CIT413068 Summer 26

Combinatorics

Lecturer: Prof. Dr. Violetta Weger

These lecture notes will be consistently updated before the lectures. If you find any typos, please send them to me via email.

Overview Many of you might wonder what combinatorics exactly is. Unfortunately, to give a precise and formal definition of combinatorics seems to be an intractable problem. To cite Tutte;

> *It is difficult to find a definition of combinatorics that is both concise and complete, unless we are satisfied with the statement "Combinatorics is what combinatorialists do."*
> William Thomas Tutte

However, all mathematicians will recognize combinatorics when they see it. If we want to give a very vague explanation, we could say that combinatorics studies the ways elements of a finite set can be arranged, typically under certain constraints imposed by the problem. Usually we either ask whether certain arrangements or objects exist, or to enumerate these objects. Whether this explanation does the field justice is questionable.

What is more agreed upon, is that the name combinatorics originates form Leibniz book "Ars Combinatoria" and that combinatorics lies at the heart of discrete mathematics and applied algebra. Personally, my goal of giving this course is to give you the right tools to *count* certain objects, *model* them and recognize *patterns* which make handling these objects easier.

There exist many directions within combinatorics: we might be interested in analytical, probabilistic, extremal, enumerative combinatorics or what we might call algebraic combinatorics.

In this lecture we will focus on the latter two. We want to provide the most essential tools for counting and modeling problems and then move on to the incidence structures studied within combinatorics. Hence we will study binomials, a few tricks for how to handle them, partitions, compositions and some graph theory. In terms of incidence structures, we will study designs, association schemes and codes - which is where I come from.

This course is complementary to the lecture "Discrete structures", where several of the topics we will study appear as well. The focus therein is in extremal combinatorics, which of course requires first the elementary tools of combinatorics.

Most of the content of these lecture notes is based on the wonderful book *A course in combinatorics,* by J.H. van Lint and R.M. Wilson [3].

Administrative Information

The lectures will be held

- Tuesdays, 16:15-17:45 in MI 00.07.014
- Wednesdays, 14:15-15:45 in MI 00.07.014

The tutorials will be held

- Thursdays, 16:15-17:45 in MI 00.09.022
- Fridays, 12:15-13:45 in MI 02.08.011 until May 15
- Wednesday, 16:15-17:45 in MI 03.08.011 from May 20 onwards

Exercises: I will upload exercise sheets on Moodle, which are voluntary to solve, but upon 70% of correct completion, I will provide a grade bonus of $+0.3$.

Exam: The exam will either be in written (120 min) or oral (30 min) form, depending on the number of students.

Prerequisites:

- MA0004 Linear Algebra 1,
- MA0005 Linear Algebra 2 and Discrete Structures.

Material: Most of the content of these lecture notes is based on

- the book *A course in combinatorics* by van Lint and Wilson [3], available online `https://tinyurl.com/lintcourse`,
- the seminal books by R. Stanley: *Enumerative Combinatorics*, volume 1 (and possibly also 2), [2] available online `https://tinyurl.com/enumerativestanley`.

Schedule

In green we marked the holidays, or days with no lecture, and in red the days I am not here. On these days we will either do a "flipped classroom", where you will be provided with video lectures to watch and we can discuss questions at the next in-person event or there will be a substitute lecturer. The dates in orange are dates where we have to switch lecture and tutorial, that is: in the week of April 20- April 24, the lecture is held on Thursday April 23, and the tutorial on Wednesday April 22. Similarly, in the week of June 15 - June 19 the lecture is held on Thursday June 18 and the tutorial on Wednesday June 17 from 14-16, so there will not be another tutorial on Wednesday June 17 from 16-18. As there is no lecture on July 15, we move the tutorial from July 15 to 14-16.

Lecture Tuesday	Lecture Wednesday	Tutorial Thursday	Tutorial Friday	Hand in	Discuss
14.04	15.04	16.04	17.04		
21.04	23.04	22.04	24.04		Sheet 1
28.04	29.04	30.04	01.05	Sheet 1	Sheet 2
05.05	06.05	07.05	08.05	Sheet 2	Sheet 3
12.5	13.5	14.05	15.05	Sheet 3	Sheet 4

Lecture Tuesday	Lecture Wednesday	Tutorial Wednesday	Tutorial Thursday	Hand in	Discuss
19.05	20.05	20.05	21.05	Sheet 4	Sheet 5
26.05	27.05	27.05	28.05		
02.06	03.06	03.06	04.06	Sheet 5	Sheet 6
09.06	10.06	10.06	11.06	Sheet 6	Sheet 7
16.06	18.06	17.06	17.06	Sheet 7	Sheet 8
23.06	24.06	24.06	25.06	Sheet 8	Sheet 9
30.06	01.07	01.07	02.07	Sheet 9	Sheet 10
07.07	08.07	08.07	09.07	Sheet 10	Sheet 11
14.07	15.07	15.07	16.07	Sheet 11	Summary

Let us begin with a bit of motivation for this course.

Before the 17th century, what we now call combinatorics largely appeared in the form of puzzles and recreational problems, long before it was formalized into a mathematical discipline. In that spirit, we will also begin with a puzzle and I promise there will be many more throughout the course.

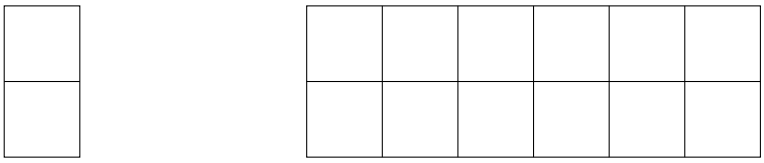
Domino Tilings: Assume you are given a box of size $2 \times n$ in which you want to place your domino tiles, which are of size 1×2 or 2×1 (depending on how you hold them).

In how many ways can you arrange the domino tiles in the box?

Think about it, before going to the next page. The page will also only be added to the notes after the first class - spoilers are not allowed here.

How do you proceed in order to solve such a problem?

A first approach could be to *draw* the problem, that is: you already try to model it.

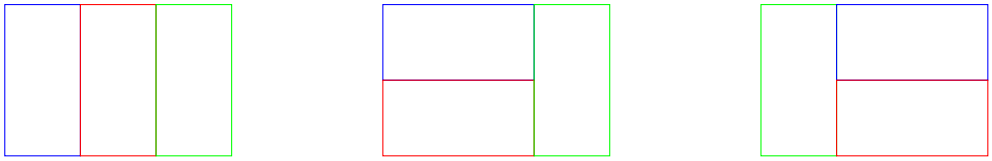


Another approach could be to solve it first for some small values of n and hope we find some pattern.

For $n = 1$, we clearly only have one way to place the domino tile inside the box.

For $n = 2$, we already have two ways:

For $n = 3$, we already get some more possibilities (namely 3):



We could continue like this, finding for $n = 4$ that there are 5 possibilities, for $n = 5$ there are 8 possibilities, etc.

Now, we might see a pattern, as the sequence $1, 2, 3, 5, 8, \ldots$ looks very familiar. In fact, it looks like

the *Fibonacci sequence*.

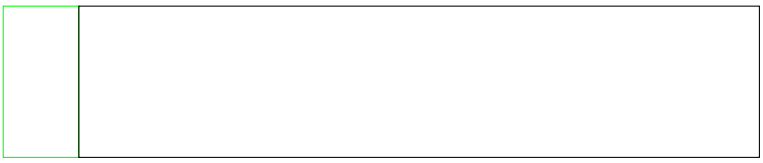
So we have a first claim:

The number of arrangements is given by f_n, where f_n denotes the Fibonacci sequence.

This clearly still needs to be proven. Knowing that the Fibonacci sequence $(f_n)_{n\in\mathbb{N}}$ follows the following recurrence relation

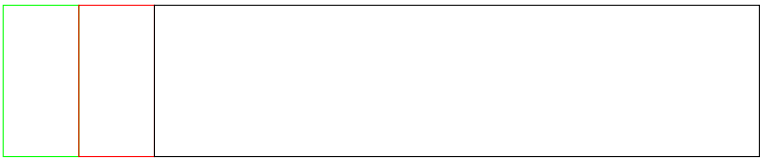
$$f_n = f_{n-1} + f_{n-2},$$

we also get an idea *how* we can prove our claim. In fact, we could give a proof by picture: instead of computing directly how many arrangements there are for n, we can use the number of arrangements for $n-1$ and add one more tile in green, thus



and we know the number of arrangements of $n-2$ and add two tiles in green and red as follows:

Note that the arrangement



does not need to be treated as well, as it is already included in the arrangement $n-1$. If we add the number of arrangements of $n-1$ and $n-2$, when then get all arrangements of n. Thus, by our previous computation for small n, we get the starting point of the recurrence $f(1)=1, f(2)=2$, we get the claim.

Of course there are more applications of this beautiful topic than domino tilings, but we can already see some of the core ideas and tools, such as recurrence, recognize patterns and trying to model problems (especially graphically).

Contents

0 Before we Start

0.1 Notation

Since we cannot include everything, we will assume a certain background. For example, we assume that modular arithmetic, the notion of field, vector space, basis and subgroup are known concepts.

Throughout these lecture notes, we will make use of the following notation:

- For a set S we denote by $|S|$ its cardinality and by S^C its complement.
- $\mathbb{Z}/n\mathbb{Z}$ denotes the integers modulo n.
- $\mathbb{F}_q$ denotes the finite field of order q, where q is a prime power.
- Id_n denotes the identity matrix of size n.
- $\mathrm{GL}_n(\mathbb{F}_q)$ denotes the general linear group of degree n in $\mathbb{F}_q$, i.e., all invertible matrices in $\mathbb{F}_q^{n\times n}$.
- For a matrix M we write $\mathrm{rk}(M)$ to denote its rank, $\det(M)$ to denote its determinant and by $M^\top$ we denote its transpose.
- For a function f we denote by $\ker(f)$ its kernel and by $\mathrm{im}(f)$ its image.
- For a vector $v \in \mathbb{F}_q^k$ and $i \in \{1,\dots,k\}$ we denote by v_i the ith entry of the vector v.
- We denote by $\mathbb{N}_0 = \{0,1,\dots\}$ all non-negative integers.
- We denote by $\mathbb{N} = \{1,\dots\} = \mathbb{N} \setminus \{0\}$ all positive integers.
- For $n \in \mathbb{N}$ we denote by $[n] = \{1,\dots,n\}$.

This list might get updated as we progress in the course.

Note that, unless otherwise stated, all sets in these notes are *finite*.

0.2 Some Tricks

While combinatorics is a research area of its own, we often use it as a provider of tools. Certain tools, you surely have seen already and recognize whenever you can apply them. This is a incomplete list of tools we already expect you to know:

- **Set operations, cardinalities and De Morgan's laws:** For sets S, T in a ground set G, we have $(S \cup T)^C = S^C \cap T^C, (S \cap T)^C = S^C \cup T^C$.

  Additionally $\cup, \cap$ are associative and distributive. To compute their cardinalities, we know $|S \cup T| = |S| + |T| - |S \cap T|$ and if S and T are disjoint, i.e., $|S \cap T| = \emptyset$, then $|S \cup T| = |S| + |T|$. On the other hand, $|S \times T| = |S||T|$.

  We say $S_1, \ldots, S_k$ are a *partition* of G, if $S_i \cap S_j = \emptyset$ for $i \neq j$ and $S_1 \cup \ldots \cup S_k = G$. We call S_i the *parts.*

- **Proof by contradiction:** To prove A, assume $\neg A$, i.e., not A, and derive a contradiction.

- **Proof by contraposition:** To prove $A \Rightarrow B$, instead prove $\neg B \Rightarrow \neg A$.

- **Induction:** To prove a statement for all $n \in \mathbb{N}_{\geq a}$: (i) prove base case $n = a$, (ii) show $n \Rightarrow n + 1$.

- **Pigeonhole principle:** If more than n objects are placed into n boxes, at least one box contains at least two objects.

- **Double counting:** Count the same object in two different ways to obtain an identity.

- **Exchange of sums (Fubini principle):** We may change the order of summation when summing over finite sets, that is

$$\sum_{i \in I} \sum_{j \in J} a_{i,j} = \sum_{j \in J} \sum_{i \in I} a_{i,j}.$$

- **Symmetric group:** S_n is the set of all permutations (bijections) of $\{1, \ldots, n\}$, with composition as operation and $|S_n| = n!$.

- **Telescoping sums:** Sums where most terms cancel, e.g. $\sum_{k=1}^{n} (a_k - a_{k+1}) = a_1 - a_{n+1}$.

- **Recurrences (e.g. Fibonacci):** Sequences defined recursively, such as the *Fibonacci sequence*, defined recursively by $f_1 = 1, f_2 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for all $n \geq 3$.

- **Harmonic numbers:** The partial sums $H_n = \sum_{k=1}^{n} \frac{1}{k}$ and their growth behavior.

- **Equivalence relations:** Sometimes we are only interested in the number of objects, up to equivalence. A relation $\sim$ is an equivalence relation if it is reflexive, symmetric, and transitive. It partitions a set into *equivalence classes* $[x] = \{y : y \sim x\}$.

- **Bijections:** Prove two sets have the same size by constructing a one-to-one correspondence. A function $f : A \to B$ is a bijection if it is injective and surjective.

1 Basic Counting

Combinatorics is not merely about counting - however, in this first chapter, we will do exactly that.

We will often encounter a statement, with several proofs - recall, that here we often care more about a technique, than a statement, thus all proofs should be treated equally important.

1.1 Arrangements and Subsets

Let us start with something we have mentioned in the motivation, without properly defining it: what even is an arrangement?

Definition 1.1. Let X be a set and k a positive integer. An *unordered arrangement* of k elements of X is a multiset of size k consisting of elements of X.

Example 1.2. *Consider $X = \{1, 2, 3, 4\}$ and $k = 3$. Then $S = \{1, 1, 2\}$ is an arrangement of size 3.*

Note that for sets, as well as multisets, the order does not matter. However, as we might want to write it more concisely, we could also use a positive integer to indicate how many times a distinct element appears. This is the *repetition number* and the distinct elements are called *types.*

Example 1.3. *In the example above, we could write $S = \{2 \cdot 1, 1 \cdot 2\}$, where $r_1 = 2$ is the repetition number for the type 1 and $r_2 = 1$ is the repetition number for the type 1.*

If the order matters, e.g., you want to call 112 and not 121, we call this an *ordered arrangement.*

If all repetition numbers are 1, we get a subset of X. How many subsets can X of size n have?

Proposition 1.4. *Let n be a non-negative integer. The number of subsets of an n-element set is 2^n.*

Proof. Let us first identify $X = [n]$. We can then encode subsets $S \subseteq X$ through sequences $(e_1, \ldots, e_n)$, where each e_i indicates whether the element i appears in S, i.e.,

$$e_i = \begin{cases} 1 & \text{if } i \in S, \\ 0 & \text{else.} \end{cases}$$

Clearly, each distinct subset S has a distinct sequence and vice versa. As for each $i \in X$ there are 2 choices for e_i, we get a total of 2^n sequences, and thus 2^n subsets. □

We can also give a second proof using an inductive argument.

Proof. We again identify $X = [n]$ and denote by $f(n)$ the number of subsets of a n-element set. Clearly, $f(0) = 1$, as the empty set only has the empty set as subset.

For each subset $S \subseteq \{1, \ldots, n\}$, we can extend to a subset of $\{1, \ldots, n + 1\}$ in two ways: we either add $n + 1$ as element or not. Thus $f(n + 1) = 2f(n)$ and since we start at $f(0) = 1 = 2^0$, we get $f(n) = 2^n$. □

To count how many subsets there are of a given size, we need some more definitions. Recall that for $n \in \mathbb{N}_0$ the *factorial* is $n! = \prod_{i=1}^{n} i = 1 \cdot 2 \cdots n$ and use the convention that $0! = 1$. In fact, as $n! = n(n-1)!$, we have

$$1 = 1! = 1 \cdot (1-1)! = 0!.$$

With this we can now define the main object of this first chapter: *the binomial coefficient.*

Definition 1.5. For $k \leq n$ non-negative integers, we define the *binomial coefficient*, to be

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k \cdot (k-1) \cdots 1},$$

whereas for $k > n$ and $k < 0$ we use the convention that $\binom{n}{k} = 0$.

Note that

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

We often denote the set of all subsets of size k of X as $\binom{X}{k}$, and if $|X| = n$, we thus get $\binom{n}{k} = |\binom{X}{k}|$. Let us show that this is indeed the case:

Proposition 1.6. *Let $k \leq n$ be non-negative integers. The number of k-element subsets of an n-element set is $\binom{n}{k}$.*

Proof. We choose k distinct elements of the n-element set X : for the first element we have n choices, for the second we are left with $n-1$ and so on. For the ith element we have $n-i+1$ choices until for the kth and last element we have $n-k+1$ choices. Multiplying all together we get $n \cdot (n-1) \cdots (n-k+1)$.

However, choosing the same elements in a different order still gives the same k-element subset. We thus have to divide by the number of different orders in which we can choose the k elements: we have k choices for the first element, $k-1$ for the second and so on. For the ith element we have $k-i+1$ choices until for the kth and last element we have $k-k+1 = 1$ choice. Multiplying all together we get $k \cdot (k-1) \cdots 1$. □

1.2 Properties of the Binomial Coefficient

There are many properties, tricks and equalities, we will only state some and whenever possible give a quick example highlighting its importance. Our first property, is the following.

Proposition 1.7. *Let $k \leq n$ be non-negative integers. Then*

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

Proof. Let us consider an n-element set X and count (again) the number of subsets. By Proposition 1.4, we have 2^n many. On the other hand, by Proposition 1.6, we know that for a fixed k, there are $\binom{n}{k}$ many subsets of size k, thus by summing over all possible k, we get all 2^n subsets. □

A possible example of this identity comes from coding theory - which we have not covered yet, but is explained without any new terms: say you want to count all vectors in $\mathbb{F}_2^n$, i.e., strings consisting of 0's and 1's. At the same time, you might be interested in counting how many vectors (or strings) there are with k 1's (we call them vectors of weight k). Then

$$|\mathbb{F}_2^n| = 2^n = \sum_{k=0}^{n} \binom{n}{k}.$$

We will revisit this kind of counting principle several times, whether we talk about permutations or partitions.

Although the definition of the binomial coefficient allows k to be zero, the value $\binom{n}{0}$ might not be well-defined, as we would divide by zero. With the property $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, and the convention $0! = 1$, we also have

$$\binom{n}{0} = \binom{n}{n} = 1$$

and in particular $\binom{0}{0} = 1$.

Again, using the coding-theoretic example, we could simply state there is only $(0, \ldots, 0)$ with 0 many 1's and only $(1, \ldots, 1)$ with n many 1's.

Together with the above edge cases $\binom{n}{0} = \binom{n}{n} = 1$, the binomial coefficients follow a recurrence relation:

Proposition 1.8 (Pascal's Formula)**.** *Let $k \leq n$ be non-negative integers. Then*

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

Proof. Let us consider the problem, where we want to count the number of k-elements sets of an n-elements set X containing a fixed element $x \in X$. We can first count the sets containing x : each must have $k-1$ of the remaining $n-1$ elements of $X \setminus \{x\}$, thus we get $\binom{n-1}{k-1}$.

On the other hand, we can count the number of sets which do not contain x : each has k elements of the $n-1$ elements in $X \setminus \{x\}$, which are $\binom{n-1}{k}$ many. Added together, these must give all k-elements subsets, which is exactly the claim. □

Proof. We can of course also proceed by simply calculating:

$$\begin{aligned}\binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-1-k)!} = \frac{(n-1)!k}{k!(n-k)!} + \frac{(n-1)!(n-k)}{k!(n-k)!} \\ &= \frac{(n-1)!k + (n-1)!(n-k)}{k!(n-k)!} = \frac{(n-1)!n}{k!(n-k)!} = \frac{n!}{k!(n-k)!}.\end{aligned}$$

□

The first proof we gave could be described as a 'combinatorial' proof, while the second has a more straightforward, calculating flavor. For educational reasons, we of course prefer the first kind, however, the second one is just as valid.

Binomial coefficients are also symmetric.

Proposition 1.9. *Let $k \leq n$ be non-negative integers. Then*

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

Exercise 1.10. *1. Prove Proposition 1.9 by finding a bijection between k-elements sets and $(n-k)$-elements sets in a set of size n.*

2. Prove Proposition 1.9 by calculation.

Why is Proposition 1.8 called Pascal's formula?[1] We can arrange the binomial coefficients in a symmetric triangular pattern, where the $(n+1)$st row contains $\binom{n}{0}, \ldots, \binom{n}{n}$. This is the so-called Pascal's triangle:

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 &
\end{array}$$

We can now see, that as in Proposition 1.8, every entry is simply computed by adding the two entries diagonally above it.

One of the most important property is captured in the Binomial Theorem, which is a generalization of Proposition 1.7.

Theorem 1.11. *Let n be a non-negative integer. Then*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Proof. We may write out $(x+y)^n = \underbrace{(x+y)\cdots(x+y)}_{n \text{ times}}$. Each term of this sum is obtained by choosing from k factors x and from the remaining $n-k$ terms we choose y, getting $x^k y^{n-k}$. Thus, the coefficient of the term $x^k y^{n-k}$ is given by the number of ways in which we can choose k many factors, which is $\binom{n}{k}$. □

[1] Although it is named after Pascal, this triangle appeared in China and India centuries before Pascal.

Proof. We can also prove this by induction on n. For $n = 0$, we clearly get

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

Assume the statement is true for n. Then

$$\begin{aligned}(x+y)^{n+1} = (x+y)(x+y)^n &= (x+y)\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k} \\ &= \sum_{k=0}^{n}\binom{n}{k}x^{k+1}y^{n-k} + \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k+1}.\end{aligned}$$

We set $j = k+1$ in the first term to get

$$\sum_{j=1}^{n+1}\binom{n}{j-1}x^j y^{n-j+1} = \sum_{j=1}^{n}\binom{n}{j-1}x^j y^{n-j+1} + \binom{n}{n}x^{n+1}y^0.$$

Whereas the second term is

$$\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k+1} = \sum_{k=1}^{n}\binom{n}{k}x^k y^{n-k+1} + \binom{n}{0}x^0 y^{n+1}.$$

Thus, their sum is

$$\binom{n}{0}x^0 y^{n+1} + \sum_{k=1}^{n} x^k y^{n-k+1}\left(\binom{n}{k-1} + \binom{n}{k}\right) + \binom{n}{n}x^{n+1}y^0.$$

Using Proposition 1.7, we get

$$\binom{n}{0}x^0 y^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}x^k y^{n-k+1} + \binom{n}{n}x^{n+1}y^0 = \sum_{k=0}^{n+1}\binom{n+1}{k}x^k y^{n+1-k}.$$

□

We know that if we sum all subsets of size k, we get 2^n many, but if we only sum over the even or odd k?

Proposition 1.12. *Let n be a positive integer. Then*

$$\sum_{k=0}^{\lfloor n/2\rfloor}\binom{n}{2k} = \sum_{k=0}^{\lfloor (n-1)/2\rfloor}\binom{n}{2k+1} = 2^{n-1}.$$

Proof. Let us denote by $S_e = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k}$, the even sum and by $S_o = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1}$, the odd sum. Since together they sum over all k, we get $S_e + S_o = 2^n$.

If we set $x = -1$ and $y = 1$ in the Binomial Theorem 1.11, we get

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

Note that in this sum the even k have a coefficient of $(-1)^k = 1$ and the odd k have a coefficient $(-1)^k = -1$, thus $S_e - S_o = 0$. And in turn, we get $S_e = S_o = 2^n/2 = 2^{n-1}$. □

If we stick to the coding-theoretic example, when interested in the number of vectors (or strings) in $\mathbb{F}_2^n$ with even numbers of 1's (i.e., of even weight), we can immediately get to 2^{n-1}.

Proposition 1.13 (ChuVandermonde identity)**.** *Let m, n, k be positive integers. Then*

$$\sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} = \binom{m+n}{k}.$$

Proof. Assume we have a set with n even numbers and m odd numbers. We have to choose a set of size k from it: in how many ways can we do so? The total number of subsets is still $\binom{m+n}{k}$ from Proposition 1.6. On the other hand, we can count in this way: To get sets with i even elements, we have $\binom{n}{i}$ choices and then choose the remaining $k-i$ from the odd elements, i.e., $\binom{m}{k-i}$. Hence we get $\sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i}$ ways. □

Proof. Let us consider $(1+x)^n(1+x)^m = (1+x)^{m+n}$. For the right hand side, we know from the Binomial Theorem 1.11 that the coefficient for the term x^k is $\binom{m+n}{k}$. For the left hand side we can choose x^i from the first factor, i.e., $(1+x)^n$, for which the Binomial Theorem tells us the coefficient is $\binom{n}{i}$, and then choose x^{k-i} from the second factor, i.e., $(1+x)^m$, for which the coefficient is $\binom{m}{k-i}$. Thus, summing over all i we get the claim. □

Note that we can plug in $m = n = k$ in Proposition 1.13 and use that $\binom{n}{i} = \binom{n}{n-i}$ from Proposition 1.9, to get

Corollary 1.14. *Let n be a positive integer. Then*

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

Proposition 1.15. *Let $k \leq n$ be non-negative integers, then*

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

Proof. Let us assume that we have n elements and need to choose k, where one element is considered special. We could first choose the k elements, giving $\binom{n}{k}$ possibilities and then out of the k chosen element choose one to be special, which gives $k\binom{n}{k}$, i.e., the right hand side. On the other hand, we could first choose the special element out of the n elements, for which we have n possibilities, and now choose $k-1$ elements from the remaining $n-1$ elements, giving the left hand side, i.e., $n\binom{n-1}{k-1}$. □

Proof. We could also prove this by simple calculations:

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

□

Proposition 1.16. *Let n be a positive integer. Then*

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}.$$

Proof. From the Binomial Theorem 1.11, we get that $(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k$. We can differentiate to get

$$n(1+x)^{n-1} = \sum_{k=1}^{n}\binom{n}{k}kx^{k-1}$$

and now plug in $x = 1$, giving the claim. □

Proof. We could again simply proceed by calculation: in fact, by Proposition 1.15 we get

$$\sum_{k=1}^{n} k\binom{n}{k} = \sum_{k=1}^{n} n\binom{n-1}{k-1} = n\sum_{\ell=0}^{n-1}\binom{n-1}{\ell},$$

which due to Proposition 1.7 gives $n2^{n-1}$. □

Proof. Note that there are $\binom{n}{k}$ subsets of size k in an n-element set X. Thus, the sum adds up the size of all these subsets. If we pair up each subset S with $X \setminus S$, their size added together is $n = |X|$. Since there are 2^n subsets in total, they fall into $2^n/2 = 2^{n-1}$ pairs. □

Proposition 1.17 (Hockey Stick Formula)**.** *Let $k \leq n$ be non-negative integers. Then*

$$\sum_{i=k}^{n}\binom{i}{k} = \binom{n+1}{k+1}.$$

Proof. The first proof is using induction on n. For $n = k$, we get

$$\sum_{i=k}^{k}\binom{i}{k} = \binom{k}{k} = 1 = \binom{k+1}{k+1}.$$

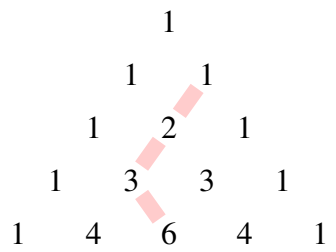
Assume the statement is true for n, then

$$\sum_{i=k}^{n+1} \binom{i}{k} = \sum_{i=k}^{n} \binom{i}{k} + \binom{n+1}{k} = \binom{n+1}{k+1} + \binom{n+1}{k},$$

which by Proposition 1.8 gives $\binom{n+2}{k+1}$. □

Proof. We can count the $(k+1)$-element subsets of a set of size $n+1$, which are $\binom{n+1}{k+1}$ many, and can partition them according to the largest element, say $i+1$. The remaining k elements can then only be chosen from $\{1, \ldots, i\}$, i.e., we have $\binom{i}{k}$ choices. Thus summing over all i gives the claim. □

The fact, that we have used Pascal's Formula for the first proof is not by coincidence: the name hockey stick is derived by the triangle as well:



Proposition 1.18 (Trinomial Revision)**.** *Let n, m, k be non-negative integers. Then*

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

Exercise 1.19. *Prove Proposition 1.18.*

As a big hint for the exercise, the name 'trinomial' stems from the fact that it splits a set into 3 groups, e.g. you may think of a team of k people, with m leaders, and the remaining $n-k$ remaining people.

Proposition 1.20 (Absorption Formula)**.** *Let n, m be positive integers. Then*

$$\binom{m}{n} = \frac{m}{n}\binom{m-1}{n-1}.$$

Exercise 1.21. *Prove Proposition 1.20.*

The name comes from the idea that one element is absorbed into the binomial coefficient.

We can also consider the binomials modulo some prime number. To do so, we will first prove one of the most important properties when computing modulo a prime: Freshman's dream.

Theorem 1.22 (Freshman's Dream)**.** *Let a, b be non-negative integers and p a prime. Then*

$$(a+b)^p \equiv a^p + b^p \mod p.$$

Proof. Due to the Binomial Theorem 1.11, we have that

$$(a+b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i},$$

however, for any $i \in \{1, \ldots, p-1\}$, we have that

$$p \mid \binom{p}{i} = \frac{p!}{i!(p-i)!},$$

only for $i = 0$ and $i = p$ do we get $\binom{p}{i} = 1$. Thus, inserting these values in the Binomial Theorem, we get the claim. □

Lemma 1.23. *Let p be a prime and n, k be non-negative integers with $n = cp + a, k = dp + b$, for some non-negative integers c, d and for $a, b \in \{0, \ldots, p-1\}$. Then*

$$\binom{n}{k} \equiv \binom{c}{d}\binom{a}{b} \mod p.$$

Proof. By Freshman's Dream 1.22, we get that $(1+x)^p \equiv 1 + x^p \mod p$, hence

$$(1+x)^n = (1+x)^{cp}(1+x)^a \equiv (1+x^p)^c(1+x)^a,$$

which due to the Binomial Theorem 1.11 is

$$\sum_{i=0}^{c} \binom{c}{i} x^{ip} \sum_{j=0}^{a} \binom{a}{j} x^j.$$

Since $a, b < p$, the only way to get the term $x^k = x^{dp+b}$ is by setting $i = d, j = b$. Thus, these coefficients have to be equal, giving the claim. □

Theorem 1.24 (Lucas Theorem)**.** *Let p be a prime number and n, k be non-negative integers. We can write n, k in the base p, say*

$$n = a_0 + a_1 p + \cdots + a_d p^d,$$
$$k = b_0 + b_1 p + \cdots + b_d p^d,$$

for some $a_i, b_i \in \{0, \ldots, p-1\}$. Then

$$\binom{n}{k} \equiv \prod_{i=0}^{d} \binom{a_i}{b_i} \mod p.$$

Proof. We proceed by induction on d. For $d = 0$, we have $n = a_0, k = b_0$ and hence

$$\binom{a_0}{b_0} = \prod_{i=0}^{0} \binom{a_0}{b_0}.$$

Assume the statement is true for $d-1$, and set $a = a_0, c = a_1 + \cdots + a_d p^{d-1}$, i.e., $n = cp + a$ and $b = b_0, d = b_1 + \cdots + b_d p^{d-1}$, i.e., $k = dp + b$. Hence we can apply Lemma 1.23 to get

$$\binom{n}{k} \equiv \binom{c}{d}\binom{a}{b}.$$

Since c, d have $d-1$ terms, the induction hypothesis can be applied to them, thus we get

$$\binom{n}{k} \equiv \binom{a_0}{b_0} \prod_{i=1}^{d} \binom{a_i}{b_i} = \prod_{i=0}^{d} \binom{a_i}{b_i}.$$

□

Lucas theorem is the tool that tells us exactly when a binomial coefficient is nonzero modulo a prime and reduces huge problems to small ones digit by digit. In fact, we get the following corollary

Corollary 1.25. *Let p be a prime number and n, k be non-negative integers. We can write n, k in the base p, say*

$$n = a_0 + a_1 p + \cdots + a_d p^d,$$
$$k = b_0 + b_1 p + \cdots + b_d p^d,$$

for some $a_i, b_i \in \{0, \ldots, p-1\}$. Then $\binom{n}{k} \not\equiv 0 \mod p$ if and only if $b_i \leq a_i$ for all i.

Exercise 1.26. *Is $\binom{30}{5}$ divisible by 2?*

Exercise 1.27. *For how many non-negative integers $k \leq 2025$ is $\binom{2025}{k}$ odd?*

1.3 Recurrence Relations

In many combinatorial problems, we do not count objects directly, but instead relate the number of objects of size n to those of smaller size.

A classical example is given by the Fibonacci numbers. Recall that f_n counts the number of ways to tile a $2 \times n$ board using tiles of length 1 and 2.

To construct such a tiling, we consider the first tile:

- If the first tile has length 1, then the remaining part can be tiled in f_{n-1} ways.

- If the first tile has length 2, then the remaining part can be tiled in f_{n-2} ways.

Thus, we obtain the recurrence

$$f_n = f_{n-1} + f_{n-2}.$$

This illustrates a general principle: instead of counting objects directly, we break them into smaller cases and derive a *recurrence relation*.

In this section, we will introduce three recurrence relations: Bell numbers and Stirling numbers of the first and second kind. Before we can do so, we need to go into the falling factorial and the Stars and Bars concept.

Instead of the factorial, we may also consider the *falling factorial*.

Definition 1.28. Let k, n be non-negative integers. The *falling factorial* is then defined as

$$(k)_n = k \cdot (k-1) \cdots (k-n+1).$$

Exercise 1.29. *Show that* $(k)_n = \frac{k!}{(k-n)!}$.

One property that we will use often is that $(k)_n = (k)_{n-1}(k-(n-1))$. In fact,

$$(k)_n = k \cdot (k-1) \cdots (k-(n-2)) \cdot (k-(n-1)) = (k)_{n-1}((k-(n-1)).$$

We are now ready for Stars and Bars: Assume we have n indistinguishable stars and want to place them into k boxes, which are marked from $\{1, \ldots, k\}$. As usual, we ask: in how many different ways can we do this?

Theorem 1.30 (Stars and Bars)**.** *Let $k \leq n$ be non-negative integers. Then, the number of solutions to $x_1 + \cdots + x_k = n$ for x_i being non-negative integers is $\binom{n+k-1}{k-1}$.*

Proof. We represent a solution $(x_1, \ldots, x_k)$ of

$$x_1 + \cdots + x_k = n, \qquad x_i \geq 0,$$

by a string consisting of n stars and $k-1$ bars. The number of stars before the first bar is x_1, the number of stars between the first and second bar is x_2, and so on, while the number of stars after the last bar is x_k.

For example, the string

$$\star\star \mid \star \mid \star\star\star$$

corresponds to the solution $(2, 1, 3)$.

Thus, every solution corresponds uniquely to a string of length $n+k-1$ containing $k-1$ bars. Hence, the number of such solutions is the number of ways to choose the positions of the $k-1$ bars among the $n+k-1$ positions, namely

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

□

Example 1.31. *Assume we have 4 balls and 2 boxes. We can visualize the* $\binom{4+2-1}{2-1} = \binom{5}{1} = 5$ *choices of placing them using stars and bars from the proof as follows:*

$$\begin{array}{ll} (0,4) & \mid \star\star\star\star \\ (1,3) & \star \mid \star\star\star \\ (2,2) & \star\star \mid \star\star \\ (3,1) & \star\star\star \mid \star \\ (4,0) & \star\star\star\star \mid \end{array}$$

If we are interested in only positive solutions x_i, we can still use the theorem above, but get a different number:

Corollary 1.32. *Let* $k \leq n$ *be non-negative integers. Then, the number of solutions to* $x_1 + \cdots + x_k = n$ *for* x_i *being positive integers is* $\binom{n-1}{k-1}$.

One should never prove a corollary, but we would still like to give a small explanation: to see that this really follows from Theorem 1.30, simply replace x_i by $y_i = x_i - 1$, then $\sum_{i=1}^{k} y_i = n - k$ and $y_i \geq 0$. Hence applying Theorem 1.30, we get $\binom{(n-k)+k-1}{k-1} = \binom{n-1}{k-1}$.

In order to introduce some interesting recurrence relations, we first recall some theory of partitions and then of permutations.

Let S be a finite set, then $S_1, \ldots, S_k$ are a *partition* of S, if $S_i \neq \emptyset$, $S_i \cap S_j = \emptyset$ for all $i \neq j$ and $\bigcup_{i=1}^{k} S_i = S$. The sets S_i are called *parts*.

In how many ways can we partition an n-element set, say $\{1, \ldots, n\}$?

Definition 1.33. Let n be a positive integer. The *Bell number*, denoted by $B(n)$, is given by the number of partitions of $\{1, \ldots, n\}$.

We additionally set $B(0) = 1$, which also makes sense, as the empty set only has one way to partition it: $\emptyset$.

Example 1.34. $B(1) = 1$ *as well, since* $\{1\}$ *can only be partitioned into* $\{1\}$.

$B(2) = 2$*, as we can readily check: the set* $\{1, 2\}$ *can be partitioned as*

$$S_1 = \{1, 2\} \quad \textit{or} \quad S_1 = \{1\}, S_2 = \{2\}.$$

For $n = 3$*, we compute all partitions of* $\{1, 2, 3\}$*:*

$$\begin{aligned} &S_1 = \{1, 2, 3\}, \\ &S_1 = \{1, 2\}, S_2 = \{3\}, \\ &S_1 = \{1, 3\}, S_2 = \{2\}, \\ &S_1 = \{2, 3\}, S_2 = \{1\}, \\ &S_1 = \{1\}, S_2 = \{2\}, S_3 = \{3\}. \end{aligned}$$

Hence $B(3) = 5$.

The Bell numbers can also be computed through the following recurrence:

Proposition 1.35. *Let n be a positive integer, then*

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

Proof. We may ask instead: how many partitions are there such that the part containing the element n consists of exactly k elements?

For this, we put n in a part and choose $k-1$ elements from the remaining $n-1$ elements: these are $\binom{n-1}{k-1}$ possibilities. Now, we are only left with partitioning the remaining $n-k$ elements, i.e., $B(n-k)$. □

Recall that $\sum_{k=0}^{n} \binom{n}{k} = 2^n$, where we split the number of all subsets into those of size k. We can do something similar here, i.e., split the number of partitions of $\{1, \ldots, n\}$ into those having k parts.

Definition 1.36. Let $k \leq n$ be positive integers. The *Stirling number of the second kind*, denoted by $S(n,k)$, is the number of partitions of $\{1, \ldots, n\}$ into k parts.

We again use the convention that $S(n,0) = 0$. Clearly, we must have that

$$\sum_{k=1}^{n} S(n,k) = B(n).$$

Example 1.37. *Recall that $B(3) = 5$, for which we have*

$$S(3,1) = 1, \quad S(3,2) = 3, \quad S(3,3) = 1.$$

These numbers also follow a recurrence relation:

Proposition 1.38. *Let $1 < k < n$ be positive integers, then $S(n,1) = S(n,n) = 1$ and*

$$S(n,k) = S(n-1,k-1) + kS(n-1,k).$$

Proof. The only way we can partition $\{1, \ldots, n\}$ into 1 part is $\{1, \ldots, n\}$ and into n parts is $\{1\}, \ldots, \{n\}$, thus $S(n,1) = S(n,n) = 1$.

For $k < n$, consider the partitions into k parts and split them into two classes: in the first class, we have all partitions where $\{n\}$ is a part and in the second, all partitions where the element n belongs to a part of size larger than 1.

For the first class, we can simply partition the remaining $n-1$ elements into $k-1$ parts, i.e., we have $S(n-1,k-1)$ ways.

For the second class, we may delete n from the part it belongs to (without getting an empty part), getting a partition of $n-1$ elements into k parts, i.e., $S(n-1,k)$ ways. However to go back from these partitions to a partition of n elements, we need to reinsert n in one of the parts, for which we have k choices. □

Additionally, we have a similar statement to the Binomial Theorem. Here the Stirling numbers of the second kind play the same role for falling factorials as binomial coefficients do for monomials.

Proposition 1.39. *Let n be a positive integer. Then*

$$x^n = \sum_{k=1}^{n} S(n,k)(x)_k.$$

Proof. Note that $(x)_{k+1} = (x)_k(x-k)$ by definition. Thus, we can prove the statement using induction on n. For $n = 1$, we get that

$$x^1 = \sum_{k=1}^{1} S(1,k)(x)_k.$$

Assume now that the statement is true for $n-1$, then

$$\begin{aligned}
x^n = x \cdot x^{n-1} &= x \sum_{k=1}^{n-1} S(n-1,k)(x)_k \\
&= \sum_{k=1}^{n-1} S(n-1,k)(x)_k(x-k+k) \\
&= \sum_{k=1}^{n-1} S(n-1,k)(x)_k(x-k) + \sum_{k=1}^{n-1} kS(n-1,k)(x)_k \\
&= \sum_{k=1}^{n-1} S(n-1,k)(x)_{k+1} + \sum_{k=1}^{n-1} kS(n-1,k)(x)_k \\
&= \sum_{\ell=2}^{n} S(n-1,\ell-1)(x)_\ell + \sum_{k=1}^{n-1} kS(n-1,k)(x)_k.
\end{aligned}$$

Hence, for $k \leq n-1$, the coefficient of $(x)_k$ is $S(n-1,k-1)+kS(n-1,k)$, which by Proposition 1.38 is $S(n,k)$. We can add $0 = nS(n-1,n)(x)_n$ and $0 = S(n-1,0)$ term as well to get

$$\begin{aligned}
x^n &= S(n-1,0) + \sum_{\ell=2}^{n} S(n-1,\ell-1)(x)_\ell + \sum_{k=1}^{n-1} kS(n-1,k)(x)_k + nS(n-1,n) \\
&= \sum_{\ell=1}^{n} S(n-1,\ell-1)(x)_\ell + \sum_{k=1}^{n} kS(n-1,k)(x)_k = \sum_{k=1}^{n} S(n,k)(x)_k.
\end{aligned}$$

□

If these were called Stirling numbers of the second kind, then clearly there must also exist Stirling numbers of the first kind - which is what we will do next!

Recall that S_n denotes all the bijections from $\{1,\ldots,n\}$ to itself, i.e., permutations. We may either denote a permutation σ using a string $i_1 \cdots i_n$ indicating that $\sigma(j) = i_j$, or use the cycle notation.

Definition 1.40. Let $\sigma \in S_n$ be a permutation. A *cycle* $(i_1\ i_2\ \ldots\ i_k)$ denotes the permutation mapping

$$i_1 \mapsto i_2,\ i_2 \mapsto i_3,\ \ldots,\ i_{k-1} \mapsto i_k,\ i_k \mapsto i_1,$$

and fixing all other elements. Every permutation can be written as a product of disjoint cycles.

Example 1.41. *The permutation $\sigma \in S_5$ given by*

$$\sigma(1) = 3,\ \sigma(3) = 5,\ \sigma(5) = 1,\ \sigma(2) = 4,\ \sigma(4) = 2$$

can be written in cycle notation as

$$\sigma = (1\ 3\ 5)(2\ 4).$$

Note that the cycle notation is unique up to ordering. Thus, we can count the number of cycles of a permutation $\sigma \in S_n$ (including fixed points) and denote this by $c(\sigma)$. The *parity* of a permutation σ is then $n - c(\sigma)$, and we say a permutation σ has *even* parity if $n - c(\sigma)$ is even and *odd* otherwise. The *sign* of a permutation σ is then $(-1)^{n-c(\sigma)}$.

Definition 1.42. The *unsigned Stirling number of the first kind* is the number of permutations of S_n with k cycles and is denoted by $u(n, k)$. The *Stirling number of the first kind* is given by

$$s(n, k) = (-1)^{n-k} u(n, k).$$

Clearly, we must have that

$$\sum_{k=1}^{n} |s(n, k)| = \sum_{k=1}^{n} u(n, k) = |S_n| = n!.$$

These Stirling numbers also follow a recurrence relation.

Proposition 1.43. *Let n be a positive integer. Then $s(n, n) = 1$ and $s(n, 1) = (-1)^{n-1}(n - 1)!$ and for $1 < k < n$ a positive integer we have*

$$s(n, k) = s(n - 1, k - 1) - (n - 1)s(n - 1, k).$$

Proof. Let us start with $s(n, 1)$, which we can compute by counting the number of permutations with a single cycle. We may start with the element 1 and can then visit any other of the $n - 1$ elements in any order, i.e., $(n - 1)!$ possibilities and each has sign $(-1)^{n-1}$.

For $s(n, n)$ we count the number of permutations with n cycles, i.e., each element of $\{1, \ldots, n\}$ is alone in the cycle, i.e., a fixed point: this is only the identity.

For $1 < k < n$, we can divide the permutations with k cycles into two classes: one where n is a fixed point, and one where n belongs to a cycle of length at least 2. In the first case, if we delete the element n, we get a permutation in S_{n-1} with $k - 1$ cycles, i.e., $|s(n - 1, k - 1)|$ many, each of sign $(-1)^{(n-1)-(k-1)} = (-1)^{n-k}$.

For the other class, we can delete the element n and get a permutation in S_{n-1} with k cycles, i.e., $|s(n - 1, k)|$ many. To get a permutation of S_n back, we can reinsert n in $n - 1$ places. For the sign, we note that $(-1)^{(n-1)-k} = -(-1)^{n-k}$. Hence for this class we get $-(n - 1)s(n - 1, k)$ and adding the two classes together we get the claim. □

Recall the analog of the Binomial Theorem counting partitions, where

$$x^n = \sum_{k=1}^{n} S(n,k)(x)_k.$$

For the Stirling numbers of the first kind, we can 'reverse' this, to write the falling factorial as sum of x^k.

Proposition 1.44. *Let n be a positive integer. Then*

$$(x)_n = \sum_{k=1}^{n} s(n,k)x^k.$$

Proof. We proceed by induction on n. For $n = 1$, we get that

$$(x)_1 = x = \sum_{k=1}^{1} s(1,k)x^k.$$

Now assume the statement is true for $n-1$, then

$$\begin{aligned}(x)_n = (x)_{n-1}(x-n+1) &= \sum_{k=1}^{n-1} s(n-1,k)x^k(x-n+1) \\ &= \sum_{k=1}^{n-1} s(n-1,k)x^{k+1} - \sum_{k=1}^{n-1} s(n-1,k)x^k(n-1) \\ &= \sum_{\ell=2}^{n} s(n-1,\ell-1)x^{\ell} - \sum_{k=1}^{n-1} s(n-1,k)x^k(n-1).\end{aligned}$$

Again, we can also add the terms $0 = s(n-1,0)x^1$ as well as $0 = s(n-1,n)x^n(n-1)$, to get the complete sums

$$\sum_{\ell=1}^{n} s(n-1,\ell-1)x^{\ell} - \sum_{k=1}^{n} s(n-1,k)x^k(n-1).$$

The coefficient of x^k for $k < n$ is $s(n-1,k-1) - (n-1)s(n-1,k)$, which by Proposition 1.43 is $s(n,k)$, hence

$$(x)_n = \sum_{k=1}^{n} s(n,k)x^k.$$

□

1.4 Multinomial Coefficients

The binomial coefficients count the number of ways to split a set into two parts of prescribed sizes. A natural generalization is to partition a set into more than two parts.

Suppose we want to divide n objects into r groups of sizes $k_1, \ldots, k_r$ with

$$k_1 + \cdots + k_r = n.$$

How many ways can this be done?

For this, we generalize the binomial coefficient to the multinomial coefficient.

Definition 1.45. Let $k_1, \ldots, k_r$ be non-negative integers such that $\sum_{i=1}^{r} k_i = n$. Then the *multinomial coefficient* is defined as

$$\binom{n}{k_1, \ldots, k_r} = \binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{n-k_1-\cdots-k_{r-1}}{k_r}.$$

Exercise 1.46. *Show that*

$$\binom{n}{k_1, \ldots, k_r} = \frac{n!}{k_1! \cdots k_r!}.$$

With this new object, we also have a generalization of the Binomial Theorem:

Theorem 1.47. *Let n be a non-negative integer. Then*

$$(x_1 + \cdots + x_r)^n = \sum_{k_1,\ldots,k_r : \sum_{i=1}^{r} k_i = n} \binom{n}{k_1, \ldots, k_r} x_1^{k_1} \cdots x_r^{k_r}.$$

Proof. The proof works similarly as for the Binomial Theorem, that is: we can multiply the left hand side out, i.e.,

$$\underbrace{(x_1 + \cdots + x_r)\cdots(x_1 + \cdots + x_r)}_{n \text{ times}},$$

and choose k_1 many times the x_1 from the factors (for which we have $\binom{n}{k_1}$ choices), then k_2 times the x_2 term from the remaining factors (for which we have $\binom{n-k_1}{k_2}$ choices) and so on. For the ith step we want to choose k_i times the x_i from the remaining factors, which are $n - k_1 - \cdots - k_{i-1}$ many. For the last term with x_r we choose k_r many times x_r from the remaining $n - k_1 - \cdots - k_{r-1} = k_r$ factors. In fact, this last choice is not really a choice as we are left with $\binom{k_r}{k_r} = 1$. Hence multiplying all choices together we get the coefficient of $x_1^{k_1} \cdots x_r^{k_r}$ is $\binom{n}{k_1,\ldots,k_r}$. □

Note that $\binom{n}{k} = \binom{n}{k,n-k}$. Hence it is indeed a generalization of the binomial coefficient.

We can also generalize Pascal's formula, as follows:

Theorem 1.48. *Let n be a positive integer and $k_1, \dots, k_r$ be non-negative integers such that $\sum_{i=1}^{r} k_i = n$. Then*

$$\binom{n}{k_1, \dots, k_r} = \sum_{i=1}^{r} \binom{n-1}{k_1, \dots, k_i - 1, \dots, k_r}.$$

Proof. By the generalized Binomial Theorem 1.47, we get that

$$(x_1 + \cdots + x_r)^n = \sum_{k_1,\dots,k_r : \sum_{i=1}^{r} k_i = n} \binom{n}{k_1, \dots, k_r} x_1^{k_1} \cdots x_r^{k_r}.$$

Computing $(x_1 + \cdots + x_r)^n$ in a different way, we get

$$\begin{aligned}
(x_1 + \cdots + x_r)^n &= (x_1 + \cdots + x_r)(x_1 + \cdots + x_r)^{n-1} \\
&= (x_1 + \cdots + x_r) \sum_{k'_1,\dots,k'_r : \sum_{i=1}^{r} k'_i = n-1} \binom{n-1}{k'_1, \dots, k'_r} x_1^{k'_1} \cdots x_r^{k'_r} \\
&= \sum_{i=1}^{r} \sum_{k'_1,\dots,k'_r : \sum_{i=1}^{r} k'_i = n-1} \binom{n-1}{k'_1, \dots, k'_r} x_1^{k'_1} \cdots x_i^{k'_i+1} \cdots x_r^{k_r}.
\end{aligned}$$

We can thus substitute the indices $k_i = k'_i + 1$ and $k_j = k'_j$ for $j \neq i$, to get

$$(x_1 + \cdots + x_r)^n = \sum_{i=1}^{r} \sum_{k_1,\dots,k_r : \sum_{i=1}^{r} k_i = n} \binom{n-1}{k_1, \dots, k_i - 1, \dots, k_r} x_1^{k_1} \cdots x_r^{k_r}.$$

By swapping the sums, we get

$$\sum_{k_1,\dots,k_r : \sum_{i=1}^{r} k_i = n} \binom{n}{k_1, \dots, k_r} x_1^{k_1} \cdots x_r^{k_r} = \sum_{k_1,\dots,k_r : \sum_{i=1}^{r} k_i = n} \sum_{i=1}^{r} \binom{n-1}{k_1, \dots, k_i - 1, \dots, k_r} x_1^{k_1} \cdots x_r^{k_r}.$$

Hence, we get the claim. □

For the binomials, we had a nice representation using Pascal's triangle. That is a bit harder now, as we would require to draw something in more than two dimensions. We can give an example however. Let $n = 4, r = 3$ and we want to compute $\binom{4}{2,1,1}$. We thus first compute the easier $n = 3$ multinomials:

$$\binom{3}{3,0,0} = 1, \binom{3}{2,1,0} = 3, \binom{3}{1,1,1} = 6.$$

And hence

$$\binom{4}{2,1,1} = \binom{3}{1,1,1} + \binom{3}{2,0,1} + \binom{3}{2,1,0} = 6 + 3 + 3 = 12.$$

This identity generalizes Pascal's rule: instead of two predecessors, each multinomial coefficient has r predecessors, corresponding to the possible choices of the first element.

Let us come back to the reason, why we introduced the multinomial coefficient, that is: we want to partition an n-set into r groups:

Theorem 1.49. *Let X be a (finite) multiset with r different types and the repetition numbers $k_1, \ldots, k_r$. If $|X| = n = \sum_{i=1}^{r} k_i$, then the number of n-permutations of X is given by $\binom{n}{k_1,\ldots,k_r}$.*

Proof. We can label the r different types as $a_1, \ldots, a_r$. In an n-permutation there exist n positions which need to be assigned a type. We may first choose k_1 positions for the first type, for which we have $\binom{n}{k_1}$ choices and then assign k_2 of the remaining positions for the second type, i.e., $\binom{n-k_1}{k_2}$. We can continue in this way and multiply all the choices together to get

$$\binom{n}{k_1} \cdots \binom{n - k_1 - \cdots - k_{r_1}}{k_r} = \binom{n}{k_1, \ldots, k_r}.$$

□

1.5 Summary of Formulas

With this we are at the end of our 'basic counting' chapter. We will see the binomials, multinomials and some of the introduced numbers again in this lecture- hence we provide here a small summary of all the formulas we have seen:

Subsets and Binomial Coefficients. For an n-element set, the number of subsets is 2^n. For $0 \leq k \leq n$, the binomial coefficient is

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

and it counts the number of k-element subsets of an n-element set.

Basic Identities for Binomial Coefficients. For $0 \leq k \leq n$, we have

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n,$$

$$\binom{n}{0} = \binom{n}{n} = 1,$$

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

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \qquad \text{(Pascal's Formula).}$$

Binomial Theorem. For $n \in \mathbb{N}_0$,

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Even and Odd Binomial Sums. For $n \geq 1$,

$$\sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1} = 2^{n-1}.$$

Chu–Vandermonde Identity. For positive integers m, n, k,

$$\sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} = \binom{m+n}{k}.$$

In particular,

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

Further Useful Binomial Identities. For $0 \leq k \leq n$,

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

$$\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1},$$

$$\sum_{i=k}^{n} \binom{i}{k} = \binom{n+1}{k+1} \qquad \text{(Hockey Stick Formula)},$$

$$\binom{n}{k}\binom{k}{m} = \binom{n}{m}\binom{n-m}{k-m} \qquad \text{(Trinomial Revision)},$$

$$\binom{m}{n} = \frac{m}{n}\binom{m-1}{n-1} \qquad \text{(Absorption Formula)}.$$

Binomial Coefficients Modulo a Prime. Let p be a prime. Then

$$(a+b)^p \equiv a^p + b^p \pmod{p} \qquad \text{(Freshman's Dream)}.$$

If $n = cp + a$ and $k = dp + b$ with $0 \leq a, b < p$, then

$$\binom{n}{k} \equiv \binom{c}{d}\binom{a}{b} \pmod{p}.$$

More generally, if

$$n = a_0 + a_1 p + \cdots + a_d p^d, \qquad k = b_0 + b_1 p + \cdots + b_d p^d,$$

then

$$\binom{n}{k} \equiv \prod_{i=0}^{d} \binom{a_i}{b_i} \pmod{p} \qquad \text{(Lucas' Theorem).}$$

In particular,

$$\binom{n}{k} \not\equiv 0 \pmod{p} \iff b_i \leq a_i \text{ for all } i.$$

Falling Factorials. For non-negative integers $n \leq k$, the falling factorial is

$$(k)_n = k(k-1)\cdots(k-n+1) = \frac{k!}{(k-n)!}.$$

Stars and Bars. The number of non-negative integer solutions to $x_1 + \cdots + x_k = n$ is

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

The number of positive integer solutions to $x_1 + \cdots + x_k = n$ is

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

Bell Numbers and Stirling Numbers of the Second Kind. The Bell number $B(n)$ counts the number of partitions of $\{1, \ldots, n\}$ and satisfies

$$B(0) = 1, \qquad B(n) = \sum_{k=1}^{n} \binom{n-1}{k-1} B(n-k).$$

The Stirling number of the second kind $S(n,k)$ counts the partitions of $\{1, \ldots, n\}$ into k parts, and

$$\sum_{k=1}^{n} S(n,k) = B(n),$$

$$S(n,1) = S(n,n) = 1, \qquad S(n,k) = S(n-1,k-1) + kS(n-1,k).$$

Moreover,

$$x^n = \sum_{k=1}^{n} S(n,k)(x)_k.$$

Stirling Numbers of the First Kind. Let $u(n,k)$ be the number of permutations in S_n with exactly k cycles, and define

$$s(n,k) = (-1)^{n-k} u(n,k).$$

Then

$$\sum_{k=1}^{n} |s(n,k)| = \sum_{k=1}^{n} u(n,k) = n!,$$

$$s(n,n) = 1, \qquad s(n,1) = (-1)^{n-1}(n-1)!,$$

$$s(n,k) = s(n-1,k-1) - (n-1)s(n-1,k),$$

and

$$(x)_n = \sum_{k=1}^{n} s(n,k) x^k.$$

Multinomial Coefficients. For non-negative integers $k_1, \ldots, k_r$ with $\sum_{i=1}^{r} k_i = n$, the multinomial coefficient is

$$\binom{n}{k_1, \ldots, k_r} = \binom{n}{k_1}\binom{n-k_1}{k_2} \cdots \binom{n-k_1-\cdots-k_{r-1}}{k_r} = \frac{n!}{k_1! \cdots k_r!}.$$

Multinomial Theorem. For $n \in \mathbb{N}_0$,

$$(x_1 + \cdots + x_r)^n = \sum_{k_1+\cdots+k_r=n} \binom{n}{k_1, \ldots, k_r} x_1^{k_1} \cdots x_r^{k_r}.$$

Generalized Pascal Formula. For $k_1, \ldots, k_r \geq 0$ with $\sum_{i=1}^{r} k_i = n$,

$$\binom{n}{k_1, \ldots, k_r} = \sum_{i=1}^{r} \binom{n-1}{k_1, \ldots, k_i - 1, \ldots, k_r}.$$

Permutations of a Multiset. Let X be a multiset with r types having repetition numbers $k_1, \ldots, k_r$, where $\sum_{i=1}^{r} k_i = n$. Then the number of distinct n-permutations of X is

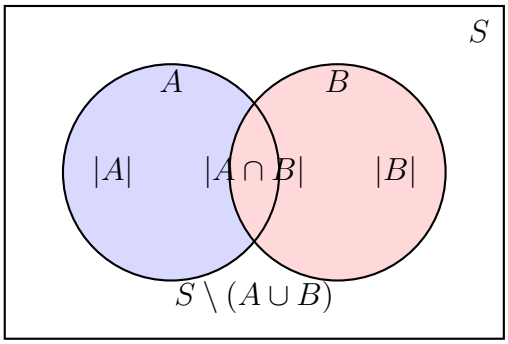
$$\binom{n}{k_1, \ldots, k_r}.$$

2 Inclusion and Exclusion

The principle of inclusion and exclusion is one of the most powerful methods for counting. The basic idea is already familiar from two sets: if $A, B \subseteq S$ and we want to count the elements in

$$(A \cup B)^C = S \setminus (A \cup B),$$

then this number is not always given by $|S| - |A| - |B|$, since the intersection $A \cap B$ is subtracted twice.

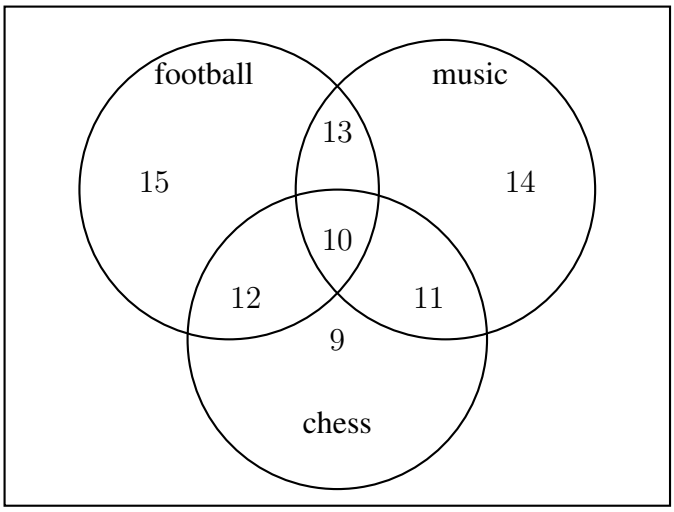


Instead, the correct formula is

$$|S \setminus (A \cup B)| = |S| - |A| - |B| + |A \cap B|.$$

Let us look at a slightly more complicated example. Assume we have 100 students in our class. Of these, 50 play football, 48 play music, and 42 play chess. Some students do more than one of these activities: 23 play both football and music, 22 play both football and chess, 21 play both music and chess, and 10 play all three.

How many students play none of the three?



By counting all students who participate in at least one activity, we obtain 84, and hence the number of students who do none of the three activities is 16.

In this chapter we develop a general formula for counting such numbers more systematically. It is called the *Principle of Inclusion and Exclusion*, or *PIE* for short.

2.1 PIE

We consider a finite *universal set* S together with subsets

$$A_1, A_2, \ldots, A_n \subseteq S.$$

In our example above, S was the set of 100 students, and A_1, A_2, A_3 were the sets of students participating in the three activities.

To simplify notation, for any subset $I \subseteq \{1, 2, \ldots, n\}$ we define

$$A_I := \bigcap_{i \in I} A_i.$$

Thus, A_I consists of all elements that lie in all sets indexed by I. By convention, we set $A_{\{i\}} = A_i$ and $A_\emptyset = S$.

Theorem 2.1 (Principle of Inclusion and Exclusion)**.** *Let n be a positive integer and S a finite set. Let $A_1, \ldots, A_n \subseteq S$. Then, the number of elements of S that lie in none of the sets $A_1, \ldots, A_n$, is given by*

$$\left| S \setminus \bigcup_{i=1}^{n} A_i \right| = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} \, |A_I|.$$

Proof. We evaluate the sum by considering the contribution of each element $x \in S$.

If x lies in none of the sets A_i, then it contributes 1 via the term $A_\emptyset = S$, and contributes nothing to any other term. Hence its total contribution is 1.

Now suppose that x lies in at least one of the sets A_i, and let

$$J := \{\, i \in \{1, \ldots, n\} \mid x \in A_i \,\}$$

be the set of indices of sets containing x. Then $x \in A_I$ if and only if $I \subseteq J$. Hence the total contribution of x is

$$\sum_{I \subseteq J} (-1)^{|I|}.$$

If $|J| = j$, then there are $\binom{j}{i}$ subsets of J of size i, and thus

$$\sum_{I \subseteq J} (-1)^{|I|} = \sum_{i=0}^{j} \binom{j}{i} (-1)^i = (1-1)^j = 0,$$

by the Binomial Theorem.

Thus, elements lying in at least one of the sets contribute 0, while elements lying in none of the sets contribute 1. It follows that the sum counts exactly the elements in none of the sets. $\square$

Let us apply this formula to the example from before. Here $n = 3$, $|A_\emptyset| = |S| = 100$, $|A_1| = 50$, $|A_2| = 48$, $|A_3| = 42$, and

$$|A_{\{1,2\}}| = 23, \qquad |A_{\{1,3\}}| = 22, \qquad |A_{\{2,3\}}| = 21, \qquad |A_{\{1,2,3\}}| = 10.$$

Thus

$$\begin{aligned}\left|S \setminus \bigcup_{i=1}^{3} A_i\right| &= |S| - |A_1| - |A_2| - |A_3| + |A_{\{1,2\}}| + |A_{\{1,3\}}| + |A_{\{2,3\}}| - |A_{\{1,2,3\}}| \\ &= 100 - 50 - 48 - 42 + 23 + 22 + 21 - 10 \\ &= 16.\end{aligned}$$

Note that a student participating in all three activities contributes

$$1 - 3 + 3 - 1 = 0,$$

while a student participating in none contributes 1. This is exactly the mechanism behind the proof of Theorem 2.1.

The theorem gives a general formula, but in concrete problems the main challenge is to choose the right sets. A useful rule of thumb is the following:

1. **Define bad properties:** Identify a universal set S containing all objects under consideration. Describe the objects you want to count as those having none of a collection of "bad" properties $P_1, \ldots, P_n$.

2. **Count the intersections:** For each i, let $A_i \subseteq S$ be the set of all objects having property P_i. Then the desired set is

$$S \setminus \bigcup_{i=1}^{n} A_i.$$

   For each subset $I \subseteq \{1, \ldots, n\}$, compute the size of

$$A_I = \bigcap_{i \in I} A_i.$$

3. **Apply PIE:** Insert these intersection sizes into the formula from Theorem 2.1.

Proposition 2.2. *Let n be a positive integer and let $A_1, \ldots, A_n \subseteq S$ be subsets of a finite set S with $|S| = m_0$. Suppose that*

$$|A_i| = m_1 \quad \textit{for all } i,$$

and more generally, that the intersection of any j distinct sets has size m_j. Then the number of elements lying in none of the sets is

$$\left| S \setminus \bigcup_{i=1}^{n} A_i \right| = \sum_{j=0}^{n} (-1)^j \binom{n}{j} m_j.$$

Proof. In the formula of the inclusion–exclusion principle, there are $\binom{n}{j}$ subsets I with $|I| = j$. By assumption, each corresponding set A_I has size m_j. Hence their total contribution is

$$(-1)^j \binom{n}{j} m_j,$$

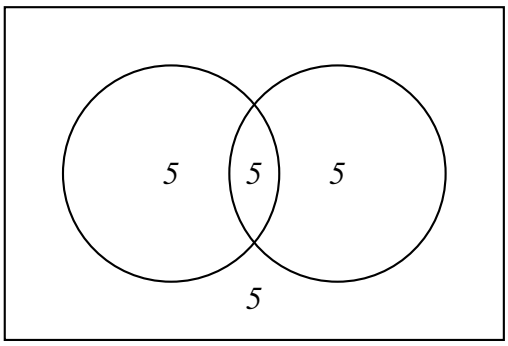
and summing over all j gives the result. □

Example 2.3. *Let us illustrate this result with an example. Let $|S| = m_0 = 20$, let $|A_1| = |A_2| = 10$, and let*

$$|A_{\{1,2\}}| = |A_1 \cap A_2| = 5.$$

Then

$$|S \setminus (A_1 \cup A_2)| = \sum_{j=0}^{2} (-1)^j \binom{2}{j} m_j = 1 \cdot 20 - 2 \cdot 10 + 1 \cdot 5 = 5.$$



Within the proof of the Principle of Inclusion and Exclusion, we have already seen another useful identity.

Lemma 2.4 (Cancellation Lemma)**.** *Let S be a finite set. Then*

$$\sum_{I \subseteq S} (-1)^{|I|} = \begin{cases} 1 & \textit{if } S = \emptyset, \\ 0 & \textit{otherwise.} \end{cases}$$

Exercise 2.5. *Prove Lemma 2.4 using the Binomial Theorem.*

As a first further application, we derive an identity for binomial coefficients.

Proposition 2.6. *Let m, n, k be non-negative integers. Then*

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{m+n-i}{k-i} = \begin{cases}\binom{m}{k} & \text{if } m \geq k,\\ 0 & \text{if } m < k.\end{cases}$$

Proof. Let $X = \{x_1, \ldots, x_n\}$ be a set of n blue points, let Y be a set of m red points, and let

$$Z = X \cup Y.$$

We count the k-element subsets of Z containing only red points.

If $m < k$, then this is impossible, so the number is 0. If $m \geq k$, then we simply choose k elements from Y, so the number is

$$\binom{m}{k}.$$

On the other hand, let S be the set of all k-element subsets of Z, and for each $i \in \{1, \ldots, n\}$ let

$$A_i = \{A \subseteq Z \mid |A| = k,\ x_i \in A\}.$$

Then a k-element subset contains only red points if and only if it lies in none of the sets A_i.

If $I \subseteq \{1, \ldots, n\}$ with $|I| = i$, then A_I consists of all k-element subsets that contain all blue points indexed by I. Hence we still need to choose the remaining $k-i$ elements from the other $m+n-i$ points, and therefore

$$|A_I| = \binom{m+n-i}{k-i}.$$

By the Principle of Inclusion and Exclusion,

$$\left|S \setminus \bigcup_{i=1}^{n} A_i\right| = \sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{m+n-i}{k-i}.$$

Comparing the two counts gives the claim. □

2.2 A first application: Surjections and Stirling Numbers

We now illustrate the general method with a first important application. The problem of counting surjective functions is difficult to solve directly, but becomes very natural once we describe non-surjective functions in terms of bad properties.

Let A and B be finite sets with $|A| = n$ and $|B| = k$.

How many functions $f : A \to B$ are there? To define such a function, we can independently choose the value $f(a)$ for each $a \in A$. Since each choice has k possibilities, we obtain k^n functions in total.

How many of these functions are injective? In this case, all values must be distinct, so we are choosing n distinct elements of B in order. Thus the number of injective functions is given by the falling factorial

$$(k)_n = k(k-1)\cdots(k-n+1).$$

In particular, this number is zero if $n > k$.

How many of these functions are surjective? This is a natural application of the Principle of Inclusion and Exclusion.

Theorem 2.7. *Let n, k be positive integers. Then, the number of surjections from an n-element set to a k-element set is*

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

Proof. We apply the three-step procedure described above.

First, let

$$S = \{\, f : \{1, \ldots, n\} \to \{1, \ldots, k\} \,\}$$

be the set of all functions. Thus

$$|S| = k^n.$$

The bad property P_i, for $i \in \{1, \ldots, k\}$, is that i does not lie in the image of f. Let $A_i \subseteq S$ denote the set of all functions with this property. Then a function is surjective if and only if it has none of the bad properties, that is, if and only if it lies in

$$S \setminus \bigcup_{i=1}^{k} A_i.$$

Now let $I \subseteq \{1, \ldots, k\}$. Then a function f lies in

$$A_I = \bigcap_{i \in I} A_i$$

if and only if none of the values in I occur as images of f. Equivalently,

$$f : \{1, \ldots, n\} \to \{1, \ldots, k\} \setminus I.$$

Since the codomain has size $k - |I|$, we obtain

$$|A_I| = (k - |I|)^n.$$

By the Principle of Inclusion and Exclusion,

$$\left| S \setminus \bigcup_{i=1}^{k} A_i \right| = \sum_{I \subseteq \{1,\ldots,k\}} (-1)^{|I|} |A_I|.$$

Grouping together all subsets I of the same size j, we get

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

This proves the formula. □

Example 2.8. *Consider functions from a* 3*-element set to a* 2*-element set. The number of surjections is*

$$\sum_{j=0}^{2} (-1)^j \binom{2}{j} (2-j)^3 = 2^3 - 2 \cdot 1^3 + 1 \cdot 0^3 = 8 - 2 = 6.$$

Indeed, these are exactly the functions where both elements of the target are used.

The formula is useful, but might seem a bit complicated. The following two corollaries provide natural sanity checks.

Corollary 2.9. *Let n be a positive integer. Then,*

$$n! = \sum_{j=0}^{n} (-1)^j \binom{n}{j} (n-j)^n.$$

Proof. If $k = n$, then a function

$$f : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

is surjective if and only if it is bijective. Hence the number of surjections is $|S_n| = n!$.

On the other hand, by Theorem 2.7, the number of surjections is

$$\sum_{j=0}^{n} (-1)^j \binom{n}{j} (n-j)^n.$$

Equating the two expressions gives the result. □

Corollary 2.10. *Let k, n be positive integers. If $k > n$, then*

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

Proof. If $k > n$, then there is no surjection from an n-element set to a k-element set, since the codomain is larger than the domain. Hence the number of surjections is 0.

On the other hand, by Theorem 2.7, the number of surjections is

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

Therefore this sum must be equal to 0. □

The formula for surjections has an immediate consequence for Stirling numbers of the second kind, since surjective functions and set partitions are closely related.

Recall that the Stirling number of the second kind, $S(n,k)$, denotes the number of ways to partition an n-element set into k non-empty parts.

Proposition 2.11. *Let $k \leq n$ be positive integers. Then, the Stirling numbers of the second kind satisfy*

$$S(n,k) = \frac{1}{k!}\sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n.$$

Proof. Let $A = \{1,\ldots,n\}$ and $B = \{1,\ldots,k\}$. We count the number of surjections $f : A \to B$ in two different ways.

On the one hand, by Theorem 2.7, the number of surjections is

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

On the other hand, every surjection $f : A \to B$ induces a partition of A into k non-empty parts given by the inverse images

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

Conversely, given a partition into k parts, we obtain a surjection by assigning distinct labels $1,\ldots,k$ to the parts.

Each partition gives rise to $k!$ different surjections, depending on how the parts are labelled. Hence the total number of surjections is

$$k!\,S(n,k).$$

Equating the two expressions and dividing by $k!$ gives the result. □

2.3 Derangements

We now turn to another classical application of the Principle of Inclusion and Exclusion: counting permutations with no fixed points.

Definition 2.12. A *derangement* is a permutation $\sigma \in S_n$ such that

$$\sigma(i) \neq i \quad \text{for all } i \in \{1, \ldots, n\}.$$

We denote the set of derangements in S_n by D_n.

Example 2.13. *For $n = 2$, there is exactly one derangement, namely the permutation swapping 1 and 2. For $n = 3$, the derangements are*

$$(1 \mapsto 2,\ 2 \mapsto 3,\ 3 \mapsto 1) \quad \textit{and} \quad (1 \mapsto 3,\ 2 \mapsto 1,\ 3 \mapsto 2),$$

so $|D_3| = 2$.

Theorem 2.14. *Let n be a positive integer. Then, the number of derangements is*

$$|D_n| = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)!.$$

Proof. We proceed as in the general method for applying the Principle of Inclusion and Exclusion.

First, let

$$S = S_n$$

be the set of all permutations of $\{1, \ldots, n\}$. Thus $|S| = n!$.

For each $i \in \{1, \ldots, n\}$, let P_i be the property that i is a fixed point, i.e. $\sigma(i) = i$. Let

$$A_i = \{\, \sigma \in S_n \mid \sigma(i) = i \,\}$$

be the set of permutations having this property.

A permutation is a derangement if and only if it has none of the properties P_i, that is,

$$D_n = S \setminus \bigcup_{i=1}^{n} A_i.$$

Now let $I \subseteq \{1, \ldots, n\}$. Then a permutation σ lies in

$$A_I = \bigcap_{i \in I} A_i$$

if and only if all elements of I are fixed points of σ.

Hence σ is an arbitrary permutation on the remaining $n - |I|$ elements. Therefore,

$$|A_I| = (n - |I|)!.$$

By the Principle of Inclusion and Exclusion, we obtain

$$|D_n| = \sum_{I \subseteq \{1,\dots,n\}} (-1)^{|I|} |A_I| = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)!,$$

where we group subsets I by their size. □

Corollary 2.15. *Let n be a positive integer. Then, we have*

$$|D_n| = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}.$$

Proof. From the theorem we have

$$|D_n| = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)!.$$

Using

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

we obtain

$$|D_n| = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}.$$

□

2.4 Euler-Totient Function and Möbius Inversion

The Principle of Inclusion and Exclusion is not only useful in combinatorics. It also appears naturally in number theory, where it leads to a formula for Euler's totient function and, more generally, to Möbius inversion.

We begin with the residue ring $\mathbb{Z}/n\mathbb{Z}$. This is a field if and only if n is prime. For composite n, there exist non-zero elements that are zero-divisors, and such elements are not invertible. The invertible elements are called *units*, and the set of all units is denoted by $(\mathbb{Z}/n\mathbb{Z})^\star$.

Example 2.16. *Let us consider $\mathbb{Z}/4\mathbb{Z}$. Here we have*

$$1 \cdot 1 = 1 \qquad \textit{and} \qquad 3 \cdot 3 = 1,$$

so 1 and 3 are units. However, 2 is a zero-divisor since

$$2 \cdot 2 = 0$$

in $\mathbb{Z}/4\mathbb{Z}$, and hence it is not a unit.

Recall that by Bézout's identity, $\gcd(k, n) = 1$ if and only if there exist integers a, b such that

$$ak + bn = 1.$$

This is exactly the condition for k to admit a multiplicative inverse modulo n.

Proposition 2.17. *Let $n > 1$ be a positive integer. Then*

$$(\mathbb{Z}/n\mathbb{Z})^\star = \{k \in \{1, \ldots, n\} \mid \gcd(k, n) = 1\}.$$

Proof. If $k \in (\mathbb{Z}/n\mathbb{Z})^\star$, then there exists $\ell \in \mathbb{Z}/n\mathbb{Z}$ such that

$$k\ell \equiv 1 \pmod{n}.$$

Hence there exists $m \in \mathbb{Z}$ such that

$$k\ell - nm = 1.$$

Thus 1 is an integer linear combination of k and n, and by Bézout's identity,

$$\gcd(k, n) = 1.$$

Conversely, if $\gcd(k, n) = 1$, then by Bézout's identity there exist $a, b \in \mathbb{Z}$ such that

$$ak + bn = 1.$$

Reducing modulo n, we obtain

$$ak \equiv 1 \pmod{n},$$

so k is invertible in $\mathbb{Z}/n\mathbb{Z}$. □

Definition 2.18. Let $n > 1$ be a positive integer. The *Euler totient function* counts the number of units in $\mathbb{Z}/n\mathbb{Z}$, that is,

$$\varphi(n) = \left|(\mathbb{Z}/n\mathbb{Z})^\star\right|.$$

For $n = 1$, we set $\varphi(n) = 1$.

Proposition 2.19. *Let $n > 1$ be a positive integer with prime factorization*

$$n = \prod_{i=1}^{r} p_i^{a_i},$$

where the p_i are distinct primes and the a_i are positive integers. Then

$$\varphi(n) = n \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right).$$

Proof. We count the elements of $\{1, \ldots, n\}$ that are coprime to n using inclusion–exclusion. Let $S = \{1, \ldots, n\}$ and for each $i \in \{1, \ldots, r\}$ define

$$A_i = \{x \in S \mid p_i \mid x\}.$$

Then an element $x \in S$ is not coprime to n if and only if it lies in at least one of the sets A_i. Since multiples of p_i in $\{1, \ldots, n\}$ occur every p_i steps, we have

$$|A_i| = \frac{n}{p_i}.$$

More generally, for distinct indices $i_1, \ldots, i_k$, we have

$$|A_{i_1} \cap \cdots \cap A_{i_k}| = \frac{n}{p_{i_1} \cdots p_{i_k}},$$

since $p_{i_1} \cdots p_{i_k} \mid n$ and we are counting multiples of this product.
By the Principle of Inclusion and Exclusion,

$$\left| S \setminus \bigcup_{i=1}^{r} A_i \right| = n - \sum_{i=1}^{r} \frac{n}{p_i} + \sum_{1 \le i < j \le r} \frac{n}{p_i p_j} - \cdots + (-1)^r \frac{n}{p_1 \cdots p_r}.$$

Factoring out n, this becomes

$$n \left(1 - \sum_{i=1}^{r} \frac{1}{p_i} + \sum_{i<j} \frac{1}{p_i p_j} - \cdots + (-1)^r \frac{1}{p_1 \cdots p_r} \right).$$

The expression in parentheses is exactly the expansion of the product

$$\prod_{i=1}^{r} \left(1 - \frac{1}{p_i} \right),$$

which proves the claim. □

One of the most important results in this direction is the following identity of Gauss.

Theorem 2.20 (Gauss Lemma)**.** *Let n be a positive integer. Then*

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

Proof. We partition the set $\{1, \ldots, n\}$ according to the value of $\gcd(m, n)$.
For each divisor $d \mid n$, define

$$S_d := \{\, m \in \{1, \ldots, n\} \mid \gcd(m, n) = d \,\}.$$

Then the sets S_d are pairwise disjoint, and their union is all of $\{1, \ldots, n\}$. Hence

$$n = \sum_{d|n} |S_d|.$$

It remains to determine $|S_d|$. Fix a divisor $d \mid n$. If $m \in S_d$, then

$$\gcd(m, n) = d,$$

so we can write

$$m = dk$$

for some integer k with $1 \leq k \leq n/d$. Moreover,

$$\gcd(dk, n) = d \quad \Longleftrightarrow \quad \gcd\Big(k, \frac{n}{d}\Big) = 1.$$

Thus the map

$$m = dk \longmapsto k$$

gives a bijection between S_d and the set

$$\left\{ k \in \left\{1, \ldots, \frac{n}{d}\right\} \;\middle|\; \gcd\Big(k, \frac{n}{d}\Big) = 1 \right\}.$$

Therefore,

$$|S_d| = \varphi\Big(\frac{n}{d}\Big).$$

Substituting into the sum gives

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

Since d runs through all divisors of n if and only if n/d does, this is equivalent to

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

as claimed. □

Example 2.21. *Let us consider again $n = 4 = 2^2$. By Proposition 2.19, we have*

$$\varphi(4) = 4\left(1 - \frac{1}{2}\right) = 2 = |\{1, 3\}| = |(\mathbb{Z}/4\mathbb{Z})^\star|.$$

Additionally, we can check that

$$\sum_{d|4} \varphi(d) = \varphi(1) + \varphi(2) + \varphi(4) = 1 + 1 + 2 = 4.$$

Definition 2.22. The *Möbius function* is the function $\mu : \mathbb{N} \to \{0, \pm 1\}$ defined by

$$\mu(d) = \begin{cases} 1 & \text{if } d \text{ is the product of an even number of distinct primes (including } d = 1), \\ -1 & \text{if } d \text{ is the product of an odd number of distinct primes,} \\ 0 & \text{if } d \text{ is not square-free.} \end{cases}$$

With this function, we can now "invert" Theorem 2.20.

Theorem 2.23. *Let n be a positive integer. Then*

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

Proof. If $n = 1$, then

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

Now let

$$n = \prod_{i=1}^{r} p_i^{a_i}$$

with distinct primes p_i and positive integers a_i. Only square-free divisors contribute to the sum, and these correspond exactly to choosing subsets of $\{p_1, \ldots, p_r\}$.

Thus,

$$\sum_{d|n} \mu(d) = \sum_{I \subseteq \{1,\ldots,r\}} (-1)^{|I|} = \sum_{i=0}^{r} \binom{r}{i} (-1)^i = (1-1)^r = 0,$$

if $r \geq 1$. □

This shows that μ acts as an inverse to summation over divisors. More generally, we obtain the following inversion formula.

Theorem 2.24 (Möbius Inversion Formula)**.** *Let f, g be functions on the positive integers such that*

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

Then

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

Proof. We compute

$$\begin{aligned}\sum_{d|n} \mu(d) f(n/d) &= \sum_{d|n} \mu(d) \sum_{d'|(n/d)} g(d') \\ &= \sum_{d'|n} g(d') \sum_{d|(n/d')} \mu(d).\end{aligned}$$

By Theorem 2.23, the inner sum is 1 if $n/d' = 1$, that is, if $d' = n$, and 0 otherwise. Hence the entire sum reduces to

$$g(n).$$

□

Example 2.25. *Let $g(n) = 1$ for all n. Then*

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

the number of divisors of n.

Applying Möbius inversion gives

$$1 = g(n) = \sum_{d|n} \mu(d)\, \tau(n/d).$$

For example, for $n = 6$:

$$\tau(6) = 4, \quad \tau(3) = 2, \quad \tau(2) = 2, \quad \tau(1) = 1,$$

and

$$\sum_{d|6} \mu(d)\, \tau(6/d) = 1 \cdot 4 - 1 \cdot 2 - 1 \cdot 2 + 1 \cdot 1 = 1.$$

Exercise 2.26. *We color the integers $\{1, \ldots, 2n\}$ either red or blue, such that if i is red, then $i - 1$ is not blue. Prove that*

$$\sum_{k=0}^{n} (-1)^k \binom{2n-k}{k} 2^{2n-2k} = 2n + 1.$$

3 Generating Functions

This chapter is written and taught by the guest lecturer Hugo Beeloo-Sauerbier Couvée

Generating functions are among the most useful tools in enumerative combinatorics. They allow us to encode a sequence into a formal power series and to translate combinatorial or recursive information into algebraic identities. In many situations, this makes it possible to derive exact formulas, but also in less fortunate situations, they might still enable us to solve recurrences or to determine asymptotic behavior. More precisely: generating functions are useful for (but not limited to) the following main reasons:

- they can lead to the *exact computation* of sequences, via deriving closed formulas or recurrence relations,
- they can help us in the *approximation* of a sequence, such as revealing its asymptotic growth,
- they allow us to *prove combinatorial identities* and *find relationships* between sequences,
- they can be used for *extracting information* or properties from a sequence, for example via Fourier- and other transforms.

For other extensive coverage on generating functions, we recommend *Analytic Combinatorics* by Flajolet and Sedgewick [1], available online `https://tinyurl.com/analyticcomb` and *generatingfunctionology* by Wilf [4], available online `https://tinyurl.com/gfology`.

3.1 A first motivating example: Fibonacci numbers

We begin with the Fibonacci sequence with $f_0 = 0, f_1 = 1$ and for $n \geq 2$

$$f_n = f_{n-1} + f_{n-2}.$$

There are several ways to solve this recurrence. One classical method is to look for solutions of the form

$$a_n = \alpha^n.$$

Substituting this into the recurrence gives

$$\alpha^n = \alpha^{n-1} + \alpha^{n-2},$$

or equivalently

$$\alpha^2 = \alpha + 1.$$

Thus α must be a root of the characteristic equation

$$x^2 = x + 1.$$

Its two roots are

$$\alpha = \frac{1+\sqrt{5}}{2}, \qquad \beta = \frac{1-\sqrt{5}}{2}.$$

Hence both sequences (α^n) and (β^n) satisfy the recurrence, and by linearity so does any linear combination

$$a_n = p\alpha^n + q\beta^n.$$

Using the initial conditions one obtains

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

We now solve the same recurrence by generating functions. Define

$$f(z) = \sum_{n\geq 0} f_n z^n.$$

Then

$$\begin{aligned} f(z) &= z + \sum_{n\geq 2} f_n z^n \\ &= z + \sum_{n\geq 2} (f_{n-1} + f_{n-2}) z^n \\ &= z + z\sum_{n\geq 2} f_{n-1} z^{n-1} + z^2 \sum_{n\geq 2} f_{n-2} z^{n-2} \\ &= z + zf(z) + z^2 f(z). \end{aligned}$$

Rearranging yields

$$(1 - z - z^2) f(z) = z,$$

hence

$$f(z) = \frac{z}{1 - z - z^2}.$$

This already illustrates the central idea: a recurrence relation for a sequence becomes an algebraic identity for the corresponding generating function.

By partial fraction decomposition, one recovers the explicit formula above. Moreover, since

$$|\beta| < 1,$$

the second term tends exponentially to zero, so

$$f_n \sim \frac{1}{\sqrt{5}}\alpha^n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

We will see later in Section 3.16 that this approximation as an exponential function is directly obtainable from $f(z)$: the exponential growth factor is determined by the reciprocal of the smallest pole of $f(z)$, which happens to be $1/\alpha$. Thus generating functions help not only with exact formulas, but also with asymptotic estimates.

3.2 Combinatorial classes and counting sequences

To use generating functions systematically, it is useful to introduce the language of combinatorial classes.

Definition 3.1. A *combinatorial class*, or simply a *class*, is a finite or countable set $\mathcal{A}$ together with a *size function*

$$|\cdot| : \mathcal{A} \to \mathbb{Z}_{\geq 0}$$

such that for every $n \geq 0$, the set

$$\mathcal{A}_n := \{\alpha \in \mathcal{A} \ : \ |\alpha| = n\}$$

is finite.

Definition 3.2. The *counting sequence* of a combinatorial class $\mathcal{A}$ is the sequence

$$A_n := |\mathcal{A}_n|, \qquad n \geq 0.$$

Example 3.3 (Binary words)**.** *Let $\mathcal{W}$ be the set of binary words over the alphabet $\{0, 1\}$, including the empty word ε. If the size of a word is its length, then*

$$W_n = 2^n,$$

since each of the n positions can be filled in two ways.

Definition 3.4. Two combinatorial classes $\mathcal{A}$ and $\mathcal{B}$ are called *isomorphic* if they have the same counting sequence. Equivalently, there exists a bijection between them that preserves size.

3.3 Ordinary generating functions

Definition 3.5. The *ordinary generating function* (OGF) of a sequence $(A_n)_{n\geq 0}$ is the formal power series

$$A(z) = \sum_{n\geq 0} A_n z^n.$$

If (A_n) is the counting sequence of a combinatorial class $\mathcal{A}$, then $A(z)$ is also called the OGF of $\mathcal{A}$.

Equivalently,

$$A(z) = \sum_{\alpha\in\mathcal{A}} z^{|\alpha|}.$$

Viewing the OGF in this way, the size function is 'encoded' as powers of the variable z.

For the class of binary words we obtain

$$W(z) = \sum_{n\geq 0} 2^n z^n = \frac{1}{1-2z}.$$

It is important to note that a generating function is in many contexts not thought of as something into which we necessarily plug numbers. Rather, it is best viewed as a *formal power series*. In other words, we manipulate it algebraically, coefficient by coefficient, without always needing to worry about convergence (more on this in Section 3.5).

At first sight, this may seem like a strange way to write a sequence. But many operations on sequences have very natural interpretations in terms of generating functions:

- shifting a sequence corresponds to multiplication by z,
- adding sequences corresponds to adding generating functions,
- convolution corresponds to multiplication of generating functions.

Coefficient extraction

We write

$$[z^n]F(z)$$

for the coefficient of z^n in the formal power series $F(z)$.

For example, if

$$F(z) = 1 + 3z + 5z^2 + \cdots,$$

then

$$[z^2]F(z) = 5.$$

3.4 Formal power series and allowed operations

We work generally in the ring $\mathbb{C}[[z]]$ of formal power series

$$F(z) = \sum_{n \geq 0} f_n z^n.$$

In this setting we may:

- add series termwise,
- multiply series using the Cauchy product,
- compose series whenever the inner series has constant term zero,
- (fomally) differentiate termwise,
- shift series to the right;
- invert a series whenever its constant term is nonzero.

Explicitly, if

$$F(z) = \sum_{n\geq 0} f_n z^n, \qquad G(z) = \sum_{n\geq 0} g_n z^n,$$

then the main operations involving the two are

$$\begin{aligned}
\text{Addition:} \quad & F(z) + G(z) = \sum_{n\geq 0} (f_n + g_n)\, z^n, \\
\text{Multiplication:} \quad & F(z)G(z) = \sum_{n\geq 0} \left(\sum_{k=0}^{n} f_k g_{n-k} \right) z^n, \\
\text{Substitution:} \quad & F(G(z)) = \sum_{n\geq 0} f_n \left(G(z)\right)^n \quad \text{if } \; g_0 = 0.
\end{aligned}$$

The multiplication rule is central in combinatorics: it expresses the fact that an object of total size n can often be decomposed into a first part of size k and a second part of size $n-k$. Some important operations involving only one function are the following:

$$\begin{aligned}
\text{Shifting to the right:} \quad & z^r F(z) = \sum_{n\geq r} f_{n-r} z^n \quad \text{for } r \in \mathbb{N}, \\
\text{Formal derivative:} \quad & F'(z) = \sum_{n\geq 0} (n+1) f_{n+1} z^n, \\
\text{Multiplicative inverse:} \quad & \frac{1}{F(z)} = \sum_{n\geq 0} a_n z^n \quad \text{if } f_0 \neq 0, \\
& \text{where } a_0 = \frac{1}{f_0} \text{ and } a_n = -\frac{1}{f_0} \sum_{k=0}^{n-1} a_k f_{n-k}.
\end{aligned}$$

Example 3.6 (Geometric series)**.** *The sequence $f_n = 1$ for all $n \geq 0$ has generating function*

$$F(z) = 1 + z + z^2 + \cdots = \sum_{n\geq 0} z^n.$$

As a formal power series, the equality

$$(1-z)F(z) = 1,$$

holds, so

$$F(z) = \frac{1}{1-z}.$$

Differentiating yields

$$F'(z) = \frac{1}{(1-z)^2} = \sum_{n\geq 0} (n+1) z^n.$$

3.5 Analytic functions and their coefficients

We can ask whether the formal derivative matches the analytically defined (complex) derivative of a function, as we implicitly used in Example 3.6. To justify this properly, we need to briefly revisit some complex analysis.

A complex function $F(z)$ is *analytic at 0* if there is an open disc centered at 0 on which $F(z)$ is representable as a convergent power series $\sum_{n\geq 0} f_n z^n$. The *radius of convergence* is given by

$$R = \sup\{r \in \mathbb{R}_{\geq 0} \mid F(z) \text{ converges for all } z \in \mathbb{C} \text{ with } |z| \leq r\},$$

and analytic functions have a strictly positive radius of convergence $R > 0$ (possibly $R = \infty$). It is well-known from complex analysis that analytic functions coincide with *complex-differentiable* functions or *holomorphic* functions, see [1, Thm IV.1]. Hence, if we know that $F(z)$ is analytic with non-zero radius of convergence, it is complex differentiable, and thus the formal and complex derivatives coincide, and unambiguously write $F'(z)$. As example, the function

$$F(z) = \frac{1}{1-z} = \sum_{n\geq 0} z^n$$

has radius of convergence $R = 1$, and hence we know that the complex derivative $\left(\frac{1}{1-z}\right)' = \frac{1}{(1-z)^2}$ coincides with the formal derivative $\left(\sum_{n\geq 0} z^n\right)' = \sum_{n\geq 0}(n+1)z^n$.

Remark 3.7. It is not always necessary to check the radius of convergence of a function, in order to apply complex derivatives. In the case above, we could also first assume that we are allowed to differentiate, and later justify our formula by checking that indeed $(1-z)^2 \sum_{n\geq 0}(n+1)z^n = 1$ holds as formal power series in $\mathbb{C}[[z]]$.

We have an explicit relationship between the coefficients of a power series and its radius of convergence, which can be used to check analyticity of a function, and will also be important later in Section 3.16 in investigating the asymptotic growth of a series' coefficients.

Theorem 3.8 (Cauchy-Hadamard theorem)**.** *The radius of convergence R of the power series $\sum_{n\geq 0} a_n z^n$ is given by*

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{1/n}}$$

where $R = 0$ if the lim sup *diverges to ∞, and $R = \infty$ if the* lim sup *is* 0.

Furthermore, one of the many nice features of analytic functions is the existence of a Taylor series:

Theorem 3.9 (Taylor series)**.** *Let $F(z)$ be analytic at 0, and denote its n-th derivative by $F^{(n)}(z)$. Then*

$$F(z) = \sum_{n\geq 0} \frac{F^{(n)}(0)}{n!} z^n.$$

Corollary 3.10 (Generalized binomial theorem)**.** *Let $\alpha \in \mathbb{C}$. Then*

$$(1+z)^\alpha = \sum_{n \geq 0} \binom{\alpha}{n} z^n,$$

where for $n \in \mathbb{Z}$ the generalized binomial coefficients are defined as

$$\binom{\alpha}{n} = \begin{cases} \frac{\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-(n-1))}{n!} \text{ if } n > 0 \\ 1 \text{ if } n = 0 \\ 0 \text{ if } n < 0 \end{cases}.$$

Remark 3.11. Note the often useful *'combinatorial reciprocity'* identity

$$\binom{-\alpha}{n} = (-1)^n \binom{\alpha+n-1}{n}$$

between binomial coefficients involving $-\alpha$ and coefficients involving α.

Lastly, we provide a famous and powerful tool for calculating the coefficients of the *compositional inverse* of a function of the form $z/\varphi(z)$.

Theorem 3.12 (Lagrange inversion formula theorem (LIFT))**.** *Let $\varphi(z)$ be analytic at 0 with $\varphi(0) \neq 0$. Then on some open disc around 0, there exists a unique function $g(z)$ that satisfies*

$$z = \frac{g(z)}{\varphi(g(z))}$$

(i.e., $g(z)$ is the compositional inverse of $z/\varphi(z)$), and $g(z)$ is analytic at 0 with coefficients

$$[z^n]g(z) = \frac{1}{n}[z^{n-1}]\varphi(z)^n$$

for $n \geq 1$ and constant term $g(0) = 0$. As generalization, if $F(z)$ is any function analytic at 0, then

$$[z^n]F(g(z)) = \frac{1}{n}[z^{n-1}]\left(F'(z)\varphi(z)^n\right).$$

This is also known as the Lagrange-Bürmann formula*.*

We will see a nice application of this theorem in the next section on Catalan numbers.

3.6 Catalan numbers

Generating functions are particularly useful for linear and non-linear recurrences. A classical example is provided by the Catalan numbers.

Let C_n be a sequence satisfying

$$C_0 = 1,$$

and for $n > 0$ the recurrence:

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

Define

$$C(z) = \sum_{n \geq 0} C_n z^n.$$

Then

$$zC(z)^2 = C(z) - 1.$$

Indeed, for $n \geq 1$, the coefficient of z^n in $zC(z)^2$, equal to the coefficient of z^{n-1} in $C(z)^2$, is precisely

$$\sum_{k=0}^{n-1} C_k C_{n-1-k} = C_n.$$

Solving the quadratic equation gives

$$C(z) = \frac{1 - \sqrt{1-4z}}{2z}.$$

Recognizing that $\sqrt{1-4z}$ is a binomial series with power $\alpha = \frac{1}{2}$, yields

$$C_n = \frac{1}{n+1}\binom{2n}{n},$$

with first few terms

$$1, 1, 2, 5, 14, 42, 132, \ldots$$

As an alternative method, since

$$z = \frac{C(z) - 1}{C(z)^2},$$

we note that $g(z) = C(z) - 1$ is the compositional inverse of the function $z/\varphi(z)$ with $\varphi(z) = (1+z)^2$ satisfying $z = g(z)/\varphi(g(z))$. Hence by the LIFT (Theorem 3.12), we immediately get

$$C_n = \frac{1}{n}[z^{n-1}]\varphi(z)^n = \frac{1}{n}[z^{n-1}](1+z)^{2n} = \frac{1}{n}\binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n}.$$

Catalan numbers also appear naturally in lattice path counting. Consider paths from $(0,0)$ to (n,n) consisting of steps $(1,0)$ and $(0,1)$. There are $\binom{2n}{n}$ such paths in total.
If we restrict to paths that never go above the diagonal $y = x$, then the number of such paths is exactly the Catalan number C_n.

Another way to view Catalan numbers is through non-crossing partitions.

Non-crossing partitions Place the elements of $[n] = \{1, \ldots, n\}$ on a circle in cyclic order. Two disjoint subsets $A, B \subseteq [n]$ are said to be *crossing* if there exist

$$i < j < k < \ell$$

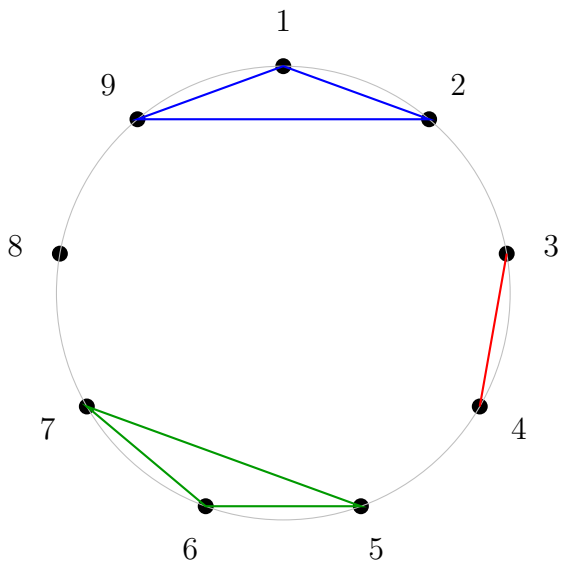
such that $\{i, k\} \subseteq A$ and $\{j, \ell\} \subseteq B$.

Definition 3.13. A partition of $[n]$ is called *non-crossing* if no two of its parts are crossing.

Example 3.14. *Let us consider the partition*

$$\big\{\{1,2,9\},\{3,4\},\{5,6,7\},\{8\}\big\}$$

of $[9]$. *Representing the numbers on a circle, we obtain the following non-crossing partition:*



Let NC_n denote the number of non-crossing partitions of $[n]$.

Theorem 3.15. *For all* $n \geq 0$,

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

Proof. We show that (NC_n) satisfies the same recursion as the Catalan numbers.

The initial values are easily checked:

$$NC_0 = 1, \quad NC_1 = 1, \quad NC_2 = 2.$$

Now consider a non-crossing partition P of $[n+1]$. Let S be the part containing $n+1$, and let k be the largest element of S other than $n+1$ (or $k = 0$ if $S = \{n+1\}$).

Then every other part of P lies entirely in $\{1, \ldots, k\}$ or entirely in $\{k+1, \ldots, n\}$, since otherwise it would cross S.

Thus P decomposes uniquely into:

- a non-crossing partition of $[k]$,
- a non-crossing partition of $\{k+1, \ldots, n\}$.

This gives the recursion

$$NC_{n+1} = \sum_{k=0}^{n} NC_k \, NC_{n-k}.$$

Since the initial values agree with those of the Catalan numbers, the claim follows. □

3.7 Combining OGFs: addition and multiplication

Generating functions do much more than merely encode a sequence. They turn combinatorial problems into algebraic and analytic ones, where powerful tools become available. One of the main strengths of generating functions we discuss in this section, is that combinatorial constructions often translate into algebraic operations of the generating functions.

Xor If (a_n) and (b_n) are counting sequences of (disjoint) classes $\mathcal{A}$ and $\mathcal{B}$ with OGFs

$$A(z) = \sum_{n \geq 0} a_n z^n, \qquad B(z) = \sum_{n \geq 0} b_n z^n,$$

then the sum $A(z) + B(z)$ has coefficient

$$[z^n](A(z) + B(z)) = a_n + b_n.$$

Thus the coefficient of z^n in the sum counts ways of building an object from either a size n object from $\mathcal{A}$ **xor** (exclusive or) a size n object from $\mathcal{B}$. This makes the disjoint union $\mathcal{A} \sqcup \mathcal{B}$ into a new combinatorial class with size function

$$|c|_{\mathcal{A}\sqcup\mathcal{B}} := \begin{cases} |c|_{\mathcal{A}} & \text{if } c \in \mathcal{A} \\ |c|_{\mathcal{B}} & \text{if } c \in \mathcal{B} \end{cases}$$

for $c \in \mathcal{A} \sqcup \mathcal{B}$, where $|\cdot|_{\mathcal{A}}$ and $|\cdot|_{\mathcal{B}}$ are the size functions for $\mathcal{A}$ and $\mathcal{B}$ respectively. The OGF corresponding to $\mathcal{A} \sqcup \mathcal{B}$ is the sum

$$A(z) + B(z).$$

And

Secondly, the product $A(z)B(z)$ has coefficient

$$[z^n]A(z)B(z) = \sum_{k=0}^{n} a_k b_{n-k}.$$

Thus the coefficient of z^n in the product counts ways of building an object of size n from a first part of size k **and** a second part of size $n-k$, for any choice of k. This makes the cartesian product $\mathcal{A} \times \mathcal{B}$, consisting of ordered pairs (a, b), into a new combinatorial class with additive size function

$$|(a, b)|_{\mathcal{A}\times\mathcal{B}} := |a|_{\mathcal{A}} + |b|_{\mathcal{B}},$$

and the OGF corresponding to $\mathcal{A} \times \mathcal{B}$ is the product

$$A(z)B(z).$$

This is directly proven by

$$\sum_{(a,b)\in\mathcal{A}\times\mathcal{B}} z^{|(a,b)|_{\mathcal{A}\times\mathcal{B}}} = \sum_{(a,b)\in\mathcal{A}\times\mathcal{B}} z^{|a|_{\mathcal{A}}+|b|_{\mathcal{B}}} = \sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}} z^{|a|_{\mathcal{A}}} z^{|b|_{\mathcal{B}}}.$$

Combinatorial construction	OGF	Combinatorial interpretation
Disjoint union $\mathcal{A} \sqcup \mathcal{B}$	$A(z) + B(z)$	$\mathcal{A}$ xor $\mathcal{B}$
Product $\mathcal{A} \times \mathcal{B}$	$A(z)B(z)$	$\mathcal{A}$ and $\mathcal{B}$

Table 1: The main constructions of disjoint union and product.

Example 3.16 (Balls in boxes)**.** *Let $a_n^{(k)}$ be the number of ways to distribute n identical balls into k labeled boxes, with at least one ball in each box. Let*

$$B^{(k)}(z) = \sum_{n\geq 0} a_n^{(k)} z^n.$$

For one box, we have

$$B^{(1)}(z) = z + z^2 + z^3 + \cdots = \frac{z}{1-z}.$$

If we split $k = s + t$, then a distribution into k boxes is obtained by first distributing some number of balls into the first s boxes and the rest into the remaining t boxes. We can say

*"distributing n balls into $s + t$ boxes equals distributing n_1 balls into s boxes **and** distributing n_2 balls into t boxes, such that $n_1 + n_2 = n$."*

Hence

$$B^{(s+t)}(z) = B^{(s)}(z)\, B^{(t)}(z),$$

and therefore

$$B^{(k)}(z) = \left(\frac{z}{1-z}\right)^k.$$

Expanding this gives

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

3.8 Combining OGF's: sequences and substitution

Tuples

Let $\mathcal{B}$ be a combinatorial class with OGF $B(z)$. The class of ***k*-tuples** of $\mathcal{B}$'s is the k-fold product:

$$\mathcal{B}^k := \underbrace{\mathcal{B} \times \mathcal{B} \times \cdots \times \mathcal{B}}_{k},$$

with OGF

$$B(z)^k.$$

For $k = 0$, we have $\mathcal{B}^0$ equal to the **'empty' class** $\mathcal{E}$, defined as the singleton $\{\bullet\}$ with size $|\bullet| = 0$ and constant OGF $\mathcal{E}(z) = 1$.

Example 3.17. *Consider the **1-object class** $\mathcal{Z} := \{\bullet\}$ with size $|\bullet| = 1$, also called the **'atom' class**. The class of k-tuples of identical objects is $\mathcal{Z}^k = \{*\}$, with $* = (\bullet, \bullet, \ldots, \bullet)$ having size $|*| = k$. The corresponding OGF is the monomial z^k.*

Sequences

In this chapter, we consider a sequence as an arbitrary length tuple, i.e., a (potentially empty) **sequence** as a 0-tuple, **xor** a 1-tuple, **xor** a 2-tuple, etc... Formally, the class of **sequences** of $\mathcal{B}$'s is given by

$$\mathrm{Seq}(\mathcal{B}) := \mathcal{B}^0 \sqcup \mathcal{B}^1 \sqcup \mathcal{B}^2 \sqcup \cdots$$

with OGF

$$\mathrm{Seq}(B(z)) := 1 + B(z) + B(z)^2 + \ldots = \frac{1}{1 - B(z)}.$$

Similarly, we can define the class of **non-empty sequences** of $\mathcal{B}$'s as

$$\mathrm{Seq}_{\geq 1}(\mathcal{B}) := \mathcal{B}^1 \sqcup \mathcal{B}^2 \sqcup \mathcal{B}^3 \sqcup \cdots$$

with OGF

$$\mathrm{Seq}_{\geq 1}(B(z)) := B(z) + B(z)^2 + B(z)^3 \ldots = \frac{B(z)}{1 - B(z)} = \frac{1}{1 - B(z)} - 1.$$

Example 3.18. *The class of sequences of identical objects is* $\mathrm{Seq}(\mathcal{Z}) = \mathcal{E} \sqcup \mathcal{Z} \sqcup \mathcal{Z}^2 \sqcup \cdots$ *with OGF the frequently occurring geometric series* $\mathrm{Seq}(z) = 1 + z + z^2 + \cdots = \frac{1}{1-z}$.

Important: when using $\mathcal{A}$ **xor** $\mathcal{B}$ in the construction of a combinatorial class, make sure that $\mathcal{A}$ and $\mathcal{B}$ are indeed *disjoint*. Similarly, when using $\mathcal{A}$ **and** $\mathcal{B}$ or $\boldsymbol{k}$**-tuples** of $\mathcal{B}$, make sure that the size of objects in $\mathcal{A} \times \mathcal{B}$ or $\mathcal{B}^k$ is indeed *additive* on the sizes of its entries. When using **sequences** of $\mathcal{B}$, make sure that both these properties hold throughout the construction.

Substitution

Tuples and sequences are the main examples of a construction called *substitution*. If $\mathcal{A}$ and $\mathcal{B}$ are two combinatorial classes (with $B_0 = 0$) with OGF's $A(z)$, $B(z)$, then the **substitution** $\mathcal{A} \circ \mathcal{B}$, thought of as *"$\mathcal{A}$'s of $\mathcal{B}$'s"*, is formally defined as the set

$$\mathcal{A} \circ \mathcal{B} := \left\{ (a, (b^{(1)}, \ldots, b^{(k)})) \mid a \in \mathcal{A},\ (b^{(1)}, \ldots, b^{(k)}) \in \mathcal{B}^k,\ k = |a|_{\mathcal{A}} \right\},$$

with size function

$$|(a, (b^{(1)}, \ldots, b^{(k)}))|_{\mathcal{A}\circ\mathcal{B}} := |b^{(1)}|_{\mathcal{B}} + \ldots + |b^{(k)}|_{\mathcal{B}}$$

and OGF

$$A(B(z)) = \sum_{k \geq 0} A_k B(z)^k.$$

In other words, we substitute every element $a \in \mathcal{A}$ with $|a|_{\mathcal{A}}$-tuples of $\mathcal{B}$'s.

Example 3.19. *The classes of k-tuples and sequences (of identical objects) $\mathcal{Z}^k$ and $\mathrm{Seq}(\mathcal{Z})$, substituted with $\mathcal{B}$, are (up to isomorphism) precisely the classes*

$$\mathcal{Z}^k \circ \mathcal{B} = \mathcal{B}^k \quad \textit{and} \quad \mathrm{Seq}(\mathcal{Z}) \circ \mathcal{B} = \mathrm{Seq}(\mathcal{B})$$

of k-tuples of $\mathcal{B}$'s and sequences of $\mathcal{B}$'s.

Combinatorial construction	OGF	Combinatorial interpretation
Disjoint union $\mathcal{A} \sqcup \mathcal{B}$	$A(z) + B(z)$	$\mathcal{A}$ xor $\mathcal{B}$
Product $\mathcal{A} \times \mathcal{B}$	$A(z)B(z)$	$\mathcal{A}$ and $\mathcal{B}$
k-tuples $\mathcal{B}^k$	$B(z)^k$	k-tuples of $\mathcal{B}$'s
Sequences $\mathrm{Seq}(\mathcal{B})$	$\frac{1}{1-B(z)}$	sequences of $\mathcal{B}$'s
Substitution $\mathcal{A} \circ \mathcal{B}$	$A(B(z))$	$\mathcal{A}$-objects of $\mathcal{B}$'s

Table 2: The main constructions of disjoint union, product, tuples, sequences, and substitution.

Besides tuples and sequences, we will for now not delve further into other examples of substitution. We refer the reader to [1] for more examples on the topic.

3.9 Combining OGF's: examples

Fibonacci revisited

Consider again domino tilings in a $2 \times n$ box, with $n \geq 0$. A vertical domino is represented as ' | ' of size 1, and two stacked horizontal dominoes as ' = ' of size 2. With our combinatorial constructions, we now have the interpretation

"a $2 \times n$ tiling equals a sequence of (a ' | ' **xor** a ' = ')."

If we denote the class of tilings with $\mathcal{F}$, the above interpretation implies the equation of classes

$$\mathcal{F} = \mathrm{Seq}(\mathcal{Z} \sqcup \mathcal{Z}^2)$$

with equation of OGF's

$$\begin{aligned} F(z) &= \mathrm{Seq}(z + z^2) \\ &= \frac{1}{1 - z - z^2} \\ &= 1 + z + 2z^2 + 3z^3 + 5z^4 + \dots . \end{aligned}$$

Alternately, we have the interpretation

"a $2 \times n$ tiling equals (an empty tiling) **xor** (a non-empty sequence of (a ' | ' **xor** a ' = '))"

which implies the equation of classes

$$\mathcal{F} = \mathcal{E} \sqcup \mathrm{Seq}_{\geq 1}(\mathcal{Z} \sqcup \mathcal{Z}^2)$$

with equation of OGF's

$$\begin{aligned} F(z) &= 1 + \mathrm{Seq}_{\geq 1}(z + z^2) \\ &= 1 + \left(\frac{1}{1 - (z + z^2)} - 1 \right) = \frac{1}{1 - z - z^2} \\ &= 1 + z + 2z^2 + 3z^2 + 5z^4 + \dots . \end{aligned}$$

Pairs of parentheses

When using parenthesis ( ) in a piece of text, they generally need to be matched in a valid way:

0 pairs : $\emptyset$
1 pair : ( )
2 pairs : (( )), ( )( )
3 pairs : ((( ))), (( )( )), (( ))( ), ( )(( )), ( )( )( )
etc.

We now ask: *how many valid ways are there of writing $n \geq 0$ pairs of parentheses?*

Note that any valid way of writing parentheses is itself a (potentially empty) sequence of pairs of parentheses, **and** with-in those pairs a valid way of writing parentheses, i.e.,

$$\text{valid way} = \underbrace{(\text{a valid way})(\text{another valid way})\cdots(\text{another valid way})}_{k\geq 0}.$$

We thus have the recursive combinatorial interpretation

"a valid way equals a sequence of (a pair **and** a valid way)."

We want n to count the total number of pairs of parentheses. We thus need to associate 1 pair with the class $\mathcal{Z}$ of an object with size 1. If we denote the class of non-empty valid ways of writing parentheses by $\mathcal{C}$, the above interpretation implies the equation of classes

$$\mathcal{C} = \mathrm{Seq}(\mathcal{Z} \times \mathcal{C})$$

with equation of OGF's

$$\begin{aligned} C(z) &= \mathrm{Seq}(z \cdot C(z)) \\ &= \frac{1}{1 - z\,C(z)}. \end{aligned}$$

This is equivalent to

$$z\,C(z)^2 - C(z) + 1 = 0.$$

We recognize this quadratic expression from the OGF $C(z) = \frac{1-\sqrt{1-4z}}{2z} = 1 + z + 2z^2 + 5z^3 + 14z^4 + \ldots$ of the Catalan numbers that were introduced in Section 3.6.

Summary of generating function method

We have seen in the previous examples a general method of solving enumerative problems. The steps are summarized here. Let $(\mathcal{A}, |\cdot|)$ be a combinatorial class with sequence $(A_n)_{n\geq 0}$ counting the number of elements of size n. The method for finding a formula for A_n is as follows:

Step 1: Define a corresponding generating function $A(z)$.

Step 2: Use recursion in A_n and combinatorial constructions from the previous sections (tuples, sequences, etc.) to find an equation in terms of $A(z)$.

Step 3: Solve this equation and obtain a function expression for $A(z)$.

Step 4: Write $A(z)$ in terms of functions with known power series expansion (e.g., binomial series).

Step 5: Use the coefficients of the known series to obtain a formula for the coefficients of $A(z)$, and hence a formula for A_n.

3.10 Multivariate generating functions

Occasionally, one wants to keep track of several parameters simultaneously.

Definition 3.20. A *multivariate generating function* is a power series in several variables, for example

$$A(u,z) = \sum_{n,k\geq 0} A_{n,k} u^k z^n.$$

Here one variable may mark total size, while another marks an auxiliary statistic.

Example 3.21 (Binomial coefficients)**.** *The binomial coefficients admit the bivariate generating function*

$$A(u,z) = \sum_{n\geq 0}\sum_{k=0}^{n} \binom{n}{k} u^k z^n,$$

counting the number of k-element subsets of an n-element set. Here, n marks the total size of the set from which the elements are picked, and k is an auxiliary statistic on the size of the subsets. Using the binomial theorem,

$$\sum_{k=0}^{n} \binom{n}{k} u^k = (1+u)^n,$$

we obtain

$$\sum_{n\geq 0}\sum_{k=0}^{n} \binom{n}{k} u^k z^n = \sum_{n\geq 0} (1+u)^n z^n = \frac{1}{1-(1+u)z}.$$

We can naturally 'forget' about the auxiliary statistic by substituting $u = 1$ into the formulas. We then obtain the OGF

$$A(1,z) = \sum_{n\geq 0}\sum_{k=0}^{n} \binom{n}{k} z^n = \sum_{n\geq 0} 2^n z^n = \frac{1}{1-2z},$$

counting arbitrary sized subsets of an n-element set.

3.11 Exponential generating functions: motivation

Ordinary generating functions are most natural for unlabeled classes. For labeled classes, which we will discuss in this section, exponential generating functions are often better adapted.

Definition 3.22. The *exponential generating function* (EGF) of a sequence $(A_n)_{n\geq 0}$ is

$$A(z) = \sum_{n\geq 0} \frac{A_n}{n!} z^n.$$

The factor $1/n!$ reflects the role of labels and turns many combinatorial constructions on labeled classes into simple algebraic operations. Without this factor, OGFs will usually have very fast growing coefficients, causing their radius of convergence to be $R = 0$, in which case they can be difficult to study with analytic functions. For example, the sequence $A_n = n!$, counting the number of permutations of $[n] = \{1, \ldots, n\}$ for all $n \geq 0$, has OGF

$$\sum_{n\geq 0} n! z^n$$

having radius of convergence $R = 0$. On the other hand, the corresponding EGF is the analytic function

$$\sum_{n\geq 0} z^n = \frac{1}{1-z}$$

with $R = 1$.

Exponential generating functions can also be useful for recurrences, especially when the coefficients depend on n. Linear recurrences of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k},$$

with *constant* coefficients c_i are easily solved with OGF's, like in the case of the Fibonacci sequence. On the other hand, a recurrence with *non-constant* coefficients (polynomial in n), such as the derangement recurrence

$$D_{n+2} = (n+1)(D_{n+1} + D_n),$$

is much more naturally handled by the EGF

$$D(z) = \sum_{n\geq 0} \frac{D_n}{n!} z^n,$$

which leads directly to the differential equation

$$D'(z) = zD'(z) + zD(z)$$

and hence to

$$D(z) = \frac{e^{-z}}{1-z}.$$

3.12 Labeled classes

In the remainder of this chapter, we will work with combinatorial classes endowed with an additional labeling structure, called *labeled classes*. Labeled classes are usually defined in terms of sets, tuples or graphs whose elements or vertices are endowed with unique labels. These labels are often the integers $[n] = \{1, \ldots, n\}$. Defining labeled classes both formally and generally is surprisingly involved, and different authors take different approaches (see e.g. [1, Ch. II] and [4, Ch. 3]). For the scope of these notes, we restrict ourselves to (tuples of) labeled graphs, but one could generalize to objects with more structure.

Definition 3.23. A *labeled graph* is a finite graph G of any type (undirected, directed, or with more structure!), together with a bijection $L : V \to [n]$ from the vertex set V of G to the natural numbers $[n] = \{1, 2, \ldots, n\}$ with $n = |V|$, called a *labeling*. The elements in $[n]$ are called *labels*. Two labeled graphs (G_1, L_1) and (G_2, L_2) are considered *isomorphic* if there is a graph isomorphism $\varphi : G_1 \to G_2$ preserving both the graph structure and the labeling, i.e., $L_1 = L_2 \circ \varphi$.

Definition 3.24. A *labeled class* is a combinatorial class $\mathcal{A}$ consisting of labeled graphs (G, L), such that the *size* of a labeled graph is the number of vertices $n = |V|$ of that graph.

Moreover, if $(G, L) \in \mathcal{A}$, then we require that also $(G, L') \in \mathcal{A}$ for any other labeling L', although we discard isomorphic copies of labeled graphs. In other words, $\mathcal{A}$ contains from every isomorphism class exactly one representative. Note that a graph G of size n has at most $n!$ different isomorphism classes of labeled graphs, since the number of bijections $V \to [n]$ is exactly $n!$.

Definition 3.25. If $(\mathcal{A}, |\cdot|)$ is a class with A_n objects of size n for each $n \geq 0$, we define the corresponding EGF as

$$A(z) = \sum_{n \geq 0} \frac{A_n}{n!} z^n = \sum_{a \in \mathcal{A}} \frac{z^{|a|}}{|a|!}.$$

Most elementary labeled classes are sets of labeled directed graphs of a certain shape or type. We discuss here a few examples.

Example 3.26 (Permutations). *The class of permutations* Perm *can be identified with the class of labeled graphs* (G, L)*, with* G *any directed line-graph of the form*

$$G = \quad \overset{v_1}{\bullet} \longrightarrow \overset{v_2}{\bullet} \longrightarrow \overset{v_3}{\bullet} \longrightarrow \cdots \longrightarrow \overset{v_{n-2}}{\bullet} \longrightarrow \overset{v_{n-1}}{\bullet} \longrightarrow \overset{v_n}{\bullet}$$

for $n \geq 0$ *and* $L : V = \{v_1, \ldots, v_n\} \to [n]$ *any labeling. Different labelings on a directed line-graph of size* n *yields different non-isomorphic labeled graphs (since* G *only has a trivial graph isomorphism). So this class contains*

$$\mathrm{Perm}_n = n!$$

labeled directed line-graphs of size $n \geq 0$ *(one for each permutation of the labels* $[n]$*), and thus the corresponding EGF is*

$$Perm(z) = \sum_{n \geq 0} \frac{n!}{n!} z^n = \sum_{n \geq 0} z^n = \frac{1}{1 - z}.$$

Example 3.27 (Sets). *The class of sets* Set *can be identified with the class of labeled graphs* (G, L) *with* G *any unconnected graph*

$$G = \quad \overset{v_1}{\bullet} \quad \overset{v_2}{\bullet} \quad \overset{v_3}{\bullet} \quad \cdots \quad \overset{v_{n-2}}{\bullet} \quad \overset{v_{n-1}}{\bullet} \quad \overset{v_n}{\bullet}$$

for $n \geq 0$ *and* $L : V = \{v_1, \ldots, v_n\} \to [n]$ *any labeling. Every labeling on an unconnected graph of size* n *yields isomorphic labeled graphs (since any permutation of the vertices is a graph isomorphism), so up to isomorphism this class contains*

$$\mathrm{Set}_n = 1$$

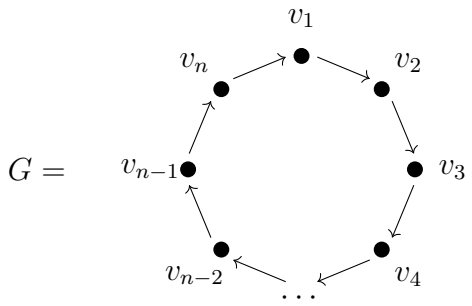
labeled unconnected graph for each size $n \geq 0$*, and the corresponding EGF is*

$$Set(z) = \sum_{n \geq 0} \frac{1}{n!} z^n = e^z.$$

We similarly define the class of non-empty *labeled sets* $\mathrm{Set}_{\geq 1}$*, which excludes the empty graph of size 0, and thus has* $\mathrm{Set}_0 = 0$*,* $\mathrm{Set}_n = 1$ *for* $n \geq 1$*, and corresponding EGF*

$$Set_{\geq 1}(z) = \sum_{n \geq 1} \frac{1}{n!} z^n = e^z - 1.$$

Example 3.28 (Cycles). *The class of cycles* Cyc *can be identified with the class of labeled graphs* (G, L)*, with* G *any directed cycle-graph of the form*



for $n \geq 1$ *and* $L : V = \{v_1, \ldots, v_n\} \to [n]$ *any labeling. Different labelings on a directed cycle-graph of size* n *yields non-isomorphic labeled graphs, except when labelings differ by a cyclic shift (since* G *is rotationally symmetric of order* n*). So up to isomorphism, this class contains*

$$\mathrm{Cyc}_n = (n-1)!$$

labeled directed cycle-graphs of size $n \geq 1$ *(one for each of the* $n!$ *permutation of the labels* $[n]$*, up to the* n *cyclic shifts), and thus the corresponding EGF is*

$$Cyc(z) = \sum_{n \geq 1} \frac{(n-1)!}{n!} z^n = \sum_{n \geq 1} \frac{1}{n} z^n = -\log(1-z).$$

3.13 Combining EGF's: addition and multiplication

One of the main strengths of exponential generating functions is that they behave very well under natural constructions on labeled classes, analogous to the constructions for OGF's in Sections 3.7 and 3.8. In the following, let $\mathcal{A}$ and $\mathcal{B}$ be labeled combinatorial classes with EGFs

$$A(z) = \sum_{n\geq 0} \frac{a_n}{n!} z^n, \qquad B(z) = \sum_{n\geq 0} \frac{b_n}{n!} z^n.$$

Xor

The disjoint union $\mathcal{A} \sqcup \mathcal{B}$ of labeled classes, interpreted as '$\mathcal{A}$ **xor** $\mathcal{B}$', is defined exactly as the disjoint union of (unlabeled) combinatorial classes, where objects in $\mathcal{A} \sqcup \mathcal{B}$ are already endowed with a labeling, as they are objects from either $\mathcal{A}$ or $\mathcal{B}$. It follows that EGF corresponding to $\mathcal{A} \sqcup \mathcal{B}$ is the sum of the EGF's

$$A(z) + B(z).$$

And

In a product construction for labeled classes, we want to count the number of ways to build an object of size n from a first part from $\mathcal{A}$ of size k **and** a second part from $\mathcal{B}$ of size $n-k$ for any choice of k, *together* with all possible choices of selecting k labels out of $[n]$ for the first part (or equivalently, all possible choices of selecting $n-k$ labels out of $[n]$ for the second part). We will formally define such a construction below.

Let $(G_A, L_A) \in \mathcal{A}$ be a labeled graph with vertex set V_A of size n_A. Likewise, let $(G_B, L_B) \in \mathcal{B}$ be a labeled graph with vertex set V_B of size n_B. The pair (G_A, G_B) is naturally interpreted as a finite graph G (with additional structure 'encoding' the separation of the left part G_A from the right part G_B) and vertex set $V = V_A \sqcup V_B$ of size $n = n_A + n_B$.

Regarding labeling, there are generally multiple ways of combining L_A and L_B into a labeling $L : V \to [n]$ on G. Namely, there are

$$\binom{n}{n_A}$$

many ways of selecting a subset $I_A \subset [n]$ of size $|I_A| = n_A$. Clearly, the complement $I_B := [n] \setminus I_A$ has then size $|I_B| = n_B$. For a given choice of (I_A, I_B), there is a unique pair of *order-preserving* bijections

$$i_A : [n_A] \to I_A \quad \text{and} \quad i_B : [n_B] \to I_B,$$

and the compositions $i_A \circ L_A : V_A \to I_A$ and $i_B \circ L_B : V_B \to I_B$ naturally induce a labeling on G:

$$(i_A \circ L_A,\ i_B \circ L_B) : V_A \sqcup V_B \to [n].$$

We now define the **labeled product** $\mathcal{A} \times_L \mathcal{B}$ to be the class of all labeled graphs of the form

$$(G, L) = ((G_A, G_B), (i_A \circ L_A,\ i_B \circ L_B)),$$

with $(G_A, L_A) \in \mathcal{A}$, $(G_B, L_B) \in \mathcal{B}$, and for all of the $\binom{n}{n_A} = \binom{n}{n_B}$ possible choices (I_A, I_B).

The size n of (G, L) is the sum of the sizes of (G_A, L_A) and (G_B, L_B), i.e., $n = n_A + n_B$.

It follows from this construction that, in the labeled product $\mathcal{A} \times_L \mathcal{B}$, the number of labeled graphs (G, L) of size n is

$$\sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k},$$

which is precisely the n-th coefficient in the product of the EGF's

$$A(z)B(z) = \sum_{n\geq 0} \sum_{k=0}^{n} \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!} z^n = \sum_{n\geq 0} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k} \right) z^n.$$

Combinatorial construction	EGF	Combinatorial interpretation
Disjoint union $\mathcal{A} \sqcup \mathcal{B}$	$A(z) + B(z)$	$\mathcal{A}$ xor $\mathcal{B}$
Labeled product $\mathcal{A} \times_L \mathcal{B}$	$A(z)B(z)$	$\mathcal{A}$ and $\mathcal{B}$

Table 3: The main constructions of disjoint union and labeled product.

Example 3.29 (Derangements). *Consider the class* Der *of all derangements on* $[n]$ *for* $n \geq 0$*. Let* D_n *be the number of derangements on* $[n]$*, and let* $D(z)$ *be the corresponding EGF. A permutation is a set of cycles, while a derangement is a set of cycles of length at least* 2*. Equivalently, a permutation is a set of some labels (the fixed points of the permutation)* ***and*** *a derangement on the remaining labels. Hence we have the equation of labeled classes*

$$\mathrm{Perm} = \mathrm{Set} \times_L \mathrm{Der}$$

and thus the EGF equation

$$Perm(z) = Set(z) \cdot D(z),$$

where

$$Perm(z) = \frac{1}{1-z} \quad \textit{and} \quad Set(z) = e^z.$$

Therefore

$$D(z) = \frac{e^{-z}}{1-z},$$

and

$$D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \sim \frac{n!}{e}.$$

3.14 Combining EGF's: tuples and sets

Let $\mathcal{B}$ be a labeled class with EGF $B(z)$.

k-tuples

The class of **k-tuples** of objects in $\mathcal{B}$ is the k-fold labeled product:

$$\mathcal{B}^k := \underbrace{\mathcal{B} \times_L \mathcal{B} \times_L \cdots \times_L \mathcal{B}}_{k},$$

with EGF

$$B(z)^k.$$

k-sets

The class of **k-sets** of objects in $\mathcal{B}$ is the quotient of the k-tuples $\mathcal{B}^k$ by the group $Sym(k)$ of all the permutations of the k entries, i.e.,

$$\mathrm{Set}_k(\mathcal{B}) := \mathcal{B}^k/_{Sym(k)}.$$

This means, we consider sets of k objects from $\mathcal{B}$, where the objects have no order among each other. Formally, for any permuation $\sigma \in Sym(k)$, the k-tuples

$$((G_1, \ldots, G_k), (i_1 \circ L_1, \ldots,\ i_k \circ L_k)) \in \mathcal{B}^k$$

and

$$((G_{\sigma(1)}, \ldots, G_{\sigma(k)}), (i_{\sigma(1)} \circ L_{\sigma(1)}, \ldots,\ i_{\sigma(k)} \circ L_{\sigma(k)})) \in \mathcal{B}^k$$

are considered the same. The class $\mathrm{Set}_k(\mathcal{B})$ of k-sets thus has EGF

$$\frac{1}{k!}B(z)^k.$$

Sets

The class of (potentially empty) **sets** of objects in $\mathcal{B}$ is given by

$$\mathrm{Set}(\mathcal{B}) := \mathrm{Set}_0(\mathcal{B}) \sqcup \mathrm{Set}_1(\mathcal{B}) \sqcup \mathrm{Set}_2(\mathcal{B}) \sqcup \cdots$$

with EGF

$$Set(B(z)) := 1 + \frac{B(z)}{1!} + \frac{B(z)^2}{2!} + \ldots = e^{B(z)}.$$

Similarly, we can define the class of **non-empty sets** of objects in $\mathcal{B}$ as

$$\mathrm{Set}_{\geq 1}(\mathcal{B}) := \mathrm{Set}_1(\mathcal{B}) \sqcup \mathrm{Set}_2(\mathcal{B}) \sqcup \mathrm{Set}_3(\mathcal{B}) \sqcup \cdots$$

with EGF

$$Set_{\geq 1}(B(z)) := \frac{B(z)}{1!} + \frac{B(z)^2}{2!} + \frac{B(z)^3}{3!} + \ldots = e^{B(z)} - 1.$$

Combinatorial construction	EGF	Combinatorial interpretation
Disjoint union $\mathcal{A} \sqcup \mathcal{B}$	$A(z) + B(z)$	$\mathcal{A}$ xor $\mathcal{B}$
Labeled product $\mathcal{A} \times_L \mathcal{B}$	$A(z)B(z)$	$\mathcal{A}$ and $\mathcal{B}$
k-tuples $\mathcal{B}^k$	$B(z)^k$	k-tuples of $\mathcal{B}$'s
k-sets $\mathrm{Set}_k(\mathcal{B})$	$\frac{1}{k!}B(z)^k$	k-sets of $\mathcal{B}$'s
Sets $\mathrm{Set}(\mathcal{B})$	$e^{B(z)}$	Sets of $\mathcal{B}$'s

Table 4: The main constructions of disjoint union, labeled product, k-tuples, k-sets, and sets.

3.15 Combining EGF's: Stirling and Bell numbers

Stirling numbers of the second kind

Recall that the Stirling number of the second kind $S(n, k)$ counts partitions of an n-element set into k non-empty blocks. We will derive the EGF of the Stirling numbers in two different ways.

Theorem 3.30. *For each $k \geq 0$,*

$$\sum_{n \geq k} S(n,k)\frac{z^n}{n!} = \frac{1}{k!}(e^z - 1)^k.$$

Proof. Let

$$F_k(z) := \sum_{n \geq k} S(n,k)\frac{z^n}{n!}.$$

Using the standard recurrence

$$S(n,k) = kS(n-1,k) + S(n-1,k-1),$$

one obtains the differential relation

$$F_k'(z) = kF_k(z) + F_{k-1}(z).$$

The claimed formula satisfies this relation and the initial conditions, so it is the unique solution. □

An alternative method to obtain the EGF is via labeled class constructions: a partition of the set $[n]$ into k non-empty blocks/subsets is equal to

*"A **k-set** of disjoint **non-empty sets** of labels, so that the union of all labels is $[n]$."*

Thus, if we denote the class of these partitions for fixed $k \geq 0$ by Stir_k^{2nd}, we find the labeled class equation

$$\mathrm{Stir}_k^{2nd} = \mathrm{Set}_k(\mathrm{Set}_{\geq 1})$$

and hence the EGF

$$Stir_k^{2nd}(z) = Set_k(Set_{\geq 1}(z)) = \frac{1}{k!}(e^z - 1)^k.$$

where

$$Stir_k^{2nd}(z) = \sum_{n \geq k} S(n,k)\frac{z^n}{n!}.$$

Bell numbers

Summing the Stirling numbers of the second kind over k gives the EGF for the Bell numbers $B(n)$:

$$\sum_{n\geq 0} B(n)\frac{z^n}{n!} = \sum_{k\geq 0} \frac{(e^z-1)^k}{k!} = e^{e^z-1}.$$

The Bell numbers are the counting sequence of the class Bell of partitions of the set $[n]$ into any number non-empty subsets. So an element in Bell is equal to

*"A **set** of disjoint **non-empty sets** of labels, so that the union of all labels is $[n]$."*

Thus we find the labeled class equation

$$\mathrm{Bell} = \mathrm{Set}(\mathrm{Set}_{\geq 1})$$

and hence the EGF

$$Bell(z) = Set(Set_{\geq 1}(z)) = e^{e^z-1}.$$

where

$$Bell(z) = \sum_{n\geq 0} B(n)\frac{z^n}{n!}.$$

Stirling numbers of the first kind

For the Stirling numbers of the first kind, the generating function is slightly less obvious.

Theorem 3.31. *For each $k \geq 0$,*

$$\sum_{n\geq k} s(n,k)\frac{z^n}{n!} = \frac{1}{k!}(\log(1+z))^k.$$

Proof. We expand $(1+z)^x$ in two different ways. First,

$$(1+z)^x = e^{x\log(1+z)} = \sum_{k\geq 0}\frac{1}{k!}(\log(1+z))^k x^k.$$

On the other hand,

$$(1+z)^x = \sum_{n\geq 0}\binom{x}{n}z^n = \sum_{n\geq 0}\frac{(x)_n}{n!}z^n.$$

Since

$$(x)_n = \sum_{r=0}^{n} s(n,r)x^r,$$

we obtain

$$(1+z)^x = \sum_{k\geq 0} x^k \sum_{n\geq k} s(n,k)\frac{z^n}{n!}.$$

Comparing coefficients of x^k proves the claim. □

An alternative method to obtain the EGF is via labeled class constructions: recall that the *unsigned* Stirling numbers of the first kind

$$u(n,k) = (-1)^{n-k} s(n,k)$$

counts the number of partitioning $[n]$ into k cycles of length ≥ 1, or in other words:

*"**k-sets** of disjoint **non-empty cycles** of labels, so that the union of all labels is $[n]$."*

If we denote the corresponding labeled classes by UStir_k^{1st}, respectively by Stir_k^{1st}, we find the labeled class equation

$$\mathrm{UStir}_k^{1st} = \mathrm{Set}_k(\mathrm{Cyc})$$

and hence the EGF

$$UStir_k^{1st}(z) = Set_k(Cyc(z)) = \frac{1}{k!}(-\log(1-z))^k.$$

where

$$UStir_k^{1st}(z) = \sum_{n \geq k} u(n,k)\frac{z^n}{n!} = (-1)^k \sum_{n \geq k} s(n,k)\frac{(-z)^n}{n!} = (-1)^k Stir_k^{1st}(-z).$$

3.16 Asymptotic analysis

We finish this chapter with another power of generating functions: the study of the asymptotic behavior of sequences. Knowing the OGF or EGF of a sequence can be useful here since the location of their singularities often determines the exponential growth of the coefficients. For rational generating functions, these singularities are precisely the zeros of the denominator. We have already seen two examples:

- Fibonacci numbers:
$$f_n \sim \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n,$$
- derangements:
$$D_n \sim \frac{n!}{e}.$$

In both cases, the asymptotic growth is determined by the singularities of the generating functions. We will explain in this section how, and refer to [1, Ch. IV-VI] for more details.

First Principle of Coefficient Asymptotics

"The location of a functions singularities dictates the exponential growth A^n of its coefficients."

In the following, let $F(z)$ be analytic at 0, and suppose the coefficients of $F(z)$ are of the form

$$[z^n]F(z) = A^n\theta(n)$$

for some $A > 0$ indicating the exponential growth, and subexponential factor $\theta(n)$. We have seen in the Cauchy-Hadamard theorem (Thm. 3.8) how to obtain the radius of convergence from the coefficients of an analytic function. Here we give a useful converse of that theorem.

Theorem 3.32 (Exponential Growth Formula, [1, Thm. IV.7])**.** *If $F(z)$ is analytic at 0 and with radius of convergence $R > 0$, then*

$$[z^n]F(z) = \left(\frac{1}{R}\right)^n \theta(n)$$

for some subexponential factor $\theta(n)$.

In practice, R is the modulus or absolute value of a singularity nearest to the origin.

Example 3.33. *The OGF for the Fibonacci numbers*

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

has singularities at the roots of $1 - z - z^2$, i.e.,

$$r_1 = \frac{1}{2}\left(-1 + \sqrt{5}\right) \approx 0.618, \qquad r_2 = -\frac{1}{2}\left(1 + \sqrt{5}\right) \approx -1.618.$$

Since r_1 is the location of the singularity nearest to the origin, $F(z)$ has radius of convergence $R = r_1$ and thus

$$[z^n]F(z) = \left(\frac{1}{r_1}\right)^n \theta(n) = \left(\frac{1 + \sqrt{5}}{2}\right)^n \theta(n).$$

Example 3.34. *The EGF for derangements*

$$D(z) = \frac{e^{-z}}{1 - z}$$

has a singularity at the root of $1 - z$, i.e.,

$$r = 1,$$

and thus $D(z)$ has radius of convergence $R = 1$, implying that

$$[z^n]D(z) = \theta(n)$$

is only subexponential in n. However, since we used the EGF, the coefficients of $D(z)$ are

$$[z^n]D(z) = \frac{1}{n!}D_n,$$

and thus

$$D_n = n!\,\theta(n).$$

For the above examples, it is possible to understand the subexponential factors better, as the nearest singularities are simple poles of the generating functions.

Theorem 3.35. *Let $R > 0$, and consider a function of the form*

$$F(z) = \frac{G(z)}{1 - z/R},$$

with $G(z)$ an analytic function with radius of convergence $R_G > R$. Then $F(z)$ has radius of convergence R, and asympotically as $n \to \infty$:

$$[z^n]F(z) \sim G(R)\left(\frac{1}{R}\right)^n.$$

Example 3.36. *The OGF of the Fibonacci numbers satisfies*

$$F(z) = \frac{z}{1 - z/r_2} \cdot \frac{1}{1 - z/r_1}$$

and thus

$$F_n = [z^n]F(z) \sim \frac{r_1}{1 - r_1/r_2}\left(\frac{1}{r_1}\right)^n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

Likewise, the EGF of derangements satisfies

$$D(z) = e^{-z}\frac{1}{1 - z}$$

and thus

$$D_n = n![z^n]D(z) \sim \frac{n!}{e}.$$

Second Principle of Coefficient Asymptotics

"The nature of a functions singularities determines the associate subexponential factor $\theta(n)$."

In general for more complicated functions, obtaining the subexponential growth factor of a series is much more involved than its exponential growth. Luckily, for functions of a standard form, we have a nice formula for the growth factors.

Theorem 3.37 (Standard function asymptotics, [1, Thm. VI.1])**.** *Let α be an arbitrary complex number in $\mathbb{C} \setminus \mathbb{Z}_{\leq 0}$. The coefficients of*

$$F(z) = (1-z)^{-\alpha}$$

satisfy asympotically as $n \to \infty$:

$$[z^n]F(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)},$$

where $\Gamma(\alpha)$ is the Gamma function.

Corollary 3.38. *The coefficients of $F(z) = (1-Az)^{-\alpha}$ satisfy asymptotically as $n \to \infty$:*

$$[z^n]F(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)} A^n.$$

Example 3.39 (Asymptotics of Catalan numbers)**.** *The OGF*

$$C(z) = \frac{1}{2z}(1 - \sqrt{1-4z})$$

of the Catalan numbers is, up to a constant term and scaling by $-2z$, of the form

$$F(z) = (1-4z)^{1/2}.$$

Its singularity is located at $R = \frac{1}{4}$, and the power of $1/2$ implies an asymptotic growth of

$$[z^n]F(z) \sim -\frac{n^{-3/2}\, 4^n}{2\sqrt{\pi}}$$

with $A = \frac{1}{R} = 4$ and $\alpha = -\frac{1}{2}$. Note that $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$. Accounting for scaling by $-2z$ we finally obtain

$$C_n = -\frac{1}{2}[z^{n+1}]F(z) \sim \frac{4^{n+1}}{4\sqrt{\pi}\,(n+1)^{3/2}} \sim \frac{4^n}{\sqrt{\pi}\, n^{3/2}}.$$

3.17 Summary

The general philosophy is:

- encode a counting sequence as a generating function,
- translate combinatorial constructions or recurrences into algebraic identities,
- manipulate the resulting formal power series,
- extract coefficients or deduce asymptotic information.

In this way, generating functions form a bridge between combinatorics, algebra, and analysis.

4 Partitions

We have already considered several partition problems, e.g. how many partitions there are of $\{1, \ldots, n\}$ or how many (positive or non-negative) integer solutions there are to $x_1 + \cdots + x_k = n$. We now come to a special partition problem, where we impose one more condition on the x_i :

Definition 4.1. Let n, k be positive integers. An *unordered partition* of n into k parts is a multiset of k positive integers whose sum is n. In particular, the order of the parts does not matter.

To avoid counting the same partition multiple times, we represent each partition uniquely by writing the parts in non-increasing order:

$$x_1 \geq x_2 \geq \cdots \geq x_k \geq 1, \quad x_1 + \cdots + x_k = n.$$

The number of such partitions is denoted by $p_k(n)$ and called *the number of partitions of n into k parts*.

Clearly, for $n < k$ we have that $p_k(n) = 0$ and for $n = k$ we get $p_k(k) = 1$, since we only have $k = 1 + \cdots + 1$. We also note that $p_1(n) = 1$, as there is only $n = n$.

Example 4.2. *Let us consider $n = 7$ and $k = 3$. We compute $7 = 5 + 1 + 1 = 4 + 2 + 1 = 3 + 3 + 1 = 3 + 2 + 2$. Thus, $p_3(7) = 4$.*

Recall, that the number of non-negative integer solutions to $x_1 + \cdots + x_k = n$ was given by $\binom{n-k+1}{k-1}$ and the trick we used to compute the number of positive solutions: replace x_i by $y_i = x_i - 1$. Using the same idea, we can show that $p_k(n)$ is the number of solutions of $n - k = y_1 + \cdots + y_k$ with $y_1 \geq y_2 \geq \cdots \geq y_k \geq 0$. If exactly s of the integers y_i are positive, then there are $p_s(n-k)$ solutions. From this we get

Proposition 4.3. *Let k, n be positive integers. Then*

$$p_k(n) = \sum_{s=1}^{k} p_s(n-k).$$

Exercise 4.4. *Let k, n be positive integers. Show that $p_k(n) = p_{k-1}(n-1) + p_k(n-k)$.*

With this we can then prove the Proposition 4.3:

Exercise 4.5. *Use Exercise 4.4 to prove Proposition 4.3.*

Together with the conditions $p_k(k) = 1, p_1(n) = 1$ and $p_k(n) = 0$ for $k > n$, we can recursively compute the number of partitions of n into k parts.

Exercise 4.6. *Let n be a positive integer. Show that $p_2(n) = \lfloor n/2 \rfloor$.*

Theorem 4.7. *Let k be a fixed positive integer. Then*

$$p_k(n) \sim \frac{n^{k-1}}{k!(k-1)!}$$

as n tends to infinity.

Proof. Let $(x_1, \ldots, x_k)$ be a solution of

$$n = \sum_{i=1}^{k} x_i \qquad \text{with} \qquad x_1 \geq x_2 \geq \cdots \geq x_k \geq 1.$$

Such a solution determines an unordered partition of n into k parts.

By permuting the entries, we obtain at most $k!$ positive integer solutions of

$$n = x_1 + \cdots + x_k.$$

Recall from Corollary 1.32 that the number of positive integer solutions of

$$n = \sum_{i=1}^{k} x_i$$

is

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

Hence

$$\binom{n-1}{k-1} \leq k!\, p_k(n).$$

Conversely, let $(x_1, \ldots, x_k)$ be such an ordered partition and define

$$y_i = x_i + (k-i) \qquad \text{for } i \in \{1, \ldots, k\}.$$

Then the numbers $y_1, \ldots, y_k$ are pairwise distinct and positive, and

$$\sum_{i=1}^{k} y_i = \sum_{i=1}^{k} x_i + \sum_{i=1}^{k} (k-i) = n + \frac{k(k-1)}{2}.$$

Since the y_i are distinct, their $k!$ permutations are all different. Therefore each partition counted by $p_k(n)$ gives rise to exactly $k!$ distinct positive integer solutions of

$$y_1 + \cdots + y_k = n + \frac{k(k-1)}{2}.$$

Again by Corollary 1.32, the number of positive integer solutions of this equation is

$$\binom{n + \frac{k(k-1)}{2} - 1}{k-1}.$$

Thus

$$k!\, p_k(n) \leq \binom{n + \frac{k(k-1)}{2} - 1}{k-1}.$$

Altogether we have shown

$$\frac{1}{k!}\binom{n-1}{k-1} \leq p_k(n) \leq \frac{1}{k!}\binom{n + \frac{k(k-1)}{2} - 1}{k-1}.$$

Now, for fixed k, both bounds are polynomials in n of degree $k-1$ with the same leading coefficient

$$\frac{1}{k!(k-1)!}.$$

Indeed,

$$\frac{1}{k!}\binom{n-1}{k-1} = \frac{n^{k-1}}{k!(k-1)!} + \text{lower order terms},$$

and likewise

$$\frac{1}{k!}\binom{n + \frac{k(k-1)}{2} - 1}{k-1} = \frac{n^{k-1}}{k!(k-1)!} + \text{lower order terms}.$$

Since $p_k(n)$ is squeezed between these two expressions, it has the same leading asymptotic term. Hence

$$p_k(n) \sim \frac{n^{k-1}}{k!(k-1)!}.$$

□

Exercise 4.8. *Let $a_1, \ldots, a_t$ be positive integers, not necessarily all distinct, with gcd$(a_1, \ldots, a_t) = 1$. Let $f(n)$ denote the number of non-negative integer solutions $(x_1, \ldots, x_t)$ for $n = a_1x_1 + \cdots + a_tx_t$. What is the generating function $F(x) = \sum f(n)x^n$?*

In fact, the number we computed in Corollary 1.32, is the number of *compositions.*

Definition 4.9. Let k, n be positive integers. Then $(x_1, \ldots, x_k)$ consisting of positive integers is called a *composition* of n into k parts if $\sum_{i=1}^{k} x_i = n$.

Recall from Proposition 1.7, that $\sum_{k=0}^{n} \binom{n}{k} = 2^n$, similarly we get the following:

Corollary 4.10. *Let n be a positive integer. The number of compositions of n is given by $\sum_{k=1}^{n} \binom{n-1}{k-1} = 2^{n-1}$.*

A different kind of proof, is to define $c_{n,k}$ as the number of compositions of n into k parts (without knowing that there are $\binom{n-1}{k-1}$ many) and to define $c_k(x) = \sum_{n=k}^{\infty} c_{n,k} x^n$. Then,

$$c_k(x) = (x + x^2 + \cdots)^k = x^k(1-x)^{-k} = \sum_{n=k}^{\infty} \binom{n-1}{k-1} x^n$$

and

$$\sum_{k=1}^{\infty} c_k(x) = \sum_{k=1}^{\infty} x^k(1-x)^{-k} = \frac{x}{1-2x}$$

from which it follows also that n has 2^{n-1} compositions.

We now come to the probably most important number in combinatorics: the *partition number*.

Definition 4.11. Let n be a positive integer. The *partition number* (or function) is then the number of unordered partitions of n and is denoted by $p(n)$.

Example 4.12. *Let us consider $n = 5$. Then we compute*

$$5 = 1+1+1+1+1 = 2+1+1+1 = 3+1+1 = 2+2+1 = 4+1 = 3+2 = 5,$$

that is $p(5) = 7$.

While we can compute the partition number for such a small example, in contrast to the number of compositions, there is *no closed formula* for the number of partitions for any n.

Theorem 4.13. *The generating function of $p(n)$ is*

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

Proof. A partition of n can be written in the form

$$n = \sum_{k \geq 1} k\, a_k,$$

where $a_k \geq 0$ denotes how many times the part k appears.

For each fixed k, the possible contributions are

$$0, k, 2k, 3k, \ldots,$$

which correspond to the generating function

$$1 + x^k + x^{2k} + x^{3k} + \cdots = \frac{1}{1-x^k}.$$

Since the choices for different k are independent, the generating function is the product

$$\prod_{k=1}^{\infty} \frac{1}{1-x^k}.$$

Expanding this product, the coefficient of x^n counts the number of ways to write n in the form $\sum ka_k$, that is, the number of partitions of n. □

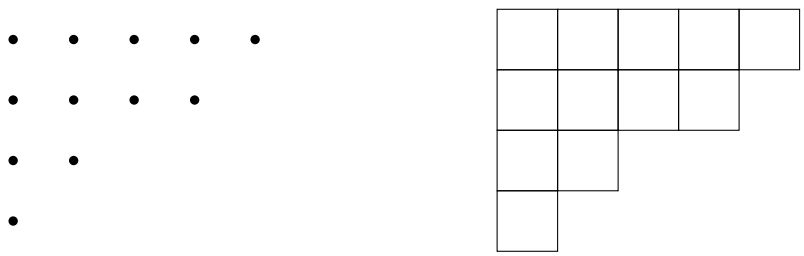
Many results on partitions are easier when we represent them visually: using a Ferrers or a Young diagram.

4.1 Ferrers diagrams

A *Ferrers diagram* of a partition of n $(x_1, \ldots, x_k)$ with $x_1 \geq \cdots \geq x_k$ is given by representing each part by a row of dots in descending order. If it is more convenient to use squares instead of boxes, the diagram is called *Young diagram*.

Example 4.14. *Let us consider $n = 12$ and the partition $(5, 4, 2, 1)$. We can represent it either as a Ferrers diagram or as a Young diagram:*

Ferrers diagram *Young diagram*



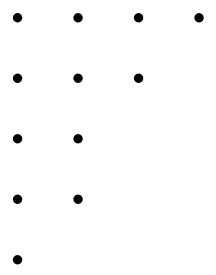
If we read the diagram using the descending columns, instead of the rows, we get another partition!

Definition 4.15. Let $(x_1, \ldots, x_k)$ be a partition of n with $x_1 \geq \cdots \geq x_k \geq 1$. The *conjugate partition* $(y_1, \ldots, y_{x_1})$ is defined by

$$y_j = \big|\{\, i \in \{1, \ldots, k\} \mid x_i \geq j \,\}\big| \qquad \text{for } j = 1, \ldots, x_1.$$

Thus, y_j counts how many parts of the original partition are at least j, i.e. how many rows reach the j-th column.

Example 4.16. *In the example above with $n = 12$ and the partition $(5, 4, 2, 1)$, we get the conjugate partition of 12 as $(4, 3, 2, 2, 1)$.*



In fact, a partition and its conjugate are both partitions of the same positive integer n.

Proposition 4.17. *Let $(x_1, \ldots, x_k)$ be a partition of a positive integer n, and let $(y_1, \ldots, y_{x_1})$ be its conjugate partition. Then*

$$\sum_{i=1}^{k} x_i = \sum_{j=1}^{x_1} y_j = n.$$

In particular, a partition and its conjugate are both partitions of the same integer.

Proof. Recall that the conjugate partition is defined by

$$y_j = \left|\{\, i \in \{1, \ldots, k\} \mid x_i \geq j \,\}\right| \qquad \text{for } j = 1, \ldots, x_1.$$

Thus y_j counts how many rows of the Ferrers diagram reach into the j-th column.

Now count the number of dots in the Ferrers diagram in two ways.

On the one hand, the i-th row contains exactly x_i dots, so the total number of dots is

$$x_1 + \cdots + x_k = n.$$

On the other hand, the j-th column contains exactly y_j dots, so the total number of dots is also

$$y_1 + \cdots + y_{x_1}.$$

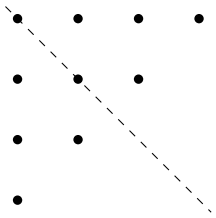
Therefore,

$$\sum_{i=1}^{k} x_i = \sum_{j=1}^{x_1} y_j = n.$$

Hence the conjugate partition is again a partition of the same integer. □

We say that a partition is *self-conjugate* if it is the same as its conjugate.

Example 4.18. *Consider $n = 10$ and the partition $(4, 3, 2, 1)$, for which the Ferrers diagram is given by*



Thus, after transposing it, we still get the same one.

Theorem 4.19. *Let k, n be positive integers. The number of partitions, where we restrict the largest part to be k, is $p_k(n)$.*

Proof. For each partition of n with largest part being k, the conjugate partition has k parts. As for every partition there is a unique conjugate partition, there are exactly $p_k(n)$. □

Exercise 4.20. *Let k, n be positive integers. Show that the number of partitions of $n + k$ into k parts is the same as the number of partitions of n into at most k parts.*

Theorem 4.21. *Let n be a positive integer. The number of partitions of n into odd parts is the same as the number of partitions of n into unequal parts.*

Proof. The generating function for the number of partitions of n into odd parts is given by

$$\prod_{i=1}^{\infty}(1 - x^{2i-1})^{-1}$$

and the generating function for the number of partitions into unequal parts is given by

$$\prod_{k=1}^{\infty}(1 + x^k).$$

We now conclude, since

$$\begin{aligned}\prod_{k=1}^{\infty}(1 + x^k) &= \prod_{k=1}^{\infty}\frac{1 - x^{2k}}{1 - x^k} = \prod_{k=1}^{\infty}(1 - x^{2k})\prod_{j=1}^{\infty}(1 - x^j)^{-1} \\ &= \prod_{i=1}^{\infty}(1 - x^{2i-1})^{-1}.\end{aligned}$$

□

We now give a combinatorial proof of the same result.

Combinatorial proof. Let

$$n = x_1 + \cdots + x_k$$

be a partition into distinct parts.

Write each part uniquely as

$$x_i = u_i 2^{a_i}$$

where u_i is odd.

Group together equal odd parts. Then we can write

$$n = \mu_1 r_1 + \mu_2 r_2 + \cdots$$

where the μ_i are distinct odd numbers and each r_i is a sum of distinct powers of 2.

Now replace each term $\mu_i r_i$ by r_i copies of μ_i. This yields a partition of n into odd parts.

Conversely, given a partition into odd parts, the multiplicities r_i can be uniquely written as sums of distinct powers of 2 (binary expansion), which reconstructs a partition into distinct parts.

This establishes a bijection. □

Example 4.22. *To illustrate the bijection, consider the partition into distinct parts*

$$26 = 12 + 6 + 4 + 3 + 1.$$

We separate the powers of 2*:*

$$12 = 3 \cdot 2^2, \quad 6 = 3 \cdot 2^1, \quad 4 = 1 \cdot 2^2, \quad 3 = 3 \cdot 2^0, \quad 1 = 1 \cdot 2^0.$$

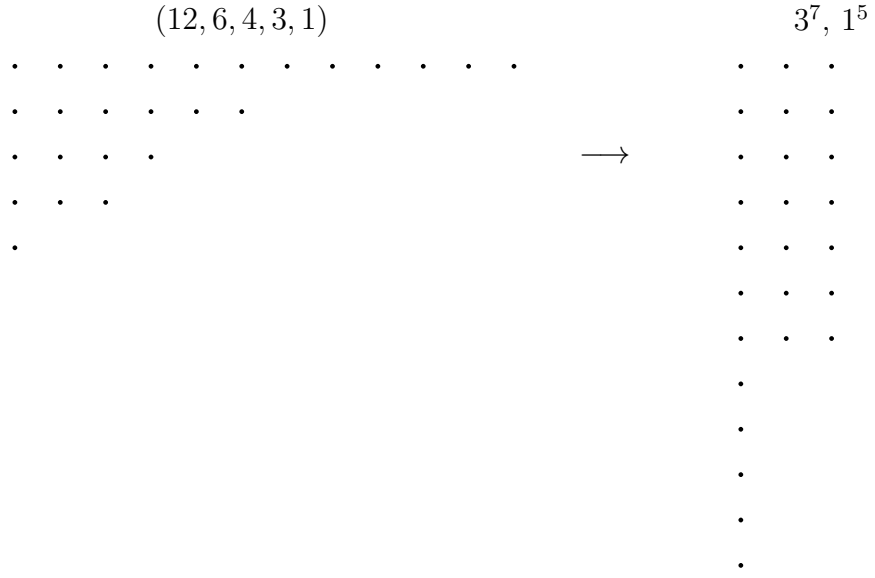
Grouping according to the odd parts, we obtain

$$\begin{aligned} 26 &= 3 \cdot 2^2 + 3 \cdot 2^1 + 1 \cdot 2^2 + 3 \cdot 2^0 + 1 \cdot 2^0 \\ &= 3(2^0 + 2^1 + 2^2) + 1(2^0 + 2^2) \\ &= 3 \cdot 7 + 1 \cdot 5. \end{aligned}$$

This yields the partition into odd parts

$$26 = \underbrace{3 + \cdots + 3}_{7 \text{ times}} + \underbrace{1 + \cdots + 1}_{5 \text{ times}}.$$

All steps are reversible, since every integer admits a unique representation as a sum of distinct powers of 2*.*



Define $p_e(n)$ and $p_o(n)$ as the number of partitions of n into an even and odd number of distinct parts, respectively. For example,

$$p_e(7) = 3 \quad (1+6,\ 2+5,\ 3+4),$$

$$p_o(7) = 2 \quad (1+2+4,\ 7).$$

4.2 Euler and Pentagonal Numbers

Let us consider the function

$$P(x)^{-1} = \prod_{k=1}^{\infty}(1-x^k).$$

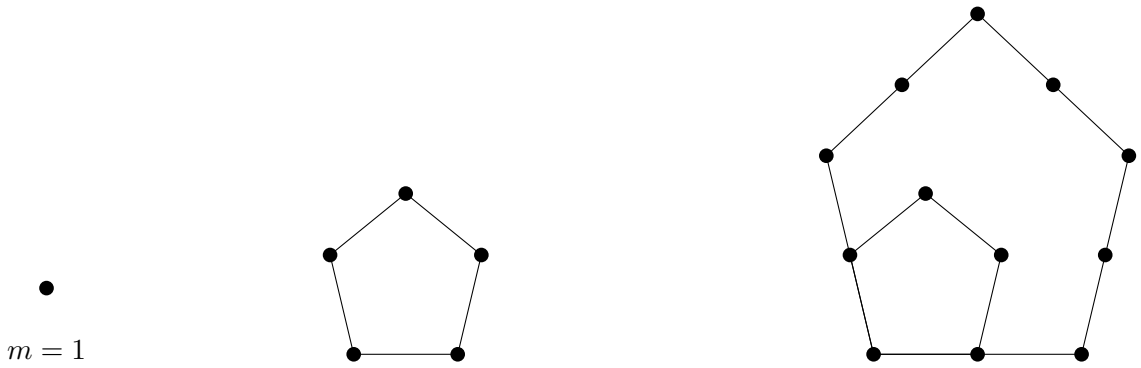
In the expansion, a partition of n into unequal parts contributes a $+1$ to the coefficient of x^n if the number of parts is even and a -1 if it is odd.

So if we denote by $p_e(n)$ the number of partitions of n into even unequal parts and by $p_o(n)$ the number of partitions into odd unequal parts, then the coefficient of x^n is $p_e(n)-p_o(n)$. It was shown by Euler that $p_e(n) = p_o(n)$, unless $n = \omega(m) = (3m^2-m)/2$ or $n = \omega(-m) = (3m^2+m)/2$, in which case we have $p_e(n)-p_o(n) = (-1)^m$. The numbers $\omega(m)$ and $\omega(-m)$ are called *pentagonal numbers*.

The name comes from the following relation (which can easily be checked):

$$\omega(m) = \sum_{k=0}^{m-1}(3k+1)$$

and the following picture:



$m = 1$
1

$m = 2$
$1 + 4 = 5$

$m = 3$
$1 + 4 + 7 = 12$

Due to the visual representation of partition, we may now also prove things by drawing.

Theorem 4.23 (Euler's identity)**.**

$$\prod_{k=1}^{\infty}(1-x^k) = 1 + \sum_{m=1}^{\infty}(-1)^m\left(x^{\omega(m)} + x^{\omega(-m)}\right).$$

Proof. Recall that the coefficient of x^n in

$$\prod_{k=1}^{\infty}(1-x^k)$$

is

$$p_e(n) - p_o(n),$$

where $p_e(n)$ and $p_o(n)$ denote the number of partitions of n into an even and odd number of unequal parts, respectively.

We shall construct a bijection between partitions of n into an even number of unequal parts and partitions of n into an odd number of unequal parts, except in certain exceptional cases.

Let

$$n = x_1 + x_2 + \cdots + x_k \qquad \text{with} \qquad x_1 > x_2 > \cdots > x_k \geq 1$$

be a partition of n into unequal parts, and consider its Young diagram.

The number of dots in the last row is called the *base* and will be denoted by b.

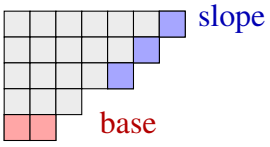
Starting from the last dot of the top row, draw the longest diagonal line of dots going downwards to the left. The number of dots in this diagonal is called the *slope* and will be denoted by s.

For example, consider the partition

$$23 = 7 + 6 + 5 + 3 + 2.$$

Its Young diagram has base length $b = 2$ and slope length $s = 3$:

$$23 = 7 + 6 + 5 + 3 + 2$$



Since here $b \leq s$, we define the following operation.

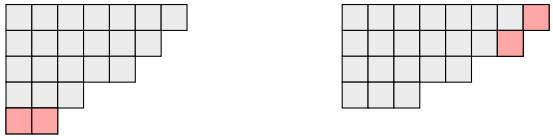
Operation A. If $b \leq s$, remove the base and attach it to the right-hand side of the diagram so that it forms a new diagonal line parallel to the old slope.

In the example above, this gives the new partition

$$23 = 8 + 7 + 5 + 3.$$

$$7+6+5+3+2 \qquad\qquad 8+7+5+3$$

$$\xrightarrow{A}$$



Conversely, if $b > s$, we define the reverse operation.

Operation B. If $b > s$, remove the slope and attach it below the diagram to form a new base.

For instance, starting from

$$23 = 8+7+5+3,$$

we have $b = 3$ and $s = 2$, so operation B applies and recovers the partition

$$23 = 7+6+5+3+2.$$

$$8+7+5+3 \qquad\qquad 7+6+5+3+2$$

$$\xrightarrow{B}$$

In all non-exceptional cases, exactly one of the operations A or B can be applied, and they are inverse to one another. Moreover, applying either operation changes the number of parts by 1, so it switches parity. Hence partitions into an even number of unequal parts are paired with partitions into an odd number of unequal parts, and therefore

$$p_e(n) - p_o(n) = 0$$

unless we are in an exceptional case.

The exceptional cases occur precisely when the base and the slope overlap.

- If $b = s$, then operation A cannot be carried out. In this case the diagram has row lengths

$$b,\ b+1,\ \ldots,\ 2b-1,$$

so

$$n = b + (b+1) + \cdots + (2b-1) = \omega(b).$$

- If $b = s+1$, then operation B cannot be carried out. In this case the diagram has row lengths

$$s+1,\ s+2,\ \ldots,\ 2s,$$

so

$$n = (s+1) + (s+2) + \cdots + 2s = \omega(-s).$$

Thus, if n is not of the form $\omega(m)$ or $\omega(-m)$, then $p_e(n) = p_o(n)$. In the exceptional cases there is exactly one unpaired partition, and its contribution is $(-1)^m$. Therefore

$$\prod_{k=1}^{\infty}(1-x^k) = 1 + \sum_{m=1}^{\infty}(-1)^m\big(x^{\omega(m)} + x^{\omega(-m)}\big).$$

□

This then lead Euler to the following recursion formula:

Theorem 4.24. *Let n be an integer and let us define $p(n) = 0$ for any $n < 0$. Then for $n \geq 1$ we have*

$$p(n) = \sum_{m=1}^{\infty}(-1)^{m+1}\{p(n-\omega(m)) + p(n-\omega(-m))\}.$$

Proof. This follows from Theorems 4.13 and 4.23. □

Example 4.25. *If we try to compute this recursion, we get*

$$n = 1:\ p(1) = 1, \qquad n = 2:\ p(2) = 2.$$

These are the same initial values as for the Fibonacci numbers. However, the partition numbers will not continue in the same way, since their recursion is different.

We give the asymptotic behavior of $p(n)$, without proving it, as it would require too much of analysis:

$$\lim_{n\to\infty} n^{-1/2}\log(p(n)) = \pi\sqrt{2/3}.$$

Euler's identity is a special case of the Jacobi identity (to which we will also omit the proof).

Theorem 4.26.

$$\prod_{n=1}^{\infty}(1-q^{2n})(1+q^{2n-1}t)(1+q^{2n-1}t^{-1}) = \sum_{r=-\infty}^{\infty} q^{r^2}t^r.$$

Exercise 4.27. *Let n be a positive integer. Show that the number of self-conjugate partitions of n is the same as the number of partitions of n into unequal parts.*

4.3 Young Tableaux

As a final problem related to partitions, we want to consider *Young tableaux* also called *standard tableaux*. A Young tableau of shape $(n_1, \ldots, n_m)$ is a Young diagram in which the integers $\{1, \ldots, n\}$ are inserted into the squares, in such a way that all rows and columns are increasing. We say T is a tableau of $(n_1, \ldots, n_m)$ or T has shape $(n_1, \ldots, n_m)$. If we do not impose any rules, we call it a tableau. There are $n!$ tableaux of shape $(n_1, \ldots, n_m)$.

Example 4.28. *Let us consider again $n = 12$ and the partition $(5, 4, 2, 1)$. We can represent it by the Young tableau:*

1	2	3	4	5
6	7	8	9	
10	11			
12				

Is the Young tableau unique?

Not at all! A Young diagram is uniquely determined by a partition. However, a Young tableau is obtained by filling the diagram with numbers, and in general there are many different tableaux of the same shape.

Example 4.29. *Let us consider again the partition $(5, 4, 2, 1)$ of $n = 12$. There are several different Young tableaux of this shape. For example:*

1	2	4	7	9
3	5	8	10	
6	11			
12				

Both tableaux have the same shape $(5, 4, 2, 1)$, but differ in their fillings. Hence Young tableaux are not unique.

We want to count the number of Young tableaux of a given shape. For this we first introduce the following function:

$$\Delta(x_1, \ldots, x_m) = \prod_{1 \le i < j \le m} (x_i - x_j).$$

You may recognize this, as the determinant of the Vandermonde matrix

$$V(x_1, \ldots, x_m) = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_m \\ \vdots & & \vdots \\ x_1^{m-1} & \cdots & x_m^{m-1} \end{pmatrix}.$$

This expression is antisymmetric in the variables $x_1, \dots, x_m$ and vanishes whenever $x_i = x_j$ for some $i \neq j$.

Lemma 4.30. *Let m be a positive integer and define*

$$g(x_1, \dots, x_m, y) = x_1\Delta(x_1 + y, x_2, \dots, x_m) + x_2\Delta(x_1, x_2 + y, \dots, x_m) + \cdots + x_m\Delta(x_1, \dots, x_m + y).$$

Then

$$g(x_1, \dots, x_m, y) = \left(\sum_{i=1}^{m} x_i + \binom{m}{2} y\right) \Delta(x_1, \dots, x_m).$$

Proof. First note that g is a homogeneous polynomial of degree

$$1 + \deg(\Delta) = 1 + \binom{m}{2}$$

in the variables $x_1, \dots, x_m, y$.

We claim that g is alternating in the variables $x_1, \dots, x_m$. Indeed, let σ be a transposition exchanging x_i and x_j. Applying σ to g exchanges the i-th and j-th summands, and since each Vandermonde determinant changes sign when two variables are swapped, the whole expression changes sign. Thus

$$g(\dots, x_i, \dots, x_j, \dots, y) = -g(\dots, x_j, \dots, x_i, \dots, y).$$

Hence $g = 0$ whenever $x_i = x_j$, so g is divisible by $x_i - x_j$ for every $1 \leq i < j \leq m$. Therefore g is divisible by

$$\Delta(x_1, \dots, x_m) = \prod_{1 \leq i < j \leq m} (x_i - x_j).$$

Since

$$\deg(g) = 1 + \deg(\Delta),$$

it follows that the quotient

$$h(x_1, \dots, x_m, y) := \frac{g(x_1, \dots, x_m, y)}{\Delta(x_1, \dots, x_m)}$$

is a homogeneous polynomial of degree 1. Thus h must be of the form

$$h(x_1, \dots, x_m, y) = a_1x_1 + \cdots + a_mx_m + by$$

for suitable constants $a_1, \dots, a_m, b$.

To determine these constants, first set $y = 0$. Then

$$g(x_1, \dots, x_m, 0) = (x_1 + \cdots + x_m)\Delta(x_1, \dots, x_m),$$

since each summand is simply $x_i\Delta(x_1,\ldots,x_m)$. Therefore

$$h(x_1,\ldots,x_m,0) = x_1+\cdots+x_m,$$

so

$$a_1 = \cdots = a_m = 1.$$

Thus

$$h(x_1,\ldots,x_m,y) = \sum_{i=1}^{m} x_i + by.$$

It remains to determine b. For this, expand the i-th summand:

$$x_i\,\Delta(x_1,\ldots,x_i+y,\ldots,x_m) = x_i\prod_{r<i}(x_r-x_i-y)\prod_{s>i}(x_i+y-x_s).$$

Factoring out $\Delta(x_1,\ldots,x_m)$, this becomes

$$x_i\,\Delta(x_1,\ldots,x_m)\prod_{r<i}\left(1+\frac{y}{x_i-x_r}\right)\prod_{s>i}\left(1+\frac{y}{x_i-x_s}\right).$$

Hence, modulo terms of degree at least 2 in y, the coefficient of y in the i-th summand is

$$x_i\left(\sum_{j\neq i}\frac{1}{x_i-x_j}\right)\Delta(x_1,\ldots,x_m).$$

Summing over i, the coefficient of y in g is therefore

$$\left(\sum_{i=1}^{m} x_i\sum_{j\neq i}\frac{1}{x_i-x_j}\right)\Delta(x_1,\ldots,x_m).$$

Now group the terms corresponding to each unordered pair $\{i,j\}$:

$$\frac{x_i}{x_i-x_j}+\frac{x_j}{x_j-x_i} = \frac{x_i-x_j}{x_i-x_j} = 1.$$

There are exactly $\binom{m}{2}$ such pairs, so the coefficient of y in g is

$$\binom{m}{2}\Delta(x_1,\ldots,x_m).$$

Thus $b=\binom{m}{2}$.

We conclude that

$$h(x_1,\ldots,x_m,y) = \sum_{i=1}^{m} x_i + \binom{m}{2}y,$$

and therefore

$$g(x_1,\ldots,x_m,y) = \left(\sum_{i=1}^{m} x_i + \binom{m}{2}y\right)\Delta(x_1,\ldots,x_m).$$

□

We now introduce a function $f(n_1, \ldots, n_m)$ with the following properties:

1. $f(n_1, \ldots, n_m) = 0$ unless $n_1 \geq n_2 \geq \cdots \geq n_m$,
2. $f(n_1, \ldots, n_m, 0) = f(n_1, \ldots, n_m)$,
3. $f(n_1, \ldots, n_m) = f(n_1 - 1, n_2, \ldots, n_m) + \cdots + f(n_1, \ldots, n_m - 1)$ if $n_1 \geq \cdots \geq n_m \geq 0$,
4. $f(n) = 1$ if $n \geq 0$.

We now show that $f(n_1, \ldots, n_m)$ counts the number of Young tableaux of shape $(n_1, \ldots, n_m)$.

The condition 1., 2. and 4. are clear. To see that the number of Young tableaux of shape $(n_1, \ldots, n_m)$ satisfies the third condition, let us consider the entry n. It must be the last entry in one of the rows and if we remove its square, we get a Young tableaux for $n - 1$. Indeed, if some rows have the same length, say $n_i = n_{i+1}$, then the last boxes of these rows lie in the same column. Since entries increase strictly down columns, the entry n can only be placed in the lowest such row.

Thus, n occupies the last box of exactly one row i, and removing this box yields a Young tableau of shape $(n_1, \ldots, n_i - 1, \ldots, n_m)$. This establishes the recursion.

Theorem 4.31. *Let m be a positive integer. The number of Young tableaux of shape $(n_1, \ldots, n_m)$ is*

$$f(n_1, \ldots, n_m) = \frac{\Delta(n_1 + m - 1, n_2 + m - 2, \ldots, n_m)\, n!}{(n_1 + m - 1)!(n_2 + m - 2)! \cdots n_m!},$$

provided

$$n_1 + m - 1 \geq n_2 + m - 2 \geq \cdots \geq n_m.$$

Proof. Let

$$F(n_1, \ldots, n_m) := \frac{\Delta(n_1 + m - 1, n_2 + m - 2, \ldots, n_m)\, n!}{(n_1 + m - 1)!(n_2 + m - 2)! \cdots n_m!},$$

where

$$n = n_1 + \cdots + n_m.$$

We show that F satisfies the four conditions that characterize f.

Condition 1. If for some i we have

$$n_i < n_{i+1},$$

then

$$n_i + m - i \leq n_{i+1} + m - (i + 1),$$

and equality occurs precisely when

$$n_i + m - i = n_{i+1} + m - (i + 1).$$

In that case two arguments of Δ coincide, so

$$\Delta(n_1 + m - 1, \ldots, n_m) = 0.$$

Hence $F(n_1, \ldots, n_m) = 0$, as required.

Condition 2. If $n_m = 0$, then the last argument of Δ is 0, and the formula reduces to the same expression with the last variable omitted. Thus

$$F(n_1, \ldots, n_m, 0) = F(n_1, \ldots, n_m).$$

Condition 4. If $m = 1$, then $\Delta = 1$, so

$$F(n) = \frac{n!}{n!} = 1 \qquad \text{for all } n \geq 0.$$

Condition 3. Set

$$x_i := n_i + m - i \qquad (i = 1, \ldots, m).$$

Then

$$x_1 + \cdots + x_m = (n_1 + \cdots + n_m) + \sum_{i=1}^{m} (m - i) = n + \binom{m}{2}.$$

Also,

$$F(n_1, \ldots, n_m) = \frac{\Delta(x_1, \ldots, x_m)\, n!}{x_1! \cdots x_m!}.$$

Now consider

$$F(n_1 - 1, n_2, \ldots, n_m), \quad F(n_1, n_2 - 1, \ldots, n_m), \quad \ldots, \quad F(n_1, \ldots, n_m - 1).$$

Decreasing n_i by 1 decreases x_i by 1, so

$$F(n_1, \ldots, n_i - 1, \ldots, n_m) = \frac{\Delta(x_1, \ldots, x_i - 1, \ldots, x_m)\, (n-1)!}{x_1! \cdots (x_i - 1)! \cdots x_m!}.$$

Since

$$(x_i - 1)! = \frac{x_i!}{x_i},$$

this becomes

$$F(n_1, \ldots, n_i - 1, \ldots, n_m) = \frac{x_i\, \Delta(x_1, \ldots, x_i - 1, \ldots, x_m)\, (n-1)!}{x_1! \cdots x_m!}.$$

Therefore

$$\begin{aligned} &F(n_1 - 1, n_2, \ldots, n_m) + \cdots + F(n_1, \ldots, n_m - 1) \\ &\quad = \frac{(n-1)!}{x_1! \cdots x_m!} \Big(x_1 \Delta(x_1 - 1, x_2, \ldots, x_m) + \cdots + x_m \Delta(x_1, \ldots, x_m - 1) \Big). \end{aligned}$$

Now apply Lemma 4.30 with $y = -1$. This gives

$$x_1\Delta(x_1-1,x_2,\ldots,x_m)+\cdots+x_m\Delta(x_1,\ldots,x_m-1)=\left(\sum_{i=1}^{m}x_i-\binom{m}{2}\right)\Delta(x_1,\ldots,x_m).$$

Since

$$\sum_{i=1}^{m}x_i=n+\binom{m}{2},$$

the bracket simplifies to

$$n.$$

Hence

$$\begin{aligned}&F(n_1-1,n_2,\ldots,n_m)+\cdots+F(n_1,\ldots,n_m-1)\\&\quad=\frac{(n-1)!\,n\,\Delta(x_1,\ldots,x_m)}{x_1!\cdots x_m!}=\frac{n!\,\Delta(x_1,\ldots,x_m)}{x_1!\cdots x_m!}=F(n_1,\ldots,n_m).\end{aligned}$$

So condition 3 holds.

Thus F satisfies conditions 1, 2, 3, and 4. Since these conditions characterize the number of Young tableaux of shape $(n_1,\ldots,n_m)$, we conclude that

$$f(n_1,\ldots,n_m)=F(n_1,\ldots,n_m).$$

□

There exists another formulation, which makes this theorem more interesting. For this we introduce the *hook* of a Young tableaux. The hook of a cell in the tableau is the union of the cell and all cells to the right of it (in the same row) and below it (in the same column). The *hook length* is the number of cells in the hook.

Example 4.32. *Let us consider again $n = 12$ and the Young tableau of shape $(5,4,2,1)$. For the cell in the second row and second column, the hook is highlighted in blue.*

1	2	3	4	5
6	7	8	9	
10	11			
12				

This cell (7) has hook length 4.

Theorem 4.33. *The number of Young tableaux of a given shape and a total of n cells is $n!$ divided by the product of all hook lengths.*

Proof. If we have a Young tableau of shape $(n_1, \ldots, n_m)$, the hook length for the cell in row 1 and column 1 is clearly $n_1 + m - 1$.

Consider the hook lengths of the cells in the first row (you may compare this with the example above). The hook length of the leftmost cell is

$$n_1 + m - 1.$$

Now move from left to right along the row. Usually, the hook length decreases by 1: we lose one cell to the right, while the column height stays the same.

However, whenever the column height drops, the hook length decreases by more than 1. Indeed, if a column has height j, then after the last column of that height the number of cells below decreases by 1, so one hook length is skipped.

More precisely, the columns of height at least j are exactly the first n_j columns. Hence the skipped hook lengths are

$$(n_1 + m - 1) - (n_j + m - j)$$

for $j \in \{2, \ldots, m\}$.

Therefore the hook lengths occurring in the first row are precisely all integers from 1 to $n_1 + m - 1$, except for the numbers

$$(n_1 + m - 1) - (n_j + m - j)$$

for $j \in \{2, \ldots, m\}$.

It follows that the product of all hook lengths is given by

$$\prod_{\text{cells}} h(c) = \frac{\prod_{i=1}^{m}(n_i + m - i)!}{\Delta(n_1 + m - 1, \ldots, n_m)}.$$

Combining this with Lemma 4.30 yields

$$f(n_1, \ldots, n_m) = \frac{n!}{\prod_{\text{cells}} h(c)}.$$

□

4.4 The Twelvefold Way

Many combinatorial problems can be phrased as distributing b balls into u urns. The answer depends on:

- whether the balls are *labeled* or *unlabeled*,
- whether the urns are *labeled* or *unlabeled*,
- and whether restrictions are imposed:

$*$ unrestricted,

≤ 1 at most one ball per urn,

≥ 1 at least one ball per urn.

This leads to $2 \times 2 \times 3 = 12$ different counting problems. Let us denote by

$$[b \leq u] = \begin{cases} 1 & \text{if } b \leq u, \\ 0 & \text{else.} \end{cases}$$

Balls	Urns	$*$	≤ 1	≥ 1
labeled	labeled	u^b	$(u)_b$	$u!\,S(b,u)$
unlabeled	labeled	$\binom{u+b-1}{b}$	$\binom{u}{b}$	$\binom{b-1}{u-1}$
labeled	unlabeled	$\sum_{i=1}^{u} S(b,i)$	$[b \leq u]$	$S(b,u)$
unlabeled	unlabeled	$\sum_{i=1}^{u} p_i(b)$	$[b \leq u]$	$p_u(b)$

1. Labeled balls, labeled urns, unrestricted number of balls per urn Each ball independently chooses one of the u urns. Since there are u choices for each of the b balls, the multiplication principle gives u^b. Equivalently, this counts all functions

$$f : \{1, \ldots, b\} \to \{1, \ldots, u\}.$$

2. Labeled balls, labeled urns, at most one ball per urn We place the balls one by one. The first ball has u choices, the second has $u - 1$ choices, and so on. Hence

$$u(u-1)\cdots(u-b+1) = (u)_b.$$

This counts injective functions from the set of balls to the set of urns.

3. Labeled balls, labeled urns, at least one ball per urn First ignore the labels on the urns. Then we are partitioning the balls into u nonempty sets, which can be done in $S(b,u)$ ways.

Since the urns are labeled, we may permute the u subsets among the u urn labels in $u!$ ways. Therefore the total number is $u!S(b,u)$. Equivalently, this counts surjective functions.

4. Unlabeled balls, labeled urns, unrestricted Since the balls are indistinguishable, a distribution is completely determined by the numbers $x_1, \ldots, x_u$ of balls in the urns. Thus we count nonnegative integer solutions of $x_1 + \cdots + x_u = b$. Using the stars-and-bars argument, the number of solutions is

$$\binom{u+b-1}{b}.$$

5. Unlabeled balls, labeled urns, at most one ball per urn Since each urn may contain at most one ball, a distribution is determined by choosing which b urns contain a ball. Therefore

$$\binom{u}{b}.$$

6. Unlabeled balls, labeled urns, at least one ball per urn We first place one ball into each urn. This guarantees the condition that every urn is nonempty. We now have $b-u$ balls remaining, which may be distributed freely among the u urns.
By the unrestricted case,

$$\binom{u+(b-u)-1}{b-u} = \binom{b-1}{u-1}.$$

Equivalently, this counts positive integer solutions of

$$x_1 + \cdots + x_u = b.$$

7. Labeled balls, unlabeled urns, unrestricted If exactly i urns are used, then we partition the balls into i nonempty unlabeled sets, which can be done in $S(b,i)$ ways.

Since the number of nonempty urns may vary between 1 and u, we sum:

$$\sum_{i=1}^{u} S(b,i).$$

8. Labeled balls, unlabeled urns, at most one ball per urn If $b > u$, then there are too many balls and the answer is 0.

If $b \leq u$, then every occupied urn contains exactly one ball. Since the urns are unlabeled, all such arrangements are indistinguishable. Hence there is exactly one possibility. Therefore the answer is

$$[b \leq u].$$

9. Labeled balls, unlabeled urns, at least one ball per urn This is exactly the definition of the Stirling numbers of the second kind: $S(b,u)$. In fact, we partition the labeled balls into u nonempty unlabeled subsets.

10. Unlabeled balls, unlabeled urns, unrestricted If exactly i urns are nonempty, then we are partitioning the integer b into i positive parts. This can be done in $p_i(b)$ ways. Summing over all possible numbers of nonempty urns gives

$$\sum_{i=1}^{u} p_i(b).$$

11. Unlabeled balls, unlabeled urns, at most one ball per urn If $b > u$, then no arrangement exists. If $b \leq u$, then every occupied urn contains exactly one ball. Since neither balls nor urns are distinguishable, all arrangements are identical. Hence the number of possibilities is $[b \leq u]$.

12. Unlabeled balls, unlabeled urns, at least one ball per urn A distribution corresponds exactly to a partition of the integer b into u positive parts:

$$b = x_1 + \cdots + x_u.$$

Therefore the number of possibilities equals $p_u(b)$.

Takeaway The Twelvefold Way unifies many classical counting problems:

- functions,
- injections,
- surjections,
- multisets,
- set partitions,
- integer partitions.

Many combinatorial formulas can be interpreted as counting balls-and-urns problems with different symmetries and restrictions.

5 Posets and Lattices

5.1 Partially Ordered Sets

This chapter is taught by the guest lecturer Jessica Bariffi

We often have objects that have different aspects, which makes them difficult to compare. Think about your reasoning of joining this lecture: You might have evaluated

- how interesting the lecture is,
- how difficult the exercises are,
- how nice the exam grading is,
- how early the lecture starts.

One course may have excellent lectures but very hard exercises, while another may have easy exercises but the lecture starts at 8am. Thus two courses are often not directly comparable. This naturally leads to the concept of a partial order.

To model such situations, we introduce partially ordered sets.

Definition 5.1. A *partially ordered set* (or *poset*) is a pair $P = (X, \leq)$, where X is a set and $\leq$ is a binary relation satisfying:

1. reflexivity: $x \leq x$ for all $x \in X$,
2. transitivity: $x \leq y$ and $y \leq z$ then $x \leq z$,
3. antisymmetry: $x \leq y$ and $y \leq x$ then $x = y$.

Note that if $x \leq y$ and $x \neq y$, we write $x < y$. We will often exchange the notion of P and X, when the binary relation $\leq$ is clear or fixed.

Example 5.2. *Examples of posets include:*

- *the integers ordered by size:* $(\mathbb{Z}, \leq)$,
- *subsets ordered by inclusion:* $(\mathcal{P}(S), \subseteq)$*, for some set* S.

A *total order* is a special case of a partial order where every pair of elements is comparable, that is: for all $x, y \in X$ we either have $x \leq y$ or $y \leq x$.

Exercise 5.3. *Consider the two posets* $(\mathbb{Z}, \leq)$*, and* $(\mathcal{P}(S), \subseteq)$*. Which one is a total order?*

One of our most used examples, is called a *Boolean lattice*:

Definition 5.4. The *Boolean lattice* B_n is the poset $B_n = (\mathcal{P}([n]), \subseteq)$, consisting of all subsets of $[n]$ ordered by inclusion.

Example 5.5. *The Boolean lattice B_3 consists of the subsets*

$$\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}.$$

If $(X, \leq)$ is not a total order, we may find two elements, which are incomparable:

Definition 5.6. Let $(X, \leq)$ be a poset. If neither $x \leq y$ nor $y \leq x$, then x and y are called *incomparable*, denoted by $x \parallel y$.

Example 5.7. *Let us consider the Boolean lattice B_3. Then the sets $X = \{1,2\}$ and $Y = \{1,3\}$ are incomparable, since $\{1,2\} \nsubseteq \{1,3\}$ and $\{1,3\} \nsubseteq \{1,2\}$.*

For a given poset $(X, \leq)$ and some element x, the elements which are directly "above" x, with no other element in between, will play a crucial role:

Definition 5.8. Let $(X, \leq)$ be a poset. If $x < y$ and there exists no z such that $x < z < y$, then y is said to *cover* x, denoted by $x \lessdot y$.

Covers, however, are not unique, as the following example shows.

Example 5.9. *Consider again the Boolean lattice B_3. The set $X = \{1\}$ is covered by $Y = \{1,2\}$, since $\{1\} \subset \{1,2\}$, and there exists no subset Z such that $\{1\} \subset Z \subset \{1,2\}$. However, X is also covered by $\{1,3\}$.*

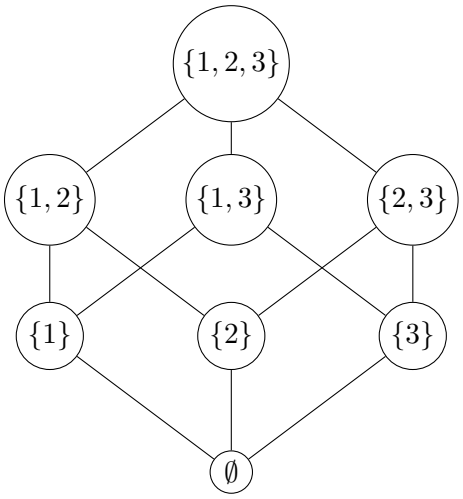
Posets are usually represented graphically by *Hasse diagrams*.

Definition 5.10. Let $(X, \leq)$ be a poset. Its *Hasse diagram* is obtained as follows:

- every element of X is represented by a point,
- if $x \lessdot y$, then we connect x and y by a line segment,
- larger elements (ordered by $\leq$) are drawn above smaller ones.

Note that transitive relations are omitted.

Example 5.11. *Consider the Boolean lattice B_3. The Hasse diagram is given by:*



We can already see in the diagram, that the covering relation can be extended to several elements, forming a chain:

Definition 5.12. Let $(X, \leq)$ be a poset.

- A *chain* is a subset of X in which every two elements are comparable.
- An *antichain* is a subset of X in which every two distinct elements are incomparable.

Example 5.13. *In the Boolean lattice B_3*

$$\emptyset \subset \{1\} \subset \{1, 2, 3\}$$

is a chain, while

$$\{1\}, \{2\}, \{3\}$$

is an antichain.

Definition 5.14. Let $P = (X, \leq)$ be a poset. A chain

$$x_0 < x_1 < \cdots < x_k$$

is called *saturated* if $x_{i-1} \lessdot x_i$ for all i, That is, every element covers the previous one.

Example 5.15. *Consider the Boolean lattice B_3. The chain*

$$\emptyset \subset \{1\} \subset \{1, 2\} \subset \{1, 2, 3\}$$

is saturated, since every set covers the previous one.

We can also speak of the *length* of a chain (or antichain), by simply counting how many elements it contains. This then gives rise to the following definition:

Definition 5.16. The *height* of a poset is the length of a largest chain.

Example 5.17. *The height of the Boolean lattice B_3 is 4. Indeed, a largest chain is*

$$\emptyset \subset \{1\} \subset \{1, 2\} \subset \{1, 2, 3\}.$$

This chain contains 4 elements.

More generally, the height of B_n is $n + 1$, since a maximal chain contains exactly one subset of each cardinality $0, 1, \ldots, n$.

Exercise 5.18. *Is a saturated chain always a chain of the largest length?*

Similarly, if we consider the longest antichain, we get the width:

Definition 5.19. The *width* of a poset is the length of a largest antichain.

Example 5.20. *The width of the Boolean lattice B_3 is 3. Indeed, the subsets*

$$\{1\}, \{2\}, \{3\}$$

form an antichain, since no one of them is contained in another. Similarly,

$$\{1, 2\}, \{1, 3\}, \{2, 3\}$$

is also an antichain of size 3.

There exists no larger antichain, so the width of B_3 is 3.

To compute the width of any Boolean lattice is not as simple as its height, we will compute it in Sperner's Theorem 5.27.

The notions of chains and antichains are dual to each other. In fact, a chain and an antichain can intersect in at most one element.

Theorem 5.21 (Dilworth)**.** *Let $P = (X, \leq)$ be a poset. Then the minimum number of disjoint chains, which together contain all elements of X, is the same as the maximum number of elements in an antichain.*

Proof. Let us denote the minimum number of chains, which together contain all elements of X by m and the maximum length of an antichain by M. Note that the chains are disjoint and contain all elements of X, hence they partition P. We first show that $m \geq M$.

Indeed, let $A = \{a_1, \ldots, a_M\}$ be an antichain of maximal length. Since no two elements of an antichain are comparable, a chain can contain at most one element of A. Hence every chain partition of P requires at least M chains. Thus $m \geq M$.

We now prove the converse inequality $m \leq M$ by induction on $|X|$.

If $|X| = 0$, there is nothing to prove. So assume the statement holds for all smaller posets Y. Let C be a maximal chain in P. If every antichain in $P \setminus C = (X \setminus C, \leq)$ has size at most $M - 1$, then by the induction hypothesis, the poset $P \setminus C$ can be partitioned into at most $M - 1$ chains. Adding the chain C gives a partition of P into at most M chains, and we are done.

Hence we may assume that $P \setminus C$ contains an antichain $\{a_1, \ldots, a_M\}$ of size M.

Now define

$$S^- := \{x \in X \mid x \leq a_i \text{ for some } i\},$$

that is, all elements lying "below" at least one of the a_i.

Similarly define

$$S^+ := \{x \in X \mid x \geq a_i \text{ for some } i\}.$$

Since C is maximal, its largest element cannot lie in S^-. Indeed, if the largest element of C satisfied $x \leq a_i$, then either $x = a_i$ or $x < a_i$. But $a_i \notin C$, so in the second case we could extend the chain C by adding a_i, contradicting its maximality.

Hence S^- is strictly smaller than P, so we may apply the induction hypothesis to it.

The antichain $\{a_1, \ldots, a_M\}$ is still maximal in S^-, so S^- can be partitioned into exactly M chains: $S_1^-, \ldots, S_M^-$, where we may assume that $a_i \in S_i^-$. We claim that each a_i is the maximal element of the chain S_i^-.

Suppose not. Then there exists some $x \in S_i^-$ such that $x > a_i$. Since $x \in S^-$, there exists some j such that $x \leq a_j$.

Combining these inequalities gives $a_i < a_j$, contradicting the fact that the a_i form an antichain. Thus every a_i is maximal in its chain.

By symmetry, we can similarly partition S^+ into chains $S_1^+, \ldots, S_M^+$, where each a_i is the minimal element of S_i^+.

Finally, for every i, we glue the chains S_i^- and S_i^+ together along the common element a_i. This produces M disjoint chains whose union is all of P. Hence $m \leq M$. □

Let us visualize this in our previous example:

Example 5.22. *Consider the Boolean lattice B_3. We already observed that*

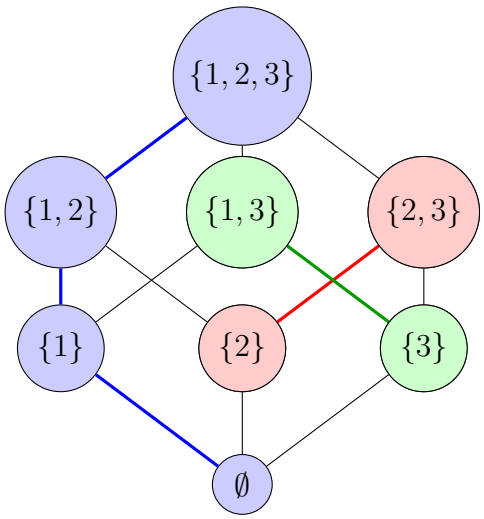
$$\{1\}, \{2\}, \{3\}$$

form an antichain of size 3. Hence the width of B_3 is at least 3. Dilworth's Theorem tells us that B_3 can therefore be partitioned into exactly 3 chains. One possible partition is:

$$\emptyset \subset \{1\} \subset \{1,2\} \subset \{1,2,3\},$$

$$\{2\} \subset \{2,3\},$$

$$\{3\} \subset \{1,3\}.$$



Thus all elements of B_3 are covered by 3 disjoint chains.

Dilworth's Theorem tells us that the width of a poset equals the minimum number of chains needed to cover it. There also exists a beautiful dual statement involving antichains. For this, we first need to define what minimal and maximal elements in a poset are.

Definition 5.23. Let $P = (X, \leq)$ be a poset and let $x \in X$. The element x is called *maximal* if there exists no element $y \in X$ such that $x < y$. Similarly, we say the element x is *minimal* if there exists no element $y \in X$ such that $y < x$.

Maximal, respectively minimal, elements do not have to be unique:

Example 5.24. *Consider the Boolean lattice B_3. The unique maximal element is $\{1, 2, 3\}$, while the unique minimal element is $\emptyset$. If we instead consider only the subsets*

$$\{1\}, \{2\}, \{1, 2\},$$

then both $\{1\}, \{2\}$ are minimal.

Theorem 5.25 (Mirsky)**.** *Let $P = (X, \leq)$ be a finite poset. If the largest chain in P has size m, then P can be partitioned into m antichains.*

Proof. We proceed by induction on m. If $m = 1$, then no two distinct elements are comparable, so P itself is an antichain. Hence the statement is true.

Now assume the theorem holds for all posets whose largest chain has size at most $m - 1$, and let P be a poset whose largest chain has size m. Let M be the set of maximal elements of P. We first claim that M is an antichain.

Indeed, suppose $x, y \in M$ with $x < y$. Then x cannot be maximal, since it lies strictly below y, a contradiction. Thus M is an antichain.

Now consider the poset $P \setminus M$. We claim that every chain in $P \setminus M$ has size at most $m - 1$.

Indeed, suppose there existed a chain $x_1 < x_2 < \cdots < x_m$ in $P \setminus M$. Since $x_m \notin M$, the element x_m is not maximal in P. Hence there exists some $y \in X$ such that $x_m < y$. But then $x_1 < x_2 < \cdots < x_m < y$ would be a chain of size $m + 1$, contradicting the assumption that the largest chain in P has size m.

Thus the largest chain in $P \setminus M$ has size at most $m - 1$.

By the induction hypothesis, $P \setminus M$ can be partitioned into $m - 1$ antichains. Adding the antichain M gives a partition of P into m antichains. □

We can again illustrate this in our previous example:

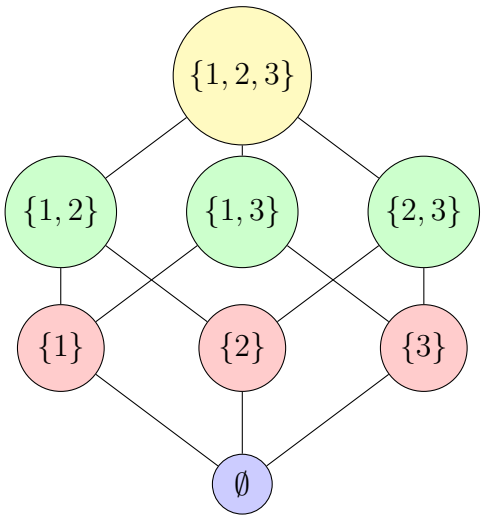
Example 5.26. *Consider again the Boolean lattice B_3. A largest chain has size 4, for example*

$$\emptyset \subset \{1\} \subset \{1,2\} \subset \{1,2,3\}.$$

Mirsky's Theorem therefore tells us that B_3 can be partitioned into 4 antichains. Indeed, one possible partition is given by the rank levels:

$$\{\emptyset\},$$

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



Each color class forms an antichain.

We can now come back to our previous question: what is the largest possible antichain in the Boolean lattice?

Theorem 5.27 (Sperner)**.** *Let B_n denote the Boolean lattice and let $\mathcal{A} \subseteq B_n$ be an antichain. Then*

$$|\mathcal{A}| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

Proof. Recall that every maximal chain has the form

$$A_0 = \emptyset \subset A_1 \subset A_2 \subset \cdots \subset A_n = [n],$$

where $|A_i| = i$ and that there are exactly $n!$ maximal chains. Indeed, every maximal chain is obtained by adding the elements of $[n]$ one-by-one. For example, the ordering

$$2, 1, 3$$

gives rise to the chain

$$\emptyset \subset \{2\} \subset \{1,2\} \subset \{1,2,3\}.$$

Conversely, every maximal chain uniquely determines the order in which the elements were added. Hence maximal chains are in bijection with permutations of $[n]$, and therefore there are exactly $n!$ maximal chains.

A fixed set A of size k is contained in exactly $k!(n-k)!$ maximal chains. To obtain a maximal chain containing a fixed set A with $|A| = k$, we must first add the elements of A in some order, and afterwards add the remaining $n-k$ elements. There are $k!$ ways to order the elements of A, and $(n-k)!$ ways to order the remaining elements. Hence A is contained in exactly $k!(n-k)!$ maximal chains.

Since an antichain contains at most one element from each maximal chain, we must have that

$$\sum_{A \in \mathcal{A}} k!(n-k)! \leq n!.$$

Dividing by $n!$ yields

$$\sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} \leq 1.$$

Since

$$\binom{n}{|A|} \leq \binom{n}{\lfloor n/2 \rfloor},$$

we obtain

$$|\mathcal{A}| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

□

We can check again from our previous example: B_3 has width 3, and in fact $\binom{3}{1} = 3$.
Recall that a minimal element, respectively a maximal element, may still be incomparable with other elements of the poset and thus they might not be unique. If we however, have a minimal element, respectively a maximal element, which is also comparable to all others, we call it the least element, respectively the greatest.

Definition 5.28. Let $P = (X, \leq)$ be a poset. An element $\hat{0} \in X$ is called the *least element* if $\hat{0} \leq x$ for all $x \in X$. Similarly, an element $\hat{1} \in X$ is called the *greatest element* if $x \leq \hat{1}$ for all $x \in X$.

Such least or greatest elements do not always exist, as the next example shows:

Example 5.29. *Consider the poset $X = \{\{1\}, \{2\}, \{1,2\}\}$ ordered by inclusion $\subseteq$. The element $\{1,2\}$ is the unique maximal element and as it is comparable to all others it is also the greatest element. The elements $\{1\}$ and $\{2\}$ are both minimal elements. However, X has no least element, since neither $\{1\} \subseteq \{2\}$ nor $\{2\} \subseteq \{1\}$.*

Exercise 5.30. *What are the least and greatest elements in the Boolean lattice B_n?*

Definition 5.31. A poset $P = (X, \leq)$ is called *graded* (or *ranked*) if there exists a function, called *rank function*, $\mathrm{rk} : X \to \mathbb{N}$ such that:

- if $x \lessdot y$, then $\mathrm{rk}(y) = \mathrm{rk}(x) + 1$,
- every minimal element has rank 0.

Example 5.32. *The Boolean lattice B_n is graded with the rank function*

$$\mathrm{rk}(A) = |A|.$$

Finally, we may also be interested in combining two posets to get a new one:

Definition 5.33. Let $P = (X, \leq_P)$ and $Q = (Y, \leq_Q)$ be posets.

- If X and Y are disjoint, we may take the *disjoint union* of P and Q, denoted by $P \bigsqcup Q$, and define it as the poset obtained by taking both posets together without introducing any new relations between their elements.
- The *product poset* $P \times Q$ is defined by

  $$(x_1, y_1) \leq (x_2, y_2)$$

  if and only if

  $$x_1 \leq_P x_2 \qquad \text{and} \qquad y_1 \leq_Q y_2.$$

- Assume that the poset P has a greatest element $\hat{1}_P$ and the poset Q has a least element $\hat{0}_Q$. The *ordinal sum* $P \oplus Q$ is obtained by identifying $\hat{1}_P = \hat{0}_Q$ and declaring $x \leq y$ for all $x \in X, y \in Y$. In other words, we place the entire poset Q above the poset P.

Let us consider also examples for these new constructions:

Example 5.34. *Consider the two posets $P = (\{a, b\}, \leq_P)$, where $a \leq_P b$, and $Q = (\{x, y\}, \leq_Q)$ where $x \leq_Q y$. Their disjoint union $P \bigsqcup Q$ with the relation $\leq$ is then given by $(\{a, b, x, y\}$ with $a \leq b, x \leq y$, but no element of P is comparable with an element of Q. Its Hasse diagram is:*

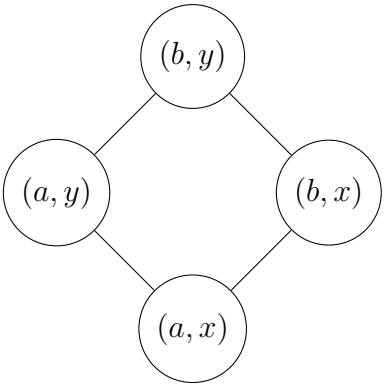
The product poset $P \times Q$ with the relation $\leq$ consists of the elements

$$(a, x),\ (a, y),\ (b, x),\ (b, y),$$

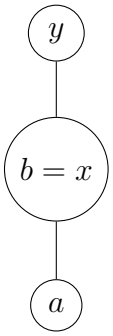
where

$$(a, x) \leq (a, y), \quad (a, x) \leq (b, x), \quad (a, x) \leq (b, y), \quad (a, y) \leq (b, y), \quad (b, x) \leq (b, y).$$

Its Hasse diagram is:



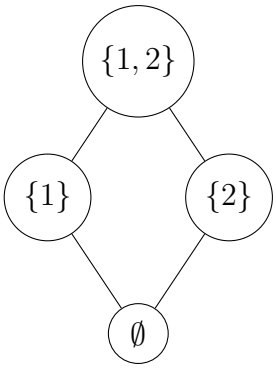
Note that this is again a Boolean lattice, namely B_2. For our third construction, we note that the greatest element of P is b, while the least element of Q is x. In the ordinal sum $P \oplus Q$, we hence identify $b = x$. The resulting poset is therefore the chain $a \leq b(= x) \leq y$ with the Hasse diagram:



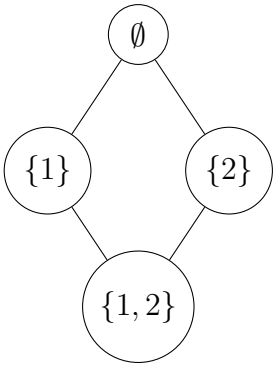
What happens if we reverse the relation $\leq$? Do we still get a poset?

Definition 5.35. Let $P = (X, \leq)$ be a poset. The *dual poset* (or *opposite poset*) of P is the poset P^{op} obtained by reversing the order relation. That is, $x \leq_{P^{\mathrm{op}}} y$ if and only if $y \leq_P x$.

Example 5.36. *Consider the Boolean lattice B_2 with the Hasse diagram:*



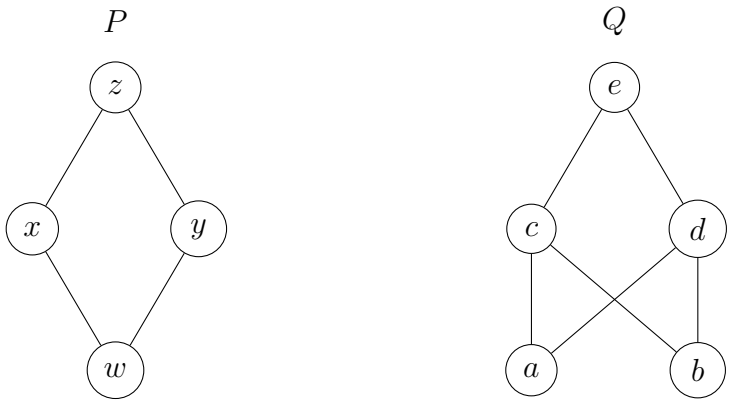
The dual poset P^{op} is obtained by turning the diagram upside down:



We see that some of our constructions contain the original posets, in fact, in this case we speak of subposets.

Definition 5.37. Let $P = (X, \leq_P)$ and $Q = (Y, \leq_Q)$ be posets. We say that P is a *subposet* of Q if there exists an injective map $f : X \to Y$ such that $x_1 \leq_P x_2$ if and only if $f(x_1) \leq_Q f(x_2)$ for all $x_1, x_2 \in X$. The map f is called an *embedding*, and the image of f is called a copy of P inside Q.

Example 5.38. *Consider the following posets.*



The map

$$w \mapsto a, \qquad x \mapsto c, \qquad y \mapsto d, \qquad z \mapsto e$$

is an embedding of P into Q. Hence P is a subposet of Q.

Similarly, we can also extend posets:

Definition 5.39. Let $P = (X, \leq_P)$ and $Q = (X, \leq_Q)$ be posets on the same ground set. We say that Q is an *extension* of P if $x \leq_P y$ implies that $x \leq_Q y$ for all $x, y \in X$.

Additionally, we call an extension Q *linear* if Q is a total order.

Example 5.40. *Consider the Boolean lattice B_2. The sets $\{1\}$ and $\{2\}$ are incomparable in P. A linear extension is obtained by forcing a relation, for example $\{1\} < \{2\}$. That is*

$$\emptyset < \{1\} < \{2\} < \{1, 2\}$$

is a linear extension of P.

Lemma 5.41. *Let $P = (X, \leq_P)$ and $Q = (X, \leq_Q)$ be posets on the same ground set. Define a relation $\leq$ on X as follows: $x \leq y$ if and only if $x \leq_P y$ and $x \leq_Q y$. Then $R = (X, \leq)$ is a poset, and both P and Q are extensions of R.*

Proof. We verify the three properties of a partial order.

- Reflexivity follows since $x \leq_P x$ and $x \leq_Q x$ for all $x \in X$.
- Transitivity follows since both $\leq_P$ and $\leq_Q$ are transitive.
- Antisymmetry follows since $x \leq y$ and $y \leq x$ imply that $x \leq_P y$, $y \leq_P x$ and $x \leq_Q y$, $y \leq_Q x$. Hence the antisymmetry of $\leq_P$ and $\leq_Q$ give that $x = y$.

Finally, every relation in $\leq$ also belongs to $\leq_P$ and $\leq_Q$, so both P and Q are extensions of R. □

5.2 Lattices

So far, we have studied posets, where not every pair of elements must be comparable. If we impose additional structure and require that every pair of elements has both a greatest lower bound and a least upper bound, we obtain a new object: lattices.

Lattices arise naturally throughout mathematics. For example:

- subsets ordered by inclusion form lattices via intersection and union,
- divisibility forms lattices via gcd and lcm,
- subspaces of a vector space form lattices via intersection and sum,
- partitions and Young diagrams also carry natural lattice structures.

Besides their order-theoretic interpretation, lattices can also be viewed algebraically through the two operations "meet" and "join", making them an important bridge between combinatorics, algebra, and geometry.

Definition 5.42. A poset $L = (X, \leq)$ is called a *lattice* if every pair of elements $x, y \in X$ has

1. a unique largest common lower bound, called the *meet* and denoted by $x \wedge y$,
2. a unique smallest common upper bound, called the *join* and denoted by $x \vee y$.

We may equivalently define a lattice as follows: for every $x, y, z \in X$, we have that

- $z \leq x$ and $z \leq y$ imply that $z \leq x \wedge y$,
- $z \geq x$ and $z \geq y$ imply that $z \geq x \vee y$.

We note that $x \wedge y = x$ if and only if $x \leq y$, and similarly $x \vee y = y$ if and only if $x \leq y$.

The meet and join operations are commutative and associative. Hence, for every finite subset $M = \{x_1, \ldots, x_n\} \subseteq X$ we may define

$$\bigwedge M = x_1 \wedge \cdots \wedge x_n$$

and

$$\bigvee M = x_1 \vee \cdots \vee x_n.$$

These are respectively the greatest lower bound and the least upper bound of all elements of M.

In particular, every finite lattice L has:

- a least element $\hat{0} = \bigwedge L$, obtained as the meet of all elements of L,
- a greatest element $\hat{1} = \bigvee L$, obtained as the join of all elements of L.

For convenience, we also define

$$\bigwedge \emptyset = \hat{1}, \qquad \bigvee \emptyset = \hat{0}.$$

Example 5.43. *The Boolean lattice is a lattice with*

$$A \wedge B = A \cap B, \qquad A \vee B = A \cup B.$$

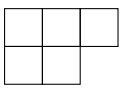
Exercise 5.44. *Show that the subspaces of a vector space form a lattice with*

$$U \wedge V = U \cap V, \qquad U \vee V = U + V.$$

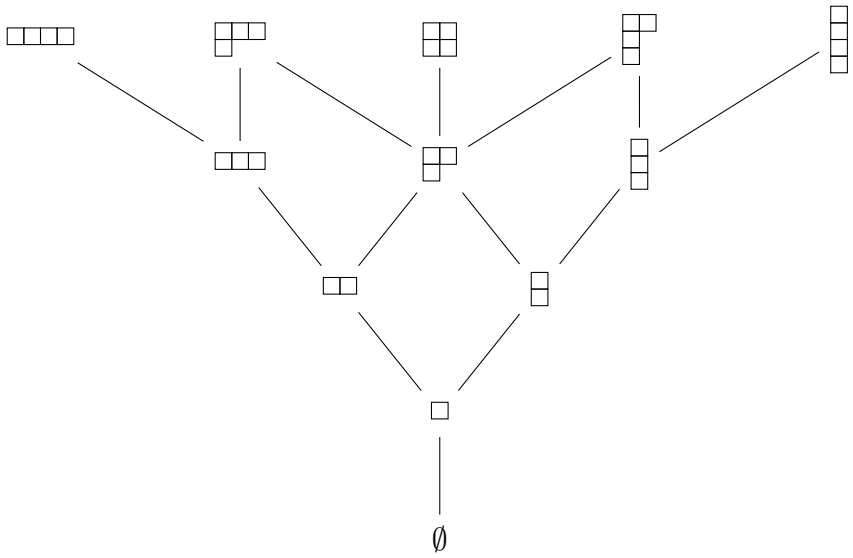
Example 5.45 (Partitions and Young's lattice)**.** *Recall that a partition $x = (x_1, \ldots, x_k)$ of n is such that $n = \sum_{i=1}^{k} x_i$ and $x_1 \geq x_2 \geq \cdots \geq x_k \geq 1$. We sometimes write $x \vdash n$ and call $n = |x|$ the size of the partition.*

We may partially order partitions x and y by setting $x \geq y$ if $x_i \geq y_i$ for all $i \geq 1$, where we extend partitions by zeros if necessary. This poset is called Young's lattice.

Recall that every partition can be represented by its Young diagram. For example, the partition $(3, 2)$ is represented by:



In terms of Young diagrams, the order relation simply means containment: $x \geq y$ if and only if the Young diagram of x contains that of y. Let us consider the Hasse diagram for the Young lattice for partitions up to $n = 4$:



The poset Y of all partitions is a ranked lattice with rank function

$$\mathrm{rk}(x) = |x| = n.$$

The meet and join are given componentwise by

$$(x \wedge y)_i = \min\{x_i, y_i\},$$

and

$$(x \vee y)_i = \max\{x_i, y_i\}.$$

Equivalently, in terms of Ferrers diagrams, we have that the meet is the intersection of diagrams, while the join is the union of diagrams.

Finally, saturated chains in Young's lattice correspond exactly to standard Young tableaux. For example, the tableau

1	2	4
3	5	

corresponds to the saturated chain

$$\emptyset \lessdot (1) \lessdot (2) \lessdot (2,1) \lessdot (3,1) \lessdot (3,2).$$

The correspondence works as follows: starting from the empty partition, we add the boxes in the order indicated by the entries $1, 2, \ldots, n$. After adding the first k boxes, we obtain the k-th partition in the chain.

We may also be interested in the elements directly above the least element $\hat{0}$, or directly below the greatest element $\hat{1}$:

Definition 5.46. Suppose a poset $P = (X, \leq)$ has a least element $\hat{0}$ and a greatest element $\hat{1}$. An *atom* is an element that covers $\hat{0}$. Similarly, a *coatom* is an element covered by $\hat{1}$.

Example 5.47. *In the Boolean lattice B_n, we have that the atoms are the singleton subsets, while the coatoms are the subsets of size $n-1$.*

Definition 5.48. A finite lattice $L = (X, \leq)$ with least element $\hat{0}$ is called *atomistic* if every element of X can be written as a join of atoms.

Proposition 5.49. *Let $L = (X, \leq)$ be a finite atomistic lattice. Then every element $x \in X$ is the join of the atoms below it:*

$$x = \bigvee \{a \in X \mid a \text{ is an atom and } a \leq x\}.$$

Proof. Since L is atomistic, we have that for each $x \in X$ there exist atoms $a_1, \ldots, a_r$ such that $x = a_1 \vee \cdots \vee a_r$.

Since each a_i appears in a join equal to x, we have $a_i \leq x$ for all i. Thus

$$x = a_1 \vee \cdots \vee a_r \leq \bigvee \{a \in L \mid a \text{ is an atom and } a \leq x\}.$$

On the other hand, every atom in the set on the right is by definition below x, so their join is also below x:

$$\bigvee \{a \in L \mid a \text{ is an atom and } a \leq x\} \leq x.$$

Hence equality holds. □

There also exists the dual notion of coatomistic:

Definition 5.50. A finite lattice L with greatest element $\hat{1}$ is called *coatomistic* if every element of L can be written as a meet of coatoms.

Proposition 5.51. *Let L be a finite coatomistic lattice. Then every element $x \in L$ is the meet of the coatoms above it:*

$$x = \bigwedge \{c \in L \mid c \text{ is a coatom and } x \leq c\}.$$

Exercise 5.52. *Prove Proposition 5.51.*

As we can talk of smaller and larger elements, we may also talk of elements in between two elements:

Definition 5.53. Let $P = (X, \leq)$ be a poset and let $x, y \in X$ with $x \leq y$. The *interval* between x and y is

$$[x, y] = \{z \in X \mid x \leq z \leq y\}.$$

Example 5.54. *Consider the Boolean lattice B_3. The interval $[\{1\}, \{1,2,3\}]$ consists of all subsets lying between these two sets, namely*

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

Proposition 5.55. *Let $L = (X, \leq)$ be a lattice and let $a, b \in X$. Then $[a, b]$ is a sublattice of L.*

Proof. Let $x, y \in [a, b]$. Then $a \leq x, y \leq b$. We show that both

$$x \wedge y \qquad \text{and} \qquad x \vee y$$

again lie in $[a, b]$.

Since

$$a \leq x \qquad \text{and} \qquad a \leq y,$$

the element a is a common lower bound of x and y. Hence

$$a \leq x \wedge y.$$

Similarly, since

$$x \leq b \qquad \text{and} \qquad y \leq b,$$

the element b is a common upper bound of x and y. Hence

$$x \vee y \leq b.$$

Moreover, by definition of meet and join,

$$x \wedge y \leq x, y \qquad \text{and} \qquad x, y \leq x \vee y.$$

Combining these inequalities gives

$$a \leq x \wedge y \leq b$$

and

$$a \leq x \vee y \leq b.$$

Therefore

$$x \wedge y,\ x \vee y \in [a, b],$$

and hence we get that the interval $[a, b]$ is closed under meet and join, and hence is a sublattice of L. □

A more algebraic (but equivalent) way of thinking of lattices is the following:

Proposition 5.56. *Let X be a set with two binary operations $\wedge$ and $\vee$ and two distinguished elements $0, 1 \in X$. Then $(X, \wedge, \vee, 0, 1)$ is a lattice if and only if the following identities hold for all $x, y, z \in X$:*

1. *associativity:*

$$x \wedge (y \wedge z) = (x \wedge y) \wedge z, \qquad x \vee (y \vee z) = (x \vee y) \vee z;$$

2. *commutativity:*
$$x \wedge y = y \wedge x, \qquad x \vee y = y \vee x;$$

3. *idempotence:*
$$x \wedge x = x, \qquad x \vee x = x;$$

4. *absorption:*
$$(x \vee y) \wedge x = x, \qquad (x \wedge y) \vee x = x;$$

5. *bounds:*
$$x \wedge \hat{0} = \hat{0}, \qquad x \vee \hat{1} = \hat{1}.$$

Proof. One direction follows directly from the definition of meet and join in a bounded lattice.

Conversely, assume that the identities hold. Define a relation on X by $x \leq y$ if and only if $x \wedge y = x$. Using the identities above, one checks that this relation is reflexive, antisymmetric and transitive, hence a partial order.

Moreover, $x \wedge y$ is the greatest lower bound of x and y, and $x \vee y$ is the least upper bound of x and y. Thus X is a lattice. Finally, the identities $x \wedge \hat{0} = \hat{0}$, and $x \vee \hat{1} = \hat{1}$ show that $\hat{0}$ is the least element and $\hat{1}$ is the greatest element. □

We will now consider two special lattices; distributive ones and modular ones.

Definition 5.57. A lattice L is called *distributive* if for all
$$x, y, z \in L$$
the following equivalent identities hold:
$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z),$$
and
$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z).$$

Moreover, in every lattice we always have
$$x \wedge (y \vee z) \geq (x \wedge y) \vee (x \wedge z),$$
and
$$x \vee (y \wedge z) \leq (x \vee y) \wedge (x \vee z).$$
Thus distributivity is precisely the condition that these inequalities are always equalities.

Example 5.58. *The Boolean lattice B_n is distributive.*

Indeed, meet and join are given by

$$A \wedge B = A \cap B, \qquad A \vee B = A \cup B,$$

and intersection and union distribute over each other:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

and

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Note that every sublattice of a distributive lattice is again distributive.

Example 5.59. *The Young's lattice is distributive, since the meet and join are given by intersection and union of Young diagrams. Hence Young's lattice can be viewed as a sublattice of a Boolean lattice, and distributivity follows.*

Exercise 5.60. *Let L be a lattice and let $x, y \in L$. Show that:*

$$x \vee (x \wedge y) = x$$

and

$$x \wedge (x \vee y) = x.$$

These identities are called the absorption laws.

Definition 5.61. A lattice $L = (X, \leq)$ is called *modular* if for all $x, y, z \in X$ with $x \leq z$, the following identity holds:

$$x \vee (y \wedge z) = (x \vee y) \wedge z.$$

This identity is called the *modular law*.

In every lattice we always have

$$x \vee (y \wedge z) \leq (x \vee y) \wedge z.$$

In fact, since $y \wedge z \leq y$, we obtain $x \vee (y \wedge z) \leq x \vee y$, and since $y \wedge z \leq z$ and $x \leq z$, we also have $x \vee (y \wedge z) \leq z$. Hence $x \vee (y \wedge z)$ is below both $x \vee y$ and z, which implies $x \vee (y \wedge z) \leq (x \vee y) \wedge z$.

Modularity says that equality always holds. Note that every sublattice of a modular lattice is again modular.

Proposition 5.62. *Every distributive lattice is modular.*

Proof. If $L = (X, \leq)$ is distributive and $x \leq z$, then $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$. Since $x \leq z$, we have $x \vee z = z$, and therefore

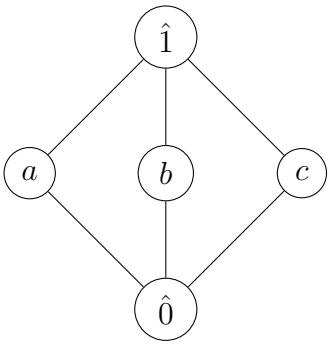
$$x \vee (y \wedge z) = (x \vee y) \wedge z.$$

Thus the modular law holds. □

Example 5.63. *The Boolean lattice B_n is modular, since it is distributive.*

The converse, however, is false: there exist modular lattices that are not distributive.

Example 5.64. *The following lattice is modular but not distributive.*



This lattice is called the diamond lattice *and is usually denoted by M_3. It is not distributive, as*

$$a \wedge (b \vee c) = a \wedge \hat{1} = a,$$

while

$$(a \wedge b) \vee (a \wedge c) = \hat{0} \vee \hat{0} = \hat{0}.$$

However, M_3 is modular.

We can generalize modularity, to semimodularity:

Definition 5.65. A lattice L is called *upper semimodular* if for all incomparable elements $x, y \in X$ the implication

$$x \wedge y \lessdot y \qquad \Longrightarrow \qquad x \lessdot x \vee y$$

holds.

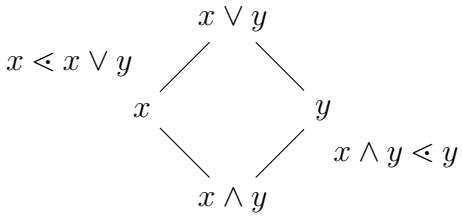
Similarly, L is called *lower semimodular* if

$$x \lessdot x \vee y \qquad \Longrightarrow \qquad x \wedge y \lessdot y.$$

Intuitively, upper semimodularity means:

If meeting with x lowers y by exactly one step, then joining with y raises x by exactly one step.

This can be visualized in the following interval:



Upper semimodularity and lower semimodularity are dual notions:

$$L \text{ is upper semimodular } \iff L^{op} \text{ is lower semimodular.}$$

Lemma 5.66. *Every modular lattice is upper and lower semimodular.*

Proof. Assume that L is modular and that $x \wedge y \lessdot y$. Then the interval $[x \wedge y, y]$ contains exactly two elements.

A basic property of modular lattices states that the intervals $[x \wedge y, y]$ and $[x, x \vee y]$ are isomorphic. Hence $[x, x \vee y]$ also contains exactly two elements, which means $x \lessdot x \vee y$. Therefore L is upper semimodular. A similar argument shows lower semimodularity. □

Definition 5.67. Let $P = (X, \leq)$ be a poset and let $a, b \in X$ with $a \leq b$. A chain

$$a = x_0 < x_1 < \cdots < x_k = b$$

is called a chain *between the bounds* a and b.

It is in addition called *maximal between* a *and* b if it cannot be refined, i.e., if every step is a cover relation:

$$x_{i-1} \lessdot x_i$$

for all $i \in \{1, \ldots, k\}$. Equivalently, it is a saturated chain in the interval $[a, b]$.

Proposition 5.68. *Let $P = (X, \leq)$ be a graded poset with rank function r and let $a \leq b \in X$. Then any two maximal chains between the same bounds have the same length. More precisely, every maximal chain from a to b has length*

$$r(b) - r(a).$$

Proof. Let

$$a = x_0 \lessdot x_1 \lessdot \cdots \lessdot x_k = b$$

be a maximal chain between a and b.

Since P is graded, every cover relation increases the rank by exactly 1. Hence

$$r(x_i) = r(x_{i-1}) + 1$$

for all $i \in \{1, \ldots, k\}$.

Therefore

$$r(b) = r(x_k) = r(x_0) + k = r(a) + k.$$

Thus

$$k = r(b) - r(a),$$

and hence the length of such a chain depends only on a and b. □

Corollary 5.69. *Let $L = (X, \leq)$ be a finite graded modular lattice with rank function r. Then for all $x, y \in X$,*

$$r(x \vee y) + r(x \wedge y) = r(x) + r(y).$$

Proof. In a modular lattice, the intervals

$$[x \wedge y, x] \qquad \text{and} \qquad [y, x \vee y]$$

are isomorphic. In fact, the isomorphism between the intervals $[x \wedge y, x]$ and $[y, x \vee y]$ is given by $z \mapsto z \vee y$, with the inverse map $w \mapsto w \wedge x$. Hence the two intervals have the same order structure, and in particular, they have maximal chains of the same length.

Since L is graded, the length of a maximal chain in an interval is the difference of ranks. Therefore

$$r(x) - r(x \wedge y) = r(x \vee y) - r(y).$$

Rearranging gives

$$r(x \vee y) + r(x \wedge y) = r(x) + r(y).$$

□

Summary In this chapter, we introduced partially ordered sets and lattices.

A *poset* is a set equipped with a reflexive, antisymmetric, and transitive relation. Unlike total orders, not every pair of elements must be comparable, leading to rich combinatorial structures such as chains, antichains, ranked posets, and Hasse diagrams.

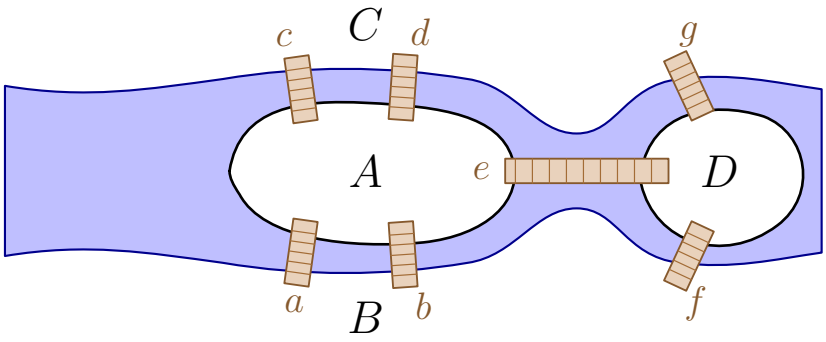
We studied important examples including Boolean lattices, Young's lattice, product posets, and ordinal sums, as well as fundamental results such as Dilworth's theorem, Mirsky's theorem, and Sperner's theorem.

By requiring that every pair of elements has a greatest lower bound (meet) and a least upper bound (join), we obtain a *lattice*. We furthermore introduced distributive, modular, and semimodular lattices.

6 Graphs

We now come to one of the most important topics in combinatorics: Graph theory! Graphs provide a language for studying relationships between objects and appear throughout mathematics, computer science, optimization, communication networks, and cryptography.

Historically, graph theory initiated with a famous puzzle known as the *Königsberg Bridge Problem*. The city of Königsberg (today Kaliningrad) was divided by a river into several land masses connected by seven bridges.

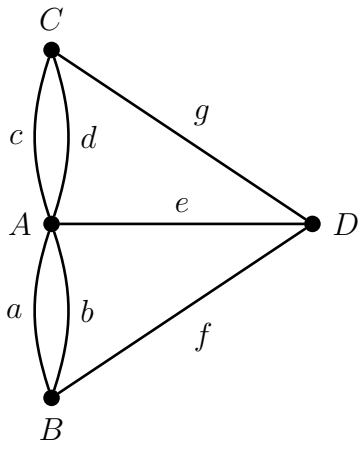


The seven bridges of Königsberg

The question was simple:

> *Is it possible to start at some location, cross each bridge exactly once, and return home?*

In 1736, Leonhard Euler observed that the precise geometric shape of the city is irrelevant. What matters is only which land masses are connected by bridges. He therefore replaced each land mass by a point and each bridge by a line connecting two points.



Euler's graph for the Königsberg bridges

This abstraction gave rise to the first graph-theoretic problem and is generally regarded as the birth of graph theory. Today, graphs are used to model a wide variety of situations:

- road, railway, and communication networks,
- social networks,
- scheduling and optimization problems,
- biological and epidemiological systems,
- error-correcting codes and cryptographic constructions.

In this chapter we introduce the basic concepts of graph theory, study important classes of graphs such as trees and bipartite graphs and investigate graph colorings and matchings.

6.1 Basic Definitions

We will start by giving some of the basic definitions, examples of graphs and some first properties.

Informally, a graph consists of points together with lines joining some of these points. The points are called *vertices* and the lines are called *edges*.

Definition 6.1. A *graph* is a pair $G = (V, E)$, where V is a non-empty finite set and

$$E \subseteq \{\{u, v\} \mid u, v \in V, u \neq v\}.$$

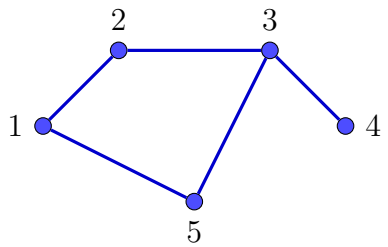
The elements of V are called *vertices*, and the elements of E are called *edges*.

To ease the notation, we will often write $\binom{V}{2}$ to denote the two-elements subsets of V, that is $E \subseteq \binom{V}{2}$ and while any labeling of vertices is fine, we will stick to $V = [n]$. However, when talking about a general vertex we will use the letter u or v.

We call $|V|$ the *order* of a graph and $|E|$ its *size*. If there are different graphs we may write $V(G)$ and $E(G)$ to specify which vertex, respectively edge set, we are considering.

We often write uv instead of $\{u, v\}$. Whenever vertices u and v are connected through an edge, i.e., $uv \in E$, we say that u and v are *adjacent*. In this case we sometimes write that $u \sim v$. If $uv \notin E$, then u and v are called *nonadjacent*. If v is a part of an edge, e.g. $e = uv \in E$, we say v is *incident* with the edge e.

Example 6.2. *Consider the graph $G = (V, E)$ drawn below.*



Here $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$.

We are often interested in vertices connected to some $v \in V$, these are its *neighbors:*

Definition 6.3. Let $G = (V, E)$ be a graph and let $v \in V$. The *neighborhood* of v is the set

$$N(v) = \{u \in V : uv \in E\}.$$

The vertices in $N(v)$ are called the *neighbors* of v.

Thus, we may also ask how many neighbors a vertex has:

Definition 6.4. Let $G = (V, E)$ be a graph and let $v \in V$. The *degree* of v, denoted by $\deg(v)$, is the number of edges incident with v.

Example 6.5. *In the graph above, i.e.,* $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$ *, we had that deg*$(1) = 2$ *and* $N(1) = \{2, 5\}$.

The way we defined the edge set E, we do not have multiple edges and we excluded the case $u = v$, i.e., an edge from u to itself. Such graphs are actually called *simple*.

In a simple graph, the degree of a vertex is equal to the number of its neighbors. Thus $\deg(v) = |N(v)|$.

One of the important tools in combinatorics is the method of counting the same set in two different ways. The following result is often the first theorem one proves in graph theory.

Theorem 6.6 (Handshaking Lemma)**.** *Let* $G = (V, E)$ *be a graph. Then*

$$\sum_{v \in V} \deg(v) = 2|E|.$$

Proof. Count the pairs (v, e) where v is a vertex, e is an edge, and v is incident with e.

On the one hand, every edge is incident with exactly two vertices, so the number of such pairs is $2|E|$. On the other hand, a fixed vertex v occurs in exactly $\deg(v)$ such pairs. Hence the number of such pairs is also

$$\sum_{v \in V} \deg(v).$$

Therefore

$$\sum_{v \in V} \deg(v) = 2|E|.$$

□

The Handshaking Lemma gets its name from the following interpretation. Consider a party where the vertices represent people and the edges represent handshakes. The degree of a vertex is then the number of handshakes made by that person. Counting all handshakes person by person yields $\sum_{v \in V} \deg(v)$. On the other hand, every handshake involves exactly two people, so counting handshakes directly yields $2|E|$. Therefore

$$\sum_{v \in V} \deg(v) = 2|E|.$$

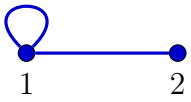
Corollary 6.7. *Every graph has an even number of vertices of odd degree.*

In fact, by the Handshaking Lemma, the sum of all degrees is even and as the sum of an odd number of odd integers would be odd, we get that the number of vertices of odd degree must be even.

6.2 Generalizations of Graphs

The definition of simple graphs is a bit restrictive so we may want to relax this a bit: if we have an edge from u to itself we call it a *loop*.

Example 6.8. *Let us consider the following graph with* $V = \{1, 2\}$ *and* $E = \{11, 12\}$.



Note that if loops are allowed, a loop contributes two to the degree of a vertex. Thus, in the example above the vertex 1 has degree 3.

Exercise 6.9. *Is the Handshaking Lemma still true for graphs with loops?*

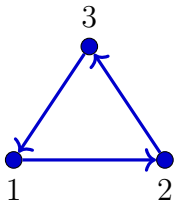
If we also allow multiple edges between the same vertices, we call them *parallel edges*. A graph containing parallel edges is often called a *multigraph*.

Example 6.10. *Let us consider the following multigraph with two parallel edges between* 1 *and* 2. *In this case we have* $V = \{1, 2\}$ *and* $E = \{12, 12\}$.

Until now, the order of an edge uv did not matter, however it may be useful to give an edge a direction:

Definition 6.11. A *directed graph*, or *digraph*, is a graph in which every edge has a direction. Thus an edge is represented by an ordered pair (u, v). The vertex u is called the *tail*, and v is called the *head*.

Example 6.12. *Let us consider* $V = \{1, 2, 3\}$ *and* $E = \{(1, 2), (2, 3), (3, 1)\}$.



Note that the edge $(1, 2)$ *is different from the edge* $(2, 1)$.

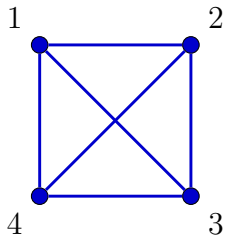
6.3 Some first Examples

We can now look at some examples of graphs. The smallest one, a graph with no edges, is called an *empty graph*.

Example 6.13. *The empty graph on four vertices is*

The opposite of an empty graph, would be a graph where all vertices are adjacent, i.e., $E = \binom{V}{2}$. These are called *complete graphs* and on n vertices, they are denoted by K_n.

Example 6.14. *The complete graph* K_4 *is*



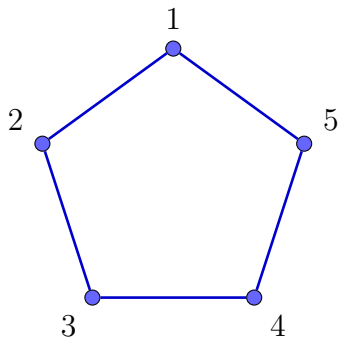
It has $\binom{4}{2} = 6$ *edges.*

Definition 6.15. For an integer $n \geq 3$, the *cycle graph* C_n is the graph with vertex set $V = [n]$ and edge set

$$E = \{12, 23, \ldots, (n-1)n, n1\}\,.$$

In other words, the vertices can be arranged in a circle so that every vertex is adjacent to exactly its two neighbors.

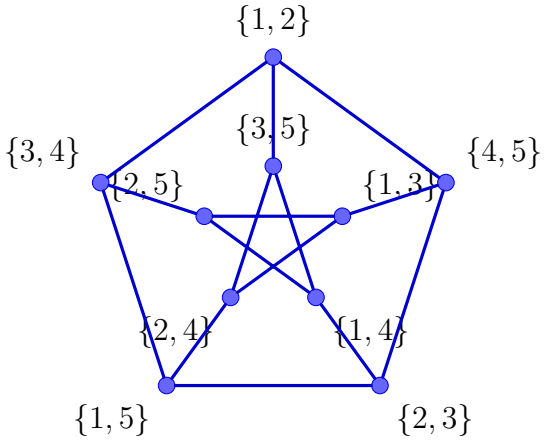
Example 6.16. *The cycle graph C_5 is shown below.*



Another important graph is called the *Petersen graph:*

Definition 6.17. The *Petersen graph* is the graph whose vertex set V consists of all the two-element subsets of $\{1, 2, 3, 4, 5\}$ and where two vertices are adjacent if and only if they are disjoint, i.e., $\{u, v\} \sim \{u', v'\}$ if $\{u, v\} \cap \{u', v'\} = \emptyset$.

Example 6.18. *The Petersen graph has $\binom{5}{2} = 10$ vertices. For instance, the vertex $\{1, 2\}$ is adjacent to the vertices $\{3, 4\}, \{3, 5\}, \{4, 5\}$. Thus every vertex has degree 3.*



The Petersen graph

By considering k-element subsets of $[n]$ we can also generalize this to the *Kneser graphs*.

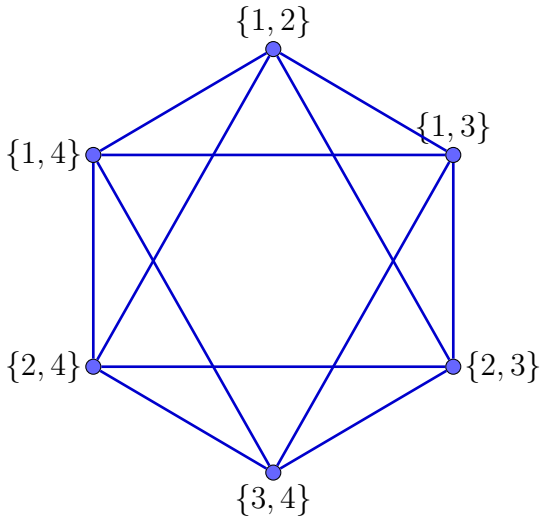
Definition 6.19. Let n and k be positive integers with $k \leq n$. The *Kneser graph* $KG(n, k)$ has vertex set V consisting of all k-element subsets of $\{1, \ldots, n\}$, where two vertices are adjacent if and only if they are disjoint.

Hence the Petersen graph is the Kneser graph $KG(5, 2)$.

We may also consider the graph connecting the subsets which are intersecting.

Definition 6.20. Let n and k be positive integers with $k \leq n$. The *Johnson graph* $J(n, k)$ is the graph whose vertex set V consists of all k-element subsets of $[n]$, where two vertices are adjacent if and only if the corresponding subsets intersect in exactly $k - 1$ elements.

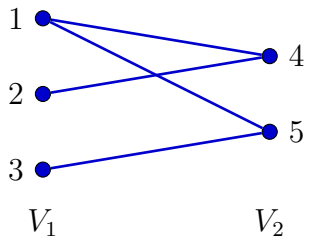
Example 6.21. *The graph* $J(4, 2)$ *has as vertices the* 2*-subsets of* $\{1, 2, 3, 4\}$.



We often encounter graphs, where we can separate the vertices into two sets, with no edges within one set.

Definition 6.22. A graph $G = (V, E)$ is called *bipartite* if its vertex set can be written as a disjoint union $V = V_1 \cup V_2$ such that every edge has one endpoint in V_1 and the other in V_2. Equivalently, no two vertices in the same part are adjacent.

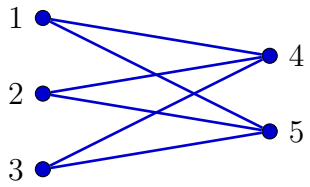
Example 6.23. *Let us consider the following bipartite graph with* $V_1 = \{1, 2, 3\}$ *and* $V_2 = \{4, 5\}$.



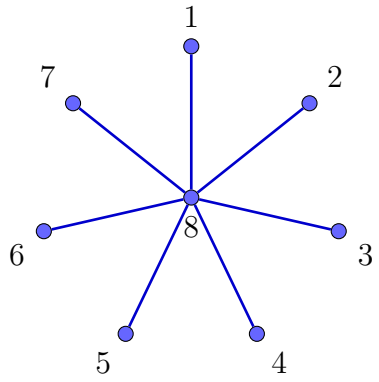
Within this class of graphs, we may again define completeness:

Definition 6.24. Let $m, n \geq 1$. The *complete bipartite graph* $K_{m,n}$ is the bipartite graph with parts V_1 and V_2 of sizes $|V_1| = m$, and $|V_2| = n$, in which every vertex of V_1 is adjacent to every vertex of V_2.

Example 6.25. *Let us consider the graph* $K_{3,2}$*, which is given by*



Example 6.26. *A special structure is the $K_{1,n}$ graph, called* star *graph. In fact, we can draw the $V_1 = \{n\}$ in the center to get a star form, e.g. $K_{1,8}$ can be drawn as*

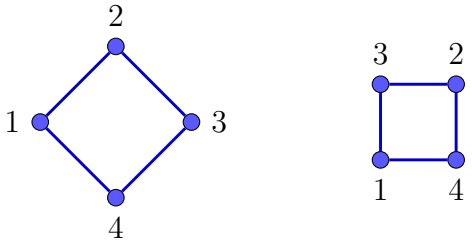


There are many ways to draw the same graph. What matters is not the geometry of the picture, but which vertices are joined by edges.

Definition 6.27. Let $G = (V, E)$ and $G' = (V', E')$ be graphs. An *isomorphism* from G to G' is a bijection $\varphi : V \to V'$ such that $uv \in E$ if and only if $\varphi(u)\varphi(v) \in E'$. If such a bijection exists, we say that G and G' are *isomorphic*.

As we may fix $V = [n]$, this asks for a permutation $\varphi \in S_n$, which respects the edges.

Example 6.28. *The following two graphs are isomorphic.*



We may then also ask whether a graph has such isomorphisms to itself.

Definition 6.29. An isomorphism from a graph to itself is called an *automorphism*. The set of all automorphisms of G is denoted by $\mathrm{Aut}(G)$.

Example 6.30. *Every permutation of the vertices of K_n is an automorphism. Hence*

$$\mathit{Aut}(K_n) = S_n.$$

Another important class of graphs, is those where all of the vertices have the same degree, we call these graphs *regular*.

Definition 6.31. A graph is called *k-regular* if every vertex has degree k.

We say G is *regular* if there exists some k such that G is k-regular.

Example 6.32. *The cycle C_5 is 2-regular.*

If the graph is not regular, it is useful to know the largest and smallest degree of any of its vertices.

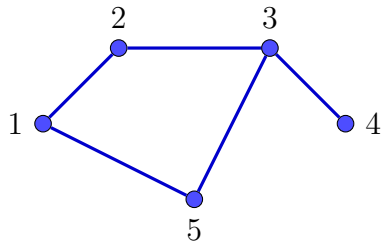
Definition 6.33. The *maximum degree* of G is

$$\Delta(G) = \max\{\deg(v) \mid v \in V\},$$

and the *minimum degree* of G is

$$\delta(G) = \min\{\deg(v) \mid v \in V\}.$$

Example 6.34. *For the graph with $V = \{1, 2, 3, 4, 5\}$ and $E = \{12, 23, 34, 15, 35\}$*



we have that $\Delta(G) = 3$ and $\delta(G) = 1$.

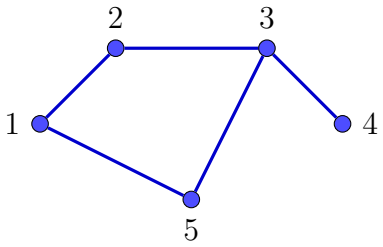
6.4 Matrix Representations

One of the most efficient ways to represent graphs is by matrices, especially if the graph is large. For this we have two main matrices; the *adjacency* matrix and the *incidence* matrix.

Definition 6.35. Let $G = (V, E)$ be a graph with $V = [n]$. The *adjacency matrix* of G is the matrix $A \in \{0, 1\}^{n \times n}$ with entries

$$a_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Example 6.36. *Let us consider again the graph with $V = \{1, 2, 3, 4, 5\}$ and $E = \{12, 23, 34, 15, 35\}$*



The adjacency matrix is

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

In the adjacency matrix each row and column are indexed by the vertices, and an entry 1 indicates that the vertices are adjacent. We could also index the columns by the edges and ask whether a vertex is incident with the edge.

Definition 6.37. Let $G = (V, E)$ be a graph and enumerate the edges as $E = \{e_1, \ldots, e_m\}$. The *incidence matrix* of G is the matrix $M \in \{0, 1\}^{n \times m}$ with the entries

$$m_{ij} = \begin{cases} 1 & \text{if } v_i \text{ is incident with } e_j, \\ 0 & \text{otherwise.} \end{cases}$$

Example 6.38. *Let us consider again the graph with* $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$. *We may order the edges as*

$$e_1 = 12, \qquad e_2 = 23, \qquad e_3 = 34, \qquad e_4 = 15, \qquad e_5 = 35.$$

Then the incidence matrix is

$$M = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Both of the matrices contain useful information about the graph.

Proposition 6.39. *Let G be a graph with adjacency matrix A. Then*

$$\deg(i) = \sum_{j=1}^{n} a_{ij}.$$

Exercise 6.40. *Prove Proposition 6.39 and use Proposition 6.39 to prove the Handshaking Lemma.*

Proposition 6.41. *Let $G = (V, E)$ be a graph with incidence matrix M Then the degree of a vertex v_i is equal to the sum of the entries in the i-th row:*

$$\deg(v_i) = \sum_{j=1}^{|E|} m_{ij}.$$

Exercise 6.42. *Prove Proposition 6.41 and prove the Handshaking Lemma using Proposition 6.41.*

The adjacency matrix is also useful to determine whether two graphs are isomorphic:

Proposition 6.43. *Let G and G' be graphs with vertex set $V = [n]$ and with adjacency matrices A and A', respectively. G and G' are isomorphic if and only if there exists a permutation matrix P such that*

$$A' = P^\top AP.$$

Proof. For the first direction: let $\sigma \in S_n$ be the isomorphism between G and G' and let P be the corresponding permutation matrix, that is,

$$P_{ij} = \begin{cases} 1, & j = \sigma(i), \\ 0, & \text{otherwise.} \end{cases}$$

The entry (i, j) of $P^\top AP$ is

$$(P^\top AP)_{ij} = A_{\sigma(i),\sigma(j)}.$$

Since σ is an isomorphism,

$$A_{\sigma(i),\sigma(j)} = 1$$

if and only if vertices i and j are adjacent in G'.

Therefore

$$(P^\top AP)_{ij} = A'_{ij}$$

for all i, j, and hence

$$A' = P^\top AP.$$

Conversely, assume that there exists a permutation matrix P such that $A' = P^\top AP$. Since P is a permutation matrix, it corresponds to a permutation $\sigma \in S_n$. We show that this permutation preserves adjacency. The (i, j)-entry of A' is

$$A'_{ij} = (P^\top AP)_{ij} = A_{\sigma(i),\sigma(j)}.$$

Therefore, $ij \in E'$ if and only if $A'_{ij} = 1$. By the equality above, this holds if and only if

$$A_{\sigma(i),\sigma(j)} = 1,$$

which is equivalent to

$$\sigma(i)\sigma(j) \in E.$$

Thus adjacency is preserved under σ. Hence G and G' are isomorphic.

□

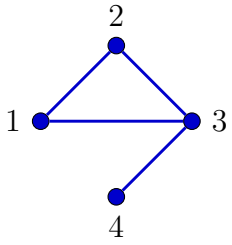
6.5 On Cliques and Friendship

We now move to some more properties of a graph, such as *cliques, independent sets* and *complements.* With these tools, we will show a nice paradox: why your friends have more friends than you.

Definition 6.44. A *clique* in a graph is a set of pairwise adjacent vertices.

While an *independent set* is a set of pairwise nonadjacent vertices.

Example 6.45. *Let us consider the graph G with $V = [4]$ and $E = \{12, 13, 23, 34\}$:*

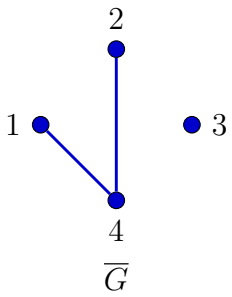


Here, the set $\{1, 2, 3\}$ is a clique, while $\{1, 4\}$ is an independent set.

We may also consider the *complement* of a graph, by exchanging adjacency with non-adjacency.

Definition 6.46. The *complement* of a graph $G = (V, E)$ is the graph $\overline{G} = \left(V, \binom{V}{2} \setminus E\right)$.

Example 6.47. *Let us consider again the graph G from above, i.e., $V = [4]$ and $E = \{12, 13, 23, 34\}$. The complement $\overline{G}$ is then given by $V = [4]$ and $\overline{E} = \{14, 24\}$:*



$\overline{G}$

In our example, we have one vertex, namely 3, which is not adjacent to any other vertex, we call such vertices *isolated*.

Exercise 6.48. *Show that $\overline{\overline{G}} = G$ for every graph G.*

Exercise 6.49. *Determine $\overline{K_n}$ and the complement of the Petersen graph.*

Note that a clique in G is an independent set in $\overline{G}$, and an independent set in G is a clique in $\overline{G}$.

We further note that every subset of an independent set is again an independent set, similarly every subset of a clique is again a clique.

Thus, we may be interested in the largest size of a clique or an independent set in our graph:

Definition 6.50. The *clique number* of a graph G, denoted by $\omega(G)$, is the maximum size of a clique in G.

The *independence number* of a graph G, denoted by $\alpha(G)$, is the maximum size of an independent set in G.

Example 6.51. *For the graph above, i.e., $V = [4]$ and $E = \{12, 13, 23, 34\}$, we have that $\omega(G) = 3$, since $\{1, 2, 3\}$ is a clique and no clique of size four exists. Moreover, $\alpha(G) = 2$, since $\{1, 4\}$ is an independent set and no independent set of size three exists.*

Also these new parameters are opposites of each other:

Theorem 6.52. *Let G be a graph. Then $\omega(G) = \alpha(\overline{G})$ and $\alpha(G) = \omega(\overline{G})$.*

Proof. A set of vertices is a clique in G if and only if every pair of vertices is adjacent in G. By definition of the complement, this is equivalent to saying that every pair of vertices is nonadjacent in $\overline{G}$, that is, the set is independent in $\overline{G}$. Therefore the cliques of G are precisely the independent sets of $\overline{G}$. Taking maximum cardinalities yields $\omega(G) = \alpha(\overline{G})$.

The second equality follows by exchanging the roles of G and $\overline{G}$. □

Exercise 6.53. *Determine the clique number and independence number of*

- K_n,
- *the empty graph on n vertices,*
- C_5.

Exercise 6.54. *Let G be a graph on n vertices. Show that $\omega(G) \cdot \alpha(G) \geq n$ does not hold in general by finding a counterexample.*

One might expect that, on average, our friends have as many friends as we do. Surprisingly, this is not true.

Theorem 6.55 (Friendship Paradox)**.** *Let G be a graph. Then the average degree of a neighbor is at least the average degree of a vertex. Equality holds if and only if the graph is regular.*

Interpreting the vertices as people and the edges as friendships, the theorem implies that, on average, the friends of a randomly chosen person have at least as many friends as a randomly chosen person.

The reason is that people with many friends appear in many friendship relations and are therefore more likely to be encountered when looking at a random friend.

Proof. The average degree of a vertex is

$$\frac{1}{|V|} \sum_{v \in V} \deg(v).$$

When we choose a random neighbor, vertices of large degree are more likely to be selected, because they appear as neighbors more often.

The average degree of a neighbor is

$$\frac{\sum_{v\in V}\deg(v)^2}{\sum_{v\in V}\deg(v)}.$$

By the Cauchy–Schwarz inequality,

$$\left(\sum_{v\in V}\deg(v)\right)^2 \le |V|\sum_{v\in V}\deg(v)^2.$$

Dividing by

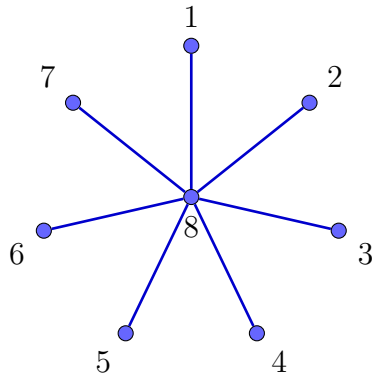
$$|V|\sum_{v\in V}\deg(v)$$

gives

$$\frac{\sum_{v\in V}\deg(v)}{|V|} \le \frac{\sum_{v\in V}\deg(v)^2}{\sum_{v\in V}\deg(v)}.$$

The left-hand side is the average degree of a vertex and the right-hand side is the average degree of a neighbor. □

Example 6.56. *Let us consider again the star graph* $K_{1,8}$



The central vertex, $v = 8$*, has degree* 7*, while each outer vertex has degree* 1*. The average degree is* $\frac{7+7\cdot 1}{8} = \frac{14}{8}$*. However, every outer vertex has a neighbor of degree* 7*. Thus the average degree of a neighbor is much larger than the average degree of a randomly chosen vertex.*

6.6 Inducing Graphs, Paths and Connectivity

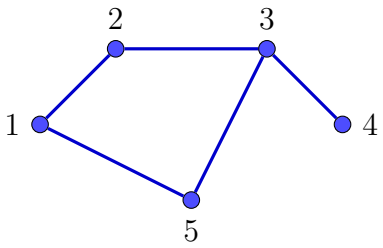
Of course, we can also consider only parts of a graph, getting a *subgraph*:

Definition 6.57. A graph H is a *subgraph* of a graph G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. In addition, we call a subgraph H of G *spanning* if $V(H) = V(G)$.

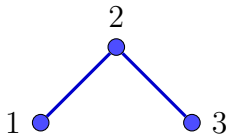
If we choose a smaller subset of the vertex set, say $S \subseteq V$ and keep all edges involving S, we get what we call an *induced* subgraph.

Definition 6.58. Let $S \subseteq V(G)$. The *induced subgraph* on S, denoted by $G[S]$, is the graph whose vertex set is S and whose edges are precisely the edges of G with both endpoints in S.

Example 6.59. *Let us consider again the graph with* $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$



Let $S = \{1, 2, 3\}$*, then* $G[S]$ *is*



We can also consider walking on our graphs, from one vertex to an adjacent one - this will give a *path*.

Definition 6.60. Let $G = (V, E)$ be a graph and $u, v \in V$. A *path* from u to v is a sequence of distinct vertices

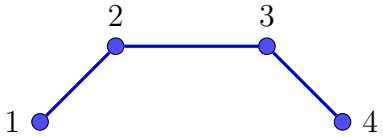
$$v_0, v_1, \ldots, v_k$$

such that $u = v_0, v = v_k$, and $v_i v_{i+1} \in E$ for all $0 \leq i \leq k - 1$. The *length* of the path is k.

Example 6.61. *For the graph above, with* $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$,

$$1, 2, 3, 4$$

is a path from 1 *to* 4 *of length* 3.



If we can start from any vertex and walk to any other vertex, we call this graph *connected.*

Definition 6.62. A graph is called *connected* if it has just one vertex, or if every pair of distinct vertices can be joined by a path.

Example 6.63. *The graph from above, with* $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$*, is connected. For instance, a path from* 4 *to* 1 *is*

$$4, 3, 5, 1.$$

Exercise 6.64. *If a graph* G *has no isolated vertex, is it connected?*

In other graphs, which are not connected, we may consider connected subgraphs. Connected subgraphs to which we cannot add any other vertex without loosing the connected property, i.e., maximal connected subgraphs, are called *connected components*.

Example 6.65. *Let us consider the following graph* G *with* $V = [4]$ *and* $E = \{12, 34\}$*, then* G *has two connected components, namely* $S = \{1, 2\}$ *and* $T = \{3, 4\}$.

Proposition 6.66. *A graph is connected if and only if it has exactly one connected component.*

Exercise 6.67. *Prove Proposition 6.66.*

Proposition 6.68. *Let* $G = (V, E)$ *be a graph. Then every vertex belongs to exactly one connected component. In particular, the connected components form a partition of* V.

Proof. Define a relation on V by $u \sim v$ if there exists a path from u to v.

This relation is reflexive, since every vertex is connected to itself by a path of length 0.

It is symmetric, because every path from u to v can be traversed in the opposite direction.

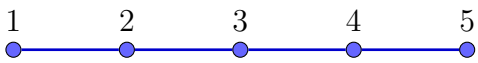
It is transitive, because if there is a path from u to v and a path from v to w, then concatenating them yields a path from u to w.

Hence $\sim$ is an equivalence relation and its equivalence classes are precisely the connected components of G. Since equivalence classes form a partition, the connected components form a partition of V. □

Definition 6.69. The *path graph* P_n is the graph with vertex set $V = [n]$ and edge set

$$E = \{12, 23, \ldots, (n-1)n\}.$$

Example 6.70. *The path graph* P_5 *is given by*



Clearly, P_n is connected.

There is a threshold on the smallest degree, after which we know that the graph must be connected.

Proposition 6.71. *Let G be a graph on n vertices. If*

$$\delta(G) \geq \frac{n-1}{2},$$

then G is connected.

Proof. Assume that G is not connected and let C be one of its connected components. Let $|C| = m$, then every vertex of C has degree at most $m-1$. Hence $\delta(G) \leq m-1$. Since we assumed that $\delta(G) \geq \frac{n-1}{2}$, we obtain

$$m \geq \frac{n+1}{2}.$$

Thus every connected component contains at least $\frac{n+1}{2}$ vertices. But a graph which is not connected has at least two connected components, which would together contain more than n vertices. This is impossible. Hence G is connected. □

If we know we have a connected graph, then we also get a bound on the number of edges.

Proposition 6.72. *Let G be a connected graph on n vertices. Then*

$$|E| \geq n-1.$$

Proof. We prove the statement by induction on n. For $n = 1$, the statement is clear.

Assume $n \geq 2$ and let v be a vertex of a connected graph G. Removing v can split the graph into at most $\deg(v)$ connected components. Since each of these components is connected, the induction hypothesis implies that together they contain at least $(n-1)-c$ edges, where c denotes the number of components.

Reinserting the vertex v contributes at least c edges, since each component must be connected to v.

Therefore $|E| \geq (n-1) - c + c = n-1$. □

6.7 Distance, Diameter and Radius

As we have a notion of "walking" on a graph, we can consider how long it will take us to go from vertex u to vertex v.

Definition 6.73. The *distance* between two vertices u and v, denoted by $d(u,v)$, is the length of a shortest path from u to v. If no such path exists, we set $d(u,v) = \infty$.

Example 6.74. *In the graph from above, with $V = \{1,2,3,4,5\}$ and $E = \{12, 23, 34, 15, 35\}$*

$$d(1,4) = 3.$$

Proposition 6.75. *Let G be a connected graph. Then the distance function satisfies*

- $d(u,v) \geq 0$, *and* $d(u,v) = 0$ *if and only if* $u = v$,
- $d(u,v) = d(v,u)$,
- $d(u,w) \leq d(u,v) + d(v,w)$.

Thus the distance is a metric on the vertex set.

Exercise 6.76. *Prove Proposition 6.75.*

Definition 6.77. The *eccentricity* of a vertex v is

$$\mathrm{ecc}(v) = \max\{d(u,v) \mid u \in V\}.$$

Having defined a distance on our graphs, we may ask for the largest and the smallest distance between two vertices:

Definition 6.78. Let G be a connected graph. The *diameter* of G is

$$\mathrm{diam}(G) = \max\{\mathrm{ecc}(v) \mid v \in V\}.$$

The *radius* of G is

$$\mathrm{rad}(G) = \min\{\mathrm{ecc}(v) \mid v \in V\}.$$

Since $\mathrm{ecc}(v) = \max\{d(u,v) \mid u \in V\}$, we also have

$$\mathrm{diam}(G) = \max\{d(u,v) \mid u, v \in V\}.$$

Example 6.79. *Consider the graph $V = [5], E = \{12, 23, 34, 15, 35\}$. The distances between the vertices are given by*

d	1	2	3	4	5
1	0	1	2	3	1
2	1	0	1	2	2
3	2	1	0	1	1
4	3	2	1	0	2
5	1	2	1	2	0

The eccentricity of a vertex is the largest entry in the corresponding row:

$$\mathrm{ecc}(1) = 3, \qquad \mathrm{ecc}(2) = 2, \qquad \mathrm{ecc}(3) = 2, \qquad \mathrm{ecc}(4) = 3, \qquad \mathrm{ecc}(5) = 2.$$

Therefore $\mathrm{rad}(G) = 2$ *and* $\mathrm{diam}(G) = 3$.

Exercise 6.80. *Determine the radius and diameter of the path graph P_n.*

As we have previously defined the largest and smallest degree in a graph, we may ask if these are related to the radius and diameter;

Proposition 6.81. *Let G be a connected graph on n vertices. Then*

$$\operatorname{diam}(G) \leq n - \delta(G).$$

Proof. Let

$$v_0, v_1, \ldots, v_d$$

be a shortest path of maximum possible length, so that $d = \operatorname{diam}(G)$. In particular, $d(v_0, v_d) = d$.

Since the path is shortest, the vertex v_0 cannot be adjacent to any vertex v_i with $i \geq 2$. Indeed, if $v_0 v_i$ were an edge for some $i \geq 2$, then we could reach v_i from v_0 in one step, contradicting the fact that $d(v_0, v_i) = i$.

Thus, among the vertices of the path, the only neighbor of v_0 is v_1. Now every neighbor of v_0 must be distinct from v_0 and must lie among the remaining $n - 1$ vertices. Moreover, none of the vertices $v_2, \ldots, v_d$ is a neighbor of v_0.

There are $d - 1$ such forbidden vertices. Hence the number of possible neighbors of v_0 is at most $(n - 1) - (d - 1) = n - d$. Therefore $\deg(v_0) \leq n - d$.

Since $\delta(G) \leq \deg(v_0)$, we obtain

$$\delta(G) \leq n - d$$

and rearranging gives

$$d \leq n - \delta(G).$$

□

Proposition 6.82. *Let G be a connected graph on n vertices. If $\Delta(G) = n - 1$, then*

$$\operatorname{diam}(G) \leq 2.$$

Proof. Since $\Delta(G) = n - 1$, there exists a vertex v adjacent to every other vertex of G.

Now let $x, y \in V$ be arbitrary. If x and y are adjacent, then $d(x, y) = 1$. If they are not adjacent, then both are adjacent to v, so x, v, y is a path of length 2. Hence $d(x, y) \leq 2$. Therefore

$$\operatorname{diam}(G) \leq 2.$$

□

Corollary 6.83. *If $\delta(G) = n - 1$, then $\operatorname{diam}(G) = 1$. Equivalently, $G = K_n$.*

Clearly, also the radius and the diameter are connected:

Proposition 6.84. *For every connected graph G, we have*

$$\operatorname{rad}(G) \leq \operatorname{diam}(G) \leq 2 \operatorname{rad}(G).$$

Proof. Since $\mathrm{rad}(G) = \min\{\mathrm{ecc}(v) \mid v \in V\}$ and $\mathrm{diam}(G) = \max\{\mathrm{ecc}(v) \mid v \in V\}$, it follows immediately that

$$\mathrm{rad}(G) \leq \mathrm{diam}(G).$$

For the second inequality, let c be a center of the graph, that is, $\mathrm{ecc}(c) = \mathrm{rad}(G)$.

Let u, v be arbitrary vertices. By the triangle inequality,

$$d(u, v) \leq d(u, c) + d(c, v)$$

and since c has eccentricity $\mathrm{rad}(G)$, we have $d(u, c) \leq \mathrm{rad}(G)$ and $d(c, v) \leq \mathrm{rad}(G)$. Hence $d(u, v) \leq 2\ \mathrm{rad}(G)$. Since this holds for every pair of vertices u, v, we obtain

$$\mathrm{diam}(G) = \max\{d(u, v) \mid u, v \in V\} \leq 2\ \mathrm{rad}(G).$$

□

Just like we have cycle graphs, we can also have cycles inside a graph:

Definition 6.85. A *cycle* in a graph is a sequence of distinct vertices $v_1, v_2, \ldots, v_k$ with $k \geq 3$ such that

$$v_1v_2, v_2v_3, \ldots, v_{k-1}v_k, v_kv_1$$

are edges of the graph.

Example 6.86. *Consider again the graph with* $V = [5]$ *and* $E = \{12, 23, 34, 15, 35\}$. *Then* $1, 2, 3, 5$ *is a cycle of length* 4.

Proposition 6.87. *Every cycle has the same number of vertices and edges.*

Exercise 6.88. *Prove Proposition 6.87.*

Definition 6.89. A graph is called *acyclic* if it contains no cycle.

In fact, such graphs are of large interest, and will be the next chapter: *trees*.

6.8 Eulerian Graphs and Trees

We now return to the question that motivated the beginning of this chapter: can one walk through the city of Königsberg and cross every bridge exactly once?

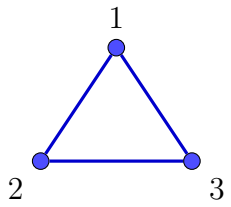
We first need some more definitions: while we have had paths, the Königsberg puzzle is really about trails:

Definition 6.90. A *trail* in a graph is a walk in which no edge is repeated.

An *Eulerian trail* is a trail that uses every edge of the graph exactly once.

While this definition seems to be similar to paths, they are different:

Example 6.91. *Consider the graph G with $V = [3]$ and $E = \{12, 23, 13\}$.*



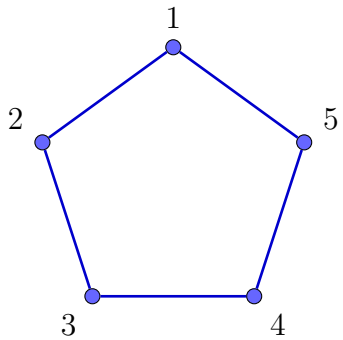
The sequence $1, 2, 3, 1$ is a trail, because no edge is repeated. However, it is not a path, because the vertex 1 appears twice.

The second question in the Königsberg puzzle was whether such trail can also return to the starting point, this would later be called an *Eulerian circuit*:

Definition 6.92. An *Eulerian circuit* is an Eulerian trail that starts and ends at the same vertex.

Finally, a graph is called *Eulerian* if it has an Eulerian circuit.

Example 6.93. *The cycle graph C_n is Eulerian. Indeed, walking once around the cycle uses every edge exactly once and returns to the starting vertex.*



Theorem 6.94 (Euler)**.** *Let G be a connected graph. Then G has an Eulerian circuit if and only if every vertex of G has even degree. Moreover, G has an Eulerian trail which is not closed if and only if exactly two vertices of G have odd degree. In that case, the Eulerian trail starts at one of these odd-degree vertices and ends at the other.*

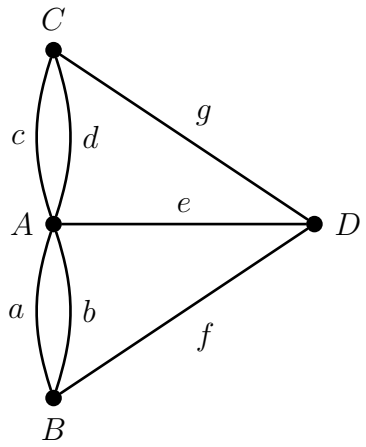
Idea of the proof. Suppose first that G has an Eulerian circuit. Whenever the circuit enters a vertex, it must also leave it again. Thus the incident edges at every vertex can be paired into an incoming edge and an outgoing edge. Hence every vertex has even degree.

Similarly, if G has an Eulerian trail that starts at s and ends at t, then every vertex different from s and t is entered and left the same number of times. Hence all vertices except possibly s and t have even degree. The vertices s and t have odd degree.

Conversely, assume that all vertices have even degree. Starting at any vertex, follow unused edges for as long as possible. Since every vertex has even degree, one cannot get stuck at a vertex different from the starting vertex. Hence this produces a closed trail. If not all edges have been used, one can start another closed trail at a vertex of the first trail that is incident to an unused edge, and splice the two closed trails together. Repeating this process gives an Eulerian circuit.

If exactly two vertices have odd degree, say s and t, then adding a new edge between s and t makes all degrees even. The resulting graph has an Eulerian circuit. Removing the new edge from this circuit gives an Eulerian trail from s to t. □

Example 6.95 (Königsberg revisited)**.** *The graph associated with the seven bridges of Königsberg has four vertices, corresponding to the four land masses. The seven bridges are represented by seven edges.*



Euler's graph for the Königsberg bridges

The degrees of the four vertices are $5, 3, 3, 3$. *Thus all four vertices have odd degree. Since a graph has an Eulerian trail only if it has either* 0 *or* 2 *vertices of odd degree, the Königsberg graph has no Eulerian trail.*

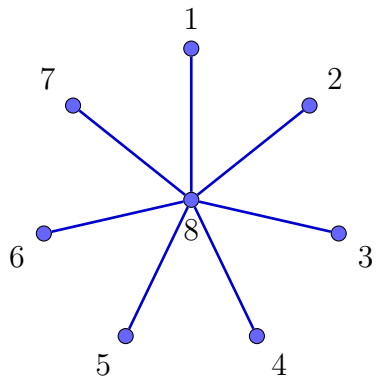
Therefore it is impossible to cross every bridge exactly once.

We next study a special class of graphs that contains no cycles, i.e., an acyclic graph or a *forest*.

Definition 6.96. A *forest* is a graph that contains no cycle. A *tree* is a connected forest.

An equivalent definition is to say that a tree is a connected graph without cycles.

Example 6.97. *The path graph P_n is a tree. The star graph $K_{1,n}$ is also a tree.*



Staying with the biological analogy, we can then also study leafs of our trees:

Definition 6.98. A vertex of degree 1 in a tree is called a *leaf*.

In the star graph $K_{1,8}$ above, all vertices, except for the vertex 8, are leaves.

Lemma 6.99. *Every tree with at least two vertices has at least two leaves.*

Proof. Let $v_1, v_2, \ldots, v_k$ be a longest path in the tree. We claim that v_1 and v_k are leaves.

Suppose that v_1 is not a leaf. Then v_1 has a neighbor different from v_2. If this neighbor were not already on the path, then we could extend the path, contradicting maximality. If the neighbor were already on the path, then we would obtain a cycle, contradicting that the graph is a tree. Thus v_1 has degree 1. Similarly, v_k has degree 1. Hence the tree has at least two leaves. □

Lemma 6.100. *Deleting a leaf from a tree with n vertices gives a tree with $n-1$ vertices.*

Proof. Let T be a tree and let u be a leaf. Since u has degree 1, it cannot lie on a path between two other vertices of T. Therefore removing u does not disconnect the remaining graph.

Also, removing vertices and edges cannot create a cycle. Hence the resulting graph is connected and acyclic, and therefore it is a tree. □

Recall that if G is connected, then $|E| \geq n-1$, for trees, we know this must be exactly $n-1$. In fact, we get several characterizations:

Theorem 6.101. *Let G be a graph on n vertices. The following statements are equivalent.*

1. *G is a tree.*
2. *G is connected and has $n-1$ edges.*

3. *G has $n-1$ edges and no cycles.*

4. *Between any two vertices of G, there is a unique path.*

Proof. We prove the implications $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1)$ and then show that (1) is equivalent to (4).

$(1) \Rightarrow (2)$: Assume that G is a tree, i.e., G is connected and has no cycles. We prove that G has $n-1$ edges by induction on n.

For $n = 1$, the graph has no edges, so the statement is true. Now let $n \geq 2$. By the previous lemma, G has a leaf u. Removing u gives a tree G' on $n-1$ vertices. By induction, G' has $n-2$ edges. Since u was incident to exactly one edge, the graph G has $n-2+1 = n-1$ edges.

$(2) \Rightarrow (3)$: Assume that G is connected and has $n-1$ edges. We show that G has no cycles.

Suppose, by contradiction, that G contains a cycle. Removing one edge from a cycle does not disconnect the graph. Hence we can remove an edge and obtain a connected graph on n vertices with $n-2$ edges.

But every connected graph on n vertices has at least $n-1$ edges, a contradiction. Thus G has no cycles.

$(3) \Rightarrow (1)$: Assume that G has $n-1$ edges and no cycles. Let $G_1, \ldots, G_k$ be the connected components of G, and let n_i be the number of vertices of G_i. Since G has no cycles, each component G_i is a tree. By the implication $(1) \Rightarrow (2)$, the component G_i has $n_i - 1$ edges.

Therefore the total number of edges of G is

$$\sum_{i=1}^{k}(n_i - 1) = \left(\sum_{i=1}^{k} n_i\right) - k = n - k.$$

But by assumption G has $n-1$ edges. Hence $n-k = n-1$, so $k = 1$. Therefore G is connected.

$(1) \Rightarrow (4)$: Assume that G is connected and has no cycles. Since G is connected, there exists at least one path between any two vertices.

Suppose that there are two distinct paths between vertices u and v. Following one path from u to v and returning along the other path produces a cycle. This contradicts the assumption that G has no cycles. Thus the path between any two vertices is unique.

$(4) \Rightarrow (1)$: Assume that between any two vertices there is a unique path. Then in particular, G is connected.

If G had a cycle, then two vertices on this cycle could be connected by going around the cycle in two different directions. This would give two distinct paths between them, contradicting uniqueness. Hence G has no cycles. □

We also know how many trees there are.

Theorem 6.102 (Cayley's Formula)**.** *The number of labeled trees on the vertex set* $[n]$ *is* n^{n-2}.

We omit the proof here. But, note that in Cayley's formula, the vertices are assumed to be labeled by $[n]$. Thus two trees are considered different whenever their edge sets differ, even if they are isomorphic as graphs.

Exercise 6.103. *List all labeled trees on the vertex set* $[3]$ *Verify Cayley's formula for* $n = 3$*. If the trees are unlabeled, then how many trees with* $n = 3$ *do we have?*

6.9 Bipartite Graphs, Matchings, and Graph Coloring

We now return to bipartite graphs and study two important types of questions: matching vertices and coloring vertices.

First, we need some results about cycles in a graph. Note that we say a cycle is odd, if it has an odd number of vertices. A walk is closed, if it returns to its starting vertex; that is however not the same as being a cycle, where we need the sequence to consist of distinct vertices.

Lemma 6.104. *Every closed odd walk contains an odd cycle.*

Proof. We prove the statement by induction on the length of the closed walk.

The smallest odd number of vertices in a closed walk is 3 and here we already have a cycle. Now assume that the statement holds for all closed odd walks of length less than k, and let W be a closed odd walk of length k.

If W has no repeated internal vertices, then W is itself an odd cycle. Otherwise, some internal vertex is repeated. Then W can be split into two shorter closed walks. Since the total length of W is odd, at least one of these two closed walks has odd length. By the induction hypothesis, this shorter closed odd walk contains an odd cycle.

Therefore W contains an odd cycle. □

Such odd cycles are very useful, in fact, they characterize whether a graph is bipartite or not:

Theorem 6.105. *A graph is bipartite if and only if it contains no odd cycle.*

Proof. Suppose first that G is bipartite with $V = V_1 \cup V_2$. Every walk in G alternates between V_1 and V_2. Hence every closed walk, and in particular every cycle, has even length. Thus G contains no odd cycle.

Conversely, assume that G contains no odd cycle. It suffices to consider one connected component at a time. Let C be a connected component and choose a vertex $u \in C$.

Define

$$A = \{v \in C \mid d(u, v) \text{ is even}\}$$

and

$$B = \{v \in C \mid d(u, v) \text{ is odd}\}.$$

Then $C = A \cup B$ and the union is disjoint.

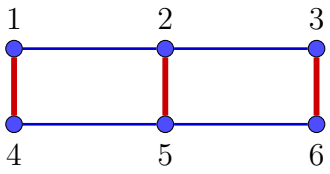
We show that both A and B are independent sets. Suppose first that $v, v' \in A$ are adjacent. Let P be a shortest path from u to v, and let P' be a shortest path from u to v'. Both paths have even length. Then following P, then the edge vv', and then following P' backwards gives a closed walk of odd length. By the previous lemma, this closed odd walk contains an odd cycle, a contradiction.

Thus no two vertices in A are adjacent. The same argument applies to B. Therefore C is bipartite. Since this can be done for every connected component, the graph G is bipartite. □

Definition 6.106. A *matching* in a graph $G = (V, E)$ is a set $M \subseteq E$ such that no two edges in M share an endpoint. A vertex is called *covered* by M if it is incident to an edge in M.

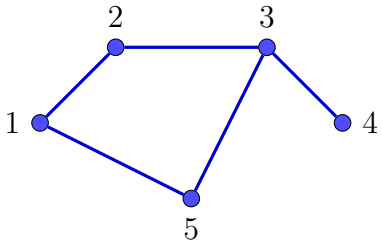
A matching is called *perfect* if every vertex is covered.

Example 6.107. *Consider the following graph with* $V = [6]$ *and* $E = \{12, 14, 23, 25, 36, 45, 56\}$:



The set $M = \{14, 25, 36\}$ *in red is a matching, since no two edges in* M *share an endpoint. In fact, this matching is perfect, because every vertex is covered by one of the edges in* M.

Exercise 6.108. *Let us consider again the graph with* $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$

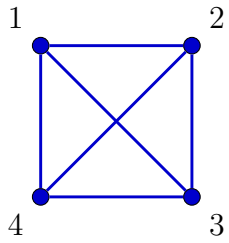


Show that the graph does not have a perfect matching.

Definition 6.109. A matching is called *maximal* if it cannot be enlarged by adding another edge. A matching is called *maximum* if it has largest possible cardinality among all matchings of the graph. The size of a maximum matching in G is denoted by $\nu(G)$.

Example 6.110. *For the complete graph* K_n*, we have*

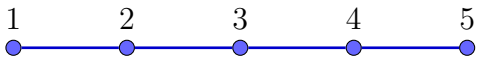
$$\nu(K_n) = \left\lfloor \frac{n}{2} \right\rfloor .$$



We can easily check in K_4, that the largest matching we can choose has size two, e.g. $M = \{12, 34\}$.

Example 6.111. *For the path graph P_n and the cycle graph C_n, we also have*

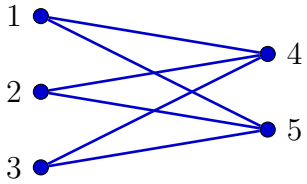
$$\nu(P_n) = \nu(C_n) = \left\lfloor \frac{n}{2} \right\rfloor .$$



In P_5 for example, we can choose $M = \{12, 34\}$ but no other edge.

Example 6.112. *For the complete bipartite graph $K_{m,n}$, we have*

$$\nu(K_{m,n}) = \min\{m, n\}.$$



In $K_{3,2}$ the largest matching is of size 2, for example $M = \{15, 24\}$.

Instead of considering the neighborhood of a single vertex, we can also consider the neighborhood of a set of vertices, that is for $S \subseteq V$. The *neighborhood* of S is

$$N(S) = \{v \in V \mid v \text{ is adjacent to at least one vertex of } S\}.$$

Theorem 6.113 (Hall's Marriage Theorem)**.** *Let $G = (A \cup B, E)$ be a bipartite graph. Then G has a matching that covers every vertex of A if and only if $|N(S)| \geq |S|$ for every subset $S \subseteq A$.*

You might wonder why it is called Marriage Theorem. Consider the vertices in A and B to represent people, and an edge indicates that two people are willing to marry each other, then Hall's theorem characterizes exactly when it is possible to find marriages so that every person in A gets a distinct partner from B.

Idea of the proof. The condition is clearly necessary: if a matching covers all vertices of A, then every subset $S \subseteq A$ is matched to $|S|$ distinct vertices in $N(S)$.

The converse is more difficult and is usually proved by induction on $|A|$. The condition says that every collection of vertices in A has enough possible partners in B. Hall's theorem states that this local condition is also sufficient to choose all partners simultaneously. □

Corollary 6.114. *Every non-empty regular bipartite graph has a perfect matching.*

Proof. Let $G = (A \cup B, E)$ be a k-regular bipartite graph. Counting edges from both sides gives

$$k|A| = |E| = k|B|.$$

Hence $|A| = |B|$. Now let $S \subseteq A$. Every vertex in S has degree k, so there are $k|S|$ edges leaving S. All these edges go to vertices in $N(S)$, and each vertex in $N(S)$ is incident to at most k of them. Thus

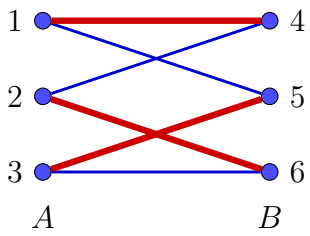
$$k|S| \leq k|N(S)|.$$

Therefore

$$|S| \leq |N(S)|.$$

By Hall's theorem, there is a matching covering A. Since $|A| = |B|$, this matching is perfect. □

Example 6.115. *Consider the bipartite graph $G = (A \cup B, E)$ with $A = \{1, 2, 3\}, B = \{4, 5, 6\}$ and $E = \{14, 15, 24, 26, 35, 36\}$.*



Every vertex has degree 2, so the graph is regular and bipartite. Hence, by the previous corollary, it has a perfect matching. One such perfect matching is shown in red: $M = \{14, 26, 35\}$.

Remember the cat game from the video? `https://graphcat.notabot.ai/` We will now reveal when Player 2 has a winning strategy.

Theorem 6.116 (The Cat Game)**.** *Consider the cat game on a graph G. Player 1 chooses the starting vertex, and afterwards the players alternately move the cat to an adjacent vertex that has not been visited before. Then Player 2 has a winning strategy for every starting vertex if G has a perfect matching.*

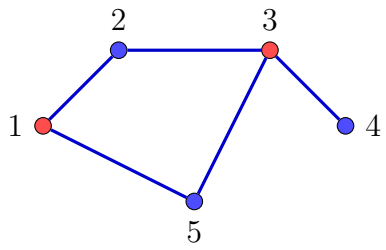
Proof. Let $M = \{e_1, \ldots, e_k\}$ be a perfect matching of G. Suppose Player 1 places the cat on a vertex v. Since M is perfect, there is a unique edge $vw \in M$ containing v. Player 2 moves the cat to w. Afterwards, whenever Player 1 moves the cat to a vertex x, Player 2 responds by moving along the matching edge of M containing x. This strategy is always possible. Indeed, if the matching partner of x had already been visited, then the cat must have visited x earlier as well, since the two vertices are paired by the matching. This is impossible because vertices may not be visited twice. Therefore Player 2 can always respond to every move of Player 1. Since every move of Player 1 is followed by a move of Player 2, Player 2 makes the last move and wins the game. □

We can also consider the opposite of a matching; a subset of vertices, which is incident to all the edges.

Definition 6.117. A *vertex cover* of a graph is a set of vertices $C \subseteq V$ such that every edge of the graph is incident to at least one vertex in C.

A vertex cover is not to be confused with a vertex being covered, which means it is covered by a matching.

Example 6.118. *Let us consider again the graph with* $V = \{1, 2, 3, 4, 5\}$ *and* $E = \{12, 23, 34, 15, 35\}$



The set $C = \{1, 3\}$ *is a vertex cover, because every edge is incident to at least one vertex in* C.

Theorem 6.119 (König)**.** *Let* G *be a bipartite graph. Then the size of a maximum matching equals the size of a minimum vertex cover.*

In every graph, the size of a matching is at most the size of a vertex cover. Indeed, every edge of the matching must be covered by a different vertex of the cover.

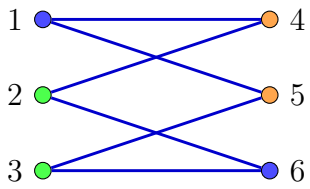
König's theorem says that for bipartite graphs this obvious lower bound is always tight. We omit the proof, and instead turn to graph colorings.

Definition 6.120. Let $G = (V, E)$ be a graph. A *t-coloring* of G is a function

$$\kappa : V \to [t].$$

A coloring is called *proper* if adjacent vertices always receive different colors, that is, if $uv \in E$ implies $\kappa(u) \neq \kappa(v)$.

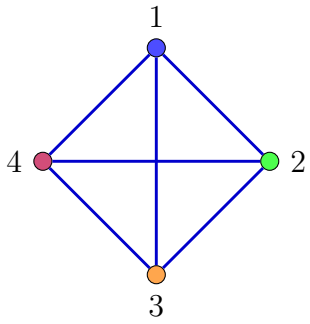
Example 6.121. *The following is a proper* 3*-coloring of the graph.*



Definition 6.122. The *chromatic number* of a graph G, denoted by $\chi(G)$, is the smallest number of colors needed for a proper coloring of G.

In our previous example, we had used three colors, but in fact, for bipartite graphs we can simply color all vertices from one part in one color and all vertices from the other part in another, getting $\chi(G) = 2$.

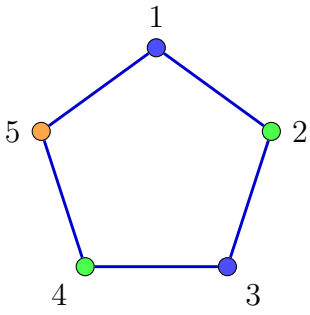
Example 6.123. *The complete graph* K_n *satisfies* $\chi(K_n) = n$.



Since every pair of vertices is adjacent, no two vertices may have the same color. Hence $\chi(K_4) = 4$.

Example 6.124. *The cycle graph* C_n *satisfies*

$$\chi(C_n) = \begin{cases} 2, & \text{if } n \text{ is even}, \\ 3, & \text{if } n \text{ is odd}. \end{cases}$$



Since C_5 is an odd cycle, it cannot be colored properly with only 2 colors. The coloring above shows that $\chi(C_5) = 3$.

Proposition 6.125 (Greedy Coloring)**.** *Let G be a graph with maximum degree $\Delta(G)$. Then*

$$\chi(G) \leq \Delta(G) + 1.$$

Proof. Order the vertices as $v_1, v_2, \ldots, v_n$. We color the vertices one by one in this order. When coloring v_i, at most $\Delta(G)$ of its neighbors have already been colored. Therefore at most $\Delta(G)$ colors are forbidden.

If we have $\Delta(G) + 1$ colors available, then at least one color remains available for v_i. Thus we obtain a proper coloring with at most $\Delta(G) + 1$ colors. □

The greedy algorithm depends on the chosen ordering of the vertices. A bad ordering may use more colors than necessary.

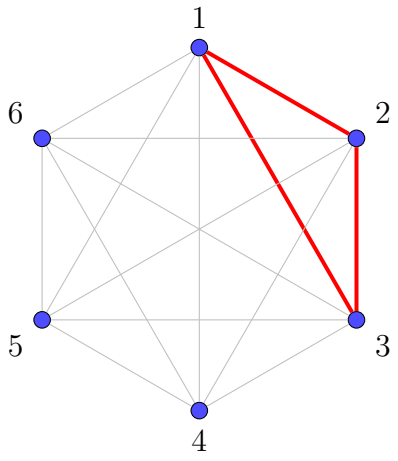
We end this subsection with a classical Ramsey-type result.

Theorem 6.126. *Among any six people, there are either three people who all know each other or three people such that no two of them know each other.*

Proof. Represent the six people by the vertices of a complete graph K_6.

Color an edge red if the corresponding two people know each other, and blue otherwise.

Choose one vertex v. It is incident to five edges. By the pigeonhole principle, at least three of these edges have the same color. Suppose first that there are three red edges from v to vertices a, b, c. Now look at the three edges among a, b, c. If at least one of these edges is red, say ab, then v, a, b form a red triangle, because va, vb, ab are all red. If none of the edges among a, b, c is red, then all three of them are blue. Hence a, b, c form a blue triangle. □



7 q-Analogues

In this chapter we study analogues of binomial coefficients where subsets of a finite set are replaced by subspaces of a finite vector space.

Let $V = \mathbb{F}_q^n$ be an n-dimensional vector space over the finite field $\mathbb{F}_q$. Our main object in this chapter will be *k-subspaces*: A *k-subspace* of V is a k-dimensional subspace of V.

We have the following dictionary translating our previous objects to their q-analogues:

sets	vector spaces
element	vector
subset	subspace
k-subset	k-dimensional subspace
cardinality	dimension
binomial coefficient	Gaussian binomial coefficient

7.1 Counting Subspaces

We define the *Gaussian binomial coefficient* $\binom{n}{k}_q$ to be the number of k-dimensional subspaces of $\mathbb{F}_q^n$.

Theorem 7.1. *The number of k-dimensional subspaces of $\mathbb{F}_q^n$ is*

$$\binom{n}{k}_q = \frac{(q^n-1)(q^{n-1}-1)\cdots(q^{n-k+1}-1)}{(q^k-1)(q^{k-1}-1)\cdots(q-1)}.$$

Proof. We count ordered bases of k-dimensional subspaces in two ways.

First choose k linearly independent vectors in $\mathbb{F}_q^n$.

There are q^n-1 choices for the first vector, q^n-q choices for the second vector, since it must not lie in the span of the first, and in general q^n-q^i choices for the $(i+1)$-st vector.

Thus the number of ordered linearly independent k-tuples in $\mathbb{F}_q^n$ is

$$(q^n-1)(q^n-q)\cdots(q^n-q^{k-1}).$$

Now fix one k-dimensional subspace U. The number of ordered bases of U is

$$(q^k-1)(q^k-q)\cdots(q^k-q^{k-1}).$$

Therefore

$$\binom{n}{k}_q = \frac{(q^n-1)(q^n-q)\cdots(q^n-q^{k-1})}{(q^k-1)(q^k-q)\cdots(q^k-q^{k-1})}.$$

Cancelling the powers of q gives

$$\binom{n}{k}_q = \frac{(q^n-1)(q^{n-1}-1)\cdots(q^{n-k+1}-1)}{(q^k-1)(q^{k-1}-1)\cdots(q-1)}.$$

□

Example 7.2. *For $k = 1$, we obtain*

$$\binom{n}{1}_q = \frac{q^n-1}{q-1} = 1 + q + \cdots + q^{n-1}.$$

This counts the 1-dimensional subspaces of $\mathbb{F}_q^n$.

If we formally set $q = 1$, then $\binom{n}{k}_q$ becomes the ordinary binomial coefficient $\binom{n}{k}$.

This is why Gaussian binomial coefficients are called q-analogues of binomial coefficients.

7.2 Recursion

The ordinary binomial coefficients satisfy Pascal's recursion

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

There is a similar recursion for Gaussian binomial coefficients.

Proposition 7.3. *For all $0 \le k \le n$, we have*

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

Proof. Let $H \subseteq \mathbb{F}_q^n$ be a fixed hyperplane, that means $\dim(H) = n-1$. We count the k-dimensional subspaces $U \subseteq \mathbb{F}_q^n$ in two classes.

First, U may be contained in H. The number of such subspaces is $\binom{n-1}{k}_q$.

Second, U may not be contained in H. In this case, $U \cap H$ has dimension $k-1$.

There are $\binom{n-1}{k-1}_q$ choices for this intersection.

For each fixed $(k-1)$-subspace $W \subseteq H$, there are q^{n-k} different k-subspaces U such that $U \cap H = W$ and $U \not\subseteq H$.

Hence the number of k-subspaces not contained in H is

$$q^{n-k}\binom{n-1}{k-1}_q.$$

Adding the two cases gives

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

□

Example 7.4. *The number of* 2*-dimensional subspaces of* $\mathbb{F}_q^4$ *is*

$$\binom{4}{2}_q = \frac{(q^4-1)(q^3-1)}{(q^2-1)(q-1)}.$$

Simplifying gives

$$\binom{4}{2}_q = q^4 + q^3 + 2q^2 + q + 1.$$

Example 7.5. *For* $q = 2$*, the number of* 2*-dimensional subspaces of* $\mathbb{F}_2^4$ *is*

$$\binom{4}{2}_2 = 2^4 + 2^3 + 2 \cdot 2^2 + 2 + 1 = 16 + 8 + 8 + 2 + 1 = 35.$$

7.3 The Poset of Subspaces

Just as subsets of a set form a partially ordered set under inclusion, the subspaces of a vector space form a partially ordered set under inclusion.

Definition 7.6. The *subspace lattice* of $\mathbb{F}_q^n$ is the partially ordered set whose elements are subspaces of $\mathbb{F}_q^n$, ordered by inclusion.

Example 7.7. *For* $\mathbb{F}_q^2$*, the subspace lattice consists of* $\{0\}$*, the* 1*-dimensional subspaces, and the whole space* $\mathbb{F}_q^2$*. There are*

$$\binom{2}{1}_q = q + 1$$

one-dimensional subspaces.

Definition 7.8. A *maximal chain* in the subspace lattice of $\mathbb{F}_q^n$ is a chain

$$0 = U_0 \subset U_1 \subset \cdots \subset U_n = \mathbb{F}_q^n$$

with $\dim(U_i) = i$ for all i.

Proposition 7.9. *The number of maximal chains in the subspace lattice of* $\mathbb{F}_q^n$ *is*

$$M(n,q) = \frac{(q^n-1)(q^{n-1}-1)\cdots(q-1)}{(q-1)^n}.$$

Proof. Starting with $U_0 = \{0\}$, suppose we have already chosen an i-dimensional subspace U_i.

To choose U_{i+1} containing U_i, we choose a vector outside U_i and take its span together with U_i.

There are $q^n - q^i$ vectors outside U_i.

However, each $(i+1)$-dimensional subspace containing U_i is obtained $q^{i+1} - q^i$ times in this way, namely from the vectors in $U_{i+1} \setminus U_i$.

Therefore the number of choices for U_{i+1} is

$$\frac{q^n - q^i}{q^{i+1} - q^i} = \frac{q^{n-i} - 1}{q - 1}.$$

Multiplying over $i \in \{0, 1, \ldots, n-1\}$ gives

$$M(n,q) = \prod_{i=0}^{n-1} \frac{q^{n-i} - 1}{q - 1} = \frac{(q^n - 1)(q^{n-1} - 1) \cdots (q - 1)}{(q-1)^n}.$$

□

7.4 A q-Analogue of Sperner's Theorem

Recall that Sperner's theorem says that the largest antichain in the Boolean lattice is obtained by taking all subsets of size $\left\lfloor \frac{n}{2} \right\rfloor$.

The q-analogue replaces subsets by subspaces.

Theorem 7.10 (q-Sperner theorem)**.** *Let $\mathcal{A}$ be an antichain in the subspace lattice of $\mathbb{F}_q^n$. Then*

$$|\mathcal{A}| \leq \binom{n}{\lfloor n/2 \rfloor}_q.$$

Idea of the proof. The proof is analogous to the proof of Sperner's theorem.

Count pairs (U, C), where $U \in \mathcal{A}$ and C is a maximal chain containing U.

Since $\mathcal{A}$ is an antichain, every maximal chain contains at most one element of $\mathcal{A}$.

Thus the number of such pairs is at most $M(n,q)$.

On the other hand, if $\dim(U) = k$, then the number of maximal chains containing U is

$$M(k,q)M(n-k,q).$$

Writing c_k for the number of k-dimensional subspaces in $\mathcal{A}$, we obtain

$$\sum_{k=0}^{n} c_k M(k,q) M(n-k,q) \leq M(n,q).$$

Equivalently,

$$\sum_{k=0}^{n} \frac{c_k}{\binom{n}{k}_q} \leq 1.$$

Since

$$\binom{n}{k}_q \leq \binom{n}{\lfloor n/2 \rfloor}_q$$

for all k, we obtain

$$|\mathcal{A}| \leq \binom{n}{\lfloor n/2 \rfloor}_q.$$

□

7.5 Partitions in a Box

Gaussian binomial coefficients are not only counting functions. They are also polynomials in q with nonnegative integer coefficients.

Theorem 7.11. *Write*

$$\binom{n}{k}_q = \sum_{\ell=0}^{k(n-k)} a_\ell q^\ell.$$

Then a_ℓ is the number of partitions of ℓ whose Ferrers diagrams fit inside a box of size $k \times (n-k)$.

Example 7.12. *We have*

$$\binom{4}{2}_q = q^4 + q^3 + 2q^2 + q + 1.$$

The coefficients $1, 1, 2, 1, 1$ count partitions fitting inside a 2×2 box.

Setting $q = 1$ shows that the total number of partitions whose Ferrers diagrams fit inside a $k \times (n-k)$ box is $\binom{n}{k}$.

8 Finite Geometry

Geometry is often described in terms of *points* and *lines*, together with the information of which points lie on which lines. In combinatorics, we abstract this idea and study the resulting incidence relations.

Definition 8.1. An *incidence structure* is a triple $(X, \mathcal{L}, I)$, where

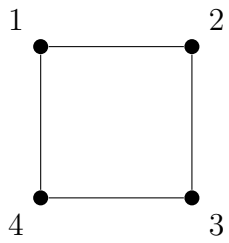
- X is a set of *points*,
- $\mathcal{L}$ is a set of *lines*,
- $I \subseteq X \times \mathcal{L}$ is the *incidence relation*.

If $(P, L) \in I$, we say that the point P is *incident* with the line L, or equivalently that P lies on L.

In practice, one often identifies a line with the set of points incident with it. In this way, an incidence structure can be viewed simply as a set of points together with a collection of subsets, called lines., i.e., $(X, \mathcal{L})$.

Example 8.2. *Consider the point set* $X = \{1, 2, 3, 4\}$ *and the collection of lines*

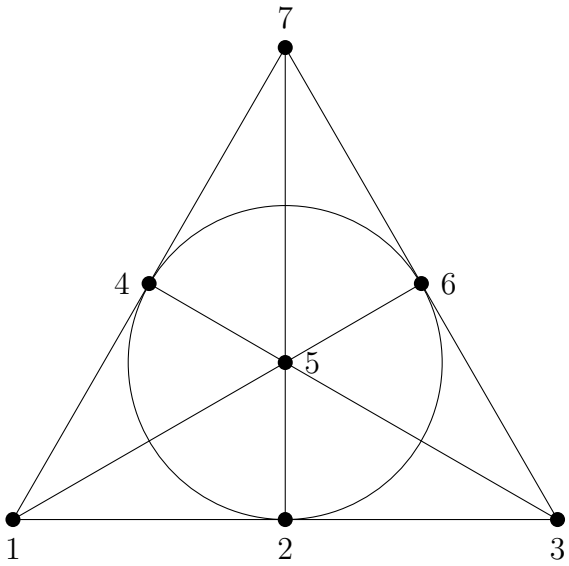
$$\mathcal{L} = \big\{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 4\}\big\}.$$



Then $(X, \mathcal{L})$ *is an incidence structure.*

As we can see in the example, a useful way to visualize an incidence structure is by drawing its points and lines.

Example 8.3. *The following incidence structure consists of seven points and seven lines.*



Each line contains exactly three points and each pair of points lies on a unique line. This structure is known as the Fano plane *and will play an important role throughout this chapter.*

Many geometric properties can be formulated entirely in terms of incidence.

Definition 8.4. A set of points is called *collinear* if all of its points lie on a common line.

A point is called an *intersection point* of two lines if it lies on both lines.

In our example of the Fano plane above, $1, 2, 3$ are collinear. The lines through $1, 2, 3$ and $1, 4, 7$ have as intersection point 1.

The incidence relation can also be encoded by a matrix.

Definition 8.5. Let $X = \{P_1, \ldots, P_v\}$ and $\mathcal{L} = \{L_1, \ldots, L_b\}$. The *incidence matrix* of the incidence structure is the matrix $M = (m_{ij}) \in \{0, 1\}^{v \times b}$ defined by

$$m_{ij} = \begin{cases} 1, & P_i \in L_j, \\ 0, & P_i \notin L_j. \end{cases}$$

Example 8.6. *In our previous example of the Fano plane, let us label our lines as*

$$\mathcal{L} = \Big\{\{1,2,3\}, \{1,4,7\}, \{3,6,7\}, \{1,5,6\}, \{2,5,7\}, \{3,4,5\}, \{2,4,6\}\Big\}.$$

Then the incidence matrix is given by

$$M = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Incidence matrices provide a bridge between combinatorics and linear algebra. They will reappear later in the study of designs and coding theory.

In the next sections we study two particularly important classes of incidence structures: projective planes and affine planes.

8.1 Projective Planes

In Euclidean geometry, two distinct points determine a unique line. However, parallel lines do not intersect. Projective geometry modifies this situation by requiring that every pair of lines meets in a point.

The resulting structures are remarkably symmetric and form one of the central objects of finite geometry.

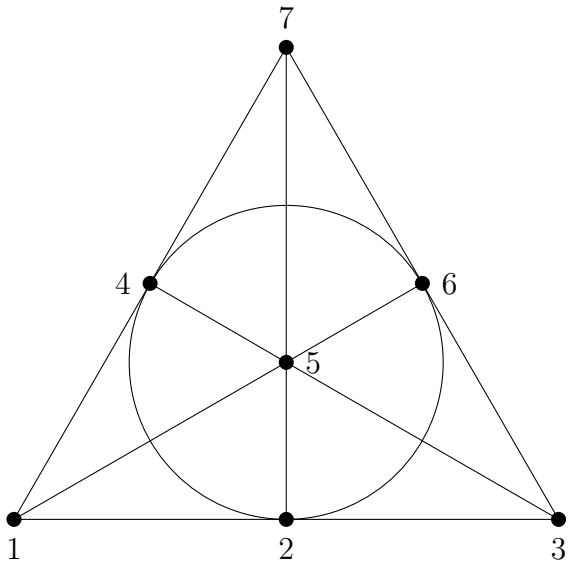
Definition 8.7. A *projective plane* is an incidence structure $(X, \mathcal{L}, I)$ satisfying:

1. Any two distinct points lie on a unique line.
2. Any two distinct lines meet in a unique point.
3. There exist four points such that no three are collinear.

The first axiom says that a line is uniquely determined by two points, while the second is the dual statement obtained by exchanging points and lines.

The third axiom excludes degenerate examples consisting of a single line.

Example 8.8. *The smallest projective plane is the Fano plane. It consists of seven points and seven lines. Every line contains three points and every point lies on three lines.*



One of the remarkable features of projective planes is that all lines contain the same number of points.

Theorem 8.9. *Let Π be a finite projective plane. Then there exists an integer $q \geq 2$ such that every line contains exactly $q+1$ points and every point lies on exactly $q+1$ lines.*

Proof. Let L be a line and let P be a point not on L. For every point $Q \in L$, there is a unique line through P and Q. Moreover, these lines are all distinct. Conversely, every line through P meets L in exactly one point, since any two lines meet in a unique point. Thus the lines through P are in bijection with the points on L. In particular, the number of lines through P is equal to the number of points on L.

Now let L_1 and L_2 be two lines. Choose a point $P \notin L_1 \cup L_2$, which exists by the non-degeneracy axiom. By the previous argument, the number of points on L_1 equals the number of lines through P, and the same is true for L_2. Hence L_1 and L_2 contain the same number of points.

Thus every line contains the same number of points. We write this number as $q+1$.

Dually, the same argument shows that every point lies on the same number of lines. Since the number of lines through any point equals the number of points on any line, every point lies on exactly $q+1$ lines. □

Definition 8.10. The integer q from the previous theorem is called the *order* of the projective plane.

The order completely determines the basic parameters of a finite projective plane.

Theorem 8.11. *A projective plane of order q contains q^2+q+1 points and equally many lines.*

Proof. Fix a point P. There are exactly $q+1$ lines through P. Each of these lines contains q further points.

By Axiom 2, two distinct lines through P intersect only in P. Hence these points are all distinct.

Therefore the total number of points is

$$1 + q(q+1) = q^2 + q + 1.$$

By duality, the same number counts the lines. □

For small values of q we obtain: The existence of a projective plane of order q is a difficult problem.

Order q	Points	Lines	Points per line
2	7	7	3
3	13	13	4
4	21	21	5

Projective planes are known whenever q is a prime power, but it is unknown whether projective planes exist for arbitrary orders.

For example, there is a projective plane of order $2, 3, 4, 5, 7, 8, 9, \ldots$ but no projective plane of order 6. The next section explains how projective planes can be constructed from finite fields.

8.2 Affine Planes

In a projective plane, any two lines intersect in a unique point. This makes the geometry highly symmetric, but it differs from the Euclidean geometry familiar from school, where parallel lines do not meet.

Affine planes provide a finite analogue of Euclidean geometry.

Definition 8.12. An *affine plane* is an incidence structure $(X, \mathcal{L}, I)$ satisfying:

1. Any two distinct points lie on a unique line.
2. Given a line L and a point $P \notin L$, there exists a unique line through P that does not intersect L.
3. There exist three points that are not collinear.

The second axiom introduces the notion of parallelism.

Definition 8.13. Two distinct lines of an affine plane are called *parallel* if they do not intersect.

The second axiom implies that parallelism behaves exactly as in ordinary Euclidean geometry.

Proposition 8.14. *Parallelism is an equivalence relation on the set of lines.*

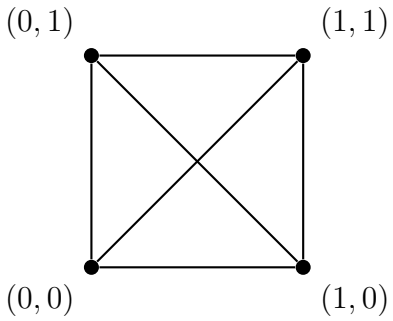
Exercise 8.15. *Prove Proposition 8.14.*

The equivalence classes are called *parallel classes*.

Example 8.16. *Consider $X = \mathbb{F}_2^2$ and the lines*

$$\{(0,0),(0,1)\},\ \{(1,0),(1,1)\},\ \{(0,0),(1,0)\},\ \{(0,1),(1,1)\},\ \{(0,0),(1,1)\},\ \{(0,1),(1,0)\}.$$

These six lines form an affine plane of order 2*.*



Just as for projective planes, all lines of a finite affine plane have the same size.

Theorem 8.17. *Let Π be a finite affine plane. Then there exists an integer $q \geq 2$ such that every line contains exactly q points and every point lies on exactly $q + 1$ lines.*

Proof. Let L be a line and let P be a point not on L. For every point $Q \in L$, there is a unique line through P and Q. These lines are all distinct. Conversely, every line through P which is not parallel to L meets L in exactly one point.

By the parallel axiom, there is exactly one line through P parallel to L. Hence the number of lines through P is $|L|+1$. In particular, the number of lines through a point is determined by the number of points on a line.

Now let L_1 and L_2 be two lines. Choose a point P not on either of them. Applying the previous argument to L_1 and to L_2 shows that $|L_1| + 1 = |L_2| + 1$. Hence all lines have the same number of points. We call this number q.

Therefore every point lies on exactly $q+1$ lines. Since the affine plane contains three non-collinear points, every line contains at least two points, so $q \geq 2$. □

Also in this case, do we call the integer q the *order* of the affine plane and the order determines the basic parameters of the geometry.

Theorem 8.18. *An affine plane of order q contains q^2 points and $q^2 + q$ lines.*

Proof. Fix a point P. There are $q + 1$ lines through P, each containing $q - 1$ further points.

Since distinct lines through P intersect only in P, these points are all distinct.

Hence the total number of points is

$$1 + (q + 1)(q - 1) = q^2.$$

Each point lies on $q + 1$ lines and each line contains q points. Counting incidences yields

$$q \cdot |\mathcal{L}| = q^2(q + 1),$$

and therefore

$$|\mathcal{L}| = q^2 + q.$$

□

Affine and projective planes are closely related. Starting with a projective plane of order q, choose a line L and remove it together with all points on it. The remaining incidence structure is an affine plane of order q.

Conversely, a projective plane can be obtained from an affine plane by adding one new line, called the *line at infinity*, whose points correspond to the parallel classes of the affine plane.

Thus projective and affine planes may be viewed as two closely related versions of the same geometric object.

8.3 Finite Fields and Geometry

In the previous sections we introduced projective and affine planes axiomatically. We now show that finite fields provide a systematic construction of these geometries.

Let q be a prime power and let $\mathbb{F}_q$ be the finite field with q elements.

We consider the vector space $\mathbb{F}_q^2$. Its elements will serve as the points of an affine plane.

Definition 8.19. The *affine plane over* $\mathbb{F}_q$, denoted by $\mathrm{AG}(2, q)$, has

- points: the vectors of $\mathbb{F}_q^2$,
- lines: the solution sets of equations $ax + by = c$, where $(a, b) \neq (0, 0)$.

Example 8.20. *For $q = 2$ the point set is $X = \mathbb{F}_2^2$ and the six lines are*

$$x = 0, \qquad x = 1, \qquad y = 0, \qquad y = 1, \qquad x + y = 0, \qquad x + y = 1.$$

These form the affine plane of order 2.

Theorem 8.21. *The incidence structure $\mathrm{AG}(2, q)$ is an affine plane of order q.*

Idea of the proof. Every line contains exactly q points.

Given two distinct points, solving a linear system yields a unique line containing them.

Furthermore, lines with the same direction are parallel, and through every point there is a unique line parallel to a given line. □

Thus finite fields give affine planes of every prime power order.

Affine geometry has parallel lines. To obtain a projective plane we add points at infinity corresponding to directions. A more elegant construction uses linear algebra.

Definition 8.22. The *projective plane over* $\mathbb{F}_q$, denoted by $\mathrm{PG}(2, q)$, is defined as follows:

- Points are the one-dimensional subspaces of $\mathbb{F}_q^3$.
- Lines are the two-dimensional subspaces o $\mathbb{F}_q^3$.
- Incidence is given by inclusion.

In other words, two non-zero vectors $v, w \in \mathbb{F}_q^3$ represent the same projective point if $v = \lambda w$ for some non-zero $\lambda \in \mathbb{F}_q$. Thus a point corresponds to an equivalence class $[x : y : z]$.

Example 8.23. *In $\mathrm{PG}(2, 2)$ there are*

$$\frac{2^3 - 1}{2 - 1} = 7$$

points.

The resulting geometry is precisely the Fano plane.

We can count the points of $\mathrm{PG}(2, q)$.

Theorem 8.24. *The projective plane* $\mathrm{PG}(2, q)$ *contains* $q^2 + q + 1$ *points.*

Proof. There are $q^3 - 1$ non-zero vectors in $\mathbb{F}_q^3$.

Each one-dimensional subspace contains exactly $q - 1$ non-zero vectors.

Therefore the number of points equals

$$\frac{q^3 - 1}{q - 1} = q^2 + q + 1.$$

□

A similar argument shows that there are also $q^2 + q + 1$ lines and that every line contains exactly $q + 1$ points.

Consequently, $\mathrm{PG}(2, q)$ is a projective plane of order q.

Finite fields therefore provide a rich source of geometric objects:

$$\mathbb{F}_q \Longrightarrow \mathrm{AG}(2, q) \Longrightarrow \mathrm{PG}(2, q).$$

In the next chapter we shall see that these geometries naturally give rise to combinatorial designs and, later, to linear codes.

9 Designs

Suppose there are seven students $S_1, S_2, \ldots, S_7$ participating in a seminar.

The instructor would like to organize study groups of size three. Moreover, every pair of students should work together exactly once during the semester.

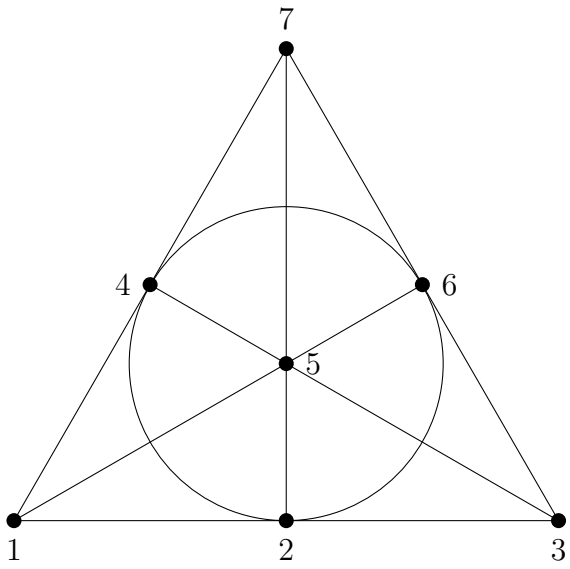
Is such an arrangement possible?

Surprisingly, yes. Consider the following seven study groups:

$$
\begin{aligned}
G_1 &= \{S_1, S_2, S_3\},\\
G_2 &= \{S_1, S_4, S_7\},\\
G_3 &= \{S_3, S_6, S_7\},\\
G_4 &= \{S_4, S_5, S_6\},\\
G_5 &= \{S_2, S_5, S_7\},\\
G_6 &= \{S_1, S_5, S_6\},\\
G_7 &= \{S_2, S_3, S_4\}.
\end{aligned}
$$

Each group contains exactly three students, and every pair of students appears together in exactly one group.

We may represent the students by points and the study groups by lines. The resulting incidence structure is the famous *Fano plane*.



This example illustrates a common theme in combinatorics: we seek collections of subsets that are as balanced and symmetric as possible. Such structures are called *designs*.

In the previous chapter we studied finite geometries, where points and lines satisfy highly regular incidence properties. For example, in a projective plane every pair of points determines a unique line, and every line contains the same number of points.

Design theory abstracts this phenomenon. Instead of lines, we consider arbitrary subsets of a point set, called *blocks*, and require that every collection of t points occurs in the same number of blocks. The resulting structures are called *designs*.

Designs arise naturally in geometry, coding theory, statistics, and combinatorics. Historically, they were introduced in the context of experimental design, where treatments had to be distributed among test groups in a balanced way. Today they play an important role in finite geometry, error-correcting codes, association schemes, and finite group theory.

Definition 9.1. Let X be a set with $|X| = v$ and let $\mathcal{B}$ be a collection of k-element subsets of X. Let t, k, v be integers satisfying $0 < t < k < v$.

The elements of X are called *points* and the elements of $\mathcal{B}$ are called *blocks*.

The pair $(X, \mathcal{B})$ is called a t-(v, k, λ) *design* if every t-subset of X is contained in exactly λ blocks.

The parameters of a design have the following meaning:

- v is the number of points,
- k is the size of each block,
- t is the size of the subsets that are balanced,
- λ is the number of blocks containing any given t-subset.

When $\lambda = 1$, every t-subset is contained in a unique block.

Definition 9.2. A t-$(v, k, 1)$ design is called a *Steiner system* and is denoted by $S(t, k, v)$.

We often write simply *t-design* if the parameters are clear from the context.

Example 9.3. *A partition of a set of v points into subsets of size k forms a $1-(v, k, 1)$ design, since every point belongs to exactly one block.*

Example 9.4. *Let X be a set of size v and let $\mathcal{B}$ be the collection of all k-subsets of X.*

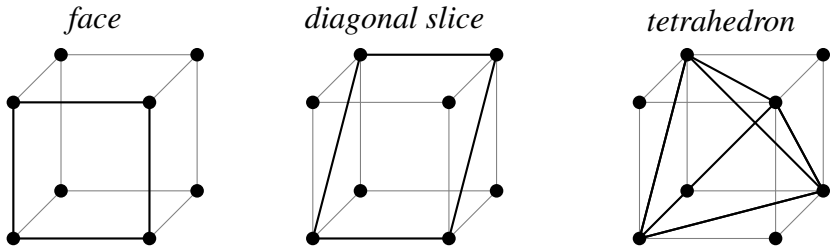
Given any t-subset $T \subseteq X$, we obtain a block containing T by choosing $k-t$ additional elements from the remaining $v-t$ points. Therefore T is contained in exactly $\binom{v-t}{k-t}$ blocks. Hence $(X, \mathcal{B})$ is a $t-\left(v, k, \binom{v-t}{k-t}\right)$ design, called the complete design.

Example 9.5. *The points and lines of a projective plane of order q form a $2-(q^2+q+1, q+1, 1)$ design.*

Indeed, every line contains $q+1$ points and every pair of points lies on a unique line.

The last example illustrates the close connection between finite geometry and design theory. Many of the incidence structures arising in finite geometry can be viewed as designs, and conversely many important designs can be constructed from geometric objects.

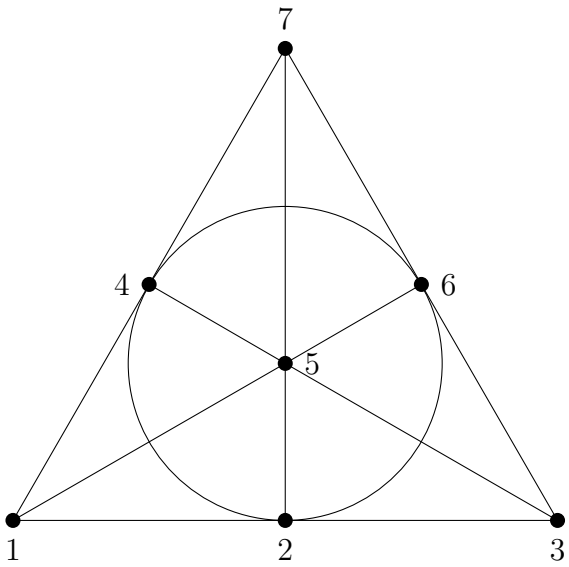
Example 9.6 (The cube design)**.** *Let X be the set of the eight vertices of a cube. We take as blocks certain 4-subsets of vertices: the six faces, the six diagonal slices, and the eight tetrahedra. Thus there are $6 + 6 + 8 = 20$ blocks. Every three vertices are contained in exactly one of these blocks. Hence this gives a $3 - (8, 4, 1)$ design.*



Example 9.7 (The Fano plane)**.** *Let $X = \{1, 2, 3, 4, 5, 6, 7\}$ and let*

$$\mathcal{B} = \Big\{\{1,2,3\}, \{1,4,7\}, \{3,6,7\}, \{2,4,6\}, \{2,5,7\}, \{1,5,6\}, \{3,4,5\}\Big\}.$$

Then every pair of points is contained in exactly one block. Hence $(X, \mathcal{B})$ is a $2 - (7, 3, 1)$ design.



The definition of a design imposes strong restrictions on its parameters. By counting the same objects in different ways, we obtain several useful relations. We write $b = |\mathcal{B}|$ for the number of blocks of a design.

Proposition 9.8. *Let $(X, \mathcal{B})$ be a t-(v, k, λ) design. Then*

$$b = \lambda \frac{\binom{v}{t}}{\binom{k}{t}}.$$

Proof. Consider all pairs (T, B) where T is a t-subset of X and B is a block containing T. On the one hand, there are $\binom{v}{t}$ choices for T, and each t-subset lies in exactly λ blocks. Thus the number of pairs equals $\lambda\binom{v}{t}$. On the other hand, each block contains $\binom{k}{t}$ t-subsets and there are b blocks. Hence the number of pairs is $b\binom{k}{t}$.

Equating the two expressions gives

$$b\binom{k}{t} = \lambda\binom{v}{t},$$

and thus the claim. □

Proposition 9.9. *Let $(X, \mathcal{B})$ be a t-(v, k, λ) design. Every point $x \in X$ is contained in the same number of blocks, namely $|\{B \in \mathcal{B} \mid x \in B\}| = \lambda \frac{\binom{v-1}{t-1}}{\binom{k-1}{t-1}}$.*

Exercise 9.10. *Prove Proposition 9.9 by a double-counting argument for incidence (x, B) with $x \in B$.*

This number has a special name, it is called *replication number*, and is denoted by r, because each point appears (is replicated) in exactly r blocks. In the language of incidence structures, r is simply the degree of a point.

Definition 9.11. Let $(X, \mathcal{B})$ be a t-(v, k, λ) design. The *replication number* of the design is the number r of blocks containing a fixed point. Equivalently,

$$r = |\{B \in \mathcal{B} \mid x \in B\}|$$

for any point $x \in X$.

Example 9.12. *The replication number of the Fano plane is*

$$r = 1 \cdot \frac{\binom{6}{1}}{\binom{2}{1}} = 3.$$

Indeed, every point lies on exactly three lines.

Counting incidences gives the fundamental relation

$$bk = vr,$$

where b is the number of blocks.

9.1 New Designs from Old

One way to obtain new designs from existing ones is to fix some points and consider only the blocks containing them.

Definition 9.13. Let $(X, \mathcal{B})$ be a t-(v, k, λ) design and let $S \subseteq X$ be a set of size $s < t$. Define $Y = X \setminus S$ and $\mathcal{C} = \{ B \setminus S \mid B \in \mathcal{B},\ S \subseteq B \}$. The incidence structure $(Y, \mathcal{C})$ is called the *derived design* of $(X, \mathcal{B})$ with respect to S.

Intuitively, we keep only those blocks containing the fixed set S and then remove the points of S from each remaining block.

Theorem 9.14. *Let $(X, \mathcal{B})$ be a t-(v, k, λ) design and let $S \subseteq X$ with $|S| = s < t$. Then the derived design is a $(t - s) - (v - s, k - s, \lambda)$ design.*

Proof. Every block of $\mathcal{C}$ has size $k - s$ and is contained in the point set Y, which has size $v - s$.

Let $Z \subseteq Y$ be a $(t - s)$-subset. Then $Z \cup S$ is a t-subset of X. Since $(X, \mathcal{B})$ is a t-(v, k, λ) design, the set $Z \cup S$ is contained in exactly λ blocks of $\mathcal{B}$.

These are precisely the blocks that give rise to blocks of $\mathcal{C}$ containing Z. Therefore Z is contained in exactly λ blocks of $\mathcal{C}$ and hence $(Y, \mathcal{C})$ is a $(t - s)$-$(v - s, k - s, \lambda)$ design. □

Example 9.15. *Recall the cube design, which is a $3 - (8, 4, 1)$ design.*

Fixing one vertex and deriving with respect to it yields a $2 - (7, 3, 1)$ design. This is the famous Fano plane, the smallest non-trivial Steiner triple system.

Derived designs provide a useful way of constructing smaller designs from larger ones. In particular, every t-design contains many $(t - 1)$-designs obtained by fixing a single point.

Another way to construct new designs from old ones is to replace each block by its complement.

Definition 9.16. Let $(X, \mathcal{B})$ be a t-(v, k, λ) design. The *complementary design* is obtained by replacing every block $B \in \mathcal{B}$ by its complement $X \setminus B$. Its block set is $\overline{\mathcal{B}} = \{X \setminus B \mid B \in \mathcal{B}\}$.

Since every original block has size k, every complementary block has size $v - k$. Thus the complementary design has parameters of the form $t - (v, v - k, \overline{\lambda})$, for a suitable value of $\overline{\lambda}$.

Theorem 9.17. *The complement of a t-(v, k, λ) design is a $t - (v, v - k, \overline{\lambda})$ design, where*

$$\overline{\lambda} = \sum_{i=0}^{t} (-1)^i \binom{t}{i} \lambda_i.$$

Here λ_i denotes the number of blocks containing a fixed i-subset of points.

Proof. Fix a t-subset $T = \{x_1, \ldots, x_t\}$. A complementary block contains T if and only if the corresponding original block contains none of the points of T. For each subset $I \subseteq T$, let

$$\mathcal{B}_I = \{B \in \mathcal{B} \mid I \subseteq B\}.$$

By the parameter formulas from the previous section, $|\mathcal{B}_I| = \lambda_{|I|}$. Using the principle of inclusion–exclusion, the number of blocks containing none of the points of T is

$$\sum_{i=0}^{t} (-1)^i \binom{t}{i} \lambda_i.$$

This number depends only on the parameters of the design and not on the particular choice of T. Hence every t-subset is contained in the same number of complementary blocks. □

For 2-designs the formula simplifies considerably.

Corollary 9.18. *The complement of a* 2-(v, k, λ) *design is a* $2 - (v, v - k, \overline{\lambda})$ *design with* $\overline{\lambda} = b - 2r + \lambda$, *where* b *is the number of blocks and* r *is the replication number.*

Example 9.19. *The Fano plane is a* $2 - (7, 3, 1)$ *design with* $b = 7, r = 3$. *Its complement is therefore a* $2 - (7, 4, 2)$ *design since* $\overline{\lambda} = 7 - 2 \cdot 3 + 12$.

Complementary designs often provide new examples from known ones and show that many parameter sets naturally occur in pairs.

9.2 Incidence Matrices

Designs can be represented by matrices in a natural way. This allows us to study designs using linear algebra and will later lead to important connections with coding theory.

Definition 9.20. Let $(X, \mathcal{B})$ be a design with $X = \{x_1, \ldots, x_v\}$ and $\mathcal{B} = \{B_1, \ldots, B_b\}$. The *incidence matrix* of the design is the matrix $M = (m_{ij}) \in \{0, 1\}^{v \times b}$ defined by

$$m_{ij} = \begin{cases} 1, & x_i \in B_j, \\ 0, & x_i \notin B_j. \end{cases}$$

Thus the rows correspond to points and the columns correspond to blocks.

Example 9.21. *Let us consider the Fano plane with* $X = \{1, 2, 3, 4, 5, 6, 7\}$ *and*

$$\mathcal{B} = \Big\{\{1,2,3\}, \{1,4,7\}, \{3,6,7\}, \{2,4,6\}, \{2,5,7\}, \{1,5,6\}, \{3,4,5\}\Big\}.$$

Note that this is a 2-(7,3,1) design. The incidence matrix is given by

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

The combinatorial parameters of a design can be read directly from its incidence matrix. In fact, if M is an incidence matrix of a $t-(v,k,\lambda)$ design, then every column of M contains exactly k ones and every row contains exactly r ones.

For 2-designs the matrix product $MM^\top$ has a particularly simple form.

Theorem 9.22. *Let M be the incidence matrix of a 2-(v,k,λ) design. Then*

$$MM^\top = (r-\lambda)Id_v + \lambda J,$$

where J denotes the all one matrix.

Proof. The (i,j)-entry of $MM^\top$ is

$$(MM^\top)_{ij} = \sum_{\ell=1}^{b} m_{i\ell} m_{j\ell}.$$

This counts the number of blocks containing both points x_i and x_j.

If $i=j$, the count equals r, since every point lies in exactly r blocks.

If $i \neq j$, the count equals λ, since every pair of points lies in exactly λ blocks.

Therefore

$$(MM^\top)_{ij} = \begin{cases} r, & i=j, \\ \lambda, & i \neq j, \end{cases}$$

which is precisely

$$(r-\lambda)\mathrm{Id}_v + \lambda J.$$

□

Example 9.23. *For the Fano plane, which is a $2-(7,3,1)$ design, we have $r=3$. Hence*

$$MM^\top = 2Id_7 + J.$$

One of the fundamental results of design theory is that a non-trivial 2-design must have at least as many blocks as points.

Theorem 9.24 (Fisher's Inequality)**.** *Let $(X, \mathcal{B})$ be a 2-(v, k, λ) design with $k < v$. Then $b \geq v$.*

Before proving the theorem, we need one more fact.

Exercise 9.25. *Let $(X, \mathcal{B})$ be a non-trivial 2-(v, k, λ) design, that is, assume $k < v$. Use Proposition 9.9 to show that $r > \lambda$.*

Proof of Fisher's Inequality. Let M be the incidence matrix of the design. Recall that

$$MM^\top = (r - \lambda)\mathrm{Id}_v + \lambda J.$$

Since $r > \lambda$, the matrix $(r - \lambda)\mathrm{Id}_v + \lambda J$ is invertible. Indeed, recall that the all-one matrix J has eigenvalue v corresponding to the eigenvector $(1, \ldots, 1)^\top$, and eigenvalue 0 with multiplicity $v-1$ on the subspace

$$\{x \in \mathbb{R}^v \mid x_1 + \cdots + x_v = 0\}.$$

Therefore:

- For $(1, \ldots, 1)$ we have

$$\big((r - \lambda)\,\mathrm{Id}_v + \lambda J\big)(1, \ldots, 1) = (r - \lambda + \lambda v)(1, \ldots, 1) = \big(r + (v - 1)\lambda\big)(1, \ldots, 1).$$

- If $x_1 + \cdots + x_v = 0$, then $Jx = 0$, and hence

$$\big((r - \lambda)\,\mathrm{Id}_v + \lambda J\big)x = (r - \lambda)x.$$

Thus the eigenvalues are $r - \lambda$ with multiplicity $v - 1$ and $r + (v - 1)\lambda$ with multiplicity 1. Since $r > \lambda$ and $\lambda > 0$, both eigenvalues are non-zero.

Therefore $\mathrm{rank}(MM^\top) = v$. Since M is a real matrix we have that

$$\mathrm{rank}(MM^\top) = \mathrm{rank}(M),$$

and hence $\mathrm{rank}(M) = v$. But M has only b columns, so $v = \mathrm{rank}(M) \leq b$. Hence $b \geq v$. □

Fisher's inequality shows that a 2-design cannot have too few blocks. Designs satisfying equality are particularly important, they are called *symmetric*.

Definition 9.26. A 2-(v, k, λ) design satisfying $b = v$ is called a *symmetric design*.

Many of the most important examples of designs are symmetric, including projective planes.

We first derive a useful relation among the parameters. Recall that for any design we have $bk = vr$, thus if we have a symmetric design, we further know $b = v$, thus also $k = r$.

Corollary 9.27. *Let $(X, \mathcal{B})$ be a symmetric 2-(v, k, λ) design. Then $r = k$.*

The following theorem gives several equivalent characterizations of symmetric designs.

Theorem 9.28. *Let $(X, \mathcal{B})$ be a 2-(v, k, λ) design with $k < v$. The following are equivalent:*

1. *$b = v$;*
2. *$r = k$;*
3. *any two distinct blocks intersect in exactly λ points;*
4. *any two distinct blocks intersect in a constant number of points.*

Proof. We already proved that (1) implies (2).

We now show (2) implies (3). Since $r = k$ we also get $b = v$, and hence the incidence matrix M is square. From

$$MM^\top = (r - \lambda)\mathrm{Id}_v + \lambda J$$

and the fact that M is invertible, one can show that

$$M^\top M = (r - \lambda)\mathrm{Id}_v + \lambda J.$$

The (i, j)-entry of $M^\top M$ counts the number of points contained in both blocks B_i and B_j.

Therefore any two distinct blocks intersect in exactly λ points. Hence (2) implies (3).

Clearly (3) implies (4).

Finally, suppose (4) holds. Interchanging the roles of points and blocks yields another 2-design, called *dual design*. Applying Fisher's inequality to this dual design gives $v \geq b$. Together with Fisher's inequality for the original design, $b \geq v$, we obtain $b = v$. Thus (4) implies (1). □

Recall that $r = \lambda \frac{\binom{v-1}{t-1}}{\binom{k-1}{t-1}}$. Thus for $t = 2$, we get $r(k-1) = \lambda(v-1)$. Combining this with $r = k$ gives a useful parameter relation.

Corollary 9.29. *The parameters of a symmetric design satisfy $k(k-1) = \lambda(v-1)$.*

Example 9.30. *The Fano plane is a $2-(7, 3, 1)$ design. Since the number of points and lines are both equal to 7, it is a symmetric design.*

Indeed,

$$3(3-1) = 6 = 1 \cdot (7-1).$$

One of the most important families of symmetric designs arises from projective planes.

Theorem 9.31. *The points and lines of a projective plane of order q form a*

$$2-(q^2+q+1, q+1, 1)$$

design.

Proof. The point set has size $v = q^2 + q + 1$. Each line contains exactly $k = q + 1$ points.

By the defining axioms of a projective plane, every pair of points lies on a unique line. Hence $\lambda = 1$. Therefore the points and lines form a

$$2 - (q^2 + q + 1, q + 1, 1)$$

design. □

Since a projective plane contains equally many points and lines, it is a symmetric design. Indeed, $v = b = q^2 + q + 1, k = r = q + 1$.

9.3 Difference Sets

Many symmetric designs can be constructed from groups. One of the most important constructions uses *difference sets*. Throughout this section, we write $\mathbb{Z}/v\mathbb{Z}$ for the integers modulo v.

Definition 9.32. Let $2 \leq k < v$ and $\lambda \geq 1$. A subset $D = \{d_1, \ldots, d_k\} \subseteq \mathbb{Z}/v\mathbb{Z}$ is called a (v, k, λ) *difference set* if every non-zero element of $\mathbb{Z}/v\mathbb{Z}$ can be written in exactly λ ways as $d_i - d_j$ with $i \neq j$.

In other words, every non-zero element occurs exactly λ times among the differences of elements of D.

Example 9.33. *The set $D = \{1, 2, 4\} \subseteq \mathbb{Z}/7\mathbb{Z}$ is a $(7, 3, 1)$ difference set.*

Indeed, the differences $1 - 2, 1 - 4, 2 - 1, 2 - 4, 4 - 1, 4 - 2$ produce each non-zero element of $\mathbb{Z}/7\mathbb{Z}$ exactly once.

Difference sets give rise to symmetric designs in a natural way.

Theorem 9.34. *Let D be a (v, k, λ) difference set in $\mathbb{Z}/v\mathbb{Z}$. For each $a \in \mathbb{Z}/v\mathbb{Z}$, define*

$$a + D = \{a + d \mid d \in D\}.$$

Then the collection $\mathcal{B} = \{a + D \mid a \in \mathbb{Z}/v\mathbb{Z}\}$ forms a symmetric $2 - (v, k, \lambda)$ design.

Proof. The point set is $\mathbb{Z}/v\mathbb{Z}$ and each block has size k. Let $x, y \in \mathbb{Z}/v\mathbb{Z}$ with $x \neq y$. We count the blocks containing both points.

The condition $x, y \in a + D$ means that there exist elements $d_i, d_j \in D$ such that $x = a + d_i, y = a + d_j$. Subtracting yields $x - y = d_i - d_j$.

Since D is a difference set, the difference $x - y$ can be written in exactly λ ways as a difference of elements of D. Hence x and y lie together in exactly λ blocks.

Therefore $(\mathbb{Z}/v\mathbb{Z}, \mathcal{B})$ is a $2 - (v, k, \lambda)$ design. □

Since there are exactly v translates of D, the resulting design has $b = v$ blocks and is therefore symmetric.

Example 9.35. *Starting from the difference set $D = \{1, 2, 4\} \subseteq \mathbb{Z}/7\mathbb{Z}$, the translates $D, 1+D, 2+D, \ldots, 6+D$ form the seven lines of the Fano plane. Thus the Fano plane can be constructed from a difference set.*

9.4 Steiner Triple Systems

Among the most important designs are those whose blocks have size three.

Definition 9.36. A *Steiner Triple System* of order v is a design with parameters $2-(v,3,1)$. Equivalently, it is a collection of triples such that every pair of points is contained in exactly one triple. The standard notation is $\mathrm{STS}(v)$.

Example 9.37. *The Fano plane is a Steiner Triple System. Indeed, it is a* $2-(7,3,1)$ *design and therefore an* $\mathrm{STS}(7)$.

Not every value of v admits a Steiner Triple System.

Theorem 9.38. *If an* $\mathrm{STS}(v)$ *exists, then* $v \equiv 1$ *or* $v \equiv 3 \pmod 6$.

Proof. Let $(X, \mathcal{B})$ be an $\mathrm{STS}(v)$. Since every pair of points lies in exactly one block, we have

$$b = \frac{\binom{v}{2}}{\binom{3}{2}} = \frac{v(v-1)}{6}.$$

As b is an integer, we obtain $6 \mid v(v-1)$.

Furthermore, $r = \frac{v-1}{2}$, so $v-1$ must be even.

Hence v is odd. Combining oddness with $3 \mid v(v-1)$ gives $v \equiv 1$ or $v \equiv 3 \pmod 6$. □

The remarkable fact is that these necessary conditions are also sufficient.

Theorem 9.39 (Kirkman)**.** *A Steiner Triple System of order* v *exists if and only if* $v \equiv 1$ *or* $v \equiv 3 \pmod 6$.

The proof is beyond the scope of this lecture.

Example 9.40. *The smallest Steiner Triple Systems are* $\mathrm{STS}(7), \mathrm{STS}(9)$. *We can easily check that* $\mathrm{STS}(13)$ *does exist since* $13 \equiv 1 \pmod 6$. *However, no Steiner Triple System exists for* $v = 5$ *because* $5 \not\equiv 1, 3 \pmod 6$.

9.5 Affine Planes and Resolvable Designs

In some designs, the blocks can be grouped into collections that partition the point set.

Definition 9.41. Let $(X, \mathcal{B})$ be a design. A set of blocks $\mathcal{P} \subseteq \mathcal{B}$ is called a *parallel class* if its blocks are pairwise disjoint and together partition the point set X.

A partition of the block set into parallel classes is called a *resolution*. A design admitting a resolution is called *resolvable*.

Example 9.42. *Consider the design on the point set $X = \{1, 2, 3, 4\}$ with blocks*

$$\{1, 2\}, \{3, 4\}, \{1, 3\}, \{2, 4\}, \{1, 4\}, \{2, 3\}.$$

The blocks can be partitioned into the three parallel classes $\big\{\{1, 2\}, \{3, 4\}\big\}$, $\big\{\{1, 3\}, \{2, 4\}\big\}$, and $\big\{\{1, 4\}, \{2, 3\}\big\}$. Thus the design is resolvable.

An important family of resolvable designs is provided by affine planes, where the notion of parallelism naturally gives rise to a resolution.

Theorem 9.43. *Every affine plane is resolvable.*

Proof. The lines of an affine plane split into parallel classes. Lines belonging to the same parallel class are pairwise disjoint, and every point lies on exactly one line of a given parallel class.

Hence each parallel class partitions the point set, and the collection of all parallel classes forms a resolution. □

9.6 Latin Squares and MOLS

We conclude our discussion of designs with another important combinatorial structure.

Definition 9.44. A *Latin square* of order n is an $n \times n$ array whose entries belong to $\{0, 1, \ldots, n-1\}$ such that every symbol occurs exactly once in each row and exactly once in each column.

Example 9.45. *The array*

$$\begin{array}{cccc} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \end{array}$$

is a Latin square of order 4.

One may think of a Latin square as a multiplication table in which each symbol appears exactly once in every row and column.

Definition 9.46. Let $A = (a_{ij})$ and $B = (b_{ij})$ be Latin squares of order n. The squares A and B are called *orthogonal* if the ordered pairs (a_{ij}, b_{ij}) are all distinct.

Equivalently, every element of $\{0, \ldots, n-1\}^2$ occurs exactly once among the pairs.

Example 9.47. *The Latin squares*

$$A = \begin{array}{ccc} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{array} \qquad B = \begin{array}{ccc} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{array}$$

are orthogonal.

Indeed, the nine pairs

$$(0,0), (1,1), (2,2), (1,2), (2,0), (0,1), (2,1), (0,2), (1,0)$$

are all distinct.

Definition 9.48. A collection of Latin squares $L_1, \dots, L_m$ is called a set of *mutually orthogonal Latin squares* (MOLS) if every pair L_i, L_j with $i \neq j$ is orthogonal.

The following result gives a fundamental limitation.

Theorem 9.49. *A set of MOLS of order n contains at most $n - 1$ squares.*

The maximum possible number is attained whenever n is a prime power.

Theorem 9.50. *If q is a prime power, then there exist $q - 1$ mutually orthogonal Latin squares of order q.*

The construction uses the finite field $\mathbb{F}_q$ and is one of the first applications of finite fields in combinatorics.

10 Codes

We now come to my personal favorite: codes!

Coding theory was developed to answer a very practical question:

How can we communicate reliably over an unreliable channel?

Suppose you want to send the binary message $m = (1, 0, 1)$. Unfortunately, during transmission, an error occurred and the receiver obtains $r = (0, 0, 1)$.

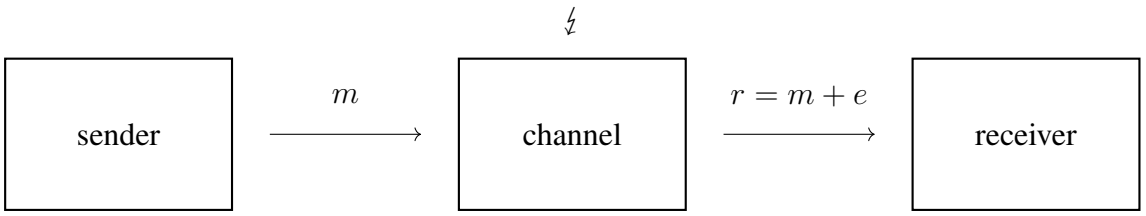
Can the receiver determine the original message?

At first glance the answer is *no*. The received word could equally well have come from

$$(1, 0, 1), \qquad (0, 0, 1), \qquad (0, 1, 1).$$

From the received vector alone, there is no way to determine which message was originally sent.

We model this noisy channel as follows:



We assume that the noise in the channel acts like an additive error, so that the received vector is given by $r = m + e$, where e is called the *error vector*. Its nonzero entries indicate the positions in which transmission errors occurred.

Since we cannot improve the communication channel itself, we instead change what we send. Rather than transmitting the message directly, we first *encode* it by adding redundancy. The receiver then uses this redundancy to recover the original message, even if transmission errors occur.

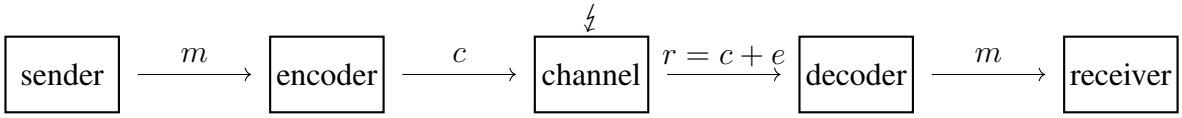
A first idea is simply to repeat the message several times. Instead of sending $m = (1, 0, 1)$, we send

$$c = (m, m, m) = (1, 0, 1, 1, 0, 1, 1, 0, 1).$$

If at most one of the three copies is corrupted, then a majority vote immediately recovers the original message. For example, if the receiver obtains

$$r = (1, 0, 0, 1, 0, 1, 1, 0, 1),$$

only one bit differs from the transmitted vector, making it easy to identify the original message.



Although repetition works, it is highly inefficient: the transmitted word is three times longer than the original message.

The goal of coding theory is therefore to design much more efficient encodings. We would like to add as little redundancy as possible while still being able to detect and correct transmission errors.

Remarkably, linear algebra provides exactly the right framework. Most of the codes studied in this chapter are simply linear subspaces of a vector space over a finite field.

Besides their applications to reliable communication, codes are also beautiful combinatorial objects. They are closely connected to several topics we have already encountered in this course. In particular,

- codes can be constructed from designs,
- many interesting codes themselves give rise to designs,
- and later we will see that both graphs and codes naturally lead to association schemes.

Our goal throughout this chapter is to develop the mathematical theory of linear error-correcting codes and understand its connections to designs.

10.1 Linear Codes

Throughout this chapter, let $\mathbb{F}_q$ denote the finite field with q elements, where q is a prime power.

Definition 10.1. An $[n,k]_q$ *linear code* is a k-dimensional linear subspace $\mathcal{C} \subseteq \mathbb{F}_q^n$. The elements of $\mathcal{C}$ are called *codewords*.

Note that we emphasize the linearity, as a *code* is simply any subset $\mathcal{C} \subseteq \mathbb{F}_q^n$.

We say that a code is *non-degenerate* if for any $i \in \{1,\dots,n\}$ there exists some $c \in \mathcal{C}$ with $c_i \neq 0$.

As $\mathcal{C}$ is linear, it must have some basis, which allows us to represent it compactly. In fact, linear codes allow for an easy representation through their *generator matrices*, which have the code as an image.

Definition 10.2 (Generator Matrix)**.** Let $k \leq n$ be positive integers and let $\mathcal{C}$ be an $[n,k]_q$ linear code. Then, a matrix $G \in \mathbb{F}_q^{k\times n}$ is called a *generator matrix* of $\mathcal{C}$ if

$$\mathcal{C} = \left\{ xG \mid x \in \mathbb{F}_q^k \right\},$$

that is, the rows of G form a basis of $\mathcal{C}$.

We will often write $\langle G \rangle$ to denote the code generated by the rows of G. Thus, we can easily check if a code is degenerate by checking whether a generator matrix has a zero column.

Example 10.3. *A repetition code $\mathcal{C} = \langle G \rangle$, is defined through the generator matrix*

$$G = \begin{pmatrix} 1 \cdots 1 \end{pmatrix} \in \mathbb{F}_q^{1 \times n}.$$

Thus, the code has dimension 1.

Given an $[n,k]_q$ linear code $\mathcal{C} = \langle G \rangle$, to *encode* a message $m \in \mathbb{F}_q^k$, we apply the map

$$\text{Enc} : \mathbb{F}_q^k \to \mathbb{F}_q^n, \quad m \mapsto mG.$$

Note that an $[n,k]_q$ linear code $\mathcal{C}$ has dimension k, that is $|\mathcal{C}| = q^k$. Thus, a generator matrix $G \in \mathbb{F}_q^{k \times n}$ of $\mathcal{C}$ has full rank k and in turn, the encoding map Enc is injective.

One can also represent a code through a matrix H, which has the code as kernel.

Definition 10.4 (Parity-Check Matrix)**.** Let $k \leq n$ be positive integers and let $\mathcal{C}$ be an $[n,k]_q$ linear code. Then, a matrix $H \in \mathbb{F}_q^{(n-k) \times n}$ is called a *parity-check matrix* of $\mathcal{C}$, if

$$\mathcal{C} = \left\{ y \in \mathbb{F}_q^n \mid yH^\top = 0 \right\}.$$

Note that there are many generator matrices G which have the same code as image. In fact, for any $U \in \text{GL}_k(q)$ we have that $\langle UG \rangle = \langle G \rangle$, as the row operation U only changes the basis but not the subspace.

Recall the rank-nullity theorem, stating that for any $A \in \mathbb{F}_q^{n \times m}$,

$$\dim(\ker(A)) = n - \text{rk}(A).$$

Since $\ker(H^\top) = \mathcal{C}$ has dimension k we have that $H \in \mathbb{F}_q^{(n-k) \times n}$ has full rank $n-k$.

Since $\mathcal{C} = \text{Im}(G) = \ker(H^\top)$, we also get a relation between the two matrices, namely

$$GH^\top = 0.$$

For $x, y \in \mathbb{F}_q^n$ let us denote by $\langle x, y \rangle$ the standard inner product, i.e.,

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

Then, we can define the dual of an $[n,k]_q$ linear code $\mathcal{C}$ as the orthogonal space of $\mathcal{C}$.

Definition 10.5 (Dual Code)**.** Let $k \leq n$ be positive integers and let $\mathcal{C}$ be an $[n,k]_q$ linear code. The *dual code* $\mathcal{C}^\perp$ is an $[n, n-k]_q$ linear code, defined as

$$\mathcal{C}^\perp = \{ x \in \mathbb{F}_q^n \mid \langle x, y \rangle = 0 \; \forall \; y \in \mathcal{C} \}.$$

We have to be careful with the term "dual code": we do not intend the space of linear forms on $\mathcal{C}$, instead we say "dual" to intend the orthogonal space with respect to the standard inner product. And even though $\dim(\mathcal{C}) + \dim(\mathcal{C}^\perp) = n$, we do not necessarily have that $\mathcal{C} \cap \mathcal{C}^\perp = \{0\}$. In fact, we can even have $\mathcal{C} = \mathcal{C}^\perp$.

Exercise 10.6. *Show that a parity-check matrix of $\mathcal{C}$ is in fact a generator matrix of $\mathcal{C}^\perp$.*

Proposition 10.7. *Let q be a prime power and $k \leq n$ be positive integers. Let $\mathcal{C}$ be an $[n,k]_q$ linear code. Then $(\mathcal{C}^\perp)^\perp = \mathcal{C}$.*

Proof. Let G be a generator matrix of $\mathcal{C}$ and H be a parity-check matrix of $\mathcal{C}$. Since H is thus a generator matrix of $\mathcal{C}^\perp$, we get that

$$(\mathcal{C}^\perp)^\perp = \{x \in \mathbb{F}_q^n \mid \langle x, y\rangle = 0 \text{ for all } y \in \mathcal{C}^\perp\}, \quad \mathcal{C}^\perp = \{y \in \mathbb{F}_q^n \mid \langle c, y\rangle = 0 \text{ for all } c \in \mathcal{C}\}.$$

Thus any $c \in \mathcal{C}$ is also in $(\mathcal{C}^\perp)^\perp$, as $\langle c, y\rangle = 0$ for all $y \in \mathcal{C}^\perp$.

Since $\mathcal{C}$ has dimension k and $(\mathcal{C}^\perp)^\perp$ has dimension k, we get the claim.

□

Given a generator matrix, one can easily find a parity-check matrix of the code.

Proposition 10.8. *Let $\mathcal{C}$ be an $[n,k]_q$ linear code and G be a generator matrix. If $G = \begin{pmatrix} Id_k & A \end{pmatrix}$, for some $A \in \mathbb{F}_q^{k\times(n-k)}$, then $H = \begin{pmatrix} -A^\top & Id_{n-k} \end{pmatrix}$ is a parity-check matrix of $\mathcal{C}$.*

Proof. We clearly have $GH^\top = -A + A = 0$, thus $\mathcal{C} = \langle G\rangle \subseteq \ker(H^\top)$. As H has rank $n-k$, the kernel of H has dimension k, and is thus equal to $\mathcal{C}$. □

Note that since G has full rank k, we may always permute the columns of G and apply a change of basis to get the identity matrix in the first k columns.

10.1.1 Hamming Metric

Now that we have described the linearity of the vectors we can send, how do we decode and how much can we decode?

Recall that we have received $r = c + e \in \mathbb{F}_q^n$, where $c \in \mathcal{C}$ was sent. To decode, we have to find the sent c. However, r could be the sum of any codeword and some other vector. In our toy example $r = (1,0,0,1,0,1,1,0,1)$ could also be written as

$$(1,1,1,1,1,1,1,1,1)+(0,1,1,0,1,0,0,1,0) \quad \text{or} \quad (1,0,0,1,0,0,1,0,0)+(0,0,0,0,0,1,0,0,1).$$

However these are less likely than

$$(1,0,1,1,0,1,1,0,1) + (0,0,1,0,0,0,0,0,0)$$

assuming that the channel only introduces a few errors. Hence, we are looking for the closest codeword c. For this, we first have to define what we mean by closest, i.e., introduce a metric to $\mathbb{F}_q^n$.

Definition 10.9. A *distance* is a function $d : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{Q}$, such that

- it is positive definite, i.e., $d(x, y) = 0$ if and only if $x = y$ and $d(x, y) \geq 0$ for all $x, y \in \mathbb{F}_q^n$,
- it is symmetric, i.e., $d(x, y) = d(y, x)$ for all $x, y \in \mathbb{F}_q^n$,
- it satisfies the triangle inequality, i.e., $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in \mathbb{F}_q^n$.

We say that a distance is translation-invariant, if for all $x, y, z \in \mathbb{F}_q^n$ we have $d(x, y) = d(x + z, y + z)$. A distance also gives raise to a weight, by defining $\mathrm{wt}(x) = d(x, 0)$.

Definition 10.10. A *weight* is a function $\mathrm{wt} : \mathbb{F}_q^n \to \mathbb{Q}$, such that

- it is positive definite, i.e., $\mathrm{wt}(x) = 0$ if and only if $x = 0$ and $\mathrm{wt}(x) \geq 0$ for all $x \in \mathbb{F}_q^n$,
- it is symmetric, i.e., $\mathrm{wt}(x) = \mathrm{wt}(-x)$ for all $x \in \mathbb{F}_q^n$,
- it satisfies the triangle inequality, i.e., $\mathrm{wt}(x + y) \leq \mathrm{wt}(x) + \mathrm{wt}(y)$ for all $x, y \in \mathbb{F}_q^n$.

In the other direction, any weight also induces a distance by defining $d(x, y) = \mathrm{wt}(x - y)$.

Exercise 10.11. *Show that if d is a translation-invariant distance function $\mathrm{wt}(x) = d(x, 0)$ is a weight function. Show that if wt is a weight function $d(x, y) = \mathrm{wt}(x - y)$ is a distance function.*

As we are interested in the amount of positions which are erroneous, the most natural weight to consider is the *Hamming metric*.

Definition 10.12 (Hamming Metric)**.** Let n be a positive integer. For $x \in \mathbb{F}_q^n$, the *Hamming weight* of x is given by the number of non-zero positions, i.e.,

$$\mathrm{wt}_H(x) = \mid \{i \in \{1, \ldots, n\} \mid x_i \neq 0\} \mid .$$

For $x, y \in \mathbb{F}_q^n$, the *Hamming distance* between x and y is given by the number of positions in which they differ, i.e.,

$$d_H(x, y) = \mid \{i \in \{1, \ldots, n\} \mid x_i \neq y_i\} \mid .$$

Note that the Hamming distance is induced from the Hamming weight, that is $d_H(x, y) = \mathrm{wt}_H(x - y)$.

Exercise 10.13. *Show that the Hamming weight is a weight function.*

Let $x \in \mathbb{F}_q^n$. We denote by $\mathrm{supp}_H(x) = \{i \in \{1, \ldots, n\} \mid x_i \neq 0\}$ the *Hamming support* of x. The Hamming weight of x is then clearly given as the size of the support: $\mathrm{wt}_H(x) = |\mathrm{supp}_H(x)|$.

Having defined a metric, one can also consider the minimum distance of a code, i.e., the smallest distance between any two distinct codewords.

Definition 10.14 (Minimum Distance)**.** Let $\mathcal{C}$ be a linear code over $\mathbb{F}_q$. The *minimum Hamming distance* of $\mathcal{C}$ is denoted by $d_H(\mathcal{C})$ and given by

$$d_H(\mathcal{C}) = \min\{d_H(x,y) \mid x,y \in \mathcal{C},\ x \neq y\} = \min\{\mathrm{wt}_H(c) \mid c \in \mathcal{C}, c \neq 0\}.$$

The minimum Hamming distance of a code turns out to be a very important parameter. Thus, whenever the minimum Hamming distance $d = d_H(\mathcal{C})$ is known, we say $\mathcal{C}$ is an $[n,k,d]_q$ linear code.

Example 10.15. *The $[n,1]_2$ repetition code has minimum distance $d_H(\mathcal{C}) = n$, whereas its $[n, n-1]_2$ dual code has minimum distance $d_H(\mathcal{C}^\perp) = 2$.*

When defining balls in a certain metric, we have to provide the radius and the center, e.g. we may define the Hamming ball of radius r and center x as

$$B_H(r,n,q,x) = \{y \in \mathbb{F}_q^n \mid d_H(x,y) \leq r\}.$$

However, to determine the size of such balls, we observe that the Hamming metric is translation invariant.

Proposition 10.16. *Let q be a prime power and $r \leq n$ be positive integers. For any $x, x' \in \mathbb{F}_q^n$ we have*

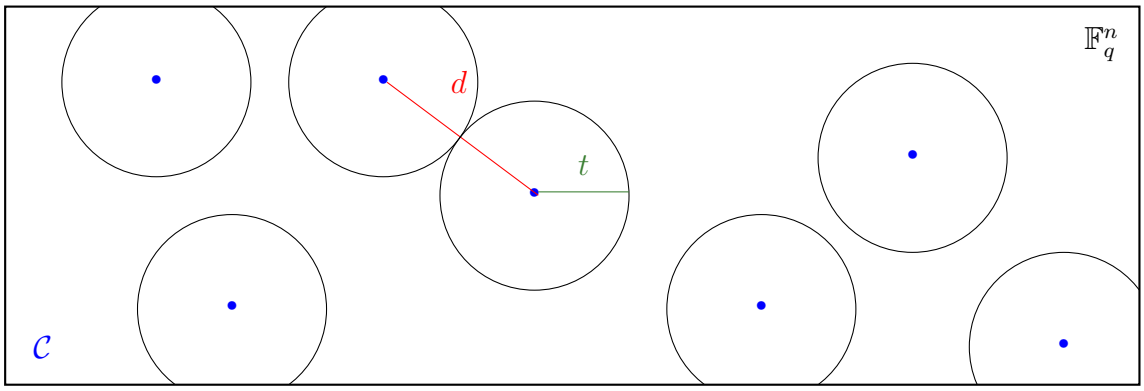
$$|B_H(r,n,q,x)| = |B_H(r,n,q,x')| = |\{y \in \mathbb{F}_q^n \mid wt_H(y) \leq r\}| = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i.$$

Exercise 10.17. *Prove Proposition 10.16*

The minimum distance of a code is an important parameter, since it is connected to the error correction capability t of the code. In fact, to decode means to find the unique closest codeword c to the received word $r = c + e$, and hence, we do not want to receive vectors r lying inside the intersection of two balls around codewords. The largest radius such that the balls around codewords do not intersect is given by

$$t = \left\lfloor \frac{d-1}{2} \right\rfloor.$$

As the intent of this course does not lie in decoding, we only give an intuition for this fact using a picture. However, if you want to learn more about this beautiful topic - do consider coming to the course "Coding Theory".



How do we determine the minimum distance without going through all codewords? One nice characterization is given by considering the independence of the columns of the parity-check matrix.

Theorem 10.18. *Let $k \leq n$ be positive integers and let $\mathcal{C}$ be an $[n,k]_q$ linear code. Let H be a parity-check matrix of $\mathcal{C}$. Then, $\mathcal{C}$ has minimum distance d if and only if every $d-1$ columns of H are linearly independent and there exist d columns, which are linearly dependent.*

Proof. For the first direction, let $c \in \mathcal{C}$ be the minimal weight codeword, i.e., $\mathrm{wt}_H(c) = d$. Since $c \in \ker(H^\top)$, we get $cH^\top = 0$ and thus the d columns h_i of H, indexed by $i \in \mathrm{supp}_H(c)$ are linearly dependent as $\sum_{i \in \mathrm{supp}_H(c)} c_i h_i = 0$. On the other hand, any $c \in \mathcal{C}$ with weight $< d$ must be the zero codeword. Thus, if a linear combination of less than d columns of H gives zero, the scalars are zero, i.e., any $d-1$ columns are linearly independent.

For the other direction, assume there exist d linearly dependent columns h_i of H for $i \in I$ and $|I| = d$, that is there exist $\lambda_i \in \mathbb{F}_q^\star$ such that $\sum_{i \in I} \lambda_i h_i = 0$. Then c with support I and entries $c_i = \lambda_i$ for $i \in I$ is such that $\mathrm{wt}_H(c) = d$ and $cH^\top = 0$, thus $c \in \mathcal{C}$ and hence $d_H(\mathcal{C}) \leq d$. Since any set of $d-1$ columns are linearly dependent, by the same argument, there is no non-zero codeword with weight less than d, thus $d_H(\mathcal{C}) \geq d$.

□

Example 10.19. *Let us consider the binary code $\mathcal{C}$ with parity-check matrix*

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

We can find 3 columns which are linearly dependent, for example the columns indexed by $\{1,4,6\}$, as

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

And any two columns are linearly independent, thus by Theorem 10.18, we get $d_H(\mathcal{C}) = 3$.

10.1.2 Bounds on Codes

When designing a code, we would like to maximize both its dimension and its minimum distance. Unfortunately, these two parameters are conflicting.

In this section we derive two of the most fundamental bounds in coding theory. The first bound is surprisingly simple.

Theorem 10.20 (Singleton Bound)**.** *Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code. Then*

$$d \leq n - k + 1.$$

Proof. Delete the last $d-1$ coordinates from every codeword. Since the minimum distance is d, no two codewords become equal after this deletion. Consequently, the resulting vectors are all distinct and lie in $\mathbb{F}_q^{n-d+1}$. Hence

$$|\mathcal{C}| = q^k \leq q^{n-d+1},$$

which immediately yields $k \leq n-d+1$. □

Codes attaining the Singleton bound are called *maximum-distance separable* (MDS) codes.

Example 10.21. *The repetition code is an MDS code, since it has parameters* $[n,1]_q$ *and minimum distance* n*. Indeed, we can easily check that* $d = n = n-1+1 = n-k+1$.

We have a construction of MDS codes for any $n \leq q+1$. You might wonder if there are MDS codes with $n > q+1$. The answer is: we do not know.

Conjecture 10.22. *Let* q *be an odd prime power and* $1 < k < n-1$ *be positive integers. Then any* $[n,k]_q$ *linear MDS code is such that* $n \leq q+1$.

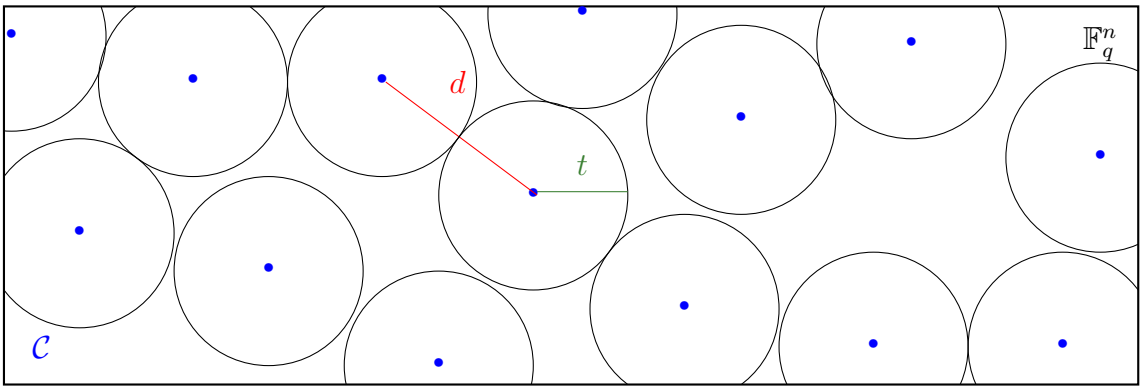
Interestingly, this conjecture is older than the Singleton bound itself. How can this be?
It was first formulated by Segre in 1955, not for codes, but for arcs - objects from finite geometry!

Apart from the Singleton bound, which is an upper bound on d, we also have a sphere-packing bound, called *Hamming bound*.

Let us denote by $V_H(r,n,q)$ the size (or volume) of a ball, i.e., for any $x \in \mathbb{F}_q^n$

$$V_H(r,n,q) = |B_H(r,n,q,x)| = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i.$$

The Hamming bound can be depicted as follows:



In this picture we cannot place any other codeword, without their balls of radius t intersecting.

Theorem 10.23 (Hamming Bound)**.** *Let* q *be a prime power and* $k \leq n$ *be positive integers. Let* $\mathcal{C}$ *be an* $[n,k]_q$ *code with* $d_H(\mathcal{C}) \geq d$ *and let* $t = \left\lfloor \frac{d-1}{2} \right\rfloor$*. Then*

$$|\mathcal{C}| \leq \frac{q^n}{V_H(t,n,q)}.$$

Proof. The balls of radius t around codewords of $\mathcal{C}$ have to be pairwise disjoint. In fact, if there exist $c \neq c' \in \mathcal{C}$ such that

$$x \in B_H(t, n, q, c) \cap B_H(t, n, q, c')$$

for some $x \in \mathbb{F}_q^n$, then by the triangle inequality

$$d_H(c, c') \leq d_H(x, c) + d_H(x, c') \leq 2t < d \leq d_H(\mathcal{C}),$$

a contradiction to $d_H(\mathcal{C})$ being the minimum distance among all distinct codewords.

Thus, we must have that the union of all balls around the codewords is disjoint and contained in $\mathbb{F}_q^n$, i.e.,

$$\bigcup_{c \in \mathcal{C}} B_H(t, n, q, c) \subseteq \mathbb{F}_q^n$$

and due to the disjointness

$$|\mathcal{C}| V_H(t, n, q) = \sum_{c \in \mathcal{C}} V_H(t, n, q) = |\bigcup_{c \in \mathcal{C}} B_H(t, n, q, c)| \leq q^n.$$

□

10.1.3 Hamming and Simplex Codes

We are again interested in codes which attain this bound.

Definition 10.24. Let q be a prime power and $k \leq n$ be positive integers. An $[n, k, d]_q$ code $\mathcal{C}$ with

$$q^{n-k} = V_H\left(\left\lfloor \frac{d-1}{2} \right\rfloor, n, q\right)$$

is called a *perfect* code.

They are called perfect, for a very special property: they can correct any received vector.

Proposition 10.25. *Let q be a prime power and $k \leq n$ be positive integers. Let $\mathcal{C}$ be an $[n, k, d]_q$ code and set $t = \lfloor \frac{d-1}{2} \rfloor$. Then, $\mathcal{C}$ is perfect if and only if for all $r \in \mathbb{F}_q^n$ there exists a unique codeword $c \in \mathcal{C}$ such that*

$$r \in B_H(t, n, q, c).$$

Exercise 10.26. *Prove Proposition 10.25.*

Example 10.27. *Let n be an odd positive integer and consider the $[n, 1]_2$ repetition code $\mathcal{C}$. Recall that $d_H(\mathcal{C}) = n = 2t + 1$, for some t and thus,*

$$|\mathcal{C}| = 2 = \frac{2^n}{2^{n-1}} = \frac{2^n}{\sum_{i=0}^{t} \binom{n}{i}}.$$

The fact, that we had to choose an odd minimum distance in this example is actually always true for perfect codes:

Proposition 10.28. *Let q be a prime power and $k \leq n$ be positive integers. Let $\mathcal{C}$ be an $[n, k, d]_q$ code. If d is even, then $\mathcal{C}$ is not perfect.*

Exercise 10.29. *Prove Proposition 10.28.*

There only exist very few non-trivial perfect codes: the Golay code $[23, 12, 7]_2$ the Golay code $[11, 6, 5]_3$ and the *Hamming codes.*

Definition 10.30. Let q be a prime power and $r \geq 2$ be a positive integer. Let $n = \frac{q^r-1}{q-1}$ and H be the $r \times n$ matrix having as columns all vectors in $\mathbb{F}_q^r$ up to non-zero scalar multiples. Then $\mathcal{C} = \ker(H^\top)$ is called *Hamming code.*

Example 10.31. *Let us consider $q = 2$ and $r = 3$ and define*

$$H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The code $\mathcal{C} = ker(H^\top)$ is a $[7, 4]_2$ Hamming code. Its generator matrix is given by

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Hamming codes are in fact perfect codes.

Proposition 10.32. *Let q be a prime power and $r \geq 2$ be a positive integer and set $n = \frac{q^r-1}{q-1}$. The $[n, n-r]_q$ Hamming code has $d_H(\mathcal{C}) = 3$ and is perfect.*

Proof. Let us first show that the construction of a Hamming code provides us with the parameters $n = \frac{q^r-1}{q-1}$ and $k = n - r$. We note that there exist $q^r - 1$ non-zero vectors in $\mathbb{F}_q^r$ and by excluding scalar multiples (which are $q - 1$ many) we get that there are $n = \frac{q^r-1}{q-1}$ many vectors in $\mathbb{F}_q^r$ up to scalar multiples, and thus $\mathcal{C}$ has length n. As the vectors $e_i \in \mathbb{F}_q^r$ with only the ith entry being 1 and the rest 0, are choices (up to scalar multiplication) for columns of H, we get that $\text{rk}(H) = r$ and hence $\dim(\mathcal{C}) = n - r$.

We now prove that $d_H(\mathcal{C}) = 3$. Since one column is not a scalar multiple of another column, any two columns must be linearly independent. Moreover, there exist 3 columns which are linearly dependent (for example choosing e_1, e_2 and $e_1 + e_2$ again up to scalar multiples). Thus, $d_H(\mathcal{C}) = 3$ by Proposition 10.18.

We can now show that Hamming codes are perfect. For this we note that $t = 1$ and

$$V_H(t, n, q) = \sum_{i=0}^{1} \binom{n}{i} (q-1)^i = 1 + n(q-1) = 1 + q^r - 1 = q^r.$$

Hence,

$$|\mathcal{C}| = q^{n-r} = \frac{q^n}{q^r}.$$

□

The dual code of the Hamming code is called *simplex* code.

Definition 10.33 (Simplex Code)**.** Let q be a prime power and $r \geq 2$ be a positive integer. Let $n = \frac{q^r-1}{q-1}$ and G be the $r \times n$ matrix having as columns all vectors in $\mathbb{F}_q^r$ up to non-zero scalar multiples. Then $\mathcal{C} = \langle G \rangle$ is called *simplex code.*

Corollary 10.34. *Let q be a prime power and $r \geq 2$ be a positive integer. Let $n = \frac{q^r-1}{q-1}$. The simplex code is a $[n, r]_q$ code.*

For example, the binary simplex code of dimension three is generated by

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Do you observe something odd when computing the weights of your codewords?

Exercise 10.35. *Show that all non-zero codewords of the simplex code have the same weight, that is q^{r-1}.*

10.2 Codes and Designs

One of the reasons coding theory plays an important role in combinatorics is its close connection with designs. We begin with the direction from designs to codes.

10.2.1 From Designs to Codes

The most natural way to define a code from a design, is to consider the space generated by the incidence matrix.

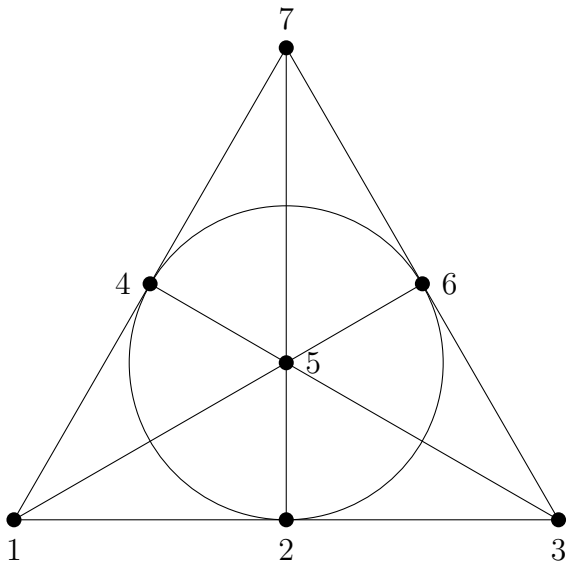
Definition 10.36. Let $\mathcal{D} = (X, \mathcal{B})$ be a design with $|X| = v, |\mathcal{B}| = b$ and incidence matrix $M \in \{0, 1\}^{v \times b}$. Let $q = p^m$, for some prime p and some positive integer m. The code $\mathcal{C}_{p^m}(\mathcal{D}) \subseteq \mathbb{F}_{p^m}^v$ *induced* from the design $\mathcal{D}$ is the code generated by $M^\top$.

While the length $n = v$ is clear, the dimension of the code is not. $M^\top \in \{0, 1\}^{b \times v}$ does not necessarily have rank b. Instead, the dimension of the induced code k is called the rank of the design and is denoted by $\text{rk}_{p^m}(\mathcal{D})$.

Example 10.37 (The Fano plane)**.** *Let us consider again the Fano plane, i.e.,* $X = \{1, 2, 3, 4, 5, 6, 7\}$ *and let*

$$\mathcal{B} = \Big\{\{1, 2, 3\}, \{1, 4, 7\}, \{3, 6, 7\}, \{2, 5, 7\}, \{1, 5, 6\}, \{3, 4, 5\}, \{2, 4, 6\}\Big\}.$$

Recall that this $(X, \mathcal{B})$ *is a* $2 - (7, 3, 1)$ *design.*



Its incidence matrix is given by

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Hence if we bring $M^\top$ *into a row reduced form, say by multiplying with an invertible matrix* U*, we get*

$$UM^\top = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence the code induced by the Fano plane is generated by

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

- this is a generator matrix of a Hamming code!

10.2.2 From Codes to Designs

The other direction, unfortunately does not always work. For some very special codes, we can however, find a design within the supports of their codewords.
For this we need the definition of a weight enumerator:

Definition 10.38 (Weight Enumerator)**.** Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a linear code. For any $w \in \{1, \ldots, n\}$, let us denote by $A_w(\mathcal{C}) = |\{c \in \mathcal{C} \mid \mathrm{wt}_H(c) = w\}|$ the *weight enumerator* of $\mathcal{C}$.

Theorem 10.39 (Assmus-Mattson)**.** *Let $\mathcal{C}$ be an $[n, k, d]_q$ code. Let $w \leq n$ be the largest integer with*

$$w - \left\lfloor \frac{w + q - 2}{q - 1} \right\rfloor < d.$$

Fix a positive integer $t < d$ and define

$$s = |\{i \in \{1, \ldots, n - t - 1\} \mid A_i(\mathcal{C}^\perp) \neq 0\}|.$$

If $s \leq d - t$, then the supports of the codewords in $\mathcal{C}$ of weight $i \leq w$ form a t-design.

We will skip the proof, as this would require either the MacWilliams identities, or association schemes - we might come back to it later.

Example 10.40. *Recall that the binary Hamming code has parameters $[7, 4, 3]_2$. Since $q = 2$, we have $w = n = 7$. The Hamming code has minimum distance $d = 3$, and its minimum-weight codewords all have weight 3.*

The supports of these codewords form a 2-$(7, 3, 1)$ design, namely the Fano plane.

A generator matrix of the Hamming code is

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

The seven minimum-weight codewords are

codeword c	$\mathrm{supp}(c)$
$(1,1,1,0,0,0,0)$	$\{1,2,3\}$
$(1,0,0,1,0,0,1)$	$\{1,4,7\}$
$(0,0,1,0,0,1,1)$	$\{3,6,7\}$
$(0,1,0,0,1,0,1)$	$\{2,5,7\}$
$(1,0,0,0,1,1,0)$	$\{1,5,6\}$
$(0,0,1,1,1,0,0)$	$\{3,4,5\}$
$(0,1,0,1,0,1,0)$	$\{2,4,6\}$

These supports are precisely the seven lines of the Fano plane.

11 Association Schemes

We arrived at the final chapter. Let us begin with a simple question.

Suppose we are given a collection of cities and we would like to describe how they are related. One natural possibility is to consider the distances between them. For every pair of cities we may ask whether they are

- the same city,
- 1 km apart,
- 2 km apart,
- ...

In this way, every pair of cities belongs to exactly one of several relations determined by their distance.

This idea appears throughout combinatorics. For example,

- in a graph, two vertices may have graph distance $0, 1, 2, \ldots$,
- in a code, two codewords may have Hamming distance $0, 1, \ldots, n$,
- in finite geometry, two points may coincide, lie on a common line, or not.

Association schemes provide a common framework for studying such partitions of the set of all pairs.

Definition 11.1. Let n be a non-negative integer. An *association scheme* with n classes is a pair $(X, \{R_0, R_1, \ldots, R_n\})$, where X is a finite, non-empty set and $R_0, R_1, \ldots, R_n$ are nonempty relations on X satisfying the following properties:

(A1) The relations $R_0, R_1, \ldots, R_n$ form a partition of $X \times X$, and

$$R_0 = \{(x, x) \mid x \in X\}.$$

(A2) For every $i \in \{0, \ldots, n\}$, the transpose relation

$$R_i^\top = \{(y, x) \mid (x, y) \in R_i\}$$

is equal to R_j for some $j \in \{0, \ldots, n\}$.

(A3) For every $i, j, k \in \{0, \ldots, n\}$ and every pair $(x, y) \in R_k$, the number of elements $z \in X$ satisfying $(x, z) \in R_i$ and $(z, y) \in R_j$ is a constant $p_{i,j}^k$ which only depends on i, j and k, but not on the particular choice of (x, y).

(A4) We also need that

$$p_{ij}^k = p_{ji}^k$$

for all i, j, k.

These constants $p_{i,j}^k$ are called *intersection numbers*, and are defined as

$$p_{ij}^k = |\{z \in X \mid (x,z) \in R_i,\ (z,y) \in R_j\}|\,.$$

We call the elements of X the *points* of the association scheme.

If every relation is symmetric, that is, $R_i = R_i^\top$ for all i, then the association scheme is called *symmetric*. In this case, Axiom (A4) follows automatically.

For every relation R_i, let $R_{i'} := R_i^\top$. The number

$$v_i := p_{ii'}^0$$

is called the *valency* of the relation R_i. Equivalently, for every point $x \in X$,

$$v_i = |\{y \in X \mid (x,y) \in R_i\}|\,,$$

that is, every point is related to exactly v_i other points through the relation R_i.

Let us have a look at some of the most important examples:

Example 11.2 (Graphs)**.** *Let $G = (V, E)$ be a graph and define*

$$\begin{aligned} R_0 &= \{(x,x) \mid x \in V\}, \\ R_1 &= \{(x,y) \mid x \sim y\}, \\ R_2 &= \{(x,y) \mid x \neq y,\ x \nsim y\}. \end{aligned}$$

These relations clearly partition $V \times V$. However, they do not always form an association scheme. Indeed, Axiom (A3) requires that the number of common neighbours of two adjacent vertices is constant, and likewise for two nonadjacent vertices.

Graphs satisfying these additional regularity conditions are called strongly regular graphs*.*

Example 11.3 (Hamming Scheme)**.** *Let n be a positive integer and let $\mathbb{F}_q$ be a finite field. The vertex set is $X = \mathbb{F}_q^n$. For $i \in \{0, \ldots, n\}$ define*

$$R_i = \{(x,y) \in X \times X \mid d_H(x,y) = i\},$$

where d_H denotes the Hamming distance. Then $(X, \{R_0, R_1, \ldots, R_n\})$ forms a symmetric association scheme, called the Hamming scheme *and denoted by $H(n,q)$.*

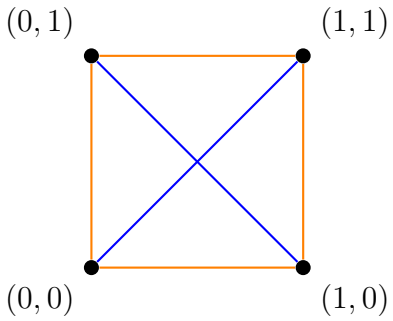
The valency of the relation R_i is

$$v_i = \binom{n}{i}(q-1)^i,$$

since we choose the i positions in which two words differ and then choose one of the $q-1$ possible symbols in each of these positions.

Note that every code $\mathcal{C} \subseteq \mathbb{F}_q^n$ can therefore be viewed as a subset of the Hamming scheme.

For example if we set $X = \mathbb{F}_2^2$, we get



With the relations R_1 and R_2.

Example 11.4 (Johnson Scheme)**.** *Let $m \geq n$ be positive integers. The vertex set consists of all n-subsets of a fixed set of size $m+n$. For $i \in \{0, \ldots, n\}$ define*

$$R_i = \{(x, y) \mid |x \cap y| = n - i\}.$$

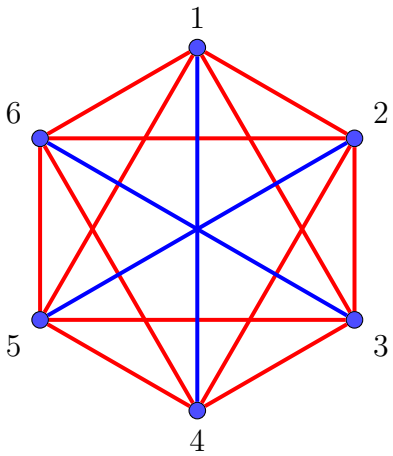
Then $(X, \{R_0, R_1, \ldots, R_n\})$ forms a symmetric association scheme, called the Johnson scheme*, and denoted by $J(n, m)$.*

The valency of the relation R_i is

$$v_i = \binom{n}{i}\binom{m}{i}.$$

Indeed, to obtain a set intersecting a given set in exactly $n-i$ elements, we remove i elements from the given set and replace them by i elements outside the set.

For example $J(2, 2)$ can be depicted as



with the relations R_1 and R_2.

Note that the Johnson scheme $J(n, m)$ can be embedded into the binary Hamming scheme $H(m+n, 2)$ as follows: Fix a n-set x in $J(n, m)$ and denote by ϕ_x its characteristic vector, that is $(\phi_x)_a = 1$ if $a \in x$ and $(\phi_x)_a = 0$ else. Then for two n-sets x, y we have that

$$n - |x \cap y| = i$$

if and only if

$$d_H(\phi_x, \phi_y) = 2i.$$

Thus in our Johnson scheme we have $(x, y) \in R_i$ if and only if $(\phi_x, \phi_y) \in R'_{2i}$ in the Hamming scheme.

11.1 Adjacency Matrices

Just as a graphs and designs can be represented by matrices, every relation of an association scheme can be as well.

Definition 11.5. Let $(X, \{R_0, \ldots, R_n\})$ be an association scheme with $|X| = v$. For each $i \in \{0, \ldots, n\}$, the *adjacency matrix* of the relation R_i is the matrix $A_i \in \{0, 1\}^{v \times v}$, where

$$a_{xy} = \begin{cases} 1, & \text{if } (x, y) \in R_i, \\ 0, & \text{otherwise.} \end{cases}$$

Example 11.6. *Let us consider the example of $H(2, 2)$. We first need to enumerate the points $\mathbb{F}_2^2$, for example as $X = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.*

The relation R_0 is simply given as $\{(x, x) \mid x \in \mathbb{F}_2^2\}$, hence its adjacency matrix is $A_0 = Id_4$.

The relation R_1 consists of all pairs vectors with distance 1, thus

$$A_1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

The relation R_2 consists of all pairs of vectors with distance 2, that is

$$A_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The adjacency matrices satisfy several nice properties.

Proposition 11.7. *Let $(X, \{R_0, \ldots, R_n\})$ be an association scheme with $|X| = v$ and adjacency matrices $A_0, A_1, \ldots, A_n$. Then*

1. *$A_0 = Id_v$.*
2. *$A_0 + A_1 + \cdots + A_n = J$, where J denotes the all-one matrix.*
3. *If the association scheme is symmetric, then $A_i^\top = A_i$ for every i.*
4. *For every i, j,*
$$A_i A_j = \sum_{k=0}^{n} p_{ij}^k A_k.$$

Proof. The first three properties follow immediately from the definition of an association scheme.

For the last identity, let $(x, y) \in X \times X$. The (x, y)-entry of $A_i A_j$ is

$$(A_i A_j)_{xy} = \sum_{z \in X} (A_i)_{xz} (A_j)_{zy}.$$

This counts the number of points $z \in X$ satisfying $(x, z) \in R_i$ and $(z, y) \in R_j$.

By Axiom (A3), this number depends only on the relation containing the pair (x, y). If $(x, y) \in R_k$, it is equal to the intersection number p_{ij}^k. Hence

$$A_i A_j = \sum_{k=0}^{n} p_{ij}^k A_k.$$

□

The vector space generated by the adjacency matrices $A_0, A_1, \ldots, A_n$ is called the *Bose–Mesner algebra* of the association scheme. We will not study this further here, but they play a fundamental role in algebraic combinatorics.

11.2 Codes in Association Schemes

Recall that a code is simply a subset $\mathcal{C} \subseteq \mathbb{F}_q^n$ of the vertex set of the Hamming scheme.

More generally, a code can be viewed as a subset of the point set of an arbitrary association scheme.

We first need to define the *inner distribution*.

Definition 11.8. Let $(X, \{R_0, \ldots, R_n\})$ be an association scheme and let $Y \subseteq X$. The *inner distribution* of Y is the tuple $D = (D_0, D_1, \ldots, D_n)$, where

$$D_i = \frac{|(Y \times Y) \cap R_i|}{|Y|},$$

Thus, D_i is the average number of points of Y that are in relation R_i with a fixed point of Y.

The inner distribution records how the points of Y are distributed among the different relations of the association scheme. A particularly important class of subsets are those whose points avoid the first few relations.

Definition 11.9. Let $1 \leq d \leq n$. A subset $Y \subseteq X$ is called a *d-code* if

$$D_1 = D_2 = \cdots = D_{d-1} = 0.$$

Equivalently, no two distinct points of Y are related by one of the relations

$$R_1, R_2, \ldots, R_{d-1}.$$

The notion of a d-code generalizes several important classes of codes from coding theory.

Example 11.10. *In the Hamming scheme $H(n, q)$, the relation R_i consists of all pairs of vectors at Hamming distance exactly i. Hence a d-code is precisely a q-ary code with minimum Hamming distance at least d.*

Example 11.11. *In the Johnson scheme $J(n, m)$, a d-code is a collection Y of n-subsets of an $(m + n)$-set such that*

$$n - |x \cap y| \geq d$$

for every two distinct sets $x, y \in Y$.

Identifying every subset x with its characteristic vector

$$\phi_x \in \mathbb{F}_2^{m+n},$$

like before, we obtain

$$d_H(\varphi_x, \varphi_y) = 2\big(n - |x \cap y|\big).$$

Hence

$$d_H(\varphi_x, \varphi_y) \geq 2d$$

for all distinct codewords.

These are precisely the binary constant-weight codes, i.e., simplex codes!

Finally, designs also give rise to association schemes.

Example 11.12. *Let $\mathcal{D} = (X, \mathcal{B})$ be a symmetric 2-(v, k, λ) design. Define the relations on the point set X by*

$$R_0 = \{(x, x) \mid x \in X\},$$
$$R_1 = \{(x, y) \mid x \neq y,\ x \text{ and } y \text{ lie in a common block}\}.$$

Since every pair of distinct points lies together in exactly λ blocks, every pair of distinct points belongs to the same relation. Thus the resulting association scheme has only one nontrivial relation.

There are many more interesting properties of association schemes, which we unfortunately do not have time to cover.

One particularly important application is the study of sphere packings and error-correcting codes. Association schemes provide the framework for Delsarte's linear programming bounds, which have led to some of the strongest known bounds on codes and sphere packings. Remarkably, ideas originating in coding theory and sphere packings played a central role in the solution of the sphere-packing problem in dimensions 8 and 24, work for which Maryna Viazovska was awarded the Fields Medal in 2022.

Another important combinatorial structure that we did not have time to study is that of *matroids*. Matroids provide an abstract notion of independence, simultaneously generalizing concepts from linear algebra, graph theory and optimization. They form yet another bridge connecting many of the topics discussed throughout this course.

Our journey through combinatorics has taken us from counting techniques, lattices, graphs and finite geometries to designs, codes and association schemes.

I hope you learned many new techniques and combinatorial objects and how they are all connected.

References

[1] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[2] R. P. Stanley. Enumerative combinatorics volume 1 second edition. *Cambridge studies in advanced mathematics*, 2011.

[3] J. H. Van Lint and R. M. Wilson. *A course in combinatorics*. Cambridge university press, 2001.

[4] H. Wilf. *Generatingfunctionology*. Academic Press, 1994.