Discrete mathematics

Contents

1. Counting problems

To read:
[1]: 1.2. Sets, 1.3. Number of subsets, 1.5. Sequences, 1.6. Permutations, 1.7. Number of The Number of Ordered Subsets, 1.8. The Number of Subsets of a Given Size, 3.1. The Binomial Theorem, 3.2. Distributing Presents, 3.5. Pascal's Triangle, 3.6. Identities in Pascal's Triangle. [3], Chapters 3.1-3.3.

1.1. **Basic results on counting sets.**

Notation. Let A be a finite set. We denote by $|A|$ the *cardinality* of A, i. e. the number of elements in the set.

Definition 1.1. Denote by $[n]$ the set of first n natural numbers: $[n] := \{1, 2, \dots n\}$.

Theorem 1.2. *If there exists a bijection between finite sets A and B then $|A| = |B|$.*

Theorem 1.3. *(Addition rule) Let A and B be finite sets. If $A \cap B = \emptyset$ then $|A \cup B| = |A| + |B|$.*

Theorem 1.4. *(Product rule) Let A and B be finite sets. Then*

$$|A \times B| = |A| \cdot |B|.$$

Recall the following formulas:

Proposition 1.5. *The number of functions from $[m]$ to $[n]$ is n^m. This is the number of m-letter words in an n-letter alphabet.*

Proposition 1.6. *The number of permutations of a set of n elements is n!*

Proof. This is likely to be familiar to you, but at any rate it follows from the multiplication rule. Call the elements $1, \dots, n$. A permutation can send 1 to any of n elements. Then 2 to any of the $n-1$ elements remaining, since 1 and 2 cannot be sent to the same. Each step leaves one less option at the next step, for a total of

$$n \times (n-1) \times \dots \times 2 \times 1$$

permutations. This is $n!$ by definition (or really, if we refuse to skip steps, by induction). □

Proposition 1.7. *The number of ways in which one can choose k objects out of n distinct objects, assuming the order of the elements matters, is $\frac{n!}{(n-k)!}$.*

Proof. It will dramatically speed up computations to note that

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

This should be calculated as a product of k numbers, not a ratio of two factorials. In fact, this form also shows how to deduce the formula from the multiplication rule. One has n choices for the first object, then $n-1$ for the second, culminating in $n-k+1$ for the last of the k objects.

Notice that when $k = n$, Propositions 1.6 and 1.7 agree. This would be clear even without the explicit formulae: an ordered choice of all n out of the n objects is simply a way to permute them.

Set-theoretically, $n(n-1)\cdots(n-k+1)$ is also the number of injective functions from $[k]$ to $[n]$. □

Proposition 1.8. *The number of ways in which one can choose k objects out of n distinct objects, assuming the order of the elements does not matter, is $\frac{n!}{(n-k)!k!} =: \binom{n}{k}$. This is the same as the number of subsets of k elements of an n-element set.*

Definition 1.9. The numbers $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ are called *binomial coefficients.*

Proof. We already know the number of ordered subsets, by Proposition 1.7. On the other hand, an ordered subset can be obtained in two steps: choose a subset, and then order it. Once the choice of k elements is made, Proposition 1.6 tells us there are $k!$ ways to do the ordering. By the multiplication rule,

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

and we complete the proof by solving for $\binom{n}{k}$.

□

As with unordered choices, there is no need to compute all the factorials. Instead, note that

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

If k is small, then we can afford to compute $k!$ in the denominator. If k is large, then it is better to exploit a basic symmetry of the binomial coefficients.

Proposition 1.10.

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

We will be convenient for us to use the following notation:

Notation. Let A be a finite set and k be a nonnegative integer. Then $\binom{A}{k}$ is the set of k-element subsets of A. We have $\left|\binom{A}{k}\right| = \binom{|A|}{k}$.

1.2. **Binomial coefficients.** The following is called Pascal's triangle

Row						
0	$\binom{0}{0} = 1$					
1	$\binom{1}{0} = 1$	$\binom{1}{1} = 1$				
2	$\binom{2}{0} = 1$	$\binom{2}{1} = 2$	$\binom{2}{2} = 1$			
3	$\binom{3}{0} = 1$	$\binom{3}{1} = 3$	$\binom{3}{2} = 3$	$\binom{3}{3} = 1$		
4	$\binom{4}{0} = 1$	$\binom{4}{1} = 4$	$\binom{4}{2} = 6$	$\binom{4}{3} = 4$	$\binom{4}{4} = 1$	
5	$\binom{5}{0} = 1$	$\binom{5}{1} = 5$	$\binom{5}{2} = 10$	$\binom{5}{3} = 10$	$\binom{5}{4} = 5$	$\binom{5}{5} = 1$

Proposition 1.11. *The following identities hold:*

(1) $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$.
(2) $\binom{n}{k}$ *is the k-th element in the n-th line of Pascal's triangle.*

Proof. Recall that $\binom{n+1}{k+1}$ is the number of subsets of cardinality $k+1$ in the set $[n+1]$. Each subset of $[n+1]$ either contains the element $n+1$ or not. The number of elements in $\binom{[n+1]}{k+1}$ containing $n+1$ is $\binom{n}{k}$ and the number of elements in $\binom{[n+1]}{k+1}$ *not* containing $n+1$ is $\binom{n}{k+1}$. Now we apply the Addition rule and finish the proof. □

Proposition 1.12. *The number of subsets of an n-element set is 2^n, since we have*

$$2^n = \binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n}.$$

The number of subsets of an n-element set having odd cardinality is 2^{n-1}. The number of subsets of an n-element set having even cardinality is 2^{n-1}.

The equalities above can be obtained using the binomial theorem.

Theorem 1.13.

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \ldots + \binom{n}{n}x^n = \sum_{i=0}^{n} \binom{n}{i} x^i.$$

Proof. To prove the binomial theorem, consider how to distribute the multiplication in

$$(1+x)^n = (1+x)(1+x)\ldots(1+x)$$

From each factor $1+x$, we can choose either the 1 or the x to form a product with the other terms. This product is x^k provided we choose x in k out of the n factors. There are $\binom{n}{k}$ such choices, and collecting terms gives the sum $\sum_k \binom{n}{k} x^k$ as claimed. □

Proof of Proposition 1.12. For $x = 1$, respectively $x = -1$, we obtain

$$\begin{aligned} 2^n =& \binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = \sum_{i=0}^{n} \binom{n}{i} \\ 0 =& \binom{n}{0} - \binom{n}{1} + \ldots + (-1)^n \binom{n}{n} = \sum_{i=0}^{n} (-1)^i \binom{n}{i}. \end{aligned}$$

Adding, respectively subtracting the two relations, and dividing each by two, one obtains

$$\begin{aligned} 2^{n-1} =& \binom{n}{0} + \binom{n}{2} + \ldots \\ 2^{n-1} =& \binom{n}{1} + \binom{n}{3} + \ldots \end{aligned}$$

which proves the statements about the number of even/odd sets.
□

Proposition 1.14. *Assume we have k identical objects and n different persons. Then, the number of ways in which one can distribute this k objects among the n persons equals*

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

Equivalently, it is a number of solutions of the equation $x_1 + \ldots + x_n = k$ in nonnegative integers or the number of k-multisets containing elements from $[n]$. If $k \geq n$ and each persons receives at least 1 object, then the number of possible ways to distribute is $\binom{k-1}{n-1}$.

Proof. Let $\mathcal{A}$ be the set of all solutions of the equation

(1) $$x_1 + \ldots + x_n = k, x_i \in \mathbb{Z}_{\geq 0}.$$

Let $\mathcal{B}$ be the set of all subsets of cardinality $n-1$ in $[k+n-1]$. We construct a bijection $\psi : \mathcal{A} \to \mathcal{B}$ in the following way: a solution $(x_1, \ldots, x_n)$ is mapped to the subset

$$B := \{x_1 + 1, x_1 + x_2 + 2, \ldots, x_1 + x_2 + \ldots + x_{n-1} + n - 1\}.$$

First, we check that B belongs to $\mathcal{B}$. Indeed, the inequalities

$$1 \leq x_1 + 1 < x_1 + x_2 + 2 < \cdots < x_1 + x_2 + \ldots x_{n-1} + n - 1 \leq k + n - 1$$

imply that the elements of B are distinct and belong to $[k+n-1]$.

Next, to show that ψ is a bijection we compute its inverse map. Let B be an element of $\mathcal{B}$. Suppose that

$$1 \leq b_1 < b_2 < \cdots < b_{n-1} \leq k + n - 1$$

are the elements of B written in the increasing order. Then the preimage $\psi^{-1}(B)$ is an n-tuple of integers $(x_1, \ldots, x_n)$ defined by

$$\begin{aligned} &x_1 = b_1 - 1 \\ &x_i = b_i - b_{i-1} - 1, \quad i = 2, \ldots, n-1 \\ &x_n = k + n - 1 - b_{n-1}. \end{aligned}$$

It is easy to see from these equations that the numbers $x_i, i = 1, \ldots n$, are non-negative integers and $x_1 + \ldots + x_n = k$.

Since there is a bijection between sets $\mathcal{A}$ and $\mathcal{B}$, their cardinalities are equal and

$$|\mathcal{A}| = |\mathcal{B}| = \binom{k+n-1}{n-1}.$$

□

2. Estimates: O, o-notation, Stirling formula, Birthday paradox and the bell curve

To read:
[1] 2.2.4. Pigeonhole principle. 2.2.5 The Twin Paradox
[3] 3.4. Estimates: an introduction - starting from 3.4.2. - Big Oh, little oh, 3.5.5. Estimate $n!$ - second proof only, 3.7. Inclusion - Exclusion.

2.1. O, o-notation.

Definition 2.1. Let $f, g : \mathbb{Z}_{\geq 0} \to \mathbb{R}$. We say that f is *big-Oh* of g and we write $f(x) = O(g(x))$ if there exist n_0 and c constants such that for all $n > n_0$, we have $|f(n)| < c \cdot |g(n)|$.

Definition 2.2. Let $f, g : \mathbb{Z}_{\geq 0} \to \mathbb{R}$. We say that f is *little-oh* of g and we write $f(x) = o(g(x))$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0.$$

Examples: $n = O(n^2)$ and also $n = o(n^2)$, $n = O(2^n)$, $n = o(2^n)$, $\sin(n) = O(1)$ and $\sin(n)$ is not $o(1)$.

2.2. Stirling's formula.

Theorem 2.3. *(Stirling's formula)*

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

where $\sim$ is used to indicate that the ratio of the two sides tends to 1 as n goes to ∞.

2.3. **Twin paradox.** Suppose that there are 50 students in a math class. What are the chances that two of them share the same birthday?

Theorem 2.4. *Suppose that $k \leq n$ are positive integers and each of k different people chooses 1 element from the set $[n]$. Their choices are uniformly random and independent. Then the probability $P = \frac{n!}{(n-k)!n^k}$ that they have chosen k different elements can be estimated as*

$$e^{\frac{-k(k-1)}{2(n-k+1)}} \leq P \leq e^{\frac{-k(k-1)}{2n}}.$$

Proof. We will use the following inequality for $\ln(x)$.

Lemma 2.5. *For $x > 0$,*

$$\frac{x-1}{x} \leq \ln(x) \leq x - 1.$$

Now we estimate

$$\ln\left(\frac{n^k}{n(n-1)\cdots(n-k+1)}\right) = \ln\left(\frac{n}{n-1}\right) + \ln\left(\frac{n}{n-2}\right) + \ldots + \ln\left(\frac{n}{n-k+1}\right)$$
$$\geq \frac{\frac{n}{n-1} - 1}{\frac{n}{n-1}} + \frac{\frac{n}{n-2} - 1}{\frac{n}{n-2}} + \ldots + \frac{\frac{n}{n-k+1} - 1}{\frac{n}{n-k+1}} = \frac{1}{n} + \frac{2}{n} + \ldots + \frac{k-1}{n}$$
$$= \frac{1}{n}(1 + 2 + \ldots + (k-1)) = \frac{k(k-1)}{2n}.$$

Also we find

$$\ln\left(\frac{n^k}{n(n-1)\cdots(n-k+1)}\right) = \ln\left(\frac{n}{n-1}\right) + \ln\left(\frac{n}{n-2}\right) + \ldots + \ln\left(\frac{n}{n-k+1}\right)$$
$$\leq \left(\frac{n}{n-1} - 1\right) + \left(\frac{n}{n-2} - 1\right) + \ldots + \left(\frac{n}{n-k+1} - 1\right) = \frac{1}{n-1} + \frac{2}{n-2} + \ldots + \frac{k-1}{n-k+1}$$
$$\leq \frac{1}{n-k+1} + \frac{2}{n-k+1} + \ldots + \frac{k-1}{n-k+1} = \frac{1}{n-k+1}(1+2+\ldots+(k-1))$$
$$= \frac{k(k-1)}{2(n-k+1)}.$$

Applying the exponential function to both sides of our estimates we get the following:

$$e^{\frac{-k(k-1)}{2(n-k+1)}} \leq \frac{n(n-1)\cdots(n-k+1)}{n^k} \leq e^{\frac{-k(k-1)}{2n}}.$$

□

So the answer to the question in the beginning of this paragraph is between 96.51% and 97.93%. More precisely, the probability is about 97.03%.

Now we will estimate the binomial coefficients. The binomial coefficients in the n-th row of the Pascal's triangle satisfy the following inequalities:

$$\binom{n}{0} < \binom{n}{1} < \cdots < \binom{n}{[n/2]}$$

and

$$\binom{n}{[n/2]+1} > \binom{n}{[n/2]+2} > \cdots > \binom{n}{1} > \binom{n}{0}.$$

Therefore, the middle binomial coefficient $\binom{n}{[n/2]}$ is the largest in the respective row. Stirling's formula implies that the largest binomial coefficient satisfies

$$\binom{n}{n/2} \sim \sqrt{\frac{2}{\pi n}} 2^n.$$

Also we have the following formula describes how binomial coefficients decrease as we move away from the middle of the Pascal's triangle.

Proposition 2.6. *Let m, t be positive integers and $t \leq m$. Then*

$$e^{-t^2/(m-t+1)} \leq \frac{\binom{2m}{m-t}}{\binom{2m}{m}} \leq e^{-t^2/(m+t)}.$$

Proof. Here we prove the lower bound. We have

$$\frac{\binom{2m}{m}}{\binom{2m}{m-t}} = \frac{(m+t)(m+t-1)\cdots(m+1)}{m(m-1)\cdots(m-t+1)}.$$

It will be convenient for us to estimate the logarithm of this quantity.

$$\ln\left(\frac{(m+t)(m+t-1)\cdots(m+1)}{m(m-1)\cdots(m-t+1)}\right) = \ln\left(\frac{m+t}{m}\right) + \ln\left(\frac{m+t-1}{m-1}\right) + \ldots + \ln\left(\frac{m+1}{m-t+1}\right)$$
$$\leq \left(\frac{m+t}{m} - 1\right) + \left(\frac{m+t-1}{m-1} - 1\right) + \ldots + \left(\frac{m+1}{m-t+1} - 1\right) = \frac{t}{m} + \frac{t}{m-1} + \ldots + \frac{t}{m-t+1}$$
$$\leq \frac{t}{m-t+1} + \frac{t}{m-t+1} + \ldots + \frac{t}{m-t+1} = \frac{t^2}{m-t+1}.$$

This finishes the proof of the first inequality. The proof of the second inequality is left to the reader. □

3. Inclusion-exclusion principle

To read:
[1] 2.2.1. Induction, 2.3. Inclusion-Exclusion.
[3] 3.7. Inclusion - Exclusion, 3.8. The hat-check lady.

3.1. **Inclusion-exclusion principle.**

Theorem 3.1. *(Inclusion-Exclusion principle). Let $A_1, \dots, A_n$ be finite sets. Then, the following holds*

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{1\le i\le n} |A_i| - \sum_{1\le i<j\le n} |A_i \cap A_j| + \sum_{1\le i<j<k\le n} |A_i \cap A_j \cap A_k| - \ldots + (-1)^{n-1}|A_1 \cap A_2 \cap \cdots \cap A_n|.$$

Proof. Suppose that an element $a \in \bigcup_{i=1}^{n} A_i$ belongs to exactly k different sets.
How many times did we count a in the inclusion-exclusion formula

$$\sum_{1\le i\le n} |A_i| - \sum_{1\le i<j\le n} |A_i \cap A_j| + \sum_{1\le i<j<k\le n} |A_i \cap A_j \cap A_k| - \ldots \quad ?$$

Element a is counted $(-1)^{\ell-1}\binom{k}{\ell}$ times in the ℓ-th sum as ℓ goes from 1 to n. By the binomial theorem we have

$$\sum_{\ell=1}^{n} (-1)^{\ell-1}\binom{k}{\ell} = 1.$$

Therefore, each element a is counted exactly once. This finishes the proof. □

3.2. **Number of permutations without fixed points.**

A hat-check girl completely loses track of which of n hats belong to which owners, and hands them back at random to their n owners as the latter leave. What is the probability p_n that nobody receives their own hat back?

This question can be reformulated in the following way: find the number of permutations of the set $\{1, 2, \dots, n\}$ without fixed points. In order to count these, we apply the inclusion-exclusion principle. Let A be the set of all permutations and A_i be the set of permutations of the set $\{1, 2, \dots, n\}$ for which i is a fixed point. The number of permutations with no fixed points is

$$|A| - \left|\bigcup_{i=1}^{n} A_i\right|.$$

We know that $|A| = n!$, so we need to count $|\bigcup_{i=1}^{n} A_i|$. We do this using the inclusion principle. Note that $A_i \cap A_j$ represents the set of all permutations for which i and j are fixed points. One can see that $|A_i| = (n-1)!$ for all i, while $|A_i \cap A_j| = (n-2)!$. Using the same idea, we obtain $|A_i \cap A_j \cap A_k| = (n-3)!$ and so on. Altogether, this gives

$$\begin{aligned}
|A| - \left|\bigcup_{i=1}^{n} A_i\right| &= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \ldots \\
&= n! - \frac{n!(n-1)!}{1!(n-1)!} + \frac{n!(n-2)!}{2!(n-2)!} - \ldots \\
&= n!(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \ldots) \\
&\approx n! \exp(-1).
\end{aligned}$$

Thus we see that the probability p_n that nobody receives their own hat back is

$$p_n = \frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \ldots + (-1)^n \frac{1}{n!}$$

As n goes to infinity this number converges to $\frac{1}{e} \approx 0.37$.

3.3. **Euler's totient function.** In number theory, Euler's totient function $\phi(n)$ counts the positive integers up to a given integer n that are relatively prime to n. For example, among the numbers $\{1, 2, 3, 4, 5, 6\}$ only 1 and 5 are coprime to 6. Therefore, we find that $\phi(6) = 2$. If p is a prime number then $\phi(p) = p - 1$ and $\phi(p^k) = p^k - p^{k-1}$.

Proposition 3.2. *Suppose that a number n has the prime factorization $n = p_1^{k_1} \cdots p_m^{k_m}$. Then by the inclusion-exclusion principle we find*

$$\phi(n) = n - \sum_{1\leq i\leq m} \frac{n}{p_i} + \sum_{1\leq i<j\leq m} \frac{n}{p_i p_j} - \sum_{1\leq i<j<k\leq m} \frac{n}{p_i p_j p_k} + \ldots = n \prod_{i=1}^{m} (1 - \frac{1}{p_i}).$$

Proof. Let A be the set of all numbers in $[n]$ *not coprime* with n.
Let A_i be the set of all numbers in $[n]$ divisible by p_i.
Then $A = \bigcup_{i=1}^m A_i$ and $|A_i| = \frac{n}{p_i}$, $|A_i \cap A_j| = \frac{n}{p_i\, p_j}$, and so on. By the inclusion-exclusion formula we find

$$\begin{aligned}\phi(n) =& n - |A| \\ =& n - \sum_{1\leq i\leq m} \frac{n}{p_i} + \sum_{1\leq i<j\leq m} \frac{n}{p_i p_j} - \sum_{1\leq i<j<k\leq m} \frac{n}{p_i p_j p_k} + \ldots = n \prod_{i=1}^{m} (1 - \frac{1}{p_i}).\end{aligned}$$

□

4. Generating functions

To read:
[3] Chapters 12.1, 12.2.

4.1. Combinatorial applications of polynomials.

Example. How many ways are there to pay the amount of 21 francs with 6 one-francs coins, 5 two-francs coins, and 4 five-francs coins? The requited number is in fact the number of solutions of the equation

$$x_1 + x_2 + x_3 = 21, \tag{2}$$

with $x_1 \in \{0,1,2,3,4,5,6\}, x_2 \in \{0,2,4,6,10\}$, and $x_3 \in \{0,5,10,15,20\}$. In order to compute this, we associate to each variable x_i a polynomial p_i as follows:

$$\begin{aligned} p_1(x) &= 1 + x + x^2 + x^3 + x^4 + x^5 + x^6, \\ p_2(x) &= 1 + x^2 + x^4 + x^6 + x^8 + x^{10}, \\ p_3(x) &= 1 + x^5 + x^{10} + x^{15} + x^{20}. \end{aligned}$$

The number of solutions of equation (2) above will be the cofficient of x^{21} in the product $p_1(x)p_2(x)p_3(x)$.

Exercise 1. A box contains 30 red, 40 blue, and 50 white balls; balls of the same color are indistinguishable. How many ways are there of selecting a collection of 70 balls from the box?

4.2. Multinomial theorem.

Theorem 4.1. *(Multinomial theorem). The following holds:*

$$(x_1 + x_2 + \ldots + x_n)^k = \sum_{\substack{i_1, i_2, \ldots, i_n \geq 0 \\ i_1 + i_2 + \ldots + i_n = k}} \frac{k!}{i_1!\, i_2! \cdots i_n!}\, x_1^{i_1}\, x_2^{i_2} \cdots x_n^{i_n}.$$

4.3. Calculation with power series.

Definition 4.2. Let $(a_0, a_1, \ldots)$ be a sequence of real numbers. Then, its *generating function* $a(x)$ is

$$a(x) = a_0 + a_1 x + a_2 x^2 + \ldots.$$

Theorem 4.3. *Let $a_0, a_1, \ldots$ be a sequence of real numbers. If $|a_k| \leq c^k$ for every k, where c is a positive real constant, then the series*

$$a_0 + a_1 x + a_2 x^2 + \ldots$$

is convergent for all x with $|x| < \frac{1}{c}$.

Proof. Since $|a_k| \leq c^k$ for ever k, we have

$$\sum_{k=0}^{\infty} \left| a_k x^k \right| = \sum_{k=0}^{\infty} |a_k|\, |x|^k \leq \sum_{k=0}^{\infty} |cx|^k .$$

Furthermore $|x| < \frac{1}{c}$, therefore $|cx| < 1$ for every k. Next we show $1 + x + x^2 + x^3 + \ldots = \frac{1}{1-x}$ for $x \in (-1,1)$: Let $s = 1 + x + x^2 + x^3 + \ldots + x^{n-1}$, then $xs = x + x^2 + x^3 + \ldots + r^n$ and therefore $s - xs = 1 - x^n$. Thus $s = \frac{1-x^n}{1-x}$ for $x \neq 1$. If $|x| < 1$ the series converges as n goes to infinity. Therefore, we have

$$1 + x + x^2 + x^3 + \ldots = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \text{ for } |x| < 1.$$

Since $|cx| < 1$, we get $\sum_{k=0}^{\infty} |cx|^k = \frac{1}{1-|cx|}$. We have shown that $\sum_{k=0}^{\infty} a_k x^k$ is absolutely convergent, hence it is convergent. □

4.4. **Examples of generating functions.** Consider the following two examples.

Example 1. Consider the sequence $a_n = n + 1$, $n \in \mathbb{Z}_{\geq 0}$. Then the generating function is

$$A(x) = 1 + 2x + 3x^2 + \ldots = \frac{d}{dx}(1 + x + x^2 + \ldots) = \frac{d}{dx}\left(\frac{1}{1-x}\right) = \frac{1}{(1-x)^2}.$$

Example 2. Consider the sequence $b_n = (n+1)^2$, $n \in \mathbb{Z}_{\geq 0}$. Arguing in a similar way, one gets that the generating function is $B(x) = \frac{d}{dx}A(x) - A(x)$.

Exercise 2. What is the generating function of the sequence $(a_0, a_1, \ldots)$ with $a_k = 2^{\lfloor k/2 \rfloor}$?

Theorem 4.4. *(Generalized binomial theorem). For every $r \in \mathbb{R}$ and every integer $n \geq 0$, let*

$$\binom{r}{n} = \frac{r(r-1)\cdots(r-n+1)}{n!}.$$

Then, the following holds:

$$(1+x)^r = \binom{r}{0} + \binom{r}{1}x + \binom{r}{2}x^2 + \cdots$$

for every x with $|x| < 1$.

Proof. Let $f(x) = (1+x)^r$, then $f^{(n)}(0) = r(r-1)(r-2)\cdots(r-n+1)$. Since $\binom{r}{n} = \frac{r(r-1)(r-2)\cdots(r-n+1)}{n!}$, we have $\binom{r}{n} = \frac{f^{(n)}(0)}{n!}$. For a series $a(x) = a_0 + a_1x + a_2x^2 + \ldots$ the element a_n is uniquely determined by $a_n = \frac{a^n(0)}{n!}$. Therefore $(1+x)^r = \binom{r}{0} + \binom{r}{1}x + \binom{r}{2}x^2 + \ldots + \binom{r}{n}x^n + \ldots$.

Next we have to show that the series converges for $|x| < 1$: The series $\sum_{n=0}^{\infty} \binom{r}{n}x^n$ converges if

$$\lim_{n\to\infty} \left| \frac{\binom{r}{n+1}x^{n+1}}{\binom{r}{n}x^n} \right| < 1.$$

This is the case if

$$\lim_{n\to\infty} \left| \frac{n+1}{n-r}x \right| < 1.$$

which holds for $|x| < 1$. □

5. Generating functions. Binary trees.

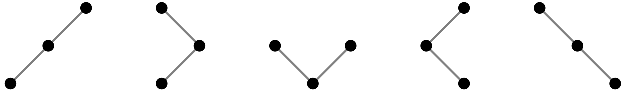
To read: [3] Chapter 12.4.

5.1. **Binary trees.**

Definition 5.1. An inductive definition of a *binary tree* can be given as follows: a binary tree either is empty (it has no vertex), or consists of one distinguished vertex called the root, plus an ordered pair of binary trees called the left subtree and right subtree.

Let b_n denote the number of binary trees with n vertices. Our goal is to find a formula for b_n.
Example. By definition we have $b_0 = 1$ and there is one empty tree. We have $b_1 = 1$, $b_2 = 2$, $b_3 = 5$.

Figure 1. Five different binary trees with three vertices.



The inductive definition of a binary tree implies the following recursive formula for b_n:

$$b_n = b_0\, b_{n-1} + b_1\, b_{n-2} + b_2\, b_{n-3} + \ldots + b_{n-1}\, b_0, \qquad n \in \mathbb{Z}_{\geq 1}. \tag{3}$$

Let $b(x) = \sum_{n=0}^{\infty} b_n x^n$ be the generating series of the sequence $\{b_n\}_{n=0}^{\infty}$. We find

$$b(x)^2 = b_0^2 + (b_1 b_0 + b_0 b_1)x + (b_2 b_0 + b_1 b_1 + b_0 b_2)x^2 + \ldots$$

The recursive relation 3 implies

$$b(x)^2 = b_1 + b_2 x + b_3 x^2 + \ldots = \frac{1}{x}(b_0 + b_1 x + b_2 x^2 + \ldots) - \frac{b_0}{x} = \frac{1}{x} b(x) - \frac{1}{x}.$$

Therefore, the generating function $b(x)$ satisfies the quadratic equation

$$x b(x)^2 - b(x) + 1.$$

This equation has two solutions

$$\frac{1 + \sqrt{1-4x}}{2x} \qquad \text{and} \qquad \frac{1 - \sqrt{1-4x}}{2x}.$$

We observe that the first solution is not bounded around $x = 0$ and the second solution is smooth around $x = 0$ tends to 1 as x tends to 0. Consider the second solution

$$\widetilde{b}(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

It has Taylor expension around $x = 0$

$$\widetilde{b}(x) = \sum_{n=0}^{\infty} \widetilde{b}_n x^n.$$

We have computed that $\widetilde{b}_0 = \widetilde{b}(0) = 1$. Moreover, the function $\widetilde{b}(x)$ satisfies the quadratic equation

$$x\, \widetilde{b}(x)^2 - \widetilde{b}(x) + 1$$

and therefore the sequence $\{\widetilde{b}_n\}_{n=0}^{\infty}$ satisfies the recursive relation (3). Since the sequences satisfy the same initial conditions $b_0 = \widetilde{b}_0$ and the same recursive relation (3) we conclude that $b_n = \widetilde{b}_n$ for all $n \in \mathbb{Z}_{\geq 0}$. The generalized binomial theorem implies

$$\sqrt{1-4x} = \sum_{k=0}^{\infty} (-4)^k \binom{1/2}{k} x^k.$$

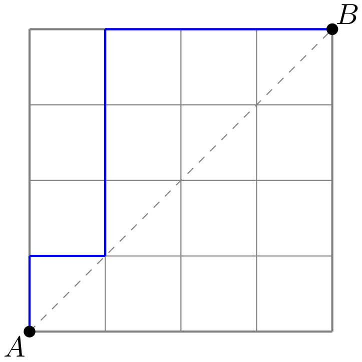
This implies $b_n = \frac{-1}{2}(-4)^{n+1}\binom{1/2}{n+1}$.

Exercise 3. Show that

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

Definition 5.2. The numbers b_n are known by the name *Catalan numbers.*

Exercise 4. Consider an $n \times n$ chessboard:



Consider the shortest paths from the corner A to the corner B following the edges of the squares (each of them consists of $2n$ edges).
(a) How many such paths are there?
(b)* Show that the number of paths that never go below the diagonal (the line AB) is exactly b_n, i.e. the Catalan number. One such path is drawn in the figure.

6. Fibonacci numbers and linear recurrence relations

6.1. **Fibonacci sequence.** The Fibonacci sequence $(F_n)_{n\geq 0}$ is defined by the following recursive formula:

$$F_0 = 0, \quad F_1 = 1, \quad F_n = F_{n-1} + F_{n-2} \quad \forall n \geq 2.$$

Another way to interpret the Fibonacci sequence is the following: let S_n denote the number of ways in which one can climb n stairs if allowed to jump one or two stairs at a time. This is the same as to count the number of the solutions of the equation $x_1 + \ldots + x_k = n$ where $x_i \in \{1, 2\}$ and the number k is not fixed. We observe that $S_1 = 1$, $S_2 = 2$ and $S_{n+2} = S_{n+1} + S_n$ for all $n \in \mathbb{Z}_{\geq 1}$. Therefore, we have $S_n = F_{n+1}$.

Identities for Fibonacci numbers. The sum of the first n numbers of the Fibonacci sequence, is

$$\sum_{k=0}^{n} F_k = F_{n+2} - 1.$$

Exercise 5. Prove the following identities for Fibonacci numbers:

$$(a) \quad F_1 + F_3 + F_5 \ldots + F_{2n-1} = F_{2n}$$
$$(b) \quad F_{2n+1} = 3F_{2n-1} - F_{2n-3}$$
$$(c)^* \quad F_{a+b+1} = F_{a+1}F_{b+1} + F_aF_b.$$

Explicit formula for Fibonacci numbers. We want to find an explicit formula for the value of the n-th Fibonacci number. We will present several possible ways to do that.

Method 1.

We will use the generating functions. Let $F(x)$ denote the generating function of the Fibonacci sequence $(F_0, F_1, \ldots)$ that is

$$F(x) = F_0 + F_1x + F_2x^2 + F_3x^3 + \ldots.$$

Note that the convergence radius of this series is at least $\frac{1}{2}$. Multiplying $F(x)$ by x, respectively x^2, we obtain that

$$xF(x) = F_0x + F_1x^2 + F_2x^3 + F_3x^4 + \ldots$$
$$x^2F(x) = F_0x^2 + F_1x^3 + F_2x^4 + F_3x^5 + \ldots.$$

Recall that for every $n \geq 2$, we have $F_n = F_{n-1} + F_{n-2}$ and consider $F(x) - xF(x) - x^2F(x)$. Grouping together the coefficients of x^k for every k, one obtains that

$$F(x) - xF(x) - x^2F(x) =$$
$$= F_0 + x(F_1 - F_0) + x^2(F_2 - F_1 - F_0) + x^3(F_3 - F_2 - F_1) + \ldots + x^k(F_k - F_{k-1} - F_{k-2}) + \ldots.$$

This implies $F(x) - xF(x) - x^2F(x) = x$ and thus

$$F(x) = \frac{x}{1 - x - x^2}$$

This means, the general term is

$$F_n = \frac{F^{(n)}(0)}{n!}$$

where $F^{(n)}(0)$ is the value in 0 of the n-th derivative of $F(x)$. We factor $1 - x - x^2$ as $-(x - x_1)(x - x_2)$, where $x_{1,2} = \frac{-1\pm\sqrt{5}}{2}$ This means

$$F(x) = \frac{x}{1 - x - x^2} = \frac{A}{x - x_1} + \frac{B}{x - x_2} = \frac{A(x - x_2) + B(x - x_1)}{-(1 - x - x^2)}$$

From this we obtain that

$$A + B = -1 \quad \text{and} \quad Ax_2 + Bx_1 = 0.$$

This is a system of two equations with A and B as unknowns, so we can obtain exact values for A and B:

$$A = \frac{x_1}{\sqrt{5}} \qquad B = \frac{-x_2}{\sqrt{5}}.$$

One can obtain that:

$$\begin{aligned}
F(x) =& \frac{A}{x - x_1} + \frac{B}{x - x_2} = -\frac{A}{x_1}\frac{1}{1 - \frac{x}{x_1}} - \frac{B}{x_2}\frac{1}{1 - \frac{x}{x_2}} = \\
=& -\frac{A}{x_1}\sum_{n=0}^{\infty} x_1^{-n}x^n - \frac{B}{x_2}\sum_{n=0}^{\infty} x_2^{-n}x^n \\
=& \frac{1}{\sqrt{5}}\sum_{n=0}^{\infty} x_1^{-n}x^n - \frac{1}{\sqrt{5}}\sum_{n=0}^{\infty} x_2^{-n}x^n. \\
=& \sum_{n=0}^{\infty} \frac{1}{\sqrt{5}}\left(x_1^{-n} - x_2^{-n}\right)x^n.
\end{aligned}$$

This implies that the general term F_n is

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Method 2.
We look first for a geometric series that satisfies $A_n = A_{n-1} + A_{n-2}$, that is $A_n = c \cdot \alpha^n$ for all $n \in \mathbb{Z}_{\geq 0}$. This implies that $c\alpha^n = c\alpha^{n-1} + c\alpha^{n-2}$ and thus $\alpha^2 - \alpha - 1 = 0$. Solving this quadratic equation, we get $\alpha_{1,2} = \frac{1 \pm \sqrt{5}}{2}$. Next, we search for F_n in the form

$$F_n = c_1\alpha_1^n + c_2\alpha_2^n = c_1\left(\frac{1+\sqrt{5}}{2}\right)^n + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n$$

for some $c_1, c_2 \in \mathbb{R}$. The initial conditions imply

$$\begin{aligned}
F_0 =& c_1 + c_2 = 0 \\
F_1 =& c_1\left(\frac{1+\sqrt{5}}{2}\right) + c_2\left(\frac{1-\sqrt{5}}{2}\right) = 1.
\end{aligned}$$

Thus, the only solution is

$$c_1 = \frac{1}{5} \quad c_2 = \frac{-1}{5}.$$

Hence we find

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

6.2. **Linear recurrence relations.** In general, to solve linear recurrence relations of the form

$$a_{n+k} = c_{k-1}a_{n+k-1} + \ldots + c_0 a_n$$

we have the following recipe. Denote by $\lambda_1, \ldots \lambda_s$ the (possibly complex) roots of the equation

$$\lambda^k = c_{k-1}\lambda^{k-1} + \ldots + c_0$$

where λ_i has multiplicity k_i and $\sum_{i=1}^{s} k_i = k$.

Theorem 6.1. *A formula for a_n is the solutions to the recurrence above if and only if it has the form $a_n = \sum_{i=1}^{s} P_i(n)\lambda_i^n$, where each $P_i(n)$ is a polynomial of degree $k_i - 1$ with coefficients chosen arbitrarily. Moreover, for any set of initial values $a_0, \ldots, a_{k-1}$ one can find coefficients of the polynomials $P_i(n)$ so that the solution fits to the initial values. Note that the number of coefficients to be determined is equal to k, the number of initial values.*

7. Möbius inversion formula

To read:
[5] Chapter 2.1.

Definition 7.1. Suppose that a positive integer n has the prime factorization

$$n = p_1^{e_1} \cdots p_r^{e_r}.$$

We define the *Möbius function* $\mu(n)$ as:

$$\mu(n) = \begin{cases} 1 \text{ for } n = 1, \\ 0 \text{ if some } e_i > 1, \\ (-1)^r \text{ if } e_1 = \ldots = e_r = 1. \end{cases}$$

Lemma 7.2. *For $n \in \mathbb{Z}_{\geq 1}$ we have*

$$\sum_{d|n} \mu(d) = \begin{cases} 1 \text{ if } n = 1, \\ 0 \text{ if } n > 1. \end{cases}$$

Here the summation is taken over all positive divisors on n.

Proof. First consider the case $n = 1$. It follows immediately from the definition

$$\sum_{d|1} \mu(d) = \mu(1) = 1.$$

Next, suppose that $n > 1$ and it has the prime decomposition $n = p_1^{e_1} \cdots p_r^{e_r}$. Set $n^* := p_1 \cdots p_r$. If $d \mid n$ and $d \nmid n^*$ then d has a prime divisor of multiplicity bigger then 1 and therefore $\mu(d) = 0$. Hence, we have

$$\sum_{d|n} \mu(d) = \sum_{d|n^*} \mu(d).$$

Now we can easily compute

$$\sum_{d|n^*} \mu(d) = 1 - \binom{r}{1} + \binom{r}{2} - \binom{r}{3} + \ldots = (1-1)^r = 0.$$

This finishes the proof. □

Theorem 7.3. *(Möbius inversion formula) Let functions $f, g : \mathbb{Z}_{\geq 1} \to \mathbb{R}$ be such that*

$$f(n) = \sum_{d|n} g(d).$$

Then

$$g(n) = \sum_{d|n} \mu(d)\, f(n/d).$$

Proof. We have

$$f(n/d) = \sum_{d'|(n/d) g(d')} \quad \text{for all } d \mid n.$$

Therefore

$$\sum_{d|n} \mu(d)\, f(n/d) = \sum_{d|n} \mu(d) \sum_{d'|(n/d)} g(d').$$

Let $n = dd'n_1$. For a fixed d', the value of d runs over all positive divisors of n/d'. Hence we get

$$\sum_{d|n} \mu(d) \sum_{d'|(n/d)} g(d') = \sum_{d'|n} g(d') \sum_{d|(n/d')} \mu(d).$$

We apply the previous lemma to the sum $\sum_{d|(n/d')} \mu(d)$ and obtain

$$\sum_{d'|n} g(d') \sum_{d|(n/d')} \mu(d) = g(n).$$

This finishes the proof. □

7.1. **Identities with Euler's totient function.**

Exercise 6. Show that for all $n \in \mathbb{Z}_{\geq 1}$ we have

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

Hint: Let Φ_n be the set all elements in $[n]$ coprime to n:

$$\Phi_n := \{m \in [n] \mid m \ \text{ is coprime to } n\}.$$

Show that $[n]$ is the disjoint union of sets $(n/d) \cdot \Phi_d$ where d runs over all divisors of n:

$$[n] = \dot{\bigcup}_{d|n} (n/d) \cdot \Phi_d.$$

Exercise 7. Show that $\phi(n) = n \sum_{d|n} \frac{\mu(d)}{d}$.

7.2. **Number of cyclic sequences.**

Definition 7.4. Let A be a set. A *linear sequence* of length n on an A is a sequence of the form

$$(a_1, \dots, a_n), \quad a_k \in A \text{ for } k = 1, \dots n.$$

In other words, a linear sequence is a function $a : [n] \to A$.

The number of linear sequences of length n on an alphabet of size r is r^n.

Consider the following equivalence relation $\sim$ on the set of linear sequences:

$$(a_1, ..., a_n) \sim (a_1, ..., a_n)$$

and

$$(a_1, ..., a_n) \sim (a_k, a_{k+1}, \dots, a_1, \dots, a_{k-1}), \quad k = 2, \dots n.$$

In other words, two linear sequences are equivalent if one of them can be obtained from another by a cyclic shift.

Example. Linear sequences of length 3 on the alphabet $\{a, b\}$:

$$(a, a, a)$$
$$(a, a, b)$$
$$(a, b, a)$$
$$(a, b, b)$$
$$(b, a, a)$$
$$(b, a, b)$$
$$(b, b, a)$$
$$(b, b, b).$$

Cyclic sequences of length 3 on the alphabet $\{a, b\}$:

$$\begin{aligned}
&(a, a, a)\\
&(a, a, b) \sim (a, b, a) \sim (b, a, a)\\
&(a, b, b) \sim (b, b, a) \sim (b, a, b)\\
&(b, b, b).
\end{aligned}$$

Definition 7.5. A *cyclic sequence* of length n on an alphabet A is an equivalence class of linear sequences with respect to the relation $\sim$.

Proposition 7.6. *The number $T(n, r)$ of cyclic sequences of of length n on an alphabet of size r is*

$$T(n, r) = \frac{1}{n} \sum_{d|n} \phi(n/d) r^d.$$

Proof. A *period* of a cyclic sequence $(a_1, \ldots, a_n)$ is a minimal number $k \in \{1, 2, \ldots, n\}$ such that $(a_1, \ldots, a_n) = (a_{1+k}, \ldots, a_n, a_1, \ldots a_k)$ (equal as linear sequences). Note that the period of a sequence is a divisor of the the sequence's length.

Let $M(d, r)$ be the number of cyclic sequences of of length d and period exactly d. It is easy to see that

$$r^n = \sum_{d|n} d\, M(d, r).$$

The Möbius inversion formula implies

$$n\, M(n, r) = \sum_{d|n} \mu(n/d)\, r^d. \tag{4}$$

We have

$$T(n, r) = \sum_{d|n} M(d, r).$$

We combine this identity with (4) and obtain

$$\begin{aligned}
T(n, r) =& \sum_{d|n} \frac{1}{d} \sum_{d'|d} \mu(d'/d)\, r^{d'}\\
&\qquad\text{(here we intoduce a new summation variable } d'' = \frac{d}{d'})\\
=& \sum_{d'|n} r^{d'} \left(\sum_{d''|\frac{n}{d'}} \frac{1}{d'd''}\, \mu(d'') \right).
\end{aligned}$$

Now we use the identity

$$\sum_{d''|\frac{n}{d'}} \frac{1}{d''}\, \mu(d'') = \frac{\phi(n/d')}{n/d'}$$

and arrive at

$$\begin{aligned}
T(n, r) =& \sum_{d'|n} r^{d'} \frac{1}{d'} \frac{\phi(n/d')}{n/d'}\\
=& \frac{1}{n} \sum_{d'|n} \phi(n/d')\, r^{d'}.
\end{aligned}$$

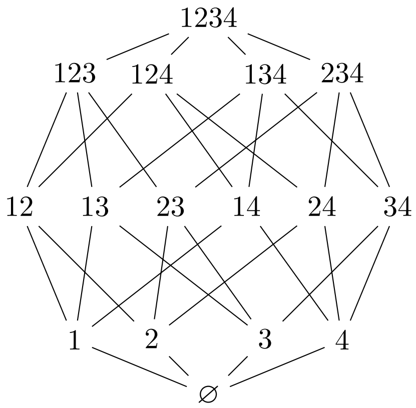
This finishes the proof. □

7.3. **Partially ordered sets (or posets).** *This section is written by Dr. Matthew de Courcy-Ireland.*

Definition 7.7. A binary *relation* on a set A is a subset $R \subseteq A \times A$. A relation is *reflexive* provided that $(x,x) \in R$ for every $x \in A$. A relation is *antisymmetric* provided that $(a,b) \in R$ and $(b,a) \in R$ together imply $a = b$. A relation is *transitive* if $(a,b) \in R$ and $(b,c) \in R$ together imply $(a,c) \in R$. A relation is *reflexive* if $(a,a) \in R$ for all $a \in R$.

Definition 7.8. (partial order) A partial order on a set A is an antisymmetric, reflexive, and transitive relation $R \subseteq A \times A$. A partially ordered set, or *poset* for short, is a set together with a partial order.

Example. The subsets of a given set A are partially ordered by inclusion.

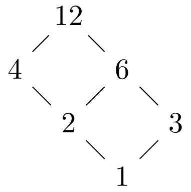


7.4. **Hasse diagram.** The *Hasse diagram* is a useful way to draw partially ordered sets. Essentially, we draw a point for each element of the poset, and a line from x to y when $x < y$. But many of these lines are redundant, in view of transitivity:

$$x < y < z \implies x < z$$

So we need only draw a line when $x < y$ and there is no element in between. By convention, we draw x lower than y.

Example. Let X be the factors of 12, ordered by divisibility. The Hasse diagram is then



7.5. **Möbius invertion for posets.** Given a function f defined on a partially ordered set $(X, \leq)$, we may form the sum

$$g(x) = \sum_{y \leq x} f(y)$$

assuming that f is real-valued, or at least that there is some way to add the values $f(y)$. To guarantee that the sum is well-defined, we assume that there are only finitely many terms y beneath any given x. For instance, this holds if X is finite. How do we recover f from g?

Theorem 7.9. *(Möbius inversion for posets) Given a partially ordered set X, there is a two-variable function $M : X \times X \to \mathbb{R}$ such that*

$$g(x) = \sum_{y \leq x} f(y) \iff f(x) = \sum_{y \leq x} g(y)M(y, x)$$

This function M is called the *Möbius function* of the poset. To show it exists for any partial order, and to compute it for specific orders, we introduce an algebraic structure that captures the order relation.

Definition 7.10. (incidence algebra) Given a partially ordered set X, the *incidence algebra* $A(X)$ is the set of all real-valued functions $f : X^2 \to \mathbb{R}$ satisfying $f(x, y) = 0$ unless $x \leq y$. More generally, for any abelian group G, we define $A_G(X)$ as the set of all G-valued functions $f : X^2 \to G$ satisfying $f(x, y) = 0$ unless $x \leq y$, where 0 now denotes the identity element of G. The elements of $A_G(X)$ are called *incidence functions*, or G-valued incidence functions.

If G is a field, in particular for $G = \mathbb{R}$, then $A_G(X)$ is a vector space over G with respect to pointwise addition and scalar multiplication. The extra structure that makes it an "algebra" is the following operation.

Definition 7.11. (convolution) Given $f, g \in A(X)$, their *convolution* $f * g$ is defined by $f * g(x, y) = 0$ unless $x \leq y$, in which case

$$f * g(x, y) = \sum_{x \leq z \leq y} f(x, z)g(z, y)$$

The sum is well-defined assuming that there are finitely many z in between x and y, which is certainly the case for finite posets and also holds for many natural infinite ones. By construction, $f * g$ is again in $A(X)$.

Definition 7.12. (locally finite) A poset X is called *locally finite* provided that for any elements x, y, there are only finitely many $z \in X$ in the interval $x \leq z \leq y$.

Example. The rational numbers, in their usual order, do not form a locally finite poset. The interval $0 < z < 1$ contains infinitely many elements $1/n$ for $n = 1, 2, 3, \ldots$

To define convolution for G-valued functions, there must be some notion of multiplication as well as addition. Thus the same concept applies for any ring instead of $\mathbb{R}$.

Example. If $X = \{1, \ldots, n\}$ with the usual order $1 < \ldots < n$, then a two-variable function $f(x, y)$ is just an $n \times n$ matrix, and the incidence condition $f(x, y) = 0$ unless $x \leq y$ says that this is a triangular matrix. Convolution is the usual notion of matrix multiplication.

In particular, the example of matrix multiplication shows that convolution is not always commutative.

Example. (delta function) Returning to a more general poset, the analogue of the identity matrix is

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if not} \end{cases}$$

Clearly $\delta(x, y)$ unless $x \leq y$, indeed unless $x = y$. Thus δ is an incidence function. For any incidence function f,

$$\delta * f = f * \delta = f$$

because there is one non-zero term $f(x,z)\delta(z,y)$ in the interval $x \le z \le y$, namely $z = y$. So the sum defining $f * \delta(x,y)$ is just $f(x,y)$. Similarly, for $\delta * f(x,y)$ we have only a single term $f(x,y)$ when $z = x$.

Proposition 7.13. *Convolution is associative: for any incidence functions f, g, h on a (locally finite) poset X*

$$(f * g) * h = f * (g * h)$$

Sketch of proof. For any $x \le y$, both sides $(f * g) * h(x,y)$ and $f * (g * h)(x,y)$ are given by the sum of $f(x,z)g(z,w)h(w,y)$ over all z and w between x and y and satisfying $z \le w$. □

Proposition 7.14. *(convolution inverses) Let X be a (locally) finite poset. An incidence function $f(x,y)$ has an inverse g satisfying $f * g = \delta$ if and only if $f(x,x) \neq 0$ for all x. In that case, the inverse works on both sides: $f * g = g * f = \delta$.*

Proof. Suppose there is an inverse. Then, for any x,

$$f * g(x,x) = \delta(x,x) = 1$$

On the other hand, there is only one term $f(x,x)g(x,x)$ in the sum over $x \le z \le x$ defining $f * g(x,x)$, namely $z = x$ (a partial order is antisymmetric!). If $f(x,x)g(x,x) = 1$, then $f(x,x) \neq 0$ or else the product would be 0.

Conversely, suppose $f(x,x) \neq 0$ for all x. We define $g(x,y)$ inductively. Note that there are only finitely many z in the interval $x \le z \le y$. For $x = y$, define $g(x,x) = 1/f(x,x)$. Then the required identity holds in the form $g(x,x)f(x,x) = 1 = \delta(x,x)$ because the sum over $x \le z \le y$ is just a single term in this case where $x = y$ (this uses the fact that $\le$ is antisymmetric to go from $x \le z \le x$ to $z = x$). If $x \neq y$, we assume inductively that $g(x,z)$ has already been defined for $z < y$, and then define

$$g(x,y) = \frac{1}{f(y,y)} \left(- \sum_{x \le z < y} g(x,z) f(z,y) \right)$$

By construction, $g * f(x,y) = 0 = \delta(x,y)$ since multiplying through gives the missing term $g(x,y)f(y,y)$ corresponding to $z = y$ in the sum. One can construct g' satisfying $f * g' = \delta$ by a similar induction, and it must be that $g = g'$ because $*$ is associative. Indeed, start from $f * g' = \delta$ and multiply on the left by g. We obtain $g * (f * g') = g * \delta = g$ because δ is the neutral element. But by associativity, the other side is

$$g * (f * g') = (g * f) * g' = \delta * g' = g'$$

appealing once again to the neutrality of δ. It follows that $g = g'$. □

Definition 7.15. The *zeta function* of a poset is defined by

$$Z(x,y) = \begin{cases} 1 & \text{if } x \le y \\ 0 & \text{if not} \end{cases}$$

Since $Z(x,x) = 1 \neq 0$ for every x, the Proposition 7.14 implies that there is an incidence function M satisfying

$$M * Z = Z * M = \delta$$

This M is called the *Möbius function* of the poset. From the construction of inverses, we have $M(x,x) = 1$ for every x and, for $x < y$,

$$M(x,y) = - \sum_{x \le z < y} M(x,z)$$

Now we can prove Theorem 7.9. The required function M is exactly the Möbius function of the poset. Recall what we have to show:

$$g(x) = \sum_{y \leq x} f(y) \iff f(x) = \sum_{y \leq x} g(y) M(y, x)$$

Define a new poset X' by adding a new element less than everything in X. In other words, let $-\infty$ be anything not already in X and extend the order by $-\infty < x$ for all $x \in X$. For any function f on X, there is a corresponding incidence function f' on X' defined by

$$f'(-\infty, x) = f(x), \qquad f'(x, y) = 0 \qquad \text{for all } x, y \in X$$

The zeta and Möbius functions of X' extend those of X by

$$Z(-\infty, y) = 1 \text{ for all } y \in X'$$

so we use the same symbols Z and M rather than Z' and M'.

Because Z and M are convolution inverses,

$$g' = f' * Z \iff f' = g' * M$$

In particular, evaluated at the pair $(-\infty, x)$, the quantity on the left is

$$g(x) = g'(-\infty, x) = \sum_{-\infty \leq y \leq x} f'(-\infty, y) Z(y, x) = \sum_{y \leq x} f(x)$$

while the quantity on the right is

$$f(x) = f'(-\infty, x) = \sum_{-\infty \leq y \leq x} g'(-\infty, y) M(y, x) = \sum_{y \leq x} g(y) M(y, x)$$

We obtain Theorem 7.9 as originally stated. □

Exercise 8. The set $\mathbb{Z}_{\geq 0}$ is partially ordered by the usual $\leq$ relation. Compute the Möbius function of this poset.

8. Elements of graph theory

To read:
[1] 8.1. How to Define Trees?,
[3] 4.1. The notion of a graph; isomorphism - only the definition of graphs, 4.3.1. Sum of the degrees, 4.3.2. Handshakes lemma, 5.1.

8.1. Definition and characterizations of trees.

Definition 8.1. A *graph* G is an ordered pair (V, E), where V is a set of elements called *vertices* and E is a set of 2-element subsets of V called *edges.*

Definition 8.2. Let $G = (V, E)$ be a graph. We call a sequence of distinct vertices $v_0, \dots, v_r$ a *path* if $\{v_i, v_{i+1}\}$ is an edge of G, for every $0 \leq i \leq r-1$.

Definition 8.3. We say that a graph $G = (V, E)$ is *connected* if for every two vertices $u, v \in V$ there exists a path in G between u and v.

Definition 8.4. For every vertex of a graph, we define its *degree* as the number of edges adjacent to it.

Definition 8.5. A *cycle* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, \dots, v_r \in V$ with $r \geq 3$ such that $\{v_i, v_{i+1}\} \in E$ for all i from 1 to $r-1$ and moreover $\{v_r, v_1\} \in E$.

Definition 8.6. A *tree* is a connected graph without cycles.

Definition 8.7. A vertex of degree one in a tree is called a *leaf.*

Lemma 8.8. *Every tree on $n \geq 2$ vertices has at least two leaves.*

Proof. Let S be the set of all the paths in the tree T. We know that every path on r vertices contains exactly $r-1$ edges. Consider now a path $v_1, \dots, v_l$ of maximum length. One can always find a path of maximum length since every path in the tree can contain at most n vertices (otherwise it will be self-intersecting, that is it will contain a cycle, which is impossible since in a tree we cannot have cycles). We prove that both v_1 and v_l (the endpoints of the path) are leafs. Assume at least one of them is not, say v_1. That means that, there is at least another edge apart from $\{v_1, v_2\}$ incident to v_1. Observe that u cannot coincide with any of the vertices of the path $v_1, \dots, v_l$ (otherwise it will close a cycle). Therefore, we can add u to the path without forming any cycle. But this is a contradiction to the maximality of the length of the path $v_1, \dots, v_l$. Thus, both v_1 and v_l must be leaves. □

Theorem 8.9. *Every tree on n vertices has exactly $n-1$ edges.*

9. Equivalent definitions of a tree, number of labeled trees

To read:
[1] 8.3. How to Count trees? 8.4. How to Store trees?
[3] 4.1 The notion of a graph; isomorphism 5.1 Definition and characterizations of trees 8.1. The number of spanning trees, 8.4. A proof using the Prüfer codes.

9.1. **Graph isomorphisms.**

Definition 9.1. Two graphs $G = (V, E)$ and $G' = (V', E')$ are called isomorphic if a bijection $f : V \to V'$ exists such that $\{x, y\} \in E$ if and only if $\{f(x), f(y)\} \in E'$ holds for all $x, y \in V$, $x \neq y$. Such an f is called an isomorphism of the graphs G and G'. The fact that G and G' are isomorphic is written $G \cong G'$.

9.2. **Characterizations of trees.**

Theorem 9.2. *The following five properties are equivalent:*

1. *T is a tree.*
2. *T is maximally acyclic, that is, it is acyclic, but if we add any edge to T, then it will contain a cycle.*
3. *Any two vertices in T are connected by a unique path.*
4. *T has one edge less than the number of vertices and it is connected.*
5. *T has one edge less than the number of vertices and it is acyclic.*

9.3. **Counting labeled trees.** In what follows, we will present a result due to Cayley. Before stating the theorem, we need the following lemma:

Lemma 9.3. *Let T be a tree on n labeled vertices and let $d_1, \dots, d_n$ be the degrees of the vertices. Then*

$$\sum_{i=1}^{n} d_i = 2|E(T)| = 2(n-1)$$

where by $E(T)$ denotes the edge set of the tree.

Now we can state Cayley's theorem.

Theorem 9.4. *(Cayley). The number of trees on n labeled vertices is n^{n-2}.*

We give two proofs to this theorem. The first one, due to Prüfer, is algorithmic.
Proof 1 of Cayley's theorem. We give now the proof, due to Prüfer. Denote the vertices by $\{1, 2, \dots, n\}$. We will define a one-to-one correspondence between the set of all trees on n labeled vertices and the set of all sequences of length $n-2$ consisting of numbers in $\{1, 2, \dots, n\}$. Since the cardinality of the latter is n^{n-2}, we obtain the desired result. The following algorithm takes a tree as input, and yields a sequence of integers:

Step 1: Find the leaf with the smallest label and write down the number of its neighbor.
Step 2: Delete this leaf, together with the only edge adjacent to it.
Step 3: Repeat until we are left with only two vertices.

We present an algorithm that reconstructs the tree from the Prüfer code.

Step 1: Draw the n nodes, and label them from 1 to n.
Step 2: Make a list of all the integers $(1, 2, \dots, n)$. This will be called the list.

Step 3: If there are two numbers left in the list, connect them with an edge and then stop. Otherwise, continue on to step 4.

Step 4: Find the smallest number in the list which is not in the sequence. Take the first number in the sequence. Add an edge connecting the nodes whose labels correspond to those numbers.

Step 5: Delete the smallest number from the list which is not in the sequence and the first number in the sequence. This gives a smaller list and a shorter sequence. Then return to step 3.

9.4. **Counting unlabeled trees.** The number of unlabeled trees, that is, classes of pairwise nonisomorphic trees is only exponential in the number of vertices. We prove the following theorem:

Theorem 9.5. *The number of pairwise nonisomorphic trees on n vertices is at most 2^{2n-4}.*

Here is a sketch of a proof: The proof uses the following encoding of trees. We think of a tree hanged from one of its vertices on a plane (we think of gravity working in the negative y-direction). We go around the tree and form a binary sequence. If we are going one edge down, we write 1 in the sequence. If we are going up - we write 0. At the end we corresponded one 0 and one 1 to each edge, which gives us a binary sequence of length $2n-2$. The last bit is always 0, and the first bit is always 1, so the total number of these sequences is at most 2^{2n-4}. □

Theorem 9.6. *The number of pairwise nonisomorphic trees on n vertices is at least $\frac{n^{n-2}}{n!}$.*

10. Kruskal's algorithm for finding a minimal spanning tree

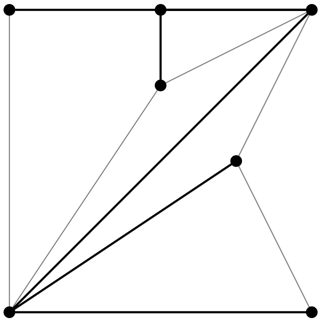
To read:
[1] 9.1. Finding the best tree
[3] 5.4. Minimum spanning tree problem

10.1. **Subgraphs, induced subgraphs, and spanning trees.**

Definition 10.1. Let G and G' be graphs. We say that G is a subgraph of G' if $V(G) \subset V(G')$ and $E(G) \subseteq E(G')$. We say that G is an induced subgraph of G' if $V(G) \subseteq V(G')$ and $E(G) = E(G') \cap \binom{V(G)}{2}$.

Definition 10.2. Let $G = (V, E)$ be a graph. We say that a tree T is a spanning tree of G if it contains all the vertices of V and is a subgraph of G, that is every edge in the tree belongs to the graph G.

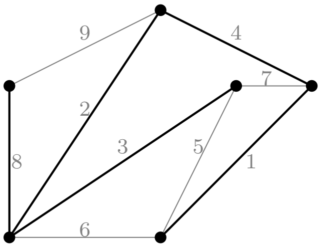
Example. Below is an example of a spanning tree:



10.2. **Weighted graphs.**

Definition 10.3. A *weighted graph* is a graph in which each edge is given a numerical weight. We define the *weight of a graph* as the sum of the weights of all its edges.

We are interested in the following problem: find a minimum weight spanning tree T for a given weighted connected graph G.
Example. A minimum weight spanning tree in a weighted connected graph.



One way to solve the problem of finding a minimum spanning tree is using Kruskal's algorithm. This works as follows:

Step 1. Start with an empty graph.
Step 2. Take all the edges that have not been selected and that would not create a cycle with the already selected edges and select it unless it creates a cycle. Add the one with the smallest weight.
Step 3. Repeat until the graph is connected.

Theorem 10.4. ***(Correctness of Kruskal's algorithm).*** *The Kruskal's algorithm solves the minimum spanning tree problem.*

Proof. The proof can be found in [1] Chapter 9.1. □

11. Counting spanning trees in a graph. Kirchhoff's theorem

11.1. **A useful fact form linear algebra.**

Theorem 11.1. *(Binet–Cauchy theorem). Let A be an arbitrary matrix with n rows and m columns. Then*

$$\det(A\,A^T) = \sum_I \det(A[I])^2,$$

where the sum is over all n-element subsets $I \subseteq \{1, 2, ..., m\}$, and where $A[I]$ denotes the matrix obtained from A by deleting all columns whose indices do not lie in I.

11.2. **Laplace matrix and incidence matrix.**

Definition 11.2. Let $G = (V, E)$ be a graph. An *orientation* $\mathfrak{o}$ on G is the choice of the ordered pair (u, v) or (v, u) for each edge $\{u, v\} \in E$. (If we choose (u, v), say,then we think of putting an arrow one pointing from u to v, and we say that $\{u, v\}$ is directed from u to v ,that u is the *initial vertex* and v the *final vertex* of $\{u, v\}$).

Definition 11.3. Let $G = (V, E)$ be a graph and $\mathfrak{o}$ be an orientation on G. The *incidence* matrix of G with respect to $\mathfrak{o}$) is the matrix $I(G, \mathfrak{o}) \in M_{|V|\times|E|}(\mathbb{Z})$, where the its entries $I_{v,e}$ for $v \in V$, $e \in E$ are given by

$$I_{v,e} = \begin{cases} 1, & \text{if the edge } e \text{ has initial vertex } v \text{ ,} \\ -1, & \text{if the edge } e \text{ has initial vertex } v \text{ ,} \\ 0, & \text{otherwise.} \end{cases}$$

Definition 11.4. The *Laplace* matrix of G is the matrix $L(G) \in M_{|V|\times|V|}(\mathbb{Z})$, where the its entries $L_{u,v}$ for $u, v \in V$ are given by

$$L_{u,v} = \begin{cases} \deg(u), & \text{if } u = v, \\ -1, & \text{if } u \neq v \text{ and } \{u, v\} \text{ is an edge,} \\ 0, & \text{otherwise.} \end{cases}$$

11.3. **Kirchhoff's theorem.**

Theorem 11.5. *(Kirchhoff) Let G be a finite connected graph with Laplace matrix $L = L(G)$. Let L_0 denote L with the last row and column removed. Then the number of spanning trees $\kappa(G)$ satisfies $\kappa(G) = \det(L_0)$.*

Proof. The proof can be found in [3] Section 8.5. □

Exercise 9. Suppose that a connected graph G has n vertices. Show that

$$\kappa(G) = \frac{1}{n}\lambda_1 \cdots \lambda_{n-1},$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $L(G)$ and $\lambda_n = 0$.

12. The probabilistic method.

12.1. **Finite probability spaces.** We denote by Ω a probability space, that is, the set consisting of some elements called elementary events equipped with a measure p such that

(1) $p(A) \geq 0$ for any event (by event we mean any union of some elementary events)
(2) $p(\Omega) = 1$.
(3) $p(A \cup B) = p(A) + p(B)$ for any disjoint events A, B.

For simplicity of exposition we work only with discrete probability here, that is, we assume that Ω is finite. A random variable $X : \Omega \to \mathbb{R}$ is just any measurable function that assigns values to elementary events. Note that the measure p does not appear in this definition. However, it appears in the next one. If X takes values $x_1, \dots x_k$, then the expectation $\mathbb{E}(X)$ of X is defined as

$$\mathbb{E}(X) = \sum_{i=1}^{k} x_i \, p(X = x_i).$$

Note that $\sum_{i=1}^{k} p(X = x_i) = 1$. Informally, it is a weighted average of X with respect to p. Some useful properties:

a) The probability of a union of events $A_1, \dots, A_n$ is at most the sum of the probabilities of the events

$$p(A_1 \cup \cdots \cup A_n) \leq p(A_1) + \ldots + p(A_n).$$

b) If $A_1, \dots, A_n$ are independent events, then

$$p(A_1 \cap \cdots \cap A_n) = p(A_1) \cdots p(A_n).$$

c) The linearity of expectation: If $X_1, \cdots, X_n$ are random variables and $a_1, \dots, a_n$ an are real numbers, then

$$\mathbb{E}(a_1 X_1 + \ldots + a_n X_n) = a_1 \mathbb{E}[X_1] + ::: + a_n \mathbb{E}[X_n].$$

d) If $\mathbb{E}(X) = m$, then there is at least one elementary event A_1 such that $X(A_1) \geq m$, and, analogously, there is at least one elementary event A_2 such that $X(A_2) \leq m$.

A general framework for the probabilistic method is the following: we are given a finite set of objects Ω and $X : \Omega \to \mathbb{R}$ is a function assigning to each object $A \in \Omega$ a real number. The goal is to show that there is at least one element $A \in \Omega$ for which $X(A)$ is at least a given value m. For this, we define a probability distribution $p : \Omega \to [0, 1]$ and consider the resulting probability space, where X becomes a random variable. Showing that the expected value of X is at least m is enough, since, if this holds, then there exists at least one event $A \in \Omega$ for which $X(A) \geq m$.

12.2. **Applications of probabilistic method.**

12.2.1. *Existence of large bipartite subgraphs.*

Theorem 12.1. *Let G be a graph with an even number, $2n$, of vertices and with $m > 0$ edges. Then the set $V = V(G)$ can be divided into two disjoint n-element subsets A and B in such a way that more than $m/2$ edges go between A and B.*

Proof. The proof can be found in [3] page 307. □

12.2.2. *Turán's theorem.*

Definition 12.2. Let $G = (V, E)$ be a graph. An *independent* set is a set of vertices $S \subseteq V$ such that no two of them are connected by an edge.

Theorem 12.3. *(Turán) For any graph G on n vertices, we have*

$$\alpha(G) \geq \frac{n^2}{2|E(G)| + n},$$

where $\alpha(G)$ denotes the size of the largest independent set of vertices in the graph G.

Proof. The proof can be found in [3] pages 308–309. □

Exercise 10. Consider the disjoint union of m copies of the complete graph K_r. Show that this graph attains the bound of Theorem 12.3.

12.2.3. *Schütte's problem.* (Not covered in the course)

Definition 12.4. A *tournament* is a directed graph obtained by assigning a direction for each edge in an undirected complete graph. In a tournament, every pair of distinct vertices is connected by a single directed edge.

Definition 12.5. We say that a tournament $T = (V, E)$ has property S_k if for any k vertices $v_1, \ldots, v_k \in V$ there exist a vertex $u \in V(T)$ such that $\overrightarrow{uv_1}, \ldots, \overrightarrow{uv_k} \in E$.

Schütte's problem can be formulated as follows: do such tournaments exist, for every k fixed? The answer is given by the following theorem:

Theorem 12.6. *For every integers $k \geq 1$ and $n \geq k^2\, 2^k(\ln(2) + o(1))$, there exist a tournament on n vertices having property S_k.*

Proof. The proof can be found in [4], pages 44–45 or [3] pages 299–300.

13. Bipartite graphs. König-Hall theorem. Sperner theorem

This material was not covered in the course. **To read:**
[1] 10.3. The Main Theorem
[3] 7.2. Sperner's theorem on independent systems: Sperner theorem and proof of Theorem 7.2.1.

Definition 13.1. A *bipartite graph* (or *bigraph*) is a graph G whose vertices can be divided into two disjoint sets A and B such that every edge of the graph connects a vertex in A to one in B (in other words, there is no edge of the graph between two vertices of A or two vertices of B.

Lemma 13.2. *A graph is bipartite if, and only if, it does not contain an odd cycle (that is, a cycle of odd length).*

Definition 13.3. Let $G = (V, E)$ be a graph. A subset $E_0 \subset E$ of pairwise disjoint edges (that is, edges which do not share any vertex) is called a *matching* in G.

Definition 13.4. A *perfect matching* is a matching where every vertex of the graph is incident to exactly one edge of the matching.

Remark. A perfect matching is therefore a matching of a graph containing $n/2$ edges (where n is the number of vertices). Thus, perfect matchings are only possible on graphs with an even number of vertices!

Theorem 13.5. *(König-Hall). Let $G = (V, E)$ be a bipartite graph with bipartition $V = A \cup B$ such that $|A| = |B|$. For every $X \subset A$, let*

$$B(X) := \{b \in B \mid \text{ there exists } x \in X \text{ such that } (b, x) \in E\}.$$

Then, a perfect matching in the graph exists if and only if $|B(X)| \geq |X|$, for all $X \subset A$.

Proof. You can find the proof in [1], chapter 10.3. The Main Theorem or [4], page 83. □

Theorem 13.6. *(Sperner). Let $X = \{1, 2, \ldots, n\}$ and $A_1, \ldots, A_m \subseteq X$, with $A_j \not\subseteq A_i$, for all $i \neq j$. Then $m \leq \binom{n}{\lfloor n/2 \rfloor}$.*

Proof. Two proofs of Sperner's theorem (one of them using LYM inequality) can be found in [3], pages 227–229. □

Acknowledgements: I thank Prof. Janos Pach for designing this course and Dr. Matthew de Courcy-Ireland for sharing his lecture notes.

References

[1] Discrete Mathematics (L. Lovasz, J. Pelikan , K. Vesztergombi);
[2] Combinatorics: Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability (B. Bollobas);
[3] Invitation to Discrete Mathematics (J. Matousek, J. Nesetril).
[4] Extremal combinatorics (S. Jukna).
[5] Combinatorial theory (M. Hall), Blaisdell publishing company, 1967.