} = S_{n,n} = 1$; since any 2-partition consists of a non-empty subset and its non-empty complement, we have $S_{n,2} = \frac{2^n-2}{2} = 2^{n-1}-1$; and since any $(n-1)$-partition consists of a 2-block and $n-1$ singleton blocks, we have $S_{n,n-1} = \binom{n}{2}$. The number $B_n := \sum_{k=0}^{n} S_{n,k} \in \mathbb{N}$, that is the number of all partitions of N, is called the associated **Bell number**.

For $n, k \geq 1$ we have the recursion $S_{n,k} = S_{n-1,k-1} + k \cdot S_{n-1,k}$, giving rise to the **Stirling triangle of the second kind** shown in Table 6: Considering the k-partitions of N, we distinguish the cases whether or not the block containing n is a singleton set; hence any k-partition of N is obtained from either a $(k-1)$-partition of $N \setminus \{n\}$ by adding the k-th block $\{n\}$, or from a k-partition of $N \setminus \{n\}$ by adding n to either of the k blocks.

c) Letting $K := \{1, \ldots, k\}$, we have already shown that $|\mathrm{Maps}(K, N)| = n^k$ and $|\mathrm{Inj}(K, N)| = n_{(k)}$. To determine the cardinality $|\mathrm{Surj}(K, N)|$ we argue as follows: Any map $f \in \mathrm{Maps}(K, N)$ is uniquely determined by the preimages $f^{-1}(1), \ldots, f^{-1}(n) \subseteq K$, where $f^{-1}(i) \neq \emptyset$ if and only if $i \in \mathrm{im}(f)$. Hence letting $f \mapsto [f^{-1}(1), \ldots, f^{-1}(n)]$ yields a bijection from $\mathrm{Surj}(K, N)$ to the set of n-compositions of K. Thus we have $|\mathrm{Surj}(K, N)| = n! \cdot S_{k,n}$.

Moreover, from $\mathrm{Maps}(K, N) = \coprod_{M \subseteq N} \mathrm{Surj}(K, M)$ we get $n^k = |\mathrm{Maps}(K, N)| = \sum_{m=0}^{n} \sum_{M \subseteq N, |M|=m} |\mathrm{Surj}(K, M)| = \sum_{m=0}^{n} \binom{n}{m} \cdot m! \cdot S_{k,m} = \sum_{m=0}^{n} S_{k,m} n_{(m)}$. Since $S_{k,m} = 0$ for $m > k$ this yields $n^k = \sum_{m=0}^{k} S_{k,m} n_{(m)}$ for all $n \in \mathbb{N}_0$, implying that $X^k = \sum_{m=0}^{k} S_{k,m} X_{(m)} \in \mathbb{Z}[X]$, relating powers, falling factorials and Stirling numbers of the second kind.

(3.2) Partitions of numbers. a) A **k-composition** of $n \in \mathbb{N}_0$, where $k \in \mathbb{N}_0$, is an **ordered sum** $n = \sum_{i=1}^{k} \lambda_i$ with k **parts** $\lambda := [\lambda_1, \ldots, \lambda_k] \in \mathbb{N}^k$; we write

$|\lambda| := n$, and only $n = 0$ has a 0-composition, namely $[]$.

For $n \geq 1$ and $k \geq 1$ we may write a k-composition of n as a tuple $[\bullet \cdots \bullet \mid \bullet \cdots \bullet \mid \ldots \mid \bullet \cdots \bullet]$, with λ_i entries '$\bullet$' in the i-th slot, for $i \in \{1, \ldots, k\}$, that is with n entries '$\bullet$' and $k-1$ entries '$\mid$' such that precisely $k-1$ of the $n-1$ spaces between adjacent '$\bullet$' are filled with a '$\mid$'. Formally, letting $\lambda = [\lambda_1, \ldots, \lambda_k] \mapsto [\lambda_1 - 1, \ldots, \lambda_k - 1]$ yields a bijection from the set of k-compositions of n to the set $\mathcal{M}_{n-k}(k)$, hence the number of k-compositions of n equals $|\mathcal{M}_{n-k}(k)| = \binom{n-1}{k-1}$. For example, for $n = 3$ the 2-compositions are given as $[\bullet \mid \bullet\bullet]$ and $[\bullet\bullet \mid \bullet]$, that is $3 = 1 + 2 = 2 + 1$.

Hence for $n \geq 1$ summation over k shows that in total there are $\sum_{k=1}^{n} \binom{n-1}{k-1} = 2^{n-1}$ compositions of n; combinatorially, starting from the n-tuple $[\bullet \ldots \bullet]$ there are $n-1$ spaces to be filled, yielding 2^{n-1} possibilities.

b) A **k-partition** of n is an **unordered sum** $n = \sum_{i=1}^{k} \lambda_i$ with k **parts** $\lambda_i \in \mathbb{N}$. Hence we may assume that the parts are in non-increasing order $\lambda_1 \geq \cdots \lambda_k \geq 1$, and we write $\lambda := [\lambda_1, \ldots, \lambda_k] \vdash n =: |\lambda|$ and $l(\lambda) := k$, where in turn $|\lambda|$ and $l(\lambda)$ are called the **size** and the **length** of λ, respectively. The partition λ can be identified with the set $\{[i,j] \in \mathbb{N}^2; i \in \{1, \ldots, k\}, j \in \{1, \ldots, \lambda_i\}\}$, which is typically depicted by a **Young diagram**, that is rectangular array of boxes consisting of $l(\lambda)$ rows, where row i contains λ_i boxes. Letting $a_i(\lambda) := |\{j \in \{1, \ldots, k\}; \lambda_j = i\}| \in \mathbb{N}_0$ be the **multiplicity** of i as a part of λ, we also write $\lambda = [n^{a_n(\lambda)}, \ldots, 1^{a_1(\lambda)}]$; we have $\sum_{i=1}^{n} i a_i(\lambda) = n$ and $\sum_{i=1}^{n} a_i(\lambda) = k$, hence in particular $[a_1(\lambda), \ldots, a_n(\lambda)] \in \mathcal{M}_k(n)$. For example, all partitions of $n = 3$, ordered **reversed lexicographically**, are given as $\{[3], [2,1], [1^3]\}$.

Let $P_k(n) := \{\lambda \in \mathbb{N}^k; \lambda \vdash n\}$ be the set of all k-partitions of n, let $P_{\leq k}(n) := \coprod_{i=0}^{k} P_i(n)$ be the set of all partitions of n with at most k parts, and let $P(n) = \coprod_{k \geq 0} P_k(n)$ be the set of all partitions of n. Hence we have $P_k(n) = \emptyset$ for $k > n$, implying $P(n) = P_{\leq n}(n) = \coprod_{k=0}^{n} P_k(n)$. We have $P_0(0) = \{[]\}$, and $P_0(n) = \emptyset$ and $P_1(n) = \{[n]\}$ and $P_2(n) = \{[n-j, j]; j \in \{1, \ldots, \lfloor \frac{n}{2} \rfloor\}\}$ and $P_{n-1}(n) = \{[2, 1^{n-2}]\}$ and $P_n(n) = \{[1^n]\}$ for $n \geq 1$.

Let $p_{n,k} := |P_k(n)| \in \mathbb{N}_0$ and $p_{n,\leq k} := \sum_{i=0}^{k} p_{n,i} = |P_{\leq k}(n)| \in \mathbb{N}_0$ and $p_n := p_{n,\leq n} = \sum_{k \geq 0} p_{n,k} = |P(n)| \in \mathbb{N}$, where $p_{n,k} = 0$ for $k > n$. For $n \geq k \geq 1$ we have the recursion $p_{n,k} = p_{n-1,k-1} + p_{n-k,k}$: Considering the k-partitions $\lambda \vdash n$, we distinguish the cases $\lambda_k = 1$ and $\lambda_k \geq 2$; hence any such λ is obtained either from a $(k-1)$-partition $[\lambda_1, \ldots, \lambda_{k-1}] \vdash n-1$ as $\lambda = [\mu_1, \ldots, \mu_{k-1}, 1] \vdash n$, or from a k-partition $[\lambda_1, \ldots, \lambda_k] \vdash n-k$ as $\lambda = [\lambda_1 + 1, \ldots, \lambda_k + 1] \vdash n$.

This yields $p_{n,k} = p_{n-k,\leq k}$ for $n \geq k \geq 0$: Proceeding by induction on $n \in \mathbb{N}_0$, where the case $n = 0$ as well as the case $k = 0$ are trivial, for $n \geq k \geq 1$ we get $p_{n,k} = p_{n-1,k-1} + p_{n-k,k} = (\sum_{i=0}^{k-1} p_{n-k,i}) + p_{n-k,k} = \sum_{i=0}^{k} p_{n-k,i} = p_{n-k,\leq k}$. Alternatively, the map $[\lambda_1, \ldots, \lambda_i] \mapsto [\lambda_1 + 1, \ldots, \lambda_i + 1, 1, \ldots, 1]$ yields a bijection $P_{\leq k}(n-k) \to P_k(n)$.

Fixing n, summation yields $p_n = \sum_{k=0}^{n} p_{n,k}$, which hence can be computed using the above recursion. The asymptotic behavior of p_n is given by the **Hardy-**

Ramanujan formula [1918] as $p_n \sim \frac{\exp(\pi\sqrt{\frac{2n}{3}})}{4n\sqrt{3}}$. The asymptotic behavior of $p_{n,k}$ will be discussed in (4.4). For example we have:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
p_n	1	1	2	3	5	7	11	15	22	30	42	56	77	101	135	176
$p_{n,2}$	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7
$p_{n,3}$	0	0	0	1	1	2	3	4	5	7	8	10	12	14	16	19

(3.3) Permutations. a) Let $n \in \mathbb{N}_0$. We consider the set of permutations $\mathcal{S}_n := \mathrm{Bij}(N, N)$ of $N := \{1, \ldots, n\}$, which with respect to concatenation of maps becomes a group, with identity element id_N and inverses given by inverses of maps, called the **symmetric group** on N. Writing permutations as tuples, we have $\mathcal{S}_0 = \{[]\}$ and $\mathcal{S}_1 = \{[1]\}$, and using **lexicographic ordering** we get $\mathcal{S}_2 = \{[1, 2], [2, 1]\}$ and $\mathcal{S}_3 = \{[1, 2, 3], [1, 3, 2], [2, 1, 3], [2, 3, 1], [3, 1, 2], [3, 2, 1]\}$.

Any permutation $\pi \in \mathcal{S}_n$ can be written as a product of **disjoint cycles**: We consider the **directed graph** with **vertex** set $\{1, \ldots, n\}$ having an **edge** $i \to j$ if $\pi(i) = j$. Since π is a map, from any vertex precisely one edge emanates; since π is surjective, at any vertex at least one edge ends, since π is injective, at any vertex at most one edge ends, thus at any vertex precisely one edge ends. Hence the **connected components** of this graph are **directed circles**, showing that the cycle decomposition of π is unique up to reordering; the number of vertices in a cycle is called its **length**; for a more formal description see (15.2).

Hence a permutation is described by its cycle decomposition, which is unique up to ordering and rotating the cycles; typically **fixed points**, that is cycles of length 1, are left out. Note that inverses are given by reading cycles backwardly. In the **standard cycle representation** cycles are ordered from right to left with increasing smallest elements, the latter being chosen as starting points. For example, we have $\pi := [11, 10, 1, 7, 9, 2, 5, 3, 4, 6, 8] = (4, 7, 5, 9)(2, 10, 6)(1, 11, 8, 3) \in \mathcal{S}_{11}$ and $\pi^{-1} = (4, 9, 5, 7)(2, 6, 10)(1, 3, 8, 11)$.

b) Proceeding even further, given $\pi \in \mathcal{S}_n$ in standard cycle representation, leaving out the pairs of parentheses ')(' yields a sequence of numbers $((\pi))$, which can be considered as an element of $\mathcal{S}_n$ again. Conversely, given such a sequence of numbers, inserting a pair of parentheses ')(' right to any successive absolute left-to-right minimum yields a permutation in standard cycle representation.

Moreover, by this insertion algorithm we recover $\pi \in \mathcal{S}_n$ from $((\pi))$, showing that the map $\mathcal{S}_n \to \mathcal{S}_n \colon \pi \to ((\pi))$ is injective, hence bijective, with inverse given by the insertion algorithm; note that hence $\pi \in \mathcal{S}_n$ has $k \in \mathbb{N}_0$ cycles if and only if $((\pi))$ has precisely k successive minima. For example, π as above yields $((\pi)) = ((\underline{4}, 7, 5, 9, \underline{2}, 10, 6, \underline{1}, 11, 8, 3))$, having the successive minima as indicated, hence the insertion algorithm recovers $(4, 7, 5, 9)(2, 10, 6)(1, 11, 8, 3)$.

c) Given $\pi \in \mathcal{S}_n$, for $i \in \mathbb{N}$ let $a_i(\pi) \in \mathbb{N}_0$ be the number of cycles of π of length i. Hence $k(\pi) = \sum_{i=1}^n a_i(\pi) \in \mathbb{N}_0$ is the number of cycles of π, and we have $\sum_{i=1}^n i a_i(\pi) = n$. Hence $\lambda(\pi) := [n^{a_n(\pi)}, \ldots, 1^{a_1(\pi)}] \in P_{k(\pi)}(n)$ is a partition of

Table 7: Cycle types in $\mathcal{S}_5$.

λ	
$[5]$	$\frac{5!}{5} = 24$
$[4,1]$	$\frac{5!}{4} = 30$
$[3,2]$	$\frac{5!}{3\cdot 2} = 20$

λ	
$[3,1^2]$	$\frac{5!}{2!\cdot 3} = 20$
$[2^2,1]$	$\frac{5!}{2!\cdot 2^2} = 15$
$[2,1^3]$	$\frac{5!}{3!\cdot 2} = 10$
$[1^5]$	$\frac{5!}{5!} = 1$

n with $k(\pi) \in \mathbb{N}_0$ parts, being called the **cycle type** of π. Note that if $\alpha \in \mathcal{S}_n$ then the cycle types of π and $\alpha\pi\alpha^{-1}$ coincide: If $(i, \pi(i), \ldots, \pi^{l-1}(i))$ is a cycle of π, for some $i \in \{1, \ldots, n\}$ and $l \in \mathbb{N}$, then $(\alpha(i), \alpha\pi(i), \ldots, \alpha\pi^{l-1}(i))$ is a cycle of $\alpha\pi\alpha^{-1}$.

Given $\lambda = [n^{a_n}, \ldots, 1^{a_1}] \in P(n)$, there are $\frac{n!}{\prod_{i=1}^n a_i!\cdot i^{a_i}}$ permutations in $\mathcal{S}_n$ with cycle type λ: Consider the pattern $(\cdots)(\cdots)\cdots(\cdots)$ consisting of $a_i \in \mathbb{N}_0$ cycles of length i, for $i \in \{1, \ldots, n\}$, written in order of non-increasing lengths. There are $n!$ possibilities to fill in pairwise distinct entries from $\{1, \ldots, n\}$. Permuting the cycles of the same lengths amongst themselves, and rotating any of the cycles, we thus get any permutation $(\prod_{i=1}^n a_i!) \cdot (\prod_{i=1}^n i^{a_i})$ times.

(3.4) Permutations with a fixed number of cycles. a) Let $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$. Then the number $s_{n,k} := |\{\pi \in \mathcal{S}_n; k(\pi) = k\}| \in \mathbb{N}_0$ of permutations in $\mathcal{S}_n$ having k cycles is called the associated **(signless) Stirling number of the first kind**. Hence we have $s_{n,k} = \sum_{\lambda\in P_k(n)} \frac{n!}{\prod_{i=1}^n a_i(\lambda)!\cdot i^{a_i(\lambda)}}$; for example, for $n = 5$ we get the figures shown in Table 7. Thus we have $s_{n,k} = 0$ for $k > n$, and $s_{0,0} = 1$. For $n \geq 1$ we have $s_{n,0} = 0$ and $s_{n,n} = 1$, as well as $s_{n,1} = (n-1)!$ and $s_{n,n-1} = \binom{n}{2}$; summing over k yields $\sum_{k=0}^n s_{n,k} = |\mathcal{S}_n| = n!$.

For $n, k \geq 1$ we have the recursion $s_{n,k} = s_{n-1,k-1} + (n-1) \cdot s_{n-1,k}$, giving rise to the **Stirling triangle of the first kind** shown in Table 8: Considering the permutations of N having k cycles, we distinguish the cases whether or not n is a fixed point; hence any such permutation is obtained from either a permutation of $N \setminus \{n\}$ having $k-1$ cycles by adding the cycle (n), or from a permutation of $N \setminus \{n\}$ having k cycles by inserting n into any of the $n-1$ positions in there.

This yields $\frac{s_{n,k}}{(n-1)!} = \sum_{m=0}^{n-1} \frac{s_{m,k-1}}{m!}$, for $n, k \geq 1$: We proceed by induction on $n \in \mathbb{N}$; the assertion being true for $n = 1$, for $n \geq 2$ the recursion $s_{n,k} = s_{n-1,k-1} + (n-1) \cdot s_{n-1,k}$ yields $\frac{s_{n,k}}{(n-1)!} = \frac{s_{n-1,k-1}}{(n-1)!} + \frac{s_{n-1,k}}{(n-2)!} = \frac{s_{n-1,k-1}}{(n-1)!} + \sum_{m=0}^{n-2} \frac{s_{m,k-1}}{m!} = \sum_{m=0}^{n-1} \frac{s_{m,k-1}}{m!}$. In particular, for $k = 2$ and $n \geq 1$ we get $s_{n,2} = (n-1)! \cdot h_{n-1}$, where $h_n := \sum_{i=1}^n \frac{1}{i} \in \mathbb{Q}$ is the n-th **harmonic number**.

b) We have $X^{(n)} = \sum_{k=0}^n s_{n,k} X^k \in \mathbb{Z}[X]$: By induction on $n \in \mathbb{N}_0$, where for $n \leq 1$ both sides are equal to $1 \in \mathbb{Z}[X]$, respectively $X \in \mathbb{Z}[X]$, we for $n \geq 2$ get

Table 8: The Stirling triangle of the first kind.

$n\backslash k$	0	1	2	3	4	5	6	7
0	1							
1	0	1						
2	0	1	1					
3	0	2	3	1				
4	0	6	11	6	1			
5	0	24	50	35	10	1		
6	0	120	274	225	85	15	1	
7	0	720	1764	1624	735	175	21	1

$X^{(n)} = (X+n-1)X^{(n-1)} = (X+n-1)\cdot\sum_{k=1}^{n-1} s_{n-1,k}X^k = \sum_{k=2}^{n} s_{n-1,k-1}X^k + \sum_{k=1}^{n-1}(n-1)s_{n-1,k}X^k = \sum_{k=1}^{n}(s_{n-1,k-1} + (n-1)s_{n-1,k})X^k = \sum_{k=1}^{n} s_{n,k}X^k$.

The reciprocity $X_{(n)} = (-1)^n(-X)^{(n)}$ yields $X_{(n)} = (-1)^n \cdot \sum_{k=0}^{n} s_{n,k}(-X)^k = \sum_{k=0}^{n}(-1)^{n-k}s_{n,k}X^k \in \mathbb{Z}[X]$. This yields the following relationship between both kinds of Stirling numbers: Letting $\mathbb{Z}[X]_{\leq n} := \{f \in \mathbb{Z}[X] \setminus \{0\}; \deg(f) \leq n\} \,\dot{\cup}\, \{0\}$, then both $\{X^k; k \in \{0,\ldots,n\}\}$ and $\{X_{(k)}; k \in \{0,\ldots,n\}\}$ are $\mathbb{Z}$-bases of $\mathbb{Z}[X]_{\leq n}$, and thus from $X^n = \sum_{k=0}^{n} S_{n,k}X_{(k)} \in \mathbb{Z}[X]$ we infer $[S_{n,k}]_{n,k}^{-1} = [(-1)^{n-k}s_{n,k}]_{n,k} \in \mathrm{GL}_{n+1}(\mathbb{Z})$.

(3.5) Permutations without fixed points. A permutation $\pi \in \mathcal{S}_n$, where $n \in \mathbb{N}_0$, having no fixed points, that is $\pi(i) \neq i$ for all $i \in \{1,\ldots,n\}$, is called a **derangement**. Let $\mathcal{D}(n) \subseteq \mathcal{S}_n$ be the set of all derangements, and let $D_n := |\mathcal{D}(n)| \in \mathbb{N}_0$; we have $D_0 = 1$ and $D_1 = 0$ and $D_2 = 1$ and $D_3 = 2$.

For $n \geq 2$ we have the recursion $D_n = (n-1)(D_{n-2} + D_{n-1})$: For $\pi \in \mathcal{D}(n)$ we consider the cycle $(1, i, \ldots)$ of π containing the letter 1, where $i \in \{2,\ldots,n\}$, and we distinguish the cases whether this cycle has length 2 or ≥ 3. In the first case, deleting the cycle $(1, i)$ from π, and leaving the other cycles unchanged, yields a derangement permuting the $n-2$ letters $\{2,\ldots,n\} \setminus \{i\}$; in the second case replacing the cycle $(1, i, j, \ldots)$ of π, where $j \in \{2,\ldots,n\} \setminus \{i\}$, by $(1, j, \ldots)$ yields a derangement permuting the $n-1$ letters $\{1,\ldots,n\} \setminus \{i\}$.

We derive a closed formula for D_n, also called the **problème des rencontres**: The above recursion can be rewritten as $D_n - nD_{n-1} = -(D_{n-1} - (n-1)D_{n-2})$, for $n \geq 2$. Hence by induction we get $D_n - nD_{n-1} = (-1)^{n-1} \cdot (D_1 - D_0) = (-1)^n$, for $n \geq 1$. Considering the proportion $\frac{|\mathcal{D}(n)|}{|\mathcal{S}_n|} = \frac{D_n}{n!}$ this yields $\frac{D_n}{n!} - \frac{D_{n-1}}{(n-1)!} = \frac{(-1)^n}{n!}$, thus $\frac{D_n}{n!} - 1 = \sum_{k=1}^{n}(\frac{D_k}{k!} - \frac{D_{k-1}}{(k-1)!}) = \sum_{k=1}^{n} \frac{(-1)^k}{k!}$. Hence we get $\frac{D_n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}$, for $n \in \mathbb{N}_0$; for more conceptual proofs see (4.5) and (10.1).

We have $\sum_{k\geq 0} \frac{(-1)^k}{k!} := \lim_{n\to\infty}(\sum_{k=0}^{n} \frac{(-1)^k}{k!}) = \frac{1}{e} \sim 0.367879$, hence the proportion of derangements amongst all permutations is approximated by $\frac{D_3}{3!} = \frac{1}{3} < \frac{D_n}{n!} < \frac{3}{8} = \frac{D_4}{4!}$, for $n \geq 5$; recall that $\frac{D_0}{0!} = 1$ and $\frac{D_1}{1!} = 0$ and $\frac{D_2}{2!} = \frac{1}{2}$.

4 Difference calculus

(4.1) Discrete differentiation. a) We consider the set of all infinite sequences $[a_0, a_1, \ldots] \subseteq \mathbb{Q}$, that is the $\mathbb{Q}$-vector space $\mathcal{F} := \mathrm{Maps}(\mathbb{N}_0, \mathbb{Q})$. Since $\mathbb{N}_0$ is infinite, the polynomial ring $\mathbb{Q}[X]$ can be identified with the $\mathbb{Q}$-subspace $\{\mathbb{N}_0 \to \mathbb{Q}\colon n \mapsto f(n); f \in \mathbb{Q}[X]\} \leq \mathcal{F}$ of **polynomial maps**. Similarly, $\mathbb{Q}(X)^\circ := \{\frac{f}{g} \in \mathbb{Q}(X); g(n) \neq 0 \text{ for all } n \in \mathbb{N}_0\}$ can be identified with the $\mathbb{Q}$-subspace of $\mathcal{F}$ of **rational maps**. Let $\mathbb{Q}[X]_{\leq k} := \{f \in \mathbb{Q}[X] \setminus \{0\}; \deg(f) \leq k\} \,\dot{\cup}\, \{0\}$ for $k \geq 0$, and $\mathbb{Q}[X]_{\leq -k} := \{0\}$ for $k \geq 1$; in particular, $\mathbb{Q}[X]_{\leq 0} = \mathbb{Q}$ consists of the **constant** maps.

We consider various maps on $\mathcal{F}$: Let $\tau \in \mathrm{End}_{\mathbb{Q}}(\mathcal{F})$ be the **shift operator** defined by $\tau f\colon \mathbb{N}_0 \to \mathbb{Q}\colon n \mapsto f(n+1)$. Let $\delta := \tau - \mathrm{id} \in \mathrm{End}_{\mathbb{Q}}(\mathcal{F})$ be the **difference operator**, that is $\delta f\colon \mathbb{N}_0 \to \mathbb{Q}\colon n \mapsto f(n+1) - f(n)$, being called the **discrete derivative** of f, which is the discrete analogue of differentiation; note that $\ker(\delta) = \mathbb{Q} \leq \mathcal{F}$.

Hence τ and δ restrict to $\mathbb{Q}$-linear maps on $\mathbb{Q}[X]$ given by $(\tau f)(X) := f(X+1)$ and $(\delta f)(X) := f(X+1) - f(X)$, respectively. Moreover, by letting $\tau(\frac{f}{g}) := \frac{\tau f}{\tau g}$, for $f, g \in \mathbb{Q}[X]$ such that $g \neq 0$, we obtain $\mathbb{Q}$-linear extensions of τ and δ to $\mathbb{Q}(X)$, which in turn restrict to $\mathbb{Q}(X)^\circ$.

We get a discrete analogue of the product rule as follows: For $f, g \in \mathcal{F}$ we have $\delta(fg)\colon n \mapsto f(n+1)g(n+1) - f(n)g(n) = f(n+1)g(n+1) - f(n)g(n+1) + f(n)g(n+1) - f(n)g(n) = (\delta f)(n) \cdot g(n+1) + f(n) \cdot (\delta g)(n)$, for $n \in \mathbb{N}_0$, hence $\delta(fg) = \delta f \cdot \tau g + f \cdot \delta g \in \mathcal{F}$.

b) Let $f \in \mathcal{F}$. Then for $i \geq 0$ we have $\delta^i f = (\tau - \mathrm{id})^i f = \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} \cdot \tau^j f \in \mathcal{F}$. Thus $(\delta^i f)(0)$ is given in terms of values of f as $(\delta^i f)(0) = \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} \cdot f(j)$; in particular, we have $(\delta^i f)(0) \in \mathbb{Z}$ whenever $f(\{0, \ldots, i\}) \subseteq \mathbb{Z}$.

Conversely, we get a discrete analogue of Taylor expansions: For $n \geq 0$ we have $\tau^n = (\delta + \mathrm{id})^n$, thus we get the **Newton expansion** $f(n) = (\tau^n f)(0) = ((\delta + \mathrm{id})^n f)(0) = \sum_{i=0}^{n} \binom{n}{i} \cdot (\delta^i f)(0)$, expressing the values of f in terms of the derivatives $(\delta^i f)(0)$, and thus inverting the above formula.

For example, for the number of derangements in $\mathcal{S}_n$, where $n \in \mathbb{N}_0$, we already know that $D_n = \sum_{k=0}^{n} (-1)^k \cdot \frac{n!}{k!} = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \cdot (n-k)! = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \cdot k!$, which in present terms says that $D_n = (\delta^n(k \mapsto k!))(0)$. Thus Newton expansion yields $n! = \sum_{i=0}^{n} \binom{n}{i} \cdot D_i$; for a combinatorial proof see (4.5).

(4.2) Polynomial functions. a) We look for polynomials providing the discrete analogue of the differentiation behavior of the power maps: To this end, we generalize the falling factorial $X_{(k)} \in \mathbb{Q}[X]_{\leq k}$, which so far is defined for

$k \geq 0$, by letting $X_{(-k)} := \frac{1}{\prod_{i=1}^{k}(X+i)} \in \mathbb{Q}(X)^\circ$, for $k \geq 1$.

Then we have $X_{(k)} = X_{(k-1)} \cdot (X-k+1) \in \mathbb{Q}(X)^\circ$ and $(X+1)_{(k)} = (X+1) \cdot X_{(k-1)} \in \mathbb{Q}(X)^\circ$, for all $k \in \mathbb{Z}$: This is immediate for $k \geq 1$, for $k = 0$ we have $X_{(0)} = 1 = X_{(-1)} \cdot (X+1)$ and $(X+1)_{(0)} = 1 = (X+1) \cdot X_{(-1)}$, while for $k \geq 1$ we have $X_{(-k)} = \frac{1}{\prod_{i=1}^{k}(X+i)} = \frac{X+k+1}{\prod_{i=1}^{k+1}(X+i)} = X_{(-k-1)} \cdot (X-(-k)+1)$ and $(X+1)_{(-k)} = \frac{1}{\prod_{i=1}^{k}(X+i+1)} = \frac{X+1}{\prod_{i=1}^{k+1}(X+i)} = (X+1) \cdot X_{(-k-1)}$.

Thus we indeed have $\delta(X_{(k)}) = (X+1)_{(k)} - X_{(k)} = (X+1) \cdot X_{(k-1)} - X_{(k-1)} \cdot (X-k+1) = kX_{(k-1)} \in \mathbb{Q}(X)^\circ$, for all $k \in \mathbb{Z}$. Hence for polynomial binomial coefficients we get $\delta\binom{X}{k} = \frac{1}{k!} \cdot \delta(X_{(k)}) = \frac{1}{(k-1)!} \cdot \delta(X_{(k-1)}) = \binom{X}{k-1} \in \mathbb{Q}[X]$, for $k \geq 1$; we have $\delta\binom{X}{0} = \delta(1) = 0 \in \mathbb{Q}[X]$. Moreover, for $k \geq 0$ we have $\delta(X_{(k)}) \in \mathbb{Q}[X]_{\leq k-1} \setminus \mathbb{Q}[X]_{\leq k-2}$, hence δ induces a surjection $\mathbb{Q}[X]_{\leq k} \to \mathbb{Q}[X]_{\leq k-1}$, and we have $\delta^{-1}(\mathbb{Q}[X]_{\leq k-1}) = \mathbb{Q}[X]_{\leq k} \subseteq \mathcal{F}$. In particular, given $f \in \mathcal{F}$ and $k \geq 0$, we have $f \in \mathbb{Q}[X]_{\leq k}$ if and only if $\delta^{k+1}(f) = 0$.

b) For $f \in \mathbb{Q}[X]_{\leq k}$, where $k \geq 0$, Newton expansion becomes $f = \sum_{i=0}^{k}(\delta^i f)(0) \cdot \binom{X}{i} = \sum_{i=0}^{k} \frac{(\delta^i f)(0)}{i!} \cdot X_{(i)} \in \mathbb{Q}[X]_{\leq k}$: Since both sides are polynomials of degree at most k, it suffices to show that the identity holds for all $x := n \in \{0, \ldots, k\}$; since $\binom{n}{i} = 0$ whenever $n < i$, we get $f(n) = \sum_{i=0}^{k}(\delta^i f)(0) \cdot \binom{n}{i} = \sum_{i=0}^{n}(\delta^i f)(0) \cdot \binom{n}{i}$, which holds by Newton expansion.

Hence we have proved **Newton's Theorem**: Given a polynomial $f \in \mathbb{Q}[X]_{\leq k}$, we have $f(\mathbb{N}_0) \subseteq \mathbb{Z}$ if and only if $(\delta^i f)(0) \in \mathbb{Z}$ for all $i \in \{0, \ldots, k\}$; thus the $\mathbb{Z}$-submodule $\{f \in \mathbb{Q}[X]_{\leq k}; f(\mathbb{N}_0) \subseteq \mathbb{Z}\} \leq \mathcal{F}$ is $\mathbb{Z}$-free with $\mathbb{Z}$-basis $\{\binom{X}{0}, \ldots, \binom{X}{k}\}$.

For example, for $X^n \in \mathbb{Q}[X]$ this yields a sum formula for the Stirling numbers of the second kind: From $X^n = \sum_{k=0}^{n} S_{n,k} X_{(k)} \in \mathbb{Z}[X]$, for $n \in \mathbb{N}_0$, we for $k \in \mathbb{N}_0$ get $k! \cdot S_{n,k} = (\delta^k X^n)(0) = (\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}(X+i)^n)(0) = \sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i} i^n$; other proofs are given in (10.1) and (13.4).

(4.3) Discrete integration. a) Let $\mathcal{F} := \mathrm{Maps}(\mathbb{N}_0, \mathbb{Q})$, and let $\sigma \in \mathrm{End}_{\mathbb{Q}}(\mathcal{F})$ be the **sum operator** given by $\sigma f \colon \mathbb{N}_0 \to \mathbb{Q} \colon n \mapsto \sum_{i=0}^{n-1} f(i)$, where $(\sigma f)(0) = 0$, which is the discrete analogue of integration with lower bound 0.

Thus for $f \in \mathcal{F}$ and $n \in \mathbb{N}_0$ we have $(\delta\sigma f)(n) = \sum_{i=0}^{n} f(i) - \sum_{i=0}^{n-1} f(i) = f(n)$, showing that $\delta\sigma = \mathrm{id}$, which is the discrete analogue of the theorem on integration-differentiation. Hence $\delta^{-1}(f) = \sigma f + \mathbb{Q} \subseteq \mathcal{F}$ are the **discrete stem functions** of f, where σf is the unique one such that $(\sigma f)(0) = 0$.

The other way around we get $(\sigma\delta f)(n) = \sum_{i=0}^{n-1}(\delta f)(i) = \sum_{i=0}^{n-1}(f(i+1) - f(i)) = f(n) - f(0)$, that is $\sigma\delta f = f - f(0) \in \mathcal{F}$, which is the discrete analogue of the theorem on differentiation-integration; hence to determine the sum $\sum_{i=a}^{b} g(i)$, for $g \in \mathcal{F}$ and $a, b \in \mathbb{N}_0$, we may just evaluate any stem function $f \in \mathcal{F}$ of g, yielding $\sum_{i=a}^{b} g(i) = (\sigma g)(b+1) - (\sigma g)(a) = f(b+1) - f(a)$.

b) For example, from $\delta(X_{(k)}) = kX_{(k-1)} \in \mathbb{Q}(X)$, for $k \in \mathbb{Z}$, we get the stem functions $\sigma(X_{(k)}) = \frac{1}{k+1} \cdot X_{(k+1)} \in \mathbb{Q}[X]$ for $k \geq 0$, and $\sigma(X_{(k)}) = \frac{1}{k+1} \cdot (X_{(k+1)} - \frac{1}{(-k-1)!}) \in \mathbb{Q}(X)^\circ$ for $k \leq -2$. But for $k = -1$ we get the discrete analogue of the logarithm $\sigma(X_{(-1)}) \colon \mathbb{N}_0 \to \mathbb{Q} \colon n \mapsto \sum_{i=0}^{n-1} \frac{1}{i+1} = h_n$, where h_n is the n-th harmonic number, and $h_0 := 0$.

In order to find the discrete analogue of the exponential map, we observe that $f \in \mathcal{F}$ fulfills $\delta f = f$ if and only if $f(n) = f(n+1) - f(n)$, that is $f(n+1) = 2f(n)$ for all $n \in \mathbb{N}_0$, which is equivalent to $f \colon \mathbb{N}_0 \to \mathbb{Q} \colon n \mapsto c \cdot 2^n$, for some $c \in \mathbb{Q}$.

For example, we determine the power sums $s_k \colon \mathbb{N}_0 \to \mathbb{N}_0 \colon n \mapsto \sum_{i=0}^{n-1} i^k$, where $k \in \mathbb{N}_0$: We have $s_k = \sigma(X^k) = \sum_{i=0}^{k} S_{k,i} \sigma(X_{(i)}) = \sum_{i=0}^{k} \frac{S_{k,i}}{i+1} \cdot X_{(i+1)} \in \mathbb{Q}[X]$, being a polynomial map of degree $k+1$ with leading coefficient $\frac{1}{k+1}$ and constant coefficient 0; in particular, we have $s_0 = X \in \mathbb{Q}[X]$ and $s_1 = \frac{1}{2} X(X-1) \in \mathbb{Q}[X]$.

c) The discrete analogue of partial integration follows from the product rule: For $f, g \in \mathcal{F}$ we have $\sigma(f \cdot \delta g) = \sigma\delta(fg) - \sigma(\delta f \cdot \tau g) = fg - (fg)(0) - \sigma(\delta f \cdot \tau g) \in \mathcal{F}$.

For example, we determine $\mathbb{N}_0 \to \mathbb{N}_0 \colon n \mapsto \sum_{i=0}^{n-1} i \cdot 2^i$: Letting $f := X \in \mathbb{Q}[X]$ and $g \colon \mathbb{N}_0 \to \mathbb{N}_0 \colon n \mapsto 2^n$, we get $\delta g = g$ and $\delta f = 1 \in \mathbb{Q}[X]$, and thus $\sum_{i=0}^{n-1} i \cdot 2^i = (\sigma(f \cdot \delta g))(n) = (fg)(n) - (fg)(0) - (\sigma(\delta f \cdot \tau g))(n) = n \cdot 2^n - \sum_{i=0}^{n-1} 2^{i+1} = n \cdot 2^n - (2^{n+1} - 2) = (n-2) \cdot 2^n + 2$.

For example, we determine $\mathbb{N}_0 \to \mathbb{N}_0 \colon n \mapsto \sum_{i=0}^{n-1} h_i$: Letting $f \colon \mathbb{N}_0 \to \mathbb{N}_0 \colon n \mapsto h_n$ and $g := X \in \mathbb{Q}[X]$, we get $\delta f = \frac{1}{X+1} \in \mathbb{Q}(X)^\circ$ and $\delta g = 1 \in \mathbb{Q}[X]$, and thus $\sum_{i=0}^{n-1} h_i = (\sigma(f \cdot \delta g))(n) = (fg)(n) - (fg)(0) - (\sigma(\delta f \cdot \tau g))(n) = nh_n - \sum_{i=0}^{n-1} \frac{i+1}{i+1} = n(h_n - 1)$.

(4.4) Example: Growth of partition numbers. The growth behavior of $p_{n,k}$, where $k \in \mathbb{N}$ is fixed while $n \in \mathbb{N}_0$ varies, is described by a **quasi-polynomial** as follows: For $k \in \mathbb{N}$ and $a \in \{0, \ldots, k!-1\}$ there are $p_{a,k}(X) \in \mathbb{Q}[X]$ such that $p_{n,k} = p_{a,k}(n)$, for all $n \in \mathbb{N}_0$ such that $n \equiv a \pmod{k!}$ and $[n,k] \neq [0,1]$. Moreover, $p_{a,k}(X)$ has degree $k-1$ and leading coefficient $\frac{1}{k!(k-1)!}$; for example, we have $p_{0,1}(X) = 1$, and $p_{a,2}(X) = \frac{X-a}{2}$ for $a \in \{0,1\}$:

We proceed by induction on $k \in \mathbb{N}$: We have $p_{n,1} = 1$ for $n \geq 1$. Moreover, since $p_{0,0} = 1$ we have $p_{n,0} + p_{n,1} = 1$ for all $n \in \mathbb{N}_0$, hence for completeness let additionally $p_{0,0} := 0 \in \mathbb{Q}[X]$. Note that $p_{n,2} = p_{n-2,2} + p_{n-2,1} + p_{n-2,0} = p_{n-2,2} + 1$ yields $p_{n,2} = \frac{n}{2}$ if $n \in \mathbb{N}_0$ is even, and $p_{n,2} = \frac{n-1}{2}$ if $n \in \mathbb{N}$ is odd.

Let $k \geq 2$. For $n \geq k$ the recursion $p_{n,k} = \sum_{i=0}^{k} p_{n-k,i}$ yields $p_{n,k} - p_{n-k,k} = \sum_{i=0}^{k-1} p_{(a-k) \pmod{i!}, i}(n-k)$, which entails $p_{n,k} - p_{n-k!,k} = \sum_{j=0}^{(k-1)!-1} (p_{n-jk,k} - p_{n-(j+1)k,k}) = \sum_{i=0}^{k-1} \sum_{j=1}^{(k-1)!} p_{(a-jk) \pmod{i!}, i}(n-jk)$, for $n \geq k!$. We apply discrete differentiation to the map $f_{a,k} \colon \mathbb{N}_0 \to \mathbb{N}_0 \colon m \mapsto p_{a+mk!,k}$:

We have $(\delta f_{a,k})(m) = \sum_{i=0}^{k-1} \sum_{j=1}^{(k-1)!} p_{(a-jk) \pmod{i!}, i}((m+1) \cdot k! + a - jk)$, hence by induction $\delta f_{a,k}$ is a polynomial map in $m \in \mathbb{N}_0$ of degree $k-2$ with leading coefficient $(k-1)! \cdot \frac{(k!)^{k-2}}{(k-1)!(k-2)!} = \frac{(k!)^{k-2}}{(k-2)!}$. Thus we infer that $f_{a,k}$ is a polynomial map in $m \in \mathbb{N}_0$ of degree $k-1$ with leading coefficient $\frac{(k!)^{k-2}}{(k-1)\cdot(k-2)!} = \frac{(k!)^{k-2}}{(k-1)!}$. Letting $n := a + mk!$ shows that $p_{n,k}$ is a polynomial map of degree $k-1$ with leading coefficient $\frac{1}{(k!)^{k-1}} \cdot \frac{(k!)^{k-2}}{(k-1)!} = \frac{1}{k!(k-1)!}$. $\sharp$

(4.5) Linear inversion. a) A sequence $[f_0, f_1, \ldots] \subseteq \mathbb{Q}[X]$ such that $\deg(f_i) = i$ for all $i \in \mathbb{N}_0$, is called **basic**. Then, for all $n \in \mathbb{N}_0$, the truncated sequence $[f_0, \ldots, f_n] \subseteq \mathbb{Q}[X]$ is a $\mathbb{Q}$-basis of $\mathbb{Q}[X]_{\leq n}$.

Thus, if $[f_0, f_1, \ldots] \subseteq \mathbb{Q}[X]$ and $[g_0, g_1, \ldots] \subseteq \mathbb{Q}[X]$ are basic sequences, then there are lower triangular matrices $A = [a_{ij}]_{ij} \in \mathbb{Q}^{\mathbb{N}_0 \times \mathbb{N}_0}$ and $B = [b_{ij}]_{ij} \in \mathbb{Q}^{\mathbb{N}_0 \times \mathbb{N}_0}$ such that $g_i = \sum_{j=0}^{i} a_{ij} f_j \in \mathbb{Q}[X]_{\leq i}$ and $f_i = \sum_{j=0}^{i} b_{ij} g_j \in \mathbb{Q}[X]_{\leq i}$, for all $i \in \mathbb{N}_0$. Thus we have $\sum_{k=j}^{i} a_{ik} b_{kj} = \delta_{ij} \in \mathbb{Q}$ and $\sum_{k=j}^{i} b_{ik} a_{kj} = \delta_{ij} \in \mathbb{Q}$, for all $i, j \in \{0, \ldots, n\}$. Note that the matrices $A_n := [a_{ij}]_{i \in \{0,\ldots,n\}, j \in \{0,\ldots,n\}} \in \mathbb{Q}^{(n+1)\times(n+1)}$ and $B_n := [b_{ij}]_{i \in \{0,\ldots,n\}, j \in \{0,\ldots,n\}} \in \mathbb{Q}^{(n+1)\times(n+1)}$ are the associated base change matrices on $\mathbb{Q}[X]_{\leq n}$, which hence are inverse to each other.

Hence, if $[x_0, x_1, \ldots] \subseteq \mathbb{Q}$ is any sequence, and the sequence $[y_0, y_1, \ldots] \subseteq \mathbb{Q}$ is given by weighted sums as $y_i := \sum_{j=0}^{i} a_{ij} x_j \in \mathbb{Q}$, for all $i \in \mathbb{N}_0$, then we may recover the original sequence by $x_i = \sum_{j=0}^{i} b_{ij} y_j \in \mathbb{Q}$, for all $i \in \mathbb{N}_0$: For all $n \in \mathbb{N}_0$ we have $[y_0, \ldots, y_n]^{\mathrm{tr}} = A_n \cdot [x_0, \ldots, x_n]^{\mathrm{tr}} \in \mathbb{Q}^{1\times(n+1)}$, which is equivalent to $[x_0, \ldots, x_n]^{\mathrm{tr}} = B_n \cdot [y_0, \ldots, y_n]^{\mathrm{tr}} \in \mathbb{Q}^{1\times(n+1)}$.

b) For the basic sequences $[X^n; n \in \mathbb{N}_0] \subseteq \mathbb{Q}[X]$ and $[(X-1)^n; n \in \mathbb{N}_0] \subseteq \mathbb{Q}[X]$ we get $X^n = (X-1+1)^n = \sum_{k=0}^{n} \binom{n}{k} (X-1)^k \in \mathbb{Q}[X]_{\leq n}$ and $(X-1)^n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} X^k \in \mathbb{Q}[X]_{\leq n}$, with base change matrices $[\binom{n}{k}]_{n,k} \in \mathbb{Z}^{\mathbb{N}_0 \times \mathbb{N}_0}$ and $[(-1)^{n-k}\binom{n}{k}]_{n,k} \in \mathbb{Z}^{\mathbb{N}_0 \times \mathbb{N}_0}$, respectively. More symmetrically, replacing the first basic sequence by $[(-X)^n; n \in \mathbb{N}_0] \subseteq \mathbb{Q}[X]$ we get the involutory base change matrix $[(-1)^n \binom{n}{k}]_{n,k} \in \mathbb{Z}^{\mathbb{N}_0 \times \mathbb{N}_0}$.

Thus, for sequences $[y_0, y_1, \ldots] \subseteq \mathbb{Q}$ and $[x_0, x_1, \ldots] \subseteq \mathbb{Q}$ **binomial inversion** says that we have $y_n = \sum_{k=0}^{n} \binom{n}{k} x_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$, if and only if we have $x_n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} y_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$. Moreover, if $y_n = \sum_{k=0}^{n} (-1)^n \binom{n}{k} x_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$, then binomial inversion becomes involutory, that is $x_n = \sum_{k=0}^{n} (-1)^n \binom{n}{k} y_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$.

In particular, considering the unit sequence $[x_0, x_1, \ldots] := [0, \ldots, 0, 1, 0, \ldots] \subseteq \mathbb{Z}$, where the non-vanishing entry is in position $m \in \mathbb{N}_0$, we get $y_n = \binom{n}{m} \in \mathbb{Z}$, for all $n \in \mathbb{N}_0$, and thus $\sum_{k=m}^{n} (-1)^{n-k} \binom{n}{k} \binom{k}{m} = \delta_{n,m} \in \mathbb{Z}$; note that for $m = 0$ we recover the identity $\sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} = \delta_{0,n}$. Moreover, Newton expansion just says that any sequence $f := [f_i; i \in \mathbb{N}_0] \subseteq \mathbb{Q}$ is related to the sequence $[(\delta^i f)(0); i \in \mathbb{N}_0] \subseteq \mathbb{Q}$ by binomial inversion.

For example, we reconsider the number D_n of derangements in $\mathcal{S}_n$, for $n \in \mathbb{N}_0$: Let $D_{n,k} \in \mathbb{N}_0$ be the number of permutations in $\mathcal{S}_n$ having precisely k fixed points, for $k \in \mathbb{N}_0$; hence we have $D_n = D_{n,0}$. Choosing any k-subset of $\{1, \ldots, n\}$ as set of fixed points, we get $n! = \sum_{k=0}^{n} \binom{n}{k} D_{n-k} = \sum_{k=0}^{n} \binom{n}{k} D_k$.

Thus binomial inversion yields $D_n = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\cdot k! = n!\cdot\sum_{k=0}^{n}\frac{(-1)^{n-k}}{(n-k)!} = n!\cdot\sum_{k=0}^{n}\frac{(-1)^k}{k!}$, and Newton expansion becomes $D_n = (\delta^n(k \mapsto k!))(0)$.

c) For the basic sequences $[X^n; n \in \mathbb{N}_0] \subseteq \mathbb{Q}[X]$ and $[X_{(n)}; n \in \mathbb{N}_0] \subseteq \mathbb{Q}[X]$ we have $X^n = \sum_{k=0}^{n} S_{n,k}X_{(k)} \in \mathbb{Q}[X]_{\leq n}$ and $X_{(n)} = \sum_{k=0}^{n}(-1)^{n-k}s_{n,k}X^k \in \mathbb{Q}[X]_{\leq n}$, with base change matrices $[S_{n,k}]_{n,k} \in \mathbb{Z}^{\mathbb{N}_0\times\mathbb{N}_0}$ and $[(-1)^{n-k}s_{n,k}]_{n,k} \in \mathbb{Z}^{\mathbb{N}_0\times\mathbb{N}_0}$, respectively. Again, replacing replacing the first basic sequence by $[(-X)^n; n \in \mathbb{N}_0] \subseteq \mathbb{Q}[X]$ we get the base change matrices $[(-1)^n S_{n,k}]_{n,k} \in \mathbb{Z}^{\mathbb{N}_0\times\mathbb{N}_0}$ and $[(-1)^n s_{n,k}]_{n,k} \in \mathbb{Z}^{\mathbb{N}_0\times\mathbb{N}_0}$, respectively.

Thus, for sequences $[y_0, y_1, \ldots] \subseteq \mathbb{Q}$ and $[x_0, x_1, \ldots] \subseteq \mathbb{Q}$ **Stirling inversion** says that we have $y_n = \sum_{k=0}^{n} S_{n,k}x_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$, if and only if we have $x_n = \sum_{k=0}^{n}(-1)^{n-k}s_{n,k}y_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$. Moreover, if $y_n = \sum_{k=0}^{n}(-1)^n S_{n,k}x_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$, then Stirling inversion becomes more symmetric inasmuch $x_n = \sum_{k=0}^{n}(-1)^n s_{n,k}y_k \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$. In particular, considering the m-th unit sequence again, where $m \in \mathbb{N}_0$, we for all $n \in \mathbb{N}_0$ recover $\sum_{k=m}^{n}(-1)^{n-k}s_{n,k}S_{k,m} = \delta_{n,m} \in \mathbb{Z}$.

II Posets

5 Partially ordered sets

(5.1) Partially ordered sets. a) A set $X \neq \emptyset$, together with a binary relation $\leq$ on X, is called a **partially ordered set** or **poset** if the following properties are fulfilled for all $x, y, z \in X$: We have **reflexivity**, that is $x \leq x$; **antisymmetry**, that is $x \leq y$ and $y \leq x$ implies $x = y$; and **transitivity**, that is $x \leq y$ and $y \leq z$ implies $x \leq z$. Elements $x, y \in X$ are called **comparable** if $x \leq y$ or $y \leq x$; and we write $x < y$ whenever $x \leq y$ and $x \neq y$. Note that we get the **dual** partially ordered set by letting $x \leq' y$ if and only if $y \leq x$, for all $x, y \in X$.

Partially ordered sets X and Y, with comparison relations $\leq_X$ and $\leq_Y$, are called **isomorphic**, if there is a bijection $\alpha\colon X \to Y$ such that $x \leq_X y$ if and only if $\alpha(x) \leq_Y \alpha(y)$, for all $x, y \in X$; we write $X \cong Y$. In particular, any subset $\emptyset \neq Y \subseteq X$ becomes a partially ordered set again, with respect to the **induced** comparison relation, which is defined by saying that the natural embedding $\iota\colon Y \to X$ yields an isomorphism $Y \cong \mathrm{im}(\iota)$.

A subset $Y \subseteq X$, such that whenever $y \in Y$ and $x \in X$ such that $x \leq y$ we already have $x \in Y$, is called an **ideal** of X. Similarly, a subset $Y \subseteq X$, such that whenever $y \in Y$ and $x \in X$ such that $x \geq y$ we already have $x \in Y$, is called a **coideal** of X. In particular, for any $x \in X$ the sets $\langle \leq x\rangle := \{y \in X; y \leq x\}$ and $\langle x \leq\rangle := \{y \in X; x \leq y\}$ are called the **principal** ideal and coideal **generated** by x, respectively.

b) An element $x \in X$ is called **minimal**, if $y \leq x \in X$ already implies $y = x$; dually, $x \in X$ is called **maximal** if $x \leq y \in X$ already implies $y = x$. In particular, if X is finite then for any $x \in X$ there is a minimal element $y \in X$ such that $y \leq x$, and a maximal element $z \in X$ such that $x \leq z$. Moreover, an element $\underline{0} \in X$ such that $\underline{0} \leq x$ for all $x \in X$ is called a **zero** element; dually, an element $\underline{1} \in X$ such that $x \leq \underline{1}$ for all $x \in X$ is called a **one** element; note that zero and one elements are the unique minimal and maximal elements, respectively, if existent.

Given a subset $\emptyset \neq Y \subseteq X$, an element $x \in X$ such that $y \leq x$ for all $y \in Y$ is called an **upper bound** of Y. If Y has an upper bound, and the set of upper bounds has a zero element x, with respect to the induced comparison relation, then $x \in X$ is called a **least upper bound** or **join** or **supremum** of Y; if existent, the least upper bound is unique, and denoted by $\bigvee_{y \in Y} y$. Dually, an element $x \in X$ such that $x \leq y$ for all $y \in Y$ is called a **lower bound** of Y. If Y has a lower bound, and the set of lower bounds has a one element x, with respect to the induced comparison relation, then $x \in X$ is called a **greatest lower bound bound** or **meet** or **infimum** of Y; if existent, the greatest lower bound is unique, and denoted by $\bigwedge_{y \in Y} y$.

c) The partially ordered set X is called a **lattice**, if for all $x, y \in X$ there is a join $x \vee y \in X$ and a meet $x \wedge y \in X$. In this case, for all $x, y, z \in X$ we have **idempotency** $x \wedge x = x$ and commutativity $x \wedge y = y \wedge x$, and dually $x \vee x = x$ and $x \vee y = y \vee x$; moreover, $(x \wedge y) \wedge z \in X$ is the greatest lower bound of $\{x, y, z\} \in X$, hence we infer associativity $(x \wedge y) \wedge z = x \wedge (y \wedge z)$, and dually $(x \vee y) \vee z = x \vee (y \vee z)$; finally $x \wedge y = x$ is equivalent to $x \leq y$, which in turn is equivalent to $x \vee y = y$.

In particular, if X is finite, then $\underline{0} := \bigwedge_{x \in X} x \in X$ and $\underline{1} := \bigvee_{x \in X} x \in X$ are a zero and a one element, respectively.

(5.2) Chains. a) Let X be a partially ordered set. Given $x \leq y \in X$, the associated **(closed) interval** is the set $[x, y] := \{z \in X; x \leq z \leq y\}$; in particular, we have $[x, x] = \{x\}$. The partially ordered set X is called **locally finite** if all its intervals are finite sets.

Given $x < y$, we say that y **covers** x if $[x, y] = \{x, y\}$, and we write $x \lessdot y$. In particular, if $\underline{0} \in X$ is a zero element, then an element $x \in X$ such that $\underline{0} \lessdot x$ is called an **atom**; similarly, if $\underline{1} \in X$ is a one element, then an element $x \in X$ such that $x \lessdot \underline{1}$ is called a **co-atom**. If X is finite, then it is typically depicted by its **Hasse diagram**, which is a **quiver**, that is a simple oriented graph, with vertex set X and **arrows** $x \leftarrow y$ whenever $x \lessdot y$.

A subset $\emptyset \neq Y \subseteq X$ is called a **chain** or **totally ordered** if all $x, y \in Y$ are pairwise comparable; in particular any chain is a lattice. Note that **Zorn's Lemma** (being equivalent to the **Axiom of Choice**) says that, if any chain $Y \subseteq X$ has an upper bound, then X has a maximal element.

A chain $Y \subseteq X$ is called **maximal** if for any chain $Y' \subseteq X$ such that $Y \subseteq Y'$ we already have $Y = Y'$. A chain $Y \subseteq X$ is called **saturated** or **unrefinable** if for all $x \leq y \in Y$ the chain $Y \cap [x, y] \subseteq [x, y]$ is maximal. Hence maximal chains are saturated, while the converse does not hold; note that a saturated chain is maximal if it contains both a minimal and a maximal element of X.

The set of all chains in X is partially ordered by set-theoretic inclusion, and any chain of chains has its union, which again is a chain, as an upper bound, hence by **Zorn's Lemma** any chain is contained in a maximal chain. Of course, if X is finite, then any chain can be refined to a maximal chain by induction.

For example, the set $\mathbb{N}_0$ is a partially ordered set with respect to the $\leq$ relation. It is a locally finite chain, thus in particular a lattice, with meet $\min\{i,j\}$ and join $\max\{i,j\}$, for $i,j \in \mathbb{N}_0$; it has zero element 0, but no maximal element. The subset $[0,n] := \{0,\ldots,n\}$, where $n \in \mathbb{N}_0$, is totally ordered of length n, with zero element 0 and one element n. Note that for any interval $[m,n]$ in $\mathbb{N}_0$, where $m \leq n$, we have $[m,n] \cong [0, m-n]$ as partially ordered sets; moreover, any finite totally ordered set X is isomorphic to $[0, |X|-1]$.

b) Let $Y = \{x_0,\ldots,x_n\} \subseteq X$ be a finite chain of **length** $l(Y) := n \in \mathbb{N}_0$, where we may assume that $x_0 < x_1 < \cdots < x_n$; in particular, x_0 and x_n are its zero and one elements, respectively, and Y is saturated if and only if $x_{i-1} \lessdot x_i$, for all $i \in \{1,\ldots,n\}$. The **length** of a partially ordered set X is defined as $l(X) := \max\{l(Y) \in \mathbb{N}_0; Y \subseteq X \text{ finite chain}\} \in \mathbb{N}_0 \,\dot\cup\, \{\infty\}$. If X is finite, then we have $l(X) \in \mathbb{N}_0$, where it of course suffices to consider the maximal chains only; if X is locally finite, then for $x \leq y$ we write $l(x,y) := l([x,y]) \in \mathbb{N}_0$. Moreover, X is said to be **graded** of **length** $n \in \mathbb{N}_0$, if all its maximal chains have finite length n; then for $x \leq y$ the interval $[x,y]$ is graded of length $l(x,y) \leq n$.

If X is graded of length $n \in \mathbb{N}_0$, then there is a well-defined **length** map $l \colon X \to \{0,\ldots,n\}$, for $x \in X$ given by $l(x) := l(x',x)$, where $x' \in X$ is any minimal element such that $x' \leq x$. Thus, in particular, if $x \in X$ is a minimal element then we have $l(x) = 0$, for $x \leq y$ we have $l(y) - l(x) = l(x,y)$, and for $x \lessdot y$ we have $l(y) = l(x) + 1$. This is seen as follows:

Note first, since any maximal chain has finite length, there is a minimal element $x' \in X$ as desired, and similarly there is a maximal element $y \in X$ such that $x \leq y$. Now, if $x'' \in X$ also is a minimal element such that $x'' \leq x$, then letting $X' \subseteq [x',x]$ and $X'' \subseteq [x'',x]$ be maximal chains, and choosing a maximal chain $Y \subseteq [x,y]$, yields maximal chains $X' \cup Y$ and $X'' \cup Y$ in X, hence $l(x',x) = l(X') = l(X' \cup Y) - l(Y) = l(X'' \cup Y) - l(Y) = l(X'') = l(x'',x)$. ♯

For example, let $X := \{\emptyset, \{1\}, \{2\}, \{2,3\}, \{1,2,3\}\}$, partially ordered by set-theoretic inclusion $\subseteq$. By inspection, X is a lattice, having $\emptyset$ and $\{1,2,3\}$ as its zero and one elements, respectively. But while meets are given by set-theoretic intersections, this is not the case for joins: we have $\{1\} \vee \{2\} = \{1,2,3\}$. Moreover, both $\emptyset \subset \{1\} \subset \{1,2,3\}$ and $\emptyset \subset \{2\} \subset \{2,3\} \subset \{1,2,3\}$ are maximal chains, of lengths 2 and 3, respectively, thus X is not graded.

(5.3) Example: Subset lattices. Let N be a set. Then the **finitary** power set $\mathcal{P}_{\text{fin}}(N) \subseteq \mathcal{P}(N)$ consisting of the finite subsets of N, and the **co-finitary** power set $\mathcal{P}_{\text{co-fin}}(N) \subseteq \mathcal{P}(N)$ consisting of the subsets of N having finite complements, are partially ordered by set-theoretic inclusion $\subseteq$. They are locally finite lattices with meet $M \cap M'$ and join $M \cup M'$. Moreover, $\mathcal{P}_{\text{fin}}(N)$ has $\emptyset$ as its zero element, and it has a one element if and only if N is finite, in this case coinciding with N; and $\mathcal{P}_{\text{co-fin}}(N)$ has N as one element, and it has a zero element if and only if N is finite, in this case coinciding with $\emptyset$.

If N is finite of cardinality $n \in \mathbb{N}_0$, that is we have $\mathcal{P}(N) = \mathcal{P}_{\text{fin}}(N) = \mathcal{P}_{\text{co-fin}}(N)$, all maximal chains are of the form $\emptyset = M_0 \subset M_1 \subset \cdots \subset M_n = N$, where $|M_i| = i$ for all $i \in \{0, \ldots, n\}$. Hence $\mathcal{P}(N)$ is graded of length n, where $M \subseteq N$ has length $|M|$. Then $\mathcal{P}(N)$ has $\binom{n}{k}$ elements of length $k \in \{0, \ldots, n\}$, that is k-subsets, and is totally ordered if and only if $n \leq 1$. For example, the Hasse diagram of $\mathcal{P}(\{1,2,3\})$ is depicted on the left hand side of Table 9, where the vertices are labeled as follows: The subsets of $\{1,2,3\}$ are identified with the indicator functions $\text{Maps}(\{1,2,3\},\{0,1\})$, and the latter in turn are identified with the 2-adic representations of the numbers $\{0, \ldots, 7\}$.

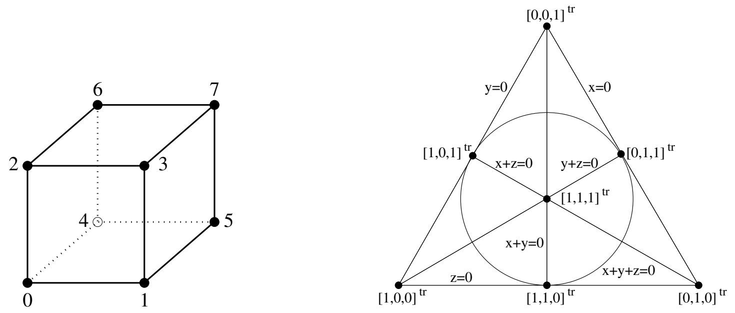
(5.4) Example: Subspace lattices. Let $\mathbb{F}_q$ be the finite field of order q, and for $n \in \mathbb{N}_0$ let $\mathbf{P}_n(q)$ be the set of $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^n$, partially ordered by set-theoretic inclusion $\subseteq$. Then $\mathbf{P}_n(q)$ is a lattice with meet $V \cap V'$ and join $V + V'$, for $V, V' \leq \mathbb{F}_q^n$, and has $\{0\}$ and $\mathbb{F}_q^n$ as its zero and one elements, respectively. Note that $\mathbf{P}_n(q)$ is totally ordered if and only if $n \leq 1$. All maximal chains are of the form $\emptyset = V_0 < V_1 < \cdots < V_n = \mathbb{F}_q^n$, where $\dim_{\mathbb{F}_q}(V_i) = i$ for all $i \in \{0, \ldots, n\}$. Hence $\mathbf{P}_n(q)$ is graded of length n, where $V \leq \mathbb{F}_q^n$ has length $\dim_{\mathbb{F}_q}(V)$. Counting sequences of $\mathbb{F}_q$-linearly independent sequences we conclude that $\mathbf{P}_n(q)$ has $\binom{n}{k}_q := \prod_{i=0}^{k-1} \frac{q^n - q^i}{q^k - q^i} = \prod_{i=0}^{k-1} \frac{q^{n-i}-1}{q^{k-i}-1} = \prod_{i=0}^{k-1} \frac{q^{n-i}-1}{q^{i+1}-1}$ $\mathbb{F}_q$-subspaces of $\mathbb{F}_q$-dimension $k \in \{0, \ldots, n\}$, that is of length k.

Counting differently, using a fixed embedding $\mathbb{F}_q^{n-1} \leq \mathbb{F}_q^n$, for $n \geq 1$, for any k-dimensional $\mathbb{F}_q$-subspace $V \leq \mathbb{F}_q^n$, for $k \in \{1, \ldots, n\}$, we have either $V \leq \mathbb{F}_q^{n-1}$ or $\dim_{\mathbb{F}_q}(V \cap \mathbb{F}_q^{n-1}) = k-1$. Fixing a $(k-1)$-dimensional $\mathbb{F}_q$-subspace $V' \leq \mathbb{F}_q^{n-1}$ and going over to $\mathbb{F}_q^n/V'$, shows that there are $\frac{q^{n-k+1}-1}{q-1} - \frac{q^{n-k}-1}{q-1} = q^{n-k}$ subspaces V as above such that $V \cap \mathbb{F}_q^{n-1} = V'$. Thus we get the **triangle identity** $\binom{n}{k}_q = \binom{n-1}{k}_q + q^{n-k} \cdot \binom{n-1}{k-1}_q$, for all $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$. Since $\binom{n}{0}_q = 1$, for all $n \in \mathbb{N}_0$, we see that $\binom{n}{k}_q$ can be considered as an element of the polynomial ring $\mathbb{Z}[q]$ in the indeterminate q, called a **Gaussian polynomial**.

Since for $i \in \mathbb{N}$ we have $q^i - 1 = (q-1) \cdot \sum_{j=0}^{i-1} q^j \in \mathbb{Z}[q]$, we may specialize $\binom{n}{k}_q = \prod_{i=0}^{k-1} \frac{q^{n-i}-1}{q^{i+1}-1} = \prod_{i=0}^{k-1} \frac{\sum_{j=0}^{n-i-1} q^j}{\sum_{j=0}^{i} q^j} \in \mathbb{Q}(q)$ at $q \mapsto 1$, yielding $\binom{n}{k}_q|_{q \mapsto 1} = \prod_{i=0}^{k-1} \frac{n-i}{i+1} = \binom{n}{k}$. Thus $\binom{n}{k}_q$ is called the ***q*-analogue** of the binomial coefficient $\binom{n}{k}$; since the latter coincides with the number of elements of length k in the partially ordered set $\mathcal{P}(N)$ where N has cardinality n, in this sense $\mathcal{P}(N)$ is called 'the n-dimensional vector space over the field with one element'.

For example, $\mathbf{P}_2(2)$ is isomorphic to $\{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2,3\}\}$, partially ordered

Table 9: The cube and the Fano plane.



by set-theoretic inclusion $\subseteq$, and $\mathbf{P}_3(2)$ is isomorphic to

$$\{\emptyset, \quad \{1\},\{2\},\{3\},\{4\},\{5\},\{6\},\{7\}, \quad \{1,\ldots,7\}, \\ \{1,2,3\},\{1,4,5\},\{2,4,6\},\{1,6,7\},\{2,5,7\},\{3,4,7\},\{3,5,6\}\}.$$

The subset $\mathbf{P}_3(2)\setminus\{\{0\},\mathbb{F}_2^3\}$ is called the **Fano plane** or the **projective plane of order** 2; it is depicted in Table 9, where 1-dimensional $\mathbb{F}_2$-subspaces are given as vertices, labeled by generating vectors, and 2-dimensional $\mathbb{F}_2$-subspaces are given as lines, labeled by defining linear equations, and where the numbers $\{1,\ldots,7\}$ are identified with vectors by going over to 2-adic representations.

(5.5) Example: Divisibility lattices. We consider the basic example from number theory: the set $\mathbb{N}$, partially ordered by divisibility $\mid$. Then $\mathbb{N}$ is a locally finite lattice, with meet $\gcd(c,d)$ and join $\operatorname{lcm}(c,d)$, for $c,d \in \mathbb{N}$; it has 1 as its zero element, but does not have a one element. Indeed, for $n \in \mathbb{N}$, the interval in $\mathbb{N}$ between 1 and n is given as $X_n := \{d \in \mathbb{N}; d \mid n\}$, that is the set of divisors of $n \in \mathbb{N}$, and the interval between $d \mid n \in \mathbb{N}$ is isomorphic to $X_{\frac{n}{d}}$.

All maximal chains in X_n are of the form $1 = d_0 \mid d_1 \mid \cdots \mid d_k = n$, where $\frac{d_i}{d_{i-1}} \in \mathbb{N}$ is a prime, for all $i \in \{1,\ldots,k\}$, and where $k \in \mathbb{N}_0$ is the length of the prime factorization of n, counting multiplicities. Hence X_n is graded of length k, where the length of $d \in X_n$ is given by the length of its prime factorization; note that X_n is totally ordered if and only if n is a prime power.

(5.6) Example: Dominance partial order. a) We consider the set $P(n)$ of partitions of $n \in \mathbb{N}_0$, where for $\lambda := [\lambda_1,\ldots,\lambda_k] \vdash n$ with $k \in \{0,\ldots,n\}$ parts we let $\lambda_i := 0$ for all $i > k$, thus we may write $\lambda = [\lambda_1,\ldots,\lambda_n]$. Then $P(n)$ is partially ordered by **dominance** $\trianglelefteq$, where $\lambda = [\lambda_1,\ldots,\lambda_n] \vdash n$ is said to **dominate** $\mu = [\mu_1,\ldots,\mu_n] \vdash n$ if $\sum_{i=1}^k \mu_i \leq \sum_{i=1}^k \lambda_i$, for all $k \in \{1,\ldots,n\}$:

It is immediate that reflexivity and transitivity hold. Moreover, from $\mu \leq \lambda$ and $\lambda \leq \mu$ we get $\sum_{i=1}^{k} \mu_i = \sum_{i=1}^{k} \lambda_i$, for all $k \in \{1, \ldots, n\}$, which successively entails $\mu_i = \lambda_i$, for all $k \in \{1, \ldots, n\}$. Hence we have antisymmetry as well, showing that dominance $\trianglelefteq$ indeed is a partial order. Actually, $P(n)$ is a lattice with respect to the partial order $\trianglelefteq$. $\sharp$

We describe the associated covering relation: Given $\mu \vdash n$, we have $\mu \lessdot \lambda$ if and only if $\lambda = [\mu_1, \ldots, \mu_{r-1}, \mu_r + 1, \mu_{r+1}, \ldots, \mu_{s-1}, \mu_s - 1, \mu_{s+1}, \ldots, \mu_n]$, where $1 \leq r < s \leq n$ such that $\mu_s > \mu_{s+1}$ and $\mu_{r-1} > \mu_r$, if $r > 1$, and such that either $s = r + 1$, or $s > r + 1$ and $\mu_r = \mu_s$:

If $\mu \lessdot \lambda$, then let $r := \min\{i \in \{1, \ldots, n\}; \mu_i \neq \lambda_i\}$ and $s := \min\{k \in \{r + 1, \ldots, n\}; \sum_{i=1}^{k} \mu_i = \sum_{i=1}^{k} \lambda_i\}$, thus $1 \leq r < s \leq n$. Hence we have $\mu_r < \lambda_r$, and $\lambda_r \leq \lambda_{r-1} = \mu_{r-1}$ if $r > 1$, as well as $\mu_s > \lambda_s \geq \lambda_{s+1} \geq \mu_{s+1}$. This yields $\mu \triangleleft \nu := [\mu_1, \ldots, \mu_{r-1}, \mu_r + 1, \mu_{r+1}, \ldots, \mu_{s-1}, \mu_s - 1, \mu_{s+1}, \ldots, \mu_n] \trianglelefteq \lambda$, hence $\nu = \lambda$. It remains to show $\mu_r = \mu_s$ whenever $s > r + 1$: Assume to the contrary that $\mu_r > \mu_s$, and let $r < t := \min\{i \in \{r + 1, \ldots, s\}; \mu_{i-1} > \mu_i\} \leq s$. If $t = s$ then $\mu \triangleleft [\mu_1, \ldots, \mu_{r-1}, \mu_r + 1, \mu_{r+1}, \ldots, \mu_{s-2}, \mu_{s-1} - 1, \mu_s, \ldots, \mu_n] \triangleleft \nu = \lambda$, while if $t < s$ then $\mu \triangleleft [\mu_1, \ldots, \mu_r, \ldots, \mu_{t-1}, \mu_t + 1, \mu_{t+1}, \ldots, \mu_{s-1}, \mu_s - 1, \mu_{s+1}, \ldots, \mu_n] \triangleleft \nu = \lambda$, a contradiction.

Let conversely λ be as asserted, and let $\nu = [\nu_1, \ldots, \nu_n] \vdash n$ such that $\mu \triangleleft \nu \trianglelefteq \lambda$. Hence for $i \not\in \{r, \ldots, s\}$ we have $\nu_i = \mu_i$. Thus if $s = r + 1$ we conclude $\nu_r = \mu_r + 1$ and $\nu_{r+1} = \mu_{r+1} - 1$, thus $\nu = \lambda$. If $s > r + 1$ and hence $\mu_r = \mu_s$, then there are $r \leq r' < s' \leq s$ such that $\nu_i = \mu_i$ for $i \not\in \{r', s'\}$ as well as $\nu_{r'} = \mu_{r'} + 1$ and $\nu_{s'} = \mu_{s'} - 1$. Since $\mu_{r'} = \nu_{r'} - 1 \leq \nu_{r'-1} - 1 = \mu_{r'-1} - 1 < \mu_{r'-1}$, whenever $r' > 1$, and $\mu_{s'} = \nu_{s'} + 1 \geq \nu_{s'+1} + 1 = \mu_{s'+1} + 1 > \mu_{s'+1}$, this implies $r' = r$ and $s' = s$, hence $\nu = \lambda$ in this case as well. $\sharp$

For example, we have $\lambda \trianglelefteq [n]$ and $[1^n] \trianglelefteq \lambda$ for all $\lambda \vdash n$ and $n \in \mathbb{N}_0$, and $[n-1, 1] \lessdot [n]$ for $n \geq 2$, and $[1^3] \lessdot [2, 1] \lessdot [3]$ and $[1^4] \lessdot [2, 1^2] \lessdot [2^2] \lessdot [3, 1] \lessdot [4]$ and $[1^4] \lessdot [2, 1^3] \lessdot [2^2, 1] \lessdot [3, 2] \lessdot [4, 1] \lessdot [5]$, and

$$[1^6] \lessdot [2, 1^4] \lessdot [2^2, 1^2] \lessdot \{[3, 1^3], [2^3]\} \lessdot [3, 2, 1] \lessdot \{[4, 1^2], [3^2]\} \lessdot [4, 2] \lessdot [5, 1] \lessdot [6],$$

where $\{[3, 1^3], [2^3]\}$ and $\{[4, 1^2], [3^2]\}$ are non-comparable.

b) Depicting $\lambda \vdash n$ by a Young diagram, the **conjugate** partition $\lambda' \vdash n$ is obtained by reflecting the diagram along its main diagonal. Formally, if $\lambda = [\lambda_1, \ldots, \lambda_n] \vdash n$ then letting $\lambda'_i := |\{j \in \mathbb{N}; \lambda_j \geq i\}| \in \mathbb{N}_0$ for all $i \in \mathbb{N}$, we have $\lambda'_1 \geq \cdots \geq \lambda'_n \geq 0$ and $\sum_{i=1}^{n} \lambda'_i = \sum_{j=1}^{n} |\{i \in \{1, \ldots, n\}; i \leq \lambda_j\}| = \sum_{j=1}^{n} \lambda_j = n$, hence we indeed may let $\lambda' := [\lambda'_1, \ldots, \lambda'_n] \vdash n$. Moreover, conjugating twice yields $\lambda'' \vdash n$, where $\lambda''_i = |\{j \in \mathbb{N}; \lambda'_j \geq i\}| = |\{j \in \mathbb{N}; |\{k \in \mathbb{N}; \lambda_k \geq j\}| \geq i\}| = |\{j \in \mathbb{N}; \{1, \ldots, i\} \subseteq \{k \in \mathbb{N}; \lambda_k \geq j\}| = |\{j \in \mathbb{N}; \lambda_i \geq j\}| = |\{1, \ldots, \lambda_i\}| = \lambda_i$, that is we indeed have $\lambda'' = \lambda$.

Alternatively, writing $\lambda' = [n^{a'_n}, \ldots, 1^{a'_1}] \vdash n$ in terms of multiplicities, we have $a'_i = |\{j \in \mathbb{N}; \lambda'_j = i\}| = |\{j \in \mathbb{N}; |\{k \in \mathbb{N}; \lambda_k \geq j\}| = i\}| = |\{j \in \mathbb{N}; \{k \in \mathbb{N}; \lambda_k \geq j\} = \{1, \ldots, i\}\}| = |\{j \in \mathbb{N}; \lambda_i \geq j, \lambda_{i+1} < j\}| = |\{\lambda_{i+1} + 1, \ldots, \lambda_i\}| = \lambda_i - \lambda_{i+1}$, for all $i \in \{1, \ldots, n\}$, providing the fastest way to compute conjugate partitions. For example, we have $[n]' = [1^n]$ for all $n \in \mathbb{N}_0$, and $[n-1,1]' = [2,1^{n-2}]$ for $n \geq 2$, as well as $[2^2]' = [2^2]$ and $[3,2]' = [2^2,1]$ and $[3,1^2]' = [3,1^2]$.

Then we have $\mu \trianglelefteq \lambda$ if and only if $\lambda' \trianglelefteq \mu'$: To show this, it suffices to assume to the contrary that $\mu \trianglelefteq \lambda$ but $\lambda' \ntrianglelefteq \mu'$. Then for some $k \in \mathbb{N}$ we have $\sum_{i=1}^{j} \lambda'_i \leq \sum_{i=1}^{j} \mu'_i$ for all $j \in \{1,\ldots,k-1\}$, and $\sum_{i=1}^{k} \lambda'_i > \sum_{i=1}^{k} \mu'_i$. Hence we have $\lambda'_k > \mu'_k$ and $\sum_{i=k+1}^{n} \lambda'_i < \sum_{i=k+1}^{n} \mu'_i$. Now we have $\sum_{i=k+1}^{n} \lambda'_i = \sum_{i=k+1}^{n} |\{j \in \mathbb{N}; i \leq \lambda_j\}| = \sum_{j=1}^{\lambda'_k} (\lambda_j - k)$ and similarly $\sum_{i=k+1}^{n} \mu'_i = \sum_{j=1}^{\mu'_k} (\mu_j - k)$; note that $\lambda_j \geq k$ for $j \in \{1,\ldots,\lambda'_k\}$. This implies $\sum_{j=1}^{\mu'_k} (\mu_j - k) > \sum_{j=1}^{\lambda'_k} (\lambda_j - k) \geq \sum_{j=1}^{\mu'_k} (\lambda_j - k)$, thus $\mu \ntrianglelefteq \lambda$, a contradiction.

(5.7) Stratification of the nilpotent variety. Recall that $\mathbb{C}$ is equipped with the usual metric topology, that subsets of topological spaces are equipped with induced topologies, and that direct products of topological spaces are equipped with product topologies. We consider the matrix algebra $\mathbb{C}^{n \times n}$ as a topological space. Then matrix addition and multiplication are continuous maps $\mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$, as well as is scalar multiplication $\mathbb{C} \times \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n}$.

Recall that, for $M \in \mathbb{C}^{n \times n}$, the rank $\mathrm{rk}(M) \in \mathbb{N}_0$ equals the smallest integer $k \in \mathbb{N}_0$ such that all $(i \times i)$-minors of M vanish, for all $i \in \{k+1,\ldots,n\}$. This shows that $\mathcal{R}_{\leq k} := \{M \in \mathbb{C}^{n \times n}; \mathrm{rk}(M) \leq k\} \subseteq \mathbb{C}^{n \times n}$ is a closed subset, for all $k \in \mathbb{N}_0$. In particular, $\mathcal{G} := \mathrm{GL}_n(\mathbb{C}) = \mathbb{C}^{n \times n} \setminus \mathcal{R}_{\leq n-1} = \{A \in \mathbb{C}^{n \times n}; \det(A) \neq 0\} \subseteq \mathbb{C}^{n \times n}$ is an open subset, and since group multiplication $\mathcal{G} \times \mathcal{G} \to \mathcal{G}$ and inversion $\mathcal{G} \to \mathcal{G} \colon A \mapsto A^{-1} = \det(A)^{-1} \cdot \mathrm{adj}(A)$ are continuous maps, $\mathcal{G}$ becomes a topological group. Finally, $\mathcal{G}$ acts continuously on $\mathbb{C}^{n \times n}$ by conjugation $\mathcal{G} \times \mathbb{C}^{n \times n} \to \mathbb{C}^{n \times n} \colon [A, M] \to AMA^{-1}$.

Let $\mathcal{N} := \{M \in \mathbb{C}^{n \times n}; M^n = 0\}$ be the **nilpotent variety**; hence $\mathcal{N} \subseteq \mathbb{C}^{n \times n}$ is a closed subset. The group $\mathcal{G}$ acts continuously on $\mathcal{N}$ by conjugation, whose orbits are parametrised by the Jordan normal forms of matrices with respect to the eigenvalue 0, that is block diagonal matrices $\bigoplus_{i=1}^{k} J_{\lambda_i} \in \mathcal{N}$, where $J_{\lambda_i} \in \mathbb{C}^{\lambda_i \times \lambda_i}$ is a Jordan block of size $\lambda_i \in \mathbb{N}$. Permuting the Jordan blocks such that $\lambda_1 \geq \cdots \geq \lambda_k$, from $\sum_{i=1}^{k} \lambda_i = n$ we infer that $\lambda := [\lambda_1,\ldots,\lambda_k]$ is a partition of n with $k \in \{0,\ldots,n\}$ parts, hence yielding the **stratification** $\mathcal{N} = \coprod_{\lambda \vdash n} \mathcal{N}_\lambda$, where $\mathcal{N}_\lambda \subseteq \mathcal{N}$ is the set of nilpotent matrices having Jordan normal form parametrised by $\lambda \vdash n$.

Since $\mathcal{G}$ acts continuously on $\mathcal{N}$ we conclude that, for any $\lambda \vdash n$, the closure $\overline{\mathcal{N}}_\lambda \subseteq \mathcal{N}$ of the $\mathcal{G}$-orbit $\mathcal{N}_\lambda$ is $\mathcal{G}$-invariant as well, thus is a union of $\mathcal{G}$-orbits. Hence we get an induced **closure relation** on the set $P(n)$ of all partitions of n, where for $\mu \vdash n$ we let $\mu \preceq \lambda$ if $\mathcal{N}_\mu \subseteq \overline{\mathcal{N}}_\lambda$. Hence $\preceq$ is reflexive, and since $\mathcal{N}_\mu \subseteq \overline{\mathcal{N}}_\lambda$ implies $\overline{\mathcal{N}}_\mu \subseteq \overline{\mathcal{N}}_\lambda$, it is transitive as well. Antisymmetry, saying that $\mathcal{N}_\mu \subseteq \overline{\mathcal{N}}_\lambda$ and $\mathcal{N}_\lambda \subseteq \overline{\mathcal{N}}_\mu$ already imply $\mathcal{N}_\lambda = \mathcal{N}_\mu$, is ensured as follows: By assumption both $\mathcal{N}_\lambda$ and $\mathcal{N}_\mu$ are dense in $\overline{\mathcal{N}}_\lambda = \overline{\mathcal{N}}_\mu$, hence the topological property that any $\mathcal{G}$-orbit is open in its closure entails that $\mathcal{N}_\lambda \cap \mathcal{N}_\mu \neq \emptyset$.

Actually, we are going to show that the closure relation $\preceq$ coincides with the dominance partial order $\trianglelefteq$ on $P(n)$, by only using reflexivity and transitivity of $\preceq$, so that this will also imply that $\preceq$ indeed is a partial order:

For a Jordan block $J_i \in \mathbb{C}^{i\times i}$, for some $i \in \mathbb{N}$, we have $\mathrm{rk}(J_i^k) = i-k$ for all $k \in \{0,\ldots,i\}$. Thus for $M \in \mathcal{N}_\lambda$, where $\lambda = [n^{a_n},\ldots,1^{a_1}] \vdash n$, we have $\mathrm{rk}(M^k) = \sum_{i=k+1}^n (i-k)a_i = \sum_{i=k+1}^n \sum_{j=i}^n a_j = \sum_{i=k+1}^n \sum_{j=i}^n (\lambda'_i - \lambda'_{i+1}) = \sum_{i=k+1}^n \lambda'_i$, for all $k \in \{0,\ldots,n\}$, implying $n - \mathrm{rk}(M^k) = \sum_{i=1}^k \lambda'_i$. Hence $\mathcal{N}_\lambda$ is uniquely determined by the rank sequence $[n - \sum_{i=1}^k \lambda'_i \in \mathbb{N}_0; k \in \{0,\ldots,n\}]$; note that we have $\mathrm{rk}(M^0) = n$ and $\mathrm{rk}(M^n) = 0$ anyway.

Moreover, for $\mu \vdash n$ and $N \in \mathcal{N}_\mu$ we have $\mu \trianglelefteq \lambda$ if and only if $\lambda' \trianglelefteq \mu'$, which holds if and only if $\mathrm{rk}(M^k) \geq \mathrm{rk}(N^k)$, for all $k \in \{0,\ldots,n\}$. Thus we have $N \in \mathcal{N}_{\trianglelefteq\lambda} := \coprod_{\mu\trianglelefteq\lambda} \mathcal{N}_\mu \subseteq \mathcal{N}$ if and only if $\mathrm{rk}(N^k) \leq n - \sum_{i=1}^k \lambda'_i$, that is $N^k \in \mathcal{R}_{n-\sum_{i=1}^k \lambda'_i}$, for all $k \in \{0,\ldots,n\}$. This implies that $\mathcal{N}_{\trianglelefteq\lambda} \subseteq \mathcal{N}$ is a closed subset, containing $\mathcal{N}_\lambda$, thus we have $\overline{\mathcal{N}}_\lambda \subseteq \mathcal{N}_{\trianglelefteq\lambda}$.

For the converse $\mathcal{N}_{\trianglelefteq\lambda} \subseteq \overline{\mathcal{N}}_\lambda$, we have to show that $\mu \trianglelefteq \lambda$ implies $\mathcal{N}_\mu \subseteq \overline{\mathcal{N}}_\lambda$. In order to do so, by the transitivity of the closure relation we may assume that $\mu := [\lambda_1,\ldots,\lambda_{r-1},\lambda_r - 1,\lambda_{r+1},\ldots,\lambda_{s-1},\lambda_s + 1,\lambda_{s+1},\ldots,\lambda_n] \lessdot \lambda$, for some $1 \leq r < s \leq n$. Letting $a := \lambda_r$ and $b := \lambda_s$, hence $a > b \geq 0$, we have $J_a \oplus J_b \oplus N \in \mathcal{N}_\lambda$ and $J_{a-1} \oplus J_{b+1} \oplus N \in \mathcal{N}_\mu$, where $N \in \mathbb{C}^{(n-a-b)\times(n-a-b)}$. Hence we may assume that $\lambda = [a,b] \vdash n$ and $\mu = [a-1,b+1] \vdash n$, and let

$$M_\epsilon = \left[\begin{array}{cccccc|cccc} . & & & & & & & & & \\ \epsilon & . & & & & & & & & \\ & 1 & . & & & & & & & \\ & & 1 & . & & & & & & \\ & & & \ddots & \ddots & & & & & \\ & & & & 1 & . & & & & \\ \hline 1 & & & & & & . & & & \\ & & & & & & 1 & . & & \\ & & & & & & & \ddots & \ddots & \\ & & & & & & & & 1 & . \end{array}\right] \in \mathbb{C}^{n\times n},$$

where the upper left and lower right hand corners have size $a \times a$ and $b \times b$, respectively. We show that $M_\epsilon \in \mathcal{N}_\lambda$ if $\epsilon \neq 0$, while $\lim_{\epsilon\to 0} M_\epsilon = M_0 \in \mathcal{N}_\mu$:

If $\epsilon \neq 0$ then, since $a > b$, the unit vector $e_1 \in \mathbb{C}^{n\times 1}$ has minimum polynomial $X^a \in \mathbb{C}[X]$ with respect to M_ϵ. Moreover, the unit vector $e_{a+1} \in \mathbb{C}^{n\times 1}$ has minimum polynomial $X^b \in \mathbb{C}[X]$; here for $b = 0$ we let $e_{a+1} := 0 \in \mathbb{C}^{n\times 1}$. From $\langle e_{a+1}\rangle_{M_\epsilon} = \langle e_{a+1},\ldots,e_n\rangle_{\mathbb{C}}$ and $\langle e_1\rangle_{M_\epsilon} \cap \langle e_{a+1},\ldots,e_n\rangle_{\mathbb{C}} = \{0\}$ we conclude $\mathbb{C}^{n\times 1} = \langle e_1\rangle_{M_\epsilon} \oplus \langle e_{a+1}\rangle_{M_\epsilon}$, hence M_ϵ has Jordan normal form $J_a \oplus J_b$.

If $\epsilon = 0$ then $e_2 \in \mathbb{C}^{n\times 1}$ and $e_1 \in \mathbb{C}^{n\times 1}$ have minimum polynomials $X^{a-1} \in \mathbb{C}[X]$ and $X^{b+1} \in \mathbb{C}[X]$, respectively, with respect to M_0. From $\langle e_2\rangle_{M_0} = \langle e_2,\ldots,e_a\rangle_{\mathbb{C}}$ and $\langle e_1\rangle_{M_0} = \langle e_1, e_{a+1},\ldots,e_n\rangle_{\mathbb{C}}$ we conclude $\mathbb{C}^{n\times 1} = \langle e_2\rangle_{M_0} \oplus \langle e_1\rangle_{M_0}$, hence M_0 has Jordan normal form $J_{a-1} \oplus J_{b+1}$. ♯

6 Modular lattices

(6.1) Modular lattices. The exposition is taken from on [19]. Let X be a lattice, having a zero element $\underline{0} \in X$, which is **locally chain-finite**, that is all chains in any interval $[\underline{0}, x] \subseteq X$ are finite; in particular, in any interval $[x, y] \subseteq X$ there are saturated chains.

Then X is called **modular** if for all $x, y, z \in X$ such that $z \leq x$ we have the **modular law** $x \wedge (y \vee z) = (x \wedge y) \vee z$. Then, for $x, y \in X$ the map $[x \wedge y, x] \rightarrow [y, x \vee y] \colon z \mapsto z \vee y$ is an isomorphism of lattices, with inverse $[y, x \vee y] \rightarrow [x \wedge y, x] \colon z \mapsto z \wedge x$: Indeed, for $z \in [x \wedge y, x]$ we have $(z \vee y) \wedge x = z \vee (y \wedge x) = z$, and for $z \in [y, x \vee y]$ we have $(z \wedge x) \vee y = z \wedge (x \vee y) = z$.

An element $z \in X \setminus \{\underline{0}\}$ is called **join-irreducible** or **local**, if whenever there are $x, y \in X$ such that $x \vee y = z$ then we already have $x = z$ or $y = z$. Hence $z \in X \setminus \{\underline{0}\}$ is join-irreducible if and only if $|\{y \in X; y \lessdot z\}| = 1$. Let $\mathcal{L} \subseteq X$ be the set of join-irreducible elements of X; note that $\mathcal{L}$ contains all atoms of X.

Similarly, if $\underline{1} \in X$ is a one element, an element $z \in X \setminus \{\underline{1}\}$ is called **meet-irreducible** or **co-local**, if whenever there are $x, y \in X$ such that $x \wedge y = z$ then we already have $x = z$ or $y = z$. Hence $z \in X \setminus \{\underline{1}\}$ is meet-irreducible if and only if $|\{y \in X; y \lessdot z\}| = 1$. Let $\mathcal{L}^* \subseteq X$ be the set of meet-irreducible elements of X; note that $\mathcal{L}^*$ contains all co-atoms of X.

(6.2) Length and rank. a) Let X be a modular lattice. Given an interval $[x, y] \subseteq X$, all saturated chains in $[x, y]$ have one and the same length; in particular, if X has a one element, then X is graded:

Let $x = x_0 \lessdot x_1 \lessdot \cdots \lessdot x_{s-1} \lessdot x_s = y$ and $x = x'_0 \lessdot x'_1 \lessdot \cdots \lessdot x'_{t-1} \lessdot x'_t = y$ be saturated chains, where $s, t \in \mathbb{N}_0$. Since $s = 0$ if and only if $x = y$ if and only if $t = 0$, and $s = 1$ if and only if $x \lessdot y$ if and only if $t = 1$, we may assume that $s \geq t \geq 2$, and proceed by induction on s. If $x_{s-1} = x'_{t-1}$ then we have $s = t$ by induction. If $x_{s-1} \neq x'_{t-1}$ then we have $x_{s-1} \vee x'_{t-1} = y$, thus for $z := x_{s-1} \wedge x'_{t-1}$ we get $[z, x_{s-1}] \cong [x'_{t-1}, y]$ and $[z, x'_{t-1}] \cong [x_{s-1}, y]$, hence $z \lessdot x_{s-1}$ and $z \lessdot x'_{t-1}$. Now extending a saturated chain $x = x''_0 < x''_1 < \cdots < x''_{r-1} < x''_r = z$ to x_{s-1} and x'_{t-1}, respectively, by induction we get $s - 1 = r + 1 = t - 1$. $\sharp$

Hence there is a well-defined **Jordan-Dedekind length** map $l \colon X \rightarrow \mathbb{N}_0$ defined by $l(x) := l(\underline{0}, x)$ for $x \in X$. As in the case of graded posets, we have $l(\underline{0}) = 0$, and $l(y) = l(x) + 1$ whenever $x \lessdot y \in X$; in particular we have $l(x) = 0$ if and only if $x = \underline{0}$. Moreover, for $x \leq y \in X$ we have $l(x, y) = l(y) - l(x)$, and for $x, y \in X$ we have the **modular equality** $l(x \vee y) - l(y) = l(x) - l(x \wedge y)$.

b) Each element $x \in X$ is the join of finitely many join-irreducible elements.

For these representations of x we have the following **exchange property**: Let $x = \bigvee_{i=1}^{r} x_i = \bigvee_{j=1}^{s} x'_j$, where $r, s \in \mathbb{N}_0$ and $x_i, x'_j \in \mathcal{L}$. Then for any $k \in \{1, \ldots, r\}$ there is $j = j_k \in \{1, \ldots, s\}$ such that $x = x'_j \vee \bigvee_{i \neq k} x_i$:

Let $y := \bigvee_{i \neq k} x_i \in X$, where we may assume that $y < x$. Hence we have $[y, x] =$

$[y, y \vee x_k] \cong [y \wedge x_k, x_k]$. Since $x_k \in \mathcal{L}_{[y \wedge x_k, x_k]}$, we infer that $x = y \vee x_k \in \mathcal{L}_{[y,x]}$. Now, for all $j \in \{1, \ldots, s\}$ we have $y \leq y \vee x'_j \leq x$, hence from $x = \bigvee_{j=1}^{s} (y \vee x'_j)$ we infer that there is $j \in \{1, \ldots, s\}$ such that $x = y \vee x'_j$. ♯

Hence we conclude that, if $x = \bigvee_{i=1}^{r} x_i \in X$ is irredundant, where $x_i \in \mathcal{L}$ for all $i \in \{1, \ldots, r\}$, then the number $r \in \mathbb{N}_0$ is independent of the particular choice of the irredundant representation, giving rise to the **rank** map $r \colon X \to \mathbb{N}_0$:

Let $x = \bigvee_{i=1}^{r} x_i = \bigvee_{j=1}^{s} x'_j$ be irredundant, where $r, s \in \mathbb{N}_0$ and $x_i, x'_j \in \mathcal{L}$. Then iterating the above exchange procedure yields $x = \bigvee_{i=1}^{r} x'_{j_i}$, hence the irredundancy of $\bigvee_{j=1}^{s} x'_j$ entails $\{j_1, \ldots, j_r\} = \{1, \ldots, s\}$, hence $s \leq r$. Similarly we conclude $r \leq s$, hence equality holds. ♯

Note that since $\underline{0} < x_1 < x_1 \vee x_2 < \cdots < \bigvee_{i=1}^{r} x_i = x$ we have $r(x) = r \leq l(x)$, and that $r(x) = 0$ if and only if $x = \underline{0}$.

(6.3) Complemented lattices. Let X be a modular lattice having a one element $\underline{1}$. Then X is called **complemented**, if for all $x \in X$ there is a **complement** $y \in X$, that is we have $x \vee y = \underline{1}$ and $x \wedge y = \underline{0}$. Note that hence X is complemented if and only if the dual of X is.

a) The following assertions are equivalent: **i)** The lattice X is complemented. **ii)** The one element $\underline{1} \in X$ is a join of atoms, that is X is **semi-simple**. **iii)** Any element $x \in X$ is a join of atoms. **iv)** The set $\mathcal{L} \subseteq X$ of join-irreducible elements consists of (all) atoms.

This is seen as follows: In order to show i)⇒ii), assume that $x := \bigvee_{\underline{0} \lessdot z} z \neq \underline{1}$, and let $\underline{0} \neq y \in X$ be a complement of x. Hence there is an atom $\underline{0} \lessdot z \in X$ such that $z \leq y$, thus $x \wedge y \neq \underline{0}$, a contradiction.

Conversely, to show ii)⇒i), let $x \in X$ and let $y \in X$ be maximal such that $x \wedge y = \underline{0}$. Assume that $x \vee y \neq \underline{1}$, thus there is an atom $\underline{0} \lessdot z \in X$ such that $z \not\leq (x \vee y)$, hence $z \wedge (x \vee y) = \underline{0}$. By the choice of y we have $x \wedge (y \vee z) \neq \underline{0}$, and thus there is an atom $\underline{0} \lessdot w \in X$ such that $w \leq x \wedge (y \vee z)$. From $x \wedge y = \underline{0}$ we infer that $w \not\leq y$, hence since $[y, y \vee z] \cong [y \wedge z, z] = [\underline{0}, z]$ we conclude that $y \vee w = y \vee z$, entailing $x \vee y \vee z = x \vee y \vee w = x \vee y$, a contradiction.

Next, as for the implication i)⇒iii), in view of the implication i)⇒ii) applied to the interval $[\underline{0}, x] \subseteq X$, it suffices to show that all intervals $[x, y] \subseteq X$ are complemented: Indeed, let $x \leq z \leq y$, let $v \in X$ be a complement of z, and let $w := (x \vee v) \wedge y = x \vee (v \wedge y) \in X$; hence we have $x \leq w \leq y$. This entails $z \wedge w = z \wedge (v \vee x) = (z \wedge v) \vee x = x$ and $z \vee w = z \vee (v \wedge y) = (z \vee v) \wedge y = y$, showing that w is a complement of z in $[x, y]$.

The implication iii)⇒ii) is trivial. The implication iii)⇒iv) follows from observing that only the atoms are join-irreducible. The implication iv)⇒iii) follows from observing that any element of X is a join of join-irreducible elements. ♯

b) We give another characterisation, in terms of the length and rank functions: The lattice X is complemented if and only if $r(\underline{1}) = l(\underline{1}) \in \mathbb{N}_0$, which holds if and only if the rank and length functions of X coincide:

The second assertion follows from the first one by recalling that if X is complemented then all intervals of X are as well. To show the first assertion, we proceed by induction on $r := r(\underline{1}) \in \mathbb{N}_0$, the cases $r = 0$ and $r = 1$ being trivial, let $r \geq 2$. Let $\underline{1} = \bigvee_{i=1}^{r} x_i \in X$ be irredundant, where $x_i \in \mathcal{L}$ for all $i \in \{1, \ldots, r\}$, and let $x := \bigvee_{i \neq j} x_i \in X$, for some $j \in \{1, \ldots, r\}$.

Now, if X is complemented, then the interval $[\underline{0}, x] \subseteq X$ is as well, hence by induction we have $l(x) = r(x) = r - 1$; since $\underline{0} \lessdot x_j \in X$ is an atom we have $x \wedge x_j = \underline{0}$, and hence $l(x_j) = 1$ implies $r = l(x) + l(x_j) = l(x \vee x_j) + l(x \wedge x_j) = l(\underline{1}) + l(\underline{0}) = l(\underline{1})$. If conversely $r = l(\underline{1})$, then we have $r - 1 = r(x) \leq l(x) < l(\underline{1}) = r$, thus $r(x) = l(x) = r - 1$, and hence by induction the interval $[\underline{0}, x] \subseteq X$ is complemented, thus all the $x_i \in X$, for $i \neq j$, are atoms; since $j \in \{1, \ldots, r\}$ was arbitrary and $r \geq 2$, we conclude that $\underline{1} \in X$ is a join of atoms. ♯

(6.4) The center. a) Let X be a modular lattice. Then the **center** $\mathcal{Z}(X)$ of X is defined as the set of all $x \in X$ such that there is $x' \in X$ such that X can be written is an **inner** direct product $X \cong [\underline{0}, x] \times [\underline{0}, x']$, that is the isomorphism is induced by the natural embeddings of the intervals $[\underline{0}, x]$ and $[\underline{0}, x']$ into X.

Then an element $z \in X$ decomposes into $z = z' \vee z''$ for uniquely defined elements $\underline{0} \leq z' \leq x$ and $\underline{0} \leq z'' \leq x'$. In particular, for $w \in \mathcal{L}$ we conclude that either $w = w' \leq x$ and $w'' = \underline{0}$, or $w = w'' \leq x'$ and $w' = \underline{0}$. Hence we have $z = \bigvee\{w \in \mathcal{L}; w \leq z\} = \bigvee\{w \in \mathcal{L}; w \leq z \wedge x\} \vee \bigvee\{w \in \mathcal{L}; w \leq z \wedge x'\} = (z \wedge x) \vee (z \wedge x')$, thus $z' = z \wedge x$ and $z'' = z \wedge x'$. In other words, the inner decomposition is given by the isomorphism $X \to [\underline{0}, x] \times [\underline{0}, x'] \colon z \mapsto [z \wedge x, z \wedge x']$.

In particular, given $x \in \mathcal{Z}(X)$, then $x' \in X$ as above is uniquely defined: If similarly $X \cong [\underline{0}, x] \times [\underline{0}, x'']$ is an inner decomposition for some $x'' \in X$, then since $x' \wedge x = \underline{0} = x'' \wedge x$ we have $x'' = x' \wedge x'' = x'$.

We have $\mathcal{Z}(X) \neq \emptyset$ if and only if X has a one element $\underline{1}$. In this case, we have $\{\underline{0}, \underline{1}\} \subseteq \mathcal{Z}(X)$; and if $\mathcal{Z}(X) = \{\underline{0}, \underline{1}\}$, then X is called **indecomposable**, otherwise X is called **decomposable**.

b) Let X have a one element $\underline{1}$. Then $\mathcal{Z}(X)$ gives rise to a canonical decomposition of X. In order to derive this, we show next that $\mathcal{Z}(X)$ is closed with respect to taking meets, that is if $x, y \in \mathcal{Z}(X)$ then we also have $x \wedge y \in \mathcal{Z}(X)$:

Let $[\underline{0}, x] \times [\underline{0}, x'] \cong X \cong [\underline{0}, y] \times [\underline{0}, y']$, where $x', y' \in X$, and let $x_{11} := x \wedge y$, $x_{12} := x \wedge y'$, $x_{21} := x' \wedge y$ and $x_{22} := x' \wedge y'$. We consider the order-preserving maps $\sigma \colon \widehat{X} := \prod_{i,j \in \{1,2\}} [\underline{0}, x_{ij}] \to X \colon [z_{ij}; i, j \in \{1, 2\}] \mapsto \bigvee_{i,j \in \{1,2\}} z_{ij}$ and $\tau \colon X \to \widehat{X} \colon z \mapsto [z \wedge x_{ij}; i, j \in \{1, 2\}]$.

For $z \in X$ we have $\bigvee_{i,j \in \{1,2\}} (z \wedge x_{ij}) = ((z \wedge x) \wedge y) \vee ((z \wedge x) \wedge y') \vee ((z \wedge x') \wedge y) \vee ((z \wedge x') \wedge y') = (z \wedge x) \vee (z \wedge x') = z$, hence $\sigma\tau = \mathrm{id}_X$. Conversely, for $[z_{ij}; i, j \in \{1, 2\}] \in \widehat{X}$ we have $(z_{12} \vee z_{21} \vee z_{22}) \wedge x_{11} \leq (x_{12} \vee x_{21} \vee x_{22}) \wedge x_{11} = (x_{12} \vee x') \wedge (x \wedge y) = (x_{12} \vee (x' \wedge x)) \wedge y = (x \wedge y') \wedge y = \underline{0}$, implying

$(\bigvee_{i,j\in\{1,2\}} z_{ij}) \wedge x_{11} = z_{11} \vee ((z_{12} \vee z_{21} \vee z_{22}) \wedge x_{11}) = z_{11}$; similarly we argue for x_{12}, x_{21} and x_{22}. Hence we have $\tau\sigma = \mathrm{id}_{\widehat{X}}$.

Thus σ and τ are a pair of mutually inverse isomorphisms of lattices, showing that we have a decomposition $X \cong [\underline{0}, x \wedge y] \times [\underline{0}, x \wedge y'] \times [\underline{0}, x' \wedge y] \times [\underline{0}, x' \wedge y]$, in particular we have $x \wedge y = x_{11} \in \mathcal{Z}(X)$. $\sharp$

Now, if $\{z_1, \ldots, z_d\}$, for some $d \in \mathbb{N}_0$, are the minimal elements of $\mathcal{Z}(X) \setminus \{\underline{0}\}$, then $X \cong \prod_{i=1}^d [\underline{0}, z_i]$ is the unique decomposition into non-trivial indecomposable intervals:

By the above, for $y \leq x \in \mathcal{Z}(X)$ we have the inner decomposition $[\underline{0}, x] \cong [\underline{0}, y] \times [\underline{0}, x \wedge y']$, hence $y \in \mathcal{Z}([\underline{0}, x]) \subseteq \mathcal{Z}(X)$. Thus for $x \in \mathcal{Z}(X) \setminus \{\underline{0}\}$ the interval $[\underline{0}, x]$ is indecomposable if and only if x is minimal in $\mathcal{Z}(X) \setminus \{\underline{0}\}$. Since for $x \in \mathcal{Z}(X)$ we have $l(\underline{1}) = l(x) + l(x')$, by induction on $l(x)$ we conclude that $X \cong \prod_{i=1}^d [\underline{0}, z_i]$, for some $d \in \mathbb{N}_0$ and certain minimal $z_i \in \mathcal{Z}(X) \setminus \{\underline{0}\}$. Moreover, for any $x \in X$ we have $x = \bigvee_{i=1}^d (x \wedge z_i)$, hence if $x \in \mathcal{Z}(X) \setminus \{\underline{0}\}$ is minimal then we have $x = x \wedge z_i = z_i$ for a unique $i \in \{1, \ldots, d\}$, implying that $\{z_1, \ldots, z_d\}$ indeed encompasses all minimal elements of $\mathcal{Z}(X) \setminus \{\underline{0}\}$. $\sharp$

In particular we conclude that $\mathcal{Z}(X)$ is a finite complemented lattice: For any $x \in \mathcal{Z}(X)$ we have $[\underline{0}, x] \cong \prod_{i=1}^d [\underline{0}, x \wedge z_i] \cong \prod_{i; z_i \leq x} [\underline{0}, z_i]$, thus $x = \bigvee \{z_i; z_i \leq x\}$, showing that x is the join of certain of the finitely many atoms in $\mathcal{Z}(X)$. $\sharp$

(6.5) Radicals. a) Let X be a modular lattice. For $\underline{0} \neq x \in X$ the element $x_* := \bigwedge \{y \in X; y \lessdot x\} \in X$ is called the **radical** of x; we let $\underline{0}_* := \underline{0}$. In particular, we have $\mathcal{L} = \{x \in X; x_* \lessdot x\}$. We collect a few properties:

i) For $x \in X$, the radical x_* is **small** in x, that is for $y \leq x \in X$ such that $y \vee x_* = x$ we already have $y = x$: Indeed, assuming that $y < x$, there is $y \leq z \lessdot x$, hence $y \vee x_* \leq z$, a contradiction.

ii) We have $x_* = \bigwedge \{y \in X; [y, x] \text{ complemented}\}$: We may assume that $x \neq \underline{0}$. From $x_* = \bigwedge \{z \in X; z \lessdot x\}$, considering the dual of the interval $[x_*, x]$, we infer that $[x_*, x]$ is complemented, hence $x_* \subseteq \bigwedge \{y \in X; [y, x] \text{ complemented}\}$. Conversely, if for $y \leq x \in X$ the interval $[y, x]$ is complemented, then again considering duals, we have $x_* = \bigwedge \{z \in X; z \lessdot x\} \leq \bigwedge \{z \in X; y \leq z \lessdot x\} = y$.

iii) We have $r(x) = r_{[x_*, x]}(x)$: Let $x = \bigvee_{i=1}^r x_i$ be irredundant, where $r = r(x) \in \mathbb{N}_0$ and $x_i \in \mathcal{L}$. Hence $x_i \not\leq x_*$, and since $[x_*, x_i \vee x_*] \cong [x_i \wedge x_*, x_i]$ is complemented, we have $x_* \lessdot x_i \vee x_*$, that is $x_i \vee x_* \in \mathcal{L}_{[x_*, x]}$. Assume that $x = \bigvee_{i \in \mathcal{I}} (x_i \vee x_*)$, for some $\mathcal{I} \subset \{1, \ldots, r\}$, then we also have $x = \bigvee_{i \in \mathcal{I}} x_i$, a contradiction. Thus $x = \bigvee_{i=1}^r (x_i \vee x_*)$ is irredundant in $[x_*, x]$, hence $r_{[x_*, x]}(x) = r$.

iv) Finally, for $x, y \in X$ we have $(x \vee y)_* = x_* \vee y_*$: Let $x = \bigvee_{i=1}^r x_i$ and $y = \bigvee_{j=1}^s y_j$ be irredundant, where $r, s \in \mathbb{N}_0$ and $x_i, y_j \in \mathcal{L}$. Since $[x_* \vee y_*, x_i \vee (x_* \vee y_*)] \cong [x_i \wedge (x_* \vee y_*), x_i]$ is complemented, we have either $x_i \leq x_* \vee y_*$ or $x_* \vee y_* \lessdot x_i \vee (x_* \vee y_*)$. A similar statement holds for the y_j, and hence $x \vee y$ is a join of atoms of $[x_* \vee y_*, x \vee y]$. Thus we have $(x \vee y)_* \leq x_* \vee y_*$.

Conversely, assume that $x_* \not\leq (x \vee y)_*$. Then $[(x \vee y)_* \wedge x, x] \cong [(x \vee y)_*, x \vee (x \vee y)_*]$ is not complemented. Since $(x \vee y)_* \leq x \vee (x \vee y)_* \leq x \vee y$, this a contradiction. Hence we have $x_* \leq (x \vee y)_*$, and similarly $y_* \leq (x \vee y)_*$. $\sharp$

b) The above considerations lead to the following subsets of X: For $r \in \mathbb{N}_0$ let $X_r := \{x \in X; r(x) = r\} = \{x \in X; r_{[x_*,x]}(x) = r\} \subseteq X$ and $\mathcal{L}_r := \{x \in X_r; [x_*, x] \text{ indecomposable}\} \subseteq X_r$. Hence in particular we have $X_0 = \mathcal{L}_0 = \{\underline{0}\}$ and $X_1 = \mathcal{L}_1 = \mathcal{L}$.

7 Benson–Conway Theorem

We have now presented the general theory of modular lattices needed to proceed towards the main result of this section, given in (7.3). We need a few more specially tailored notions, inspired by [14]:

(7.1) Dotted-lines. a) Let X be a modular lattice. For an element $z \in X$ we have $z \in X_2$, that is $r(z) = 2$, if and only if there are $x_1, x_2 \in \mathcal{L}$ such that $x_1 \not\leq x_2 \not\leq x_1$ and $z = x_1 \vee x_2$.

In this case, we have $l_{[z_*,z]}(z) = r_{[z_*,z]}(z) = 2$. Hence the elements of $[z_*, z]$ are its zero element z_*, its one element z, and its atoms $z_* \lessdot z_i \lessdot z$, for $i \in \mathcal{I}_z$, where $\mathcal{I}_z$ is a suitable index set. In particular, since $z_* < z$ is small, for $z_i := x_i \vee z_* \in X$ we have $z_* \lessdot z_i \lessdot z$ and $z_1 \neq z_2$, hence $|\mathcal{I}_z| \geq 2$.

If $|\mathcal{I}_z| = 2$, then we have $[z_*, z] \cong [z_*, z_1] \times [z_*, z_2]$, where $z_1 \neq z_* \neq z_2$, hence $[z_*, z]$ is decomposable. Conversely, let $[z_*, z] \cong [z_*, z'] \times [z_*, z'']$, where $z' \neq z_* \neq z''$. Since $l_{[z_*,z]}(z) = 2$, we have $l_{[z_*,z]}(z') = l_{[z_*,z]}(z'') = 1$, hence both $z_* \lessdot z' \lessdot z$ and $z_* \lessdot z'' \lessdot z$, and thus $|\mathcal{I}_z| = 2$. Hence we conclude that $[z_*, z]$ is indecomposable if and only if $|\mathcal{I}_z| \geq 3$, and we have $\mathcal{L}_2 = \{z \in X_2; |\mathcal{I}_z| \geq 3\}$.

Now, given $z \in \mathcal{L}_2$, a set $\mathcal{D} = \{x_i \in \mathcal{L}; i \in \mathcal{I}_z\}$, such that $x_i \vee z_* = z_i \in X$ for all $i \in \mathcal{I}_z$, is called a **dotted-line** for z; in particular, the $x_i \in \mathcal{D}$ are pairwise distinct and thus $|\mathcal{D}| = |\mathcal{I}_z| \geq 3$. Actually, dotted-lines always exist: Since each z_i is the join of the join-irreducible elements it contains, and since $z_* \lessdot z_i$, we may choose $x_i \in \mathcal{L}$ such that $x_i \leq z_i$ and $x_i \not\leq z_*$, entailing $x_i \vee z_* = z_i$.

b) We collect a few properties of dotted-lines: Let $\mathcal{D} = \{x_i \in \mathcal{L}; i \in \mathcal{I}_z\}$ be a dotted-line for $z \in \mathcal{L}_2$. Then we have $x_i \not\leq x_j$; moreover, we have $x_i \vee x_j = z$, for all $i \neq j \in \mathcal{I}_z$, and $\mathcal{D} \subseteq \mathcal{L}$ is maximal having this property. In particular, $\bigvee \mathcal{D} = z$ is well-defined, even if $\mathcal{D}$ is infinite.

For all $i \in \mathcal{I}_z$ let $z_i = x_i \vee z_* \lessdot z$. As $z_i \not\leq z_j$, for all $i \neq j \in \mathcal{I}_z$, we also have $x_i \not\leq x_j$. Moreover, since $z = z_i \vee z_j = x_i \vee x_j \vee z_*$ we have $x_i \vee x_j = z$.

Assume that there is $x_0 \in \mathcal{L} \setminus \mathcal{D}$ such that $x_0 \vee x_i = z$ for all $i \in \mathcal{I}_z$, and let $z_0 := x_0 \vee z_*$. Assuming $z_0 = z_*$ yields $z = x_0 \vee x_i = x_i \vee z_* = x_i$, for all $i \in \mathcal{I}_z$, a contradiction; and assuming $z_0 = z$ yields $z = x_0 \vee z_* = x_0$, a contradiction. Hence we conclude that $z_* \lessdot z_0 \lessdot z$. Thus there is $i \in \mathcal{I}_z$ such that $z_0 = z_i$, hence $x_0 \vee z_* = x_i \vee z_*$. This entails $z = x_0 \vee x_i \vee z_* = x_0 \vee z_* = x_0$, a contradiction. $\sharp$

c) We proceed to give a characterisation of dotted-lines, which actually is the original definition given in [14]: Let $\mathcal{D} = \{x_i \in \mathcal{L}; i \in \mathcal{I}\} \subseteq \mathcal{L}$, where $\mathcal{I}$ is an index set such that $|\mathcal{I}| \geq 3$, the $x_i \in \mathcal{L}$ are pairwise distinct such that $x_i \vee x_j \in X$ is independent of the choice of $i \neq j \in \mathcal{I}$, and $\mathcal{D} \subseteq \mathcal{L}$ is maximal having this property. Then $z := \bigvee \mathcal{D} \in \mathcal{L}_2 \subseteq X$ is well-defined and $\mathcal{D}$ is a dotted-line for z:

Assume that we have $x_i < x_j$, for some $i \neq j \in \mathcal{I}$, and let $k \in \mathcal{I}$ such that $i \neq k \neq j$. Hence we have $x_i \vee x_k = x_i \vee x_j = x_j \in \mathcal{L}$, implying that $x_i = x_j$ or $x_k = x_j$, a contradiction. Thus we have $x_i \not\leq x_j$ for all $i \neq j \in \mathcal{I}$. Hence for $z := x_i \vee x_j = \bigvee \mathcal{D}$ we have $r(z) = l_{[z_*, z]}(z) = 2$, that is $z \in X_2$.

Let $z_i := x_i \vee z_*$ for all $i \in \mathcal{I}$. Assuming $z_i = z_*$ yields $z = x_i \vee x_j = x_j \vee z_* = x_j$, for all $i \neq j \in \mathcal{I}$, a contradiction; and assuming $z_i = z$ yields $z = x_i \vee z_* = x_i$, a contradiction. Hence we conclude that $z_* \lessdot z_i \lessdot z$. Assume that $z_i = z_j$ for some $i \neq j \in \mathcal{I}$, then we have $z = x_i \vee x_j \vee z_* = z_i \vee z_j = z_i$, a contradiction. Thus we have $\mathcal{I} \subseteq \mathcal{I}_z$, and hence $z \in \mathcal{L}_2$.

Assume that there is $k \in \mathcal{I}_z \setminus \mathcal{I}$, and choose $x_k \in \mathcal{L}$ such that $z_k = x_k \vee z_*$. Hence for all $i \in \mathcal{I}$ we have $x_k \neq x_i$, and $x_i \vee x_k \vee z_* = z_i \vee z_k = z$ and thus $x_i \vee x_k = z$, contradicting the maximality property. Hence we have $\mathcal{I} = \mathcal{I}_z$. $\sharp$

(7.2) Completeness. a) Let X be a modular lattice, and let $\mathcal{X} \subseteq \mathcal{L}$ be an ideal. Then $\mathcal{X}$ is called **complete**, if $\mathcal{X}$ has an upper bound in X, and for each dotted-line $\mathcal{D} \subseteq \mathcal{L}$, for any $z \in \mathcal{L}_2$, fulfilling $|\mathcal{D} \cap \mathcal{X}| \geq 2$ we already have $\mathcal{D} \subseteq \mathcal{X}$.

Let $X(\mathcal{L})$ be the partially ordered set of complete ideals of $\mathcal{L}$, where the partial order is given by set-theoretic inclusion. Hence $X(\mathcal{L})$ is closed under taking set-theoretic intersections, and thus becomes a lattice by letting $\mathcal{X} \wedge \mathcal{X}' := \mathcal{X} \cap \mathcal{X}'$ and $\mathcal{X} \vee \mathcal{X}' := \bigwedge \{\mathcal{Y} \in X(\mathcal{L}); \mathcal{X} \cup \mathcal{X}' \subseteq \mathcal{Y}\} \in X(\mathcal{L})$. Moreover, by the local chain-finiteness of X, an ideal of $\mathcal{L}$ having a bound in X has only finite chains, hence $X(\mathcal{L})$ is locally chain-finite as well.

b) For each $z \in \mathcal{L}_2$ choose a dotted-line $\mathcal{D}_z \subseteq \mathcal{L}$ for z. Then an ideal $\mathcal{X} \subseteq \mathcal{L}$ is called **weakly complete** with respect to $\{\mathcal{D}_z; z \in \mathcal{L}_2\}$, if $\mathcal{X}$ has an upper bound in X, and for each $z \in \mathcal{L}_2$ such that $|\mathcal{D}_z \cap \mathcal{X}| \geq 2$ we already have $\mathcal{D}_z \subseteq \mathcal{X}$.

The notion of weak completeness is due to the observation that in general there is more than one dotted-line for a given element of $\mathcal{L}_2$; typically there are many, for example in the applications in [18]. We will show in (7.4) that in order to decide whether a given ideal of $\mathcal{L}$ is complete, it suffices to check for weak completeness, reducing considerably the necessary amount of checking.

c) If X is finite, then the complete ideals of $\mathcal{L}$ are found as follows: The principal ideals of $\mathcal{L}$, being in bijection with $\mathcal{L}$, are complete. Taking these as initialisation, we iterate the following procedure: Picking in turn any of the complete ideals already found, we add a further ideal generator from $\mathcal{L}$, and then determine the smallest complete overideal of $\mathcal{L}$, by iteratively completing with respect to dotted-lines and the ideal property.

For example, let X be the modular lattice whose Hasse diagram is depicted in

Table 10: A strange modular lattice.

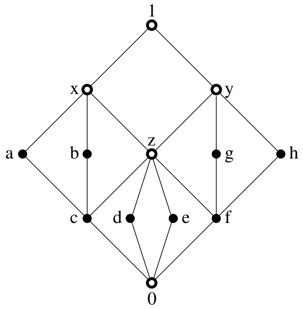


Table 10: We have $\mathcal{L} = \{a,b,c,d,e,f,g,h\}$ and $\mathcal{L}_2 = \{x,y,z\}$ and $X_2 \backslash \mathcal{L}_2 = \{\underline{1}\}$, where $|\mathcal{I}_x| = 3 = |\mathcal{I}_y|$ and $|\mathcal{I}_z| = 4$. The unique dotted-line for z is $\{c,d,e,f\}$, while x has three dotted-lines $\{\{a,b,d\},\{a,b,e\},\{a,b,f\}\}$, and y also has three dotted-lines $\{\{c,g,h\},\{d,g,h\},\{e,g,h\}\}$. Apart from the principal ideals of $\mathcal{L}$, there are precisely the complete ideals

$$\emptyset,\ \langle \leq c,d,e,f\rangle,\ \langle \leq a,b,d,e,f\rangle,\ \langle \leq c,d,e,g,h\rangle,\ \langle \leq a,b,g,h\rangle,$$

where $\langle \leq S\rangle$ denotes the ideal generated by $S \subseteq \mathcal{L}$; here, it suffices to consider the ideals generated by two elements.

(7.3) Theorem: Benson–Conway [1985]. Let X be a modular lattice. Then the following maps are a pair of mutually inverse isomorphisms of lattices:

$$\beta\colon X \to X(\mathcal{L})\colon x \mapsto \{y \in \mathcal{L}; y \leq x\} \quad \text{and} \quad \beta^{-1}\colon X(\mathcal{L}) \to X\colon \mathcal{X} \mapsto \bigvee \mathcal{X}.$$

Proof. The maps β and β^{-1} are well-defined and order-preserving, and we have $\beta^{-1} \circ \beta = \mathrm{id}_X$. Hence we have to show that $\beta \circ \beta^{-1} = \mathrm{id}_{X(\mathcal{L})}$ also holds:

Assume to the contrary that there are $\mathcal{X} \in X(\mathcal{L})$ and $y \in \mathcal{L} \setminus \mathcal{X}$ such that $y \leq \bigvee \mathcal{X}$. Let $n := \min\{|\mathcal{Y}| \in \mathbb{N}; \mathcal{Y} \subseteq \mathcal{X} \text{ finite}, y \leq \bigvee \mathcal{Y}\}$; note that $n \in \mathbb{N}$ is well-defined. Assume that $n = 1$, then we have $y \leq x$ for some $x \in \mathcal{X}$, and since $\mathcal{X} \subseteq \mathcal{L}$ is an ideal we infer $y \in \mathcal{X}$, a contradiction. Hence we have $n \geq 2$, and we may choose y such that n is minimal. Let $\mathcal{Y} = \{y_1, \ldots, y_n\} \subseteq \mathcal{X}$, where we may choose $\mathcal{Y}$ such that $y \vee y_1 \in X$ is minimal.

By modularity we have $y \vee y_1 = (y \vee y_1) \wedge \bigvee_{i=1}^n y_i = (y \vee y_1) \wedge (y_1 \vee \bigvee_{i=2}^n y_i) = y_1 \vee \big((y \vee y_1) \wedge \bigvee_{i=2}^n y_i\big) \in X$. Let $z := (y \vee y_1) \wedge \bigvee_{i=2}^n y_i \in X$, and let $z_1, \ldots, z_r \in \mathcal{L}$ such that $z = \bigvee_{j=1}^r z_j \in X$, for some $r \in \mathbb{N}_0$. Assume that $z_j \notin \mathcal{X}$ for some $j \in \{1, \ldots, r\}$, then we have $z_j \leq y \vee y_1 \leq \bigvee \mathcal{X}$ and $z_j \leq \bigvee_{i=2}^n y_i$, contradicting the minimality of n. Hence we have $z_j \in \mathcal{X}$ for all $j \in \{1, \ldots, r\}$.

By definition we have $y \vee y_1 = y_1 \vee z = y_1 \vee \bigvee_{j=1}^r z_j$, hence since $y \not\leq y_1$ the exchange property entails $y \vee y_1 = y_1 \vee z_j$, for some $j \in \{1, \ldots, r\}$. In particular we have $y \leq y_1 \vee z_j$, and thus $n = 2$ and $y_1 \neq z_j$. Moreover, since $z_j \leq y \vee y_1$ we have $y \vee z_j \leq y \vee y_1$, and by the minimality of $y \vee y_1$ we conclude that $y \vee z_j = y \vee y_1$. Hence, by the maximality property of dotted-lines, there is a dotted-line $\mathcal{D}$ such that $\{y, y_1, z_j\} \subseteq \mathcal{D}$. Since both $y_1, z_j \in \mathcal{X}$, completeness implies that $y \in \mathcal{X}$, the final contradiction. ♯

(7.4) Theorem. Let X be a modular lattice, and for each $z \in \mathcal{L}_2$ we pick a dotted-line $\mathcal{D}_z \subseteq \mathcal{L}$ for z. Then an ideal $\mathcal{X} \subseteq \mathcal{L}$ is complete if and only if it is weakly complete with respect to $\{\mathcal{D}_z; z \in \mathcal{L}_2\}$.

Proof. We only have to show that a weakly complete ideal $\mathcal{X} \subseteq \mathcal{L}$ with respect to $\{\mathcal{D}_z; z \in \mathcal{L}_2\}$ is already complete with respect to all dotted-lines $\mathcal{D} \subseteq \mathcal{L}$. To do so, we proceed by induction on $l(\bigvee \mathcal{X})$. If $l(\bigvee \mathcal{X}) \leq 1$, then $|\mathcal{X}| \leq 1$, thus $\mathcal{X}$ is complete. Hence let $l(\bigvee \mathcal{X}) \geq 2$, and let $x_1, x_2 \in \mathcal{X} \cap \mathcal{D}$ such that $z := x_1 \vee x_2 \in \mathcal{L}_2$. We show that for all $x \in \mathcal{L}$ such that $x \leq z$ we have $x \in \mathcal{X}$:

For $i \in \{1, 2\}$ let $z_i := x_i \vee z_*$. Assuming $z_i \leq z_*$ yields $z = x_1 \vee x_2 \vee z_* = z_1 \vee z_2 = x_{3-i} \vee z_* = x_{3-i}$, a contradiction; and assuming $z_i = z$ yields $z = x_i \vee z_* = x_i$, a contradiction. Hence we have $z_* \lessdot z_i \lessdot z$. We consider the ideals $\mathcal{Z}_* := \{y \in \mathcal{X}; y \leq z_*\}$ and $\mathcal{Z}_i := \{y \in \mathcal{X}; y \leq z_i\}$. Since $\mathcal{X}$ is an ideal and $z_* = (x_1)_* \vee (x_2)_*$, we have $\bigvee \mathcal{Z}_* = z_*$, and hence we have $\bigvee \mathcal{Z}_i = z_i \lessdot z \leq \bigvee \mathcal{X}$.

For any distinguished dotted-line $\mathcal{D}_w \subseteq \mathcal{L}$, for $w \in \mathcal{L}_2$, such that $|\mathcal{D}_w \cap \mathcal{Z}_i| \geq 2$ we have $w = \bigvee \mathcal{D}_w \leq z_i$, and by weak completeness we have $\mathcal{D}_w \subseteq \mathcal{X}$, hence we conclude that $\mathcal{D}_w \subseteq \mathcal{Z}_i$. Thus $\mathcal{Z}_i$ is weakly complete, and since $l(\bigvee \mathcal{Z}_i) < l(\bigvee \mathcal{X})$ it is complete by induction.

Let $y_i \in \mathcal{D}_z$ such that $z_i = y_i \vee z_*$. Since the $\mathcal{Z}_i$ are complete such that $\bigvee \mathcal{Z}_i = z_i$, by (7.3) we have $y_i \in \mathcal{Z}_i \subseteq \mathcal{X}$, and as $\mathcal{X}$ is weakly complete we have $\mathcal{D}_z \subseteq \mathcal{X}$. Now we consider $k \in \mathcal{I}_z$ such that $x \leq z_k \lessdot z$, and let $\mathcal{Z}_k := \{y \in \mathcal{X}; y \leq z_k\}$. Since we have $\mathcal{D}_z \subseteq \mathcal{X}$ and $\bigvee \mathcal{Z}_* = z_*$, we obtain $\bigvee \mathcal{Z}_k = z_k \lessdot z \leq \bigvee \mathcal{X}$. Moreover, $\mathcal{Z}_k$ is weakly complete, and since $l(\bigvee \mathcal{Z}_k) < l(\bigvee \mathcal{X})$ it is complete by induction. As $\bigvee \mathcal{Z}_k = z_k$, by (7.3) we finally have $x \in \mathcal{Z}_k \subseteq \mathcal{X}$. ♯

(7.5) Distributive lattices. a) Let X be a locally chain-finite lattice, having a zero element $\underline{0} \in X$. Then X is called **distributive** if the **distributive law** $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ holds for all $x, y, z \in X$. In this case any interval in X is distributive as well. Note that for $z \leq x$ from the distributive law we get $x \wedge (y \vee z) = (x \wedge y) \vee z$, that is the modular law is fulfilled automatically, hence X is modular.

Here is the natural example for a lattice fulfilling the distributive law: Let N be a set. Then $\mathcal{P}(N)$, partially ordered by set-theoretic inclusion $\subseteq$ and having the zero element $\emptyset$, becomes a lattice with meet $M \cap M'$ and join $M \cup M'$, where for $M_0, M_1, M_2 \subseteq N$ we have $M_0 \cap (M_1 \cup M_2) = (M_0 \cap M_1) \cup (M_0 \cap M_2)$. But note that $\mathcal{P}(N)$ is locally chain-finite if and only if N is finite.

b) The basic structure theorem for distributive lattices is **Birkhoff's Representation Theorem**, which is a special case of (7.3), shedding some further light on the significance of the set $\mathcal{L}_2$:

Let X be a modular lattice. Then X is distributive if and only if $\mathcal{L}_2 = \emptyset$. In this case, we have $X(\mathcal{L}) = \{\mathcal{X} \subseteq \mathcal{L}; \mathcal{X} \text{ finite ideal}\}$, and the following maps are a pair of mutually inverse isomorphisms of lattices:

$$\beta \colon X \to X(\mathcal{L}) \colon x \mapsto \{y \in \mathcal{L}; y \leq x\} \quad \text{and} \quad \beta^{-1} \colon X(\mathcal{L}) \to X \colon \mathcal{X} \mapsto \bigvee \mathcal{X}.$$

Moreover, the length function on $X(\mathcal{L})$ is given by $l(\mathcal{X}) := |\mathcal{X}|$, and any interval $[\underline{0}, x]$, for $x \in X$, is finite. In particular, if X has a one element, then X is finite.

This is seen as follows: If $z \in \mathcal{L}_2 \neq \emptyset$, then let $z_* \lessdot z_i \lessdot z$, for $i \in \{0, 1, 2\}$. Then we have $z_0 \wedge (z_1 \vee z_2) = z_0 \neq z_* = (z_0 \wedge z_1) \vee (z_0 \wedge z_2)$, hence $[z_*, z]$ is not distributive, hence neither X is. Conversely, if $\mathcal{L}_2 = \emptyset$ then $X(\mathcal{L})$ is the set of ideals of $\mathcal{L}$ having a bound in X, and whose join and meet operations are given by taking set theoretic unions and intersections, respectively, hence $X(\mathcal{L})$ is distributive, and thus X also is.

Now, letting X be distributive, in order to determine the length function on $X(\mathcal{L})$, we show that $|\beta(x)| = l(x)$, for $x \in X$. We proceed by induction on $l := l(x) \in \mathbb{N}_0$; the case $l = 0$ being trivial, we assume that $l \geq 1$, and let $\mathcal{X} := \beta(x) \in X(\mathcal{L})$. Then there are $y \in X$ and $z \in \mathcal{L}$, such that $y \lessdot x$ and $x = y \vee z$. Letting $\mathcal{Y} := \beta(y) \in X(\mathcal{L})$, and $\mathcal{Z} := \beta(z) = \langle \leq z \rangle \in X(\mathcal{L})$ be the principal ideal generated by z, we get $\mathcal{X} = \mathcal{Y} \cup \mathcal{Z}$. We have $l(y) = l(x) - 1$, and thus by induction we get $|\mathcal{Y}| = l(y)$. Moreover, we have $[y \wedge z, z] \cong [y, y \vee z] = [y, x]$, implying that $y \wedge z = z_* \lessdot z$, in other words $\mathcal{Z} \setminus \{z\} = \beta(z_*) = \beta(y \wedge z) = \mathcal{Y} \cap \mathcal{Z}$. This yields $\mathcal{Z} = (\mathcal{Y} \cap \mathcal{Z}) \,\dot\cup\, \{z\}$, and thus $\mathcal{X} = \mathcal{Y} \cup \mathcal{Z} = \mathcal{Y} \,\dot\cup\, \{z\}$, saying that $|\mathcal{X}| = |\mathcal{Y}| + 1 = l(y) + 1 = l(x)$.

In particular this shows that $X(\mathcal{L})$ consists of the finite ideals of $\mathcal{L}$. Finally, since for any $x \in X$ the ideal $\beta(x) \subseteq \mathcal{L}$ is finite, there are only finitely many subideals of $\beta(x)$, in other words there are only finitely many $y \in X$ such that $y \leq x$, that is the interval $[\underline{0}, x]$ is finite. ♯

(7.6) The block graph. Let X be a modular lattice having a one element $\underline{1}$, and let β be as in (7.3). We proceed to derive a description of the centre of X:

a) For $x \in X$ let $\mathcal{X} := \beta(x) \in X(\mathcal{L})$. Then we have $x \in \mathcal{Z}(X)$ if and only if $\mathcal{X}' := \mathcal{L} \setminus \mathcal{X} \in X(\mathcal{L})$. In this case, letting $x' := \beta^{-1}(\mathcal{X}') \in X$, we have $X \cong [\underline{0}, x] \times [\underline{0}, x']$ as inner direct product:

Let $x \in \mathcal{Z}(X)$, let $X \cong [\underline{0}, x] \times [\underline{0}, x']$, where $x' \in X$, and let $\mathcal{X}' := \beta(x') \in X(\mathcal{L})$. Hence for $z \in \mathcal{L}$ we have either $z \leq x$ or $z \leq x'$, and thus $\mathcal{L} = \{z \in \mathcal{L}; z \leq x\} \,\dot\cup\, \{z \in \mathcal{L}; z \leq x'\} = \mathcal{X} \,\dot\cup\, \mathcal{X}'$, hence $\mathcal{X}' = \mathcal{L} \setminus \mathcal{X}$.

Let conversely $\mathcal{X}' := \mathcal{L} \setminus \mathcal{X} \in X(\mathcal{L})$ and $x' := \beta^{-1}(\mathcal{X}') \in X$. Now let $\mathcal{Y}, \mathcal{Y}' \in X(\mathcal{L})$ such that $\mathcal{Y} \subseteq \mathcal{X}$ and $\mathcal{Y}' \subseteq \mathcal{X}'$. Then $\mathcal{Y} \,\dot\cup\, \mathcal{Y}' \subseteq \mathcal{L}$ is an ideal which is bounded in X. Let $\mathcal{D} \subseteq \mathcal{L}$ be a dotted-line such that $|\mathcal{D} \cap (\mathcal{Y} \,\dot\cup\, \mathcal{Y}')| \geq 2$. Since

$|\mathcal{D}| \geq 3$ we have $|\mathcal{D}\cap\mathcal{X}| \geq 2$ or $|\mathcal{D}\cap\mathcal{X}'| \geq 2$, and hence either $\mathcal{D} \subseteq \mathcal{X}$ or $\mathcal{D} \subseteq \mathcal{X}'$. Thus we have either $|\mathcal{D} \cap \mathcal{Y}| \geq 2$ or $|\mathcal{D} \cap \mathcal{Y}'| \geq 2$, and hence either $\mathcal{D} \subseteq \mathcal{Y}$ or $\mathcal{D} \subseteq \mathcal{Y}'$. Thus $\mathcal{Y} \,\dot\cup\, \mathcal{Y}' \in X(\mathcal{L})$, coinciding with the join of $\mathcal{Y}$ and $\mathcal{Y}'$ in $X(\mathcal{L})$.

We consider the order-preserving maps $\sigma\colon [\underline{0}, x] \times [\underline{0}, x'] \to X\colon [y, y'] \mapsto y \vee y'$ and $\tau\colon X \to [\underline{0}, x] \times [\underline{0}, x']\colon z \mapsto [z \wedge x, z \wedge x']$. By the above we have $\beta(y \vee y') = \beta(y) \cup \beta(y')$, while for $z \in X$ we have $\beta(z \wedge x) = \beta(z) \cap \mathcal{X} \in X(\mathcal{L})$ and $\beta(z \wedge x') = \beta(z) \cap \mathcal{X}' \in X(\mathcal{L})$ anyway. Thus we have $\beta(\sigma\tau(z)) = \beta((z \wedge x) \vee (z \wedge x')) = \beta(z \wedge x) \cup \beta(z \wedge x') = (\beta(z) \cap \mathcal{X}) \cup (\beta(z) \cap \mathcal{X}') = \beta(z)$, by (7.3) entailing $\sigma\tau(z) = z$. Moreover, for $\tau\sigma([y, y']) = [(y \vee y') \wedge x, (y \vee y') \wedge x']$ we get $\beta((y\vee y')\wedge x) = \beta(y\vee y')\cap\mathcal{X} = (\beta(y)\cup\beta(y'))\cap\mathcal{X} = (\beta(y)\cap\mathcal{X})\cup(\beta(y')\cap\mathcal{X}) = \beta(y)$ and similarly $\beta((y \vee y') \wedge x') = \beta(y \vee y') \cap \mathcal{X}' = (\beta(y) \cup \beta(y')) \cap \mathcal{X}' = (\beta(y) \cap \mathcal{X}') \cup (\beta(y') \cap \mathcal{X}') = \beta(y')$, which by (7.3) entails $\tau\sigma([y, y']) = [y, y']$. Hence σ and τ are a pair of mutually inverse isomorphisms of lattices. ♯

b) Let the **block graph** of X be defined as the undirected simple graph having vertex set $\mathcal{L}$, where vertices $x, y \in \mathcal{L}$ are adjacent if and only if $x < y$ or $y < x$ or $x \vee y \in \mathcal{L}_2$. Vertices being in the same connected component of the block graph are called to be in the same **block**, giving rise to the disjoint union $\mathcal{L} = \coprod_{i=1}^{d} \mathcal{Z}_i$, where $d \in \mathbb{N}_0$ is the number of blocks occurring.

Then $\mathcal{Z}_i \subseteq \mathcal{L}$ is an ideal, for $i \in \{1, \ldots, d\}$, and for any dotted-line $\mathcal{D} \subseteq \mathcal{L}$ we have either $\mathcal{D} \cap \mathcal{Z}_i = \emptyset$ or $\mathcal{D} \subseteq \mathcal{Z}_i$. Since X has a one element, $\mathcal{Z}_i$ has an upper bound in X, hence $\mathcal{Z}_i \in X(\mathcal{L})$, and we let $z_i := \beta^{-1}(\mathcal{Z}_i) \in X$.

c) Given $x \in X$, we have $x \in \mathcal{Z}(X)$ if and only if $x = \bigvee\{z_i \in X; i \in \{1, \ldots, d\},\ z_i \leq x\}$. In particular, $X \cong \prod_{i=1}^{d}[\underline{0}, z_i]$ is the unique decomposition of X into non-trivial indecomposable intervals:

Let $x \in \mathcal{Z}(X)$ and $\mathcal{X} := \beta(x) \in X(\mathcal{L})$, and $\mathcal{X}' := \mathcal{L} \setminus \mathcal{X} \in X(\mathcal{L})$ and $x' := \beta^{-1}(\mathcal{X}') \in X$. We consider $\mathcal{Z}_i$ such that $\mathcal{X} \cap \mathcal{Z}_i \neq \emptyset$, and show that $\mathcal{Z}_i \subseteq \mathcal{X}$:

To this end let $y \in \mathcal{X} \cap \mathcal{Z}_i$, and $y' \in \mathcal{Z}_i$ be connected to y in the block graph. If $y' < y$, then by completeness we have $y' \in \mathcal{X}$ as well. If $y' > y$, assume that $y' \notin \mathcal{X}$, hence we have $y' \in \mathcal{X}'$, thus by completeness $y \in \mathcal{X}'$, a contradiction; this shows $y' \in \mathcal{X}$. If $y \vee y' \in \mathcal{L}_2$, assume that $y' \notin \mathcal{X}$, hence $y' \in \mathcal{X}'$; now there is a dotted-line $\mathcal{D} \subseteq \mathcal{L}$ such that $\{y, y'\} \subseteq \mathcal{D}$, entailing $\mathcal{D} \cap \mathcal{X} \neq \emptyset$ and $\mathcal{D} \cap \mathcal{X}' \neq \emptyset$; since $|\mathcal{D}| \geq 3$ we have $|\mathcal{D} \cap \mathcal{X}| \geq 2$ or $|\mathcal{D} \cap \mathcal{X}'| \geq 2$, and hence by completeness either $\mathcal{D} \subseteq \mathcal{X}$ or $\mathcal{D} \subseteq \mathcal{X}'$, a contradiction; this shows $y' \in \mathcal{X}$.

Thus we have $\bigvee\{z_i \in X; i \in \{1, \ldots, d\},\ \mathcal{X} \cap \mathcal{Z}_i \neq \emptyset\} \leq x$ and similarly we get $\bigvee\{z_j \in X; j \in \{1, \ldots, d\},\ \mathcal{X} \cap \mathcal{Z}_j \neq \emptyset\} \leq x'$. Since $X \cong [\underline{0}, x] \times [\underline{0}, x']$ and $\bigvee_{i=1}^{d} z_i = \underline{1} = x \vee x'$ we conclude $x = \bigvee\{z_i \in X; i \in \{1, \ldots, d\},\ \mathcal{X} \cap \mathcal{Z}_i \neq \emptyset\}$..

Let conversely $\{1, \ldots, d\} = \mathcal{I} \,\dot\cup\, \mathcal{J}$, and $x := \bigvee_{i\in\mathcal{I}} z_i$ and $x' := \bigvee_{j\in\mathcal{J}} z_j$. Since for any dotted-line $\mathcal{D} \subseteq \mathcal{L}$ we have $\mathcal{D} \cap \mathcal{Z}_i = \emptyset$ or $\mathcal{D} \subseteq \mathcal{Z}_i$, for any $i \in \{1, \ldots, d\}$, we conclude that $\beta(x) = \bigcup_{i\in\mathcal{I}} \mathcal{Z}_i = \coprod_{i\in\mathcal{I}} \mathcal{Z}_i \in X(\mathcal{L})$ and $\beta(x') = \bigcup_{j\in\mathcal{J}} \mathcal{Z}_j = \coprod_{j\in\mathcal{J}} \mathcal{Z}_j \in X(\mathcal{L})$. Since $\mathcal{L} = (\coprod_{i\in\mathcal{I}} \mathcal{Z}_i) \,\dot\cup\, (\coprod_{j\in\mathcal{J}} \mathcal{Z}_j)$ we have $x \in \mathcal{Z}(X)$, where $X \cong [\underline{0}, x] \times [\underline{0}, x']$. ♯

(7.7) Maeda's Theorem. a) Let X be a modular lattice. Then we have the followung transitivity property for atoms belonging to the same connected component of the block graph: Let $x, y, z \in \mathcal{L}$ be atoms such that $x \neq y$ and $x \vee z \in \mathcal{L}_2$ as well as $y \vee z \in \mathcal{L}_2$. Then we also have $x \vee y \in \mathcal{L}_2$:

We may assume that $x \vee z \neq y \vee z$ holds. Hence $(x \vee z) \wedge (y \vee z) = z$, and thus $l(x \vee y \vee z) = l(x \vee z) + l(y \vee z) - l((x \vee z) \wedge (y \vee z)) = 2+2-1 = 3$. Since $x \vee z \in \mathcal{L}_2$ and $y \vee z \in \mathcal{L}_2$ there are atoms $v, w \in X \setminus \{x, y, z\}$ such that $0 \lessdot v \lessdot x \vee z$ and $0 \lessdot w \lessdot y \vee z$. Assume that $v = w$, hence $0 \lessdot v = w \leq (x \vee z) \wedge (y \vee z) = z$, thus $v = w = z$, a contradiction. Thus we have $v \neq w$.

Let $u := (v \vee w) \wedge (x \vee y) \in X$. Since $(v \vee w) \vee (x \vee y) = x \vee y \vee z$, we get $l(u) = l(v \vee w) + l(x \vee y) - l(x \vee y \vee z) = 2 + 2 - 3 = 1$. Hence $0 \lessdot u \lessdot x \vee y$. Assume that $u = x$, then we have $x \leq v \vee w$, hence $v \vee w = x \vee v = x \vee z = w \vee z = y \vee z$, a contradiction. Similarly, assuming $u = y$ implies $y \leq v \vee w$, thus $v \vee w = y \vee w = y \vee z = v \vee z = x \vee z$, a contradiction. Hence we have $x \neq u \neq y$, showing $x \vee y \in \mathcal{L}_2$. $\sharp$

Note that the above transitivity property does not hold for local elements in general: For the example depicted in Table 10, see (7.2), we have $b \vee d \in \mathcal{L}_2$ and $c \vee d \in \mathcal{L}_2$, but $b \vee c \notin \mathcal{L}_2$, actually even $c < d$.

b) We now consider complemented lattices. In this case the local elements coincide with the atoms, thus the block graph just encodes the $\mathcal{L}_2$ property; in view of the above observation the connected components of the block graph are complete graphs: Let X be complemented; recall that X has a one element $\underline{1}$. Then for atoms $x \neq y \in X$ we have $x \vee y \in \mathcal{L}_2$ if and only if x and y have a common complement in X:

Let $z \in X$ be a common complement for x and y. From $[(x \vee y) \wedge z, x \vee y] \cong [z, x \vee y \vee z] = [z, \underline{1}] = [z, x \vee z] \cong [z \wedge x, x] = [\underline{0}, x]$ we get $\underline{0} \lessdot (x \vee y) \wedge z \lessdot x \vee y$. Since $x \neq (x \vee y) \wedge z \neq y$ we have $x \vee y \in \mathcal{L}_2$.

Let conversely $x \vee y \in \mathcal{L}_2$, let $z \in X \setminus \{x, y\}$ such that $\underline{0} \lessdot z \lessdot x \vee y$, and let $w \in X$ be a complement for $x \vee y$. Hence we have $x \vee (w \vee z) = w \vee (x \vee y) = \underline{1}$. Moreover, we get $[x \wedge (w \vee z), x] \cong [w \vee z, x \vee (w \vee z)] = [w \vee z, w \vee (x \vee y)] \cong [z, x \vee y]$, thus $x \wedge (w \vee z) \lessdot x$ and hence $x \wedge (w \vee z) = \underline{0}$. Thus $w \vee z \in X$ is a complement for x. A similar argument shows that $w \vee z$ also is a complement for y. $\sharp$

c) Let still X be complemented, and let $X \cong \prod_{i=1}^{d} [\underline{0}, z_i]$, where $d \in \mathbb{N}_0$, be the decomposition of X into non-trivial indecomposable intervals. Then we have **Maeda's Theorem**, saying that for atoms $x \neq y \in X$ we have $x \vee y \in \mathcal{L}_2$ if and only if there is $i \in \{1, \ldots, d\}$ such that both $x \leq z_i$ and $y \leq z_i$:

There are unique $i, j \in \{1, \ldots, d\}$ such that $x \leq z_i$ and $y \leq z_j$. If $i \neq j$, then since $[0, z_i \vee z_j] \cong [0, z_i] \times [0, z_j]$ we have $[0, x \vee y] = [0, x] \times [0, y]$, thus $x \vee y \notin \mathcal{L}_2$. If $i = j$, then let $\mathcal{Z}_i := \beta(z_i) \subseteq \mathcal{L}$ be the associated block. As $\mathcal{Z}_i$ is a connected component of the block graph, there is a chain $x = x_0, x_1, \ldots, x_s = y$ in $\mathcal{L}$ such that $x_{i-1} \vee x_i \in \mathcal{L}_2$, for all $i \in \{1, \ldots, s\}$. Hence we have $x \vee y \in \mathcal{L}_2$ as well. $\sharp$

(7.8) Submodule lattices. Although these have been the original motivation for the developments in [19], we only briefly consider submodule lattices, in order to collect the facts needed later on, see (11.3):

a) Let K be a field, let $\mathcal{A}$ be a finite-dimensional K-algebra $\mathcal{A}$, and let V be a finitely generated $\mathcal{A}$-module; hence V is finite-dimensional as well. Then the set $\mathcal{M}(V)$ of all $\mathcal{A}$-submodules of V becomes a partially ordered set with respect to set-theoretic inclusion $\subseteq$, and thus a lattice with meet $U \cap U'$ and join $U + U'$.

If $X, Y, Z \leq V$ are $\mathcal{A}$-submodules such that $Z \leq X$, then we have $X \cap (Y + Z) = (X \cap Y) + Z$, which holds algebraically already for K-vector spaces, that is $\mathcal{M}(V)$ fulfills the modular law; note that the natural isomorphism of the lattice intervals $X \cap Y \leq X$ and $Y \leq X + Y$ is induced by the isomorphism of $\mathcal{A}$-modules $X/(X \cap Y) \to (X + Y)/Y \colon x + (X \cap Y) \mapsto x + Y$. Moreover, $\mathcal{M}(V)$ is locally chain-finite, hence becomes a modular lattice with zero element $\{0\}$ and one element V; note that $\mathcal{M}(V)$ is not necessarily locally finite.

In particular, $\mathcal{M}(V)$ is graded, the length function on $\mathcal{A}$-submodules being given by **composition length**. Note that, by the **Jordan-Hölder Theorem**, whenever $\{0\} = V_0 \lessdot V_1 \lessdot \cdots \lessdot V_l = V$ is a saturated chain of $\mathcal{A}$-submodules, where $l = l(V) \in \mathbb{N}_0$, then the **multiplicity** $[V \colon S] \in \mathbb{N}_0$ with which a simple $\mathcal{A}$-module S occurs, up to isomorphism, as a subquoient X_i/X_{i-1}, for $i \in \{1, \ldots, l\}$, is independent of the particular saturated chain chosen; if $[V \colon S] \geq 1$ then S is called a **constituent** of V.

b) The lattice theoretic radical of an $\mathcal{A}$-submodule $U \leq V$ coincides with its **Jacobson radical** $\mathrm{rad}(U)$, hence $U/\mathrm{rad}(U)$ being complemented just says that it is a **semi-simple** $\mathcal{A}$-module. Given a simple $\mathcal{A}$-module S, then U is called **S-local** if $U/\mathrm{rad}(U) \cong S$. Letting $\mathcal{L}_S(V) \subseteq \mathcal{M}(V)$ be the set of all S-local $\mathcal{A}$-submodules of V, then $\mathcal{L}(V) := \coprod_{S \text{ simple}} \mathcal{L}_S(V) \subseteq \mathcal{M}(V)$, the disjoint union running over the finitely many isomorphism types of simple $\mathcal{A}$-modules, is the set of all local $\mathcal{A}$-submodules of V; we have $\mathcal{L}(V) \neq \emptyset$ whenever $V \neq \{0\}$.

We have $\mathcal{L}_S(V) \neq \emptyset$ if and only if S is a constituent of V: If $L \in \mathcal{L}_S(V)$ then $L/\mathrm{rad}(L) \cong S$ shows that S is a constituent of V. Conversely, if $U' < U \leq V$ are $\mathcal{A}$-submodules such that $U/U' \cong S$, then letting $L_1, \ldots, L_r \in \mathcal{L}$ such that $U = \sum_{i=1}^r L_i$ is irredundant, where $r = r(U) = r(U/\mathrm{rad}(U)) = l(U/\mathrm{rad}(U)) \in \mathbb{N}$, then $\mathrm{rad}(U) = \sum_{i=1}^r \mathrm{rad}(L_i)$ shows that $U/\mathrm{rad}(U) \cong \bigoplus_{i=1}^r L_i/\mathrm{rad}(L_i)$, hence $\mathrm{rad}(U) \leq U'$ implies that $L_i/\mathrm{rad}(L_i) \cong S$ for some $i \in \{1, \ldots, r\}$.

In particular, we have $|\mathcal{L}_S(V)| = 1$ if and only if $[V \colon S] = 1$: Assume that there are $L', L \in \mathcal{L}_S(V)$ such that $L' \neq L$; if $L' < L$ then from $\mathrm{rad}(L') \lessdot L' \leq \mathrm{rad}(L) \lessdot L$ we conclude that $[V \colon S] \geq 2$; if $L' \not\leq L \not\leq L'$ then we have $(L + L')/\mathrm{rad}(L + L') \cong L/\mathrm{rad}(L) \oplus L'/\mathrm{rad}(L') \cong S \oplus S$, which again entails $[V \colon S] \geq 2$. Conversely, let $\mathcal{L}_S(V) = \{L\}$, and assume that $[V \colon S] \geq 2$; then there are $\mathcal{A}$-submodules $U_1' \lessdot U_1 \leq U_2' \lessdot U_2 \leq V$ such that $U_i/U_i' \cong S$, for $i \in \{1, 2\}$; letting $U_2 = \sum_{i=1}^r L_i$, where $L_1, \ldots, L_r \in \mathcal{L}$ for some $r \in \mathbb{N}$, then there is $i \in \{1, \ldots, r\}$ such that $L_i \not\leq U_2'$, hence $L_i/(L_i \cap U_2') \cong (L_i + U_2')/U_2' \cong U_2/U_2' \cong S$ shows that $L_i = L$, thus we have $L \leq U_2$ but $L \not\leq U_2'$; similarly we infer that $L \leq U_1$, a contradiction.

c) We proceed to describe $\mathcal{L}_2(V) \subseteq \mathcal{M}(V)$. In order to do so, we first we first describe the $\mathcal{A}$-submodules of $V := S \oplus T$, where S and T are simple $\mathcal{A}$-modules: To this end, let $\{0\} \lessdot U \lessdot V$ be an $\mathcal{A}$-submodule, where we assume that $U \neq \{0\} \oplus T$, hence there is $[u, u'] \in U$ such that $u \neq 0$.

If $S \not\cong T$, then by **Wedderburn's Theorem** there is $a \in \mathcal{A}$ such that $a|_S = \mathrm{id}_S$ and $a|_T = 0$. Thus we have $0 \neq [u, 0] \in U$, hence from S being simple we conclude that $S \oplus \{0\} \leq U$. Thus we infer that $S \oplus \{0\}$ and $\{0\} \oplus T$ are the only non-trivial proper $\mathcal{A}$-submodules of V.

If $S \cong T$, then we may assume that $S = T$, and let $E = E_S := \mathrm{End}_{\mathcal{A}}(S)$ be the ring of **$\mathcal{A}$-endomorphisms** of S, where by **Schur's Lemma** E is a skew field over K. Hence S can be considered as an E-vector space, and by Wedderburn's Theorem again $\mathcal{A}|_S \cong E^{n \times n}$, where $n := \dim_E(S) = \frac{\dim_K(S)}{\dim_K(E)}$. If $\{u, u'\}$ is E-linearly independent, then there is $a \in \mathcal{A}$ such that $ua = u$ and $u'a = 0$, hence we conclude that $0 \neq [u, 0] \in U$ and thus $U = S \oplus \{0\}$. If $\{u, u'\}$ is E-linearly dependent, then $[u, u'] = [u, u\alpha]$ for some $\alpha \in E$, and thus projecting onto the first component shows that $U = [u, u\alpha] \cdot \mathcal{A} \cong S$. Moreover, if $U = [u, u\alpha] \cdot \mathcal{A} = [u, u\beta] \cdot \mathcal{A}$, for some $\alpha \neq \beta \in E$, then $[0, u(\alpha - \beta)] \in U$, where since $\alpha - \beta \in E^*$ we have $u(\alpha - \beta) \neq 0$, hence $U = \{0\} \oplus T$, a contradiction. Hence the $\mathcal{A}$-submodules of V different from $\{0\} \oplus S$ are in bijection with E via $\alpha \mapsto [u, u\alpha] \cdot \mathcal{A}$, for some fixed $0 \neq u \in S$.

We conclude that the submodule lattice of $V \cong S \oplus T$ is indecomposable if and only if $S \cong T$. In this case, if moreover $K = \mathbb{F}_q$ is the finite field with q elements, then E is the field with $q^{\dim_K(E)}$ elements, and there are precisely $|E| + 1 = q^{\dim_K(E)} + 1$ non-trivial proper $\mathcal{A}$-submodules of $V = S \oplus S$. $\sharp$

We are now prepared to describe $\mathcal{L}_2(V)$, where V is arbitrary again: If $L, L' \in \mathcal{L}(V)$ such that $L \not\leq L' \not\leq L$, then letting $Z := L + L' \leq V$ we have $Z/\mathrm{rad}(Z) \cong L/\mathrm{rad}(L) \oplus L'/\mathrm{rad}(L')$, hence we have $Z \in \mathcal{L}_2(V)$ if and only if $L/\mathrm{rad}(L) \cong L'/\mathrm{rad}(L')$, that is there is a simple $\mathcal{A}$-module S such that both $L, L' \in \mathcal{L}_S(V)$; in this case we may let $\mathcal{I}_Z := E_S \,\dot\cup\, \{\infty\}$. In conclusion, accompanying the disjoint decomposition of $\mathcal{L}(V)$, we have a finite disjoint union $\mathcal{L}_2(V) := \coprod_{S \text{ simple}} \mathcal{L}_{2,S}(V) \subseteq \mathcal{M}(V)$, where in turn we have $\mathcal{L}_{2,S}(V) = \{L + L' \in \mathcal{M}(V); L, L' \leq \mathcal{L}_S(V), L \not\leq L' \not\leq L\}$.

In particular, we have $\mathcal{L}_{2,S}(V) = \emptyset$ if and only if $\mathcal{L}_S(V) = \emptyset$ or $\mathcal{L}_S(V)$ is a chain. Moreover, dotted-lines can be computed by considering each simple $\mathcal{A}$-module in turn, and all dotted-lines belonging to one and the same simple $\mathcal{A}$-module have the same cardinality. Finally, letting $\mathcal{S}(V)$ be the set of simple $\mathcal{A}$-modules, up to isomorphism, which occur as constituents of V, and letting $\mathcal{L}(V) = \coprod_{i=1}^{d} \mathcal{Z}_i$ be the block decomposition, where $d \in \mathbb{N}_0$, from the definition of the block graph we infer that there is a corresponding partition $\mathcal{S}(V) = \coprod_{i=1}^{d} \mathcal{S}_i$.

d) In particular, by Birkhoff's Representation Theorem, $\mathcal{M}(V)$ is a distributive lattice if and only if $\mathcal{L}_2(V) = \emptyset$, that is $\mathcal{L}_{2,S}(V) = \emptyset$ for all simple $\mathcal{A}$-modules

S, which in turn holds if and only if $\mathcal{L}_S(V)$ is a chain for all constituents S of V. In particular, this is the case if $[V\colon S] \leq 1$ for all simple $\mathcal{A}$-modules S, that is V is **multiplicity-free**. Recall that if $\mathcal{M}(V)$ is distributive then it is finite, being in natural bijection with the partially ordered set of ideals of $\mathcal{L}(V)$.

If $\mathcal{M}(V)$ is distributive, then the indecomposable direct summands of the $\mathcal{A}$-module V are precisely given by the block graph; note that since $\mathcal{L}_2(V) = \emptyset$, the latter just encodes the mutual inclusion of local $\mathcal{A}$-submodules: Let $\mathcal{L}(V) = \coprod_{i=1}^{d} \mathcal{Z}_i$ be the block decomposition, where $d \in \mathbb{N}_0$. Then letting $Z_i := \beta^{-1}(\mathcal{Z}_i) = \mathcal{Z}_i \cdot \mathcal{A} \leq V$ be the $\mathcal{A}$-submodule generated by $\mathcal{Z}_i$, where β is as in (7.3), we have the lattice block decomposition $V = \bigoplus_{i=1}^{d} Z_i$ as $\mathcal{A}$-modules. Hence it remains to be shown that the Z_i are indecomposable:

Assume that $Z_i = Z_i' \oplus Z_i''$ as $\mathcal{A}$-modules, where $Z_i' \neq \{0\} \neq Z_i''$. Then, assume there is a $S \in \mathcal{S}_i$ being a constituent of both Z_i' and Z_i''; then there are S-local submodules $L' \leq Z_i'$ and $L'' \leq Z_i''$, which hence fulfill $L' \not\leq L'' \not\leq L'$, implying that $\mathcal{L}_{2,S}(V) \neq \emptyset$, a contradiction. Thus we have a corresponding partition $\mathcal{S}_i = \mathcal{S}_i' \,\dot{\cup}\, \mathcal{S}_i''$, where $\mathcal{S}_i' \neq \emptyset \neq \mathcal{S}_i''$, entailing a non-trivial partition $\coprod_{S\in\mathcal{S}_i} \mathcal{L}_S(V) = \coprod_{S\in\mathcal{S}_i'} \mathcal{L}_S(V) \,\dot{\cup}\, \coprod_{S\in\mathcal{S}_i''} \mathcal{L}_S(V)$, where there are no inclusions between the local $\mathcal{A}$-submodules in the left hand and right hand parts. From this we conclude that the latter are contained in distinct connected components of the block graph, a contradiction. $\sharp$

III Incidence algebras

8 Incidence algebras

(8.1) The incidence algebra. a) Let X be a locally finite partially ordered set, let $R \neq \{0\}$ be a commutative ring, and let $\mathcal{A}_R(X) := \{f \in \mathrm{Maps}(X \times X, R); f(x,y) = 0 \text{ if } x \not\leq y\}$. Then $\mathcal{A}_R(X)$ is an R-module respect to pointwise addition $(f+g)(x,y) := f(x,y) + g(x,y)$ and pointwise scalar multiplication $(cf)(x,y) := c \cdot f(x,y)$, for all $f, g \in \mathcal{A}_R(X)$ and $x, y \in X$ and $c \in R$.

Convolutional multiplication on $\mathcal{A}_R(X)$ is well-defined by letting $(fg)(x,y) := \sum_{z\in X} f(x,z)g(z,y)$, for all $f, g \in \mathcal{A}_R(X)$ and $x, y \in X$: If $x \not\leq y$ then all summands vanish, hence $(fg)(x,y) = 0$, while if $x \leq y$ we get $(fg)(x,y) = \sum_{x\leq z\leq y} f(x,z)g(z,y)$, which by local finiteness is a finite sum. Then we have $(fg)h = f(gh)\colon [x,y] \mapsto \sum_{z,z'\in X} f(x,z)g(z,z')h(z',y)$, for all $f, g, h \in \mathcal{A}_R(X)$, that is associativity holds. Moreover, for the map $\delta\colon X \times X \to R\colon [x,y] \mapsto \delta_{x,y}$, where $\delta_{x,y}$ is the Kronecker symbol, we have $(f\delta)(x,y) = \sum_{z\in X} f(x,z)\delta(z,y) = f(x,y)$ and $(\delta f)(x,y) = \sum_{z\in X} \delta(x,z)f(z,y) = f(x,y)$, for all $x, y \in X$, hence $\delta \in \mathcal{A}_R(X)$ is a neutral element with respect to convolutional multiplication. Since we have distributivity, and $c(fg) = (cf)g = f(cg) \in \mathcal{A}_R(X)$, for all $c \in R$, we conclude that $\mathcal{A}_R(X)$ is a non-commutative R-algebra, called the **incidence algebra** associated with the partially ordered set X.

b) Let $f \in \mathcal{A}_R(X)$. We show that f is right invertible, if and only if f is left invertible, which holds if and only if $f(x,x) \in R^*$ for all $x \in X$:

The map $g \in \mathcal{A}_R(X)$ is a right inverse of f if and only if $\delta_{x,y} = (fg)(x,y) = \sum_{z \in X} f(x,z)g(z,y)$, for all $x, y \in X$. Hence, for $x = y$, from the existence of g we infer $f(x,x) \in R^*$. Conversely, if $f(x,x) \in R^*$ for all $x \in X$, we define $g \in \mathcal{A}_R(X)$ using local finiteness as follows: We let $g(x,x) := f(x,x)^{-1} \in R$ for all $x \in X$, and for $x < y \in X$, by induction on $l(x,y) \in \mathbb{N}_0$ assuming that $g(z,y)$ has already been found for all $x < z \leq y \in X$, we let $g(x,y) = -f(x,x)^{-1} \cdot \sum_{x<z\leq y} f(x,z)g(z,y) \in R$. Then we have $\sum_{x\leq z\leq y} f(x,z)g(z,y) = f(x,x)g(x,y) + \sum_{x<z\leq y} f(x,z)g(z,y) = 0$.

Similarly, the map $g \in \mathcal{A}_R(X)$ is a left inverse of f if and only if $\delta_{x,y} = (gf)(x,y) = \sum_{z \in X} g(x,z)f(z,y)$, for all $x, y \in X$. Hence, for $x = y$, from the existence of g we infer $f(x,x) \in R^*$. Conversely, if $f(x,x) \in R^*$ for all $x \in X$, we define $g \in \mathcal{A}_R(X)$ using local finiteness as follows: We let $g(x,x) := f(x,x)^{-1} \in R$ for all $x \in X$, and for $x < y \in X$, by induction on $l(x,y) \in \mathbb{N}_0$ assuming that $g(x,z)$ has already been found for all $x \leq z < y \in X$, we let $g(x,y) = -f(y,y)^{-1} \cdot \sum_{x\leq z<y} g(x,z)f(z,y) \in R$. Then we have $\sum_{x\leq z\leq y} g(x,z)f(z,y) = g(x,y)f(y,y) + \sum_{x\leq z<y} g(x,z)f(z,y) = 0$. ♯

Note that, in this case, right and left inverses coincide and are uniquely defined, giving rise to the **unit group** $\mathcal{A}_R^*(X)$ of $\mathcal{A}_R(X)$. Note that for $f \in \mathcal{A}_R^*(X)$ the value $f^{-1}(x,y) \in R$, for $x \leq y \in X$, only depends on the interval $[x,y] \subseteq X$.

c) Let $\mathrm{rad}(R)$ denote the **Jacobson radical** of R. We show that $\mathrm{rad}(\mathcal{A}_R(X)) = \mathcal{J}_R(X) := \{f \in \mathcal{A}_R(X); f(x,x) \in \mathrm{rad}(R) \text{ for all } x \in X\}$:

Let $f \in \mathcal{J}_R(X)$. Then for any $x \in X$ and any $a \in R$ we have $1 + a \cdot f(x,x) \in R^*$. Hence for any $g \in \mathcal{A}_R(X)$ we have $(\delta + fg)(x,x) = 1 + f(x,x)g(x,x) \in R^*$, for all $x \in X$, thus $\delta + fg \in \mathcal{A}_R^*(X)$, implying that $f \in \mathrm{rad}(\mathcal{A}_R(X))$. Conversely, if $f \in \mathcal{A}_R(X)$ such that $f(x,x) \notin \mathrm{rad}(R)$ for some $x \in X$, then there is $a \in R$ such that $1 + a \cdot f(x,x) \notin R^*$. Hence picking $g \in \mathcal{A}_R(X)$ such that $g(x,x) = a$, we infer $(\delta + fg)(x,x) \notin R^*$, thus $\delta + fg \notin \mathcal{A}_R^*(X)$, hence $f \notin \mathrm{rad}(\mathcal{A}_R(X))$. ♯

Hence we have $\mathcal{A}_R(X)/\mathcal{J}_R(X) \cong \prod_{x \in X} R/\mathrm{rad}(R)$. Moreover, letting $\mathcal{U}_R(X) := \{f \in \mathcal{A}_R(X); f(x,x) = 0 \text{ for all } x \in X\} \subseteq \mathcal{J}_n(R)$, then from $(fg)(x,x) = f(x,x)g(x,x)$, for all $f, g \in \mathcal{A}_R(X)$ and $x \in X$, we infer that $\mathcal{U}_R(X) \trianglelefteq \mathcal{A}_R(X)$ is an ideal, and we have $\mathcal{A}_R(X)/\mathcal{U}_R(X) \cong \prod_{x \in X} R$.

(8.2) Zeta and Möbius functions. a) Let X be a locally finite partially ordered set, and let $\zeta \in \mathcal{A}(X) = \mathcal{A}_{\mathbb{Z}}(X)$ be the **zeta function** of X, that is the indicator function of the partial order, given by $\zeta(x,y) = 1$ whenever $x \leq y \in X$, and $\zeta(x,y) = 0$ whenever $x \not\leq y \in X$; in particular, we have $\zeta \in \mathcal{A}^*(X)$. Note that restricting ζ to an interval in X just yields the zeta function of the interval. The zeta function is related to chains:

i) Induction on $k \in \mathbb{N}$ yields $\zeta^k(x,y) = \sum_{x=z_0\leq z_1\leq\cdots\leq z_k=y} \prod_{i=1}^{k} \zeta(z_{i-1}, z_i) = \sum_{x=z_0\leq z_1\leq\cdots\leq z_k=y} 1$, for $x \leq y \in X$. Thus for $k \in \mathbb{N}_0$ we infer that $\zeta^k(x,y) \in \mathbb{N}_0$

is the number of **multichains** of length k, that is chains with repeated entries, between x and y; in particular, $\zeta^2(x,y) = |[x,y]|$ for all $x \leq y \in X$.

ii) We have $(\zeta - \delta)(x,x) = 0$, and $(\zeta - \delta)(x,y) = 1$ for all $x < y \in X$; in particular, we have $\zeta - \delta \in \mathcal{J}(X) = \mathcal{J}_{\mathbb{Z}}(X)$. For this function, induction on $k \in \mathbb{N}$ yields $(\zeta - \delta)^k(x,y) = \sum_{x=z_0 \leq z_1 \leq \cdots \leq z_k = y} \prod_{i=1}^{k} (\zeta - \delta)(z_{i-1}, z_i) = \sum_{x=z_0 < z_1 < \cdots < z_k = y} 1 \in \mathbb{N}_0$, for all $x \leq y \in X$, saying that for $k \in \mathbb{N}_0$ there are precisely $(\zeta - \delta)^k(x,y) \in \mathbb{N}_0$ chains of length k between x and y. In particular, we have $(\zeta - \delta)^k(x,y) = 0$ if and only if $k \geq l+1$, where $l := l(x,y) \in \mathbb{N}_0$, saying that $\zeta - \delta$ is **locally nilpotent**, that is nilpotent on any interval $[x,y]$.

iii) We have $(2\delta - \zeta)(x,x) = 1$, and $(2\delta - \zeta)(x,y) = -1$ for all $x < y \in X$, hence again $2\delta - \zeta \in \mathcal{A}^*(X)$. Then we have $(2\delta - \zeta)^{-1} = \sum_{i \geq 0} (\zeta - \delta)^i \in \mathcal{A}(X)$, which is indeed well-defined inasmuch it says that $(2\delta - \zeta)^{-1}(x,y) \in \mathbb{N}_0$ coincides with the number of all chains between $x, y \in X$: We may assume that $x \leq y$, and letting $l := l(x,y) \in \mathbb{N}_0$ we have $(\zeta - \delta)^{l+1}|_{[x,y]} = 0$, hence upon restriction to $[x,y]$ we have $(2\delta - \zeta) \cdot \sum_{i=0}^{l} (\zeta - \delta)^i = (\delta - (\zeta - \delta)) \cdot \sum_{i=0}^{l} (\zeta - \delta)^i = \delta - (\zeta - \delta)^{l+1} = \delta$, and thus $(2\delta - \zeta)^{-1}(x,y) = \sum_{i=0}^{l} (\zeta - \delta)^i(x,y)$.

b) The zeta function $\zeta \in \mathcal{A}(X)$ being invertible, we let $\mu := \zeta^{-1} \in \mathcal{A}(X)$ be the **Möbius function** of X. More precisely, we have $\mu(x,x) = 1$ for all $x \in X$, and for $x < y \in X$ using the right inversion formula we get $\mu(x,y) = -\sum_{x < z \leq y} \mu(z,y)$, or equivalently $\sum_{x \leq z \leq y} \mu(z,y) = 0$, while the left inversion formula yields $\mu(x,y) = -\sum_{x \leq z < y} \mu(x,z)$, or equivalently $\sum_{x \leq z \leq y} \mu(x,z) = 0$; in particular, $\mu(x,y) \in \mathbb{Z}$ can be computed by induction on $l(x,y) \in \mathbb{N}_0$.

Letting $x \leq y \in X$ and $l := l(x,y) \in \mathbb{N}_0$ we have $(\zeta - \delta)^{l+1}|_{[x,y]} = 0$, hence upon restriction to $[x,y]$ we have $\zeta \cdot \sum_{i=0}^{l} (\delta - \zeta)^i = (\delta - (\delta - \zeta)) \cdot \sum_{i=0}^{l} (\delta - \zeta)^i = \delta - (\delta - \zeta)^{l+1} = \delta$. Hence we have **Hall's Theorem [1939]** $\mu(x,y) = \zeta^{-1}(x,y) = \sum_{i=0}^{l} (\delta - \zeta)^i(x,y) = \sum_{i=0}^{l} (-1)^i \cdot (\zeta - \delta)^i(x,y) = \sum_{i \geq 0} (-1)^i \cdot (\zeta - \delta)^i(x,y)$, expressing $\mu(x,y)$ as the alternating sum over the number of chains of lengths $i \geq 0$ between $x, y \in X$, in other words $\mu = \sum_{i \geq 0} (-1)^i \cdot (\zeta - \delta)^i \in \mathcal{A}(X)$.

c) Given $f \in \mathcal{A}(X)$, multiplication with ζ yields $f^+ := f\zeta \in \mathcal{A}(X)$ and $f_+ := \zeta f \in \mathcal{A}(X)$ given by $(f^+)(x,y) = \sum_{x \leq z \leq y} f(x,z)\zeta(z,y) = \sum_{x \leq z \leq y} f(x,z)$ and $(f_+)(x,y) = \sum_{x \leq z \leq y} \zeta(x,z) f(z,y) = \sum_{x \leq z \leq y} f(z,y)$, for all $x \leq y \in X$. Then we have the **Möbius inversion** formulae **[Rota, 1964]** $f^+\mu = f \in \mathcal{A}(X)$, that is $f(x,y) = \sum_{x \leq z \leq y} f^+(x,z)\mu(z,y)$ for all $x \leq y \in X$, and $\mu f_+ = f \in \mathcal{A}(X)$, that is $f(x,y) = \sum_{x \leq z \leq y} \mu(x,z) f_+(z,y)$ for all $x \leq y \in X$.

(8.3) Modules for the incidence algebra. Let X be a locally finite partially ordered set, let $R \neq \{0\}$ be a commutative ring. We are looking for $\mathcal{A}_R(X)$-modules, apart from those coming from the quotients $\mathcal{A}_R(X)/\mathcal{J}_R(X)$ and $\mathcal{A}_R(X)/\mathcal{U}_R(X)$ considered earlier. We mainly restrict ourselves to right $\mathcal{A}_R(X)$-modules, which we usually just call $\mathcal{A}_R(X)$-modules, and where for left $\mathcal{A}_R(X)$-modules we typically have the 'dual' picture. The first natural choice is to look for right and left ideals in $\mathcal{A}_R(X)$:

a) For a subset $Y \subseteq X$ let $\mathcal{I}_R(Y,X) := \{f \in \mathcal{A}_R(X); f(x,?) \neq 0 \text{ only if } x \in Y\}$. Then $\mathcal{I}_R(Y,X) \leq \mathcal{A}_R(X)$ is a right ideal: For $f \in \mathcal{I}_R(Y,X)$ and $g \in \mathcal{A}_R(X)$, from $(fg)(x,y) = \sum_{x \leq z \leq y} f(x,z)g(z,y)$, for all $x \leq y \in X$, we get $(fg)(x,y) = 0$ if $x \notin Y$, thus $fg \in \mathcal{I}_R(Y,X)$. We have $\mathcal{A}_R(X) \cong \mathcal{I}_R(Y,X) \oplus \mathcal{I}_R(X \setminus Y, X)$, thus $\mathcal{I}_R(Y,X)$ is a **projective** $\mathcal{A}_R(X)$-module. Its Jacobson radical is $\mathrm{rad}(\mathcal{I}_R(Y,X)) = \mathcal{I}_R(Y,X) \cap \mathcal{J}_R(X) = \{f \in \mathcal{I}_R(Y,X); f(x,x) \in \mathrm{rad}(R) \text{ for } x \in Y\}$, hence we have $\mathcal{I}_R(Y,X)/\mathrm{rad}(\mathcal{I}_R(Y,X)) \cong \prod_{x \in Y} R/\mathrm{rad}(R)$. Moreover, we have $\mathcal{I}_R(Y,X) \cong \prod_{x \in Y} \mathcal{I}_R(\{y\},X)$ as $\mathcal{A}_R(X)$-modules.

Similarly, let $\mathcal{I}_R(X,Y) := \{f \in \mathcal{A}_R(X); f(?,y) \neq 0 \text{ only if } y \in Y\}$. Then $\mathcal{I}_R(X,Y) \leq \mathcal{A}_R(X)$ is a left ideal: For $f \in \mathcal{I}_R(X,Y)$ and $g \in \mathcal{A}_R(X)$, from $(gf)(x,y) = \sum_{x \leq z \leq y} g(x,z)f(z,y)$, for all $x \leq y \in X$, we get $(gf)(x,y) = 0$ if $y \notin Y$, thus $gf \in \mathcal{I}_R(X,Y)$. We have $\mathcal{A}_R(X) \cong \mathcal{I}_R(X,Y) \oplus \mathcal{I}_R(X, X \setminus Y)$, thus $\mathcal{I}_R(X,Y)$ is a projective left $\mathcal{A}_R(X)$-module. Its Jacobson radical is $\mathrm{rad}(\mathcal{I}_R(X,Y)) = \mathcal{I}_R(X,Y) \cap \mathcal{J}_R(X) = \{f \in \mathcal{I}_R(X,Y); f(x,x) \in \mathrm{rad}(R) \text{ for } x \in Y\}$, hence we have $\mathcal{I}_R(X,Y)/\mathrm{rad}(\mathcal{I}_R(X,Y)) \cong \prod_{x \in Y} R/\mathrm{rad}(R)$. Moreover, we have $\mathcal{I}_R(X,Y) \cong \prod_{x \in Y} \mathcal{I}_R(X,\{x\})$ as left $\mathcal{A}_R(X)$-modules.

b) For a coideal $Y \subseteq X$ the left ideal $\mathcal{I}_R(X,Y)$ also is a right ideal: For $f \in \mathcal{I}_R(X,Y)$ and $g \in \mathcal{A}_R(X)$, we have $(fg)(x,y) = \sum_{x \leq z \leq y} f(x,z)g(z,y)$, for all $x \leq y \in X$, hence, since for $y \notin Y$ we also have $z \notin Y$ for all $z \leq y$, we conclude $(fg)(x,y) = 0$ if $y \notin Y$, thus $fg \in \mathcal{I}_R(X,Y)$.

Combining these constructions, for any subset $Z \subseteq X$ and any coideal $Y \subseteq X$ we get the right ideal $\mathcal{I}_R(Z,Y) := \mathcal{I}_R(Z,X) \cap \mathcal{I}_R(X,Y) \leq \mathcal{A}_R(X)$. Moreover, we have $\mathcal{I}_R(Z,Y) \cong \prod_{x \in Z} \mathcal{I}_R(\{x\},Y)$ as well as $\mathcal{I}_R(Z,X)/\mathcal{I}_R(Z,Y) \cong \prod_{x \in Z} \mathcal{I}_R(\{x\},X)/\mathcal{I}_R(\{x\},Y)$ as $\mathcal{A}_R(X)$-modules. In this sense, coideals naturally provide submodules and quotient modules of the projective $\mathcal{A}_R(X)$-modules $\mathcal{I}_R(Z,X)$. Note that for any ideal $Y \subseteq X$ the complement $X \setminus Y \subseteq X$ is a coideal, so that this essentially also applies to ideals.

(8.4) The Möbius space. Let X be a locally finite partially ordered set, let $R \neq \{0\}$ be a commutative ring, and let $\mathcal{F}_R(X) := \mathrm{Maps}(X,R)$ be the associated **Möbius space**, which is an R-module with respect to pointwise addition and scalar multiplication.

a) Let X additionally have a zero element $\underline{0}$. Then we get an injective R-linear map $\mathcal{F}_R(X) \to \mathcal{A}_R(X) \colon f \mapsto f^\circ$, where $f^\circ(x,y) := \delta(\underline{0},x)f(y)$, for all $x,y \in X$. Its image $\mathcal{F}_R^\circ(X) := \{f \in \mathcal{A}_R(X); f(x,?) \neq 0 \text{ only if } x = \underline{0}\} = \mathcal{I}_R(\{\underline{0}\},X) \leq \mathcal{A}_R(X)$ is a right ideal and a projective $\mathcal{A}_R(X)$-module, whose Jacobson radical is given as $\mathrm{rad}(\mathcal{F}_R^\circ(X)) = \{f \in \mathcal{F}_R^\circ(X); f(\underline{0},\underline{0}) \in \mathrm{rad}(R)\}$.

We have $(f^\circ g)(\underline{0},y) = \sum_{z \leq y} f(z)g(z,y)$, for $f \in \mathcal{F}_R(X)$ and $g \in \mathcal{A}_R(X)$, and all $y \in X$. This shows that $\mathcal{F}_R(X)$ becomes an $\mathcal{A}_R(X)$-module by letting $f * g \in \mathcal{F}_R(X)$ be defined as $(f * g)(x) := \sum_{z \leq x} f(z)g(z,x)$, for $x \in X$. Then $\mathcal{F}_R(X) \to \mathcal{A}_R(X) \colon f \mapsto f^\circ$ is an $\mathcal{A}_R(X)$-module isomorphism, that is $(f*g)^\circ = f^\circ g \in \mathcal{A}_R(X)$. In particular, $\mathcal{F}_R(X)$ is a projective $\mathcal{A}_R(X)$-module, such that $\mathrm{rad}(\mathcal{F}_R(X)) = \{f \in \mathcal{F}_R(X); f(\underline{0}) \in \mathrm{rad}(R)\}$.

b) If X instead has a one element $\underline{1}$, we get an injective R-linear map $\mathcal{F}_R(X) \to \mathcal{A}_R(X)\colon f \mapsto f^\bullet$, where now $f^\bullet(x,y) := \delta(y,\underline{1})f(x)$, for all $x,y \in X$. Its image $\mathcal{F}_R^\bullet(X) = \{f \in \mathcal{A}_R(X); f(?,y) \neq 0 \text{ only if } y = \underline{1}\} = \mathcal{I}_R(X,\{\underline{1}\}) \leq \mathcal{A}_R(X)$ is a left ideal and a projective left $\mathcal{A}_R(X)$-module, whose Jacobson radical is given as $\operatorname{rad}(\mathcal{F}_R^\bullet(X)) = \mathcal{F}_R^\bullet(X) \cap \mathcal{J}_R(X) = \{f \in \mathcal{F}_R^\bullet(X); f(\underline{1},\underline{1}) \in \operatorname{rad}(R)\}$.

We have $(gf^\bullet)(x,\underline{1}) = \sum_{z\geq x} g(x,z)f(z)$, for $f \in \mathcal{F}_R(X)$ and $g \in \mathcal{A}_R(X)$, and all $x \in X$. This shows that $\mathcal{F}_R(X)$ also becomes a left $\mathcal{A}_R(X)$-module by letting $g * f \in \mathcal{F}_R(X)$ be defined as $(g*f)(x) := \sum_{z\geq x} g(x,z)f(z)$, for $x \in X$. Then $\mathcal{F}_R(X) \to \mathcal{A}_R(X)\colon f \mapsto f^\bullet$ is an $\mathcal{A}_R(X)$-module isomorphism, that is $(g*f)^\bullet = gf^\bullet \in \mathcal{A}_R(X)$. In particular, $\mathcal{F}_R(X)$ is a projective left $\mathcal{A}_R(X)$-module, such that $\operatorname{rad}(\mathcal{F}_R(X)) = \{f \in \mathcal{F}_R(X); f(\underline{1}) \in \operatorname{rad}(R)\}$.

But note that $\mathcal{F}_R(X)$ in general does not become an $\mathcal{A}_R(X)$-$\mathcal{A}_R(X)$-bimodule.

c) In terms of $\mathcal{F}(X) = \mathcal{F}_\mathbb{Z}(X)$ Möbius inversion reads as follows:

If X has a zero element, using the $\mathcal{A}(X)$-module structure on $\mathcal{F}(X)$, we let $f^+ := f * \zeta \in \mathcal{F}(X)$, hence $f^+(x) = \sum_{z\leq x} f(z)$, for all $x \in X$; note that $(f^+)^\circ = (f^\circ)^+ \in \mathcal{A}(X)$. Thus Möbius inversion yields $f^+ * \mu = (f*\zeta)*\mu = f*(\zeta\mu) = f*\delta = f \in \mathcal{F}(X)$, hence $f(x) = \sum_{z\leq x} f^+(z)\mu(z,x)$, for all $x \in X$.

If X has a one element, using the left $\mathcal{A}(X)$-module structure on $\mathcal{F}(X)$, we let $f_+ := \zeta * f \in \mathcal{F}(X)$, hence $f_+(x) = \sum_{z\geq x} f(z)$, for all $x \in X$; note that $(f_+)^\bullet = (f^\bullet)_+ \in \mathcal{A}(X)$. Thus Möbius inversion yields $\mu * f_+ = \mu * (\zeta * f) = (\mu\zeta) * f = \delta * f = f \in \mathcal{F}(X)$, hence $f(x) = \sum_{z\geq x} \mu(x,z)f_+(z)$, for all $x \in X$.

(8.5) Finite partially ordered sets. a) We consider the case of X being finite. In order to do so, let $X = \{x_1,\ldots,x_n\}$, where $n := |X| \in \mathbb{N}$, where we assume that $x_i \leq x_j$ only if $i \leq j \in \{1,\ldots,n\}$; note that this amounts to **refining** the given partial order on X to a total order. This lends to the following description of $\mathcal{A}_R(X)$, where $R \neq \{0\}$ is a commutative ring:

For $n \in \mathbb{N}$ let $\mathcal{T}_n(R) := \{A = [a_{ij}]_{ij} \in R^{n\times n}; a_{ij} = 0 \text{ for all } 1 \leq j < i \leq n\}$ be the set of all upper triangular $(n\times n)$-matrices over R, and let $\mathcal{U}_n(R) := \{A \in \mathcal{T}_n(R); a_{ii} = 0 \text{ for all } i\} \trianglelefteq \mathcal{T}_n(R)$ be the ideal consisting of all strictly upper triangular matrices. Then the injective R-linear map $\mathcal{A}_R(X) \to \mathcal{T}_n(R)\colon f \mapsto [f(x_i,x_j)]_{ij}$ translates convolutional multiplication into matrix multiplication, hence is a homomorphism of R-algebras. Thus we may and will view $\mathcal{A}_R(X)$ as an R-subalgebra of $\mathcal{T}_n(R)$.

b) In particular, we have $\mathcal{A}_R(X) = \mathcal{T}_n(R)$ if and only if X is totally ordered. Thus $\mathcal{T}_n(R)$ can be considered as an incidence algebra, too. Doing so, we infer that $\mathcal{T}_n^*(R) = \{A \in \mathcal{T}_n(R); a_{ii} \in R^* \text{ for all } i\}$, and that $\operatorname{rad}(\mathcal{T}_n(R)) = \mathcal{J}_n(R) := \{A \in \mathcal{T}_n(R); a_{ii} \in \operatorname{rad}(R) \text{ for all } i\}$. Moreover, we have $\mathcal{A}_R^*(X) = \mathcal{A}_R(X) \cap \mathcal{T}_n^*(R)$, and $\mathcal{J}_R(X) = \mathcal{A}_R(X) \cap \mathcal{J}_n(R)$, and $\mathcal{U}_R(X) := \mathcal{A}_R(X) \cap \mathcal{U}_n(R)$.

We show that $\mathcal{U}_R(X) \trianglelefteq \mathcal{A}_R(X)$ actually is a nilpotent ideal:

For $i,j \in \{1,\ldots,n\}$ let $E_{ij} \in R^{n\times n}$ be the associated **matrix unit**, that is having $[k,l]$-entry $\delta_{ik}\delta_{jl} \in R$; recall the multiplication rule $E_{ij}E_{kl} = \delta_{jk}\cdot E_{il}$.

Then we have $\mathcal{A}_R(X) = \langle E_{xy}; x, y \in X, x \leq y\rangle_R$ and $\mathcal{U}_R(X) = \langle E_{xy}; x, y \in X, x < y\rangle_R$, where the given generating sets are even R-bases. For $k \in \mathbb{N}_0$ we have $\mathcal{U}_R(X)^k \leq \langle E_{ij}; 1 \leq i \leq j-k \leq n\rangle_R = \{A \in \mathcal{T}_n(R); a_{ij} = 0 \text{ for all } i > j-k\}$, and we have $E_{xy} \in \mathcal{U}_R(X)^k$, where $x \leq y$, if and only if there is a chain of length k between x and y, that is $l(x,y) \geq k$. Thus we have $\mathcal{U}_R(X)^k = \{0\}$ if and only if $k \geq l(X)+1$. ♯

As for the Möbius space, writing $f \in \mathcal{F}_R(X)$ as $[f(x_1), \ldots, f(x_n)] \in R^n$, the embedding $\mathcal{F}_R^\circ(X) \subseteq \mathcal{A}_R(X) \subseteq \mathcal{T}_n(R)$ is given by associating $f \in \mathcal{F}_R(X)$ with the matrix $\sum_{i=1}^n f(x_i)E_{1i} \in \mathcal{T}_n(R)$, which has $[f(x_1), \ldots, f(x_n)]$ as its first row and zero entries otherwise. Similarly, the embedding $\mathcal{F}_R^\bullet(X) \subseteq \mathcal{A}_R(X) \subseteq \mathcal{T}_n(R)$ is given by associating $f \in \mathcal{F}_R(X)$ with the matrix $\sum_{i=1}^n f(x_i)E_{in} \in \mathcal{T}_n(R)$, which has $[f(x_1), \ldots, f(x_n)]^{\mathrm{tr}}$ as its last column and zero entries otherwise.

c) Now the Möbius function $\mu = \zeta^{-1} \in \mathcal{A}(X) \subseteq \mathcal{T}_n = \mathcal{T}_n(\mathbb{Z})$ can be computed in matrix terms, by inverting the matrix associated with ζ. In particular, if $\underline{0} \in X$ is a zero element, then the values $\mu(\underline{0}, ?)$ are given in the first row of the matrix of μ, and if $\underline{1} \in X$ is a one element, then the values $\mu(?, \underline{1})$ are given in the last column of the matrix of μ.

For example, the partially ordered set $X := [\emptyset, \{1\}, \{2\}, \{2,3\}, \{1,2,3\}]$, see (5.2), yields

$$\mathcal{A}_R(X) \cong \begin{bmatrix} * & * & * & * & * \\ & * & . & . & * \\ & & * & * & * \\ & & & * & * \\ & & & & * \end{bmatrix} \subseteq \mathcal{T}_5(R).$$

The matrices of ζ and μ are given as

$$\zeta \mapsto \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & . & . & 1 \\ & & 1 & 1 & 1 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}, \qquad \mu \mapsto \begin{bmatrix} 1 & -1 & -1 & . & 1 \\ & 1 & . & . & -1 \\ & & 1 & -1 & . \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix}.$$

(8.6) The Möbius algebra. To compute Möbius functions of lattices, the conventional recursive procedure can be adjusted to involve much fewer terms than for partially ordered sets in general. To this end, it suffices to assume only that X is a locally finite partially ordered set having meets and a one element. If X additionally has a zero element, then by local finiteness X is finite, and conversely if X is finite then it has a zero element, thus in this case by the existence of meets X has joins as well, hence is a lattice; see Exercise (20.3).

a) Let $R \neq \{0\}$ be a commutative ring. Then we define **Möbius multiplication** on $\mathcal{F}_R(X)$ as follows: For $f, g \in \mathcal{F}_R(X)$ let $f \wedge g \colon X \to R \colon y \mapsto \sum_{x,x' \in X, x \wedge x' = y} f(x)g(x')$. From local finiteness and the existence of a one element we infer that this is indeed well-defined. Moreover, we have $f \wedge g = g \wedge f$, and $(f \wedge g) \wedge h = f \wedge (g \wedge h) \colon y \mapsto \sum_{x,x',x'' \in X, x \wedge x' \wedge x'' = y} f(x)g(x')h(x'')$, for all $f, g, h \in \mathcal{F}_R(X)$, that is commutativity and associativity hold. For the map $\delta_{\underline{1}} \colon X \to R \colon x \mapsto \delta(x, \underline{1})$, where $\delta \in \mathcal{A}_R(X)$ is as above, we have $(f \wedge \delta_{\underline{1}})(y) = \sum_{x,x' \in X, x \wedge x' = y} f(x)\delta(x', \underline{1}) = f(y)$, for all $y \in X$, hence $\delta_{\underline{1}} \in \mathcal{A}(X)$ is a neutral element with respect to Möbius multiplication. Since we have distributivity, and $c(f \wedge g) = (cf) \wedge g \in \mathcal{F}_R(X)$, for all $c \in R$, we conclude that $\mathcal{F}_R(X)$ is a commutative R-algebra, called the associated **Möbius algebra**.

More generally, for $y \in X$ let $\delta_y \colon X \to R \colon x \mapsto \delta(x, y)$ and $\mu_y \colon X \to R \colon x \mapsto \mu(x, y)$, where $\mu \in \mathcal{A}_R(X)$ is the Möbius function of X as above; recall that $\mu(x, y) \in \mathbb{Z}$ for all $x, y \in X$, hence we may consider μ as a function with values in R. Then for any fixed $z \in X$ we have $\sum_{y \leq z} \mu_y(x) = \sum_{y \leq z} \mu(x, y) = \sum_{x \leq y \leq z} \mu(x, y) = \delta(x, z) = \delta_z(x)$, for all $x \in X$, saying that $\sum_{y \leq z} \mu_y = \delta_z$; in particular we have $\sum_{y \in X} \mu_y = \delta_{\underline{1}}$.

We have $(\mu_y \wedge \mu_{y'})(w) = \sum_{x \leq y, x' \leq y', x \wedge x' = w} \mu(x, y)\mu(x', y')$, for $w, y, y' \in X$, hence $(\mu_y \wedge \mu_{y'})(w) \neq 0$ only if $w \leq y \wedge y'$. In this case, for all $v \leq y \wedge y'$ we have $\sum_{v \leq w \leq y \wedge y'} (\mu_y \wedge \mu_{y'})(w) = \sum_{v \leq w \leq y \wedge y'} \sum_{x \leq y, x' \leq y', x \wedge x' = w} \mu(x, y)\mu(x', y')$, hence changing the order of summation the right hand side can be written as $\sum_{v \leq x \leq y, v \leq x' \leq y'} \mu(x, y)\mu(x', y') = \sum_{v \leq x \leq y} \mu(x, y) \cdot \sum_{v \leq x' \leq y'} \mu(x', y') = \delta_{v,y} \cdot \delta_{v,y'}$. Thus, if $y \neq y'$ then by induction on $l(v, y \wedge y') \in \mathbb{N}_0$ we infer $(\mu_y \wedge \mu_{y'})(v) = 0$, and hence $\mu_y \wedge \mu_{y'} = 0 \in \mathcal{F}_R(X)$.

Similarly, for $y = y'$ we get $\sum_{v \leq w \leq y} (\mu_y \wedge \mu_y)(w) = \delta_{v,y} = \sum_{v \leq w \leq y} \mu(w, y) = \sum_{v \leq w \leq y} \mu_y(w)$, which by induction on $l(v, y) \in \mathbb{N}_0$ entails $(\mu_y \wedge \mu_y)(v) = \mu_y(v)$, hence $\mu_y \wedge \mu_y = \mu_y \in \mathcal{F}_R(X)$. Thus, in conclusion we have $\mu_y \wedge \mu_{y'} = \delta_{y,y'}\mu_y \in \mathcal{F}_R(X)$, saying that the $\{\mu_y \in \mathcal{F}_R(X); y \in X\}$ form a decomposition of $\delta_{\underline{1}}$ into pairwise orthogonal idempotents.

In particular, if K is a field and X is finite, then from $\dim_K(\mathcal{F}_K(X)) = |X|$ we infer that $\{\mu_y \in \mathcal{F}_K(X); y \in X\}$ are the (centrally) primitive idempotents, thus $\mathcal{F}_K(X)$ is a split semisimple K-algebra of shape $\mathcal{F}_K(X) \cong \bigoplus_{x \in X} K$.

b) For $\underline{1} \neq z \in X$ we have $\delta_z \wedge \mu_{\underline{1}} = \sum_{y \leq z} \mu_y \wedge \mu_{\underline{1}} = 0 \in \mathcal{F}_R(X)$; note that for $z = \underline{1}$ we just get the trivial formula $\delta_{\underline{1}} \wedge \mu_{\underline{1}} = \sum_{y \in X} \mu_y \wedge \mu_{\underline{1}} = \mu_{\underline{1}}$. Thus, we have **Weisner's Theorem [1935]**, saying that for $\underline{1} \neq z \in X$ and all $y \in X$ we have $0 = (\delta_z \wedge \mu_{\underline{1}})(y) = \sum_{w,x \in X, w \wedge x = y} \delta(w, z)\mu(x, \underline{1}) = \sum_{x \in X, x \wedge z = y} \mu(x, \underline{1})$.

In particular, if X is finite such that $\underline{0} \neq \underline{1}$, letting $y = \underline{0}$, for $z = \underline{0}$ we recover $\sum_{x \in X} \mu(x, \underline{1}) = 0$, and whenever $\underline{0} \not\lessdot \underline{1}$ for $\underline{0} \lessdot z$ we get $\sum_{z \not\leq x} \mu(x, \underline{1}) = 0$.

Note that, similarly, if X has joins and a zero element, for $\underline{0} \neq z \in X$ and all $y \in X$ we have $\sum_{x \in X, x \vee z = y} \mu(\underline{0}, x) = 0$. In particular, again, if X is finite such that $\underline{0} \neq \underline{1}$, letting $y = \underline{1}$, for $z = \underline{0}$ we recover $\sum_{x \in X} \mu(\underline{0}, x) = 0$, and whenever $\underline{0} \not\lessdot \underline{1}$ for $z \lessdot \underline{1}$ we get $\sum_{x \not\leq z} \mu(\underline{0}, x) = 0$.

(8.7) Möbius functions of lattices. We now proceed to the promised simplification to compute Möbius functions of lattices, where it still suffices to assume that X is a locally finite partially ordered set having meets and a one element.

Let $M \subseteq X \setminus \{\underline{1}\}$ such that for all $\underline{1} \neq x \in X$ there is $y \in M$ such that $x \leq y$;

in particular M contains all $x \in X$ such that $x \lessdot \underline{1}$. Then for $y \in X$ we have $\delta_{\underline{1}} - \delta_y = \sum_{x \in X} \mu_x - \sum_{x \leq y} \mu_x = \sum_{x \not\leq y} \mu_x$. Hence, using the fact that the μ_x are pairwise orthogonal idempotents and the special choice of M, this yields $\bigwedge_{y \in M}(\delta_{\underline{1}} - \delta_y) = \bigwedge_{y \in M}(\sum_{x \not\leq y} \mu_x) = \sum_{x \not\leq y \text{ for all } y \in M} \mu_x = \mu_{\underline{1}} \in \mathcal{F}_R(X)$. Thus from $\bigwedge_{y \in M}(\delta_{\underline{1}} - \delta_y) = \sum_{L \subseteq M} (-1)^{|L|} \cdot \bigwedge_{y \in L} \delta_y \in \mathcal{F}_R(X)$, where as usual we define the empty product to be the identity $\delta_{\underline{1}}$, by evaluating at $x \in X$, we get $\mu(x, \underline{1}) = \sum_{L \subseteq M} (-1)^{|L|} \cdot (\bigwedge_{y \in L} \delta_y)(x) = \sum_{L \subseteq M, \bigwedge L = x} (-1)^{|L|}$. In particular, for any $x \in X$ such that $\bigwedge\{y \in X; y \lessdot \underline{1}\} \not\leq x$ we have $\mu(x, \underline{1}) = 0$.

If X is finite, then for the zero element we get the **cross-cut theorem** $\mu(\underline{0}, \underline{1}) = \sum_{L \subseteq M, \bigwedge L = \underline{0}} (-1)^{|L|}$; if $\bigwedge\{x \in X; x \lessdot \underline{1}\} \neq \underline{0}$ then we have $\mu(\underline{0}, \underline{1}) = 0$.

Note that, similarly, if X has joins and a zero element, and if $M \subseteq X \setminus \{\underline{0}\}$ is such that for all $\underline{0} \neq x \in X$ there is $y \in M$ such that $y \leq x$, then for all $x \in X$ we have $\mu(\underline{0}, x) = \sum_{L \subseteq M, \bigvee L = x} (-1)^{|L|}$. In particular, for any $x \in X$ such that $x \not\leq \bigvee\{y \in X; \underline{0} \lessdot y\}$ we have $\mu(\underline{0}, x) = 0$; if X is finite, then $\mu(\underline{0}, \underline{1}) = \sum_{L \subseteq M, \bigvee L = \underline{1}} (-1)^{|L|}$, and if $\bigvee\{x \in X; \underline{0} \lessdot x\} \neq \underline{1}$ then $\mu(\underline{0}, \underline{1}) = 0$.

9 Möbius functions

(9.1) Möbius functions of products. a) Let X' and X'' be partially ordered sets. Then the Cartesian product $X := X' \times X''$ becomes a partially ordered set, called the **direct product** of X' and X'', with respect to componentwise comparison, that is we let $x := [x', x''] \leq [y', y''] =: y \in X$ if and only if $x' \leq y' \in X'$ and $x'' \leq y'' \in X''$.

Then intervals in X are direct products again, namely $[x, y] = [x', y'] \times [x'', y'']$ whenever $x \leq y \in X$. Moreover, X is a lattice if X' and X'' are; X is graded if X' and X'' are; X is locally finite if X' and X'' are; and if X' and X'' are finite then we have $l(X) = l(X') + l(X'')$.

Let X' and X'' be locally finite. We aim at finding the Möbius function of $X = X' \times X''$: Given the Möbius functions μ' and μ'' of X' and X'', respectively, we get $\sum_{x \leq z \leq y} \mu'(x', z') \cdot \mu''(x'', z'') = \sum_{x' \leq z' \leq y'} \sum_{x'' \leq z'' \leq y''} \mu'(x', z') \cdot \mu''(x'', z'') = \sum_{x' \leq z' \leq y'} \mu'(x', z') \cdot \sum_{x'' \leq z'' \leq y''} \mu''(x'', z'') = \delta_{x', y'} \cdot \delta_{x'', y''} = \delta_{x, y}$, for $x, y \in X$, hence the Möbius function of X is given as the product $\mu(x, y) = \mu'(x', y') \cdot \mu''(x'', y'')$ of the component Möbius functions, for all $x, y \in X$.

b) We consider the direct product $X := \prod_{i=1}^k [0, n_i]$ of the totally ordered intervals $[0, n_i]$, where $n_1, \ldots, n_k \in \mathbb{N}$ and $k \in \mathbb{N}$. We provide to ways to determine its Möbius function μ, in particular showing that the cross-cut theorem indeed helps to avoid too many explicit computations:

i) Firstly, since μ is determined locally, it suffices to find $\mu(\underline{0}, [n_1, \ldots, n_k]) = \prod_{i=1}^k \mu(0, n_k)$. We show that $\mu(\underline{0}, [n_1, \ldots, n_k]) \neq 0$ if and only if $n_i \leq 1$ for all $i \in \{1, \ldots, k\}$, in which case $\mu(\underline{0}, [n_1, \ldots, n_k]) = (-1)^\epsilon$, where $\epsilon := \sum_{i=1}^k n_i$:

For the totally ordered set $[0,n]$, where $n \in \mathbb{N}_0$, the matrices of ζ and μ are

$$\zeta \mapsto \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ & 1 & 1 & \dots & 1 \\ & & \ddots & \ddots & \vdots \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}, \qquad \mu \mapsto \begin{bmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix}.$$

Hence we have $\mu(0,0) = 1$ and $\mu(0,1) = -1$, as well as $\mu(0,n) = 0$ for $n \geq 2$. $\sharp$

Note that this also describes the Möbius function of the totally ordered set $\mathbb{N}_0$.

ii) Secondly, we apply the cross-cut theorem: We have $M := \{[m_1,\dots,m_k] \in X; \underline{0} \lessdot [m_1,\dots,m_k]\} = \{e_1,\dots,e_k\}$, where $e_i \in \mathbb{Z}^k$ denotes the i-th **unit vector**, and hence $\bigvee M = [1,\dots,1] \in X$. Thus we have $\mu(\underline{0},[m_1,\dots,m_k]) = 0$ whenever $\{m_1,\dots,m_k\} \not\subseteq \{0,1\}$. while if $m_i \leq 1$ for all $i \in \{1,\dots,k\}$ then we recover $\mu(\underline{0},[m_1,\dots,m_k]) = \sum_{L \subseteq M, \bigvee L = [m_1,\dots,m_k]} (-1)^{|L|} = (-1)^\epsilon$, where $\epsilon = |\{e_i; m_i = 1\}| = |\sum_{i=1}^k m_i|$.

(9.2) Example: Divisibility lattices. a) To determine the Möbius function of $\mathbb{N}$, partially ordered by divisibility $|$, using the fact that Möbius functions are determined locally, and since for $c \mid d \in \mathbb{N}$ we have $[c,d] \cong [\frac{c}{c},\frac{d}{c}] \cong [1,\frac{d}{c}]$ as partially ordered sets, hence $\mu(c,d) = \mu(1,\frac{d}{c})$, it suffices to compute the **number-theoretic Möbius function** $\mu(n) := \mu(1,n)$ for $n \in \mathbb{N}$:

If $n = \prod_{i=1}^k p_i^{n_i}$ is the prime factorization of n, where $p_1,\dots,p_k \in \mathbb{N}$ are pairwise distinct primes and $n_1,\dots,n_k \in \mathbb{N}$, for some $k \in \mathbb{N}_0$, then we have $c = \prod_{i=1}^k p_i^{c_i} \mid d = \prod_{i=1}^k p_i^{d_i} \mid n$ if and only if $0 \leq c_i \leq d_i \leq n_i$ for all $i \in \{1,\dots,k\}$, showing that, by going over to the multiplicities of the various primes in the prime factorization of n, the interval between 1 and n is isomorphic to the direct product $\prod_{i=1}^k [0,n_i]$ of the totally ordered intervals $[0,n_i]$ as partially ordered sets. Hence we have $\mu(n) = 0$ if $n_i \geq 2$ for some $i \in \{1,\dots,k\}$, that is if n is not squarefree; and if $n = \prod_{i=1}^k p_i$ is squarefree then we have $\mu(n) = (-1)^k$, where k is the number of distinct prime divisors of n. Another related lattice is discussed in Exercise (19.36).

b) We present a few applications of Möbius inversion in number theory: Let $f \in \mathcal{F}(\mathbb{N})$ be a **number-theoretic function**. Since $\mathbb{N}$ has a zero element, $\mathcal{F}(\mathbb{N})$ becomes a right $\mathcal{A}(\mathbb{N})$-module, hence we get $f^+ := f * \zeta \in \mathcal{F}(\mathbb{N})$, where $f^+(n) = \sum_{d \mid n} f(d) = \sum_{d \mid n} f(\frac{n}{d})$, for all $n \in \mathbb{N}$, and thus $f = f^+ * \mu \in \mathcal{F}(\mathbb{N})$, where $(f^+ * \mu)(n) = \sum_{d \mid n} f^+(d)\mu(d,n) = \sum_{d \mid n} f^+(d)\mu(\frac{n}{d}) = \sum_{d \mid n} \mu(d) f^+(\frac{n}{d})$.

For the constant map $\epsilon \in \mathcal{F}(\mathbb{N})$, given by $\epsilon(n) = 1$ for all $n \in \mathbb{N}$, we get $\tau := \epsilon^+ \in \mathcal{F}(\mathbb{N})$, where $\tau(n) = \sum_{d \mid n} 1$ is the **number of divisors** of n; this yields $\tau * \mu = \epsilon \in \mathcal{F}(\mathbb{N})$, that is $\sum_{d \mid n} \tau(d)\mu(\frac{n}{d}) = \sum_{d \mid n} \mu(d)\tau(\frac{n}{d}) = 1$. For the identity map $\mathrm{id}_\mathbb{N} \in \mathcal{F}(\mathbb{N})$ we get $\sigma := \mathrm{id}_\mathbb{N}^+ \in \mathcal{F}(\mathbb{N})$, where $\sigma(n) = \sum_{d \mid n} d$

is the **sum of the divisors** of n; this yields $\sigma * \mu = \mathrm{id}_{\mathbb{N}} \in \mathcal{F}(\mathbb{N})$, that is $\sum_{d \,|\, n} \sigma(d)\mu(\frac{n}{d}) = \sum_{d \,|\, n} \mu(d)\sigma(\frac{n}{d}) = n$.

The other way around, for $n \in \mathbb{N}$ we get $(\epsilon * \mu)(n) = \sum_{d \,|\, n} \mu(\frac{n}{d}) = \sum_{d \,|\, n} \mu(d) = \delta(1, n) = \delta_1(n)$, that is $\epsilon * \mu = \delta_1 \in \mathcal{F}(\mathbb{N})$; this yields $(\delta_1)^+ = \epsilon \in \mathcal{F}(\mathbb{N})$, that is $\sum_{d \,|\, n} \delta_1(d) = \sum_{d \,|\, n} \delta_1(\frac{n}{d}) = 1$. More interestingly, let $\varphi := \mathrm{id}_{\mathbb{N}} * \mu \in \mathcal{F}(\mathbb{N})$ be the **Euler totient function**, where $\varphi(n) = \sum_{d \,|\, n} d \cdot \mu(\frac{n}{d}) = \sum_{d \,|\, n} \mu(d) \cdot \frac{n}{d}$; this yields $\varphi^+ = \mathrm{id}_{\mathbb{N}} \in \mathcal{F}(\mathbb{N})$, that is $\sum_{d \,|\, n} \varphi(d) = \sum_{d \,|\, n} \varphi(\frac{n}{d}) = n$.

We show that $\varphi(n)$ can be interpreted as a cardinality, and in particular has positive values: Let $\widetilde{\varphi} \in \mathcal{F}(\mathbb{N})$ be given as $\widetilde{\varphi}(n) := |\{k \in \{1, \ldots, n\}; \gcd(k, n) = 1\}|$. Then we have $\sum_{d \,|\, n} \widetilde{\varphi}(\frac{n}{d}) = \sum_{d \,|\, n} |\{k \in \{1, \ldots, \frac{n}{d}\}; \gcd(k, \frac{n}{d}) = 1\}| = \sum_{d \,|\, n} |\{k \in \{1, \ldots, n\}; \gcd(k, n) = d\}| = |\coprod_{d \,|\, n} \{k \in \{1, \ldots, n\}; \gcd(k, n) = d\}| = n$, implying that $\widetilde{\varphi} * \zeta = \mathrm{id}_{\mathbb{N}}$, and thus $\widetilde{\varphi} = \mathrm{id}_{\mathbb{N}} * \mu = \varphi \in \mathcal{F}(\mathbb{N})$.

We determine φ explicitly, using prime factorizations: For $n = \prod_{i=1}^k p_i^{n_i} \in \mathbb{N}$, where $p_1, \ldots, p_k \in \mathbb{N}$ are pairwise distinct primes, $n_1, \ldots, n_k \in \mathbb{N}$ and $k \in \mathbb{N}_0$, we get $\varphi(n) = \sum_{[a_1, \ldots, a_k] \in \prod_{i=1}^k [0, \ldots, n_i]} \left(\mu(\prod_{i=1}^k p_i^{a_i}) \cdot \prod_{i=1}^k p_i^{n_i - a_i}\right) = \sum_{[\epsilon_1, \ldots, \epsilon_k] \in \{0,1\}^k} \left((-1)^{\sum_{i=1}^k \epsilon_i} \cdot \prod_{i=1}^k p_i^{n_i - \epsilon_i}\right)$, which thus implies $\varphi(n) = n \cdot \sum_{[\epsilon_1, \ldots, \epsilon_k] \in \{0,1\}^k} \prod_{i=1}^k (\frac{-1}{p_i})^{\epsilon_i} = n \cdot \prod_{i=1}^k (1 - \frac{1}{p_i}) = \prod_{i=1}^k p_i^{n_i - 1}(p_i - 1)$.

c) We present a q-analogue of the formula $(\delta_1)^+ = \epsilon \in \mathcal{F}(\mathbb{N})$: Let $\mathbb{F}_q$ be the finite field of order q, and let $I_n(q) \in \mathbb{N}_0$ be the number of monic irreducible polynomials over $\mathbb{F}_q$ of degree $n \in \mathbb{N}$. Then since $\mathbb{F}_{q^d}$ is the separable splitting field of any irreducible polynomial over $\mathbb{F}_q$ of degree $d \in \mathbb{N}$, for $d \in \{1, \ldots, n\}$, and $\mathbb{F}_{q^d} \subseteq \mathbb{F}_{q^n}$ if and only if $d \mid n$, we get $q^n = \sum_{d \,|\, n} d \cdot I_d(q)$

Thus letting $f_q \in \mathcal{F}(\mathbb{N})$ be defined by $f_q(n) := n \cdot I_n(q)$, for $n \in \mathbb{N}$, for $f_q^+ \in \mathcal{F}(\mathbb{N})$ we have $(f_q^+)(n) = \sum_{d \,|\, n} d \cdot I_d(q) = q^n$, and thus $n \cdot I_n(q) = f_q(n) = (f_q^+ * \mu)(n) = \sum_{d \,|\, n} q^d \cdot \mu(\frac{n}{d})$, which is a combinatorial formula to compute $I_n(q)$.

Indeed, for all $n \in \mathbb{N}$ we may view $f_q(n)$ and $f_q^+(n)$ as a polynomial in $\mathbb{Z}[q]$, hence specializing at $q \mapsto 1$ yields $f_q^+(n)|_{q \mapsto 1} = q^d|_{q \mapsto 1} = 1$, saying that $f_q^+|_{q \mapsto 1} = \epsilon \in \mathcal{F}(\mathbb{N})$. Moreover, we get $f_q(n)|_{q \mapsto 1} = \sum_{d \,|\, n} \mu(\frac{n}{d}) = \sum_{d \,|\, n} \mu(d, n) = \delta(1, n)$, saying that $f_q|_{q \mapsto 1} = \delta_1 \in \mathcal{F}(\mathbb{N})$.

(9.3) Example: Subspace lattices. To compute the Möbius function of $\mathbf{P}_n(q)$, where q is a prime power and $n \in \mathbb{N}_0$, we observe that for $V \leq W \leq \mathbb{F}_q^n$ we have $[V, W] \cong [V/V, W/V] \cong [\{0\}, \mathbb{F}_q^{\dim_{\mathbb{F}_q}(W) - \dim_{\mathbb{F}_q}(V)}]$ as partially ordered sets, hence, since μ is determined locally, it suffices to determine $\mu_n(q) := \mu(\{0\}, \mathbb{F}_q^n)$. By induction on $n \in \mathbb{N}$ we show $\mu_n(q) = -q^{n-1}\mu_{n-1}(q)$:

Letting $L \leq \mathbb{F}_q^n$ be a 1-dimensional $\mathbb{F}_q$-subspace, by Weisner's Theorem we have $\sum_{U \leq \mathbb{F}_q^n, U + L = \mathbb{F}_q^n} \mu(\{0\}, U) = 0$. The condition $U + L = \mathbb{F}_q^n$ is equivalent to either $U = \mathbb{F}_q^n$, or $U \lessdot \mathbb{F}_q^n$ such that $L \not\leq U$. Hence we get $\mu_n(q) = \mu(\{0\}, \mathbb{F}_q^n) = -\sum_{L \not\leq U \lessdot \mathbb{F}_q^n} \mu(\{0\}, U) = -\mu_{n-1}(q) \cdot |\{U \lessdot \mathbb{F}_q^n; L \not\leq U\}|$. Since there are $\frac{q^n - 1}{q - 1}$

maximal proper subspaces of $\mathbb{F}_q^n$, of which $\frac{q^{n-1}-1}{q-1}$ contain L, we infer $\mu_n(q) = -\mu_{n-1}(q) \cdot (\frac{q^n-1}{q-1} - \frac{q^{n-1}-1}{q-1}) = -q^{n-1}\mu_{n-1}(q)$. $\sharp$

This entails $\mu_n(q) = (-1)^n q^{\binom{n}{2}}$ for $n \in \mathbb{N}_0$: We have $\mu_0(q) = 1$, and by induction on $n \in \mathbb{N}$ we get $\mu_n(q) = -q^{n-1}\mu_{n-1}(q) = -(-1)^{n-1}q^{\binom{n-1}{1}}q^{\binom{n-1}{2}} = (-1)^n q^{\binom{n}{2}}$. For example, we have $\mu_1(q) = -1$ and $\mu_2(q) = q$ and $\mu_3(q) = -q^3$.

Moreover, we observe the following: Considering $\mu_n(q)$ as a polynomial in $\mathbb{Z}[q]$, we may specialize at $q \mapsto 1$, yielding $\mu_n(q)|_{q\mapsto 1} = (-1)^n$, for all $n \in \mathbb{N}_0$. This coincides with the Möbius function μ_n for the partially ordered set $\mathcal{P}(N)$, where N has cardinality n, to be discussed below. Hence $\mu_n(q)$ is a q-analogue of μ_n. Indeed, since we have already noted that $\mathbf{P}_n(q)$ can be seen as a q-analogue of $\mathcal{P}(N)$, this is not too surprising: Both $\mathbf{P}_n(q)$ and $\mathcal{P}(N)$ are graded of rank n, and for the respective number of elements of rank k, for $k \in \{0, \ldots, n\}$, we have $\binom{n}{k}_q|_{q\mapsto 1} = \binom{n}{k}$. Since Möbius functions in general are determined locally by the recursion $\sum_{x\leq z\leq y} \mu(x, y) = 0$ for $x < y$, where $\mu(x, x) = 1$, in the present cases it follows that $\mu_n(q)|_{q\mapsto 1} = \mu(\{0\}, \mathbb{F}_q^n)|_{q\mapsto 1} = \mu(\emptyset, N) = \mu_n$.

(9.4) Example: Subset lattices. a) Let N be a set. To compute the Möbius function of the finitary power sets $\mathcal{P}_{\text{fin}}(N)$ and $\mathcal{P}_{\text{co-fin}}(N)$, since Möbius functions are determined locally, we infer that is suffices to consider the case of a finite set N, where we have $\mathcal{P}(N) = \mathcal{P}_{\text{fin}}(N) = \mathcal{P}_{\text{co-fin}}(N)$. Hence let now $N := \{1, \ldots, n\}$ be a finite set of cardinality $n \in \mathbb{N}_0$. Since for $L \subseteq M \subseteq N$ we have $[L, M] \cong [L \setminus L, M \setminus L] \cong [\{\}, \{1, \ldots, |M| - |L|\}]$ as partially ordered sets, we conclude that it suffices to compute $\mu_n := \mu(\emptyset, N)$:

We may identify $\mathcal{P}(N)$ with $\{0, 1\}^n$ by using indicator functions, that is $M \subseteq N$ is associated with $[a_1, \ldots, a_n]$, where $a_i = 1$ if $i \in M$, and $a_i = 0$ if $i \in N \setminus M$. Moreover, $\{0, 1\}^n$ becomes a partially ordered set with respect to the n-fold direct product of the totally ordered set $\{0, 1\}$ with itself. Then the chosen identification actually is an isomorphism $\mathcal{P}(N) \cong \{0, 1\}^n$ of partially ordered sets; for the example of $n = 3$ see Table 9. Hence we conclude that $\mu_n = \mu(\emptyset, N) = \mu([0, \ldots, 0], [1, \ldots, 1]) = (-1)^n$. Alternatively, we may proceed by induction on $n \in \mathbb{N}_0$: For $n = 0$ we have $\mu_0 = 1$; and for $n \geq 1$ the set N has $\binom{n}{k}$ subsets of cardinality k, for all $k \in \{0, \ldots, n\}$, hence we get $\mu_n = -\sum_{k=0}^{n-1}(-1)^k\binom{n}{k} = (-1)^n$.

b) Let N be an arbitrary set again, and let $R \neq \{0\}$ be a commutative ring. Then Möbius inversion on $\mathcal{P}_{\text{fin}}(N)$ and $\mathcal{P}_{\text{co-fin}}(N)$ reads as follows:

Since $\mathcal{P}_{\text{fin}}(N)$ has the zero element $\emptyset$, for $f \in \mathcal{F}_R(\mathcal{P}_{\text{fin}}(N))$ we get $f^+(M) = (f * \zeta)(M) = \sum_{L\subseteq M} f(L)$, for all finite $M \subseteq N$. Similarly, $(f * \mu)(M) = \sum_{L\subseteq M} f(L)\mu(L, M) = \sum_{L\subseteq M}(-1)^{|M\setminus L|}f(L)$, for all finite $M \subseteq N$. Thus Möbius inversion yields $f^+ * \mu = f = (f * \mu)^+$ for all $f \in \mathcal{F}_R(\mathcal{P}_{\text{fin}}(N))$.

Since, $\mathcal{P}_{\text{co-fin}}(N)$ has the one element N, for $f \in \mathcal{F}_R(\mathcal{P}_{\text{co-fin}}(N))$ we get $f_+(M) = (\zeta * f)(M) = \sum_{M\subseteq L\subseteq N} f(L)$, for all co-finite $M \subseteq N$, and $(\mu * f)(M) =$

$\sum_{M\subseteq L\subseteq N}\mu(M,L)f(L)=\sum_{M\subseteq L\subseteq N}(-1)^{|L\setminus M|}f(L)$, for all co-finite $M\subseteq N$. Thus Möbius inversion yields $\mu*f_+=f=(\mu*f)_+$ for all $f\in\mathcal{F}_R(\mathcal{P}_{\text{co-fin}}(N))$.

10 Inclusion-exclusion

(10.1) Binomial inversion. a) Keeping the notation of (9.4), for finite N and functions $f\in\mathcal{F}_R(\mathcal{P}(N))$ only depending on the cardinality of the subsets of N, but not on the particular subsets considered, we obtain **binomial inversion**:

Let $a:=[a_0,\ldots,a_n]\in R^{n+1}$, where $n:=|N|\in\mathbb{N}_0$, and let $f\in\mathcal{F}_R(\mathcal{P}(N))$ be defined by $f(M)=a_m$, whenever $M\subseteq N$ such that $|M|=m\in\{0,\ldots,n\}$. Then we have $f^+(M)=\sum_{L\subseteq M}f(L)=\sum_{i=0}^m\binom{m}{i}a_i=:b_m$, saying that $f^+\in\mathcal{F}_R(\mathcal{P}(N))$ only depends on cardinalities, and is described by $b:=[b_0,\ldots,b_n]\in R^{n+1}$. This yields $a_m=f(M)=(f^+*\mu)(M)=\sum_{L\subseteq M}(-1)^{|M\setminus L|}\cdot(f^+)(L)=\sum_{L\subseteq M}(-1)^{|M\setminus L|}b_{|L|}=\sum_{j=0}^m(-1)^{m-j}\binom{m}{j}b_j$.

For example, the i-th unit vector $a:=e_i\in R^{n+1}$, where $i\in\{0,\ldots,n\}$, is mapped to $b=[b_0,\ldots,b_n]$, where $b_j=\binom{j}{i}$, hence binomial inversion yields $\delta_{i,k}=\sum_{j=0}^k(-1)^{k-j}\binom{k}{j}\binom{j}{i}$, for $k\in\{0,\ldots,n\}$; in particular, $i=0$ yields $\sum_{j=0}^k(-1)^{k-j}\binom{k}{j}=\delta_{0,k}$. Moreover, for $a:=[1,\ldots,1]$ we get $b_j=\sum_{i=0}^j\binom{j}{i}=(1+1)^j=2^j$, for $j\in\{0,\ldots,n\}$, thus $b=[2^0,2^1,\ldots,2^n]$; hence binomial inversion yields $1=\sum_{j=0}^k(-1)^{k-j}\binom{k}{j}\cdot 2^j$, for $k\in\{0,\ldots,n\}$, where the right hand side indeed coincides with the expansion of $(2-1)^k$.

b) We present a couple of more interesting examples: Firstly, we reconsider the number D_n of derangements in $\mathcal{S}_n$, for $n\in\mathbb{N}_0$. Letting still $D_{n,k}\in\mathbb{N}_0$ be the number of permutations in $\mathcal{S}_n$ having precisely $k\in\mathbb{N}_0$ fixed points, choosing any k-subset of $\{1,\ldots,n\}$ as set of fixed points, we get $n!=\sum_{k=0}^n\binom{n}{k}D_{n-k}=\sum_{k=0}^n\binom{n}{k}D_k$. Thus by binomial inversion we recover $D_n=\sum_{k=0}^n(-1)^{n-k}\binom{n}{k}\cdot k!=n!\cdot\sum_{k=0}^n\frac{(-1)^{n-k}}{(n-k)!}=n!\cdot\sum_{k=0}^n\frac{(-1)^k}{k!}$.

Secondly, we reconsider the Stirling numbers of the second kind; recall that we have $n!\cdot S_{k,n}=|\text{Surj}(K,N)|$, where $K:=\{1,\ldots,k\}$ and $k\in\mathbb{N}_0$. From $\text{Maps}(K,N)=\coprod_{I\subseteq N}\text{Surj}(K,I)$ we get $n^k=\sum_{i=0}^n\binom{n}{i}\cdot i!\cdot S_{k,i}$. Thus by binomial inversion we get $|\text{Surj}(K,N)|=n!\cdot S_{k,n}=\sum_{j=0}^n(-1)^{n-j}\binom{n}{j}\cdot j^k$.

(10.2) Inclusion-exclusion. a) As a further application, we obtain the **principle of inclusion-exclusion**: Let X be a finite set, and let $X_1,\ldots,X_n\subseteq X$, where $n\in\mathbb{N}_0$. Letting $N:=\{1,\ldots,n\}$ we aim at computing the cardinality $|X\setminus\bigcup_{i\in N}X_i|=|\bigcap_{i\in N}(X\setminus X_i)|\in\mathbb{N}_0$; note that intersections over empty index sets are set equal to X throughout.

For $I\subseteq J\subseteq N$ let $X_I:=\bigcap_{i\in I}X_i$, and $X_{I\subseteq J}:=X_I\cap\bigcap_{j\in J\setminus I}(X\setminus X_j)$; in particular we have $X_{\emptyset\subseteq N}=\bigcap_{j\in N}(X\setminus X_j)$. Hence for $x\in X$ we have $x\in X_{I\subseteq N}$ if and only if $I=\{i\in N;x\in X_i\}$, and thus $X=\coprod_{I\subseteq N}X_{I\subseteq N}$. We have $X_I=$

$X_I \cap X = (\bigcap_{i\in I} X_i) \cap (\coprod_{J\subseteq N\setminus I} X_{J\subseteq N\setminus I}) = \coprod_{J\subseteq N\setminus I}((\bigcap_{i\in I} X_i) \cap X_{J\subseteq N\setminus I}) = \coprod_{J\subseteq N\setminus I}((\bigcap_{i\in I\dot\cup J} X_i) \cap (\bigcap_{j\in N\setminus(I\dot\cup J)}(X \setminus X_j))) = \coprod_{I\subseteq J\subseteq N} X_{J\subseteq N}$.

Let $f\colon \mathcal{P}(N) \to \mathbb{Z}\colon I \mapsto |X_{I\subseteq N}|$. Then we get $(f_+)(I) = \sum_{I\subseteq J\subseteq N} |X_{J\subseteq N}| = |\coprod_{I\subseteq J\subseteq N} X_{J\subseteq N}| = |X_I|$, implying that $|X_{I\subseteq N}| = f(I) = (\mu * f_+)(I) = \sum_{I\subseteq J\subseteq N} (-1)^{|J\setminus I|} \cdot f_+(J) = \sum_{I\subseteq J\subseteq N} (-1)^{|J\setminus I|} \cdot |X_J|$, for all $I \subseteq N$. In particular, $I = \emptyset$ yields the set-theoretical inclusion-exclusion formula $|\bigcap_{i\in N}(X\setminus X_i)| = \sum_{J\subseteq N}(-1)^{|J|} \cdot |X_J| = \sum_{J\subseteq N}(-1)^{|J|} \cdot |\bigcap_{j\in J} X_j|$.

b) If the cardinality $|X_I|$ only depends on the cardinality of I, for all $I \subseteq N$, or equivalently if this holds for all the $|X_{I\subseteq N}|$, then the principle of inclusion-exclusion boils down to binomial inversion. Hence the cases of interest here are those where $|X_I|$, and hence X_I, genuinely depend on I. Indeed, the couple of examples in (10.1) is not too interesting in the present context:

Firstly, we reconsider the number D_n of derangements in $X := \mathcal{S}_n$. Letting $X_i := \{\pi \in \mathcal{S}_n; \pi(i) = i\}$, for $i \in N$, for any $I \subseteq N$ we have $|X_I| = |\{\pi \in \mathcal{S}_n; \pi(i) = i \text{ for all } i \in I\}| = (n-|I|)!$, and $|X_{I\subseteq N}| = |\{\pi \in \mathcal{S}_n; \mathrm{Fix}_N(\pi) = I\}| = D_{n-|I|}$. The principle of inclusion-exclusion says $D_{n-|I|} = \sum_{I\subseteq J\subseteq N}(-1)^{|J\setminus I|} \cdot (n-|J|)!$. Although this entails $D_{n,k} = \binom{n}{k} \cdot D_{n-k} = \binom{n}{k} \cdot \sum_{j=0}^{n-k}(-1)^j\binom{n-k}{j} \cdot (n-k-j)! = \frac{n!}{k!} \cdot \sum_{j=0}^{n-k}\frac{(-1)^j}{j!}$, for $k \in \{0, \ldots, n\}$, it is just equivalent to the case $I = \emptyset$, which is the set-theoretic inclusion-exclusion formula, saying $D_n = \sum_{J\subseteq N}(-1)^{|J|} \cdot (n-|J|)! = \sum_{j=0}^{n}(-1)^j\binom{n}{j} \cdot (n-j)! = n! \cdot \sum_{j=0}^{n}\frac{(-1)^j}{j!}$.

Secondly, we reconsider the Stirling numbers of the second kind, or equivalently $|\mathrm{Surj}(K,N)|$, where $K := \{1, \ldots, k\}$ and $k \in \mathbb{N}_0$. Let $X := \mathrm{Maps}(K,N)$ and $X_i := \{f \in X; f^{-1}(i) = \emptyset\}$, for $i \in N$. Then for any $I \subseteq N$ we have $|X_I| = |\{f \in X; f^{-1}(I) = \emptyset\}| = |\mathrm{Maps}(K, N\setminus I)| = (n-|I|)^k$ and $|X_{I\subseteq N}| = |\mathrm{Surj}(K, N\setminus I)|$. The principle of inclusion-exclusion says $|\mathrm{Surj}(K, N\setminus I)| = \sum_{I\subseteq J\subseteq N}(-1)^{|J\setminus I|} \cdot (n-|J|)^k$, which again is equivalent to the case $I = \emptyset$, in which the set-theoretic inclusion-exclusion formula says $|\mathrm{Surj}(K,N)| = \sum_{J\subseteq N}(-1)^{|J|} \cdot (n-|J|)^k = \sum_{j=0}^{n}(-1)^j\binom{n}{j} \cdot (n-j)^k = \sum_{j=0}^{n}(-1)^{n-j}\binom{n}{j} \cdot j^k$.

c) Here is a more interesting example: For $r \geq 2$ let $P_{r'}(n) \subseteq P(n)$ be the set of partitions consisting of parts not divisible by r, and let $P_{r\text{-reg}}(n) \subseteq P(n)$ be the set of ***r*-regular partitions**, that is those all of whose parts have multiplicity less than r; for $r = 2$ we get the sets $O(n)$ and $D(n)$ of partitions consisting of odd parts and of pairwise distinct parts, respectively, see Table 11. Then we have $|P_{r'}(n)| = |P_{r\text{-reg}}(n)| \in \mathbb{N}_0$:

For $i \in N$ let $X_i \subseteq P(n) =: X$ be the set of partitions having ri as a part, and let $Y_i \subseteq P(n)$ be the set of partitions containing the part i at least r times. Then for any subset $I \subseteq N$ we have $|\bigcap_{i\in I} X_i| = |P(n - \sum_{i\in I} ri)| = |P(n - r \cdot \sum_{i\in I} i)| = |\bigcap_{i\in I} Y_i|$. Hence applying the set-theoretic inclusion-exclusion formula to both $P_{r'}(n) = P(n) \setminus \bigcup_{i=1}^{n} X_i$ and $P_{r\text{-reg}}(n) = P(n) \setminus \bigcup_{i=1}^{n} Y_i$ yields the assertion. ♯

Table 11: Odd-part and distinct-part partitions.

n	$O(n)$	$D(n)$
0	$[]$	$[]$
1	$[1]$	$[1]$
2	$[1^2]$	$[2]$
3	$[3], [1^3]$	$[3], [2,1]$
4	$[3,1], [1^4]$	$[4], [3,1]$
5	$[5], [3,1^2], [1^5]$	$[5], [4,1], [3,2]$
6	$[5,1], [3^2], [3,1^3], [1^6]$	$[6], [5,1], [4,2], [3,2,1]$
7	$[7], [5,1^2], [3^2,1], [3,1^4], [1^7]$	$[7], [6,1], [5,2], [4,3], [4,2,1]$

(10.3) Grouping index sets. a) We keep the notation of (10.2), and let $s_k := \sum_{J\subseteq N, |J|=k} |X_J| \in \mathbb{N}_0$, for $k \in \{0,\ldots,n\}$; in particular, we have $s_0 = |X|$ and $s_1 = \sum_{i=1}^n |X_i|$. Summing over the subsets of cardinality k, the principle of inclusion-exclusion translates as $\sum_{I\subseteq N,|I|=k} |X_{I\subseteq N}| = \sum_{I\subseteq J\subseteq N,|I|=k} (-1)^{|J|-k} \cdot |X_J| = \sum_{l=k}^n (-1)^{l-k}\binom{l}{k} \cdot (\sum_{J\subseteq N,|J|=l} |X_J|) = \sum_{l=k}^n (-1)^{l-k}\binom{l}{k} s_l$. In particular, the case $k=0$ allows to reformulate the set-theoretical inclusion-exclusion formula as $|\bigcap_{i\in N}(X \setminus X_i)| = \sum_{l=0}^n (-1)^l s_l$.

Summing over all subsets of cardinality at least k yields $\sum_{I\subseteq N,|I|\geq k} |X_{I\subseteq N}| = \sum_{m=k}^n \sum_{l=m}^n (-1)^{l-m}\binom{l}{m} s_l = \sum_{l=k}^n s_l \cdot (\sum_{m=k}^l (-1)^{l-m}\binom{l}{m})$. For $k=0$ from $\sum_{m=0}^l (-1)^{l-m}\binom{l}{m} = \delta_{0,l}$ we recover $\sum_{I\subseteq N,|I|\geq 0} |X_{I\subseteq N}| = \sum_{l=0}^n \delta_{0,l} s_l = s_0 = |X|$. Moreover, for $k \geq 1$ the partial alternating row sum formula yields $\sum_{m=k}^l (-1)^{l-m}\binom{l}{m} = \sum_{m=0}^{l-k} (-1)^m \binom{l}{m} = (-1)^{l-k}\binom{l-1}{l-k} = (-1)^{l-k}\binom{l-1}{k-1}$, from which we get $\sum_{I\subseteq N,|I|\geq k} |X_{I\subseteq N}| = \sum_{l=k}^n (-1)^{l-k}\binom{l-1}{k-1} s_l$.

b) We show **Bonferoni's Theorem [1936]**, for $k \in \{0,\ldots,n\}$ saying that $(-1)^k \cdot \sum_{l=k}^n (-1)^l s_l \geq 0$:

The set-theoretical inclusion-exclusion formula implying the case $k=0$, we may assume $k \geq 1$. Then we have $\sum_{l=k}^n (-1)^{k-l} s_l = \sum_{J\subseteq N,|J|\geq k} (-1)^{|J|-k} \cdot |X_J| = \sum_{J\subseteq N,|J|\geq k} \sum_{J\subseteq I\subseteq N} (-1)^{|J|-k} \cdot |X_{I\subseteq N}|$. Next, changing the order of summation yields $\sum_{l=k}^n (-1)^{k-l} s_l = \sum_{I\subseteq N,|I|\geq k} |X_{I\subseteq N}| \cdot (\sum_{J\subseteq I,|J|\geq k} (-1)^{|J|-k}) = \sum_{I\subseteq N,|I|\geq k} |X_{I\subseteq N}| \cdot (\sum_{j=k}^{|I|} (-1)^{j-k}\binom{|I|}{j})$. The partial alternating row sum formula yields $\sum_{j=k}^{|I|} (-1)^{j-k}\binom{|I|}{j} = \sum_{j=0}^{|I|-k} (-1)^{|I|-k-j}\binom{|I|}{j} = \binom{|I|-1}{|I|-k} = \binom{|I|-1}{k-1}$, from which we finally get $\sum_{l=k}^n (-1)^{k-l} s_l = \sum_{I\subseteq N,|I|\geq k} |X_{I\subseteq N}| \cdot \binom{|I|-1}{k-1} \geq 0$. ♯

Hence, letting $s := |\bigcap_{i\in N}(X \setminus X_i)| = \sum_{l=0}^n (-1)^l s_l$, we have $s \leq \sum_{l=0}^k (-1)^l s_l$ if k is even, and $s \geq \sum_{l=0}^k (-1)^l s_l$ if k is odd. Thus, the partial sums $\sum_{l=0}^k (-1)^l s_l$ successively are lower and upper estimates of s; note that the inequalities $s \leq s_0 = |X|$ and $s \geq s_0 - s_1 = |X| - \sum_{i=1}^n |X_i|$ are obvious anyway.

(10.4) Improving inclusion-exclusion. We get an improvement of the principle of inclusion-exclusion by taking the intersection configuration into account. To do so, keeping the notation of (10.2), we consider the set $\mathcal{X} := \{X_I \in \mathcal{P}(N); I \subseteq N\}$, partially ordered by set-theoretic inclusion $\subseteq$, and having zeta and Möbius functions $\zeta \in \mathcal{A}(\mathcal{X})$ and $\mu \in \mathcal{A}(\mathcal{X})$, respectively.

For $M \in \mathcal{X}$ let $\widehat{M} := M \setminus \bigcup\{L \in \mathcal{X}; L \subset M\} = \bigcap\{M \setminus L; L \in \mathcal{X}, L \subset M\} \subseteq X$ be the set of all elements of M which are not contained in a strictly smaller set in $\mathcal{X}$; hence we have $M = \coprod_{L \subseteq M \in \mathcal{X}} \widehat{L}$. In particular, we have $X = X_\emptyset \in \mathcal{X}$ and $\widehat{X} = X \setminus \bigcup_{i \in N} X_i = \bigcap_{i \in N}(X \setminus X_i)$, where our aim hence is to compute $|\widehat{X}|$.

Let now $f \colon \mathcal{X} \to \mathbb{Z} \colon M \mapsto |\widehat{M}|$. Then for $f^+ = f * \zeta \in \mathcal{F}(\mathcal{X})$ we have $f^+(M) = \sum_{L \subseteq M \in \mathcal{X}} f(L) = \sum_{L \subseteq M \in \mathcal{X}} |\widehat{L}| = |\coprod_{L \subseteq M \in \mathcal{X}} \widehat{L}| = |M|$, for all $M \in \mathcal{X}$. Thus we indeed get a shorter inclusion-exclusion formula $|\widehat{X}| = f(X) = (f^+ * \mu)(X) = \sum_{L \in \mathcal{X}} f^+(L) \cdot \mu(L, X) = \sum_{L \in \mathcal{X}} \mu(L, X) \cdot |L|$, of course at the expense of having to compute the Möbius function μ of $\mathcal{X}$.

For example, for $n = 3$, let $A, B, C \subseteq X$ such that $A \cap B = A \cap C = B \cap C$; hence $A \cap B \cap C = A \cap B$ as well. Then the set-theoretic inclusion-exclusion formula yields $|X \setminus (A \cup B \cup C)| = |X| - |A| - |B| - |C| + |A \cap B| + |A \cap C| + +|B \cap C| - |A \cap B \cap C| = |X| - |A| - |B| - |C| + 2 \cdot |A \cap B \cap C|$. Now we have $\mathcal{X} = \{A \cap B \cap C, A, B, C, X\} \cong \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2, 3\}\} \cong \mathbf{P}_2(2)$ as partially ordered sets; hence the matrices of ζ and μ are

$$\zeta \mapsto \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & . & . & 1 \\ & & 1 & . & 1 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}, \qquad \mu \mapsto \begin{bmatrix} 1 & -1 & -1 & -1 & 2 \\ & 1 & . & . & -1 \\ & & 1 & . & -1 \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix}.$$

Reading off the last column of the matrix of μ yields again $|X \setminus (A \cup B \cup C)| = |X| - |A| - |B| - |C| + 2 \cdot |A \cap B \cap C|$.

(10.5) Example: Problème des ménages. How many ways are there to place $n \geq 2$ couples at a circular table, such that men and women are in alternate places, but such that no husband sits at either side of his wife? Allowing for arbitrary renumbering of the couples, and for exchanging the roles of men and women, up to a factor of $2 \cdot n!$ the number sought is given as follows: We may assume that women 1 to n are seated in clockwise order. Numbering the places for the men in clockwise order as well, where seat 1 is the one left to woman 1, we have to determine the number $m_n \in \mathbb{N}_0$ of all **discordant permutations** $\pi \in \mathcal{S}_n$ such that $\pi(i) \not\equiv i, i+1 \pmod{n}$ for all $i \in N := \{1, \ldots, n\}$.

To this end, we apply the principle of inclusion-exclusion to the sets $X_i := \{\pi \in \mathcal{S}_n; \pi(i) = i \text{ or } \pi(i) \equiv i + 1 \pmod{n}\}$, for $i \in N$. Since in this case $|X_I|$ indeed depends on the particular choice of $I \subseteq N$, not just on $|I|$ alone, we are going to determine $\sum_{I \subseteq N, |I| = k} |X_I|$, for $k \in \{0, \ldots, n\}$, in a single step:

To do so, we consider the permutation matrices associated with $\mathcal{S}_n$, that is the $(n \times n)$-matrices with entries in $\{0, 1\}$, such that any row and any column contains precisely one entry 1. Then, given $\pi \in \mathcal{S}_n$, we have $\pi \in X_I$, for some $I \subseteq N$, if and only if the entries 1 in the columns indexed by I are amongst the positions marked by '$*$' in the following $(n \times n)$-matrix, while after fixing the entries in the columns indexed by I for the columns indexed by $N \setminus I$ all $(n - |I|)!$ possible choices left are allowed:

$$\begin{bmatrix} * & . & . & \cdots & * \\ * & * & . & & . \\ . & * & * & & . \\ \vdots & & \ddots & \ddots & \vdots \\ . & . & \cdots & * & * \end{bmatrix}$$

Allowing for all subsets $I \subseteq N$ of a fixed cardinality k, the fact that any row and column of the above matrix contains at most one entry 1 translates into the task to determine the number of ways to choose k pairwise non-adjacent positions on a circle of length $2n$.

To do so, we consider a circular array with positions $\{1, \ldots, m\}$, where $m \in \mathbb{N}$, which has to be filled with $k \leq \lfloor \frac{m}{2} \rfloor$ entries '$\,|\,$' and $m - k$ entries '$\bullet$', such that the entries '$\,|\,$' are pairwise non-adjacent. Thinking of the $m-k$ entries '$\bullet$' to be arranged already, the valid configurations are described by the k-subset of the $m - k$ spaces between the entries '$\bullet$' hosting the entries '$\,|\,$'. In order to count these configurations, we distinguish the cases which entry occupies position 1:

If this is an entry '$\bullet$', we may choose any k-subset of the $m - k$ spaces; if $k \geq 1$ and this is an entry '$\,|\,$', then there are only $m - k - 1$ spaces left, from which we may choose any $(k - 1)$-subset. Thus we get a total of $\binom{m-k}{k} + \binom{m-k-1}{k-1} = \binom{m-k}{k} + \frac{k}{m-k} \cdot \binom{m-k}{k} = \frac{m}{m-k} \cdot \binom{m-k}{k}$ possibilities; this also holds for $k = 0$.

Thus we get $\sum_{I \subseteq N, |I|=k} |X_I| = (n - k)! \cdot \frac{2n}{2n-k} \cdot \binom{2n-k}{k}$. Hence the set-theoretic inclusion-exclusion formula says that $m_n = \sum_{I \subseteq N} (-1)^{|I|} \cdot |X_I| = \sum_{k=0}^{n} (-1)^k \cdot (\sum_{I \subseteq N, |I|=k} |X_I|) = \sum_{k=0}^{n} \frac{(-1)^k \cdot (n-k)! \cdot 2n}{2n-k} \cdot \binom{2n-k}{k}$ **[Touchard, 1934]**. Singling out the case $k = n$, index shifting, and simplifying the summands yields $m_n = (-1)^n \cdot 2 + n \cdot \sum_{k=1}^{n} (-1)^{k-1} \cdot (n - k)! \cdot \binom{2n-k}{k-1}$. For example, we have $m_2 = 0$, $m_3 = 1$, $m_4 = 2$, $m_5 = 13$, $m_6 = 80$, $m_7 = 579$, $m_8 = 4738$, $m_9 = 439792$.

11 Finite incidence algebras

We consider finite dimensional incidence algebras from the perspective of representation theory; for more on the background needed see [3, Ch.I–III].

(11.1) Finite dimensional incidence algebras. Keeping the notation of (8.5), let X be a finite partially ordered set, and let $\mathcal{A} := \mathcal{A}_K(X) = \langle E_{ij}; x_i \leq x_j \rangle_K \subseteq \mathcal{T}_n(K)$, where K is a field.

Then we have $\mathcal{J} := \mathcal{J}_K(X) = \mathcal{U}_K(X) = \langle E_{ij}; x_i < x_j \rangle_K$. Hence we have $\mathcal{A}/\mathcal{J} \cong \bigoplus_{i=1}^n K$, thus $\mathcal{A}$ is split and has precisely n simple modules $\{S_1, \ldots, S_n\}$, up to isomorphism. We have $\dim_K(S_i) = 1$, for $i \in \{1, \ldots, n\}$, hence $\mathcal{A}$ is a **basic algebra**. Moreover, we get $\mathcal{J}^2 = \langle E_{ik}; x_i < x_j < x_k \text{ for some } j \rangle_K$, and thus $\mathcal{J}/\mathcal{J}^2 \cong \langle E_{ij}; x_i \lessdot x_j \rangle_K$. The **radical length** of $\mathcal{A}$ equals $l(X) + 1$, that is $\mathcal{J}^k = \{0\}$ if and only if $k \geq l(X) + 1$; note that $l(X) \leq n - 1$.

Next, $E_n = \sum_{i=1}^n E_{ii} \in \mathcal{A}$ is a decomposition into pairwise **orthogonal primitive idempotents**. We assume notation chosen such that E_{ii} acts as the identity on S_i; hence $S_j E_{ii} = \{0\}$ whenever $i \neq j \in \{1, \ldots, n\}$. Then the **projective cover** of S_i is given as $P_i = E_{ii}\mathcal{A}$, and we have $\mathcal{A} \cong \bigoplus_{i=1}^n P_i$. Using the notation of (8.3), we have $P_i = E_{ii}\mathcal{A} = \mathcal{I}_K(\{x_i\}, X) = \langle E_{ij}; x_i \leq x_j \rangle_K$, for $i \in \{1, \ldots, n\}$, thus $\dim_K(P_i) = |\{j \in \{1, \ldots, n\}; x_i \leq x_j\}| = |\langle x_i \leq \rangle|$, where $\langle x_i \leq \rangle \subseteq X$ denotes the principal coideal generated by x_i.

Given $j \in \{1, \ldots, n\}$ such that $x_i \leq x_j$, the K-linear map $\alpha_{ij} \colon P_j \to P_i \colon E_{jk} \to E_{ik}$, for all $k \in \{1, \ldots, n\}$ such that $x_j \leq x_k$, is an embedding of $\mathcal{A}$-modules, where $\alpha_{ij}(P_j) = \langle E_{ik}; x_j \leq x_k \rangle_K = \mathcal{I}_K(\{x_i\}, \langle x_j \leq \rangle)$. Indeed, a consideration of matrices shows that applying α_{ij} amounts to left multiplication with E_{ij}, hence $\alpha_{ij}(P_j) = E_{ij} \cdot E_{jj}\mathcal{A} = E_{ij}\mathcal{A}$, and $\alpha_{ij} \circ \alpha_{jk} = \alpha_{ik}$ whenever $x_i \leq x_j \leq x_k$.

Hence S_j is a constituent of P_i whenever $x_i \leq x_j$, and a comparison with $\dim_K(P_i)$ shows that it occurs with multiplicity $[P_i \colon S_j] = 1$ if $x_i \leq x_j$, and $[P_i \colon S_j] = 0$ otherwise; thus P_i is multiplicity-free. Thus in turn we have $\mathrm{Hom}_{\mathcal{A}}(P_j, P_i) = \langle \alpha_{ij} \rangle_K$ whenever $x_i \leq x_j$, and $\mathrm{Hom}_{\mathcal{A}}(P_j, P_i) = \{0\}$ otherwise; in particular, P_i is simple if and only if $x_i \in X$ is maximal.

Thus the **Cartan matrix** $\big[[P_i \colon S_j]\big]_{ij} \in \mathbb{Z}^{n \times n}$ of $\mathcal{A}$ is coincides with the matrix of the zeta function in $\mathcal{A}$. In particular, we conclude that the partially ordered set X can be recovered from its incidence algebra $\mathcal{A}$, in other words two incidence algebras are isomorphic if and only if the underlying partially ordered sets are.

(11.2) Path algebra description. We proceed to describe $\mathcal{A}$ as a path algebra quotient. To do so, we first determine the **Ext quiver** of $\mathcal{A}$, which has the simple $\mathcal{A}$-modules $\{S_1, \ldots, S_n\}$ as its vertices, and $[\mathrm{rad}(P_i)/\mathrm{rad}^2(P_i) \colon S_j]$ arrows $S_j \leftarrow S_i$. Actually, the latter coincides with the Hasse diagram $\widehat{X}$ of X; note that $\widehat{X}$ only has simple arrows and is **acyclic**, that is without oriented loops:

We have $\mathrm{rad}(P_i) = E_{ii}\mathcal{A} \cdot \mathcal{J} = E_{ii}\mathcal{J} = \langle E_{ij}; x_i < x_j \rangle_K$ and $\mathrm{rad}^2(P_i) = E_{ii}\mathcal{J}^2 = \langle E_{ik}; x_i < x_j < x_k \text{ for some } j \rangle_K$, thus $\mathrm{rad}(P_i)/\mathrm{rad}^2(P_i) \cong \langle E_{ij}; x_i \lessdot x_j \rangle_K$. Now we have $\alpha_{ij}(P_j) \leq \mathrm{rad}(P_i)$ whenever $x_i < x_j$, which entails $\alpha_{ik}(P_k) = \alpha_{ij}\alpha_{jk}(P_k) \leq \mathrm{rad}^2(P_i)$ whenever $x_i < x_j < x_k$, saying that in this case S_k is not a constituent of $\mathrm{rad}(P_i)/\mathrm{rad}^2(P_i)$. Comparing with the K-dimension of $\mathrm{rad}(P_i)/\mathrm{rad}^2(P_i)$ we thus infer that indeed $\mathrm{rad}(P_i)/\mathrm{rad}^2(P_i) \cong \bigoplus_{j; x_i \lessdot x_j} S_j$. ♯

Let $\widehat{\mathcal{A}}$ be the **path algebra** associated with $\widehat{X}$: A K-basis of $\widehat{\mathcal{A}}$ is given by the set of all **oriented paths** $\rho = (x_{i_k} \leftarrow x_{i_{k-1}} \leftarrow \cdots \leftarrow x_{i_1} \leftarrow x_{i_0})$ of **length** $k \in \mathbb{N}_0$ in $\widehat{X}$; in particular, this encompasses the **empty** paths ρ_i at each vertex

x_i, and the paths $\rho_{ij} := (x_i \leftarrow x_j)$ whenever $x_i \lessdot x_j$. Here, $s(\rho) = i_0$ and $t(\rho) = i_k$ are called the **source** and **target** vertices of ρ, respectively. For paths $\sigma, \rho \in \widehat{\mathcal{A}}$, the product of $\sigma\rho$ is defined as the concatenation (to the left) of σ and ρ whenever $s(\sigma) = t(\rho)$, otherwise let $\sigma \cdot \rho := 0$; we have $\rho\rho_i = \rho$ if $s(\rho) = i$, and $\rho_j \rho = \rho$ if $t(\rho) = j$. Hence $\widehat{\mathcal{A}}$ is generated as a K-algebra by $\{\rho_1, \ldots, \rho_n\} \,\dot{\cup}\, \{\rho_{ij}; x_i \lessdot x_j\}$, and since $\widehat{X}$ is acyclic $\widehat{\mathcal{A}}$ is finite dimensional.

Now, the concatenation rule for paths and the multiplication rule for the matrices E_{ij} entail that $\pi\colon \widehat{\mathcal{A}} \to \mathcal{A}\colon \rho_i \mapsto E_{ii}$, $\rho_{ij} \mapsto E_{ij}$, whenever $x_i \lessdot x_j$, extends to an epimorphism of K-algebras. We determine a finite set of defining relations for $\mathcal{A}$ as a quotient of $\widehat{\mathcal{A}}$, that is an ideal generating set of $\ker(\pi) \trianglelefteq \widehat{\mathcal{A}}$:

If $\rho = (x_{i_k} \leftarrow x_{i_{k-1}} \leftarrow \cdots \leftarrow x_{i_1} \leftarrow x_{i_0}) \in \widehat{\mathcal{A}}$ is a path, then we have $\pi(\rho) = E_{i_k i_{k-1}} E_{i_{k-1} i_{k-2}} \cdots E_{i_1 i_0} = E_{i_k i_0} = E_{t(\alpha), s(\alpha)} \in \mathcal{A}$. Hence, if $\sigma, \rho \in \widehat{\mathcal{A}}$ are paths such that $s(\sigma) = s(\rho)$ and $t(\sigma) = t(\rho)$, then we have $\pi(\sigma) = \pi(\rho)$. Thus, letting $\mathcal{K} \trianglelefteq \widehat{\mathcal{A}}$ be the ideal generated by all differences $\sigma - \rho$ of paths such that $s(\sigma) = s(\rho)$ and $t(\sigma) = t(\rho)$, we have $\mathcal{K} \subseteq \ker(\pi)$. Conversely, choosing a path $x_i \leftarrow \cdots \leftarrow x_j$, for any pair $[i,j]$ such that $x_i \leq x_j$, yields a K-generating set of $\widehat{\mathcal{A}}/\mathcal{K}$ of cardinality $|\{[i,j]; x_i \leq x_j\}| = \dim_K(\langle E_{ij}; x_i \leq x_j \rangle_K) = \dim_K(\mathcal{A})$. Hence we conclude that $\dim_K(\widehat{\mathcal{A}}/\ker(\pi)) \leq \dim_K(\widehat{\mathcal{A}}/\mathcal{K}) \leq \dim_K(\mathcal{A}) = \dim_K(\widehat{\mathcal{A}}/\ker(\pi))$, implying equality throughout, and thus $\ker(\pi) = \mathcal{K}$. $\sharp$

Note that since for any i there is only one empty path at x_i, namely ρ_i, and for any $x_i \lessdot x_j$ there only one path from x_j to x_i, namely ρ_{ij}, all paths involved in elements of $\ker(\pi)$ have length at least 2, saying that $\ker(\pi)$ indeed is an **admissible** ideal, that is $\ker(\pi) \subseteq \operatorname{rad}^2(\widehat{\mathcal{A}})$; note that $\operatorname{rad}^k(\widehat{\mathcal{A}})$ has the set of all paths of length at least k as a K-basis, for $k \in \mathbb{N}_0$.

(11.3) Projective modules. a) We proceed to examine the submodule lattice $\mathcal{M}(P_i)$ of P_i, using the terminology introduced in (7.8): Since P_i is multiplicity-free, its submodule lattice $\mathcal{M}(P_i)$ is distributive, in particular finite, being described by the set of ideals of $\mathcal{L}(P_i)$, where for any j such that $x_i \leq x_j$ there is a unique S_j-local submodule. In this case, since P_j is S_j-local and α_{ij} is an embedding, we conclude that $P_j \cong \alpha_{ij}(P_j) = E_{ij}\mathcal{A} \leq E_{ii}\mathcal{A} = P_i$ is the S_j-local submodule of P_i. In particular, any local submodule of P_i is projective.

Hence the submodules of $P_i = \mathcal{I}_K(\{x_i\}, X)$ are precisely given as $\mathcal{I}_K(\{x_i\}, Y)$, where Y runs through the coideals of X contained in the principal coideal $\langle x_i \leq \rangle$. Thus the submodule lattice of P_i is naturally isomorphic to the lattice of coideals mentioned, where the latter is partially ordered by set-theoretic inclusion. In particular, for the **socle** of P_i we have $\operatorname{soc}(P_i) \cong \bigoplus_{j; x_i \leq x_j \in X \text{ maximal}} S_j$.

Moreover, the **blocks** of $\mathcal{A}$, that is the smallest direct summands of $\mathcal{A}$ as a K-algebra, are given by the **connected components** of X, that is the connected components of the unoriented graph underlying the Hasse diagram $\widehat{X}$ of X; in particular, $\mathcal{A}$ is a block if and only if X is connected.

b) We show that all submodules of P_i are projective, not just the local ones, if and only if for all $x_j \geq x_i$ there is a unique path in the Hasse diagram $\widehat{X}$ between x_j and x_i, that is a unique saturated chain in X between x_i and x_j:

Assume that there are distinct paths bewtween x_j and x_i. Then there are x_k and x_l such that $x_i \leq x_k \leq x_j$ and $x_i \leq x_l \leq x_j$, but $x_k \not\leq x_l \not\leq x_k$. Thus we have $\alpha_{ik}(P_k) \leq P_i$ and $\alpha_{il}(P_l) \leq P_i$, such that $\alpha_{ik}(P_k) \not\leq \alpha_{il}(P_l) \not\leq \alpha_{ik}(P_k)$ and $\alpha_{ij}(P_j) \leq \alpha_{ik}(P_k) \cap \alpha_{il}(P_l)$, thus $U := \alpha_{ik}(P_k) + \alpha_{il}(P_l)$ fulfills $U/\mathrm{rad}(U) \cong S_k \oplus S_l$, but its constituent S_j has multiplicity 1, hence U is not projective.

Conversely, assume that the above uniqueness property holds, let $U \leq P_i$ be a submodule, and let $Y = \{x_j \in X; \alpha_{ij}(P_j) \leq U\} \subseteq \langle x_i \leq \rangle \subseteq X$ be the associated coideal of X. Then letting $x_{j_1}, \ldots, x_{j_r} \in Y$ be the minimal elements of Y, we have $U = \sum_{k=1}^r \alpha_{i,j_k}(P_{j_k})$, where $r = r(U) \in \mathbb{N}_0$ coincides with the rank of $U \in \mathcal{M}(P_i)$. Thus, by the uniqueness property assumed, for any $x_j \in Y$ there is a unique $k = k(j) \in \{1, \ldots, r\}$ such that $x_{j_k} \leq x_j$, implying that for any constituent S_j of U there is a unique $k = k(j)$ such that S_j is a constituent of $\alpha_{i,j_k}(P_{j_k})$. Hence we infer that $\alpha_{i,j_k}(P_{j_k}) \cap \sum_{l \neq k} \alpha_{i,j_l}(P_{j_l}) = \{0\}$ for all $k \in \{1, \ldots, r\}$, that is $U = \bigoplus_{k=1}^r \alpha_{i,j_k}(P_{j_k}) \cong \bigoplus_{k=1}^r P_{j_k}$ is projective. ♯

In particular, in this case $\mathrm{rad}(P_i) = \sum_{x_i \lessdot x_j} \alpha_{i,j}(P_j) = \bigoplus_{x_i \lessdot x_j} \alpha_{i,j}(P_j) \cong \bigoplus_{x_i \lessdot x_j} P_j$ is projective. Moreover, running over all $i \in \{1, \ldots, n\}$, we conclude that any submodule of any projective-indecomposable module is projective again, if and only if all connected components of X are trees; in particular, in this case $\mathcal{J} = \bigoplus_{i=1}^n \mathrm{rad}(P_i) \cong \bigoplus_{i=1}^n (\bigoplus_{x_i \lessdot x_j} P_j)$ is projective.

For example, let $X := [\emptyset, \{1\}, \{2\}, \{2,3\}, \{1,2,3\}]$ be partially ordered by set-theoretic inclusion $\subseteq$, whose Cartan matrix is given as the matrix of its zeta function in (8.5). Then we have $\mathrm{rad}(P_\emptyset) = \alpha_{\emptyset,\{1\}}(P_{\{1\}}) + \alpha_{\emptyset,\{2\}}(P_{\{2\}})$, where $\alpha_{\emptyset,\{1\}}(P_{\{1\}}) \cap \alpha_{\emptyset,\{2\}}(P_{\{2\}}) = \alpha_{\emptyset,\{1,2,3\}}(P_{\{1,2,3\}}) = \mathrm{soc}(P_\emptyset) \cong S_{\{1,2,3\}}$, hence $\mathrm{rad}(P_\emptyset)$ is not projective.

(11.4) Injective modules. a) Using the projective cover $P_i = E_{ii}\mathcal{A}$ of S_i, its **$\mathcal{A}$-dual** $P_i^\vee := \mathrm{Hom}_\mathcal{A}(P_i, \mathcal{A}) \cong \mathrm{Hom}_\mathcal{A}(E_{ii}\mathcal{A}, \mathcal{A}) \cong \mathcal{A}E_{ii} = \langle E_{ji}; x_j \leq x_i \rangle_K = \mathcal{I}_K(X, \{x_i\})$ is a projective left $\mathcal{A}$-module; hence we have $\dim_K(P_i^\vee) = |\{j \in \{1, \ldots, n\}; x_j \leq x_i\}| = |\langle \leq x_i \rangle|$, where $\langle \leq x_i \rangle \subseteq X$ denotes the principal ideal generated by x_i. Letting $\{T_1, \ldots, T_n\}$ be the simple left $\mathcal{A}$-modules, where we assume notation chosen such that E_{ii} acts as the identity on T_i, we conclude that $P_i^\vee$ is the projective cover of T_i. Analogous to the right module case we infer that $[P_i^\vee : T_j] = 1$ if $x_j \leq x_i$, and $[P_i^\vee : T_j] = 0$ otherwise.

The submodules of $P_i^\vee = \mathcal{I}_K(X, \{x_i\})$ are precisely given as $\mathcal{I}_K(Y, \{x_i\})$, where Y runs through the ideals of X contained in the principal ideal $\langle \leq x_i \rangle$. Hence the submodule lattice of $P_i^\vee$ is naturally isomorphic to the lattice of ideals mentioned, where the latter is partially ordered by set-theoretic inclusion. In particular, we have $\mathrm{soc}(P_i^\vee) \cong \bigoplus_{j; x_j \leq x_i \in X \text{ minimal}} T_j$ and $\mathrm{rad}(P_i^\vee)/\mathrm{rad}^2(P_i^\vee) \cong \bigoplus_{j; x_j \lessdot x_i} T_j$, where $P_i^\vee$ is simple if and only if $x_i \in X$ is minimal.

Then the **K-dual** $I_i := (P_i^\vee)^* := \mathrm{Hom}_K(P_i^\vee, K)$ is an injective $\mathcal{A}$-module, which since $\mathrm{soc}(I_i) \cong (P_i^\vee/\mathrm{rad}(P_i^\vee))^* \cong T_i^* \cong S_i$ is the **injective hull** of S_i. The properties of $P_i^\vee$ by dualising translate into properties of I_i: We have $\dim_K(I_i) = \dim_K(P_i^\vee) = |\langle \leq x_i \rangle|$, where more precisely $[I_i \colon S_j] = 1$ if $x_j \leq x_i$, and $[I_i \colon S_j] = 0$ otherwise; in particular, I_i is simple if and only if $x_i \in X$ is minimal. The quotient modules of I_i are given by the ideals of X contained in the principal ideal $\langle \leq x_i \rangle$; in particular, we have $I_i/\mathrm{rad}(I_i) \cong \bigoplus_{j; x_j \leq x_i \in X \text{ minimal}} S_j$.

b) In general, **injective-indecomposable** modules are rather intractable. But this changes for connected components of X having a unique minimal element:

Let x_i be the unique minimal element of the connected componens of X under consideration. For $x_i \leq x_j$ we have $\mathrm{soc}(P_j^\vee) \cong T_i$, implying $I_j/\mathrm{rad}(I_j) \cong (\mathrm{soc}(P_j^\vee))^* \cong T_i^* \cong S_i$. Hence I_j is an epimorphic image of P_i. More precisely, since $P_i \cong \mathcal{I}_K(\{x_i\}, X)$ is multiplicity-free, and I_j has precisely the constituents S_k for $x_k \leq x_j$, this determines I_j uniquely as the quotient $I_j \cong \mathcal{I}_K(\{x_i\}, X)/\mathcal{I}_K(\{x_i\}, X \setminus \langle \leq x_j \rangle)$; note that $X \setminus \langle \leq x_j \rangle \subseteq X$ is a coideal. ♯

In particular, this applies if X is a lattice: In this case X is connected and has x_1 as its unique minimal element. Hence all the modules I_j occur as epimorphic images of the **projective-indecomposable** module P_1.

c) Similarly we are able to determine the indecomposable **projective-injective** modules: If $P_i \cong I_j$, then we have $\mathrm{soc}(P_i) \cong S_j$, where x_j is the unique maximal element such that $x_i \leq x_j$, and we have $I_j/\mathrm{rad}(I_j) \cong S_i$, where x_i is the unique minimal element such that $x_i \leq x_j$. Hence the associated connected component of X has x_i and x_j as its unique minimal and maximal elements, respectively. Conversely, given the latter property, we conclude that I_j fulfills $I_j/\mathrm{rad}(I_j) \cong S_i$, that is I_j is an epimorphic image of P_i, and from $\dim_K(I_j) = |\{k \in \{1, \ldots, n\}; x_i \leq x_k \leq x_j\}| = \dim_K(P_i)$ we conclude that $I_j \cong P_i$; note that in this case the constituents S_k occurring encompass the full connected component of X under consideration.

In particular, if X is a lattice, in which case x_1 and x_n are the unique minimal and maximal elements, respectively, then the projective-indecomposable module P_1 is the only projective-injective one, and it coincides with I_n.

(11.5) The natural module. The **natural** $\mathcal{T}_n(K)$-module $M := K^n$ gives rise to a **faithful** representation of $\mathcal{A}$, that is for the associated **annihilator** we have $\mathrm{ann}_{\mathcal{A}}(M) = \{0\}$. Moreover, we have $\dim_K(\mathrm{Hom}_{\mathcal{A}}(E_{ii}\mathcal{A}, M)) = \dim_K(ME_{ii}) = 1$ for all $i \in \{1, \ldots, n\}$, hence we conclude that $[M \colon S_i] = 1$, thus M is multiplicity-free.

Hence the submodule lattice $\mathcal{M}(M)$ of M is described by the set of ideals of $\mathcal{L}(M)$, where for any i there is a unique S_i-local submodule $L_i \leq M$. Since $ME_{ii}\mathcal{A} \leq M$ is a non-zero epimorphic image of $P_i = E_{ii}\mathcal{A}$, we conclude that $L_i = ME_{ii}\mathcal{A}$. Moreover, we have $L_i \leq L_j$ if and only if $ME_{ii} \subseteq ME_{jj}\mathcal{A}$, which holds if and only if $ME_{jj}\mathcal{A}E_{ii} \neq \{0\}$, which in turn since M is faithful holds if and only if $\mathrm{Hom}_{\mathcal{A}}(P_i, P_j) = \mathrm{Hom}_{\mathcal{A}}(E_{ii}\mathcal{A}, E_{jj}\mathcal{A}) \cong E_{jj}\mathcal{A}E_{ii} \neq \{0\}$, where the latter is equivalent to $x_j \leq x_i$.

Thus as partially ordered sets $\mathcal{L}(M)$ is the dual of X, and the ideals of $\mathcal{L}(M)$ are in bijection with the coideals of X. The block components of M are indecomposable, coinciding with the lattice block components, and are given by the connected components of X; thus M is indecomposable if and only if X is connected. To describe when a block components of M is projective or injective, we may assume that X is connected: Then M is a projective-indecomposable module if and only if $M \cong P_1$, which holds if and only if x_1 is the unique minimal element of X; similarly, M is an injective-indecomposable module if and only if $M \cong I_n$, which holds if and only if x_n is the unique maximal element.

(11.6) Characterizing incidence algebras. We proceed to give a representation theoretic characterization of finite dimensional incidence algebras; this is inspired by [16]. The starting point is the above observation that all local submodules of all the projective-indecomposable $\mathcal{A}$-modules P_i are projective again, that is to say that $\mathcal{A}$ is **locally hereditary** [**Bautista, 1981**]. We first discuss a few general related properties:

a) Any locally hereditary basic finite-dimensional K-algebra $\mathcal{A}$ necessarily has an acyclic Ext quiver: Any arrow $T \leftarrow S$ in the Ext quiver, where S and T are simple $\mathcal{A}$-modules, gives rise to an embedding $P_T \to P_S$ of the associated projective covers; hence a putative oriented cycle in the Ext quiver through S would entail a non-surjective embedding $P_S \to P_S$, a contradiction.

Moreover, acyclicity of the Ext quiver entails $[P_S\colon S] = 1$ for all simple $\mathcal{A}$-module S: Assume to the contrary that $[P_S\colon S] \geq 2$, then there is a k-fold extension $\{0\} \to S \to V_k \to \cdots \to V_1 \to S \to \{0\}$, for some $k \in \mathbb{N}$, where the V_l are indecomposable $\mathcal{A}$-modules of composition length 2; hence there is a cycle $S \leftarrow \cdots \leftarrow S$ of length k in the Ext quiver, a contradiction.

b) A finite-dimensional K-algebra $\mathcal{A}$ is called **hereditary** if all its right ideals are projective. Then $\mathcal{A}$ is hereditary if and only if its Jacobson radical is a projective module, which holds if and only if any submodule of any projective-indecomposable module is projective again. In particular, hereditary algebras are indeed locally hereditary, but the converse does not hold.

Hence, if $\mathcal{A}$ is an incidence algebra, then the discussion in (11.3) shows that it is hereditary if and only if all connected components of the underlying partially ordered set X are trees.

(11.7) Theorem: Iovanov–Koffi [2017]. Let $\mathcal{A}$ be a basic finite-dimensional K-algebra. Then the following are are equivalent:
i) $\mathcal{A}$ is an incidence algebra.
ii) $\mathcal{A}$ has a faithful multiplicity-free representation.
iii) $\mathcal{A}$ has a faithful representation with distributive submodule lattice, and we have $[P_S\colon S] = 1$ for all simple $\mathcal{A}$-modules S.

Proof. We have already seen that the natural module of an incidence algebra fulfills the properties of ii), hence the implication ‘i)⇒ii)’ holds.

a) We show the equivalence 'ii)$\Leftrightarrow$iii)' without reference to incidence algebras. To this end, let M be a faithful $\mathcal{A}$-module. Since for any primitive idempotent $e \in \mathcal{A}$ we have $Me \neq \{0\}$, we infer that $[M\colon S] \neq 0$ for all simple $\mathcal{A}$-modules S.

Let now M be multiplicity-free, that is $[M\colon S] = 1$ for all simple $\mathcal{A}$-modules S, which implies that $\mathcal{M}(M)$ is distributive. In order to show the second statement, let S be any simple $\mathcal{A}$-module, and let $e \in \mathcal{A}$ be a primitive idempotent associated with S. Then we have $\dim_K(\mathrm{Hom}_{\mathcal{A}}(eA, M)) = \dim_K(Me) = 1$, and letting $0 \neq \varphi \in \mathrm{Hom}_{\mathcal{A}}(eA, M)$ we have $Me = \langle \varphi(e) \rangle_K$, and hence $M \cong M(1-e) \oplus \langle \varphi(e) \rangle_K$. For $ea \in \ker(\varphi)$ we have $\varphi(e) \cdot ea = \varphi(ea) = 0$, and since $M(1-e) \cdot ea = \{0\}$ anyway, we infer that $ea \in \mathrm{ann}_A(M) = \{0\}$. Thus φ is injective, implying that $eA \cong \varphi(eA) \leq M$ is multiplicity-free, in particular we have $[P_S\colon S] = 1$. This shows the implication 'ii)$\Rightarrow$iii)'.

Conversely, let $\mathcal{M}(M)$ be distributive. Letting S be a simple $\mathcal{A}$-module, $\mathcal{L}_S(M)$ is non-empty and a chain. Assume that $|\mathcal{L}_S(M)| \geq 2$, then there is an S-local submodule $L \leq M$ such that $[L\colon S] \geq 2$, and L being an epimorphic image of P_S, we infer $[P_S\colon S] \geq 2$, a contradiction. Hence $\mathcal{L}_S(M)$ is a singleton set, that is $[M\colon S] = 1$. Thus M is multiplicity-free, showing the implication 'iii)$\Rightarrow$ii)'.

b) To show the implication 'ii)$\Rightarrow$i)', let M be a faithful multiplicity-free $\mathcal{A}$-module, such that $n := \dim_K(M) \in \mathbb{N}$. Hence let $\{S_1, \ldots, S_n\}$ be the simple $\mathcal{A}$-modules, up to isomorphism, and let $1 = \sum_{i=1}^n e_i \in \mathcal{A}$ be a decomposition into pairwise orthogonal primitive idempotents, where we assume notation chosen such that e_i acts as the identity on S_i.

As we have seen in the proof of 'ii)$\Rightarrow$iii)' above, $e_i\mathcal{A}$ is multiplicity-free, hence we have $\dim_K(e_i\mathcal{A}e_j) = \dim_K(\mathrm{Hom}_{\mathcal{A}}(e_j\mathcal{A}, e_i\mathcal{A})) = [e_i\mathcal{A}\colon S_j] \leq 1$ for all $i, j \in \{1, \ldots, n\}$. Whenever $[e_i\mathcal{A}\colon S_j] = 1$, we let $0 \neq \varphi_{ij} \in \mathrm{Hom}_{\mathcal{A}}(e_j\mathcal{A}, e_i\mathcal{A})$ and $e_{ij} := \varphi_{ij}(e_j) \in e_i\mathcal{A}e_j$; for $i = j$ we may assume that $\varphi_{ii} = \mathrm{id}_{e_i\mathcal{A}}$, that is $e_{ii} = e_i$, while φ_{ij} is not surjective whenever $i \neq j$. Hence we have the **Pierce decomposition** $\mathcal{A} \cong \bigoplus_{[e_i\mathcal{A}\colon S_j]=1} e_i\mathcal{A}e_j = \bigoplus_{[e_i\mathcal{A}\colon S_j]=1} \langle e_{ij} \rangle_K$ as K-vector spaces.

Moreover, any non-zero homomorphism $e_i\mathcal{A} \to M$ is injective. Hence choosing $0 \neq \varphi_i \in \mathrm{Hom}_{\mathcal{A}}(e_i\mathcal{A}, M)$, from $\varphi_{ij} \neq 0$ we infer that $\varphi_i\varphi_{ij} \in \mathrm{Hom}_{\mathcal{A}}(e_j\mathcal{A}, M)$ is non-zero, and hence injective, implying that φ_{ij} is injective as well. Hence, whenever $[e_i\mathcal{A}\colon S_j] = 1 = [e_j\mathcal{A}\colon S_k]$, we have $\varphi_{ij}\varphi_{jk} \neq 0$, implying that $[e_i\mathcal{A}\colon S_k] = 1$ as well, and there is $a_{ijk} \in K^*$ such that $\varphi_{ij}\varphi_{jk} = a_{ijk}\varphi_{ik}$; in particular we have $a_{ijj} = a_{jjk} = a_{iii} = 1$. Moreover, for $k = i$ from $\varphi_{ij}\varphi_{ji} \neq 0$ we infer that this is only possible for $j = i$.

Letting $X := \{x_1, \ldots, x_n\}$, we may define a relation $\leq$ on X by $x_i \leq x_j$ if and only if $[e_i\mathcal{A}\colon S_j] = 1$. By the above considerations this relation is reflexive, antisymmetric and transitive, that is a partial order. We show that $\mathcal{A} \cong \mathcal{A}_K(X)$: We have already seen that $\mathcal{A} \cong \bigoplus_{x_i \leq x_j} \langle e_{ij} \rangle_K$, where multiplication is given by $e_{ij}e_{jk} = a_{ijk}e_{ik}$, whenever $x_i \leq x_j \leq x_k$. In other words, multiplication in $\mathcal{A}$ is 'twisted' compared to $\mathcal{A}_K(X)$, but in a trivial sense as follows:

Recalling that $Me_i = \langle \varphi_i(e_i) \rangle_K$, we get $\varphi_i(e_i) \cdot e_{ij} = \varphi_i(e_{ij}) = \varphi_i(e_{ij}) \cdot e_j =$

$b_{ij} \cdot \varphi_j(e_j) \in M$, for suitable $b_{ij} \in K^*$; note that $b_{ii} = 1$. This yields $\varphi_i(e_i) \cdot e_{ij} \cdot e_{jk} = b_{ij} \cdot \varphi_j(e_j) \cdot e_{jk} = b_{ij}b_{jk} \cdot \varphi_k(e_k) \in M$ and $\varphi_i(e_i) \cdot e_{ij}e_{jk} = a_{ijk} \cdot \varphi_i(e_i) \cdot e_{ik} = a_{ijk}b_{ik} \cdot \varphi_k(e_k) \in M$, hence equating yields $a_{ijk} = \frac{b_{ij}b_{jk}}{b_{ik}}$. Thus let $e'_{ij} := \frac{1}{b_{ij}} \cdot e_{ij} \in \mathcal{A}$, for all $i, j \in \{1, \ldots, n\}$ such that $x_i \leq x_j$; note that $e'_{ii} = e_{ii}$. Then we get $\mathcal{A} \cong \bigoplus_{x_i \leq x_j} \langle e'_{ij} \rangle_K$, such that $e'_{ij}e'_{jk} = \frac{1}{b_{ij}b_{jk}} \cdot a_{ijk} \cdot (b_{ik} \cdot e'_{ik}) = e'_{ik}$. This proves $\mathcal{A} \cong \mathcal{A}_K(X)$, and hence the implication 'ii)⇒i)' holds. $\sharp$

IV Generating functions

12 Power series

(12.1) Formal power series. a) Let K be a field. We consider the K-vector space $K[[X]] := \mathrm{Maps}(\mathbb{N}_0, K)$, with pointwise addition and scalar multiplication. We write the elements of $K[[X]]$, that is sequences $[f_n \in K; n \in \mathbb{N}_0]$, as **(ordinary) generating series** or **formal power series** $f := \sum_{n \geq 0} f_n X^n$, where X is an indeterminate. The principle of **comparison of coefficients** holds, saying that $f, g \in K[[X]]$ are equal if and only if $f_n = g_n$ for all $n \in \mathbb{N}_0$.

Then $K[[X]]$ is a commutative K-algebra with respect to **convolutional multiplication** $(\sum_{i \geq 0} f_i X^i) \cdot (\sum_{j \geq 0} g_j X^j) := \sum_{n \geq 0} \sum_{k=0}^{n} (f_k g_{n-k}) X^n \in K[[X]]$, the neutral element being $1 := X^0 \in K[[X]]$; note that to determine any fixed coefficient only finitely many arithmetical operations in K are necessary, and that the polynomial ring $K[X] \subseteq K[[X]]$ is a subring.

Let $\nu(f) = \nu_X(f) := \min\{n \in \mathbb{N}_0; f_n \neq 0\} \in \mathbb{N}_0$, for any $0 \neq f \in K[[X]]$, be the **order** or **(discrete) valuation** of f at X; for completeness we let $\nu(0) = \infty$. Then for any $0 \neq f, g \in K[[X]]$, letting $n := \nu(f) \in \mathbb{N}_0$ and $m := \nu(g) \in \mathbb{N}_0$, we have $fg = f_n g_m X^{n+m} + \sum_{k \geq n+m+1} h_k X^k \in K[[X]]$, for suitable $h_k \in K$, implying that $\nu(fg) = n + m \in \mathbb{N}_0$. In particular, we have $fg \neq 0 \in K[[X]]$, implying that $K[[X]]$ is an integral domain.

Moreover, $\nu_X \colon (K[[X]] \setminus \{0\}, \cdot) \to (\mathbb{Z}, +)$ is a monoid homomorphism into a totally ordered abelian group. The set $\mathcal{I}_n := \{f \in K[[X]]; \nu(f) \geq n\} = X^n K[[X]] \trianglelefteq K[[X]]$ is an ideal, for all $n \in \mathbb{N}_0$. We have $K[[X]] = \mathcal{I}_0 \supset \mathcal{I}_1 \supset \cdots$, where $\bigcap_{n \in \mathbb{N}_0} \mathcal{I}_n = \{0\}$, and $K[[X]]/\mathcal{I}_1 \cong K$ as K-algebras, and $\mathcal{I}_{n+1}/\mathcal{I}_n \cong K$ as K-vector spaces, for all $n \in \mathbb{N}$. In particular, we conclude that $K[[X]]$ is a **complete discrete valuation ring**.

Let $K[[X]]^* \subseteq K[[X]]$ be the set of all invertible elements. Then we have $K[[X]]^* = K[[X]] \setminus XK[[X]] = \{f \in K[[X]]; f_0 \neq 0\} = \{f \in K[[X]]; \nu(f) = 0\}$: If $f \in K[[X]]$ has inverse $g \in K[[X]]$, then from $fg = 1 \in K[[X]]$ we get $f_0 g_0 = 1 \in K$, thus $f_0 \neq 0$; conversely, if $f \in K[[X]]$ such that $f_0 \neq 0$, then letting $g_0 := f_0^{-1} \in K$, and for $n \in \mathbb{N}$ by induction letting $g_n := -f_0^{-1} \cdot \sum_{k=0}^{n-1} f_{n-k} g_k \in K$, for $g := \sum_{n \geq 0} g_n X^n \in K[[X]]$ we get $fg = \sum_{n \geq 0} (\sum_{k=0}^{n} f_{n-k} g_k) X^n = 1 \in K[[X]]$, hence g is the inverse of f.

The field of fractions $K((X))$ of $K[[X]]$ can be identified with the set of **formal Laurent series** $f := \sum_{n\geq m} f_n X^n$, where $m \in \mathbb{Z}$ and $f_n \in K$, for all $n \geq m$: Indeed, convolutional multiplication extends to $K((X))$, and the valuation extends to $K((X))$ by letting $\nu(f) := \min\{n \in \mathbb{Z}; f_n \neq 0\} \in \mathbb{Z}$, for $0 \neq f \in K((X))$. Then for $0 \neq f \in K((X))$ we have $X^{-\nu(f)} f \in K[[X]]^*$, and hence letting $g := (X^{-\nu(f)} f)^{-1} \in K[[X]]^*$ we get $f^{-1} = X^{-\nu(f)} g \in K((X))$. Note that the field $K(X)$ of **rational functions**, that is the field of fractions of $K[X]$, is a subfield of $K((X))$.

b) Then for all $f \in XK[[X]]$, that is $f_0 = 0$, we have $\nu(f) \geq 1$, and hence $\nu(f^n) = n\nu(f) \geq n$, for all $n \in \mathbb{N}_0$. Thus for all $g \in K[[X]]$ and $f \in XK[[X]]$, the **composition** $g(f) := \sum_{n\geq 0} g_n f(X)^n \in K[[X]]$ is well-defined.

Similarly, given a sequence $[F_i \in K[[X]]; i \in \mathbb{N}]$ such that $\lim_{i\to\infty} \nu(F_i) = \infty$, the infinite sum $\sum_{i\geq 1} F_i \in K[[X]]$ is well-defined; and given a sequence $[1+F_j \in 1+XK[[X]]; j \in \mathbb{N}]$ such that $\lim_{j\to\infty} \nu(F_j) = \infty$, the infinite product $\prod_{j\geq 1}(1+F_j) \in K[[X]]$ is well-defined.

The **formal derivative** defined as $\frac{\partial}{\partial X}\colon K[[X]] \to K[[X]]\colon \sum_{n\geq 0} f_n X^n \mapsto \sum_{n\geq 1} n f_n X^{n-1} = \sum_{n\geq 0}(n+1) f_{n+1} X^n$ is K-linear, and we have the **product rule** $\frac{\partial}{\partial X}(fg) = (\frac{\partial}{\partial X} f)g + f(\frac{\partial}{\partial X} g) \in K[[X]]$, for $f, g \in K[[X]]$, and the **chain rule** $\frac{\partial}{\partial X}(g(f)) = (\frac{\partial}{\partial X} g)(f) \cdot (\frac{\partial}{\partial X} f) \in K[[X]]$, for $g \in K[[X]]$ and $f \in XK[[X]]$:

By K-linearity, for the product rule it suffices to note that $\frac{\partial}{\partial X}(X^m X^n) = (m+n)X^{m+n-1} = mX^{m-1}X^n + nX^m X^{n-1} = \frac{\partial}{\partial X}(X^m) \cdot X^n + X^m \cdot \frac{\partial}{\partial X}(X^n) \in K[[X]]$, for $m, n \in \mathbb{N}$. Similarly, for the chain rule, it suffices to show, by induction on $n \in \mathbb{N}$, that $\frac{\partial}{\partial X}(f^n) = nf^{n-1} \cdot \frac{\partial}{\partial X}(f) \in K[[X]]$, for $f \in XK[[X]]$: The case $n = 1$ being trivial, let $n \geq 2$. Then by induction and using the product rule we have $\frac{\partial}{\partial X}(f^n) = \frac{\partial}{\partial X}(f^{n-1} f) = \frac{\partial}{\partial X}(f^{n-1}) \cdot f + f^{n-1} \cdot \frac{\partial}{\partial X}(f) = nf^{n-1} \cdot \frac{\partial}{\partial X}(f)$. ♯

(12.2) Taylor series. a) For $K = \mathbb{C}$ this is related to Taylor series expansions around $x = 0$, generalizing the connection between polynomials and polynomial maps: Let $\mathbb{C}[[X]]^\infty \subseteq \mathbb{C}[[X]]$ be the $\mathbb{C}$-subalgebra of formal power series $f = \sum_{n\geq 0} f_n X^n \in \mathbb{C}[[X]]$ such that the associated Taylor series $\widehat{f}\colon \mathcal{O} \to \mathbb{C}\colon x \mapsto \sum_{n\geq 0} f_n x^n$ converges on an open disc $\mathcal{O} \subseteq \mathbb{C}$ of positive radius centered at $x = 0$. Hence, letting $\mathcal{H}$ be the $\mathbb{C}$-algebra of **holomorphic function germs** around $x = 0$, by the principle of comparison of coefficients the map $\mathbb{C}[[X]]^\infty \to \mathcal{H}\colon f \mapsto \widehat{f}$ is an isomorphism of $\mathbb{C}$-algebras. Moreover, for $f \in \mathbb{C}[[X]]^\infty$ the valuation $\nu(f) \in \mathbb{N}_0$ coincides with the order of $x = 0$ as a zero of $\widehat{f}$.

b) For example, let the **exponential series** be defined as $\exp := \sum_{n\geq 0} \frac{1}{n!} X^n \in 1 + X\mathbb{Q}[[X]] \subseteq \mathbb{Q}[[X]]$. Then for the associated Taylor series we have $\widehat{\exp}(x) = \exp(x)$, for all $x \in \mathbb{C}$, hence we have $\exp \in \mathbb{C}[[X]]^\infty$. Moreover, we have $\frac{\partial}{\partial X}\exp = \sum_{n\geq 1} \frac{1}{(n-1)!} X^{n-1} = \sum_{n\geq 0} \frac{1}{n!} X^n = \exp \in \mathbb{Q}[[X]]$.

If K is a field of characteristic 0, then for $f \in XK[[X]]$ we have $\exp(f) := \sum_{n\geq 0} \frac{1}{n!} f^n \in 1 + XK[[X]] \subseteq K[[X]]$, fulfilling the identity $\exp(f+g) =$

$\sum_{n\geq 0}\frac{1}{n!}(f+g)^n = \sum_{n\geq 0}\sum_{k=0}^n \frac{1}{k!(n-k)!}f^k g^{n-k} = (\sum_{i\geq 0}\frac{1}{i!}f^i)\cdot(\sum_{j\geq 0}\frac{1}{j!}g^j) = \exp(f)\cdot\exp(g) \in K[[X]]$, for all $f,g \in XK[[X]]$, hence $\exp(f)^{-1} = \exp(-f) \in K[[X]]$. In particular, we have $\exp(X)\cdot\exp(-X) = 1 \in \mathbb{Q}[[X]]$, from which we recover the formula $\sum_{k=0}^n \frac{(-1)^{n-k}}{k!(n-k)!} = \delta_{0,n}$, that is $\sum_{k=0}^n (-1)^{n-k}\binom{n}{k} = \delta_{0,n}$, for $n \in \mathbb{N}_0$. Note that going over to the associated Taylor series yields the identity $\exp(x)\cdot\exp(-x) = 1 \in \mathbb{C}$, for all $x \in \mathbb{C}$; but conversely the identity $\exp(x+1) = e\cdot\exp(x) \in \mathbb{C}$, yielding $\sum_{n\geq 0}\frac{1}{n!}(x+1)^n = e\cdot\sum_{m\geq 0}\frac{1}{m!}x^m$ for all $x \in \mathbb{C}$, cannot be translated to the formal setting.

Let $\log := \sum_{n\geq 1}\frac{(-1)^{n-1}}{n}X^n \in X\mathbb{Q}[[X]] \lhd \mathbb{Q}[[X]]$ be the **logarithm series**. Then for the associated Taylor series we have $\widehat{\log}(x) = \sum_{n\geq 1}\frac{(-1)^{n-1}}{n}x^n = \ln(x+1)$, for all $x \in \mathbb{C}$ such that $|x| < 1$, hence $\log \in \mathbb{C}[[X]]^\infty$. Moreover, we have $\frac{\partial}{\partial X}\log = \sum_{n\geq 1}(-1)^{n-1}X^{n-1} = \frac{1}{1+X} \in \mathbb{Q}[[X]]$, and going over to the associated Taylor series we get $\exp(\log) = 1+X \in \mathbb{Q}[[X]]$ and $\log(\exp -1) = X \in \mathbb{Q}[[X]]$.

If K is a field of characteristic 0, then for $f \in XK[[X]]$ we have $\log(f) := \sum_{n\geq 0}\frac{(-1)^{n-1}}{n}f^n \in 1+XK[[X]] \subseteq K[[X]]$, fulfilling the identity $\log((f+1)(g+1)-1) = \log(f)+\log(g) \in K[[X]]$, for all $f,g \in XK[[X]]$: We have $\exp(\log((f+1)(g+1)-1)) = (f+1)(g+1) = \exp(\log(f))\cdot\exp(\log(g)) = \exp(\log(f)+\log(g)) \in K[[X]]$, thus $\log((f+1)(g+1)-1) = \log(\exp(\log((f+1)(g+1)-1))-1) = \log(\exp(\log(f)+\log(g))-1) = \log(f)+\log(g) \in K[[X]]$. Note that this for $f,g \in 1+XK[[X]]$ yields $\log(fg-1) = \log(f-1)+\log(g-1) \in K[[X]]$.

(12.3) Binomial series. a) The simplest generating series are the polynomials, being associated to finite sequences. For example, for $n \in \mathbb{N}_0$, for the sequence of binomial coefficients we have $\sum_{k=0}^n \binom{n}{k}X^k = (1+X)^n \in \mathbb{Z}[X]$, and for the Stirling numbers of the first kind we have $\sum_{k=0}^n s_{n,k}X^k = X^{(n)} \in \mathbb{Z}[X]$.

For the sequence of Stirling numbers of the second kind we similarly have $X^n = \sum_{k=0}^n S_{n,k}X_{(k)} \in \mathbb{Z}[X]$, which does not translate into a simple formula for the associated generating series; but recalling that both $\{X^k; k \in \{0,\ldots,n\}\}$ and $\{X_{(k)}; k \in \{0,\ldots,n\}\}$ are $\mathbb{Z}$-bases of $\mathbb{Z}[X]_{\leq n}$ shows that this differs from a proper generating series only by a base change.

The other way around, fixing $k \in \mathbb{N}_0$, for the generating series of binomial coefficients we get $\sum_{n\geq 0}\binom{n}{k}X^n = \frac{X^k}{k!}\cdot\sum_{n\geq k}n_{(k)}X^{n-k} = \frac{X^k}{k!}\cdot(\frac{\partial}{\partial X})^k(\sum_{n\geq 0}X^n) = \frac{X^k}{k!}\cdot(\frac{\partial}{\partial X})^k(\frac{1}{1-X}) = \frac{X^k}{(1-X)^{k+1}} \in \mathbb{Q}[[X]]$.

b) We proceed towards a generalization of the generating series of binomial coefficients: Let K be a field. Then we have $1-X \in K[X] \cap K[[X]]^*$ such that $(1-X)^{-1} = \frac{1}{1-X} = \sum_{n\geq 0}X^n \in K(X) \cap K[[X]]$. Using the combinatorial reciprocity $|\mathcal{M}_k(n)| = \binom{k+n-1}{k} = (-1)^k\cdot\binom{-n}{k}$, for $n \in \mathbb{N}$, this yields the **binomial series** $\frac{1}{(1-X)^n} = \sum_{k\geq 0}|\mathcal{M}_k(n)|\cdot X^k = \sum_{k\geq 0}\binom{-n}{k}\cdot(-X)^k \in K[[X]]$. Thus for all $n \in \mathbb{Z}$ we have $(1+X)^n = \sum_{k\geq 0}\binom{n}{k}X^k \in K[[X]]$, being a polynomial if and only if $n \geq 0$. In particular, the generating series of binomial coefficients considered above equals the binomial series $\sum_{n\geq 0}\binom{n}{k}X^n = \frac{X^k}{(1-X)^{k+1}} =$

$\sum_{n\geq 0}(-1)^n\binom{-k-1}{n}X^{n+k} \in \mathbb{Q}[[X]]$, for $k \in \mathbb{N}_0$.

Let K be a field of characteristic 0. Letting Z be an indeterminate, we let $(1+X)^Z := \sum_{k\geq 0}\binom{Z}{k}X^k \in K[Z][[X]]$. Thus by evaluating for $z \in K$ we let $(1+X)^z := \sum_{k\geq 0}\binom{z}{k}X^k \in K[[X]]$; in particular for $n \in \mathbb{Z}$ we recover the above expression for $(1+X)^n$. Then, letting Y be an indeterminate, the Vandermonde identity implies $(1+X)^Y \cdot (1+X)^Z = (\sum_{i\geq 0}\binom{Y}{i}X^i) \cdot (\sum_{j\geq 0}\binom{Z}{j}X^j) = \sum_{n\geq 0}(\sum_{k=0}^n \binom{Y}{k}\binom{Z}{n-k})X^n = \sum_{n\geq 0}\binom{Y+Z}{n}X^n = (1+X)^{Y+Z} \in K[Y,Z][[X]]$; thus for all $y, z \in K$ we get $(1+X)^y \cdot (1+X)^z = (1+X)^{y+z} \in K[[X]]$. Note that for $K := \mathbb{C}$ we have $(1+X)^z \in \mathbb{C}[[X]]^\infty$ with Taylor series $(1+x)^z = \sum_{k\geq 0}\binom{z}{k}x^k \in \mathbb{C}$, for all $x \in \mathbb{C}$ such that $|x| < 1$.

c) As an application of the above machinery we consider the infinite product $f := \prod_{n\geq 1}(1-X^n)^{\frac{-\mu(n)}{n}} \in \mathbb{Q}[[X]]$, where $\mu\colon \mathbb{N} \to \mathbb{N}$ is the number theoretic Möbius function; since $\nu((1-X^n)^{\frac{-\mu(n)}{n}} - 1) = \nu(\sum_{k\geq 1}\binom{\frac{-\mu(n)}{n}}{k} \cdot (-1)^k X^{kn}) = n$, for $n \in \mathbb{N}$, the infinite product is well-defined. Then we get $\log(f-1) = -\sum_{n\geq 1}\frac{\mu(n)}{n} \cdot \log(-X^n) = \sum_{n\geq 1}(\frac{\mu(n)}{n} \cdot \sum_{k\geq 1}\frac{X^{kn}}{k}) = \sum_{n\geq 1}(\sum_{d\,|\,n}\mu(d)) \cdot \frac{X^n}{n} = \sum_{n\geq 1}\delta_{1,n} \cdot \frac{X^n}{n} = X \in \mathbb{Q}[[X]]$, hence $\exp = \exp(\log(f-1)) = f \in \mathbb{Q}[[X]]$.

(12.4) Rational functions. a) Let $d \in \mathbb{N}$ and $q_1, \ldots, q_d \in \mathbb{C}$ such that $q_d \neq 0$, let $q = 1 + \sum_{i=1}^d q_i X^i = \prod_{j=1}^k (1-a_jX)^{d_j} \in \mathbb{C}[X] \cap \mathbb{C}[[X]]^*$, for some $k \in \mathbb{N}$ and $d_j \in \mathbb{N}$, and pairwise distinct $0 \neq a_1, \ldots, a_k \in \mathbb{C}$; hence we have $\sum_{j=1}^k d_j = d$. Letting $f := \sum_{n\geq 0} f_n X^n \in \mathbb{C}[[X]]$, the following are equivalent:

i) The series f is rational $f = \frac{p}{q} \in \mathbb{C}(X) \cap \mathbb{C}[[X]]$, where $p \in \mathbb{C}[X]_{\leq d-1}$.

ii) We have the **partial fraction decomposition** $f = \sum_{j=1}^k \frac{g_j}{(1-a_jX)^{d_j}} \in \mathbb{C}(X) \cap \mathbb{C}[[X]]$, where $g_j \in \mathbb{C}[X]_{\leq d_j-1}$ for all $j \in \{1, \ldots, k\}$.

iii) The sequence $[f_n \in \mathbb{C}; n \in \mathbb{N}_0]$ is a **linear recurrent sequence** of **degree** d, that is $f_{n+d} + \sum_{i=1}^d q_i f_{n+d-i} = 0$ for all $n \in \mathbb{N}_0$.

iv) We have $f_n = \sum_{j=1}^k h_j(n)a_j^n$, for all $n \in \mathbb{N}_0$, where $h_j \in \mathbb{C}[X]_{\leq d_j-1}$.

In order to see this let $V_{(i)}, V_{(ii)}, V_{(iii)}, V_{(iv)} \leq \mathbb{C}[[X]]$ be the $\mathbb{C}$-subspaces of all formal power series fulfilling property (i), (ii), (iii) and (iv), respectively. We show that these $\mathbb{C}$-subspaces actually coincide:

For $f \in V_{(ii)}$, letting $q_j := (1-a_jX)^{d_j} \in \mathbb{C}[X]$ and $r_j := \frac{q}{q_j} \in \mathbb{C}[X]$, for $j \in \{1, \ldots, k\}$, yields $f = \frac{1}{q} \cdot \sum_{j=1}^k g_jr_j \in \mathbb{C}[[X]]$, where $g_j = 0$ or $\deg(g_j) + \deg(r_j) < d_j + (d-d_j) = d$, showing that $V_{(ii)} \leq V_{(i)}$. Moreover, recalling that $\mathbb{C}[X]$ is an Euclidean domain, let $s_j \in \mathbb{C}[X]$ such that $r_js_j = 1 \in \mathbb{C}[X]/q_j\mathbb{C}[X]$; note that in particular $s_j \in \mathbb{C}[X]/q_j\mathbb{C}[X]$ is invertible. Then, since the $q_j \in \mathbb{C}[X]$ are pairwise coprime, the Chinese remainder theorem shows that the map $\bigoplus_{j=1}^k \mathbb{C}[X]/q_j\mathbb{C}[X] \to \mathbb{C}[X]/q\mathbb{C}[X]\colon [g_1, \ldots, g_k] \to \sum_{j=1}^k g_js_jr_j$ is an isomorphism of $\mathbb{C}$-algebras, thus $\bigoplus_{j=1}^k \mathbb{C}[X]/q_j\mathbb{C}[X] \to \mathbb{C}[X]/q\mathbb{C}[X]\colon [g_1, \ldots, g_k] \to \sum_{j=1}^k g_jr_j$ is as well, in particular is injective. Hence choosing the coefficients of the $g_j \in \mathbb{C}[X]_{\leq d_j-1}$, for $j \in \{1, \ldots, k\}$, shows that $\dim_{\mathbb{C}}(V_{(ii)}) = \sum_{j=1}^k d_j = d$. Similarly, choosing the coefficients of $p \in \mathbb{C}[X]_{\leq d-1}$ shows that $\dim_{\mathbb{C}}(V_{(i)}) = d$, thus we conclude that $V_{(i)} = V_{(ii)}$. Note that, as an alternative, for $f \in V_{(i)} \cup V_{(ii)}$ we may consider the associated Taylor series, which converges for all $x \in \mathbb{C}$ such that $|x| < \frac{1}{|\alpha_j|}$, for all $j \in \{1, \ldots, k\}$, hence $V_{(i)} = V_{(ii)}$ also follows from considering the partial fraction decomposition of **rational maps**.

Since $f \in V_{(iii)}$ is uniquely determined by the initial sequence $[f_0, \ldots, f_{d-1}]$, we have $\dim_{\mathbb{C}}(V_{(iii)}) \leq d$. For $f \in V_{(i)}$, from $(\sum_{n\geq 0} f_n X^n)(1 + \sum_{i=1}^d q_i X^i) = fq = p \in \mathbb{C}[[X]]$, the right hand side being in $\mathbb{C}[X]_{\leq d-1}$, we get $f_{n+d} + \sum_{i=1}^d q_i f_{n+d-i} = 0$, for $n \in \mathbb{N}_0$, showing that $V_{(i)} \leq V_{(iii)}$, hence $V_{(i)} = V_{(iii)}$.

Choosing the coefficients of the h_j, for all $j \in \{1, \ldots, k\}$, yields $\dim_{\mathbb{C}}(V_{(iv)}) \leq \sum_{j=1}^k d_j = d$. Letting $f \in V_{(ii)}$, we may assume that $f = \frac{X^l}{(1-aX)^d} \in \mathbb{C}[[X]]$, where $l \in \{0, \ldots, d-1\}$ and $0 \neq a \in \mathbb{C}$, thus $f = \sum_{n\geq 0} \binom{n+d-1}{d-1}(aX)^n X^l = a^{-l} \cdot \sum_{n\geq 0} \binom{n-l+d-1}{d-1} a^n X^n \in \mathbb{C}[[X]]$, where the polynomial $\binom{X-l+d-1}{d-1} \in \mathbb{C}[X]$ has degree $d-1$, showing that $V_{(ii)} \leq V_{(iv)}$, hence $V_{(ii)} = V_{(iv)}$. ♯

Note that, if f is given by a linear recursion of degree d, then the coefficients of $q \in \mathbb{C}[X]$ can be read off directly, while the coefficients of $p \in \mathbb{C}[X]_{\leq d-1}$ can be determined from $[f_0, \ldots, f_{d-1}]$ using $(\sum_{n\geq 0} f_n X^n)(1 + \sum_{i=1}^d q_i X^i) = p = \sum_{j=0}^{d-1} p_j X^j \in \mathbb{C}[[X]]$ as $p_j = f_j + \sum_{i=1}^j q_i f_{j-i} \in \mathbb{C}$, for $j \in \{0, \ldots, d-1\}$.

b) We have the equivalence of the following assertions, characterizing the degree of recurrence:

i) We have $\gcd(p, q) = 1 \in \mathbb{C}[X]$.
ii) We have $\deg(h_j) = d_j - 1$ for all $j \in \{1, \ldots, k\}$.
iii) The number $d \in \mathbb{N}$ is the smallest degree of recurrence of f.

c) Finally, we have the following characterization of polynomial maps, see (4.1): Given a map $h\colon \mathbb{N}_0 \to \mathbb{C}$, then the following assertions are equivalent:

i) The map h is polynomial of degree $d-1$.
ii) For the generating series we have $\sum_{n\geq 0} h(n)X^n = \frac{p}{(1-X)^d} \in \mathbb{C}(X) \cap \mathbb{C}[[X]]$, where $p \in \mathbb{C}[X]_{\leq d-1}$ such that $1 - X \not| \, p \in \mathbb{C}[X]$, that is $p(1) \neq 0 \in \mathbb{C}$.

In this case $p = \sum_{j=0}^{d-1} p_j X^j \in \mathbb{C}[X]_{\leq d-1}$ is called the **Euler polynomial** of h, its coefficients $[p_0, \ldots, p_{d-1}]$ are called the associated **Euler numbers**.

13 Generating functions

(13.1) Example: Fibonacci numbers. We consider the number $u_n \in \mathbb{N}_0$ of tilings of a $(2 \times n)$-rectangle by (1×2)-rectangles, for $n \in \mathbb{N}_0$. Hence we have $u_0 := 1$ and $u_1 := 1$, and the linear recursion $u_{n+2} = u_{n+1} + u_n$ for $n \in \mathbb{N}_0$. Hence we have $u_n = F_{n+1}$ for $n \in \mathbb{N}_0$, where the $F_n \in \mathbb{N}_0$ are the Fibonacci numbers, see (0.1).

To determine a closed formula for F_n and u_n, we determine the generating series $F := \sum_{n\geq 0} F_n X^n \in \mathbb{Q}[[X]]$ and $u := \sum_{n\geq 0} u_n X^n \in \mathbb{Q}[[X]]$: In view of the recursion we consider the polynomial $1 - X - X^2 = (1-\rho_+ X)(1-\rho_- X) \in \mathbb{R}[X]$, where $\rho_\pm := \frac{1}{2}(1 \pm \sqrt{5}) \in \mathbb{R}$. We have $F = \frac{a+bX}{1-X-X^2} \in \mathbb{Q}[[X]]$, where $a = F_0 = 0$ and $b = F_1 - F_0 = 1$, hence $F = \frac{X}{1-X-X^2} \in \mathbb{Q}[[X]]$; thus we have $u = \sum_{n\geq 0} F_{n+1} X^n = \sum_{n\geq 1} F_n X^{n-1} = \frac{1}{1-X-X^2} \in \mathbb{Q}[[X]]$.

Partial fraction decomposition $\frac{1}{1-X-X^2} = \frac{\rho_+}{\sqrt{5}(1-\rho_+ X)} - \frac{\rho_-}{\sqrt{5}(1-\rho_- X)} \in \mathbb{R}[[X]]$ yields $F = \frac{1}{\sqrt{5}} \cdot \sum_{n\geq 0}(\rho_+^{n+1} - \rho_-^{n+1})X^{n+1} = \frac{1}{\sqrt{5}} \cdot \sum_{n\geq 0}(\rho_+^n - \rho_-^n)X^n \in \mathbb{R}[[X]]$, that is $F_n = \frac{1}{\sqrt{5}} \cdot (\rho_+^n - \rho_-^n) \in \mathbb{R}$, for all $n \in \mathbb{N}_0$; similarly, we get $u = \frac{1}{\sqrt{5}} \cdot \sum_{n\geq 0}(\rho_+^{n+1} - \rho_-^{n+1})X^n$, thus $u_n = \frac{1}{\sqrt{5}} \cdot (\rho_+^{n+1} - \rho_-^{n+1})$, coinciding with F_{n+1}.

(13.2) Example: Catalan numbers. For $n \in \mathbb{N}_0$ let $c_n \in \mathbb{N}$ be the n-th **Catalan number**, which can be defined in numerous ways, for example as the number of non-associative words which can be formed from a sequence of $n+1$ letters, also called **Schröder's (first) problem**. We show that $c_n = \frac{1}{n+1} \cdot \binom{2n}{n}$:

Let $b_n := c_{n-1} \in \mathbb{N}$ be the number of such words containing $n \in \mathbb{N}$ letters. Indicating the sequence of products taken by placing brackets, a word containing $n \geq 2$ letters is of the form $(\cdots)(\cdots)$, with factors containing k and $n-k$ letters, respectively, where $k \in \{1, \ldots, n-1\}$. This yields the recursion $b_n = \sum_{k=1}^{n-1} b_k b_{n-k}$, for $n \geq 2$, where $b_1 = 1$. For example, we have $b_1 = b_2 = 1$ and $b_3 = 2$ and $b_4 = 5$: For $n = 1$ we have $\{\cdot\}$, for $n = 2$ we have $\{\cdot\cdot\}$, for $n = 3$ we have $\{\cdot(\cdot\cdot), (\cdot\cdot)\cdot\}$, and for $n = 4$ we have $\{\cdot(\cdot(\cdot\cdot)), \cdot((\cdot\cdot)\cdot), (\cdot\cdot)(\cdot\cdot), (\cdot(\cdot\cdot))\cdot, ((\cdot\cdot)\cdot)\cdot\}$.

Let $b := \sum_{n\geq 0} b_n X^n \in \mathbb{Q}[[X]]$ be the associated generating series, where we let $b_0 := 0$. Then the above recursion yields $b^2 = \sum_{n\geq 0}(\sum_{k=0}^{n} b_k b_{n-k})X^n = \sum_{n\geq 2}(\sum_{k=1}^{n-1} b_k b_{n-k})X^n = \sum_{n\geq 2} b_n X^n = b - X \in \mathbb{Q}[[X]]$, that is $b^2 - b + X = 0 \in \mathbb{Q}[[X]]$. We look for a holomorphic map $\widehat{b}$, on a suitable open disc of positive radius centered at $x = 0$, fulfilling the functional equation $\widehat{b}(x)^2 - \widehat{b}(x) + x = 0$. Solving the quadratic equation yields $\widehat{b}(x) = \frac{1}{2}(1 - \sqrt{1-4x})$, where the sign is chosen so that $\widehat{b}(0) = 0$. Hence $\widehat{b}$ indeed is holomorphic for all $x \in \mathbb{C}$ such that $|x| < \frac{1}{4}$, and has the Taylor series expansion $\widehat{b}(x) = \sum_{n\geq 0} b_n x^n$ around $x = 0$.

We have $(1-4x)^{\frac{1}{2}} = \sum_{n\geq 0}\binom{\frac{1}{2}}{n}(-4x)^n$ for all $x \in \mathbb{C}$ such that $|x| < \frac{1}{4}$. For $n \in \mathbb{N}_0$ we have $\binom{-\frac{1}{2}}{n} = \frac{(-1)^n \cdot \prod_{i=1}^{n}(2i-1)}{2^n \cdot n!} = \frac{(-1)^n \cdot (2n)!}{2^{2n} \cdot (n!)^2} = (\frac{-1}{4})^n \cdot \binom{2n}{n}$; see also Exercise (19.12). This for $n \geq 1$ yields $\binom{\frac{1}{2}}{n} = \frac{1}{2n} \cdot \binom{-\frac{1}{2}}{n-1} = \frac{(-1)^{n-1}}{2^{2n-1}n} \cdot \binom{2n-2}{n-1}$. Thus we get $\widehat{b}(x) = -\frac{1}{2} \cdot \sum_{n\geq 1} \frac{(-1)^{n-1}(-4)^n}{2^{2n-1}n} \cdot \binom{2n-2}{n-1} x^n = \sum_{n\geq 1} \frac{(2n-2)!}{n!(n-1)!} x^n$; note that it is not at all obvious how to show directly the convergence of the right hand side on an open disc of positive radius centered at $x = 0$. Hence we have $b = \sum_{n\geq 1} \frac{(2n-2)!}{n!(n-1)!} X^n = \sum_{n\geq 1} \frac{1}{n} \cdot \binom{2n-2}{n-1} X^n \in \mathbb{Q}[[X]]$.

(13.3) Exponential generating series. a) Let K be a field of characteristic 0. For a sequence $[f_n \in K; n \in \mathbb{N}_0]$ let $\widetilde{f} := \sum_{n\geq 0} \frac{f_n}{n!} X^n \in K[[X]]$ be the associated **exponential generating series**. In particular, differentiation yields $\frac{\partial}{\partial X}\widetilde{f} = \sum_{n\geq 1} \frac{nf_n}{n!} X^{n-1} = \sum_{n\geq 0} \frac{f_{n+1}}{n!} X^n \in K[[X]]$, which amounts to forming the exponential generating series of a shift of the given sequence.

Given $\widetilde{f} = \sum_{n\geq 0} \frac{f_n}{n!} X^n \in K[[X]]$ and $\widetilde{g} = \sum_{n\geq 0} \frac{g_n}{n!} X^n \in K[[X]]$, and letting $\widetilde{h} := \widetilde{f}\widetilde{g} = \sum_{n\geq 0} \frac{h_n}{n!} X^n \in K[[X]]$, convolutional multiplication becomes $\widetilde{h} = \sum_{n\geq 0}(\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} f_k g_{n-k}) \cdot \frac{1}{n!} X^n$, saying that $h_n = \sum_{k=0}^{n} \binom{n}{k} f_k g_{n-k} \in K$, for all $n \in \mathbb{N}_0$, that is the sequence $[h_n \in K; n \in \mathbb{N}_0]$ is given as **binomial convolution** of the sequences $[f_n \in K; n \in \mathbb{N}_0]$ and $[g_n \in K; n \in \mathbb{N}_0]$.

For example, considering the sequence $[a_{(n)}; n \in \mathbb{N}_0]$ of falling factorials associated with $a \in K$ we get $\sum_{n\geq 0} \frac{a_{(n)}}{n!} X^n = \sum_{n\geq 0} \binom{a}{n} X^n = (1+X)^a \in K[[X]]$; hence from the identity $(1+X)^{a+b} = (1+X)^a(1+X)^b \in K[[X]]$, for $a, b \in K$, we recover the formula $(a+b)_{(n)} = \sum_{k=0}^{n} \binom{n}{k} a_{(k)} b_{(n-k)} \in K$, that is the Vandermonde identity $\binom{a+b}{n} = \sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} \in K$, for all $n \in \mathbb{N}_0$.

b) The exponential generating series of the geometric series associated with $a \in K$ is given as $\sum_{n\geq 0} \frac{a^n}{n!} X^n = \exp(aX) \in K[[X]]$; hence from the identity $\exp((a+b)X) = \exp(aX)\exp(bX) \in K[[X]]$, for $a, b \in K$, we recover the binomial formula $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \in K$, for all $n \in \mathbb{N}_0$.

In particular, using the constant series associated with $a := 1$ we get the following: For sequences $[f_n \in K; n \in \mathbb{N}_0]$ and $[g_n \in K; n \in \mathbb{N}_0]$ being related by $g_n = \sum_{k=0}^{n} \binom{n}{k} f_k \in K$, for all $n \in \mathbb{N}_0$, we have $\widetilde{g} = \widetilde{f} \cdot \exp \in K[[X]]$, implying $\widetilde{f} = \widetilde{g} \cdot \exp(X)^{-1} = \widetilde{g} \cdot \exp(-X) \in K[[X]]$, from which we in turn recover the binomial inversion formula $f_n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} g_k \in K$, for all $n \in \mathbb{N}_0$.

For example, to determine the exponential generating series $\widetilde{D} \in \mathbb{Q}[[X]]$ of the sequence $[D_n; n \in \mathbb{N}_0]$ of derangement numbers we proceed as follows: We have $\sum_{k=0}^{n} \binom{n}{k} D_{n-k} = n!$, for all $n \in \mathbb{N}_0$, hence $\widetilde{D} \cdot \exp = \sum_{n\geq 0} \frac{n!}{n!} X^n = (1-X)^{-1} \in \mathbb{Q}[[X]]$, thus $\widetilde{D} = \frac{\exp(-X)}{1-X} \in \mathbb{Q}[[X]]$; in particular, from this we recover the formula $D_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \in \mathbb{Q}$, for all $n \in \mathbb{N}_0$.

(13.4) Example: Stirling numbers of the second kind. a) For $n, k \in \mathbb{N}_0$ let $S_{n,k} \in \mathbb{N}_0$ be the associated Stirling number of the second kind, and let $S_k := \sum_{n\geq 0} S_{n,k} X^n \in \mathbb{Q}[[X]]$ be the associated generating series, where hence $S_0 = 1$. We show that $S_k = \prod_{l=1}^{k} \frac{X}{1-lX} \in \mathbb{Q}[[X]]$, for all $k \in \mathbb{N}_0$:

We proceed by induction on $k \in \mathbb{N}_0$; the assertion being true for $k = 0$, let $k \geq 1$. The recursion $S_{n,k} = S_{n-1,k-1} + k \cdot S_{n-1,k}$, for $n, k \geq 1$, yields $S_k = X \cdot (\sum_{n\geq k} S_{n-1,k-1} X^{n-1} + k \cdot \sum_{n\geq k+1} S_{n-1,k} X^{n-1}) = X(S_{k-1} + kS_k) \in \mathbb{Q}[[X]]$, thus $(1-kX)S_k = XS_{k-1} \in \mathbb{Q}[[X]]$, implying $S_k = \frac{X}{1-kX} \cdot S_{k-1} = \frac{X}{1-kX} \cdot \prod_{l=1}^{k-1} \frac{X}{1-lX} = \prod_{l=1}^{k} \frac{X}{1-lX} \in \mathbb{Q}[[X]]$. ♯

b) Let $\widetilde{S}_k := \sum_{n\geq 0} \frac{S_{n,k}}{n!} X^n \in \mathbb{Q}[[X]]$ be the associated exponential generating series, where $\widetilde{S}_0 = 1$. We show that $\widetilde{S}_k = \frac{1}{k!}(\exp - 1)^k \in \mathbb{Q}[[X]]$, for all $k \in \mathbb{N}_0$:

The assertion being true for $k = 0$, let $k \geq 1$. The recursion $S_{n,k} = S_{n-1,k-1} + k \cdot S_{n-1,k}$, for $n, k \geq 1$, yields $\widetilde{S}_k = \sum_{n\geq k} \frac{S_{n-1,k-1}}{n!} X^n + k \cdot \sum_{n\geq k+1} \frac{S_{n-1,k}}{n!} X^n \in \mathbb{Q}[[X]]$, leading to the **differential equation** $\frac{\partial}{\partial X}\widetilde{S}_k = \sum_{n\geq k} \frac{S_{n-1,k-1}}{(n-1)!} X^{n-1} + k \cdot \sum_{n\geq k+1} \frac{S_{n-1,k}}{(n-1)!} X^{n-1} = \widetilde{S}_{k-1} + k\widetilde{S}_k \in \mathbb{Q}[[X]]$. Since $\widetilde{S}_{k-1} \in \mathbb{Q}[[X]]$ is known, and the constant term of $\widetilde{S}_k$ is $S_{0,k} = 0$, by successively comparing coefficients we conclude that $\widetilde{S}_k \in \mathbb{Q}[[X]]$ is uniquely determined; note that this is also follows from the fact that the differential equation $\frac{\partial}{\partial X}\widehat{\widetilde{S}}_k(x) = \widehat{\widetilde{S}}_{k-1}(x) + k\widehat{\widetilde{S}}_k(x)$, given any initial value $\widehat{\widetilde{S}}_k(0) \in \mathbb{C}$, has a unique solution in an open disc of positive radius centered at $x = 0$. Hence it suffices to show that $\widetilde{T}_k := \frac{1}{k!}(\exp - 1)^k \in \mathbb{Q}[[X]]$ is a solution of the above equation: We have $\frac{\partial}{\partial X}\widetilde{T}_k = \frac{1}{(k-1)!}(\exp - 1)^{k-1} \exp = \frac{1}{(k-1)!}(\exp - 1)^{k-1} + \frac{k}{k!}(\exp - 1)^k = \widetilde{T}_{k-1} + k\widetilde{T}_k \in \mathbb{Q}[[X]]$. ♯

Note that $\widetilde{S}_k = \sum_{n\geq 0} \frac{S_{n,k}}{n!} X^n = \frac{1}{k!}(\exp - 1)^k = \frac{1}{k!} \cdot \sum_{l=0}^{k} (-1)^{k-l}\binom{k}{l} \exp(lX) = \frac{1}{k!} \cdot \sum_{l=0}^{k}((-1)^{k-l}\binom{k}{l} \cdot \sum_{n\geq 0} \frac{l^n}{n!} X^n) = \frac{1}{k!} \cdot \sum_{n\geq 0}(\sum_{l=0}^{k} (-1)^{k-l}\binom{k}{l} l^n)\frac{X^n}{n!} \in \mathbb{Q}[[X]]$ yields the sum formula $S_{n,k} = \frac{1}{k!} \cdot \sum_{l=0}^{k} (-1)^{k-l}\binom{k}{l} l^n \in \mathbb{Q}$ again.

c) Let $B_n := \sum_{k=0}^{n} S_{n,k} \in \mathbb{N}$ be the n-th Bell number. The associated exponential generating series is $\widetilde{B} := \sum_{n\geq 0} \frac{B_n}{n!} X^n = \sum_{k\geq 0} \widetilde{S}_k = \sum_{k\geq 0} \frac{1}{k!}(\exp - 1)^k = \exp(\exp - 1) \in \mathbb{Q}[[X]]$. Thus we get $\sum_{n\geq 0} \frac{B_{n+1}}{n!} X^n = \frac{\partial}{\partial X}\widetilde{B} = \exp(\exp - 1) \cdot \exp = \widetilde{B} \cdot \exp = \sum_{n\geq 0}(\sum_{k=0}^{n} \frac{B_k}{k!(n-k)!})X^n \in \mathbb{Q}[[X]]$, and comparing coefficients yields $B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$, for all $n \in \mathbb{N}_0$; for a combinatorial proof see (19.23).

The Taylor series $\widehat{\widetilde{B}}(x)$ converges for all $x \in \mathbb{C}$. This yields $\sum_{n\geq 0} B_n \cdot \frac{x^n}{n!} = \widehat{\widetilde{B}}(x) = \exp(\exp(x) - 1) = \frac{1}{e} \cdot \exp(\exp(x)) = \frac{1}{e} \cdot \sum_{k\geq 0} \frac{\exp(x)^k}{k!} = \frac{1}{e} \cdot \sum_{k\geq 0} \frac{\exp(kx)}{k!} = \frac{1}{e} \cdot \sum_{k\geq 0}(\frac{1}{k!} \cdot \sum_{n\geq 0} \frac{(kx)^n}{n!}) = \frac{1}{e} \cdot \sum_{n\geq 0}(\sum_{k\geq 0} \frac{k^n}{k!}) \cdot \frac{x^n}{n!} \in \mathbb{C}$, for all $x \in \mathbb{C}$, hence $B_n = \frac{1}{e} \cdot \sum_{k\geq 0} \frac{k^n}{k!} \in \mathbb{C}$; note that this cannot be translated to the formal setting.

(13.5) Example: Stirling numbers of the first kind. For $n, k \in \mathbb{N}_0$ let $s_{n,k} \in \mathbb{N}_0$ be the associated Stirling number of the first kind, and let $\widetilde{s}_k := \sum_{n\geq 0} \frac{s_{n,k}}{n!} X^n \in \mathbb{Q}[[X]]$ be the associated exponential generating series, where hence $\widetilde{s}_0 = 1$. We show that $\widetilde{s}_k = \frac{(-1)^k}{k!} \log(-X)^k \in \mathbb{Q}[[X]]$, for all $k \in \mathbb{N}_0$:

We proceed by induction on $k \in \mathbb{N}_0$; the assertion being true for $k = 0$, let $k \geq 1$. The recursion $s_{n,k} = s_{n-1,k-1} + (n-1) \cdot s_{n-1,k}$ yields $\widetilde{s}_k = \sum_{n\geq k} \frac{s_{n-1,k-1}}{n!} X^n + \sum_{n\geq k+1} \frac{(n-1)s_{n-1,k}}{n!} X^n \in \mathbb{Q}[[X]]$, thus we get $\frac{\partial}{\partial X}\widetilde{s}_k = \sum_{n\geq k} \frac{s_{n-1,k-1}}{(n-1)!} X^{n-1} + \sum_{n\geq k+1} \frac{s_{n-1,k}}{(n-2)!} X^{n-1} \in \mathbb{Q}[[X]]$. Recall that the above recursion entails the identity $\frac{s_{n,k}}{(n-1)!} = \sum_{m=0}^{n-1} \frac{s_{m,k-1}}{m!}$, thus we get $\widetilde{s}_{k-1} \cdot (\frac{1}{1-X} - 1) = (\sum_{n\geq 0} \frac{s_{n,k-1}}{n!} X^n) \cdot (\sum_{l\geq 1} X^l) = \sum_{n\geq 1}(\sum_{m=0}^{n-1} \frac{s_{m,k-1}}{m!})X^n = \sum_{n\geq 1} \frac{s_{n,k}}{(n-1)!} X^n \in \mathbb{Q}[[X]]$, hence we arrive at the differential equation $\frac{\partial}{\partial X}\widetilde{s}_k = \widetilde{s}_{k-1} + \widetilde{s}_{k-1} \cdot (\frac{1}{1-X} - 1) = \frac{1}{1-X} \cdot \widetilde{s}_{k-1} \in \mathbb{Q}[[X]]$, which determines $\widetilde{s}_k$ uniquely, since its constant term is known to be $s_{0,k} = 0$. Thus it suffices to show that $\widetilde{t}_k := \frac{(-1)^k}{k!} \log(-X)^k \in \mathbb{Q}[[X]]$ is a solution: We have $\frac{\partial}{\partial X}\widetilde{t}_k = \frac{(-1)^k}{(k-1)!} \log(-X)^{k-1} \cdot \frac{-1}{1-X} = \frac{1}{1-X} \cdot \widetilde{t}_{k-1} \in \mathbb{Q}[[X]]$. ♯

14 Partition identities

(14.1) Partitions. a) Let $p_{n,k} := |P_k(n)| \in \mathbb{N}_0$ be the number of partitions of $n \in \mathbb{N}_0$ into $k \in \mathbb{N}_0$ parts, let $p_{n,\leq k} := |P_{\leq k}(n)| = \sum_{l=0}^{k} p_{n,l} \in \mathbb{N}_0$ be the number of partitions of n into at most k parts, and let $p_n := |P(n)| = \sum_{k=0}^{n} p_{n,k} \in \mathbb{N}_0$ be the number of all partitions of n; recall that $p_{n,k} = 0$ for $k > n$. These numbers are encoded into the **weighted generating series** $P := \sum_{n\geq 0}(\sum_{k=0}^{n} p_{n,k}Y^k)X^n \in \mathbb{Q}[Y][[X]]$, hence the ordinary generating series $p := \sum_{n\geq 0} p_n X^n \in \mathbb{Q}[[X]]$ is obtained by evaluating P at $y := 1$. We let $P_k := \sum_{n\geq k} p_{n,k}X^n \in X^k\mathbb{Q}[[X]]$ and $P_{\leq k} := \sum_{n\geq 0} p_{n,\leq k}X^n = \sum_{l=0}^{k} P_k \in \mathbb{Q}[[X]]$, hence $P = \sum_{k\geq 0}(\sum_{n\geq k} p_{n,k}X^n)Y^k = \sum_{k\geq 0} P_k Y^k \in \mathbb{Q}[[X]][[Y]]$.

Writing partitions in terms of multiplicities as $\lambda = [n^{a_n(\lambda)}, \ldots, 1^{a_1(\lambda)}] \vdash n$, the map $\lambda \mapsto [a_1(\lambda), a_2(\lambda), \ldots]$ is a bijection from the set $\coprod_{n\geq 0} P(n)$ of all partitions to the set $\mathcal{M} := \{[a_1, a_2, \ldots] \in \mathrm{Maps}(\mathbb{N}, \mathbb{N}_0); a_i = 0 \text{ for almost all } i \in \mathbb{N}\}$; hence $a \in \mathcal{M}$ corresponds to a partition $\lambda(a)$ of $|\lambda(a)| = \sum_{i\geq 1} ia_i$ with $l(\lambda(a)) = \sum_{i\geq 1} a_i$ parts. Using this we get $P = \sum_{n\geq 0}(\sum_{\lambda\vdash n} Y^{l(\lambda)})X^n = \sum_{a\in\mathcal{M}} Y^{l(\lambda(a))}X^{|\lambda(a)|} = \sum_{a\in\mathcal{M}} Y^{\sum_{i\geq 1} a_i}X^{\sum_{i\geq 1} ia_i} = \sum_{a\in\mathcal{M}} \prod_{i\geq 1}(YX^i)^{a_i} = \prod_{i\geq 1}(\sum_{j\geq 0}(YX^i)^j) = \prod_{i\geq 1} \frac{1}{1-YX^i} \in \mathbb{Q}[Y][[X]]$. Hence evaluating at $y := 1$ yields $p = \prod_{i\geq 1} \frac{1}{1-X^i} \in \mathbb{Q}[[X]]$.

It turns out that $p \in \mathbb{Q}[[X]]$ has a close relation to number theory: We have $\log(p-1) = -\sum_{n\geq 1} \log(-X^n) \in \mathbb{Q}[[X]]$, hence we get $X \cdot \frac{\partial}{\partial X}(\log(p-1)) = \sum_{n\geq 1} \frac{nX^n}{1-X^n} = \sum_{n\geq 1}(n \cdot \sum_{k\geq 1} X^{kn}) = \sum_{n\geq 1}(\sum_{d\,|\,n} d)X^n = \sum_{n\geq 1} \sigma(n)X^n =: \sigma \in \mathbb{Q}[[X]]$, where $\sigma(n) := \sum_{d\,|\,n} d \in \mathbb{N}$ is the sum of the divisors of $n \in \mathbb{N}$. Since $X \cdot \frac{\partial}{\partial X}(\log(p-1)) = \frac{X}{p} \cdot \frac{\partial}{\partial X}(p) \in \mathbb{Q}[[X]]$ is a shifted **logarithmic derivative** of p, we obtain $X \cdot \frac{\partial}{\partial X}(p) = p \cdot \sigma \in \mathbb{Q}[[X]]$, yielding $\sum_{n\geq 1} np_nX^n = (\sum_{n\geq 0} p_nX^n) \cdot (\sum_{n\geq 1} \sigma(n)X^n) = \sum_{n\geq 1}(\sum_{i=1}^{n} \sigma(i)p_{n-i})X^n \in \mathbb{Q}[[X]]$. Thus we have proved the recursion $n \cdot p_n = \sum_{i=1}^{n} \sigma(i) \cdot p_{n-i}$ for all $n \in \mathbb{N}_0$.

b) We determine $P_{\leq k} \in \mathbb{Q}[[X]]$ and $P_k \in \mathbb{Q}[[X]]$. To this end, we write $\lambda \in P_{\leq k}(n)$ as $\lambda = [\lambda_1, \ldots, \lambda_k]$ where $\lambda_1 \geq \cdots \geq \lambda_k \geq 0$, and let $P_{\leq k}^{\geq j}(n) := \{\lambda \in P_{\leq k}(n); \lambda_k \geq j\}$, for $j \in \mathbb{N}_0$. For the associated generating series we show by induction on $k \in \mathbb{N}_0$ that $P_{\leq k}^{\geq j} = \prod_{i=1}^{k} \frac{X^j}{1-X^i} \in \mathbb{Q}[[X]]$: The case $k = 0$ being trivial, let $k \geq 1$. Then we have $P_{\leq k}^{\geq j} = \sum_{\lambda_1\geq\cdots\geq\lambda_k\geq j} X^{\sum_{i=1}^{k} \lambda_i} = \sum_{\lambda_k\geq j}(X^{\lambda_k} \cdot \sum_{\lambda_1\geq\cdots\geq\lambda_{k-1}\geq\lambda_k} X^{\sum_{i=1}^{k-1} \lambda_i})$, thus by induction $P_{\leq k}^{\geq j} = \sum_{\lambda_k\geq j}(X^{\lambda_k} \cdot \prod_{i=1}^{k-1} \frac{X^{\lambda_k}}{1-X^i}) = (\prod_{i=1}^{k-1} \frac{1}{1-X^i}) \cdot \sum_{\lambda_k\geq j} X^{k\lambda_k} = (\prod_{i=1}^{k-1} \frac{1}{1-X^i}) \cdot \frac{X^{jk}}{1-X^k} = \prod_{i=1}^{k} \frac{X^j}{1-X^i}$.

Thus in particular we get $P_{\leq k} = P_{\leq k}^{\geq 0} = \prod_{i=1}^k \frac{1}{1-X^i} \in \mathbb{Q}[[X]]$ and $P_k = P_{\leq k}^{\geq 1} = \prod_{i=1}^k \frac{X}{1-X^i} \in \mathbb{Q}[[X]]$. Note that this implies $p = \lim_{k\to\infty} P_{\leq k} = \prod_{i\geq 1} \frac{1}{1-X^i} \in \mathbb{Q}[[X]]$. Moreover, for $k \geq 1$ we get $P_k = P_{\leq k} - P_{\leq k-1} = (\frac{1}{1-X^k} - 1) \cdot \prod_{i=1}^{k-1} \frac{1}{1-X^i} = X^k \cdot \prod_{i=1}^k \frac{1}{1-X^i} = X^k \cdot P_{\leq k} \in \mathbb{Q}[[X]]$, while $P_0 = P_{\leq 0} = 1 \in \mathbb{Q}[[X]]$ anyway, which is reminiscent of the bijection $P_k(n) \to P_{\leq k}(n-k)\colon [\lambda_1, \ldots, \lambda_k] \mapsto [\lambda_1 - 1, \ldots, \lambda_k - 1]$. Finally, this implies the partition identity $\prod_{i\geq 1} \frac{1}{1-YX^i} = P = \sum_{k\geq 0} P_k Y^k = \sum_{k\geq 0}(\prod_{i=1}^k \frac{X}{1-X^i})Y^k = \sum_{k\geq 0} \prod_{i=1}^k \frac{XY}{1-X^i} = \sum_{k\geq 0}(\prod_{i=1}^k \frac{1}{1-X^i})X^k Y^k \in \mathbb{Q}[[X]][[Y]]$.

(14.2) Regular partitions. a) Let $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$. For $r \geq 2$ let $P_{r',k}(n) \subseteq P_k(n)$ be the set of partitions of n consisting of k parts none of which is divisible by r, and let $P_{r\text{-reg},k}(n) \subseteq P_k(n)$ be the set of r-regular k-part partitions of n, that is those all of whose parts have multiplicity less than r. Then we have $P_{r'}(n) = \coprod_{k=0}^n P_{r',k}(n)$ and $P_{r\text{-reg}}(n) = \coprod_{k=0}^n P_{r\text{-reg},k}(n)$. For $r = 2$ we get the sets $O_k(n) := P_{2',k}(n)$ of partitions of n consisting of k odd parts, and $D_k(n) := P_{2\text{-reg},k}(n)$ of partitions of n having k pairwise distinct parts; hence $O(n) = \coprod_{k=0}^n O_k(n)$ and $D(n) = \coprod_{k=0}^n D_k(n)$.

Then $\coprod_{n\geq 0} P_{r'}(n)$ corresponds to $\mathcal{M}_{r'} := \{a \in \mathcal{M}; a_{ri} = 0 \text{ for all } i \in \mathbb{N}\}$, which yields $P_{r'} := \sum_{n\geq 0}(\sum_{k=0}^n |P_{r',k}(n)|Y^k)X^n = \sum_{a\in\mathcal{M}_{r'}} \prod_{i\geq 1}(YX^i)^{a_i} = \prod_{i\geq 1, r\nmid i}(\sum_{j\geq 0}(YX^i)^j) = \prod_{i\geq 1, r\nmid i} \frac{1}{1-YX^i} = \prod_{i\geq 1} \frac{1-YX^{ri}}{1-YX^i} \in \mathbb{Q}[Y][[X]]$ and thus $p_{r'} := \sum_{n\geq 0} |P_{r'}(n)|X^n = \prod_{i\geq 1} \frac{1-X^{ri}}{1-X^i} \in \mathbb{Q}[[X]]$ as weighted and ordinary generating series, respectively; for $r = 2$ we get $p_{2'} = \prod_{i\geq 1}(1 + X^i) \in \mathbb{Q}[[X]]$.

Then $\coprod_{n\geq 0} P_{r\text{-reg}}(n)$ corresponds to $\mathcal{M}_{r\text{-reg}} := \{a \in \mathcal{M}; a_i < r \text{ for all } i \in \mathbb{N}\}$, thus $P_{r\text{-reg}} := \sum_{n\geq 0}(\sum_{k=0}^n |P_{r\text{-reg},k}(n)|Y^k)X^n = \sum_{a\in\mathcal{M}_{r\text{-reg}}} \prod_{i\geq 1}(YX^i)^{a_i} = \prod_{i\geq 1}(\sum_{j=0}^{r-1}(YX^i)^j) = \prod_{i\geq 1} \frac{1-(YX^i)^r}{1-YX^i} \in \mathbb{Q}[Y][[X]]$ and hence we get $p_{r\text{-reg}} := \sum_{n\geq 0} |P_{r\text{-reg}}(n)|X^n = \prod_{i\geq 1} \frac{1-X^{ir}}{1-X^i} \in \mathbb{Q}[[X]]$ as weighted and ordinary generating series, respectively. Thus we recover $|P_{r'}(n)| = |P_{r\text{-reg}}(n)|$ for all $n \in \mathbb{N}_0$.

b) In particular, for $r = 2$ wet get $D := P_{2\text{-reg}} = \prod_{i\geq 1} \frac{1-(YX^i)^2}{1-YX^i} = \prod_{i\geq 1}(1 + YX^i) \in \mathbb{Q}[Y][[X]]$. The bijection $D_k(n) \to P_{\leq k}(n - \binom{k+1}{2})\colon [\lambda_1, \ldots, \lambda_k] \mapsto [\lambda_1 - k, \lambda_1 - (k-1), \ldots, \lambda_k - 1]$ yields $D_k := \sum_{n\geq 0} |D_k(n)|X^n = X^{\binom{k+1}{2}} \cdot P_{\leq k} = \prod_{i=1}^k \frac{X^i}{1-X^i} \in \mathbb{Q}[[X]]$, yielding the identity $\prod_{i\geq 1}(1+YX^i) = D = \sum_{k\geq 0} D_k Y^k = \sum_{k\geq 0}(\prod_{i=1}^k \frac{X^i}{1-X^i})Y^k = \sum_{k\geq 0}(\prod_{i=1}^k \frac{1}{1-X^i})X^{\binom{k+1}{2}}Y^k \in \mathbb{Q}[[X]][[Y]]$.

Moreover, let $C_k(n) := D_k(n) \cap O_k(n)$ be the set of partitions of n consisting of k pairwise distinct odd parts, and let $C(n) := \coprod_{k\geq 0} C_k(n)$. Then $\coprod_{n\geq 0} C(n)$ corresponds to $\mathcal{M}_C := \{a \in \mathcal{M}; a_{2i} = 0, a_{2i-1} \leq 1 \text{ for all } i \in \mathbb{N}\}$. Hence we get $C := \sum_{n\geq 0}(\sum_{k=0}^n |C_k(n)|Y^k)X^n = \sum_{a\in\mathcal{M}_C} \prod_{i\geq 1}(YX^i)^{a_i} = \prod_{i\geq 1}(1 + YX^{2i-1}) \in \mathbb{Q}[Y][[X]]$, thus $c := \sum_{n\geq 0} |C(n)|X^n = \prod_{i\geq 1}(1 + X^{2i-1}) \in \mathbb{Q}[[X]]$. Note that evaluating $D \in \mathbb{Q}[[X]][[Y]]$ at $X \mapsto X^2$ and $Y \mapsto \frac{Y}{X}$ yields $C = \prod_{i\geq 1}(1 + YX^{2i-1}) = \sum_{k\geq 0}(\prod_{i=1}^k \frac{1}{1-X^{2i}})X^{k^2}Y^k \in \mathbb{Q}((X))[[Y]]$.

(14.3) Conjugate partitions. a) Let $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$. Let $P^k(n) := \{\lambda \in P(n); \lambda_1 = k\}$ be the set of all partitions of n whose largest part is k, and $P^{\leq k}(n) := \coprod_{l=0}^k P^l(n)$ be the those whose largest part is at most k, and let $p_n^k := |P^k(n)| \in \mathbb{N}_0$ and $p_n^{\leq k} := |P^{\leq k}(n)| = \sum_{l=0}^k p_n^l \in \mathbb{N}_0$.

Then $\coprod_{n \geq 0} P^{\leq k}(n)$ corresponds to $\mathcal{M}^{\leq k} := \text{Maps}(\{1, \ldots, k\}, \mathbb{N}_0)$, thus $P^{\leq k} := \sum_{n \geq 0} p_n^{\leq k} X^n = \sum_{a \in \mathcal{M}^{\leq k}} \prod_{i=1}^k X^{ia_i} = \prod_{i=1}^k \frac{1}{1-X^i} \in \mathbb{Q}[[X]]$. And $\coprod_{n \geq 0} P^k(n)$ corresponds to $\mathcal{M}^{=k} := \{a \in \mathcal{M}^{\leq k}; a_k \geq 1\}$, thus $P^{=k} := \sum_{n \geq 0} p_n^k X^n = \sum_{a \in \mathcal{M}^{=k}} \prod_{i=1}^k X^{ia_i} = X^k \cdot \prod_{i=1}^k \frac{1}{1-X^i} = \prod_{i=1}^k \frac{X}{1-X^i} \in \mathbb{Q}[[X]]$; alternatively, we get $P^{=k} = P^{\leq k} - P^{\leq (k-1)} = (\frac{1}{1-X^k} - 1) \cdot \prod_{i=1}^{k-1} \frac{1}{1-X^i} = X^k \cdot \prod_{i=1}^k \frac{1}{1-X^i} \in \mathbb{Q}[[X]]$, for $k \geq 1$. Hence for all $k \in \mathbb{N}_0$ we have $P^{=k} = P_k \in \mathbb{Q}[[X]]$ and $P^{\leq k} = P_{\leq k} \in \mathbb{Q}[[X]]$, for all $k \in \mathbb{N}_0$, that is $p_n^k = p_{n,k}$ and $p_n^{\leq k} = p_{n,\leq k}$, for all $n \in \mathbb{N}_0$.

This can also be deduced combinatorially from taking conjugate partitions: If $\lambda \vdash n$ has k parts then we get $\lambda'_1 = k$; and if $\lambda_1 = k$ then we get $\lambda'_k = |\{j \in \mathbb{N}; \lambda_j \geq k\}| > 0$ and $\lambda'_{k+1} = |\{j \in \mathbb{N}; \lambda_j \geq k+1\}| = 0$, saying that λ' has k parts. Hence conjugation yields bijections $P_k(n) \to P^k(n)$ and $P_{\leq k}(n) \to P^{\leq k}(n)$.

b) Let $P_{\text{sc}}(n) := \{\lambda \in P(n); \lambda = \lambda'\}$ be the set of **self-conjugate** partitions of n. We proceed to determine their number: To this end, given any $\lambda \vdash n$, let $r(\lambda) := \max\{i \in \mathbb{N}; \lambda_i \geq i\} \in \mathbb{N}$ for $n \geq 1$, and $r([]) := 0$, be the **rank** of λ; note that the largest square fitting into the upper left hand corner of the Young diagram of λ has edge length $r(\lambda)$, it is called the **Durfee square** of λ. Moreover, let $h_i(\lambda) := (\lambda_i - i) + (\lambda'_i - i) + 1 \in \mathbb{N}$, for $i \in \{1, \ldots, r(\lambda)\}$, be the associated **(diagonal) hook lengths**; note that $h_i(\lambda)$ is the length of the **hook** centered at the i-th box on main diagonal of the Young diagram of λ, where $\lambda_i - i \in \mathbb{N}_0$ and $\lambda'_i - i \in \mathbb{N}_0$ are the associated **arm** and **leg lengths**, respectively. Hence we get the partition $h(\lambda) := [h_1(\lambda), \ldots, h_{r(\lambda)}(\lambda)]$ of n, consisting of $r(\lambda)$ pairwise distinct parts, giving rise to the map $P(n) \to D(n) \colon \lambda \mapsto h(\lambda)$.

If $\lambda \in P_{\text{sc}}(n)$, then $h(\lambda) = [2(\lambda_i - i) + 1 \in \mathbb{N}; i \in \{1, \ldots, r(\lambda)\}]$ consists of odd parts, that is $h(\lambda) \in O_{r(\lambda)}(n) \cap D_{r(\lambda)}(n) = C_{r(\lambda)}(n)$, and thus we get the restriction $P_{\text{sc}}(n) \to C(n) \colon \lambda \mapsto h(\lambda)$, which moreover is injective. We show surjectivity: Given $\mu = [\mu_1, \ldots, \mu_n] \in C(n)$, let $\lambda_i := \frac{\mu_i - 1}{2} + i$ for $i \in \{1, \ldots, l(\mu)\}$, and $\lambda_i := |\{j \in \{1, \ldots, l(\mu)\}; \lambda_j \geq i\}|$ for $i > l(\mu)$. Then, since $\mu_i - \mu_{i+1} \geq 2$, for all $i \in \{1, \ldots, l(\mu) - 1\}$, we have $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{l(\mu)} \geq l(\mu) \geq \lambda_{l(\mu)+1} \geq \lambda_{l(\mu)+2} \geq \cdots \geq 0$, and from $\sum_{i=1}^{l(\mu)} \lambda_i = \frac{1}{2} \cdot \sum_{i=1}^{l(\mu)} \mu_i - \frac{l(\mu)}{2} + \binom{l(\mu)+1}{2} = \frac{n}{2} + \frac{l(\mu)^2}{2}$ and $\sum_{i > l(\mu)} \lambda_i = \sum_{i > l(\mu)} |\{j \in \{1, \ldots, l(\mu)\}; \lambda_j \geq i\}| = \sum_{j=1}^{l(\mu)} |\{i > l(\mu); i \leq \lambda_j\}| = \sum_{j=1}^{l(\mu)} (\lambda_j - l(\mu)) = \sum_{j=1}^{l(\mu)} \lambda_j - l(\mu)^2$ we get $\sum_{i \geq 1} \lambda_i = 2 \cdot \sum_{i=1}^{l(\mu)} \lambda_i - l(\mu)^2 = n$. Hence $\lambda := [\lambda_1, \ldots, \lambda_n]$ is a partition of n, such that $r(\lambda) = l(\mu)$, which by construction is self-conjugate such that $h(\lambda) = \mu$.

Thus we conclude $P_{\text{sc}} := \sum_{n \geq 0} |P_{\text{sc}}(n)| X^n = c = \prod_{i \geq 1} (1 + X^{2i-1}) \in \mathbb{Q}[[X]]$.

(14.4) Euler's identity. We reconsider the weighted generating series $D = \sum_{n \geq 0} (\sum_{k=0}^n |D_k(n)| Y^k) X^n = \prod_{i \geq 1} (1 + YX^i) \in \mathbb{Q}[Y][[X]]$. Thus evaluating at $y := -1$ yields the identity $\sum_{n\geq 0}(|D_{\text{even}}(n)| - |D_{\text{odd}}(n)|)X^n = \prod_{i\geq 1}(1 - X^i) \in \mathbb{Q}[[X]]$, where $D_{\text{even}}(n) := \{\lambda \in D(n); l(\lambda) \text{ even}\}$ and $D_{\text{odd}}(n) := \{\lambda \in D(n); l(\lambda) \text{ odd}\}$ are the sets of distinct-part partitions of $n \in \mathbb{N}_0$ having an even respectively an odd number of parts; note that $\prod_{i\geq 1}(1 - X^i) = p^{-1} \in \mathbb{Q}[[X]]$.

We proceed to show the **pentagonal number theorem**, saying that $d_n := |D_{\text{even}}(n)| - |D_{\text{odd}}(n)| \neq 0$ only if $n \in \{0\} \dot\cup \{n_k^{\pm} \in \mathbb{N}; k \in \mathbb{N}\}$, where $n_k^{\pm} := \frac{1}{2}k(3k \pm 1) \in \mathbb{N}$, and that $d_{n_k^{\pm}} = (-1)^k$; a few examples are given in Table 12. The relevant numbers are $n^+ = [2, 7, 15, 26, 40, 57, 77, 100, 126, 155, \ldots]$ and $n^- = [1, 5, 12, 22, 35, 51, 70, 92, 117, 145, \ldots]$, where the latter sequence explains the terminology inasmuch this is the number of vertices in a configuration of nested 5-gons; note that the numbers $n_k^{\pm}$, for $k \in \mathbb{N}$, are pairwise distinct.

We consider the following surgical procedure on partitions, leading to an essentially sign-reversing involution on the distinct-part ones: Given a partition $\lambda = [\lambda_1, \ldots, \lambda_k]$ of $n \in \mathbb{N}$ having $k \in \mathbb{N}$ parts, then λ_k is the number of boxes in the last row of the Young diagram of λ, and $d = d(\lambda) := \max\{i \in \{1, \ldots, k\}; \lambda_i = \lambda_1 - i + 1\} \in \mathbb{N}$ is the length of the maximal anti-diagonal at its upper right hand corner. Now let $\lambda \in D(n)$, and we distinguish two cases:

Firstly, if $d < \lambda_k$ we obtain $\mu = [\lambda_1 - 1, \ldots, \lambda_d - 1, \lambda_{d+1}, \ldots, \lambda_k, d]$, whose Young diagram is obtained from that of λ by removing the maximal anti-diagonal and gluing a row of the same length underneath. Thus, if $d < k$ then by definition of d we have $\lambda_d - 1 > \lambda_{d+1}$, and if $d = k$ then we have $\lambda_d - 1 = \lambda_k - 1 \geq d$. Hence μ is a partition of n, having $k+1$ parts such that $d(\mu) \geq d = \mu_{k+1}$. Moreover, μ has pairwise distinct parts, except we have $d = k$ and $\lambda_k = k+1$, which happens if and only if $\lambda = \lambda^{k,+} := [2k, 2k-1, \ldots, k+1] \vdash \frac{1}{2}k(3k+1) = n_k^+$. For example, for $\lambda := [8, 7, 5, 3] \vdash 23$ we get $\mu := [7, 6, 5, 3, 2] \vdash 23$, but for $\lambda := [8, 7, 6, 5] \vdash 26$ we get $\mu := [7, 6, 5, 4, 4] \vdash 26$.

Secondly, if $c := \lambda_k \leq d$ we aim at reversing this procedure, by removing the last row of the Young diagram of λ and gluing an anti-diagonal of the same length to its upper right hand corner. This works well if $c < k$, yielding $\mu = [\lambda_1 + 1, \ldots, \lambda_c + 1, \lambda_{c+1}, \ldots, \lambda_d, \ldots, \lambda_{k-1}]$, which is a partition of n, having $k - 1$ pairwise distinct parts such that $d(\mu) = c = \lambda_k < \lambda_{k-1} = \mu_{k-1}$. But if $c = k$ the gluing step fails, yielding $\mu = [\lambda_1 + 1, \ldots, \lambda_{c-1} + 1, 1]$, which is a partition of n having k distinct parts. From $c = \lambda_k \leq d \leq k$ we infer that this occurs if and only if $\lambda = \lambda^{k,-} := [2k - 1, 2k - 2, \ldots, k] \vdash \frac{1}{2}k(3k - 1) = n_k^-$. For example, for $\lambda := [7, 6, 5, 3, 2] \vdash 23$ we get $\mu := [8, 7, 5, 3] \vdash 23$, but for $\lambda := [7, 6, 5, 4] \vdash 22$ we get $\mu := [8, 7, 6, 1] \vdash 22$.

Thus letting $D_{\text{exc}} := \{\lambda^{k,\pm} \in D_k(n_k^{\pm}); k \in \mathbb{N}\}$ this defines an involutory map on $D(n) \setminus D_{\text{exc}}$, where $n \in \mathbb{N}$, which changes the length of partitions by ±1. Hence if $n \notin \{n_k^{\pm} \in \mathbb{N}; k \in \mathbb{N}\}$ we get $d_n = 0$, while for $n = n_k^{\pm}$ we get $d_n = (-1)^k$. Hence we have proved the **Euler identity** $\prod_{i\geq 1}(1 - X^i) = 1 + \sum_{k\geq 1}(-1)^k(X^{\frac{1}{2}k(3k-1)} + X^{\frac{1}{2}k(3k+1)}) \in \mathbb{Q}[[X]]$.

Recalling that $\prod_{i\geq 1}(1 - X^i) = p^{-1} \in \mathbb{Q}[[X]]$, we get $(\sum_{n\geq 0} p_n X^n) \cdot (1 +$

Table 12: Distinct-part partitions.

n	$D_{\text{even}}(n)$	$D_{\text{odd}}(n)$	d_n
0	$[]$		1
1		$[1]$	-1
2		$[2]$	-1
3	$[2,1]$	$[3]$	.
4	$[3,1]$	$[4]$	.
5	$[4,1],[3,2]$	$[5]$	1
6	$[5,1],[4,2]$	$[6],[3,2,1]$	.
7	$[6,1],[5,2],[4,3]$	$[7],[4,2,1]$	1
8	$[7,1],[6,2],[5,3]$	$[8],[5,2,1],[4,3,1]$	.

$\sum_{k\geq 1}(-1)^k(X^{\frac{1}{2}k(3k-1)} + X^{\frac{1}{2}k(3k+1)})) = 1 \in \mathbb{Q}[[X]]$, and thus convolutional multiplication for $n \in \mathbb{N}$ says $p_n = \sum_{k\geq 1}(-1)^{k-1}(p_{n-n_k^+} + p_{n-n_k^-})$, where we agree on letting $p_i := 0$ for $i < 0$. Actually, this is the most efficient tool known to determine all partition numbers $p_1, \ldots, p_n$ at the same time.

V Group actions

15 Actions

(15.1) Actions. a) Let G be a group, and let N be a set. Then G is said to **act** on the **G-set** N, if there is an **action map** $G \times N \to N\colon [g,x] \mapsto gx$ such that $1x = x$, and $(gh)x = g(hx)$, for all $g, h \in G$ and $x \in N$. If M and N are G-sets, then a map $\alpha\colon M \to N$ such that $\alpha(gx) = g \cdot \alpha(x)$, for all $g \in G$ and $x \in M$, is called a **homomorphism of G-sets**; a bijective homomorphism is called an **isomorphism**, and isomorphic G-sets are called **equivalent**.

Given an action of G on N, for $g \in G$ let $\varphi_g\colon N \to N\colon x \mapsto gx$. Hence from $\varphi_g\varphi_{g^{-1}} = \varphi_{g^{-1}}\varphi_g = \varphi_1 = \mathrm{id}_N$ we get $\varphi_g \in \mathcal{S}_N := \mathrm{Bij}(N,N)$, for all $g \in G$, and since $\varphi_g\varphi_h = \varphi_{gh}$, for all $g, h \in G$, we have an **action homomorphism** $G \to \mathcal{S}_N\colon g \mapsto \varphi_g$. Conversely, if $\varphi\colon G \to \mathcal{S}_N\colon g \mapsto \varphi_g$ is a group homomorphism, then $G \times N \to N\colon [g,x] \mapsto \varphi_g(x)$ is an action of G on N: We have $\varphi_1 = \mathrm{id}_N \in \mathcal{S}_N$, and $\varphi_g\varphi_h = \varphi_{gh}$ implies $(gh)x = g(hx)$ for all $g, h \in G$ and $x \in N$. Moreover, if $\alpha\colon N \to N$ is an isomorphism of G-sets, then the action $\psi_g \in \mathcal{S}_N$ of $g \in G$ on N is given as $\psi_g = \alpha\varphi_g\alpha^{-1}$.

Any group G acts **trivially** on any set N by letting $\varphi_g = \mathrm{id}_N$ for all $g \in G$. If G acts **faithfully** on the set N, that is the action homomorphism $\varphi\colon G \to \mathcal{S}_N$ is injective, then G is called a **permutation group** on N, and can be identified with a subgroup of $\mathcal{S}_N$. In particular, $\mathcal{S}_N$ acts **naturally** on N by $\varphi_\pi\colon N \to$

$N\colon i \mapsto \pi(i)$, for all $\pi \in \mathcal{S}_N$, the action homomorphism being $\mathrm{id}_{\mathcal{S}_N}$.

If N is a finite G-set, then for any $g \in G$ we get a permutation $\varphi_g \in \mathcal{S}_N$, and thus a cycle type $\lambda(g) = \lambda_\varphi(g) := \lambda(\varphi_g)$, a partition of $|N|$. If $\psi\colon G \to \mathcal{S}_N$ is an equivalent action induced by $\alpha \in \mathcal{S}_N$, that is we have $\psi_g = \alpha\varphi_g\alpha^{-1}$ for all $g \in G$, then we have $\lambda_\psi(g) = \lambda(\psi_g) = \lambda(\varphi_g) = \lambda_\varphi(g)$ for all $g \in G$, that is the cycle type of g only depends on the equivalence class of G-actions considered.

(15.2) Orbits. Let G be a group, and let N be a G-set. The relation $\mathcal{R} := \{[x,y] \in N \times N; y = gx \text{ for some } g \in G\}$ is an equivalence relation on N: From $1x = x$ we infer that $\mathcal{R}$ is reflexive; from $y = gx$ we get $g^{-1}y = x$, thus $\mathcal{R}$ is symmetric; and from $y = gx$ and $z = hy$ we get $z = hgx$, thus $\mathcal{R}$ is transitive.

Given $x \in N$, its equivalence class $Gx := \{gx \in N; g \in G\}$ again is a G-set, called the **G-orbit** of x; its cardinality $|Gx|$ is called its **length**, and a subset $T \subseteq G$ such that $T \to Gx\colon t \mapsto tx$ is a bijection is called a **transversal** of Gx with respect to x. Let $G\backslash N := \{Gx \subseteq N; x \in N\}$ be the set of G-orbits; a subset $S \subseteq N$ such that $S \to G\backslash N\colon x \mapsto Gx$ is a bijection is called a set of **orbit representatives** of N, hence we have $N = \coprod_{x\in S} Gx$. If $N \neq \emptyset$ and $N = Gx$ for any and thus all $x \in N$, then N is called a **transitive** G-set. Note that transversals and orbit representatives always exist by the Axiom of Choice.

For $x \in N$ let $G_x = \mathrm{Stab}_G(x) := \{g \in G; gx = x\} \leq G$, being called the **stabilizer** of x in G: We have $1 \in G_x$, and for $g, h \in G_x$ we have $g^{-1}hx = x$, hence $g^{-1}h \in G_x$ as well. Moreover, the stabilizers of elements in the same orbit are **conjugate** in G, more precisely for $g \in G$ we have $G_{gx} = gG_xg^{-1}$: For $h \in G_x$ we have $ghg^{-1} \cdot gx = gx$, hence $gG_xg^{-1} \leq G_{gx}$, thus we also have $g^{-1}G_{gx}g \leq G_{g^{-1}gx} = G_x$, implying $G_{gx} \leq gG_xg^{-1}$.

For example, for the trivial G-set the orbits are the singleton subsets of N, and we have $G_x = G$ for all $x \in N$. Moreover, the natural action of $\mathcal{S}_n$, where $n \in \mathbb{N}$, is transitive such that $\mathrm{Stab}_{\mathcal{S}_n}(n) = \mathcal{S}_{n-1}$; and for any $\pi \in \mathcal{S}_n$ the orbits of $\langle\pi\rangle \leq \mathcal{S}_n$ in the natural action are just the cycles of π.

(15.3) Dihedral groups. Let $\mathbb{R}^{2\times 1}$ be the Euclidean plane equipped with the standard scalar product, and let $\mathrm{GO}_2(\mathbb{R}) := \{g \in \mathrm{GL}_2(\mathbb{R}); gg^{\mathrm{tr}} = E_2\} \leq \mathrm{GL}_2(\mathbb{R})$ be the associated **orthogonal group**. We have $\mathrm{GO}_2(\mathbb{R}) = \{g \in \mathrm{GO}_2(\mathbb{R}); \det(g) = 1\} \,\dot\cup\, \{g \in \mathrm{GO}_2(\mathbb{R}); \det(g) = -1\}$, where the elements of the **special orthogonal group** $\mathrm{SO}_2(\mathbb{R}) := \{g \in \mathrm{GO}_2(\mathbb{R}); \det(g) = 1\} \leq \mathrm{GO}_2(\mathbb{R})$ are called **rotations**, while those of $\mathrm{GO}_2(\mathbb{R}) \setminus \mathrm{SO}_2(\mathbb{R})$ are called **reflections**.

For $n \geq 3$ let $\mathcal{R}_n \subseteq \mathbb{R}^{2\times 1}$ be a regular n-gon centered at the origin, and let $G := \{g \in \mathrm{GO}_2(\mathbb{R}); g(\mathcal{R}_n) = \mathcal{R}_n\}$ be its **group of symmetries**; then $G \cap \mathrm{SO}_2(\mathbb{R})$ is called its **group of rotations**. Hence G acts transitively on the n vertices of $\mathcal{R}_n$, thus numbering them counterclockwise yields an action homomorphism $\varphi\colon G \to \mathcal{S}_n$, which since the vertices contain an $\mathbb{R}$-basis of $\mathbb{R}^{2\times 1}$ is injective; the image $D_{2n} := \mathrm{im}(\varphi) \leq \mathcal{S}_n$ is called the associated **dihedral group**.

We describe the elements of D_{2n}, showing that $|D_{2n}| = 2n$; see Table 13 for

Table 13: Dihedral groups D_6 and D_8.

$\pi \in D_6$	$\lambda(\pi)$	$k(\pi)$
$()$	$[1^3]$	3
$(1,2)$	$[2,1]$	2
$(1,3)$	$[2,1]$	2
$(2,3)$	$[2,1]$	2
$(1,2,3)$	$[3]$	1
$(1,3,2)$	$[3]$	1

$\pi \in D_8$	$\lambda(\pi)$	$k(\pi)$
$()$	$[1^4]$	4
$(1,3)$	$[2,1^2]$	3
$(2,4)$	$[2,1^2]$	3
$(1,2)(3,4)$	$[2^2]$	2
$(1,3)(2,4)$	$[2^2]$	2
$(1,4)(2,3)$	$[2^2]$	2
$(1,2,3,4)$	$[4]$	1
$(1,4,3,2)$	$[4]$	1

$D_6 = \mathcal{S}_3$ and D_8: Since rotations in $\mathrm{GO}_2(\mathbb{R})$ are determined by their rotation angle, the rotations in D_{2n} are those with angle $\frac{2k\pi}{n}$, for $k \in \{0,\ldots,n-1\}$. Thus D_{2n} contains precisely n rotations, given as $\langle \tau_n \rangle = \{\tau_n^k \in \mathcal{S}_n; k \in \{0,\ldots,n-1\}\}$, where $\tau_n := (1,2,\ldots,n) \in \mathcal{S}_n$. Since reflections in $\mathrm{GO}_2(\mathbb{R})$ are determined by their reflection axis, we distinguish the cases n odd and n even:

For n odd the axis runs through one of the vertices of $\mathcal{R}_n$ and the edge opposite; thus there are precisely n reflections, one of them being $\sigma_n := (2,n)(3,n-1)\cdots(\frac{n+1}{2},\frac{n+3}{2}) \in \mathcal{S}_n$. For n even the axis either runs through a pair of opposite vertices, or runs through a pair of opposite edges; thus again there are precisely $\frac{n}{2}+\frac{n}{2} = n$ reflections, one of the former being $\sigma_n := (2,n)(3,n-1)\cdots(\frac{n}{2},\frac{n+4}{2}) \in \mathcal{S}_n$ and one of the latter being $(1,2)(3,n)(4,n-1)\cdots(\frac{n+2}{2},\frac{n+4}{2}) \in \mathcal{S}_n$. In both cases we have $\mathrm{Stab}_{D_{2n}}(1) = \langle \sigma_n \rangle$.

(15.4) Cosets. a) Let G be a group, and $U \leq G$. Then U acts on G by **right multiplication** $\rho_u \colon G \to G \colon x \mapsto xu^{-1}$, for all $u \in U$: We have $\rho_1(x) = x1^{-1} = x$ and $\rho_{uv}(x) = x(uv)^{-1} = xv^{-1}u^{-1} = \rho_u(\rho_v(x))$, for all $x \in G$ and $u,v \in U$. Hence the U-orbit of $x \in G$ is the **(left) coset** $xU := \{xu \in G; u \in U\} \subseteq G$.

The group G acts transitively on $G/U := \{xU \subseteq G; x \in G\}$ by **left multiplication** $\lambda_g \colon G/U \to G/U \colon xU \mapsto gxU$, for all $g \in G$: We have $1(xU) = xU$ and $gh(xU) = g(hxU)$, for all $g,h,x \in G$, and $xU = x(1 \cdot U)$. Since $gu \in U$ if and only if $g \in U$, for $g \in G$, we have $G_U = U$. A transversal $T \subseteq G$ of the action on G/U is called a **(left) transversal** of U in G; then we have $G = \coprod_{t \in T} tU$, and $[G \colon U] := |G/U| = |T| \in \mathbb{N} \,\dot\cup\, \{\infty\}$ is called the **index** of U in G.

b) Let N be a transitive G-set, and $x \in N$. Then $\alpha \colon G/G_x \to N \colon gG_x \mapsto gx$ is an isomorphism of G-sets: For $g \in G$ and $u \in G_x$ we have $gux = gx$, hence α is well-defined. Since N is transitive we infer that α is surjective. For $g,h \in G$ such that $gx = hx$ we have $h^{-1}g \in G_x$, that is $g \in hG_x$, thus α is injective. Finally, we have $\alpha(ghG_x) = ghx = g(hx) = g \cdot \alpha(hG_x)$, for all $g,h \in G$.

c) Let now G be finite, and let $U \leq G$. Then we have $[G \colon U] = \frac{|G|}{|U|}$: Let $T \subseteq G$

be a transversal of U in G, hence $G = \coprod_{t \in T} tU$. For the U-orbit tU we have $U_t = \{u \in U; tu^{-1} = t\} = \{1\}$. Hence $U \to U/\{1\} \to tU \colon u \mapsto u\{1\} \mapsto tu^{-1}$ is a bijection, thus $|tU| = |U|$ and $|G| = |T| \cdot |U|$.

In particular, we have **Lagrange's Theorem** saying that $|U| \mid |G|$. Moreover, for $U := G_x \leq G$, where $x \in N$, this implies the **orbit theorem** saying that $|N| = [G \colon G_x] = \frac{|G|}{|G_x|}$; in particular we have $|N| \mid |G|$.

(15.5) The Cauchy-Frobenius-Burnside Lemma. a) Let G be a finite group, let N be a finite G-set, and for $g \in G$ let $\mathrm{Fix}_N(g) := \{x \in N; gx = x\}$ be its set of **fixed points**. Then we have the **Cauchy-Frobenius-Burnside Lemma** $|G \backslash N| = \frac{1}{|G|} \cdot \sum_{g \in G} |\mathrm{Fix}_N(g)|$; note that $|G \backslash N|$ and $|\mathrm{Fix}_N(g)|$ only depend on the equivalence class of G-actions considered:

Letting $\mathcal{A} := \{[g, x] \in G \times N; gx = x\}$ we use **double counting** to determine $|\mathcal{A}|$ in two ways: On the one hand we have $|\mathcal{A}| = \sum_{g \in G} |\{x \in N; gx = x\}| = \sum_{g \in G} |\mathrm{Fix}_N(g)|$. On the other hand we have $|\mathcal{A}| = \sum_{x \in N} |\{g \in G; gx = x\}| = \sum_{x \in N} |G_x|$. For $y \in Gx$, where $x \in N$, we have $|Gy| = |Gx|$, and thus $|G_x| = |G_y|$. Letting $S \subseteq N$ be a set of orbit representatives, we get $\sum_{x \in N} |G_x| = \sum_{x \in S} \sum_{y \in Gx} |G_y| = \sum_{x \in S} |Gx| \cdot |G_x| = \sum_{x \in S} |G| = |G \backslash N| \cdot |G|$. ♯

Note that enumerating $G = \{g_1, g_2, \ldots\}$ and $N = \{x_1, x_2, \ldots\}$, the set $\mathcal{A}$ can be represented by the indicator matrix $A := [a_{ij}]_{ij} \in \{0, 1\}^{|G| \times |N|}$, where $a_{ij} := 1$ if $g_i x_j = x_j$, and $a_{ij} := 0$ if $g_i x_j \neq x_j$. Hence $|\mathcal{A}|$ is the sum of the entries of A, whose i-th row sum equals $|\mathrm{Fix}_N(g_i)|$ and whose j-th column sum equals $|G_{x_j}|$. The two ways of counting are then computing the sum of the rows sums, and computing the sum of the column sums, respectively.

b) The group G acts **diagonally** on $N \times N$ by $g \colon [y, z] \mapsto [gy, gz]$, for all $g \in G$. Note that the **diagonal** $\{[y, y] \in N \times N; y \in N\}$ is a union of orbits, and consists of a single orbit if and only if G acts transitively.

If G acts transitively on N, then for all $x \in N$ the map $G_x \backslash N \to G \backslash (N \times N) \colon G_x y \mapsto G[x, y]$ is a bijection, and we have $|G_x \backslash N| = |G \backslash (N \times N)| = \frac{1}{|G|} \cdot \sum_{g \in G} |\mathrm{Fix}_N(g)|^2$, being called the **rank** of N:

Let $z \in G_x y$, then there is $g \in G_x$ such that $z = gy$, and hence we have $[x, z] = [x, gy] = [gx, gy] = g[x, y] \in G[x, y]$, implying that the above map is well-defined. Because of transitivity, for any $z \in N$ there is $g \in G$ such that $gz = x$, and hence we have $G[z, y] = G[x, gy]$, showing that the above map is surjective. Finally, letting $z \in N$ such that $G[x, y] = G[x, z]$, that is there is $g \in G$ such that $[x, z] = g[x, y] = [gx, gy]$, we infer that $g \in G_x$, hence $z \in G_x y$, implying injectivity. Finally we have $\mathrm{Fix}_{N \times N}(g) = \mathrm{Fix}_N(g) \times \mathrm{Fix}_N(g)$. ♯

(15.6) Fixed-point-free elements. a) Let G act transitively on N, where $n := |N| \in \mathbb{N}$ and $r := |G \backslash (N \times N)| \in \mathbb{N}$, and let $\mathcal{D} = \mathcal{D}_0(G, N) := \{g \in G; \mathrm{Fix}_N(g) = \emptyset\} \subseteq G$ be the set fixed-point-free elements. Letting $x \in N$ we have the **Cameron-Cohen inequality [1992]** saying $|\mathcal{D}| \geq (r-1) \cdot |G_x| \geq n-1$:

We have $|G| = \sum_{g\in G} |\mathrm{Fix}_N(g)| = \sum_{g\in G\setminus\mathcal{D}} |\mathrm{Fix}_N(g)|$, as well as $(r-1)\cdot|G_x| = \sum_{g\in G_x} |\mathrm{Fix}_{N\setminus\{x\}}(g)|$, where $G_x\cap\mathcal{D} = \emptyset$ anyway. This yields $|G|-(r-1)\cdot|G_x| = \sum_{g\in G_x}(|\mathrm{Fix}_N(g)| - |\mathrm{Fix}_{N\setminus\{x\}}(g)|) + \sum_{g\in G\setminus(G_x\cup\mathcal{D})} |\mathrm{Fix}_N(g)|$. Since $|\mathrm{Fix}_N(g)| - |\mathrm{Fix}_{N\setminus\{x\}}(g)| = 1$ for all $g \in G_x$, and $|\mathrm{Fix}_N(g)| \geq 1$ for all $g \in G\setminus\mathcal{D}$, we infer $|G| - (r-1)\cdot|G_x| \geq |G\setminus\mathcal{D}| = |G| - |\mathcal{D}|$. Moreover, since any of the $r-1$ orbits of G_x on $N\setminus\{x\}$ has length at most $|G_x|$, we get $(r-1)\cdot|G_x| \geq n-1$. ♯

Thus for $n \geq 2$, that is $r \geq 2$, we have the **Lenstra inequality [1991]** saying $\frac{|\mathcal{D}|}{|G|} \geq \frac{(r-1)\cdot|G_x|}{n\cdot|G_x|} = \frac{(r-1)}{n} \geq \frac{1}{n}$, only depending on n, but being independent of $|G|$; In particular, this also implies **Jordan's Theorem [1872]** saying that for $n \geq 2$ we have $\mathcal{D} \neq \emptyset$.

b) Assuming that G acts faithfully, we determine the cases where the Lenstra inequality actually is an equality; in particular, it will turn out that then n is a prime power, and that for any such n there actually is a group achieving equality:

Now $\frac{|\mathcal{D}|}{|G|} = \frac{1}{n}$ implies $r = 2$, that is G acts transitively on the pairs of distinct elements of N, in other words G acts **2-fold** transitively on N. Moreover, we have equality $|\mathcal{D}| = |G_x| = (r-1)\cdot|G_x|$, from which the estimates in the proof of the Cameron-Cohen inequality yield $|\mathrm{Fix}_N(g)| = 1$ for all $g \in G\setminus(G_x\cup\mathcal{D})$, that is any such g is contained in G_y for some unique $y \in N\setminus\{x\}$. Since any $1 \neq g \in G$ acts non-trivially, we conclude that $G = \{1\} \,\dot\cup\, \mathcal{D} \,\dot\cup\, \coprod_{x\in N}(G_x \setminus \{1\})$. Hence we have $G_{x,y} = G_x \cap G_y = \{1\}$ for all $x \neq y \in N$, that is G acts **sharply** 2-fold transitively on N, and thus $|G_x| = |\mathcal{D}| = n-1$ and $|G| = n\cdot|G_x| = n(n-1)$.

Conversely, for any sharply 2-fold transitive group we have $|G_x| = n-1$ and $|G| = n(n-1)$, as well as a decomposition as above, so that $|\mathcal{D}| = |G| - 1 - n\cdot(|G_x|-1) = n-1 = \frac{|G|}{n}$. The sharply 2-fold transitive permutation groups have been classified by **Zassenhaus [1936]**, using the notion of **near-fields**; we indicate the elementary steps, at least showing that n is a prime power:

Let $g \in \mathcal{D}$. Since g acts fixed point freely by conjugation on $\{G_x; x \in N\}$, we infer that g cannot possibly centralize any element of $\coprod_{x\in N}(G_x\setminus\{1\})$, which entails that $C_G(g) \subseteq \{1\} \,\dot\cup\, \mathcal{D}$. On the other hand, for any $x \in N$ the group G_x acts on $\mathcal{D}$, where from $C_{G_x}(g) = C_G(g)\cap G_x = \{1\}$ and $|G_x| = |\mathcal{D}|$ we infer that G_x acts regularly. Thus $\mathcal{D} \subseteq G$ is a single conjugacy class, and hence we have $|C_G(g)| = \frac{|G|}{|\mathcal{D}|} = n = |\mathcal{D}|+1$. In conclusion, we get $C_G(g) = \{1\} \,\dot\cup\, \mathcal{D} =: K$.

This shows that $K \trianglelefteq G$ is a subgroup, which automatically is normal. (Note that this elegantly avoids the subtle theory of Frobenius groups.) From $|K| = n$ we get $G \cong K \rtimes G_x$, where since K acts regularly on N we may identify K with N via $K \to N\colon g \mapsto gx$. All elements of $\mathcal{D}$ being conjugate, we infer that these have one and the same order, which hence is a prime, p say. Furthermore, K is the centralizer of any element of $\mathcal{D}$, hence K is abelian. Thus K is an elementary abelian p-group, which we may write as $K \cong \mathbb{Z}_p^d$, for some $d \in \mathbb{N}$; in particular $n = p^d$. Using the above identification we may describe G as an affine group $G \cong \mathbb{Z}_p^d \rtimes G_0$, where $G_0 \leq \mathrm{GL}_d(p)$ is a subgroup of order $p^d - 1$.

The classification of sharply 2-fold transitive permutation groups now is given as follows, in particular showing that for any prime power there is such a group; see [5, Table 7.3]: **i)** Generically, $\mathbb{Z}_p^d$ can be identified with $\mathbb{F}_{p^d}^+$, where p is a prime and $d \in \mathbb{N}$, and $G_0 \leq \Gamma L_1(p^d)$ is a group of **semilinear** maps; note that $|\Gamma L_1(p^d)| = d \cdot (p^d - 1)$, where $\mathrm{GL}_1(p^d)$ can be identified with a cyclic subgroup $T_{\mathrm{Singer}} \leq \mathrm{GL}_d(p)$ of order $p^d - 1$ generated by a **Singer cycle**, and $\Gamma L_1(p^d)$ can be identified with the normalizer of T_{Singer} in $\mathrm{GL}_d(p)$. **ii)** Apart from that, G is one of 7 exceptional cases, where $d = 2$, and either $p \in \{5, 7, 11, 23\}$ and $\mathrm{SL}_2(3) \leq G_0$, or $p \in \{11, 19, 29, 59\}$ and $\mathrm{SL}_2(5) \leq G_0$; indeed, in the former cases G_0 has shape $\mathrm{SL}_2(3)$, $\mathrm{SL}_2(3) \cdot 2$, $\mathrm{SL}_2(3) \times 5$, $(\mathrm{SL}_2(3) \cdot 2) \times 5$, while in the latter cases we get $\mathrm{SL}_2(5)$, $\mathrm{SL}_2(5) \times 3$, $\mathrm{SL}_2(5) \times 7$, $\mathrm{SL}_2(5) \times 29$, respectively. $\sharp$

In particular, the above discussion implies that the estimate $\frac{|\mathcal{D}|}{|G|} \geq \frac{1}{n}$ is best possible, at least if the class of all permutation groups is considered. Since none of the sharply 2-transitive groups is simple, actually apart from finitely many exceptions they are all solvable, the question arises whether there are stronger bounds for certain classes of permutation groups, for example for simple groups acting primitively. Similarly, we might wonder whether there always are fixed-point-free elements fulfilling further conditions, for example concerning their order. But all these kinds of questions are under ongoing discussion.

(15.7) The fixed point index. a) Let G be a finite group, acting on $N := \{1, \ldots, n\}$, where $n \in \mathbb{N}_0$. For $k \in \{0, \ldots, n\}$ let $\mathcal{D}_k = \mathcal{D}_k(G, N) := \{g \in G; |\mathrm{Fix}_N(g)| = k\}$, and $f_k = f_k(G, N) := |\mathcal{D}_k(G, N)| \in \mathbb{N}_0$. Then the polynomial $f_G = f_{G,N} := \frac{1}{|G|} \cdot \sum_{k=0}^n f_k X^k \in \mathbb{Q}[X]$ is called the **fixed point index** of the action of G on N; note that f_G only depends on the equivalence class of G-actions considered. Hence the fixed point index has degree n and leading coefficient $\frac{f_n}{|G|} \geq \frac{1}{|G|}$; note that $f_n = 1$ if and only if G acts faithfully.

Evaluating yields $f_G(1) = \frac{1}{|G|} \cdot \sum_{k=0}^n f_k = 1$ and $f_G(0) = \frac{f_0}{|G|}$, where $f_0 = |\mathcal{D}_0|$ is the number of the elements of G acting fixed point freely. Evaluating the formal derivative $\frac{\partial}{\partial X} f_G = \frac{1}{|G|} \cdot \sum_{k=1}^n k f_k X^{k-1} \in \mathbb{Q}[X]$, the Cauchy-Frobenius-Burnside Lemma yields $(\frac{\partial}{\partial X} f_G)(1) = \frac{1}{|G|} \cdot \sum_{k=1}^n k f_k = \frac{1}{|G|} \cdot \sum_{k=0}^n k f_k = |G \backslash N|$. Moreover, if G acts transitively, then for any $x \in N$ we have $\frac{\partial}{\partial X} f_{G,N} = f_{G_x, N \setminus \{x\}} \in \mathbb{Q}[X]$, the fixed point index of the action of G_x on $N \setminus \{x\}$:

For $k \geq 1$ we have $\mathcal{D}_k(G, N) = \bigcup_{y \in N} \mathcal{D}_{k-1}(G_y, N \setminus \{y\})$, where any $g \in \mathcal{D}_k(G, N)$ belongs to precisely k of the n sets $\mathcal{D}_{k-1}(G_y, N \setminus \{y\})$. Stabilizers being G-conjugate we infer that $f_{k-1}(G_y, N \setminus \{y\}) = |\mathcal{D}_{k-1}(G_y, N \setminus \{y\})|$ does not depend on $y \in N$. Thus we get $k \cdot f_k(G, N) = n \cdot f_{k-1}(G_x, N \setminus \{x\})$. Using $|G| = n \cdot |G_x|$ this yields $\frac{\partial}{\partial X} f_{G,N} = \frac{1}{|G|} \cdot \sum_{k=1}^n k f_k(G, N) X^{k-1} = \frac{1}{|G_x|} \cdot \sum_{k=1}^n f_{k-1}(G_x, N \setminus \{x\}) X^{k-1} = \frac{1}{|G_x|} \cdot \sum_{k=0}^{n-1} f_k(G_x, N \setminus \{x\}) X^k = f_{G_x, N \setminus \{x\}}$. $\sharp$

b) For $k \in \{0, \ldots, n\}$, the group G acts diagonally on N^k by $g \colon [x_1, \ldots, x_k] \mapsto [gx_1, \ldots, gx_k]$, for all $g \in G$. Thus the set $\mathcal{S}_k(N) := \mathrm{Inj}(\{1, \ldots, k\}, N) = \{[x_1, \ldots, x_k] \in N^k; x_i \neq x_j \text{ for } i \neq j\} \subseteq N^k$ of k-arrangements of N is a union of orbits, hence also becomes a G-set; note that $\mathcal{S}_0(N) = N^0 = \{[]\}$.

Letting $t_k = t_k(G, N) := |G \backslash \mathcal{S}_k(N)| \in \mathbb{N}_0$, the polynomial $t_G = t_{G,N} := \sum_{k=0}^{n} \frac{t_k}{k!} X^k \in \mathbb{Q}[X]$ is called the **transitivity index** of the action of G on N; note that t_G only depends on the equivalence class of G-actions considered. Note that we always have $t_0 = 1$, while G acts transitively if and only if $t_1 = 1$. If $t_k = 1$, for some $k \geq 1$, then G is said to act k**-fold** transitively; note that in this case we have $t_i = 1$ for all $i \in \{0, \ldots, k\}$.

Then we have **Boston's Theorem [1993]** $t_G(X) = f_G(X+1) \in \mathbb{Q}[X]$, that is $t_G = \tau f_G \in \mathbb{Q}[X]$, in terms of difference calculus: If $g \in \mathcal{D}_j$, for some $j \in \{0, \ldots, n\}$, then g fixes precisely $|\mathrm{Inj}(\{1, \ldots, k\}, \mathrm{Fix}_N(g))| = j_{(k)}$ elements of $\mathcal{S}_k(N)$. Hence the Cauchy-Frobenius-Burnside Lemma yields $t_k = \frac{1}{|G|} \cdot \sum_{j=0}^{n} f_j j_{(k)}$, and thus $t_G(X) = \sum_{k=0}^{n} \frac{t_k}{k!} X^k = \frac{1}{|G|} \cdot \sum_{k=0}^{n} \sum_{j=0}^{n} \frac{f_j j_{(k)}}{k!} X^k = \frac{1}{|G|} \cdot \sum_{j=0}^{n} f_j (\sum_{k=0}^{n} \binom{j}{k} X^k) = \frac{1}{|G|} \cdot \sum_{j=0}^{n} f_j (X+1)^j = f_G(X+1) \in \mathbb{Q}[X]$.

In particular, we have $((\frac{\partial}{\partial X})^k f_G)(1) = ((\frac{\partial}{\partial X})^k t_G)(0)$, where in turn $(\frac{\partial}{\partial X})^k t_G = (\frac{\partial}{\partial X})^k (\sum_{j=0}^{n} \frac{t_j}{j!} X^j) = \sum_{j=k}^{n} \frac{t_j}{(j-k)!} X^{j-k} \in \mathbb{Q}[X]$ yields $((\frac{\partial}{\partial X})^k t_G)(0) = t_k$, showing that t_G can be easily computed from f_G.

c) For example, for the trivial group we get $f_{\{1\}} = X^n \in \mathbb{Q}[X]$. Hence, if G acts regularly, then we have $G_x = \{1\}$, and formal integration yields $f_G = \frac{1}{n} \cdot (X^n + (n-1)) \in \mathbb{Q}[X]$, reflecting the fact that 1_G is the only element G having a fixed point. By the way, this also shows that the proportion of fixed-point-free elements of G can indeed be as large as possible, namely $\frac{n-1}{n}$.

Iterating this, we obtain the fixed point indices of **sharply** k-transitive groups, for $k \in \mathbb{N}$, that is G acts k-fold transitively such that $G_{n,n-1,\ldots,n-k+1} = \{1\}$; note that this is independent of the shape of the group considered. For example, for $n = 12$ and $k = 5$ we successively get, writing f_i for the fixed point index of the stabilizer of an i-tuple, for $i \in \{0, \ldots, 5\}$:

$$\begin{array}{rcl}
f_5 & = & X^7, \\
f_4 & = & \frac{1}{8} \cdot (X^8 + 7), \\
f_3 & = & \frac{1}{72} \cdot (X^9 + 63X + 8), \\
f_2 & = & \frac{1}{720} \cdot (X^{10} + 315X^2 + 80X + 324), \\
f_1 & = & \frac{1}{7920} \cdot (X^{11} + 1155X^3 + 440X^2 + 3564X + 2760), \\
f_0 & = & \frac{1}{95040} \cdot (X^{12} + 3465X^4 + 1760X^3 + 21384X^2 + 33120X + 35310)
\end{array}$$

Prominent examples of the above situation are the sporadic simple **Mathieu groups** M_{11} and M_{12}, acting sharply 4-transitive on 11 and sharply 5-transitive on 12 points, respectively; see also (17.3). Hence from $f_{M_{11}} = f_1$ and $f_{M_{12}} = f_0$ we get the transitivity indices $t_{M_{11}} = f_{M_{11}}(X+1) \in \mathbb{Q}[X]$ and $t_{M_{12}} = f_{M_{12}}(X+1) \in \mathbb{Q}[X]$, and thus for M_{11} we get $[t_0, \ldots, t_{11}] = [1, 1, 1, 1, 1, 7, 42, 210, 840, 2520, 5040, 5040]$ and for M_{12} we get $[t_0, \ldots, t_{12}] = [1, 1, 1, 1, 1, 1, 7, 42, 210, 840, 2520, 5040, 5040]$, reflecting the facts that M_{11} acts

4-fold, but not 5-fold transitively, and that M_{12} acts 5-fold, but not 6-fold transitively, respectively.

Formal integration might serve as a tool to proof the non-existence of transitive extensions: For example, for the natural action of the dihedral group D_8 on 4 points, by Table 13 we have $f_{D_8} = \frac{1}{8} \cdot (X^8 + 2X^2 + 5) \in \mathbb{Q}[X]$, from which formal integration yields $f = \frac{1}{40} \cdot (X^5 + \frac{10}{3}X^3 + 25X + \frac{32}{3}) \in \mathbb{Q}[X]$, showing that there cannot be a group of order $5 \cdot 8 = 40$ acting transitively on 5 points, having point stabilizer D_8.

(15.8) Example: Derangements. For the natural action of the symmetric group $\mathcal{S}_n$ we get the following: Let $D_{n,k} \in \mathbb{N}_0$ be the number of permutations in $\mathcal{S}_n$ having precisely $k \in \{0, \ldots, n\}$ fixed points. Since $\mathcal{S}_n$ acts n-fold transitively on N, we have $t_j = 1$ for all $j \in \{0, \ldots, n\}$, thus for the associated fixed point index we get $\frac{1}{n!} \cdot \sum_{k=0}^{n} D_{n,k}X^k = f_{\mathcal{S}_n} = t_{\mathcal{S}_n}(X-1) = \sum_{j=0}^{n} \frac{1}{j!}(X-1)^j = \sum_{j=0}^{n} \sum_{k=0}^{j} \frac{1}{j!} \cdot (-1)^{j-k} \binom{j}{k} X^k = \sum_{k=0}^{n} \sum_{j=k}^{n} \frac{(-1)^{j-k}}{k!(j-k)!} X^k = \sum_{k=0}^{n} (\sum_{j=0}^{n-k} \frac{(-1)^j}{j!}) \frac{X^k}{k!} \in \mathbb{Q}[X]$. From this we recover the formula $D_n = D_{n,0} = n! \cdot \sum_{j=0}^{n} \frac{(-1)^j}{j!}$ for derangement numbers, as well as $D_{n,k} = \frac{n!}{k!} \cdot \sum_{j=0}^{n-k} \frac{(-1)^j}{j!} = \frac{n!}{k!} \cdot \frac{D_{n-k}}{(n-k)!} = \binom{n}{k} \cdot D_{n-k}$. Moreover, for $n \to \infty$ the proportion of elements having precisely k fixed points amongst all permutations is given as $\frac{D_{n,k}}{n!} \to \frac{1}{e \cdot k!}$, that is a Poisson distribution with parameter 1.

(15.9) Example: The structure of benzene. The aromatic hydrocarbon compound **benzene**, which was discovered around 1825, and is nowadays known to have the chemical sum formula C_6H_6, consists of 6 carbon atoms and 6 hydrogen atoms. Due to its chemical stability it was conjectured that the carbon atoms form the vertices of a highly symmetric polyhedron, in which the vertices cannot be distinguished from each other; in particular any vertex is adjacent to the same number of vertices, all edges have the same length, and at any vertex the same types of faces meet. There are three geometrically sensible possible configurations: **i)** the regular 6-gon, **ii)** the rectangular prism over an equilateral triangle, and **iii)** the regular octahedron; see Table 14.

We try to distinguish these cases by producing compounds where two of the hydrogen atoms are replaced by other, distinct compounds, yielding molecules having chemical sum formula C_6H_4XY, where $X \neq Y$. We ask ourselves how many distinct compounds can be formed like this, assuming either of the above three configurations, where compounds are the same if they can be transformed into each other by rotations.

Thus we have to consider the group of symmetries induced by the special orthogonal group $\mathrm{SO}_3(\mathbb{R})$. Its action on the vertices $N := \{1, \ldots, 6\}$ yields transitive subgroups $G^{(i)} \leq \mathcal{S}_6$, for $i \in \{1, \ldots, 3\}$. Hence the new compounds are given as the orbits of $G^{(i)}$ on the set $(N \times N) \setminus \{[j,j] \in N \times N; j \in N\}$ of pairs of distinct vertices, where the Cauchy-Frobenius-Burnside Lemma yields

$|G^{(i)}\backslash(N\times N)|=\frac{1}{|G^{(i)}|}\cdot\sum_{\pi\in G^{(i)}}|\mathrm{Fix}_N(\pi)|^2$.

To determine $G^{(i)}$ we proceed as follows: We have $[G^{(i)}\colon G_1^{(i)}]=6$, hence we first find $G_1^{(i)}$, and then find further elements of $G^{(i)}$ until we generate a transitive subgroup of $\mathcal{S}_6$, then the group generated is contained in $G^{(i)}$ and contains $G_1^{(i)}$ of index at least 6, hence coincides with $G^{(i)}$. Having found $G^{(i)}$ explicitly, we determine $|\mathrm{Fix}_N(\pi)|$ for all $\pi\in G^{(i)}$:

i) For the regular 6-gon we have the dihedral group $G^{(1)}=D_{12}$ of order 12, see (15.3). Hence $G_1^{(1)}=\langle\sigma_6\rangle=\langle(2,6)(3,5)\rangle$ has order 2, and thus $G^{(1)}=\langle\sigma_6,\tau_6\rangle=\langle(2,6)(3,5),(1,2,3,4,5,6)\rangle$. Moreover, the only non-identity elements having fixed points are the three reflections with axis running through a pair of opposite vertices, the latter being the respective fixed points. Hence we find $|G^{(1)}\backslash(N\times N)|=\frac{1}{12}\cdot(6^2+3\cdot 2^2)=4$, leading to three distinct compounds; in terms of the positions at which X and Y are located, since rotations are distance-preserving, orbit representatives are given by $\{[1,2],[1,3],[1,4]\}$.

ii) For the prism we have $G_1^{(2)}=\{1\}$, hence $|G^{(2)}|=6$, and thus $G^{(2)}=\langle(1,2,3)(4,5,6),(1,4)(2,6)(3,5)\rangle$. Moreover, since $G_j^{(2)}=\{1\}$ for all $j\in N$ we conclude that no non-identity element has a fixed point, implying $|G^{(2)}\backslash(N\times N)|=\frac{6^2}{6}=6$, leading to five distinct compounds; similar to the above argument we find that orbit representatives are given by $\{[1,2],\ldots,[1,6]\}$.

iii) For the regular octahedron we have $G_1^{(3)}=\langle(2,3,4,5)\rangle$, having order 4, hence $|G^{(3)}|=24$, and thus $G^{(3)}:=\langle(2,3,4,5),(1,3,6,5)\rangle$. Moreover, we have $G_1^{(3)}=G_6^{(3)}$ and $G_2^{(3)}=G_4^{(3)}$ and $G_3^{(3)}=G_5^{(3)}$, where any of the non-identity elements of the various stabilizers has precisely two fixed points, showing that $|G^{(2)}\backslash(N\times N)|=\frac{1}{24}\cdot(6^2+3\cdot 3\cdot 2^2)=3$, leading to two distinct compounds; similarly we find that orbit representatives are given by $\{[1,2],[1,6]\}$.

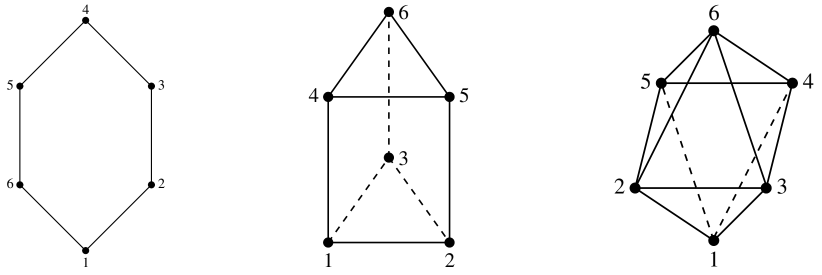
In practice, it is found that there are three distinct compounds arising this way, thus giving a strong indication that case (i) is correct. Nowadays it is commonly believed that case (i), the KEKULÉ structure [1865], describes benzene correctly. Case (ii) was suggested by LADENBURG as the structure of benzene, and has been synthesized as **Prisman** as late as in 1973. Case (iii) does not make sense chemically; nowadays there are various other compounds with chemical sum formula C_6H_6 known which possess less symmetries.

16 Action on maps

(16.1) Maps. a) Let $N:=\{1,\ldots,n\}$ and $K:=\{1,\ldots,k\}$, for some $n\in\mathbb{N}_0$ and $k\in\mathbb{N}_0$. The symmetric group $\mathcal{S}_k$ acts on the set $\mathrm{Maps}(N,K)$ by **post-multiplication**, that is $f\mapsto\sigma(f)\colon i\mapsto\sigma(f(i))$, for all $\sigma\in\mathcal{S}_k$: For $\tau\in\mathcal{S}_k$ we have $(\sigma\tau)(f)=\sigma(\tau(f))\colon i\mapsto(\sigma\tau)(f(i))=\sigma(\tau(f(i)))$. Writing $f=[f(1),\ldots,f(n)]$, then $\sigma\in\mathcal{S}_k$ acts by **renaming** the entries.

The symmetric group $\mathcal{S}_n$ acts on $\mathrm{Maps}(N,K)$ by **pre-multiplication**, that

Table 14: Possible structures of benzene.



is $f \mapsto {}^{\pi}f\colon i \mapsto f(\pi^{-1}(i))$, for all $\pi \in \mathcal{S}_n$: For $\mu \in \mathcal{S}_n$ we have $({}^{\pi\mu}f)(i) = f((\pi\mu)^{-1}(i)) = f((\mu^{-1}\pi^{-1})(i)) = f(\mu^{-1}(\pi^{-1}(i))) = ({}^{\mu}f)(\pi^{-1}(i)) = ({}^{\pi}({}^{\mu}f))(i)$, for all $i \in N$, that is ${}^{\pi\mu}f = {}^{\pi}({}^{\mu}f)$. Writing $f = [f(1), \ldots, f(n)]$, then $\pi \in \mathcal{S}_n$ acts by **reordering** the entries.

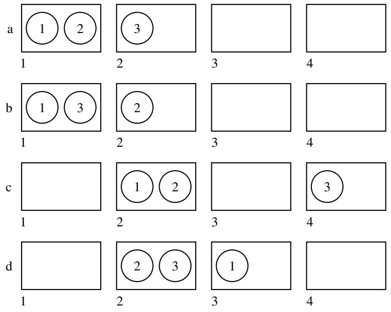
The Cartesian product $\mathcal{S}_n \times \mathcal{S}_k$ becomes a group with componentwise multiplication and inversion, having neutral element $[1, 1]$. Hence $\mathcal{S}_n$ and $\mathcal{S}_k$ can be considered as subgroups of $\mathcal{S}_n \times \mathcal{S}_k$ via the embeddings $\mathcal{S}_n \to \mathcal{S}_n \times \mathcal{S}_k\colon \pi \mapsto [\pi, 1]$ and $\mathcal{S}_k \to \mathcal{S}_n \times \mathcal{S}_k\colon \sigma \mapsto [1, \sigma]$, respectively, and as such we have $\pi\sigma = \sigma\pi$ for all $\pi \in \mathcal{S}_n$ and $\sigma \in \mathcal{S}_k$. Moreover we have ${}^{\pi}(\sigma(f)) = \sigma({}^{\pi}f)\colon i \mapsto \sigma(f(\pi^{-1}(i)))$, for all $f \in \mathrm{Maps}(N, K)$, thus defining an action of $\mathcal{S}_n \times \mathcal{S}_k$ on $\mathrm{Maps}(N, K)$.

Hence we may consider the orbits of either of the groups $\{1\} \times \{1\}$ and $\mathcal{S}_n \times \{1\}$ as well as $\{1\} \times \mathcal{S}_k$ and $\mathcal{S}_n \times \mathcal{S}_k$ on $\mathrm{Maps}(N, K)$. The elements of N are called **indistinguishable** if we let $\mathcal{S}_n$ act, otherwise they are called **distinguishable**, and similarly for K and the action of $\mathcal{S}_k$. Maps being in the same orbit with respect to either of these groups are called **equivalent**; if the elements of both N and K are distinguishable, then all equivalence classes are singleton sets.

b) For example, for $n = 3$ and $k = 4$ let $a, b, c, d \in \mathrm{Maps}(N, K)$ be the pairwise distinct maps $a := [1, 1, 2]$ and $b := [1, 2, 1]$ and $c := [2, 2, 4]$ and $d := [3, 2, 2]$; see Table 15. Equivalence with respect to $\mathcal{S}_n$, amounting to arbitrary reordering, is given by $\{\{a, b\}, \{c\}, \{d\}\}$; equivalence with respect to $\mathcal{S}_k$, amounting to arbitrary renaming, is given by $\{\{a, c\}, \{b\}, \{d\}\}$; with respect to $\mathcal{S}_n \times \mathcal{S}_k$, that is up to arbitrary reordering and renaming, $\{a, b, c, d\}$ are pairwise equivalent.

More intuitively, we may think of N as numbered balls, of K as numbered boxes, and of $f \in \mathrm{Maps}(N, K)$ as a recipe to put the balls into the boxes. This describes the case where both N and K are distinguishable, while if we ignore the numbers on the balls then N becomes indistinguishable, and if we ignore the numbers on the boxes then K becomes indistinguishable.

Table 15: Balls in boxes.



(16.2) The 16-fold way. We continue to consider the action of $\mathcal{S}_n \times \mathcal{S}_k$ on $\operatorname{Maps}(N, K)$, where $N := \{1, \ldots, n\}$ and $K := \{1, \ldots, k\}$, for some $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$. Then a map $f \in \operatorname{Maps}(N, K)$ is injective if and only if $\sigma({}^{\pi}f)$ is, and similarly f is surjective if and only if $\sigma({}^{\pi}f)$ is, for all $\pi \in \mathcal{S}_n$ and $\sigma \in \mathcal{S}_k$. Hence $\operatorname{Inj}(N, K) \subseteq \operatorname{Maps}(N, K)$ and $\operatorname{Surj}(N, K) \subseteq \operatorname{Maps}(N, K)$ as well as $\operatorname{Bij}(N, K) \subseteq \operatorname{Maps}(N, K)$ are also acted on by $\mathcal{S}_n \times \mathcal{S}_k$.

The number of equivalence classes in $\operatorname{Maps}(N, K)$ and the subsets mentioned, with respect to the action of the obvious subgroups of $\mathcal{S}_n \times \mathcal{S}_k$, are given in Table 16: The entries for $\{1\} \times \{1\}$ have already been determined in (2.1) and (3.1). From this we also get the entries in the column for $|\operatorname{Bij}(N, K)|$.

For $\mathcal{S}_n \times \{1\}$ we get: Since the $\mathcal{S}_n$-action amounts to arbitrary reordering of the entries of n-tuples, the equivalence classes of $\operatorname{Maps}(N, K)$ and of $\operatorname{Inj}(N, K)$ are in bijection with the n-multisets $\mathcal{M}_n(k)$ and the n-subsets $\mathcal{P}_n(k)$ of K, respectively; see (2.2) and (2.1). If $k \leq n$, then a map $f \in \operatorname{Maps}(N, K)$ is surjective if and only if it represents an n-multiset $\mu = [\mu_1, \ldots, \mu_k] \in \mathcal{M}_n(k)$ of K such that $\mu_i \geq 1$ for all $i \in \{1, \ldots, k\}$, where the latter are in bijection with $\mathcal{M}_{n-k}(k)$ via $\mu \mapsto [\mu_1 - 1, \ldots, \mu_k - 1]$.

For $\{1\} \times \mathcal{S}_k$ we get: Since the $\mathcal{S}_k$-action amounts to arbitrary renaming of the entries of tuples over K, the equivalence classes of $\operatorname{Maps}(N, K)$ and of $\operatorname{Surj}(N, K)$ are in bijection, via $f \mapsto \coprod_{i=1}^{k} f^{-1}(i)$, with the partitions of N into at most k blocks and into precisely k blocks, respectively; see (3.1). Moreover, if $n \leq k$, then $\operatorname{Inj}(N, K)$ consists of a single equivalence class.

For $\mathcal{S}_n \times \mathcal{S}_k$ we get: Since the $(\mathcal{S}_n \times \mathcal{S}_k)$-action amounts to arbitrary reordering and renaming of tuples in K^n, the equivalence classes of $\operatorname{Maps}(N, K)$ and of $\operatorname{Surj}(N, K)$ are in bijection, via $f \mapsto [|f^{-1}(1)|, \ldots, |f^{-1}(k)|]$, with the sets $P_{\leq k}(n)$ and $P_k(n)$ of partitions of n into at most k parts and into precisely k parts, respectively; see (3.2). Moreover, if $n \leq k$, then $\operatorname{Inj}(N, K)$ still consists of a single equivalence class.

Table 16: The 16-fold way.

	$\lvert\mathrm{Maps}(N,K)\rvert$	$\lvert\mathrm{Inj}(N,K)\rvert$	$\lvert\mathrm{Surj}(N,K)\rvert$	$\lvert\mathrm{Bij}(N,K)\rvert$
$\{1\} \times \{1\}$	k^n	$k_{(n)}$	$k! \cdot S_{n,k}$	$n!$, if $n = k$ 0, if $n \neq k$
$\mathcal{S}_n \times \{1\}$	$\binom{n+k-1}{n}$	$\binom{k}{n}$	$\binom{n-1}{n-k}$, if $n \geq k$ 0, if $n < k$	1, if $n = k$ 0, if $n \neq k$
$\{1\} \times \mathcal{S}_k$	$\sum_{l=0}^{k} S_{n,l}$	1, if $n \leq k$ 0, if $n > k$	$S_{n,k}$	1, if $n = k$ 0, if $n \neq k$
$\mathcal{S}_n \times \mathcal{S}_k$	$p_{n,\leq k}$	1, if $n \leq k$ 0, if $n > k$	$p_{n,k}$	1, if $n = k$ 0, if $n \neq k$

(16.3) Example: Shuffles. We consider shuffles of a deck of $n \in \mathbb{N}_0$ of cards. Usually a deck of cards is viewed as the sequence of cards appearing while running through it, from top to bottom say. Thus we let a **deck** be a bijection $\delta\colon N \to K$ from the set of positions $N := \{1, \ldots, n\}$ to a set K of cardinality n, hence can be identified with the tuple $[\delta(1), \ldots, \delta(n)]$.

a) We consider the **perfect Riffle shuffles**: For $n \in \mathbb{N}$ even, divide the deck into its top and bottom halves of the same size, and then interleave the halves perfectly. Then the top card of either the top or the bottom half ends up at the top of the resulting deck, being called the **out-shuffle** and the **in-shuffle**, respectively. In terms of the positions N this yields permutations $\omega_n \in \mathcal{S}_n$ and $\iota_n \in \mathcal{S}_n$, respectively, writing maps from top to bottom, where $m \in \mathbb{N}$:

$$\omega_{2m} = \begin{bmatrix} 1 & m+1 & 2 & m+2 & \ldots & m-1 & 2m-1 & m & 2m \\ 1 & 2 & 3 & 4 & \ldots & 2m-3 & 2m-2 & 2m-1 & 2m \end{bmatrix}$$

$$\iota_{2m} = \begin{bmatrix} m+1 & 1 & m+2 & 2 & \ldots & 2m-1 & m-1 & 2m & m \\ 1 & 2 & 3 & 4 & \ldots & 2m-3 & 2m-2 & 2m-1 & 2m \end{bmatrix}$$

$$\omega_{2m-1} = \begin{bmatrix} 1 & m+1 & 2 & m+2 & \ldots & 2m-2 & m-1 & 2m-1 & m \\ 1 & 2 & 3 & 4 & \ldots & 2m-4 & 2m-3 & 2m-2 & 2m-1 \end{bmatrix}$$

$$\iota_{2m-1} = \begin{bmatrix} m & 1 & m+1 & 2 & \ldots & m-2 & 2m-2 & m-1 & 2m-1 \\ 1 & 2 & 3 & 4 & \ldots & 2m-4 & 2m-3 & 2m-2 & 2m-1 \end{bmatrix}$$

For example, for $n = 8$ we get:

$$\omega_8 = \begin{bmatrix} 1 & 5 & 2 & 6 & 3 & 7 & 4 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} = (2,3,5)(4,7,6)$$

$$\iota_8 = \begin{bmatrix} 5 & 1 & 6 & 2 & 7 & 3 & 8 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} = (1,2,4,8,7,5)(3,6)$$

Iterating the shuffling corresponds to multiplying the associated permutations:

For example, for $n = 8$ performing an in-shuffle followed by an out-shuffle yields $\omega_8\iota_8 = (2,3,5)(4,7,6) \cdot (1,2,4,8,7,5)(3,6) = (1,3,4,8,6,5)(2,7)$, while the other way around we get $\iota_8\omega_8 = (1,2,4,8,7,5)(3,6) \cdot (2,3,5)(4,7,6) = (1,2,6,8,7,3)(4,5)$. This translates back into decks of cards as follows:

$$\begin{aligned} \omega_8\iota_8 &= (1,3,4,8,6,5)(2,7) = \begin{bmatrix} 5 & 7 & 1 & 3 & 6 & 8 & 2 & 4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} \\ \iota_8\omega_8 &= (1,2,6,8,7,3)(4,5) = \begin{bmatrix} 3 & 1 & 7 & 5 & 4 & 2 & 8 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} \end{aligned}$$

Considering the Riffle shuffle group $\mathcal{R}_n := \langle \omega_n, \iota_n \rangle \leq \mathcal{S}_n$, for $n \geq 2$, we experimentally find the following pattern: $\mathcal{R}_n$ is transitive; $\mathcal{R}_n$ is tiny compared to $\mathcal{S}_n$ whenever n is odd, of order at most $n(n-1)$; $\mathcal{R}_n$ is small compared to $\mathcal{S}_n$ whenever n is even, where for $m \notin \{6,12\} \dot\cup \{2^k; k \geq 0\}$ we get $|\mathcal{R}_{2m}| = 2^{m-1} \cdot \frac{m!}{2}$ whenever $m \equiv 0 \pmod 4$, and $|\mathcal{R}_{2m}| = 2^m \cdot \frac{m!}{2}$ whenever $m \equiv 1 \pmod 4$, and $|\mathcal{R}_{2m}| = 2^m \cdot m!$ whenever $m \equiv 2 \pmod 4$, and $|\mathcal{R}_{2m}| = 2^{m-1} \cdot m!$ whenever $m \equiv 3 \pmod 4$, while for $k \geq 1$ we find $|\mathcal{R}_{2^k}| = 2^k \cdot k$, and $|\mathcal{R}_{12}| = 2^6 \cdot 120$ and $|\mathcal{R}_{24}| = 2^{11} \cdot 95040$; note that this indicates a close relationship between $\mathcal{R}_{2m}$ and the Mongean shuffle groups $\mathcal{M}_m$ discussed now:

b) The **Mongean shuffles** are given as follows: Start with the topmost card, and then put every other card on the top and on the bottom, respectively on the bottom and on the top. Then if $n \in \mathbb{N}$ is even the last card ends up at the top and the bottom, respectively, and the other way around if $n \in \mathbb{N}$ is odd. In terms of the positions N this yields permutations $\mu_n \in \mathcal{S}_n$ and $\mu'_n \in \mathcal{S}_n$, respectively, where $m \in \mathbb{N}$:

$$\begin{aligned} \mu_{2m} &= \begin{bmatrix} 2m & 2m-2 & \ldots & 4 & 2 & 1 & 3 & \ldots & 2m-1 \\ 1 & 2 & \ldots & m-1 & m & m+1 & m+2 & \ldots & 2m \end{bmatrix} \\ \mu'_{2m} &= \begin{bmatrix} 2m-1 & 2m-3 & \ldots & 3 & 1 & 2 & 4 & \ldots & 2m \\ 1 & 2 & \ldots & m-1 & m & m+1 & m+2 & \ldots & 2m \end{bmatrix} \\ \mu_{2m-1} &= \begin{bmatrix} 2m-2 & 2m-4 & \ldots & 2 & 1 & 3 & \ldots & 2m-1 \\ 1 & 2 & \ldots & m-1 & m & m+1 & \ldots & 2m-1 \end{bmatrix} \\ \mu'_{2m-1} &= \begin{bmatrix} 2m-1 & 2m-3 & \ldots & 3 & 1 & 2 & \ldots & 2m-2 \\ 1 & 2 & \ldots & m-1 & m & m+1 & \ldots & 2m-1 \end{bmatrix} \end{aligned}$$

For example, for $n = 8$ we get:

$$\begin{aligned} \mu_8 &= \begin{bmatrix} 8 & 6 & 4 & 2 & 1 & 3 & 5 & 7 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} = (1,5,7,8)(2,4,3,6) \\ \mu'_8 &= \begin{bmatrix} 7 & 5 & 3 & 1 & 2 & 4 & 6 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} = (1,4,6,7)(2,5) \end{aligned}$$

Considering the Mongean shuffle group $\mathcal{M}_n := \langle \mu_n, \mu'_n \rangle \leq \mathcal{S}_n$, for $n \geq 2$, we experimentally find the following pattern: If $n \notin \{6,12\} \dot\cup \{2^k; k \geq 3\}$, then

$\mathcal{M}_n = \mathcal{S}_n$ whenever $n \equiv \{2,3\} \pmod 4$, and $\mathcal{M}_n = \mathcal{A}_n$ whenever $n \equiv \{0,1\}$ (mod 4), while for $k \geq 3$ we find $|\mathcal{M}_{2^k}| = 2^k \cdot (k+1)$, and we get $|\mathcal{M}_6| = 120 = \frac{6!}{6}$ and $|\mathcal{M}_{12}| = 95040 = \frac{12!}{5040}$.

17 Polya's Theorem

If a counting problem is translated into a question concerning equivalence classes of maps, where symmetries of the objects considered have to be taken into account, the groups acting by pre- oder post-multiplication might be smaller than the full symmetric groups. We are going to consider the action of a group $G \cong G \times \{1\} \leq \mathcal{S}_n \times \mathcal{S}_k$ on the full set $\mathrm{Maps}(N, K)$.

(17.1) The cycle index. Let $n \in \mathbb{N}_0$, and let $X_1, \ldots, X_n$ be indeterminates. For $\pi \in \mathcal{S}_n$, letting $\lambda(\pi) = [n^{a_n(\pi)}, \ldots, 1^{a_1(\pi)}] \in P_{k(\pi)}(n)$ be its cycle type, let $c_\pi := \prod_{i=1}^n X_i^{a_i(\pi)} \in \mathbb{Z}[X_1, \ldots, X_n]$ be its **cycle monomial**. Hence c_π is monic of degree $\sum_{i=1}^n a_i(\pi) = k(\pi)$, thus $c_\pi(X, \ldots, X) = X^{k(\pi)} \in \mathbb{Z}[X]$; but if X_i is given degree i, for all $i \in \{1, \ldots, n\}$, then c_π has weighted degree n.

Let G be a finite group acting faithfully on the set $N := \{1, \ldots, n\}$ via $\varphi\colon G \to \mathcal{S}_n$, where $n \in \mathbb{N}_0$. Then letting $c_g := c_{\varphi(g)} \in \mathbb{Z}[X_1, \ldots, X_n]$ be the associated cycle monomials, the polynomial $c_G := \frac{1}{|G|} \cdot \sum_{g \in G} c_g \in \mathbb{Q}[X_1, \ldots, X_n]$ is called the **cycle index** of G; note that c_G only depends on the equivalence class of G-actions considered, and that $c_G(1, \ldots, 1) = 1$. The cycle index has degree n, the only monomial of degree n appearing is X_1^n, with coefficient $\frac{1}{|G|}$, and evaluating yields $c_G(1, \ldots, 1) = 1$; but for the weighted degree we observe that c_G is homogeneous of degree n. For example, for $\{1\} \leq \mathcal{S}_n$ we get $c_{\{1\}} = X_1^n \in \mathbb{Q}[X_1, \ldots, X_n]$.

Hence using the $\mathbb{Q}$-algebra homomorphism $\mathbb{Q}[X_1, \ldots, X_n] \to \mathbb{Q}[X]$ given by $X_1 \mapsto X$, and $X_i \mapsto 1$ for $i \in \{2, \ldots, n\}$, we recover the fixed point index $f_G = c_G(X, 1, \ldots, 1) \in \mathbb{Q}[X]$. More interestingly, using the $\mathbb{Q}$-algebra homomorphism $\mathbb{Q}[X_1, \ldots, X_n] \to \mathbb{Q}[X]\colon X_i \mapsto X$, for all $i \in \{1, \ldots, n\}$, the **cycle-number index** is given as $k_G = k_{G,N} := c_G(X, \ldots, X) = \frac{1}{|G|} \cdot \sum_{g \in G} X^{k(g)} \in \mathbb{Q}[X]$, where $k(g) := k(\varphi(g)) \in \mathbb{N}_0$ is the number of cycles of g.

We show that the cycle index indeed depends on the G-action considered, not just only on the isomorphism class of G: By way of example, let $G := \langle (1,2)(3,4), (1,4)(2,3) \rangle = \{(), (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$ as well as $H := \langle (1,3), (2,4) \rangle = \{(), (1,3), (2,4), (1,3)(2,4)\}$; actually we have $G, H \leq D_8 \leq \mathcal{S}_4$, see Table 13. Then $G \cong H$ as groups via $(1,2)(3,4) \mapsto (1,3)$ and $(1,4)(2,3) \mapsto (2,4)$, hence $(1,3)(2,4) \mapsto (1,3)(2,4)$, but we have $c_G = \frac{1}{4}(X_1^4 + 3X_2^2) \neq \frac{1}{4}(X_1^4 + 2X_2X_1^2 + X_2^2) = c_H \in \mathbb{Q}[X_1, \ldots, X_4]$.

Moreover, groups having one and the same cycle index are not necessarily equivalent as permutation groups, and actually are not necessarily isomorphic: By way of example, let p be an odd prime, let G be the elementary-abelian group of order p^3, and let H be the extra-special group of order p^3 of exponent p, then with respect to their regular representations both groups have cycle index $c_G = c_H = \frac{1}{p^3}(X_1^{p^3} + (p^3-1)X_p^3) \in \mathbb{Q}[X_1, \ldots, X_{p^3}]$.

(17.2) Example: Cycle index of $\mathcal{S}_n$. For $n \in \mathbb{N}_0$ and a partition $\lambda = [n^{a_n}, \ldots, 1^{a_1}] \vdash n$ there are $\frac{n!}{\prod_{i=1}^n (a_i! \cdot i^{a_i})}$ permutations in $\mathcal{S}_n$ with cycle type λ, implying $c_n := c_{\mathcal{S}_n} = \sum_{\lambda=[n^{a_n},\ldots,1^{a_1}]\vdash n} (\prod_{i=1}^n \frac{1}{a_i! \cdot i^{a_i}} X_i^{a_i}) \in \mathbb{Q}[X_1, \ldots, X_n]$. We aim at considering the cycle indices c_n for all $n \in \mathbb{N}_0$ simultaneously:

To this end, let $\mathbb{Q}[\mathcal{X}]$ be the polynomial ring in countably infinitely many indeterminates $\mathcal{X} := \{X_i; i \in \mathbb{N}\}$; hence we have $\mathbb{Q} \subseteq \mathbb{Q}[X_1] \subseteq \mathbb{Q}[X_1, X_2] \subseteq \cdots \subseteq \mathbb{Q}[\mathcal{X}]$. Let $c := \sum_{n \geq 0} c_n X^n \in \mathbb{Q}[\mathcal{X}][[X]]$ be the generating series associated to the sequence $[c_n; n \in \mathbb{N}_0] \in \mathrm{Maps}(\mathbb{N}_0, \mathbb{Q}[\mathcal{X}])$. Hence we have $c = \sum_{n\geq 0}(\sum_{\lambda=[n^{a_n},\ldots,1^{a_1}]\vdash n} \prod_{i=1}^n \frac{1}{a_i! \cdot i^{a_i}} X_i^{a_i}) X^n \in \mathbb{Q}[\mathcal{X}][[X]]$.

Identifying the set $\coprod_{n \geq 0} P(n)$ of all partitions with the set $\mathcal{M} := \{[a_1, a_2, \ldots] \in \mathrm{Maps}(\mathbb{N}, \mathbb{N}_0); a_i = 0$ for almost all $i \in \mathbb{N}\}$, where hence $a = [a_1, a_2, \ldots] \in \mathcal{M}$ represents a partition of $n = \sum_{i \geq 1} i a_i$, see (14.1), the generating series becomes $c = \sum_{a \in \mathcal{M}} (\prod_{i \geq 1} \frac{1}{a_i! \cdot i^{a_i}} X_i^{a_i} X^{i a_i}) = \prod_{i \geq 1}(\sum_{a_i \geq 0} \frac{1}{a_i!}(\frac{1}{i} X_i X^i)^{a_i}) = \prod_{i \geq 1} \exp(\frac{1}{i} X_i X^i) = \exp(\sum_{i \geq 1} \frac{1}{i} X_i X^i) \in \mathbb{Q}[\mathcal{X}][[X]]$. ♯

(17.3) Cycle index of Galois groups. a) Let $f = X^n + \sum_{i=1}^n a_{n-i} X^{n-i} \in \mathbb{Z}[X]$ be an irreducible monic polynomial of degree $n = \deg(f) \geq 2$, with zeroes $\{x_1, \ldots, x_n\} \subseteq \mathbb{C}$, and let $K = \mathbb{Q}[x_1, \ldots, x_n] \subseteq \mathbb{C}$ be its splitting field with Galois group $G := \mathrm{Gal}(K/\mathbb{Q})$; then G acts faithfully and transitively on $\{x_1, \ldots, x_n\}$, hence we may view G as a subgroup of $\mathcal{S}_n$. Let $\mathrm{disc}(f) := \prod_{1 \leq i < j \leq n} (x_i - x_j)^2 \in \mathbb{Z} \setminus \{0\}$ denote the **discriminant** of f. Then, for any prime $p \nmid \mathrm{disc}(f)$ the reduced polynomial $f_p \in \mathbb{F}_p[X]$ is separable of degree n, and factors as $f_p = \prod_{i=1}^l f_{p,i} \in \mathbb{F}_p[X]$, where $l = l(p) \in \mathbb{N}$ and the $f_{p,i} \in \mathbb{F}_p[X]$ are pairwise distinct irreducible monic polynomials, of degree $d_{p,i} \in \mathbb{N}$ say; assuming the $d_{p,i}$ to be suitably ordered, we get $\lambda(p) := [d_{p,1}, \ldots, d_{p,l}] \vdash n$.

For any prime $p \nmid \mathrm{disc}(f)$, choose a discrete valuation ring $R_p \subseteq K$ with (finite) residue field $\mathbb{F}_q$ of characteristic p. Then $\mathbb{F}_q \subseteq \overline{\mathbb{F}}_p$ is the splitting field of f_p, hence we have $[\mathbb{F}_q \colon \mathbb{F}_p] = \log_p(q) = d_p = \mathrm{lcm}(d_{p,1}, \ldots, d_{p,l}) \in \mathbb{N}$, and thus $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p) = \langle \varphi_p \rangle \cong C_{d_p}$, where $\varphi_p \colon \mathbb{F}_q \to \mathbb{F}_q \colon x \mapsto x^p$ is the associated Frobenius automorphism. Then, since $p \nmid \mathrm{disc}(K/\mathbb{Q}) \mid \mathrm{disc}(f)$ implies that $K/\mathbb{Q}$ is unramified at the prime p, we conclude that φ_p lifts to a unique element $\sigma_p \in G$. Taking the choice made into account this determines a unique conjugacy class $\mathcal{C}_p \subseteq G$; viewing $G \leq \mathcal{S}_n$ the latter consists of elements of cycle type $\lambda(p)$. Now, by **Chebotarev's Density Theorem**, for any conjugacy class $\mathcal{C} \subseteq G$ the set of primes such that $\mathcal{C}_p = C$ has **density** $\lim_{x\to\infty} \frac{|\{p \text{ prime}; p \leq x, \mathcal{C}_p = C\}|}{|\{p \text{ prime}; p \leq x\}|} = \frac{|C|}{|G|}$; note that this in particular says that the limit actually exists.

b) Now, the original question behind the Lenstra inequality, see (15.6), was (in the context of the asymptotically fast **Number Field Sieve** algorithm

Table 17: Cycle types for f_1 and f_2.

λ	proportion
$[1^{11}]$	~ 0.0000
$[2^4, 1^3]$	~ 0.0538
$[3^3, 1^2]$	~ 0.1616
$[5^2, 1]$	~ 0.4431
$[6, 3, 2]$	~ 0.1616
$[11]$	~ 0.1796

λ	proportion
$[1^{11}]$	~ 0.0000
$[2^4, 1^3]$	~ 0.006
$[3^3, 1^2]$	~ 0.0363
$[4^2, 1^3]$	~ 0.0969
$[5^2, 1]$	~ 0.1878
$[6, 3, 2]$	~ 0.1878
$[8, 2, 1]$	~ 0.2848
$[11]$	~ 0.2

for factoring integers), whether there are infinitely many primes p such that $f_p \in \mathbb{F}_p[X]$ does not have any zeroes in $\mathbb{F}_p$, and to give a lower bound for their frequency, independent of the particular Galois group of f:

By the above considerations, $p \nmid \operatorname{disc}(f)$ is as desired if and only if $\mathcal{C}_p \subseteq G$ consists of fixed-point-free elements, that is $\mathcal{C}_p \subseteq \mathcal{D}$, where we infer that $\lim_{x\to\infty} \frac{|\{p \text{ prime};\, p\leq x, \mathcal{C}_p\subseteq\mathcal{D}\}|}{|\{p \text{ prime};\, p\leq x\}|} = \frac{|\mathcal{D}|}{|G|} \geq \frac{1}{n} = \frac{1}{\deg(f)}$.

c) This yields a randomized algorithm to determine the cycle index of $G \leq \mathcal{S}_n$ with respect to its transitive action on the (unknown) set $\{x_1, \ldots, x_n\}$, and thus to narrow down the conjugacy classes of transitive subgroups of $\mathcal{S}_n$ to be considered as possible candidates for G. We explain this by an example:

We consider the case $n = 11$ and the (sensibly chosen) polynomials

$$\begin{aligned} f_1 &= X^{11} - 2X^{10} + 3X^9 + 2X^8 - 5X^7 + 16X^6 \\ &\quad -10X^5 + 10X^4 + 2X^3 - 3X^2 + 4X - 1 \in \mathbb{Z}[X], \\ f_2 &= X^{11} - 3X^{10} + 8X^9 - 6X^8 - 9X^7 + 51X^6 \\ &\quad -96X^5 + 108X^4 - 66X^3 - 14X^2 + 42X - 40 \in \mathbb{Z}[X]; \end{aligned}$$

we have $\operatorname{disc}(f_1) = 47^2 \cdot 1831^4 \sim 2 \cdot 10^{15}$ and $\operatorname{disc}(f_2) = 2^{18} \cdot 3^{18} \cdot 97^4 \sim 9 \cdot 10^{21}$. Checking the admissible primes not exceeding 1000, that is 167 and 165 primes, respectively, we find the proportion of cycle types as given in Table 17. In particular, since in both cases the cycle type $[11]$ occurs, we conclude that f_1 and f_2 actually are irreducible.

Now there are 8 conjugacy classes of transitive subgroups of $\mathcal{S}_{11}$, namely: C_{11}, D_{22}, $C_{11}\colon C_5$, $C_{11}\colon C_{10}$, $\mathrm{PSL}_2(11)$, M_{11}, $\mathcal{A}_{11}$, $\mathcal{S}_{11}$. From the cycle types occurring we immediately exclude the first four cases. Moreover, from the discriminants of f_1 and f_2 being squares, we conclude the Galois groups are contained in $\mathcal{A}_{11}$, hence this excludes $\mathcal{S}_{11}$ as well. Next, as far as $\mathcal{A}_{11}$ is concerned, the

Table 18: Cycle types of $L_2(11)$ and M_{11}.

$\lambda(g)$	$\lvert g^G\rvert$	proportion
$[1^{11}]$	1	$\frac{1}{660}$ ~ 0.001
$[2^4, 1^3]$	55	$\frac{1}{12}$ ~ 0.083
$[3^3, 1^2]$	110	$\frac{1}{6}$ ~ 0.166
$[5^2, 1]$	264	$\frac{2}{5}$ ~ 0.4
$[6, 3, 2]$	110	$\frac{1}{6}$ ~ 0.166
$[11]$	120	$\frac{2}{11}$ ~ 0.181

$\lambda(g)$	$\lvert g^G\rvert$	proportion
$[1^{11}]$	1	$\frac{1}{7920}$ ~ 0.0001
$[2^4, 1^3]$	165	$\frac{1}{48}$ ~ 0.0208
$[3^3, 1^2]$	440	$\frac{1}{18}$ ~ 0.0555
$[4^2, 1^3]$	990	$\frac{1}{8}$ ~ 0.125
$[5^2, 1]$	1584	$\frac{1}{5}$ ~ 0.2
$[6, 3, 2]$	1320	$\frac{1}{6}$ ~ 0.1666
$[8, 2, 1]$	$990 + 990$	$\frac{1}{4}$ ~ 0.25
$[11]$	$720 + 720$	$\frac{2}{11}$ ~ 0.1818

proportion of elements containing a 7-cycle is ~ 0.1429, and of those containing a 9-cycle is ~ 0.1111, hence the absence of these elements from the sample taken is a strong indication that the Galois groups looked for are proper subgroups of $\mathcal{A}_{11}$. Hence there are only two groups remaining:

The simple **projective special linear group** $\mathrm{PSL}_2(11)$, acting 2-transitively but not 3-transitively, has order $|\mathrm{PSL}_2(11)| = 660 = 11 \cdot 10 \cdot 6$, and conjugacy classes as given in Table 18, which yields the cycle index $c_{\mathrm{PSL}_2(11)} = \frac{1}{660}(X_1^{11} + 55X_1^3X_2^4 + 110X_1^2X_3^3 + 264X_1X_5^2 + 110X_2X_3X_6 + 120X_{11}$. The sporadic simple **Mathieu group** M_{11}, acting sharply 4-transitively, has order $|M_{11}| = 7920 = 11 \cdot 10 \cdot 9 \cdot 8$, and conjugacy classes as given in Table 18, which yields the cycle index $c_{M_{11}} = \frac{1}{7920}(X_1^{11} + 165X_1^3X_2^4 + 990X_1^3X_4^2 + 440X_1^2X_3^3 + 1980X_1X_2X_8 + 1584X_1X_5^2 + 1320X_2X_3X_6 + 1440X_{11}$; note that thus we may also cross-check the fixed point index $f_{M_{11}}$, which was already determined in (15.7).

For f_2, the existence of elements of cycle type $[8, 2, 1]$ and $[4^2, 1^3]$ excludes the former possibility, so that we have a strong indication that its Galois group is isomorphic to M_{11}. Similarly, for f_1, the absence of elements of the cycle types just mentioned is a strong indication that its Galois group is not isomorphic to M_{11}, hence that it is isomorphic to $\mathrm{PSL}_2(11)$.

(17.4) The weight index. Let $N := \{1, \ldots, n\}$ and $K := \{1, \ldots, k\}$, for some $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$. Let G be a finite group acting faithfully on N, with cycle index $c_G \in \mathbb{C}[X_1, \ldots, X_n]$. The G-orbits $G\backslash\mathrm{Maps}(N, K) := \{{}^G f \subseteq \mathrm{Maps}(N, K); f \in \mathrm{Maps}(N, K)\}$ are called the **patterns** with respect to the G-action, and let $S \subseteq \mathrm{Maps}(N, K)$ be a set of orbit representatives. We aim at determining the number $|G\backslash\mathrm{Maps}(N, K)|$ of patterns:

To this end, let $R := \mathbb{Q}[Y_1, \ldots, Y_k]$, where $Y_1, \ldots, Y_k$ are indeterminates, being called the associated **(generic) weight algebra**. Let the **(generic) weight** of $f \in \mathrm{Maps}(N, K)$ be defined as $w(f) := \prod_{i=1}^n Y_{f(i)} = \prod_{j=1}^k Y_j^{|f^{-1}(j)|} \in R$; hence $w(f)$ is monic of degree n. Then for all $\pi \in \mathcal{S}_n$ we have $w({}^{\pi}f) = \prod_{i=1}^n Y_{({}^{\pi}f)(i)} = \prod_{i=1}^n Y_{f(\pi^{-1}(i))} = \prod_{i=1}^n Y_{f(i)} = w(f)$. Thus weights are constant on $\mathcal{S}_n$-orbits, hence the weight $w({}^{G}f) := w(f) \in R$ of ${}^{G}f \in G\backslash\mathrm{Maps}(N, K)$ is well-defined.

Then $\sum_{f\in S} w(f) \in R$ is called the **(generic) weight index** of $G\backslash\mathrm{Maps}(N, K)$. For $a := [a_1, \ldots, a_k] \in \mathbb{N}_0^k$ we let $\omega_a := |\{f \in S; |f^{-1}(j)| = a_j \text{ for } j \in \{1, \ldots, k\}\}| \in \mathbb{N}_0$ be the number of patterns in which j occurs precisely a_j times for all $j \in \{1, \ldots, k\}$; note that by the above this is independent from the choice of representatives. Hence we have $\sum_{f\in S} w(f) = \sum_{a=[a_1,\ldots,a_k]\in\mathbb{N}_0^k}(\omega_a \cdot \prod_{j=1}^k Y_j^{a_j}) \in R$, that is ω_a is the coefficient of the monomial $\prod_{j=1}^k Y_j^{a_j}$ in the weight index; note that the weight index is homogeneous of degree n.

The main result now is the following polynomial identity, relating the weight index of $G\backslash\mathrm{Maps}(N, K)$ to the cycle index of the G-action on N:

(17.5) Theorem: Polya [1937]. Keeping the notation of (17.4), we have:

$$\sum_{f\in S} w(f) = c_G\left(\sum_{j=1}^k Y_j, \sum_{j=1}^k Y_j^2, \ldots, \sum_{j=1}^k Y_j^n\right) \in \mathbb{Q}[Y_1, \ldots, Y_k] =: R,$$

using the $\mathbb{Q}$-algebra homomorphism $\mathbb{Q}[X_1, \ldots, X_n] \to R\colon X_i \mapsto \sum_{j=1}^k Y_j^i$.

Proof. Given $f \in S$, then G acts transitively on the G-orbit ${}^{G}f \subseteq \mathrm{Maps}(N, K)$, and the Cauchy-Frobenius-Burnside Lemma yields $1 = \frac{1}{|G|} \cdot \sum_{g\in G} |\mathrm{Fix}_{({}^{G}f)}(g)| = \frac{1}{|G|} \cdot \sum_{g\in G} \sum_{e\in\mathrm{Fix}(g)\cap({}^{G}f)} 1 \in \mathbb{Q}$, where we write $\mathrm{Fix}(g) := \mathrm{Fix}_{\mathrm{Maps}(N,K)}(g)$. Since weights are constant on G-orbits, multiplication with $w(f) \in R$ yields $w(f) = \frac{1}{|G|} \cdot \sum_{g\in G} \sum_{e\in\mathrm{Fix}(g)\cap({}^{G}f)} w(e) \in R$, and thus summation over all $f \in S$ yields the weight index $\sum_{f\in S} w(f) = \frac{1}{|G|} \cdot \sum_{g\in G} \sum_{e\in\mathrm{Fix}(g)} w(e) \in R$.

Given $g \in G$, we consider the inner sum $\sum_{e\in\mathrm{Fix}(g)} w(e) \in R$: For any map $e \in \mathrm{Maps}(N, K)$ and any $\pi \in \mathcal{S}_n$ we have ${}^{\pi}e = e$ if and only if $e(\pi^{-1}(i)) = e(i)$ for all $i \in N$, which holds if and only if e is constant on the $k(\pi) \in \mathbb{N}_0$ disjoint cycles of π. Thus we have $e \in \mathrm{Fix}(g)$ if and only if e is constant on the $l := k(g) \in \mathbb{N}_0$ disjoint cycles $N_1, \ldots, N_l \subseteq N$ of g, where the values $e_1, \ldots, e_l \in K$ of e on the various cycles can be chosen arbitrarily. Thus we get $\sum_{e\in\mathrm{Fix}(g)} w(e) = \sum_{e\in\mathrm{Fix}(g)} \prod_{i=1}^n Y_{e_i} = \sum_{e\in\mathrm{Fix}(g)} \prod_{s=1}^l Y_{e_s}^{|N_s|} = \sum_{[e_1,\ldots,e_l]\in K^l} \prod_{s=1}^l Y_{e_s}^{|N_s|} = \prod_{s=1}^l (\sum_{j=1}^k Y_j^{|N_s|})$.

Hence if g has cycle type $\lambda(g) = [n^{a_n(g)}, \ldots, 1^{a_1(g)}] \in P_l(n)$ where $l = k(g) = \sum_{i=1}^n a_i(g)$, using the cycle monomial $c_g = \prod_{i=1}^n X_i^{a_i(g)} \in \mathbb{Z}[X_1, \ldots, X_n]$ we get $\sum_{e\in\mathrm{Fix}(g)} w(e) = \prod_{i=1}^n (\sum_{j=1}^k Y_j^i)^{a_i(g)} = c_g(\sum_{j=1}^k Y_j, \ldots, \sum_{j=1}^k Y_j^n) \in R$. ♯

(17.6) Specialized weights. Specializations of the generic weight algebra give rise to further weight maps, and consequently yield specializations of the above polynomial identity. Keeping the notation of (17.5), the typical specializations are given by replacing some of the indeterminates Y_j by 1, for example:

a) The $\mathbb{Q}$-algebra homomorphism $R \to \mathbb{Q}\colon Y_j \mapsto 1$ yields the number of patterns as $|G\backslash\mathrm{Maps}(N,K)| = \sum_{f\in S} 1 = c_G(k,\ldots,k) = k_G(k) \in \mathbb{Q}$, where $k_G = c_G(X,\ldots,X) = \frac{1}{|G|}\cdot\sum_{g\in G} X^{k(g)} \in \mathbb{Q}[X]$ is the cycle-number index. Note that this shows that, varying the cardinality of K, the number $k_G(k)$ of patterns is given by a polynomial map in $k \in \mathbb{N}_0$. In particular, for $G=\{1\}$ we have $k_{\{1\}} = X^n$, and thus we recover $|\mathrm{Maps}(N,K)| = k_{\{1\}}(k) = k^n$.

More interestingly, let $G = \mathcal{S}_n$. Then we have $k_{\mathcal{S}_n} = c_{\mathcal{S}_n}(X,\ldots,X) = \frac{1}{n!}\cdot\sum_{\pi\in\mathcal{S}_n} X^{k(\pi)} = \frac{1}{n!}\cdot\sum_{l=0}^{n} s_{n,l}X^l \in \mathbb{Q}[X]$. Recalling that $\mathcal{S}_n\backslash\mathrm{Maps}(N,K)$ is in bijection with the n-multisets $\mathcal{M}_n(k)$ on K, we get $\frac{1}{n!}\cdot\sum_{l=0}^{n} s_{n,l}k^l = c_{\mathcal{S}_n}(k,\ldots,k) = |\mathcal{M}_n(k)| = \binom{k+n-1}{n} = \frac{1}{n!}\cdot\prod_{i=0}^{n-1}(k+i)$, for all $k \in \mathbb{N}_0$, recovering the polynomial identity $\sum_{l=0}^{n} s_{n,l}X^l = X^{(n)} \in \mathbb{Q}[X]$.

b) For $k=2$, letting $K=\{0,1\}$, the generic weight algebra becomes $\mathbb{Q}[Y_0,Y_1]$, and the $\mathbb{Q}$-algebra homomorphism $\mathbb{Q}[Y_0,Y_1] \to \mathbb{Q}[Y]$ given by $Y_0 \mapsto 1$ and $Y_1 \mapsto Y$ yields $\sum_{l\in\{0,\ldots,n\}} \omega_l Y^l = c_G(1+Y,\ldots,1+Y^n) \in \mathbb{Q}[Y]$, where $\omega_l = |\{f \in S; |f^{-1}(1)| = l\}| \in \mathbb{N}_0$, for $l \in \{0,\ldots,n\}$, is the number of patterns with precisely l occurrences of $1 \in K$.

This allows for another interpretation: Viewing the elements of $\mathrm{Maps}(N,\{0,1\})$ as indicator functions yields a bijection $\mathrm{Maps}(N,\{0,1\}) \to \mathcal{P}(N)\colon f \mapsto f^{-1}(1)$. Moreover, the group G acts on $\mathcal{P}(N)$ via $g\colon M \mapsto gM := \{gx \in N; x \in M\} = \{x \in N; g^{-1}x \in M\}$, for all $g \in G$ and $M \subseteq N$; note that $|M| = |gM|$. Then we have $({}^g f)^{-1}(1) = \{x \in N; f(g^{-1}(x)) = 1\} = g(f^{-1}(1)) \subseteq N$, for all $g \in G$ and $f \in \mathrm{Maps}(N,\{0,1\})$, showing that the above map $\mathrm{Maps}(N,\{0,1\}) \to \mathcal{P}(N)$ is an isomorphism of G-sets.

Hence $\omega_l = |G\backslash\mathcal{P}_l(N)| \in \mathbb{N}_0$ is the number of G-orbits on the set of l-subsets of N, and thus we get $\sum_{l\in\{0,\ldots,n\}} |G\backslash\mathcal{P}_l(N)|\cdot Y^l = c_G(1+Y,\ldots,1+Y^n) \in \mathbb{Q}[Y]$. In particular, the $\mathbb{Q}$-algebra homomorphism $\mathbb{Q}[Y] \to \mathbb{Q}\colon Y \mapsto 1$ yields $|G\backslash\mathcal{P}(N)| = \sum_{l\in\{0,\ldots,n\}} |G\backslash\mathcal{P}_l(N)| = c_G(2,\ldots,2) = k_G(2) = \frac{1}{|G|}\cdot\sum_{g\in G} 2^{k(g)}$.

Thus for $G=\{1\}$ we recover $\sum_{l\in\{0,\ldots,n\}} \binom{n}{l}\cdot Y^l = \sum_{l\in\{0,\ldots,n\}} |\mathcal{P}_l(N)|\cdot Y^l = c_{\{1\}}(1+Y,\ldots,1+Y^n) = (1+Y)^n \in \mathbb{Q}[Y]$; this specializes further yielding $\sum_{l\in\{0,\ldots,n\}} \binom{n}{l} = |\mathcal{P}(N)| = 2^n$. For $G = \mathcal{S}_n$ we have $|G\backslash\mathcal{P}_l(N)| = 1$ for all $l \in \{0,\ldots,n\}$, implying $\sum_{l\in\{0,\ldots,n\}} Y^l = c_{\mathcal{S}_n}(1+Y,\ldots,1+Y^n) \in \mathbb{Q}[Y]$, from which by specializing further we obtain $n+1 = c_{\mathcal{S}_n}(2,\ldots,2) = k_{\mathcal{S}_n}(2) = \frac{1}{n!}\cdot\sum_{\pi\in\mathcal{S}_n} 2^{k(\pi)} = \frac{1}{n!}\cdot\sum_{l=0}^{n} s_{n,l}2^l$, recovering a special case of the polynomial identity for Stirling numbers of the first kind mentioned above.

18 Counting patterns

(18.1) Example: Necklaces. A **necklace** with $n \geq 3$ pearls having at most $k \in \mathbb{N}_0$ distinct colors is a map $f\colon N := \{1, \ldots, n\} \to \{1, \ldots, k\} =: K$. The set N may be considered as the set of vertices of a regular n-gon $\mathcal{R}_n$, and necklaces are equivalent if they arise from each other by a symmetry of $\mathcal{R}_n$. Hence we consider the dihedral group $D_{2n} \leq \mathcal{S}_n$ acting on the set $\mathrm{Maps}(N, K)$ by pre-multiplication, and the equivalence classes are the associated orbits. Thus their number is $\eta_n(k) := |D_{2n}\backslash\mathrm{Maps}(N, K)| = k_{D_{2n}}(k) = c_{D_{2n}}(k, \ldots, k) = \frac{1}{2n} \cdot \sum_{\pi \in D_{2n}} k^{k(\pi)} \in \mathbb{N}$.

For example, for $n \in \{3, 4\}$ from Table 13 we get the cycle indexes $c_{D_6} = \frac{1}{6}(X_1^3 + 3X_2X_1 + 2X_3) \in \mathbb{Q}[X_1, \ldots, X_3]$ and $c_{D_8} = \frac{1}{8}(X_1^4 + 2X_2X_1^2 + 3X_2^2 + 2X_4) \in \mathbb{Q}[X_1, \ldots, X_4]$. Thus we have the cycle-number indices $k_{D_6} = \frac{1}{6}(X^3 + 3X^2 + 2X) \in \mathbb{Q}[X]$ and $k_{D_8} = \frac{1}{8}(X^4 + 2X^3 + 3X^2 + 2X) \in \mathbb{Q}[X]$, yielding $\eta_3(k) = k_{D_6}(k)$ and $\eta_4(k) = k_{D_8}(k)$.

Moreover, letting $\eta_{n,a} \in \mathbb{N}_0$ be the number of equivalence classes of necklaces with $n \geq 3$ pearls which are either black or white, such that there are precisely $a \in \{0, \ldots, n\}$ black pearls, we get $\sum_{a \in \{0,\ldots,n\}} \eta_{n,a} Y^a = c_{D_{2n}}(1 + Y, 1 + Y^2, \ldots, 1 + Y^n) \in \mathbb{Q}[Y]$, for $n \in \{3, 4\}$ yielding $\sum_{a \in \{0,\ldots,3\}} \eta_{3,a} Y^a = \frac{1}{6}((1 + Y)^3 + 3(1 + Y^2)(1 + Y) + 2(1 + Y^3)) = Y^3 + Y^2 + Y + 1 \in \mathbb{Q}[Y]$ and $\sum_{a \in \{0,\ldots,4\}} \eta_{4,a} Y^a = \frac{1}{8}((1+Y)^4 + 2(1+Y^2)(1+Y)^2 + 3(1+Y^2)^2 + 2(1+Y^4)) = Y^4 + Y^3 + 2Y^2 + Y + 1 \in \mathbb{Q}[Y]$, respectively, where in both cases the numbers $\eta_{n,a}$ are combinatorially obvious.

(18.2) Example: The structure of benzene, revisited. Using the notation of (15.9), finding the number of orbits of $G = G^{(i)}$, for $i \in \{1, 2, 3\}$, on the set of pairs of distinct elements of $N := \{1, \ldots, 6\}$, can be rephrased in terms of a pattern counting problem: Letting $K := \{0, 1, 2\}$, we have to find the number of patterns ${}^{G}f \in G\backslash\mathrm{Maps}(N, K)$ such that $|f^{-1}(1)| = |f^{-1}(2)| = 1$. Thus, specializing the associated generic weight index using the $\mathbb{Q}$-algebra homomorphism $\mathbb{Q}[Y_0, Y_1, Y_2] \to \mathbb{Q}[Y_1, Y_2]\colon Y_0 \mapsto 1, Y_1 \mapsto Y_1, Y_2 \mapsto Y_2$, we have to determine the coefficient $\omega_{[1,1]}$ of the monomial Y_1Y_2 in the specialized weight index. To do so, we compute the cycle index $c(X_1, \ldots, X_6) \in \mathbb{Q}[X_1, \ldots, X_6]$, and by Polya's Theorem specialize $c(1+Y_1+Y_2, , \ldots, , 1+Y_1^6+Y_2^6) \in \mathbb{Q}[Y_1, Y_2]$:

i) We have $G^{(1)} \cong D_{12}$. It is immediate from the description in (15.3) that $c^{(1)}(X_1, \ldots, X_6) = \frac{1}{12} \cdot (2X_6 + 2X_3^2 + 4X_2^3 + 3X_2^2X_1^2 + X_1^6) \in \mathbb{Q}[X_1, \ldots, X_6]$. In order to determine the coefficient $\omega_{[1,1]}$ we proceed as follows: Expanding $c^{(1)}(1 + Y_1 + Y_2, , \ldots, , 1 + Y_1^6 + Y_2^6)$ shows that Y_1Y_2 only occurs in $(1 + Y_1^2 + Y_2^2)^2(1 + Y_1 + Y_2)^2$, with coefficient 2, and in $(1 + Y_1 + Y_2)^6$, with coefficient $\frac{6!}{4!\cdot 1!\cdot 1!} = 30$. Hence we get $\omega_{[1,1]} = \frac{1}{12} \cdot (3 \cdot 2 + 30) = 3$.

ii) We have $G^{(2)} \cong \mathcal{S}_3$. Since $G^{(2)}$ acts transitively on 6 points, this action is equivalent to the regular action of $\mathcal{S}_3$. Hence we get $c^{(2)}(X_1, \ldots, X_6) = \frac{1}{6} \cdot (2X_3^2 + 3X_2^3 + X_1^6) \in \mathbb{Q}[X_1, \ldots, X_6]$, where similarly, Y_1Y_2 only occurs in

$(1+Y_1+Y_2)^6$, with coefficient 30, yielding $\omega_{[1,1]} = \frac{30}{6} = 5$.

iii) We have $G^{(3)} \cong \mathcal{S}_4$. Since $G^{(3)}$ acts transitively and faithfully on 6 points, this action is equivalent to the action of $\mathcal{S}_4$ on the cosets of $\langle (1,2),(3,4)\rangle \cong V_4$. Hence we get $c^{(3)}(X_1,\ldots,X_6) = \frac{1}{24} \cdot (6X_4X_1^2 + 8X_3^2 + 6X_2^3 + 3X_2^2X_1^2 + X_1^6) \in \mathbb{Q}[X_1,\ldots,X_6]$, where similarly Y_1Y_2 only occurs in $(1+Y_1^4+Y_2^4)(1+Y_1+Y_2)^2$, with coefficient 2, in $(1+Y_1^2+Y_2^2)^2(1+Y_1+Y_2)^2$, with coefficient 2, and in $(1+Y_1+Y_2)^6$, with coefficient 30, yielding $\omega_{[1,1]} = \frac{1}{24} \cdot (6 \cdot 2 + 3 \cdot 2 + 30) = 2$. ♯

(18.3) Example: Isomerism of alcohols. An **alcohol** is a aliphatic hydrocarbon compound having the chemical sum formula $C_nH_{2n+1}(OH)$, where $n \in \mathbb{N}$, consisting of a connected skeleton without circles of n carbon atoms, one of which carries a hydroxy group OH, and the $2n+1$ hydrogen atoms are attached to the carbon atoms so that any of the latter is adjacent to precisely 4 neighbors. To determine the number $t_n \in \mathbb{N}$ of distinct **isomers** of alcohols having n carbon atoms (where here we consider **stereo isomers** as identical), we consider the generating series $t := \sum_{n\geq 0} t_nY^n \in \mathbb{Q}[[Y]]$, where we let $t_0 := 1$, corresponding to the degenerate alcohol $H(OH)$, that is water.

The carbon atom skeleton with the distinguished atom carrying the hydroxy group can be considered as a **rooted tree** Γ, that is a connected finite **simple graph** without **circles** with an exceptional vertex; the **empty graph** is a rooted tree, and we let $|\Gamma| \in \mathbb{N}_0$ be the number of vertices of Γ. Note that a tree having $n \geq 1$ vertices has precisely $n-1$ edges, showing that there are indeed precisely $4n-2(n-1)-1 = 2n+1$ free slots for hydrogen atoms. Hence we are interested in the set $\mathcal{T}$ of rooted trees all of whose vertices have **valency** at most 4, and whose exceptional vertex has valency at most 3. Hence we have $\mathcal{T} = \coprod_{n\in\mathbb{N}_0} \mathcal{T}_n$, where $\mathcal{T}_n := \{\Gamma \in \mathcal{T}; |\Gamma| = n\}$; in particular, $\mathcal{T}_0$ is the singleton set containing the empty graph. Thus we have $t_n = |\mathcal{T}_n|$, for $n \in \mathbb{N}_0$.

Given $\Gamma \in \mathcal{T}_n$, for some $n \geq 1$, allowing for the empty graph as a subtree, there are precisely 3 subtrees $\Gamma_1, \Gamma_2, \Gamma_3 \in \mathcal{T}$ attached to the exceptional vertex of Γ, where the vertex of Γ_i being attached to the exceptional vertex of Γ is designated as exceptional. Since the order of the Γ_i is irrelevant, we conclude that Γ is completely described by the 3-multiset $[\Gamma_1, \Gamma_2, \Gamma_3] \subseteq \mathcal{T}$. Thus $\mathcal{T} \setminus \mathcal{T}_0$ can be identified with $\mathcal{S}_3\backslash\text{Maps}(N, \mathcal{T})$, where $N := \{1,2,3\}$, by associating $f \in \text{Maps}(N,\mathcal{T})$ with the rooted tree spliced together as described above from $[f(1), f(2), f(3)]$. Hence t_n coincides with the number of $\mathcal{S}_3$-orbits on the set $\{f \in \text{Maps}(N,\mathcal{T}); |f(1)| + |f(2)| + |f(3)| = n-1\}$, for $n \in \mathbb{N}$.

The generic weight algebra for this setting is as follows: Let $\mathcal{Y} := \{Y_\Gamma; \Gamma \in \mathcal{T}\}$ be indeterminates. Then for any map $f \in \text{Maps}(N,\mathcal{T})$ the associated weight $w(f)$ is a polynomial of degree 3 in $\mathbb{Q}[\mathcal{Y}]$, but we will have to allow for infinite sums like $\sum_{f\in\text{Maps}(N,\mathcal{T})} w(f)$. Hence the generic weight algebra is defined as the **inverse limit** $R := \varprojlim_{n\in\mathbb{N}_0} \{\mathbb{Q}[\mathcal{Y}]/(Y_\Gamma; |\Gamma| \geq n)\}$, with respect to the natural quotient maps; note that the proof of Polya's Theorem still works for weight maps with values in R. We consider the following specialization: Letting

Y be an indeterminate, the ring of formal power series in Y can be seen as the inverse limit $\mathbb{Q}[[Y]] = \varprojlim_{n\in\mathbb{N}_0}\{\mathbb{Q}[Y]/(Y^n)\}$, with respect to the natural quotient maps. Now, for any $n \in \mathbb{N}_0$ there is a $\mathbb{Q}$-algebra homomorphism $\mathbb{Q}[\mathcal{Y}]/(Y_\Gamma; |\Gamma| \geq n) \to \mathbb{Q}[Y]/(Y^n)\colon Y_\Gamma \mapsto Y^{|\Gamma|}$. Since these maps commute with the respective quotient maps used to define the inverse limits considered, this induces a $\mathbb{Q}$-algebra homomorphism $R \to \mathbb{Q}[[Y]]\colon Y_\Gamma \mapsto Y^{|\Gamma|}$.

Hence, with respect to this specialization, any map $f \in \mathrm{Maps}(N, \mathcal{T})$ has associated weight $w(f) := Y^{|f(1)|+|f(2)|+|f(3)|} \in \mathbb{Q}[Y]$, we get $\sum_{f\in\mathcal{S}_3\backslash\mathrm{Maps}(N,\mathcal{T})} w(f) = \sum_{n\geq 0} t_{n+1}Y^n = \frac{1}{Y}(t(Y)-1) \in \mathbb{Q}[[Y]]$. We evaluate the left hand side using Polya's Theorem: The cycle index of $\mathcal{S}_3$ is given as $c_{\mathcal{S}_3} = \frac{1}{6}(X_1^3 + 3X_2X_1 + 2X_3) \in \mathbb{Q}[X_1, X_2, X_3]$, see (18.1), and since the sum $\sum_{\Gamma\in\mathcal{T}} Y_\Gamma \in R$ specializes to $t = \sum_{n\geq 0} t_nY^n \in \mathbb{Q}[[Y]]$, the left hand side equals $c_{\mathcal{S}_3}(t(Y), t(Y^2), t(Y^3)) = \frac{1}{6}(t(Y)^3 + 3t(Y^2)t(Y) + 2t(Y^3)) \in \mathbb{Q}[[Y]]$.

Hence we get the **functional equation** $\frac{1}{Y}(t(Y)-1) = \frac{1}{6}(t(Y)^3 + 3t(Y^2)t(Y) + 2t(Y^3))$, or equivalently $t(Y) = 1 + \frac{Y}{6}(t(Y)^3 + 3t(Y^2)t(Y) + 2t(Y^3))$, allowing to solve recursively for t_n, for $n \in \mathbb{N}$, given $t_0 = 1$:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
t	1	1	1	2	4	8	17	39	89	211	507	1238	3057	7639	19241

Hence in particular there are a unique **methanol** and a unique **ethanol**, as well as two **propanols**, four **butanols**, eight **pentanols**, 17 **hexanols**, and so on.

A closed formula for t_n currently is not known. Considering the power series $\sum_{n\geq 0} t_nz^n$ as a complex function in $z \in \mathbb{C}$, it can be shown that it has (analytically defined) convergence radius $0.35 < \rho \sim 0.355181742 < 0.36$, and that the growth behavior of the t_n is described by $\lim_{n\to\infty}(t_n \cdot \rho^n \cdot n^{\frac{3}{2}}) = \gamma$, for some (analytically defined) constant $\gamma \sim 0.5179$, that is $t_n \sim (\frac{1}{\rho})^n \cdot (\frac{1}{n})^{\frac{3}{2}}$. Actually, approximations to ρ and γ can be found experimentally as follows: Varying ρ within the above bounds, the sequence of the numbers $t_n \cdot \rho^n \cdot n^{\frac{3}{2}}$, for $n \in \{1, \ldots, 700\}$ say, turns out to be increasing if ρ is chosen too large, while from a certain turning point on it becomes decreasing if ρ is chosen too small, hence we may just try to push this turning point to larger and larger n.

b) Similarly, for any $d \geq 2$, we may count the set $\mathcal{T}^{(d)}$ of **d-ary** rooted trees, that is rooted trees all of whose vertices have valency at most d and whose exceptional vertex has valency at most $d-1$, using the action of the symmetric group $\mathcal{S}_d$ on $\mathrm{Maps}(\{1, \ldots, d\}, \mathcal{T}^{(d)})$.

For example, for $d = 2$ we get the cycle index $c_{\mathcal{S}_2} = \frac{1}{2}(X_1^2 + X_2) \in \mathbb{Q}[X_1, X_2]$, and hence the associated generating series $\tau := \sum_{n\geq 0} \tau_nY^n \in \mathbb{Q}[[Y]]$, where $\tau_n := |\{\Gamma \in \mathcal{T}^{(2)}; |\Gamma| = n\}|$ for $n \in \mathbb{N}_0$, fulfills the functional equation $\frac{1}{Y}(\tau(Y)-1) = \frac{1}{2}(\tau(Y)^2 + \tau(Y^2))$. Using $\tau_0 = 1$, the latter recursively yields the **Wedderburn-Etherington numbers** as follows:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
τ_n	1	1	1	2	3	6	11	23	46	98	207	451	983	2179	4850

A closed formula for the τ_n currently is not known. Considering the power series $\sum_{n\geq 0} \tau_n z^n$ as a complex function in $z \in \mathbb{C}$, it can be shown that it has convergence radius $\rho \sim 0.402697504$, and that the growth behavior of the τ_n is described by $\lim_{n\to\infty}(\tau_n \cdot \rho^{n-1} \cdot n^{\frac{3}{2}}) = \gamma \sim 0.7916$.

(18.4) Example: The card game SET. a) The card game SET was invented by Falco [1974] as a translation of observables used in a study of the genetics of epilepsy in dogs, German shepherds to be precise. The translation consists of four attributes, each of which can assume three distinct values, of the symbols on specially tailored playing cards; this amounts to a total of $3^4 = 81$ cards:

attribute	0	1	2
quantity	three	one	two
color	green	red	purple
shading	open	striped	solid
shape	diamond	squiggle	oval

Given a subset of cards, a **'set'** in the sense of this game is a subset of three cards, such that with respect to any of the attributes, independently, the three cards either all have the same value or are pairwise distinct. 'Set'-free subsets of cards are also called **caps**, where those which are not properly contained in a larger cap are called **maximal caps**.

The game then is played as follows: Twelve cards are laid out face-up on the table. The player first detecting a 'set' collects these three cards, and the latter are replaced by three new ones. If there is no 'set' detected, then three more cards are laid out. This is iterated until the deck is empty, and the remaining cards laid out form a cap. The player with the most 'sets' detected wins.

This leads, at least, to the following questions: Firstly, with respect to all subsets of cards: When are two subsets of cards equivalent, with respect to the 'set' property? How many inequivalent k-subsets of cards are there? (In particular, for $k = 12$?) Secondly, with respect to caps: What is the largest $k = k_{\max}$ such that there is a k-cap of cards? (In particular, is $k_{\max} \geq 12$?) For $k \leq k_{\max}$, how many inequivalent k-caps of cards are there? What is the proportion of the k-caps amongst all k-subsets of cards? How do maximal caps look like? (In particular, for $k = k_{\max}$?)

In order to attack this algebraically, the values of the attributes are associated with $\mathbb{F}_3 := \{0, 1, 2\}$, as is already indicated above. Hence the cards can be identified with the vectors in $\mathbb{F}_3^4$. Then the condition that, for a fixed attribute, three values either are all the same or are pairwise distinct translates into saying that their sum vanishes. Hence a 'set' of cards just is a 3-subset $\{v, v', v''\} \subseteq \mathbb{F}_3^4$ such that $v + v' + v'' = 0 \in \mathbb{F}_3^4$. Moreover, the admissible symmetries are the bijections on $\mathbb{F}_3^4$ mapping 'sets' to 'sets'.

The condition of being a 'set' can be rephrased as follows: If $\{v, v', v''\}$ is a 'set', we have $v' = v + (v' - v)$ and $v'' = -(v + v') = v + 2(v' - v)$, thus

$\{v, v', v''\} = v + \mathbb{F}_3 \cdot (v' - v)$, where $0 \neq (v' - v) \in \mathbb{F}_3^4$. Conversely, for any $0 \neq u \in \mathbb{F}_3^4$, the 3-subset $v + \mathbb{F}_3 \cdot u \subseteq \mathbb{F}_3^4$ fulfills $v + (v+u) + (v+2u) = 0$, hence is a 'set'. Thus 'sets' are just the 'lines' in $\mathbb{F}_3^4$ in the following geometrical sense:

b) Letting $d \in \mathbb{N}$ and $\mathbb{F}_q$ be the field with q elements, we consider the **affine space** with underlying **points** $V := \mathbb{F}_q^d$, whose **lines** are the subsets $v + \langle u \rangle_{\mathbb{F}_q} \subseteq V$, for some $v \in V$ and $0 \neq u \in V$. Then the associated symmetries (in the sense of **affine geometry**) are given by the **affine general linear group** $A := \mathrm{AGL}_d(q) \cong T \rtimes \mathrm{GL}_d(q)$, where $\mathrm{GL}_d(q) \leq A$ is the subgroup of affine maps fixing the origin, that is the subgroup of $\mathbb{F}_q$-linear maps, and $T \cong V$ is the normal subgroup of **translations**; hence we have $|A| = q^d \cdot \prod_{i=0}^{d-1}(q^d - q^i)$. Then A acts faithfully on V via $[t, g] \cdot v := t + gv$, for $v \in V$ and $g \in \mathrm{GL}_d(q)$ and $t \in T$. Embedding $A \to \mathrm{GL}_{d+1}(q)$ via $[t, g] \mapsto \left[\begin{array}{c|c} g & t \\ \hline . & 1 \end{array}\right]$, and identifying $v \in V$ with the extended vector $[v^{\mathrm{tr}}|1]^{\mathrm{tr}} \in \mathbb{F}_q^{d+1}$ we get $\left[\begin{array}{c|c} g & t \\ \hline . & 1 \end{array}\right] \cdot \left[\begin{array}{c} v \\ \hline 1 \end{array}\right] = \left[\begin{array}{c} gv + t \\ \hline 1 \end{array}\right]$, implying that affine maps can be considered as $\mathbb{F}_q$-linear maps.

This paves the way to find the cycle index of A; note that using the structure of A as a semidirect product it is possible to use the well-understood conjugacy classes of $\mathrm{GL}_d(q)$ to obtain a conceptual description of the conjugacy classes of A, but instead we pursue a purely computational approach: Considering the action of A on extended vectors, we obtain a faithful permutation representation, which can be used to determine its conjugacy classes, together with their cardinality, and the cycle type of its elements. Given the cycle index $c_A \in \mathbb{Q}[X_1, \ldots, X_{q^d}]$ of A, the number of orbits of A on the k-subsets of V, for $k \in \{0, \ldots, q^d\}$, is readily given by the coefficient of Y^k in $c_A(1+Y, \ldots, 1+Y^{q^d}) \in \mathbb{Q}[Y]$.

However, it is computationally much more difficult to find a transversal of the A-orbits on the k-subsets of V. Moreover, the picture changes completely when it comes to finding the number of A-orbits on the k-caps in V, where Polya's Theorem does not at all help: Actually, neither $k_{\max}$ is known, apart from a very small cases, nor the growth behavior of the number of orbits, for fixed k and growing d, let alone explicit formulae for them, are known. But since any subset of a cap again is a cap, it it possible to use recursion with respect to k to explicitly find a transversal of the A-orbits on k-caps, and hence in particular their number, and consequently $k_{\max}$, for a few non-trivial cases.

c) Now let $q := 3$. For $d = 1$ we get $A \cong \mathbb{F}_3 \rtimes \mathbb{F}_3^* \cong \mathcal{S}_3$, acting naturally, which (obviously) yields $c_A(1+Y, \ldots, 1+Y^3) = Y^3 + Y^2 + Y + 1 \in \mathbb{Q}[Y]$. For $d \in \{2, 3, 4\}$, where $|A| = 432$ and $|A| = 303264$ and $|A| = 1965150720$, respectively, proceeding as described above yields $c_A(1+Y, \ldots, 1+Y^9) = Y^9 + Y^8 + Y^7 + 2Y^6 + 2Y^5 + 2Y^4 + 2Y^3 + Y^2 + Y + 1 \in \mathbb{Q}[Y]$ for $d = 2$, and $c_A(1+Y, \ldots, 1+Y^{27}) = Y^{27} + Y^{26} + Y^{25} + 2Y^{24} + 3Y^{23} + 5Y^{22} + 10Y^{21} + 16Y^{20} + 28Y^{19} + 47Y^{18} + 68Y^{17} + 91Y^{16} + 114Y^{15} + 127Y^{14} + 127Y^{13} + 114Y^{12} + 91Y^{11} + 68Y^{10} + 47Y^9 + 28Y^8 + 16Y^7 + 10Y^6 + 5Y^5 + 3Y^4 + 2Y^3 + Y^2 + Y + 1 \in \mathbb{Q}[Y]$ for $d = 3$, and $c_A(1+Y, \ldots, 1+Y^{81}) \in \mathbb{Q}[Y]$ for $d = 4$ is given in Table 19;

hence there are 41407 inequivalent 12-subsets of cards. Note that, while for $d \leq 3$ transversals of the A-orbits on the k-subsets of V are easily found for $k \in \{1, \ldots, q^d\}$, for $d = 4$ this is computationally tractable only up to $k \sim 14$.

As for caps and the A-orbits thereon, the knowledge is very poor. The current state of the art for $k_{\max}$ is as follows:

d	1	2	3	4	5	6
$k_{\max}$	2	4	9	20	45	$\{112, 113, 114\}$

Transversals and the number of A-orbits are known for $d \in \{1, \ldots, 4\}$ and all $k \in \{0, \ldots, k_{\max}\}$, from explicit computations, see Table 20. This verifies the above results on $k_{\max}$ in this range. In particular, it turns out that there always is a single A-orbit of caps of cardinality $k_{\max}$; there is no a-priori-reason known why this should be the case. Moreover, for $d = 4$ we indicate the proportion of k-caps amongst all k-subsets of V; a cap of cardinality $k_{\max} = 20$ (actually the lexicographically smallest one) is reproduced in Table 21, together with its translation into real cards according to the table given earlier.

Table 19: Number of $\mathrm{AGL}_4(3)$-orbits on subsets of $\mathbb{F}_3^4$.

$$\begin{gathered}
Y^{81}+\\
Y^{80}+Y^{79}+\\
2Y^{78}+3Y^{77}+6Y^{76}+\\
15Y^{75}+34Y^{74}+105Y^{73}+\\
384Y^{72}+1658Y^{71}+8135Y^{70}+\\
41407Y^{69}+205211Y^{68}+963708Y^{67}+\\
4231059Y^{66}+17295730Y^{65}+65807588Y^{64}+\\
233346408Y^{63}+772518828Y^{62}+2392611091Y^{61}+\\
6946116261Y^{60}+18937468347Y^{59}+48568206996Y^{58}+\\
117356752981Y^{57}+267548687984Y^{56}+576222904363Y^{55}+\\
1173737365919Y^{54}+2263568972663Y^{53}+4136780036942Y^{52}+\\
7170309576688Y^{51}+11796184561289Y^{50}+18431386920534Y^{49}+\\
27367649303603Y^{48}+38636503940897Y^{47}+51883126670392Y^{46}+\\
66294936428615Y^{45}+80628826002618Y^{44}+93359571424793Y^{43}+\\
102934827016066Y^{42}+108081525023972Y^{41}+108081525023972Y^{40}+\\
102934827016066Y^{39}+93359571424793Y^{38}+80628826002618Y^{37}+\\
66294936428615Y^{36}+51883126670392Y^{35}+38636503940897Y^{34}+\\
27367649303603Y^{33}+18431386920534Y^{32}+11796184561289Y^{31}+\\
7170309576688Y^{30}+4136780036942Y^{29}+2263568972663Y^{28}+\\
1173737365919Y^{27}+576222904363Y^{26}+267548687984Y^{25}+\\
117356752981Y^{24}+48568206996Y^{23}+18937468347Y^{22}+\\
6946116261Y^{21}+2392611091Y^{20}+772518828Y^{19}+\\
233346408Y^{18}+65807588Y^{17}+17295730Y^{16}+\\
4231059Y^{15}+963708Y^{14}+205211Y^{13}+\\
41407Y^{12}+8135Y^{11}+1658Y^{10}+\\
384Y^{9}+105Y^{8}+34Y^{7}+\\
15Y^{6}+6Y^{5}+3Y^{4}+\\
2Y^{3}+Y^{2}+\\
Y+1
\end{gathered}$$

Table 20: Number of $\mathrm{AGL}_d(3)$-orbits on k-caps.

k	1	2	3	4	prop.
1	1	1	1	1	
2		1	1	1	
3		1	1	1	0.99
4		1	2	2	0.95
5			2	3	0.88
6			3	7	0.76
7			2	11	0.62
8			3	33	0.45
9			1	91	0.30
10				267	0.17
11				670	$8 \cdot 10^{-2}$
12				1437	$3 \cdot 10^{-2}$
13				2225	$1 \cdot 10^{-2}$
14				2489	$2 \cdot 10^{-3}$
15				1756	$4 \cdot 10^{-4}$
16				748	$3 \cdot 10^{-5}$
17				143	$1 \cdot 10^{-6}$
18				20	$1 \cdot 10^{-8}$
19				1	$9 \cdot 10^{-12}$
20				1	$1 \cdot 10^{-13}$

Table 21: A 20-cap for $d = 4$.

$[0, 0, 0, 0]$	three-green-open-diamond
$[0, 0, 0, 1]$	three-green-open-squiggle
$[0, 0, 1, 0]$	three-green-striped-diamond
$[0, 0, 1, 1]$	three-green-striped-squiggle
$[0, 1, 0, 0]$	three-red-open-diamond
$[0, 1, 0, 1]$	three-red-open-squiggle
$[0, 1, 1, 0]$	three-red-striped-diamond
$[0, 1, 1, 1]$	three-red-striped-squiggle
$[1, 0, 0, 0]$	one-green-open-diamond
$[1, 0, 0, 1]$	one-green-open-squiggle
$[1, 0, 1, 2]$	one-green-striped-oval
$[1, 0, 2, 2]$	one-green-solid-oval
$[1, 1, 0, 2]$	one-red-open-oval
$[1, 2, 0, 2]$	one-purple-open-oval
$[2, 0, 1, 2]$	two-green-striped-oval
$[2, 1, 0, 2]$	two-red-open-oval
$[2, 1, 1, 0]$	two-red-striped-diamond
$[2, 1, 1, 1]$	two-red-striped-squiggle
$[2, 1, 2, 2]$	two-red-solid-oval
$[2, 2, 1, 2]$	two-purple-striped-oval

VI Exercises and references

19 Exercises for Part I (in German)

(19.1) Aufgabe: Potenzmenge.
Es seien M eine beliebige Menge und $\mathcal{P}(M)$ ihre Potenzmenge. Man zeige: Es gibt keine Surjektion $M \to \mathcal{P}(M)$.

Hinweis. Für $f\colon M \to \mathcal{P}(M)$ surjektiv betrachte man $\{x \in M; x \notin f(x)\}$.

(19.2) Aufgabe: Potenzmengen endlicher Mengen.
Für $n \in \mathbb{N}_0$ bestimme man (ohne die Benutzung von Binomialkoeffizienten) die Anzahl der Teilmengen von $\{1, \ldots, n\}$ mit gerader bzw. ungerader Kardinalität.

(19.3) Aufgabe: Aufzählung von Potenzmengen.
Es sei $n \in \mathbb{N}_0$. Zur **Gray-Aufzählung** aller Teilmengen von $\{1, \ldots, n\}$ werden diese als n-Tupel mit Einträgen in $\{0, 1\}$ dargestellt. Ist $\mathcal{G}(n) = [x_1, \ldots, x_{2^n}]$ die Aufzählung für n, so sei $\mathcal{G}(n+1) := [0x_1, 0x_2, \ldots, 0x_{2^n}, 1x_{2^n}, 1x_{2^n-1}, \ldots, 1x_1]$, wobei $\mathcal{G}(0) := [[]]$ die Liste mit dem leeren Tupel sei. Man zeige:

a) Je zwei benachbarte Tupel in $\mathcal{G}(n)$ unterscheiden sich an genau einer Stelle.

b) Für $0 \leq k \leq n$ sei $\mathcal{G}_k(n)$ die Unterfolge von $\mathcal{G}(n)$ der Tupel mit genau k Einträgen 1. Man zeige: Je zwei benachbarte Tupel in $\mathcal{G}_k(n)$ unterscheiden sich an genau zwei Stellen.

(19.4) Aufgabe: Parlamentswahlen.
Das Parlament eines Landes habe $n \in \mathbb{N}$ Sitze, und es seien die rote, die grüne und die gelbe Partei vertreten. Wieviele mögliche Sitzverteilungen gibt es, so daß keine Partei die absolute Mehrheit hat?

(19.5) Aufgabe: Kartenmischen.
Ein Stapel von 52 üblichen Spielkarten wird gemischt. Wieviele Ergebnisse gibt es, bei denen sowohl die oberste als auch die unterste Karte ein Ass ist?

(19.6) Aufgabe: Fakultäten.
Man zeige: Jede natürliche Zahl $n \in \mathbb{N}_0$ hat eine eindeutige Darstellung der Form $n = \sum_{k \geq 1} a_k \cdot k!$ mit $a_k \in \{0, \ldots, k\}$.

(19.7) Aufgabe: Teilmengen ohne benachbarte Elemente.
a) Für $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$ seien $a_k(n) \in \mathbb{N}_0$ die Anzahl der k-elementigen Teilmengen $M \subseteq \{1, \ldots, n\}$ mit $|x - y| \geq 2$ für alle $x \neq y \in M$, sowie $a(n) = \sum_{k=0}^{n} a_k(n)$ die Anzahl aller solcher Teilmengen von $\{1, \ldots, n\}$.

Man zeige kombinatorisch: Es gelten $a_k(n) = \binom{n-k+1}{k}$ sowie $a(n) = F_{n+2}$, die $(n+2)$-te Fibonacci-Zahl. (Dies ist also ein kombinatorischer Beweis der Identität $\sum_{k=0}^{n} \binom{n-k+1}{k} = F_{n+2}$.)

b) Nun sei $j \in \mathbb{N}$. Man bestimme die Anzahl der k-elementigen Teilmengen $M \subseteq \{1, \ldots, n\}$ mit $|x - y| \geq j$ für alle $x \neq y \in M$.

(19.8) Aufgabe: Multimengen.
Es seien $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$. Man zeige: Es gilt $|\mathcal{M}_k(n+1)| = |\mathcal{M}_n(k+1)|$; dazu gebe man einen algebraischen und einen kombinatorischen Beweis an.

(19.9) Aufgabe: Teiler.
Für $n \in \mathbb{N}$ sei $\tau(n) := |\{d \in \mathbb{N}; d \mid n\}| \in \mathbb{N}$ die Anzahl der Teiler von n. Man zeige: Die Anzahl $\tau(n)$ ist genau dann ungerade, wenn n eine Quadratzahl ist.

(19.10) Aufgabe: Gefangenenbefreiung.
Der Gefängniswärter eines Gefängnisses mit 100 Zellen will einige Gefangenen freilassen, indem er folgendes Verfahren anwendet: Zunächst schließt er alle Türen auf, dann der Reihe nach jede zweite wieder zu, danach ändert er den Zustand jeder dritten Tür von offen nach geschlossen und umgekehrt, und so fährt er mit jeder vierten, fünften, bis hundertsten Tür fort. Am Ende dürfen die Gefangenen gehen, deren Tür offen ist. Welche Gefangenen sind das?

(19.11) Aufgabe: Warteschlangen.
Die Personen $\{1, \ldots, n\}$, für $n \in \mathbb{N}$, mögen eine Warteschlage bilden, so daß sich, mit Ausnahme der vordersten Person, vor jeder Person i die Person $i-1$ oder $i+1$ befinde. Wieviele mögliche Warteschlangen gibt es?

(19.12) Aufgabe: Fallende Faktorielle.
a) Es sei $n \in \mathbb{N}_0$. Man zeige: Es gilt $(2X)_{(2n)} = 2^{2n} X_{(n)} (X - \frac{1}{2})_{(n)} \in \mathbb{Z}[X]$.

b) Daraus folgere man: Es gilt $\binom{-\frac{1}{2}}{n} = (-\frac{1}{4})^n \cdot \binom{2n}{n}$.

(19.13) Aufgabe: Binomialkoeffizienten.
Es sei $n \in \mathbb{N}_0$. Man zeige:

a) Es gilt **Unimodalität** $\binom{n}{0} < \binom{n}{1} < \cdots < \binom{n}{\lfloor \frac{n}{2} \rfloor} = \binom{n}{\lceil \frac{n}{2} \rceil} > \cdots > \binom{n}{n-1} > \binom{n}{n}$.

b) Es gelten $\sum_{k=0}^{n} k \cdot \binom{n}{k} = n \cdot 2^{n-1}$ und $\sum_{k=0}^{n} k^2 \cdot \binom{n}{k} = n(n+1) \cdot 2^{n-2}$.

c) Es gilt $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$.

(19.14) Aufgabe: Identitäten für Binomialkoeffizienten.
Es seien $m \leq n \in \mathbb{N}_0$. Man zeige:

a) Es gilt $\binom{n}{m}\binom{m}{k} = \binom{n}{k}\binom{n-k}{m-k}$ und $\sum_{k=0}^{m} \binom{n}{k}\binom{n-k}{m-k} = 2^m \cdot \binom{n}{m}$. Welche kombinatorische Interpretation hat diese Identität?

b) Es gilt $\sum_{k=0}^{m} (-1)^k \cdot \binom{n}{k}\binom{n-k}{m-k} = \begin{cases} 1, & \text{falls } m = 0, \\ 0, & \text{falls } m \geq 1. \end{cases}$

c) Es gilt $\sum_{k=0}^{m}(-1)^k \cdot \binom{n}{k}\binom{n}{m-k} = \begin{cases} 0, & \text{falls } m \text{ ungerade}, \\ (-1)^{\frac{m}{2}} \cdot \binom{n}{\frac{m}{2}}, & \text{falls } m \text{ gerade.} \end{cases}$

Hinweis zu c). Man betrachte $(X+1)^n(X-1)^n = (X^2-1)^n \in \mathbb{Z}[X]$.

(19.15) Aufgabe: Pascal-Dreieck.
Es seien $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$. Man zeige: Für **partielle Antidiagonalsummen** gilt $\sum_{i=0}^{k}\binom{n+i}{k-i} = F_{n+k+1}$, die $(n+k+1)$-te Fibonacci-Zahl.

(19.16) Aufgabe: Verschobenes Pascal-Dreieck.
Man betrachte das verschobene Pascal-Dreieck, in dem die $n+1$ nicht-verschwindenden Einträge in Zeile $n \in \mathbb{N}_0$ in den Spalten $2n, \ldots, 3n$ stehen, und die Einträge eingerahmt werden, die Vielfache von n sind:

$n\backslash k$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	1															
1			$\boxed{1}$	$\boxed{1}$												
2					1	$\boxed{2}$	1									
3							1	$\boxed{3}$	$\boxed{3}$	1						
4									1	$\boxed{4}$	6	$\boxed{4}$	1			
5											1	$\boxed{5}$	$\boxed{10}$	$\boxed{10}$	$\boxed{5}$	1

Man zeige: Es ist $k \geq 2$ genau dann eine Primzahl, wenn alle nicht-verschwindenden Einträge in Spalte k eingerahmt sind.

(19.17) Aufgabe: Gitterwege.
Für $m, n \in \mathbb{N}_0$ betrachte man ein Gitter mit den Kantenlängen m und n. Durch Klassifikation von Wegen gebe man kombinatorische Beweise der folgenden Identitäten an: Für $k \in \mathbb{N}_0$ gilt $\sum_{i=0}^{k}\binom{n+i}{i} = \binom{n+k+1}{k}$ und $\binom{m+n}{k} = \sum_{i=0}^{k}\binom{m}{i}\binom{n}{k-i}$.

Hinweis. Man betrachte das erste Eintreffen auf der rechten Vertikalen bzw. den Schnittpunkt mit einer Antidiagonalen.

(19.18) Aufgabe: Schachbrett.
Für $n \in \mathbb{N}_0$ betrachte man ein $(n \times n)$-Schachbrett. Man bestimme die Anzahl der kürzesten Wege, auf denen ein König von der linken unteren zur rechten oberen Ecke gelangen kann.

(19.19) Aufgabe: Bertrands Problem.
Zwei Kandidaten A und B erhalten bei einer Wahl $a > b \in \mathbb{N}_0$ Stimmen. Man zeige, daß die Stimmzettel auf genau $\frac{a-b}{a+b} \cdot \binom{a+b}{a}$ Weisen geordnet werden können, so daß bei der sukzessiven Auszählung A stets mehr Stimmen als B hat.

Hinweis. Man betrachte geeignete Gitterwege.

(19.20) Aufgabe: Kleiner Satz von Fermat.
Es seien $n \in \mathbb{N}$ und $p \in \mathbb{N}$ eine Primzahl. Man zeige: Es gilt $n^p \equiv n \pmod{p}$; dazu gebe man einen algebraischen und einen kombinatorischen Beweis an.

(19.21) Aufgabe: Folgen von Teilmengen.
Es seien $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$. Man bestimme jeweils die Anzahl der Folgen $[N_1, \ldots, N_k]$ von Teilmengen $N_i \subseteq N := \{1, \ldots, n\}$ mit:

a) $N_1 \subseteq N_2 \subseteq \cdots \subseteq N_k$;

b) $N_i \cap N_j = \emptyset$ für alle $i \neq j \in \{1, \ldots, k\}$;

c) $N_1 \cup N_2 \cup \cdots \cup N_k = N$.

(19.22) Aufgabe: Stirling-Zahlen zweiter Art.
a) Für $n \in \mathbb{N}_0$ zeige man: Es gilt $S_{n,0} < S_{n,1} < \cdots < S_{n,m_n-1} \leq S_{n,m_n} > \cdots > S_{n,n-1} > S_{n,n}$, für ein $m_n \in \{0, \ldots, n\}$ mit $m_n = m_{n-1} + \epsilon$, wobei $\epsilon \in \{0, 1\}$.

b) Für $n \geq 2$ zeige man: Es gilt $S_{n,n-2} = \binom{n}{3} + 3 \cdot \binom{n}{4}$.

c) Es seien $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$. Man zeige: Es gilt $S_{n+1,k+1} = \sum_{i=0}^{n} \binom{n}{i} S_{i,k}$; dazu gebe man einen algebraischen und einen kombinatorischen Beweis an.

d) Man bestimme die Anzahl der Tupel $[a_1, \ldots, a_n] \in \mathbb{N}^n$, für $n \in \mathbb{N}_0$, deren Einträge genau die Zahlen $\{1, \ldots, k\}$ sind (eventuell mehrfach), so daß der erste Eintrag i vor dem ersten Eintrag $i + 1$ liegt, für alle $i \in \{1, \ldots, k - 1\}$.

(19.23) Aufgabe: Bell-Zahlen.
a) Es sei $n \in \mathbb{N}_0$. Man zeige: Es gilt $B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k$.

b) Es sei $n \in \mathbb{N}$. Man zeige: Es gibt genau B_{n-1} Partitionen von $\{1, \ldots, n\}$, deren sämtliche Blöcke keine zwei aufeinanderfolgende Zahlen enthalten.

(19.24) Aufgabe: Teile in Kompositionen.
a) Es sei $n \in \mathbb{N}$. Man zeige: Die Gesamtanzahl aller Teile in allen Kompositionen von n ist gleich $(n + 1) \cdot 2^{n-2}$.

b) Es sei $k \in \mathbb{N}$ mit $k < n$. Man zeige: In allen Kompositionen von n kommt der Teil k insgesamt genau $(n - k + 3) \cdot 2^{n-k-2}$ mal vor.

(19.25) Aufgabe: Kompositionen.
a) Es seien $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$. Man bestimme jeweils die Anzahl der Tupel $[a_1, \ldots, a_k]$ positiver bzw. nicht-negativer ganzer Zahlen mit $\sum_{i=1}^{k} a_i \leq n$.

b) Unter Benutzung der Fibonacci-Zahlen zeige man:
i) Es gibt genau F_{n-1} Kompositionen von n in Teile ≥ 2;
ii) es gibt genau F_{n+1} Kompositionen von n in Teile ≤ 2;
iii) es gibt genau F_n Kompositionen von n in ungerade Teile.

(19.26) Aufgabe: Partitionen.
Es sei $p_n \in \mathbb{N}$ die Anzahl der Partitionen von $n \in \mathbb{N}_0$. Man zeige: Es gilt $p_{n+1} - 2p_{n+1} + p_n \geq 0$.

Hinweis. Es ist $p_n - p_{n-1}$ die Anzahl der Partitionen von n ohne den Teil 1.

(19.27) Aufgabe: Partitionen mit mehrfachen Teilen.
Es sei $n \in \mathbb{N}_0$. Für eine Partition $\lambda = [\lambda_1, \ldots, \lambda_k] = [n^{a_n(\lambda)}, \ldots, 1^{a_1(\lambda)}] \in P_k(n)$ mit $k \in \mathbb{N}_0$ Teilen seien $\alpha(\lambda) := \prod_{i=1}^k \lambda_i \in \mathbb{N}$ und $\beta(\lambda) := \prod_{i=1}^n a_i(\lambda)! \in \mathbb{N}$. Man zeige: Es gilt $\prod_{\lambda \in P(n)} \alpha(\lambda) = \prod_{\lambda \in P(n)} \beta(\lambda)$.

Hinweis. Für $j \in \mathbb{N}$ setze man $b_j(\lambda) := |\{i \in \{1, \ldots, n\}; a_i(\lambda) \geq j\}| \in \mathbb{N}_0$ und zeige $\sum_{\lambda \in P(n)} a_j(\lambda) = \sum_{\lambda \in P(n)} b_j(\lambda)$.

(19.28) Aufgabe: Partitionen mit vielen Teilen.
Es sei $k \in \mathbb{N}_0$. Man zeige: Es gibt $n_k \geq k$ und $\rho_k \in \mathbb{N}$ mit $p_{n,n-k} = \rho_k$ für alle $n \geq n_k$. Man bestimme ρ_k, und gebe das kleinstmögliche n_k an.

(19.29) Aufgabe: Partitionen mit drei Teilen.
Es sei $n \in \mathbb{N}_0$. Man zeige: Ist $n \equiv a \pmod 6$, wobei $a \in \{0, \ldots, 5\}$, so ist die Anzahl der Partitionen von n in drei Teile gegeben als $p_{a,3}(n) \in \mathbb{N}_0$, wobei

$$p_{a,3}(X) = \begin{cases} \frac{X^2}{12}, & \text{für } a = 0, \\ \frac{X^2-1}{12}, & \text{für } a \in \{1, 5\}, \\ \frac{X^2-4}{12}, & \text{für } a \in \{2, 4\}, \\ \frac{X^2+3}{12}, & \text{für } a = 3. \end{cases}$$

(19.30) Aufgabe: Aufzählung von Permutationen.
Für $n \in \mathbb{N}_0$ sei $\mathcal{S}_n = \{[1, 2, \ldots, n], \ldots, [n, n-1, \ldots, 1]\}$ in lexikographischer Anordnung gegeben. Für $\pi = [\pi_1, \ldots, \pi_n] \in \mathcal{S}_n$ sei $\alpha_n(\pi) + 1 \in \{1, \ldots, n!\}$ die Position von π in $\mathcal{S}_n$, etwa $\alpha_n([1, 2, \ldots, n]) = 0$ und $\alpha_n([n, n-1, \ldots, 1]) = n! - 1$.

a) Man gebe einen Algorithmus an, der zu einer Permutation $\pi \in \mathcal{S}_n$ den Nachfolger berechnet, also die Permutation $\sigma \in \mathcal{S}_n$ mit $\alpha_n(\sigma) = \alpha_n(\pi) + 1$.

b) Für $n \geq 1$ zeige man: Es gilt $\alpha_n(\pi) = (\pi_1 - 1)(n-1)! + \alpha_{n-1}(\pi')$, wobei π' aus π entsteht, indem man π_1 entfernt und alle $\pi_i > \pi_1$ durch $\pi_i - 1$ ersetzt.

c) Für $l \in \{0, \ldots, n! - 1\}$ bestimme man die Permutation $\pi \in \mathcal{S}_n$ mit $\alpha_n(\pi) = l$.

(19.31) Aufgabe: Bell-Zahlen und Permutationen.
Für $n \in \mathbb{N}$ seien $\mathcal{I}_n^+ \subseteq \mathcal{S}_n$ und $\mathcal{I}_n^- \subseteq \mathcal{S}_n$ die Mengen aller Permutationen $\pi = [\pi_1, \ldots, \pi_n] \in \mathcal{S}_n$ mit folgender Eigenschaft: Es gibt kein $i < j \in \{1, \ldots, n-1\}$ mit $\pi_i < \pi_j < \pi_{j+1}$. bzw. mit $\pi_i < \pi_{j+1} < \pi_j$.

a) Man zeige: Es gilt $|\mathcal{I}_n^+| = B_n = |\mathcal{I}_n^-|$.

b) Man beschreibe die Menge $\mathcal{I}_n^+ \cap \mathcal{I}_n^- \subseteq \mathcal{S}_n$.

(19.32) Aufgabe: Inversionen.
Es seien $n \in \mathbb{N}_0$ und $\pi = [\pi_1, \ldots, \pi_n] \in \mathcal{S}_n$. Ein Paar $\pi_i > \pi_j$, wobei $i < j$ ist, heißt eine **Inversion** von π. Es sei $l(\pi) \in \mathbb{N}_0$ die Anzahl der Inversionen von π, und für $i \in \{1, \ldots, n\}$ sei $\iota_i \in \mathbb{N}_0$ die Anzahl der $j \in \{i+1, \ldots, n\}$, sodaß $[j, i]$ eine Inversion von π ist; die Folge $[\iota_1, \ldots, \iota_n]$ heißt die **Inversionstafel** von π.

a) Es sei $I_{n,k} \in \mathbb{N}_0$ die Anzahl der Permutationen in $\mathcal{S}_n$ mit $k \in \mathbb{N}_0$ Inversionen. Man zeige: Es gilt $I_{n,0} = 1$, und $I_{n,k} = I_{n-1,k} + I_{n,k-1}$ für $1 \leq k < n$.

b) Man zeige: Es gilt $I_{n,k} = 0$ für $k > \binom{n}{2}$, und $I_{n,k} = I_{n,\binom{n}{2}-k}$ für $k \leq \binom{n}{2}$, sowie $\sum_{k=0}^{\binom{n}{2}} (-1)^k I_{n,k} = 0$ für $n \geq 2$. Außerdem gilt $l(\pi) = l(\pi^{-1})$.

c) Man zeige: Für die Einträge von Inversionstafeln gilt $\iota_i \in \{0, \ldots, n-i\}$, für alle $i \in \{1, \ldots, n\}$. Umgekehrt ist jede Folge $[\iota_1, \ldots, \iota_n]$ mit $\iota_i \in \{0, \ldots, n-i\}$, für alle $i \in \{1, \ldots, n\}$, die Inversionstafel genau einer Permutation in $\mathcal{S}_n$.

(19.33) Aufgabe: Stirling-Zahlen erster Art.
a) Für $n \in \mathbb{N}_0$ zeige man: Es gilt $s_{n,0} < s_{n,1} < \cdots < s_{n,m_n-1} \leq s_{n,m_n} > \cdots > s_{n,n-1} > s_{n,n}$, für ein $m_n \in \{0, \ldots, n\}$ mit $m_n = m_{n-1} + \epsilon$, wobei $\epsilon \in \{0, 1\}$.

b) Für $n \geq 2$ zeige man: Es gilt $s_{n,n-2} = 2 \cdot \binom{n}{3} + 3 \cdot \binom{n}{4}$.

c) Es seien $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$. Man zeige: Es gilt $s_{n+1,k+1} = \sum_{i=0}^{n} \binom{i}{k} s_{n,i}$; dazu gebe man einen algebraischen und einen kombinatorischen Beweis an.

(19.34) Aufgabe: Fixpunktfreie Permutationen.
Für $n \geq 2$ sei $D^+(n) \in \mathbb{N}_0$ die Anzahl der fixpunktfreien Permutationen in der **alternierenden Gruppe** $\mathcal{A}_n := \{\pi \in \mathcal{S}_n; \operatorname{sgn}(\pi) = 1\} < \mathcal{S}_n$. Man zeige: Es gilt $D^+(n) = (-1)^{n-1}(n-1) + \frac{n!}{2} \cdot \sum_{k=0}^{n-2} \frac{(-1)^k}{k!}$.

(19.35) Aufgabe: Die Theatergarderobe.
Die Garderobenpersonal des 'Theater 99' ist unaufmerksam. (Das Theater ist übrigens nach der Anzahl seiner Sitzplätze benannt.) Nach der Vorstellung werden die Mäntel zufällig an die Besucher zurückgegeben. Wie groß ist die Wahrscheinlichkeit (ungefähr), daß dabei keiner den eigenen Mantel erhält?

(19.36) Aufgabe: Derangement-Zahlen.
Für $n \in \mathbb{N}$ betrachte man die folgenden Zahlen: Man beginne mit 1, subtrahiere 1, multipliziere mit 2, addiere 1, multipliziere mit 3, subtrahiere 1, …, multipliziere mit n und addiere $(-1)^n$. Welche bekannte Folge erhält man?

(19.37) Aufgabe: Türme von Hanoi.
Auf drei senkrechten Stäben A, B und C sind $n \in \mathbb{N}_0$ Scheiben unterschiedlichen Durchmessers gestapelt, so daß jede nicht zuunterst liegende Scheibe auf einer mindestens ebenso großen Scheibe liegt. Durch einen Zug kann man die oberste Scheibe eines Stabes auf einen anderen Stab bewegen, falls dabei diese Bedingung nicht verletzt wird. Zunächst befinden sich alle Scheiben auf Stab A.

a) Wieviele Züge sind nötig, die Scheiben nach Stab B zu bewegen, wenn die Scheiben paarweise verschiedenen Durchmesser haben?

b) Wieviele Züge sind nötig, die Scheiben nach Stab B zu bewegen, wenn die Scheiben paarweise verschiedenen Durchmesser haben, und nur Züge zwischen den Stäben A und C sowie B und C erlaubt sind?

c) Wieviele Züge sind nötig, die Scheiben nach Stab B zu bewegen, wenn Scheiben gleichen Durchmessers jeweils paarweise vorkommen?

(19.38) Aufgabe: Summation.
Für $n \in \mathbb{N}_0$ berechne man $\sum_{k=1}^{n} \frac{2k+1}{k(k+1)} \in \mathbb{Q}$ und $\sum_{k=1}^{n} (-1)^k \cdot k \in \mathbb{Q}$ sowie $\sum_{k=1}^{n} (-1)^k \cdot k^2 \in \mathbb{Q}$.

(19.39) Aufgabe: Reelle Differentiation.
Für $a \in \mathbb{R} \setminus \{1\}$ und $n \in \mathbb{N}_0$ seien $g_a(n) := \sum_{k=1}^{n} a^k$ und $f_a(n) := \sum_{k=1}^{n} k \cdot a^k$. Man gebe geschlossene Formeln für $g_a(n)$ und $f_a(n)$ an. Welche bekannte Formel erhält man für $f_2(n)$? Was passiert für $a = 1$?

(19.40) Aufgabe: Potenzsummen.
Für $k \geq 1$ sei $s_k \in \mathbb{Q}[X]$, so daß $s_k(n) = \sum_{i=0}^{n-1} i^k$ für alle $n \in \mathbb{N}_0$.

a) Man zeige: Es gilt $s_k(1 - X) = (-1)^{k+1} s_k(X) \in \mathbb{Q}[X]$.

b) Man bestimme $s_k(1) \in \mathbb{Q}$, sowie $s_k(\frac{1}{2}) \in \mathbb{Q}$ für k gerade.

(19.41) Aufgabe: Partielle Summation.
Es sei $n \in \mathbb{N}_0$, und für $k \in \mathbb{N}_0$ sei $h_k \in \mathbb{Q}$ die k-te harmonische Zahl.

a) Man berechne $\sum_{k=1}^{n} \frac{h_k}{k} \in \mathbb{Q}$ und $\sum_{k=1}^{n} \frac{h_k}{(k+1)(k+2)} \in \mathbb{Q}$.

b) Man berechne $\sum_{k=1}^{n} h_k^2 \in \mathbb{Q}$.

c) Für $m \in \mathbb{N}_0$ berechne man $\sum_{k=1}^{n} \binom{k}{m} \cdot h_k \in \mathbb{Q}$ und $\sum_{k=1}^{n} (-1)^k \binom{m}{k} \cdot h_k \in \mathbb{Q}$.

(19.42) Aufgabe: Diskrete Integration.
Man gebe eine geschlossene Formel für $\sum_{k=0}^{n} \frac{(-1)^k}{\binom{n}{k}}$, wobei $n \in \mathbb{N}_0$, an.

(19.43) Aufgabe: Summen fallender Faktorieller.
Für $n \in \mathbb{N}_0$ sei $f_n := \sum_{k=0}^{n-1} n_{(k)} \in \mathbb{N}$. Man bestimme eine Rekursionsformel und daraus eine geschlossene Formel für f_n.

(19.44) Aufgabe: Binomialinversion.
Man zeige: Es gibt eine Folge $[a_n \in \mathbb{Q}; n \in \mathbb{N}_0]$, so daß $n! = \sum_{k \geq 0} a_k n_{(k)}$, für alle $n \in \mathbb{N}_0$, gilt. Man gebe eine geschlossene Formel für die Zahlen a_n an.

(19.45) Aufgabe: Lineare Inversion.
a) Man zeige, daß die Basisfolgen $[X_{(n)} \in \mathbb{Q}[X]; n \in \mathbb{N}_0]$ und $[(-1)^n X^{(n)} \in \mathbb{Q}[X]; n \in \mathbb{N}_0]$ durch die **Lah-Zahlen** $L_{n,k} := (-1)^n \cdot \frac{n!}{k!} \cdot \binom{n-1}{k-1}$, für $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$, zusammenhängen.

Daraus folgere man die **Lah-Inversion**: Sind $[x_n \in \mathbb{Q}; n \in \mathbb{N}_0]$ und $[y_n \in \mathbb{Q}; n \in \mathbb{N}_0]$, so gilt genau dann $y_n = \sum_{k=0}^{n} L_{n,k} x_k \in \mathbb{Q}$, für alle $n \in \mathbb{N}_0$, wenn $x_n = \sum_{k=0}^{n} L_{n,k} y_k \in \mathbb{Q}$, für alle $n \in \mathbb{N}_0$, gilt.

b) Unter Benutzung der Basisfolge $\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \frac{n}{n-k} \binom{n-k}{k} X^{n-2k} \in \mathbb{Q}[X]$, für $n \in \mathbb{N}_0$, beweise man die **Tschebyscheff-Inversion**: Sind $[x_n \in \mathbb{Q}; n \in \mathbb{N}_0]$ und $[y_n \in \mathbb{Q}; n \in \mathbb{N}_0]$, so gilt genau dann $y_n = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{k} x_{n-2k} \in \mathbb{Q}$, für alle $n \in \mathbb{N}_0$, wenn $x_n = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \frac{n}{n-k} \binom{n-k}{k} y_{n-2k} \in \mathbb{Q}$, für alle $n \in \mathbb{N}_0$, gilt.

20 Exercises for Part II (in German)

(20.1) Aufgabe: Endliche partiell geordnete Mengen.
a) Man bestimme bis auf Isomorphie alle endlichen partiell geordneten Mengen der Kardinalität ≤ 4.

b) Man bestimme bis auf Isomorphie alle Verbände der Kardinalität ≤ 6.

(20.2) Aufgabe: Partiell geordnete Mengen.
Es sei X eine partiell geordnete Menge.

a) Man zeige: Es gibt eine Menge N und eine Teilmenge $\mathcal{X} \subseteq \mathcal{P}(N)$, so daß $X \cong \mathcal{X}$ als partiell geordnete Mengen. Ist zudem X endlich, so kann auch N endlich gewählt werden.

b) Es seien X endlich und $f \colon X \to X$ eine **ordnungserhaltende** Bijektion, das heißt, für $x \leq y \in X$ ist stets auch $f(x) \leq f(y)$. Man zeige: Die Abbildung f ist ein Automorphismus. Kann man auf die Endlichkeitsvoraussetzung verzichten?

b) Es sei X ein Verband, und $f \colon X \to X$ eine ordnungserhaltende Bijektion, die auch kleinste obere Schranken und größte untere Schranken erhält. Man zeige: Die Abbildung f ist ein Automorphismus.

(20.3) Aufgabe: Endliche Verbände.
a) Es sei X eine endliche partiell geordnete Menge, die ein Eins-Element besitzt, und für die je zwei Elemente eine größte untere Schranke besitzen. Man zeige: X ist ein Verband. Man formuliere und beweise die dazu duale Aussage.

b) Es sei X ein endlicher Verband. Man zeige die Äquivalenz der folgenden Aussagen: **i)** X ist graduiert, und für die Längen-Funktion gilt $d(x) + d(y) \geq d(x \vee y) + d(x \wedge y)$, für alle $x, y \in X$. **ii)** Für alle $x, y \in X$ mit $x \wedge y \lessdot x$ und $x \wedge y \lessdot y$ gilt stets auch $x \lessdot x \vee y$ und $y \lessdot x \vee y$.

(20.4) Aufgabe: Maximale Ketten.
Es sei X eine lokal-endliche partiell geordnete Menge, und für $x, y \in X$ sei $m(x, y) \in \mathbb{N}_0$ die Anzahl der maximalen Ketten zwischen x und y.

a) Es sei $\eta \in \mathcal{A}(X)$ definiert durch $\eta(x, y) := 1$ für $x \lessdot y$, und $\eta(x, y) := 0$ sonst. Man zeige: Die Funktion $\delta - \eta \in \mathcal{A}(X)$ ist invertierbar, und für $x, y \in X$ gilt $(\delta - \eta)^{-1}(x, y) = m(x, y)$.

b) Es seien X zudem graduiert, und $f \in \mathcal{A}(X)$ definiert durch $f(x, y) := \frac{m(x,y)}{l(x,y)!}$ für $x \leq y$, und $f(x, y) := 0$ sonst. Man zeige: Die Funktion $f \in \mathcal{A}(X)$ ist invertierbar, und für $x \leq y$ gilt $f^{-1}(x, y) = (-1)^{l(x,y)} \cdot f(x, y)$.

(20.5) Aufgabe: Dominanz-Ordnung.
Man betrachte die durch $\trianglelefteq$ gegebene partielle Ordnung auf $P(n)$, für $n \in \mathbb{N}_0$.

a) Man zeige: Die durch $\trianglelefteq$ gegebene partielle Ordnung auf $P(n)$ ist ein Verband. Für welche $n \in \mathbb{N}_0$ ist $\trianglelefteq$ eine Ordnung? Für welche $n \in \mathbb{N}_0$ ist $\trianglelefteq$ graduiert? Man zeichne die Hasse-Diagramme von $\trianglelefteq$ für $n \leq 8$.

b) Für $\lambda = [\lambda_1, \dots, \lambda_n] \vdash n$ und $\mu = [\mu_1, \dots, \mu_n] \vdash n$ schreibe $\mu \leq \lambda$, falls es $k \in \{1, \dots, n\}$ gibt mit $\mu_i = \lambda_i$ für alle $i \in \{1, \dots, k-1\}$, und $\mu_k < \lambda_k$. Man zeige: Dies definiert eine Ordnung, die **lexikographische Ordnung**, auf $P(n)$, die die partielle Ordnung $\trianglelefteq$ verfeinert.

21 Exercises for Part III (in German)

(21.1) Aufgabe: Möbius-Funktion.
Es sei X eine endliche partiell geordnete Menge mit Null-Element $\underline{0}$ und Eins-Element $\underline{1}$. Für die Möbius-Funktion μ von X zeige man:

a) Es gilt $\sum_{x,y \in X, x \leq y} \mu(x, y) = 1$.

b) Es gilt $\sum_{k \geq 0} \sum_{\underline{0}=x_0<x_1<\cdots<x_k=\underline{1}} (-1)^k \cdot \prod_{i=1}^k \mu(x_{i-1}, x_i) = 1$.

(21.2) Aufgabe: Möbius-Inversion.
Es seien X ein endlicher Verband X mit Null-Element $\underline{0}$, sowie $g\colon X \to \mathbb{N}_0\colon x \mapsto |\{y \in X; x \leq y\}|$ und $f_k\colon X \to \mathbb{N}_0\colon x \mapsto |\{|x_1, \dots, x_k] \in X^k; x_1 \wedge \cdots \wedge x_k = x\}|$, für $k \in \mathbb{N}_0$. Man zeige: Es gilt $f_k(x) = \sum_{x \in X} \mu(\underline{0}, x) \cdot g(x)^k$; dazu gebe man sowohl einen Beweis mittels Möbius-Inversion an, also auch einen, der die Möbius-Algebra benutzt.

(21.3) Aufgabe: Zahlentheoretische Möbius-Funktion.
Für $n \in \mathbb{N}$ sei $\mathcal{U}_n := \{d \in \mathbb{N}; d \mid n, \mathrm{ggT}(d, \frac{n}{d}) = 1\}$ partiell geordnet durch Teilbarkeit. Man zeige: $\mathcal{U}_n$ ist ein Verband, und bestimme seine Möbius-Funktion.

(21.4) Aufgabe: Inklusion-Exklusion.
Es seien $m \leq k \leq n$. Man zeige: Es gilt $\binom{n-m}{k-m} = \sum_{i=0}^m (-1)^i \cdot \binom{m}{i}\binom{n-i}{k}$.

Hinweis. Es ist $\binom{n-m}{k-m}$ die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge, die eine feste m-elementige Teilmenge umfassen.

(21.5) Aufgabe: Dezimalzahlen.
Wieviele Zahlen mit genau $k \in \mathbb{N}$ Dezimalstellen gibt es, in deren Dezimaldarstellung keine benachbarten Ziffern gleich sind?

(21.6) Aufgabe: Primzahlen.
a) Unter Benutzung von Inklusion-Exklusion gebe man ein Verfahren an, um aus der Kenntnis aller Primzahlen $\leq \sqrt{n}$, wobei $n \in \mathbb{N}$, die Anzahl der Primzahlen $\leq n$ bestimmen kann, ohne diese zu berechnen. Als Anwendung bestimme man die Anzahl der Primzahlen ≤ 100 aus der Kenntnis der Primzahlen $\{2, 3, 5, 7\}$, und verifiziere das Ergebnis explizit.

b) Für $n \in \mathbb{N}$ sei $\varphi(n) := |\{k \in \{1, \ldots, n\}; \mathrm{ggT}(k, n) = 1\}| \in \mathbb{N}$ die Euler-Funktion. Mittels Inklusion-Exklusion zeige man: Sind $\{p_1, \ldots, p_l\}$ die Primteiler von n, so gilt $\frac{\varphi(n)}{n} = \prod_{i=1}^{l}(1 - \frac{1}{p_i})$.

(21.7) Aufgabe: Partitionen in ungerade und verschiedene Teile.
Für $n \in \mathbb{N}_0$ seien $O(n)$ und $D(n)$ die Mengen aller Partitionen von n in lauter ungerade bzw. in paarweise verschiedene Teile. Man gebe eine explizite Bijektion $O(n) \to D(n)$ an.

Hinweis. Man betrachte die folgende Abbildung $O(n) \to D(n)\colon \lambda \mapsto \mu$: Hat λ den Teil i genau a-mal, so habe μ den Teil $i \cdot 2^b$ genau dann, wenn 2^b in der 2-adischen Darstellung von a vorkommt.

(21.8) Aufgabe: Menage-Zahlen.
Es sei $u_n := \sum_{k=0}^{n} \frac{(-1)^k \cdot (n-k)! \cdot 2n}{2n-k} \cdot \binom{2n-k}{k} \in \mathbb{N}_0$ die **Menage-Zahl**, für $n \geq 2$.

a) Man zeige: Für $n \geq 4$ gilt $(n-2)u_n = n(n-2)u_{n-1} + nu_{n-2} + 4 \cdot (-1)^{n+1}$.

b) Man zeige: Es gilt $\lim_{n\to\infty} \frac{u_n}{n!} = \frac{1}{e^2}$.

(21.9) Aufgabe: Folgen mit verschiedenen Nachbarn.
Es sei $\mathcal{X}_n$ die Menge aller Folgen der Länge $2n$ über einer Menge X der Kardinalität $n \in \mathbb{N}_0$, die jedes Element von X genau zweimal enthalten, und so daß benachbarte Folgenglieder jeweils verschieden sind. Man zeige: Es gilt $|\mathcal{X}_n| = \sum_{k=0}^{n} \frac{(-1)^{n-k} \cdot (n+k)!}{2^k} \cdot \binom{n}{k}$.

(21.10) Aufgabe: Äquivalenzklassen von Permutationen.
Es sei $n \in \mathbb{N}_0$. Zwei Permutationen der Menge $X_n := \{a_1, \ldots, a_n, b_1, \ldots, b_n\}$, geschrieben als Listen, heißen äquivalent, falls sie durch Vertauschung zweier benachbarter Einträge $a_i b_i$ oder $b_i a_i$, für ein $i \in \{1, \ldots, n\}$, auseinander hervorgehen. Man zeige: Die davon erzeugte Äquivalenzrelation auf $\mathcal{S}_{X_n}$ besteht aus genau $\sum_{i=0}^{n}(-1)^i \binom{n}{i}(2n-i)!$ Äquivalenzklassen.

(21.11) Aufgabe: Paarungen.
An einem runden Tisch sitzen $2n$ Personen, wobei $n \in \mathbb{N}_0$. Auf wieviele Weisen kann man daraus n Paare bilden, wenn sich darunter keine zwei benachbarte Personen befinden dürfen?

(21.12) Aufgabe: Zykel mit eingeschränkten Nachbarn.
Für $n \in \mathbb{N}_0$ sei $f(n) \in \mathbb{N}_0$ die Anzahl der n-Zykel $(a_1, \ldots, a_n) \in \mathcal{S}_n$, so daß $a_{i+1} \not\equiv a_i + 1 \pmod{n}$, für alle $i \in \{1, \ldots, n\}$, wobei $a_{n+1} := a_1$.

a) Man zeige: Es gilt $f(n) = (-1)^n + \sum_{i=0}^{n-1} (-1)^i \binom{n}{i} (n-1-i)!$, und folgere daraus $\lim_{n\to\infty} \frac{f(n)}{(n-1)!} = \frac{1}{e}$.

b) Für die n-te Derangement-Zahl zeige man: Es gilt $D_n = f(n) + f(n+1)$.

(21.13) Aufgabe: Zykellängen.
a) Für $n \in \mathbb{N}_0$ und $k \in \mathbb{N}$ sei $f_k(n) \in \mathbb{N}_0$ die Anzahl der Permutationen in $\mathcal{S}_n$, die keinen Zykel der Länge k enthalten. Man zeige: Es gilt $f_k(n) = n! \cdot \sum_{i=0}^{\lfloor \frac{n}{k} \rfloor} \frac{(-1)^i}{i! \cdot k^i}$, und folgere daraus $\lim_{n\to\infty} \frac{f_k(n)}{n!} = e^{-\frac{1}{k}}$.

b) Für $n \in \mathbb{N}_0$ sei $f(n) \in \mathbb{N}_0$ die Anzahl der Permutationen in $\mathcal{S}_n$, die einen Zykel der Länge $> \lfloor \frac{n}{2} \rfloor$ enthalten. Man zeige: Es gilt $\lim_{n\to\infty} \frac{f(n)}{n!} = \ln(2)$.

(21.14) Aufgabe: Numerierte Kisten.
In einem Raum sind 100 verschlossene Kisten fest montiert. Der Boden jeder Kiste trägt eine Zahl zwischen 1 und 100, wobei jede dieser Zahlen genau einmal vorkommt. Außerhalb des Raumes sind 100 Personen, numeriert von 1 bis 100. Ohne miteinander zu sprechen gehen die Personen nun einzeln in den Raum, öffnen jeweils genau 50 Kisten, schauen hinein, schließen sie und kommen heraus.

Die Wahrscheinlichkeit, daß eine Person zufällig die Kiste mit ihrer Nummer findet, ist $\frac{1}{2}$. Die Wahrscheinlichkeit, daß alle Personen jeweils ihre Nummer finden, ist also $(\frac{1}{2})^{100} \sim 10^{-30}$. Können die Personen eine Strategie vereinbaren, so daß die Wahrscheinlichkeit, daß alle Personen ihre Nummer finden, $> \frac{3}{10}$ ist?

22 Exercises for Part IV (in German)

(22.1) Aufgabe: Erzeugende Funktionen vom Dirichlet-Typ.
Für eine Folge $[f_n \in \mathbb{C}; n \in \mathbb{N}_0]$ heißt die formale Reihe der Form $f(s) := \sum_{n\geq 1} \frac{f_n}{n^s}$, wobei $s \in \mathbb{C}$, die zugehörige **Dirichlet-Reihe**. Dann kann $\mathcal{D} := \mathrm{Abb}(\mathbb{N}, \mathbb{C})$ als die Menge der Dirichlet-Reihen aufgefaßt werden, und ist ein $\mathbb{C}$-Vektorraum mit punktweiser Addition und Skalarmultiplikation. Auf $\mathcal{D}$ werde eine Multiplikation wie folgt definiert: Für $f = \sum_{n\geq 1} \frac{f_n}{n^s}$ und $g = \sum_{n\geq 1} \frac{g_n}{n^s}$ sei $(fg)(s) \in \mathcal{D}$ gegeben durch $(fg)(s) := \sum_{n\geq 1} (\sum_{d \mid n} \frac{f_d g_{\frac{n}{d}}}{d}) \frac{1}{n^s}$.

a) Man zeige: $\mathcal{D}$ ist eine kommutative $\mathbb{C}$-Algebra mit Eins $\frac{1}{1^s} + \sum_{n\geq 2} \frac{0}{n^s} \in \mathcal{D}$, und $f(s) = \sum_{n\geq 1} \frac{f_n}{n^s} \in \mathcal{D}$ ist genau dann invertierbar, wenn $f_1 \neq 0$ ist.

b) Es sei $\zeta(s) := \sum_{n\geq 1} \frac{1}{n^s} \in \mathcal{D}$ die **Riemann-Funktion**. Man zeige: Es ist $\zeta^{-1}(s) := \sum_{n\geq 1} \frac{\mu(n)}{n^s}$, wobei μ die zahlentheoretische Möbius-Funktion ist.

(22.2) Aufgabe: Harmonische Zahlen.
Für $n \in \mathbb{N}$ seien $h_n := \sum_{i=1}^{n} \frac{1}{i} \in \mathbb{Q}$ die n-te harmonische Zahl, und $h_0 := 0$.

a) Man bestimme die erzeugende Funktion $\sum_{n\geq 0} h_n X^n \in \mathbb{Q}[[X]]$.

b) Für $n \in \mathbb{N}_0$ berechne man $\sum_{k=0}^{n} h_k h_{n-k} \in \mathbb{Q}$.

(22.3) Aufgabe: Erzeugende Funktionen.
Es sei $[f_n \in \mathbb{Z}; n \in \mathbb{N}_0]$ die durch $f_0 := 1$ und $\sum_{k=0}^{n} f_k f_{n-k} = 1$, für $n \geq 1$, rekursiv definierte Folge. Man gebe eine geschlossene Formel für f_n an.

(22.4) Aufgabe: Linear rekursive Folgen.
Es seien $d \in \mathbb{N}$ sowie $q_1, \ldots, q_d \in \mathbb{C}$ mit $q_d \neq 0$, und $F := [f_n \in \mathbb{C}; n \in \mathbb{N}_0]$ die durch Startwerte $f_0, \ldots, f_{d-1} \in \mathbb{C}$ und $f_{n+d} + \sum_{i=1}^{d} q_i f_{n+d-i} = 0$, für alle $n \in \mathbb{N}_0$, definierte linear rekursive Folge.

a) Man zeige: Die Folge F kann eindeutig zu einer Folge $[f_n \in \mathbb{C}; n \in \mathbb{Z}]$ fortgesetzt werden, die die obige Rekursion für alle $n \in \mathbb{Z}$ erfüllt.

b) Für die erzeugende Funktion $f := \sum_{n\geq 0} f_n X^n \in \mathbb{Q}(X) \cap \mathbb{Q}[[X]]$ zeige man: Es gilt $\sum_{n\geq 1} f_{-n} X^n = -f(\frac{1}{X}) \in \mathbb{Q}(X)$. Gilt diese Identität auch in $\mathbb{Q}[[X]]$?

(22.5) Aufgabe: Rationale Funktionen.
Man zeige: Für $f := \sum_{n\geq 0} f_{n+1} X^n \in \mathbb{C}[[X]]$ sind äquivalent:

i) Es gibt $g \in \mathbb{C}(X) \cap \mathbb{C}[[X]]$ mit $f = \frac{\partial}{\partial X} g \cdot \frac{1}{g} \in \mathbb{C}(X)$.

ii) Es gilt $\exp(\sum_{n\geq 1} \frac{f_n}{n} X^n) \in \mathbb{C}(X) \cap \mathbb{C}[[X]]$.

iii) Es gibt nicht notwendig verschiedene $a_1, \ldots, a_k, b_1, \ldots, b_l \in \mathbb{C}$, für geeignete $k, l \in \mathbb{N}_0$, so daß $a_n = \sum_{i=1}^{k} a_i^n - \sum_{j=1}^{l} b_j^n$, für alle $\in \mathbb{N}$, gilt.

(22.6) Aufgabe: Quasipolynome.
Eine Abbildung $h \colon \mathbb{N}_0 \to \mathbb{C}$ heißt ein **Quasipolynom** vom Grad $d \in \mathbb{N}$ und **Quasiperiode** $k \in \mathbb{N}$, falls es für alle $a \in \{0, \ldots, k-1\}$ ein Polynom $h_a \in \mathbb{Q}[X]_{\leq d}$ gibt, so daß $h(n) = h_a(n)$ für alle $n \in \mathbb{N}_0$ mit $n \equiv a \pmod{k}$ gilt; dabei habe mindestens eines der Polynome h_a den Grad d. Man zeige die Äquivalenz der folgenden Aussagen:

i) Die Abbildung h ist ein Quasipolynom vom Grad $d-1$ und Quasiperiode k.

ii) Es ist $\sum_{n\geq 0} h(n) X^n = \frac{p}{q} \in \mathbb{C}(X) \cap \mathbb{C}[[X]]$, wobei $q \in \mathbb{C}[X]$ den Grad d und nur Nullstellen $a \in \mathbb{C}$ mit $a^k = 1$ hat, und $p \in \mathbb{C}[X]_{\leq d-1}$ teilerfremd zu q ist.

(22.7) Aufgabe: Rekursion.
Man gebe geschlossene Formeln für die wie folgt rekursiv definierten $a_n \in \mathbb{Z}$ an:

a) Es seien $a_0 := 2$ und $a_1 := 3$ sowie $a_{n+2} := 3a_{n+1} - 2a_n$ für $n \in \mathbb{N}_0$.

b) Es seien $a_0 := 0$ und $a_1 := 2$ sowie $a_{n+2} := 4a_{n+1} - 4a_n$ für $n \in \mathbb{N}_0$.

c) Es seien $a_0 := 5$ und $a_1 := 12$ sowie $a_{n+2} := 4a_{n+1} - 3a_n - 2^n$ für $n \in \mathbb{N}_0$.

(22.8) Aufgabe: Lineare Rekursion.
Für $n \in \mathbb{N}_0$ sei $r_n := (1+\sqrt{3})^n + (1-\sqrt{3})^n \in \mathbb{R}$. Man zeige: Es ist $r_n \in \mathbb{N}$ und es gilt $2^{\lfloor \frac{n}{2} \rfloor + 1} \mid r_n$.

(22.9) Aufgabe: Quadratwurzeln.
Es sei $1 \neq d \in \mathbb{N}$ quadratfrei.

a) Man zeige: Es gibt eindeutig bestimmte Folgen $[a_n \in \mathbb{Z}; n \in \mathbb{N}_0]$ und $[b_n \in \mathbb{Z}; n \in \mathbb{N}_0]$, so daß $(1+\sqrt{d})^n = a_n + b_n \cdot \sqrt{d} \in \mathbb{R}$ für alle $n \in \mathbb{N}_0$ gilt.

b) Man zeige: Die zugehörigen erzeugenden Funktionen $\sum_{n \geq 0} a_n X^n \in \mathbb{Q}[[X]]$ und $\sum_{n \geq 0} b_n X^n \in \mathbb{Q}[[X]]$ sind rationale Funktionen, und schreibe sie explizit als Quotient von Polynomen.

(22.10) Aufgabe: Drohnen und Königinnen.
Die Männchen der Honigbiene werden als Drohnen bezeichnet. Eine Drohne enwickelt sich aus einem unbefruchteten Ei, das ausschließlich Erbgut einer Bienenkönigin enhält, während eine Königin sich aus einem befruchteten Ei entwickelt, das das Erbgut einer Drohne und einer anderen Königin enthält. Wieviele Vorfahren n-ter Stufe (also Eltern, Großeltern, ...) hat eine Drohne?

(22.11) Aufgabe: Fibonacci-Zahlen.
Für $n \in \mathbb{N}_0$ sei $F_n \in \mathbb{N}_0$ die n-te Fibonacci-Zahl. Man zeige:

a) Für $n \in \mathbb{N}_0$ gelten $\sum_{k=0}^{n} F_k = F_{n+2} - 1$ und $\sum_{k=1}^{n} F_{2k-1} = F_{2n}$, sowie $\sum_{k=0}^{n} F_k^2 = F_n F_{n+1}$ und $\sum_{k=0}^{n} F_k F_{n-k} = \frac{1}{5}(2nF_{n+1} - (n+1)F_n)$.

b) Für $n \in \mathbb{N}$ gilt $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$.

c) Für $k \in \mathbb{N}$ und $n \in \mathbb{N}_0$ gelten $F_{n+k} = F_k F_{n+1} + F_{k-1} F_n$ und $F_n \mid F_{kn}$.

d) Man zeige: Jede natürliche Zahl $1 \neq m \in \mathbb{N}$ hat eine eindeutige Darstellung der Form $m = \sum_{n \geq 1} a_n F_n$, wobei $a_n \in \{0, 1\}$ und $a_n a_{n+1} = 0$ für alle $n \in \mathbb{N}$.

(22.12) Aufgabe: Lucas-Zahlen.
Die Folge $[L_n \in \mathbb{N}_0; n \in \mathbb{N}]$ der **Lucas-Zahlen** ist definiert durch $L_{n+2} := L_n + L_{n+1}$, für alle $n \in \mathbb{N}_0$, wobei $L_0 := 2$ und $L_1 := 1$.

a) Man gebe eine geschlossene Formel für L_n an.

b) Man zeige: Es gilt $L_n = F_{n-1} + F_{n+1}$, für alle $n \in \mathbb{N}$, und $F_{2n} = F_n L_n$, für alle $n \in \mathbb{N}_0$; dabei ist $F_n \in \mathbb{N}_0$ die n-te Fibonacci-Zahl.

c) An einem runden Tisch sitzen $n \geq 2$ Personen. Wieviele Teilmengen dieser Personen gibt es, unter denen sich keine zwei benachbarte Personen befinden?

(22.13) Aufgabe: Erzeugende Funktion der Fibonacci-Zahlen.
a) Für $n \in \mathbb{N}_0$ sei $F_n \in \mathbb{N}_0$ die n-te Fibonacci-Zahl. Man zeige: Die erzeugende Funktion $F' := \sum_{n\geq 0} F_{2n}X^n \in \mathbb{Q}[[X]]$ ist eine rationale Funktion, und schreibe sie explizit als Quotient von Polynomen.

b) Es sei $[g_n \in \mathbb{N}_0; n \in \mathbb{N}_0]$ die durch $g_0 := 1$ und $g_n := \sum_{i=0}^{n-1}(n-i)g_i$, für $n \geq 1$, rekursiv definierte Folge. Man zeige: Die erzeugende Funktion $\sum_{n\geq 0} g_n X^n \in \mathbb{Q}[[X]]$ ist eine rationale Funktion, schreibe sie explizit als Quotient von Polynomen, und gebe eine geschlossene Formel für g_n an.

(22.14) Aufgabe: Schachbrett und Domino-Steine.
Es sei $a_n \in \mathbb{N}_0$ die Anzahl der Überdeckungen eines $(3 \times n)$-Schachbrettes mit (2×1)-Domino-Steinen, für $n \in \mathbb{N}_0$. Man bestimme die zugehörige erzeugende Funktion, und gebe eine geschlossene Formel für a_n an.

Hinweis. Man betrachte auch die Anzahl der Überdeckungen, bei denen ein Eckfeld frei bleibt.

(22.15) Aufgabe: Ballot-Problem.
Vor einer Theaterkasse warten $2n$ Personen, für $n \in \mathbb{N}_0$, um jeweils eine Karte für 10 Euro zu kaufen. Von ihnen haben n Personen einen 10-Euro-Schein, die anderen nur einen 20-Euro-Schein Wechselgeld zur Verfügung. Wieviele Warteschlangen gibt es, so daß der Kassierer stets passend herausgeben kann, wenn er zu Beginn kein Wechselgeld zur Verfügung hat?

(22.16) Aufgabe: Dyck-Wege.
Für $n \in \mathbb{N}_0$ betrachte man ein quadratisches Gitter mit der Kantenlänge n. Ein kürzester Weg im Gitter von der linken unteren zur rechten oberen Ecke, der nie die Diagonale übertritt, heißt ein **Dyck-Weg**. Man bestimme die Anzahl der Dyck-Wege in Abhängigkeit von n.

(22.17) Aufgabe: Ebene binäre Bäume.
Ein **ebener binärer Baum** ist eine endliche Menge V von **Ecken**, zusammen mit einer **Wurzel** $v \in V$ und einer geordneten Partition $V \setminus \{v\} = V_1 \,\dot{\cup}\, V_2$, wobei die **Unterbäume** V_1 und V_2 wiederum ebene binäre Bäume sind. Ein ebener binärer Baum heißt **strikt** binär, falls entweder $V_1 = \emptyset = V_2$ oder $V_1 \neq \emptyset \neq V_2$ gilt, und dies auch für alle Unterbäume gilt. Ebene binäre Bäume mit Eckenmengen V bzw. W heißen isomorph, falls es eine Bijektion $\pi\colon V \to W$ gibt, die Isomorphismen $\pi\colon V_i \to W_i$, für $i \in \{1, 2\}$, induziert.

a) Man bestimme die Anzahl der Isomorphieklassen ebener binärer Bäume mit $n \in \mathbb{N}_0$ Ecken.

b) Man bestimme ebenso die Anzahl der Isomorphieklassen ebener strikt binärer Bäume mit $2n+1$ Ecken.

(22.18) Aufgabe: Triangulierung.
Auf wieviele Weisen kann man ein konvexes n-Eck, für $n \geq 3$, durch Einfügen von sich im Inneren nicht schneidenden Diagonalen in $n-2$ Dreiecke zerlegen?

(22.19) Aufgabe: Schröder-Problem.
Für $n \in \mathbb{N}$ ist eine **vollständige binäre Partition** einer n-elementigen Menge N wie folgt rekursiv definiert: Für $n \geq 2$ ist eine vollständige binäre Partition von N eine eine disjunkte Zerlegung $N = N' \mathbin{\dot\cup} N''$ mit $N' \neq \emptyset \neq N''$, zusammen mit jeweils einer vollständigen binären Partition von N' und N''; für $n = 1$ ist N die einzige vollständige binäre Partition von N.

Es seien $s(n) \in \mathbb{N}_0$ die Anzahl der vollständigen binären Partitionen einer n-elementigen Menge, und $s := \sum_{n\geq 1} \frac{s(n)}{n!} X^n \in \mathbb{Q}[[X]]$ die zugehörige exponentiell erzeugende Funktion. Man zeige: Es gilt $s(x) = 1 - \sqrt{1-2x}$, für alle $x \in \mathbb{C}$ in einer hinreichend kleinen offenen Umgebung von 0. Daraus bestimme man eine geschlossene Formel für $s(n)$.

(22.20) Aufgabe: Algebraische Funktionen.
a) Man zeige: Es gilt $\sum_{n\geq 0} \binom{2n}{n} x^n = \frac{1}{\sqrt{1-4x}}$, für alle $x \in \mathbb{C}$ in einer hinreichend kleinen offenen Umgebung von 0.

b) Man gebe eine analytische Beschreibung von $\sum_{n\geq 0} \binom{2n+1}{n} x^n$ an.

c) Man zeige: Für alle $n \in \mathbb{N}_0$ gilt $\sum_{k=0}^{n} \binom{2k}{k}\binom{2(n-k)}{n-k} = 4^n$.

(22.21) Aufgabe: Exponentiell erzeugende Funktionen.
Es sei $[g_n \in \mathbb{Z}; n \in \mathbb{N}_0]$ die durch $g_n + 2ng_{n-1} = \sum_{k=0}^{n} \binom{n}{k} g_k g_{n-k}$, für $n \geq 1$, sowie $g_0 := 0$ und $g_1 := 1$, rekursiv definierte Folge. Man gebe eine geschlossene Formel für g_n an.

(22.22) Aufgabe: Involutionen.
Für $n \in \mathbb{N}_0$ sei $i_n := |\{\pi \in \mathcal{S}_n; \pi^2 = 1\}|$ die Anzahl der **Involutionen** in $\mathcal{S}_n$.

a) Man zeige: Es gilt die Rekursion $i_{n+1} := i_n + n \cdot i_{n-1}$, für alle $n \in \mathbb{N}$.

b) Man zeige: Es gilt $\sum_{n\geq 0} \frac{i_n}{n!} X^n = \exp(X + \frac{X^2}{2}) \in \mathbb{Q}[[X]]$.

c) Man zeige: Für $n \in \mathbb{N}_0$ gilt $i_n = \sum_{k\geq 0} \binom{n}{2k} \cdot \frac{(2k)!}{2^k \cdot k!} \in \mathbb{Q}$.

d) Man gebe eine geschlossene Formel für $\sum_{k=0}^{n} (-1)^k \binom{n}{k} i_k \in \mathbb{Z}$ an.

Hinweis zu c). Man benutze $\exp(X + \frac{X^2}{2}) = \exp(X)\exp(\frac{X^2}{2})$.

(22.23) Aufgabe: Arrangements.
Es seien $n, k \in \mathbb{N}_0$, und $a_{n,k} \in \mathbb{N}_0$ die Anzahl der k-Arrangements einer n-elementigen Menge. Weiter sei $b_n := \sum_{k \geq 0} a_{n,k} \in \mathbb{N}_0$ die Anzahl aller Arrangements einer n-elementigen Menge. Für die zugehörigen exponentiell erzeugenden Funktionen zeige man:

a) Für $n \in \mathbb{N}_0$ gilt $\sum_{k \geq 0} \frac{a_{n,k}}{k!} X^k = (1+X)^n \in \mathbb{Q}[[X]]$.

b) Für $k \in \mathbb{N}_0$ gilt $\sum_{n \geq 0} \frac{a_{n,k}}{n!} X^n = X^k \cdot \exp \in \mathbb{Q}[[X]]$.

c) Es gilt $\sum_{n \geq 0} \frac{b_n}{n!} X^n = \frac{\exp}{1-X} \in \mathbb{Q}[[X]]$.

(22.24) Aufgabe: Mengen von Abbildungen.
Für $n \in \mathbb{N}_0$ sei $a_n \in \mathbb{N}_0$ die Anzahl aller Abbildungen $f \colon N \to N$, wobei $N := \{1, \ldots, n\}$, mit $f(N) = \{1, \ldots, i\}$ für ein $i \in N$.

a) Man zeige: Für die exponentiell erzeugende Funktion $\widetilde{A} := \sum_{n \geq 0} \frac{a_n}{n!} X^n \in \mathbb{Q}[[X]]$ gilt $\widetilde{A} = \frac{1}{2-\exp} \in \mathbb{Q}[[X]]$.

b) Man zeige: Es gilt $a_n = \sum_{k \geq 0} \frac{k^n}{2^{k+1}}$, für alle $n \in \mathbb{N}_0$.

(22.25) Aufgabe: Bernoulli-Zahlen.
Es sei $[B_n \in \mathbb{Q}; n \in \mathbb{N}_0]$ die durch $B_0 := 1$ und $\sum_{k=0}^{n} \binom{n+1}{k} B_k = 0$, für $n \in \mathbb{N}$, definierte Folge der **Bernoulli-Zahlen**.

a) Man zeige: Es gilt $\widetilde{B}(X) := \sum_{n \geq 0} \frac{B_n}{n!} X^n = \frac{X}{\exp(X)-1} \in \mathbb{Q}[[X]]$. Daraus folgere man: Es ist $\widetilde{B}(X) \in \mathbb{C}[[X]]^\infty$, und es gilt $B_{2n+1} = 0$ für alle $n \in \mathbb{N}_0$.

b) Es sei $b_n(T) := \sum_{k=0}^{n} \binom{n}{k} B_k T^{n-k} \in \mathbb{Q}[T]$ das n-te **Bernoulli-Polynom**, für $n \in \mathbb{N}_0$. Man zeige: Für $n \in \mathbb{N}_0$ gilt $\frac{\partial}{\partial T}(b_{n+1}(T)) = (n+1) \cdot b_n(T) \in \mathbb{Q}[T]$.

c) Für die exponentiell erzeugende Funktion $\widetilde{B}(T, X) := \sum_{n \geq 0} \frac{b_n(T)}{n!} X^n \in \mathbb{Q}[T][[X]]$ zeige man: Es gilt $\widetilde{B}(T, X) := \widetilde{B}(X) \cdot \exp(TX) = \frac{X \cdot \exp(TX)}{\exp(X)-1}$.

(22.26) Aufgabe: Erzeugende Funktion der Potenzsummen.
Für $k \geq 1$ sei $s_k \in \mathbb{Q}[T]$, so daß $s_k(n) = \sum_{i=0}^{n-1} i^k$ für alle $n \in \mathbb{N}_0$.

a) Für die exponentiell erzeugende Funktion $\widetilde{S}(T, X) := \sum_{k \geq 0} \frac{s_k(T)}{k!} X^k \in \mathbb{Q}[T][[X]]$ zeige man: Es gilt $\widetilde{S}(T, X) = \frac{\exp(TX)-1}{\exp(X)-1} \in \mathbb{Q}[T][[X]]$, und folglich $X \cdot \widetilde{S}(T, X) = \widetilde{B}(T, X) - \widetilde{B}(0, X) \in \mathbb{Q}[T][[X]]$; dabei ist $\widetilde{B}(T, X) \in \mathbb{Q}[T][[X]]$ die exponentiell erzeugende Funktion der Bernoulli-Polynome.

b) Daraus folgere man: Es gilt $s_k(T) = \frac{1}{k+1} \sum_{i=0}^{k} \binom{k+1}{i} B_i T^{k+1-i} \in \mathbb{Q}[T]$.

(22.27) Aufgabe: Konjugierte Partitionen.
Für eine Partition $\lambda = [\lambda_1, \lambda_2, \ldots] \vdash n \in \mathbb{N}_0$ und die zugehörige konjugierte Partition $\lambda' = [\lambda'_1, \lambda'_2, \ldots] \vdash n$ zeige man:

a) Es gilt $\sum_{i\geq 1}(i-1)\lambda_i = \sum_{i\geq 1}\binom{\lambda_i'^2}{2}$.

b) Es gelten $\sum_{i\geq 1}\lceil\frac{\lambda_{2i-1}}{2}\rceil = \sum_{i\geq 1}\lceil\frac{\lambda_{2i-1}}{2}\rceil$ und $\sum_{i\geq 1}\lfloor\frac{\lambda_{2i-1}}{2}\rfloor = \sum_{i\geq 1}\lceil\frac{\lambda_{2i}}{2}\rceil$ sowie $\sum_{i\geq 1}\lfloor\frac{\lambda_{2i}}{2}\rfloor = \sum_{i\geq 1}\lfloor\frac{\lambda_{2i}}{2}\rfloor$.

(22.28) Aufgabe: Durfee-Quadrat.
a) Man zeige: Es gilt die Identität

$$\prod_{i\geq 1}\frac{1}{1-YX^i} = \sum_{k\geq 0}\left(\prod_{i=1}^{k}\frac{1}{(1-X^i)(1-YX^i)}\right)X^{k^2}Y^k \in \mathbb{Q}[[X]][[Y]].$$

b) Man gebe einen kombinatorischen Beweis für die folgende Identität an:

$$\prod_{i\geq 1}(1+YX^{2i-1}) = \sum_{k\geq 0}\left(\prod_{i=1}^{k}\frac{1}{1-X^{2i}}\right)X^{k^2}Y^k \in \mathbb{Q}[[X]][[Y]]$$

(22.29) Aufgabe: Selbst-konjugierte Partitionen.
Für $n \in \mathbb{N}_0$ sei $e_n \in \mathbb{N}_0$ die Anzahl der selbst-konjugierten Partitionen von n in lauter gerade Teile. Man gebe eine geschlossene Formel für die erzeugende Funktion $\sum_{n\geq 0} e_n X^n \in \mathbb{Q}[[X]]$ an.

(22.30) Aufgabe: Erzeugende Funktionen für Partitionsanzahlen.
Für $n \in \mathbb{N}_0$ seien $f_n \in \mathbb{N}_0$ die Anzahl der Partitionen von n, in denen der Teil i höchstens i-mal vorkommt, für alle $i \in \{1,\ldots,n\}$, und $g_n \in \mathbb{N}_0$ die Anzahl der Partitionen von n, die keinen Teil der Form $i(i+1)$ haben, für ein $i \in \mathbb{N}$. Man zeige: Es gilt $f_n = g_n$, für alle $\in \mathbb{N}_0$, und gebe ein geschlossene Formel für die erzeugende Funktion $\sum_{n\geq 0} f_n X^n = \sum_{n\geq 0} g_n X^n \in \mathbb{Q}[[X]]$ an.

23 Exercises for Part V (in German)

(23.1) Aufgabe: Double Counting.
Für $n \in \mathbb{N}$ seien $\tau(n) := |\{i \in \mathbb{N}; i \mid n\}| \in \mathbb{N}$ die Anzahl der Teiler von n, und $\overline{\tau}(n) := \frac{1}{n}\cdot\sum_{i=1}^{n}\tau(i) \in \mathbb{Q}$. Man zeige: Es gilt $\lim_{n\to\infty}\frac{\overline{\tau}(n)}{\ln(n)} = 1$.

Hinweis. Man zeige $\lim_{n\to\infty}\frac{\overline{\tau}(n)}{h_n} = 1$, wobei $h_n := \sum_{i=1}^{n}\frac{1}{i} \in \mathbb{Q}$ die n-te harmonische Zahl bezeichne, und benutze $\lim_{n\to\infty}\frac{h_n}{\ln(n)} = 1$.

(23.2) Aufgabe: Parkers Lemma.
Es sei $G \leq \mathcal{S}_n$, wobei $n \in \mathbb{N}$. Für $k \in \mathbb{N}$ sei $\mathcal{C}_k \subseteq \mathcal{S}_n$ die Menge der k-Zykel, die in einem beliebigen Element von G vorkommen.

a) Man zeige: $\mathcal{S}_n$ operiert auf sich durch **Konjugation** $\mathcal{S}_n \to \mathcal{S}_n\colon \sigma \mapsto \pi\sigma\pi^{-1}$, für alle $\pi \in \mathcal{S}_n$. Daraus folgere man: G operiert auf $\mathcal{C}_k$ durch Konjugation.

b) Man zeige: Die G-Mengen $\mathcal{C}_1$ und $\{1, \ldots, n\}$ sind äquivalent. Was hat also die Aussage in (c) mit dem Cauchy-Frobenius-Burnside-Lemma zu tun?

c) Für die Anzahl der G-Bahnen auf $\mathcal{C}_k$ zeige man $|G \backslash \mathcal{C}_k| = \frac{1}{|G|} \cdot \sum_{\pi \in G} k a_k(\pi)$; dabei sei $a_k(\pi) \in \mathbb{N}_0$ die Anzahl der Teile k im Zykeltyp von π.

Hinweis zu c). Man bestimme den Stabilisator eines Zykels in $\mathcal{C}_k$, und betrachte eine geeignete $(|G| \times n)$-Matrix wie im Beweis des Cauchy-Frobenius-Burnside-Lemmas.

(23.3) Aufgabe: Türme auf dem Schachbrett.
Für $n \in \{1, \ldots, 5\}$ betrachte man ein $(n \times n)$-Schachbrett, dessen Vorder- und Rückseite nicht unterschieden werden können. Wieviele Muster gibt es, n Türme so zu verteilen, daß sie sich nicht gegenseitig schlagen können? Können Sie eine allgemeine Formel für diese Anzahl angeben?

(23.4) Aufgabe: Konstruktionen mit Permutationsgruppen.
Es seien $G \leq \mathcal{S}_n$ und $H \leq \mathcal{S}_m$ Permutationsgruppen, für $n, m \in \mathbb{N}$, mit zugehörigen Zykelindizes $c_G \in \mathbb{Q}[X_1, \ldots, X_n]$ bzw. $c_H \in \mathbb{Q}[X_1, \ldots, X_m]$. Man bestimme jeweils den Zykelindex für die natürliche Operation
a) des direkten Produkts $G \times H$ auf $n + m$ Punkten, und
b) des Kranzprodukts $G \wr H$ auf nm Punkten.

(23.5) Aufgabe: Zykelindex symmetrischer Gruppen.
Unter Benutzung des Zykelindexes für $\mathcal{S}_n$, für $n \in \mathbb{N}_0$, bestimme man (erneut) die exponentiell erzeugende Funktion $\sum_{n \geq 0} \frac{i_n}{n!} X^n \in \mathbb{Q}[[X]]$ der Anzahl $i_n := |\{\pi \in \mathcal{S}_n; \pi^2 = 1\}| \in \mathbb{N}$ der Involutionen in $\mathcal{S}_n$.

(23.6) Aufgabe: Zykelindex alternierender Gruppen.
Für $n \geq 2$ bestimme man den Zykelindex der alternierenden Gruppe $\mathcal{A}_n$.

Hinweis. Eine Permutation $\pi \in \mathcal{S}_n$ ist genau dann ungerade, wenn für ihren Zykeltyp $[n^{a_n(\pi)}, \ldots, 1^{a_1(\pi)}]$ gilt: $1 + (-1)^{\sum_{i \geq 1} a_{2i}(\pi)} = 0$.

(23.7) Aufgabe: Zykelindex von Diedergruppen.
a) Für den Zykelindex der **zyklischen Gruppe** $C_n := \langle (1, \ldots, n) \rangle \leq \mathcal{S}_n$, wobei $n \in \mathbb{N}$, zeige man: Es gilt $c_{C_n} = \frac{1}{n} \cdot \sum_{d \mid n} \varphi(d) X_d^{\frac{n}{d}} \in \mathbb{Q}[X_1, \ldots, X_n]$; dabei sei $\varphi(n) := |\{k \in \{0, \ldots, n-1\}; \mathrm{ggT}(k, n) = 1\}| \in \mathbb{N}$ die Euler-Funktion.

b) Für den Zykelindex $c_{D_{2n}} \in \mathbb{Q}[X_1, \ldots, X_n]$ der Diedergruppe $D_{2n} \leq \mathcal{S}_n$, wobei $n \geq 3$, zeige man: Es gilt

$$c_{D_{2n}} = \frac{1}{2} c_{C_n} + \begin{cases} \frac{1}{2} X_1 X_2^{\frac{n-1}{2}}, & \text{falls } n \text{ ungerade,} \\ \frac{1}{4}(X_2^{\frac{n}{2}} + X_1^2 X_2^{\frac{n}{2}-1}), & \text{falls } n \text{ gerade.} \end{cases}$$

(23.8) Aufgabe: Kleiner Satz von Fermat II.
Es seien $n \in \mathbb{N}$ und $p \in \mathbb{N}$ eine Primzahl. Man zeige: Es gilt $n^p \equiv n \pmod{p}$.

Hinweis. Man benutze den Zykelindex der zyklischen Gruppe C_p.

(23.9) Aufgabe: Würfel.
Man bestimme die Drehgruppe eines Würfels im Euklidischen Raum $\mathbb{R}^{3\times 1}$ jeweils als Gruppe von Permutationen seiner **i)** acht Ecken, **ii)** zwölf Kanten, **iii)** sechs Flächen, **iv)** vier Raumdiagonalen, und gebe die jeweiligen Zykelindizes an. Zu welcher bekannten Gruppe ist die Drehgruppe isomorph?

(23.10) Aufgabe: Tetraeder.
a) Man bestimme die Symmetriegruppe eines gleichseitigen Tetraeders im Euklidischen Raum $\mathbb{R}^{3\times 1}$ jeweils als Gruppe von Permutationen seiner **i)** vier Ecken, **ii)** sechs Kanten, **iii)** vier Flächen, und gebe die jeweiligen Zykelindizes an. Wieviele Drehungen gibt es? Zu welchen bekannten Gruppen sind die Symmetriegruppe und die Drehgruppe isomorph?

b) Man bestimme die Anzahl der verschiedenen Färbungen der vier Ecken des Tetraeders mit bis zu vier Farben, bezüglich der vollen Symmetriegruppe und der Drehgruppe. Wie kann man das Ergebnis geometrisch interpretieren?

(23.11) Aufgabe: Prisma.
Man bestimme die Symmetriegruppe eines geraden Prismas über einem gleichseitigen Dreieck im Euklidischen Raum $\mathbb{R}^{3\times 1}$ jeweils als Gruppe von Permutationen seiner **i)** sechs Ecken, **ii)** neun Kanten, **iii)** fünf Flächen, und gebe die jeweiligen Zykelindizes an. Wieviele Drehungen gibt es? Zu welchen bekannten Gruppen sind die Symmetriegruppe und die Drehgruppe isomorph?

(23.12) Aufgabe: Gewichtsindex.
Es seien G eine endliche Gruppe, die treu auf $N := \{1, \ldots, n\}$ operiere, und $K := \{1, \ldots, k\}$, wobei $n \in \mathbb{N}_0$ und $k \in \mathbb{N}_0$. Weiter seien $S \subseteq \mathrm{Inj}(N, K)$ eine Repräsentantenmenge für $G\backslash\mathrm{Inj}(N, K)$ und $e_n(Y_1, \ldots, Y_k) := \sum_{J\subseteq K, |J|=n} \prod_{j\in J} Y_j \in R := \mathbb{Z}[Y_1, \ldots, Y_k]$ das **elementar-symmetrische Polynom** vom Grad n.

Für den Gewichtsindex von $G\backslash\mathrm{Inj}(N, K)$ zeige man: Es gilt $\sum_{f\in S} \prod_{i=1}^n Y_{f(i)} = \frac{n!}{|G|} \cdot e_n(Y_1, \ldots, Y_k) \in R$. Welche bekannten Formeln erhält man daraus für die Spezialisierung $R \to \mathbb{Z}$: $Y_j \mapsto 1$, für alle $j \in K$, und $G = \{1\}$ bzw. $G = \mathcal{S}_N$?

(23.13) Aufgabe: Graphen.
Ein **(endlicher einfacher) Graph** Γ besteht aus einer endlichen Menge $\mathcal{V}$ von **Ecken** und einer Menge von 2-elementigen Teilmengen von $\mathcal{V}$, die **Kanten** genannt werden. Ein Graph heißt **zusammenhängend**, falls man von jeder Ecke auf einem **Weg**, also einer Kantenfolge, zu jeder anderen Ecke gelangen kann. Graphen Γ und Γ' heißen **isomorph**, falls es eine Bijektion zwischen den Ecken von Γ und Γ' gibt, die eine Bijektion der Kanten von Γ und Γ' bewirkt.

a) Es seien $r_n \in \mathbb{N}$ und $s_n \in \mathbb{N}$ die Anzahl der Isomorphieklassen von Graphen bzw. von zusammenhängenden Graphen mit $n \in \mathbb{N}$ Ecken. Man zeige: Für die erzeugenden Funktionen $r := \sum_{n\geq 1} r_n Y^n \in \mathbb{Q}[[Y]]$ und $s := \sum_{n\geq 1} s_n Y^n \in \mathbb{Q}[[Y]]$ gilt die Funktionalgleichung $r(Y) = \exp(\sum_{k\geq 1} \frac{s(Y^k)}{k}) \in \mathbb{Q}[[Y]]$.

b) Es sei $c^{(2)}_{\mathcal{S}_n} \in \mathbb{Q}[X_1, \ldots, X_{\binom{n}{2}}]$ der Zykelindex der Operation von $\mathcal{S}_n$ auf den 2-elementigen Teilmengen von $\{1, \ldots, n\}$. Man zeige: Die Anzahl $r_{n,k} \in \mathbb{N}_0$ der Isomorphieklassen von Graphen mit $n \in \mathbb{N}$ Ecken und $k \in \{0, \ldots, \binom{n}{2}\}$ Kanten ist der Koeffizient von Y^k im Polynom $c^{(2)}_{\mathcal{S}_n}(1+Y, \ldots, 1+Y^{\binom{n}{2}}) \in \mathbb{Q}[Y]$. Daraus folgere man: Es gilt $r_n = c^{(2)}_{\mathcal{S}_n}(2, \ldots, 2)$.

c) Man bestimme die Anzahl der Isomorphieklassen von Graphen bzw. von zusammenhängenden Graphen mit $n \in \{1, \ldots, 5\}$ Ecken und $k \in \{0, \ldots, \binom{n}{2}\}$ Kanten, und stelle die zugehörigen Graphen graphisch dar.

(23.14) Aufgabe: Bäume mit Wurzel.
Ein **Baum** ist ein zusammenhängender kreisfreier Graph; eine **Wurzel** ist eine ausgezeichnete Ecke. Zwei Bäume mit Wurzel heißen **isomorph**, falls es einen Graphenisomorphismus zwischen ihnen gibt, der die Wurzeln respektiert.

Es sei $t_n \in \mathbb{N}$ die Anzahl der Isomorphieklassen von Bäumen mit Wurzel auf $n \in \mathbb{N}$ Ecken. Man zeige: Die erzeugende Funktion $t := \sum_{n\geq 1} t_n Y^n \in \mathbb{Q}[[Y]]$ erfüllt die Funktionalgleichung $t(Y) = Y \cdot \exp(\sum_{k\geq 1} \frac{t(Y^k)}{k}) \in \mathbb{Q}[[Y]]$. Daraus bestimme man $t_1, \ldots, t_5$, und stelle die zugehörigen Bäume graphisch dar.

(23.15) Aufgabe: Selbst-duale Muster.
a) Es seien G eine endliche Gruppe, die treu auf der endlichen Menge N operiere, und c_G der zugehörige Zykelindex. Eine Abbildung $f \in \mathrm{Abb}(N, \{1, 2\})$ heißt **selbst-dual**, falls das zugehörige Muster sich bei der Umbenennung durch $\sigma := (1, 2) \in \mathcal{S}_2$ nicht ändert, das heißt, wenn $\sigma(f) \in {}^G f$ gilt. Man zeige:

Die Menge $\mathrm{Abb}(N, \{1, 2\})^+$ der selbst-dualen Abbildungen ist eine Vereinigung von G-Bahnen, und es gilt $|G\backslash\mathrm{Abb}(N, \{1, 2\})^+| = c_G(0, 2, 0, 2, \ldots) \in \mathbb{N}_0$.

b) Man bestimme die Anzahl selbst-dualer Halsketten mit n Perlen, die zwei verschiedene Farben haben können, für **i)** $n \geq 3$ ungerade, **ii)** $n = 2^k$ mit $k \geq 2$.

24 References

Textbooks

[1] M. Aigner: Diskrete Mathematik, 6. Aufl., Vieweg Studium Aufbaukurs Mathematik, 2006.

[2] M. Aigner: Discrete mathematics, translated from the 2004 German original by David Kramer, American Mathematical Society, 2007.

[3] M. Auslander, I. Reiten, S. Smalø: Representation theory of Artin algebras, corrected reprint of the 1995 original, Cambridge Studies in Advanced Mathematics 36, Cambridge University Press, 1997.

[4] G. Birkhoff: Lattice theory, 3rd ed., AMS Colloquium Publications 25, 1984.

[5] P. Cameron: Permutation groups, Cambridge University Press, 1999.

[6] M. Hall: Combinatorial theory, 2nd ed., Wiley, 1998.

[7] K. Jacobs, D. Jungnickel: Einführung in die Kombinatorik, 2. Aufl., de Gruyter, 2004.

[8] A. Kerber: Algebraic combinatorics via finite group actions, Bibliographisches Institut Mannheim, 1991.

[9] D. Marcus: Number fields, Springer, 1977.

[10] G. Polya, R. Tarjan, D. Woods: Notes on introductory combinatorics, Modern Birkhäuser Classics, 2010.

[11] R. Stanley: Enumerative combinatorics, vol. I, Cambridge Studies in Advanced Mathematics 49, Cambridge University Press, 1997.

[12] R. Stanley: Enumerative combinatorics, vol. II, Cambridge Studies in Advanced Mathematics 62, Cambridge University Press, 1999.

[13] A. Tucker: Applied combinatorics, 3rd ed., Wiley, 1995.

Research papers

[14] D. Benson, J. Conway: Diagrams for modular lattices, J. Pure Appl. Algebra 37, 1985, 111–116.

[15] B. Coleman, K. Hartshorn: Game, set, math, Math. Mag. 85, 2012, 83–96.

[16] M. Iovanov, G. Koffi: On incidence algebras and their representations, Preprint, 2017, arXiv:1702.03356 [math.RT].

[17] J. Klüners, G. Malle: A database for number fields, http://galoisdb.math.upb.de, 2011.

[18] K. Lux, J. Müller, M. Ringe: Peakword condensation and submodule lattices: an application of the MeatAxe, J. Symbolic Comput. 17, 1994, 529–544.

[19] J. Müller: On a theorem by Benson and Conway, J. Pure Appl. Algebra 208, 2007, 89–100.

[20] G. Polya: Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen, Acta Math. 68, 1937, 145–254.

[21] G. Rota: On the foundations of combinatorial theory, I. Theory of Möbius functions, Z. Wahrscheinlichkeit 2, 1964, 340–368.

[22] J. Serre: On a theorem of Jordan, Bull. Amer. Math. Soc. 40, 2003, 429–440.