Lecture 1. Counting problems

Denote by \([n]\) the set of first \(n\) natural numbers: \([n] := \{1, 2, \ldots, n\}\).
Recall the following formulas:
- the number of functions from \([m]\) to \([n]\) is \(n^m\). This is the number of \(m\)-letter words in an \(n\)-letter alphabet.
- the number of permutations of a set of \(n\) elements is \(n!\).
- the number of ways in which one can choose \(k\) objects out of \(n\) distinct objects, assuming the order of the elements matters, is \(\frac{n!}{(n-k)!}\).
- 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.

The following is called Pascal’s triangle
The following identities hold:
1. \(\binom{n}{k}\) is the \(k\)-th element in the \(n\)-th line of Pascal’s triangle.
2. \(\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}\).
3. All the diagonals of Pascal’s triangle are strictly increasing.

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.

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

For \(x = 1\), respectively \(x = -1\), we obtain

\[
2^n = \binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = \sum_{i=0}^{n} \binom{n}{i},
\]

\[
0 = \binom{n}{0} - \binom{n}{1} + \ldots + (-1)^n \binom{n}{n} = \sum_{i=0}^{n} (-1)^i \binom{n}{i},
\]

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

\[
2^{n-1} = \binom{n}{0} + \binom{n}{2} + \ldots
\]

\[
2^{n-1} = \binom{n}{1} + \binom{n}{3} + \ldots,
\]

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

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 \(\binom{k-1}{n-1}\).

We will make use of the following basic facts from algebra and calculus.
**Theorem 1** (Multinomial theorem). For any positive integer \( k \), we have
\[
(x_1 + \ldots + x_n)^k = \sum_{i_1, i_2, \ldots, i_n \geq 0} \frac{k!}{i_1! \ldots i_n!} x_1^{i_1} \ldots x_n^{i_n},
\]
where, the coefficients \( \frac{k!}{i_1! \ldots i_n!} \) are all positive integers.

**Theorem 2** (Bernoulli’s inequality).
\[
(1 + x)^r \geq 1 + rx
\]
for every real \( r \geq 1 \) and every real number \( x \geq -1 \). If \( r \) is an even integer, then it holds for all real \( x \).

The following generalization of Bernoulli’s inequality when \( r \) is a positive integer is easy to prove by induction on \( r \).

**Theorem 3** (Generalized Bernoulli’s inequality). Let \( x_1, \ldots, x_r > -1 \) be real numbers, all with the same sign. Then
\[
\prod_{i=1}^{r} (1 + x_i) \geq 1 + \sum_{i=1}^{r} x_i.
\]

**Lecture 2. “big oh” and “little oh”, estimates for \( n! \), \( \binom{n}{k} \), Stirling formula**

To read:
[Lov] 2.1. Induction, 2.2. Comparing and estimating umbers, 2.4. Pigeonhole principle
[Mat] 3.4. Estimates: an introduction - starting from 3.4.2. - Big Oh, little oh, 3.5.5. Estimate \( n! \) - second proof only,

**Definition 4.** Let \( f, g : \mathbb{Z}_+ \to \mathbb{R} \). 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 5.** Let \( f, g : \mathbb{Z}_+ \to \mathbb{R} \). We say that \( f \) is little-o of \( g \) and we write \( f(x) = o(g(x)) \) if
\[
\lim_{n \to \infty} \frac{f(n)}{g(n)} = 0.
\]

**Examples:**
- If \( f(n) = o(g(n)) \) then \( f(n) = O(g(n)) \).
- If \( f(n) = O(g(n)) \) and \( g(n) = O(h(n)) \) then \( f(n) = O(h(n)) \) (and if \( g(n) = o(h(n)) \) then \( f(n) = o(h(n)) \))
- \( n + 1 = o(n^2) \) but \( n^2 \) is not \( o(n+1) \)
- \( n^2 + 2 = o(2^n) \), \( n^4 + 5n^3 - 2n - 10 = o(2^n) \)
**Estimating \( n! \)**

Easy observation: for all \( n \geq 1 \) we have

\[
2^n \leq n! \leq n^{n-1}.
\]

Improved bounds:

\[
e \left( \frac{n}{e} \right)^n \leq n! \leq en \left( \frac{n}{e} \right)^n
\]

This proof uses estimates of \( \ln 1 + \ldots \ln n \) using integrals.

**Theorem 6** (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 \( \infty \).

Binomial coefficients:

- can be estimated using Stirling’s formula
- or the binomial theorem: \( \binom{n}{k} < \left( \frac{ne}{k} \right)^k \)
Lecture 3. Inclusion-exclusion, permutations without fixed points, Euler’s $\varphi$ function, introduction to graph theory

Number of permutations without fixed points

To read: [Lov] 2.3. Inclusion-Exclusion

For two finite sets $A_1$ and $A_2$, we have $|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$.

For three finite sets $A_1, A_2, A_3$, we have

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$ 

This works more generally:

**Theorem 7** (Inclusion-Exclusion). Let $A_1, \ldots, A_n$ be finite sets. Then, the following holds

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

The Hatcheck Lady problem

$n$ people receive their hats randomly permuted. What is the probability that no one gets back his own hat?

Let us look at the permutations of the set $\{1, 2, \ldots, n\}$ 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 $\{1, 2, \ldots, n\}$ for which $i$ is a fixed point. The number of permutations with no fixed points is

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

We know that $|A| = n!$, so we need to count $|\bigcup_{i=1}^{n} A_i|$. We do this using the inclusion-exclusion 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

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{i=1}^{n} (n-1)! - \sum_{1 \leq i < j \leq n} (n-2)! + \sum_{1 \leq i < j < k \leq n} (n-3)! - \ldots =$$

$$= \binom{n}{1} (n-1)! - \binom{n}{2} (n-2)! + \binom{n}{3} (n-3)! - \ldots = n!(\sum_{k=0}^{n} (-1)^{k+1} \frac{1}{k!}) \sim n!(1 - \frac{1}{e}).$$
Thus, the number of permutations without fixed points is $\sim n! - n!(1 - \frac{1}{e}) = n!/e$, as $n \to \infty$.

To get the desired probability, we need to divide by the total number of permutations, $n!$, and we conclude that the probability tends to $1/e$.

**Euler’s $\phi$ function**

For every positive integer $n$ we define $\phi(n)$ as the number of positive integers that are relatively prime with $n$. Formally, one writes

$$\phi(n) = \{m \in \{1, 2, \ldots, n\} | \gcd(m, n) = 1\},$$

where $\gcd(m, n)$ denotes the greatest common divisor of $m$ and $n$.

Observe that if $n = p^\alpha$, with $p$ prime, we have

$$\phi(n) = \phi(p^\alpha) = p^\alpha - p^{\alpha-1}.$$

We prove following generalization of this observation using the inclusion-exclusion principle.

**Theorem 8.** Let $n = \prod_{i=1}^m p_i^{\alpha_i}$ be a positive integer with $p_i$ prime factors. Then

$$\phi(n) = n \cdot \prod_{i=1}^m \left(1 - \frac{1}{p_i}\right).$$

**Proof.** For $i \in [m]$ let $A_i := \{x \in [n] : p_i \text{ divides } x\}$. Clearly, $\phi(n) = \left| \bigcap_{i \in [m]} ([n] \setminus A_i) \right|$. For $S \subset [m]$, let $A_S := \bigcap_{i \in S} A_i$. Note that $|A_S| = \frac{n}{\prod_{i \in S} p_i}$. Thus, by the inclusion-exclusion principle, we obtain

$$\phi(n) = \sum_{S \subset [m]} (-1)^{|S|} \frac{n}{\prod_{i \in S} p_i},$$

which is equal to what is stated in the theorem.

**Elements of graph theory**

To read:


**Definition 9.** 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 10.** Let $G = (V, E)$ be a graph. We call a sequence of vertices $v_0, \ldots, v_r$ a **walk**, if $(v_i, v_{i+1})$ is an edge of $G$, for every $0 \leq i \leq r - 1$. A **simple walk** or a **path** is a walk without any repeated vertices.
**Definition 11.** A graph $H$ is a *subgraph* of $G = (V, E)$ (denoted $H \subseteq G$), if $H$ can be obtained from $G$ by deleting vertices and edges.

**Definition 12.** We say that a graph $G = (V, E)$ is *connected*, if for every two vertices $v, u \in V$ there exists a walk in $G$ between $u$ and $v$. Equivalently, $G$ is connected, if for every two vertices $v, u \in V$, there exists a path in $G$ between $u$ and $v$.

**Definition 13.** For every vertex of a graph, we define its *degree* as the number of edges adjacent to it.

The following lemma gives us a relation between the degrees of the vertices of a graph and the total number of edges:

**Lemma 14.** *For every graph $G = (V, E)$, the following holds:*

$$2|E| = \sum_{v \in V} d(v),$$

*where $d(v)$ represents the degree of the vertex $v$.*

**Proof.** The proof is based on a double counting of the total number of edges. On the one hand, we know that this is $|E|$. On the other hand, when counting the sum of degrees of all vertices, we count every edge twice. □

**Definition 15.** A *cycle* in a graph $G = (V, E)$ is a sequence of distinct vertices $v_1, \ldots, v_r \in V$ such that $v_r = v_0$ and $\{v_i, v_{i+1}\} \in E$ for all $i$ from 0 to $r - 1$.

**Definition 16.** A *tree* is a connected graph without cycles.
Lecture 4. Definitions of a tree, counting trees, Prüfer codes

Equivalent definitions of a tree


**Theorem 17.** Let $T$ be a graph. The following properties are equivalent:
1. $T$ is a tree.
2. Any two vertices in $T$ are connected by a unique path.
3. $T$ is minimally connected, that is, it is connected, but after deleting any edge, it becomes disconnected.
4. $T$ is maximally acyclic, that is, it is acyclic, but if we add any edge to $T$, then it will contain a cycle.
5. $T$ has one edge less than the number of vertices and it is connected.
6. $T$ has one edge less than the number of vertices and it is acyclic.

**Definition 18.** A vertex of degree one in a tree is called a leaf.

**Lemma 19.** Every tree on $n \geq 2$ vertices has at least one leaf.

**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 $v_1, \ldots, v_l$ 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 $v_1v_2$ incident to $v_1$. Observe that $u$ cannot coincide with any of the vertices of the path $v_1, \ldots, v_l$ (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 $v_1, \ldots, v_l$. Thus, both $v_1$ and $v_l$ must be leaves. □

**Lemma 20.** If $T$ is a tree and $v$ is a leaf of $T$ then $T - v$ is also a tree. Here $T - v$ denotes the graph obtained from $T$ by deleting $v$ and the edge incident to it.

**Theorem 21.** Every tree on $n$ vertices has exactly $n - 1$ edges.

**Proof.** We do induction on the number of vertices. The statement holds for $n = 1, 2$ so assume that every graph on $n$ vertices contains $n - 1$ edges.

We want to prove that the statement is true for $n + 1$, that is, every tree on $n + 1$ vertices has exactly $n$ edges. Let $T_{n+1} = (V, E)$ be an arbitrary tree on $n + 1$ vertices. By the first lemma above, we know that $T$ contains at least one leaf $v$. We remove $v$ and its unique incident edges from the tree $T_{n+1}$. By the second lemma, we know that this leaves us with a tree $T_n$ on $n$ vertices, which, by the induction hypothesis has exactly $n - 1$ edges. Therefore, $T_{n+1}$ contains exactly $n - 1 + 1 = n$ edges, which completes the proof. □
Counting labeled trees via Prüfer codes

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

Lemma 22. Let $T$ be a tree on $n$ labeled vertices and let $d_1,\ldots,d_n$ 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 23 (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 $\{1,2,\ldots,n\}$. 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 $\{1,2,\ldots,n\}$. 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.

Claim 24. The labels not occurring in the sequence are exactly the leaves of the tree. Moreover, a vertex of degree $d$ occurs exactly $d-1$ times in the sequence.

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,\ldots,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 and the first number in the sequence. This gives a smaller list and a shorter sequence. Then return to step 3.

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 25.** The number of pairwise non-isomorphic trees on \( n \) vertices is at most \( 2^{2n-4} \).

We sketch a proof.

The proof uses the following encoding of trees. We think of a tree hanged from one of its vertices on the 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’re 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 \( 2^{2n-4} \). It is not difficult to see that any such sequence determines a tree.

**Theorem 26.** The number of pairwise non-isomorphic trees on \( n \) vertices is at least \( 2^{cn} \) for some constant \( c > 0 \).

The proof is a simple combination of Cayley’s theorem and the Stirling-formula. There are \( n^{n-2} \) labeled trees. For each one, there are \( n! \) ways to relabel it. Some of these relabelings are, of course, not isomorphisms of the two unlabeled trees, but any isomorphism is a relabeling. Thus, the number of pairwise non-isomorphic trees is \( n^{n-2}/n! \), which, by Stirling’s formula is exponentially large in \( n \).

**Counting trees revisited**

Recall Cayley’s formula:

**Theorem 27** (Cayley). The number of trees on \( n \) labeled vertices is \( n^{n-2} \).

We give another proof of it. It uses induction.

*Proof 2 of Cayley’s theorem.* The proof is based on the following lemma

**Lemma 28.** The number of trees on \( n \) vertices labeled with 1, 2, \ldots, \( n \) with degrees \( d_1, \ldots, d_n \) equals

\[
\frac{(n - 2)!}{(d_1 - 1)! \cdots (d_n - 1)!}.
\]

[Two proofs for the lemma: one using Prüfer codes and another using induction.]

By the lemma, we obtain that the total number of trees on \( n \) labeled vertices is the sum of the number of trees over all possible values of degrees, that is

\[
\sum_{\substack{d_1, \ldots, d_n \geq 1 \\ d_1 + \cdots + d_n = 2n-2}} \frac{(n - 2)!}{(d_1 - 1)! \cdots (d_n - 1)!}.
\]

Let now \( r_i = d_i - 1 \), for all \( 1 \leq i \leq n \). By substitution, we obtain that the number of trees on \( n \) labeled vertices is

\[
\sum_{\substack{r_1, \ldots, r_n \geq 1 \\ r_1 + \cdots + r_n = n-2}} \frac{(n - 2)!}{(r_1)! \cdots (r_n)!}.
\]
which, by the multinomial theorem equals $n^{n-2}$.

[Lov] 8.3. How to Count trees? 8.4. How to Store trees?
[Mat] 5.1 Definition and characterizations of trees 8.1. The number of spanning trees, 8.4. A proof using the Prüfer codes.
Lecture 5. Minimum weight spanning trees: Kruskal’s algorithm

Weighted spanning trees

Definition 29. A weighted graph is a graph and a mapping of the set of edges to the real numbers, that is, a real number (weight) is assigned to each edge. The weight of a weighted graph is the sum of the weights of all its edges.

Definition 30. 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:

Our problem is the following. Given weighted connected graph $G$. We want to find a minimum weight spanning tree $T$ of $G$.

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 haven’t 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 31. For any connected weighted graph $G$, Kruskal’s algorithm outputs a spanning tree of $G$, and this tree has minimal weight among all spanning trees of $G$.

[Lov] 8.3. How to Count trees?

Lecture 6. Partial orders, Dilworth’s theorem

Definition 32. A partially ordered set (or simply poset) is a pair \((X, \preceq)\), where \(X\) is a set and \(\preceq\) is a binary relation over \(X\), which is reflexive, antisymmetric, and transitive, i.e., which satisfies the following relations, for all \(a, b,\) and \(c\) in \(X\):

a. \(a \preceq a\) (reflexivity);

b. if \(a \preceq b\) and \(b \preceq a\) then \(a = b\) (antisymmetry);

c. if \(a \preceq b\) and \(b \preceq c\) then \(a \preceq c\) (transitivity).

We write \(a < b\), if \(a \preceq b\) and \(a \neq b\).

Definition 33. Let \((X, \preceq)\) be a partially ordered set.

A chain in \(X\) is a sequence \(x_1, \ldots, x_t \in X\) such that
\[ x_1 \prec x_2 \prec \ldots \prec x_t. \]

An antichain in \(X\) is a subset \(\{x_1, \ldots, x_l\}\) of \(X\) such that no two elements are comparable.

Recall, that for a set \(Z\), \(2^Z\) denotes the power set of \(Z\), that is the set of all subsets of \(X\).

Example. Consider the partially ordered set \((2^{\{1,2,3\}}, \subseteq)\).

The sequence \(\emptyset \subset \{1\} \subset \{1, 2, 3\}\) is a chain.
\(\{1, 2\}, \{1, 3\}\) is an antichain.

A good way to visualize small partially ordered sets is to draw their Hasse diagram: one represents each element of \(X\) as a vertex in the plane and draws a line segment or curve that goes upward from \(x\) to \(y\) whenever \(x \prec y\) and there is no \(z\) such that \(x \prec z \prec y\).

Theorem 34 (Dilworth). Let \((X, \preceq)\) be a partially ordered set.

a. If the maximum size of an antichain is \(k\), then \(X\) can be decomposed into \(k\) chains.

b. If the maximum size of a chain is \(k\), then \(X\) can be decomposed into \(k\) antichains.

Proof. The proof can be found in [Juk], page 108.

First, we prove the easier part, b. We define the height of an element \(x\) as the length of longest chain whose maximum element is \(x\). For \(i \in [k]\), let \(A_i := \{x \in X : \text{height}(x) = i\}\). It is easy to see that each \(A_i\) is an antichain, and that their union is \(X\).

To prove part a., we use induction on the cardinality of \(X\). Let \(a\) be a maximal element of \(X\), and let \(k\) be the size of a largest antichain in \(X' = X \setminus \{a\}\). Then \(X'\) is the union of \(k\) disjoint chains \(C_1, \ldots, C_k\). We have to show that \(X\) either contains an \((k + 1)\)-element antichain or else is the union of \(k\) disjoint chains. Now, every \(k\)-element antichain in \(X'\) consists of one element from each \(C_i\). Let \(a_i\) be the maximal element in \(C_i\) which belongs to some \(k\)-element antichain in \(X'\). It is easy to see that \(A = \{a_1, \ldots, a_k\}\) is an antichain in \(X'\). If \(A \cup \{a\}\) is an antichain in \(X\), we are done: we have found an antichain of size \(k + 1\). Otherwise, we have \(a > a_i\) for some \(i\). Then \(K = \{a\} \cup \{x \in C_i : x \preceq a_i\}\) is a chain in \(X\), and there are no \(k\)-element antichains in \(X \setminus K\) (since \(A_i\) was the maximal element of \(C_i\) participating in such an antichain), whence \(X \setminus K\) is the union of \(k - 1\) chains. \(\Box\)
Corollary 35. Let \((X, \preceq)\) be a partially ordered set. Then, the following hold:

a. The maximum size of an chain is equal to the minimum number of antichains that cover \(X\).

b. The maximum size of an antichain is equal to the minimum number of chains that cover \(X\).

Refer as well to the following:
Lecture 7. König-Hall theorem, LYM inequality, Sperner’s theorem

Definition 36. A bipartite graph 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 37. A graph is bipartite if, and only if, it does not contain an odd cycle (that is, a cycle of odd length).

Definition 38. Let $G = (V, E)$ be a graph. A subset $E' \subseteq E$ of pairwise disjoint edges (that is, edges which do not share any vertex) is called a matching in $G$.

Definition 39. A perfect matching is a matching where every vertex of the graph is incident to exactly one edge of the matching.

Remark 40. 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!

Example 41. The blue edges in the following graph form a matching: The matching is not a perfect matching though.

What obstacles prevent the existence of a perfect matching in a bipartite graph $G(A \cup B, E)$?

Example 42. If a vertex set $X$ in $A$ has fewer than $|X|$ neighbors in $B$ then $X$ cannot be covered by any matching in $G$. In particular, if such a set $X$ exists then $G$ has no perfect matching.

Interestingly, the converse is also true:

Theorem 43 (König-Hall). Let $G = (V, E)$ be a bipartite graph with bipartition $V = A \cup B$. For every $X \subseteq A$, let

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

Then there is a matching $M$ in $G$ such that every vertex of $A$ belongs to some edge in $M$ if and only if
\[ |B(X)| \geq |X|, \text{ for all } X \subseteq A \quad (\text{Hall's condition}). \]

**Note:** On the lecture, this theorem was proved in the special case when \(|A| = |B|\). But the same proof, as you will see, yields this more general statement as well.

**Proof.** We will deduce this from Dilworth’s theorem. Define the poset \((P, \prec)\) where \(P = V\) is the vertex set of the graph and \(u \prec v\) if \(u \in A, v \in B\) and \(uv \in E\). By Dilworth’s theorem, we know that the maximum size of an antichain is equal to the maximum number of chains that \(P\) can be partitioned into.

**Claim 44.** The maximum size of an antichain is \(|B|\).

**Proof.** Let \(D\) be an antichain in \(P\). Then we know that \(X = D \cap A\) satisfies the condition (Hall’s condition) above, so \(|B(X)| \geq |X|\). As \(D\) is an antichain, we also know that no vertex of \(B(X)\) appears in \(D\), and hence \(B(X)\) and \(D \cap B\) are disjoint. But then

\[ |B| \geq |B(X)| + |D \cap B| \geq |X| + |D \cap B| = |D \cap A| + |D \cap B| = |D| \]

which is what we wanted to show.

By Dilworth’s theorem, \(P\) can be partitioned into \(|B|\) chains. The following observation is easy to see from the definition of our poset:

**Claim 45.** There is no chain longer than 2.

Thus each chain is either an edge or a vertex.

Let us take a partition of the vertices into \(|B|\) chains. We know that no chain can contain two vertices from \(B\), so each vertex in \(B\) has a separate chain containing it. This means that at least \(|B|\) (that is: all) of our chains contain a vertex from \(B\). In particular, there is no chain that only contains a single vertex of \(A\). But then the chains containing the vertices of \(A\) form a matching that we were looking for.

**Theorem 46** (Sperner). Let \(X = \{1, 2, \ldots, n\}\) and \(A_1, \ldots, A_m \subseteq X\), with \(A_j \not\subset A_i\), for all \(i \neq j\). Then \(m \leq \binom{n}{\lfloor n/2 \rfloor}\).

Theorem 46 can be easily deduced from the *Lubell-Yamamoto-Meshalkin* inequality, more commonly known as the *LYM inequality*.

**Theorem 47** (LYM inequality). Let \(X = \{1, 2, \ldots, n\}\), and let \(\mathcal{F}\) be a family of subsets \(A_1, \ldots, A_m \subseteq X\) such that \(A_j \not\subset A_i\), for all \(i \neq j\), that is, \(\mathcal{F}\) is an antichain. Let \(m_k = |\{A \in \mathcal{F} : |A| = k\}|\), that is, the number of sets in \(\mathcal{F}\) containing \(k\) elements. (One can see that \(m_1 + \ldots + m_n = m\). Then, the following holds

\[ \sum_{i=0}^{n} \frac{m_i}{\binom{n}{i}} \leq 1. \]
Proof of the LYM inequality. Notice that in the power set $\mathcal{P}(X)$ of $X$, the number of maximal chains is $n!$. Moreover, for each subset $A$, exactly $|A|!(n - |A|)!$ maximal chains contain $A$ (convince yourself!).

Since none of the $n!$ maximal chains meet $\mathcal{F}$ more than once, we have $\sum_{A \in \mathcal{F}} |A|!(n - |A|)! \leq n!$. Dividing this inequality by $n!$ yields the desired result. \hfill \square

Proof of Sperner’s theorem. Recall that for fixed $n$, $\binom{n}{k}$ is maximal when $k = \lfloor \frac{n}{2} \rfloor$. The theorem now easily follows. \hfill \square

Refer as well to the following:

Lecture 8. Sperner’s theorem proved with symmetric chains, Erdős-Ko-Rado Theorem

Second proof of Sperner’s theorem. Define the poset on the power set $2^X$ of $X$ ($2^X$ contains all subsets of $X$), where $A \prec B$ is $A \subset B$. By Dilworth’ theorem, it is enough to decompose this poset into ${n \choose \lfloor n/2 \rfloor}$ chains. In fact, we will partition $2^X$ into symmetric chains, that is, maximal chains connecting level $k$ of $2^X$ (that is, the family of size $k$ sets) with level $n-k$ of $2^X$ for some $k$.

The construction is based on induction as follows: For $n = 0$ the poset has 1 element, so we can cover it by $0$ chains. Now suppose we have a chain decomposition for some $n$. Then if for each of our chains of the form $A_k \subset A_{k+1} \subset \cdots \subset A_{n-k}$ we add the chain $A_k \cup \{n+1\} \subset A_{k+1} \cup \{n+1\} \subset \cdots \subset A_{n-k} \cup \{n+1\}$ to our partition, then we get a chain decomposition for $n+1$.

However, these chains are not symmetric for $n+1$. So instead we move $A_k$ to the front of the other chain. Then $A_{k+1} \subset \cdots \subset A_{n-k}$ and $A_k \subset A_{k+1} \cup \{n+1\} \subset \cdots \subset A_{n-k} \cup \{n+1\}$ are symmetric, and they cover the same sets, so they form a symmetric chain decomposition, proving the theorem.

Note that some of our new chains can become empty and so are not really chains. This happens if $n$ is even and $k = n/2$.

Definition 48. Let $X$ be a set and $F$ be a family of subsets of $X$, that is $F \subseteq 2^X$. We say that $F$ is intersecting, if any two members of $F$ intersect.

Lemma 49. Let $X$ be a finite set with $|X| = n$ and $F \subseteq 2^X$ an intersecting family. Then $|F| \leq 2^{n-1}$ and this bound is sharp (that is, there exists an intersecting family of size $2^{n-1}$).

Proof. In total, $X$ has $2^n$ subsets, which can be arranged in pairs of sets that are complements of each other. We have $2^{n-1}$ such pairs. Out of each pair, $F$ may contain at most one.

To see that the bound is sharp, take an arbitrary element of $X$, and consider all subsets of $X$ containing this element.

Theorem 50 (Erdős-Ko-Rado). Let $k \leq \frac{n}{2}$, $X$ be a set with $|X| = n$ and $F$ an intersecting family of $k$-element subsets of $X$. Then

$$|F| \leq \binom{n - 1}{k - 1}.$$

Proof. The proof can be found in [Juk], pages 100-101.

Remark. The theorem above does not hold in the case when $k > n/2$. Indeed, consider the family of all $k$-element subsets of an $n$-element set $X$. This will be intersecting by the pigeonhole principle. The cardinality of this family is $\binom{n}{k}$, which is larger than $\binom{n-1}{k-1}$.

Corollary 51 (Generalized Erdős-Ko-Rado). Let $F = A_1, \ldots, A_m \subseteq \{1, \ldots, n\}$ be an intersecting set family with $|A_i| \leq k$ for every $i$, and suppose $F$ is a Sperner family, i.e., $A_i \nsubseteq A_j$ for every $i \neq j$. Then $m \leq \binom{n-1}{k-1}$.
Proof. The $A_i$ correspond to elements of the Boolean lattice (the poset of $2^{[n]}$). We can assume that there are sets in $\mathcal{F}$ strictly below the $k$'th level. Let $l$ be the lowest level that contains such a set. Levels $l$ and $l + 1$ induce a bipartite graph that has a matching that contains every vertex on the $l$'th level (see also the second proof of Sperner’s theorem in Lecture 8). Note that no two elements of $\mathcal{F}$ are connected by an edge (because it is a Sperner family). So we can replace each element of $\mathcal{F}$ on the $l$'th level with its neighbor in the matching, and the resulting family will still satisfy all our conditions. 

Also refer to the following:

- [Mat] 7.2. Sperner’s theorem on independent systems: Sperner theorem and proof of Theorem 7.2.1.
- [Juk] 7.2 Erdős-Ko-Rado theorem.
Lecture 9. The Littlewood-Offord problem; A strengthening of the Erdős-Ko-Rado theorem

The Littlewood-Offord problem:

**Theorem 52.** Let $a_1, \ldots, a_n \geq 1$ be fixed real numbers and

$$A = \left\{ \sum_{i=1}^{n} \epsilon_i a_i, \ \epsilon_i = -1 \text{ or } 1 \right\}.$$  

Then, for any length one interval $I$, we have

$$|A \cap I| \leq \left( \binom{n}{\lfloor n/2 \rfloor} \right).$$

**Proof.** For every sum $\sum_{i=0}^{n} \epsilon_i a_i$, we assign a unique “characteristic” set, $\{ i \mid \epsilon_i = 1 \} \subseteq \{1, 2, \ldots, n\}$. One can easily see that this is an injection.

We prove that the characteristic sets of sums contained in the same interval satisfy the condition in Sperner’s theorem, that is no one is contained in another. We assume the contrary, that is, there exist two characteristic sets $B, B'$ with $B \subseteq B'$, $B \neq B'$ such that the corresponding sums lie in the interval $I$

$$\sum_{i \in B} \epsilon_i a_i \in I \text{ and } \sum_{i \in B'} \epsilon_i a_i \in I.$$

Since the length of $I$ is less than 2, their difference has to be less than two. On the other hand, since $B \subseteq B'$ and $B \neq B'$, we obtain that

$$\sum_{i \in B'} \epsilon_i a_i - \sum_{i \in B} \epsilon_i a_i = \sum_{i \in B' \setminus B} 2a_i < 2,$$

where the sum is a sum of non-negative numbers, containing at least one non-zero term. On the other hand, since every $a_i \geq 1$, we obtain that the sum must be at least two, which is a contradiction.

Therefore, the condition of Sperner’s theorem is satisfied, so we obtain that the number of sums inside the interval $I$ cannot exceed $\left( \binom{n}{\lfloor n/2 \rfloor} \right)$, which completes the proof. 

We prove the following variation of the Erdős-Ko-Rado theorem.

**Theorem 53 (Erdős-Ko-Rado).** Let $k \leq \frac{n}{2}$, $X$ be a set with $|X| = n$ and $\mathcal{F}$ an intersecting family of subsets of $X$, each of size at most $k$. Assume that $\mathcal{F}$ is an antichain in $2^X$, that is, no member of $\mathcal{F}$ contains another member of $\mathcal{F}$. Then

$$|\mathcal{F}| \leq \binom{n-1}{k-1}.$$  

**Proof.** We can “move up” each member of $\mathcal{F}$ to become a $k$-element set using exercise 8-4, and then, apply the Erdős-Ko-Rado theorem. 


Lecture 10. The probabilistic method

Basics of Probability theory.

We define a finite probability space. Let $\Omega$ be a finite set, which we call the set of elementary events. Any subset $A \subset \Omega$ is called an event. Let $p$ be

$$p : 2^\Omega \rightarrow [0, 1]$$

a function from the family of subsets of $\Omega$ to the $[0, 1]$ interval with the following properties.

1. $p(\emptyset) = 0$,
2. $p(\Omega) = 1$,
3. $p(A \cup B) = p(A) + p(B)$ for any $A, B$ with $A \cap B = \emptyset$, that is, when $A$ and $B$ are disjoint events. Now, the pair $(\Omega, p)$ is called a discrete probability space.

A random variable is a function $X : \Omega \rightarrow \mathbb{R}$. Note that the measure $p$ does not appear in this definition.

If $X$ takes values $x_1, \ldots, x_k$, then the expectation (or, expected value) $E(X)$ of $X$ is defined as

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

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

Two events $A, B \subseteq \Omega$ are said to be independent if $p(A \cap B) = p(A)p(B)$. Events $A_1, \ldots, A_k$ are said to be independent if $p(\cap_{i \in I} A_i) = \prod_{i \in I} p(A_i)$ for any set $I \subseteq [k]$.

A set $\{X_1, \ldots, X_k\}$ of discrete random variables is said to be independent if the events $X_i = a_i$ are independent for any choice of the $a_i$’s.

Some useful properties:

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

$$P(A_1 \cup \ldots \cup A_n) \leq P(A_1) + \ldots + P(A_n).$$

b. If $A_1, \ldots, A_n \subseteq \Omega$ are independent events, then

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdot \ldots \cdot P(A_n).$$

c. The linearity of expectation: Let $X_1, \ldots, X_n$ be random variables (not necessary independent), and $a_1, \ldots, a_n \in \mathbb{R}$. Then

$$E(a_1 X_1 + \ldots + a_n X_n) = a_1 E[X_1] + \ldots + a_n E[X_n].$$

d. If $E[X] = m$, 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 $X(A_2) \leq m$.

Probably, the simplest form of the probabilistic method is the following. We are given a finite set $\Omega$ and $X : \Omega \rightarrow \mathbb{R}$ is a function assigning to each object $\omega \in \Omega$ a real number. The goal is to show that there is at least one element $\omega \in \Omega$ for which $X(\omega)$ is at least a
given value \( m \). For this, we define a probability distribution \( P : \Omega \rightarrow [0,1] \), and consider the resulting probability space, where \( X \) becomes a random variable. If we can show that the expected value of \( X \) is at least \( m \), then we proved that there is an \( \omega \in \Omega \) for which \( X(\omega) \geq m \).

The proof of the following theorem demonstrates this idea.

**Theorem 54.** 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 [Mat], page 307.

We translate the beginning of the proof to the language introduced above. Now, \( \Omega \) is the set of all partitions of \( V \) into two \( n \)-element subsets. For any partition \( \omega \), let \( X(\omega) \) denote the number of edges going between the two parts of \( \omega \).

The probability \( p \) on \( \Omega \) is simply the uniform probability, that is, each \( \omega \in \Omega \) has the same probability. Since \( |\Omega| = \binom{2n}{n} \), that means, that \( p(\omega) = 1/\binom{2n}{n} \) for each \( \omega \in \Omega \).

The interesting part of the proof is now to show that \( \mathbb{E}[X] \geq m/2 \), see the book. \( \Box \)

We will need one more notion from probability: Let \( B \) be an event with \( p(B) > 0 \). The **conditional probability** of an event \( A \) conditioned on event \( B \) is defined as

\[
\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}.
\]

**Proof of LYM inequality using the probabilistic method.**

Let \( \mathcal{F} \subseteq 2^X \) be a family of subsets of \( X \), where \( |X| = n \), such that \( A \nsubseteq B, \forall A, B \in \mathcal{F}, A \neq B \). Let \( m_k = |\{A \in \mathcal{F} : |A| = k\}| \), that is the number of sets in \( \mathcal{F} \) containing \( k \) elements. (One can see that \( m_1 + \ldots + m_n = m \). We want to prove that

\[
\sum_{i=0}^{n} \frac{m_i}{\binom{n}{i}} \leq 1.
\]

Recall that \( (2^X, \subseteq) \) is a poset.

First, let us observe that the number of maximal chains is \( n! \). This is because a maximum length chain has in it exactly one \( k \)-element set for each \( 0 \leq k \leq n \). Assume you start from the empty set. There are \( n \) ways to continue the chain to a 1-element set (think of the number of ways in which we can add one element). For each one-element set, there are \( n - 1 \) ways to continue the chain to a 2-element set, and so on.

We choose randomly with equal probability \( 1/n! \) one such maximum chain and we define a random variable \( Y \) as follows

\[
Y = \sum_{F \in \mathcal{F}} Y_F, \text{ where } Y_F = \begin{cases} 1 \text{ if } F \text{ is in the chain} \\ 0 \text{ otherwise} \end{cases}.
\]

We know that \( \mathbb{E}(Y_F) = P(Y_F = 1) = \frac{|F|!(n-|F|)!}{n!} = 1/\binom{n}{|F|} \). In order to see this, we just count the number of chains containing \( F \). Below \( F \), there are \( |F|! \) possibilities, while
above $F$, there are $(n - |F|)!$ possibilities. On the other hand, the total number of chains is $n!$, which gives the statement above.

By the linearity of expectation, we obtain

$$
\mathbb{E}(Y) = \sum_{F \in \mathcal{F}} \mathbb{E}(Y_F) = \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}}.
$$

On the other hand, a chain can contain at most one element of $\mathcal{F}$, so $\mathbb{E}(Y) \leq 1$. Recalling that $m_i$ is the number of $i$-element subsets in $\mathcal{F}$, we obtain that

$$
\mathbb{E}(Y) = \sum_{i=0}^{n} \frac{m_i}{\binom{n}{i}} \leq 1.
$$

Also refer to the following:


We sketch a probabilistic proof of the Erdős-Ko-Rado theorem (Theorem 50). As in the non-probabilistic proof, we use the following observation. For any permutation $\pi$ on $[n]$, the number of sets $F \in \mathcal{F}$ of the form $\{\pi(s + 1), \pi(s + 1), \ldots, \pi(s + k)\}$, where $s \in [n]$ and indices are taken modulo $n$, is at most $k$.

Now, let $\pi$ be a random permutation of $[n]$, that is, choose one of the $n!$ permutations, each one with the same probability. Choose an integer $s \in [n]$, again uniformly. By the observation above, the probability that the set $\{\pi(s + 1), \pi(s + 1), \ldots, \pi(s + k)\}$ belongs to the family $\mathcal{F}$ is at most $\frac{k}{n}$. On the other hand, this probability is clearly the same as the probability of a randomly selected $k$-element set belongs to $\mathcal{F}$, which is $|\mathcal{F}|/\binom{n}{k}$, finishing the proof.

**Theorem 55.** Every graph $G = (V,E)$ contains a bipartite subgraph with at least $|E|/2$ edges.

**Proof.** The proof can be found in [Mat], page 307.

**Schütte’s problem: Tournaments**

**Definition 56.** 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 57.** We say that a tournament $T = (V(T),E(T))$ has property $S_k$ if for any $k$ vertices $v_1, \ldots, v_k \in V(T)$, there exist a vertex $u \in V(T)$ such that $\vec{uv}_1, \vec{uv}_2, \ldots, \vec{uv}_k \in E(T)$.

Schütte’s problem can be formulated as follows: do such tournaments exist, for every $k$ fixed? A probabilistic argument shows that such tournaments exist. We describe it below.

**Random graphs**

The random graph $G(n,p)$ is a graph on $n$ vertices in which any two vertices are connected by an edge with probability $p$ independently of others. Thus, one may consider a random graph a probability space defined on the set of all labeled graphs on $n$ vertices, in which every graph is assigned a probability which depends on the number of edges in it.

$$\Pr(\text{every pair of vertices is connected}) = p^{\binom{n}{2}}.$$ 

We can define a random tournament $T_{n,1/2}$ in a similar way: each edge of the complete graph $K_n$ is oriented in one direction or the other with equal probabilities.
**Definition 58.** We say that a tournament \( T = (V(T), E(T)) \) has property \( S_k \) if for any \( k \) vertices \( v_1, \ldots, v_k \in V(T) \), there exist a vertex \( u \in V(T) \) such that \( \overrightarrow{uv_1}, \overrightarrow{uv_2}, \ldots, \overrightarrow{uv_k} \in E(T) \).

**Theorem 59.** \( \Pr(T_{n,1/2} \text{ has property } S_k) \to 1 \) as \( n \to \infty \).

Also refer to the following:


Generating functions

Example 1. Consider the following combinatorial problem: How many ways are there to pay the amount of 21 francs if we have 6 one-francs coins, 5 two-francs coins, and 4 five-francs coins?

The required number is the number of solutions of the equation

\[ x_1 + x_2 + x_3 = 21, \]

with \( x_1 \in \{0, 1, 2, 3, 4, 5, 6\} \), \( x_2 \in \{0, 2, 4, 6, 8, 10\} \), and \( x_3 \in \{0, 5, 10, 15, 20\} \).

In order to compute this, we associate to each variable \( x_i \) above a polynomial \( p_i \) as follows:

\[
p_1(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6, \quad p_2(x) = 1 + x^2 + x^4 + x^6 + x^8 + x^{10},
\]

\[
p_3(x) = 1 + x^5 + x^{10} + x^{15} + x^{20}.
\]

The number of solutions of the equation above will be the coefficient of \( x^{21} \) in the product \( p_1(x) \cdot p_2(x) \cdot p_3(x) \).

Example 2. We prove that \( \sum_{k=0}^{n} k \binom{n}{k} = n \cdot 2^{n-1} \). By the binomial theorem, we know that \( (1 + x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \).

Consider the derivative of this polynomial. This gives

\[
n(1 + x)^{n-1} = \sum_{k=1}^{n} \binom{n}{k} k x^{k-1}.
\]

Considering \( x = 1 \) in the above, we obtain the desired identity.

Example 3. We prove that

\[
\sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right)^2 = \left( \begin{array}{c} 2n \\ n \end{array} \right).
\]

We know that \( (1 + x)^n (1 + x)^n = (1 + x)^{2n} \). Consider the coefficient of \( x^n \) in this expression. On the one hand, it is \( \binom{2n}{n} \). On the other hand, from the left, we obtain that this coefficient is \( \sum_{k=0}^{n} \binom{n}{k} \binom{n}{n-k} \), and since \( \binom{n}{k} = \binom{n}{n-k} \), we obtain the desired identity.

Theorem 60. 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 \( |x| < 1/c \).
Definition 61. Let \((a_0, a_1, \ldots)\) be a sequence of real numbers. Then, its generating function \(a(x)\) is
\[
a(x) = a_0 + a_1 x + a_2 x^2 + \ldots
\]

Example 1. The generating function of the sequence \((0, 1/2, 1/3, \ldots)\) is
\[
a(x) = 0 + x + \frac{1}{2} x^2 + \frac{1}{3} x^3 + \ldots = -\ln(1 - x).
\]

We can compute the result of certain operations on generating functions. Let \(a(x)\) and \(b(x)\) be generationg functions of the sequences \(a_0, a_1, \ldots\) and \(b_0, b_1, \ldots\).

- The generating function of \(a_0 + b_0, a_1 + b_1, \ldots\) is simply \(a(x) + b(x)\).
- The generating function of sequence \(\alpha a_0, \alpha a_1, \ldots\) is \(\alpha a(x)\), for any real number \(\alpha\).
- The generating function of sequence \(a_1, 2a_2, 3a_3, \ldots\) is the derivative, \(a'(x)\).
- The generating function of sequence \(0, a_0, a_1, 2a_2, 3a_3, \ldots\) is the primitive function, \(x \mapsto \int_0^x a(t) dt\).

Example 2. The generating function of the sequence \((1, 1/2, 1/3, \ldots)\) is
\[
a(x) = 1 + x + \frac{1}{2!} x^2 + \frac{1}{3!} x^3 + \ldots = e^x.
\]

Theorem 62 (Generalized binomial theorem). For every \(r \in \mathbb{R}\) and every integer \(k \geq 0\), let
\[
\binom{r}{k} = \frac{r(r-1) \ldots (r-k+1)}{k!}.
\]
Then, the following holds:
\[
(1 + x)^r = \binom{r}{0} + \binom{r}{1} x + \binom{r}{2} x^2 + \ldots
\]
for every \(x\) with \(|x| < 1\).

Lecture 13. Generating functions, Fibonacci’s sequence

Example 1. Let \(a_{i-1} = i\). Then the generating function \(A(x) = 1 + 2x + 3x^2 + \ldots = (1 + x + x^2 + \ldots)' = (\frac{1}{1-x})' = \frac{1}{(1-x)^2}\).

Example 2. Let \(b_{i-1} = i^2\). Arguing in a similar way, one gets that \(B(x) = A'(x) - A(x)\).

Example 3. Combinatorial application. Given 10 red balls, 20 blue balls and 30 green balls, in how many ways we can choose 40 balls out of it? It is a coefficient in front of \(x^{40}\) in the polynomial
\[
p(x) = (1 + x + x^2 + \ldots + x^{10})(1 + x + x^2 + \ldots + x^{20})(1 + x + x^2 + \ldots + x^{30})
\]. Using geometric series formula, we rewrite it in the form
\[
p(x) = \frac{1 - x^{11}}{1-x} \frac{1 - x^{21}}{1-x} \frac{1 - x^{31}}{1-x} = \frac{1}{(1-x)^3}(1 - x^{11} - x^{21} - x^{31} + x^{42} + \ldots)
\]
and, using the generalized binomial formula for \((1 - x)^{-3}\), find the coefficient in front of \(x^{40}\).

**Fibonacci sequence**

The Fibonacci sequence \((F_n)_{n\geq0}\) is defined by the following recursive formula:

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

One 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.

Note that by an easy proof using induction on \(n\), we obtain that the sum of the first \(n\) members of the Fibonacci sequence, is

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

We want to find an **explicit formula** for the value of the \(n\)-th Fibonacci number. We present two ways to find it.

**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

\[
F(x) = F_0 + F_1 x + F_2 x^2 + F_3 x^3 + \ldots
\]

Multiplying \(F(x)\) by \(x\), respectively \(x^2\), we obtain that

\[
x F(x) = F_0 x + F_1 x^2 + F_2 x^3 + F_3 x^4 + \ldots
\]

\[
x^2 F(x) = F_0 x^2 + F_1 x^3 + F_2 x^4 + F_3 x^5 + \ldots
\]

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

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

This implies \(F(x) - x F(x) - x^2 F(x) = x\) 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 \(F(X)\). Finding these derivatives directly is not easy. We will use a trick from analysis: **partial fraction decomposition**.

We factor \(1 - x - x^2\) as \(- (x - x_1)(x - x_2)\), where \(x_{1,2} = \frac{-1 \pm \sqrt{1 + 4}}{2} = \frac{-1 \pm \sqrt{5}}{2}\).

This means

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

28
for appropriately chosen $A$ and $B$.

From this we obtain that

$$A + B = -1 \text{ and } Ax_2 + Bx_1 = 0.$$  \hspace{1cm} (1)

This is a system of two equations with $A$ and $B$ as unknowns, so we can obtain exact values for $A$ and $B$. However, we prefer writing these fractions in a nicer form, for which we know the Taylor series expansion.

$$\frac{A}{x - x_1} = \frac{-A/x_1}{1 - x/x_1} = -\frac{A}{x_1} \sum_{n=0}^{\infty} \left(\frac{x}{x_1}\right)^n,$$

and

$$\frac{B}{x - x_2} = \frac{-B/x_2}{1 - x/x_2} = -\frac{B}{x_2} \sum_{n=0}^{\infty} \left(\frac{x}{x_2}\right)^n.$$

Note that $x_1x_2 = -1$. Finally, using the solution of (1), and the fact that $F_0 = 0$, $F_1 = 1$, we obtain

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

**Method 2.**

We look first for a geometric sequence that satisfies $F_n = F_{n-1} + F_{n-2}$, that is $F_n = c \cdot \alpha^n, \forall n$.

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 $\alpha_{1,2} = \frac{1 \pm \sqrt{5}}{2}$. Clearly, both $F_n = c\alpha_1^n$, nor $F_n = c\alpha_2^n$ sequence satisfy the recurrence relation, but they do not satisfy the initial value conditions, $F_0 = 0$, $F_1 = 1$, no matter how we choose $c$.

On the other hand, if two sequences satisfy the recurrence relation, then so do their linear combinations.

Thus, we write

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

and we know that it also satisfies $F_n = F_{n-1} + F_{n-2}$ for any $c_1, c_2 \in \mathbb{R}$.

To obtain the values of $c_1$ and $c_2$, we substitute $n = 0$ and $n = 1$.

$$0 = F_0 = c_1 + c_2, \quad 1 = F_1 = c_1 \frac{1 + \sqrt{5}}{2} + c_2 \frac{1 - \sqrt{5}}{2},$$

and obtain $c_1 = 1/\sqrt{5}$, $c_2 = -1/\sqrt{5}$.

We did not learn it on the lecture, but we note the following. To solve a linear recurrence relations of the form

$$a_{n+k} = c_{k-1}a_{n+k-1} + \ldots + c_0a_n,$$

where $c_0, \ldots, c_{k-1}$ are constants, we can use the method described in Section 12.3 in [Mat].

29