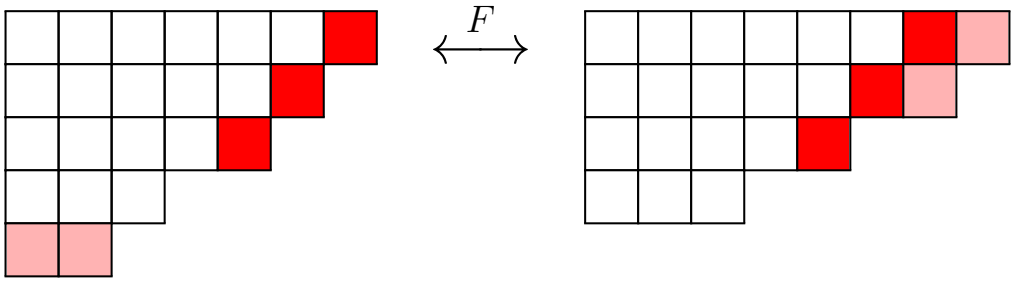
Combinatorics

Lecture notes winter semester 2018/19

Benjamin Sambale
Friedrich-Schiller-Universität Jena

Version: March 21, 2026

$$\prod_{k=1}^{\infty}(1-X^k)=\sum_{k=-\infty}^{\infty}(-1)^k X^{\frac{3k^2+k}{2}}$$



Contents

Preface 3

1. Finite Sets 3

2. Permutations and Partitions 10

3. Möbius Inversion 20

4. Power series 24

5. Generating Functions 34

6. Polynomials 49

7. Polynomials in several variables 58

8. Bernoulli Numbers 65

9. Catalan Numbers 71

10. Groups 74

11. Graphs 86

12. Exercises 93

A. GAP commands 101

Index 102

Warning: This is an AI-translated version of my German lectures notes, performed by *Gemini 3 Flash Preview*. I have not checked whether Gemini introduced errors. Use with care!

Preface

The present notes originated from a $3+1$ lecture in the winter semester 2018/19 (15 weeks) at the Friedrich Schiller University Jena and is primarily aimed at students of the following degree programs:

- Mathematics Teacher Training for Gymnasium
- B. Sc. Mathematics, Business Mathematics, Computer Science

Knowledge of Linear Algebra 1 and Analysis 1 is assumed. Some parts were not presented (especially the last chapter). In 2020 and 2021, extensive changes and additions were made, including the Ramanujan congruences and the Rogers-Ramanujan identities.

Literature:

- P. Tittmann, *Einführung in die Kombinatorik*, 2. Auflage, Springer Spektrum, Heidelberg, 2014, `https://link.springer.com/book/10.1007/978-3-642-54589-4`.
- R. P. Stanley, *Enumerative Combinatorics Vol. I, II*, 2nd edition, Cambridge University Press, Cambridge, 2012, `http://www-math.mit.edu/~rstan/ec/ec1.pdf`.
- G. E. Andrews, K. Eriksson, *Integer Partitions*, Cambridge University Press, Cambridge, 2004, `https://doi.org/10.1017/CBO9781139167239`

1. Finite Sets

Remark 1.1. Combinatorics is the study of counting discrete objects:

- (easy) The number of k-element subsets of an n-element set is $\binom{n}{k}$.
- (medium) The number of fixed-point-free permutations on $\{1,\dots,n\}$ is $[n!/e]$.
- (hard) The number of partitions of $5n+4$ is divisible by 5.
- (very hard) Every map can be colored with four colors such that adjacent countries have different colors.
- (open) How many magic squares of size 6×6 are there?

Definition 1.2.

- Empty set: $\varnothing$.
- Natural numbers: $\mathbb{N}=\{1,2,\dots\}$ $\mathbb{N}_0=\{0,1,\dots\}$.
- Prime numbers: $\mathbb{P}=\{2,3,5,7\dots\}$.
- Integers: $\mathbb{Z}=\{\dots,-1,0,1,\dots\}$.
- Rational numbers: $\mathbb{Q}=\{\frac{a}{b} : a,b\in\mathbb{Z},\ b\neq 0\}$.
- Real numbers: $\mathbb{R}$ (Analysis).

- Complex numbers: $\mathbb{C} = \{a + b\mathrm{i} : a, b \in \mathbb{R}\}$.
- For a set A, let $|A|$ be the cardinality of A. A is called *finite* if $|A| < \infty$ and *infinite* otherwise. We do not distinguish between cardinalities (countable, uncountable etc.) with the notation $|A| = \infty$. Two sets A and B are called *equinumerous* if a bijection $A \to B$ exists.
- If A_i $(i \in I)$ are sets, then so is their *Cartesian product* $\bigtimes_{i\in I} A_i = \{(a_i : i \in I) : a_i \in A_i\}$. In the case $A = A_i$ for all $i \in I$, we also write $A^I := \bigtimes_{i\in I} A$. For $I = \{1, \ldots, n\}$, we write $A_1 \times \ldots \times A_n$ and $A^n = A \times \ldots \times A$ (n factors).
- If A_i $(i \in I)$ are sets, then so is their *disjoint union*

$$\bigsqcup_{i\in I} A_i := \bigcup_{i\in I} \{(a, i) : a \in A_i\} \subseteq \Big(\bigcup_{i\in I} A_i\Big) \times I.$$

  For $I = \{1, \ldots, n\}$, we write $A_1 \sqcup \ldots \sqcup A_n$.
- For a set A, $2^A := \{B \subseteq A\}$ is the *power set* of A. For $k \in \mathbb{N}_0$, let

$$\binom{A}{k} := \{B \subseteq A : |B| = k\} \subseteq 2^A$$

  be the set of k-element subsets of A.

Remark 1.3. For sets A and I, one can identify A^I with the set of all mappings $I \to A$ by replacing $(a_i)_{i\in I} \in A^I$ with $f : I \to A$ where $f(i) := a_i$.

Theorem 1.4. *For finite sets $A, B, A_1, \ldots, A_n$, the following holds:*

(i) $|A_1 \times \ldots \times A_n| = |A_1| \ldots |A_n|$ *and* $|A^n| = |A|^n$.

(ii) $|A_1 \sqcup \ldots \sqcup A_n| = |A_1| + \ldots + |A_n|$.

(iii) *A and B are equinumerous if and only if* $|A| = |B|$.

(iv) $|2^A| = 2^{|A|}$.

Proof.

(i) For each element $(a_1, \ldots, a_n) \in A_1 \times \ldots \times A_n$ there are $|A_1|$ possibilities to choose a_1, $|A_2|$ possibilities for a_2 and so on. Conversely, each such choice yields exactly one element of $A_1 \times \ldots \times A_n$.

(ii) Each element in $A_1 \sqcup \ldots \sqcup A_n$ lies in exactly one of the sets $\{(a, i) : a \in A_i\}$. Here, $|\{(a, i) : a \in A_i\}| = |A_i|$ holds.

(iii) Let $A = \{a_1, \ldots, a_n\}$ and $f : A \to B$ be a bijection. Then $B = \{f(a_1), \ldots, f(a_n)\}$ with $f(a_i) \neq f(a_j)$ for $i \neq j$. This shows $|B| = n = |A|$. Conversely, let $|A| = |B|$ and $A = \{a_1, \ldots, a_n\}$ as well as $B = \{b_1, \ldots, b_n\}$. Then $f : A \to B$, $a_i \mapsto b_i$ is a bijection.

(iv) Let $A = \{a_1, \ldots, a_n\}$. For $M \subseteq A$ let $f(M) := (x_1, \ldots, x_n) \in \{0, 1\}^n$ with $x_i = 1 \iff a_i \in M$. Then $f : 2^A \to \{0, 1\}^n$, $M \mapsto f(M)$ is a bijection. According to (iii) and (i) it follows

$$|2^A| = |\{0, 1\}^n| = |\{0, 1\}|^n = 2^n = 2^{|A|}. \qquad \square$$

Definition 1.5.

- For $n \in \mathbb{N}_0$, $n! := \prod_{k=1}^{n} k$ is the *factorial* of n. Note: $0! = 1$ (empty product).

- For $a \in \mathbb{C}$ and $k \in \mathbb{N}_0$ one defines the *binomial coefficient*

$$\boxed{\binom{a}{k} := \frac{a(a-1)\dots(a-k+1)}{1 \cdot 2 \cdot \ldots \cdot k}.}$$

- For $n, k_1, \dots, k_s \in \mathbb{N}_0$ with $n = k_1 + \ldots + k_s$ let

$$\boxed{\binom{n}{k_1, \dots, k_s} := \frac{n!}{k_1! \dots k_s!}}$$

be the *multinomial coefficient* of n and $k_1, \dots, k_s$.

Remark 1.6. It holds that $\binom{a}{0} = 1$ (empty product) and $\binom{n}{k} = 0$ for $k > n \in \mathbb{N}_0$. For $k \leq n \in \mathbb{N}_0$ it holds that

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k} = \binom{n}{k, n-k}$$

and

$$\boxed{\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.}$$

One can therefore calculate the binomial coefficients with the *Pascal's triangle*:

$$\begin{array}{ccccccccccccccccccccc}
 & & & & & & \binom{0}{0} & & & & & & & & 1 & & & & & & \\
 & & & & \binom{1}{0} & & & & \binom{1}{1} & & & & & & 1 & & 1 & & & & \\
 & & & \binom{2}{0} & & & \binom{2}{1} & & & & \binom{2}{2} & & & & 1 & & 2 & & 1 & & \\
 & & \binom{3}{0} & & & & \binom{3}{1} & & & & \binom{3}{2} & & & & \binom{3}{3} & & 1 & & 3 & & 3 & & 1 \\
\binom{4}{0} & & & & \binom{4}{1} & & & & \binom{4}{2} & & & & \binom{4}{3} & & & & \binom{4}{4} & 1 & & 4 & & 6 & & 4 & & 1
\end{array}$$

Remark 1.7 ("Variation with repetition")**.** For finite sets A and B, there exist exactly $|B^A| = |B|^{|A|}$ maps $A \to B$ according to Theorem 1.4.

Example 1.8. A 4-digit combination lock has $10^4 = 10,000$ possible states (choose $A = \{1, 2, 3, 4\}$ and $B = \{0, 1, \dots, 9\}$ in Remark 1.7). If a thief checks one state per second, he needs on average approx. 83 minutes to crack the lock.

Theorem 1.9 ("Variation without repetition")**.** *For finite sets A and B, there exist exactly $\binom{|B|}{|A|}|A|!$ injective maps $A \to B$.*

Proof. In the case $|A| > |B|$, there are no injective maps $A \to B$ and indeed $\binom{|B|}{|A|} = 0$. Now let $k := |A| \leq |B| =: n$ and $A = \{a_1, \dots, a_k\}$. For every injective map $f : A \to B$, there are n possibilities for $f(a_1)$. Once $f(a_1)$ is fixed, there remain $n-1$ possibilities for $f(a_2) \in B \setminus \{f(a_1)\}$ etc. The number of injective maps is thus $n(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!} = \binom{n}{k}k!$. □

Example 1.10 (Birthday Paradox)**.** For persons $P_1, \ldots, P_n$ we consider the mapping $f : \{1, \ldots, n\} \to \{1, \ldots, 365\}$, which maps i to the birthday of P_i (leap years, twins etc. neglected). According to Remark 1.7 there are 365^n such mappings, of which $\binom{365}{n} n!$ are injective. The probability that at least two persons have their birthday on the same day is therefore

$$1 - \binom{365}{n} \frac{n!}{365^n}$$

(Laplace formula). For $n = 23$ one already obtains $> 50\%$.

Remark 1.11.

(i) The case $|A| > |B|$ in Theorem 1.9 yields the *Dirichlet pigeonhole principle*: If one distributes n objects into $k < n$ drawers, then at least one drawer must contain several objects. Example: In Leipzig there are two persons with the same number of hairs on their head (nobody has more hairs than Leipzig has inhabitants ($> 500,000$)).

(ii) In the case $|A| = |B|$, every injective mapping $A \to B$ is also bijective (provided $|A| < \infty$). Bijections $A \to A$ are called *permutations* on A. As is well known, the permutations on A form the *symmetric group* $\mathrm{Sym}(A)$ with respect to composition of mappings. The neutral element is id_A and the inverse to $f \in \mathrm{Sym}(A)$ is the inverse mapping f^{-1}. We set $S_n := \mathrm{Sym}(\{1, \ldots, n\})$. According to Theorem 1.9

$$|\mathrm{Sym}(A)| = |S_{|A|}| = \binom{|A|}{|A|} |A|! = |A|!.$$

Example 1.12.

$$S_3 := \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}.$$

Theorem 1.13 ("Combination without repetition")**.** *For every finite set A and $k \in \mathbb{N}_0$ holds*

$$\boxed{\left| \binom{A}{k} \right| = \binom{|A|}{k}.}$$

Proof. Let A_k be the set of injective mappings $\{1, \ldots, k\} \to A$. Then the mapping

$$\begin{aligned} F : A_k &\to \binom{A}{k}, \\ f &\mapsto \{f(1), \ldots, f(k)\} \end{aligned}$$

is surjective. For $\sigma \in S_k$ holds $F(f \circ \sigma) = \{f(\sigma(1)), \ldots, f(\sigma(k))\} = F(f)$. For $B \in \binom{A}{k}$ even $F^{-1}(B) = \{f \circ \sigma : \sigma \in S_k\}$, where $f \in A_k$ is a fixed preimage of B under F. In particular, every $B \in \binom{A}{k}$ has exactly $|S_k| = k!$ preimages. It follows $\left|\binom{A}{k}\right| = \frac{|A_k|}{k!} = \binom{|A|}{k}$ according to Theorem 1.9. □

Example 1.14. In the lottery "6 out of 49" there are $\binom{49}{6} = 13,983,816$ possibilities and the probability for a 4-match is

$$\frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}} = \frac{645}{665896} \approx 0.1\%.$$

Remark 1.15.

(i) Theorem 1.13 provides a combinatorial interpretation of the identity $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$: For $a \in A$ there are exactly $\binom{|A\setminus\{a\}|}{k-1}$ subsets $B \in \binom{A}{k}$ that contain a and $\binom{|A\setminus\{a\}|}{k}$ subsets $B \in \binom{A}{k}$ that do not contain a.

(ii) According to Theorem 1.4 and Theorem 1.13

$$2^n = |2^{\{1,\dots,n\}}| = \sum_{k=0}^{n} \binom{n}{k}.$$

This is a special case of the well-known *binomial theorem* (set $a = b = 1$)

$$\boxed{(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}} \qquad (a, b \in \mathbb{R}).$$

Theorem 1.16 (VANDERMONDE identity)**.** *For $n, a_1, \dots, a_n \in \mathbb{N}$ and $k \in \mathbb{N}_0$ we have*

$$\boxed{\binom{a_1 + \dots + a_n}{k} = \sum_{\substack{(k_1,\dots,k_n)\in\mathbb{N}_0^n \\ k_1+\dots+k_n=k}} \binom{a_1}{k_1} \cdots \binom{a_n}{k_n}.}$$

Proof. Let $A_1, \dots, A_n$ be sets with $|A_i| = a_i$ for $i = 1, \dots, n$. We determine $\left|\binom{A_1 \sqcup \dots \sqcup A_n}{k}\right|$ in two ways. According to Theorem 1.13, on the one hand

$$\left|\binom{A_1 \sqcup \dots \sqcup A_n}{k}\right| = \binom{|A_1 \sqcup \dots \sqcup A_n|}{k} = \binom{a_1 + \dots + a_n}{k}.$$

On the other hand, every k-element subset of $A_1 \sqcup \dots \sqcup A_n$ is composed of k_i-element subsets of A_i for $i = 1, \dots, n$ and $k_1 + \dots + k_n = k$. For each of these subsets there are $\left|\binom{A_i}{k_i}\right| = \binom{a_i}{k_i}$ possibilities. This shows

$$\left|\binom{A_1 \sqcup \dots \sqcup A_n}{k}\right| = \sum_{\substack{(k_1,\dots,k_n)\in\mathbb{N}_0^n \\ k_1+\dots+k_n=k}} \binom{a_1}{k_1} \cdots \binom{a_n}{k_n}. \qquad \square$$

Example 1.17. The special case $n = 2$ and $a_1 = a_2 = k$ in Theorem 1.16 yields

$$\binom{2k}{k} = \sum_{i=0}^{k} \binom{k}{i}\binom{k}{k-i} = \sum_{i=0}^{k} \binom{k}{i}^2.$$

Theorem 1.18 ("Variation with repetition" II)**.** *Let $A = \{a_1, \dots, a_n\}$ and B be finite sets and $k_1, \dots, k_n \in \mathbb{N}_0$ with $|B| = k_1 + \dots + k_n$. Then there exist exactly $\binom{|B|}{k_1,\dots,k_n}$ mappings $f : B \to A$ with $|f^{-1}(a_i)| = k_i$ for $i = 1, \dots, n$.*

Proof. Let $|B| = k$ and $f : B \to A$ with $|f^{-1}(a_i)| = k_i$ for $i = 1, \dots, n$. According to Theorem 1.13 there are $\binom{k}{k_1}$ possibilities for $f^{-1}(a_1)$. Once $f^{-1}(a_1)$ is fixed, there remain $\binom{k-k_1}{k_2}$ possibilities for $f^{-1}(a_2)$ etc. Thus there are

$$\binom{k}{k_1}\binom{k-k_1}{k_2} \cdots \binom{k-k_1-\dots-k_{n-1}}{k_n} = \frac{k!(k-k_1)!\dots(k-k_1-\dots-k_{n-1})!}{k_1!(k-k_1)!k_2!(k-k_1-k_2)!\dots k_n!} = \binom{k}{k_1,\dots,k_n}$$

possibilities for f. $\square$

Example 1.19.

(i) An *anagram* is a permutation of the letters of a word. According to Theorem 1.18 there are $\binom{5}{2,1,1,1} = 60$ anagrams of EULER (choose $A = \{E, U, L, R\}$, $B = \{1, 2, 3, 4, 5\}$, $k_1 = 2$, $k_2 = k_3 = k_4 = 1$). For example REUEL, LUREE etc.

(ii) There are $\binom{32}{10,10,10,2} = 2.753.294.408.504.640$ possibilities to distribute 32 skat cards to three players.

Remark 1.20. According to Remark 1.7 and Theorem 1.18

$$n^k = |\{1, \ldots, n\}^k| = \sum_{\substack{(k_1,\ldots,k_n)\in\mathbb{N}_0^n \\ k_1+\ldots+k_n=k}} \binom{k}{k_1, \ldots, k_n}.$$

This is a special case of the *multinomial theorem*

$$\boxed{(a_1 + \ldots + a_n)^k = \sum_{\substack{(k_1,\ldots,k_n)\in\mathbb{N}_0^n \\ k_1+\ldots+k_n=k}} \binom{k}{k_1, \ldots, k_n} a_1^{k_1} \ldots a_n^{k_n}} \qquad (a_1, \ldots, a_n \in \mathbb{R})$$

(Exercise 7). For $n = 2$ one obtains the binomial theorem.

Definition 1.21. For an arbitrary set A, the elements from $\mathbb{N}_0^A$ are called *multisets* over A. One can interpret a multiset $m := (n_a)_{a\in A}$ as a "subset" of A, where each $a \in A$ occurs exactly n_a times (in the case $n_a \leq 1$ for all $a \in A$, m is thus a true set). Accordingly, one sets $|m| := \sum_{a\in A} n_a$. We will often denote multisets in the form $\{a, a, b, c, c, c, \ldots\}$, where, as with sets, the order does not matter.

Theorem 1.22 ("Combination with replacement")**.** *An n-element set possesses exactly*

$$\boxed{\left(\binom{n}{k}\right) := \binom{n+k-1}{k}}$$

many k-element multisets ($n, k \in \mathbb{N}_0$).

Proof. Wlog. let $A = \{1, \ldots, n\}$. One can then identify the k-element multisets over A with the tuples $(a_1, \ldots, a_k) \in A^k$ with $a_1 \leq \ldots \leq a_k$. Let A_k be the set of these k-tuples. Obviously, then

$$f : A_k \to \binom{\{1, \ldots, n+k-1\}}{k},$$
$$(a_1, \ldots, a_k) \mapsto \{a_1, a_2 + 1, \ldots, a_k + k - 1\}$$

is bijective. From Theorem 1.13 it follows that $|A_k| = |f(A_k)| = \binom{n+k-1}{k}$. □

Example 1.23. When throwing three identical dice simultaneously, there are $\left(\binom{6}{3}\right) = \binom{8}{3} = 56$ possible events, which, however, are not all equally probable.

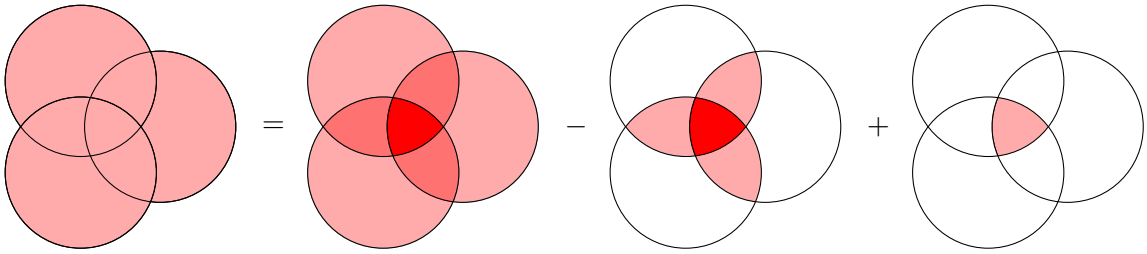
Remark 1.24.

(i) For $1 \leq k \leq n$, it holds that

$$\left(\binom{n+1}{k}\right) = \binom{n+k}{k} = \binom{n+k-1}{k-1} + \binom{n+k-1}{k} = \left(\binom{n+1}{k-1}\right) + \left(\binom{n}{k}\right).$$

(ii) For finite sets A and B, it is well known that $|A \cup B| = |A| + |B| - |A \cap B|$. Apparently, it also holds that

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$



This can be generalized as follows.

Theorem 1.25 (Inclusion-Exclusion Principle)**.** *For finite sets $A_1, \ldots, A_n$, it holds that*

$$\boxed{|A_1 \cup \ldots \cup A_n| = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < \ldots < i_k \le n} |A_{i_1} \cap \ldots \cap A_{i_k}|.}$$

Proof. We count how many times an element $a \in A_1 \cup \ldots \cup A_n$ is taken into account on the right side. For this, let wlog. $a \in A_1 \cap \ldots \cap A_l$ and $a \notin A_i$ for $i > l$. Then a is counted if and only if $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, l\}$ holds. In the k-th summand, a is thus counted $(-1)^{k+1}\binom{l}{k}$ times. In total, a is counted on the right side exactly

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

times. This shows the claim. □

Definition 1.26. As usual, $a, b \in \mathbb{N}$ are called *coprime*, if 1 is the only common positive divisor of a and b, i. e. $\gcd(a, b) = 1$. One calls

$$\varphi(n) := |\{1 \le a \le n : \gcd(a, n) = 1\}| \qquad (n \in \mathbb{N})$$

the *Euler φ-function*.

Theorem 1.27. *Let $n = p_1^{a_1} \ldots p_k^{a_k}$ be the prime factorization of $n \in \mathbb{N}$. Then*

$$\boxed{\varphi(n) = \prod_{i=1}^{k} (p_i^{a_i} - p^{a_i - 1}).}$$

Proof. For $i = 1, \ldots, k$ let $A_i := \{1 \le a \le n : p_i \mid a\}$. Then $A := \{1 \le a \le n : \gcd(a, n) \ne 1\} = A_1 \cup \ldots \cup A_k$. For $1 \le i_1 < \ldots < i_l \le k$ is

$$A_{i_1} \cap \ldots \cap A_{i_l} = \left\{ j p_{i_1} \ldots p_{i_l} : j = 1, \ldots, \frac{n}{p_{i_1} \ldots p_{i_l}} \right\}.$$

With Theorem 1.25 it follows

$$\varphi(n) = |\{1,\dots,n\} \setminus A\}| = n - |A| = n + \sum_{l=1}^{k} (-1)^l \sum_{1 \le i_1 < \dots < i_l \le k} \frac{n}{p_{i_1} \cdots p_{i_l}}$$
$$= n\Big(1 - \frac{1}{p_1}\Big) \dots \Big(1 - \frac{1}{p_k}\Big) = (p_1^{a_1} - p_1^{a_1 - 1}) \dots (p_k^{a_k} - p_k^{a_k - 1}). \qquad \square$$

Remark 1.28. In algebra, one proves Theorem 1.27 using the Chinese Remainder Theorem.

Theorem 1.29 (Erdős-Szekeres)**.** *Every real sequence of pairwise distinct numbers $a_1, \dots, a_{n^2+1}$ with $n \in \mathbb{N}$ possesses a monotone subsequence of length $n+1$.*

Proof. For $i = 1, \dots, n^2 + 1$ let α_i (resp. β_i) be the maximal length of a monotonically increasing (resp. decreasing) subsequence that ends with a_i. Let us assume that all monotone subsequences have length $\le n$. Then there are at most n^2 pairs (α_i, β_i). By the pigeonhole principle, there exist $1 \le i < j \le n^2 + 1$ with $(\alpha_i, \beta_i) = (\beta_j, \beta_j)$. In the case $a_i < a_j$ (resp. $a_i > a_j$), one can extend the monotonically increasing (resp. decreasing) sequence ending with a_i by a_j. But then $\alpha_i < \alpha_j$ or $\beta_i < \beta_j$. Contradiction. $\square$

Example 1.30. The sequence

$$n, n-1, \dots, 1, 2n, 2n-1, \dots, n+1, \dots, n^2, n^2-1, \dots, n^2-n+1$$

of length n^2 possesses no monotone subsequence of length $n+1$.

2. Permutations and Partitions

Definition 2.1.

- Let A be a set and $\sigma \in \mathrm{Sym}(A)$. One calls $a \in A$ a *fixed point* of σ if $\sigma(a) = a$. If σ has no fixed points, then σ is called *fixed-point-free*.
- For $x \in \mathbb{R}$ let $[x] \in \mathbb{Z}$ with $|x - [x]| < \frac{1}{2}$ or $[x] = x + \frac{1}{2}$ ("rounding").

Theorem 2.2 (Montmort)**.** *The number of fixed-point-free permutations in S_n is $[n!/e]$, where e is Euler's number.*

Proof. For $i = 1, \dots, n$ let $F_i := \{\sigma \in S_n : \sigma(i) = i\}$. The number f_n of fixed-point-free permutations of S_n is then $f_n = |S_n \setminus (F_1 \cup \dots \cup F_n)| = n! - |F_1 \cup \dots \cup F_n|$. For $1 \le i_1 < \dots < i_k \le n$ is

$$|F_{i_1} \cap \dots \cap F_{i_k}| = |\mathrm{Sym}(\{1,\dots,n\} \setminus \{i_1, \dots, i_k\})| = (n-k)!.$$

Theorem 1.25 shows

$$f_n = n! + \sum_{k=1}^{n} (-1)^k \sum_{1 \le i_1 < \dots < i_k \le n} (n-k)! = n! + \sum_{k=1}^{n} (-1)^k \binom{n}{k} (n-k)! = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Now

$$\Big|\frac{n!}{e} - f_n\Big| = \Big|n! \sum_{k=n+1}^{\infty} \frac{(-1)^k}{k!}\Big| = \frac{1}{n+1} - \frac{1}{(n+1)(n+2)} \pm \dots < \frac{1}{n+1} \le \frac{1}{2}$$

and $f_n = [n!/e]$. $\square$

Example 2.3.

(i) The fixed-point-free permutations of S_4 are

$$\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix},\begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix},\begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix},\begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix},\begin{pmatrix}1&2&3&4\\2&4&1&3\end{pmatrix},$$
$$\begin{pmatrix}1&2&3&4\\3&4&2&1\end{pmatrix},\begin{pmatrix}1&2&3&4\\3&1&4&2\end{pmatrix},\begin{pmatrix}1&2&3&4\\4&3&1&2\end{pmatrix},\begin{pmatrix}1&2&3&4\\4&1&2&3\end{pmatrix}.$$

(ii) In the pre-Christmas "Secret Santa" (Wichteln), n people give each other gifts by drawing lots beforehand that indicate to whom the gift is to be directed. This describes a permutation on $\{1,\dots,n\}$, which is fixed-point-free if and only if no person draws their own lot. The probability that at least one person draws their own lot is therefore $1-\frac{[n!/e]}{n!}\approx 1-\frac{1}{e}\approx 63\%$.

(iii) The encryption machine *Enigma* used in World War II permutes the 26 letters of the Latin alphabet. To supposedly make the encryption more secure, only fixed-point-free permutations were used. However, this was a crucial weakness that allowed the Allies to decrypt the Enigma.[1]

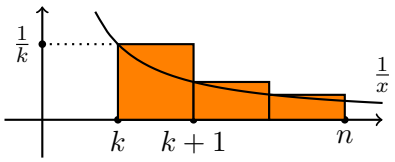
Example 2.4 (Secretary Problem)**.** There are n applicants for a vacant position invited one after another for an interview. Immediately after each interview, the applicant must be informed whether they have been hired or rejected. In the first case, the process is finished and no further applicants are considered. With which strategy does one find the best possible applicant?

One first rejects the first $k<n$ applicants consistently and chooses among the remaining $n-k$ the first one who is better than the first k applicants (possibly one has to reject all applicants, in which case the strategy has failed). The order of the applicants describes a permutation $\sigma\in S_n$, where $\sigma(1)$ is the position of the best applicant and $\sigma(2)$ is the position of the second best, etc. Let

$$m:=\min\{i\le n:\sigma(i)<\sigma(1)\}.$$

The above strategy finds the best applicant if and only if $\sigma(1)>k$ and $\sigma(m)\le k$ holds. The probability that $\sigma(1)$ is at position l is $1/n$. The probability for $\sigma(m)\le k$ is then $\frac{k}{l-1}$. The success probability of the strategy is therefore

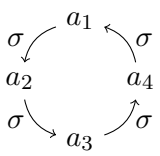
$$\sum_{l=k+1}^{n}\frac{1}{n}\frac{k}{l-1}=\frac{k}{n}\sum_{l=k}^{n-1}\frac{1}{l}\ge\frac{k}{n}\int_k^n\frac{1}{x}\,\mathrm{d}x=\frac{k}{n}(\log n-\log k).$$



The function $f(x)=\frac{x}{n}(\log n-\log x)$ has derivative $f'(x)=\frac{1}{n}(\log n-\log x-1)$ and therefore takes its maximum at $x=n/e$. For $k=[n/e]$, the success probability is thus approx. $f(n/e)=1/e\approx 37\%$ (for "large" n). One can show that this is the best strategy. For $n=20$, this results in $k=7$ and approx. 38%.

Definition 2.5. For a set A, one calls $\sigma\in\mathrm{Sym}(A)$ a (k-)*cycle* (or cycle of *length* k), if pairwise distinct $a_1,\dots,a_k\in A$ exist, such that

$$\sigma(x)=\begin{cases}a_{i+1} & \text{if } x=a_i \text{ with } i<k,\\ a_1 & \text{if } x=a_k,\\ x & \text{otherwise.}\end{cases}$$



[1] https://en.wikipedia.org/wiki/Cryptanalysis_of_the_Enigma#The_Enigma_machine

One then writes $\sigma = (a_1, \ldots, a_k)$. This notation is unique up to "rotation", i. e.

$$\sigma = (a_2, \ldots, a_k, a_1) = \ldots = (a_k, a_1, \ldots, a_{k-1}).$$

The only 1-cycle is id_A. To make formulations uniform, we will nevertheless formally distinguish the 1-cycles $(1), (2), \ldots, (n)$. Furthermore, we regard id_A as the product of all 1-cycles. Cycles of length 2 are called *transpositions*. Cycles $\sigma = (a_1, \ldots, a_k)$ and $\tau = (b_1, \ldots, b_l)$ are called *disjoint*, if

$$\{a_1, \ldots, a_k\} \cap \{b_1, \ldots, b_l\} = \varnothing.$$

Remark 2.6.

(i) It holds that $(a_1, \ldots, a_k)^{-1} = (a_k, a_{k-1}, \ldots, a_1)$.

(ii) Disjoint cycles $\sigma, \tau \in \mathrm{Sym}(A)$ commute, i. e. $\sigma \circ \tau = \tau \circ \sigma$. In the following, we will often omit the composition symbol $\circ$.

Lemma 2.7. *Every permutation σ of a finite set A is a composition of pairwise disjoint cycles $\sigma = \sigma_1 \ldots \sigma_k$ of length > 1 and these are uniquely determined up to their order.*

Proof. Existence: Let $A_\sigma := \{a \in A : \sigma(a) \neq a\}$. We argue by induction on $|A_\sigma|$. In the case $A_\sigma = \varnothing$, $\sigma = \mathrm{id}_A$ is the empty product. So let $a \in A_\sigma \neq \varnothing$. Because $|A_\sigma| \leq |A| < \infty$, the elements $a, \sigma(a), \sigma^2(a), \ldots \in A_\sigma$ cannot all be distinct. So let $0 \leq k < l$ with $\sigma^k(a) = \sigma^l(a)$. Then $\sigma^{l-k}(a) = a$. Let $s \in \mathbb{N}$ be minimal with $\sigma^s(a) = a$. Then $a, \sigma(a), \ldots, \sigma^{s-1}(a)$ are pairwise distinct and $\sigma_1 = (a, \sigma(a), \ldots, \sigma^{s-1}(a))$ is an s-cycle with $s > 1$. For $\tau := \sigma_1^{-1}\sigma \in \mathrm{Sym}(A_\sigma)$ and $i = 0, \ldots, s-1$, it then holds that

$$\tau(\sigma^i(a)) = \sigma_1^{-1}\sigma^{i+1}(a) = \sigma^i(a).$$

This shows $A_\tau = A_\sigma \setminus A_{\sigma_1}$. By induction, there exist pairwise disjoint cycles $\sigma_2, \ldots, \sigma_k \in \mathrm{Sym}(A_\tau)$ with length > 1 and $\tau = \sigma_2 \ldots \sigma_k$. Obviously, $\sigma_1, \ldots, \sigma_k$ are also pairwise disjoint and $\sigma = \sigma_1 \ldots \sigma_k$.

Uniqueness: Let $\sigma = \sigma_1 \ldots \sigma_k = \tau_1 \ldots \tau_l$ be two representations with pairwise disjoint cycles $\sigma_1, \ldots, \sigma_k$ as well as $\tau_1, \ldots, \tau_l$. Let $a \in A$ with $\sigma_1(a) \neq a$. Then there exists exactly one τ_i with $\tau_i(a) = \sigma_1(a)$. Furthermore, $\sigma_1^2(a) = \tau_i^2(a)$ and so on. This shows $\sigma_1 = \tau_i$. By multiplying both sides by σ_1^{-1}, one obtains $\sigma_2 \ldots \sigma_k = \tau_1 \ldots \tau_{i-1}\tau_{i+1} \ldots \tau_l$. The claim now follows easily by induction on k. □

Remark 2.8.

(i) One can make the notation in disjoint cycles

$$\sigma = (a_1, \ldots, a_s)(b_1, \ldots, b_t) \ldots$$

completely unique by requiring $a_1 = \min\{a_1, \ldots, a_s\} < b_1 = \min\{b_1, \ldots, b_t\} < \ldots$. This is implemented in the computer algebra system GAP.

(ii) In the following, we say that $\sigma \in S_n$ *contains* a cycle τ if τ occurs in the disjoint cycle representation. In doing so, we want to count the fixed points as 1-cycles.

(iii) As is well known (Linear Algebra), every permutation can also be written as a product of transpositions, although these are generally not disjoint.

Example 2.9.

(i) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 6 & 2 & 5 & 3 \end{pmatrix} = (1,4,2)(3,6)$.

(ii) $(2,5,3,1)(3,1,6) = (1,6)(2,5,3)$ (maps are evaluated from right to left).

(iii) $S_3 = \{(), (1,2), (1,3), (2,3), (1,2,3), (1,3,2)\}$.

Theorem 2.10. *For $1 \leq k \leq n$, the following holds:*

(i) The number of k-cycles of S_n is $\frac{n!}{k(n-k)!}$.

(ii) If $z_k(\sigma)$ is the number of k-cycles of σ, then

$$\frac{1}{n!} \sum_{\sigma \in S_n} z_k(\sigma) = \frac{1}{k}.$$

(iii) The average number of cycles of a permutation $\sigma \in S_n$ is the n-th harmonic number

$$H_n := \sum_{l=1}^{n} \frac{1}{l}.$$

Proof.

(i) Each k-cycle permutes a k-element set $\{a_1, \ldots, a_k\} \subseteq \{1, \ldots, n\}$. For the choice of this set, there are $\binom{n}{k}$ possibilities (Theorem 1.13). Each k-cycle on this set can be uniquely written in the form $(a_1, b_2, \ldots, b_k)$ with $\{b_2, \ldots, b_k\} = \{a_2, \ldots, a_k\}$. This yields $(k-1)!$ cycles, because the digits $b_2, \ldots, b_k$ can be permuted arbitrarily. In total, there are

$$\binom{n}{k}(k-1)! = \frac{n!(k-1)!}{k!(n-k)!} = \frac{n!}{k(n-k)!}$$

cycles of length k.

(ii) Let $C_k \subseteq S_n$ be the set of k-cycles. Each k-cycle is contained in $(n-k)!$ many permutations, because one can permute the $n-k$ digits outside the cycle arbitrarily. It follows that

$$\sum_{\sigma \in S_n} z_k(\sigma) = |\{(\sigma, c) \in S_n \times C_k : \sigma \text{ contains } c\}| = \sum_{c \in C_k} (n-k)! = |C_k|(n-k)! \overset{\text{(i)}}{=} \frac{n!(n-k)!}{k(n-k)!} = \frac{n!}{k}.$$

(iii) The average number of cycles is

$$\frac{1}{n!} \sum_{k=1}^{n} \sum_{\sigma \in S_n} z_k(\sigma) = \sum_{k=1}^{n} \frac{1}{k}$$

according to (ii). □

Remark 2.11. As is well known (Analysis),

$$\gamma := \lim_{n \to \infty} (H_n - \log n) = 0{,}577\ldots$$

is the *Euler-Mascheroni constant*. For large n, $H_n \approx \log(n) + \gamma$ therefore holds. It is not yet known whether γ is rational.

Example 2.12.

(i) The average number of cycles of $\sigma \in S_8$ is $H_8 = \frac{761}{280} \approx 2{,}71$.

(ii) (100 prisoners problem) The names of 100 prisoners are kept in 100 closed numbered envelopes. The prisoners are asked one after another to open 50 envelopes of their choice with the goal of finding their own name. If every prisoner succeeds in finding their own name, they all receive freedom. They may decide on a strategy beforehand, but may not communicate during the experiment. What is a good strategy? Without a strategy (i. e. everyone opens 50 random envelopes), the probability of success is only

$$2^{-100} = (2^{10})^{-10} = 1024^{-10} < 1000^{-10} = 10^{-30}.$$

The prisoners are numbered, so that the distribution of names into the envelopes describes a permutation $\sigma \in S_{100}$. When it is the turn of the prisoner with number a, they first open envelope a and find therein the name of prisoner $\sigma(a)$. After that, they open envelope $\sigma(a)$ and find therein the name of $\sigma^2(a)$ etc. In this way, they find their own name if and only if the cycle of σ containing a has length ≤ 50. The procedure is therefore successful if and only if σ contains no cycle of length > 50. Obviously, σ can contain at most one such cycle. The number of permutations with a cycle of length $k > 50$ is therefore

$$\sum_{k=51}^{100} \sum_{\sigma \in S_n} z_k(\sigma).$$

The probability that the strategy fails is consequently

$$\frac{1}{n!} \sum_{k=51}^{100} \sum_{\sigma \in S_n} z_k(\sigma) \overset{2.10}{=} \sum_{k=51}^{100} \frac{1}{k} \leq \int_{50}^{100} \frac{1}{x} \, \mathrm{d}x = \log(2 \cdot 50) - \log(50) = \log(2) < 0{,}7$$

(cf. Example 2.4). The probability of success is therefore greater than 30% (independent of the number of prisoners).

Definition 2.13. The number of permutations of S_n with exactly k cycles is called *Stirling number of the first kind* and is written as $\begin{bmatrix} n \\ k \end{bmatrix}$. If one considers the identity on the empty set as a product of 0 cycles, one obtains $\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1$.

Remark 2.14. For $n \in \mathbb{N}_0$ it holds that

$$n! = |S_n| = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}.$$

Example 2.15.

(i) By definition, $\begin{bmatrix} n \\ k \end{bmatrix} = 0$ if $k = 0 < n$ or $k > n$. Since id is the only permutation in S_n with n cycles, $\begin{bmatrix} n \\ n \end{bmatrix} = 1$ holds. In contrast to the binomial coefficient, in general $\begin{bmatrix} n \\ k \end{bmatrix} \neq \begin{bmatrix} n \\ n-k \end{bmatrix}$.

(ii) A permutation with only one cycle is an n-cycle. From Theorem 2.10 it follows that $\begin{bmatrix} n \\ 1 \end{bmatrix} = (n-1)!$.

(iii) Obviously, $\begin{bmatrix} n \\ n-1 \end{bmatrix}$ is the number of transpositions and thus also the number of 2-element subsets of $\{1, \ldots, n\}$. This shows $\begin{bmatrix} n \\ n-1 \end{bmatrix} = \binom{n}{2}$.

(iv) According to Remark 2.14, $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 4! - \begin{bmatrix} 4 \\ 1 \end{bmatrix} - \begin{bmatrix} 4 \\ 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 24 - 6 - 6 - 1 = 11$. The corresponding permutations are $(1,2,3)$, $(1,3,2)$, $(1,2,4)$, $(1,4,2)$, $(1,3,4)$, $(1,4,3)$, $(2,3,4)$, $(2,4,3)$, $(1,2)(3,4)$, $(1,3)(2,4)$, $(1,4)(2,3)$.

Lemma 2.16. *For $k, n \in \mathbb{N}$, it holds that*

$$\boxed{\begin{bmatrix} n \\ k-1 \end{bmatrix} + n \begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n+1 \\ k \end{bmatrix}.}$$

Proof. Let $\sigma \in S_n$ with exactly $k-1$ cycles. By appending the 1-cycle $(n+1)$, one obtains a permutation in S_{n+1} with exactly k cycles. Now let $\sigma \in S_n$ with exactly k cycles. Then the digit $n+1$ can be inserted at n positions in the cycle notation of σ (Example: inserting 4 into $(1,2)(3)$ yields $(4,1,2)(3)$, $(1,4,2)(3)$, $(1,2)(4,3)$). In this way, one obtains n different permutations in S_{n+1} with exactly k cycles. Obviously, every permutation of S_{n+1} with exactly k cycles arises in exactly one of the two ways. This shows the claim. □

Theorem 2.17. *For $0 \leq k < n$, it holds that*

$$\boxed{\begin{bmatrix} n \\ k \end{bmatrix} = \sum_{0<a_1<\ldots<a_{n-k}<n} a_1 \ldots a_{n-k}.}$$

Proof. Induction on n: For $k = 0$, one obtains the empty sum, because there are only $n-1$ natural numbers between 1 and $n-1$. Indeed, $\begin{bmatrix} n \\ 0 \end{bmatrix} = 0$. In particular, the claim holds for $n = 1$. Now let $k > 0$ and the claim for n be already proven. According to Lemma 2.16,

$$\begin{aligned} \begin{bmatrix} n+1 \\ k \end{bmatrix} = \begin{bmatrix} n \\ k-1 \end{bmatrix} + n \begin{bmatrix} n \\ k \end{bmatrix} &= \sum_{0<a_1<\ldots<a_{n-k+1}<n} a_1 \ldots a_{n-k+1} + n \sum_{0<a_1<\ldots<a_{n-k}<n} a_1 \ldots a_{n-k} \\ &= \sum_{0<a_1<\ldots<a_{n+1-k}<n+1} a_1 \ldots a_{n+1-k}. \end{aligned}$$

□

Example 2.18. For $n \in \mathbb{N}$, it holds that

$$\begin{bmatrix} n \\ n-1 \end{bmatrix} = \sum_{k=1}^{n-1} k = \binom{n}{2}$$

(cf. Example 2.15).

Definition 2.19.

- A *partition* of a (finite) set A is a set of pairwise disjoint, non-empty subsets $\{A_1, \ldots, A_k\} \subseteq 2^A$ with $A = A_1 \cup \ldots \cup A_k$. We denote the set of all partitions of A by $P(A)$. One calls $b(n) := |P(\{1, \ldots, n\})|$ the *n-th Bell number*.
- A *partition* of $n \in \mathbb{N}_0$ is a multiset $\lambda := \{k_1, \ldots, k_s\} \subseteq \mathbb{N}$ with $n = k_1 + \ldots + k_s$. One calls $k_1, \ldots, k_s$ the *parts* of λ. Let the set of all partitions of n be $P(n)$ and $p(n) := |P(n)|$.

Example 2.20. The partitions of $\{1,2,3\}$ are $\{\{1,2,3\}\}$, $\{\{1\},\{2,3\}\}$, $\{\{2\},\{1,3\}\}$, $\{\{3\},\{1,2\}\}$ and $\{\{1\},\{2\},\{3\}\}$. The partitions of 3 are $3 = 1+2 = 1+1+1$. Thus $b(3) = 5$ and $p(3) = 3$.

Remark 2.21.

(i) Note: $b(0) = 1 = p(0)$, because the empty (multi)set is a partition of $\varnothing$ (or 0).

(ii) If $\{A_1, \ldots, A_k\}$ is a partition of a finite set A, then $\{|A_1|, \ldots, |A_k|\}$ is a partition of $|A|$. Conversely, from every partition of $n \in \mathbb{N}$ one can construct a partition of $\{1, \ldots, n\}$. Therefore $b(n) \geq p(n)$ and $b(n) > p(n)$, if $n \geq 3$.

(iii) We will often write partitions of numbers in the form $(k_1, \ldots, k_s)$ with $k_1 \geq \ldots \geq k_s$ or in the form $(1^{m_1}, \ldots, n^{m_n}) := (\underbrace{1, \ldots, 1}_{m_1}, \ldots, \underbrace{n, \ldots, n}_{m_n})$ with $m_1, \ldots, m_n \in \mathbb{N}_0$.

(iv) A *relation* on a set A is a subset $\sim \subseteq A \times A$. One writes $a \sim b :\Longleftrightarrow (a, b) \in \sim$. One calls $\sim$ an *equivalence relation*, if for all $a, b, c \in A$ the following holds:

- $a \sim a$ (reflexive),
- $a \sim b \Longrightarrow b \sim a$ (symmetric),
- $a \sim b \sim c \Longrightarrow a \sim c$ (transitive).

For $a \in A$ one calls $[a] := \{b \in A : a \sim b\}$ the *equivalence class* of a. Let

$$A/\sim := \{[a] : a \in A\}$$

be the set of equivalence classes. Because of $a \in [a]$, we have $\bigcup_{a \in A}[a] = A$. Now let $x \in [a] \cap [b]$ for $a, b \in A$. With $a \sim x$ it also holds that $x \sim a$. For an arbitrary $c \in A$ it follows that

$$c \in [a] \Longrightarrow x \sim a \sim c \Longrightarrow b \sim x \sim c \Longrightarrow b \sim c \Longrightarrow c \in [b],$$

i. e. $[a] \subseteq [b]$. For reasons of symmetry, $[b] \subseteq [a]$ also holds, thus $[a] = [b]$. Any two equivalence classes are therefore either equal or disjoint. Thus $A/\sim$ is a partition of A.

(v) Conversely, if a partition $\{A_1, \ldots, A_k\}$ of A is given, then one obtains by

$$a \sim b :\Longleftrightarrow \exists i : a, b \in A_i$$

an equivalence relation on A. In this case $A/\sim = \{A_1, \ldots, A_k\}$. In this way, partitions and equivalence relations correspond to each other.

Example 2.22. The equality relation $=$ is an equivalence relation on every set A. The corresponding partition is $\{\{a\} : a \in A\}$.

Theorem 2.23. *Let $(1^{a_1}, \ldots, n^{a_n})$ be a partition of n. Then every n-element set has exactly*

$$\boxed{\frac{n!}{(1!)^{a_1} \ldots (n!)^{a_n} a_1! \ldots a_n!}}$$

partitions of the form $\{A_1, \ldots, A_l\}$ with $\{|A_1|, \ldots, |A_l|\} = (1^{a_1}, \ldots, n^{a_n})$.

Proof. Wlog. let $A = \{1, \ldots, n\}$. One can transform every arrangement $b_1, \ldots, b_n$ of the numbers $1, \ldots, n$ into a partition of the desired type by distributing corresponding braces { and }. We can first brace the a_1 1-element subsets, then the a_2 2-element subsets, etc.:

$$\{b_1\}, \{b_2\}, \ldots, \{b_i, b_{i+1}\}, \ldots.$$

Of the $n!$ possible arrangements $b_1, \ldots, b_n$, however, several lead to the same partition. On the one hand, one can arbitrarily permute the elements of each k-element subset without changing the partition. On the other hand, one can permute the a_k k-element subsets among themselves without changing the partition. Therefore, each $\prod_{k=1}^{n}(k!)^{a_k} a_k!$ arrangements lead to the same partition. This shows the claim. □

Example 2.24. The number of partitions of $\{1,2,3,4\}$ of type $(2,2) = (1^0, 2^2)$ is $\frac{4!}{(2!)^2 2!} = \frac{24}{8} = 3$. These are $\{\{1,2\},\{3,4\}\}$, $\{\{1,3\},\{2,4\}\}$ and $\{\{1,4\},\{2,3\}\}$.

Definition 2.25. If $\sigma \in S_n$ is a disjoint product of $a_i \geq 0$ cycles of length i, then $(1^{a_1}, \ldots, n^{a_n})$ is called the *cycle type* of σ. According to Lemma 2.7, this is a well-defined partition of n. The number of fixed points of σ is a_1.

Theorem 2.26. *The number of permutations of S_n with cycle type $(1^{a_1}, \ldots, n^{a_n})$ is*

$$\boxed{\frac{n!}{1^{a_1} \ldots n^{a_n} a_1! \ldots a_n!}.}$$

Proof. If one regards cycles as subsets of $\{1, \ldots, n\}$, then every permutation corresponds to a partition of $\{1, \ldots, n\}$. According to Theorem 2.23, the permutations with cycle type $(1^{a_1}, \ldots, n^{a_n})$ correspond to exactly

$$\frac{n!}{\prod_{k=1}^{n} (k!)^{a_k} a_k!}$$

partitions. It remains to count how many permutations yield the same partition. Since every k-cycle can be uniquely written in the form $(b_1, \ldots, b_k)$ with $b_1 := \min\{b_1, \ldots, b_k\}$, exactly $(k-1)!$ cycles yield the same set $\{b_1, \ldots, b_k\}$ (one can arbitrarily permute the $b_2, \ldots, b_k$). The number of the desired permutations is therefore

$$\frac{n!}{\prod_{k=1}^{n} (k!)^{a_k} a_k!} \prod_{k=1}^{n} ((k-1)!)^{a_k} = \frac{n!}{\prod_{k=1}^{n} k^{a_k} a_k!}. \qquad \square$$

Example 2.27. The k-cycles of S_n have cycle type $(1^{n-k}, k^1)$. Their number is $\frac{n!}{1^{n-k}(n-k)!k^1 1!} = \frac{n!}{k(n-k)!}$ in accordance with Theorem 2.10.

Definition 2.28. The number of k-element partitions of an n-element set is called the *Stirling number of the second kind* and is written as $\left\{ {n \atop k} \right\}$.

Remark 2.29.

(i) Since every permutation with k cycles defines a partition with k subsets, $\left\{ {n \atop k} \right\} \leq \left[{n \atop k} \right]$ for all $k, n \in \mathbb{N}$.

(ii) It holds that $b(n) = |P(\{1, \ldots, n\})| = \sum_{k=0}^{n} \left\{ {n \atop k} \right\}$.

Example 2.30.

(i) As usual, $\left\{ {0 \atop 0} \right\} = 1$ and $\left\{ {n \atop k} \right\} = 0$ for $k = 0 < n$ or $k > n$. Furthermore, $\left\{ {n \atop 1} \right\} = 1 = \left\{ {n \atop n} \right\}$ and $\left\{ {n \atop n-1} \right\} = \left[{n \atop n-1} \right] = \binom{n}{2}$.

(ii) Every 2-element partition of A has the form $\{B, A \setminus B\}$ with $B \in 2^A \setminus \{\varnothing, A\}$. This shows $\left\{ {n \atop 2} \right\} = \frac{1}{2}\left(|2^{\{1,\ldots,n\}}| - 2\right) \overset{1.4}{=} 2^{n-1} - 1$.

(iii) According to Remark 2.29, $b(4) = \left\{ {4 \atop 1} \right\} + \left\{ {4 \atop 2} \right\} + \left\{ {4 \atop 3} \right\} + \left\{ {4 \atop 4} \right\} = 1 + 2^3 - 1 + \binom{4}{2} + 1 = 15$.

Lemma 2.31. *For $k, n \in \mathbb{N}$ it holds that*

$$\boxed{\left\{ {n \atop k-1} \right\} + k \left\{ {n \atop k} \right\} = \left\{ {n+1 \atop k} \right\}.}$$

Proof. Let $A = \{1, \ldots, n\}$ and $\{A_1, \ldots, A_{k-1}\}$ be a partition of A. Then $\{A_1, \ldots, A_{k-1}, \{n+1\}\}$ is a k-element partition of $\{1, \ldots, n+1\}$. Now let $\{A_1, \ldots, A_k\}$ be a partition of A. Then one can add the number $n+1$ to each of the sets $A_1, \ldots, A_k$ and obtains in this way a k-element partition of $\{1, \ldots, n+1\}$. Obviously, every k-element partition of $\{1, \ldots, n+1\}$ arises in exactly one of the two ways. This shows the claim. □

Theorem 2.32. *For $0 \leq k < n$ it holds that*

$$\boxed{\left\{ {n \atop k} \right\} = \sum_{1 \leq a_1 \leq \ldots \leq a_{n-k} \leq k} a_1 \ldots a_{n-k}.}$$

Proof. Induction on n: For $k = 0$ one obtains the empty sum in accordance with $\left\{ {n \atop 0} \right\} = 0$. In particular, the claim holds for $n = 1$. Now let $k > 0$ and the claim be already proven for n. According to Lemma 2.31

$$\begin{aligned} \left\{ {n+1 \atop k} \right\} = \left\{ {n \atop k-1} \right\} + k \left\{ {n \atop k} \right\} &= \sum_{1 \leq a_1 \leq \ldots \leq a_{n-k+1} \leq k-1} a_1 \ldots a_{n-k+1} + k \sum_{1 \leq a_1 \leq \ldots \leq a_{n-k} \leq k} a_1 \ldots a_{n-k} \\ &= \sum_{1 \leq a_1 \leq \ldots \leq a_{n+1-k} \leq k} a_1 \ldots a_{n+1-k}. \end{aligned}$$

□

Example 2.33. For $n \in \mathbb{N}$ it holds that

$$\left\{ {n \atop 2} \right\} = \sum_{1 \leq a_1 \leq \ldots \leq a_{n-2} \leq 2} a_1 \ldots a_{n-2} = \sum_{k=0}^{n-2} 2^k = 2^{n-1} - 1$$

(cf. Example 2.30).

Remark 2.34. One should compare the following result with Theorem 1.9.

Theorem 2.35. *For finite sets A and B there exist exactly $\left\{ {|A| \atop |B|} \right\} |B|!$ surjective mappings $A \to B$.*

Proof. Wlog. let $B = \{1, \ldots, k\}$. Every surjective mapping $f \colon A \to B$ yields a k-element partition $\{f^{-1}(1), \ldots, f^{-1}(k)\}$ of A. For $\sigma \in \mathrm{Sym}(B)$, the mapping $\sigma \circ f \colon A \to B$ is also surjective and leads to the partition

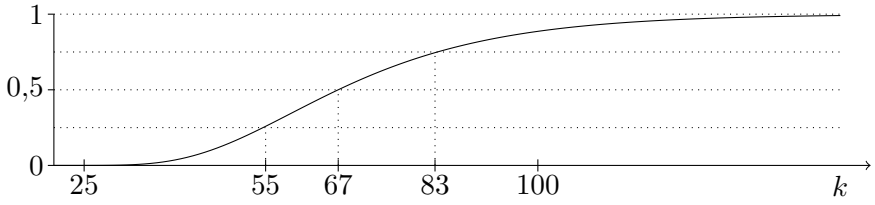
$$\{(\sigma \circ f)^{-1}(1), \ldots, (\sigma \circ f)^{-1}(k)\} = \{f^{-1}(\sigma^{-1}(1)), \ldots, f^{-1}(\sigma^{-1}(k))\} = \{f^{-1}(1), \ldots, f^{-1}(k)\}.$$

One easily sees that these are the only mappings that lead to the same partition. The number of surjective mappings is therefore $\left\{ {|A| \atop k} \right\} |\mathrm{Sym}(B)| = \left\{ {|A| \atop k} \right\} k!$. □

Example 2.36 (Coupon collector's problem)**.** With every purchase at the supermarket, you receive one of n different trading cards (randomly and uniformly distributed). What is the probability that you own all trading cards after k purchases? The k purchases provide a mapping $\{1,\dots,k\} \to \{1,\dots,n\}$. There are n^k such mappings, of which $\left\{ {k \atop n} \right\} n!$ are surjective. The probability is therefore

$$\frac{n!}{n^k} \left\{ {k \atop n} \right\}.$$

For $n = 20$ one obtains:



Theorem 2.37. *For $k, n \in \mathbb{N}_0$, it holds that*

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

Proof. Let $A := \{1,\dots,n\}$, $B := \{1,\dots,k\}$ and M be the set of surjective mappings from A to B. According to Theorem 2.35, it suffices to show $|M| = \sum_{l=0}^{k} (-1)^{k-l} \binom{k}{l} l^n$. For $i = 1,\dots,k$ let

$$M_i := \big\{ f : A \to B : i \notin f(A) \big\}.$$

For $1 \le i_1 < \ldots < i_l \le k$, then $M_{i_1} \cap \ldots \cap M_{i_l}$ is the set of all mappings from A to $B \setminus \{i_1,\dots,i_l\}$. In particular, $|M_{i_1} \cap \ldots \cap M_{i_l}| = (k-l)^n$ according to Remark 1.7. Theorem 1.25 shows

$$|M| = |B^A \setminus (M_1 \cup \ldots \cup M_k)| = k^n + \sum_{l=1}^{k} (-1)^l \binom{k}{l} (k-l)^n = \sum_{l=0}^{k} (-1)^l \binom{k}{l} (k-l)^n.$$

The claim follows from $\binom{k}{l} = \binom{k}{k-l}$. □

Remark 2.38. From Theorem 2.37 it follows that

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

Asymptotically, the *Stirling formula* holds

$$n! \sim \sqrt{2\pi n} \Big(\frac{n}{e} \Big)^n,$$

i. e.

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}(n/e)^n} = 1$$

(without proof). Example: $100! \approx 9.333 \cdot 10^{157}$ and $\sqrt{200\pi}(100/e)^{100} \approx 9.325 \cdot 10^{157}$.

Theorem 2.39 (DOBIŃSKI formula)**.** *For $n \in \mathbb{N}_0$ is*

$$\boxed{b(n) = \frac{1}{e}\sum_{k=0}^{\infty}\frac{k^n}{k!}.}$$

Proof. Because $\left\{{n \atop k}\right\} = 0$ for $k > n$, it holds that

$$b(n) \overset{2.29}{=} \sum_{k=0}^{\infty}\left\{{n \atop k}\right\} \overset{2.37}{=} \sum_{k=0}^{\infty}\frac{1}{k!}\sum_{l=0}^{k}(-1)^{k-l}\binom{k}{l}l^n = \sum_{k=0}^{\infty}\sum_{l=0}^{k}(-1)^{k-l}\frac{l^n}{l!(k-l)!}$$
$$= \sum_{k=0}^{\infty}\sum_{l=0}^{k}\frac{(-1)^l}{l!}\frac{(k-l)^n}{(k-l)!} \overset{(*)}{=} \Big(\sum_{l=0}^{\infty}\frac{(-1)^l}{l!}\Big)\Big(\sum_{k=0}^{\infty}\frac{k^n}{k!}\Big) = \frac{1}{e}\sum_{k=0}^{\infty}\frac{k^n}{k!},$$

where in $(*)$ the Cauchy product formula for absolutely convergent series is used (Analysis). □

Theorem 2.40. *For $n \in \mathbb{N}_0$ is*

$$b(n+1) = \sum_{k=0}^{n}\binom{n}{k}b(k).$$

Proof. Let $\mathcal{A}$ be a partition of $\{1, \dots, n+1\}$ and $n+1 \in A \in \mathcal{A}$ with $k := |A| - 1 \geq 0$. Then there are $\binom{n}{k}$ possibilities for A and $\mathcal{A} \setminus \{A\}$ is a partition of $\{1, \dots, n\} \setminus A$. For $\mathcal{A} \setminus \{A\}$ there are thus $b(n-k)$ possibilities. It follows

$$b(n+1) = \sum_{k=0}^{n}\binom{n}{k}b(n-k) = \sum_{k=0}^{n}\binom{n}{k}b(k).$$ □

Remark 2.41. No simple formula for $p(n)$ is known. However, Hardy and Ramanujan have proved

$$p(n) \sim \frac{e^{\pi\sqrt{2n/3}}}{4n\sqrt{3}}$$

Example: $p(10^4) \approx 3.617 \cdot 10^{106}$ and $\frac{e^{\pi\sqrt{20000/3}}}{40000\sqrt{3}} \approx 3.633 \cdot 10^{106}$.

3. Möbius Inversion

Definition 3.1. A relation $\leq$ on a set A is called an *order relation* (or *partial order*), if for all $a, b, c \in A$ the following holds:

- $a \leq a$ (reflexive),
- $a \leq b \leq a \Longrightarrow a = b$ (antisymmetric),
- $a \leq b \leq c \Longrightarrow a \leq c$ (transitive).

If applicable, $(A, \leq)$ is called an *ordered* set. One also writes $a \geq b$ if $b \leq a$, and $a < b$ (resp. $a > b$) if $a \leq b \neq a$ (resp. $b \leq a \neq b$).

Example 3.2.

(i) The usual "less than or equal to relation" $\leq$ on $\mathbb{R}$.

(ii) The subset relation $\subseteq$ on 2^A for every set A.

(iii) The divisibility relation $\mid$ on $\mathbb{N}$, but not on $\mathbb{Z}$, because $1 \mid -1 \mid 1$.

(iv) For every ordered set $(A, \leq)$, $(A, \geq)$ is also an ordered set.

(v) For ordered sets $(A_1, \leq_1), \ldots, (A_n, \leq_n)$, $A_1 \times \ldots \times A_n$ is *lexicographically* ordered by

$$(a_1, \ldots, a_n) < (b_1, \ldots, b_n) :\Longleftrightarrow \exists k \in \mathbb{N}_0 : a_1 = b_1, \ldots, a_k = b_k, a_{k+1} <_{k+1} b_{k+1}.$$

Definition 3.3. For an ordered set $(A, \leq)$ and $a, b \in A$, let

$$[a, b] := \{c \in A : a \leq c \leq b\}.$$

$(A, \leq)$ is called *locally finite* if $|\{b \in A : b \leq a\}| < \infty$ holds for all $a \in A$. If applicable, the *Möbius function* $\mu_A \colon A \times A \to \mathbb{Z}$ is defined recursively by

$$\mu_A(a, b) := \begin{cases} 1 & \text{if } a = b, \\ -\sum_{a \leq x < b} \mu_A(a, x) & \text{if } a \neq b. \end{cases}$$

Remark 3.4. In the situation of Definition 3.3, $\sum_{x \in [a,b]} \mu_A(a, x) = 0$ holds if $a \neq b$. We show $\sum_{x \in [a,b]} \mu_A(x, b) = 0$ by induction on $k := |[a, b]| \geq 2$. For $k = 2$, we have

$$\sum_{x \in [a,b]} \mu_A(x, b) = \mu_A(a, b) + \mu_A(b, b) = \mu_A(a, b) + \mu_A(a, a) = \sum_{x \in [a,b]} \mu_A(a, x) = 0.$$

Now let the claim be already proven for $k - 1$. Then

$$\begin{aligned} \sum_{x \in [a,b]} \mu_A(x, b) &= \mu_A(b, b) - \sum_{a \leq x < b} \sum_{x \leq y < b} \mu_A(x, y) = \mu_A(a, a) - \sum_{a \leq y < b} \sum_{x \in [a,y]} \mu_A(x, y) \\ &= \mu_A(a, a) - \mu_A(a, a) = 0. \end{aligned}$$

Theorem 3.5 (Möbius Inversion)**.** *Let $(A, \leq)$ be locally finite. For $f, F \colon A \to \mathbb{R}$, the following are equivalent:*

(1) $\boxed{F(a) = \sum_{x \leq a} f(x)}$ *for all $a \in A$.*

(2) $\boxed{f(a) = \sum_{x \leq a} \mu_A(x, a) F(x)}$ *for all $a \in A$.*

Proof. Since $(A, \leq)$ is locally finite, the sums are well-defined. Let $F(a) = \sum_{x \leq a} f(x)$ for all $a \in A$. Then it holds that

$$\sum_{x \leq a} \mu_A(x, a) F(x) = \sum_{x \leq a} \mu_A(x, a) \sum_{y \leq x} f(y) = \sum_{y \leq a} f(y) \sum_{x \in [y,a]} \mu_A(x, a) \overset{3.4}{=} f(a).$$

Conversely, let $f(a) = \sum_{x \leq a} \mu_A(x, a) F(x)$ for all $a \in A$. Then it follows that

$$\sum_{x \leq a} f(x) = \sum_{x \leq a} \sum_{y \leq x} \mu_A(y, x) F(y) = \sum_{y \leq a} F(y) \sum_{x \in [y,a]} \mu_A(y, x) \overset{3.4}{=} F(a). \qquad \square$$

Remark 3.6. Theorem 3.5 is particularly useful when μ_A has a simple form.

Example 3.7.

(i) For every set A, $(A, =)$ is locally finite. The Möbius function is the Kronecker delta $\mu_A(a, b) = \delta_{ab}$ and the Möbius inversion reduces to $f = F$.

(ii) Obviously $(\mathbb{N}, \leq)$ is locally finite. For $a \in \mathbb{N}$ we have $\mu_{\mathbb{N}}(a, a) = 1$, $\mu_{\mathbb{N}}(a, a+1) = -\mu_{\mathbb{N}}(a, a) = -1$ and $\mu_{\mathbb{N}}(a, a+2) = -1 + 1 = 0$. By induction it is easy to show $\mu_{\mathbb{N}}(a, b) = 0$ for $b \notin \{a, a+1\}$. Theorem 3.5 yields in this case

$$F(n) = \sum_{k=1}^{n} f(k) \iff f(n) = F(n) - F(n-1).$$

This is a discrete version of the fundamental theorem of calculus (F corresponds to the integral of f and f corresponds to the derivative of F).

(iii) For every finite set A, $(2^A, \subseteq)$ is locally finite. We show

$$\mu_{2^A}(X, Y) = \begin{cases} (-1)^{|Y \setminus X|} & \text{if } X \subseteq Y, \\ 0 & \text{if } X \nsubseteq Y. \end{cases}$$

The cases $X = Y$ and $X \nsubseteq Y$ are clear. So let $X \subsetneq Y$ and $k := |Y \setminus X| \geq 1$. By induction we assume that the statement already holds for $k - 1$. Then

$$\begin{aligned} \mu_{2^A}(X, Y) &= -\sum_{X \subseteq Z \subsetneq Y} \mu_{2^A}(X, Z) = -\sum_{X \subseteq Z \subsetneq Y} (-1)^{|Z \setminus X|} \\ &= -\sum_{l=0}^{k-1} \binom{k}{l} (-1)^l = -(1-1)^k + (-1)^k = (-1)^k. \end{aligned}$$

For $f, F \colon 2^A \to \mathbb{R}$ it therefore holds that

$$F(B) = \sum_{X \subseteq B} f(X) \iff f(B) = \sum_{X \subseteq B} (-1)^{|B \setminus X|} F(X).$$

Replacing $\subseteq$ by $\supseteq$, one analogously obtains

$$F(B) = \sum_{X \supseteq B} f(X) \iff f(B) = \sum_{X \supseteq B} (-1)^{|X \setminus B|} F(X). \tag{3.1}$$

(iv) Let $A_1, \ldots, A_n$ be finite sets and $N := \{1, \ldots, n\}$. We define $f, F \colon 2^N \to \mathbb{R}$ by

$$\begin{aligned} f(I) &:= \Big| \bigcap_{i \in I} A_i \setminus \bigcup_{j \in N \setminus I} A_j \Big|, \\ F(I) &:= \Big| \bigcap_{i \in I} A_i \Big| \end{aligned}$$

for $I \subseteq N$. For $a \in \bigcap_{i \in I} A_i$ there exists exactly one $J \supseteq I$ with $a \in \bigcap_{j \in J} A_j \setminus \bigcup_{l \in N \setminus J} A_l$. This shows $F(I) = \sum_{J \supseteq I} f(J)$ and (3.1) yields

$$0 = f(\varnothing) = \sum_{I \supseteq \varnothing} (-1)^{|I|} \Big| \bigcap_{i \in I} A_i \Big| = |A_1 \cup \ldots \cup A_n| + \sum_{\varnothing \neq I \subseteq N} (-1)^{|I|} \Big| \bigcap_{i \in I} A_i \Big|.$$

This is exactly the inclusion-exclusion principle.

(v) Also $(\mathbb{N}, |)$ is locally finite. Let $n = p_1^{a_1} \dots p_s^{a_s}$ be the prime factorization of $n \in \mathbb{N}$. We define the *classical Möbius function* $\mu\colon \mathbb{N} \to \mathbb{R}$ by

$$\mu(n) := \begin{cases} (-1)^s & \text{if } a_1 = \ldots = a_s = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Note: $\mu(1) = (-1)^0 = 1$. It then holds that

$$\sum_{d|n} \mu(d) = \sum_{k=0}^{s} \sum_{q_1,\ldots,q_k \in \{p_1,\ldots,p_s\}} \mu(q_1 \dots q_k) = \sum_{k=0}^{s} \sum_{\substack{M \subseteq \{p_1,\ldots,p_s\} \\ |M|=k}} (-1)^k$$

$$= \sum_{k=0}^{s} (-1)^k \binom{s}{k} = (1-1)^s = 0,$$

if $n \neq 1$. We show $\mu_{\mathbb{N}}(a,b) = \mu(b/a)$ if $a \mid b$. This is clear for $a = b$. We now assume $a \neq b$ and argue by induction on b/a. Then

$$\mu_{\mathbb{N}}(a,b) = -\sum_{\substack{a|x|b \\ x \neq b}} \mu_{\mathbb{N}}(a,x) = -\sum_{\substack{a|x|b \\ x \neq b}} \mu(x/a) = \mu(b/a) - \sum_{y|\frac{b}{a}} \mu(y) = \mu(b/a).$$

Theorem 3.5 thus has the following form

$$\boxed{F(n) = \sum_{d|n} f(d) \iff f(n) = \sum_{d|n} \mu(n/d) F(d).} \tag{3.2}$$

(vi) Let $n = p_1^{a_1} \dots p_s^{a_s}$ be the prime factorization of $n \in \mathbb{N}$. Then

$$\sum_{d|n} \mu(n/d) d = \sum_{t=0}^{s} \sum_{1 \leq i_1 < \ldots < i_t \leq s} (-1)^t \frac{n}{p_{i_1} \dots p_{i_t}} = n\Big(1 - \frac{1}{p_1}\Big) \dots \Big(1 - \frac{1}{p_s}\Big) = \varphi(n)$$

according to Theorem 1.27. Equation 3.2 shows

$$\boxed{\sum_{d|n} \varphi(d) = n.}$$

Remark 3.8. One can generalize Theorem 3.5 by replacing $\mathbb{R}$ with an arbitrary abelian group $(G, \cdot)$. One then obtains

$$\boxed{F(a) = \prod_{x \leq a} f(x) \iff f(a) = \prod_{x \leq a} F(x)^{\mu(x,a)}}$$

(proof is exactly the same).

Definition 3.9. Let $(A, \leq)$ be a locally finite ordered set and $a, b \in A$. A *path* of *length* $l \leq 0$ between a and b is a sequence of the form $a = x_1 < \dots < x_{k+1} = b$ with $x_1, \dots, x_{k+1} \in A$.

Theorem 3.10 (Hall)**.** *Let $(A, \leq)$ be a locally finite ordered set and $a, b \in A$. Let w_l be the number of paths of length l between a and b. Then*

$$\mu_A(a,b) = w_0 - w_1 + w_2 \mp \dots.$$

Proof. Induction on $k := |[a,b]| < \infty$. In the sum, only finitely many summands are non-zero, because $c_l = 0$ for $l \geq k$. In the case $k = 1$, there is only the path of length 0 between a and a, i. e. $\mu_A(a,a) = 1 = c_0$. Now let $k \geq 2$ and assume the claim is already proven for $k-1$. Let $a = x_1 < \cdots < x_{l+1} = b$ and $c := x_l$. Then $x_1 < \cdots < x_l$ is a path of length $l-1$ between a and c. By induction, this path provides the contribution $(-1)^l$ to $\mu_A(a,c)$. Because of $\mu_A(a,b) = -\sum_{a \leq c < b} \mu_A(a,c)$, $x_1 < \cdots < x_{l+1}$ provides the contribution $(-1)^{l+1}$ to $\mu_A(a,b)$. Conversely, every path of length $l-1$ between a and c can be extended to a path of length l between a and b. This shows the claim. □

4. Power series

Remark 4.1. For many counting problems, no simple formulas are known (think of $p(n)$). Often it is more favorable to consider the sequence of the desired count values in its entirety. Through clever algebraic transformations, one can thereby generate new identities. In this section, the foundations for this are laid.

Definition 4.2. For a field K, let $K[[X]] := K^{\mathbb{N}_0} = \{(a_0, a_1, \ldots) : a_i \in K\}$. Two elements $\alpha := (a_0, \ldots)$ and $\beta := (b_0, \ldots)$ of $K[[X]]$ can be added and multiplied as follows:

$$\begin{aligned} \alpha + \beta &:= (a_0 + b_0, a_1 + b_1, \ldots) \in K[[X]], \\ \alpha \cdot \beta &:= (a_0 b_0, a_1 b_0 + a_0 b_1, \ldots, \sum_{i=0}^{n} a_i b_{n-i}, \ldots) \in K[[X]]. \end{aligned}$$

We set $0 := (0, 0, \ldots) \in K[[X]]$ and $1 := (1, 0, 0, \ldots) \in K[[X]]$.

Lemma 4.3. *For $\alpha, \beta, \gamma \in K[[X]]$ the following hold:*

$$\begin{array}{ccc} (\alpha+\beta)+\gamma = \alpha+(\beta+\gamma) & \alpha+\beta = \beta+\alpha & \alpha+0=\alpha \\ (\alpha\cdot\beta)\cdot\gamma = \alpha\cdot(\beta\cdot\gamma) & \alpha\cdot\beta = \beta\cdot\alpha & \alpha\cdot 1 = \alpha \\ \alpha\cdot(\beta+\gamma) = (\alpha\cdot\beta)+(\alpha\cdot\gamma) & \exists \delta \in K[[X]] : \alpha+\delta = 0, & \alpha\beta = 0 \Longrightarrow \alpha = 0 \vee \beta = 0. \end{array}$$

Proof. The first three statements follow directly from the corresponding axioms in K. Now let $\alpha = (a_0, \ldots)$, $\beta = (b_0, \ldots)$ and $\gamma = (c_0, \ldots)$. For $\delta := (-a_0, -a_1, \ldots)$, it then holds that $a + \delta = 0$. The n-th entry of $\alpha \cdot (\beta \cdot \gamma)$ is

$$\sum_{i=0}^{n} a_i \sum_{j=0}^{n-i} b_j c_{n-i-j} = \sum_{i+j+k=n} a_i b_j c_k = \sum_{i=0}^{n} \Big(\sum_{j=0}^{i} a_j b_{i-j} \Big) c_{n-i}.$$

This shows $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$. Because of $\sum_{i=0}^{n} a_i b_{n-i} = \sum_{i=0}^{n} b_i a_{n-i}$, we have $\alpha \cdot \beta = \beta \cdot \alpha$. The equation $\alpha \cdot 1 = \alpha$ is easy to see. The n-th entry of $\alpha \cdot (\beta + \gamma)$ is

$$\sum_{i=0}^{n} a_i (b_{n-i} + c_{n-i}) = \sum_{i=0}^{n} a_i b_{n-i} + \sum_{i=0}^{n} a_i c_{n-i}$$

and $\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$ follows. Finally, let $\alpha\beta = 0$. Assume indirectly $\alpha \neq 0 \neq \beta$. Let $k := \min\{n \in \mathbb{N}_0 : a_n \neq 0\}$ and $l := \min\{n \in \mathbb{N}_0 : b_n \neq 0\}$. The $(k+l)$-th entry of $\alpha\beta$ is then $\sum_{i=0}^{k+l} a_i b_{k+l-i} = a_k b_l \neq 0$. This contradiction shows $\alpha = 0$ or $\beta = 0$. □

Remark 4.4.

(i) Lemma 4.3 states that one can calculate in $K[[X]]$ as in $\mathbb{Z}$. One calls $K[[X]]$ the *ring of (formal) power series*. Its elements are also written in the form $\sum_{n=0}^{\infty} a_n X^n = a_0 + a_1 X + a_2 X^2 + \ldots$, where $X = (0, 1, 0, 0, \ldots)$ is an *indeterminate*. It holds that

$$\sum_{n=0}^{\infty} a_n X^n = \sum_{n=0}^{\infty} b_n X^n \iff a_n = b_n \ \forall n \in \mathbb{N}_0.$$

When multiplying power series, one multiplies term by term and subsequently groups identical powers of X. One calls a_0 the *constant term* of α. If the summation range is clear, we write more briefly $\sum a_n X^n$. Furthermore, we will omit the multiplication symbol $\cdot$ and use "order of operations" (multiplication before addition), i.e., $\alpha\beta + \gamma := (\alpha \cdot \beta) + \gamma$. Let the inverse of α with respect to $+$ be $-\alpha$. As usual, we write $\alpha - \beta$ instead of $\alpha + (-\beta)$.

(ii) The meaning of the word "formal" lies in the fact that, in contrast to analysis, we do not consider convergence, since X is always an indeterminate and not a real number (hence also the use of the capital letter). Instead, we introduce a much simpler metric on $K[[X]]$ in Definition 4.11.

(iii) For $\alpha \in K[[X]]$, we define $\alpha K[[X]] := \{\alpha\beta : \beta \in K[[X]]\}$. For example, $XK[[X]]$ is the set of power series with constant term 0.

(iv) One can extend $K[[X]]$ to a field $K((X))$ by replacing power series with (formal) *Laurent series* of the form $\sum_{n=k}^{\infty} a_n X^n$ with $k \in \mathbb{Z}$ and $a_n \in K$ (Exercise 25).

Example 4.5. For every field K, there exist $\sum_{n=0}^{\infty} X^n$, $\sum nX^n$ and $\sum (-1)^n X^n \in K[[X]]$. Furthermore,

$$\boxed{\exp(X) := \sum_{n=0}^{\infty} \frac{X^n}{n!} = 1 + X + \frac{X^2}{2} + \frac{X^3}{6} + \ldots \in \mathbb{Q}[[X]]} \qquad \text{((formal) } \textit{exponential function}\text{)}.$$

It holds that

$$(1 - X) \sum_{n=0}^{\infty} X^n = \sum_{n=0}^{\infty} X^n - \sum_{n=1}^{\infty} X^n = 1.$$

Definition 4.6. One calls $\alpha \in K[[X]]$ *invertible*, if there exists a $\beta \in K[[X]]$ with $\alpha\beta = 1$.

Remark 4.7.

(i) If α, β, γ are such that $\alpha\beta = 1 = \alpha\gamma$, then $\alpha(\beta - \gamma) = 0$ and it follows that $\beta = \gamma$, because otherwise $1 = \alpha\beta = 0\beta = 0$. Thus, there exists at most one β with $\alpha\beta = 1$. One calls β the *inverse* of α and writes $\alpha^{-1} := \beta$ or $1/\alpha$. More generally, we set

$$\alpha^k := \begin{cases} \underbrace{\alpha \ldots \alpha}_{k \text{ times}} & \text{if } k > 0, \\ 1 & \text{if } k = 0, \\ (\alpha^{-1})^{-k} & \text{if } k < 0. \end{cases}$$

for $k \in \mathbb{Z}$.

(ii) For $\alpha, \beta, \gamma \in K[[X]]$ with $\alpha\beta = \gamma$, we write $\frac{\gamma}{\beta} := \alpha$, if $\beta \neq 0$ (as in (i), this is well-defined).

Lemma 4.8. *Let $\alpha = \sum a_n X^n \in K[[X]]$.*

(i) α is invertible if and only if $a_0 \neq 0$ holds.

(ii) If there exists an $m \in \mathbb{N}$ with $\alpha^m = 1$, then $\alpha \in K$.

Proof.

(i) Let $\beta = \sum b_n X^n \in K[[X]]$ with $\alpha\beta = 1$. Then $a_0 b_0 = 1$ and $a_0 \neq 0$. Conversely, let $a_0 \neq 0$. We define $b_0, b_1, \ldots \in K$ inductively by $b_0 := 1/a_0$ and

$$b_k := -\frac{1}{a_0} \sum_{i=1}^{k} a_i b_{k-i} \in K$$

for $k \in \mathbb{N}$. It then holds that

$$\sum_{i=0}^{k} a_i b_{k-i} = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{if } k > 0. \end{cases}$$

This shows $\alpha\beta = 1$ for $\beta := \sum b_n X^n$.

(ii) We can assume $m > 1$. For a prime divisor p of m, we have $(\alpha^{m/p})^p = 1$. By induction on m, we may therefore assume $m = p$. Suppose indirectly that $\alpha \notin K$ and let $n \in \mathbb{N}$ be minimal with $a_n \neq 0$. The n-th coefficient of $\alpha^p = 1$ is $p a_0^{p-1} a_n = 0$. Since α is invertible ($\alpha^{-1} = \alpha^{m-1}$), $a_0 \neq 0$ holds and it follows that $p = 0$ in K (this holds, for example, for $|K| = p$). We now examine the coefficient of X^{np} in α^p. This depends only on $a_0, \ldots, a_{np}$. According to the multinomial theorem,

$$(a_0 + \ldots + a_{np} X^{np})^p = \sum_{\substack{(k_0,\ldots,k_{np}) \in \mathbb{N}_0^{np+1} \\ k_0 + \ldots + k_{np} = p}} \binom{p}{k_0, \ldots, k_{np}} a_0^{k_0} \ldots a_{np}^{k_{np}} X^{k_1 + 2k_2 + \ldots + npk_{np}}.$$

For $k_0, \ldots, k_{np} < p$, the multinomial coefficient $\binom{p}{k_0,\ldots,k_{np}}$ is obviously divisible by p and therefore vanishes in K. Thus, only the multisets $\{k_0, \ldots, k_{np}\} = \{0, \ldots, 0, p\}$ remain, i.e.

$$(a_0 + \ldots + a_{np} X^{np})^p = a_0^p + a_n^p X^{np} + a_{n+1}^p X^{(n+1)p} + \ldots + a_{np}^p X^{np^2}.$$

The np-th coefficient of α^p is thus $a_n^p \neq 0$, in contradiction to $\alpha^p = 1$. □

Remark 4.9. If $\alpha, \beta \in K[[X]]$ are invertible, then so are α^{-1} and $\alpha\beta$ according to Lemma 4.8. In this case, $(\alpha^{-1})^{-1} = \alpha$ and $(\alpha\beta)^{-1} = \alpha^{-1}\beta^{-1}$. The invertible power series form an abelian group with respect to multiplication with identity element 1 (cf. Definition 10.2). It is called the *unit group* of $K[[X]]$ and is written as $K[[X]]^\times$. According to Lemma 4.8, $K[[X]]^\times = K[[X]] \setminus XK[[X]]$ holds.

Example 4.10.

(i) According to Example 4.5, $\frac{1}{1-X} = \sum X^n$ is the (formal) *geometric series*. More generally,

$$\frac{1}{a - X} = \sum a^{-n-1} X^n$$

for $a \in K \setminus \{0\}$ and

$$\sum_{k=0}^{n-1} \alpha^k = \frac{\alpha^n - 1}{\alpha - 1}$$

for $\alpha \in K[[X]] \setminus \{1\}$ and $n \in \mathbb{N}$.

(ii) For distinct $a, b \in K \setminus \{0\}$, $X + a$ and $X + b$ are invertible and the *partial fraction decomposition*

$$\frac{1}{(X+a)(X+b)} = \frac{1}{b-a}\Big(\frac{1}{X+a} - \frac{1}{X+b}\Big)$$

holds (bring the right side to a common denominator).

Definition 4.11. For $\alpha = \sum a_n X^n \in K[[X]]$ let

$$|\alpha| := 2^{-\inf\{k \in \mathbb{N}_0 : a_k \neq 0\}} \in \mathbb{R}$$

be the *norm* of α, where $|0| = 2^{-\infty} = 0$.

Example 4.12. $\alpha \in K[[X]]$ is invertible if and only if $|\alpha| = 1$.

Lemma 4.13. *For $\alpha, \beta \in K[[X]]$ we have $|\alpha\beta| = |\alpha||\beta|$ and $|\alpha + \beta| \leq \max\{|\alpha|, |\beta|\}$ with equality if $|\alpha| \neq |\beta|$* (ultrametric inequality).

Proof. Wlog. let $\alpha = \sum a_n X^n \neq 0 \neq \beta = \sum b_n X^n$. Let $|\alpha| = 2^{-k}$, $|\beta| = 2^{-l}$ and wlog. $k \geq l$. Then

$$\alpha\beta = a_k b_l X^{k+l} + \sum_{n=k+l+1}^{\infty} c_n X^n$$

for suitable $c_n \in K$. Because of $a_k b_l \neq 0$ it follows that $|\alpha\beta| = 2^{-k-l} = |\alpha||\beta|$.

From $a_n + b_n \neq 0$ it follows that $a_n \neq 0$ or $b_n \neq 0$. Because of $k \geq l$ we then have $n \geq l$ and $|\alpha + \beta| \leq 2^{-l} = \max\{|\alpha|, |\beta|\}$. For $k > l$ we have $a_l + b_l = b_l \neq 0$ and $|\alpha + \beta| = 2^{-l}$. □

Theorem 4.14. *By $d(\alpha, \beta) := |\alpha - \beta|$, $K[[X]]$ becomes a complete metric space.*

Proof. Clearly $d(\alpha, \beta) = d(\beta, \alpha) \geq 0$ with equality if and only if $\alpha = \beta$. Therefore d is symmetric and positive definite. The triangle inequality follows from the ultrametric inequality

$$d(\alpha, \gamma) = |\alpha - \gamma| = |\alpha - \beta + \beta - \gamma| \leq \max\{|\alpha - \beta|, |\beta - \gamma|\} \leq |\alpha - \beta| + |\beta - \gamma| = d(\alpha, \beta) + d(\beta, \gamma).$$

Let $\alpha_1, \alpha_2, \ldots \in K[[X]]$ be a Cauchy sequence with $\alpha_m = \sum a_{m,n} X^n$ for $m \in \mathbb{N}$. For each $k \in \mathbb{N}$ there exists $M = M(k) \geq 1$ with $|\alpha_m - \alpha_M| < 2^{-k}$ for all $m \geq M$. This shows $a_{m,n} = a_{M,n}$ for all $m \geq M$ and $n \leq k$. We define

$$a_k := a_{M(k),k}$$

and $\alpha = \sum a_n X^n$. Then $|\alpha - \alpha_{M(k)}| < 2^{-k} \to 0$, i. e. $\lim_{m\to\infty} \alpha_m = \alpha$. Thus $K[[X]]$ is complete with respect to d. □

Lemma 4.15. *If $\alpha_1, \alpha_2, \ldots \in K[[X]]$ is a sequence converging to zero, then $\sum_{k=1}^{\infty} \alpha_k$ and $\prod_{k=1}^{\infty}(1 + \alpha_k)$ converge.*

Proof. According to Theorem 4.14 it suffices to show that the partial sums are Cauchy sequences. For $\epsilon > 0$ let $N \geq 0$ with $|\alpha_k| < \epsilon$ for all $k \geq N$. For $k > l \geq N$ we have

$$\Big|\sum_{i=1}^{k} \alpha_i - \sum_{i=1}^{l} \alpha_i\Big| = \Big|\sum_{i=l+1}^{k} \alpha_i\Big| \overset{4.13}{\leq} \max\big\{|\alpha_i| : i = l+1, \dots, k\big\} < \epsilon,$$

$$\Big|\prod_{i=1}^{k}(1+\alpha_i) - \prod_{i=1}^{l}(1+\alpha_i)\Big| = \prod_{i=1}^{l}\underbrace{|1+\alpha_i|}_{\leq 1}\Big|\prod_{i=l+1}^{k}(1+\alpha_i) - 1\Big| = \Big|\sum_{\varnothing \neq I \subseteq \{l+1,\dots,k\}} \prod_{i \in I} \alpha_i\Big|$$
$$\leq \max\big\{|\alpha_i| : i = l+1, \dots, k\big\} < \epsilon. \qquad \square$$

Remark 4.16.

(i) Let $\alpha_1, \dots \in K[[X]]$ be a null sequence and $\alpha_k = \sum a_{k,n} X^n$. For each n, only finitely many of the coefficients $a_{1,n}, a_{2,n}, \dots$ are then different from 0. This shows

$$\sum \alpha_k = \sum_{n=0}^{\infty}\Big(\sum_{k=1}^{\infty} a_{k,n}\Big) X^n,$$

i. e. for the calculation of the coefficient of X^n, one only needs to evaluate finitely many terms. The same applies to $\prod_{k=1}^{\infty}(1+\alpha_k)$. For example,

$$(1+X)(1+X^2)(1+X^3)(1+X^4)\dots = 1 + X + X^2 + 2X^3 + 2X^4 + \dots.$$

(ii) For $\gamma \in K[[X]]$ and null sequences $\alpha_1, \dots, \beta_1, \dots$, it holds as usual that $\sum \alpha_k + \sum \beta_k = \sum(\alpha_k + \beta_k)$ and $\gamma \sum \alpha_k = \sum \gamma\alpha_k$.

(iii) It holds that

$$\Big(\sum a_n X^n\Big)\Big(\sum b_n X^n\Big) = \sum_{k,l \geq 0} a_k b_l X^{k+l},$$

because the right side converges regardless of the order in which the pairs (k, l) are summed.

Example 4.17. For $\alpha \in XK[[X]]$, we have $|\alpha^n| = |\alpha|^n \leq 2^{-n} \to 0$ and $\sum \alpha^n = \frac{1}{1-\alpha}$. We have thus replaced X by α in the geometric series.

Definition 4.18. For $\alpha = \sum a_n X^n \in K[[X]]$ and $\beta \in XK[[X]]$, one defines

$$\alpha \circ \beta := \alpha(\beta) := \sum_{n=0}^{\infty} a_n \beta^n.$$

Remark 4.19. For arbitrary $\alpha, \beta \in K[[X]]$, the 0-th coefficient $\sum_{k=0}^{\infty} a_0 b_0^k$ of $\alpha(\beta)$ would in general not be well-defined.

Example 4.20. For $\alpha = \sum a_n X^n \in K[[X]]$, it holds that $\alpha(X^2) = \sum a_n X^{2n}$ and $\alpha(0) = a_0$.

Lemma 4.21. *For $\alpha, \beta, \gamma \in K[[X]]$ and every null sequence $\alpha_1, \ldots \in K[[X]]$, it holds (if well-defined):*

$$X \circ \alpha = \alpha = \alpha \circ X, \tag{4.1}$$

$$\Big(\sum \alpha_k\Big) \circ \beta = \sum (\alpha_k \circ \beta), \tag{4.2}$$

$$(\alpha\beta) \circ \gamma = (\alpha \circ \gamma)(\beta \circ \gamma), \tag{4.3}$$

$$\alpha \circ (\beta \circ \gamma) = (\alpha \circ \beta) \circ \gamma. \tag{4.4}$$

Proof. Equation 4.1 is trivial. With the notation from Remark 4.16, it holds that

$$\Big(\sum \alpha_k\Big) \circ \beta = \sum_{n=0}^{\infty} \Big(\sum_{k=1}^{\infty} a_{k,n}\Big) \beta^n = \sum_{k=1}^{\infty} \Big(\sum_{n=0}^{\infty} a_{k,n} \beta^n\Big) = \sum (\alpha_k \circ \beta).$$

Equation 4.3 is obtained by

$$(\alpha\beta) \circ \gamma = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k} \gamma^n = \sum_{n=0}^{\infty} \sum_{k=0}^{n} (a_k \gamma^k)(b_{n-k} \gamma^{n-k}) = (\alpha \circ \gamma)(\beta \circ \gamma).$$

In (4.4), we may assume $\alpha = X^n$ according to (4.2). With (4.3), it follows that

$$\alpha \circ (\beta \circ \gamma) = (\beta \circ \gamma)^n = \beta^n \circ \gamma = (\alpha \circ \beta) \circ \gamma. \qquad \square$$

Remark 4.22. In general, $\alpha \circ \beta \neq \beta \circ \alpha$, $\alpha \circ (\beta\gamma) \neq (\alpha \circ \beta)(\alpha \circ \gamma)$ and $\alpha \circ (\beta + \gamma) \neq \alpha \circ \beta + \alpha \circ \gamma$ (Exercise 23). The last equation can be corrected for the exponential function.

Lemma 4.23 (Functional equation)**.** *For every null sequence $\alpha_1, \alpha_2, \ldots \in X\mathbb{Q}[[X]]$, it holds that*

$$\boxed{\exp\Big(\sum \alpha_k\Big) = \prod \exp(\alpha_k).} \tag{4.5}$$

In particular, $\exp(kX) = \exp(X)^k$ *for* $k \in \mathbb{Z}$.

Proof. Due to $\sum \alpha_k \in X\mathbb{Q}[[X]]$ and $\exp(\alpha_k) \in 1 + \alpha_k + \frac{\alpha_k^2}{2} + \ldots$, both sides of (4.5) are well-defined (Lemma 4.15). For two summands $\alpha, \beta \in X\mathbb{Q}[[X]]$, it holds that

$$\begin{aligned} \exp(\alpha + \beta) &= \sum \frac{(\alpha + \beta)^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} \frac{\alpha^k \beta^{n-k}}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{\alpha^k \beta^{n-k}}{k!(n-k)!} = \sum \frac{\alpha^n}{n!} \cdot \sum \frac{\beta^n}{n!} = \exp(\alpha) \exp(\beta). \end{aligned}$$

Inductively, one obtains (4.5) for finitely many summands. Finally,

$$\Big|\prod \exp(\alpha_k) - \exp\Big(\sum_{k=1}^{n} \alpha_k\Big)\Big| = \prod_{k=1}^{n} |\exp(\alpha_k)| \Big| \prod_{k=n+1}^{\infty} \exp(\alpha_k) - 1\Big| \to 0.$$

For $k \in \mathbb{N}_0$, $\exp(kX) = \exp(X + \ldots + X) = \exp(X)^k$. Due to $\exp(kX)\exp(-kX) = \exp(kX - kX) = \exp(0) = 1$, it holds that $\exp(-kX) = \exp(kX)^{-1} = \exp(X)^{-k}$. Therefore, the last assertion holds for all $k \in \mathbb{Z}$. $\square$

Theorem 4.24. *For every $\alpha = \sum a_n X^n \in K[[X]]$ with $a_0 = 0$ and $a_1 \neq 0$, there exists exactly one $\beta \in K[[X]]$ with $\beta(\alpha) = \alpha(\beta) = X$. Therefore, $K[[X]]^\circ := XK[[X]] \setminus X^2K[[X]]$ is a group with respect to $\circ$ with identity element X.*

Proof. Let $\alpha^k = \sum_{n=0}^{\infty} a_{kn} X^n$ for $k \in \mathbb{N}_0$. Due to $a_0 = 0$, $a_{kn} = 0$ for $n < k$ and $a_{nn} = a_1^n \neq 0$. We define inductively $b_0 := 0$, $b_1 := \frac{1}{a_1} \neq 0$ and

$$b_n := -\frac{1}{a_{nn}} \sum_{k=0}^{n-1} a_{kn} b_k$$

for $n \geq 2$. For $\beta := \sum b_n X^n \in K[[X]]^\circ$, it then holds that

$$\beta(\alpha) = \sum_{k=0}^{\infty} b_k \alpha^k = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} b_k a_{kn} X^n = \sum_{n=0}^{\infty} \Big(\sum_{k=0}^{n} b_k a_{kn} \Big) X^n = X.$$

Interchanging the roles of α and β, one obtains $\gamma \in K[[X]]^\circ$ with $\gamma(\beta) = X$. According to Lemma 4.21, it holds that

$$\alpha(\beta) = X \circ (\alpha \circ \beta) = (\gamma \circ \beta) \circ (\alpha \circ \beta) = \gamma \circ (\beta \circ \alpha) \circ \beta = \gamma \circ X \circ \beta = \gamma(\beta) = X.$$

Thus β is the inverse of α with respect to $\circ$. In particular, β is uniquely determined and $K[[X]]^\circ$ is a group. □

Remark 4.25. In the situation of Theorem 4.24, β is called the *inverse function* of α. Note: $\beta \neq \alpha^{-1}$ (we will not introduce a notation for the inverse function).

Example 4.26. Let α be the inverse function of $X + X^2 + \ldots = \frac{X}{1-X}$. Then

$$X = \frac{\alpha}{1 - \alpha}$$

and it follows that $\alpha = \frac{X}{1+X} = X - X^2 + X^3 - \ldots$.

Definition 4.27. For $\alpha = \sum a_n X^n \in K[[X]]$,

$$\alpha' := \sum_{n=1}^{\infty} n a_n X^{n-1} \in K[[X]]$$

is called the (formal) *derivative* of α. Furthermore, let $\alpha^{(0)} := \alpha$ and $\alpha^{(n)} := (\alpha^{(n-1)})'$ be the *n-th derivative* for $n \in \mathbb{N}$.

Example 4.28. It holds that $1' = 0$, $X' = 1$ as well as

$$\exp(X)' = \sum_{n=1}^{\infty} n \frac{X^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{X^n}{n!} = \exp(X).$$

Remark 4.29. With derivatives, the coefficients of $\alpha = \sum a_n X^n \in \mathbb{C}[[X]]$ can be calculated: $\alpha^{(0)}(0) = \alpha(0) = a_0$, $\alpha'(0) = a_1$, $\alpha''(0) = 2a_2, \ldots, \alpha^{(n)}(0) = n!a_n$. Therefore

$$\boxed{\alpha = \sum_{n=0}^{\infty} \frac{\alpha^{(n)}(0)}{n!} X^n} \qquad (Taylor series).$$

Over arbitrary fields K, one cannot always divide by $n!$. As a substitute, one can define the k-th *Hasse derivative*

$$H^k(\alpha) := \sum_{n=k}^{\infty} \binom{n}{k} a_n X^{n-k}$$

for $\alpha = \sum a_n X^n \in K[[X]]$. Analogously, it now holds that $\alpha = \sum_{n=0}^{\infty} H^n(\alpha)(0) X^n$.

Lemma 4.30. *For $\alpha, \beta \in K[[X]]$ and every null sequence $\alpha_1, \alpha_2, \ldots \in K[[X]]$, it holds that*

$$\begin{aligned}
\Big(\sum \alpha_k\Big)' &= \sum \alpha_k' && (\textit{sum rule}),\\
(\alpha\beta)' &= \alpha'\beta + \alpha\beta' && (\textit{product rule}),\\
\Big(\prod (1+\alpha_k)\Big)' &= \prod (1+\alpha_k) \sum \frac{\alpha_k'}{1+\alpha_k}, &&\\
\Big(\frac{\alpha}{\beta}\Big)' &= \frac{\alpha'\beta - \alpha\beta'}{\beta^2} && (\textit{quotient rule}),\\
(\alpha \circ \beta)' &= \alpha'(\beta)\beta' && (\textit{chain rule}).
\end{aligned}$$

Proof.

(i) With the notation from Remark 4.16, it holds that

$$\Big(\sum \alpha_k\Big)' = \Big(\sum_{n=0}^{\infty}\sum_{k=1}^{\infty} a_{k,n} X^n\Big)' = \sum_{n=0}^{\infty}\sum_{k=1}^{\infty} n a_{k,n} X^{n-1} = \sum_{k=1}^{\infty}\Big(\sum_{n=0}^{\infty} n a_{k,n} X^{n-1}\Big) = \sum \alpha_k'.$$

(ii) According to (i), one may assume $\alpha = X^k$ and $\beta = X^l$. Then

$$(\alpha\beta)' = (X^{k+l})' = (k+l)X^{k+l-1} = kX^{k-1}X^l + lX^{l-1}X^k = \alpha'\beta + \beta'\alpha.$$

(iii) Wlog. let $\alpha_k \neq -1$ for all $k \in \mathbb{N}$ (otherwise both sides are 0). From (ii) one obtains inductively

$$\Big(\prod_{k=1}^{n} (1+\alpha_k)\Big)' = \prod_{k=1}^{n} (1+\alpha_k) \sum_{k=1}^{n} \frac{\alpha_k'}{1+\alpha_k}$$

for all $n \in \mathbb{N}$. The claim follows with $n \to \infty$.

(iv) From (ii) it follows that

$$\alpha' = \Big(\frac{\alpha}{\beta}\beta\Big)' = \Big(\frac{\alpha}{\beta}\Big)'\beta + \frac{\alpha\beta'}{\beta}.$$

(v) According to (iii), $(\alpha^n)' = n\alpha^{n-1}\alpha'$ holds for $n \in \mathbb{N}_0$. From the sum rule it follows that

$$(\alpha \circ \beta)' = \Big(\sum a_n \beta^n\Big)' = \sum a_n (\beta^n)' = \sum_{n=1}^{\infty} n a_n \beta^{n-1}\beta' = \alpha'(\beta)\beta'. \qquad \square$$

Remark 4.31. The product rule also implies the *constant multiple rule* $(\lambda\alpha)' = \lambda\alpha'$ for $\lambda \in K$ and $\alpha \in K[[X]]$.

Example 4.32. We define the (formal) *logarithm* by the *Mercator series*

$$\boxed{\log(1+X) := \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} X^n = X - \frac{X^2}{2} + \frac{X^3}{3} \mp \ldots \in \mathbb{Q}[[X]].}$$

According to Theorem 4.24, $\alpha := \exp(X) - 1$ has an inverse function and $\log(\exp(X)) = \log(1+\alpha) \in \mathbb{Q}[[X]]^\circ$. Because of

$$\log(1+X)' = 1 - X + X^2 \mp \ldots = \sum (-X)^n = \frac{1}{1+X}$$

it follows from the chain rule that

$$\log(\exp(X))' = \frac{\alpha'}{1+\alpha} = \frac{\exp(X)}{1+\alpha} = \frac{1+\alpha}{1+\alpha} = 1.$$

This shows $\log(\exp(X)) = X$. Thus $\log(1+X)$ is the inverse function of $\alpha = \exp(X) - 1$ as in analysis. Furthermore, $\log(1-X) = -\sum_{n=1}^{\infty} \frac{X^n}{n}$ holds.

Lemma 4.33 (Functional equation)**.** *For every null sequence $\alpha_1, \alpha_2, \ldots \in X\mathbb{Q}[[X]]$ it holds that*

$$\boxed{\log\Big(\prod(1+\alpha_k)\Big) = \sum \log(1+\alpha_k).}$$

Proof.

$$\begin{aligned} \log\Big(\prod(1+\alpha_k)\Big) &= \log\Big(\prod \exp(\log(1+\alpha_k))\Big) \overset{4.23}{=} \log\big(\exp\Big(\sum \log(1+\alpha_k)\Big)\big) \\ &= \sum \log(1+\alpha_k). \end{aligned}$$

□

Example 4.34. According to Lemma 4.33 it holds that

$$\begin{aligned} \log\Big(\frac{1}{1-X}\Big) &= \log\Big(\frac{1}{1-X}\Big) + \log(1-X) - \log(1-X) = \log\Big(\frac{1}{1-X}(1-X)\Big) - \log(1-X) \\ &= \log(1) - \log(1-X) = -\log(1-X) = \sum_{n=1}^{\infty} \frac{X^n}{n}. \end{aligned}$$

Definition 4.35. For $c \in \mathbb{C}$ and $\alpha \in X\mathbb{C}[[X]]$ we define

$$(1+\alpha)^c := \exp(c \log(1+\alpha)).$$

In the case $c = 1/k$ with $k \in \mathbb{N}$ we write $\sqrt[k]{1+\alpha} := (1+\alpha)^{1/k}$ and specifically $\sqrt{1+\alpha} := \sqrt[2]{1+\alpha}$.

Remark 4.36.

(i) According to the functional equation it holds that

$$(1+\alpha)^c (1+\alpha)^d = \exp(c\log(1+\alpha) + d\log(1+\alpha)) = (1+\alpha)^{c+d}$$

for $c, d \in \mathbb{C}$ as usual. For $k \in \mathbb{N}$ it is therefore $\sqrt[k]{1+\alpha}^k = 1+\alpha$, i. e. $\sqrt[k]{1+\alpha}$ is a k-th *root* of $1+\alpha$ with constant term 1. Let also $\beta \in \mathbb{C}[[X]]$ with $\beta^k = 1+\alpha$. Then $\sqrt[k]{1+\alpha}\beta^{-1}$ has order $\leq k$ in $\mathbb{C}[[X]]^\times$. From Lemma 4.8 it follows that $\sqrt[k]{1+\alpha}\beta^{-1}$ is constant, i. e. $\beta = \beta(0)\sqrt[k]{1+\alpha}$. Therefore $\sqrt[k]{1+\alpha}$ is the unique k-th root of $1+\alpha$ with constant term 1.

(ii) The following theorem generalizes both the binomial theorem ($c \in \mathbb{N}$) and the geometric series ($c = -1$).

Theorem 4.37 (Newton's Binomial Theorem)**.** *For $\alpha \in X\mathbb{C}[[X]]$ and $c \in \mathbb{C}$ it holds that*

$$\boxed{(1+\alpha)^c = \sum_{k=0}^{\infty} \binom{c}{k} \alpha^k.}$$

Proof. It suffices to prove the claim for $\alpha = X$. According to the chain rule it holds that

$$((1+X)^c)' = (\exp(c\log(1+X))' = c\frac{(1+X)^c}{1+X} = c(1+X)^{c-1}$$

and inductively it follows that $((1+X)^c)^{(k)} = c(c-1)\dots(c-k+1)(1+X)^{c-k}$. The claim now follows from the Taylor series. □

Example 4.38. Let $\zeta \in \mathbb{C}$ with $\zeta^n = 1$ (cf. Definition 6.29) and $\alpha := (1+X)^\zeta - 1 \in X\mathbb{C}[[X]]$. Then $\alpha \circ \alpha = (1 + (1+X)^\zeta - 1)^\zeta - 1 = (1+X)^{\zeta^2} - 1$ and inductively $\alpha \circ \ldots \circ \alpha = (1+X)^{\zeta^n} - 1 = X$, i. e. the order of α in the group $\mathbb{C}[[X]]^\circ$ divides n. In contrast to Lemma 4.8, $\mathbb{C}[[X]]^\circ$ thus possesses "interesting" elements of finite order.

Definition 4.39. For $n \in \mathbb{N}_0$ let $X^n! := (1-X)(1-X^2)\dots(1-X^n)$. For $0 \le k \le n$ one calls

$$\left\langle {n \atop k} \right\rangle := \frac{X^n!}{X^k! X^{n-k}!} = \frac{1-X^n}{1-X^k} \cdots \frac{1-X^{n-k+1}}{1-X} \in K[[X]]$$

Gaussian binomial coefficients. For $k < 0$ or $k > n$ let $\left\langle {n \atop k} \right\rangle := 0$.

Remark 4.40. As always, $\left\langle {n \atop 0} \right\rangle = \left\langle {n \atop n} \right\rangle = 1$ and $\left\langle {n \atop k} \right\rangle = \left\langle {n \atop n-k} \right\rangle$ for all $n \in \mathbb{N}_0$ and $k \in \mathbb{Z}$. Furthermore, $\left\langle {n \atop 1} \right\rangle = \frac{1-X^n}{1-X} = 1 + X + \ldots + X^{n-1}$ and

$$\left\langle {4 \atop 2} \right\rangle = \frac{(1-X^4)(1-X^3)}{(1-X^2)(1-X)} = (1+X^2)\frac{1-X^3}{1-X} = (1+X^2)(1+X+X^2) = 1 + X + 2X^2 + X^3 + X^4.$$

Lemma 4.41. *For $n \in \mathbb{N}_0$ and $k \in \mathbb{Z}$, it holds that*

$$\left\langle {n+1 \atop k} \right\rangle = X^k \left\langle {n \atop k} \right\rangle + \left\langle {n \atop k-1} \right\rangle = \left\langle {n \atop k} \right\rangle + X^{n+1-k} \left\langle {n \atop k-1} \right\rangle.$$

Proof. For $k > n+1$ or $k < 0$, all terms are 0. For $k = n+1$ or $k = 0$, both sides are 1. For $1 \le k \le n$, it holds that

$$\begin{aligned} X^k \left\langle {n \atop k} \right\rangle + \left\langle {n \atop k-1} \right\rangle &= \Big(X^k \frac{1-X^{n-k+1}}{1-X^k} + 1\Big) \frac{X^n!}{X^{k-1}! X^{n-k+1}!} = \frac{1-X^{n+1}}{1-X^k} \frac{X^n!}{X^{k-1}! X^{n+1-k}!} \\ &= \left\langle {n+1 \atop k} \right\rangle = \left\langle {n+1 \atop n+1-k} \right\rangle = X^{n+1-k} \left\langle {n \atop n+1-k} \right\rangle + \left\langle {n \atop n-k} \right\rangle \\ &= \left\langle {n \atop k} \right\rangle + X^{n+1-k} \left\langle {n \atop k-1} \right\rangle. \end{aligned}$$

□

Remark 4.42. One may compare the recursion formulas[2]

$$\begin{aligned}
\binom{n+1}{k} &\overset{1.6}{=} \binom{n}{k-1} + \binom{n}{k},\\
\left(\!\binom{n+1}{k}\!\right) &\overset{1.24}{=} \left(\!\binom{n+1}{k-1}\!\right) + \left(\!\binom{n}{k}\!\right),\\
\begin{bmatrix} n+1 \\ k \end{bmatrix} &\overset{2.16}{=} \begin{bmatrix} n \\ k-1 \end{bmatrix} + n \begin{bmatrix} n \\ k \end{bmatrix},\\
\left\{ {n+1 \atop k} \right\} &\overset{2.31}{=} \left\{ {n \atop k-1} \right\} + k \left\{ {n \atop k} \right\},\\
\left\langle {n+1 \atop k} \right\rangle &\overset{4.41}{=} \left\langle {n \atop k-1} \right\rangle + X^k \left\langle {n \atop k} \right\rangle.
\end{aligned}$$

Thus, one can calculate $\left(\!\binom{n}{k}\!\right)$, $\begin{bmatrix} n \\ k \end{bmatrix}$, $\left\{ {n \atop k} \right\}$ and $\left\langle {n \atop k} \right\rangle$ with a modified Pascal's triangle.

Theorem 4.43 (Gauss's Binomial Theorem)**.** *For $n \in \mathbb{N}$ and $\alpha \in K[[X]]$, it holds that*

$$\boxed{\prod_{k=0}^{n-1} (1 + \alpha X^k) = \sum_{k=0}^{n} \left\langle {n \atop k} \right\rangle \alpha^k X^{\binom{k}{2}}.}$$

Proof. Induction on n: For $n = 1$, both sides are $1 + \alpha$. For the induction step, we let all sums run from $-\infty$ to ∞ (this makes index shifts easier):

$$\begin{aligned}
\prod_{k=0}^{n} (1 + \alpha X^k) &= (1 + \alpha X^n) \sum_{k=-\infty}^{\infty} \left\langle {n \atop k} \right\rangle \alpha^k X^{\binom{k}{2}}\\
&= \sum \left\langle {n \atop n-k} \right\rangle \alpha^k X^{\binom{k}{2}} + \sum \left\langle {n \atop n-k} \right\rangle \alpha^{k+1} X^{n-k} X^{\overbrace{\binom{k}{2}+k}^{\binom{k+1}{2}}}\\
&= \sum \left\langle {n \atop n-k} \right\rangle \alpha^k X^{\binom{k}{2}} + \sum X^{n-k+1} \left\langle {n \atop n-k+1} \right\rangle \alpha^k X^{\binom{k}{2}}\\
&\overset{4.41}{=} \sum \left\langle {n+1 \atop n-k+1} \right\rangle \alpha^k X^{\binom{k}{2}} = \sum \left\langle {n+1 \atop k} \right\rangle \alpha^k X^{\binom{k}{2}}.
\end{aligned}$$

□

Remark 4.44. There is also a Vandermonde identity for $\left\langle {n \atop k} \right\rangle$.

5. Generating Functions

Definition 5.1. In this chapter, let always $\mathbb{Q} \subseteq K$. For a sequence of numbers $a_0, a_1, \ldots \in \mathbb{C}$, one calls $\sum_{n=0}^{\infty} a_n X^n \in \mathbb{C}[[X]]$ the *generating function* of $a_0, a_1, \ldots$.

Example 5.2.

(i) The generating function of the constant sequence $1, 1, \ldots$ is $(1 - X)^{-1}$.

[2]All formulas can be unified: [J. Konvalina, *A unified interpretation of the binomial coefficients, the Stirling numbers, and the Gaussian coefficients*, Amer. Math. Monthly 107, (2000), 901–910]

(ii) If $\alpha \in \mathbb{C}[[X]]$ is the generating function of $a_0, a_1, \ldots$, then $\alpha(-X)$ is the generating function of $a_0, -a_1, a_2, \ldots$.

(iii) If α is the generating function of $a_0, a_1, \ldots$, then α' is the generating function of $0, a_1, 2a_2, 3a_3, \ldots$. For example, $\left(\frac{1}{1-X}\right)' = \frac{1}{(1-X)^2}$ (quotient rule) is the generating function of $0, 1, 2, 3, \ldots$.

(iv) Every k-element multiset $A \subseteq \{1, \ldots, n\}$ corresponds to exactly one partition $k = k_1 + \ldots + k_n$, where $k_i \in \mathbb{N}_0$ denotes the multiplicity of i in A. This shows

$$\frac{1}{(1-X)^n} = \Big(\sum_{k=0}^{\infty} X^k\Big)^n = \sum_{k=0}^{\infty}\Big(\sum_{\substack{(k_1,\ldots,k_n)\in\mathbb{N}_0^n \\ k_1+\ldots+k_n=k}} 1\Big)X^k = \sum_{k=0}^{\infty}\left(\binom{n}{k}\right)X^k$$

for $n \in \mathbb{N}$ (cf. Theorem 4.37).

(v) We consider the recursively defined *Fibonacci sequence* $f_0 := 0$, $f_1 := 1$ and $f_{n+1} := f_{n-1} + f_n$ for $n \in \mathbb{N}$ (i.e., $0, 1, 1, 2, 3, 5, 8, \ldots$). For $\alpha := \sum f_n X^n \in \mathbb{R}[[X]]$, it then holds that

$$\alpha = X + \sum_{n=2}^{\infty} f_n X^n = X + \sum_{n=2}^{\infty}(f_{n-2} + f_{n-1})X^n = X + X^2\alpha + X\alpha.$$

Thus

$$\alpha = \frac{X}{1 - X - X^2}.$$

Theorem 5.3 (BINET formula)**.** *For $n \in \mathbb{N}_0$, it holds that*

$$\boxed{f_n = \frac{1}{\sqrt{5}}\Big(\frac{1+\sqrt{5}}{2}\Big)^n - \frac{1}{\sqrt{5}}\Big(\frac{1-\sqrt{5}}{2}\Big)^n.}$$

Proof. Let α be as in Example 5.2. If $\varphi := \frac{1+\sqrt{5}}{2} \in \mathbb{R}$ is the *golden ratio* and $\psi := \frac{1-\sqrt{5}}{2} \in \mathbb{R}$, then $(X + \varphi)(X + \psi) = X^2 + X - 1$. Partial fraction decomposition (Example 4.10) yields

$$\alpha = \frac{-X}{(X+\varphi)(X+\psi)} = \frac{X}{\varphi - \psi}\Big(\frac{1}{X+\varphi} - \frac{1}{X+\psi}\Big) = \frac{X}{\sqrt{5}}\Big(\frac{1}{X+\varphi} - \frac{1}{X+\psi}\Big).$$

Now $\varphi^{-1} = \frac{2}{1+\sqrt{5}} = -\psi$ and

$$\frac{X}{X+\varphi} \overset{4.10}{=} -\sum(-\varphi)^{-n-1}X^{n+1} = -\sum_{n=1}^{\infty}\psi^n X^n.$$

It follows that

$$\alpha = \frac{1}{\sqrt{5}}\Big(\sum_{n=1}^{\infty}\varphi^n X^n - \sum_{n=1}^{\infty}\psi^n X^n\Big) = \sum_{n=0}^{\infty}\frac{1}{\sqrt{5}}(\varphi^n - \psi^n)X^n.$$

Comparison of coefficients yields the assertion. □

Remark 5.4.

(i) Since the second summand in Theorem 5.3 is smaller than $\frac{1}{2}$, one obtains the simpler formula by rounding

$$f_n = \left[\frac{1}{\sqrt{5}}\Big(\frac{1+\sqrt{5}}{2}\Big)^n\right].$$

(ii) One can also interpret f_n combinatorially: Let g_n be the number of $\{1,2\}$-sequences whose sum yields n. Certainly $g_0 = 0$ and $g_1 = 1$. If $(a_1, \ldots, a_k) \in \{1,2\}^k$ with $\sum a_i = n+1$, then $(a_1, \ldots, a_{k-1})$ is a sequence with sum n or $n-1$, depending on whether $a_k = 1$ or $a_k = 2$ holds. This shows $g_{n+1} = g_n + g_{n-1}$ and it follows that $g_n = f_n$ for all $n \in \mathbb{N}_0$.

Theorem 5.5. *It holds that*

$$(i) \qquad \boxed{\sum_{n=0}^{\infty} p(n) X^n = \prod_{k=1}^{\infty} \frac{1}{1-X^k}} \qquad \text{(Euler)},$$

$$(ii) \qquad \boxed{\sum_{n=0}^{\infty} \frac{b(n)}{n!} X^n = \exp(\exp(X) - 1).}$$

Proof.

(i) Because of $(1-X^k)^{-1} = \sum_{n=0}^{\infty} (X^k)^n \in 1 + X^k K[[X]]$, the infinite product is well-defined (Lemma 4.15). The partition $(1^{a_1}, \ldots, n^{a_n})$ of $n \in \mathbb{N}_0$ satisfies $a_1 + 2a_2 + \ldots + na_n = n$. Therefore $p(n)$ is the number of all solutions $(a_1, \ldots, a_n) \in \mathbb{N}_0^n$ with $a_1 + 2a_2 + \ldots + na_n = n$. This is exactly the n-th coefficient of

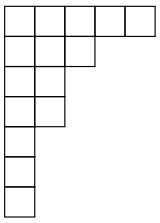
$$(1 + X + X^2 + X^3 + \ldots)(1 + X^2 + X^{2\cdot 2} + X^{2\cdot 3} + \ldots)(1 + X^3 + X^{3\cdot 2} + X^{3\cdot 3} + \ldots) \ldots = \prod_{k=1}^{\infty} \frac{1}{1-X^k}.$$

(ii) Let $\alpha := \exp(\exp(X) - 1) = \sum \frac{a_n}{n!} X^n$. Then $a_0 = \exp(\exp(0) - 1) = \exp(0) = 1 = b(0)$. The chain rule yields

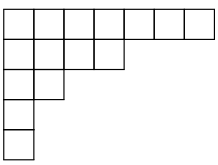
$$\begin{aligned} \sum_{n=0}^{\infty} \frac{a_{n+1}}{n!} X^n = \alpha' &= \exp(X) \exp(\exp(X) - 1) \\ &= \Big(\sum_{k=0}^{\infty} \frac{1}{k!} X^k \Big) \Big(\sum_{k=0}^{\infty} \frac{a_k}{k!} X^k \Big) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{a_k}{k!(n-k)!} X^n. \end{aligned}$$

Therefore $a_{n+1} = \sum_{k=0}^{n} \binom{n}{k} a_k$ for $n \geq 0$ and the assertion follows from Theorem 2.40. □

Remark 5.6. One can visualize partitions $\lambda = (\lambda_1, \ldots, \lambda_k)$ (with $\lambda_1 \geq \ldots \geq \lambda_k$) of $n \in \mathbb{N}$ by *Young diagrams* (also called *Ferrers diagrams*). For example:

$(5, 3, 2^2, 1^3) =$ 

By reflection across the diagonal, one obtains the Young diagram of the *conjugate* partition $\lambda' = (\lambda'_1, \ldots, \lambda'_l)$ of n. For example:

$(5, 3, 2^2, 1^3)' =$  $= (7, 4, 2, 1^2)$

In general, $\lambda'_i = |\{j : \lambda_j \geq i\}|$ for $i = 1, \ldots, l$. Furthermore, $\lambda'' = \lambda$. One calls λ *symmetric*, if $\lambda' = \lambda$.

Theorem 5.7. *Let $n, k \in \mathbb{N}_0$.*

(i) (Euler) *The number of partitions of n into distinct parts is equal to the number of partitions into odd parts.*

(ii) The number of partitions of n into k parts is equal to the number of partitions with largest part k.

(iii) (Sylvester) *The number of symmetric partitions of n is equal to the number of partitions into distinct, odd parts.*

Proof.

(i) If $u(n)$ is the number of partitions into distinct parts, then

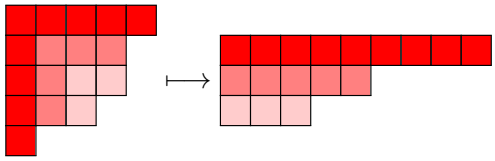
$$\begin{aligned}\sum u(n)X^n &= (1+X)(1+X^2)(1+X^3)\ldots = \frac{(1-X^2)(1-X^4)}{(1-X)(1-X^2)}\ldots \\ &= \frac{1}{(1-X)(1-X^3)(1-X^5)\ldots} \\ &= (1+X+X^2+\ldots)(1+X^3+X^6+\ldots)\ldots.\end{aligned}$$

On the right side stands the generating function of the number of partitions into odd parts.

(ii) The map $\lambda \mapsto \lambda'$ provides a bijection between the specified sets.

(iii) The following map provides the desired bijection:

$$\begin{aligned}\{\text{symmetric partitions of } n\} &\longrightarrow \{\text{partitions into distinct, odd parts}\}, \\ (\lambda_1, \ldots, \lambda_k) &\longmapsto (2\lambda_1 - 1, 2\lambda_2 - 3, 2\lambda_3 - 5, \ldots)\end{aligned}$$



□

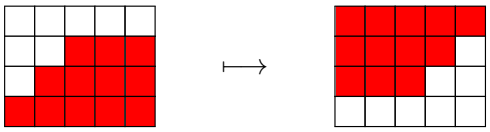
Example 5.8. For $n = 8$ one obtains

Partitions into distinct parts:	(8), $(7,1)$, $(6,2)$, $(5,3)$, $(5,2,1)$, $(4,3,1)$
Partitions into odd parts:	$(7,1)$, $(5,3)$, $(5,1^3)$, $(3^2,1^2)$, $(3,1^5)$, (1^8)
Partitions into distinct, odd parts:	$(7,1)$, $(5,3)$
Symmetric partitions:	$(3^2,2)$, $(4,2,1^2)$
Partitions into 4 parts:	$(5,1^3)$, $(4,2,1^2)$, $(3^2,1^2)$, $(3,2^2,1)$, (2^4)
Partitions with largest part 4:	(4^2), $(4,3,1)$, $(4,2^2)$, $(4,2,1^2)$, $(4,1^4)$

Remark 5.9. Let $p_k(n)$ be the number of partitions of n into parts $\leq k$ (according to Theorem 5.7 this is also the number of partitions into at most k parts). The proof of Theorem 5.5 shows

$$\sum_{n=0}^{\infty} p_k(n)X^n = (1+X+X^2+\ldots)(1+X^2+X^4+\ldots)\ldots(1+X^k+X^{2k}+\ldots) = \frac{1}{X^{k!}}.$$

We now study the number $p_{k,l}(n) = p_{l,k}(n)$ of partitions of n with at most k parts and each part $\leq l$. These are exactly the partitions whose Young diagram fits into a rectangle of size $k \times l$. The remaining part of the rectangle rotated by 180° yields a partition of $kl - n$ of the same format:



This shows $p_{k,l}(n) = p_{k,l}(kl - n)$.

Theorem 5.10. *For $k, l \geq 0$ it holds that*

$$\sum_{n=0}^{\infty} p_{k,l}(n) X^n = \left\langle \begin{matrix} k+l \\ k \end{matrix} \right\rangle.$$

Proof. Induction on $k + l$. For $k = 0$ or $l = 0$ both sides are equal to 1. So let $k, l \geq 1$. Let $\lambda = (\lambda_1, \lambda_2, \ldots) \in P(n)$ with at most k parts and each part $\leq l$. If $\lambda_1 = l$, then $(\lambda_2, \lambda_3, \ldots) \in P(n-l)$ with at most $k-1$ parts. Otherwise, every part of λ is at most $l-1$. This shows $p_{k,l}(n) = p_{k,l-1}(n) + p_{k-1,l}(n-l)$. For $P(k, l) := \sum p_{k,l}(n) X^n$ it therefore holds that

$$P(k, l) = P(k, l-1) + X^l P(k-1, l).$$

The claim now follows by induction and Lemma 4.41. □

Remark 5.11. For $k, N \geq 0$ it holds that

$$\sum_{n=0}^{\infty} (p_k(n) - p_{k,N-k}(n)) X^n = \sum_{n=N-k+1}^{\infty} (p_k(n) - p_{k,N-k}(n)) X^n \longrightarrow 0 \qquad (N \to \infty)$$

and

$$\lim_{N \to \infty} \left\langle \begin{matrix} N \\ k \end{matrix} \right\rangle = \sum_{n=0}^{\infty} p_k(n) X^n = \frac{1}{X^k!}.$$

For $k, l \in \mathbb{Z}$ it holds analogously

$$\lim_{N \to \infty} \left\langle \begin{matrix} 2N+k \\ N+l \end{matrix} \right\rangle = \lim_{N \to \infty} \sum_{n=0}^{\infty} p_{N+l,N+k-l}(n) X^n = \sum_{n=0}^{\infty} p(n) X^n = \prod_{m=1}^{\infty} \frac{1}{1 - X^m}.$$

Both limits can also be derived from the definition of $\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle$.

Theorem 5.12 (Erdős-Turán)**.** *Let $n, d \in \mathbb{N}$. The number of permutations of S_n whose cycle lengths are not divisible by d is*

$$n! \prod_{k=1}^{\lfloor n/d \rfloor} \frac{kd - 1}{kd}.$$

Proof (Pólya). The number of permutations of type $(1^{l_1}, \ldots, n^{l_n})$ is

$$\frac{n!}{1^{l_1} \ldots n^{l_n} l_1! \ldots l_n!}$$

according to Theorem 2.26. The sought number, divided by $n!$, is therefore the coefficient of X^n in

$$\prod_{\substack{k=1\\ d\mid k}}^{\infty}\sum_{l=0}^{\infty}\frac{1}{l!}\Big(\frac{X^k}{k}\Big)^l = \prod_{d\mid k}\exp(X^k/k) \overset{4.23}{=} \exp\Big(\sum_{d\mid k}\frac{X^k}{k}\Big) = \exp\Big(\sum_{k=1}^{\infty}\frac{X^k}{k} - \sum_{k=1}^{\infty}\frac{X^{dk}}{dk}\Big)$$
$$= \exp(-\log(1-X) + \frac{1}{d}\log(1-X^d)) \overset{4.33}{=} \sqrt[d]{1-X^d}\,\frac{1}{1-X}$$
$$= \frac{1-X^d}{1-X}(1-X^d)^{\frac{1-d}{d}} \overset{4.37}{=} \Big(\sum_{r=0}^{d-1}X^r\Big)\Big(\sum_{q=0}^{\infty}\binom{(1-d)/d}{q}(-X^d)^q\Big).$$

In this, X^n occurs if and only if $n = qd + r$ with $0 \le r < d$ and $q = \lfloor n/d \rfloor$ (division with remainder). The coefficient is then

$$(-1)^q\binom{(1-d)/d}{q} = (-1)^q\prod_{k=1}^{q}\frac{\frac{1}{d}-k}{k} = \prod_{k=1}^{q}\frac{kd-1}{kd}. \qquad \square$$

Example 5.13. A permutation has odd order if and only if it consists only of cycles of odd length. The number of permutations in S_n with odd order is therefore

$$n!\prod_{k=1}^{\lfloor n/2\rfloor}\frac{2k-1}{2k} = \begin{cases} 1^2\cdot 3^2\cdot\ldots\cdot(n-1)^2 & \text{if } n \text{ is even,}\\ 1^2\cdot 3^2\cdot\ldots\cdot(n-2)^2\cdot n & \text{if } n \text{ is odd.}\end{cases}$$

Theorem 5.14 (Euler's Pentagonal Number Theorem)**.**

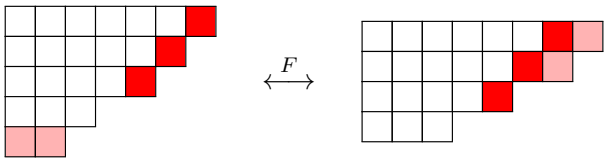
$$\prod_{k=1}^{\infty}(1-X^k) = 1+\sum_{k=1}^{\infty}(-1)^k\Big(X^{\frac{3k^2-k}{2}} + X^{\frac{3k^2+k}{2}}\Big) = \sum_{k=-\infty}^{\infty}(-1)^k X^{\frac{3k^2+k}{2}}$$
$$= 1 - X - X^2 + X^5 + X^7 - X^{12} - X^{15} + \ldots.$$

Proof (Franklin). Let $n \in \mathbb{N}$ and Λ_n be the set of partitions of n into distinct parts. For $\lambda \in \Lambda_n$ let $|\lambda|$ be the number of parts of λ. The n-th coefficient of $(1-X)(1-X^2)\ldots$ is then $\sum_{\lambda\in\Lambda_n}(-1)^{|\lambda|}$ (cf. proof of Theorem 5.7). Suppose first $n \ne (3k^2 \pm k)/2$ for all $k \in \mathbb{N}$. We construct a permutation F on Λ_n with $|F(\lambda)| = |\lambda| \pm 1$ for all $\lambda \in \Lambda_n$. Then it follows

$$\sum_{\lambda\in\Lambda_n}(-1)^{|\lambda|} = \sum_{\lambda\in\Lambda_n}(-1)^{|F(\lambda)|} = -\sum_{\lambda\in\Lambda_n}(-1)^{|\lambda|} = 0$$

as desired. Let $\lambda = (\lambda_1, \ldots, \lambda_l) \in \Lambda_n$ with $\lambda_1 > \ldots > \lambda_l$ and $s := \max\{1 \le i \le l : \lambda_i = \lambda_1 - i + 1\}$. We define

$$F(\lambda) := \begin{cases}(\lambda_1 - 1, \ldots, \lambda_s - 1, \lambda_{s+1}, \ldots, \lambda_l, s) & \text{if } s < \lambda_l,\\ (\lambda_1 + 1, \ldots, \lambda_{\lambda_l} + 1, \lambda_{\lambda_l+1}, \ldots, \lambda_{l-1}) & \text{if } s \ge \lambda_l.\end{cases}$$



This fails only in two cases: In the first case $\lambda = (2k-1, 2k-2, \ldots, k)$ and

$$n = \sum_{i=k}^{2k-1} i = \binom{2k}{2} - \binom{k}{2} = \frac{3k^2-k}{2}.$$

In the second case $\lambda = (2k, 2k-1, \ldots, k+1)$ and

$$n = \sum_{i=k+1}^{2k} i = \binom{2k+1}{2} - \binom{k+1}{2} = \frac{3k^2+k}{2}.$$

Both were excluded. Thus F is well-defined and $|F(\lambda)| = |\lambda| \pm 1$ for all $\lambda \in \Lambda_n$. Since $F^2 = \mathrm{id}$, F is a permutation.

If now $n = (3k^2 \pm k)/2$, then F can still be defined on $\Lambda_n \setminus \{\mu\}$, where μ is one of the two partitions mentioned above. One then obtains

$$\sum_{\lambda \in \Lambda_n} (-1)^{|\lambda|} = (-1)^{|\mu|} + \sum_{\lambda \in \Lambda_n \setminus \{\mu\}} (-1)^{|F(\lambda)|} = (-1)^k - \sum_{\lambda \in \Lambda_n \setminus \{\mu\}} (-1)^{|\lambda|} = (-1)^k$$

as desired. □

Remark 5.15. From Theorems 5.5 and 5.14 it follows that

$$\sum_{n=0}^{\infty} p(n) X^n \cdot \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{3k^2+k}{2}} = 1$$

and

$$\sum_{k=-n}^{n} (-1)^k p\Big(n - \frac{3k^2+k}{2}\Big) = 0$$

for $n \in \mathbb{N}$, where $p(k) := 0$ for $k < 0$. One obtains a recursion formula:

$$\begin{aligned} &p(0) = 1, \\ &p(n) = p(n-1) + p(n-2) - p(n-5) - p(n-7) + \ldots \qquad (n \in \mathbb{N}). \end{aligned}$$

Example 5.16. It holds that

$$\begin{aligned} p(1) &= p(0) = 1, \\ p(2) &= p(1) + p(0) = 2, \\ p(3) &= p(2) + p(1) = 3, \\ p(4) &= p(3) + p(2) = 3 + 2 = 5, \\ p(5) &= p(4) + p(3) - p(0) = 5 + 3 - 1 = 7, \\ p(6) &= p(5) + p(4) - p(1) = 7 + 5 - 1 = 11 \end{aligned}$$

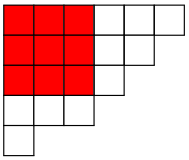
(cf. `https://oeis.org/A000041`).

Theorem 5.17 (Durfee's Square Theorem)**.** *It holds that*

$$\sum_{n=0}^{\infty} p(n) X^n = \sum_{k=0}^{\infty} \frac{X^{k^2}}{(X^{k!})^2}.$$

Proof. Let $\lambda \in P(n)$ and $k \in \mathbb{N}$ be maximal with $\lambda_k \geq k$. This means that the Young diagram of λ contains a square Q with side length k, but none with side length $k+1$. (One calls Q the *Durfee square*

of λ). Below Q there is a partition with largest part $\leq k$. To the right of Q there is a partition with at most k parts.



The number of both partitions is counted by p_k. Therefore

$$\sum_{n=0}^{\infty} p(n)X^n = \sum_{k=0}^{\infty} X^{k^2}\Big(\sum_{n=0}^{\infty} p_k(n)X^n\Big)^2 \overset{5.9}{=} \sum_{k=0}^{\infty} \frac{X^{k^2}}{(X^k!)^2}. \qquad \square$$

Lemma 5.18 (Hirschhorn). *For $n \in \mathbb{N}_0$ it holds that*

$$\prod_{k=1}^{n}(1-X^k)^2 = \sum_{k=0}^{n}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}\left\langle \begin{matrix} 2n+1 \\ n-k \end{matrix} \right\rangle. \tag{5.1}$$

Proof. Induction on n: For $n = 0$ both sides are 1. Now let $n \geq 1$. Let Q_n be the right side of (5.1). Repeated application of Lemma 4.41 yields

$$\begin{aligned}
Q_n &= X^n \sum_{k=0}^{n}(-1)^k(2k+1)X^{\frac{k^2-k}{2}}\left\langle \begin{matrix} 2n \\ n-k \end{matrix} \right\rangle + \sum_{k=0}^{n-1}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}\left\langle \begin{matrix} 2n \\ n-k-1 \end{matrix} \right\rangle \\
&= X^n \sum_{k=0}^{n}(-1)^k(2k+1)X^{\frac{k^2-k}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k \end{matrix} \right\rangle + X^{2n} \sum_{k=0}^{n-1}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle \\
&\quad + \sum_{k=0}^{n-1}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle + X^n \sum_{k=0}^{n-2}(-1)^k(2k+1)X^{\frac{k^2+3k+2}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k-2 \end{matrix} \right\rangle \\
&= (1+X^{2n})Q_{n-1} + X^n\Bigg(\left\langle \begin{matrix} 2n-1 \\ n \end{matrix} \right\rangle - 3\left\langle \begin{matrix} 2n-1 \\ n-1 \end{matrix} \right\rangle + \sum_{k=2}^{n}(-1)^k(2k+1)X^{\frac{k^2-k}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k \end{matrix} \right\rangle \\
&\quad + \sum_{k=0}^{n-2}(-1)^k(2k+1)X^{\frac{k^2+3k+2}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k-2 \end{matrix} \right\rangle\Bigg) \\
&= (1+X^{2n})Q_{n-1} + X^n\Bigg(-2\left\langle \begin{matrix} 2n-1 \\ n-1 \end{matrix} \right\rangle + \sum_{k=1}^{n-1}(-1)^{k+1}(2k+3)X^{\frac{k^2+k}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle \\
&\quad + \sum_{k=1}^{n-1}(-1)^{k-1}(2k-1)X^{\frac{k^2+k}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle\Bigg) \\
&= (1+X^{2n})Q_{n-1} - 2X^n\Bigg(\left\langle \begin{matrix} 2n-1 \\ n-1 \end{matrix} \right\rangle + \sum_{k=1}^{n-1}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}\left\langle \begin{matrix} 2n-1 \\ n-k-1 \end{matrix} \right\rangle\Bigg) \\
&= (1+X^{2n})Q_{n-1} - 2X^nQ_{n-1} = (1-X^n)^2Q_{n-1} = \prod_{k=1}^{n}(1-X^k)^2.
\end{aligned}$$

$\square$

Theorem 5.19 (Jacobi). *It holds that*

$$\prod_{k=1}^{\infty}(1-X^k)^3 = \sum_{k=0}^{\infty}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}.$$

Proof. The coefficient of X^n in $\prod_{k=1}^{\infty}(1-X^k)^3$ obviously depends only on the first n factors. According to Lemma 5.18, it holds that

$$\begin{aligned}\prod_{k=1}^{n}(1-X^k)^3 &= \prod_{k=1}^{n}(1-X^k)\sum_{k=0}^{n}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}\left\langle\begin{matrix}2n+1\\ n-k\end{matrix}\right\rangle \\ &= \sum_{k=0}^{n}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}(1-X^{n-k+1})\dots(1-X^n)(1-X^{n+k+2})\dots(1-X^{2n+1}).\end{aligned}$$

Because $\frac{1}{2}(k^2+k)+n-k+1 = n+\frac{1}{2}(k^2-k)+1 > n$ for $k \geq 0$, there exists an $\alpha \in K[[X]]$ with

$$\prod_{k=1}^{n}(1-X^k)^3 = \sum_{k=0}^{n}(-1)^k(2k+1)X^{\frac{k^2+k}{2}} + \alpha$$

and $|\alpha| < 2^n$. In particular, X^n does not occur in α. □

Remark 5.20.

(i) For $\alpha = \sum a_n X^n, \beta = \sum b_n X^n \in \mathbb{Z}[[X]]$ and $d \in \mathbb{N}$ we write $\alpha \equiv \beta \pmod{d}$ if $a_k \equiv b_k \pmod{d}$ holds for all $k \in \mathbb{N}_0$.

(ii) If $\alpha \equiv \beta \pmod{d}$ and $\gamma \equiv \delta \pmod{d}$, then also $\alpha+\gamma \equiv \beta+\delta \pmod{d}$ and $\alpha\gamma \equiv \beta\delta \pmod{d}$, because $\sum_{k=0}^{n} a_k c_{n-k} \equiv \sum_{k=0}^{n} b_k d_{n-k} \pmod{d}$ for $n \in \mathbb{N}_0$.

(iii) If $a_0 = b_0 = 1$, then $\alpha^{-1}, \beta^{-1} \in \mathbb{Z}[[X]]$ follows from the proof of Lemma 4.8. Furthermore, $\alpha \equiv \beta \pmod{d}$ is equivalent to $\alpha^{-1} \equiv \beta^{-1} \pmod{d}$.

(iv) For every prime p it holds that

$$(\alpha+\beta)^p = \sum_{k=0}^{p}\binom{p}{k}\alpha^k\beta^{p-k} \equiv \alpha^p + \beta^p \pmod{p},$$

since $\binom{p}{k}$ is always divisible by p for $0 < k < p$.

Theorem 5.21 (Ramanujan). *For $n \in \mathbb{N}_0$, $5 \mid p(5n+4)$ and $7 \mid p(7n+5)$.*

Proof. Let $\alpha := \prod(1-X^k)$. According to Remark 5.20, $\alpha^5 = \prod(1-X^k)^5 \equiv \prod(1-X^{5k}) \equiv \alpha(X^5) \pmod{5}$ and $\alpha^{-5} \equiv \alpha(X^5)^{-1} \pmod{5}$. For $k \in \mathbb{Z}$, it holds that

$$\frac{1}{2}(k^2+k) \equiv \begin{cases} 0 & \text{if } k \equiv 0,-1 \pmod{5}, \\ 1 & \text{if } k \equiv 1,-2 \pmod{5}, \\ 3 & \text{if } k \equiv 2 \pmod{5}. \end{cases}$$

We can therefore write Jacobi's identity in the form

$$\begin{aligned}\alpha^3 &= \sum_{k\,\equiv\,0,-1 \pmod 5}(-1)^k(2k+1)X^{\frac{k^2+k}{2}} + \sum_{k\,\equiv\,1,-2 \pmod 5}(-1)^k(2k+1)X^{\frac{k^2+k}{2}} + \sum_{k\,\equiv\,2 \pmod 5}(-1)^k(2k+1)X^{\frac{k^2+k}{2}} \\ &\equiv \alpha_0 + \alpha_1 \pmod{5}\end{aligned}$$

where α_i is formed from the terms $a_k X^k$ with $k \equiv i \pmod 5$. From Theorem 5.5 it now follows that

$$\sum_{n=0}^{\infty} p(n) X^n = \alpha^{-1} = \frac{(\alpha^3)^3}{(\alpha^5)^2} \equiv \frac{(\alpha_0 + \alpha_1)^3}{\alpha (X^5)^2} \pmod 5. \tag{5.2}$$

In $(\alpha_0+\alpha_1)^3$, only the X^k with $k \equiv 0, 1, 2, 3 \pmod 5$ occur, while in $\alpha(X^5)^{-2}$ only X^{5k} occur. Therefore, the powers of the form X^{5k+4} do not occur at all on the right side of (5.2). Thus, $p(5k+4) \equiv 0 \pmod 5$ must hold.

For the second statement, we observe $\frac{1}{2}(k^2+k) \equiv 0, 1, 3, 6 \pmod 7$, where the last case only occurs for $k \equiv 3 \pmod 7$ and then $2k+1 \equiv 0 \pmod 7$. As above, we can therefore write $\alpha^3 \equiv \alpha_0 + \alpha_1 + \alpha_3 \pmod 7$. Then

$$\sum_{n=0}^{\infty} p(n) X^n = \alpha^{-1} = \frac{(\alpha^3)^2}{\alpha^7} \equiv \frac{(\alpha_0 + \alpha_1 + \alpha_3)^2}{\alpha(X^7)} \pmod 7.$$

Again, X^{7k+5} does not occur on the right side. □

Remark 5.22. Ramanujan also proved $11 \mid p(11n+6)$ for all $n \in \mathbb{N}_0$, but this is more involved to show. The relation $5 \mid p(5n+4)$ can be specified as

$$\sum_{n=0}^{\infty} p(5n+4) X^n = 5 \prod_{k=1}^{\infty} \frac{(1-X^{5k})^5}{(1-X^k)^6},$$

Ramanujan's "most beautiful" formula.[3] Ono has proven that for every prime $p \geq 5$ there is a corresponding relation. These are, however, significantly more complex, such as

$$13 \mid p(11^3 \cdot 13n + 237).$$

Theorem 5.23 (Jacobi's triple product)**.** *For every $\alpha \in K[[X]] \setminus X^2 K[[X]]$, it holds that*

$$\prod_{k=1}^{\infty} (1-X^{2k})(1+\alpha X^{2k-1})(1+\alpha^{-1} X^{2k-1}) = \sum_{k=-\infty}^{\infty} \alpha^k X^{k^2}.$$

Proof (Wright). Since $\alpha \notin X^2 K[[X]]$, $\alpha^{-1} X$ is indeed well-defined according to Lemma 4.8. Therefore, $\alpha^{-1} X^{2k-1}$ is also well-defined for all $k \in \mathbb{N}$. Likewise, $\alpha^k X^{k^2} = (\alpha^{-1} X)^{-k} X^{k^2+k}$ is well-defined for $k < 0$. As usual, the infinite products and sums are also well-defined. With $\beta := \alpha X$ and $\gamma := \alpha^{-1} X$, we must show

$$\prod_{k=1}^{\infty} (1+\beta^k \gamma^{k-1})(1+\beta^{k-1}\gamma^k) = \sum_{k=-\infty}^{\infty} \beta^{\frac{k^2+k}{2}} \gamma^{\frac{k^2-k}{2}} \prod_{n=1}^{\infty} \frac{1}{1-(\beta\gamma)^n}. \tag{5.3}$$

According to Theorem 5.5, we have

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

On the other hand,

$$\prod_{k=1}^{\infty} (1+\beta^k \gamma^{k-1})(1+\beta^{k-1}\gamma^k) = \sum_{n,m=0}^{\infty} t(n,m) \beta^n \gamma^m,$$

[3] see [M. Hirschhorn, *The power of q*, Chapter 5]

where $t(n,m)$ is the number of partitions of the pair $(n,m) \in \mathbb{N}_0^2$ into pairwise distinct parts of the form $(a, a-1)$ and $(b-1, b)$ with $a, b \in \mathbb{N}$ (Example: $(3,4) = (1,0)+(0,1)+(2,3) = (2,1)+(0,1)+(1,2)$, so $t(3,4) = 3$). The term $\beta^n \gamma^m$ appears on the right side of (5.3) only for the $(n-m)$-th summand. The corresponding coefficient is then $p\big(n-(n-m)(n-m+1)/2\big)$, where we assume $p(k) = 0$ for $k < 0$. It thus suffices to show

$$t(n,m) = p\big(n-(n-m)(n-m+1)/2\big) \tag{5.4}$$

for all $n, m \in \mathbb{N}_0$. Because of $t(n,m) = t(m,n)$ and

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

we can assume $n \geq m$. Let $k := n-m$. Every partition of (n,m) then corresponds to a representation of the form

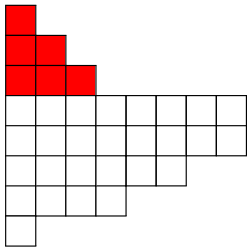
$$n = \sum_{i=1}^{k+s} a_i + \sum_{i=1}^{s} (b_i - 1) \qquad (s \geq 0,\ 1 \leq a_1 < \ldots < a_{k+s},\ 1 \leq b_1 < \ldots < b_s). \tag{5.5}$$

Let $N := n - k(k+1)/2$. For $N < 0$,

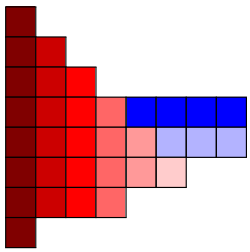
$$\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} i = \frac{k(k+1)}{2} > n$$

and (5.5) has no solution. Because of $p(N) = 0$, (5.4) is proven in this case. In the case $N = 0$, (5.5) has only the solution $s = 0$ and $a_i = i$ for $i = 1, \ldots, k$. Because of $p(0) = 1$, we can now assume $N > 0$.

We construct a bijection between $P(N)$ and the representations (5.5). Let $\lambda \in P(N)$. Above the Young diagram of λ, we place a right-angled triangle with side length k. Example $(n,m) = (33,30)$, $k = 3$, $N = 27$ and $\lambda = (8^2, 6, 4, 1)$:



In total, one obtains $N + k(k+1)/2 = n$ boxes. One now extends the diagonal of the triangle and divides the boxes below and above it into columns and rows respectively:



If $s \geq 0$ is the number of resulting rows, then $k+s$ is the number of resulting columns. Let a_i be the number of boxes in the i-th column and let $b_i - 1 \geq 0$ be the number of boxes in the i-th row. Then $1 \leq a_1 < \ldots < a_{k+s}$, $1 \leq b_1 < \ldots < b_s$ and $\sum_{i=1}^{k+s} a_i + \sum_{i=1}^{s}(b_i - 1) = n$ hold. We have thus found a representation (5.5) (in the example $s = 3$, $(a_1, \ldots, a_6) = (8,6,5,4,2,1)$ and $(b_1, b_2, b_3) = (5,4,1)$). Conversely, one can start with such a representation, draw the corresponding diagram and remove the upper triangle. In this way, one always obtains a Young diagram of a partition of N. These two processes are obviously inverse to each other, so that one obtains a bijection between $P(N)$ and the representations (5.5). Thus (5.4) is proven. □

Example 5.24.

(i) For $\alpha \in \{\pm 1, X\}$, Theorem 5.23 becomes

$$\prod_{k=1}^{\infty}(1-X^{2k})(1+X^{2k-1})^2 = \sum_{k=-\infty}^{\infty} X^{k^2}, \tag{5.6}$$

$$\prod_{k=1}^{\infty}\frac{(1-X^k)^2}{1-X^{2k}} = \prod_{k=1}^{\infty}(1-X^{2k})(1-X^{2k-1})^2 = \sum_{k=-\infty}^{\infty}(-1)^k X^{k^2}, \tag{5.7}$$

$$\prod_{k=1}^{\infty}(1-X^{2k})(1+X^{2k})^2 = \frac{1}{2}\sum_{k=-\infty}^{\infty} X^{k^2+k} = \sum_{k=0}^{\infty} X^{k^2+k}, \tag{5.8}$$

where the bijection $k \mapsto -k-1$ on $\mathbb{Z}$ was used in the third formula. In there, X only occurs with even exponents. By comparing coefficients, one may halve the exponents and obtains

$$\prod_{k=1}^{\infty}(1-X^{2k})(1+X^k) = \prod_{k=1}^{\infty}(1-X^k)(1+X^k)^2 = \sum_{k=0}^{\infty} X^{\frac{k^2+k}{2}}$$

similar to Theorem 5.19.

(ii) According to Definition 4.18, we may replace X by X^3 in Theorem 5.23. If one then sets $\alpha = -X$, it follows that

$$\prod_{k=1}^{\infty}(1-X^{6k})(1-X^{6k-2})(1-X^{6k-4}) = \sum_{k=-\infty}^{\infty}(-1)^k X^{3k^2+k}.$$

Again, one may halve the exponents and obtains

$$\prod_{k=1}^{\infty}(1-X^k) = \prod_{k=1}^{\infty}(1-X^{3k})(1-X^{3k-1})(1-X^{3k-2}) = \sum_{k=-\infty}^{\infty}(-1)^k X^{\frac{3k^2+k}{2}},$$

i. e. Theorem 5.14.

(iii) Replacing X by X^5 and choosing $\alpha \in \{-X, -X^3\}$, one obtains in a similar way

$$\prod_{k=1}^{\infty}(1-X^{5k})(1-X^{5k-2})(1-X^{5k-3}) = \sum_{k=-\infty}^{\infty}(-1)^k X^{\frac{5k^2+k}{2}}, \tag{5.9}$$

$$\prod_{k=1}^{\infty}(1-X^{5k})(1-X^{5k-1})(1-X^{5k-4}) = \sum_{k=-\infty}^{\infty}(-1)^k X^{\frac{5k^2+3k}{2}}. \tag{5.10}$$

Theorem 5.25 (Lagrange-Jacobi)**.** *Every natural number is the sum of four squares. More precisely,*

$$q(n) := \big|\{(a,b,c,d) \in \mathbb{Z}^4 : a^2+b^2+c^2+d^2 = n\}\big| = 8\sum_{4\nmid d \mid n} d$$

holds for $n \in \mathbb{N}$.

Proof. Obviously, it suffices to prove the second statement (by Jacobi). Since the summands $(-1)^k(2k+1)X^{\frac{k^2+k}{2}}$ in Theorem 5.19 are invariant under the transformation $k \mapsto -k-1$, it holds that

$$\prod_{k=1}^{\infty}(1-X^k)^3 = \frac{1}{2}\sum_{k=-\infty}^{\infty}(-1)^k(2k+1)X^{\frac{k^2+k}{2}}.$$

Squaring yields

$$\alpha := \prod_{k=1}^{\infty}(1-X^k)^6 = \frac{1}{4}\sum_{k,l=-\infty}^{\infty}(-1)^{k+l}(2k+1)(2l+1)X^{\frac{k^2+k+l^2+l}{2}}.$$

We transform the pairs (k,l) with $k \equiv l \pmod 2$ using $(k,l) \mapsto (s,t) := \frac{1}{2}(k+l, k-l)$, while we transform the pairs with $k \not\equiv l \pmod 2$ using $(s,t) := \frac{1}{2}(k-l-1, k+l+1)$. It holds that $k = s+t$ and $l = s-t$ or $l = t-s-1$, respectively. Therefore,

$$\begin{aligned}
\alpha &= \frac{1}{4}\sum_{s,t=-\infty}^{\infty}(2s+2t+1)(2s-2t+1)X^{\frac{(s+t)^2+s+t+(s-t)^2+s-t}{2}} \\
&\quad - \frac{1}{4}\sum_{s,t=-\infty}^{\infty}(2s+2t+1)(2t-2s-1)X^{\frac{(s+t)^2+s+t+(t-s-1)^2+t-s-1}{2}} \\
&= \frac{1}{4}\sum_{s,t}\big((2s+1)^2-(2t)^2\big)X^{s^2+s+t^2} - \frac{1}{4}\sum_{s,t}\big((2t)^2-(2s+1)^2\big)X^{s^2+s+t^2} \\
&= \frac{1}{2}\sum_{s,t}\big((2s+1)^2-(2t)^2\big)X^{s^2+s+t^2} \\
&= \frac{1}{2}\sum_{t=-\infty}^{\infty}X^{t^2}\sum_{s=-\infty}^{\infty}(2s+1)^2X^{s^2+s} - \frac{1}{2}\sum_{s=-\infty}^{\infty}X^{s^2+s}\sum_{t=-\infty}^{\infty}(2t)^2X^{t^2}.
\end{aligned}$$

For $\beta := \sum X^{t^2}$ and $\gamma := \frac{1}{2}\sum X^{s^2+s}$, it holds that $\gamma + 4X\gamma' = \frac{1}{2}\sum(2s+1)^2X^{s^2+s}$ and it follows that

$$\alpha = \beta(\gamma + 4X\gamma') - 4X\beta'\gamma.$$

We apply the product rule to (5.6) and (5.8):

$$\begin{aligned}
\beta' &= \Big(\prod_{k=1}^{\infty}(1-X^{2k})(1+X^{2k-1})^2\Big)' = \beta\sum_{k=1}^{\infty}\Big(2\frac{(2k-1)X^{2k-2}}{1+X^{2k-1}} - \frac{2kX^{2k-1}}{1-X^{2k}}\Big) \\
\gamma' &= \Big(\prod_{k=1}^{\infty}(1-X^{2k})(1+X^{2k})^2\Big)' = \gamma\sum_{k=1}^{\infty}\Big(2\frac{2kX^{2k-1}}{1+X^{2k}} - \frac{2kX^{2k-1}}{1-X^{2k}}\Big)
\end{aligned}$$

Substitution yields:

$$\alpha = \beta\gamma\Big(1 + 8\sum_{k=1}^{\infty}\Big(\frac{2kX^{2k}}{1+X^{2k}} - \frac{(2k-1)X^{2k-1}}{1+X^{2k-1}}\Big)\Big).$$

Here, $\beta\gamma = \prod(1-X^{2k})^2(1+X^{2k-1})^2(1+X^{2k})^2 = \prod(1-X^{2k})^2(1+X^k)^2 = \prod(1-X^{2k})^4(1-X^k)^{-2}$. After balancing this with α, there remains

$$\Big(\sum_{k=-\infty}^{\infty}(-1)^kX^{k^2}\Big)^4 \overset{(5.7)}{=} \prod_{k=1}^{\infty}\frac{(1-X^k)^8}{(1+X^{2k})^4} = \frac{\alpha}{\beta\gamma} = 1 + 8\sum_{k=1}^{\infty}\Big(\frac{2kX^{2k}}{1+X^{2k}} - \frac{(2k-1)X^{2k-1}}{1+X^{2k-1}}\Big).$$

Finally, we replace X with $-X$:

$$\begin{aligned}\sum q(n)X^n &= \Big(\sum_{k=-\infty}^{\infty} X^{k^2}\Big)^4 = 1 + 8\sum_{k=1}^{\infty}\Big(\frac{2kX^{2k}}{1+X^{2k}} + \frac{(2k-1)X^{2k-1}}{1-X^{2k-1}}\Big)\\ &= 1 + 8\sum_{k=1}^{\infty}\Big(\frac{(2k-1)X^{2k-1}}{1-X^{2k-1}} + \frac{2kX^{2k}}{1-X^{2k}} - \frac{2kX^{2k}}{1-X^{2k}} + \frac{2kX^{2k}}{1+X^{2k}}\Big)\\ &= 1 + 8\sum_{k=1}^{\infty}\Big(\frac{kX^k}{1-X^k} - \frac{4kX^{4k}}{1-X^{4k}}\Big) = 1 + 8\sum_{4\nmid k}\frac{kX^k}{1-X^k}\\ &= 1 + 8\sum_{4\nmid k} k\sum_{l=1}^{\infty} X^{kl} = 1 + 8\sum_{n=1}^{\infty}\sum_{4\nmid d\,|\,n} dX^n. \qquad \square\end{aligned}$$

Example 5.26.

(i) For $n = 30$, we have $\sum_{4\nmid d\,|30} d = 1 + 2 + 3 + 5 + 6 + 10 + 15 + 30 = 72$. Therefore, there are $8 \cdot 72 = 576$ possibilities to write 30 as a sum of four squares. These arise by permutation and choice of signs from

$$30 = 1^2 + 2^2 + 3^2 + 4^2 = 0^2 + 1^2 + 2^2 + 5^2.$$

(ii) Obviously, 7 is not a sum of three squares. Because $a^2 + b^2 + c^2 \not\equiv 7 \mod 8$, more generally every number $n \equiv 7 \pmod 8$ is not a sum of three squares.

Remark 5.27.

(i) If $n, m \in \mathbb{N}$ are sums of four squares, then so is nm, because Euler's identity holds:

$$\begin{aligned}&(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) = (a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4)^2\\ &+(a_1b_2 - a_2b_1 + a_3b_4 + a_4b_3)^2 + (a_1b_3 - a_3b_1 + a_4b_2 - a_2b_4)^2 + (a_1b_4 - a_4b_1 + a_2b_3 - a_3b_2)^2\end{aligned}$$

This reduces the first statement (by Lagrange) in Theorem 5.25 to the case $n \in \mathbb{P}$.

(ii) The *Waring problem* for $k \in \mathbb{N}$ asks for the smallest number $w(k) \in \mathbb{N}$ such that every natural number is the sum of $w(k)$ k-th powers. Hilbert proved $w(k) < \infty$. It holds that $w(1) = 1$, $w(2) = 4$ (Theorem 5.25), $w(3) = 9$, $w(4) = 19$ and it is conjectured in general that

$$w(k) = \Big\lfloor \Big(\frac{3}{2}\Big)^k \Big\rfloor + 2^k - 2.$$

Interestingly, only the numbers $23 = 2 \cdot 2^3 + 7 \cdot 1^3$ and $239 = 2 \cdot 4^3 + 4 \cdot 3^3 + 3 \cdot 1^3$ are not the sum of eight cubes. Furthermore, there are only 15 numbers that are not the sum of seven cubes. It is conjectured more generally that every sufficiently large number is the sum of four cubes.

(iii) Since every odd number has the form $\pm 1 + 4k$, Theorem 5.7(i) can be formulated as follows: The number of partitions into distinct parts is equal to the number of partitions into parts of the form $\pm 1 + 4k$. We now replace $\pm 1 + 4k$ by $\pm 1 + 5k$.

Theorem 5.28 (Rogers-Ramanujan identities)**.** *It holds that*

$$\prod_{k=1}^{\infty} \frac{1}{(1-X^{5k-1})(1-X^{5k-4})} = \sum_{k=0}^{\infty} \frac{X^{k^2}}{X^{k!}}, \tag{5.11}$$

$$\prod_{k=1}^{\infty} \frac{1}{(1-X^{5k-2})(1-X^{5k-3})} = \sum_{k=0}^{\infty} \frac{X^{k^2+k}}{X^{k!}}. \tag{5.12}$$

Proof (Chapman). For $n \in \mathbb{N}_0$ let

$$\alpha_n := \sum_{k=0}^{\infty} X^{k^2} \left\langle {n \atop k} \right\rangle \qquad \beta_n := \sum_{k=0}^{\infty} X^{k^2+k} \left\langle {n \atop k} \right\rangle$$

$$\widetilde{\alpha}_n := \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} \left\langle {2n \atop n+2k} \right\rangle \qquad \widetilde{\beta}_n := \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2-3k}{2}} \left\langle {2n+1 \atop n+2k} \right\rangle$$

(all sums are finite). It holds that $\alpha_0 = \beta_0 = \widetilde{\alpha}_0 = \widetilde{\beta}_0 = 1$. For $n \geq 1$ we have

$$\begin{aligned}
\alpha_n &\overset{(4.41)}{=} \sum_{k=0}^{\infty} X^{k^2} \left(\left\langle {n-1 \atop k} \right\rangle + X^{n-k} \left\langle {n-1 \atop k-1} \right\rangle \right) = \alpha_{n-1} + X^n \sum_{k=1}^{\infty} X^{k(k-1)} \left\langle {n-1 \atop k-1} \right\rangle \\
&= \alpha_{n-1} + X^n \sum_{k=0}^{\infty} X^{k(k+1)} \left\langle {n-1 \atop k} \right\rangle = \alpha_{n-1} + X^n \beta_{n-1}, \\
\beta_n - X^n \alpha_n &= \sum_{k=0}^{\infty} X^{k^2+k} \left\langle {n \atop k} \right\rangle (1 - X^{n-k}) = \sum_{k=0}^{\infty} X^{k^2+k} \frac{X^n!}{X^k! X^{n-k}!} (1 - X^{n-k}) \\
&= (1 - X^n) \sum_{k=0}^{\infty} X^{k^2+k} \left\langle {n-1 \atop k} \right\rangle = (1 - X^n) \beta_{n-1}.
\end{aligned}$$

By these recursion equations, α_n and β_n are uniquely determined. We now show that an "index shift" does not change $\widetilde{\alpha}_n$:

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} \left\langle {2n+1 \atop n+2k+1} \right\rangle &\overset{(4.41)}{=} \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} \left(\left\langle {2n \atop n+2k} \right\rangle + X^{n+2k+1} \left\langle {2n \atop n+2k+1} \right\rangle \right) \\
&= \widetilde{\alpha}_n + X^{n+1} \left(\sum_{k=0}^{\infty} (-1)^k X^{\frac{5k^2+5k}{2}} \left\langle {2n \atop n+2k+1} \right\rangle + \sum_{k=-\infty}^{-1} (-1)^k X^{\frac{5k(k+1)}{2}} \left\langle {2n \atop n+2k+1} \right\rangle \right) \\
&= \widetilde{\alpha}_n + X^{n+1} \left(\sum_{k=0}^{\infty} (-1)^k X^{\frac{5k^2+5k}{2}} \left\langle {2n \atop n+2k+1} \right\rangle + \sum_{k=0}^{\infty} (-1)^{-k-1} X^{\frac{5(-k-1)(-k)}{2}} \left\langle {2n \atop n-2k-1} \right\rangle \right) \\
&= \widetilde{\alpha}_n + X^{n+1} \left(\sum_{k=0}^{\infty} (-1)^k X^{\frac{5k^2+5k}{2}} \left\langle {2n \atop n+2k+1} \right\rangle - \sum_{k=0}^{\infty} (-1)^k X^{\frac{5k^2+5k}{2}} \left\langle {2n \atop n+2k+1} \right\rangle \right) = \widetilde{\alpha}_n.
\end{aligned}$$

Therefore it holds that

$$\begin{aligned}
\widetilde{\alpha}_n - \widetilde{\alpha}_{n-1} &= \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} \left\langle {2n \atop n+2k} \right\rangle - \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} \left\langle {2n-1 \atop n+2k} \right\rangle \\
&\overset{(4.41)}{=} \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} X^{n-2k} \left\langle {2n-1 \atop n+2k-1} \right\rangle = X^n \widetilde{\beta}_{n-1}, \\
\widetilde{\beta_n} - X^n \widetilde{\alpha}_n &= \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2-3k}{2}} \left(\left\langle {2n+1 \atop n+2k} \right\rangle - X^{n+2k} \left\langle {2n \atop n+2k} \right\rangle \right) = \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2-3k}{2}} \left\langle {2n \atop n+2k-1} \right\rangle \\
&= \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2-3k}{2}} \left(\left\langle {2n-1 \atop n+2k-1} \right\rangle + X^{n-2k+1} \left\langle {2n-1 \atop n+2k-2} \right\rangle \right) \\
&= \widetilde{\beta}_{n-1} + X^n \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2-7k+2}{2}} \left\langle {2n-1 \atop n+2k-2} \right\rangle
\end{aligned}$$

$$= \widetilde{\beta}_{n-1} + X^n \sum_{k=-\infty}^{\infty} (-1)^{1-k} X^{\frac{5(1-k)^2-7(1-k)+2}{2}} \left\langle {2n-1 \atop n-2k} \right\rangle$$
$$= \widetilde{\beta}_{n-1} - X^n \sum_{k=-\infty}^{\infty} (-1)^{k} X^{\frac{k^2-3k}{2}} \left\langle {2n-1 \atop n+2k-1} \right\rangle = (1-X^n)\widetilde{\beta}_{n-1}.$$

Thus $\widetilde{\alpha}_n$ and $\widetilde{\beta}_n$ satisfy the same recursion equations. Inductively one obtains $\alpha_n = \widetilde{\alpha}_n$ and $\beta_n = \widetilde{\beta}_n$ for all $n \in \mathbb{N}_0$. Now it holds that

$$\Big|\sum_{k=0}^{n} X^{k^2}\Big(\frac{1}{X^k!} - \left\langle {n \atop k} \right\rangle\Big)\Big| \overset{5.11}{\leq} \max_{k=0,\ldots,n} 2^{-k^2-n+k-1} \leq 2^{-n} \to 0 \qquad (n \to \infty).$$

This shows

$$\sum_{k=0}^{\infty} \frac{X^{k^2}}{X^k!} = \sum_{k=0}^{\infty} X^{k^2} \lim_{n\to\infty} \left\langle {n \atop k} \right\rangle = \lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \widetilde{\alpha}_n = \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2+k}{2}} \lim_{n\to\infty} \left\langle {2n \atop n+2k} \right\rangle$$
$$\overset{(5.9)+5.11}{=} \frac{\prod_{k=1}^{\infty}(1-X^{5k})(1-X^{5k-2})(1-X^{5k-3})}{\prod_{k=1}^{\infty}(1-X^k)} = \prod_{k=1}^{\infty} \frac{1}{(1-X^{5k-1})(1-X^{5k-4})},$$
$$\sum_{k=0}^{\infty} \frac{X^{k^2+k}}{X^k!} = \sum_{k=0}^{\infty} X^{k^2+k} \lim_{n\to\infty} \left\langle {n \atop k} \right\rangle = \lim_{n\to\infty} \beta_n = \lim_{n\to\infty} \widetilde{\beta}_n = \sum_{k=-\infty}^{\infty} (-1)^k X^{\frac{5k^2-3k}{2}} \lim_{n\to\infty} \left\langle {2n+1 \atop n+2k} \right\rangle$$
$$\overset{(5.10)+5.11}{=} \frac{\prod_{k=1}^{\infty}(1-X^{5k})(1-X^{5k-1})(1-X^{5k-4})}{\prod_{k=1}^{\infty}(1-X^k)} = \prod_{k=1}^{\infty} \frac{1}{(1-X^{5k-2})(1-X^{5k-3})}. \qquad \square$$

Remark 5.29. The coefficient of X^n on the left side of (5.11) is the number of partitions of n into parts of the form $\pm 1 + 5k$. The right side of (5.11) is

$$\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} p_k(n) X^{n+k^2} = \sum_{n=0}^{\infty}\sum_{k=0}^{n} p_k(n-k^2) X^n.$$

If $(\lambda_1, \ldots, \lambda_k) \in P(n-k^2)$ with at most k parts, then $(\lambda_1 + 2k - 1, \lambda_2 + 2k - 3, \ldots, \lambda_k + 1)$ is a partition of $n - k^2 + 1 + 3 + \ldots + 2k - 1 = n$ with exactly k parts, which all differ by at least 2. Conclusion: The number of partitions of n into parts that differ by at least 2 is equal to the number of partitions into parts of the form $\pm 1 + 5k$. Using $k^2 + k = 2 + 4 + \ldots + 2k$, one obtains the following interpretation of (5.12): The number of partitions of n into parts that differ by at least 2 and are greater than 1 is equal to the number of partitions with parts of the form $\pm 2 + 5k$.

6. Polynomials

Definition 6.1. A (formal) power series $\alpha = \sum a_n X^n \in K[[X]]$ with only finitely many non-zero summands is called a *polynomial* of *degree* $\deg(\alpha) := \sup\{n \in \mathbb{N}_0 : a_n \neq 0\}$ (where $\deg(0) = \sup \varnothing = -\infty$). The set of polynomials is denoted by $K[X]$. One calls α *monic*, if $\alpha \neq 0$ and $a_{\deg(\alpha)} = 1$. In general, $a_{\deg(\alpha)}$ is the *leading coefficient* of α.

Remark 6.2.

(i) In contrast to power series, polynomials are often written in reverse order starting with the highest X-power. For example $X^2 + 1 \in \mathbb{R}[X]$.

(ii) Every polynomial $\alpha \in K[X] \setminus \{0\}$ can be *normalized* by multiplying with $a_{\deg(\alpha)}^{-1}$.

(iii) Every power series can be interpreted as a Cauchy sequence of polynomials, i. e. $K[[X]]$ is the *completion* of $K[X]$ with respect to the metric in Lemma 4.13.

Lemma 6.3. *For $\alpha, \beta \in K[X]$ we have $\deg(\alpha+\beta) \leq \max\{\deg(\alpha), \deg(\beta)\}$ and $\deg(\alpha\beta) = \deg(\alpha) + \deg(\beta)$. In particular, $\alpha+\beta$ and $\alpha\beta$ are polynomials.*

Proof. Wlog. let $\alpha = \sum_{n=0}^{\infty} a_n X^n \neq 0$ and $\beta = \sum_{n=0}^{\infty} b_n X^n \neq 0$ with $d := \deg(\alpha)$ and $e := \deg(\beta)$. Because of $a_k + b_k = 0$ for $k > \max\{d, e\}$, we have $\deg(\alpha+\beta) \leq \max\{d, e\}$. Similarly, $\sum_{k=0}^{d+e} a_k b_{d+e-k} = a_d b_e \neq 0$ and $\sum_{k=0}^{n} a_k b_{n-k} = 0$ for $n > d + e$. This shows $\deg(\alpha\beta) = \deg(\alpha) + \deg(\beta)$. □

Remark 6.4.

(i) One can therefore calculate in $K[X]$ just as in $\mathbb{Z}$ (note: $0, 1 \in K[X]$). If $\alpha \in K[X]$ is invertible in $K[[X]]$, it does not necessarily follow that $\alpha^{-1} \in K[X]$! For example, $(1-X)^{-1} \notin K[X]$.

(ii) One can embed K into $K[X]$ via the *constant* polynomials KX^0, i. e. $K \subseteq K[X] \subseteq K[[X]]$. $\alpha \in K$ holds if and only if $\deg(\alpha) \leq 0$.

(iii) For polynomials $\alpha = \sum a_n X^n \in K[X]$ and $\beta \in K[X]$, the composition $\alpha(\beta) = \sum a_n \beta^n$ is always well-defined (even if the constant term of β does not vanish, cf. Definition 4.18).

Example 6.5.

(i) For $\alpha = a_n X^n + \ldots + a_1 X + a_0 \in K[X]$ and $b \in K \subseteq K[[X]]$, we have $\alpha(b) = a_n b^n + \ldots + a_1 b + a_0 \in K$. As usual, b is called a *root* of α if $\alpha(b) = 0$ holds.

(ii) According to Theorem 5.10, $\left\langle {n \atop k} \right\rangle$ is a monic polynomial of degree $k(n-k)$. For $X = 1$, $\left\langle {n \atop k} \right\rangle$ and $\binom{n}{k}$ have the same recursion formula according to (4.41). Because of $\left\langle {0 \atop 0} \right\rangle = 1 = \binom{0}{0}$, $\left\langle {n \atop k} \right\rangle$ and $\binom{n}{k}$ even coincide for $X = 1$. We calculate further values.

Remark 6.6. In the following, let K be a finite field. In algebra, it is shown that $q := |K|$ is always a prime power. Conversely, for every prime power $q > 1$, there exists essentially exactly one field with q elements. This is denoted by $\mathbb{F}_q$.

Example 6.7.

(i) For every prime number p, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, i. e. one identifies the elements of $\mathbb{F}_p$ with the numbers $0, \ldots, p-1$ and calculates modulo p (see Definition 6.20). For example, $1 + 1 = 0$ in $\mathbb{F}_2$ and $3 \cdot 5 = 1$ in $\mathbb{F}_7$.

(ii) The operation tables for $\mathbb{F}_4 = \{0, 1, a, b\}$ are given as follows:

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

$\cdot$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

Theorem 6.8. *There are exactly $(q^n - 1)(q^n - q) \ldots (q^n - q^{n-1})$ invertible $n \times n$ matrices over every field with q elements.*

Proof. Let K be a field with q elements and $A \in K^{n\times n}$. As is well known, A is invertible if and only if the rows of A are linearly independent. The first row a_1 of A can be chosen arbitrarily from $K^n \setminus \{0\}$. There are $q^n - 1$ possibilities for this. The second row a_2 must not lie in the span of a_1, i. e. $a_2 \in K^n \setminus Ka_1$. There are $q^n - q$ possibilities for this. For the third row, $a_3 \in K^n \setminus (Ka_1 + Ka_2)$ holds, and so on. □

Theorem 6.9. *The value of $\left\langle {n \atop k} \right\rangle$ at $X = q$ is the number of k-dimensional subspaces of an n-dimensional vector space over a field with q elements.*

Proof. The proof is similar to Theorem 6.8. There are exactly $(q^n - 1)(q^n - q)\dots(q^n - q^{k-1})$ linearly independent k-tuples in K^n. We now count how many of these tuples span the same subspace $U \subseteq K^n$. This is clearly the number of linearly independent k-tuples in U, i. e. $(q^k - 1)(q^k - q)\dots(q^k - q^{k-1})$. The number of k-dimensional subspaces is therefore

$$\frac{(q^n - 1)(q^n - q)\dots(q^n - q^{k-1})}{(q^k - 1)(q^k - q)\dots(q^k - q^{k-1})} = \frac{(1 - q^n)\dots(1 - q^{n-k+1})}{(1 - q^k)\dots(1 - q)}.$$ □

Definition 6.10. A monic polynomial $\alpha \in K[X] \setminus K$ is called *irreducible*, if it cannot be written in the form $\alpha = \beta\gamma$ with $\beta, \gamma \in K[X] \setminus K$. Otherwise α is called *reducible*.

Example 6.11.

(i) Monic polynomials of degree 1 are always irreducible, because $1 = \deg(\alpha) = \deg(\beta\gamma) = \deg(\beta) + \deg(\gamma)$ implies $\deg(\beta) = 0$ or $\deg(\gamma) = 0$.

(ii) $X^2 - 2$ is irreducible in $\mathbb{Q}[X]$, because the approach $X^2 - 2 = (X + a)(X + b)$ leads to $a + b = 0$ and $ab = -2$, i. e. $a^2 = 2$ and $a = \pm\sqrt{2} \notin \mathbb{Q}$. Due to $X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2})$, $X^2 - 2$ is however reducible in $\mathbb{R}[X]$.

(iii) $X^2 + 1$ is irreducible in $\mathbb{R}[X]$, but not in $\mathbb{C}[X]$, because $X^2 + 1 = (X - \mathrm{i})(X + \mathrm{i}) \in \mathbb{C}[X]$.

Theorem 6.12 (Division with remainder)**.** *For $\alpha \in K[X]$ and $\beta \in K[X] \setminus \{0\}$ there exist $\gamma, \delta \in K[X]$ with $\alpha = \beta\gamma + \delta$ and $\deg\delta < \deg\beta$.*

Proof. Choose $\delta = \sum a_i X^i \in \{\alpha - \mu\beta : \mu \in K[X]\}$ with minimal possible degree d. Let $\beta = \sum b_i X^i$. If $d \geq \deg\beta =: e$ holds, then $\deg\big(\delta - a_d b_e^{-1} X^{d-e}\beta\big) < d$ in contradiction to the choice of δ. Thus $d < e$ and the claim follows. □

Lemma 6.13. *If $\alpha, \beta \in K[X]$ are distinct irreducible polynomials, then there exist $\tilde{\alpha}, \tilde{\beta} \in K[X]$ with $\alpha\tilde{\alpha} + \beta\tilde{\beta} = 1$.*

Proof. Let $\tilde{\alpha}, \tilde{\beta} \in K[X]$ such that $\rho := \alpha\tilde{\alpha} + \beta\tilde{\beta} \neq 0$ has minimal degree and is monic. Division with remainder yields $\gamma, \delta \in K[X]$ with $\alpha = \gamma\rho + \delta$ and $\deg\delta < \deg\rho$. Thus

$$\delta = \alpha - \gamma\rho = \alpha(1 - \gamma\tilde{\alpha}) - \beta(\gamma\tilde{\beta})$$

and the choice of ρ shows $\delta = 0$, i. e. $\alpha = \gamma\rho$. Analogously, we obtain $\beta = \tau\rho$ for some $\tau \in K[X]$. Since α, β and ρ are monic, so are γ and τ. In the case $\rho \neq 1$, we would have $\alpha = \rho = \beta$, because α and β are irreducible. This contradicts $\alpha \neq \beta$. Thus $\alpha\tilde{\alpha} + \beta\tilde{\beta} = \rho = 1$. □

Theorem 6.14 (Prime factorization in $K[X]$)**.** *For every polynomial $\alpha \in K[X]\setminus\{0\}$, there exist uniquely determined irreducible polynomials $\sigma_1, \ldots, \sigma_n \in K[X]$ up to their order and a uniquely determined constant $c \in K \setminus \{0\}$ such that $\alpha = c\sigma_1 \ldots \sigma_n$.*

Proof. Existence: Because $\alpha \neq 0$, there exists $c \in K \setminus \{0\}$ such that $c^{-1}\alpha$ is monic. We can therefore assume that α is monic. Induction on $d := \deg \alpha$: In the case $d = 0$, we have $\alpha = 1$ and we choose $n = 0$ (empty product). If α is irreducible (for example $d = 1$), then we are also finished. Otherwise $\alpha = \beta\gamma$ with $\beta, \gamma \in K[X] \setminus K$. Because $d = \deg(\beta\gamma) = \deg(\beta) + \deg(\gamma)$, we have $\deg \beta, \deg \gamma < d$. By induction, β and γ are products of irreducible polynomials and constants, and therefore so is α.

Uniqueness: Obviously c is uniquely determined as the leading coefficient of α. We can therefore again assume that α is monic. Let $\alpha = \sigma_1 \ldots \sigma_n = \tau_1 \ldots \tau_m$ with irreducible $\sigma_1, \ldots, \sigma_n, \tau_1, \ldots, \tau_m \in K[X]$. Induction on m. In the case $m = 1$, we have $n = 1$ and $\sigma_1 = \alpha = \tau_1$. Now let $m \geq 2$. In the case $\sigma_1 = \tau_1$, we have $\sigma_2 \ldots \sigma_n = \tau_2 \ldots \tau_m$ and induction yields the claim. Now let $\sigma_1 \neq \tau_1$. Then there exist $\tilde{\sigma}, \tilde{\tau} \in K[X]$ with $\sigma_1\tilde{\sigma} + \tau_1\tilde{\tau} = 1$ according to Lemma 6.13. It follows

$$\sigma_1(\tilde{\sigma}\tau_2 \ldots \tau_m + \sigma_2 \ldots \sigma_n\tilde{\tau}) = \sigma_1\tilde{\sigma}\tau_2 \ldots \tau_m + \tau_1 \ldots \tau_m\tilde{\tau} = (\sigma_1\tilde{\sigma} + \tau_1\tilde{\tau})(\tau_2 \ldots \tau_m) = \tau_2 \ldots \tau_m.$$

Inductively, one obtains $\sigma_1 = \tau_i$ for some $i \in \{1, \ldots, m\}$. Then $\sigma_2 \ldots \sigma_n = \tau_1 \ldots \tau_{i-1}\tau_{i+1} \ldots \tau_m$ and induction yields the claim. □

Remark 6.15. Over a finite field with q elements, there are obviously exactly q^d monic polynomials of degree d. We want to count how many of them are irreducible.

Definition 6.16. Let $I_d(K)$ be the number of irreducible polynomials of degree $d \geq 1$ over a field K.

Theorem 6.17 (Gauss)**.** *For every field K with $q < \infty$ elements, it holds that*

$$\boxed{I_d(K) = \frac{1}{d}\sum_{e|d} \mu(e) q^{d/e}}$$

with the classical Möbius function μ. In particular, $I_d(K)$ depends only on $|K|$.

Proof. The q^d monic polynomials of degree d can be uniquely written as a product of irreducible polynomials $\sigma_1 \ldots \sigma_n$ according to Theorem 6.14. Here, $d = \deg(\sigma_1) + \ldots + \deg(\sigma_n)$ holds. Thus, q^d is the d-th coefficient of

$$(1 + X + X^2 + \ldots)^{I_1(K)}(1 + X^2 + X^4 + \ldots)^{I_2(K)}(1 + X^3 + X^6 + \ldots)^{I_3(K)} \ldots = \prod_{e=1}^{\infty}\Big(\frac{1}{1 - X^e}\Big)^{I_e(K)}.$$

On the other hand, $\frac{1}{1-qX} = 1 + qX + q^2X^2 + \ldots$ and it follows that

$$\frac{1}{1 - qX} = \prod_{e=1}^{\infty}\Big(\frac{1}{1 - X^e}\Big)^{I_e(K)}.$$

We apply log to both sides (taking Lemma 4.33 and Example 4.34 into account):

$$\sum_{n=1}^{\infty} \frac{q^n X^n}{n} = \sum_{e=1}^{\infty} I_e(K) \sum_{f=1}^{\infty} \frac{X^{ef}}{f} = \sum_{n=1}^{\infty}\Big(\sum_{e|n} \frac{I_e(K)}{n/e}\Big)X^n.$$

A comparison of coefficients yields

$$q^n = \sum_{e|n} I_e(K)e.$$

Möbius inversion (Example 3.7(v)) now implies the claim. □

Example 6.18. According to Theorem 6.17, we can define $I_d(q) := I_d(K)$ with $|K| = q$. Then $I_1(q) = \mu(1)q^1 = q$ holds (every monic polynomial of degree 1 is irreducible). Furthermore,

$$I_2(q) = \frac{1}{2}(q^2 - q), \qquad I_3(q) = \frac{1}{3}(q^3 - q), \qquad I_4(q) = \frac{1}{4}(q^4 - q^2).$$

Because $I_2(2) = 1$, $X^2 + X + 1$ is the only irreducible polynomial of degree 2 in $\mathbb{F}_2[X]$.

Theorem 6.19. *For every finite field K and every $d \in \mathbb{N}$, there exists an irreducible polynomial in $K[X]$ of degree d.*

Proof. According to Theorem 6.17, $dI_d(K) = q^d \pm q^{e_1} \pm \ldots \pm q^{e_k}$ with $d > e_1 > \ldots > e_k > 0$. This shows $q^{e_k} \mid dI_d(K)$, but $q^{e_k+1} \nmid dI_d(K)$. In particular, $I_d(K) > 0$. □

Definition 6.20. For $a, b \in \mathbb{Z}$ and $d \in \mathbb{N}$, we write $a \equiv b \pmod{d}$ if $d \mid a - b$. One then says a is *congruent* to b *modulo* d.

Example 6.21. It holds that $7 \equiv -11 \pmod{2}$ and $100 \equiv 0 \pmod{10}$.

Remark 6.22 (Division with remainder)**.** For $a \in \mathbb{Z}$ and $d \in \mathbb{N}$, $a \pm d \equiv a \pmod{d}$ holds. Therefore, there exists a $b \in \{0, \ldots, d-1\}$ with $a \equiv b \pmod{d}$.

Lemma 6.23. *The congruence modulo $d \in \mathbb{N}$ is an equivalence relation on $\mathbb{Z}$ with*

$$\left.\begin{array}{ll} a \equiv a' & \pmod{d} \\ b \equiv b' & \pmod{d} \end{array}\right\} \Longrightarrow a \dot{+} b \equiv a' \dot{+} b' \pmod{d}.$$

Proof. Reflexive: $d \mid 0 = a - a$.
Symmetric: $d \mid a - b \Longrightarrow d \mid -(a - b) = b - a$.
Transitive: $(d \mid a - b) \wedge (d \mid b - c) \Longrightarrow d \mid (a - b) + (b - c) = a - c$.
For the last statement, let $d \mid a - a'$ and $d \mid b - b'$. Then $d \mid (a - a') + (b - b') = (a + b) - (a' + b')$ and $d \mid (a - a')b + (b - b')a' = ab - a'b'$. □

Example 6.24. Lemma 10.8 simplifies many calculations. We want to check whether $5^{100} + 2^7$ is divisible by 3:

$$5^{100} + 2^7 \equiv 2^{100} + (-1)^7 \equiv 4^{50} - 1 \equiv 1^{50} - 1 \equiv 0 \pmod{3}.$$

Lemma 6.25 (Fermat's "little" theorem)**.** *For $a \in \mathbb{Z}$ and $p \in \mathbb{P}$, it holds that* $\boxed{a^p \equiv a \pmod{p}.}$

Proof. For $p = 2$, it holds that $p \mid a(a-1) = a^p - a$, because a or $a-1$ is even. So let $p > 2$. Due to

$$a^p \equiv a \pmod{p} \iff (-a)^p \equiv -a^p \equiv -a \pmod{p}$$

we can assume $a \geq 0$. The claim certainly holds for $a \in \{0, 1\}$. So let $a \geq 2$. We now argue by induction on the number of prime factors of a. First, let $a \in \mathbb{P}$ and $K := \mathbb{F}_a$. From Theorem 6.17 it follows that

$$\frac{1}{p}(a^p - a) = \frac{1}{p}(\mu(1)a^p + \mu(p)a^1) = I_p(K) \in \mathbb{N}.$$

Now let $a = bc$ with $1 < b < a$. By induction, it holds that $b^p \equiv b \pmod{p}$ and $c^p \equiv c \pmod{p}$. With Lemma 6.23 it follows that $a^p = (bc)^p = b^p c^p \equiv bc \equiv a \pmod{p}$. □

Remark 6.26. If a is not divisible by p, then $a^{p-1} \equiv 1 \pmod{p}$ follows from $p \mid (a^p - a) = a(a^{p-1} - 1)$.

Lemma 6.27. *Every $\alpha \in K[X] \setminus \{0\}$ has at most $\deg(\alpha)$ roots.*

Proof. Induction on $d := \deg \alpha$. In the case $d = 0$, α is constant and therefore has no root. Now let $d > 0$ and $a \in K$ be a root of α. Division with remainder yields $\alpha = (X - a)\beta + r$ with $\beta, r \in K[X]$ and $\deg(r) < \deg(X - a) = 1$, i. e. $r \in K$. Then $r = \alpha(a) = 0$, so $\alpha = (X - a)\beta$ with $\deg(\beta) = d - 1$. For every further root $b \neq a$ of α, it holds that $0 = \alpha(b) = (b - a)\beta(b) = \beta(b)$. Thus b is also a root of β. By induction, β has at most $d - 1$ roots. In total, α thus has at most d roots. □

Example 6.28. For $k \in \mathbb{N}_0$ we define

$$\binom{X}{k} := \prod_{i=1}^{k} \frac{X - i + 1}{i} \in \mathbb{C}[X].$$

Let $\alpha := \binom{2X}{k} \in \mathbb{C}[X]$ and $\beta := \sum_{i=0}^{k} \binom{X}{i}\binom{X}{k-i} \in \mathbb{C}[X]$. According to the Vandermonde identity 1.16 it then holds that $(\alpha - \beta)(n) = 0$ for all $n \in \mathbb{N}$. Lemma 6.27 shows $\alpha = \beta$. Therefore

$$\binom{2x}{k} = \alpha(x) = \beta(x) = \sum_{i=0}^{k} \binom{x}{i}\binom{x}{k-i}$$

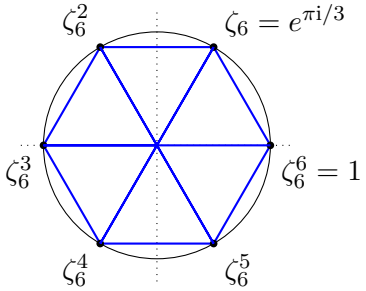
holds even for all $x \in \mathbb{C}$ and $k \in \mathbb{N}_0$. For example,

$$\sum_{i=0}^{k} \binom{1/2}{i}\binom{1/2}{k-i} = \binom{1}{k} = \begin{cases} 1 & \text{if } k \in \{0, 1\}, \\ 0 & \text{if } k \geq 2. \end{cases}$$

Definition 6.29. Let $n \in \mathbb{N}$ and $\zeta_n := e^{2\pi \mathrm{i}/n} \in \mathbb{C}$. The numbers $\zeta_n^k := e^{2\pi \mathrm{i}k/n} \in \mathbb{C}$ with $k = 1, \ldots, n$ are called *n-th roots of unity* (keyword: polar coordinates). For $\gcd(k, n) = 1$, ζ_n^k is called *primitive*. One calls

$$\Phi_n := \prod_{\substack{1 \leq k \leq n \\ \gcd(k,n)=1}} (X - \zeta_n^k) \in \mathbb{C}[X]$$

the *n-th cyclotomic polynomial*.



Remark 6.30. With Euler's φ-function, $\deg(\Phi_n) = \varphi(n)$ holds.

Example 6.31. It holds that

- $\zeta_1 = 1$ and $\Phi_1 = X - 1$,
- $\zeta_2 = -1$ and $\Phi_2 = X + 1$,
- $\zeta_3 = \frac{1}{2}(-1 + \sqrt{3}\mathrm{i})$, $\zeta_3^2 = \overline{\zeta_3}$ and
$$\Phi_3 = (X - \zeta_3)(X - \overline{\zeta_3}) = X^2 - (\zeta_3 + \overline{\zeta_3})X + 1 = X^2 + X + 1,$$
- $\zeta_4 = \mathrm{i}$, $\zeta_4^2 = -1 = \zeta_2$, $\zeta_4^3 = -\mathrm{i}$ and
$$\Phi_4 = (X - \mathrm{i})(X + \mathrm{i}) = X^2 + 1.$$

Theorem 6.32. *The n-th roots of unity are the roots of $X^n - 1$, i. e.*
$$X^n - 1 = \prod_{k=1}^{n} (X - \zeta_n^k)$$
is the prime factorization of $X^n - 1$ in $\mathbb{C}[X]$.

Proof. Because of $(\zeta_n^k)^n = e^{2k\pi\mathrm{i}} = (e^{2\pi\mathrm{i}})^k = 1$, $\zeta_n, \zeta_n^2, \ldots, \zeta_n^n$ are pairwise distinct roots of $X^n - 1$. Therefore, also
$$\alpha := X^n - 1 - \prod_{k=1}^{n} (X - \zeta_n^k)$$
has n roots, but $\deg \alpha < n$. Lemma 6.27 shows $\alpha = 0$. □

Example 6.33. A comparison of coefficients (or a geometric series) shows $1 + \zeta_3 + \zeta_3^2 = 0$. For $n \in \mathbb{N}_0$ it holds that
$$1 + \zeta_3^n + \zeta_3^{2n} = \begin{cases} 1 + \zeta_3 + \zeta_3^2 = 0 & \text{if } 3 \nmid n, \\ 1 + 1 + 1 = 3 & \text{if } 3 \mid n. \end{cases}$$
For $\alpha = \sum a_n X^n \in \mathbb{C}[[X]]$ one obtains
$$\frac{1}{3}\big(\alpha + \alpha(\zeta_3 X) + \alpha(\zeta_3^2 X)\big) = \frac{1}{3} \sum a_n \big(1 + \zeta_3^n + \zeta_3^{2n}\big) X^n = \sum a_{3n} X^{3n}.$$

Theorem 6.34. *For $n \in \mathbb{N}$, it holds that*
$$X^n - 1 = \prod_{d \mid n} \Phi_d.$$
In particular, Φ_n has integer coefficients.

Proof. For $d \mid n$, $\zeta_n^d = e^{2\pi \mathrm{d}\mathrm{i}/n} = \zeta_{n/d}$ is a primitive $\frac{n}{d}$-th root of unity. This shows
$$X^n - 1 \overset{6.32}{=} \prod_{k=1}^{n} (X - \zeta_n^k) = \prod_{d \mid n} \prod_{\substack{1 \le k \le n/d, \\ \gcd(k, n/d) = 1}} (X - \zeta^{dk}) = \prod_{d \mid n} \Phi_{n/d} = \prod_{d \mid n} \Phi_d.$$

For the second assertion, we argue by induction on n. For $n = 1$, $\Phi_1 = X - 1$ has integer coefficients. Now let $n > 1$ and assume the assertion is proven for $d < n$. Then $\alpha := \prod_{d \mid n,\, d < n} \Phi_d$ is monic with integer coefficients. Since the polynomial division $\Phi_n = (X^n - 1)/\alpha$ has no remainder in $\mathbb{C}[X]$, it also has no remainder in $\mathbb{Q}[X]$, i.e., $\Phi_n \in \mathbb{Q}[X]$. Since α is monic, no denominators occur in the process and the assertion follows. □

Remark 6.35. Gauss showed that the cyclotomic polynomials are irreducible in $\mathbb{Q}[X]$, i.e., $X^n - 1 = \prod_{d|n} \Phi_d$ is the prime factorization of $X^n - 1$ in $\mathbb{Q}[X]$ (without proof).

Example 6.36. One can use Theorem 6.34 to calculate Φ_n recursively:

$$\begin{aligned}
\Phi_1 &= X - 1, \\
\Phi_2 &= \frac{X^2 - 1}{\Phi_1} = X + 1, \\
\Phi_3 &= \frac{X^3 - 1}{\Phi_1} = X^2 + X + 1, \\
\Phi_4 &= \frac{X^4 - 1}{\Phi_1 \Phi_2} = \frac{X^4 - 1}{X^2 - 1} = X^2 + 1.
\end{aligned}$$

For $p \in \mathbb{P}$, one generally obtains $\Phi_p = \frac{X^p - 1}{X - 1} = X^{p-1} + X^{p-2} + \ldots + X + 1$ and inductively

$$\Phi_{p^n} = \frac{X^{p^n} - 1}{\Phi_1 \Phi_p \ldots \Phi_{p^{n-1}}} = \frac{X^{p^n} - 1}{X^{p^{n-1}} - 1} = X^{p^{n-1}(p-1)} + X^{p^{n-1}(p-2)} + \ldots + X^{p^{n-1}} + 1 = \Phi_p(X^{p^{n-1}}).$$

Theorem 6.37. *For $n \in \mathbb{N}$, it holds that*

$$\boxed{\Phi_n = \prod_{d|n} (X^{\frac{n}{d}} - 1)^{\mu(d)}}$$

with the classical Möbius function μ.

Proof. The polynomials $X^{n/d} - 1$ lie in the abelian group $\mathbb{Q}[[X]]^\times$ according to Lemma 4.8. The assertion therefore follows from Theorem 6.34 and the multiplicative version of the Möbius inversion (Remark 3.8). □

Example 6.38. It holds that

$$\Phi_6 = \frac{(X^6 - 1)(X - 1)}{(X^2 - 1)(X^3 - 1)} = \frac{X^3 + 1}{X + 1} = X^2 - X + 1.$$

Theorem 6.39.

(i) Let $n = p_1^{a_1} \ldots p_s^{a_s}$ be the prime factorization of $n \in \mathbb{N}$ and $q := p_1 \ldots p_s$. Then $\Phi_n = \Phi_q(X^{\frac{n}{q}})$.

(ii) For $n \in \mathbb{N}$ and $p \in \mathbb{P}$ with $p \nmid n$, it holds that $\Phi_{pn} = \frac{\Phi_n(X^p)}{\Phi_n}$.

(iii) For odd $n \geq 3$, it holds that $\Phi_{2n} = \Phi_n(-X)$.

Proof.

(i) According to Theorem 6.37, it holds that

$$\Phi_n = \prod_{d|n} (X^{\frac{n}{d}} - 1)^{\mu(d)} = \prod_{d|q} ((X^{\frac{n}{q}})^{\frac{q}{d}} - 1)^{\mu(d)} = \Phi_q(X^{\frac{n}{q}}).$$

(ii) It holds that

$$\Phi_{pn} = \prod_{d|n} (X^{\frac{pn}{d}} - 1)^{\mu(d)} (X^{\frac{pn}{pd}} - 1)^{\mu(pd)} = \prod_{d|n} \frac{((X^p)^{\frac{n}{d}} - 1)^{\mu(d)}}{(X^{\frac{n}{d}} - 1)^{\mu(d)}} = \frac{\Phi_n(X^p)}{\Phi_n}.$$

(iii) Because $n \geq 3$, $\varphi(n)$ is even according to Theorem 1.27. Thus $\Phi_n(-X)$ is monic. As in (ii), we have

$$\begin{aligned}\Phi_{2n} &= \prod_{d|n} \Big(\frac{(X^{\frac{n}{d}})^2 - 1}{X^{\frac{n}{d}} - 1} \Big)^{\mu(d)} = \prod_{d|n} (X^{\frac{n}{d}} + 1)^{\mu(d)} = \pm \prod_{d|n} (-X^{\frac{n}{d}} - 1)^{\mu(d)} \\ &= \pm \prod_{d|n} ((-X)^{\frac{n}{d}} - 1)^{\mu(d)} = \pm \Phi_n(-X) = \Phi_n(-X).\end{aligned}$$

□

Example 6.40. Theorem 6.39 allows an efficient calculation of Φ_n. For example:

$$\begin{aligned}\Phi_{24} &\overset{\text{(i)}}{=} \Phi_6(X^4) \overset{\text{(iii)}}{=} \Phi_3(-X^4) \overset{6.36}{=} (-X^4)^2 - X^4 + 1 = X^8 - X^4 + 1, \\ \Phi_{300} &= \Phi_{2^2 \cdot 3 \cdot 5^2} \overset{\text{(i)}}{=} \Phi_{30}(X^{10}) \overset{\text{(iii)}}{=} \Phi_{15}(-X^{10}) \overset{\text{(ii)}}{=} \frac{\Phi_3(X^5)}{\Phi_3}(-X^{10}) \\ &= \frac{X^{100} - X^{50} + 1}{X^{20} - X^{10} + 1} = X^{80} + X^{70} - X^{50} - X^{40} - X^{30} + X^{10} + 1.\end{aligned}$$

Remark 6.41. In the following, we investigate the coefficients of the cyclotomic polynomials.

Theorem 6.42. *Let $n \geq 2$ and $\Phi_n = \sum_{k=0}^{\varphi(n)} a_k X^k$. Then $a_k = a_{\varphi(n)-k}$ for $k = 0, \ldots, \varphi(n)$, i. e. the coefficients of Φ_n are "symmetric".*

Proof. Because $\Phi_2 = X + 1$, we may assume $n \geq 3$. According to Theorem 1.27, $\varphi(n)$ is then even. With ζ_n^k, $\zeta_n^{-k} = \zeta_n^{n-k} \neq \zeta_n^k$ is also a primitive n-th root of unity. This shows

$$a_0 = \Phi_n(0) = \prod_{\substack{1 \leq k \leq n \\ \gcd(k,n)=1}} \zeta_n^k = \prod_{\substack{1 \leq k \leq n/2 \\ \gcd(k,n)=1}} \zeta_n^k \zeta_n^{-k} = 1 = a_{\varphi(n)}.$$

In particular, $\alpha := \sum_{k=0}^{\varphi(n)} a_k X^{\varphi(n)-k}$ is monic. For a primitive n-th root of unity ζ_n^l, it holds that

$$\alpha(\zeta_n^l) = \sum_{k=0}^{\varphi(n)} a_k \zeta_n^{l(\varphi(n)-k)} = \zeta_n^{l\varphi(n)} \sum_{k=0}^{\varphi(n)} a_k (\zeta_n^{-l})^k = \zeta_n^{l\varphi(n)} \Phi_n(\zeta_n^{-l}) = 0.$$

Therefore α has the same roots as Φ_n and it follows that

$$\alpha = \prod_{\substack{1 \leq k \leq n \\ \gcd(k,n)=1}} (X - \zeta_n^k) = \Phi_n.$$

Thus $a_k = a_{\varphi(n)-k}$. □

Theorem 6.43. *It holds that $\Phi_n = X^{\varphi(n)} - \mu(n) X^{\varphi(n)-1} + \ldots - \mu(n) X + 1$ for $n \geq 2$.*

Proof. For $\alpha = \sum a_n X^n \in \mathbb{Q}[X] \setminus \mathbb{Q}$ let $f(\alpha) = a_{\deg(\alpha)-1}$. For $\alpha, \beta \in \mathbb{Q}[X]$ it then follows that $f(\alpha\beta) = f(\alpha) + f(\beta)$. With the Kronecker delta δ_{ij} it follows that

$$\sum_{d|n} f(\Phi_d) = f\Big(\prod_{d|n} \Phi_d\Big) = f(X^n - 1) = -\delta_{1n} = -\sum_{d|n} \mu(d)$$

for all $n \in \mathbb{N}$ (see Example 3.7). Möbius inversion shows $f(\Phi_n) = -\mu(n)$. The other coefficients result from Theorem 6.42. □

Remark 6.44. If one calculates $\Phi_1, \Phi_2, \ldots, \Phi_{104}$, one finds that all coefficients are 0 or ± 1. However, Suzuki has shown that all integers occur as coefficients of cyclotomic polynomials.

7. Polynomials in several variables

Definition 7.1. In this section, let $n \in \mathbb{N}$ be fixed. A *polynomial* in the variables $X_1, \ldots, X_n$ over a field K is a formal sum of the form $\alpha = \sum a_{k_1,\ldots,k_n} X_1^{k_1} \ldots X_n^{k_n}$, where only finitely many of the coefficients $a_{k_1,\ldots,k_n} \in K$ are non-zero. The set of these polynomials is denoted by $K[X_1, \ldots, X_n]$. One calls $\deg(\alpha) := \sup\{i_1 + \ldots + i_n : a_{i_1,\ldots,i_n} \neq 0\}$ the *degree* of α, where $\deg 0 = \sup \varnothing = -\infty$.

Remark 7.2.

(i) One easily sees that the calculation rules from Lemma 4.3 also hold in $K[X_1, \ldots, X_n]$. In fact, one can also regard every polynomial $\alpha \in K[X_1, \ldots, X_n]$ as a polynomial in X_i with coefficients in $K[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n]$. The rules $\deg(\alpha + \beta) \leq \max\{\deg \alpha, \deg \beta\}$ and $\deg(\alpha\beta) = \deg \alpha + \deg \beta$ from Lemma 6.3 also hold in $K[X_1, \ldots, X_n]$.

(ii) Although there is no division with remainder in $K[X_1, \ldots, X_n]$ for $n > 1$ (try to divide X_1 by X_2), Gauss has shown that $K[X_1, \ldots, X_n]$ nevertheless possesses a unique prime factorization (without proof).

(iii) Lemma 6.27 does not hold in $K[X_1, \ldots, X_n]$ for $n > 1$: For example, $\alpha = X - Y \in \mathbb{R}[X, Y]$ has infinitely many roots of the form $(x, x) \in \mathbb{R}^2$, although $\deg \alpha = 1$. We prove a substitute for this statement.

Theorem 7.3 (Interpolation)**.** *Let $L \subseteq K^n$ and $d \in \mathbb{N}$ with $|L| < \binom{n+d}{d}$. Then there exists an $\alpha \in K[X_1, \ldots, X_n] \setminus \{0\}$ with $\deg \alpha \leq d$ and $\alpha(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in L$.*

Proof. Let V_d be the K-vector space of all polynomials $\alpha \in K[X_1, \ldots, X_n]$ with $\deg \alpha \leq d$. Obviously, the monomials $X_1^{a_1} \ldots X_n^{a_n}$ with $a_1 + \ldots + a_n \leq d$ form a basis of V_d. If one sets $a_0 := d - a_1 + \ldots + a_n$, then $(a_0, \ldots, a_n)$ is a d-element multisubset of $\{0, \ldots, n\}$ (a_i is the multiplicity of i). From Theorem 1.22 it follows that $\dim V_d = \binom{n+1}{d} = \binom{n+d}{d}$. The linear map

$$V_d \to K^{|L|}, \qquad \alpha \mapsto (\alpha(x))_{x \in L},$$

has a non-trivial kernel because $\dim K^{|L|} = |L| < \dim V_d$. Thus there exists $\alpha \in V_d \setminus \{0\}$ with $\alpha(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in L$. □

Lemma 7.4 (Schwartz-Zippel)**.** *A polynomial $\alpha \in K[X_1, \ldots, X_n] \setminus \{0\}$ has at most $|K|^{n-1} \deg \alpha$ zeros in K^n.*

Proof. Induction on n. The case $n = 1$ is exactly Lemma 6.27. So let $n \geq 2$ and wlog. $|K| < \infty$. For $x \in K$ let $\alpha_x(X_1, \ldots, X_{n-1}) := \alpha(X_1, \ldots, X_{n-1}, x) \in K[X_1, \ldots, X_{n-1}]$. Let us assume $\alpha_x = 0$. Then $\beta(X_1, \ldots, X_n) := \alpha(X_1, \ldots, X_{n-1}, X_n + x) \in K[X_1, \ldots, X_n]$ has the form $\beta = X_n\gamma$ with $\gamma \in K[X_1, \ldots, X_n]$. This shows $\alpha = \beta(X_1, \ldots, X_{n-1}, X_n - x) = (X_n - x)\widetilde{\gamma}$ (one can thus split off linear factors as usual). Let $L := \{x \in K : \alpha_x = 0\}$. By iteration one obtains

$$\alpha = \prod_{x \in L} (X_n - x)\gamma$$

with $\deg \gamma \leq \deg(\alpha) - |L|$ and $\gamma_x \neq 0$ for $x \in K \setminus L$. Let

$$Z(\alpha) := \{(x_1, \ldots, x_n) \in K^n : \alpha(x_1, \ldots, x_n) = 0\}$$

be the set of zeros of α. Then

$$Z(\alpha) \subseteq (K^{n-1} \times L) \cup \bigcup_{x \in K \setminus L} (Z(\gamma_x) \times \{x\}).$$

By induction $|Z(\gamma_x)| \leq |K|^{n-2} \deg \gamma_x \leq |K|^{n-2}(\deg(\alpha) - |L|)$. In total it follows

$$|Z(\alpha)| \leq |K|^{n-1}|L| + |K||K|^{n-2}(\deg(\alpha) - |L|) = |K|^{n-1} \deg \alpha. \qquad \square$$

Lemma 7.5. *Let K be an infinite field and $\alpha, \beta \in K[X_1, \ldots, X_n]$ with $\alpha(x_1, \ldots, x_n) = \beta(x_1, \ldots, x_n)$ for all $x_1, \ldots, x_n \in K$. Then $\alpha = \beta$.*

Proof. For $n = 1$, $\alpha - \beta$ has infinitely many zeros and it follows that $\alpha = \beta$. Now let $n \geq 2$ and $\alpha - \beta = \sum_{k=0}^{d} \gamma_k X_n^k$ with $\gamma_0, \ldots, \gamma_d \in K[X_1, \ldots, X_{n-1}]$. For all $x_1, \ldots, x_{n-1} \in K$, $\sum_{k=0}^{d} \gamma_k(x_1, \ldots, x_{n-1}) X_n^k \in K[X_n]$ has infinitely many zeros and it follows again that $\gamma_k(x_1, \ldots, x_{n-1}) = 0$. By induction on n, $\gamma_0 = \ldots = \gamma_d = 0$ and thus $\alpha = \beta$. $\square$

Lemma 7.6. *Let $\alpha \in K[X_1, \ldots, X_n]$. Let d_i be the degree of α as a polynomial in X_i. Let $L_1, \ldots, L_n \subseteq K$ with $|L_i| > d_i$ for $i = 1, \ldots, n$. If $\alpha(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in L_1 \times \ldots \times L_n$, then $\alpha = 0$.*

Proof. For $n = 1$, $\deg \alpha = d_1 < |L_1|$ and the claim follows from Lemma 6.27. Now let $n \geq 2$ and assume the claim has already been proven for $n - 1$. We write $\alpha = \sum_{i=0}^{d_n} \alpha_i X_n^i$ with $\alpha_0, \ldots, \alpha_{d_n} \in K[X_1, \ldots, X_{n-1}]$. Obviously, the degree of α_i as a polynomial in X_j is at most d_j. Let $(x_1, \ldots, x_{n-1}) \in L_1 \times \ldots \times L_{n-1}$ be fixed. The polynomial $\beta := \alpha(x_1, \ldots, x_{n-1}, X_n) \in K[X_n]$ has at least the roots $x_n \in L_n$. Because $\deg \beta \leq d_n < |L_n|$, $\beta = 0$ by Lemma 6.27. This shows $\alpha_i(x_1, \ldots, x_{n-1}) = 0$ for $i = 0, \ldots, d_n$. By induction, it follows that $\alpha_0 = \ldots = \alpha_{d_n} = 0$ and $\alpha = 0$. $\square$

Lemma 7.7. *Let $L_1, \ldots, L_n \subseteq K$ be non-empty, finite subsets and $\alpha \in K[X_1, \ldots, X_n]$ with $\alpha(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in L_1 \times \ldots \times L_n$. Then there exist $\beta_1, \ldots, \beta_n \in K[X_1, \ldots, X_n]$ with $\deg \beta_i \leq \deg(\alpha) - |L_i|$ for $i = 1, \ldots, n$ and $\alpha = \sum_{i=1}^{n} \beta_i \prod_{x_i \in L_i} (X_i - x_i)$.*

Proof. For $1 \leq i \leq n$ let $d_i := |L_i| - 1$ and

$$\gamma_i = \prod_{x_i \in L_i} (X_i - x_i) = X_i^{d_1+1} - \sum_{j=0}^{d_i} c_{ij} X_i^j \in K[X_i].$$

For $x_i \in L_i$, $x_i^{d_i+1} = \sum_{j=0}^{d_i} c_{ij} x_i^j$ holds. By repeatedly replacing every term of the form X_i^e with $e > d_i$ in α by $X_i^{e-d_i-1} \sum_{j=0}^{d_i} c_{ij} X_i^j$, we obtain a polynomial $\widetilde{\alpha} \in K[X_1, \ldots, X_n]$ with the following properties:

- The degree of $\widetilde{\alpha}$ in X_i is at most d_i.

- For $(x_1, \ldots, x_n) \in L_1 \times \ldots \times L_n$, $\widetilde{\alpha}(x_1, \ldots, x_n) = \alpha(x_1, \ldots, x_n) = 0$ holds.
- $\alpha - \widetilde{\alpha} = \sum_{i=1}^{n} \beta_i \gamma_i$ with $\beta_1, \ldots, \beta_n \in K[X_1, \ldots, X_n]$ and $\deg \beta_i \leq \deg(\alpha) - |L_i|$ for $i = 1, \ldots, n$.

From Lemma 7.6 it follows that $\widetilde{\alpha} = 0$ and thus the claim. □

Theorem 7.8 (Combinatorial Nullstellensatz)**.** *Let $\alpha \in K[X_1, \ldots, X_n]$ with $\deg \alpha = d_1 + \ldots + d_n$, such that the coefficient of $X_1^{d_1} \ldots X_n^{d_n}$ in α does not vanish. Let $L_1, \ldots, L_n \subseteq K$ with $|L_i| > d_i$ for $i = 1, \ldots, n$. Then there exists $(x_1, \ldots, x_n) \in L_1 \times \ldots \times L_n$ with $\alpha(x_1, \ldots, x_n) \neq 0$.*

Proof. Wlog. let $|L_i| = d_i + 1$ for $i = 1, \ldots, n$. Assume indirectly $\alpha(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in L_1 \times \ldots \times L_n$. According to Lemma 7.7 there exist $\beta_1, \ldots, \beta_n \in K[X_1, \ldots, X_n]$ with $\deg \beta_i \leq \deg(\alpha) - |L_i| = \sum_{j \neq i} d_j - 1$ and

$$\alpha = \sum_{i=1}^{n} \beta_i \prod_{x_i \in L_i} (X_i - x_i).$$

By assumption there exists $i \in \{1, \ldots, n\}$, such that the coefficient of $X_1^{d_1} \ldots X_n^{d_n}$ in $\beta_i \prod_{x_i \in L_i} (X_i - x_i)$ does not vanish. But then $\deg \beta_i \geq \sum_{j \neq i} d_i$. From this contradiction the claim follows. □

Remark 7.9.

(i) Hilbert's (original) Nullstellensatz states: Let K be an algebraically closed field and $\alpha_1, \ldots, \alpha_k, \beta \in K[X_1, \ldots, X_n]$, such that β vanishes on all common zeros of $\alpha_1, \ldots, \alpha_k$. Then there exist $\beta_1, \ldots, \beta_k \in K[X_1, \ldots, X_n]$ and $s \in \mathbb{N}$ with $\beta^s = \alpha_1 \beta_1 + \ldots + \alpha_k \beta_k$.

(ii) In algebra one shows that for every prime number p a *primitive root* $x \in \mathbb{F}_p$ exists, i. e. $x^k \neq 1$ for $k = 1, \ldots, p-2$ (cf. Remark 6.26). If applicable it holds that

$$x^k \sum_{y \in \mathbb{F}_p} y^k = \sum_{y \in \mathbb{F}_p} (xy)^k = \sum_{y \in \mathbb{F}_p} y^k.$$

From this follows $\sum_{y \in \mathbb{F}_p} y^k = 0$ for $k = 0, \ldots, p-2$ (note $0^0 = 1$). This will be used in the next proof.

Theorem 7.10 (Chevalley-Warning)**.** *Let $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_p[X_1, \ldots, X_n]$ with $\sum_{i=1}^{k} \deg \alpha_i < n$. Then the number of common zeros of $\alpha_1, \ldots, \alpha_k$ in $\mathbb{F}_p^n$ is divisible by p. Thus, there can never be exactly one common zero.*

Proof. According to Remark 6.26 is

$$\beta(x) := \prod_{i=1}^{k} (1 - \alpha_i(x)^{p-1}) = \begin{cases} 1 & \text{if } x \text{ is a common zero of } \alpha_1, \ldots, \alpha_k, \\ 0 & \text{otherwise} \end{cases}$$

for all $x \in \mathbb{F}_p^n$. For the number N of common zeros it therefore holds $N \equiv \sum_{x \in \mathbb{F}_p^n} \beta(x) \pmod{p}$. Multiplying out β, one obtains a linear combination of monomials of the form $\prod_{i=1}^{n} x_i^{a_i}$ with

$$\sum_{i=1}^{n} a_i \leq (p-1) \sum_{j=1}^{k} \deg \alpha_i < (p-1)n$$

by assumption. In every such monomial there must therefore be an i with $a_i < p-1$. Wlog. let $i = 1$. According to Remark 7.9 it holds that

$$\sum_{x \in \mathbb{F}_p^n} \prod_{i=1}^{n} x_i^{a_i} = \sum_{x_1 \in \mathbb{F}_p} x_1^{a_1} \sum_{x_2,\ldots,x_n \in \mathbb{F}_p} \prod_{i=2}^{n} x_i^{a_i} = 0.$$

This shows $N \equiv \sum_{x \in \mathbb{F}_p^n} \beta(x) \equiv 0 \pmod{p}$. The second claim is clear because of $p \geq 2$. □

Theorem 7.11 (Erdős-Ginzburg-Ziv)**.** *Every multiset of $2n-1$ integers has n elements whose sum is divisible by n.*

Proof. Let $k := 2n-1$ and $a_1, \ldots, a_k \in \mathbb{Z}$. Let us first assume that $n = p$ is a prime number. Since we are only interested in divisibility by p, we can assume $a_1, \ldots, a_k \in \mathbb{F}_p$. The polynomials

$$\alpha := \sum_{i=1}^{k} X_i^{p-1}, \qquad \beta := \sum_{i=1}^{k} a_i X^{p-1}$$

in $\mathbb{F}_p[X_1, \ldots, X_k]$ have degree $p-1$ and a common zero $(0, \ldots, 0)$. Because $2(p-1) < k$, there must be another common zero $x := (x_1, \ldots, x_k)$ by Chevalley-Warning. Let $I := \{1 \leq i \leq k : a_i \neq 0\} \neq \varnothing$. By Fermat, it holds that

$$|I| \equiv \sum_{i=1}^{k} x_i^{p-1} = \alpha(x) = 0, \qquad \sum_{i \in I} a_i = \beta(x) = 0.$$

From $|I| \leq k$ it follows that $|I| = p$ and $\sum_{i \in I} a_i$ is divisible by p.

Now let us assume $n = pm$ with $p \in \mathbb{P}$ and $m > 1$. According to the first part of the proof, there exists $I_1 \subseteq \{1, \ldots, k\}$ with $|I_1| = p$ and $\sum_{i \in I_1} a_i \equiv 0 \pmod{p}$. Furthermore, there exists $I_2 \subseteq \{1, \ldots, k\} \setminus I_1$ with $\sum_{i \in I_2} a_i \equiv 0 \pmod{p}$. Because $k - 2(m-1)p = 2p - 1$, we find in the same way disjoint subsets $I_1, \ldots, I_{2m-1} \subseteq \{1, \ldots, k\}$ with $|I_j| = p$ and $b_j := \frac{1}{p} \sum_{i \in I_j} a_i \in \mathbb{Z}$ for $j = 1, \ldots, 2m-1$. By induction on n, there exists $J \subseteq \{1, \ldots, 2m-1\}$ with $|J| = m$ and $\sum_{j \in J} b_j \equiv 0 \pmod{m}$. For $I := \bigcup_{j \in J} I_j$ it holds that $|I| = mp = n$ and $\sum_{j \in J} a_j \equiv 0 \pmod{n}$. □

Example 7.12. Theorem 7.11 is optimal, because from the $(2n-2)$-element multiset

$$\{\underbrace{0, \ldots, 0}_{n-1}, \underbrace{1, \ldots, 1}_{n-1}\}$$

no n elements can be chosen whose sum is divisible by n.

Definition 7.13. A polynomial $\alpha \in K[X_1, \ldots, X_n]$ is called *symmetric* if

$$\alpha(X_{\pi(1)}, \ldots, X_{\pi(n)}) = \alpha(X_1, \ldots, X_n)$$

holds for all $\pi \in S_n$. The *elementary symmetric* polynomials of order n are $\sigma_0 := 1$ and

$$\sigma_k := \sum_{1 \leq i_1 < \ldots < i_k \leq n} X_{i_1} \ldots X_{i_k} \qquad (k \geq 1).$$

The *complete symmetric* polynomials of order n are $\tau_0 := 1$ and

$$\tau_k := \sum_{1 \leq i_1 \leq \ldots \leq i_k \leq n} X_{i_1} \ldots X_{i_k} \qquad (k \geq 1).$$

The *power sum polynomials* are $\rho_k := X_1^k + \ldots + X_n^k$ for $k \geq 0$.

Remark 7.14.

(i) The symmetric polynomials form a multiplicatively closed subspace of $K[X_1, \ldots, X_n]$ with basis

$$\beta_{(a_1,\ldots,a_n)} := \frac{1}{m_1! \ldots m_n!} \sum_{\pi \in S_n} X_{\pi(1)}^{a_1} \ldots X_{\pi(n)}^{a_n} \qquad (a_1 \geq \ldots \geq a_n \geq 0),$$

where m_i is the multiplicity of a_i in the sequence $(a_1, \ldots, a_n)$ (all coefficients are 1, so that the characteristic of the field plays no role). One obtains

$$\begin{aligned} \sigma_k &= \alpha_{(1^k, 0^{n-k})}, \\ \tau_k &= \sum_{a_1+\ldots+a_n=k} \alpha_{(a_1,\ldots,a_n)}, \\ \rho_k &= \alpha_{(k, 0^{n-1})}. \end{aligned}$$

(ii) Note: $\sigma_k = 0$ for $k > n$ and $\rho_0 = n$. It holds that $\deg \sigma_k = \deg \tau_k = \deg \rho_k = k$ for $k = 1, \ldots, n$.

Theorem 7.15 (Vieta)**.** *In the ring of formal power series $K[X_1, \ldots, X_n][[Y]]$ with coefficients in $K[X_1, \ldots, X_n]$, it holds that*

$$\begin{aligned} \prod_{k=1}^{n} (1 + X_k Y) &= \sum_{k=0}^{n} \sigma_k Y^k, \\ \prod_{k=1}^{n} \frac{1}{1 - X_k Y} &= \sum_{k=0}^{\infty} \tau_k Y^k. \end{aligned}$$

Proof. The first equation results from expanding the product. The second equation follows from

$$\prod_{k=1}^{n} \frac{1}{1 - X_k Y} = \prod_{k=1}^{n} \sum_{l=0}^{\infty} (X_k Y)^l = \sum_{k=0}^{\infty} \Big(\sum_{l_1+\ldots+l_n=k} X_1^{l_1} \ldots X_n^{l_n} \Big) Y^k = \sum_{k=0}^{\infty} \tau_k Y^k. \qquad \square$$

Theorem 7.16 (Girard-Newton identity)**.** *For $k \in \mathbb{N}$, it holds that*

$$\begin{aligned} \sum_{i=0}^{k} (-1)^i \sigma_i \tau_{k-i} &= 0, \\ \sum_{i=0}^{\min\{k,n\}} (-1)^i \sigma_i \rho_{k-i} &= \begin{cases} 0 & \text{if } k \geq n, \\ (-1)^k (n-k) \sigma_k & \text{if } k < n \end{cases} \end{aligned}$$

Proof.

(i) According to Vieta,

$$1 = \Big(\sum_{i=0}^{\infty} (-1)^i \sigma_i Y^i \Big) \Big(\sum_{j=0}^{\infty} \tau_j Y^j \Big) = \sum_{k=0}^{\infty} \Big(\sum_{i=0}^{k} (-1)^i \sigma_i \tau_{k-i} \Big) Y^k.$$

A comparison of coefficients yields the first assertion.

(ii) For $k \in \mathbb{N}_0$ and $0 \leq l \leq n-1$, let

$$\alpha(k,l) := \sum_{i=1}^{n} X_i^k \sigma_l(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n).$$

Then it holds that

$$\sigma_i \rho_{k-i} = \begin{cases} \rho_k = \alpha(k,0) & \text{if } i = 0, \\ \alpha(k-i,i) + \alpha(k-i+1, i-1) & \text{if } 1 \leq i \leq k < n, \\ \alpha(k-n+1, n-1) & \text{if } i = n \leq k. \end{cases}$$

As a telescoping sum, one obtains the second assertion (note $\alpha(0,l) = (n-l)\sigma_l$). □

Theorem 7.17 (Fundamental theorem on symmetric polynomials)**.** *For every symmetric $\alpha \in K[X_1, \ldots, X_n]$, there exists exactly one $\gamma \in K[X_1, \ldots, X_n]$ with $\alpha = \gamma(\sigma_1, \ldots, \sigma_n)$.*

Proof. Existence: Let wlog.

$$\alpha = \sum_{i_1, \ldots, i_n} a_{i_1, \ldots, i_n} X_1^{i_1} \ldots X_n^{i_n} \neq 0.$$

We order the tuples $(i_1, \ldots, i_n)$ lexicographically and argue by induction on

$$f(\alpha) := \max\big\{(i_1, \ldots, i_n) : a_{i_1, \ldots, i_n} \neq 0\big\}.$$

In the case $f(\alpha) = (0, \ldots, 0)$, then $\gamma := \alpha = a_{0,\ldots,0} \in K$. Now let $f(\alpha) = (d_1, \ldots, d_n) > (0, \ldots, 0)$. Because of $\alpha = \alpha(X_{\pi(1)}, \ldots, X_{\pi(n)})$ for all $\pi \in S_n$, it follows that $d_1 \geq \ldots \geq d_n$. Let

$$\beta := a_{d_1, \ldots, d_n} \sigma_1^{d_1 - d_2} \sigma_2^{d_2 - d_3} \ldots \sigma_{n-1}^{d_{n-1} - d_n} \sigma_n^{d_n}.$$

It holds that $f(\sigma_k^{d_k - d_{k+1}}) = (d_k - d_{k+1}) f(\sigma_k) = (d_k - d_{k+1}, \ldots, d_k - d_{k+1}, 0, \ldots, 0)$ and

$$f(\beta) = f(\sigma_1^{d_1 - d_2}) + \ldots + f(\sigma_n^{d_n}) = (d_1, \ldots, d_n).$$

The symmetric polynomial $\alpha - \beta$ therefore satisfies $f(\alpha - \beta) < (d_1, \ldots, d_n)$ and the existence of γ follows by induction.

Uniqueness: Let $\gamma, \delta \in K[X_1, \ldots, X_n]$ with $\gamma(\sigma_1, \ldots, \sigma_n) = \delta(\sigma_1, \ldots, \sigma_n)$. For $\rho := \gamma - \delta$ we then have $\rho(\sigma_1, \ldots, \sigma_n) = 0$ and we must show $\rho = 0$. Suppose indirectly $\rho \neq 0$. Let $d_1 \geq \ldots \geq d_n$ be the lexicographically largest n-tuple such that the coefficient of $X_1^{d_1 - d_2} X_2^{d_2 - d_3} \ldots X_n^{d_n}$ in ρ does not vanish. As above, $f(\sigma_1^{d_1 - d_2} \ldots \sigma_n^{d_n}) = (d_1, \ldots, d_n)$ holds. For every other summand $X_1^{e_1 - e_2} \ldots X_n^{e_n}$ of ρ, we have $f(\sigma_1^{e_1 - e_2} \ldots \sigma_n^{e_n}) < (d_1, \ldots, d_n)$. This yields $f(\rho(\sigma_1, \ldots, \sigma_n)) = (d_1, \ldots, d_n)$ in contradiction to $\rho(\sigma_1, \ldots, \sigma_n) = 0$. □

Example 7.18. We consider $\alpha = XY^3 + X^3Y - X - Y \in K[X,Y]$. With the notation from the proof, $f(\alpha) = (3,1)$ and

$$\beta := \sigma_1^2 \sigma_2 = (X+Y)^2 XY = X^3Y + 2X^2Y^2 + XY^3.$$

It follows that $\alpha - \beta = -2X^2Y^2 - X - Y$. In the next step, $f(\alpha - \beta) = (2,2)$ and

$$\beta_2 := -2\sigma_2^2 = -2X^2Y^2.$$

There remains $\alpha - \beta - \beta_2 = -X - Y = -\sigma_1$. Finally,

$$\alpha = \beta + \beta_2 - \sigma_1 = \sigma_1^2 \sigma_2 - 2\sigma_2^2 - \sigma_1 = \gamma(\sigma_1, \sigma_2)$$

with $\gamma = X^2Y - 2Y^2 - X$.

Theorem 7.19 (WARING formula)**.** *For $K = \mathbb{C}$ and $k \in \mathbb{N}$ it holds that*

$$\begin{aligned}\rho_k &= -k \sum_{(1^{a_1},\dots,k^{a_k})\in P(k)} \frac{(a_1+\ldots+a_k-1)!}{a_1!\ldots a_k!}(-\tau_1)^{a_1}\ldots(-\tau_k)^{a_k}\\ &= (-1)^k \sum_{(1^{a_1},\dots,k^{a_k})\in P(k)} \frac{(a_1+\ldots+a_k-1)!}{a_1!\ldots a_k!}(-\sigma_1)^{a_1}\ldots(-\sigma_k)^{a_k}.\end{aligned}$$

Proof. We introduce a new variable Y and manipulate the generating function:

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{\rho_k}{k}Y^k &= \sum_{i=1}^{n}\sum_{k=1}^{\infty}\frac{(X_iY)^k}{k} = \sum_{i=1}^{n}\log((1-X_iY)^{-1}) = \log\Big(\prod_{i=1}^{n}\frac{1}{1-X_iY}\Big)\\ &\overset{7.15}{=} \log\Big(1+\sum_{i=1}^{n}\tau_iY^i\Big) = \sum_{l=1}^{\infty}\frac{(-1)^{l-1}}{l}\Big(\sum_{i=1}^{n}\tau_iY^i\Big)^l\\ &\overset{1.20}{=} -\sum_{l=1}^{\infty}\frac{1}{l}\sum_{a_1+\ldots+a_n=l}\binom{l}{a_1,\dots,a_n}(-\tau_1)^{a_1}\ldots(-\tau_n)^{a_n}Y^{a_1+2a_2+\ldots+na_n}\\ &= -\sum_{k=1}^{\infty}\sum_{(1^{a_1},\dots,k^{a_k})\in P(k)}\frac{1}{a_1+\ldots+a_k}\binom{a_1+\ldots+a_k}{a_1,\dots,a_k}(-\tau_1)^{a_1}\ldots(-\tau_n)^{a_n}Y^k\\ \sum_{k=1}^{\infty}(-1)^k\frac{\rho_k}{k}Y^k &= -\sum_{i=1}^{n}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{(X_iY)^k}{k} = -\sum_{i=1}^{n}\log(1+X_iY) = -\log\Big(\prod_{i=1}^{n}(1+X_iY)\Big)\\ &\overset{7.15}{=} -\log\Big(1+\sum_{i=1}^{n}\sigma_iY^i\Big) = \sum_{l=1}^{\infty}\frac{(-1)^l}{l}\Big(\sum_{i=1}^{n}\sigma_iY^i\Big)^l\\ &\overset{1.20}{=} \sum_{l=1}^{\infty}\frac{1}{l}\sum_{a_1+\ldots+a_n=l}\binom{l}{a_1,\dots,a_n}(-\sigma_1)^{a_1}\ldots(-\sigma_n)^{a_n}Y^{a_1+2a_2+\ldots+na_n}\\ &= \sum_{k=1}^{\infty}\sum_{(1^{a_1},\dots,k^{a_k})\in P(k)}\frac{1}{a_1+\ldots+a_k}\binom{a_1+\ldots+a_k}{a_1,\dots,a_k}(-\sigma_1)^{a_1}\ldots(-\sigma_n)^{a_n}Y^k. \qquad \square\end{aligned}$$

Remark 7.20. The formula proven in the next theorem has an external similarity to the Leibniz formula for determinants:

$$\det(A) = \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{1\sigma(1)}\ldots a_{n\sigma(n)} \qquad (A=(a_{ij})\in K^{n\times n}).$$

Theorem 7.21. *For $A \in K^{n\times n}$ and $\sigma \in S_n$ let $\operatorname{tr}_\sigma(A) := \operatorname{tr}(A^{c_1})\ldots\operatorname{tr}(A^{c_k}) \in K$, where $c_1,\dots,c_k$ are the cycle lengths of σ (including cycles of length one). Then*

$$\det(A)n! = \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)\operatorname{tr}_\sigma(A).$$

Proof. We consider the entries of A as independent variables in the polynomial ring $K[X_{ij} : 1 \le i,j \le n]$. According to the Leibniz formula, the claim is then an equation of polynomials in $\mathbb{Z}[X_{ij}]$. According to Lemma 7.5, these polynomials are already equal if they coincide at all points $x_{ij} \in \mathbb{C}$. It therefore suffices to prove the claim for $K = \mathbb{C}$. Let $\lambda_1,\dots,\lambda_n \in \mathbb{C}$ be the eigenvalues of A with multiplicities. According to the Jordan normal form, $\operatorname{tr}(A^c) = \lambda_1^c+\ldots+\lambda_n^c$ holds for $c \ge 0$. For permutations $\sigma,\tau \in S_n$ of the same

cycle type, $\mathrm{tr}_\sigma(A) = \mathrm{tr}_\tau(A)$ certainly holds. We therefore sum over the partitions $(1^{a_1}, \ldots, n^{a_n}) \in P(n)$ using Theorem 2.26. For a corresponding permutation σ, $\mathrm{sgn}(\sigma) = \prod_{i=1}^{n}(-1)^{(i-1)a_i} = (-1)^{n+a_1+\ldots+a_n}$ holds. The right side is thus

$$\begin{aligned}
&n!(-1)^n \sum_{(1^{a_1},\ldots,n^{a_n})\in P(n)} \frac{(-1)^{a_1+\ldots+a_n}}{1^{a_1}a_1!\ldots n^{a_n}a_n!}(\lambda_1+\ldots+\lambda_n)^{a_1}\ldots(\lambda_1^n+\ldots+\lambda_n^n)^{a_n} \\
&= n!(-1)^n \sum_{(1^{a_1},\ldots,n^{a_n})\in P(n)} \prod_{l=1}^{n} \frac{(-\rho_l(\lambda_1,\ldots,\lambda_n))^{a_l}}{l^{a_l}a_l!} \overset{7.5}{=} n!\sigma_n(\lambda_1,\ldots,\lambda_n) \\
&= n!\lambda_1\ldots\lambda_n = \det(A)n!. \qquad \square
\end{aligned}$$

8. Bernoulli Numbers

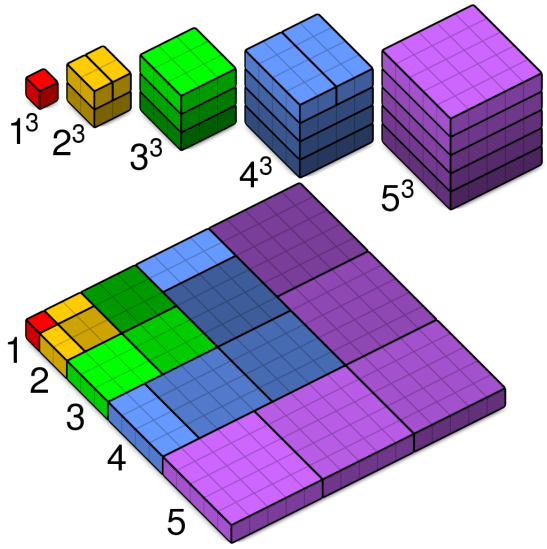
Remark 8.1. One easily shows (Exercise 1):

$$\begin{aligned}
1+2+\ldots+n-1 &= \binom{n}{2}, \\
1^2+2^2+\ldots+(n-1)^2 &= \frac{1}{4}\binom{2n}{3}, \\
1^3+2^3+\ldots+(n-1)^3 &= \binom{n}{2}^2 = (1+2+\ldots+(n-1))^2 \qquad (\textit{Nicomachus} \text{ identity}).
\end{aligned}$$

Gauss found the first formula as a child through

$$2(1+\ldots+n-1) = (1+n-1)+(2+n-2)+\ldots+(n-1+1) = n(n-1).$$

The third one can be seen as follows[4]:



[4]Source: https://math.stackexchange.com/questions/61482/proving-the-identity-sum-k-1n-k3-big-sum-k-1n-k-big2-without-i

We are looking for a general formula for $\sum_{k=1}^{n} k^m$ with $m \in \mathbb{N}$. For this we use that

$$\frac{\exp(X)-1}{X} = \sum_{n=1}^{\infty} \frac{X^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{X^n}{(n+1)!} \in \mathbb{Q}[[X]]$$

is invertible (Lemma 4.8).

Definition 8.2. The numbers $B_0, B_1, \ldots \in \mathbb{Q}$ with

$$\frac{X}{\exp(X)-1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} X^n$$

are called *Bernoulli numbers*.

Example 8.3. Because of

$$\Big(B_0 + B_1 X + \frac{B_2}{2} X^2 + \ldots\Big)\Big(1 + \frac{1}{2}X + \frac{1}{6}X^2 + \ldots\Big) = 1$$

it holds that $B_0 = 1$, $B_1 = -1/2$ and $B_2 = 1/6$. With the following lemma, B_n can be calculated recursively from B_k with $k < n$.

Lemma 8.4. *For $n \geq 2$ it holds that*

$$\sum_{k=0}^{n-1} \binom{n}{k} B_k = 0.$$

Proof. A comparison of coefficients as in Example 8.3 yields

$$0 = \sum_{k=0}^{n-1} \frac{B_k}{k!} \frac{1}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n-1} \binom{n}{k} B_k. \qquad \square$$

Lemma 8.5. *For $n \in \mathbb{N}$, $B_{2n+1} = 0$ holds.*

Proof. Let

$$\alpha = 1 + \sum_{k=2}^{\infty} \frac{B_k}{k!} X^k = \frac{X}{\exp(X)-1} + \frac{1}{2}X = \frac{X}{2} \frac{\exp(X)+1}{\exp(X)-1} \overset{4.23}{=} \frac{X}{2} \frac{\exp(\frac{1}{2}X) + \exp(-\frac{1}{2}X)}{\exp(\frac{1}{2}X) - \exp(-\frac{1}{2}X)}.$$

Because of $\alpha(X) = \alpha(-X)$, the claim follows by comparison of coefficients. $\square$

Example 8.6. Because of

$$\binom{5}{0} B_0 + \binom{5}{1} B_1 + \binom{5}{2} B_2 + \binom{5}{4} B_4 = 0$$

it follows that

$$B_4 = -\frac{1}{5}\Big(1 - \frac{5}{2} + \frac{10}{6}\Big) = -\frac{1}{30}.$$

Although the first Bernoulli numbers are relatively small, $|B_{2k}| \sim \frac{2(2k)!}{(2\pi)^{2k}} \to \infty$ holds (without proof).

Lemma 8.7. *For $n \in \mathbb{N}$, $(-1)^n B_{2n} < 0$ holds.*

Proof. According to (the proof of) Lemma 8.5,

$$\alpha := \frac{X}{\exp(X)-1} + \frac{1}{2}X = \sum_{k=0}^{\infty} \frac{B_{2k}}{(2k)!} X^{2k}.$$

Differentiation with the quotient rule yields

$$\alpha' = \frac{\exp(X) - 1 - X\exp(X)}{(\exp(X)-1)^2} + \frac{1}{2} = \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k-1)!} X^{2k-1}.$$

On the other hand,

$$\alpha^2 = \frac{X^2}{(\exp(X)-1)^2} + \frac{X^2}{\exp(X)-1} + \frac{1}{4}X^2 = \frac{X^2\exp(X)}{(\exp(X)-1)^2} + \frac{1}{4}X^2 = \alpha - X\alpha' + \frac{1}{4}X^2.$$

Substitution yields

$$\sum_{n=0}^{\infty}\Big(\sum_{k=0}^{n} \frac{B_{2k}B_{2(n-k)}}{(2k)!(2(n-k))!}\Big)X^{2n} = 1 + \frac{1}{4}X^2 + \sum_{n=1}^{\infty} B_{2n}\Big(\frac{1}{(2n)!} - \frac{1}{(2n-1)!}\Big)X^{2n}.$$

By comparison of coefficients,

$$\sum_{k=0}^{n} \binom{2n}{2k} B_{2k}B_{2(n-k)} = (2n)!B_{2n}\frac{(2n-1)! - (2n)!}{(2n-1)!(2n)!} = (1-2n)B_{2n}$$

for $n \geq 2$. We subtract $2B_{2n}$ on both sides and obtain

$$(2n+1)B_{2n} = -\sum_{k=1}^{n-1} \binom{2n}{2k} B_{2k}B_{2(n-k)} \tag{8.1}$$

for $n \geq 2$. We now prove the claim by induction on n. For $n = 1$, $B_2 = \frac{1}{6} > 0$. If the claim is already proven for all $k < n$, then it follows that

$$(-1)^n(2n+1)B_{2n} = -\sum_{k=1}^{n-1} \binom{2n}{2k} \underbrace{(-1)^k B_{2k}(-1)^{n-k}B_{2(n-k)}}_{>0} < 0$$

according to (8.1). □

Theorem 8.8 (Faulhaber's formula)**.** *For $n, m \in \mathbb{N}_0$,*

$$\boxed{\sum_{k=0}^{n-1} k^m = \frac{1}{m+1}\sum_{k=0}^{m} \binom{m+1}{k} B_k n^{m-k+1}}$$

holds.

Proof. According to Lemma 4.23, it holds that

$$\begin{aligned}\sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{n-1}k^mX^m &= \sum_{k=0}^{n-1}\sum_{m=0}^{\infty}\frac{k^mX^m}{m!} = \sum_{k=0}^{n-1}\exp(kX) = \sum_{k=0}^{n-1}\exp(X)^k\\ &= \frac{\exp(X)^n-1}{\exp(X)-1} = \frac{X}{\exp(X)-1}\frac{\exp(nX)-1}{X}\\ &= \sum_{k=0}^{\infty}\frac{B_k}{k!}X^k\cdot\sum_{k=0}^{\infty}\frac{n^{k+1}X^k}{(k+1)!} = \sum_{m=0}^{\infty}\Big(\sum_{k=0}^{m}\frac{B_k}{k!}\frac{n^{m-k+1}}{(m-k+1)!}\Big)X^m.\end{aligned}$$

Comparison of coefficients yields

$$\sum_{k=0}^{n-1}k^m = m!\sum_{k=0}^{m}\frac{B_k}{k!}\frac{n^{m-k+1}}{(m-k+1)!} = \frac{1}{m+1}\sum_{k=0}^{m}\binom{m+1}{k}B_kn^{m-k+1}$$ □

Example 8.9. For $m=4$, we have

$$\begin{aligned}1^4+2^4+\ldots+(n-1)^4 &= \frac{1}{5}\left(\binom{5}{0}B_0n^5+\binom{5}{1}B_1n^4+\binom{5}{2}B_2n^3+\binom{5}{4}B_4n\right)\\ &= \frac{1}{5}n^5-\frac{1}{2}n^4+\frac{1}{3}n^3-\frac{1}{30}n = \frac{6n^5-15n^4+10n^3-n}{30}.\end{aligned}$$

Remark 8.10.

(i) Faulhaber's formula expresses the sum $\sum_{k=0}^{n-1}k^m$ as a polynomial in n of degree $m+1$. Pascal showed the existence of this polynomial inductively, without explicitly calculating the coefficients.[5]

(ii) Bernoulli numbers also appear in analysis:

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}, \qquad \sum_{k=1}^{\infty}\frac{1}{k^4} = \frac{\pi^4}{90}, \qquad \sum_{k=1}^{\infty}\frac{1}{k^{2n}} = \frac{(2\pi)^{2n}(-1)^{n+1}B_{2n}}{2(2n)!}$$

(without proof). On the other hand, no formula is known for the *Apéry constant* $\sum_{k=1}^{\infty}\frac{1}{k^3} = 1.202\ldots$.

Theorem 8.11. *For $n\in\mathbb{N}_0$, it holds that*

$$B_n = \sum_{k=0}^{n}\frac{(-1)^k}{k+1}k!\left\{{n\atop k}\right\}.$$

Proof. For $k\in\mathbb{N}_0$, according to the functional equation, it holds that

$$(\exp(X)-1)^k = \sum_{l=0}^{k}(-1)^{k-l}\binom{k}{l}\exp(lX) = \sum_{n=0}^{\infty}\sum_{l=0}^{k}(-1)^{k-l}\binom{k}{l}\frac{l^n}{n!}X^n \overset{2.37}{=} \sum_{n=0}^{\infty}\frac{k!}{n!}\left\{{n\atop k}\right\}X^n.$$

[5] see [Beardon, *Sums of Powers of Integers*, Amer. Math. Monthly 103 (1996), 201–213]

It follows that

$$\begin{aligned}\sum \frac{B_n}{n!}X^n &= \frac{X}{\exp(X)-1} = \frac{\log(1+(\exp(X)-1))}{\exp(X)-1} = \frac{1}{\exp(X)-1}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{(\exp(X)-1)^k}{k}\\ &= \sum_{k=0}^{\infty}(-1)^k\frac{(\exp(X)-1)^k}{k+1} = \sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}\sum_{n=0}^{\infty}\frac{k!}{n!}\left\{{n \atop k}\right\}X^n\\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\Big(\sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}k!\left\{{n \atop k}\right\}\Big)X^n.\end{aligned}$$

The claim follows by comparison of coefficients. □

Theorem 8.12 (Clausen, von Staudt)**.** *For $n \in \mathbb{N}$, there exists a $z_n \in \mathbb{Z}$ with*

$$B_{2n} = z_n + \sum_{\substack{p\in\mathbb{P}\\(p-1)\,|\,2n}} \frac{1}{p}.$$

In particular, the (fully reduced) denominator of B_{2n} is the product of all prime numbers p with $p-1 \mid 2n$.

Proof. We analyze the summands $\frac{(-1)^k}{k+1}k!\left\{{2n \atop k}\right\}$ with $0 \le k \le 2n$ in Theorem 8.11. Because of $(-1)^k\left\{{2n \atop k}\right\} \in \mathbb{Z}$, we are only interested in $\frac{k!}{k+1}$. Let us first assume that $k+1$ is not a prime number. Then there exist $a, b \in \mathbb{Z}$ with $k+1 = ab$ and $a, b \le k$. In the case $a \ne b$, it follows that $k+1 = ab \mid k!$ and thus $\frac{(-1)^k}{k+1}k!\left\{{2n \atop k}\right\} \in \mathbb{Z}$. Now let $a = b$. In the case $2a \le k$, it again holds that $k+1 = ab \mid a(2a) \mid k!$ and $\frac{(-1)^k}{k+1}k!\left\{{2n \atop k}\right\} \in \mathbb{Z}$. In the case $2a > k$, we have $k+1 = a^2 \ge 2a \ge k+1$. It follows that $a = 2$ and $k = 3$. Then

$$(-1)^k k!\left\{{2n \atop k}\right\} \overset{2.37}{=} \sum_{l=0}^{3}(-1)^l\binom{3}{l}l^{2n} = -3 + 3\cdot 2^{2n} - 3^{2n} \overset{6.23}{\equiv} 1 + 3\cdot 4^n - 9^n \equiv 1 - 1 \equiv 0 \pmod 4.$$

Thus $\frac{(-1)^3}{4}3!\left\{{2n \atop 3}\right\} \in \mathbb{Z}$. The summands in Theorem 8.11 with $k+1 \notin \mathbb{P}$ are therefore all integers.

Now let $p := k+1 \in \mathbb{P}$. In the case $(p-1) \mid 2n$, it holds that

$$l^{2n} = (l^{p-1})^{\frac{2n}{p-1}} \overset{6.26}{\equiv} 1^{\frac{2n}{p-1}} \equiv 1 \pmod p$$

for $l = 1, \dots, p-1 = k$. This yields

$$(-1)^k k!\left\{{2n \atop k}\right\} \overset{2.37}{=} \sum_{l=1}^{k}(-1)^l\binom{k}{l}l^{2n} \equiv -1 + \sum_{l=0}^{k}(-1)^l\binom{k}{l} \equiv -1 + (1-1)^k \equiv -1 \pmod p.$$

Therefore

$$\frac{1}{p} + \frac{(-1)^k}{k+1}k!\left\{{2n \atop k}\right\} \in \mathbb{Z}.$$

Finally, let $(p-1) \nmid 2n$. Then there exist $m \in \{1, \dots, p-2\}$ and $q \in \mathbb{Z}$ with $2n = m + q(p-1)$ (division with remainder). For $l = 1, \dots, p-1$, it holds that $l^{2n} = l^m(l^{p-1})^q \equiv l^m 1^q \equiv l^m \pmod p$ by Fermat. This shows

$$(-1)^k k!\left\{{2n \atop k}\right\} = \sum_{l=1}^{k}(-1)^l\binom{k}{l}l^{2n} \equiv \sum_{l=1}^{k}(-1)^l\binom{k}{l}l^m \equiv (-1)^k k!\left\{{m \atop k}\right\} \equiv 0 \pmod p$$

because of $m < p-1 = k$. We again obtain $\frac{(-1)^k}{k+1} k! \left\{ {2n \atop k} \right\} \in \mathbb{Z}$. This proves the first assertion.

For the second assertion, let α be the denominator of B_{2n} and $\{p \in \mathbb{P} : p-1 \mid 2n\} = \{p_1, \ldots, p_s\}$. Because of $B_{2n} p_1 \ldots p_s \in \mathbb{Z}$, α is a divisor of $p_1 \ldots p_s$. Furthermore, it holds that

$$\sum_{i=1}^{s} \frac{1}{p_i} = \frac{p_2 \ldots p_s + p_1 p_3 \ldots p_s + \ldots + p_1 \ldots p_{s-1}}{p_1 \ldots p_s}$$

with

$$p_2 \ldots p_s + p_1 p_3 \ldots p_s + \ldots + p_1 \ldots p_{s-1} \equiv p_1 \ldots p_{i-1} p_{i+1} \ldots p_s \not\equiv 0 \pmod{p_i}$$

for $i = 1, \ldots, s$. Because of $\alpha \sum \frac{1}{p_i} \in \mathbb{Z}$, it follows that $p_1 \ldots p_s \mid \alpha$. Overall, we therefore have $\alpha = p_1 \ldots p_s$. □

Example 8.13. From Theorem 8.12 it follows that the denominator of B_{2n} is always divisible by 6. The denominator of B_{12}, for example, is $2 \cdot 3 \cdot 5 \cdot 7 \cdot 13 = 2730$.

Theorem 8.14 (Lucas)**.** *Let p be a prime and $n = \sum_{i \geq 0} n_i p^i$ as well as $k = \sum_{i \geq 0} k_i p^i$ with $0 \leq n_i, k_i \leq p-1$ for $i \geq 0$ (p-adic expansion). Then*

$$\boxed{\binom{n}{k} \equiv \prod_{i \geq 0} \binom{n_i}{k_i} \pmod{p}.}$$

In particular, $\binom{n}{k}$ is divisible by p if and only if there exists an $i \geq 0$ with $n_i < k_i$.

Proof. For $0 \leq m < p$, we have $\binom{p}{m} = \frac{p(p-1)\ldots(p-m+1)}{m!} \equiv 0 \pmod{p}$. In $\mathbb{F}_p[X]$, it therefore holds that

$$(1+X)^p = \sum_{m=0}^{p} \binom{p}{m} X^m = 1 + X^p.$$

Inductively, one obtains $(1+X)^{p^i} = 1 + X^{p_i}$ for all $i \geq 0$. It follows that

$$\sum_{m=0}^{n} \binom{n}{m} X^m = (1+X)^n = \prod_{i \geq 0} \left((1+X)^{p^i}\right)^{n_i} = \prod_{i \geq 0} \left(1 + X^{p^i}\right)^{n_i} = \prod_{i \geq 0} \sum_{m_i=0}^{n_i} \binom{n_i}{m_i} X^{m_i p^i}.$$

Multiplying out the product on the right side results in summands of the form

$$X^{\sum_{i \geq 0} m_i p^i} \prod_{i \geq 0} \binom{n_i}{m_i}.$$

Since the p-adic expansion of $m = \sum_{i \geq 0} m_i p^i$ is uniquely determined, X^m occurs only once in the sum. A comparison of coefficients at $m = k$ yields the first claim.

Because $n_i < p$, we have $\binom{n_i}{k_i} \not\equiv 0 \pmod{p}$ if $k_i \leq n_i$. In the case $k_i > n_i$, we have $\binom{n_i}{k_i} = 0$. From this, the second claim follows. □

Example 8.15. Because $2^n - 1 = 1 + 2 + \ldots + 2^{n-1}$, $\binom{2^n-1}{k}$ is odd for $k = 0, \ldots, 2^n - 1$.

9. Catalan Numbers

Remark 9.1.

(i) A *magma* is a set M together with a map (*binary operation*) $\cdot\colon M \times M \to M$, $(x, y) \mapsto x \cdot y$. For $x_1, \ldots, x_n \in M$, the product $x_1 \cdot \ldots \cdot x_n$ is then generally not well-defined, because different parenthesizations can yield different results (M is not necessarily associative). We investigate how many possible parenthesizations there are.

(ii) Even if M is associative, one can save computation time through suitable parenthesization. If, for example, A, B, and C are (real) matrices of format 10×20, 20×5, and 5×100, then for $(AB)C$ one only needs

$$10 \cdot 20 \cdot 5 + 10 \cdot 5 \cdot 100 = 6000$$

multiplications (of real numbers), while for $A(BC)$ one already needs

$$20 \cdot 5 \cdot 100 + 10 \cdot 20 \cdot 100 = 30,000$$

multiplications.

Example 9.2. The integers $\mathbb{Z}$ form a magma with respect to subtraction. It holds that

$$(1 - 2) - 3 = -4 \neq 2 = 1 - (2 - 3).$$

Definition 9.3. For $n \in \mathbb{N}_0$, let C_n be the number of possible parenthesizations of $x_1 \cdot \ldots \cdot x_{n+1}$, where $x_1, \ldots, x_{n+1}$ are elements of a magma. C_n is called the n-th *Catalan number*.

Example 9.4. Certainly $C_0 = C_1 = 1$ and $C_2 = 2$. Furthermore, $C_3 = 5$ holds because of

$$x_1(x_2(x_3x_4)), \qquad x_1((x_2x_3)x_4), \qquad (x_1x_2)(x_3x_4), \qquad ((x_1x_2)x_3)x_4, \qquad (x_1(x_2x_3))x_4.$$

Lemma 9.5 (Segner)**.** *For $n \in \mathbb{N}_0$, it holds that*

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

Proof. Every parenthesization of $x_1 \ldots x_{n+2}$ has the form $(x_1 \ldots x_{k+1})(x_{k+2} \ldots x_{n+2})$ for some $k \in \{0, \ldots, n\}$. Within the first pair of parentheses $(x_1 \ldots x_{k+1})$ there are C_k many parenthesizations and within the second pair of parentheses there are C_{n-k} possible parenthesizations. This shows the claim. □

Example 9.6.

(i) If $n \geq 2$ is even, then also $C_n = 2\sum_{k=0}^{n/2-1} C_k C_{n-1-k}$. In particular,

$$C_4 = 2(C_0C_3 + C_1C_2) = 2(5 + 2) = 14.$$

(ii) Let $\alpha = \sum a_n X^n \in \mathbb{Q}[[X]]$ be the inverse function of $X - X^2$ (Exercise 21). Then

$$X = \alpha - \alpha^2 = \sum_{n=0}^{\infty} a_n X^n - \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k a_{n-k} X^n = \sum_{n=0}^{\infty} \Big(a_n - \sum_{k=0}^{n} a_k a_{n-k} \Big) X^n$$

and it follows that $a_0 = 0$ (by Theorem 4.24), $a_1 = 1$ as well as $a_n = \sum_{k=1}^{n-1} a_k a_{n-k}$ for $n \geq 2$. Thus $a_n = C_{n-1}$ for $n \in \mathbb{N}$.

Theorem 9.7 (Euler)**.** *The generating function of C_n is*

$$\frac{1 - \sqrt{1 - 4X}}{2X}.$$

Proof. Let $\alpha = \sum C_n X^n$. Then

$$\begin{aligned}(1 - 2X\alpha)^2 &= (1 - 2C_0 X - 2C_1 X^2 - 2C_2 X^3 - \ldots)^2 \\ &= 1 - 4X + \sum_{n=2}^{\infty} \Big(-4C_{n-1} + 4 \sum_{k=0}^{n-2} C_k C_{n-2-k} \Big) X^n \overset{9.5}{=} 1 - 4X.\end{aligned}$$

This shows $1 - 2X\alpha = \sqrt{1 - 4X}$ and the claim follows. □

Theorem 9.8 (Catalan)**.** *For $n \in \mathbb{N}_0$, it holds that*

$$\boxed{C_n = \frac{1}{n+1} \binom{2n}{n}.}$$

Proof. According to Newton's binomial theorem,

$$\frac{1 - \sqrt{1 - 4X}}{2X} = \frac{1 - \sum \binom{1/2}{n} (-1)^n 4^n X^n}{2X} = \frac{1}{2} \sum (-1)^n \binom{1/2}{n+1} 4^{n+1} X^n.$$

A comparison of coefficients according to Theorem 9.7 shows

$$\begin{aligned}C_n &= (-1)^n \frac{1}{2} \binom{1/2}{n+1} 4^{n+1} = (-1)^n \frac{2^{n+1}}{2} \cdot \frac{2(1/2) \cdot 2(1/2-1) \cdot \ldots \cdot 2(1/2-n)}{(n+1)!} \\ &= \frac{2^n}{n+1} \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)}{n!} = \frac{1}{n+1} \frac{2 \cdot 4 \cdot \ldots \cdot 2n}{n!} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)}{n!} \\ &= \frac{1}{n+1} \frac{(2n)!}{(n!)^2} = \frac{1}{n+1} \binom{2n}{n}.\end{aligned}$$

□

Example 9.9. It holds that

$$C_5 = \frac{1}{6} \binom{10}{5} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{120} = 2 \cdot 3 \cdot 7 = 42.$$

Definition 9.10. Let $n \in \mathbb{N}$ and $\{a, b, c\} = \{1, 2, 3\}$. A permutation $\sigma \in S_n$ *possesses* the *pattern abc* if there exist $1 \leq i_1 < i_2 < i_3 \leq n$ with $\sigma(i_a) < \sigma(i_b) < \sigma(i_c)$. Otherwise, one says: σ *avoids* the pattern *abc*.

Example 9.11. The permutations in S_4 that possess the pattern 123 are:

$$\begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix},\begin{pmatrix}1&2&3&4\\1&2&4&3\end{pmatrix},\begin{pmatrix}1&2&3&4\\1&3&2&4\end{pmatrix},\begin{pmatrix}1&2&3&4\\1&3&4&2\end{pmatrix},\begin{pmatrix}1&2&3&4\\1&4&2&3\end{pmatrix},$$
$$\begin{pmatrix}1&2&3&4\\2&1&3&4\end{pmatrix},\begin{pmatrix}1&2&3&4\\2&3&1&4\end{pmatrix},\begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix},\begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix},\begin{pmatrix}1&2&3&4\\4&1&2&3\end{pmatrix}.$$

Thus, there are $|S_4| - 10 = 14$ permutations that avoid 123.

Theorem 9.12 (MacMahon)**.** *The number of permutations in S_n that avoid a pattern abc is C_n.*

Proof (Simion-Schmidt). Let $S_n(abc)$ be the set of permutations in S_n that avoid abc. With the convention $S_0 = \{\mathrm{id}_\varnothing\}$, it holds that $|S_0(abc)| = |S_1(abc)| = 1 = C_0 = C_1$ and $|S_2(abc)| = 2 = C_2$. So let $n \geq 3$ and assume the claim is already proven for smaller n.

We first show $|S_n(132)| = C_n$. Let $\sigma \in S_n(132)$ and $y := \sigma^{-1}(n)$. For $1 \leq x < y < z \leq n$, it then holds that $\sigma(x) > \sigma(z)$, because otherwise we would have $\sigma(x) < \sigma(z) < \sigma(y) = n$ and σ would possess the pattern 132. Thus $\{\sigma(x) : 1 \leq x < y\} = \{n-1, n-2, \ldots, n-y+1\}$ and $\{\sigma(z) : y < z \leq n\} = \{1, 2, \ldots, n-y\}$. The permutations

$$\begin{pmatrix}1&2&\cdots&y-1\\ \sigma(1)-n+y&\sigma(2)-n+y&\cdots&\sigma(y-1)-n+y\end{pmatrix} \in S_{y-1},$$
$$\begin{pmatrix}1&2&\cdots&n-y\\ \sigma(y+1)&\sigma(y+2)&\cdots&\sigma(n)\end{pmatrix} \in S_{n-y}$$

likewise avoid 132. For a fixed $y = \sigma^{-1}(n)$, there are thus $|S_{y-1}(132)||S_{n-y}(132)| = C_{y-1}C_{n-y}$ possible σ (induction). In total, one obtains

$$|S_n(132)| = \sum_{y=1}^{n} C_{y-1}C_{n-y} = \sum_{y=0}^{n-1} C_y C_{n-y-1} = C_n$$

according to Segner.

The bijections

$$\Gamma\colon S_n(abc) \to S_n(cba), \qquad \Delta\colon S_n(abc) \to S_n(4-a, 4-b, 4-c),$$
$$\begin{pmatrix}1&\cdots&n\\a_1&\cdots&a_n\end{pmatrix} \mapsto \begin{pmatrix}1&\cdots&n\\a_n&\cdots&a_1\end{pmatrix}, \qquad \begin{pmatrix}1&\cdots&n\\a_1&\cdots&a_n\end{pmatrix} \mapsto \begin{pmatrix}1&\cdots&n\\n+1-a_1&\cdots&n+1-a_n\end{pmatrix}$$

show $|S_n(132)| \overset{\Gamma}{=} |S_n(231)| \overset{\Delta}{=} |S_n(213)| \overset{\Gamma}{=} |S_n(312)|$ and $|S_n(123)| \overset{\Gamma}{=} |S_n(321)|$. It therefore suffices to construct a bijection $f\colon S_n(132) \to S_n(123)$. For $\sigma \in S_n(132)$, let

$$M(\sigma) := \{(k, \sigma(k)) : \sigma(k) < \sigma(i) \text{ for } i = 1, \ldots, k-1\}$$

be the set of *left-to-right minima*. We construct $f(\sigma)$ from σ by keeping $M(\sigma)$ fixed and sorting the remaining images of σ in descending order. For example:

$$\begin{pmatrix}1&2&3&4&5&6&7&8\\ \boxed{6}&7&\boxed{3}&4&\boxed{1}&2&5&8\end{pmatrix} \xrightarrow{f} \begin{pmatrix}1&2&3&4&5&6&7&8\\ \boxed{6}&8&\boxed{3}&7&\boxed{1}&5&4&2\end{pmatrix}.$$

Then $f(\sigma)$ is the unique permutation in $S_n(123)$ with $M(f(\sigma)) = M(\sigma)$. For $\sigma \in S_n(123)$, we conversely construct $\tau \in S_n(132)$ as follows: For $i = 1, \ldots, n$, let $\tau(i) := \sigma(i)$ if $(i, \sigma(i)) \in M(\sigma)$ and otherwise let

$\tau(i)$ be the smallest not yet used number that is larger than the next left-to-right minimum to the left of i. This yields a map $g\colon S_n(123) \to S_n(132)$, $\sigma \mapsto \tau$. For example:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \boxed{6} & 8 & \boxed{3} & 7 & \boxed{1} & 5 & 4 & 2 \end{pmatrix} \xrightarrow{g} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \boxed{6} & 7 & \boxed{3} & 4 & \boxed{1} & 2 & 5 & 8 \end{pmatrix}.$$

Again, $g(\sigma)$ is the unique permutation in $S_n(132)$ with $M(g(\sigma)) = M(\sigma)$. It follows that $f \circ g = \mathrm{id}_{S_n(123)}$ and $g \circ f = \mathrm{id}_{S_n(132)}$. In particular, f is a bijection and $|S_n(132)| = |S_n(123)|$. □

Remark 9.13. No simple formulas are known for the number of permutations that avoid 1234 (see `https://oeis.org/A005802`).

10. Groups

Remark 10.1. Many counting problems simplify when symmetries are taken into account. Symmetries are modeled by groups.

Definition 10.2. A *group* is a set G with an operation $\cdot\colon G \times G \to G$, $(x, y) \mapsto x \cdot y$, such that:

- $\forall x, y, z \in G : (x \cdot y) \cdot z = x \cdot (y \cdot z)$ (associative),
- $\exists e \in G : x \cdot e = e \cdot x = x$ (neutral element),
- $\forall x \in G : \exists y \in G : x \cdot y = y \cdot x = e$ (inverse elements).

One calls $|G|$ the *order* of G. In the case $|G| < \infty$, one calls G *finite*. If additionally

- $\forall x, y \in G : x \cdot y = y \cdot x$,

then G is called *abelian*.

Remark 10.3. As usual, one shows that the neutral element $e \in G$ is uniquely determined. We then write $e = 1_G = 1$ or also $e = 0$ if the operation is $+$. Furthermore, $x \in G$ possesses exactly one inverse element, which we denote by x^{-1} or $-x$. Then $(x^{-1})^{-1} = x$ and $(x \cdot y)^{-1} = y^{-1} \cdot x^{-1}$ (Attention!). Often we will omit the symbol $\cdot$ and write xy instead.

Example 10.4.

(i) The *trivial* group $G = \{1_G\}$.

(ii) The integers $\mathbb{Z}$ form an abelian group with respect to $+$. On the other hand, $(\mathbb{N}, +)$ is *not* a group, because for example the identity element is missing.

(iii) We had already mentioned that $\mathrm{Sym}(A)$ for a set A is a group with respect to composition of maps. For $|A| \geq 3$, $\mathrm{Sym}(A)$ is non-abelian (consider $(1,2)(1,3) = (1,3,2) \neq (1,2,3) = (1,3)(1,2)$ in S_3).

(iv) Let K be a field and V a finite-dimensional K-vector space. The invertible linear maps $V \to V$ form the *general linear group* $\mathrm{GL}(V)$ with respect to composition of maps.

(v) We consider the Euclidean space $\mathbb{R}^n$ with the standard inner product $(x,y) := \sum_{i=1}^n x_i y_i$ for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. In linear algebra, one studies the *orthogonal group*

$$\mathrm{O}(\mathbb{R}^n) := \big\{ f \in \mathrm{GL}(\mathbb{R}^n) : \forall x, y \in \mathbb{R}^n : (f(x), f(y)) = (x,y) \big\}.$$

The equation $\|f(x)\| \sqrt{(f(x), f(x))} = \sqrt{(x,x)} = \|x\|$ means that $f \in \mathrm{O}(\mathbb{R}^n)$ preserves lengths. As is well known, a map is orthogonal if and only if its matrix A satisfies the equation $AA^{\mathrm{t}} = 1_n$, where A^{t} denotes the transpose of A. In particular, $\det(A) = \pm 1$.

(vi) Let $\Delta \subseteq \mathbb{R}^2$ be a regular n-gon ($n \geq 3$) with center $(0,0)$ and let

$$G := \big\{ f \in \mathrm{O}(\mathbb{R}^2) : f(\Delta) = \Delta \big\}.$$

Certainly $\mathrm{id}_{\mathbb{R}^2} \in G$ and for $f, g \in G$ we have $(f \circ g)(\Delta) = f(g(\Delta)) = f(\Delta) = \Delta$, so $f \circ g \in G$. Furthermore, $f^{-1}(\Delta) = \Delta$. This shows that G is a group. One calls G the *symmetry group* of Δ. Since orthogonal maps preserve lengths, $f \in G$ must map vertices of Δ to vertices. Since two adjacent vertices x and y form a basis of $\mathbb{R}^2$ ($n \geq 3$), f is already uniquely determined as a linear map by $f(x)$ and $f(y)$. Because of $\|f(x) - f(y)\| = \|f(x-y)\| = \|x - y\|$, $f(x)$ and $f(y)$ are also adjacent vertices. There are thus n possibilities for $f(x)$ and subsequently two more possibilities for $f(y)$. This implies $|G| \leq 2n$. We show $|G| = 2n$. Obviously, the rotations by $\frac{2\pi k}{n}$ for $k = 1, \ldots, n$ all lie in G. If n is even, then G possesses $\frac{n}{2}$ reflections through two opposite vertices and just as many reflections through two opposite midpoints of sides. If n is odd, then G possesses exactly n reflections through one vertex and one midpoint of a side. In total, we have found $2n$ elements. One calls $D_{2n} := G$ therefore also the *dihedral group* of order $2n$. If we label the vertices with $1, \ldots, n$, then each $f \in D_{2n}$ describes a permutation in S_n. The rotations are the powers of the n-cycle $(1, \ldots, n)$. For $n = 3$ one obtains S_3.

Definition 10.5. A non-empty subset H of a group G is called a *subgroup*, if $xy^{-1} \in H$ for all $x, y \in H$. We then write $H \leq G$ or $H < G$ if $H \neq G$.

Remark 10.6. Let $H \leq G$. Then there exists an $x \in H$. Thus $1_G = xx^{-1} \in H$ and $x^{-1} = 1_G x^{-1} \in H$. For $x, y \in H$ we also have $xy = x(y^{-1})^{-1} \in H$. This shows that H with the restricted operation is itself a group.

Example 10.7.

(i) Every group G possesses the subgroups $\{1_G\}$ and G.

(ii) For every family of subgroups $H_i \leq G$ ($i \in I$), the intersection $\bigcap_{i \in I} H_i \leq G$ is also a subgroup (verify).

(iii) For $x \in G$ and $k \in \mathbb{Z}$ we define

$$x^k := \begin{cases} 1_G & \text{if } k = 0, \\ x \ldots x \ (k \text{ factors}) & \text{if } k > 0, \\ (x^{-1})^{-k} & \text{if } k < 0. \end{cases}$$

Then certainly $x^m x^n = x^{m+n}$ and $(x^m)^n = x^{mn}$ for $n, m \in \mathbb{Z}$. Furthermore, $\langle x \rangle := \{x^k : k \in \mathbb{Z}\}$ is a subgroup of G because of $x^k (x^l)^{-1} = x^{k-l} \in \langle x \rangle$. One calls $\langle x \rangle$ the subgroup *generated* by x. It is always abelian. Moreover, one calls $|\langle x \rangle|$ the *order* of x.

(iv) For $x_1, \dots, x_n \in G$ one defines more generally

$$\langle x_1, \dots, x_n \rangle := \bigcap_{\substack{H \leq G \\ x_1, \dots, x_n \in H}} H \leq G.$$

Obviously $\langle x_1, \dots, x_n \rangle$ contains all elements of the form $x_{i_1}^{\pm 1} \dots x_{i_k}^{\pm 1}$. Conversely, these elements themselves form a subgroup, which then must coincide with $\langle x_1, \dots, x_n \rangle$. In the case $G = \langle x_1, \dots, x_n \rangle$, one calls $x_1, \dots, x_n$ a *generating system* of G.

(v) It holds that $D_{2n} \leq \mathrm{O}(\mathbb{R}^2) \leq \mathrm{GL}(\mathbb{R}^2) \leq \mathrm{Sym}(\mathbb{R}^2)$. Furthermore, the n rotations in D_{2n} form a subgroup, which is called a *cyclic group* and is denoted by C_n. Obviously C_n is generated by the rotation by $2\pi/n$. In general, rotations lie in the *special orthogonal group*

$$\mathrm{SO}(\mathbb{R}^n) := \{f \in \mathrm{O}(\mathbb{R}^n) : \det(f) = 1\} = \mathrm{SL}(\mathbb{R}^n) \cap \mathrm{O}(\mathbb{R}^n).$$

Lemma 10.8. *For every group G and $x \in G$ it holds that*

$$|\langle x \rangle| = \inf\{n \in \mathbb{N} : x^n = 1_G\}$$

with the convention $\inf \varnothing = \infty$.

Proof. First, let $n \in \mathbb{N}$ be minimal with $x^n = 1$. For $k \in \mathbb{Z}$ there exist $q \in \mathbb{Z}$ and $r \in \mathbb{N}_0$ with $k = qn + r$ and $0 \leq r < n$ (division with remainder). Then $x^k = x^{qn+r} = (x^n)^q x^r = 1^q x^r = x^r$ and it follows that $\langle x \rangle = \{1, x, x^2, \dots, x^{n-1}\}$. Suppose there exist $0 \leq k < l < n$ with $x^k = x^l$. Then $x^{l-k} = 1$ with $1 \leq l - k < n$ in contradiction to the choice of n. This shows $|\langle x \rangle| = n$.

Now let $x^n \neq 1$ for all $n \in \mathbb{N}$. For $1 \leq k < l$ we then have $x^k \neq x^l$, because otherwise $x^{l-k} = 1$. This shows that the elements $x, x^2, \dots$ are pairwise distinct. In particular, $|\langle x \rangle| = \infty$. □

Definition 10.9. An *action* of a group G on a set Ω is a map $G \times \Omega \to \Omega$, $(g, \omega) \mapsto {}^g\omega$ with the following properties:

- $\forall \omega \in \Omega : {}^1\omega = \omega$,
- $\forall g, h \in G \, \forall \omega \in \Omega : {}^g({}^h\omega) = {}^{gh}\omega$.

For $\omega \in \Omega$ one calls ${}^G\omega := \{{}^g\omega : g \in G\}$ the *orbit* of ω and $G_\omega := \{g \in G : {}^g\omega = \omega\}$ the *stabilizer* of ω in G. If there exists only one orbit, the action is called *transitive*. One calls $|{}^G\omega|$ the *length* of the orbit.

Example 10.10.

(i) For every set A, $G := \mathrm{Sym}(A)$ acts on A by ${}^\sigma a := \sigma(a)$, because $\mathrm{id}(a) = a$ and ${}^\sigma({}^\tau a) = \sigma(\tau(a)) = (\sigma\tau)(a) = {}^{\sigma\tau}a$ for $a \in A$ and $\sigma, \tau \in G$. This action is transitive, because for distinct $a, b \in A$ the transposition (a, b) lies in G and ${}^{(a,b)}a = b$. This shows ${}^G a = A$ for all $a \in A$. The length of a cycle $\sigma \in \mathrm{Sym}(A)$ is simultaneously the length of an orbit of $\langle \sigma \rangle$.

(ii) For every K-vector space V, $G := \mathrm{GL}(V)$ acts on V by ${}^f v := f(v)$ for $f \in \mathrm{GL}(V)$ and $v \in V$. This follows from (i) because $G \leq \mathrm{Sym}(V)$. The orbit of the zero vector is ${}^G 0 = \{0\}$. In particular, the action is only transitive if $V = \{0\}$.

(iii) The groups $D_{2n} \leq \mathrm{O}(\mathbb{R}^n) \leq \mathrm{GL}(\mathbb{R}^n)$ act on $\mathbb{R}^n$. The action of D_{2n} can be restricted to the regular n-gon Δ. Furthermore, D_{2n} acts transitively on the set of vertices of Δ.

Remark 10.11.

(i) In the situation of Definition 10.9, $1 \in G_\omega \neq \varnothing$ holds and for $x, y \in G_\omega$ we have

$$^{xy^{-1}}\omega = {}^{x}({}^{y^{-1}}\omega) = {}^{x}({}^{y^{-1}}({}^{y}\omega)) = {}^{x}({}^{y^{-1}y}\omega) = {}^{x}({}^{1}\omega) = {}^{x}\omega = \omega.$$

This shows $xy^{-1} \in G_\omega$ and G_ω is a subgroup of G.

(ii) For $g \in G$, the map $\omega \mapsto {}^{g}\omega$ is a bijection on Ω with inverse map $\omega \mapsto {}^{g^{-1}}\omega$, because ${}^{g}({}^{g^{-1}}\omega) = {}^{gg^{-1}}\omega = {}^{1}\omega = \omega = \ldots = {}^{g^{-1}}({}^{g}\omega)$. Therefore, each g determines a permutation in $\mathrm{Sym}(\Omega)$.

(iii) We show that

$$\alpha \sim \beta :\Longleftrightarrow \exists g \in G : {}^{g}\alpha = \beta$$

defines an equivalence relation on Ω. Because of ${}^{1}\alpha = \alpha$, $\sim$ is reflexive. From ${}^{g}\alpha = \beta$ it follows that ${}^{g^{-1}}\beta = {}^{g^{-1}}({}^{g}\alpha) = {}^{g^{-1}g}\alpha = {}^{1}\alpha = \alpha$. Thus $\sim$ is symmetric. Finally, let ${}^{g}\alpha = \beta$ and ${}^{h}\beta = \gamma$ for $g, h \in G$ and $\alpha, \beta, \gamma \in \Omega$. Then ${}^{hg}\alpha = {}^{h}({}^{g}\alpha) = {}^{h}\beta = \gamma$. Therefore $\sim$ is transitive and an equivalence relation. The equivalence classes are exactly the orbits. In particular, the orbits form a partition of Ω.

(iv) Let $H \leq G$. Then H operates on G by ${}^{h}x := xh^{-1}$ for $h \in H$ and $x \in G$, because ${}^{1}x = x1^{-1} = x$ and ${}^{g}({}^{h}x) = {}^{g}(xh^{-1}) = (xh^{-1})g^{-1} = x(h^{-1}g^{-1}) = x(gh)^{-1} = {}^{gh}x$ for $g, h \in H$. The orbits have the form $xH := \{xh : h \in H\}$ for $x \in G$. They are called *left cosets*. Let $G/H := \{xH : x \in G\}$ and $|G : H| := |G/H|$. One calls $|G : H|$ the *index* of H in G.

Theorem 10.12 (Lagrange)**.** *For $H \leq G$, $\boxed{|G| = |G : H||H|}$ holds. In particular, $|H|$ and $|G : H|$ are divisors of $|G|$ if $|G| < \infty$.*

Proof. For $x \in G$, the map $H \to xH$, $h \to xh$ is bijective with inverse map $g \mapsto x^{-1}g$. Thus all left cosets of H have the cardinality $|H|$. The assertion follows because G is the disjoint union of the left cosets. □

Lemma 10.13. *For $H \leq G$ and $x, y \in G$, $xH = yH \Longleftrightarrow y^{-1}x \in H$ holds.*

Proof.

$$\begin{aligned} xH = yH &\Longleftrightarrow xH \cap yH \neq \varnothing \Longleftrightarrow \exists h, k \in H : xh = yk \\ &\Longleftrightarrow \exists h, k \in H : y^{-1}x = kh^{-1} \Longleftrightarrow y^{-1}x \in H. \end{aligned}$$

□

Theorem 10.14 (Orbit-Stabilizer Theorem)**.** *For every operation of G on a set Ω,*

$$\boxed{|{}^{G}\omega| = |G : G_\omega|}$$

holds for all $\omega \in \Omega$.

Proof. It suffices to show that the map $F\colon G/G_\omega \to {}^{G}\omega$, $xG_\omega \mapsto {}^{x}\omega$ is a bijection. Because of

$$xG_\omega = yG_\omega \overset{10.13}{\Longleftrightarrow} y^{-1}x \in G_\omega \Longleftrightarrow {}^{y^{-1}x}\omega = \omega \Longleftrightarrow {}^{x}\omega = {}^{y}\omega$$

for $x, y \in G$, F is well-defined and injective. The surjectivity follows from the definition of the orbit. □

Remark 10.15.

(i) Theorem 10.14 and Lagrange show that the orbit lengths are always divisors of the group order, if $|G| < \infty$.

(ii) If Δ is a system of representatives for the orbits of G on Ω, then one obtains the *orbit equation*

$$\boxed{|\Omega| = \sum_{\delta \in \Delta} |{}^G\delta| = \sum_{\delta \in \Delta} |G : G_\delta|.}$$

(iii) Let $|G| = 77$ and $|\Omega| = 23$. According to the orbit equation, there exist $a, b, c \in \mathbb{N}_0$ with $23 = a + 7b + 11c$. It follows that $a > 0$, i. e. G always has a fixed point on Ω.

(iv) One can prove Lucas's Theorem using group actions: We decompose $N := \{1, \ldots, n\}$ into a partition of the form

$$N = \bigcup_{i \geq 0} \bigcup_{j=1}^{n_i} A_{ij}$$

with $|A_{ij}| = p^i$ for $j = 1, \ldots, n_i$. For $A_{ij} = \{\alpha_1, \ldots, \alpha_{p^i}\}$ let $G_{ij} := \langle (\alpha_1, \ldots, \alpha_{p^i}) \rangle$. Then

$$G := \bigtimes_{i \geq 0} \bigtimes_{j=1}^{n_i} G_{ij} \leq S_n$$

is a p-group acting on N. Certainly G also acts on $\binom{N}{k}$. The orbit lengths are always p-powers according to Theorem 10.14. Therefore, the number of fixed points of G on $\binom{N}{k}$ is congruent to $\binom{n}{k}$ modulo p. A subset $K = \{\beta_1, \ldots, \beta_k\} \subseteq N$ is a fixed point under G if and only if K is the union of certain A_{ij}. For each $i \geq 0$, one must choose exactly k_i of the sets $A_{i1}, \ldots, A_{in_i}$. The number of fixed points is therefore $\prod_{i \geq 0} \binom{n_i}{k_i}$.

Theorem 10.16 (Burnside's Lemma)**.** *Let G be a finite group acting on a set Ω. For $g \in G$ let $f(g) := |\{\omega \in \Omega : g \in G_\omega\}|$ be the number of fixed points of g on Ω. Then*

$$\boxed{\frac{1}{|G|} \sum_{g \in G} f(g)}$$

is the number of orbits of G on Ω.

Proof. If $\alpha, \beta \in \Omega$ lie in the same orbit, then $|G : G_\alpha| = |{}^G\alpha| = |{}^G\beta| = |G : G_\beta|$. By Lagrange it follows that $|G_\alpha| = |G_\beta|$ (note: $|G| < \infty$). Let $\Delta \subseteq \Omega$ be a system of representatives for the orbits of G on Ω. Then

$$\begin{aligned} \sum_{g \in G} f(g) &= |\{(g, \omega) \in G \times \Omega : {}^g\omega = \omega\}| = \sum_{\omega \in \Omega} |G_\omega| = \sum_{\delta \in \Delta} |{}^G\delta||G_\delta| \\ &= \sum_{\delta \in \Delta} |G : G_\delta||G_\delta| = \sum_{\delta \in \Delta} |G| = |\Delta||G|. \end{aligned}$$

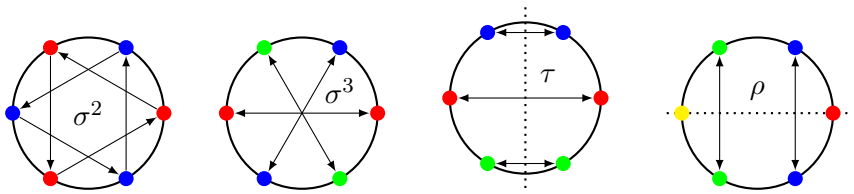
This shows the claim. □

Example 10.17.

(i) Let G be a finite group acting transitively on Ω with $|\Omega| > 1$. According to Burnside's Lemma, 1 is the average number of fixed points of elements from G. On the other hand, $f(1) = |\Omega| > 1$. There must therefore always be fixed-point-free elements in G. This generalizes Theorem 2.2.

(ii) We want to count necklaces with six beads, where beads in three colors are available. Naively, there are initially 3^6 such necklaces, some of which, however, are identical. We arrange the necklace such that the beads form a regular 6-gon. Rotation by $\pi/3$ will not change the necklaces. Likewise, we can rotate the necklace in space and thereby realize a reflection of the 6 vertices. Two necklaces are thus identical if and only if they lie in the same orbit under $G := D_{12}$. We apply Burnside's Lemma to the set Ω of the 3^6 necklaces.

Certainly $f(1) = 3^6$. A rotation $\sigma \in G$ by $\pi/3$ fixes only the three monochromatic necklaces, i. e. $f(\sigma) = 3$. The rotation σ^2 by $2\pi/3$ fixes the monochromatic necklaces and the necklaces with alternating colors. There are $f(\sigma^2) = 3^2$ of these. Analogously, one shows $f(\sigma^3) = 3^3$. Furthermore, $f(\sigma^4) = f(\sigma^{-2}) = 3^2$, $f(\sigma^5) = f(\sigma^{-1}) = 3$ as well as $\sigma^6 = 1$. Now let τ be one of the three reflections through two midpoints of sides. Then $f(\tau) = 3^3$. Finally, let ρ be one of the three reflections through two vertices. Then $f(\rho) = 3^4$.



According to Burnside's Lemma there are

$$\frac{1}{12}\left(3^6 + 2\cdot 3 + 2\cdot 3^2 + 3^3 + 3\cdot 3^3 + 3\cdot 3^4\right) = \frac{1}{4}\left(3^4(3+1) + 3^2(1+3) + 2 + 6\right)$$
$$= 81 + 9 + 2 = 92$$

different necklaces.

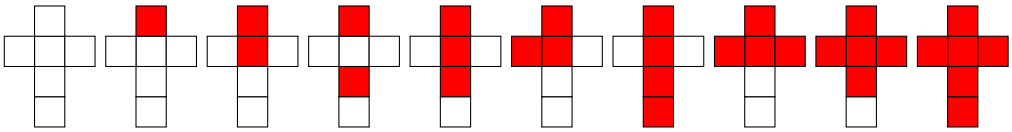
(iii) In how many ways can one color the six faces of a cube $W \subseteq \mathbb{R}^3$ if n colors are available? Naively: n^6. Rotations in space do not change W essentially. Reflections, however, do. We therefore seek the number of orbits under the rotation group of W (as a subgroup of $\mathrm{SO}(\mathbb{R}^3)$).

Rotation axis	Angle	Number of rotations	Number of fixed points
opposite side midpoints	0°	1	n^6
opposite side midpoints	±90°	6	n^3
opposite side midpoints	180°	3	n^4
opposite edge midpoints	180°	6	n^3
space diagonal	±120°	8	n^2
hline		24	

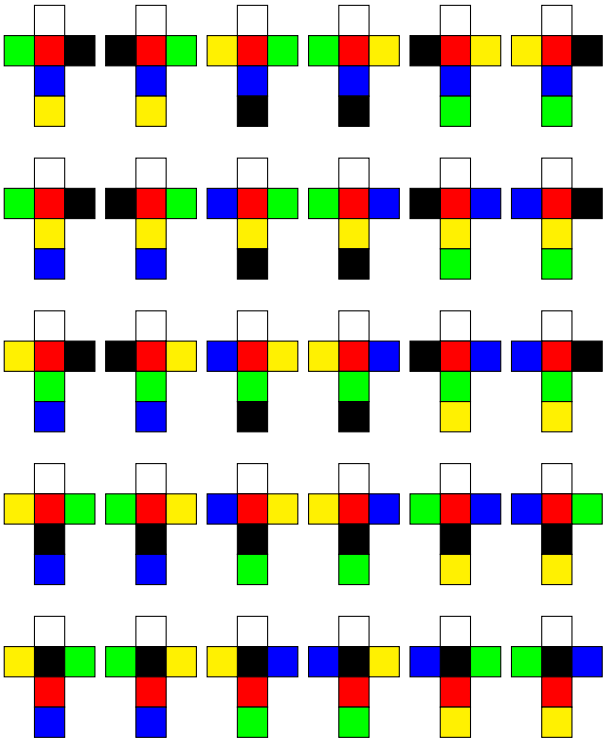
According to Burnside's Lemma, the number of colored cubes is given by

$$\frac{1}{24}\left(n^6 + 6n^3 + 3n^4 + 6n^3 + 8n^2\right) = \frac{n^2}{24}\left(n^4 + 3n^2 + 12n + 8\right).$$

For $n = 2$ one obtains the following ten cubes:



If one wants only cubes with pairwise distinct side colors, then one initially has $n(n-1)\dots(n-5)$ possibilities (variations without repetition). Since every non-trivial rotation is now fixed-point-free, Burnside's Lemma simplifies to $\frac{1}{24}n(n-1)\dots(n-5)$. For $n = 6$ one obtains the 30 MacMahon cubes:



(iv) There are

$$3^{14} \cdot 5^3 \cdot 7^2 \cdot 11 = 43,252,003,274,489,856,000$$

states of the $3 \times 3 \times 3$ Rubik's Cube, many of which, however, can be transformed into each other by spatial rotation and reflection. With Burnside's Lemma, the number reduces to

$$901,083,404,981,813,616$$

essentially different states.[6] With this, it was possible to show in 2010 that every state can be solved in at most 20 "moves" (*god's number*, see `https://cube20.org/`).

(v) With Burnside's Lemma, one can also show that there are

$$5,472,730,538$$

essentially different (filled) 9×9 Sudokus.[7]

[6] see [Sambale, Endliche Permutationsgruppen, Springer, 2017]
[7] see [Russell-Jarvis, Mathematics of Sudoku II, Mathematical Spectrum 39 (2006), 54–58]

Remark 10.18. In the following, we refine Burnside's Lemma to count, for example, necklaces with a certain *value* (the bead colors are no longer necessarily equivalent). For this, we consider an operation of G on Ω and another finite set Δ. Then G operates on $\Delta^\Omega = \{f\colon \Omega \to \Delta\}$ by $({}^g f)(\omega) := f({}^{g^{-1}}\omega)$ for $g \in G$, $\omega \in \Omega$ and $f \in \Delta^\Omega$, because $({}^1 f)(\omega) = f(\omega)$ and

$$({}^g({}^h f))(\omega) = ({}^h f)({}^{g^{-1}}\omega) = f({}^{h^{-1}}({}^{g^{-1}}\omega)) = f({}^{h^{-1}g^{-1}}\omega) = f({}^{(gh)^{-1}}\omega) = ({}^{gh} f)(\omega)$$

for $g, h \in G$. For a *weight function* $w\colon \Delta \to \mathbb{N}_0$, we define $w_i := |w^{-1}(i)|$ for $i \in \mathbb{N}_0$ and

$$W(X) := \sum_{i=0}^{\infty} w_i X^i \in \mathbb{Q}[X].$$

Finally, let

$$(\Delta^\Omega)_k := \Big\{ f \in \Delta^\Omega : \sum_{\alpha \in \Omega} w(f(\alpha)) = k \Big\}.$$

Because of

$$\sum_{\alpha \in \Omega} w(({}^g f)(\alpha)) = \sum_{\alpha \in \Omega} w(f({}^{g^{-1}}\alpha)) = \sum_{\alpha \in \Omega} w(f(\alpha))$$

for $g \in G$, G also operates on $(\Delta^\Omega)_k$. Finally, let $(1^{z_1(g)}, 2^{z_2(g)}, \ldots)$ be the cycle type of g as an element of $\mathrm{Sym}(\Omega)$.

Theorem 10.19 (Pólya)**.** *With the notation from Remark 10.18, the number of orbits of G on $(\Delta^\Omega)_k$ is the coefficient of X^k in*

$$\frac{1}{|G|} \sum_{g \in G} \prod_{i=1}^{\infty} W(X^i)^{z_i(g)}. \tag{10.1}$$

Proof. Let $f_k(g)$ be the number of fixed points of $g \in G$ on $(\Delta^\Omega)_k$. According to Burnside's Lemma, we must show that

$$\sum_{k=0}^{\infty} \Big(\frac{1}{|G|} \sum_{g \in G} f_k(g) \Big) X^k = \frac{1}{|G|} \sum_{g \in G} \sum_{k=0}^{\infty} f_k(g) X^k$$

coincides with (10.1). It thus suffices to prove

$$\sum_{k=0}^{\infty} f_k(g) X^k = \prod_{i=1}^{\infty} W(X^i)^{z_i(g)}$$

for $g \in G$. If g has cycle type $(l_1, \ldots, l_s)$ (i. e. cycles of length $l_1, l_2, \ldots, l_s$), then

$$\prod_{i=1}^{\infty} W(X^i)^{z_i(g)} = \prod_{i=1}^{s} (w_0 + w_1 X^{l_i} + w_2 X^{2l_i} + \ldots).$$

Every fixed point $f \in (\Delta^\Omega)_k$ of g is constant on the cycles of g. For the l_i-cycle σ of g, there are $|\Delta|$ possibilities for the assignment of f on the elements of σ. Exactly w_j of these possibilities contribute jl_i to k. The claim follows. □

Example 10.20.

(i) We consider once again the necklaces with six beads of three colors (red, blue and green). Let the red beads be worth 3 €, the blue ones 2 € and the green ones 1 €. How many necklaces worth 12 €

can be produced? Let $\Omega := \{1, \ldots, 6\}$, $\Delta := \{r, b, g\}$ and $w(r) := 3$, $w(b) := 2$ and $w(g) := 1$. Then $W(X) = X + X^2 + X^3$ and we are looking for the number of orbits of $G := D_{12}$ on $(\Delta^\Omega)_{12}$. The trivial element of G has cycle type (1^6). The rotation σ by $\pi/3$ has cycle type (6^1). Analogously one obtains $z_3(\sigma^2) = 2 = z_3(\sigma^4)$ and $z_2(\sigma^3) = 3$. For the reflections $\tau \in G$ through midpoints of sides, $z_2(\tau) = 3$ and the remaining three reflections through vertices have cycle type $(1^2, 2^2)$. Equation 10.1 now has the form

$$\begin{aligned}\frac{1}{12}\Big(W(X)^6 + 2W(X^6) + 2W(X^3)^2 + W(X^2)^3 + 3W(X^2)^3 + 3W(X)^2W(X^2)^2\Big)\\ = \ldots = X^{18} + X^{17} + 4X^{16} + 6X^{15} + 12X^{14} + 13X^{13}\\ +18X^{12} + 13X^{11} + 12X^{10} + 6X^9 + 4X^8 + X^7 + X^6\end{aligned}$$

Thus there are 18 necklaces worth 12€.

(ii) Pólya determined the number of isomers of alcohols and paraffins with Theorem 10.19.

Definition 10.21. Two groups G and H are called *isomorphic*, if a bijection $f\colon G \to H$ with $f(xy) = f(x)f(y)$ for all $x, y \in G$ exists. We then write $G \cong H$.

Remark 10.22. Obviously, the isomorphism of groups is an equivalence relation. The equivalence classes are called *isomorphism classes*. Isomorphic groups G and H differ only by the naming of their elements. In particular, G and H have the same properties (e. g. $|G| = |H|$, G abelian $\iff H$ abelian etc.).

Definition 10.23. Let $g(n)$ be the number of isomorphism classes of groups of order n.

Remark 10.24. Since every group of order n is uniquely determined by its multiplication table, $g(n) \leq n^{n^2} < \infty$ holds. The existence of cyclic groups shows $g(n) \geq 1$ for all $n \in \mathbb{N}$.

Example 10.25.

(i) Let G be a group with prime order $p = |G|$ and let $x \in G \setminus \{1\}$. Then $|\langle x\rangle| > 1$ and Lagrange shows $G = \langle x\rangle$. According to Lemma 10.8, $G = \{1, x, \ldots, x^{p-1}\}$ holds. Because of $x^i x^j = x^{i+j \pmod p}$, the multiplication table of G is already uniquely determined. One obtains an isomorphism $G \cong C_p$ by mapping x^k to the rotation by $2\pi k/p$. Therefore, C_p is the only group of order p up to isomorphism, i. e. $g(p) = 1$.

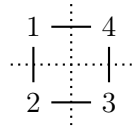
(ii) Let G be a group with four elements. If there exists an $x \in G$ with $G = \langle x\rangle$, then $G \cong C_4$ holds again. According to Lagrange, we can thus assume $x^2 = 1$ for all $x \in G$. Then $x = x^{-1}$ for all $x \in G$. It follows

$$xy = (xy)^{-1} = y^{-1}x^{-1} = yx$$

for $x, y \in G$, i. e. G is abelian. Obviously, G has the form $G = \{1, x, y, xy\}$ for certain $x, y \in G$. The multiplication table is thereby uniquely determined and G is isomorphic to the *Klein four-group*

$$V_4 := \{1, (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\} \leq S_4.$$

In particular, $g(4) = 2$. One also finds V_4 as a subgroup of D_8:



Theorem 10.26. *It holds that*

$$g(n) \leq \binom{[n!/e]}{\lambda(n)} \leq n^{n\lambda(n)} \leq n^{n\log_2(n)},$$

where $\lambda(n)$ is the number of prime factors of $n \in \mathbb{N}$ with multiplicities.

Proof. According to Example 10.25, we may assume $n \geq 4$. Let G be a group of order n and let $x_1, \ldots, x_d \in G$ be a minimal generating system of G, i. e. G cannot be generated with $d-1$ elements. In particular, $x_i \neq 1$ for $i = 1, \ldots, d$. By ${}^g x := gx$ for $g, x \in G$, G acts on itself. In particular, each $g \in G$ determines an element $f_g \in \mathrm{Sym}(G)$ with $f_g(x) = gx$ (Remark 10.11(ii)). Since every element in G can be written as a product of the x_i, the isomorphism type of G is already uniquely determined by $f_{x_1}, \ldots, f_{x_d}$. For $g \in G$, $f_{x_i}(g) = x_i g \neq g$, i. e. f_{x_i} is fixed-point free. Because of $f_{x_i}(1) = x_i$, $f_{x_1}, \ldots, f_{x_d}$ are pairwise distinct. According to Theorem 2.2, there are thus at most

$$\binom{[n!/e]}{d}$$

possibilities for G. Now let $H_i := \langle x_1, \ldots, x_i \rangle$ for $i = 1, \ldots, d$. Then $1 < H_1 < \ldots < H_d = G$ holds, because otherwise G could be generated with fewer than d elements. According to Lagrange,

$$n = |G| = |H_d : H_{d-1}||H_{d-1} : H_{d-2}| \ldots |H_1|$$

and it follows that $d \leq \lambda(n)$. Since every prime divisor of n is at least 2, it also holds that $\lambda(n) \leq \log_2(n) \leq \frac{[n!/e]}{2}$ because of $n \geq 4$. This shows

$$g(n) \leq \binom{[n!/e]}{d} \overset{\text{Exercise 3}}{\leq} \binom{[n!/e]}{\lambda(n)} \leq (n!)^{\lambda(n)} \leq (n^n)^{\lambda(n)} = n^{n\lambda(n)} \leq n^{n\log_2(n)}. \qquad \square$$

Example 10.27. It holds that

$$g(6) \leq \binom{[720/e]}{\lambda(6)} = \binom{265}{2} = 34.980.$$

In fact, there are only two groups of order 6 up to isomorphism, namely C_6 and S_3 (without proof). On the other hand, $g(2^{10}) = 49.487.365.422$ and $g(2^{11})$ is unknown. Over 99% of all groups of order ≤ 2000 have order $2^{10} = 1024$ (see `https://oeis.org/A000001`). The number of abelian groups of order 2^{10} is only $p(10) = 42$. This follows from the fundamental theorem of finite abelian groups (see Algebra 1).

Remark 10.28.

(i) Every group G acts on itself by *conjugation*, i. e. ${}^g x := gxg^{-1}$ for $g, x \in G$, because ${}^1 x = 1x1^{-1} = x$ and ${}^g({}^h x) = g(hxh^{-1})g^{-1} = (gh)x(gh)^{-1} = {}^{gh}x$ for $g, h, x \in G$. The stabilizer of $x \in G$ is then called the *centralizer* and is written as $\mathrm{C}_G(x) := \{g \in G : gx = xg\}$. The orbits are called *conjugacy classes* of G and their number $k(G)$ is called the *class number* of G. Certainly $k(G) \leq |G|$. The orbit equation becomes the *class equation*

$$\boxed{|G| = \sum_{x \in R} |G : \mathrm{C}_G(x)|,}$$

where R is a system of representatives for the conjugacy classes of G.

(ii) Let Ω be a set on which G acts. For $\omega \in \Omega$ and $g, x \in G$ it holds that

$$ {}^{x}\omega = \omega \iff {}^{gxg^{-1}}({}^{g}\omega) = {}^{g}\omega. $$

Therefore, the map $\omega \mapsto {}^{g}\omega$ is a bijection between the set of fixed points of x and the set of fixed points of gxg^{-1}. In particular, conjugate elements have the same number of fixed points on Ω. In Burnside's Lemma, one therefore only needs to sum over a system of representatives R of the conjugacy classes of G, i. e.

$$ \frac{1}{|G|} \sum_{x \in R} |G : \mathrm{C}_G(x)| f(x) = \sum_{x \in R} \frac{f(x)}{|\mathrm{C}_G(x)|} $$

is the number of orbits of G on Ω.

Lemma 10.29. *A finite group G is abelian if and only if $k(G) = |G|$ holds.*

Proof. If G is abelian and $g, x \in G$, then ${}^{g}x = gxg^{-1} = gg^{-1}x = x$, i. e. every conjugacy class has only one element. This shows $k(G) = |G|$. Conversely, if $k(G) = |G|$, then $\{x\} = {}^{G}x = \{gxg^{-1} : g \in G\}$ for all $x \in G$. This shows that G is abelian. □

Theorem 10.30 (Landau)**.** *Up to isomorphism, there are only finitely many finite groups with a given class number.*

Proof. Let G be a finite group with class number k and let $x_1, \ldots, x_k \in G$ be a system of representatives for the conjugacy classes of G with $x_k = 1_G$. Let $n_i := |\mathrm{C}_G(x_i)|$ for $i = 1, \ldots, k$. wlog. let $n_1 \leq \ldots \leq n_k = |G|$. The class equation shows

$$ 1 = \frac{1}{n_1} + \ldots + \frac{1}{n_k} \leq \frac{k}{n_1} $$

and $n_1 \leq k$. In particular, there are only finitely many possibilities for n_1. Now

$$ \frac{n_1 - 1}{n_1} = \frac{1}{n_2} + \ldots + \frac{1}{n_k} \leq \frac{k-1}{n_2} $$

and $n_2 \leq \frac{n_1(k-1)}{n_1 - 1}$. Thus there are also only finitely many possibilities for n_2. Continuing in this manner, one sees that there are only finitely many possibilities for $n_k = |G|$. The assertion now follows from Remark 10.24. □

Remark 10.31. Erdős and Turán have shown that $|G| < 2^{2^{k(G)}}$ holds.

Example 10.32. A group with class number 1 is trivial, because $\{1_G\}$ is always a conjugacy class. Now let $k(G) = 2$. As in Theorem 10.30, $n_1 \leq 2$ and there is only the solution $n_1 = 2 = n_2 = |G|$. Then $G \cong C_2$ holds. Finally, we assume $k(G) = 3$. In the case $n_1 = 3$, then $n_1 \leq n_2 \leq 3$ as in Theorem 10.30. This leads to $n_1 = n_2 = n_3 = 3 = |G|$ and $G \cong C_3$ according to Example 10.25. The case $n_1 = 2$ remains. Here $n_2 \leq 4$ and one has the solutions $(n_1, n_2, n_3) \in \big\{(2, 3, 6), (2, 4, 4)\big\}$. The possibility $n_2 = 4 = |G|$ is excluded, because then G would be abelian (Example 10.25) and would have four conjugacy classes according to Lemma 10.29. The first possibility leads to $G \cong S_3$. There are therefore exactly two groups with class number 3 (up to isomorphism). All groups with class number ≤ 14 are known (see `https://oeis.org/A073043`).

Theorem 10.33. *For $n \in \mathbb{N}$, $k(S_n) = p(n)$ (number of partitions of n). In particular, $p(n) \leq n!$.*

Proof. It suffices to show that the permutations with a given cycle type form a conjugacy class of S_n. Let $\sigma = (a_1, a_2, \ldots)(b_1, b_2, \ldots) \ldots \in S_n$ be a product of pairwise disjoint cycles and $\tau \in S_n$ arbitrary. Then

$$\tau\sigma\tau^{-1} = (\tau(a_1), \tau(a_2), \ldots)(\tau(b_1), \tau(b_2), \ldots) \ldots.$$

Since τ is injective, the cycles $(\tau(a_1), \tau(a_2), \ldots)$, $(\tau(b_1), \tau(b_2), \ldots), \ldots$ are also pairwise disjoint. In particular, σ and $\tau\sigma\tau^{-1}$ have the same cycle type.

Conversely, let $\sigma' = (a'_1, a'_2, \ldots)(b'_1, b'_2, \ldots) \ldots \in S_n$ be a permutation with the same cycle type as σ. For

$$\tau := \begin{pmatrix} a_1 & a_2 & \cdots & b_1 & b_2 & \cdots \\ a'_1 & a'_2 & \cdots & b'_1 & b'_2 & \cdots \end{pmatrix}$$

then $\tau\sigma\tau^{-1} = \sigma'$ holds. This shows the claim. □

Example 10.34. The conjugacy classes of S_5 are represented by 1, $(1,2)$, $(1,2,3)$, $(1,2,3,4)$, $(1,2,3,4,5)$, $(1,2)(3,4)$ and $(1,2)(2,3,4)$ ($p(5) = 7$).

Remark 10.35. A group G also acts by conjugation on the set of all subgroups by means of $gHg^{-1} := \{ghg^{-1} : h \in H\}$ for $g \in G$ and $H \leq G$. We write $H \sim K$ if $H, K \leq G$ are conjugate. The stabilizer of H is the *normalizer* $\mathrm{N}_G(H) := \{g \in G : gHg^{-1} = H\}$. According to the orbit-stabilizer theorem, $|G : \mathrm{N}_G(H)|$ is the length of the conjugacy class of H.

Theorem 10.36. *Let G be a finite group acting on a set Ω. For $H \leq G$, let $f(H) := |\{\omega \in \Omega : H \leq G_\omega\}|$ be the number of fixed points of H on Ω. Let μ be the Möbius function on the set of all subgroups of G (ordered by $\subseteq$). Then*

$$\frac{1}{|G|} \sum_{H \leq G} \mu(1, H) f(H)$$

is the number of orbits of G on Ω with length $|G|$.

Proof. For $g \in G$ and $\omega \in \Omega$, we have

$$x \in gG_\omega g^{-1} \iff g^{-1}xg \in G_\omega \iff {}^{g^{-1}xg}\omega = \omega \iff {}^{xg}\omega = {}^{g}\omega \iff x \in G_{g\omega},$$

i. e. $gG_\omega g^{-1} = G_{g\omega}$. Every orbit of G on Ω therefore determines a conjugacy class of stabilizers. For $H \leq G$, let $\rho(H)$ be the number of orbits whose stabilizers are conjugate to H. The length of these orbits is $|G : H|$. Now we have

$$\begin{aligned} f(H) &= \sum_{\substack{\omega \in \Omega \\ H \leq G_\omega}} 1 = \sum_{K \geq H} \sum_{\substack{\omega \in \Omega \\ G_\omega = K}} 1 = \sum_{K \geq H} \frac{1}{|G : \mathrm{N}_G(K)|} \sum_{\substack{\omega \in \Omega \\ G_\omega \sim K}} 1 \\ &= \sum_{K \geq H} \frac{|G : K|}{|G : \mathrm{N}_G(K)|} \rho(H) = \sum_{K \geq H} |\mathrm{N}_G(K) : K| \rho(K). \end{aligned}$$

Möbius inversion yields

$$|\mathrm{N}_G(H) : H| \rho(H) = \sum_{K \geq H} \mu(H, K) f(K).$$

The claim follows from the special case $H = 1$. □

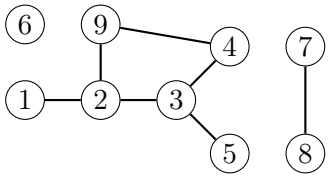
11. Graphs

Definition 11.1. A *graph* $\Omega = (\Omega_E, \Omega_K)$ of *order* $n \in \mathbb{N}$ consists of an n-element set Ω_E of *vertices* and a set $\Omega_K \subseteq \binom{\Omega_E}{2}$ of *edges*. We set $|\Omega| := |\Omega_E| = n$. Vertices $\alpha, \beta \in \Omega_E$ are called *adjacent*, if $\{\alpha, \beta\} \in \Omega_K$. The *degree* of a vertex is the number of its adjacent vertices. More generally, vertices $\alpha, \beta \in \Omega_E$ are called *connected*, if a *path* $\alpha = \alpha_1, \ldots, \alpha_m = \beta \in \Omega_E$ with $\{\alpha_i, \alpha_{i+1}\} \in \Omega_K$ for $i = 1, \ldots, m-1$ exists. This describes a partition on Ω_E, whose parts are called (*connected*) *components* of Ω. If there is only one component, then Ω is called *connected* and otherwise *disconnected*.

Remark 11.2. As usual, we will illustrate graphs with diagrams and number the vertices with natural numbers.

Example 11.3.

(i) The components of

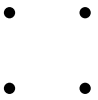


are $\{1, 2, 3, 4, 5, 9\}$, $\{6\}$ and $\{7, 8\}$.

(ii) The *trivial* graph $\mathcal{T}_n := \big(\{1, \ldots, n\}, \varnothing\big)$ without edges and the *complete* graph

$$\mathcal{V}_n := (\{1, \ldots, n\}, \binom{\{1, \ldots, n\}}{2})$$

of order $n \in \mathbb{N}$.

$\mathcal{T}_4$: 

$\mathcal{V}_4$: 

(iii) For every graph Ω there exists the *complementary* graph $\Omega^C = (\Omega_E, \binom{\Omega_E}{2} \setminus \Omega_K)$. A vertex in Ω of degree k corresponds to a vertex in Ω^C of degree $|\Omega| - k - 1$. Obviously $\mathcal{T}_n = \mathcal{V}_n^C$.

Ω : 

Ω^C : 

(iv) For graphs Ω and Δ, the (*disjoint*) *union* $\Omega \sqcup \Delta := (\Omega_E \sqcup \Delta_E, \Omega_K \sqcup \Delta_K)$ is also a graph with $|\Omega \sqcup \Delta| = |\Omega| + |\Delta|$. Obviously, every graph is the union of its components.

Remark 11.4. A graph Ω with $\Omega_E = \{1, \ldots, n\}$ is obviously uniquely determined by Ω_K. The number of all graphs of order n is therefore $|2^{\binom{\Omega_E}{2}}| = 2^{\binom{n}{2}}$. However, many of these graphs look the same. For example, there are six graphs of order 4 with only one edge.

Definition 11.5. Graphs Ω and Δ are called *isomorphic*, if a bijection $\sigma\colon \Omega_E \to \Delta_E$ with

$$\{x, y\} \in \Omega_K \iff \{\sigma(x), \sigma(y)\} \in \Delta_K$$

exists. We then write $\Omega \cong \Delta$.

Remark 11.6. As with groups, the isomorphism of graphs is an equivalence relation. Isomorphic graphs differ only by the labeling of the vertices and therefore have the same properties (same order, same number of edges, etc.).

Definition 11.7. Let $\mathfrak{g}(n)$ be the number of isomorphism classes of graphs of order n.

Remark 11.8. To determine $\mathfrak{g}(n)$, it suffices to consider the set $\mathcal{G}_n$ of all graphs Ω with $\Omega_E = \{1, \dots, n\} =: N$. Thus $\mathfrak{g}(n) \leq |\mathcal{G}_n| = 2^{\binom{n}{2}}$. We count the isomorphism classes in $\mathcal{G}_n$. For this, we consider that S_n acts on $\binom{N}{2}$ by

$$^{\sigma}\{a, b\} := \{\sigma(a), \sigma(b)\}$$

for $\sigma \in S_n$ and $\{a, b\} \in \binom{N}{2}$. This induces an action of S_n on $2^{\binom{N}{2}}$. Therefore S_n also acts on $\mathcal{G}_n$ by $^{\sigma}\Omega := (N, {}^{\sigma}\Omega_K)$.

Example: Ω : 1 — 2, 4 — 3 $\xrightarrow{\sigma = (1,2,3)}$ $^{\sigma}\Omega$: 1 2, 4 3

Two graphs in $\mathcal{G}_n$ are isomorphic if and only if they lie in the same orbit of S_n. To calculate the number of these orbits, we use Burnside's Lemma. For this, we must count how many fixed points $\sigma \in S_n$ has on $\mathcal{G}_n$. Let $\widetilde{\sigma}$ be the permutation on $\binom{N}{2}$ induced by σ. A graph $\Omega \in \mathcal{G}_n$ remains fixed under σ if and only if Ω_K is the union of orbits of $\widetilde{\sigma}$. The number $f(\sigma)$ of these graphs is thus $2^{z(\widetilde{\sigma})}$, where $z(\widetilde{\sigma})$ is the number of cycles of $\widetilde{\sigma}$. Burnside's Lemma shows

$$\mathfrak{g}(n) = \frac{1}{n!} \sum_{\sigma \in S_n} 2^{z(\widetilde{\sigma})}. \tag{11.1}$$

It is easy to see that $f(\sigma)$ depends only on the cycle type of σ (cf. Remark 10.28(ii)). Furthermore, we know from Theorem 2.26 how many elements of each cycle type exist. In (11.1), one therefore "only" needs to sum over the partitions of n. No explicit formula for $\mathfrak{g}(n)$ is known (cf. `https://oeis.org/A000088`).

Example 11.9.

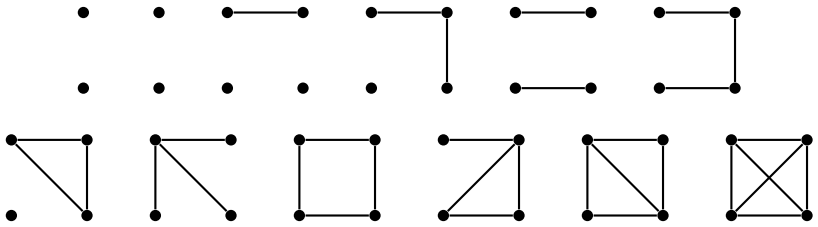
(i) Certainly $\mathfrak{g}(1) = 1$, $\mathfrak{g}(2) = 2$ and $\mathfrak{g}(3) = 4$:

$\mathcal{T}_3$: $\quad$ $\mathcal{V}_2 \sqcup \mathcal{T}_1$: $\quad$ $(\mathcal{V}_2 \sqcup \mathcal{T}_1)^C$: $\quad$ $\mathcal{V}_3$:

(ii) We consider the case $n = 4$ in Remark 11.8. Obviously $\sigma = 1$ has exactly $z(\widetilde{\sigma}) = \binom{4}{2} = 6$ cycles (of length 1) on $\binom{N}{2}$, i. e. $f(\sigma) = 2^6$. Furthermore $\sigma = (1, 2)$ has the cycles $(\{1, 2\})$, $(\{3, 4\})$, $(\{1, 3\}, \{2, 3\})$ and $(\{1, 4\}, \{2, 4\})$ on $\binom{N}{2}$ and it follows $f(\sigma) = 2^4$. Analogously one shows $f((1, 2, 3)) = 2^2$, $f((1, 2)(3, 4)) = 2^4$ and $f((1, 2, 3, 4)) = 2^2$. According to Theorem 2.26 there are six permutations of cycle type (2), eight of type (3), three of type (2^2) and six of type (4). With (11.1) one obtains

$$\begin{aligned}\mathfrak{g}(4) &= \frac{1}{4!}(f(1) + 6f((1, 2)) + 8f((1, 2, 3)) + 3f((1, 2)(3, 4)) + 6f((1, 2, 3, 4))) \\ &= \frac{1}{24}(2^6 + 6 \cdot 2^4 + 8 \cdot 2^2 + 3 \cdot 2^4 + 6 \cdot 2^2) = \frac{1}{24}(2^5(2 + 3 + 1) + 3 \cdot 2^3(2 + 1)) = 11.\end{aligned}$$

Representatives of these graphs are:



(iii) With Pólya's Theorem we can more precisely count the graphs with a given number of vertices and edges. For this, let $\Gamma := \binom{\{1,\dots,n\}}{2}$ and $\Delta := \{0,1\}$. Every graph Ω of order n corresponds to a mapping $f \in \Delta^\Gamma$ with $f(a) = 1$ if a is an edge of Ω and 0 otherwise. As in (i), S_n acts on Δ^Γ. With the notation from Theorem 10.19, let $w\colon \Delta \to \mathbb{N}_0$, $w(0) = 0$, $w(1) = 1$. Then $W(X) = 1 + X$. The number of orbits of S_n on $(\Delta^\Gamma)_k$ is exactly the number of graphs with n vertices and k edges up to isomorphism. With the calculations from (i) one obtains the polynomial

$$\begin{aligned}\frac{1}{4!}\sum_{\sigma\in S_4}\prod_{i=1}^{4}(1+X^i)^{c_i(\widetilde{\sigma})} &= \frac{1}{24}\Big((1+X)^6 + 6(1+X)^2(1+X^2)^2 \\ &\quad + 8(1+X^3)^2 + 3(1+X)^2(1+X^2)^2 + 6(1+X^2)(1+X^4)^1\Big) \\ &= \ldots = X^6 + X^5 + 2X^4 + 3X^3 + 2X^2 + X + 1.\end{aligned}$$

There are thus exactly three graphs of order 4 with three edges up to isomorphism (cf. figure above). The symmetry in the coefficients is explained by the bijection $\Omega \mapsto \Omega^C$.

Definition 11.10. The *automorphism group* $\mathrm{Aut}(\Omega)$ of a graph Ω consists of the isomorphisms from Ω to itself, i. e.

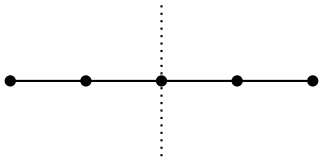
$$\mathrm{Aut}(\Omega) := \{\sigma \in \mathrm{Sym}(\Omega_E) : \{x,y\} \in \Omega_K \iff \{\sigma(x),\sigma(y)\} \in \Omega_K\} \leq \mathrm{Sym}(\Omega_E).$$

In the case $|\mathrm{Aut}(\Omega)| \neq 1$, Ω is called *symmetric* and otherwise *asymmetric*.

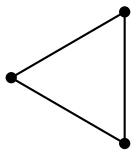
Example 11.11.

(i) For every graph Ω, $\mathrm{Aut}(\Omega) = \mathrm{Aut}(\Omega^C)$ holds. In particular, $\mathrm{Aut}(\mathcal{T}_n) = \mathrm{Aut}(\mathcal{V}_n) \cong S_n$.

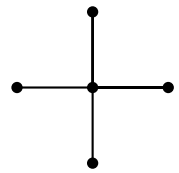
(ii) Let $\mathcal{G}_n := \big(\{1,\dots,n\}, \{\{i,i+1\} : i = 1,\dots,n-1\}\big)$ be a *path* of order $n \geq 2$. Then $\mathrm{Aut}(\mathcal{G}_n) \cong C_2$, where the non-trivial automorphism describes a reflection at the center of the path.

$\mathcal{G}_5$: 

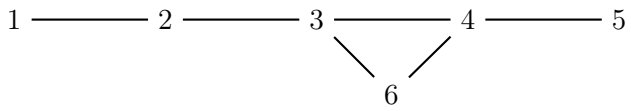
(iii) Let $\mathcal{K}_n := \big(\{1,\dots,n\}, \{\{i,i+1\} : i = 1,\dots,n-1\} \cup \{1,n\}\big)$ be a *cycle* of order $n \geq 3$. Then $\mathrm{Aut}(\mathcal{K}_n) \cong D_{2n}$, because every automorphism must preserve the distances of the vertices and thus corresponds to a symmetry of the regular n-gon.

$\mathcal{K}_3 = \mathcal{V}_3$: 

(iv) Let $\mathcal{S}_n := (\{1,\dots,n\}, \{\{i,n\} : i = 1,\dots,n-1\})$ be a *star* of order $n \geq 2$. Then $\mathcal{S}_n^C \cong \mathcal{T}_1 \sqcup \mathcal{V}_{n-1}$ and $\mathrm{Aut}(\mathcal{S}_n) \cong S_{n-1}$.

$\mathcal{S}_4$: 

(v) We examine the automorphism group of the graph

Ω : 

Every $\alpha \in \mathrm{Aut}(\Omega)$ must permute the two vertices 3 and 4 of degree 3, i. e. $\{\alpha(3), \alpha(4)\} = \{3,4\}$. The only common neighbor of 3 and 4 must therefore remain fixed, so $\alpha(6) = 6$. Besides 6, only vertex 2 has degree 2. This shows $\alpha(2) = 2$. It now follows easily that $\alpha = \mathrm{id}$ and $\mathrm{Aut}(\Omega) = \{\mathrm{id}\}$, i. e. Ω is asymmetric. This is the smallest asymmetric graph. Surprisingly, however, almost all graphs are asymmetric.

Theorem 11.12 (Erdős-Rényi)**.** *It holds that*

$$\mathfrak{g}(n) \sim \frac{1}{n!} 2^{\binom{n}{2}},$$

i. e. almost all graphs are asymmetric.

Proof. With (11.1) from Remark 11.8 we must show

$$\lim_{n\to\infty} \frac{1}{2^m} \sum_{\sigma \in S_n} 2^{z(\widetilde{\sigma})} = 1.$$

In the following, let n therefore always be "large enough". Let $2 \leq k < n$ and $\Sigma_k \subseteq S_n$ be the set of permutations with at most k fixed points. For $\sigma \in \Sigma_k$, $\widetilde{\sigma}$ then has at most $\binom{k}{2} + \frac{n-k}{2}$ fixed points (where equality only holds if σ is a disjoint product of $\frac{n-k}{2}$ transpositions). All other orbits of $\widetilde{\sigma}$ have at least length 2. Thus

$$\begin{aligned} z(\widetilde{\sigma}) &\leq m - \frac{1}{2}\Big(m - \binom{k}{2} - \frac{n-k}{2}\Big) = m - \frac{n(n-1) - k(k-1) - n + k}{4} \\ &= m - \frac{n(n-2) - k(k-2)}{4} \leq m - \frac{n(n-k)}{4} \qquad \text{(replace } k(k-2) \text{ by } n(k-2)\text{)}. \end{aligned}$$

For $1 \neq \sigma \in S_n \setminus \Sigma_k$ it holds that $\sigma \in \Sigma_{n-2}$ and

$$\begin{aligned} z(\widetilde{\sigma}) &\leq m - \frac{1}{2}\Big(m - \binom{n-2}{2} - 1\Big) \\ &= m - \frac{n(n-1) - (n-2)(n-3) - 2}{4} = m - n + 2. \end{aligned}$$

Because of $|\Sigma_k| \leq n! \leq n^n$ and $|S_n \setminus \Sigma_k| \leq \binom{n}{k}(n-k)! = \frac{n!}{k!} \leq n^{n-k}$ it follows

$$\begin{aligned} \sum_{\sigma \in S_n} 2^{z(\widetilde{\sigma})} &= 2^{z(1)} + \sum_{\sigma \in \Sigma_k} 2^{z(\widetilde{\sigma})} + \sum_{1 \neq \sigma \in S_n \setminus \Sigma_k} 2^{z(\widetilde{\sigma})} \\ &\leq 2^m + n^n 2^{m - n(n-k)/4} + n^{n-k} 2^{m-n+2}. \end{aligned}$$

It suffices to show that $n^n 2^{-n(n-k)/4}$ and $n^{n-k}2^{2-n}$ tend to 0 if one chooses k appropriately. We take the logarithm to the base 2 and then set $k = n - 5[\log n]$ (which is allowed for large n). Then

$$n\log(n) - \frac{n(n-k)}{4} \longrightarrow -\frac{n\log(n)}{4} \longrightarrow -\infty$$
$$(n-k)\log(n) + 2 - n \longrightarrow 5\log^2(n) - n \longrightarrow -\infty$$

for $n \to \infty$.

For the second assertion, let $t(n)$ be the number of asymmetric graphs of order n up to isomorphism. The orbits of these graphs under the above operation then have length $n!$. All other orbits have at most length $n!/2$. This shows

$$2^m \leq \frac{n!}{2}\big(\mathfrak{g}(n) + t(n)\big).$$

Dividing by $\mathfrak{g}(n)n!$ and letting n tend to ∞, it follows $\frac{t(n)}{\mathfrak{g}(n)} \to 1$. □

Remark 11.13. Frucht has shown that every finite group G is the automorphism group of a graph Ω. Except in the cases $G \in \{C_3, C_4, C_5\}$, one can choose $|\Omega| \leq 2|G|$. The smallest graph with automorphism group C_5 has order 15 (without proof).

Definition 11.14. A connected graph Ω is called a *tree*, if Ω contains no cycle, i. e. any two vertices of Ω are connected by exactly one path. Vertices of degree 1 are then (appropriately) called *leaves*.

Example 11.15. Lines and stars are always trees. In contrast, $\mathcal{T}_n$, $\mathcal{V}_n$ and $\mathcal{K}_n$ for $n \geq 3$ are not trees.

Theorem 11.16. *A connected graph Ω of order n is a tree if and only if $|\Omega_K| = n - 1$ holds. In particular, every connected graph of order n has at least $n - 1$ edges.*

Proof. Induction on n: For $n = 1$ the assertion is clear. So let Ω be a tree with order $n \geq 2$. Let $\omega = \omega_1, \ldots, \omega_k$ be a path of maximal length in Ω. Then ω is a leaf, because otherwise one could extend the path by one vertex. If one removes ω and the edge containing ω, one obtains a tree of order $n - 1$. By induction, this tree has exactly $n - 2$ edges. Thus $|\Omega_K| = n - 1$.

Conversely, let Ω be a connected graph of order n with k edges. Suppose that Ω contains a cycle Δ (i. e. $\Delta_E \subseteq \Omega_E$ and $\Delta_K \subseteq \Omega_K$). Then one can remove an edge of Δ such that Ω is still connected. This can be repeated until one obtains a tree. According to the first part, $k \geq n - 1$ holds, where equality occurs if and only if Ω is already a tree. □

Remark 11.17. We first count trees without considering isomorphism.

Theorem 11.18 (Cayley formula)**.** *There are exactly n^{n-2} trees with vertex set $\{1, \ldots, n\}$.*

Proof[8](Prüfer). wlog. let $n \geq 3$. For an n-element subset $M \subseteq \mathbb{N}$, let $B(M)$ be the set of all trees Ω with $\Omega_E = M$. We construct mutually inverse bijections

$$f\colon B(M) \to M^{n-2},$$
$$g\colon M^{n-2} \to B(M)$$

[8]alternative proofs can be found in [Aigner-Ziegler, Proofs from THE BOOK, Springer, 2014]

by induction on n. For $n = 3$, $\Omega \in B(M)$ is a path and we define $f(\Omega)$ as the center of Ω (the only vertex of degree 2). Conversely, if $\alpha \in M = M^{n-2}$ is given, we define $g(\alpha)$ as the path with center α. Then surely $f \circ g = \mathrm{id}_{M^{n-2}}$ and $g \circ f = \mathrm{id}_{B(M)}$.

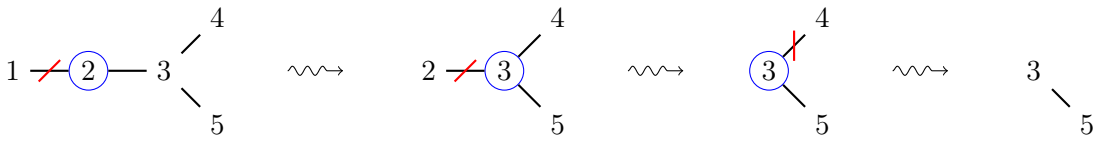
Now let $n \geq 4$ and $\Omega \in B(M)$ be given. Let $\alpha \in \Omega_E = M$ be the leaf with the smallest value in M and let β be the unique neighbor of α. We remove α and the edge $\{\alpha, \beta\}$ from Ω and thereby obtain a tree $\Delta \in B\big(M \setminus \{\alpha\}\big)$. Then we define $f(\Omega) := \big(\beta, f(\Delta)\big)$. Conversely, let $(\alpha_1, \ldots, \alpha_{n-2}) \in M^{n-2}$ be given. We set $\alpha := \min M \setminus \{\alpha_1, \ldots, \alpha_{n-2}\}$. Inductively, $g(\alpha_2, \ldots, \alpha_{n-2}) \in B\big(M \setminus \{\alpha\}\big)$ already exists and we can define

$$g(\alpha_1, \ldots, \alpha_{n-2}) := \big(M, g(\alpha_2, \ldots, \alpha_{n-2})_K \cup \{\alpha_1, \alpha\}\big) \in B(M).$$

To calculate $g \circ f$, we choose $\Omega \in B(M)$, the smallest leaf $\alpha \in \Omega_E$ and $\Delta \in B(M \setminus \{\alpha\})$ as above. By construction, $M \setminus f(\Omega)$ is the set of leaves of Ω. In particular, $\min M \setminus f(\Omega) = \alpha$. Inductively, $g(f(\Delta)) = \Delta$ holds and it follows that $g(f(\Omega)) = \Omega$. Conversely, let $(\alpha_1, \ldots, \alpha_{n-2}) \in M^{n-2}$ and $\alpha := \min M \setminus \{\alpha_1, \ldots, \alpha_{n-2}\}$. Inductively, $f(g(\alpha_2, \ldots, \alpha_{n-2})) = (\alpha_2, \ldots, \alpha_{n-2})$. Thus $M \setminus \{\alpha, \alpha_2, \ldots, \alpha_{n-2}\}$ are the leaves of $g(\alpha_2, \ldots, \alpha_{n-2})$ and $M \setminus \{\alpha_1, \ldots, \alpha_{n-2}\}$ are the leaves of $g(\alpha_1, \ldots, \alpha_{n-2})$. Therefore α is the smallest leaf of $g(\alpha_1, \ldots, \alpha_{n-2})$ and it follows that $f(g(\alpha_1, \ldots, \alpha_{n-2})) = (\alpha_1, \ldots, \alpha_{n-2})$. Thus f and g are mutually inverse bijections. In particular, f is bijective and $|B(M)| = |M^{n-2}| = n^{n-2}$. □

Remark 11.19. Using the notation from the above proof, $f(\Omega)$ is called the *Prüfer code* of a tree Ω.

Example 11.20. The Prüfer code of



is $(2, 3, 3)$.

Remark 11.21.

(i) Let Ω be a tree with $\Omega_E = \{1, \ldots, n\}$ and let d_i be the degree of the vertex $i \in \{1, \ldots, n\}$. According to Theorem 11.16, Ω has exactly $n - 1$ edges and it follows that $\sum_{i=1}^{n} d_i = 2(n-1)$. The bijection in the proof of Theorem 11.18 maps Ω to a sequence $(a_1, \ldots, a_{n-2})$ with $|\{1 \leq i \leq n-2 : a_i = k\}| = d_k - 1$ for $k = 1, \ldots, n$. Conversely, every such sequence arises from a tree with vertex degrees $d_1, \ldots, d_n$. The number of these trees is thus

$$\binom{n-2}{d_1 - 1, \ldots, d_n - 1}$$

according to Theorem 1.18.

(ii) The multiset $\{d_1 - 1, \ldots, d_n - 1\}$ describes a partition of $n - 2$ (if one omits zeros) and conversely, for every partition of $n - 2$ there is a corresponding tree. Isomorphic trees obviously yield the same multiset $\{d_1, \ldots, d_n\}$. The number of isomorphism classes of trees of order n is therefore at least $p(n-2)$.

(iii) Let $e_1, \ldots, e_s$ be the multiplicities in the multiset $\{d_1, \ldots, d_n\}$ (so $e_1 + \ldots + e_s = n$). Then one can arrange the numbers $d_1, \ldots, d_n$ in

$$\binom{n}{e_1, \ldots, e_s}$$

ways. The number of trees whose vertex degrees yield the multiset $\{d_1, \ldots, d_n\}$ is thus

$$\binom{n-2}{d_1 - 1, \ldots, d_n - 1}\binom{n}{e_1, \ldots, e_s}.$$

For $n = 7$ and $\{d_1, \ldots, d_7\} = \{1, 1, 1, 2, 2, 2, 3\}$ one obtains

$$\binom{5}{1, 1, 1, 2}\binom{7}{3, 3, 1} = \frac{5!7!}{2!3!3!} = 5! \cdot 2 \cdot 5 \cdot 7 = 120 \cdot 70 = 8.400.$$

Theorem 11.22. *The number of trees of order n with exactly k leaves is $\left\{ {n-2 \atop n-k} \right\} \frac{n!}{k!}$.*

Proof. For the choice of the leaves $e_1, \ldots, e_k \in \{1, \ldots, n\}$ of Ω there are $\binom{n}{k}$ possibilities. Let $e_{k+1}, \ldots, e_n$ be the remaining vertices. The Prüfer code of Ω then corresponds to a surjective mapping $\{1, \ldots, n-2\} \to \{e_{k+1}, \ldots, e_n\}$. The sought number of trees is therefore

$$\binom{n}{k}\left\{ {n-2 \atop n-k} \right\}(n-k)! = \left\{ {n-2 \atop n-k} \right\}\frac{n!}{k!}$$

according to Theorem 2.35. □

Example 11.23. According to Theorem 11.18 there are $5^3 = 125$ trees with vertex set $\{1, \ldots, 5\}$. We determine the isomorphism classes by going through the partitions of 3:

- The partition (1^3) yields the $\binom{5}{2,3} = 10$ possible degree sequences $(1, 1, 2, 2, 2), \ldots, (2, 2, 2, 1, 1)$. There are therefore

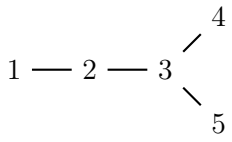
  $$10\binom{3}{1, 1, 1} = 60$$

  such trees, which are all isomorphic to the path $\mathcal{G}_5$.

- The partition $(2, 1)$ yields $\binom{5}{1,1,3} = 20$ possible degree sequences $(1, 1, 1, 2, 3), \ldots, (3, 2, 1, 1, 1)$. This results in

  $$20\binom{3}{1, 2} = 60$$

  trees, which are all isomorphic to

  (the vertex of degree 3 forms a star with three other vertices).

- The partition (3) yields $\binom{5}{4,1} = 5$ degree sequences $(1, 1, 1, 1, 4), \ldots, (4, 1, 1, 1, 1)$ and 5 trees, which are all isomorphic to $\mathcal{S}_5$.

There are therefore only three trees of order 5 up to isomorphism (cf. `https://oeis.org/A000055`).

Remark 11.24. In contrast to Theorem 11.12, Erdős and Rényi also proved that almost all trees are symmetric. For this, they showed that most trees possess *cherries*. These are two leaves with a common neighbor. Swapping these leaves yields a non-trivial automorphism.

12. Exercises

Exercise 1 (2 + 2 + 2 + 2 + 2 points)**.** Let $n \in \mathbb{N}$. Find "combinatorial" proofs (i. e. if possible without induction) for the following identities:

(a) $1 + 2 + \ldots + n - 1 = \binom{n}{2}$.

(b) $1 + 3 + 5 + \ldots + 2n - 1 = n^2$.

(c) $1^2 + 2^2 + \ldots + (n-1)^2 = \frac{1}{4}\binom{2n}{3}$.
Hint: Determine the cardinality of $\{(a, b, c) \in \mathbb{N}^3 : a, b < c \leq n\}$ in two ways.

(d) $1\binom{n}{1} + 2\binom{n}{2} + \ldots + n\binom{n}{n} = n2^{n-1}$.

(e) $2^0\binom{n}{0} + 2^1\binom{n}{1} + \ldots + 2^n\binom{n}{n} = 3^n$.

Exercise 2 (2 + 2 + 2 points)**.** Answer the following questions with justification:

(a) How many subsets of $\{1, \ldots, 10\}$ contain at least one odd number?

(b) How many possibilities are there to arrange the letters of MISSISSIPPI such that the four S are not all next to each other?

(c) How many possibilities are there to place 7 people at a round table, where two possibilities are considered the same if every person has the same two neighbors?

Exercise 3 (3 points)**.** Show

$$\binom{n}{0} < \binom{n}{1} < \ldots < \binom{n}{\lfloor n/2 \rfloor} = \binom{n}{\lceil n/2 \rceil} > \binom{n}{\lceil n/2 \rceil + 1} > \ldots > \binom{n}{n}$$

for $n \in \mathbb{N}$, where $\lfloor n/2 \rfloor$ denotes rounding down and $\lceil n/2 \rceil$ denotes rounding up.

Exercise 4 (2 points)**.** Let M be a non-empty set. Show that M has as many subsets with even cardinality as with odd cardinality. Determine the number of these subsets.

Exercise 5 (2 points)**.** How many six-digit decimal numbers are there whose digits are strictly monotonically increasing?

Exercise 6 (2 points)**.** Show

$$\sum_{k=n}^{m} \binom{k}{n} = \binom{m+1}{n+1}$$

for $n, m \in \mathbb{N}_0$.
Hint: Count the $(n+1)$-element subsets of $\{0, \ldots, m\}$ with a given maximum.
Remark: Since the involved binomial coefficients in Pascal's triangle take the shape of a hockey stick, this is called the *hockey-stick identity*.

Exercise 7 (3 points)**.** Prove without induction the multinomial theorem

$$(a_1 + \ldots + a_n)^k = \sum_{\substack{(k_1,\ldots,k_n)\in\mathbb{N}_0^n \\ k_1+\ldots+k_n=k}} \binom{k}{k_1,\ldots,k_n} a_1^{k_1} \ldots a_n^{k_n}$$

for $a_1, \ldots, a_n \in \mathbb{C}$ and $k \in \mathbb{N}$.

Exercise 8 $(2 + 2 + 2 + 2$ points)**.**

(a) Calculate $\varphi(1500)$.

(b) Determine $\{n \in \mathbb{N} : \varphi(n) = 8\}$.

(c) Write $(1, 2, 3, 4)(7, 6, 5, 4, 3)$ as a product of disjoint cycles.

(d) Write $(1, 4, 6, 3)(2, 7, 9)$ as a product of transpositions. How many transpositions are needed for this?

Exercise 9 (3 points)**.** Every permutation $\sigma \in S_n$ can be uniquely written as a product of disjoint cycles

$$\sigma = (a_1, \ldots, a_k)(b_1, \ldots, b_l) \ldots$$

if one requires $a_1 = \max\{a_1, \ldots, a_k\} < b_1 = \max\{b_1, \ldots, b_l\} < \ldots$ (cf. Remark 2.8). Here we also count the 1-cycles. Show that the mapping

$$\begin{aligned} \Psi\colon S_n &\to S_n, \\ \sigma &\mapsto \begin{pmatrix} 1 & 2 & \cdots & k & k+1 & k+2 & \cdots & k+l & \cdots \\ a_1 & a_2 & \cdots & a_k & b_1 & b_2 & \cdots & b_l & \cdots \end{pmatrix} \end{aligned}$$

is bijective. Ψ is called the Foata *transformation.*

Exercise 10 $(2 + 2 + 2$ points)**.** How many possibilities are there to place n different figures on an $n \times n$ chessboard such that

(a) there is one figure in each horizontal row?

(b) there is one figure in each horizontal and vertical row?

(c) there are no restrictions?

Exercise 11 $(4 + 2 + 2 + 2$ points)**.**

(a) Calculate $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$, $\left\{ {7 \atop 4} \right\}$, $p(7)$ and $b(6)$.

(b) Show

$$\begin{bmatrix} n+1 \\ 2 \end{bmatrix} = n! H_n$$

for $n \in \mathbb{N}$, where H_n is the n-th harmonic number.

(c) Show

$$\begin{bmatrix} n \\ n-2 \end{bmatrix} = 2\binom{n}{3} + 3\binom{n}{4}$$

for all $n \geq 2$.
Hint: Consider which cycle types are relevant and apply Theorem 2.26.

(d) Find and prove an analogous formula for $\left\{ {n \atop n-2} \right\}$.

Exercise 12 $(2+1$ points$)$**.** Show that

$$x(x+1)\dots(x+n-1) = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k$$

for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$. Conclude:

$$\sum_{k=1}^{n} (-1)^k \begin{bmatrix} n \\ k \end{bmatrix} = 0.$$

Exercise 13 $(1+1$ points$)$**.** How many summands do the two formulas

$$\begin{bmatrix} n \\ k \end{bmatrix} = \sum_{0<a_1<\ldots<a_{n-k}<n} a_1 \dots a_{n-k},$$
$$\left\{ {n \atop k} \right\} = \sum_{1\le a_1\le\ldots\le a_{n-k}\le k} a_1 \dots a_{n-k}$$

from Theorems 2.17 and 2.32 have?

Exercise 14 (2 points)**.** Show that S_{2n} has exactly $1 \cdot 3 \cdot \ldots \cdot (2n-1)$ fixed-point-free permutations of order 2.

Exercise 15 (2 points)**.** Show that among six people there are always three who all know each other or all do not know each other.

Exercise 16 (2 points)**.** Show that there exist two prime numbers whose difference is divisible by 2019. *Hint:* Using the pigeonhole principle, one does not need to specify the prime numbers explicitly. *Addition:* Are there two prime numbers whose sum is divisible by 2019?

Exercise 17 (2 points)**.** For every $n \in \mathbb{N}$ there exists an $m \in \mathbb{N}$ such that mn consists only of the digits 0 and 1.

Exercise 18 $(2+2$ points$)$**.** For partitions $\lambda = (\lambda_1, \dots, \lambda_s)$ and $\mu = (\mu_1, \dots, \mu_t)$ of $n \in \mathbb{N}$ we write $\lambda \le \mu$ if λ is a *refinement* of μ, i.e., with suitable numbering $\mu_1 = \lambda_1 + \ldots + \lambda_{i_1}$, $\mu_2 = \lambda_{i_1+1} + \ldots + \lambda_{i_2}$ etc. Show that $(P(n), \le)$ is a locally finite ordered set. Calculate the corresponding Möbius function $\mu_{P(4)}((1,1,1,1),(4))$.

Exercise 19. Let V be a finite-dimensional vector space over a field with $q < \infty$ elements. The set $\mathcal{U}$ of subspaces of V is ordered by $\subseteq$. Show that

$$\mu_{\mathcal{U}}(U,W) = (-1)^{\dim(W/U)} q^{\binom{\dim(W/U)}{2}}$$

for $U \le W \le V$.

Exercise 20. Let $a_0, a_1 \ldots, b_0, b_1, \ldots \in \mathbb{C}$. Show:

(a) (Binomial inversion)

$$\forall n \in \mathbb{N}_0 : a_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} b_k \iff \forall n \in \mathbb{N}_0 : b_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} a_k$$

(b) (Stirling inversion)

$$\forall n \in \mathbb{N}_0 : a_n = \sum_{k=0}^{n} (-1)^k \begin{bmatrix} n \\ k \end{bmatrix} b_k \iff \forall n \in \mathbb{N}_0 : b_n = \sum_{k=0}^{n} (-1)^k \begin{Bmatrix} n \\ k \end{Bmatrix} a_k$$

Exercise 21 (2 + 2 points)**.**

(a) Calculate the coefficients of $(1 + X + X^2)^{-1} \in \mathbb{Q}[[X]]$.

(b) Calculate the first six coefficients of the inverse function of $X - X^2 \in \mathbb{Q}[[X]]$.
Hint: Set up the inverse function in the form $\sum a_n X^n$ and establish equations for $a_1, a_2, \ldots$.

Exercise 22 (2 points)**.** For every prime number p, the power series of the form $X + a_2 X^2 + \ldots \in \mathbb{F}_p[[X]]$ form a subgroup of $\mathbb{F}_p[[X]]^\circ$, which is called the *Nottingham group* N_p. Construct an element of order p in N_p.

Exercise 23 (3 points)**.** Construct for every field K suitable power series $\alpha, \beta, \gamma \in K[[X]]$, such that $\alpha \circ \beta \neq \beta \circ \alpha$, $\alpha \circ (\beta + \gamma) \neq \alpha \circ \beta + \alpha \circ \gamma$ and $\alpha \circ (\beta\gamma) \neq (\alpha \circ \beta)(\alpha \circ \gamma)$

Exercise 24. We consider the following power series in $\mathbb{C}[[X]]$:

$$\sin(X) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} X^{2n+1}, \qquad \cos(X) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} X^{2n},$$

$$\tan(X) := \frac{\sin(X)}{\cos(X)}, \qquad \arctan(X) := \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} X^{2k+1},$$

$$\arcsin(X) := \sum_{n=0}^{\infty} \frac{(2n)!}{(2^n n!)^2} \frac{X^{2n+1}}{2n+1}.$$

Show:

(a) (Euler's formula) $\exp(\mathrm{i}X) = \cos(X) + \mathrm{i}\sin(X)$, where $\mathrm{i} = \sqrt{-1} \in \mathbb{C}$.

(b) $\sin(2X) = 2\sin(X)\cos(X)$ and $\cos(2X) = \cos(X)^2 - \sin(X)^2$.

(c) ("trigonometric Pythagoras") $\cos(X)^2 + \sin(X)^2 = 1$.

(d) $\sin(X)' = \cos(X)$ and $\cos(X)' = -\sin(X)$.

(e) $\arctan \circ \tan = X$.
Hint: First check $\tan \in \mathbb{C}[[X]]^\circ$. Then differentiate.

(f) $\arctan(X) = \frac{\mathrm{i}}{2} \log\left(\frac{\mathrm{i}+X}{\mathrm{i}-X}\right)$ with $\mathrm{i} = \sqrt{-1} \in \mathbb{C}$.
Hint: Check that $\log\left(\frac{\mathrm{i}+X}{\mathrm{i}-X}\right)$ is well-defined.

(g) $\arcsin(X)' = \frac{1}{\sqrt{1-X^2}}$.
Hint: Newton's binomial theorem.

(h) $\arcsin \circ \sin = X$.

Remark: There are countless other trigonometric identities, but not all can be proven via formal power series. For example, cos has no formal inverse function (nevertheless, one can give the analytical Taylor series for arccos).

Exercise 25 (3 points)**.** Let K be a field. Verify that

$$K((X)) := \Big\{ \sum_{n=k}^{\infty} a_n X^n : k \in \mathbb{Z},\, a_n \in K \Big\}$$

with the operations

$$\sum a_n X^n + \sum b_n X^n = \sum (a_n + b_n) X^n,$$
$$\sum a_n X^n \cdot \sum b_n X^n = \sum_{n=-\infty}^{\infty} \Big(\sum_{k=-\infty}^{\infty} a_k b_{n-k} \Big) X^n$$

is a field.
Hint: Do not forget to check the well-definedness of addition and multiplication.

Exercise 26 (2 points)**.** Find and prove an explicit formula for the recursively defined sequence $a_0 := 1$, $a_1 := 2$, $a_{n+1} := 3a_n - a_{n-1}$ for $n \in \mathbb{N}$.

Exercise 27 (2 points)**.** Determine all invertible 2×2-matrices over the field $\mathbb{F}_2$.

Exercise 28 ($2 + 3 + 2$ points)**.** Let $n \in \mathbb{N}$. Prove:

(a) Let $p_1(n)$ be the number of partitions of n with even parts. Let $p_2(n)$ be the number of partitions of n whose parts have even multiplicity. Show that $p_1(n) = p_2(n)$ for all $n \in \mathbb{N}$.

(b) Let $p_+(n)$ (resp. $p_-(n)$) be the number of partitions of n with an even (resp. odd) number of parts. Show that $(-1)^n(p_+(n) - p_-(n))$ is the number of symmetric partitions of n.

(c) Let $q_+(n)$ (resp. $q_-(n)$) be the number of partitions of n with an even (resp. odd) number of even parts. Show that $q_+(n) - q_-(n)$ is the number of symmetric partitions of n.

Exercise 29 ($3 + 3 + 3$ points)**.** Let $n, d \in \mathbb{N}$. Prove:

(a) (Glaisher) The number of partitions of n whose parts are not divisible by d is equal to the number of partitions of n in which no part occurs d times (or more).
Hint: The case $d = 2$ corresponds to Theorem 5.7(i).

(b) (Subbarao) The number of partitions of n where each part occurs exactly twice, three times, or five times is equal to the number of partitions of n into parts of the form $\pm 2 + 12k$, $\pm 3 + 12k$, or $6 + 12k$.

(c) (MacMahon) The number of partitions of n where each part occurs at least twice is equal to the number of partitions of n into parts that do not have the form $\pm 1 + 6k$.

Exercise 30 (2 + 2 points)**.**

(a) Determine all irreducible polynomials in $\mathbb{F}_3[X]$ of degree ≤ 3.

(b) Calculate the cyclotomic polynomial Φ_{396}.

Exercise 31 (3 points)**.** Let $n \in \mathbb{N}$. How many pairwise distinct numbers $a, b, c \in \{1, \ldots, 2n\}$ are there such that a is the arithmetic mean of b and c?

Christmas puzzle (+3 bonus points)**.** From a bag with r red and b blue beads, two are drawn at random. If both beads have the same color, they are removed and a new red bead is put into the bag instead. If both beads have different colors, the red one is removed and the blue one is put back into the bag. If this process is repeated often enough, exactly one bead remains in the bag at the end. How can the color of this remaining bead be determined from r and b?
Note: It has nothing to do with probabilities.

Exercise 32 (2 + 2 + 2 points)**.** Let Z be the set of all rational polynomials with integer values, i.e.,

$$Z := \{\alpha \in \mathbb{Q}[X] : \alpha(z) \in \mathbb{Z} \ \forall z \in \mathbb{Z}\}.$$

Show:

(a) For $\alpha, \beta \in Z$, it holds that $\alpha + \beta, \alpha\beta \in Z$.

(b) The polynomials $\binom{X}{k}$ defined in Example 6.28 lie in Z for $k \in \mathbb{N}_0$.

(c) Every $\alpha \in Z$ can be uniquely written in the form $\alpha = \sum_{k=0}^{\infty} a_k \binom{X}{k}$ with $a_k \in \mathbb{Z}$.

Exercise 33 (3+3 points)**.** For $\alpha = \sum a_{k_1,\ldots,k_n} X_1^{k_1} \ldots X_n^{k_n} \in K[X_1, \ldots, X_n]$, we define the $(l_1, \ldots, l_n)$-th Hasse derivative by

$$H^{(l_1,\ldots,l_n)}(\alpha) = \sum a_{k_1,\ldots,k_n} \binom{k_1}{l_1} \ldots \binom{k_n}{l_n} X_1^{k_1-l_1} \ldots X_n^{k_n-l_n} \in K[X_1, \ldots, X_n].$$

Let the *multiplicity* of $x = (x_1, \ldots, x_n) \in K^n$ as a root of α be the smallest number $m_\alpha(x) \in \mathbb{N}_0 \cup \{\infty\}$ with $H^{(l_1,\ldots,l_n)}(x) = 0$ for all $(l_1, \ldots, l_n)$ with $l_1 + \ldots + l_n < m_\alpha(x)$ (in the case $m_\alpha(x) = 0$, x is not a root of α). Show:

(a) For $\alpha \neq 0$, it holds that $\sum_{x \in K^n} m_\alpha(x) \leq \deg(\alpha)|K|^{n-1}$.

(b) Let $c\colon K^n \to \mathbb{N}_0$ and $d \in \mathbb{N}_0$ with $\sum_{x \in K^n} \binom{c(x)+n-1}{n} < \binom{d+n}{n}$. Then there exists an $\alpha \in K[X_1, \ldots, X_n] \setminus \{0\}$ with $\deg(\alpha) \leq d$ and $m_\alpha(x) \geq c(x)$ for all $x \in K^n$.

Exercise 34 (2 + 2 points)**.** Prove the "inverse" Waring formulas

$$\tau_k = \sum_{(1^{a_1},\ldots,k^{a_k}) \in P(k)} \prod_{i=1}^{k} \frac{\rho_i^{a_i}}{i^{a_i} a_i!},$$

$$\sigma_k = (-1)^k \sum_{(1^{a_1},\ldots,k^{a_k}) \in P(k)} \prod_{i=1}^{k} \frac{(-\rho_i)^{a_i}}{i^{a_i} a_i!}.$$

for $K = \mathbb{C}$ and $k \in \mathbb{N}$.
Note: Proof of Theorem 5.12.

Exercise 35 (2 + 2 points)**.**

(a) Calculate the Bernoulli numbers $B_5, \ldots, B_{10}$.

(b) For which $n < 100$ does the n-th Bernoulli number have the form $B_n = \frac{z}{6}$ with $z \in \mathbb{Z}$?

Exercise 36 (2 points)**.** Apparently, the set $M := \mathbb{Q} \setminus \{0\}$ is a magma w.r.t. division. How many results are obtained if one brackets $1 : 2 : 3 : 4$ in all possible ways?

Exercise 37 (2 + 2 points)**.**

(a) Which patterns does the permutation $(2, 4, 5, 6, 1, 3) \in S_6$ possess (or avoid)?

(b) Let $n \geq 4$. Which permutations in S_n possess only one pattern?

Exercise 38 (2 + 2 points)**.** Let $n, k \in \mathbb{N}$. Show that

$$\left|\left\{(n_1, \ldots, n_k) \in \mathbb{N}^k : \sum_{i=1}^{k} n_i = n\right\}\right| = \binom{n-1}{k-1},$$

$$\left|\left\{(n_1, \ldots, n_k) \in \mathbb{N}_0^k : \sum_{i=1}^{k} n_i = n\right\}\right| = \left(\binom{n+1}{k-1}\right) = \left(\binom{k}{n}\right).$$

Hint: Consider the map $(n_1, \ldots, n_k) \mapsto (n_1, n_1 + n_2, \ldots, n_1 + \ldots + n_{k-1})$.

Exercise 39 (2 + 2 points)**.** Let $n \in \mathbb{N}$ and $A \subseteq \{1, \ldots, 2n\}$ with $|A| = n + 1$. Show:

(a) A contains two coprime numbers.

(b) There exist $a, b \in A$ with $a \neq b$ and $a \mid b$.

Exercise 40 (2 points)**.** Show that $\sum_{n=1}^{\infty} \frac{C_n}{4^n} = 1$.

Exercise 41 (2 points)**.** Let $H \leq G$ be finite groups. Prove or disprove $k(H) \leq k(G)$.

Exercise 42 (2 points)**.** Determine the automorphism group of the graph $\mathcal{V}_2 \sqcup \mathcal{V}_2$.

Exercise 43 (2 + 2 + 2 points)**.** Determine the trees with Prüfer code $(1, 2, 2, 3, 3, 3)$, $(1, 2, \ldots, n)$ and $(1, 1, \ldots, 1)$.

Exercise 44 (2 points)**.** Construct all trees of order 5 up to isomorphism.

Exercise 45 (2 + 2 + 2 points)**.** Let K be a field. For $\alpha = \sum a_n X^n \in K[[X]]$ let

$$m(\alpha) := \inf\{n \in \mathbb{N}_0 : a_n \neq 0\},$$

where $m(0) = \inf \varnothing = \infty$. Show:

(a) For $\alpha, \beta \in K[[X]]$, $m(\alpha\beta) = m(\alpha) + m(\beta)$ and $m(\alpha + \beta) \geq m(\alpha) + m(\beta)$ hold.

(b) For $\alpha, \beta \in K[[X]]$ with $\beta \neq 0$, there exist uniquely determined $\delta, \gamma \in K[[X]]$ with $\alpha = \beta\gamma + \delta$ and ($\gamma = 0$ or $m(\delta) > m(\beta)$).

(c) For $\alpha, \beta \in K[[X]]$, there exists a $\gamma \in K[[X]]$ with $\alpha = \beta\gamma$ if and only if $m(\alpha) \geq m(\beta)$ holds (this generalizes Lemma 4.8).

Exercise 46 (3 + 3 + 3 points)**.** Prove for all $\alpha \in \mathbb{C}[[X]]$:

(a)
$$\prod_{k=0}^{\infty}(1 + \alpha X^k) = \sum_{k=0}^{\infty} \frac{\alpha^k X^{\binom{k}{2}}}{X^k!}.$$

(b) For $n \in \mathbb{N}$,
$$\prod_{k=1}^{n} \frac{1}{1 - \alpha X^k} = \sum_{k=0}^{\infty} \alpha^k \left\langle {n+k-1 \atop k} \right\rangle X^k$$
holds. *Hint:* Interpret the coefficient of $\alpha^k X^l$.

(c)
$$\prod_{k=1}^{\infty} \frac{1}{1 - \alpha X^k} = \sum_{k=0}^{\infty} \frac{\alpha^k X^k}{X^k!}.$$

Exercise 47 (3 points)**.** Let $\alpha = \sum a_n X^n \in \mathbb{C}[[X]]$, $k \in \mathbb{N}$ and $\zeta_k = e^{2\pi \mathrm{i}/k}$. Show that
$$\frac{1}{k} \sum_{l=1}^{k} \alpha(\zeta_k^l X) = \sum_{n=0}^{\infty} a_{kn} X^{kn}.$$

A. GAP commands

Most combinatorial objects can be calculated in the free computer algebra system GAP[9]:

Object	Code
$n!$	`Factorial(n);`
2^A	`Combinations(A);`
Variations with repetition	`Tuples([1,2,3],2);`
Variations without repetition	`Arrangements([1,2,3],2);`
Combinations with repetition	`UnorderedTuples([1,2,3],2);`
Combinations without repetition	`Combinations([1,2,3],2);`
S_n	`SymmetricGroup(n);`
$\binom{n}{k}$	`Binomial(n,k);`
$\left\langle {n \atop k} \right\rangle$	`GaussianCoefficient(n,k,X(Integers,"q"));`
$\left(\binom{n}{k}\right)$	`NrUnorderedTuples(n,k);`
$\left[{n \atop k} \right]$	`Stirling1(n,k);`
$\left\{ {n \atop k} \right\}$	`Stirling2(n,k);`
$P(n)$	`Partitions(n);`
Partitions into odd parts	`RestrictedPartitions(n,[1,3..2*n+1]);`
Partitions into distinct parts	`RestrictedPartitionsWithoutRepetitions(n,[1..n]);`
Partitions into k parts	`Filtered(Partitions(n),p->Size(p)=k);`
Symmetric partitions	`Filtered(Partitions(n),p->AssociatedPartition(p)=p);`
$p(n)$	`NrPartitions(n);`
$P(A)$	`PartitionsSet(A);`
$b(n)$	`Bell(n);`
$\varphi(n)$	`Phi(n);`
Φ_n	`CyclotomicPolynomial(Rationals,n);`
Factorization in $K[X]$	`x:=X(Rationals,"X");; Factors(x^4+3*x^3-7*x+1);`
B_n	`Bernoulli(n);`
f_n	`Fibonacci(n);`

[9] `https://www.gap-system.org/`

Index

Symbols

$\varnothing$, 3
2^A, 4
$\left\langle {n \atop k} \right\rangle$, 33
$|A|$, 4
A^I, 4
$\binom{A}{k}$, 4
$\binom{a}{k}$, 5
$a \leq b$, 20
α^{-1}, 25
$|\alpha|$, 27
α', 30
$\alpha(\beta)$, 28
$\alpha \circ \beta$, 28
$\alpha^{(n)}$, 30
A^n, 4
$\mathrm{Aut}(\Omega)$, 88
B_n, 66
$b(n)$, 15
$\mathbb{C}$, 4
$\mathrm{C}_G(x)$, 83
C_n, 71, 76
D_{2n}, 75
$\deg(\alpha)$, 49, 58
$\exp(X)$, 25
$\mathbb{F}_q$, 50
$\mathrm{GL}(V)$, 74
$\mathcal{G}_n$, 88
$g(n)$, 82
$\mathfrak{g}(n)$, 87
$H^k(\alpha)$, 31
$H^{(l_1,\ldots,l_n)}(\alpha)$, 98
H_n, 13
$I_d(K)$, 52
$k(G)$, 83
$k(S_n)$, 85
$K[X]$, 49
$K[[X]]$, 24
$K[[X]]^\times$, 26
$K((X))$, 97
$K[[X]]^\circ$, 30
$\mathcal{K}_n$, 88
λ', 36
$\log(1+X)$, 32
$m(\alpha)$, 99
$\mu_A(a,b)$, 21
$n!$, 4
$\mathrm{N}_G(H)$, 85
$\mathbb{N}$, 3
$\mathbb{N}_0$, 3
$\binom{n}{k_1,\ldots,k_s}$, 5
$\Omega \sqcup \Delta$, 86
$\Omega \cong \Delta$, 86
$\mathrm{O}(\mathbb{R}^n)$, 75
Ω^C, 86
$\mathbb{P}$, 3
$p_\pm(n)$, 97
$P(A)$, 15
$\varphi(n)$, 9
$p_{k,l}(n)$, 38
$p_k(n)$, 37
$p(n)$, *15*, 20, 40
$\mathbb{Q}$, 3
$q_\pm(n)$, 97
$\mathbb{R}$, 3
ρ_k, 61
$\mathrm{SO}(\mathbb{R}^n)$, 76
σ_k, 61
S_n, 6
$\mathcal{S}_n$, 89
$\begin{bmatrix} n \\ k \end{bmatrix}$, 14
$\left\{ {n \atop k} \right\}$, 17
$\mathrm{Sym}(A)$, 6
τ_k, 61
$\mathcal{T}_n$, 86
V_4, 82
$\mathcal{V}_n$, 86
$\bigtimes_{i \in I} A_i$, 4
$\langle x_1, \ldots, x_n \rangle$, 76
$\binom{X}{k}$, *54*, 98
$\mathbb{Z}$, 3
ζ_n, 54

A

action, 76
    transitive, 76
anagram, 8
Apéry-Konstante, 68
asymmetric, 88
automorphism group, 88

B

Bell number, 15
Bernoulli number, 66
binary operation, 71
Binet formula, 35
binomial coefficient, 5
binomial inversion, 96
binomial theorem, 7
Birthday Paradox, 6
Burnside's Lemma, 78

C

Cartesian product, 4
Catalan, 72
Catalan number, 71

Cayley formula, 90
centralizer, 83
chain rule, 31
Chevalley-Warning, 60
class equation, 83
class number, 83
Clausen, 69
combination
    with replacement, 8
    without repetition, 6
combination lock, 5
Combinatorial Nullstellensatz, 60
component, 86
congruence, 53
conjugacy class, 83
conjugation, 83
connected, 86
connected component, 86
constant multiple rule, 32
constant term, 25
coprime, 9
Coupon collector's problem, 19
cube, 79
cycle, 11, 88
    disjoint, 12
cycle type, 17
cyclotomic polynomial, 54

D
degree, 58
    polynomial, 49
    vertex, 86
derivative, 30
    n-th, 30
dice, 8
Dihedral group, 75
disjoint union, 4
division with remainder, 51, 53
Dobiński formula, 20
Durfee's Square Theorem, 40

E
edge, 86
Enigma, 11
equivalence class, 16
equivalence relation, 16
Erdős-Ginzburg-Ziv, 61
Erdős-Rényi, 89
Erdős-Szekeres, 10
Erdős-Turán, 38, 84
Euler, 36, 72
Euler φ-function, 9
Euler's formula, 96
Euler's Pentagonal Number Theorem, 39
Euler-Mascheroni constant, 13
exponential function, 25

F
factorial, 4
Faulhaber's formula, 67
Fermat's little theorem, 53
Fibonacci sequence, 35
fixed point, 10
fixed-point-free, 10
Foata transformation, 94
Franklin, 39
functional equation
    for $\exp(X)$, 29
    for $\log(1+X)$, 32
Fundamental theorem on symmetric polynomials, 63

G
Gauss, 52
Gauss's Binomial Theorem, 34
Gaussian binomial coefficient, 33
generating function, 34
    for $b(n)$, 36
    for $p(n)$, 36
generating system, 76
geometric series, 26
Girard-Newton identity, 62
Glaisher, 97
graph, 86
    asymmetric, 88
    complementary, 86
    complete, 86
    connected, 86
    symmetric, 88
    trivial, 86
    union, 86
group, 74
    abelian, 74
    cyclic, 76
    finite, 74
    general linear, 74
    orthogonal, 75
    special orthogonal, 76
    symmetric, 6
    trivial, 74

H
Hall, 23
Hardy-Ramanujan, 20
Hasse derivative, 31, 98
Hilbert's Nullstellensatz, 60
Hirschhorn, 41
hockey-stick identity, 93

I
Inclusion-Exclusion Principle, 9
indeterminate, 25
interpolation, 58
inverse function, 30
irreducible, 51

isomorphism
of graphs, 86
of groups, 82
isomorphism class, 82

J
Jacobi, 41
Jacobi's triple product, 43

K
Klein four-group, 82

L
Lagrange, 77
Lagrange-Jacobi, 45
Landau, 84
Laurent series, 25
leading coefficient, 49
leaf, 90
left coset, 77
left-to-right minimum, 73
length
of a cycle, 11
of an orbit, 76
locally finite, 21
logarithm, 32
lottery, 6
Lucas, 70

M
MacMahon, 73, 97
MacMahon cube, 80
magma, 71
Mercator series, 32
modulo, 53
Montmort, 10
multinomial coefficient, 5
multinomial theorem, *8*, 94
multiset, 8
Möbius function, 21
classical, 23
Möbius inversion, 21
multiplicative, 23

N
necklace, 79, 81
Newton's Binomial Theorem, 33
Nicomachus identity, 65
norm, 27
normalizer, 85
Nottingham group, 96
number
complex, 4
harmonic, 13
integer, 3
k-cycles, 13
natural, 3
of partitions, 15
partitions into distinct parts, 37
partitions with fixed type, 16
partitions with largest part k, 37
permutations with cycle type, 17
prime, 3
rational, 3
real, 3
symmetric partitions, 37

O
Ono, 43
orbit, 76
orbit equation, 78
Orbit-Stabilizer Theorem, 77
order
element, 75
graph, 86
group, 74
order relation, 20
locally finite, 21

P
partial fraction decomposition, 27
partition
conjugate, 36
of a number, 15
of a set, 15
parts, 15
symmetric, 36
Pascal's triangle, 5
path, 86, 88
pattern, 72
permutation, 6
fixed-point-free, 10
pigeonhole principle, 6
Pólya, 81
polynomial, 49
constant, 50
in several variables, 58
monic, 49
reducible/irreducible, 51
power series, 25
inverse, 25
invertible, 25
norm, 27
root, 32
power set, 4
Prime factorization in $K[X]$, 52
primitive root, 60
prisoners problem, 14
product rule, 31
Prüfer, 90
Prüfer code, 91

Q
quotient rule, 31

R
Ramanujan, 42
reducible, 51
relation, 16
    antisymmetric, 20
    reflexive, 16
    symmetric, 16
    transitive, 16
Rogers-Ramanujan identities, 47
root, 50
root of unity, 54
    primitive, 54
Rubik's Cube, 80

S
Schwartz-Zippel, 58
Secretary problem, 11
Segner, 71
set
    equinumerous, 4
    finite/infinite, 4
    ordered, 20
Simion-Schmidt, 73
skat cards, 8
stabilizer, 76
star, 89
Stirling formula, 19
Stirling inversion, 96
Stirling number
    of the first kind, 14
    of the second kind, 17
Subbarao, 97
subgroup, 75
    generated, 75
Sudoku, 80
sum rule, 31
Symmetry group, 75

T
Taylor series, 31
transposition, 12
tree, 90
trigonometric Pythagoras, 96

U
unit group, 26

V
Vandermonde identity, 7, 54
variation
    with repetition, 5, 7
    without repetition, 5
vertex, 86
    adjacent, 86
    connected, 86
    degree, 86
Vieta, 62
von Staudt, 69

W
Waring formula, 64
    inverse, 98
Waring problem, 47
Wright, 43

Y
Young diagram, 36